Spanning structures in random graphs and hypergraphs

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 12
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von **Olaf Parczyk**
aus Frankfurt am Main

Frankfurt 2017
(D 30)

vom Fachbereich 12 der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Andreas Bernig

Gutachter: Jun.-Prof. Dr. Yury Person
Prof. David Conlon

Datum der Disputation: 18.12.2017

Abstract

Embedding spanning structures into the random graph $\mathcal{G}(n,p)$ is a well-studied problem in random graph theory, but when one turns to the random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$ much less is known. In this thesis we will examine this topic from different perspectives, providing insights into various aspects of the theory of random graphs. Our results cover the determination of existence thresholds in two models, as well as an algorithmic approach. For the embeddings, we work with random and pseudorandom structures.

Together with Person [93, 94] we first notice that a general result of Riordan [97] can be adapted from random graphs to hypergraphs and provide sufficient conditions for when $\mathcal{H}^{(r)}(n,p)$ contains a given spanning structure asymptotically almost surely. As applications, we discuss several spanning structures such as cubes, lattices, spheres, and Hamilton cycles in hypergraphs.

Moreover, we study universality, i.e. when does an r-uniform hypergraph contain *every* hypergraph on n vertices with maximum vertex degree bounded by Δ? For $\mathcal{H}^{(r)}(n,p)$, it is shown with Person [94] that this holds for $p = \omega(\ln n/n)^{1/\Delta}$ asymptotically almost surely by combining approaches taken by Dellamonica, Kohayakawa, Rödl, and Ruciński [41], of Ferber, Nenadov, and Peter [56], and of Kim and Lee [73].

Any hypergraph that is universal for the family of bounded degree r-uniform hypergraphs has to contain $\Omega(n^{r-r/\Delta})$ edges. With Hetterich and Person [64] we exploit constructions of Alon and Capalbo [11, 12] to obtain universal r-uniform hypergraphs with the optimal number of edges $O(n^{r-r/\Delta})$ when r is even, $r \mid \Delta$, or $\Delta = 2$. Furthermore, we generalise the result of Alon and Asodi [8] about optimal universal graphs for the family of graphs with at most m edges and no isolated vertices to hypergraphs.

In an r-uniform hypergraph on n vertices a tight Hamilton cycle consists of n edges such that there exists a cyclic ordering of the vertices where the edges correspond to consecutive segments of r vertices. In collaboration with Allen, Koch, and Person [6] we provide a first deterministic polynomial time algorithm, which finds asymptotically almost surely tight Hamilton cycles in random r-uniform hypergraphs with edge probability at least $C\log^3 n/n$. This result partially answers a question of Nenadov and Škorić [92] and of Dudek and Frieze [44] who proved that tight Hamilton cycles exist already for $p = \omega(1/n)$ for $r = 3$ and $p \geq (e + o(1))/n$ for $r \geq 4$ using a second moment argument. Moreover our algorithm is superior to previous results of Allen, Böttcher, Kohayakawa, and Person [4] and Nenadov and Škorić [92].

Lastly, we study the model of randomly perturbed dense graphs introduced by Bohman, Frieze and Martin [24], that is, the union of any n-vertex graph G_α with minimum degree at least αn and $\mathcal{G}(n,p)$. For any fixed $\alpha > 0$, and $p = \omega(n^{-2/(\Delta+1)})$, we show with Böttcher, Montgomery, and Person [32, 31] that $G_\alpha \cup \mathcal{G}(n,p)$ almost surely contains any single spanning graph with maximum degree Δ, where $\Delta \geq 5$. As in previous results concerning this model, the bound used for p is lower by a $\log$-term in comparison to the conjectured threshold for the general appearance of such subgraphs in $\mathcal{G}(n,p)$ alone. The new techniques we introduce also give simpler proofs of related results in the literature on trees [82] and factors [19].

Zusammenfassung

Das Finden von aufspannenden Strukturen im zufälligen Graphen $\mathcal{G}(n,p)$ is ein viel studiertes Problem in der Theorie der zufälligen Graphen, aber sobald man sich dem zufälligen r-uniformen Hypergraphen $\mathcal{H}^{(r)}(n,p)$ zuwendet ist noch deutlich weniger bekannt. In dieser Arbeit beschäftigen wir uns mit diesem Thema aus verschiedenen Blickwinkeln und geben dabei einen Einblick in viele Aspekte des Studiums von zufälligen Graphen. Zu unseren Ergebnissen gehören sowohl die Bestimmung von Schwellenwerten in verschiedenen Modellen als auch ein algorithmischer Zugang. Für die Einbettungen arbeiten wir mit zufälligen und pseudozufälligen Strukturen.

Zusammen mit Person [93, 94] stellen wir zuerst fest, dass sich ein allgemeines Ergebnis von Riordan [97] von zufälligen Graphen auf Hypergraphen verallgemeinern lässt, und zeigen eine hinreichende Bedingung dafür, dass $\mathcal{H}^{(r)}(n,p)$ eine gegebene aufspannende Struktur asymptotisch fast sicher enthält. Als Anwendung diskutieren wir verschiedene Strukturen, wie Würfel, Gitter und Hamiltonkreise in Hypergraphen.

Desweiteren studieren wir Universalität, also die Frage, wann ein r-uniformer Hypergraph *alle* Hypergraphen auf n Knoten mit maximalem Knotengrad höchstens Δ enthält. Für $\mathcal{H}^{(r)}(n,p)$ zeigen wir mit Person [94], dass dies für $p = \omega(\ln n/n)^{1/\Delta}$ asymptotisch fast sicher stimmt, indem wir Ideen von Dellamonica, Kohayakawa, Rödl and Ruciński [41], von Ferber, Nenadov and Peter [56] und von Kim und Lee [73] kombinieren.

Jeder Hypergraph, der universal für die Familie der gradbeschränkten Hypergraphen ist, muss mindestens $\Omega(n^{r-r/\Delta})$ Kanten besitzen. Mit Hetterich und Person [64] nutzen wir Konstruktionen von Alon und Capalbo [11, 12] aus, um daraus universale r-uniforme Hypergraphen mit optimaler Kantenanzahl $O(n^{r-r/\Delta})$ zu konstruieren, falls r gerade ist, $r \mid \Delta$ oder $\Delta = 2$. Darüberhinaus verallgemeinern wir ein Resultat von Alon und Asodi [8] über optimale universale Graphen für die Familie der Graphen mit m Kanten und ohne isolierte Knoten auf Hypergraphen.

In einem r-uniformen Hypergraphen auf n Knoten besteht ein enger Hamiltonkreis aus n Kanten, so dass es eine zyklsiche Anordnung der Knoten gibt, in der die Kanten zu aufeinanderfolgenden Segmenten gehören. In Kollaboration mit Allen, Koch und Person [6] finden wir einen ersten deterministischen Polynomialzeitalgorithmus, der asymptotisch fast sicher einen engen Hamiltonkreis in $\mathcal{H}^{(r)}(n,p)$ findet für $p \geq C \log^3 n/n$. Damit beantworten wir teilweise eine Frage von Nenadov und Škorić [92] und von Dudek und Frieze [44], die zeigten, dass enge Hamiltonkreise bereits für $p \geq (e + o(1))/n$ exisiteren für $r \geq 4$ ($p = \omega(1/n)$ für $r = 3$), indem sie die Methode des zweiten Moments anwendeten. Desweiteren verbessern wir zuvorige Algorithmen von Allen, Böttcher, Kohayakawa und Person [4] und Nenadov und Škorić [92].

Zuletzt widmen wir uns dem Modell der zufällig manipulierten dichten Graphen, dass von Bohman, Frieze und Martin [24] eingeführt wurde. In diesem Modeel betrachten wir die Vereinigung von einem Graphen G_α auf n Knoten mit Minimalgrad αn und $\mathcal{G}(n,p)$. Für ein fixiertes $\alpha > 0$, und $p = \omega(n^{-2/(\Delta+1)})$ zeigen wir mit Böttcher, Montgomery und Person [32, 31], dass $G_\alpha \cup \mathcal{G}(n,p)$ asymptotisch fast sicher einen beliebigen aufspannenden Graphen auf n Knoten mit Maximalgrad Δ

ernhält, falls $\Delta \geq 5$. Ebenso wie in vorherigen Ergebnissen in diesem Modell ist die Schranke an p um einen $\log$-Faktor kleiner als der vermutete Schwellenwert für das Auftreten dieser Strukturen in $\mathcal{G}(n,p)$ alleine. Unsere neue Methode ergibt auch einfachere Beweise für einige verwandte Probleme über Bäume [82] und Faktoren [19].

Contents

Chapter 1

Introduction

In the last 60 years, the area of discrete mathematics grew enormously in importance, not least because of the development of computers. In particular, graph theory did benefit significantly from this development since networks were used as models in all natural sciences. Moreover, it was simultaneously discovered that randomness is a powerful tool for many applications in discrete structures, including randomised algorithms and the probabilistic method.

The theory of random graphs is a fascinating field inside of discrete mathematics, which lies at the intersection of probability theory, graph theory, and combinatorics. Besides these intersections, there are numerous connections to other areas, for example, theoretical computer science, information theory, and statistical physics, which are mutually beneficial for either side. A major motivation for understanding the fundamental nature of random objects was to conceive the behaviour of typical instances in real-world applications. However, it was later on discovered that random instances are usually much harder and generate tough benchmarks for algorithms.

In the next section we give a short history of the evolution of random graph theory. Afterwards in Section 1.2 there will be a brief summary of the results of this thesis without precise statements. We then conclude the introduction with preliminary remarks and notation.

1.1 Random graphs

The first appearance of a random graph is often devoted to a 1947 paper of Erdős [51], where the bounds on diagonal Ramsey numbers are improved by showing the existence of a certain Ramsey graph[1]. The binomial random graph model $\mathcal{G}(n,p)$, which is a probability space on all graphs on n vertices, where edges are drawn uniformly and independent with probability p, was first introduced by Gilbert [62], who studied connectedness in this model. However, this graph model is often attributed to Erdős and Rényi [48], who in fact started working on the same problem in the hypergeometric model $\mathcal{G}(n,M)$[2], which is choosing a graph uniformly at random from all graphs on n vertices and M edges. In the following years Erdős and Rényi [49, 50] set the groundwork for the emerging field of random graphs. One of their main interests[3] was the emergence of the giant component as p passes $1/n$.

Besides more profound work on this phase transition, some of the major achievements in the theory of random graphs are the transference of results from extremal combinatorics [18, 35, 101, 102] , the

[1] This can also be seen as one of the first conscious applications of the *probabilistic method* [15] which utilises the simple observation, that if an event has non-zero probability, then there exists an instance where this event occurs.

[2] There is a close relation between the two models and in many regards they are equivalent, cf. Łuczak [86]. For convenience, we will mostly work in $\mathcal{G}(n,p)$.

[3] We will discuss some of their other results in Section 2.1.

advances on the KŁR Conjecture[4] in random graphs [37], the solution of Ramsey-type questions [98], the analysis of the chromatic number[5] in the dense [27] and sparse [1] regime, and the embedding of general factors [67], which we will discuss later.

Apart from this, there has been a lot of development leading to many other great results and thus it is impossible to give an exhaustive survey. In particular, we would like to emphasise three monographs on random graphs [30, 60, 65], each roughly 15 years apart, which reflect the evolution of the field. Besides $\mathcal{G}(n,p)$ and $\mathcal{G}(n,M)$ there are several other random graph models, e.g. random regular graphs, intersection graphs, and preferential attachment models with a power law degree distribution, which are motivated by observations from real-world networks. It might be due to its simplicity, that $\mathcal{G}(n,p)$ is the most studied model of these and still there are many interesting problems and phenomena which are not well understood.

Typical questions in $\mathcal{G}(n,p)$ deal with the investigation of graph parameters and the structure of the graph. For example, there is a lot of research on the size and structure of the largest connected component, the size of the largest independent set, the chromatic number, and the embedding of various substructures. The spanning version of the latter is the central topic of this thesis. To be a little more precise, we are interested in finding the smallest p such that we can embed[6] a specific graph on n vertices into $\mathcal{G}(n,p)$. Embedding spanning subgraphs is well studied for various kinds of graphs such as perfect matchings, Hamilton cycles, trees, factors, and to some extent general bounded degree graphs. Nonetheless, many questions remain open despite years of extensive research.

In the case of hypergraphs even less is known and it is natural to study the corresponding problems for random hypergraphs. The random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$ is the model, where on n vertices any r-set is an edge with probability p independent of all the others. For $r = 2$ this reduces to $\mathcal{G}(n,p)$.

1.2 Summary of results

In this thesis, we examine spanning structures in random graphs from different perspectives. We obtain a pure existence statement, embed many structures simultaneously, investigate an algorithmic approach, and analyse the combination of random with deterministic properties. This versatility enables us to shed light on various aspects and properties of random graph theory. We now give a brief summary of the results, without precise statements. For a more detailed exposition and discussion, we refer to Chapter 2. Here we follow the chronological order of submission, whereas in the rest of the thesis the results are sorted by topic.

First, we prove a general result for the embedding of (spanning) hypergraphs into $\mathcal{H}^{(r)}(n,p)$ (Theorem 2.5), which is a generalisation of a theorem by Riordan [97] from the graph case. The proof uses detailed second moment calculations and we present several applications giving asymptotically optimal results for some classes of hypergraphs, such as cubes, lattices, spheres, and Hamilton cycles. This result was obtained together with Person [94] and the proof is given in Chapter 3.

In the same paper together with Person [94] we also studied universality for bounded degree

[4]This conjecture by Kohayakawa, Łuczak and Rödl [74] roughly says that almost all graphs satisfy a sparse counting lemma.
[5]The typical chromatic number of $\mathcal{G}(n,p)$ is the last remaining open question from the early paper by Erdős and Rényi [48].
[6]By embedding we mean an injective map between the vertex sets, which respects the edges of the graph we are embedding.

graphs in $\mathcal{H}^{(r)}(n,p)$. A graph is universal for a family of graphs if it contains every graph from the family as a subgraph. This is a much stronger statement than just containing one graph from the family and most results cannot be easily extended. We obtain Theorem 2.9, which generalises a result of Dellamonica, Kohayakawa, Rödl, and Rucinski [41] on universality in $\mathcal{G}(n,p)$ to hypergraphs. For the proof given in Chapter 6 we employ ideas of Ferber, Nenadov, and Peter [56] and of Kim and Lee [73] to find a deterministic, pseudorandom structure in $\mathcal{H}^{(r)}(n,p)$, which enables us to embed any bounded degree graph.

Even further, starting in this paper with Person [94] and afterwards continued with Hetterich and Person [64], we work on the existence and explicit construction of universal hypergraphs. By exploiting constructions of Alon and Capalbo [9, 10] we manage to obtain universal hypergraphs for a wide range of parameters, which are even sparser than the random hypergraphs mentioned before. Further, we generalise a result of Alon and Asodi [8] about optimal universal graphs for the family of graphs with at most m edges and no isolated vertices to hypergraphs. See Chapter 7 for the proofs of these results.

We also obtain an algorithmic result together with Allen, Koch, and Person [6]. Many of the results for finding subgraphs do not yield any meaningful algorithms and are purely existence statements. Improving on previous results by Allen, Böttcher, Kohayakawa, and Person [4] and Nenadov and Škorić [92] we present a first deterministic polynomial time algorithm for finding tight Hamilton cycles in $\mathcal{H}^{(r)}(n,p)$ with probability a small polylog-factor away from the optimal bound. The proof of this result (Theorem 2.6) uses the so-called absorber technique and is given in Chapter 4.

The last result is in a slightly different model of randomly perturbed graphs introduced by Bohman, Frieze, and Martin [24], where we take a graph G_α of minimum degree αn and then add $\mathcal{G}(n,p)$ on top of it. This combines the elements of random and extremal graph theory and opens many possibilities. Typically in this model one can save some $\log$-terms compared to the probability in $\mathcal{G}(n,p)$ purely, as already shown for Hamilton cycles in the above mentioned paper. Together with Böttcher, Montgomery, and Person [32] we prove the corresponding analog for the family of bounded degree graphs, Theorem 2.7. Our method, which uses ideas from Ferber, Luh, and Nguyen [54], also reproves results on bounded degree trees by Krivelevich, Kwan, and Sudakov [82] and factors by Balogh, Treglown, and Wagner [19]. We discuss these implications together with the proof of our result in Chapter 5.

1.3 Preliminaries and notation

This section gives a short outline of the structure of the remainder of the thesis as well as an introduction to the notation and syntax used therein. In Chapter 2 we introduce the main concepts, explain related work, and then present, discuss, and analyse our results in detail. Afterwards, in Chapters 3–7 we give the proofs of our theorems. Concluding remarks and a brief discussion of open problems complete the thesis in Chapter 8.

We want to remark at this point that this thesis is a cumulation of the results from four different papers [4, 32, 64, 94], each with a different set of coauthors. The extension of Riordan's theorem from [94] also appeared previously as an extended abstract [93], the main result from [32] appeared in [31], and the result from [6] in [5]. In all four papers, the author of this thesis contributed significantly to all stages, beginning from the research conducted, until the preparation of the final paper. Large parts of this thesis will be verbatim copies from the papers, in particular Chapter 3 from [94], Chapter 4 from [4], Chapter 5 from [32], Chapter 6 from [94] and Chapter 7 from [64] with a short paragraph from [94]. Moreover, the following notation, the abstract, German summary, some paragraphs explaining the respective results in Chapter 2 and short parts of the Conclusion in Chapter 8 are close adaptations from the corresponding parts in [4, 32, 64, 94]. All these results do not appear in any other thesis.

To state the results in detail, avoid too much repetition, and have a point of reference for the reader we now collect the basic notation. We mostly follow the standard notation from [30, 60, 65] and thus the reader who is familiar with the basics can skip the rest of this section. We state the definitions for hypergraphs and remark that with $r = 2$ this gives the corresponding definitions for graphs, in which case we usually omit the superscript. An *r-uniform hypergraph* H is a tuple (V, E), where $V(H) := V$ is its vertex set and $E(H) := E \subseteq \binom{V}{r}$ is the set of edges in H. We write $v(H)$ for $|V(H)|$ and $e(H)$ for $|E(H)|$. By $K_n^{(r)}$ we denote the *complete r-uniform hypergraph* $\left([n], \binom{n}{r}\right)$ on the vertex set $[n] := \{1, 2, \ldots, n\}$. The *random r-uniform hypergraph* $\mathcal{H}^{(r)}(n, p)$ is the probability space of all labelled r-uniform hypergraphs on the vertex set $[n]$, where each edge $e \in \binom{[n]}{r}$ is chosen independently of all the other edges with probability p.

We say that a graph H contains a graph G as a *subgraph* if there exists a map ϕ from $V(G)$ to $V(H)$ such that edges are preserved, i.e. for all $e \in E(G)$ we have that $\phi(e) \in E(H)$. If G is a subgraph of H we write $G \subseteq H$. The subgraph *induced* by a subset of the vertices $W \subseteq V$ in H is denoted by $H[W] := \left(W, E(H) \cap \binom{W}{r}\right)$ and we define $H - W = H[V \setminus W]$. We denote by $\deg_H(f) := |\{e\colon f \subseteq e\}|$ the *degree* of a set of vertices f of size $1 \leq |f| \leq r - 1$ in H, i.e. the number of edges f is contained in. Given a set $W \subseteq V$, we write $\deg_H(f, W)$ for the degree into W, that is, we count only edges e satisfying $e \setminus f \subset W$. Further, $\Delta_\ell(H)$ is defined to be the *maximum ℓ-degree* in H, i.e. $\Delta_\ell(H) := \max\left\{\deg_H(f)\colon f \in \binom{V}{\ell}\right\}$. We usually omit the subscript if $\ell = 1$. With $d(H) := \frac{e(H)}{v(H)}$ and $d_1(H) := \frac{e(H)}{v(H)-1}$ we define the *density* $m(H) := \max\{d(H')\colon H' \subseteq H\}$ and the *one-density* $m_1(H) := \max\{d_1(H')\colon H' \subseteq H\}$. The *shadow graph* H' is obtained from H by replacing every edge $e \in E(H)$ by all possible $\binom{r}{2}$ subsets of cardinality two (ignoring multiple edges).

An alternating sequence of vertices and edges $v_1, e_1, v_2, e_2, \ldots, v_t, e_t, v_{t+1}$ is called a *path*[7] of length t from v_1 to v_{t+1} if $v_i, v_{i+1} \in e_i$ for all $i \in [t]$. If there is a path from u to v, then we say that u and v are *connected*. This defines an equivalence relation on V. We say that a hypergraph H is *connected* if there is a path between any two vertices of H. A *component* in an r-uniform hypergraph is a maximally connected subgraph. The *distance* between two vertices u and v in H is the minimal length over all paths from u to v, and if they are in different components then we set it to infinity.

The *neighbourhood* $N_H(v)$ of a vertex v is the set of vertices which are contained in an edge together with v

$$N_H(v) := \{w \in V \setminus \{v\} : \exists\, e \in E \text{ s.t. } \{w, v\} \subseteq e\}.$$

For a subset of the vertices $W \subseteq V$, the neighbourhood in H is $N_H(W) = \bigcup_{w \in W} N_H(w) \setminus W$. If there is

[7]In our definition of a path, we do not mind repetitions of vertices and edges. Usually, this is referred to as a *walk* rather than a path.

no risk of confusion, we sometimes omit the graph in the subscript. The set W is called *t-independent* in a hypergraph H, if the distance between $v \in W$ and $w \in W$ in H is at least $t + 1$. A 1-independent set is independent in the usual sense.

Let f and g be real valued functions, where we usually omit the dependencies on variables. We write $g = O(f)$[8] if g is not growing much faster than f, i.e. there exist $C > 0$ and $n_0 > 0$ such that for all $n > n_0$ we have $|f(n)| \leq C \cdot |g(n)|$. If $f = O(g)$, then vice versa g is not decreasing much faster than f and thus we can also write $g = \Omega(f)$, and if both statements hold $g = \Theta(f)$. We usually make no effort in optimising the constant C hidden in this notation. Sometimes we include other constants as a subscript to point out that C depends on them. By $\omega(f)$ we denote any function g that is growing faster than f, i.e. $g(n)/f(n) \to \infty$ as $n \to \infty$, and similarly $o(g)$ denotes any function f such that $g = \omega(f)$. For brevity we will often use $\omega(1)$ instead of saying that there exists a large enough constant C, even though this is is general.

With $\ln n$ we denote the natural logarithm, bute we usually use $\log n$ if the base does not matter. With a polylog-factor or $\operatorname{polylog} n$ we refer to any polynomial in $\log n$. To simplify readability, we will omit in the calculations floor and ceiling signs whenever they are not crucial for the arguments.

[8]Formally it should be $g \in O(f)$ but it is common pratice to abuse the notation in this way.

Chapter 2

Results, discussion and outline

This is the main chapter of this thesis as all our results are motivated and explained. We will also give brief sketches of the proofs and include a short discussion on the methods and techniques that are involved, but the rigorous proofs will be found in later chapters of the thesis. In order to embed the theorems into a wider context, we will introduce some established concepts and well-known results which will eventually lead to our main results. We will guide the reader from single spanning structures, via algorithmic results for Hamilton cycles and the randomly perturbed graphs model, to universality. The last part of this chapter is used to introduce standard tools that we will make extensive use of in the remainder of the thesis.

2.1 Thresholds

A graph property $\mathcal{F}$ is a set of graphs. This set could for example consist of all graphs with a specific subgraph, special structure, fixed chromatic number or any other graph parameter. Consider, for example, the graph property $\mathcal{F}_{HAM}$ of having a *Hamilton cycle* as a subgraph, then a graph H is in $\mathcal{F}_{HAM}$ if there exists a cyclic ordering of the vertices of H such that neighbouring vertices are adjacent. It is a classical result of Dirac [42] that every graph on n vertices with minimum degree at least $n/2$ satisfies this property and thus is Hamiltonian. This result is sharp in the sense that there are graphs with minimum degree slightly below $n/2$ that do not contain a Hamilton cycle and thus minimum degree $n/2$ is a distinguished point.

In random graph theory one of the most natural objectives is to study similar graph properties in $\mathcal{G}(n,p)$ as previously in deterministic graphs. For example we can ask when does $\mathcal{G}(n,p)$ contain a Hamilton cycle. More precisely, we can ask for which values of p does $\mathbb{P}[\mathcal{G}(n,p) \in \mathcal{F}]$ tend to 1 as n tends to infinity for some property $\mathcal{F}$. If this is true, then we say that $\mathcal{G}(n,p)$ has the property $\mathcal{F}$ *asymptotically almost surely* (a.a.s.).

Containing a subgraph is a *monotone* property, which means that adding edges cannot destroy the property and thus with larger p the probability that $\mathcal{G}(n,p)$ has this property increases. With this observation in mind it makes sense to ask for a value $\hat{p}$ below which $\mathcal{G}(n,p)$ does not have the property a.a.s. but above it has. It turns out that often there is a very abrupt change in behaviour and thus we say that $\hat{p} : \mathbb{N} \to [0,1]$ is a *threshold function* for a graph property $\mathcal{F}$ if

$$\mathbb{P}\big[\mathcal{G}(n,p) \in \mathcal{F}\big] \begin{cases} \to 0 & \text{if } p = o(\hat{p}) \\ \to 1 & \text{if } p = \omega(\hat{p}). \end{cases}$$

Sometimes this kind of threshold is referred to as *coarse*, where for a *sharp* threshold we require for any $\varepsilon > 0$ that $p \le (1-\varepsilon)\hat{p}$ and $p \ge (1+\varepsilon)\hat{p}$ already are sufficient for the convergence. For a criterion and discussion of which thresholds are sharp see Friedgut [57, 58]. We will mostly focus on coarse thresholds and thus for simplicity refer to them as thresholds. In the random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$ thresholds are defined analogously.

It was shown by Bollobás and Thomason [29] that all nontrivial monotone properties admit a threshold function. As mentioned before containing a subgraph is a monotone property and thus it makes sense to study the thresholds of these properties. There are many other interesting graph properties admitting a threshold behaviour, but we do not go into details here. Note that a function $\hat{p}$ satisfying $\mathbb{P}[\mathcal{G}(n,\hat{p}) \in \mathcal{F}] = 1/2$ always is a threshold if $\mathcal{F}$ is a nontrivial monotone property.

As a first example for a subgraph, we consider a fixed small graph G. It is necessary that the expected number of copies of any subgraph G' of G in $\mathcal{G}(n,p)$, which is roughly $n^{v(G')}p^{e(G')}$, does not tend to zero[9]. From this we easily obtain with the density $m(G) = \max\left\{\frac{e(G')}{v(G')} : G' \subseteq G\right\}$ that p has to be at least $n^{-1/m(G)}$. In their early, seminal work in 1960 Erdős and Rényi [49] proved that this in fact gives the threshold if G is balanced, which means that F itself is not sparser than any subgraph, i.e. $m(G) = \frac{e(G)}{v(G)}$. This was much later extended by Bollobás [25] to all graphs G and also extends to hypergraphs.

2.2 Single spanning structures

Advancing to spanning subgraphs, a first example is the *perfect matching*, which is the disjoint union of $n/2$ edges (n even). The expected number of perfect matchings in $\mathcal{G}(n,p)$ is larger than 1 already shortly after p passing $1/n$, but a.a.s. $\mathcal{G}(n,p)$ still contains many isolated vertices at this range of p. Thus, there have to be some events, which are not to rare, containing many perfect matchings, and therefore push up the expected number, even though most graphs do not contain a single perfect matching. The threshold was determined in another paper by Erdős and Rényi [50][10] at $\log n/n$. This function is also a threshold for the property of the minimum degree being larger than a given constant and in particular for minimum degree 1, which is necessary for a perfect matching. Łuczak and Ruciński [87] proved that in the graph process, where we start with an empty graph and add edges uniformly at random, at the precise moment where the graph has minimum degree 1 it already has a perfect matching a.a.s. Also note that for the connectivity property, which basically is the containment of any spanning tree and was the first property studied in $\mathcal{G}(n,p)$ by Gilbert [62], the same is true and $\log n/n$ is a threshold [49][10].

Now recall the example $\mathcal{F}_{HAM}$, where we needed minimum degree at least $n/2$ to guarantee Hamiltonicity in any graph. In the random setting we definitely require connectivity and minimum degree 2, which both hold in $\mathcal{G}(n,p)$ a.a.s. for $p = \omega(\log n/n)$. Pósa [96] and Korŝunov [78] independently showed that we do not need much more and Hamiltonicity also has the threshold $\log n/n$. This corresponds to an expected number of $\omega(n \log n)$ edges[11]. Note that, again the expected number of Hamilton cycles in $\mathcal{G}(n,p)$ already gets large after $1/n$. Their result was improved by Komlós and Szemerédi [77][10] who showed that the Hamiltonicity threshold really coincides with the threshold for

[9] It follows from Markov's inequality that the probability that the number of copies of G' is at least one tends to zero. This is called a first moment argument and we similarly obtain lower bounds for other graphs as well.

[10] In fact they proved that this property has a sharp threshold at $\ln n/n$ and even more precise results are known.

[11] Equivalently one can derive that $\mathcal{G}(n,M)$ is Hamiltonian if $M \gg n \log n$.

minimum degree 2, which lead to more precise results. Even further Bollobás [26] demonstrated that this is even true for the hitting times of these two properties in the associated graph process. Furthermore, we want to remark that the results presented so far, only guarantee existence and do not give us any meaningful algorithm for finding the structures. We will get back to this issue, in particular addressing Hamilton cycles, in Section 2.3.

As discussed, among the first spanning structures considered in graphs were perfect matchings and Hamilton cycles. More recently, the thresholds for the appearance of (bounded degree) spanning trees [14, 63, 66, 69, 80] were studied as well. The current best-known bound due to Montgomery [90, 91] is $p \geq \Delta \log^5 n/n$, where a lower bound is again given by $\log n/n$.

Riordan [97] gave a general result for embedding any graph using second moment arguments, which is non-constructive. To state it precisely, consider the following density-parameter[12] $\gamma(H) := \max\{e(H')/(v(H')-2))\colon H' \subseteq H \text{ and } v(H') \geq 3\}$, which will be responsible for the upper bound on the threshold.

Theorem 2.1 (Riordan [97][13])**.** *Let H be a graph on n vertices with $\Delta = \Delta(H)$. If H has a vertex of degree at least 2 and the following condition is satisfied*

$$np^{\gamma(H)}\Delta^{-4} \to \infty,$$

then a.a.s. the random graph $\mathcal{G}(n,p)$ contains a copy of H.

The motivation for this result was to determine the threshold functions for the appearance of cubes and lattices. Even though the general statement was known for a while, only in recent years its full potential and applicability has been realised.

A generalisation of cycles is finding the k-th power of a Hamilton cycle in $\mathcal{G}(n,p)$, where $k \geq 2$. In general, the k-th power of a graph G is the graph obtained from G by connecting all vertices at distance at most k. While Theorem 2.1 already shows that the threshold for $k \geq 3$ is given by $n^{-1/k}$ (as observed in [84]), the threshold for $k = 2$ is still open, where the best known upper bound is a polylog-factor away [92] from the conjecture $n^{-1/2}$.

Apart from cycles and trees another interesting class of graphs are factors, as a natural generalisation of matchings. The *G-factor* on n vertices for a fixed graph G consists of $n/v(G)$ vertex-disjoint copies of G (assuming that $v(G) \mid n$). Finding thresholds for spanning factors of graphs and hypergraphs was an open problem for a long time (cf. intermediate results for the triangle factor [72, 79]) until a breakthrough was achieved by Johansson, Kahn, and Vu [67]. With $d_1(G) := e(G)/(v(G)-1)$ we state their result for future reference.

Theorem 2.2 (Johansson, Kahn, and Vu [67])**.** *Let G be a strictly balanced graph, i.e. $d_1(G) > d_1(G')$ for all $G' \subseteq G$. Then the threshold for the appearance of a G-factor in $\mathcal{G}(n,p)$ is $n^{-1/d_1(G)} \log^{1/e(G)} n$.*

[12] Note that this is different from the *two-density* $m_2(H)$, where the quotient of $e(H')-1$ and $v(H')-2$ is maximised over all $H' \subseteq H$.

[13] In [97] there are some additional technical conditions imposed on H, which are in fact not needed. We refer to the discussion after Theorem 2.5, which generalises this result to hypergraphs. Note that Riordan already mentions that these assumptions are probably not crucial.

In particular, this implies that the threshold for a $K_{\Delta+1}$-factor is given by

$$p_\Delta := \left(n^{-1}\ln^{1/\Delta} n\right)^{\frac{2}{\Delta+1}}.$$

For not strictly balanced graphs G, they show that the threshold is at most $O(n^{-1/m_1(G)+o(1)})$, which is optimal up to the $o(1)$-term. Gerke and McDowell [61] removed the $o(1)$-term when G is non-vertex balanced, i.e., there exists a vertex in G which is not contained in a subgraph G' of 1-density $m_1(G)$. Furthermore, Theorem 2.2 also holds for hypergraphs and this resolved the question on the threshold for perfect matchings in hypergraphs, which was a long standing open problem.[14]

The Kahn-Kalai Conjecture

Besides many others, these results support a general conjecture of Kahn and Kalai [68] on the appearance of a given structure. It states that the threshold $\hat{p}$ is always within a factor of $O(\log n)$ from p_E, the so-called *expectation threshold*, which is the smallest p_E such that the expected number of copies of any subgraph G' of G in $\mathcal{G}(n, p_E)$ is at least 1. In the results discussed above we observe two types of behaviour that are responsible for the threshold of the appearance of bounded degree spanning structures (cf. [68] for more details).

For the example of matchings, Hamilton cycles and G-Factors, where we need some extra $\log$-terms to overcome a local obstruction, which in the first two cases is the minimum degree and in the latter is that every vertex has to lie in a copy of G. In all similar cases there is some local reason for p_E not being enough and then it is plausible that also a hitting time result might be true. This says that the structure appears at the precise moment when the last local obstructions disappeared. In the case of G-factors we are thus waiting until the very last vertex lies in a copy of G.

On the other hand there are structures, where p_E also is a threshold for the containment property and we do not need extra $\log$-terms. This usually is justified by the absence of a local obstruction. Examples are higher powers of Hamilton cycles and other applications of Riordan's results. This is also highly correlated with the applicability of the second moment method[15], which fails for Hamilton cycles, but is sufficient to determine the thresholds for higher powers and to prove Theorem 2.1. We will come across this phenomenon in more examples and also in hypergraphs.

If we only require an *almost spanning embedding*, which is for any $\varepsilon > 0$ an embedding of e.g. a matching, cycle or almost G-factor on at least $(1-\varepsilon)n$ vertices, then we usually do not need the $\log$-terms and p_E is enough. The probability p_E is large enough to ensure that only a small fraction of the vertices has the obstruction. Usually, these embeddings are easier, even without the $\log$-terms. For matchings it is almost trivial, for Hamilton cycles [83] it can be shown using Depth-first search, it was shown for trees by Alon, Krivelevich, and Sudakov [14] and is an easy application of Janson's inequality (Theorem 2.18) for factors (Theorem 2.19).

[14] It was called Shamir's problem and first explicitly stated in [39].

[15] In the second moment method we use the variance of a random variable to bound the probability that it deviates much from its expectation. A standard example is Chebyshev's inequality and from this thesis the proof of Theorem 2.5, Theorems 2.16, and 2.18 and Lemma 2.17.

General bounded degree graphs[16]

Turning to a much more general class of graphs, let $\mathcal{F}(n,\Delta)$ be the family of graphs on n vertices with maximum degree at most Δ. Alon and Füredi [13] studied the question of when the random graph $\mathcal{G}(n,p)$ contains a given graph from $\mathcal{F}(n,\Delta)$, proving the bound $p \geq C(\log n/n)^{1/\Delta}$ for some absolute constant $C > 0$. This is far from optimal and mainly relies on the fact that at this probability any set of Δ vertices is expected to lie in roughly $p^{\Delta}n = \Omega(\log n)$ edges (a.a.s. by Chernoff's inequality, Theorem 2.16). The proof in [13] uses a greedy strategy together with a matching argument to finish the embedding and is constructive.

Since the clique-factor is widely believed to be the hardest graph in $\mathcal{F}(n,\Delta)$ to embed and has threshold $p_{\Delta} = (n^{-1}\ln 1/\Delta n)^{2/(\Delta+1)}$, it is natural to state the following, well-known, conjecture.

Conjecture 2.3. *Let $\Delta > 0$, $F \in \mathcal{F}(n,\Delta)$ and $p = \omega(p_{\Delta})$. Then a.a.s. $\mathcal{G}(n,p)$ contains a copy of F.*

For $\Delta = 2$, this conjecture was recently solved by Ferber, Kronenberg, and Luh [53], who in fact showed a stronger universality statement, which we will discuss later in Section 2.5. For larger Δ, Theorem 2.1 implies a probability within a factor of $n^{\Theta(1/\Delta^2)}$ from p_{Δ}. The currently best result in this direction is the following almost spanning version by Ferber, Luh, and Nguyen [54]. Note, that most of the previously mentioned results are pure existence statements and do not help very much in finding a copy.

Theorem 2.4 (Ferber, Luh, and Nguyen [54])**.** *Let $\varepsilon > 0$ and $\Delta \geq 5$. Then, for every $F \in \mathcal{F}((1-\varepsilon)n,\Delta)$ and $p = \omega(p_{\Delta})$, a.a.s. $\mathcal{G}(n,p)$ contains a copy of F.*

In [54] the authors split the graph F into a sparse part F' with $\gamma(F') < (\Delta+1)/2$ and many small graphs with higher density. The sparse part is embedded with Riordan's result [97] (Theorem 2.1) and for the dense spots their approach is based on ideas from Conlon, Ferber, Nenadov, and Škorić [36] who proved a stronger universality statement for the almost spanning case while using the edge probability $p = \omega(n^{-1/(\Delta-1)}\log^5 n)$. Therefore, Theorem 2.4 for $\Delta = 3$ was already known (up to some $\log$-terms), whereas the case for $\Delta = 4$ remains open.

In the almost spanning case again the $\log$-term in p_{Δ} is expected to be redundant [54], but this remains unproven. In Section 2.4 we will discuss our result showing that the $\log$-term in p_{Δ} is redundant, even in the spanning case, if we add $\mathcal{G}(n,p)$ to a deterministic graph with linear minimum degree.

Single spanning structures in hypergraphs[17]

When one turns to hypergraphs, apart from perfect matchings, general factors [67], and Hamilton cycles (which we will discuss in details later) not very much was known. Together with Person [94] we extended the result of Riordan [97] to the setting of r-uniform hypergraphs. Let $e_H(v) = \max\{e(F) : F \subseteq H, v(F) = v\}$ and observe that the following is an extension of the previous definition of the density introduced by Riordan

$$\gamma(H) := \max_{r+1 \leq v \leq n} \left\{ \frac{e_H(v)}{v-2} \right\}.$$

[16] Some parts of this section closely follow [32].
[17] Some parts of this section are close adaptations from [94].

Theorem 2.5. *Let $r \geq 2$ be an integer and H be an r-uniform hypergraph on n vertices with $\Delta = \Delta(H)$. If H has a vertex of degree at least 2 and the following condition is satisfied*

$$np^{\gamma(H)}\Delta^{-4} \to \infty, \tag{2.1}$$

then a.a.s. the random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$ contains a copy of H.

We remark, that for $r = 2$ this is Theorem 2.1, the result by Riordan [97, Theorem 2.1], except that some technical conditions are omitted. By examining carefully the proof in [97], one can verify that there too these technical conditions are not in fact needed. Instead, it is sufficient to only assume (2.1) and that $\Delta(H) \geq 2$. In fact, the proof for hypergraphs will follow along the lines of Riordan's original argument, but requires adaptations at various places. We provide the details of the proof in Chapter 3 and in Section 3.5 we discuss its applications to some particular spanning structures such as Hamilton cycles, hypercubes, lattices, spheres, and powers of Hamilton cycles in hypergraphs.

The only other spanning structures that were studied more recently in hypergraphs are Hamilton cycles. There are various notions of Hamilton cycles in hypergraphs: weak Hamilton cycles, Berge Hamilton cycles, ℓ-offset Hamilton cycles (for $1 \leq \ell \leq r/2$), and ℓ-overlapping Hamilton cycles (for $1 \leq \ell \leq r-1$). The most attention was attracted by *ℓ-overlapping Hamilton cycles*, where one seeks to cyclically order the vertex set such that edges are consecutive segments and neighbouring edges intersect in ℓ vertices. We say that a hypergraph is ℓ-Hamiltonian if it contains an ℓ-overlapping Hamilton cycle. An ℓ-overlapping Hamilton cycle requires that $r - \ell$ divides n and thus a ℓ-overlapping Hamilton cycle has $n/(r-\ell)$ edges. It is customary to refer to an ℓ-overlapping cycle as a *tight cycle* for $\ell = r-1$ and a *loose cycle* for $\ell = 1$.

The study of Hamilton cycles in random hypergraphs was initiated by Frieze [59] who determined the threshold for the appearance of loose 3-uniform Hamilton cycles to be $\log n/n^2$ (when $4|n$). Dudek and Frieze [43] extended the result to higher uniformities with threshold $\log n/n^{r-1}$ (when $2(r-1)|n$). The divisibility requirement was improved to the optimal one ($(r-1)|n$) by Dudek, Frieze, Loh, and Speiss [45], see also Ferber [52]. Loose Hamilton cycles closely resemble the properties of Hamilton cycles from the graph case, in the sense that the expectation threshold is not enough and we need some extra $\log$-factor to avoid isolated vertices. The hitting time results are still open.

Subsequently, Dudek and Frieze [44] determined thresholds for general ℓ-overlapping Hamilton cycles purely relying on the second moment method. Generally $\omega(n^{\ell-r})$ is the threshold for an ℓ-overlapping Hamilton cycle for $\ell \geq 2$, but for most values more precise results are known (cf. the table at the end of [44]). Note that Theorem 2.5 gives back these results for $\ell \geq 2$ in a slightly weaker form (cf. Corollary 3.8). In particular, in [44] they proved for $r \geq 4$ that e/n is the sharp threshold function for containment of a tight cycle. An easy first moment calculation shows that if $p \leq (1-\varepsilon)e/n$ then a.a.s. $\mathcal{H}^{(r)}(n,p)$ does not contain a tight Hamilton cycle. A general result of Friedgut [57] readily shows that the threshold for the appearance of an ℓ-overlapping cycle in $\mathcal{H}^{(r)}(n,p)$ is sharp. We want to remark, that all these results were nonconstructive, relying either on Theorem 2.2 by Johansson, Kahn, and Vu [67] or the second moment method.

The case of weak Hamilton cycles (any two consecutive vertices lie in a hyperedge) was studied by Poole in [95], ℓ-offset Hamilton cycles (neighbouring edges intersect in ℓ and $r-\ell$ vertices alternatingly) by Dudek and Helenius [47][18], and Berge Hamilton cycles (any two consecutive vertices lie in some chosen hyperedge and no hyperedge is chosen twice) by Clemens, Ehrenmüller, and Person in [34], the latter one being algorithmic.

2.3 Algorithms for Hamilton cycles[19]

We will now take a step back and discuss algorithmic questions, especially for the Hamilton Cycle Problem. Proving the existence of a Hamilton cycle, does not necessary help very much in finding one. The general problem of deciding whether *any* given graph contains a Hamilton cycle, is one of the 21 classical NP-complete problems due to Karp [70]. The best currently known algorithm is due to Björklund [22]: a Monte-Carlo algorithm with worst case running time $O^*(1.657^n)$[20], without false positives and false negatives occurring only with exponentially small probability. But what about *typical* instances? In other words, when the input is a random graph sampled from some specific distribution, e.g. $\mathcal{G}(n,p)$. Is there an algorithm which finds a Hamilton cycle in polynomial time with small error probabilities?

The previously mentioned results for the appearance of Hamilton Cycles [26, 77, 78, 96] do not allow one to actually find any Hamilton cycle in polynomial time. The first polynomial time randomised algorithms for finding Hamilton cycles in $\mathcal{G}(n,p)$ are due to Angluin and Valiant [16] and Shamir [103]. Subsequently, Bollobás, Fenner, and Frieze [28] developed a deterministic algorithm, whose success probability (for input sampled from $\mathcal{G}(n,p)$) matches the probability of $\mathcal{G}(n,p)$ being Hamiltonian in the limit as $n \to \infty$. Thus, the problem is quite well understood in the graph case.

But what about hypergraphs? At the end of [44], Dudek and Frieze posed the question of finding algorithmically various ℓ-overlapping Hamilton cycles in $\mathcal{H}^{(r)}(n,p)$ at the respective thresholds. Together with Allen, Koch, and Person [6] we study tight Hamilton cycles and provide a first deterministic polynomial time algorithm, which works for p only slightly above the threshold.

Theorem 2.6. *For each integer $r \geq 3$ there exists $C > 0$ and a deterministic polynomial time algorithm with runtime $O(n^r)$ which for any $p \geq Cn^{-1}\log^3 n$ a.a.s. finds a tight Hamilton cycle in the random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$.*

The probability is only a polylog-factor away from the best known bounds, which are $p \geq (e + o(1))/n$ for $r \geq 4$ and $p = \omega(1/n)$ for $r = 3$. Prior to our work there were two algorithms known that dealt with finding tight cycles. The first algorithmic proof was given by Allen, Böttcher, Kohayakawa, and Person [4], who presented a randomised polynomial time algorithm which could find tight cycles a.a.s. at the edge probability $p \geq n^{-1+\varepsilon}$ for any fixed $\varepsilon \in (0, 1/6r)$ and running time n^{20/ε^2}. The second result is a randomised quasipolynomial time algorithm of Nenadov and Škorić [92], which works for $p \geq Cn^{-1}\log^8 n$.

Our result builds on the adaptation of the absorbing technique of Rödl, Ruciński and Szemerédi [100] to sparse random (hyper-)graphs. This technique was actually used earlier by Krivelevich in [79] in the context of random graphs. However, the first results that provided essentially optimal thresholds

[18] They obtain a sharp threshold and observe that the coarse threshold follows already from our Theorem 2.5.
[19] Large parts of this section are an almost verbatim copy from [6].
[20] Writing O^* means we ignore polylogarithmic factors.

(for other problems) are proved in [4] mentioned above in the context of random hypergraphs and independently by Kühn and Osthus in [84], who studied the threshold for the appearance of powers of Hamilton cycles in random graphs. The probability of $p \geq C(\log n)^3 n^{-1}$ results in the use of so-called reservoir structures of polylogarithmic size, as first used by Montgomery to find spanning trees in random graphs [91], and later in [92]. Our improvements result in the combination of the two algorithmic approaches [4, 92] and in the analysis of a simpler algorithm that we provide.

The general idea for the algorithm is as follows. In a hypergraph $H = (V, E)$ our algorithm finds a long tight path with the property that from some Reservoir set $R \subseteq V$ of polylogarithmic size every subset $R' \subseteq R$ can be absorbed into the path. This path is then extended until it covers $V \setminus R$ and possibly some vertices of R. Using the leftover vertices of R we can close the path to a cycle and then absorb the remaining vertices from R' from R into the cycle, because of the property described above.

We give the details of the algorithm in Chapter 4. There we first provide an informal overview of our algorithm and then two key lemmas and the proof of Theorem 2.6 which rests on these lemmas. In the subsequent sections we prove these main lemmas: the Connecting Lemma and the Reservoir Lemma.

2.4 Randomly perturbed graphs[21]

We now leave the algorithmic perspective and change the setup. In most of the examples discussed so far the appearance of spanning structures in random graphs are influenced by local properties such as the minimum degree required. On the other hand, in extremal graph theory minimum degree conditions are studied that force given spanning structures in any deterministic graph. Typically, as in Dirac's Theorem, the required minimum degree is rather large for trivial reasons such as connectedness.

Randomly peturbed graphs combine both of these worlds. The randomly perturbed graphs we consider are obtained as the union of a deterministic graph satisfying a certain minimum degree condition, and a random graph. The question then is how small one can choose the minimum degree of the deterministic graph and the edge probability of the random graph while still compelling the given spanning subgraph. It turns out that, typically, both quantities can be chosen smaller than in the corresponding pure setting, because the minimum degree condition of the deterministic graph helps to guarantee stronger local properties, while the random graph warrants stronger connectedness properties.

The following model of randomly perturbed graphs was first suggested by Bohman, Frieze and Martin [24]. For $\alpha \in (0, 1)$ and an integer n, we first let G_α be any n-vertex graph with minimum degree at least αn. We then reveal more edges among the vertices of this graph independently at random with probability p. The resulting graph $G_\alpha \cup G(n, p)$ is a *randomly perturbed graph* and we shall be interested in its properties. In particular, research has focused on comparing thresholds in $G_\alpha \cup G(n, p)$ to thresholds in $G(n, p)$.

Again, we concentrate on spanning subgraphs. Note that the existence of such subgraphs in $G_\alpha \cup G(n, p)$ is a monotone property (in $G(n, p)$), and thus has a threshold. For $\alpha \in (0, 1/2)$, Bohman, Frieze, and Martin [24] showed that, if $p = \omega(1/n)$, then, for any G_α, there is a Hamilton cycle in

[21] This section is a close adaption of the introduction from [32].

$G_\alpha \cup G(n,p)$ a.a.s. They also proved that this is optimal: for $p = o(1/n)$ there are graphs G_α such that $G_\alpha \cup G(n,p)$ is not Hamiltonian a.a.s. Comparing this threshold to the threshold for Hamiltonicity in $G(n,p)$ we note an extra factor of $\log n$ in the latter. This $\log n$ term is necessary to guarantee minimum degree at least 2 in $G(n,p)$ – otherwise clearly no Hamilton cycle exists. But in the model $G_\alpha \cup G(n,p)$ already G_α gives this minimum degree. The result also shows that for smaller α, a large linear number of random edges can compensate for the loss in minimum degree.

Krivelevich, Kwan, and Sudakov [82] studied the corresponding problem for the containment of spanning bounded degree trees in $G_\alpha \cup \mathcal{G}(n,p)$. For $p = \omega(1/n)$ it is already possible to find any almost spanning bounded degree tree in $\mathcal{G}(n,p)$ [14]. The addition of G_α then ensures there are no isolated vertices, and Krivelevich, Kwan, and Sudakov [82] showed that this indeed allows every vertex to be incorporated into the embedding. They thus proved that for $\alpha > 0$ and $p = \omega(1/n)$ every spanning bounded degree tree is contained in $G_\alpha \cup G(n,p)$ (cf. Theorem 5.7).

Quite recently Balogh, Treglown, and Wagner [19] managed to show that also for factors p_E is enough. To be precise they showed that the threshold for any G-factor in $\mathcal{G}(n,p) \cup G_\alpha$ is $n^{-1/m_1(G)}$ (cf. Theorem 5.8). For strictly balanced G this saves the extra $\log$-factors, but for non-vertex balanced G there is no benefit from G_α. For the proof they use Szemerédi's Regularity lemma [105] and a result of Komlós [75], which gives optimal bounds on α such that any G_α contains an almost spanning factor. The addition of random edges allows them to make the result of Komlós spanning with a much smaller minimum degree.

Apart from this example the general strategy for embedding spanning graphs into $G_\alpha \cup \mathcal{G}(n,p)$ is to first find an almost spanning embedding in $\mathcal{G}(n,p)$ and then, in a second step, complete the embedding using G_α and more random edges from $\mathcal{G}(n,p)$. For this it is very convenient to split the random graph into several independent rounds.

Other monotone properties considered in this model include containing a fixed sized clique, having small diameter, k-connectivity [23], and non-2-colorability [104]. With Böttcher, Montgomery, and Person [32], we analyse the model $G_\alpha \cup \mathcal{G}(n,p)$ with respect to the containment of spanning bounded degree graphs and obtain the following result.

Theorem 2.7. *Let $\alpha > 0$ be a constant, $\Delta \geq 5$ an integer, and G_α a graph on n vertices with minimum degree at least αn. Then, for every $F \in \mathcal{F}(n,\Delta)$ and $p = \omega\left(n^{-2/(\Delta+1)}\right)$, a.a.s. $G_\alpha \cup \mathcal{G}(n,p)$ contains a copy of F.*

Observe that the bound on p is best possible. Indeed, in the case where F is a $K_{\Delta+1}$-factor on n vertices and $G_\alpha = K_{\alpha n,(1-\alpha)n}$, we need to find an almost spanning $K_{\Delta+1}$-factor of size $(1-\alpha(\Delta+1))n$ in $\mathcal{G}(n,p)$. Finally, note that the edge probability p_Δ used in Theorem 2.7 is lower by a $\log$-term in comparison to the anticipated threshold for the graph F to appear in $\mathcal{G}(n,p)$ (see Conjecture 2.3).

We provide a new method, which combines the edges of G_α with those of $\mathcal{G}(n,p)$ to obtain a spanning embedding. Our proof also uses the approach by Ferber, Luh, and Nguyen [54] explained above to decompose the graph and find an embedding of almost the whole graph by only using edges of $\mathcal{G}(n,p)$. The crucial observation is, that this embedding maps uniformly at random onto the vertex set G_α, since $\mathcal{G}(n,p)$ is purely random. This enables us to find for every remaining vertex v a large so-called *reservoir set* of vertices $B(v)$ which can replace v without harming the current embedding. For the rest we again follow a similar embedding approach as before using Janson's inequality (Theorem 2.18) and a theorem of Aharoni and Haxell (Lemma 2.20), where we now want to embed into the sets $B(v)$. Owing to the large choice of vertices in all $B(v)$ and some extra edges taken from G_α, we manage to finish the embedding at the given probability. The details of the proof are given in Chapter 5.

The methods we introduce, in particular our novel techniques for creating a reservoir set, give rise to simpler proofs for the results on spanning trees [82] and factors [19]. We give the short proofs, after the proof of Theorem 2.7 in Chapter 5. A very intriguing question is if the behaviour we observed in this model in comparison to $\mathcal{G}(n,p)$ is always true. That is, can we always save the extra $\log$ terms when the threshold differs from the expectation threshold?

The model can be easily generalised to r-uniform hypergraphs, where we have to decide which kind of minimum degree condition we want to require from G_α. Krivelevich, Kwan, and Sudakov [81] also considered matchings and loose cycles in hypergraphs. Their generalised minimum degree condition in G_α is that all $(r-1)$-sets are contained in at least αn edges. Here, revealing additional random edges with probability $\omega(n^{-r+1})$ is sufficient to almost surely create both matchings and loose cycles in $G_\alpha \cup \mathcal{G}(n,p)$. Note that, comparing this to the threshold for matchings and loose cycles in random hypergraphs (which are both $\log n/n^{r-1}$ [43, 59, 67]), we again have a difference of $\log n$.

We remark that there is a lot of research on the corresponding Dirac-type questions for perfect matchings and ℓ-overlapping cycles in hypergraphs with different minimum degree conditions. For example for loose Hamilton cycles minimum vertex degree $\frac{n}{2(k+1)} + o(n)$ is sufficient as shown by Keevash, Kühn, Mycroft, and Osthus [71]. For further details we refer to the survey article by Rödl and Ruciński [99].

Interestingly McDowell and Mycroft [89] managed to show that for ℓ-overlapping cycles ($\ell \geq 2$) it is possible to save a polynomial factor n^ε in comparison to the threshold in $\mathcal{H}^{(r)}(n,p)$ under the assumption of high ℓ and $r-\ell$ degree in G_α. This result was extended by Bedenknecht, Han, Kohayakawa, and Mota [20] to powers of tight Hamilton cycles, where they required even higher minimum degree conditions. This gives rise to the questions whether this is also possible in the graph case and if there is some structure where we can save more than some log-factors. We will discuss this in more details in the concluding remarks in Chapter 8.

2.5 Universality[22]

So far all mentioned results deal with the containment of one structure. What happens if we want to find more graphs simultaneously? For a family of graphs $\mathcal{F}$, we call a graph $\mathcal{F}$-universal if it contains all $F \in \mathcal{F}$ as a subgraph. Note that most of the general results mentioned above do not imply the analogous universality statement, because there are too many graphs and we can not apply a union bound. We are mainly interested in bounds on the threshold for universality for the family of bounded degree graphs and hypergraphs in $\mathcal{G}(n,p)$ and $\mathcal{H}^{(r)}(n,p)$ respectively.

Universality in random graphs

Universality properties were first studied by Alon, Capalbo, Kohayakawa, Rödl, Ruciński, and Szemerédi [11]. They showed that for any $\varepsilon > 0$ there exists a C, such that for $p \geq C(\log n/n)^{1/\Delta}$ the

[22] Large parts of the first two subsections are taken verbatim from [94] and the last two from [64].

$\mathcal{G}(n,p)$ is $\mathcal{F}((1-\varepsilon)n,\Delta)$-universal a.a.s. that is, it contains with high probability any graph with degree bounded by Δ on $(1-\varepsilon)n$ vertices as a subgraph. Then, Dellamonica, Kohayakawa, Rödl, and Ruciński [40] showed that for $\Delta \geq 2$ the random graph $\mathcal{G}(n,p)$ is $\mathcal{F}(n,\Delta)$-universal a.a.s. provided that $p \geq C\left(\log^2 n/n\right)^{1/(2\Delta)}$, where $C > 0$ is some absolute constant. The same authors subsequently improved in [41] the bound on p to $C\left(\log n/n\right)^{1/\Delta}$ for the $\mathcal{F}(n,\Delta)$-universality of $\mathcal{G}(n,p)$ for any given $\Delta \geq 3$. Later, Kim and Lee [73] dealt with the missing case $\Delta = 2$. As mentioned before at this probability every set of Δ vertices has many common neighbours, which is very helpful for embedding bounded degree graphs. Thus it forms a natural barrier for the methods used up to this point.

Bringing the density of a graph into the statement, Ferber, Nenadov, and Peter [56] showed that for universality of all graphs with maximum degree Δ and maximum density m the probability $p = \omega(\Delta^{12} n^{-1/(4m)} \log^3 n)$ suffices. By embedding some cycles separately and using the previous result, Conlon, Ferber, Nenadov, and Škorić [36] were able to show, that for $p = \omega(n^{-1/(\Delta-1)} \log^5 n)$ the random graph $\mathcal{G}(n,p)$ is $\mathcal{F}((1-\varepsilon)n,\Delta)$-universal. Moeover, Ferber and Nenadov [55] proved very recently that $\mathcal{G}(n,p)$ is $\mathcal{F}(n,\Delta)$-universal provided that $p \geq C(n^{-1}\log^3 n)^{1/(\Delta-1/2)}$ using an embedding technique by Conlon and Nenadov [38] together with the ideas from [36] and absorbers.

The lower bound again comes from the $K_{\Delta+1}$-factor and thus Conjecture 2.3 can be generalised in the following way.

Conjecture 2.8. *Let $\Delta > 0$ and $p = \omega(p_\Delta)$. Then a.a.s. $\mathcal{G}(n,p)$ is $\mathcal{F}(n,\Delta)$-universal.*

For $\Delta = 2$ this conjecture was solved in [53] by Ferber, Kronenberg, and Luh. For $\Delta = 3$ the almost spanning version from [36] is optimal up to the $\log$-terms. For larger Δ the gap between p_Δ and the currently best known bound $(n^{-1}\log^3 n)^{1/(\Delta-1/2)}$ obtained by Ferber and Nenadov [55] is polynomial in n. Again in the almost spanning version of Conjecture 2.8 the $\log$-terms should be redundant.

For the embedding of a family of graphs the core of the approach [11, 36, 41, 53, 56, 73] is to find a deterministic structure with nice expansion properties. These pseudo-random structures appear in random graphs and admit an embedding without any further randomness, because they behave in a random-like way. As for single spanning structures it proved to be helpful to remove specific structures from the graphs which are embedded in the end. For instance to complete the embedding in [36], they used Janson's inequality [65] to find previously removed cycles and a matching trick [2] to allocate them correctly.

Universality in random hypergraphs

As for single spanning structures in the hypergraph case, much less is known. For a family $\mathcal{F}$ of r-uniform hypergraphs we say that an r-uniform hypergraph H is $\mathcal{F}$-universal if every hypergraph $F \in \mathcal{F}$ occurs as a copy in H. Let $\mathcal{F}^{(r)}(n,\Delta)$ denote the family of all r-uniform hypergraphs F of maximum vertex degree at most Δ on n vertices.

Together with Person, we were able to generalise the result of Dellamonica, Kohayakawa, Rödl, and Ruciński [41] to hypergraphs. We prove universality of $\mathcal{H}^{(r)}(n,p)$ for the family $\mathcal{F}^{(r)}(n,\Delta)$, where we show that a natural bound on $p \geq C(\log n/n)^{1/\Delta}$ suffices.

Theorem 2.9. *For every $r \geq 2$ and any integer $\Delta \geq 1$, there exists a constant $C > 0$, such that for $p \geq C(\log n/n)^{1/\Delta}$ the random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$ is $F^{(r)}(n,\Delta)$-universal a.a.s.*

In the proof of Theorem 2.9 (see Chapter 6) we employ the strategy of Dellamonica, Kohayakawa, Rödl, and Ruciński [41], but a shortcut will be obtained by using similar notions of good properties that were used by Kim and Lee [73] and by Ferber, Nenadov, and Peter [56].

The bound on $p = \Omega\left((\log n/n)^{1/\Delta}\right)$ in Theorem 2.9 is presumably not optimal and similar improvements as in [36, 55] would be interesting. As for the lower bound, if $\binom{t-1}{r-1} \leq \Delta$, the $K_t^{(r)}$-factor is a member of $\mathcal{F}^{(r)}(n, \Delta)$. Since by the hypergraph version of Theorem 2.2 the threshold probability for the appearance of such factor is $\Theta\left((\log n)^{1/\binom{t}{r}} n^{-(t-1)/\binom{t}{r}}\right)$ and in view of $n^{-(t-1)/\binom{t}{r}} < n^{-1/\binom{t-1}{r-1}}$, the best lower bound (if $\Delta = \binom{t-1}{r-1}$) we are aware of is $p = \Omega\left((\log n)^{1/\binom{t}{r}} n^{-(t-1)/\binom{t}{r}}\right)$.

Explicit constructions of universal graphs

A closely related natural problem in the non-random setting is the *existence* and *explicit construction* of graphs that are universal for some family of graphs. For an excellent survey on this problem see Alon [7] and the references therein. The first nearly optimal universal graphs for $\mathcal{F}(n, \Delta)$ with $O(n)$ vertices and $O(n^{2-2/\Delta} \log^{1+8/\Delta} n)$ edges for $\Delta \geq 3$ were given by Alon, Capalbo, Kohayakawa, Rödl, Ruciński, and Szemerédi [12]. It was also noted by the same authors in [11] with simple calculations that any such universal graph has to contain $\Omega(n^{2-2/\Delta})$ edges. As mentioned in [9] the square of a Hamilton cycle is $\mathcal{F}(n, 2)$-universal and thus $2n$ edges are enough in this case. In two subsequent papers, Alon and Capalbo [9, 10] further pursued this line of research and improved upon the result of [12] obtaining $\mathcal{F}(n, \Delta)$-universal graphs with the optimal number of edges and only $O(n)$[23] vertices and also providing $\mathcal{F}(n, \Delta)$-universal graphs on n vertices with *almost optimal* number of edges.

Theorem 2.10 (Alon and Capalbo [9, 10])**.** *For any $\Delta \geq 2$ there exist explicitly constructible $\mathcal{F}(n, \Delta)$-universal graphs on $O(n)$ vertices with $O(n^{2-2/\Delta})$ edges and on n vertices with $O(n^{2-2/\Delta} \log^{4/\Delta} n)$ edges.*

Started together with Person [94] and then further developed with Hetterich and Person [64] we worked on extending these constructions to hypergraphs. It follows from the asymptotic number of Δ-regular r-graphs on n vertices (see e.g. [46]) that *any* $\mathcal{F}^{(r)}(n, \Delta)$-universal hypergraph must possess $\Omega(n^{r-r/\Delta})$[24] edges. In [94] we derive explicit constructions of $\mathcal{F}^{(r)}(n, \Delta)$-universal hypergraphs on $O(n)$ vertices with $O(n^{r-2/\Delta})$ edges and on n vertices with $O(n^{r-2/\Delta} \log^{4/\Delta} n)$ edges from Theorem 2.10. Furthermore, we obtain the existence of even sparser universal hypergraphs from the results on universality of random graphs [36, 41]. For example, it is shown that there exist $\mathcal{F}^{(r)}(n, \Delta)$-universal hypergraphs with n vertices and $\Theta\left(n^{r-\frac{r}{2\Delta}} (\log n)^{\frac{r}{2\Delta}}\right)$ edges, which shows that the best known lower and upper bounds are at most the multiplicative factor $n^{\frac{r}{2\Delta}} \cdot \mathrm{polylog}\, n$ apart.

Subsequently, with Hetterich and Person in [64], we worked on further improving this approach. We proved the following statements that allow us to construct r-uniform universal hypergraphs from universal hypergraphs of smaller uniformity. We use universal graphs from Theorem 2.10 with carefully chosen parameters to provide the best known $\mathcal{F}^{(r)}(n, \Delta)$-universal hypergraphs.

Theorem 2.11. *Let $r, r' \geq 2$ and $\Delta \geq 2$ be integers. If $r' \mid r$ and H' is an $\mathcal{F}^{(r')}(n, \Delta)$-universal hypergraph, then there exists an $\mathcal{F}^{(r)}(n, \Delta)$-universal hypergraph H on the same vertex set as H' and $e(H) \leq e(H')^{r/r'}$.*

[23] In all cases $O(n)$ vertices can be reduced to $(1 + \varepsilon)n$ vertices for any $\varepsilon > 0$ by using a concentrator as done in [12]. We give details at the end of Section 7.2.

[24] We give the short calculations at the beginning of Chapter 6.

This implies that, whenever $r' \mid r$ and almost optimal $\mathcal{F}^{(r')}(n,\Delta)$-universal hypergraphs are known as is, for example, the case when r is even due to Theorem 2.10, this leads to constructions of almost optimal $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraphs.

Corollary 2.12. *Let $r, r' \geq 2$ and $\Delta \geq 2$ be integers. If $r' \mid r$ and there exists an $\mathcal{F}^{(r')}(n,\Delta)$-universal hypergraph H' with $O(n^{r'-r'/\Delta})$ edges, then there exists an $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraph on the same vertex set $V(H')$ with $O(n^{r-r/\Delta})$ edges. In particular, if r is even then there exist explicitly constructible $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraphs on $O(n)$ vertices with $O(n^{r-r/\Delta})$ edges and on n vertices with $O(n^{r-r/\Delta}\log^{2r/\Delta}(n))$ edges.*

In the case of odd r we cannot apply Theorem 2.11 and we prove the following.

Theorem 2.13. *Let $r \geq 3$ and $\Delta \geq 2$ be integers. Then with $\Delta' = \lceil (r+1)\Delta/r \rceil$ there exist explicitly constructible $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraphs on $O(n)$ vertices with $O(n^{r-(r+1)/\Delta'})$ edges and on n vertices with $O(n^{r-(r+1)/\Delta'}\log^{2(r+1)/\Delta'}(n))$ edges. In particular, if $r \mid \Delta$ this leads to almost optimal $O(n^{r-r/\Delta}\mathrm{polylog}(n))$ edges.*

By estimating Δ' we see that in any case the lower and upper bounds on the edge densities of optimal universal hypergraphs differ by at most a factor of n^{r/Δ^2}. By generalising their construction and a graph decomposition result of Alon and Capalbo [9, Lemma 3.3] (cf. Lemmas 7.3, 7.5, and 7.7) we obtain yet another case when constructed universal hypergraphs match the lower bound.

Theorem 2.14. *Let r be an integer. Then there exists an explicitly constructible $\mathcal{F}^{(r)}(n,2)$-universal hypergraph on $O(n)$ vertices and $O(n^{r/2})$ edges.*

Thus, for the cases where $2 \mid r$, $r \mid \Delta$, or $\Delta = 2$ we present constructions of universal hypergraphs, that are as optimal as the best-known constructions in the graph case. There are several open cases, with the smallest ones being $r = 3$ and $\Delta = 4$ or 5.

The proofs of these results will be given in Chapter 7. In the first section we introduce a very useful concept of hitting graphs, which we use in Section 7.2 to prove Theorem 2.11, Corollary 2.12, and Theorem 2.13 and in Section 7.3 along with a graph decomposition result from [9] to prove Theorem 2.14.

$\mathcal{E}^{(r)}(m)$-universal hypergraphs

Another family of graphs that received attention is the family $\mathcal{E}^{(r)}(m)$ of r-graphs with at most m edges and without isolated vertices. Babai, Chung, Erdős, Graham, and Spencer [17] proved that any $\mathcal{E}^{(2)}(m)$-universal graph must contain $\Omega(m^2/\log^2 m)$ edges and there exists a graph on $O(m^2 \log\log m/\log m)$ edges. Alon and Asodi [8] closed this gap by proving the existence of an $\mathcal{E}^{(2)}(m)$-universal graph with $O(m^2/\log^2 m)$ edges.

We briefly study $\mathcal{E}^{(r)}(m)$-universal hypergraphs. It can be shown for fixed $r \geq 3$ that any $\mathcal{E}^{(r)}(m)$-universal hypergraph must contain at least $\Omega(m^r/\log^r m)$ edges. This can be seen by a simple counting argument as in [17] or by counting $(r\log m)$-regular r-graphs on $m/\log m$ vertices as was done in the graph case in [8]. We prove that the optimal existence result of Alon and Asodi gives rise to optimal $\mathcal{E}^{(r)}(m)$-universal hypergraphs.

Theorem 2.15. *There exist $\mathcal{E}^{(r)}(m)$-universal hypergraphs with $O(m^r/\log^r m)$ edges.*

In Section 7.4 we discuss $\mathcal{E}^{(r)}(m)$-universal hypergraphs and prove Theorem 2.15.

2.6 Tools

In this section we briefly collect some of the tools mentioned earlier, which we need for the proofs. This is mostly intended for referencing and thus we advice the reader to consult this only when necessary. We will make use of the following version of Chernoff's inequality, see e.g. [65, Theorem 2.8].

Theorem 2.16 (Chernoff's inequality). *Let X be the sum of independent binomial random variables, then for any $\gamma \in (0,1)$*

$$\mathbb{P}\big[|X - \mathbb{E}[X]| \leq \gamma\,\mathbb{E}[X]\big] \leq 2\exp\left(-\frac{\gamma^2\mathbb{E}[X]}{3}\right).$$

Furthermore, we use the following submartingale-type inequality to deal with a sum of Bernoulli random variables which are not independent. A proof can be found in [3, Lemma 2.2].

Lemma 2.17 (Sequential dependence lemma). *Let Ω be a finite probability space, and let $\mathcal{F}_0, \dots, \mathcal{F}_m$ be partitions of Ω, with $\mathcal{F}_{i-1}$ refined by $\mathcal{F}_i$ for each $i \in [m]$. For each $i \in [m]$, let Y_i be a Bernoulli random variable on Ω which is constant on each part of $\mathcal{F}_i$. Let δ be a real number, $\gamma \in (0,1)$, and $X = Y_1 + \dots + Y_m$. If $\mathbb{E}[Y_i|\mathcal{F}_{i-1}] \geq \delta$ holds for all $i \in [m]$, then*

$$\mathbb{P}\big[X \leq (1-\gamma)\delta m\big] \leq \exp\left(\frac{-\gamma^2\delta m}{3}\right).$$

For embedding small graphs, Janson's inequality proved to be a very useful tool. The following variant is adapted from [15, Chapter 8] and [65, Theorem 2.18].

Theorem 2.18 (Janson's inequality). *Let $p \in (0,1)$ and consider a family $\{H_i\}_{i\in\mathcal{I}}$ of subgraphs of the complete r-uniform hypergraph on the vertex set $[n] = \{1,\dots,n\}$. For each $i \in \mathcal{I}$, let X_i denote the indicator random variable for the event that $H_i \subseteq \mathcal{H}^{(r)}(n,p)$ and, write $H_i \sim H_j$ for each ordered pair $(i,j) \in \mathcal{I}\times\mathcal{I}$ with $i \neq j$ if $E(H_i)\cap E(H_j) \neq \emptyset$. Then, for $X = \sum_{i\in\mathcal{I}} X_i$, $\mathbb{E}[X] = \sum_{i\in\mathcal{I}} p^{e(H_i)}$,*

$$\delta = \sum_{H_i\sim H_j} \mathbb{E}\big[X_iX_j\big] = \sum_{H_i\sim H_j} p^{e(H_i)+e(H_j)-e(H_i\cap H_j)}$$

we have

$$\mathbb{P}\big[X = 0\big] \leq \exp\left(-\frac{\mathbb{E}[X]^2}{\mathbb{E}[X]+\delta}\right).$$

As explained earlier, finding almost spanning embeddings is much easier. In Section 2.5 we will use the following theorem that guarantees an almost spanning factor.

Theorem 2.19. *For every r-uniform hypergraph G and every $\varepsilon > 0$ there is a $C > 0$ such that with $p \geq Cn^{-1/m_1(G)}$, the random hypergraph $\mathcal{H}^{(r)}(n,p)$ contains a.a.s. an almost G-factor on at least $(1-\varepsilon)n$ vertices.*

This is a generalisation from [65, Theorem 4.9] to hypergraphs. For completeness and to give a demonstration of an easy application of Theorem 2.18, we give the short proof here.

Proof. Let V_0 be a set of at least εn vertices and $t = v(G)$. We label the vertices of G with $v_1, \dots, v_t$ and split V_0 into t sets $V_1, \dots, V_t$ of equal size. Let $\{H_i\}_{i \in \mathcal{I}}$ be the set of copies of G with $v_i \in V_i$ and X be the number of these copies appearing in $\mathcal{H}^{(r)}(n,p)$. Then

$$\mu := \mathbb{E}[X] = \left(\frac{\varepsilon n}{t}\right)^t p^{e(G)} \geq \left(\frac{\varepsilon n}{t}\right)^t C n^{-e(G)/m_1(G)} \geq \left(\frac{\varepsilon}{t}\right)^t C n.$$

On the other hand we have that

$$\begin{aligned}
\delta &= \sum_{H_i \sim H_j} p^{e(H_i)+e(H_j)-e(H_i \cap H_j)} = \sum_{J \subset G} \sum_{H_i \cap H_j \cong J} p^{2e(G)-e(J)} \\
&\leq \sum_{J \subset G} \sum_{H_i \cap H_j \cong J} p^{2e(G)-m_1(G)\cdot(v(J)-1)} \\
&\leq \sum_{j=r}^{t-1} \binom{t}{j} \left(\frac{\varepsilon n}{t}\right)^{2t-j} p^{2e(G)-m_1(G)\cdot(j-1)} \\
&= \mu^2 \sum_{j=r}^{t-1} \binom{t}{j} \left(\frac{\varepsilon n}{t}\right)^{-j} p^{-m_1(G)\cdot(j-1)} \\
&\leq \mu^2 \sum_{j=r}^{t-1} \binom{t}{j} \left(\frac{\varepsilon n}{t}\right)^{-j} \frac{1}{C} n^{j-1} \leq \mu^2 \left(\frac{t^2}{\varepsilon}\right)^t \frac{1}{Cn},
\end{aligned}$$

where we used that $e(J) \leq m_1(G)(v(J)-1)$ for all $J \subseteq G$. Using Theorem 2.18 we then get

$$\mathbb{P}\big[X = 0\big] \leq \exp\left(-\frac{\mu^2}{8(\mu+\delta)}\right) \leq e^{-n}$$

for large enough C.

Now assume that the statement of the theorem is false. Then there would be a set V_0 inside of $\mathcal{H}^{(r)}(n,p)$ of size at least εn which does not contain a copy of G. There are at most 2^n choices for V_0 and thus a union bound reveals that this happens with probability at most $2^n \cdot e^{-n} = o(1)$. □

A classical theorem of Hall gives an easy criterion for matchings in bipartite graphs. A bipartite graph $G = (A \cup B, E)$, which satisfies the condition that $|N(S)| \geq |S|$ for all $S \subseteq A$, contains a matching saturating all of A. This criterion is very useful in many aspects. We will use the following generalisation to hypergraphs by Aharoni and Haxell [2] in an auxiliary graph to finish our embedding in Chapter 5. This Hall-type theorem for hypergraphs was first applied in this regard in [36].

Theorem 2.20 (Aharoni and Haxell [2])**.** *Let $\{L_1, \dots, L_t\}$ be a family of s-uniform hypergraphs on the same vertex set. If, for every $\mathcal{I} \subseteq [t]$, the hypergraph $\bigcup_{i \in \mathcal{I}} L_i$ contains a matching of size greater than $s(|\mathcal{I}|-1)$, then there exists a function $g : [t] \to \bigcup_{i=1}^t E(L_i)$ such that $g(i) \in E(L_i)$ and $g(i) \cap g(j) = \emptyset$ for $i \neq j$.*

Apart from the tools collected here, which we will use repeatedly, there will be more in the next chapters, tailored to fit the respective proofs.

Chapter 3

Riordan's theorem for hypergraphs

In this chapter we give the proof[25] obtained together with Person [94] for Theorem 2.5. We start with an outline of the proof, repeat some arguments that extend verbatim from Riordan [97] and then give the detailed second moment calculation for hypergraphs. In the last section we give some applications of this theorem as mentioned in the introduction. Besides Hamilton cycles and powers thereof, we give two general corollaries and extend the results of Riordan [97] on cubes and lattices to hypergraphs.

3.1 Proof outline

The overall proof strategy of Theorem 2.5 is the same as Riordan's in [97], which is an elegant second moment argument. In fact, a large part of the proof proceeds along the same lines and we are thus going to use the same notation, in particular providing the references at several places to [97] for comparison. With the exception of the first steps which we summarise in Lemma 3.2, since these can be performed verbatim for hypergraphs as well, we will give full details so that the reader will be able to follow the argument without looking up in [97]. We try to be brief anyway.

The actual proof deals instead of $\mathcal{H}^{(r)}(n,p)$ with the related model $\mathcal{H}^{(r)}(n,p\binom{n}{r})$, which is the probability space of all labelled r-uniform hypergraphs with the vertex set $[n]$ and exactly $p\binom{n}{r}$ edges with a uniform measure. Thus, for $r=2$ this is the standard model $\mathcal{G}(n,M)$. As we are dealing with monotone properties, a corresponding statement in the model $\mathcal{H}^{(r)}(n,p)$ can be obtained by conditioning on the number of edges in $\mathcal{H}^{(r)}(n,p)$. We omit the standard argument and refer to Łuczak [86].

One considers the random variable X which counts copies of H in $\mathcal{H}^{(r)}(n,p\binom{n}{r})$ and analyses the quantity $f := \mathbb{E}[X^2]/\mathbb{E}[X]^2$. It is enough to show that $f = 1+o(1)$, since then one infers by Chebyshev's inequality:

$$\mathbb{P}[X=0] \leq \mathbb{P}[|X-\mathbb{E}X| \geq \mathbb{E}X] \leq \frac{\operatorname{Var}(X)}{\mathbb{E}[X]^2} = f-1 = o(1). \tag{3.1}$$

All this is, of course, reminiscent of the second moment argument of Erdős and Rényi [49] for the appearance of H when H is small. However, the details of estimating f and thus the variance of X are more involved as H can now occupy the whole vertex set of the random hypergraph.

Before we get to the actual proof, let us briefly state the steps that are geared towards the estimation of f as done in [97], since this is the path we are going to pursue as well. At this point we aim to only give a flavour of this technical proof and thus postpone the definitions of S_H, S'_H, T'_H, T''_H, and of good hypergraphs to a later point:

[25] The proof given in this chapter is a close adaption of [94].

1. pondering f, Lemma 3.2 gives us that $f \leq (1+o(1))e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}}S_H$, where S_H is a sum that depends on all subhypergraphs of H, which will be introduced below;

2. then it is shown that $S_H \leq (1+o(1))e^{\frac{1-p}{p}\alpha^2\binom{n}{r}}S'_H$, where S'_H runs only over certain good subhypergraphs of H – this step requires most adaptation and we provide full details in Lemma 3.3 below;

3. one can further simplify S'_H and bound it above by another quantity T'_H – this is done in Lemma 3.4;

4. in the penultimate step, cf. equation (3.9), T'_H is bounded by $e^{T''_H}$, where T''_H is the sum over all good connected hypergraphs;

5. finally, using Lemma 3.5, it is shown that $T''_H = o(1)$ and combining the estimates, the desired bound on f follows:

$$f \leq \big(1+o(1)\big)e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}}S_H \leq \big(1+o(1)\big)e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}}e^{\frac{1-p}{p}\alpha^2\binom{n}{r}}S'_H \leq \\ \big(1+o(1)\big)T'_H \leq \big(1+o(1)\big)e^{T''_H} \leq \big(1+o(1)\big)e^{o(1)} = 1+o(1).$$

The somewhat mysterious appearance of the factor $e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}}$ in the first step and the appearance of $e^{\frac{1-p}{p}\alpha^2\binom{n}{r}}$ in the second step have good reasons that will be explained right before the corresponding lemmas, Lemma 3.2 and Lemma 3.3 respectively. For more intuition, we refer to the concluding remarks of [97]. After all, in many applications, the term $\frac{1-p}{p}\alpha^2\binom{n}{r}$ is $o(1)$ and, thus, the exponential term is not *seen* at all.

3.2 Technical estimates and first steps

Let us first collect some useful estimates that involve α and p for future reference.

Lemma 3.1. *Suppose that H has a vertex of degree at least 2 and condition (2.1) holds. Then, we have*

$$\begin{aligned} &n^{\frac{3-2r}{2}}p^{-1}\Delta^{4r-6} \to 0, && \Delta = o(n^{1/4}), \\ &\alpha^3\binom{n}{r}p^{-2} \to 0, && \alpha p^{-1}\Delta = o(n^{-1/2}), \\ &\alpha = o(p), && p\binom{n}{r} \to \infty, \\ &p^{-1}\Delta^2 n^{2-r} = o(n^{1/2}) \ \textit{and} && p^{-1}\alpha^2\binom{n}{r} = o(n^{1/2}). \end{aligned}$$

Proof. Since H contains a vertex of degree 2 we have $\gamma(H) \geq \frac{2}{2r-3}$. With $p \leq 1$, it follows from (2.1) that $\Delta = o(n^{1/4})$, and rearranging yields with $\gamma(H) \geq \frac{2}{2r-3}$ that $n^{\frac{3-2r}{2}}p^{-1}\Delta^{4r-6} \to 0$.

Now we notice immediately that $p = \omega\left((\Delta^4/n)^{\frac{2r-3}{2}}\right)$. Since $\alpha\binom{n}{r} \leq \Delta n/r$ it follows $\alpha \leq \Delta\binom{n-1}{r-1}^{-1}$. The combination of the two estimates yields $\alpha^3\binom{n}{r}p^{-2} \to 0$.

From $\alpha^3\binom{n}{r}p^{-2} \to 0$ and $\alpha\binom{n}{r} \geq 2$ we immediately get $\alpha = o(p)$, and from $n^{\frac{3-2r}{2}}p^{-1}\Delta^{4r-6} \to 0$ we obtain $p\binom{n}{r} \to \infty$.

To obtain the remaining estimates, we combine the lower bound on $p = \omega\left((\Delta^4/n)^{\frac{2r-3}{2}}\right)$ with $\alpha \leq \Delta\binom{n-1}{r-1}^{-1}$, thus obtaining $\alpha p^{-1}\Delta \leq p^{-1}\Delta^2\binom{n-1}{r-1}^{-1} = o(n^{-1/2})$ and $p^{-1}\alpha^2\binom{n}{r} = o(n^{1/2})$. □

The first steps towards the proof of Theorem 2.5 aim at estimating f, which was defined as $f := \mathbb{E}[X^2]/\mathbb{E}[X]^2$, where X is the random variable that counts the number of copies of H in $\mathcal{H}^{(r)}(n, p\binom{n}{r})$. This is done by writing X as the sum $\sum_i X_{H_i}$, where each X_{H_i} is the indicator random variable for a copy H_i of H in $K_n^{(r)}$, and then rewriting $\mathbb{E}[X^2]/\mathbb{E}[X]^2$ (in a non-obvious way) as a weighted sum over all possible subhypergraphs F of H involving $X_F(H)$ and $X_F(K_n^{(r)})$, which are the number of subhypergraphs of H (resp. of $K_n^{(r)}$) that are isomorphic to F. The subhypergraphs F of H should be thought of as contained in the intersection of two copies of H.

The following lemma summarises the desired first estimate on f. The appearance of the factor $e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}}$ is due to working in the model $\mathcal{H}^{(r)}(n, p\binom{n}{r})$ where the appearance of some edges in $\mathcal{H}^{(r)}(n, p\binom{n}{r})$ makes others slightly less likely to be chosen.

Lemma 3.2. *Suppose that $\alpha = o(p)$, $p\binom{n}{r} \to \infty$ and that $\alpha^3\binom{n}{r}p^{-2} \to 0$ (as n tends to infinity) hold. Then with $c = \frac{1-p}{p-2\alpha}$ we get*

$$f \leq \big(1 + o(1)\big)e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}} \sum_{F \subseteq H} c^{e(F)} \frac{X_F(H)}{X_F(K_n^{(r)})}. \tag{3.2}$$

Proof. This lemma has exactly the same proof as in [97]. In fact, the desired inequality (3.2) follows from [97, Inequality (4.2)] immediately because all statements and lemmas up to this point only involve manipulations with binomial coefficients, which do not take into account the uniformity of a graph. More precisely, the inequality (3.2) follows by repeating verbatim the same steps from [97, Section 3] to [97, Inequality (4.2) in Section 4], where the only difference is that the condition $\alpha\binom{n}{r} \to \infty$ in [97, Lemma 4.1] is not needed[26]. □

3.3 Generalisation to hypergraphs

Notice that every component in an r-uniform hypergraph has either one vertex (isolated vertex) or at least r vertices. We define the function $r(F) := n - k_1(F) - (r-1)k_r(F)$ where n is the number of vertices in F, $k_1(F)$ is the number of isolated vertices in F and $k_r(F)$ is the number of components in F, which are not isolated vertices. We have $\alpha p^{-1}\Delta = o(n^{-1/2})$ (which follows from Lemma 3.1) which allows us to estimate

$$\left(\frac{1-p}{p-2\alpha}\right)^{e(F)} = (p^{-1}-1)^{e(F)}\left(\frac{p}{p-2\alpha}\right)^{e(F)} \leq (p^{-1}-1)^{e(F)}(1+n^{-\frac{1}{2}})^{r(F)},$$

where we also used $\log\left(\frac{p}{p-2\alpha}\right) \leq \frac{3\alpha}{p}$ and $e(F) \leq \Delta r(F)$. By substituting the above estimate in (3.2) we can upper bound f via Lemma 3.2 as follows:

$$f \leq \big(1 + o(1)\big)e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}} S_H,$$

[26] In fact it is only used to show $\alpha = o(p)$ which follows from Lemma 3.1 anyway.

where, as in [97], we set

$$S_H = \sum_{F \subseteq H} \left(p^{-1} - 1\right)^{e(F)} \left(1 + n^{-\frac{1}{2}}\right)^{r(F)} \frac{X_F(H)}{X_F(K_n^{(r)})}. \tag{3.3}$$

Next, instead of going over all F in the sum above, we consider only those F whose components do not consist of an isolated edge. More precisely, we say that a hypergraph F is *good* if none of the components of F consists of a single (isolated) edge. In the following lemma below we seek to estimate the sum S_H from (3.3) by the following sum:

$$S'_H = {\sum_{F \subseteq H}}' \left(p^{-1} - 1\right)^{e(F)} 2^{r(F)} \frac{X_F(H)}{X_F(K_n^{(r)})}, \tag{3.4}$$

where $\sum'$ is the sum over good hypergraphs $F \subseteq H$.

Lemma 3.3 below has the same conclusion as [97, Lemma 4.3]. In the case of r-uniform hypergraphs ($r \geq 3$) one needs to be more careful and the estimates are somewhat different from those in [97]. Therefore we provide its full proof. As for intuition, so recall that the hypergraphs F in the sum S_H (defined in (3.3)) correspond to the possible intersections of two copies of H. Thus, two typical copies of H in $K_n^{(r)}$ are expected to intersect in $\alpha^2\binom{n}{r}$ edges which is, by Lemma 3.1, $o(n^{1/2})$, moreover such edges are likely to be disjoint and their contribution to the sum will be shown to be roughly $e^{\frac{1-p}{p}\alpha^2\binom{n}{r}}$.

Lemma 3.3. *If H is any r-uniform hypergraph with maximum degree $\Delta \geq 2$ and condition* (2.1) *holds, then*

$$S_H \leq \left(1 + o(1)\right) e^{\frac{1-p}{p}\alpha^2\binom{n}{r}} S'_H.$$

Proof. Let F be some good hypergraph from the sum $\sum'$ in S'_H. Thus, F is an r-uniform hypergraph with v isolated vertices and no isolated edges. We define $S'[F]$ to be the contribution to S'_H that comes from the isomorphism class of this good hypergraph $F \subseteq H$, i.e. $S'[F] = (p^{-1} - 1)^{e(F)} 2^{r(F)} \frac{X_F(H)^2}{X_F(K_n^{(r)})}$. We write F_t for a hypergraph obtained from (a good) F with v isolated vertices by adding $t \leq v/r$ isolated edges to it. We define $S[F]$ be the contribution to S_H of all subhypergraphs of H that are isomorphic to F_i for some i, where $0 \leq i \leq v/r$. Thus, $S[F] = \sum_{i=0}^{v/r} (p^{-1} - 1)^{e(F_i)} (1 + n^{-\frac{1}{2}})^{r(F_i)} \frac{X_{F_i}(H)^2}{X_{F_i}(K_n^{(r)})}$.

Every hypergraph from the sum in S_H can be reduced to a good F by deleting all isolated edges. Therefore, we have $S'_H = \sum S'[F]$ and $S_H = \sum S[F]$, where the sums are over all isomorphism classes of good subhypergraphs F of H. To prove the lemma it is sufficient to bound $S[F]/S'[F]$ for every good $F \subseteq H$ by $(1 + o(1)) e^{\frac{1-p}{p}\alpha^2\binom{n}{r}}$. Then, summing over all isomorphism classes of good subhypergraphs, we obtain: $S_H \leq (1 + o(1)) e^{\frac{1-p}{p}\alpha^2\binom{n}{r}} S'_H$, as desired.

Let $F \subseteq H$ be a good hypergraph with v isolated vertices, then

$$X_{F_t}(K_n^{(r)}) = X_F(K_n^{(r)}) \cdot \frac{1}{t!} \binom{v}{r} \binom{v-r}{r} \cdot \ldots \cdot \binom{v - rt + r}{r}$$

and

$$X_{F_t}(H) \leq X_F(H) \cdot \frac{1}{t!} e_H(v) e_H(v - r) \cdot \ldots \cdot e_H(v - rt + r).$$

Setting $\beta_w = e_H(w)^2/\binom{w}{r}$ we obtain

$$\frac{X_{F_t}(H)^2}{X_{F_t}(K_n^{(r)})} \le \frac{X_F(H)^2}{X_F(K_n^{(r)})} \frac{\beta_v \beta_{v-r} \dots \beta_{v-rt+r}}{t!}.$$

Since $e(F_t) = e(F) + t$ and $r(F_t) = r(F) + t$ we have

$$\frac{S[F]}{S'[F]} = \frac{\sum_{i=0}^{v/r} \left(p^{-1}-1\right)^{e(F_i)} \left(1+n^{-\frac{1}{2}}\right)^{r(F_i)} \frac{X_{F_i}(H)^2}{X_{F_i}(K_n^{(r)})}}{\left(p^{-1}-1\right)^{e(F)} 2^{r(F)} \frac{X_F(H)^2}{X_F(K_n^{(r)})}}$$

$$\le 2^{-r(F)}(1+n^{-\frac{1}{2}})^{r(F)} \sum_{t=0}^{v/r} \left(p^{-1}-1\right)^t \left(1+n^{-\frac{1}{2}}\right)^t \frac{\beta_v \beta_{v-r} \dots \beta_{v-rt+r}}{t!}. \tag{3.5}$$

Next we take a closer look at the β_w terms. Since $\Delta(H) \le \Delta$ we can bound $e_H(w) \le w\Delta/r$ and $\beta_w \le \left(\frac{w\Delta}{r}\right)^2 \binom{w}{r}^{-1} \le \frac{\Delta^2 r^{r-2}}{w^{r-2}}$. Therefore we estimate the product of all β_w as follows

$$\prod_{i=0}^{t-1} \beta_{v-ir} \le \left(\Delta^2 r^{r-2}\right)^t \left(\prod_{i=0}^{t-1} (v-ir)\right)^{-(r-2)} \le \Delta^{2t} \left(\frac{(\lfloor v/r \rfloor - t)!}{\lfloor v/r \rfloor !}\right)^{r-2}.$$

By applying induction one can show that $\frac{(s-t)!}{s!} \le \left(\frac{e}{s}\right)^t$ for all $0 \le t \le s$ and thus we obtain

$$\prod_{i=0}^{t-1} \beta_{v-ir} \le \left(\frac{e^{r-2}\Delta^2}{\lfloor v/r \rfloor^{r-2}}\right)^t.$$

Thus, we further upper bound $S[F]/S'[F]$, using (3.5), by

$$\frac{S[F]}{S'[F]} \le 2^{-r(F)} \left(1+n^{-\frac{1}{2}}\right)^{r(F)} \sum_{t=0}^{v/r} \left(p^{-1}-1\right)^t \left(1+n^{-\frac{1}{2}}\right)^t \left(\frac{e^{r-2}\Delta^2}{\lfloor v/r \rfloor^{r-2}}\right)^t \frac{1}{t!}. \tag{3.6}$$

From Lemma 3.1 it follows that $p^{-1}\Delta^2 n^{2-r} = o(\sqrt{n})$ and therefore

$$p^{-1}\Delta^2 v^{2-r} = o((n/v)^{r-2}\sqrt{n}) \tag{3.7}$$

and in the following we will distinguish four cases.

Suppose $0 \le v \le n/(100r \ln n)$. Then we use (3.7) to upper bound each term in the sum from (3.6) by $n^{(r-1)t} \le n^v \le \exp(n/(100r))$. On the other hand we have $r(F) \ge \frac{n-v}{r} > n/(2r)$. It follows that $2^{-r(F)}$ dominates each of the at most n/r terms in the sum and the factor $(1+n^{-\frac{1}{2}})^{r(F)}$ as well. This gives us $S[F]/S'[F] = o(1)$. If $v = 0$ then we trivially have $S[F]/S'[F] = o(1)$ as well.

Next we assume that $n/(100r \ln n) < v \le n - (\ln n)^{r-2}\sqrt{n}$. We can interpret the sum in (3.6) as the first $v/r + 1$ terms in the expansion of $\exp\left((p^{-1}-1)(1+n^{-1/2})\frac{e^{r-2}\Delta^2}{\lfloor v/r \rfloor^{r-2}}\right)$, which leads to

$$\frac{S[F]}{S'[F]} \le 2^{-r(F)} \left(1+n^{-\frac{1}{2}}\right)^{r(F)} \exp\left(2p^{-1}\frac{e^{r-2}\Delta^2}{\lfloor v/r \rfloor^{r-2}}\right). \tag{3.8}$$

Again we have $r(F) \geq \frac{n-v}{r} \geq \frac{(\ln n)^{r-2}\sqrt{n}}{r}$, whereas $\exp\left(2p^{-1}\frac{e^{r-2}\Delta^2}{\lfloor v/r\rfloor^{r-2}}\right) = \exp\left(o((\ln n)^{r-2}\sqrt{n})\right)$ by (3.7). Thus, we have $S[F]/S'[F] = o(1)$.

Assume now that $n-(\ln n)^{r-2}\sqrt{n} < v \leq n-\sqrt{n}$. Similarly as in the previous case one gets $r(F) \geq \sqrt{n}/r$ and $\exp\left(2p^{-1}\frac{e^{r-2}\Delta^2}{\lfloor v/r\rfloor^{r-2}}\right) = \exp\left(o(\sqrt{n})\right)$. Again one gets $S[F]/S'[F] = o(1)$ as before.

Finally, let $v > n-\sqrt{n}$ and we are going to use the inequality (3.5) to estimate $S[F]/S'[F]$. We bound β_w with $e(H)^2\binom{w}{r}^{-1} = \alpha^2\binom{n}{r}^2\binom{w}{r}^{-1}$ which is $\alpha^2\binom{n}{r}(1+O(n^{-1/2}))$ for $w \geq n-(r+1)\sqrt{n}$. This gives us

$$\sum_{t=0}^{\sqrt{n}}\left(p^{-1}-1\right)^t\left(1+n^{-\frac{1}{2}}\right)^t\frac{\left(\alpha^2\binom{n}{r}\left(1+O(n^{-1/2})\right)\right)^t}{t!} \leq \exp\left(\frac{1-p}{p}\alpha^2\binom{n}{r}\left(1+O(n^{-1/2})\right)\right).$$

By Lemma 3.1 we have $\frac{1-p}{p}\alpha^2\binom{n}{r}n^{-1/2} = o(1)$. Thus,

$$\exp\left(\frac{1-p}{p}\alpha^2\binom{n}{r}\left(1+O(n^{-1/2})\right)\right) \leq \left(1+o(1)\right)\exp\left(\frac{1-p}{p}\alpha^2\binom{n}{r}\right).$$

As for $t > \sqrt{n}$, we estimate the rest by (3.6) and using (3.7) it follows:

$$\sum_{t=\sqrt{n}}^{v/r}\left(p^{-1}-1\right)^t\left(1+n^{-\frac{1}{2}}\right)^t\left(\frac{e^{r-2}\Delta^2}{\lfloor v/r\rfloor^{r-2}}\right)^t\frac{1}{t!} \leq \sum_{t=\sqrt{n}}^{v/r} o(1)^t = o(1).$$

Combining together we obtain: $\frac{S[F]}{S'[F]} \leq (1+o(1))e^{\alpha^2\binom{n}{r}\frac{1-p}{p}} + o(1) = (1+o(1))e^{\alpha^2\binom{n}{r}\frac{1-p}{p}}$.

Thus, in the first three cases we get $S[F] = o(S'[F])$ and the fourth case implies $S[F] \leq (1+o(1))e^{\alpha^2\binom{n}{r}\frac{1-p}{p}}S'[F]$. Therefore, for every good F, we get in any of the four possible cases that

$$S[F] \leq \left(1+o(1)\right)e^{\alpha^2\binom{n}{r}\frac{1-p}{p}}S'[F].$$

Building the sums over all isomorphism classes of good subhypergraphs F of H we obtain $S_H \leq (1+o(1))e^{\frac{1-p}{p}\alpha^2\binom{n}{r}}S'_H$ which completes the proof. ☐

So far we have $f \leq (1+o(1))e^{-\frac{1-p}{p}\alpha^2\binom{n}{r}}S_H$ and $S_H \leq (1+o(1))e^{\frac{1-p}{p}\alpha^2\binom{n}{r}}S'_H$, thus $f \leq (1+o(1))S'_H$.

As a next step we give an upper bound on $X_F(H)/X_F(K_n^{(r)})$.

Lemma 3.4. *Let H be any r-uniform hypergraph with maximum degree Δ and $F \subseteq H$, then*

$$\frac{X_F(H)}{X_F(K_n^{(r)})} \leq \frac{\left(e(r-1)!\Delta\right)^{r(F)}e^{r(F)+(r-2)k_r(F)}}{n^{r(F)+(r-2)k_r(F)}}$$

Proof. The proof is a straightforward adaptation of [97, Lemma 4.4]. For the sake of completeness, we provide full details. Let $Y_F(H)$ be the number of labelled copies of F in H.

As every unlabeled copy of F corresponds to $|\mathrm{aut}(F)|$ labelled copies, where $|\mathrm{aut}(F)|$ is the number of automorphisms of F, one verifies that $Y_F(H)/Y_F(K_n^{(r)}) = X_F(H)/X_F(K_n^{(r)})$ holds.

Since $Y_F(K_n^{(r)}) = n!$, one needs to estimate $Y_F(H)$. We will embed first exactly one vertex from each of the $k_r(F)$ nontrivial components. This can be done in $(n)_{k_r(F)}$ ways. Next, we can embed $(r-1)$ vertices of each component by embedding one particular edge. This can be done in at most $\Delta(r-1)!$ ways into H. This gives at most $(\Delta(r-1)!)^{k_r(F)}$ possibilities in total. Finally, all the remaining $r(F)-k_r(F)$ vertices from the nontrivial components can be embedded in at most $((r-1)\Delta)^{r(F)-k_r(F)}$ ways. The isolated vertices can be embedded in at most $k_1(F)!$ ways. We estimate $Y_F(H) \le (n)_{k_r(F)} (\Delta(r-1)!)^{k_r(F)} ((r-1)\Delta)^{r(F)-k_r(F)} k_1(F)!$. We obtain (using $(n)_m \ge (n/e)^m$ for all $0 \le m \le n$, $k_r(F) \le n/r$):

$$\frac{Y_F(H)}{Y_F(K_n^{(r)})} \le \frac{(n)_{k_r(F)}\big(\Delta(r-1)!\big)^{k_r(F)}\big((r-1)\Delta\big)^{r(F)-k_r(F)}k_1(F)!}{n!} \le$$
$$\frac{\big((r-1)!\Delta\big)^{r(F)}}{(n-k_r(F))_{n-k_r(F)-k_1(F)}} \le \frac{\big((r-1)!\Delta\big)^{r(F)}}{((n-k_r(F))/e)^{n-k_r(F)-k_1(F)}} \le \frac{\big((r-1)!\Delta\big)^{r(F)}}{((1-1/r)n/e)^{n-k_r(F)-k_1(F)}},$$

and we further estimate this last term from above, using $r(F) \ge k_r(F)$ and $r(F)+(r-2)k_r(F) = n-k_r(F)-k_1(F)$, as follows:

$$\le \frac{\big((r-1)!\Delta\big)^{r(F)}(r/(r-1))^{(r-1)r(F)}e^{r(F)+(r-2)k_r(F)}}{n^{r(F)+(r-2)k_r(F)}} \le \frac{\big(e(r-1)!\Delta\big)^{r(F)}e^{r(F)+(r-2)k_r(F)}}{n^{r(F)+(r-2)k_r(F)}}.$$

□

The lemma above bounds S'_H as follows:

$$S'_H \le {\sum_{F\subseteq H}}' (p^{-1}-1)^{e(F)} \frac{\big(2e(r-1)!\Delta\big)^{r(F)}e^{r(F)+(r-2)k_r(F)}}{n^{r(F)+(r-2)k_r(F)}} =: T'_H.$$

Proceeding exactly as in [97], we introduce the following function ψ:

$$\psi(F) := (p^{-1}-1)^{e(F)} \frac{\big(2e(r-1)!\Delta\big)^{r(F)}e^{r(F)+(r-2)k_r(F)}}{n^{r(F)+(r-2)k_r(F)}}$$

This function is *multiplicative*, i.e. $\psi(F_1 \cup F_2) = \psi(F_1)\psi(F_2)$ for any two hypergraphs F_1 and F_2, where vertices from $V(F_1)\cap V(F_2)$ are isolated both in F_1 and F_2 (in other words: F_1 and F_2 are vertex-disjoint where we don't take into account isolated vertices). Since every good hypergraph is a union of such *disjoint* connected good hypergraphs this yields

$$T'_H = {\sum_{F\subseteq H}}' \psi(F) \le 1 + \sum_{t=1}^{\infty}\frac{1}{t!}\left({\sum_{F\subseteq H}}'' \psi(F)\right)^t, \tag{3.9}$$

where $\sum''$ is the sum over connected good hypergraphs F. We set $T''_H = \sum''_{F\subseteq H}\psi(F)$, thus the above shows $T'_H \le e^{T''_H}$. Lastly, we prove the following estimate on T''_H.

Lemma 3.5. *For every r-uniform hypergraph H on $[n]$ we have*

$$T''_H \le ne^{2r}\sum_{s=r+1}^{n}\left(\frac{12r!^2\Delta^2}{n}\right)^{s-1}p^{-e_H(s)}, \tag{3.10}$$

where Δ is the maximum degree of H.

Proof. The proof is similar to that of Lemma 4.5 from [97]. One rewrites T''_H by going over all good connected hypergraphs F on s vertices (then $r(F) = s - (r-1)$ and $k_r(F) = 1$) and upper bounds the sum as follows:

$$T''_H \leq \sum_{s=r+1}^{n} \frac{\big(2e(r-1)!\Delta\big)^{s-r+1} e^{s-1}}{n^{s-1}} \sum_V \sum_{m=0}^{e_H(s)} \binom{e_H(s)}{m} (p^{-1}-1)^m$$
$$\leq e^{r-2} \sum_{s=r+1}^{n} \frac{(8r!\Delta)^{s-r+1}}{n^{s-1}} \sum_V p^{-e_H(s)},$$

where the second sum is over all s-element sets V such that $H[V]$ is connected.

We consider the shadow graph H' of H. Now every $V \subseteq [n]$ as above also induces a subgraph in H' which is connected and therefore contains a spanning tree. We can estimate the number of such V by estimating the number of labelled trees in H' on s vertices and then unlabelling these.

Given a labelled tree G on s vertices, there are at most $n(\Delta(r-1)!)(\Delta(r-1))^{s-r}$ ways of mapping it into H': n accounts for the first vertex of G, then (in H) we can choose next $(r-1)$ vertices at once in $\Delta(r-1)!$ ways, and finally every remaining vertex in at most $\Delta(r-1)$ ways since $\Delta(H') \leq \Delta(r-1)$. We get at most $n(\Delta(r-1)!)(\Delta(r-1))^{s-r}$ mappings of G into H'. Since there are, by Cayley's formula, s^{s-2} labelled trees on s vertices, there will be at most $n(\Delta(r-1)!)(\Delta(r-1))^{s-r}s^{s-2}$ labellings of s-element sets V such that $H[V]$ is connected. Unlabelling every s-set V gives us at most

$$n\big(\Delta(r-1)!\big)\big(\Delta(r-1)\big)^{s-r} s^{s-2}/s! \leq n\big(\Delta(r-1)!\big)^{s-r+1} e^s$$

sets V. This implies $T''_H \leq ne^{2r} \sum_{s=r+1}^{n} \left(\frac{12r!^2\Delta^2}{n}\right)^{s-1} p^{-e_H(s)}$. □

3.4 Finishing the argument

Now we are in a position to finish the argument. We further estimate T''_H using (3.10) as follows:

$$T''_H \leq ne^{2r} \sum_{s=r+1}^{n} \left(\frac{12r!^2\Delta^2}{n}\right)^{s-1} p^{-e_H(s)} \leq 12e^{2r} r!^2 \sum_{s=r+1}^{n} \left(12r!^2\Delta^4 p^{-e_H(s)/(s-2)} n^{-1}\right)^{s-2}.$$

Therefore we get $T''_H \leq 12e^{2r}r!^2 \sum_{s=r+1}^{n} \left(12r!^2\Delta^4 p^{-\gamma(H)} n^{-1}\right)^{s-2}$, which by condition (2.1) tends to zero as n goes to infinity. Thus, $T''_H = o(1)$, and with $T'_H \leq e^{T''_H}$ and $f \leq (1+o(1))S'_H \leq T'_H$ we obtain $f \leq 1 + o(1)$ and then by Chebyshev's inequality (3.1) the statement of Theorem 2.5 follows for $\mathcal{H}^{(r)}(n, p\binom{n}{r})$.

3.5 Applications

First we obtain the following two corollaries.

Corollary 3.6. *Let r, $\Delta \geq 2$ be integers and H is an r-uniform hypergraph with n vertices, $\Delta(H) \leq \Delta$, $e(H) > n/r$ and $\gamma(H) = e(H)/(n-2)$. Then for $p = \omega\left(n^{-1/\gamma(H)}\right)$ the random graph $\mathcal{H}^{(r)}(n,p)$ contains a copy of H a.a.s., while for every $\varepsilon > 0$ we have for $p \leq (1-\varepsilon)(e/n)^{1/\gamma}$ that $\mathbb{P}[H \subseteq \mathcal{H}^{(r)}(n,p)] \to 0$.*

Proof. Since Δ is fixed and $\gamma(H) \leq (1+o(1))\Delta$, condition (2.1) is satisfied. Moreover, $e(H) > n/r$ implies $\Delta(H) \geq 2$. Theorem 2.5 yields the first part of the claim.

Let X be the number of copies of H in $\mathcal{H}^{(r)}(n,p)$ and we estimate its expectation $\mathbb{E}[X]$ as follows:

$$\mathbb{E}[X] \leq n!\, p^{e(H)} \leq 3\sqrt{n}(1-\varepsilon)^{e(H)}(n/e)^2 = o(1).$$

Now Markov's inequality $\mathbb{P}[X \geq 1] \leq \mathbb{E}[X]$ yields the second part of the corollary. □

We call a hypergraph H d-regular if every vertex of H has degree d.

Corollary 3.7. *Let $r \geq 2$ be an integer and H be an Δ-regular r-uniform hypergraph, where $\Delta = o(n^{1/4})$ but $\Delta = \omega(\log^{1-1/r} n)$. Then for every $\varepsilon > 0$ we have that $\mathcal{H}^{(r)}(n,p)$ contains a.a.s. H if $p = (1+\varepsilon)n^{-r/\Delta}$. Furthermore $\mathbb{P}[H \subseteq \mathcal{H}^{(r)}(n,p)] \to 0$ for $p \leq n^{-r/\Delta}$, i.e. $p = n^{-r/\Delta}$ is a sharp threshold for the appearance of copies of H in $\mathcal{H}^{(r)}(n,p)$.*

Proof. Let X count the copies of H in $\mathcal{H}^{(r)}(n,p)$ and for $p \leq n^{-r/\Delta}$ we have

$$\mathbb{P}[X \geq 1] \leq \mathbb{E}[X] \leq n!\, n^{-re(H)/\Delta} = n!\, n^{-n} = o(1).$$

Next we bound $\gamma(H)$ as follows: $\Delta/r \leq \gamma(H) \leq \frac{\Delta}{r}\frac{(\Delta^{1/(r-1)}+1)}{(\Delta^{1/(r-1)}-1)}$. This is obtained from the estimate $e_H(v) \leq \min\{\Delta v/r, \binom{v}{r}\}$ by considering two cases whether $v \leq \Delta^{1/(r-1)}+1$ or not. Let $\varepsilon \in (0,1)$, then

$$n\Big((1+\varepsilon)n^{-r/\Delta}\Big)^{\gamma(H)}\Delta^{-4} \geq \Big((1+\varepsilon)n^{1/\gamma(H)-r/\Delta}\Delta^{-4r(1+o(1))/\Delta}\Big)^{\gamma(H)} \geq \Big((1+\varepsilon)n^{-2r/(\Delta^{1+1/(r-1)})}\big(1+o(1)\big)\Big)^{\gamma(H)} \to \infty,$$

holds and therefore Theorem 2.5 is applicable and the statement follows. □

Thus, Theorem 2.5 (Corollaries 3.6 and 3.7) states that under some technical conditions the threshold for the appearance of the spanning structure comes from the expectation threshold p_E. In the following we derive asymptotically optimal thresholds for the appearance of various spanning structures in $\mathcal{H}^{(r)}(n,p)$ which are consequences of the Corollaries 3.6 and 3.7.

Hamilton Cycles

The following is a slightly weaker version of Dudek and Frieze [44]. Recall that an r-uniform hypergraph is ℓ-Hamiltonian if it contains $n/(r-\ell)$ edges which form consecutive segments of some cyclic ordering of all vertices and two consecutive edges overlap in ℓ vertices.

Corollary 3.8. *For all integers $r > \ell \geq 2$, $(r-\ell)|n$ and $p = \omega(n^{\ell-r})$ the random hypergraph $\mathcal{H}^{(r)}(n,p)$ is ℓ-Hamiltonian a.a.s.*

Proof. Denote by $C_n^{(r,\ell)}$ an ℓ-overlapping Hamilton cycle on n vertices. It is not difficult to see that $\gamma(C_n^{(r,\ell)}) = \frac{n}{(r-\ell)(n-2)}$. Indeed, let $V \subseteq [n]$ be a set of size $v < n$. Then $C_n^{(r,\ell)}[V]$ is a union of vertex-disjoint ℓ-overlapping paths, where an ℓ-overlapping path of length s consists of $s(r-\ell)+\ell$ ordered vertices and edges are consecutive segments intersecting in ℓ vertices. This gives $e(C_n^{(r,\ell)}[V]) \leq (v-$

$\ell)/(r-\ell)$ and from $\frac{v-\ell}{(r-\ell)(v-2)} \leq \frac{n}{(r-\ell)(n-2)}$ we get that the optimal value for $\gamma(C_n^{(r,\ell)})$ is obtained by the whole cycle, i.e. $\gamma(C_n^{(r,\ell)}) = \frac{n}{(r-\ell)(n-2)}$.

Since $e(C_n^{(r,\ell)}) > n/r$, $\Delta(C_n^{(r,\ell)}) = \lceil \frac{r}{r-\ell} \rceil$ and $n^{2(r-\ell)/n} \to 1$, Corollary 3.6 implies the statement. □

Cube-hypergraphs

The r-uniform d-dimensional cube-hypergraph $Q^{(r)}(d)$ was studied in [33] and its vertex set is $V := [r]^d$ and its hyperedges are r-sets of the vertex set V that all differ in one coordinate. Thus, $Q^{(r)}(d)$ has r^d vertices, dr^{d-1} edges and is d-regular. In the case $r = 2$ this is the usual definition of the (graph) hypercube. The following corollary is a direct consequence of Corollary 3.7.

Corollary 3.9. *For all integers $r \geq 2$, $\varepsilon > 0$ and $p = r^{-r} + \varepsilon$ it holds $\mathbb{P}[Q^{(r)}(d) \subseteq \mathcal{H}^{(r)}(r^d, p)]$ tends to 1 as d tends to infinity. On the other hand, $\mathbb{P}[Q^{(r)}(d) \subseteq \mathcal{H}^{(r)}(r^d, r^{-r})] \to 0$ as $d \to \infty$.*

We remark that, in the case $r = 2$, Riordan [97] proved even better dependence of ε on d, and similar dependence can be shown for $r > 2$.

Lattices

Another example considered in [97] was the graph of the lattice L_k, whose vertex set is $[k]^2$ and two vertices are adjacent if their Euclidean distance is one. There it is shown that $p = n^{-1/2}$ is asymptotically the threshold. One can view L_k as the cubes $Q^{(2)}(2)$ (these are cycles C_4) glued *along* the edges.

We define the ℓ-overlapping hyperlattice $L^{(r)}(\ell, k)$ as the r-uniform hypergraph on the vertex set $[(k-2)(r-\ell)+r]^2$ and the hyperedges being either of the form $\{(x,i), \ldots, (x, i+r-1)\}$ or $\{(j,y), \ldots, (j+r-1, y)\}$, where $x, y \in [(k-2)(r-\ell)+r]$ and $i, j \equiv 1 \mod (r-\ell)$. This hypergraph thus arises if we glue together $(k-1)^2$ copies of $Q^{(r)}(2)$ that overlap on ℓ hyperedges accordingly. Thus, $L^{(2)}(1,k)$ is just the usual graph lattice L_k.

Corollary 3.10. *Let $r \geq 2$ and k be an integer. For $p = \omega\left(n^{-1/2}\right)$ (where $n = (k-2+r)^2$) the random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$ contains a copy of $L^{(r)}(r-1,k)$ a.a.s. Moreover, for $p = n^{-1/2}$, $\mathbb{P}[L^{(r)}(r-1,k) \subseteq \mathcal{H}^{(r)}(n,p)] \to 0$ as k (and thus n) tends to infinity.*

Proof. Observe that $L := L^{(r)}(r-1,k)$ has $(k-2+r)^2$ vertices (which can be associated with $[k-2+r]^2$) and $2(k-1)(k-2+r)$ edges.

We aim to show that $e_L(v) \leq 2(v-r)$ for all $v \geq r+1$. We argue similarly as in [97]. Observe that $e_L(v) \leq 2$ for $v = r+1$. Let now L' be an arbitrary subhypergraph of L on $v+1 \leq (k-2+r)^2$ vertices such that $e(L') = e_L(v+1)$. It is easy to see that there is a vertex of degree 2 in L' (take (i,j) such that $(i+1,j), (i,j+1) \notin V(L')$). It follows that then $e_L(v+1) \leq e_L(v) + 2$ for $v > r+1$ giving $e_L(v) \leq 2(v-r)$ for all $v \geq r+1$.

It follows that $\gamma := \gamma(L) \leq 2$ and applying Theorem 2.5 with $np^\gamma = \omega(1)$ yields the first part. Markov's inequality yields the second part. □

Spheres

Let $r \geq 3$ and let G be a connected planar graph on n vertices with a drawing all of whose faces are cycles of length r. We define a sphere S_n^r as an r-uniform hypergraph all of whose edges correspond to the faces of that particular drawing (note that a sphere is not unique) and the vertex set being $V(G)$. Since every edge of G lies in 2 faces and there are r edges in every face, we obtain $2e(G) = rf(G)$, where $f(G)$ is the number of faces of G. We thus get from Euler's formula for planar graphs the condition $2v(S_n^r) - 4 = (r-2)e(S_n^r)$.

Corollary 3.11. *Let $r \geq 3$ and S be some sphere S_n^r with $\Delta = \Delta(S_n^r)$. Then for $p = \omega\left(\Delta^{2r-4} n^{-(r-2)/2}\right)$ the random r-uniform hypergraph $\mathcal{H}^{(r)}(n,p)$ contains a copy of S_n^r a.a.s.*

Proof. Let G be a planar graph, out of which the sphere S_n^r arose. Let $V \subseteq V(S_n^r)$ and $v = |V|$. We can assume that $G' := G[V]$ is connected[27]. Therefore, we get from Euler's formula: $f(G') = 2 + e(G') - v$. On the other hand, by counting edge-face incidences we get: $rf(G') \leq 2e(G')$ and we obtain: $(r-2)f(G') \leq 2v - 4$, which yields an upper bound $e_{S_n^r}(v) \leq \frac{2v-4}{r-2}$. It follows that $\gamma(S_n^r) \leq 2/(r-2)$ holds and this upper bound is attained by the sphere S_n^r itself. Thus, we immediately get $\gamma(S_n^r) = 2/(r-2)$. Since (2.1) holds, the statement follows now directly from Theorem 2.5. □

Powers of tight Hamilton cycles

Consider a tight Hamilton cycle $C_n^{(r,r-1)}$ with n vertices which are ordered cyclically. Given an integer k, we define the k-th power $C_n^{(r)}(k)$ of $C_n^{(r,r-1)}$ to consist of all r-tuples e such that the maximum distance in this cyclic ordering between any two vertices in e is at most $r+k-2$. Recall that in the graph case, the threshold for the appearance of $C_n^{(2)}(k)$ is known to be $n^{-1/k}$ for $k \geq 3$ [84]. If we count the edges of $C_n^{(r)}(k)$ by their leftmost vertex we get $e(C_n^{(r)}(k)) = n\binom{r+k-2}{r-1}$ for $n \geq r + 2k - 2$.

Corollary 3.12. *Let $r \geq 3$ and $k \geq 2$ be integers. Suppose that $p = \omega(n^{-1/\binom{r+i-2}{r-1}})$, then the random hypergraph $\mathcal{H}^{(r)}(n,p)$ contains a.a.s. a copy of $C_n^{(r)}(k)$. This threshold is asymptotically optimal.*

Proof. One can argue similarly to Proposition 8.2 in [84] to show $\gamma(C_n^{(r)}(k)) \leq \binom{r+k-2}{r-1} + O_{r,k}(1/n)$. The statement follows from Theorem 2.5. We omit the details. □

[27] Since we can add each time one edge to connect two components and this doesn't create any further cycles.

Chapter 4

Finding tight Hamilton cycles in random hypergraphs

In this chapter we prove[28] Theorem 2.6, which resulted from collaboration with Allen, Koch, and Person [6]. We first give a short overview of the algorithm, then in Section 4.2 we prove the theorem, where the proofs of the Connecting and Reservoir Lemma are given in Section 4.3 and 4.4 respectively.

4.1 An informal algorithm overview

We briefly give some additional notation. An *s-tuple* $(u_1, \dots, u_s)$ of vertices is an ordered set of distinct vertices. We often denote tuples by bold symbols, and occasionally also omit the brackets and write $\mathbf{u} = u_1, \dots, u_s$. Additionally, we may also use a tuple as a set and write for example, if S is a set, $S \cup \mathbf{u} := S \cup \{u_i : i \in [s]\}$. The *reverse* of the s-tuple $\mathbf{u}$ is the s-tuple $\overleftarrow{\mathbf{u}} := (u_s, \dots, u_1)$.

In an r-uniform hypergraph $\mathcal{G}$ the tuple $P = (u_1, \dots, u_\ell)$ forms a *tight path* if the set $\{u_{i+1}, \dots, u_{i+r}\}$ is an edge for every $0 \le i \le \ell - r$. For any $s \in [\ell]$ we say that P *starts* with the s-tuple $(u_1, \dots, u_s) =: \mathbf{v}$ and *ends* with the s-tuple $(u_{\ell-(s-1)}, \dots, u_\ell) =: \mathbf{w}$. We also call $\mathbf{v}$ the *start s-tuple* of P, $\mathbf{w}$ the *end s-tuple* of P, and P a $\mathbf{v} - \mathbf{w}$ path. The *interior* of P is formed by all its vertices but its start and end $(r-1)$-tuples. Note that the interior of P is not empty if and only if $\ell > 2(r-1)$.

Overview of the algorithm

We start with the given sample of the random hypergraph $\mathcal{H}^{(r)}(n,p)$ and we will reveal the edges as we proceed. First, using the Reservoir Lemma (Lemma 4.1 below), we construct a tight path P_{res} which covers a small but bounded away from zero fraction of $[n]$, which has the *reservoir property*, namely that there is a set $R \subset V(P_{\text{res}})$ of size $2Cp^{-1}\log n \le 2n/\log^2 n$ such that for any $R' \subset R$, there is a tight path covering exactly the vertices $V(P_{\text{res}}) \setminus R'$ whose ends are the same as those of P_{res}, and this tight path can be found given P_{res} and R' in time polynomial in n a.a.s.

We now greedily extend P_{res}, choosing new vertices when possible and otherwise vertices in R. We claim that a.a.s. this strategy produces a structure P_{almost} which is almost a tight path extending P_{res} and covering $[n]$. The reason it is only *almost* a tight path is that some vertices in R may be used twice. We denote the set of vertices used twice by R'_1. But we will succeed in covering $[n]$ with high probability. Recall that, due to the reservoir property, we can dispense with the vertices from R'_1 in the part P_{res} of the almost tight Hamilton path P_{almost}.

[28]The proof presented in this chapter is a close adaption from [6].

Finally, we apply the Connecting Lemma (Lemma 4.2 below) to find a tight path in $R \setminus R'_1$ joining the ends of P_{almost}, and using the reservoir property this gives the desired tight Hamilton cycle.

This approach is similar to that in [4]. The main difference is the way we prove the Reservoir Lemma (Lemma 4.1). In both [4] and this paper, we first construct many small, identical, vertex-disjoint *reservoir structures* [29]. A reservoir structure contains a spanning tight path, and a second tight path with the same ends which omits one *reservoir vertex*. We then use Lemma 4.2 to join the ends of all these reservoir structures together into the desired P_{res}. In [4], reservoir structures are of constant size (depending on the ε) and they are found by using brute-force search. This is slow and is also the cause of the algorithm in [4] being randomised: there it is necessary to simulate exposure in rounds of the random hypergraph since the brute-force search reveals all edges. In this paper, by contrast, we construct reservoir structures by a local search procedure which is both much faster and reveals much less of the random hypergraph.

We will perform all the constructions in this paper by using local search procedures. At each step, we reveal all the edges of $\mathcal{H}^{(r)}(n,p)$ which include a specified $(r-1)$-set, the *search base*. The number of such edges will always be in expectation of the order of pn, so that by Chernoff's inequality and the union bound, with high probability at every step in the algorithm the number of revealed edges is close to the expected number. Of course, what we may not do is attempt to reveal a given edge twice: therefore, we keep track of an *exposure hypergraph* $\mathcal{E}$, which is the $(r-1)$-uniform hypergraph consisting of all the $(r-1)$-sets which have been used as search bases up to a given time in the algorithm. We will show that $\mathcal{E}$ remains quite sparse, which means that at each step we have almost as much freedom as at the start when no edges are exposed.

For concreteness, we use a doubly linked list of vertices as the data structure representing a tight (almost-) path. However, this choice of data structure is not critical to the proof and we will not further comment on it. The reader can easily verify that the various operations we describe can be implemented in the claimed time using this data structure.

4.2 Two key lemmas and the main proof

Two Key Lemmas

Recall the definition of the *reservoir path* P_{res}. It is an r-uniform hypergraph with a special subset $R \subsetneq V(P_{\text{res}})$ and some start and end $(r-1)$-tuples $\mathbf{v}$ and $\mathbf{w}$ respectively, such that:

(1) P_{res} contains a tight path with the vertex set $V(P_{\text{res}})$ and the *end tuples* $\mathbf{v}$ and $\mathbf{w}$, and

(2) for *any* $R' \subseteq R$, P_{res} contains a tight path with the vertex set $V(P_{\text{res}}) \setminus R'$ and the *end tuples* $\mathbf{v}$ and $\mathbf{w}$.

We first give the lemma which constructs P_{res}. In addition to with high probability returning P_{res}, we also need to describe the likely resulting exposure hypergraph.

Lemma 4.1 (Reservoir Lemma)**.** *For each $r \geq 3$ and $p \in (0,1]$ there exists $C > 0$ and a deterministic $O(n^r)$-time algorithm whose input is an n-vertex r-uniform hypergraph G and whose output is either* ***Fail*** *or*

[29] in some part of the literature, mostly in the dense case, this structure is called an absorber.

a reservoir path P_{res} with ends $\mathbf{u}$ and $\mathbf{v}$ and an $(r-1)$-uniform exposure hypergraph $\mathcal{E}$ on vertex set $V(G)$ with the following properties.

(i) All vertices of P_{res} and edges of $\mathcal{E}$ are contained in a set S of size at most $\frac{n}{4}$.

(ii) The reservoir $R \subset V(P_{\mathrm{res}})$ has size $2Cp^{-1}\log n$.

(iii) There are no edges of $\mathcal{E}$ contained in $R \cup \mathbf{u} \cup \mathbf{v}$.

(iv) All r-sets in $V(G)$ which have been exposed contain at least one edge of $\mathcal{E}$.

*When G is drawn from the distribution $\mathcal{H}^{(r)}(n,p)$ and $p \geq Cn^{-1}\log^3 n$, the algorithm returns **Fail** with probability at most n^{-2}.*

Furthermore, we need a lemma which allows us to connect two given tuples with a not too long path. This lemma is the engine behind the proof and behind the Reservoir Lemma.

Lemma 4.2 (Connecting Lemma)**.** *For each $r \geq 3$ there exist $c, C > 0$ and a deterministic $O(n^{r-1})$-time algorithm whose input is an n-vertex r-uniform hypergraph G, a pair of distinct $(r-1)$-tuples $\mathbf{u}$ and $\mathbf{v}$, a set $S \subset V(G)$ and an $(r-1)$-uniform exposure hypergraph $\mathcal{E}$ on the same vertex set $V(G)$. The output of the algorithm is either **Fail** or a tight path of length $o(\log n)$[30] in G whose ends are $\mathbf{u}$ and $\mathbf{v}$ and whose interior vertices are in S, and an exposure hypergraph $\mathcal{E}' \supset \mathcal{E}$. We have that all the edges $E(\mathcal{E}') \setminus E(\mathcal{E})$ are contained in $S \cup \mathbf{u} \cup \mathbf{v}$.*

*Suppose that G is drawn from the distribution $\mathcal{H}^{(r)}(n,p)$ with $p \geq C(\log n)^3/n$, that $\mathcal{E}$ does not contain any edges intersecting both S and $\mathbf{u} \cup \mathbf{v}$. If furthermore $|S| = Cp^{-1}\log n$ and $|e(\mathcal{E}[S])| \leq c|S|^{r-1}$ then $e(\mathcal{E}') \leq e(\mathcal{E}) + O(|S|^{r-2})$ and the algorithm returns **Fail** with probability at most n^{-5}.*

Overview continued: more details

We now describe the algorithm claimed by Theorem 2.6, which we state in a high-level overview as Algorithm 1 and explain somewhat informally some of the arguments.

Algorithm 1: Find a tight Hamilton cycle in $\mathcal{H}^{(r)}(n,p)$

1. use subroutine from Lemma 4.1 to either construct P_{res} (with ends $\mathbf{u}$, $\mathbf{v}$ and exposure hypergraph $\mathcal{E}$ on S) or halt with **failure**;
   $L := V(G) \setminus S$;
   $U := S \setminus V(P_{\mathrm{res}})$;
2. extend P_{res} greedily from v to cover all vertices of U and using up to $n/2$ vertices of L, otherwise halt with **failure**;
3. extend P_{res} further greedily to P_{almost} by covering all vertices of L and using up to $|R|/2$ vertices of R, otherwise halt with **failure**;
4. use subroutine of Lemma 4.2 to connect the ends of P_{almost} using the unused at least $|R|/2$ vertices of R, otherwise halt with **failure**;

[30] We will make this more precise later. For now the reader can replace this by at most $Cn/\log\log n$.

Step 1. Given G drawn from the distribution $\mathcal{H}^{(r)}(n,p)$, we begin by applying Lemma 4.1 to a.a.s. find a reservoir path $P_{\rm res}$ with ends $\mathbf{u}$ and $\mathbf{v}$ contained in a set S of size $\frac{n}{4}$. Let $L = V(G) \setminus S$, and $U = S \setminus V(P_{\rm res})$. Recall that by Lemma 4.1 (i) and (iii), all edges of $\mathcal{E}$ are contained in S; and $R \cup \mathbf{u} \cup \mathbf{v}$ contains no edges of $\mathcal{E}$. By (iv) all exposed r-sets contain an edge of $\mathcal{E}$; by choosing a little carefully where to expose edges (see Step 2 below), we will not need to worry about what exactly the edges of $\mathcal{E}$ are beyond the above information.

Step 2. We extend $P_{\rm res} := P_0$ greedily, one vertex at a time, from its end $\mathbf{u} = \mathbf{u}_0$, to cover all of U. At each step i, we simply expose the edges of G which contain the end $\mathbf{u}_{i-1}$ of P_{i-1} and whose other vertex is not in $V(P_{i-1})$, choose one of these edges e and add the vertex from $e \setminus \mathbf{u}_{i-1}$ to P_{i-1} to form P_i. The rule we use for choosing e is the following: if i is congruent to 1 or 2 modulo 3, we choose e such that $e \setminus \mathbf{u}_{i-1}$ is in L, and if i is congruent to 0 modulo 3 we choose e such that $e \setminus \mathbf{u}_{i-1}$ is in U if it is possible; if not we choose e such that $x_i := e \setminus \mathbf{u}_{i-1}$ is in L. The point of this rule is that at each step we want to choose an edge which contains at least two vertices of L, because no such r-set can contain an edge of $\mathcal{E}$ since all the edges of $\mathcal{E}$ are contained in S (Property (i)). We will see that while $U \setminus V(P_{i-1})$ is large, we always succeed in choosing a vertex in U when i is congruent to 0 modulo 3. When it becomes small, we do not, but a.a.s. we succeed often enough to cover all of U while using not more than $\frac{5n}{8}$ vertices of L.

Step 3. Next, we continue the greedy extension, this time choosing a vertex in L when possible and in R when not, until we cover all of L. It follows from the first two steps and Properties (i) and (iii) that no edge of $\mathcal{E}$ is in $L \cup R$. Thus, at each step we choose from newly exposed edges and again we a.a.s. succeed in covering L using only a few vertices of R. Let the final almost-path (which uses some vertices $R_1' \subseteq R$ twice) be $P_{\rm almost}$, and R_1 the subset of R consisting of vertices we did not use in the greedy extension, i.e. $R_1 = R \setminus R_1'$.

Step 4. At last, $P_{\rm almost}$ covers $V(G) = L \cup U \cup V(P_{\rm res})$. Its ends, together with the vertices of R_1, satisfy the conditions of Lemma 4.2, which we apply to a.a.s. complete $P_{\rm almost}$ to an almost-tight cycle H' in which some vertices of R_1 are used twice. The reservoir property of R now gives a tight Hamilton cycle H.

Runtime. Our applications of Lemmas 4.1 and 4.2 take time polynomial in n by the statements of those lemmas; the greedy extension procedure is trivially possible in $O(n^2)$ time[31]. Finally the construction of $P_{\rm res}$ allows us to obtain H from H' in time $O(n^2)$: we scan through $P_{\rm res}$, for each vertex r of R we scan the remainder of H' to see if it appears a second time, and if so locally reorder $V(P_{\rm res})$ to remove r from $P_{\rm res}$.

To prove Theorem 2.6, what remains is to justify our claims that various procedures above a.a.s. succeed.

[31] Since at each extension step we just need to look at the neighbourhood of an $(r-1)$-tuple, and there are $O(n)$ steps.

Setup for the proof of Theorem 2.6

We choose $C \geq \max\{C_{L_{4.1}}, C_{L_{4.2}}, 10^8\}$ large enough for Lemmas 4.1 and 4.2. For this proof we do not need to know the value of c required for Lemma 4.2. We suppose that n is large enough to make $\log\log n$ larger than any constant appearing in the following proof. We will use Chernoff's inequality, Theorem 2.16, at various occasions.

Constructing P_{res}

Let G be drawn from the distribution $\mathcal{H}^{(r)}(n,p)$. Lemma 4.1 states that with probability at least $1-n^{-2}$, a reservoir path P_{res} in G is found in polynomial time. From this point on, at each step except the final connection, when we expose edges at an $(r-1)$-set $\mathbf{x}$, that $(r-1)$-set will be included in the path we construct. Hence in future steps we will not examine edges containing $\mathbf{x}$. Thus while we should keep updating $\mathcal{E}$, in fact we will never need to know which edges are added after generating P_{res}.

Extending P_{res} to cover all of U

We next aim to prove that with high probability the greedy extension of P_{res} to cover U succeeds, with at least $n/8$ vertices of L remaining uncovered at the end. Recall that we chose $|S| = \frac{n}{4}$ and thus $|L| = \frac{3n}{4}$. We choose the next vertex from L when i is congruent to 1 or 2 modulo 3 or when we fail to extend into U. At each step i where at least $n/8$ vertices of L are uncovered, we expose all the r-sets in $V(G)$ which contain the end u_{i-1} of P_{i-1} and a vertex of L. The greedy algorithm can only fail to complete step i if none of these r-sets turn out to be edges, which happens with probability at most $(1-p)^{n/8} \leq \exp\big(-\frac{pn}{8}\big) < n^{-4}$ (since the edges of the random hypergraph are independent). Taking the union bound, the greedy algorithm to cover U fails before covering $\frac{5}{8}n$ vertices of L with probability at most n^{-3}.

Similarly, for any i such that $\big|U \setminus V(P_{i-1})\big| \geq Cp^{-1}\log n$, if i is divisible by 3 the probability that no edge containing u_{i-1} and a vertex of $U \setminus V(P_{i-1})$ is in G is at most $\exp\big(-C\log n\big) < n^{-4}$. It follows that with probability at most n^{-3} the greedy algorithm chooses a vertex of L when i is divisible by 3 and $U \setminus V(P_{i-1})$ has size at least $Cp^{-1}\log n$. Let t_1 be the first time in the greedy extension procedure when $U \setminus V(P_{t_1})$ has size less than $Cp^{-1}\log n$.

It remains to show that while the last $Cp^{-1}\log n$ vertices of U are covered, at most $n/8$ vertices of L are used. We split these last $Cp^{-1}\log n$ vertices into the last $\frac{1}{2}p^{-1}$ vertices and the rest. When x vertices of U remain uncovered with $x \geq \frac{1}{2}p^{-1}$, then the probability of choosing a vertex of U for the vertex x_i extending P_{i-1} (when i is divisible by 3) is at least $1-(1-p)^x \geq \frac{1}{3}$. By Chernoff's inequality, the probability that at time $t_2 := t_1 + 6Cp^{-1}\log n$ there are more than $\frac{1}{2}p^{-1}$ vertices of U remaining uncovered is at most $\exp\big(-\frac{1}{6}Cp^{-1}\log n\big) \leq n^{-3}$. Next, we show that we cover all but at most $\log n$ vertices of U in not too much more time.

To see this, consider the following event. For $1 \leq j \leq 7n/8$ and $\log n \leq x \leq \frac{1}{2}p^{-1}$, let $A(x,j)$ be the event that we have $\big|U \setminus V(P_j)\big| = x$ and $\big|U \setminus V(P_{j-3000p^{-1}})\big| \leq 2x$. We claim that the probability for any of these events to hold is at most n^{-3}. Indeed, if for some given x and j the event $A(x,j)$ occurs, then at each of the at least $500p^{-1}$ values of i with $j-3000p^{-1} \leq i \leq j$, an edge containing $\mathbf{u}_{i-1}$ and a vertex of U appears with probability at least $1-(1-p)^x \geq px/2$ (since $x \leq \frac{1}{2}p^{-1}$). Thus for $A(x,j)$ to hold, it is necessary that a sum of at least $500p^{-1}$ Bernoulli random variables, each with probability at least $px/2$, is at most x. Chernoff's inequality states that this probability is at most $\exp\left(-\frac{250x}{12}\right) \le n^{-5}$, and taking the union bound over all $A(x,j)$ the claim follows. Taking in particular $x = 2^{-k}n/\log n$ for $k \ge 1$ such that $2^{-k}n \log n \ge \log n$ (so $k \le \log n$) we see that with probability at least $1 - n^{-3}$, at time $t_3 := t_2 + 3000p^{-1}\log n$ there are at most $\log n$ vertices of U remaining uncovered.

While at least one vertex of U remains uncovered, the probability that when i is divisible by three we choose a vertex of U is at least p. Applying Chernoff's inequality, the probability that at time $t_4 := t_3 + 300p^{-1}\log n$ we still have not covered all of U is at most $\exp(-\frac{100\log n}{12}) \le n^{-3}$. Putting all this together, the probability that $V(P_{t_4})$ does not cover U is at most $4n^{-3}$. Since $t_1 \le 3|U|$, since $|U| \le |S| \le n/4$, and since $t_4 - t_1 \le n/16$, we conclude that with probability at least $1 - 4n^{-3}$ the greedy extension procedure indeed covers U with at least $n/8$ vertices of L left uncovered. Let t_5 be the first time at which P_{t_5} covers U.

Extending P_{res} further to P_{almost} by covering all of L

We now repeat a similar procedure to use up all of $L \setminus V(P_{t_5})$ while not using too many vertices in R. Since no edges of $\mathcal{E}$ are contained in $R \cup L$, at each time t, all the r-sets containing the end $\mathbf{u}_{t-1}$ of P_{t-1} and a vertex of $L \cup R \setminus V(P_{t-1})$ are unrevealed. In particular, provided that at each step we have $\left|R \setminus V(P_{t-1})\right| \ge \frac{1}{2}|R|$, by Chernoff's inequality with probability at least $1 - n^{-4}$ at least one edge of G is found consisting of $\mathbf{u}_{t-1}$ and a vertex of $R \setminus V(P_{t-1})$. Taking the union bound, the probability of the extension procedure failing when $\left|R \setminus V(P_{t-1})\right| \ge \frac{1}{2}|R|$ is at most n^{-3}.

As long as $\left|L \setminus V(P_{t-1})\right| \ge \frac{C}{100}p^{-1}\log n$, we have by Chernoff's inequality with probability at most $\exp\left(-\frac{C}{300}\log n\right) \le n^{-4}$ that there is no edge of G containing $\mathbf{u}_{t-1}$ and a vertex of $L \setminus V(P_{t-1})$; in particular with probability at least $1 - n^{-3}$ the greedy extension covers all but at most $\frac{C}{100}p^{-1}\log n$ vertices of L before using any vertex of R. Let t_6 be the time at which all but at most $\frac{C}{100}p^{-1}\log n$ vertices of L are covered. Again, we now consider the time taken to cover all but $\frac{1}{2}p^{-1}$ vertices of L. At each time the probability of being able to choose a vertex of L to extend our path with is at least $\frac{1}{3}$, so that with probability at least $1 - n^3$ we cover all but at most $\frac{1}{2}p^{-1}$ vertices of L by time $t_7 \le t_6 + \frac{C}{25}p^{-1}\log n$. In particular we use at most $\frac{C}{25}p^{-1}\log n$ vertices of R in this time.

By the same analysis as before, the total time taken to go from covering all but at most $\frac{1}{2}p^{-1}$ vertices of L to covering all but at most $\log n$ vertices of L and then all vertices of L is with probability at least $1 - 2n^{-3}$ not more than $3000p^{-1}\log n + 300p^{-1}\log n$. Putting this together, provided all these good events hold we succeed in covering all of L having used at most

$$\tfrac{C}{25}p^{-1}\log n + 3300p^{-1}\log n < Cp^{-1}\log n = \tfrac{1}{2}|R|$$

vertices of R.

In sum, with probability at least $1 - n^{-2} - 8n^{-3}$, the algorithm succeeds in generating P_{almost}, with the property that the set $R' \subset R$ of vertices not used in the greedy extension has size at least $\frac{1}{2}|R|$.

Connecting the end tuples of P_{almost} and getting the tight Hamilton cycle

Applying Lemma 4.2 to connect the end tuples of P_{almost} in a subset of R' of size $Cp^{-1}\log n$ (which is possible since R' together with the ends of P_{almost} contains no edges of $\mathcal{E}$ and since $|R'| \geq n/\log^2 n$), with probability at least $1 - n^{-4}$ we find the desired almost-tight cycle H', which gives us deterministically the desired tight Hamilton cycle H. Thus as desired the probability that our algorithm fails to find a tight Hamilton cycle is at most n^{-1}. $\square$

4.3 Proof of the Connecting Lemma

In this section we prove Lemma 4.2 and a very similar lemma (Lemma 4.5) dealing with *spike-paths* which we will require for Lemma 4.1. A spike-path is similar to a tight path but after $(r-1)$-steps the direction of the last $(r-1)$-tuple is inverted.

Preliminaries

Definition 4.3 (Spike path). *In an r-uniform hypergraph, a spike path of length t consists of a sequence of t pairwise disjoint $(r-1)$-tuples $\mathbf{a}_1, \ldots, \mathbf{a}_t$, where $\mathbf{a}_i = (a_{i,1}, \ldots, a_{i,r-1})$ for all i, with the property, that the edges $\{a_{i,r-j}, \ldots, a_{i,1}, a_{i+1,1}, \ldots, a_{i+1,j}\}$ are present for all $i = 1, \ldots, t-1$ and $j = 1, \ldots, r-1$. We call $\mathbf{a}_i$ the i-th spike.*

This is the same as taking t tight paths of length $2(r-1)$, where the end $(r-1)$-tuples of path i are $\mathbf{x}_i$ and $\mathbf{y}_i$, and identifying $\overleftarrow{\mathbf{x}_i}$ with $\mathbf{y}_{i+1}$ for all $i = 1, \ldots, t-1$. The proofs of Lemmas 4.2 and the spike-path version Lemma 4.5 are essentially identical, so we give the details of the former and then explain how to modify it to obtain the latter.

For an $(r-1)$-tuple $\mathbf{u}$ and an integer i we define a *fan* $\mathcal{F}_i(\mathbf{u})$ in an r-uniform hypergraph $\mathcal{H}$ as a set $\{P_1, \ldots, P_s\}$ of tight paths in $\mathcal{H}$, of length i or $i+1$, starting in $\mathbf{u}$. For any set or tuple $\mathbf{a}$, let $\{P_j\}_{j \in I}$ be the subcollection of tight paths from $\mathcal{F}_i(\mathbf{u})$ in which $\mathbf{a}$ appears as a consecutive interval (in arbitrary order). The *leaves* or *ends* of $\mathcal{F}_i(\mathbf{u})$ are the ending $(r-1)$-tuples of alle the paths $P_1, \ldots, P_s$. We denote by $\text{mult}(\mathbf{a})$ the number of different paths we see in $\{P_j\}_{j\in I}$ after truncating behind $\mathbf{a}$.

Idea and further notation

The basic idea is that, starting with the $\mathbf{u}$ and $\mathbf{v}$ and the empty fans $\mathcal{F}_0(\mathbf{u})$ and $\mathcal{F}_0(\mathbf{v})$, we want to fan out. That is, for each path in $\mathcal{F}_i(\mathbf{u})$ we will find a large collection of ways to extend by one vertex and all the resulting paths form $\mathcal{F}_{i+1}(\mathbf{u})$. We do this until we have fans $\mathcal{F}_t(\mathbf{u})$ and $\mathcal{F}_t(\mathbf{v})$ with

$$Q := p^{-(r-1)/2} \log n$$

leaves each. This happens roughly when we have

$$t := 2 \cdot \left\lceil \frac{\log(Q)}{\log(\log n)} \right\rceil \leq (r-1) \cdot \left\lceil \frac{\log(p^{-1})}{\log(\log n)} \right\rceil + 2 = o(\log n).$$

A complication is that in this process we have to avoid the edges of $\mathcal{E}$ when expanding the fans. In order to make the modifications for the promised spike-path variation easy (cf. Lemma 4.5 below), we will do something a little more complicated. We split into expansion and continuation phases, each of length $r-1$. The first phase is an expansion phase, so when forming $\mathcal{F}_1(\mathbf{u})$, ..., $\mathcal{F}_{r-1}(\mathbf{u})$ we find many ways to extend each path by one vertex and put all of them into the next fan. The second phase is a continuation phase, so when forming $\mathcal{F}_r(\mathbf{u}), \ldots, \mathcal{F}_{2r-2}(\mathbf{u})$ we choose only one way to extend each path. As soon as we have a collection of paths with the desired Q leaves, we cease expanding (even if we are still in an expansion phase) and simply continue each path such that each has the same length. We construct fans from $\mathbf{v}$ similarly, and we continue construction up to $\mathcal{F}_t(\mathbf{v})$.

In the final step we find $r-1$ further edges connecting two of the leaves, giving us a tight path connecting $\mathbf{u}$ to $\mathbf{v}$. Again there is a complication here: some pairs of leaves $(\mathbf{w}, \mathbf{x})$ may be *blocked* by edges of $\mathcal{E}$, meaning that inside some r consecutive vertices of the concatenation $\mathbf{w}\overleftarrow{\mathbf{x}}$ there is an edge of $\mathcal{E}$. If a pair of leaves is blocked, then trying to reveal $(r-1)$ edges connecting the pair would mean revealing an edge of the random hypergraph twice (and if a pair is not blocked then doing so does not reveal any edge twice). We need to take this into account in our analysis, and we need to construct $\mathcal{F}_t(\mathbf{v})$ carefully to avoid creating *dangerous* leaves for which a large fraction of the pairs is blocked.

To make this precise, we use the following algorithm.

Algorithm 2: Find a connecting path from $\mathbf{u}$ to $\mathbf{v}$

---

split S into equal parts $S_1, \ldots, S_{4(r-1)}, S'_1, \ldots, S'_{4(r-1)}$;
$\mathcal{F}_t(\mathbf{u}) := \mathrm{BuildFan}(\mathbf{u}, S_1, \ldots, S_{4(r-1)}, \emptyset)$;
set $D := \{\mathbf{x} \in S^{r-1} : (\mathbf{w}, \mathbf{x})$ is blocked for at least $\xi' Q$ leaves $\mathbf{w}$ of $\mathcal{F}_t(\mathbf{u})\}$;
$\mathcal{F}_t(\mathbf{v}) := \mathrm{BuildFan}(\mathbf{v}, S'_1, \ldots, S'_{4(r-1)}, D)$;
find $r-1$ edges connecting a leaf of $\mathcal{F}_t(\mathbf{u})$ to the reverse of one of $\mathcal{F}_t(\mathbf{v})$;
return tight path P connecting $\mathbf{u}$ to $\mathbf{v}$;

The subroutine BuildFan takes as input a starting tuple, the sets in which to build a fan, and a *danger hypergraph* D which is important for the construction of the second fan: it is an $(r-1)$-uniform hypergraph which records the tuples in $S'_1, \ldots, S'_{4(r-1)}$ which we cannot easily connect to the leaves of $\mathcal{F}_t(\mathbf{u})$. The algorithm ensures that no leaf of a fan will be a dangerous tuple. Though we only need this for the leaves of the final fan, it is convenient to maintain this property throughout. For convenience, we write S_i for the set $S_{i \bmod 4(r-1)} \in \{S_1, \ldots, S_{4(r-1)}\}$ with $S_0 = S_{4(r-1)}$; the point of these sets is that we choose the ith vertex of each path in S_i, which is helpful in the analysis. Finally, we need to ensure that we always choose *good* vertices which allow us to continue our construction and prove various probabilistic statements. To that end, we define a vertex b to be *good* with respect to an exposure hypergraph $\mathcal{E}$, a set $\mathcal{F}$ of paths with distinct ends, a danger hypergraph D and a $(r-1)$-tuple $\mathbf{a}$ if none of the following statements hold for any (possibly empty) tuple $\mathbf{c}$ whose vertices are contained in those of $\mathbf{a}$ (not necessarily in the same order).

(i) b appears somewhere on the unique path $P(\mathbf{a})$ ending in $\mathbf{a}$,

(ii) $|\mathbf{c}| \leq r-2$ and $\deg_{\mathcal{E}}(\{\mathbf{c}, b\}, S) > \xi^{r-|\mathbf{c}|-1}|S|^{r-|\mathbf{c}|-2}$,

(iii) $\mathrm{mult}(\{\mathbf{c}, b\}) > \xi^{r-|\mathbf{c}|-1} Q \cdot |S|^{-|\mathbf{c}|-1} \cdot \log^{|\mathbf{c}|+1} n$, and

(iv) $|\mathbf{c}| \le r-2$ and $\deg_D(\{\mathbf{c}, b\}, S) > (\xi'|S|)^{r-|\mathbf{c}|-2}$.

Normally $\mathcal{E}$, $\mathcal{F}$ and D will be clear from the context and we will simply say good for $\mathbf{a}$. We are finally ready to give the BuildFan subroutine.

Algorithm 3: $\mathrm{BuildFan}(\mathbf{s}, T_1, \ldots, T_{4(r-1)}, D)$

$\mathcal{F}_0 := \{\mathbf{s}\}$;
foreach $i = 1, \ldots, t$ **do**
  if $i \mod 2(r-1) \in \{1, \ldots, r-1\}$ **then**
    $\mathrm{phase} =$ *expand*;
  else
    $\mathrm{phase} =$ *continue*;
  end
  $\mathrm{NumPaths} := |\mathcal{F}_{i-1}|$;
  $\mathcal{F}_i := \mathcal{F}_{i-1}$;
  foreach $P \in \mathcal{F}_{i-1}$ **do**
    let the $(r-1)$-tuple $\mathbf{a}$ be the end of P ;
    reveal the edges of G containing $\mathbf{a}$ and add $\mathbf{a}$ to $\mathcal{E}$;
    let $T \subseteq T_i$ be the set of vertices b which are good for $\mathbf{a}$ and $\{\mathbf{a}, b\}$ is an edge;
    if $\mathrm{phase} =$ *expand* **then**
      $\mathrm{Add} := \min\big(\log n, Q + 1 - \mathrm{NumPaths}\big)$;
      choose Add vertices $b_1, \ldots, b_{\mathrm{Add}} \in T$;
      $\mathcal{F}_i := \mathcal{F}_i \cup \{(P, b_1), \ldots, (P, b_{\mathrm{Add}})\} \setminus \{P\}$;
      $\mathrm{NumPaths} := \mathrm{NumPaths} + \mathrm{Add} - 1$;
    else
      choose a vertex $b \in T$;
      $\mathcal{F}_i := \mathcal{F}_i \cup \{(P, b)\} \setminus \{P\}$;
    end
  end
end
return $\mathcal{F}_t$;

Setup

We set

$$\xi' = \frac{1}{100^r}\,, \quad \xi = (\xi')^r/(2r2^{20r})\,, \quad \delta = 8^r\xi + \xi'\,, \quad C = 10^{8r} \quad \text{and} \quad c = 10^{-r}\xi^r\,. \tag{4.1}$$

The proof amounts to showing two things. First, BuildFan is likely to succeed—that is, that it does not fail for lack of good vertices before returning a fan, that the returned fan does have size Q, and that it does not add too many tuples to $\mathcal{E}$. Second, the required extra $r-1$ edges which should connect the fans can be found.

Creating the fans

We begin by showing that, whether we choose $\mathbf{s} = \mathbf{u}$, $T_i = S_i$ and $D = \emptyset$ or we choose $\mathbf{s} = \mathbf{v}$, $T_i = S'_i$ and D as given in Algorithm 2, the subroutine $\mathrm{BuildFan}(\mathbf{s}, T_1, \ldots, T_{4(r-1)}, D)$ is likely to succeed, using the following claim.

We define L_i to be the leaves of $\mathcal{F}_i$.

Claim 4.4. *If step i was successful, then step $i+1$ is successful with probability at least $1 - n^{-3r}$ and the following holds throughout step $i+1$ for each $\mathbf{a} \in L_{i+1}$ and each non-empty $\mathbf{c}$ whose vertices are chosen from $\mathbf{a}$, not necessarily in the same order.*

P1 *Each path in $\mathcal{F}_i$ extends to at least one path in $\mathcal{F}_{i+1}$; if $2(r-1)\ell < i \le 2(r-1)\ell + r - 1$ and $|\mathcal{F}_{i+1}| < Q$ then each path in $\mathcal{F}_i$ extends to at least $\log n$ paths in $\mathcal{F}_{i+1}$. In both cases, all leaves are not in $\mathcal{E}$.*

P2 $e(\mathcal{E}[S]) \le c|S|^{r-1} + 20rQ$.

P3 *If $|\mathbf{c}| < r-1$ we have $\deg_{\mathcal{E}}(\mathbf{c}, S) \le \xi^{r-|\mathbf{c}|}|S|^{r-1-|\mathbf{c}|} + 1$.*

P4 *We have mult$(\mathbf{c}) \le \xi^{r-|\mathbf{c}|}Q \cdot |S|^{-|\mathbf{c}|} \cdot \log^{|\mathbf{c}|} n + 1$.*

P5 *If $1 \le |\mathbf{c}| \le r-2$ we have $\deg_D(\mathbf{c}, S) \le (\xi'|S|)^{r-|\mathbf{c}|-1}$.*

Proof of Claim 4.4. Observe that $\mathcal{F}_0$ trivially satisfies the conditions of Claim 4.4, modulo Chernoff's inequality for **P1**. Suppose that for some $0 \le i < t$, at each step $0 \le j \le i$ of Algorithm 3 the conditions of Claim 4.4 are satisfied. In particular, by **P4**, the ends of the paths $\mathcal{F}_i$ are distinct as for $|\mathbf{c}| = r-1$ we have mult$(\mathbf{c}) < 2$, and by **P1** we have $|\mathcal{F}_i| \ge \min\left(\log^{i/2} n, Q\right)$.

To begin with, we show that $\mathcal{E}$ cannot have to many edges. At each step j with $1 \le j \le i$, we add $|\mathcal{F}_{j-1}|$ edges to $\mathcal{E}$, so that we want to upper bound $\sum_{j=1}^{t} |\mathcal{F}_{j-1}|$. Definitely $\mathcal{F}_t$ has size at most Q and $\mathcal{F}_{j-4(r-1)}$ always has size less than half of $\mathcal{F}_j$, so that this sum is dominated by $4r\sum_{i=1}^{\ell} 2^i$ where $\ell = \log_2 Q$. We conclude that $\sum_{j=1}^{t} |\mathcal{F}_{j-1}| \le 8rQ$. Since we create two fans, in total we obtain the claimed bound **P2**.

We now show that, for each choice of $P \in \mathcal{F}_i$ with end $\mathbf{a}$, the total number of vertices in T_{i+1} which are not good for $\mathbf{a}$ is at most $\delta|S|$. This will allow us to prove **P1**. First, since P has at most t vertices, at most t vertices are excluded by (i).

For each $\mathbf{c}$ of size at most $r-2$ with vertices chosen from $\mathbf{a}$, there are at most $2r\xi|S|$ vertices fulfilling (ii). To see this for $|\mathbf{c}| = 0$, observe that otherwise we have $e(\mathcal{E}[S]) > 2\xi^r|S|^{r-1} > 2c|S|^{r-1}$, contradicting **P2** as $Q \le \frac{1}{C}|S|^{r-1}$. Assume that it fails for some non-empty $\mathbf{c}$. Then there are more than $2r\xi|S|$ vertices $x \in T_{i+1}$ with

$$\deg_{\mathcal{E}}(\{\mathbf{c}, x\}, S) > \xi^{r-|\mathbf{c}|-1}|S|^{r-|\mathbf{c}|-2}$$

which implies that

$$\deg_{\mathcal{E}}(\mathbf{c}, S) > 2\xi^{r-|\mathbf{c}|}|S|^{r-|\mathbf{c}|-1}$$

in contradiction to **P3**.

Furthermore there are at most $2r\xi|S|$ vertices b fulfilling (iii) for each $\mathbf{c}$. Again for $|\mathbf{c}| = 0$ it is enough to note that there are at most Q paths in total and thus there are at most

$$\frac{Q}{\xi^{r-1}Q \cdot |S|^{-1} \cdot \log n} \le \xi|S|$$

vertices b with $\mathrm{mult}(b) > \xi^{r-1}Q \cdot |S|^{-1} \cdot \log n$. Now suppose $\mathbf{c}$ is not empty. Every path in $\mathcal{F}_{i+1}$ whose end contains $\{\mathbf{c}, b\}$ was constructed by the expansion of some path in $\mathcal{F}_i$ whose end contains $\mathbf{c}$. Note that every path expands at most by a factor of $\log n$ and by **P3** there are at most $\xi^{r-|\mathbf{c}|}Q \cdot |S|^{-|\mathbf{c}|}\log^{|\mathbf{c}|} n + 1$ paths in $\mathcal{F}_i$ whose end contains $\mathbf{c}$. If this bound is less than two, then there are at most $\log n$ vertices b with $\mathrm{mult}(\{\mathbf{c}, b\}) \geq 1$. Otherwise there are at most

$$\frac{2\xi^{r-|\mathbf{c}|}Q \cdot |S|^{-|\mathbf{c}|}\log^{|\mathbf{c}|+1} n}{\xi^{r-|\mathbf{c}|-1}Q \cdot |S|^{-|\mathbf{c}|-1}\log^{|\mathbf{c}|+1} n} = 2\xi|S|$$

vertices $x \in S_i$ with $\mathrm{mult}(\{\mathbf{c}, b\}) > \xi^{r-|\mathbf{c}|-1}Q \cdot |S|^{-|\mathbf{c}|-1} \cdot \log^{|\mathbf{c}|+1} n$.

Finally, we want to show that for each $\mathbf{c}$ there are at most $\xi'|S|$ vertices b in T_i which satisfy (iv). This is trivial for $D = \emptyset$, so we may assume that D is as given in Algorithm 2.

First suppose $|\mathbf{c}| = 0$. If a vertex b satisfies (iv), then it is in $(\xi'|S|)^{r-2}$ edges of D, so if there are $\xi'|S|$ such vertices then there are at least $(\xi'|S|)^{r-1}$ edges in D using vertices of T_i (note that edges of D only intersect T_i in one vertex). In other words, the number of blocked pairs $(\mathbf{a}, \mathbf{b})$ with $\mathbf{a} \in \mathcal{F}_t(\mathbf{u})$ and $\mathbf{b} \in S^{r-1}$ is at least

$$(\xi'|S|)^{r-1} \cdot \xi' Q \geq 2r \cdot 2^{2r}\xi|S|^{(r-1)} \cdot Q$$

using our choice of parameters (4.1). We conclude that there is a leaf $\mathbf{a}$ of $\mathcal{F}_t(\mathbf{u})$ that is in at least $2r \cdot 2^{2r}\xi|S|^{r-1}$ blocked pairs with tuples $\mathbf{b} \in S^{r-1}$. Fix this leaf. Now **P3** holds for $\mathbf{a}$, and we will show that this gives a contradiction. Consider the following property of tuples $\mathbf{b}$. For any sets A and B with vertices in $\mathbf{a}$ and $\mathbf{b}$ respectively, if $|A|+|B| = r-1$ then $A \cup B$ is not in $\mathcal{E}$, while if $|A|+|B| < r-1$ then we have $\deg_{\mathcal{E}}(A \cup B, S) \leq 2\xi^{r-|A|-|B|}|S|^{r-1-|A|-|B|}$. Trivially if $\mathbf{b}$ has the property, then $(\mathbf{a}, \mathbf{b})$ is not blocked. If $\mathbf{b}$ does not have the property, then let $B_{\mathbf{b}}$ be a set of minimal size witnessing the property's failure. Since $A \notin \mathcal{E}$ by **P1**, and by **P3**, we do not have $|B_{\mathbf{b}}| = 0$.

We now count the ways to create $\mathbf{b}$ which does not have the property. For this we choose vertices $b_1, \ldots, b_{r-1}$ one at a time until we create a witness $B \neq \emptyset$ that $\mathbf{b}$ cannot have the property. When we come to choose b_j, we have at most $|S|$ ways to choose it without creating a witness. If we are to choose b_j which witnesses the property's failure, then there are sets A and B' contained respectively in $\mathbf{a}$ and $\{b_1, \ldots, b_{j-1}\}$ such that $(A, B' \cup \{b_j\})$ fails the property. There are at most 2^{2r} choices for A and B'. Since (A, B') does not witness the property failing, by definition for each choice of A and B' there are at most $\xi|S|$ choices of b_j. Summing up, there are at most $r \cdot 2^{2r}\xi|S|^{r-1}$ tuples $\mathbf{b}$ which do not have the property. As all blocked pairs use a tuple from this set, this is the desired contradiction.

Now suppose $\mathbf{c}$ is a tuple for which there are at least $\xi'|S|$ vertices b satisfying (iv). In other words, there are more than $\xi'|S|$ vertices $b \in T_{i+1}$ with $\deg_D(\{\mathbf{c}, b\}, S) > (\xi'|S|)^{r-|\mathbf{c}|-2}$, which implies that

$$\deg_D(\mathbf{c}, S) > (\xi'|S|)^{r-|\mathbf{c}|-1}$$

in contradiction to **P5**.

Putting all this together we conclude that there are at most $\delta|S|$ vertices b such that $\mathbf{c}$ exists satisfying any one of the conditions (i)–(iv), as desired.

Now let $\mathbf{a}$ be a leaf of $\mathcal{F}_i$. We now reveal all r-sets containing $\mathbf{a}$ which were not revealed before and which use a vertex x of T_{i+1} which is good for $\mathbf{a}$. Let X be the number of edges $\{\mathbf{a}, x\}$ which appear. Then the expected value of X is at least $p(1-\delta)|T_{i+1}| \geq \frac{C}{20r}\log n$. Applying the Chernoff bound, Theorem 2.16, we get that $X < \frac{C}{40r}\log n$ with probability at most $2\exp(-C\log n/(240r)) \leq n^{-4r}$. Let us suppose that $X \geq \log n$. Then Algorithm 3 does not fail to create the required number of paths from $\mathbf{a}$. Taking a union bound over the at most $|S|^{r-1}t$ such events, we obtain the stated success probability of Claim 4.4.

It remains to prove that **P3**, **P4** and **P5** also hold in $\mathcal{F}_{i+1}(\mathbf{u})$. But this is immediate since we avoided choosing vertices which could cause their failure. □

Taking a union bound over the $2t$ steps, we conclude that with probability at most n^{-2r} there is a failure to construct either of the desired fans $\mathcal{F}_t(\mathbf{u})$ and $\mathcal{F}_t(\mathbf{v})$.

Connecting the fans

By construction, as set up in line 6 of Algorithm 3, all leaves of $\mathcal{F}_t(\mathbf{v})$ are not edges of D and thus not dangerous. Let L be the leaves from $\mathcal{F}_t(\mathbf{u})$ and L' the leaves from $\mathcal{F}_t(\mathbf{v})$ reversed. We now want to reveal more edges to connect a leaf from L with one from L'.

For $\mathbf{a} \in L$ and $\mathbf{b} \in L'$ let P be the tight path with $r-1$ edges on the vertices $(\mathbf{a}, \mathbf{b})$. There are $|L'| \cdot (1-\xi')|L| = (1-\xi')Q^2$ many such paths P, which are not blocked, because $\mathbf{b}$ is not dangerous. Let $\mathcal{P}$ be the set of all these paths which are not blocked.

Let I_P be the indicator random variable for the event that the path P appears, which occurs with probability p^{r-1}. Further let X be the random variable counting the number of paths which we obtain and note $X = \sum_{P\in\mathcal{P}} I_P$. With Janson's inequality, Theorem 2.18, we want to bound the probability that $X = 0$. First, let us estimate the expected value of X. By the observation from above, we have $\mathbb{E}[X] = |\mathcal{P}|p^{r-1} \geq (1-\xi')(cC)^{r-1}\log^{r-1} n \geq \log n$.

Now consider two distinct paths $P = (\mathbf{a}, \mathbf{b})$ and $P' = (\mathbf{a}', \mathbf{b}')$, which share at least one edge. It follows from property **P4** of Claim 4.4 and the quantities Q and $|S|$, that two paths are identical if they share at least $r/2$ vertices in their end tuple. Since either the start or end $r/2$-tuple of one of the $(r-1)$-tuples from P has to agree with P', we can assume without loss of generality that $\mathbf{a} = \mathbf{a}'$. Furthermore, we can assume that for some $1 \leq j < r/2$, $\mathbf{b}$ and $\mathbf{b}'$ agree on the first j entries, but not in the $(j+1)$-st. They can not share another $r/2$ or more entries as this would imply $\mathbf{b} = \mathbf{b}'$. Thus P and P' share precisely an interval of length $r-1+j$ and thus j edges. With this we can bound $\mathbb{E}[I_P I_{P'}] \leq p^{2r-2-j}$.

Let $N_{P,j}$ be the number of paths P' such that P and P' share precisely j edges. The above shows that for fixed $P = (\mathbf{a}, \mathbf{b})$, $N_{P,j}$ is at most the number of choices of leaves $\mathbf{b}' \in L'$ such that $\mathbf{b}$ and $\mathbf{b}'$ only differ in the ending $(r-1-j)$-tuple, plus the number of choices of leaves $\mathbf{a}' \in L$ such that $\mathbf{a}$ and $\mathbf{a}'$ only differ in the start $(r-1-j)$-tuple. It follows from property **P4** of Claim 4.4, that the start j-tuple of $\mathbf{b}'$ and the end j-tuple of $\mathbf{a}'$ are the ends of at most $\xi^{r-j}Q \cdot |S|^{-j}\log^j n + 1$ many paths. This implies that $N_{P,j} \leq Q \cdot |S|^{-j}\log^j n$, because $j < r/2$.

We can now obtain for $P, P' \in \mathcal{P}$

$$\delta = \sum_{P\sim P'} \mathbb{E}[I_P I_{P'}] = \sum_{P\in\mathcal{P}} \sum_{1\leq j<r/2} \Big(\sum_{|P'\cap P|=j} \mathbb{E}[I_P I_{P'}] \Big).$$

With the above we get

$$\begin{aligned}\delta &\leq \sum_{P\in\mathcal{P}} \sum_{1\leq j<r/2} N_{P,j} \cdot p^{2r-2-j} \\ &\leq |\mathcal{P}|^2 p^{2r-2} \sum_{1\leq j<r/2} |\mathcal{P}|^{-1} \cdot Q \cdot |S|^{-j} \log^j n \cdot p^{-j} \\ &\leq \mathbb{E}[X]^2 \cdot 2Q^{-1} \sum_{1\leq j<r/2} 3C^{-j} \leq \mathbb{E}[X]^2 C^{-1} \log^{-1} n,\end{aligned}$$

where we used that $|S| \geq Cp^{-1}\log n$ and $Q \geq \log n$. Hence, Theorem 2.18 implies that $\mathbb{P}[X=0] \leq \exp(-\mathbb{E}[X]^2/(E[X]+\delta)) \leq \exp(-\frac{C}{6}\log n)$. Thus we find some connection with probability at least $1-n^{-2r}$.

But we do not want to reveal all the $O(Q^2)$ edges for all paths from $\mathcal{P}$, since this would add way to many edges to the exposure hypergraph $\mathcal{E}$. The above argument proves that it is very likely that the desired connecting path exists and we will argue how to find such a path in an *economic* way. We find it by the following procedure. First, we reveal all the edges at each leaf in L and L'. This entails adding $2Q$ edges to $\mathcal{E}$ and if $r=3$ then we are already done and we have added $2Q \leq |S|$ edges to $\mathcal{E}$.

For $r \geq 4$ we then construct from each leaf of L all possible tight paths in S with $\lfloor (r-2)/2 \rfloor$ edges and similarly from each leaf of L' all tight paths of length $\lfloor (r-3)/2 \rfloor$. We do this by the obvious breadth-first search procedure, revealing at each step all edges at the end of each currently constructed path with less than $\lfloor (r-2)/2 \rfloor$ (or $\lfloor (r-3)/2 \rfloor$ respectively) edges which have not so far been revealed and adding each end to $\mathcal{E}$. Trivially, if the desired path exists, then two of these constructed paths will link up, so that this procedure succeeds in finding a connecting path with probability $1-n^{-2r}$.

The expected number of edges in S containing any given $(r-1)$-set in S is $p(|S|-r+1)$, which is between $\frac{C}{2}\log n$ and $C\log n$. Thus by Chernoff's inequality and the union bound, with probability at least $1-n^{-3r}$ no such $(r-1)$-set is in more than $2C\log n$ edges contained in S. It follows that the number of edges we add to $\mathcal{E}$ in this procedure is with probability at least $1-n^{-3r}$ not more than

$$\begin{aligned}2Q \sum_{i=0}^{\lfloor (r-2)/2 \rfloor} (2C\log n)^i &\leq 2p^{-(r-1)/2} \log n \cdot r(2C\log n)^{(r-2)/2} \\ &= O\left(p^{-(r-2)} \log^{r-2} n\right) = O(|S|^{r-2}),\end{aligned}$$

for $r \geq 4$. Putting this together with property **P2** of Claim 4.4 we see that the final exposure graph $\mathcal{E}'$ has at most $O(|S|^{r-2})$ edges more than $\mathcal{E}$, as desired.

Probability and runtime

Altogether we have that our algorithm for the Connecting Lemma fails with probability at most $n^{-2r} + n^{-2r} + n^{-3r} \leq n^{-5}$.

We now estimate the running time of our algorithm. In total we added $O(|S|^{r-2})$ many $(r-1)$-tuples to $\mathcal{E}$. For every $(r-1)$-tuple exposed, we have to go through at most n vertices until we find all new edges. This gives at most $O(n^{r-1})$ steps. We can easily keep track of the bounds for Claim 4.4 and update them after each event. Since there is nothing else to take care of, we have a total number of at most $O(n^{r-1})$ steps.

Spike path version

The statement of the lemma is almost the same as for the tight path version, Lemma 4.2.

Lemma 4.5 (Spike path Lemma)**.** *For each $r \geq 3$ there exist $c, C > 0$ and a deterministic $O(n^{r-1})$-time algorithm whose input is an n-vertex r-uniform hypergraph G, a pair of distinct $(r-1)$-tuples $\mathbf{u}$ and $\mathbf{v}$, a set $S \subset V(G)$ and a $(r-1)$-uniform exposure hypergraph $\mathcal{E}$ on the same vertex set. The output of the algorithm is either **Fail** or a spike path of even length $o(\log n)$ in G whose ends are $\mathbf{u}$ and $\mathbf{v}$ and whose interior vertices are in S, and an exposure hypergraph $\mathcal{E}' \supset \mathcal{E}$. We have $e(\mathcal{E}') \leq e(\mathcal{E}) + O(|S|^{r-2})$ and all the edges $E(\mathcal{E}') \setminus E(\mathcal{E})$ are contained in $S \cup \mathbf{u} \cup \mathbf{v}$.*

*Suppose that G is drawn from the distribution $\mathcal{H}^{(r)}(n,p)$ with $p \geq C(\log n)^3/n$, that $\mathcal{E}$ does not contain any edges intersecting both S and $\mathbf{u} \cup \mathbf{v}$. If furthermore we have $|S| = Cp^{-1}\log n$ and $|e(\mathcal{E}[S])| \leq c|S|^{r-1}$ then the algorithm returns **Fail** with probability at most n^{-5}.*

Sketch proof. We modify the proof of Lemma 4.2 in the following simple ways. First, we will maintain fans of spike paths rather than tight paths, and we change Algorithm 3 line 5 so that the tuple $\mathbf{a}$ to be extended is the (unique) one whose extension continues to give us a spike path. Note that whenever we have a spike path ending in $\mathbf{a}$ and we extend the spike path by adding one vertex b then the end of the new spike path is an $(r-1)$-set whose vertices are contained in $(\mathbf{a}, b)$ (though in general not the last $r-1$ vertices nor in the same order). This is all we need to make our analysis of the fan construction work; it is not necessary to change anything in this part of the proof or the constants. Second, when we come to connect fans, we let L be the reverses of the end tuples of $\mathcal{F}_t(\mathbf{u})$ and L' be the end tuples of $\mathcal{F}_t(\mathbf{v})$, and (again) look for a tight path connecting a tuple in L to one in L'. This has no effect on the proof that a connecting path from some member of L to some member of L' exists, and the result is the desired spike path. The resulting spike path is of even length as both fans have the same size. □

4.4 Proof of the Reservoir Lemma

Idea

The reservoir path P_{res} will consist of absorbing structures (each *carrying* one vertex from R). More precisely, these absorbing structures can be seen as small reservoir path with reservoir of cardinality 1. Each of these small absorbers consists of a cyclic spike path plus the reservoir vertex, where pairs of spikes are additionally connected with tight paths (cf. Figure 4.1).

First choose the reservoir set R and disjoint sets U_1, U_2 and U_3. For every vertex in R we will reveal the necessary path segment in U_1. From the endpoints of these path we fan out and also close the backbone structure of the reservoir inside U_2. Finally, we use U_3 and Lemma 4.2 to get the missing connections in the reservoir structures and connect all structures to one path P_{res}. In each step the relevant edges of the exposure graph $\mathcal{E}$ are solely coming from the same step.

Proof

We arbitrarily fix the reservoir set R of size $2Cp^{-1}\log n$ and disjoint sets U_1, U_2 and U_3 of the same size such that $S = R \cup U_1 \cup U_2 \cup U_3$ is of size $\frac{n}{4}$. First we want to build the absorbing structures for every $a \in R$, which have size roughly $t^2 = o(\log^2 n)$. There is a sketch of this structure for some $a \in R$ in Figure 4.1.

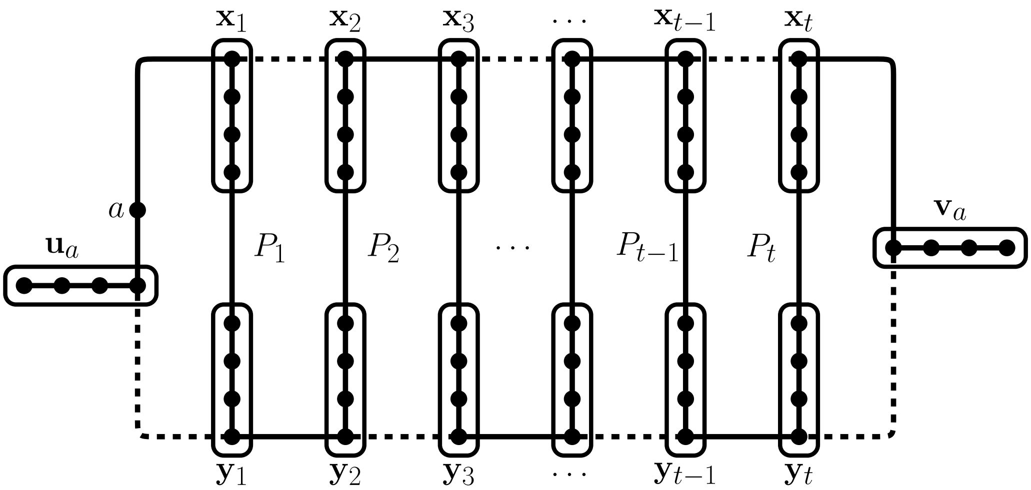

Figure 4.1: Illustration of the absorber for one vertex $a \in R$ and $r = 5$ with the path, which contains the vertex a.

So we fix $a \in R$. We want to construct the following tight path on $2r - 1$ vertices containing a in the middle. The end tuples are $\mathbf{x}_1 = (x_1, \ldots, x_{r-1})$ and $\mathbf{u}_a = (u_1, \ldots, u_{r-1})$ and together with a we require that all the edges $\{x_{r-j}, \ldots, x_1, a, u_1, \ldots, u_{j-1}\}$ are present for $j = 1, \ldots, r$. We build this path by first choosing $x_1, \ldots, x_{r-2}$ arbitrarily from U_1. Then we expose all edges containing $\{x_1, \ldots, x_{r-2}, a\}$ to get x_{r-1}. We continue by exposing all edges containing the $(r-1)$ set $\{x_{r-j-1}, \ldots, x_1, a, u_1, \ldots, u_{j-1}\}$ to get u_j for $j = 1, \ldots, r-1$. The probability that in any of these cases we fail to find a new vertex inside a subset of U_1 of size at least $|U_1|/2$ is at most n^{-5} by Chernoff's inequality. A union bound over all r edges and over all $a \in R$ reveals that with probability at most n^{-3} we fail to construct the small starting graph for any a.

Recall that when adding edges, we always expose all edges containing one $(r-1)$-tuple and then add this to $\mathcal{E}$. All exposed $(r-1)$-tuples from this step are contained in $U_1 \cup R$ and none of them contains more than one vertex from R. Furthermore we did at most $O(|R| \cdot |U_1|) = O(n^2)$ many steps so far.

Now we want to build the absorbing structure for a. We partition each of U_2 and U_3 into parts of size $Cp^{-1}\log n$ (plus perhaps a smaller left-over set). We apply Lemma 4.5 to the $(r-1)$-tuples $\overleftarrow{\mathbf{x}_1}$ and $\overleftarrow{\mathbf{u}_a}$ and connect them with a spike path of even length $2t+2$ in some part of U_2, with $t = o(\log n)$. At each step we use a part of U_2 in which we have so far built the least spike paths for the application of Lemma 4.5, which is necessary to control the edges of $\mathcal{E}$ within this set. We use U_2 as both tuples are contained in U_1 and thus we have no problem with edges from $\mathcal{E}$ intersecting both U_2 and the end tuples. Let the spikes after $\mathbf{x}_1$ and $\mathbf{u}_a$ be called $\mathbf{x}_2, \ldots, \mathbf{x}_t$ and $\mathbf{y}_1, \ldots, \mathbf{y}_t$ respectively. The last remaining spike opposite of $\mathbf{u}_a$ we call $\mathbf{v}_a$. We apply Lemma 4.2 to find paths P_i connecting the tuples $\mathbf{x}_i$ and $\mathbf{y}_i$ for $i = 1, \ldots, t$ in a part of U_3. Again, we choose a part of U_3 which was used for building the least connecting paths so far. We use parts of U_3 for these connections because all the spikes are contained in $U_1 \cup U_2$ and thus there are no edges of $\mathcal{E}$ intersecting U_3 and the spikes. This finishes the absorbing structure for a. It has end-tuples $\mathbf{u}_a$ and $\mathbf{v}_a$.

To finish P_{res} we enumerate the vertices in R increasingly $a_1, \ldots, a_{|R|}$. Then we use Lemma 4.2 repeatedly, again at each step using a part of U_3 which has been used least often previously, to connect the tuples $\mathbf{v}_{a_i}$ to $\mathbf{u}_{a_{i+1}}$ for $i = 1, \ldots, |R| - 1$ with tight paths. Thus, we have obtained the path P_{res} with end tuples $\mathbf{u} = \mathbf{u}_{a_1}$ and $\mathbf{v} = \mathbf{v}_{a_{|R|}}$.

The absorbing works in the following way for the structure of a single vertex $a \in R$. It relies on the fact, that the paths P_i can be traversed in both directions and that we can walk from any spike to its neighbouring spikes using a tight path. The path which uses a (Figure 4.1) starts with $\mathbf{u}_a$, goes through a to $\mathbf{x}_1$ and then uses the path P_1 to $\mathbf{y}_1$. From there it goes via a tight path to $\mathbf{y}_2$ and uses P_2 to go back to $\mathbf{x}_2$. Going from $\mathbf{x}_i$ via path P_i to $\mathbf{y}_i$ and back from $\mathbf{y}_{i+1}$ through P_{i+1} to $\mathbf{x}_{i+1}$ for $i = 2, \ldots, t-1$ the path ends up in $\mathbf{v}_a$ and uses all vertices. To avoid a (Figure 4.2) the path starting in $\mathbf{u}_a$ goes immediately to $\mathbf{y}_1$, then uses the path P_1 to go to $\mathbf{x}_1$. Alternating as above and traversing all the paths P_i in opposite direction we again end up in $\mathbf{v}_a$ and used all vertices but a.

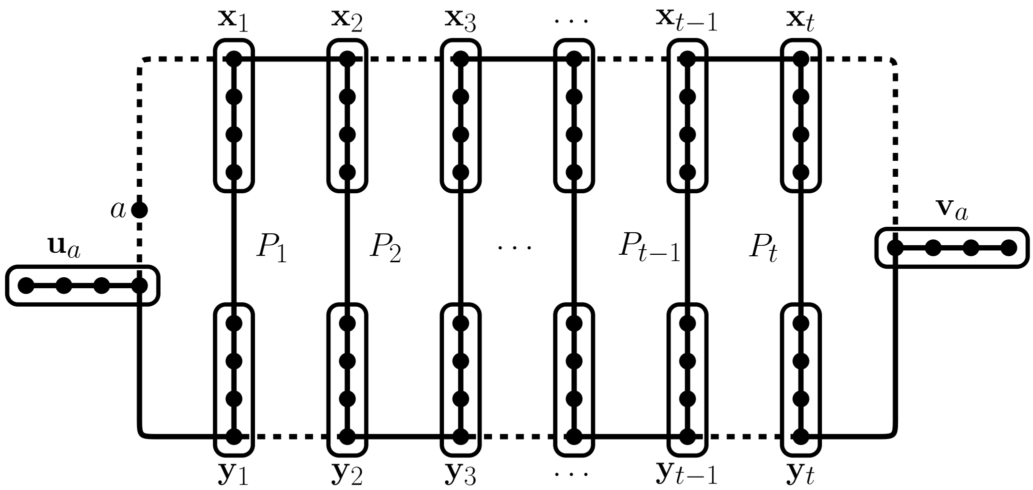

Figure 4.2: Illustration of the absorber for one vertex $a \in R$ and $r = 5$ with the path, which does not contain the vertex a.

For the proof of the lemma, it remains to check that we obtain the right probability and we are indeed able to apply Lemma 4.2 and 4.5 as we described. It is immediate from the construction, that no edges of $\mathcal{E}$ are contained in $R \cup \mathbf{u} \cup \mathbf{v}$.

In total we are performing $|R|$ many connections with spike-paths and $|R| \cdot t + |R| - 1$ many connections with tight-paths. Thus altogether we have $o\left(p^{-1}\log^2 n\right)$ executions of Lemma 4.2 and Lemma 4.5. In each application we add $O\big(Cp^{-1}\log n\big)^{r-2}$ edges to $\mathcal{E}$ in some part of U_2 or U_3. Since each part initially contains no edges of $\mathcal{E}$, provided a given part has been used at most p^{-1} times the total number of edges of $\mathcal{E}$ in it is $o\big(Cp^{-1}\log n\big)^{r-1}$, and therefore we can apply Lemma 4.2 or 4.5

at least one more time with that part. Since $|U_2|$ and $|U_3|$ are of size linear in n, they each contain $\Omega\big(pn/\log n\big)$ parts. Thus, we can perform in total $\Omega(n/\log n) = \Omega\big(p^{-1}\log^2 n\big)$ applications of either Lemma 4.2 or Lemma 4.5 before all parts have been used p^{-1} times and thus might acquire too many edges of $\mathcal{E}$. Since we do not need to perform that many applications, we conclude that the conditions of each of Lemma 4.2 and Lemma 4.5 are met each time we apply them.

Since the connecting lemma fails with probability at most n^{-5} the construction of this absorber fails with probability at most n^{-3}. In every connection there are at most $O(n^{r-1})$ steps performed and thus we need $o(n^{r-1}p^{-1}\log^2 n) = O(n^r)$ many steps for the construction of the absorber. $\square$

Chapter 5

Randomly perturbed graphs

Now we come to the paper with Böttcher, Montgomery, and Person [32] and the proof[32] of Theorem 2.7. We first give a brief outline of the steps and explain a decomposition result from Ferber, Luh, and Nguyen [54], which we will use. Then the proof of Theorem 2.7 is presented in Section 5.2, with the proofs of some auxiliary lemmas postponed to Section 5.3.

5.1 Overview of the proof

Step 1. We first obtain an *almost spanning embedding* of all but εn vertices of F, using only the edges of the random graph $\mathcal{G}(n, p)$. For this we adapt the strategy of Ferber, Luh, and Nguyen [54] to decompose F, and embed it using the theorem of Riordan [97] (Theorem 2.1) together with Janson's inequality (Theorem 2.18). A major difference to previous methods is that we do not choose which large subgraph of F to embed, only seeking to embed *some* almost spanning subgraph of F which covers the sparser parts of F.

Step 2. A key part in the remainder of our proof is obtaining a *reservoir set*. This will enable us to complete the partial embedding to an embedding of all of F. The idea behind such a reservoir set is as follows. The reservoir set will contain vertices already covered in the partial embedding of F obtained in the first step. The properties of the reservoir allow us to reuse some of these vertices for embedding new F-vertices later in the proof, and *swap* the image of F-vertices already embedded there to some other vertex in $G_\alpha \cup \mathcal{G}(n, p)$ that was not used in the embedding so far. For these swaps we crucially use the deterministic graph G_α, and that in the first step we did not use G_α but only $\mathcal{G}(n, p)$.

Reservoir structures of similar nature were used for embedding tight Hamilton cycles in random hypergraphs in [4], cycle-powers in random graphs in [84], bounded degree trees in random graph in [90]. However, we use the interplay of the random and deterministic graphs in a new way to create our reservoir structure.

Step 3. Using additional edges of G_α and $\mathcal{G}(n, p)$, we then *complete the embedding* of F, utilising the reservoir. The approach for this completion again uses ideas from [54], relying on Janson's inequality and the Hall-type matching argument for hypergraphs by Aharoni and Haxell [2] (Theorem 2.20). The use of edges from G_α in this step is crucial in gaining the $\log$-term in comparison to p_Δ.

[32]The proof given here is similar the on in [32].

Decomposition of bounded degree graphs

As outlined above we will adapt much of the embedding strategy of Ferber, Luh, and Nguyen [54]. We therefore briefly sketch their approach here, while mentioning the tools we need. In [54], each graph $F \in \mathcal{F}(n, \Delta)$ is decomposed into a sparse part and many dense spots. For this recall that $\gamma(H) := \max\{e(H')/(v(H') - 2)) \colon H' \subseteq H \text{ and } v(H') \geq 3\}$, and call a graph G *dense* if $\gamma(H) > \frac{\Delta+1}{2}$ and *sparse* otherwise.

Say that two graphs $S, S' \subseteq F$ are *isomorphic in* F, if there exist labellings $V(S) = \{v_1, \ldots, v_s\}$ and $V(S') = \{v'_1, \ldots, v'_s\}$, such that $v_j \mapsto v'_j$ is an isomorphism between S and S' and, for each $1 \leq j \leq r$, $|N_F(v_j) \setminus V(S)| = |N_F(v'_j) \setminus V(S')|$.

Definition 5.1 (ε-good decomposition)**.** *Let $\varepsilon > 0$, $F \in \mathcal{F}(n, \Delta)$ and let $\mathcal{S}_1, \ldots, \mathcal{S}_k$ be families of induced subgraphs of F. For $F' = F - (\bigcup_h \bigcup_{S \in \mathcal{S}_h} V(S))$ we say that $(F', \mathcal{S}_1, \ldots, \mathcal{S}_k)$ is an ε-*good decomposition *if the following hold.*

P1 *F' is sparse, that is, $\gamma(F') \leq \frac{\Delta+1}{2}$.*

P2 *Each $S \in \bigcup_h \mathcal{S}_h$ is minimally dense, that is, $\gamma(S) > \frac{\Delta+1}{2}$ and S' is sparse for all $S' \subseteq S$ with $3 \leq v(S') < v(S)$.*

P3 *For each $1 \leq h \leq k$, all the graphs in $\mathcal{S}_h$ are isomorphic in F.*

P4 *Every $\mathcal{S}_h$ contains graphs on at most εn vertices, that is $|\bigcup_{S \in \mathcal{S}_h} V(S)| \leq \varepsilon n$.*

P5 *All the graphs in $\bigcup_i \mathcal{S}_i$ are vertex disjoint and, for each $1 \leq h \leq k$ and $S, S' \in \mathcal{S}_h$ with $S \neq S'$, there are no edges between S and S' in F, and S and S' share no neighbours in F.*

We call the graphs in $\mathcal{S}_1, \ldots, \mathcal{S}_k$ the dense spots *of the decomposition.*

An ε-good decomposition can easily be found using a greedy algorithm. The following lemma is proved in [54].

Lemma 5.2 (Lemma 2.2 in [54])**.** *For each $\varepsilon > 0$ and $\Delta > 0$, there exists some k_0 such that, for each $F \in \mathcal{F}(n, \Delta)$, there is some $k \leq k_0$ and an ε-good decomposition $(F', \mathcal{S}_1, \ldots, \mathcal{S}_k)$ of F.*

The sparse part F' can be embedded using Theorem 2.1, the result of Riordan [97]. The embedding of F' is then extended in [54] step by step to include the graphs in $\mathcal{S}_h$, for $1 \leq h \leq k$. This can be done using Janson's inequality (Theorem 2.18) and a hypergraph analog of Hall's theorem due to Aharoni and Haxell [2] (Theorem 2.20). These are the tools used to prove Theorem 2.4. To construct our reservoir structure in the second step of our proof, we will also need the concentration inequalities Theorem 2.16 and 2.17.

5.2 The embedding

Let $\alpha > 0$ and $\Delta \geq 5$, and take $\varepsilon = (\frac{\alpha}{4\Delta})^{2\Delta}$. Let k_0 be large enough for the result of Lemma 5.2 to hold with ε and Δ. Let $F \in \mathcal{F}(n, \Delta)$, and, for some $k \leq k_0$, using the property from Lemma 5.2, let $(F', \mathcal{S}_1, \ldots, \mathcal{S}_k)$ be an ε-good decomposition of F. For each $1 \leq h \leq k$, let s_h be the size of the graphs in $\mathcal{S}_h$ (possible by **P3**), and, picking some representative $S \in \mathcal{S}_h$, note that, by **P2** and as $\Delta(S) \leq \Delta$, we have

$$(\Delta+1)(s_h-2) < 2e(S) \leq \Delta s_h,$$

so that $s_h < 2\Delta+2$. Thus, we may consider α, Δ, ε, $k \leq k_0$, and the maximum size of each dense spot $(2\Delta+1)$ to be constant, while n tends to infinity.

We will expose the graph $\mathcal{G}(n,p)$ in a total of $2k+1$ rounds, revealing $G_h = G(n,q)$ for $0 \leq h \leq k$ and $G'_h = G(n,q)$ for $1 \leq h \leq k$, where $(1-q)^{2k+1} = 1-p$ and thus $q = \Theta(p)$. Every edge is thus present with probability p in $(\cup_h G_h) \cup (\cup_h G'_h)$. We use $G_0, \ldots, G_k$ to embed most of F while embedding all of F', and then use $G'_1, \ldots, G'_k$ to finish the embedding. We let

$$G := \bigcup_h G_h\,.$$

Embedding most of the graph

Our goal here is to embed all but at most εn vertices of the graph F while ensuring we embed all the vertices in F'. Since, by **P1**, $\gamma(F') \leq \frac{\Delta+1}{2}$, and thus $q = \omega(n^{-\frac{1}{\gamma(F')}})$, by Theorem 2.1, we can almost surely embed F' into G_0. Let $f_0 : V(F') \to V(G_0)$ be such an embedding and let $F'_0 = F'$.

For $1 \leq h \leq k$, we want to (almost surely) use edges from G_h to extend the embedding f_{h-1} to cover all but at most $\frac{\varepsilon n}{s_h^2 k}$ graphs from $\mathcal{S}_h$. We then let f_h be the extended embedding and let F'_h be the subgraph of F embedded by f_h. We use the following lemma, which allows us to extend the current embedding to one more dense spot $S \in \mathcal{S}_h$, even if we restrict its image to small but linearly sized set U, using only random edges of G_h. This lemma is proved along with the other lemmas from this section in Section 5.3.

Lemma 5.3. *For each $1 \leq h \leq k$, the following holds a.a.s. for any $\mathcal{S} \subseteq \mathcal{S}_h$ and $U \subseteq V(G_\alpha)$ with $|\mathcal{S}| \geq \frac{\varepsilon n}{s_h^2 k}$ and $|U| \geq \frac{\varepsilon n}{s_h k}$. There is some $S \in \mathcal{S}$ and a copy S' of S in $G_h[U]$ with an embedding $\pi : V(S) \to V(S')$ such that, for each $v \in V(S)$,*

$$f_{h-1}(N_F(v) \cap V(F'_{h-1})) \subseteq N_{G_h}(\pi(v)). \tag{5.1}$$

So, we start with f_0 and F'_0. For each $1 \leq h \leq k$, we construct f_h and F'_h, as follows. The property in Lemma 5.3 almost surely holds for h. We extend the embedding f_{h-1} to f_h using edges from G_h to cover as many of the graphs in $\mathcal{S}_h$ as possible (with any edges to F'_{h-1} correctly embedded), and call the resulting graph F'_h. By Lemma 5.3 this leaves at most $\frac{\varepsilon n}{s_h^2 k}$ graphs in $\mathcal{S}_h$ unembedded. Indeed, if there is a set $\mathcal{S}$ of at least $\frac{\varepsilon n}{s_h^2 k}$ unembedded graphs in $\mathcal{S}_h$, then, let $U = V(G_\alpha) \setminus V(F'_h)$ and note that $|U| \geq s_h \cdot |\mathcal{S}| \geq \frac{\varepsilon n}{s_h k}$. There then exists some $S \in \mathcal{S}$ and a copy S' of S in $G_h[U]$ with isomorphism $\pi : V(S) \to V(S')$ such that (5.1) holds for each $v \in V(S)$. As, by **P5**, no two subgraphs in $\mathcal{S}_h$ have an edge between them, π can be used to embed S and extend the embedding f_h, a contradiction.

From this we (almost surely) obtain the embedding f_k of a subgraph of F, covering F' and all but at most $\frac{\varepsilon n}{s_h^2 k}$ graphs from each S_h, $1 \leq h \leq k$, into $G = \bigcup_h G_h$.

Independence of f_k and G_α

Until now, we have only used random edges for our embedding, and, therefore, we can consider this embedding to be independent of the edges of G_α, which we make precise as follows. Let $\mathcal{F}$ be the set of possible induced subgraphs of F which cover F' and, for each $1 \le h \le k$, all but at most $\frac{\varepsilon n}{s_h^2 k}$ of the graphs from $\mathcal{S}_h$. Let r be the number of graph isomorphism classes in $\mathcal{F}$, and let $F_1^*, \dots, F_r^* \in \mathcal{F}$ be representatives of each class. In Section 5.2 we proved that $\mathbb{P}[\exists i$ with some copy of F_i^* in $G] = 1 - o(1)$.

For each $1 \le i \le r$, let E_i be the event that there is a copy of F_i^* in G, but no copy of F_j^* for any $j < i$. For each $1 \le i \le r$, let $\hat{F}_i$ be a random copy of F_i^* in G with vertex set in $V(G_\alpha)$. Then, we have,

$$\mathbb{P}\Big[\exists \text{ an } F\text{-copy in } G_\alpha \cup G \cup \Big(\bigcup_h G'_h\Big) | E_i\Big] \ge \mathbb{P}\Big[\exists \text{ an } F\text{-copy in } G_\alpha \cup \hat{F}_i \cup \Big(\bigcup_h G'_h\Big)\Big],$$

since G is independent of $G_\alpha \cup (\bigcup_h G'_h)$ and on the left hand side, by conditioning on E_i, G contains some copy of F_i^*. The second part of the definition of E_i is to ensure that $\sum_{i=1}^r \mathbb{P}[E_i]$ equals the probability that there exists and i with some copy of F_i^* in G. We will show in Section **??**, for each $1 \le i \le r$, that

$$\mathbb{P}\Big[\exists \text{ an } F\text{-copy in } G_\alpha \cup \hat{F}_i \cup \Big(\bigcup_h G'_h\Big)\Big] = 1 - o(1)\,. \tag{5.2}$$

It follows that,

$$\begin{aligned}
&\mathbb{P}\Big[\exists \text{ an } F\text{-copy in } G_\alpha \cup G \cup \Big(\bigcup_h G'_h\Big)\Big] \\
&\qquad \ge \sum_{i=1}^r \mathbb{P}\Big[\exists \text{ an } F\text{-copy in } G_\alpha \cup G \cup \Big(\bigcup_h G'_h\Big) | E_i\Big] \cdot \mathbb{P}[E_i] \\
&\qquad \ge \sum_{i=1}^r \mathbb{P}\Big[\exists \text{ an } F\text{-copy in } G_\alpha \cup \hat{F}_i \cup \Big(\bigcup_h G'_h\Big)\Big] \cdot \mathbb{P}[E_i] = (1 - o(1)) \cdot \sum_{i=1}^r \mathbb{P}[E_i] \\
&\qquad = (1 - o(1)) \cdot \mathbb{P}\Big[\exists i \text{ with some copy of } F_i^* \text{ in } G\Big] = 1 - o(1).
\end{aligned}$$

Hence, it remains to show (5.2). Before we can turn to this, we first need to prepare our reservoir structure.

Preparing the reservoir

Let us fix an arbitrary $F^* \in \mathcal{F}$ and let $\hat{F}$ be a random copy of F^* in G with vertices in $V(G_\alpha)$. Let g_0 be the embedding of F^* to $\hat{F}$. For each $1 \le h \le k$, let $\mathcal{S}'_h \subseteq \mathcal{S}_h$ be those dense parts not in F^*, so that $|\mathcal{S}'_h| \le \frac{\varepsilon n}{s_h^2 k}$. We have, for each $1 \le h \le k$, that the graphs in $\mathcal{S}'_h$ are isomorphic in F, minimally dense, disjoint, neither have edges between them nor share any neighbours. Furthermore, the sets in $\{V(F^*)\} \cup \{V(S) : S \in \mathcal{S}'_h, 1 \le h \le k\}$ form a partition of $V(F)$. Note that $|V(F) \setminus V(F^*)| \le \varepsilon n$.

Let $V_0 \subseteq V(F^*)$ be a maximal independent set in F^* of vertices with no neighbours in $V(F) \setminus V(F^*)$ in F. Note that $|V_0| \ge (|F^*| - \Delta|V(F) \setminus V(F^*)|)/(\Delta + 1) \ge n/2\Delta$ and let $W_0 := g_0(V_0) \subseteq V(\hat{F})$. For each vertex $v \in V(G_\alpha)$, let $B(v) \subseteq W_0$ be the set of vertices $w \in W_0$ such that every neighbour of w in

$\hat{F}$ is also a neighbour of v in G_α. That is,

$$B(v) = \{w \in W_0 : N_{\hat{F}}(w) \subset N_{G_\alpha}(v)\}. \tag{5.3}$$

We shall later, in the proof of Lemma 5.4, show that each $B(v)$ is a random set, which entails that it has properties convenient for our embedding strategy. Crucially, by the definition of$B(v)$ we can *switch* any vertex from $B(v)$ in $\hat{F}$ with v and still get a copy of F^*. Hence, these sets form our *reservoir structure*. Furthermore, the sets $B(v)$ all lie in W_0, a large independent set in $\hat{F}$, hence their preimages need no further neighbours added to extend the embedding g_0 to F. Therefore, we can switch different vertices in this manner without creating conflicts. The following lemma states that almost surely u has linearly many G_α-neighbours in $B(v)$ for each $u, v \in V(G_\alpha)$. This, in particular, implies that almost surely for each $v \in V(G_\alpha)$ the set $B(v)$ is linear in size.

Lemma 5.4. *A.a.s., for each $u, v \in V(G_\alpha)$ we have $|N_{G_\alpha}(u) \cap B(v)| \geq 4\varepsilon n$.*

Again, we defer the proof of Lemma 5.4 to Section 5.3. We note that it is not difficult, but constitutes a crucial new idea of our proof. It relies on the fact that the sets $B(v)$ are random sets.

Let us now briefly indicate how we will use the reservoir structure in the next subsection to finish the embedding. Essentially, we pair each vertex $w \in V(F) \setminus V(F^*)$ with a different vertex, v_w say, in $V(G_\alpha) \setminus V(\hat{F})$. Then, we 'embed' each $w \in V(F) \setminus V(F^*)$ to some vertex $z_w \in B(v_w)$, where it is important here that $B(v_w)$ has linear size. The only problem is that z_w already has a vertex embedded to it, but by switching z_w out of $\hat{F}$ and replacing it with v_w, we can shift part of the original copy of F^* to fix this. If we ensure that the vertices z_w with $w \in V(F) \setminus V(F^*)$ are distinct then these switchings can be carried out simultaneously, completing the embedding.

Finishing the embedding

We want to show that $\mathbb{P}[\exists$ a copy of F in $G_\alpha \cup \hat{F} \cup (\bigcup_h G'_h)] = 1 - o(1)$, which is precisely (5.2) and completes the proof of Theorem 2.7. We will follow the approach of Ferber, Luh, and Nguyen [54] and use our reservoir structure as outlined above. Roughly speaking, in comparison to [54], the minimum degree condition into the sets $B(v)$ guaranteed by Lemma 5.4 gives us one edge between $\hat{F}$ and each dense spot *for free*, allowing the use of a lower edge probability in our result.

So, let $F_0 = F^*$ and recall that we have an embedding $g_0 : F_0 \to G_\alpha \cup G$. Now, for each $0 \leq h \leq k$ let $F_h = F[V(F^*) \cup (\bigcup_{h' \leq h} \bigcup_{S \in \mathcal{S}_{h'}} V(S))]$. Noting that $|V(G_\alpha) \setminus V(\hat{F})| = |\bigcup_h \bigcup_{S \in \mathcal{S}'_h} V(S)|$, label the vertices in $V(G_\alpha) \setminus V(\hat{F})$ as $\{v_{S,i} : 1 \leq h \leq k, S \in \mathcal{S}'_h, 1 \leq i \leq s_h\}$.

Starting with g_0, for each $1 \leq h \leq k$ in turn, we will (almost surely) find a function

$$g_h : V(F_h) \to V(\hat{F}) \cup \{v_{S,i} : 1 \leq h' \leq h, S \in \mathcal{S}_{h'}, 1 \leq i \leq s_h\}$$

such that

Q1 g_h is an embedding of F_h into $G_\alpha \cup G \cup (\bigcup_{h' \leq h} G'_{h'})$, and

Q2 for each vertex v of F_{h-1}, except for at most $\frac{\varepsilon h n}{k}$ vertices in V_0, we have $g_h(v) = g_{h-1}(v)$.

Note that g_0 satisfies these properties, and that, once we a.a.s. find g_k, we will have an embedding of $F = F_k$ into $G_\alpha \cup G \cup (\bigcup_h G'_h)$, as required.

Suppose then that $1 \leq h \leq k$ and we have already found the function g_{h-1}. Then we define the set $W_{h-1} = \{g_0(v) : g_{h-1}(v) = g_0(v), v \in V_0\}$, that is, the vertices of $\hat{F}$ in W_0 that have not been switched. Note that $|W_0 \setminus W_{h-1}| \leq \frac{\varepsilon(h-1)n}{k}$ by **Q2**. For each $S \in \mathcal{S}'_h$, label $V(S) = \{z_{S,1}, \ldots, z_{S,s_h}\}$, and let L_S be the s_h-uniform auxiliary hypergraph with vertex set W_{h-1}, where e is an edge of L_S if, for some labelling $e = \{w_{S,1}, \ldots, w_{S,s_h}\}$, the map $z_{S,i} \mapsto w_{S,i}$ is an embedding of S into $G_\alpha \cup G'_h$, where, for each $1 \leq i \leq s_h$,

$$w_{S,i} \in B(v_{S,i}) \quad \text{and} \quad g_{h-1}(N_{F_h}(z_{S,i}) \cap V(F_{h-1})) \subseteq N_{G_\alpha \cup G'_h}(w_{S,i}). \tag{5.4}$$

Each hyperedge $e = \{w_{S,1}, \ldots, w_{S,s_h}\}$ of L_S then corresponds to a possible extension of g_{h-1} to cover $S \in \mathcal{S}'_h$, subject to the switching of the vertices $v_{S,i}$ and $w_{S,i}$, for $1 \leq i \leq s_h$.

We wish to show that there a.a.s. exists a function $\pi : \mathcal{S}'_h \mapsto \bigcup_{S \in \mathcal{S}'_h} E(L_S)$ such that $\pi(S) \in E(L_S)$ for each $S \in \mathcal{S}'_h$, and the edges in $\pi(\mathcal{S}'_h)$ are pairwise vertex disjoint. This is possible, as shown below, using Theorem 2.20 and the following lemma.

Lemma 5.5. *For each $1 \leq h \leq k$, $1 \leq r \leq |\mathcal{S}'_h|$, $\mathcal{S} \subseteq \mathcal{S}'_h$ and $U \subseteq W_{h-1}$, with $|\mathcal{S}| = r$ and $|U| \leq s_h^2 r$, the following holds with probability at least $1 - \exp(-\omega(r \ln(\frac{n}{r})))$. There exists some $S \in \mathcal{S}$ and an edge $e \in E(L_S)$ with $V(e) \subseteq W_{h-1} \setminus U$.*

The property in Lemma 5.5 then holds for each $1 \leq h \leq k$, $1 \leq r \leq |\mathcal{S}'_h|$, $\mathcal{S} \subseteq \mathcal{S}'_h$ and $U \subseteq W_{h-1}$, with $|\mathcal{S}| = r$ and $|U| \leq s_h^2 r$ with probability at least

$$1 - k \cdot n \cdot \sum_{r=1}^{|\mathcal{S}'_h|} \binom{n}{r} \cdot \binom{n}{s_h^2 r} \cdot \exp\Big(- \omega\big(r \ln(\tfrac{n}{r})\big)\Big) = 1 - o(1).$$

To apply Theorem 2.20, we need to show that, for every $\mathcal{S} \subseteq \mathcal{S}'_h$, the hypergraph $\bigcup_{S \in \mathcal{S}} L_S$ contains a matching with size greater than $s_h(|\mathcal{S}| - 1)$. Let $\mathcal{S} \subseteq \mathcal{S}'_h$ and $r = |\mathcal{S}|$, and let U be the vertex set of a maximal matching in $\bigcup_{S \in \mathcal{S}} L_S$. This means that there is no graph $S \in \mathcal{S}$ and edge $e \in E(L_S)$ with $V(e) \subseteq W_{h-1} \setminus U$. Thus, by the property from Lemma 5.5, we have $|U| \geq s_h^2 |\mathcal{S}|$, so that $\bigcup_{S \in \mathcal{S}} L_S$ contains a matching with size at least $s_h|\mathcal{S}|$. Therefore, we can apply Theorem 2.20, and conclude that a function π as described above exists.

For each $S \in \mathcal{S}'_h$, label $\pi(S) = \{w_{S,i} : 1 \leq i \leq s_h\}$ so that $z_{S,i} \mapsto w_{S,i}$ is an embedding of S into $G_\alpha \cup G'_h$ and (5.4) holds. For each $S \in \mathcal{S}$, switch the vertices in $\{v_{S,i} : 1 \leq i \leq s_h\}$ with those in $\pi(S)$ in F_{h-1} and use π to embed S. That is, define $g_h : V(F_h) \to V(\hat{F}) \cup \{v_{S,i} : 1 \leq h' \leq h, S \in \mathcal{S}'_{h'}, 1 \leq i \leq s_{h'}\}$ by, for each $v \in V(F_h)$, letting

$$g_h(v) = \begin{cases} w_{S,i} & \text{if } v = z_{S,i} \text{ for some } S \in \mathcal{S}'_h, 1 \leq i \leq s_h \\ v_{S,i} & \text{if } g_{h-1}(v) = w_{S,i} \text{ for some } S \in \mathcal{S}'_h, 1 \leq i \leq s_h \\ g_{h-1}(v) & \text{otherwise.} \end{cases}$$

Note that g_h agrees with g_{h-1} throughout $V(F_0)$, except for the at most $\frac{\varepsilon n}{k}$ vertices in $\{v : g_{h-1}(v) = w_{S,i}, S \in \mathcal{S}'_h, 1 \leq i \leq s_h\}$, and therefore **Q2** holds for g_h.

We claim that g_h is an embedding of F_h into $G_\alpha \cup G \cup (\bigcup_{h' \le h} G'_{h'})$, so that also **Q1** holds. Let

$$Z_0 := \{v : v = z_{S,i} \text{ or } g_{h-1}(v) = w_{S,i} \text{ for some } S \in \mathcal{S}'_h, 1 \le i \le s_h\}.$$

Note that g_h agrees with g_{h-1} outside of Z_0, so that g_h (appropriately restricted) is an embedding of $F_h - Z_0$. By the definition of Z_0 and **P5**, the only edges in $F_h[Z_0]$ are those within each $S \in \mathcal{S}'_h$. For each $S \in \mathcal{S}'_h$ we have that $z_{S,i} \mapsto w_{S,i}$ is an embedding of S into $G_\alpha \cup G'_h$. It follows that g_h (appropriately restricted) is an embedding of $F_h[Z_0]$. It remains only to check that the edges between Z_0 and $V(F_h) \setminus Z_0$ are appropriately embedded by g_h. That is, we wish to show for each $v \in Z_0$ that

$$g_h(N_{F_h}(v) \setminus Z_0) \subset N_{G_\alpha \cup G \cup (\bigcup_{h' \le h} G'_{h'})}(g_h(v)).$$

We consider two cases. Firstly, for each $v = z_{S,i}$ with $S \in \mathcal{S}_h$ and $1 \le i \le s_h$ the vertex v has no neighbours in V_0, and hence

$$g_h(N_{F_h}(v) \setminus Z_0) = g_h(N_{F_h}(z_{S,i}) \setminus Z_0) \overset{\mathbf{Q2}}{=} g_{h-1}(N_{F_h}(z_{S,i}) \setminus Z_0) \overset{(5.4)}{\subseteq} N_{G_\alpha \cup G'_h}(w_{S,i}) = N_{G_\alpha \cup G'_h}(g_h(v))\,.$$

Secondly, for each $v \in Z_0$ with $g_{h-1}(v) = w_{S,i}$ for some $S \in \mathcal{S}'_h$ and $1 \le i \le s_h$, we have $g_h(v) = w_{S,i}$. Moreover, by the choice of V_0 we have $N_{F_h}(v) \subset V(F_0)$. By **Q2** it follows that

$$g_h(N_{F_h}(v) \setminus Z_0) = g_0(N_{F_h}(v)) = N_{\hat{F}}(w_{S,i}) \subset N_{G_\alpha}(v_{S,i}) = N_{G_\alpha}(g(v)),$$

where we have used the definition of $B(v)$ in (5.3) and that $w_{S,i} \in B(v_{S,i})$ by (5.4).

Thus, we can a.a.s. extend the embedding g_0 to an embedding g_k of F in $G_\alpha \cup G \cup (\bigcup_h G'_h)$, completing the proof of the theorem.

5.3 Proofs of the lemmas

In this section, we give the proofs of the lemmas from Section 5.2. We prove Lemma 5.3 (and later Lemma 5.5) with Janson's inequality, using similar calculations to Ferber, Luh, and Nguyen [54].

Proof of Lemma 5.3

Fixing h, note that there are certainly at most $2^n \cdot 2^n$ choices for $\mathcal{S}$ and U. Therefore, it is sufficient to prove, for fixed $\mathcal{S} \subseteq \mathcal{S}_h$ and $U \subseteq V(G_\alpha) \setminus f_{h-1}(V(F'_{h-1}))$ with $|\mathcal{S}| \ge \frac{\varepsilon n}{s_h^2 k}$ and $|U| \ge \frac{\varepsilon n}{s_h k}$, the property in the lemma holds with probability $1 - e^{-\omega(n)}$.

Let $s = s_h$. Pick some $S_0 \in \mathcal{S}$, so that, by **P3**, each graph in $\mathcal{S}$ is isomorphic to S_0, and label $V(S_0) = \{v_1, \ldots, v_s\}$. Let $\mathcal{H}$ be the set of copies of S_0 in the complete graph with vertex set in U. Note that $|U| = \Omega(n)$ and $|\mathcal{H}| = \Omega(n^s)$.

For each $S \in \mathcal{S}$ and $H \in \mathcal{H}$, label $V(S) = \{z_{S,1}, \ldots, z_{S,s}\}$ and $V(H) = \{v_{H,1}, \ldots, v_{H,s}\}$ so that $v_i \mapsto z_{S,i}$ and $v_i \mapsto v_{H,i}$ are embeddings of S_0. We now distinguish two cases: Case I where there is some edge between S_0 and F'_{h-1} in F'_h and Case II where there is no such edge.

Let us assume first that we are in Case I. For each $S \in \mathcal{S}$, let $W_S = f_{h-1}(\bigcup_{v \in V(S)} N_{F'_{h-1}}(v))$ be the images of the already embedded neighbours of vertices in S. Note that these sets W_S are non-empty by the definition of Case I and by **P5** disjoint. For each $H \in \mathcal{H}$ and $S \in \mathcal{S}$, let $H \oplus W_S$ be the graph with vertex set $V(H) \cup W_S$ containing exactly those edges that we would need in order to extend the partial embedding we have so far to S embedded into H. That is, $H \oplus W_S$ has edge set

$$E(H) \cup \{v_{H,i}v : 1 \leq i \leq s, v \in f_{h-1}(N_{F'_{h-1}}(z_{S,i}))\}.$$

For each $S \in \mathcal{S}$, $H \in \mathcal{H}$ and $J \subseteq H$, let $J \oplus W_S = (H \oplus W_S)[V(J) \cup W_S]$. Let $\mathcal{H}^+ = \{H \oplus W_S : H \in \mathcal{H}, S \in \mathcal{S}\}$, and note that if any graph from $\mathcal{H}^+$ appears in G_h then we can indeed extend our current embedding to one more dense spot in $\mathcal{S}$, and hence are done.

Let $\mathcal{J} = \{H \cap H' : H, H' \in \mathcal{H}, e(H \cap H') > 0\}$ and $\mathcal{J}' = \{H \cap H' : H, H' \in \mathcal{H}, H \neq H'\} \setminus \emptyset$. We will show that $\mathbb{P}[\exists H \in \mathcal{H}^+ \text{ with } H \subseteq G_h] = 1 - \exp(-\omega(n))$ follows from Theorem 2.18 and the following claim, which we then prove.

Claim 5.6.

(i) For each $J \in \mathcal{J}$, $2e(J) < (\Delta+1)(|J|-1)$.

(ii) For each $H \in \mathcal{H}$ and $S \in \mathcal{S}$, $2e(H \oplus W_S) \leq (\Delta+1)s$.

(iii) For each $S \in \mathcal{S}$ and $J \in \mathcal{J}'$, $2e(J \oplus W_S) < (\Delta+1)|J|$.

Indeed, using (ii) of Claim 5.6, and that $q = \omega(n^{-\frac{2}{\Delta+1}})$, let

$$\mu := \sum_{S \in \mathcal{S}} \sum_{H \in \mathcal{H}} q^{e(H \oplus W_S)} = \Omega(n^{s+1} q^{(\Delta+1)s/2}) = \omega(n).$$

Note that, by **P3** and (ii) of Claim 5.6, each graph in $\mathcal{H}^+$ has the same number of edges, say $m \leq (\Delta+1)s/2$. Let

$$\begin{aligned}
\delta &:= \sum_{S,S' \in \mathcal{S}} \sum_{\substack{H,H' \in \mathcal{H} \\ H \oplus W_S \sim H' \oplus W_{S'}}} q^{e(H \oplus W_S) + e(H' \oplus W_{S'}) - e((H \oplus W_S) \cap (H' \oplus W_{S'}))} \\
&= q^{2m} \sum_{S,S' \in \mathcal{S}} \sum_{\substack{H,H' \in \mathcal{H} \\ H \oplus W_S \sim H' \oplus W_{S'}}} q^{-e((H \oplus W_S) \cap (H' \oplus W_{S'}))} \\
&\leq q^{2m} \sum_{J \in \mathcal{J}} \sum_{\substack{S,S' \in \mathcal{S} \\ S \neq S'}} \sum_{\substack{H,H' \in \mathcal{H} \\ H \cap H' = J}} q^{-e(J)} + q^{2m} \sum_{J \in \mathcal{J}'} \sum_{S \in \mathcal{S}} \sum_{\substack{H,H' \in \mathcal{H} \\ H \cap H' = J}} q^{-e(J \oplus W_S)} \\
&\leq q^{2m} \sum_{J \in \mathcal{J}} |\mathcal{S}|^2 n^{2s-2|J|} q^{-e(J)} + q^{2m} \sum_{J \in \mathcal{J}'} \sum_{S \in \mathcal{S}} n^{2s-2|J|} q^{-e(J \oplus W_S)}.
\end{aligned} \tag{5.5}$$

Then, using (i) and (iii) of Claim 5.6, and as $\mu = \Omega(n^{s+1}q^m)$, we have

$$\begin{aligned}
\frac{\delta}{\mu^2} &= O\Big(\sum_{J \in \mathcal{J}} |\mathcal{S}|^2 n^{-2|J|-2} q^{-e(J)} + \sum_{J \in \mathcal{J}'} \sum_{S \in \mathcal{S}} n^{-2|J|-2} q^{-e(J \oplus W_S)} \Big) \\
&= O\Big(\sum_{J \in \mathcal{J}} n^{-2|J|} q^{-(\Delta+1)(|J|-1)/2} + \sum_{J \in \mathcal{J}'} |\mathcal{S}| \cdot n^{-2|J|-2} q^{-(\Delta+1)|J|/2} \Big)
\end{aligned}$$

$$= o\Big(\sum_{J \in \mathcal{J}} n^{-2|J|} n^{|J|-1} + \sum_{J \in \mathcal{J}'} n^{-2|J|-1} n^{|J|} \Big) = o(n^{-1}).$$

Therefore, as $\mu = \omega(n)$ and $\frac{\delta}{\mu^2} = o(n^{-1})$, by Theorem 2.18, the probability that there is no graph in $\mathcal{H}^+$ in G_h is at most $\exp(-\frac{\mu^2}{4(\mu+\delta)}) = \exp(-\omega(n))$, as required. For Case I, it is left then only to prove Claim 5.6.

Proof of Claim 5.6. For (i) let $H, H' \in \mathcal{H}$ be such that $H \cap H' = J$. If $J \neq H$, and $|J| \geq 3$, then, by **P2**, we have $2e(J) \leq (\Delta+1)(|J|-2) < (\Delta+1)(|J|-1)$, as required. If $|J| = 2$, then $(\Delta+1)(|J|-1) = \Delta + 1 > 2 \geq e(J)$.

Suppose then that $J = H$, so $|J| = s$. If $s \leq \Delta$, then $2e(J) \leq s(s-1) < (s+1)(s-1) \leq (\Delta+1)(s-1)$, and if $s > \Delta + 1$, then $2e(J) \leq s\Delta < s\Delta + s - \Delta + 1 = (\Delta+1)(s-1)$, as required. If $s = \Delta + 1$, note that, as there is some edge between S_0 and F_{h-1} in F_h, we have that S_0, and hence $J = H$, is not a clique with $\Delta + 1$ vertices. Thus, $2e(J) < s(s-1) = (\Delta+1)(s-1)$.

For (ii) suppose $s \geq \Delta + 1$. As H is dense we have $2e(H) > (\Delta+1)(s-2)$, and thus

$$2e(H \oplus S) \leq 2\Delta s - 2e(H) < 2\Delta s - (\Delta+1)(s-2) = (\Delta+1)s + 2(\Delta+1-s) \leq (\Delta+1)s,$$

as required.

So suppose that $s \leq \Delta$. If $4 \leq s \leq \Delta - 1$, then, as $2e(H) > (\Delta+1)(s-2)$, we must have

$$s(s-1) > (\Delta+1)(s-2) \geq (s+2)(s-2) = s(s-1) + s - 4 \geq s(s-1),$$

a contradiction. If $s = 3$, then $2e(H) > \Delta + 1$ contradicts $\Delta \geq 5$.

Finally, if $s = \Delta$, then H must be the clique on Δ vertices because $2e(H) > (\Delta+1)(\Delta-2) = \Delta(\Delta-1) - 2$. Therefore,

$$2e(H \oplus S) \leq 2\Delta^2 - 2e(H) = \Delta(\Delta+1) = (\Delta+1)s.$$

For (iii) let $H, H' \in \mathcal{H}$ be such that $H \cap H' = J$ and $H \neq H'$, which exist by the definition of $\mathcal{J}'$. Observe that $|J| < s$. Let $I = H - V(J)$, and let $e(I, J)$ be the number of edges between I and J in H. Then,

$$\begin{aligned} 2e(J \oplus W_S) \leq 2(\Delta|J| - e(J) - e(J, I)) &= 2(\Delta|J| - e(H) + e(I)) \\ &= (\Delta+1)|J| + (\Delta-1)|J| - 2e(H) + 2e(I). \end{aligned}$$

Thus, to prove the claim it is sufficient to show that $(\Delta-1)|J| < 2e(H) - 2e(I)$.

As H is dense, we have $2e(H) > (\Delta+1)(|J| + |I| - 2)$. If $|I| \geq 3$, then, from **P2**, we have $2e(H) > (\Delta+1)|J| + 2e(I)$. If $|I| = 2$, then $2e(H) > (\Delta+1)|J| \geq (\Delta-1)|J| + 2e(I)$. Finally, if $|I| = 1$ then $e(I) = 0$ and, since by (ii), $|J| = s - 1 \geq \Delta - 1$ and $\Delta \geq 3$, we have $2e(H) > (\Delta-1)|J| + 2e(I)$.

So in each case, $2e(H) - 2e(I) > (\Delta-1)|J|$ as required. □

It remains to consider Case II. In this case the graphs in $\mathcal{S}_h$ have no edges to F_{h-1}. Therefore, it is sufficient for some graph in $\mathcal{H}$ to exist. Let m be the size of each (isomorphic) graph in $\mathcal{H}$, and note that $2m \le \min\{s\Delta, s(s-1)\} \le (s-1)(\Delta+1)$. Thus, we may take

$$\mu := \sum_{H\in\mathcal{H}} q^m = \Omega(n^s q^m) = \Omega(n^s q^{(s-1)(\Delta+1)/2}) = \omega(n).$$

Let $\mathcal{J} = \{H \cap H' : H, H' \in \mathcal{H}, e(H\cap H') > 0, H \neq H'\}$ and note that, if $J \in \mathcal{J}$ and $|J| \ge 3$, then $2e(J) \le (\Delta+1)(|J|-2)$ by **P2**. Let

$$\begin{aligned}\delta &:= \sum_{\substack{H,H'\in\mathcal{H}\\ H\sim H', H\neq H'}} q^{e(H)+e(H')-e(H\cap H')} = q^{2m}\sum_{J\in\mathcal{J}}\sum_{\substack{H,H'\in\mathcal{H}\\ H\cap H'=J}} q^{-e(J)} \le q^{2m}\sum_{J\in\mathcal{J}} n^{2s-2|J|}q^{-e(J)}\\ &\le q^{2m-1}n^{2s-2} + q^{2m}\sum_{J\in\mathcal{J}:|J|\ge 3} n^{2s-2|J|}q^{-(\Delta+1)(|J|-2)/2}.\end{aligned}$$

Then, as $\mu = \Omega(n^s q^m)$, we have

$$\frac{\delta}{\mu^2} = O\Big(q^{-1}n^{-2} + \sum_{J\in\mathcal{J}:|J|\ge 3} n^{-2|J|}q^{-(\Delta+1)(|J|-2)/2}\Big) = o\Big(n^{-1} + \sum_{J\in\mathcal{J}:|J|\ge 3} n^{-|J|-2}\Big) = o(n^{-1}).$$

Therefore, as $\mu = \omega(n)$, and $\frac{\delta}{\mu^2} = o(n^{-1})$, by Lemma 2.18, the probability that there is no graph in $\mathcal{H}$ in G_h is at most $\exp(-\frac{\mu^2}{2(\mu+\delta)}) = \exp(-\omega(n))$, as required. □

Proof of Lemma 5.4

The proof of this lemma is based on the fact that the sets $B(v)$ are *random* sets. The reason for this is that $\hat{F}$ is a random copy of F^* in G and that G is itself a random graph.

Recall that we have $F^* \subset F$, and a copy $\hat{F}$ of F^* chosen randomly in G. Since G is the union of random graphs it follows that $V(\hat{F}) \subseteq V(G_\alpha)$ is chosen uniformly at random in $V(G_\alpha)$, which will be crucial in the following. Recall also that we call the resulting embedding $g_0 : F^* \to \hat{F}$. Further, recall that we have an independent set $V_0 \subseteq V(F^*)$ with $|V_0| \ge \frac{n}{2\Delta}$, and $W_0 = g_0(V_0)$. Again, W_0 is a uniformly random set of size $|V_0|$ in $V(G_\alpha)$.

Now, greedily pick a set $V_1 \subseteq V_0$ of at least $\frac{n}{4\Delta^3}$ vertices in V_0 which pairwise have no common neighbours in $\hat{F}$, and let $W_1 = g_0(V_1)$. Fix $u, v \in V(G_\alpha)$. For each $w \in W_1$, let Z_w be the set of neighbours of w in $\hat{F}$, and note that $|Z_w| \le \Delta$ and that the sets $Z_w \cup \{w\}$ are all disjoint. For each $w \in W_1$, let I_w be the indicator variable for the event $w \in N_{G_\alpha}(u)$ and $Z_w \subseteq N_{G_\alpha}(v)$. Since $B(v) = \{w \in W_0 : N_{\hat{F}}(w) \subset N_{G_\alpha}(v)\}$ by (5.3), it follows that

$$|N_{G_\alpha}(u) \cap B(v)| \ge \sum_{w\in W_1} I_w. \tag{5.6}$$

Let $r = \frac{\alpha n}{4\Delta^3} \le |W_1|$ and pick distinct vertices $w_1, \dots, w_r$ in W_1. Consider revealing $\hat{F}$ by, first revealing the vertices in $\{w_1\} \cup Z_{w_1}$, then revealing the vertices in $\{w_2\} \cup Z_{w_2}$, and so on, until $\{w_r\} \cup Z_{w_r}$, before finally revealing the rest of the vertices in $\hat{F}$. Note that, for each $1 \le i \le r$, when the location of the vertices in $\{w_i\} \cup Z_{w_i}$ is revealed there will be at least $\alpha n/2$ vertices both in $N_{G_\alpha}(v)$ and $N_{G_\alpha}(u)$ which are not occupied by a vertex in $\{w_j\} \cup Z_{w_j}$ with $j < i$. Therefore, for each $1 \le i \le r$, if $m = |Z_{w_i}|$ and F_i is a random variable recording the location of the vertices in $\{w_j\} \cup Z_{w_j}$ with $j < i$, then

$$\mathbb{E}[I_{w_i}|F_i] \geq \frac{\alpha n/2 \cdot \binom{\alpha n/2-1}{m}}{(m+1)n\binom{n}{m}} \geq \frac{(\frac{\alpha}{4})^{m+1}}{(m+1)^2} \geq \frac{\alpha^{\Delta+1}}{4^{\Delta+1}(\Delta+1)^2} =: \delta. \tag{5.7}$$

Therefore, by Lemma 2.17 and (5.6), with probability $1 - \exp(-\Omega(\delta r)) = 1 - o(n^{-2})$ we have

$$|N_{G_\alpha}(u) \cap B(v)| \geq \delta r/2 \geq \frac{\alpha^{\Delta+2} n}{4^{\Delta+2}(\Delta+1)^5} \geq 4\Big(\frac{\alpha}{4\Delta}\Big)^{2\Delta} n = 4\varepsilon n.$$

Therefore, with probability $1 - o(1)$, $|N_{G_\alpha}(u) \cap B(v)| \geq 4\varepsilon n$ for each $u, v \in V(G_\alpha)$. □

Proof of Lemma 5.5

Again, we use Janson's inequality and similar calculations to Ferber, Luh and Nguyen [54].

Recall that $W_{h-1} = \{g_0(v) : g_{h-1}(v) = g_0(v), v \in V_0\}$ is the set of vertices of $\hat{F}$ in W_0 that have not been switched. Let $1 \leq h \leq k$, $1 \leq r \leq |\mathcal{S}'_h| \leq \varepsilon n/s_h^2 k$, $\mathcal{S} \subseteq \mathcal{S}'_h$ and $U \subseteq W_{h-1}$ with $|\mathcal{S}| = r$ and $|U| \leq s_h^2 r$. Note that, as $|U| \leq s_h^2 r \leq \varepsilon n$, we have $|U \cup (W_0 \setminus W_{h-1})| \leq 2\varepsilon n$. Therefore, by the property from Lemma 5.4, for each $u, v \in V(G_\alpha)$, we have

$$|N_{G_\alpha}(u, B(v)) \cap (W_{h-1} \setminus U)| \geq 2\varepsilon n, \tag{5.8}$$

and, in particular, $|B(v) \cap (W_{h-1} \setminus U)| \geq 2\varepsilon n$.

Let $s = s_h$. Pick $S_0 \in \mathcal{S}$, so that, by **P3**, each graph in $\mathcal{S}$ is isomorphic to S_0. As in the proof of Lemma 5.3, we will consider two cases: Case I where there is some edge between S_0 and F_{h-1} in F_h, and Case II where there is no such edge.

Suppose first we are in Case I. Label $V(S_0) = \{v_1, \ldots, v_s\}$ so that v_1 has a neighbour in F_{h-1}. Recall that for $S \in \mathcal{S}$ we labelled $V(S) = \{z_{S,1}, \ldots, z_{S,s_h}\}$. Without loss of generality, we can assume for each $S \in \mathcal{S}$ that $v_i \mapsto z_{S,i}$ is an isomorphism from S_0 into S. Let $\mathcal{H}$ be the set of copies of S_0 in the complete graph with vertex set $W_{h-1} \setminus U$. For each $H \in \mathcal{H}$, label $V(H) = \{v_{H,1}, \ldots, v_{H,s}\}$ so that $v_i \mapsto v_{H,i}$ is an isomorphism of S_0 to H.

For each $S \in \mathcal{S}$, pick the image w_S of an already embedded neighbour of the vertex $z_{S,1}$ corresponding to v_1, that is, pick $w_S \in g_{h-1}(N_{F_h}(z_{S,1}))$. For each $S \in \mathcal{S}$, let

$$\mathcal{H}_S = \{H \in \mathcal{H} : v_{H,1} \in N_{G_\alpha}(w_S) \text{ and } v_{H,i} \in B(v_{S,i}) \text{ for each } 1 \leq i \leq h\}.$$

These are precisely those copies of S_0 where the vertex previously embedded to $v_{H,i}$ can be replaced by $v_{S,i}$ for $1 \leq i \leq h$ and the edge from $v_{H,1}$ to w_S is already present in G_α. For each $S \in \mathcal{S}$, note that, from (5.8), we have $|\mathcal{H}_S| = \Omega(n^s)$. For each $S \in \mathcal{S}$, let $W_S = g_{h-1}(\bigcup_{v \in V(S)} N_{F_h}(v))$ be the set of images of already embedded neighbours of vertices in S.

For each $H \in \mathcal{H}_S$ and $S \in \mathcal{S}$, let $H \oplus W_S$ be the graph with vertex set $V(H) \cup W_S$ and edge set

$$E(H) \cup (\{v_{H,i}v : 1 \leq i \leq s, v \in g_{h-1}(N_{F_h}(w_{S,i}))\} \setminus \{v_{H,1}w_S\}).$$

These are exactly the edges we need in order to extend our embedding of F'_{h-1} to contain S embedded into H. For each $S \in \mathcal{S}$, $H \in \mathcal{H}$ and $J \subseteq H$, let $J \oplus W_S = (H \oplus W_S)[V(J) \cup W_S]$. Let $\mathcal{H}^+ = \{H \oplus W_S : S \in \mathcal{S}, H \in \mathcal{H}_S\}$, and note that if any graph from $\mathcal{H}^+$ appears in G'_h then we are done.

Let $\mathcal{J} = \{H \cap H' : H, H' \in \mathcal{H}, e(H \cap H') > 0\}$ and $\mathcal{J}' = \{H \cap H' : H, H' \in \mathcal{H}, H \neq H'\} \setminus \emptyset$. Note that (i) and (iii) of Claim 5.6 hold here as well. For each $H \in \mathcal{H}$ and $S \in \mathcal{S}$, $E(H \oplus W_S)$ does not include $v_{H,1}w_S$, and, therefore, in place of (ii), the following holds.

(ii') For each $S \in \mathcal{S}$ and $H \in \mathcal{H}_S$, $2e(H \oplus W_S) \leq (\Delta + 1)s - 2$.

Note that, by **P3**, each graph in $\mathcal{H}^+$ has the same number of edges, m say. Note that, as the property we are looking for is monotone, we may assume that $q^{-1/2} = \omega(\ln n)$. Using (ii') from above, let

$$\mu := \sum_{S \in \mathcal{S}} \sum_{H \in \mathcal{H}_S} q^m = \Omega(rn^s q^m) = \Omega(rn^s q^{(\Delta+1)s/2-1}) = \Omega(rq^{-1}) = \omega(r \ln n).$$

We remark, that this is the only place where we use that the edge $v_{H,1}w_S$ is not included in $H \oplus W_S$, since it is already present in G_α.

Defining δ as follows, and using similar deductions to those used to reach (5.5), we have

$$\begin{aligned}\delta &:= \sum_{S,S' \in \mathcal{S}} \sum_{\substack{H \in \mathcal{H}_S, H' \in \mathcal{H}_{S'} \\ H \oplus W_S \sim H' \oplus W_{S'}}} q^{e(H \oplus W_S) + e(H' \oplus W_{S'}) - e((H \oplus W_S) \cap (H' \oplus W_{S'}))} \\ &\leq q^{2m} r^2 \sum_{J \in \mathcal{J}} n^{2s-2|J|} q^{-e(J)} + q^{2m} \sum_{J \in \mathcal{J}'} \sum_{S \in \mathcal{S}} n^{2s-2|J|} q^{-e(J \oplus W_S)}.\end{aligned}$$

Then, using (i) and (iii) of Claim 5.6, and that $\mu = \Omega(rn^s q^m)$, we have

$$\begin{aligned}\frac{\delta}{\mu^2} &= O\Big(\sum_{J \in \mathcal{J}} n^{-2|J|} q^{-e(J)} + r^{-2} \sum_{J \in \mathcal{J}'} \sum_{S \in \mathcal{S}} n^{-2|J|} q^{-e(J \oplus W_S)}\Big) \\ &= O\Big(\sum_{J \in \mathcal{J}} n^{-2|J|} q^{-((\Delta+1)(|J|-1)-1)/2} + r^{-2} \sum_{J \in \mathcal{J}'} \sum_{S \in \mathcal{S}} n^{-2|J|} q^{-((\Delta+1)|J|-1)/2}\Big) \\ &= o\Big(q^{1/2} \sum_{J \in \mathcal{J}} n^{-2|J|} n^{|J|-1} + q^{1/2} r^{-2} \sum_{J \in \mathcal{J}'} \sum_{S \in \mathcal{S}} n^{-2|J|} n^{|J|}\Big) = o(q^{1/2} n^{-1} + q^{1/2} r^{-1}) = o(r^{-1} \ln^{-1} n).\end{aligned}$$

Therefore, as $\mu = \omega(r \ln n)$, and $\frac{\delta}{\mu^2} = o(r^{-1} \ln^{-1} n)$, by Lemma 2.18, the probability that there is no graph in $\mathcal{H}^+$ in G_h is at most $\exp(-\frac{\mu^2}{2(\mu+\delta)}) = \exp(-\omega(r \ln n))$, completing the proof of Lemma 5.5 in Case I.

Let us assume now we are in Case II, with no edges between S_0 and F_{h-1} in F_h. Label $V(S_0) = \{v_1, \ldots, v_s\}$. Let $\mathcal{H}$ be the set of copies of S_0 in the complete graph with vertex set $W_{h-1} \setminus U$. For each $H \in \mathcal{H}$, label $V(H) = \{v_{H,1}, \ldots, v_{H,s}\}$ so that $v_i \mapsto v_{H,i}$ is an embedding of S_0. Let

$$\mathcal{H}' = \{H \in \mathcal{H} : \exists S \in \mathcal{S} \text{ such that } v_{H,i} \in B(v_{S,i}) \text{ for each } 1 \leq i \leq h\}.$$

Note that if we have some graph from $\mathcal{H}'$ in G'_h, then we are done. From (5.8), we have $|\mathcal{H}'| = \Omega(n^s)$, so, with very similar calculations to Case II in the proof of Lemma 5.3, we have that the probability that there exists no graph from $\mathcal{H}'$ in G'_h is at most $\exp(-\omega(n)) \leq \exp(-\omega(r \ln(n/r)))$, as required. □

5.4 Proofs for trees and factors

Theorem 2.7 can be shown to extend to $\Delta \leq 3$ using basically the same approach as presented here. The definition of the *dense spots*, however, has to be slightly adapted to each case. For $\Delta = 4$, there is one problematic dense spot, a triangle attached to the rest of the graph with two pendant edges at each vertex, which prevents our methods extending to this case. A similar gap exists for extensions of Theorem 2.4.

With simple modifications, our methods are applicable more generally to the determination of appearance thresholds for spanning structures in the randomly perturbed graph model. This allows us to give simpler, non-regularity, proofs of results already found in the literature. In particular, we can reprove the recent results concerning bounded degree spanning trees and factors. We will show how our methods imply these results in the following. As the calculations and arguments are very similar to our proof, we shall be brief.

Spanning trees

Theorem 5.7 (Krivelevich, Kwan and Sudakov [82]). *For any $\alpha, \Delta > 0$, if $p = \omega(1/n)$ and T is an n-vertex tree with maximum degree at most Δ, then $G_\alpha \cup \mathcal{G}(n,p)$ contains a copy of T a.a.s.*

Applying our new approach to the randomly perturbed graph model, we can give a proof of this result, as follows.

Proof. Fixing $\alpha > 0$ and $\Delta > 0$, let $\varepsilon = \varepsilon(\alpha, \Delta)$ be a small constant. Let $p = \omega(n^{-1})$ and let T be a tree with n vertices and maximum degree at most Δ. Clearly T contains some subtree T' with $\lceil (1-\varepsilon)n \rceil$ vertices. By the work of Alon, Krivelevich and Sudakov[14], we know that $\mathcal{G}(n,p)$ almost surely contains a copy S' of T'. As before, we observe that this copy is independent of G_α and placed uniformly at random on top of it. Using the same methods as we used for Lemma 5.4, we obtain that, a.a.s. each uncovered vertex v in G_α can be switched for a set $B(v)$ of vertices in the copy of T' so that each vertex in G_α has at least $(\Delta+1)\varepsilon n$ neighbours in G_α in $B(v)$.

We then greedily extend the copy of T' to a copy of T using the following deterministic strategy. Picking an uncovered vertex v, let B_v be obtained from $B(v)$ by removing those vertices which have been switched in S' or whose neighbours in S' have been switched and note that we removed less than $(\Delta+1)\varepsilon n$ vertices. Picking a vertex u in T which needs to be embedded as a leaf of the partial embedding, we use that its already embedded parent has at least one neighbour w in B_v. We switch v with w and, by embedding u onto w, gain an extended partial embedding of T. When complete, this gives the required copy of T. □

Factors

Recall that by Theorem 2.2 $n^{-1/d_1(G)} \log^{1/e(G)} n$ gives the threshold for factors of strictly balanced graphs G. Gerke and McDowell [61], on the other hand, showed that for vertex balanced graphs G, this threshold is $n^{-1/m_1(G)}$. The result for general factors in the model $G_\alpha \cup \mathcal{G}(n,p)$ is the following.

Theorem 5.8 (Balogh, Treglown and Wagner [19]). *For every G, if $p = \omega(n^{-1/m_1(G)})$, then $G_\alpha \cup \mathcal{G}(n,p)$ contains a G-factor a.a.s.*

Our methods give a simpler, non-regularity, proof of this result, as follows.

Proof. Fixing G, let $t = v(G)$, $\varepsilon = \varepsilon(\alpha, t) > 0$ be small and $p = \omega(n^{-1/m_1(G)})$. It follows from Theorem 2.19 that $\mathcal{G}(n,p)$ a.a.s. contains an almost G-factor covering at least $(1-\varepsilon t)n$ vertices, which we call F^*. As before, we observe that this copy is independent of G_α and placed uniformly at random on top of it. Using the very same methods we used for Lemma 5.4, we get that, a.a.s. each uncovered vertex v in $V(G_\alpha)$ can be switched for one of at least $3\varepsilon t^3 n$ vertices in different copies of G in F^* to get the same number of disjoint copies of G in $G_\alpha \cup \mathcal{G}(n,p)$. We call this set of switchable vertices $B(v)$.

We can then iteratively extend our embedding as follows. We pick t uncovered vertices $v_1, \ldots, v_t$ in $V(G_\alpha)$ and pick disjoint sets $B_i \subset B(v_i)$ with $|B_i| \geq \varepsilon n$ so that the vertices in B_i are switchable with v_i, the vertices in $\bigcup_i B_i$ appear in different copies of G, and these copies of G have had no vertices switched with them. This is possible as there are at most εn copies of G added, which together switched at most $\varepsilon t n$ vertices and thus blocking at most $\varepsilon t^2 n$ vertices. Furthermore, the previously chosen B_i can block at most another $\varepsilon t^2 n$ many vertices and since at most t vertices appear in the same copy of G this leaves us with $\varepsilon t^2 n$ switchable vertices.

It easily follows from the proof of Theorem 2.19 that, for any t disjoint vertex subsets with at least εn vertices in each subset, with probability at least $1 - n^{-2}$ there is a copy of G in $\mathcal{G}(n,p)$ with one vertex in each subset. Therefore $\mathcal{G}(n,p)$ contains a copy of G with one vertex in each set B_i and thus we can use this copy, along with switchings, to increase the number of disjoint copies of G. As there are at most εn steps until completion, this process finds a G-factor in $G_\alpha \cup \mathcal{G}(n,p)$ a.a.s. □

Chapter 6

Universality in random hypergraphs

Next, we present the proof[33] of Theorem 2.9 on $\mathcal{F}^{(r)}(n,\Delta)$-universality in $\mathcal{H}^{(r)}(n,p)$ obtained together with Person [94]. After the proof outline and some more definitions we state two lemmas which readily imply the theorem. The first lemma guarantees pseudorandom properties in $\mathcal{H}^{(r)}(n,p)$, whereas the second shows that these are sufficient for embedding any graph from $\mathcal{F}^{(r)}(n,\Delta)$. Afterwards, we prove both lemmas.

6.1 Proof outline

Our proof follows a similar strategy as the one of Dellamonica, Kohayakawa, Rödl, and Ruciński [41] for universality of random graphs which we combine with the approach of Kim and Lee [73] and of Ferber, Nenadov and Peter [56].

We will embed any bounded degree hypergraph $F \in \mathcal{F}^{(r)}(n,\Delta)$ into the random hypergraph $\mathcal{H} = \mathcal{H}^{(r)}(n,p)$ with $p = C(\log n/n)^{1/\Delta}$ by verifying certain deterministic pseudorandom properties. More precisely, we introduce the notion of an $(n,r,p,t,\varepsilon,\Delta)$-good hypergraph H (see Definition 6.2 below), and prove that the random hypergraph $\mathcal{H}^{(r)}(n,p)$ is $(n,r,p,t,\varepsilon,\Delta)$-good a.a.s. This reduces our task to find an embedding of any $F \in \mathcal{F}^{(r)}(n,\Delta)$ into such an $(n,r,p,t,\varepsilon,\Delta)$-good hypergraph H.

Roughly speaking, such good hypergraph H admits a partition of its vertices into sets V_0, V_1, …, V_t, so that certain *extension* properties hold. Next we partition most of the vertices of F into 3-independent sets X_1,…, X_t plus an additional set $X_0 = N_F(X_t)$ (recall, that a set is 3-independent if any two of its vertices are at distance at least 4). These 3-independent sets X_1,…, X_t are constructed by colouring greedily the third power of the shadow graph of F. The set X_t has the property that the $(r-1)$-uniform link hypergraph of every $x \in X_t$ in F looks the same together with possibly some further edges of F contained in $N_F(x)$. Thus, we think of $F[X_0]$ as a collection of vertex-disjoint copies of isomorphic pairs (E_1,E_2), which we call profiles, of the edge set E_1 of some r-uniform hypergraph and of the edge set E_2 of some $(r-1)$-uniform hypergraph (isomorphic to every link of $x \in X_t$). One of the properties of H asserts then that $F[X_0]$ can be embedded into $H[V_0]$, no matter which $F \in \mathcal{F}^{(r)}(n,\Delta)$ we consider. Then we extend, using other properties of an $(n,r,p,t,\varepsilon,\Delta)$-good hypergraph H, our embedding in t rounds to the whole $V(F)$ by embedding in the i-th round the vertices from X_i into $V_0 \cup V_1 \ldots \cup V_i$. For this we will verify Hall's condition for the existence of a matching in an appropriately defined bipartite graph that allows us to carry on with our embedding. The role of the sets V_i is technical – it allows us to verify Hall's condition for small subsets, which we cannot simply do in V_0.

[33]The proof in this chapter is a close adaption of [94].

6.2 Auxiliary results and definitions

Let $H = (V, E)$ be an r-uniform hypergraph. The link of v in H is a subset of $\binom{V}{r-1}$ consisting of all $(r-1)$-sets of vertices which form an edge together with v

$$\mathrm{link}_H(v) = \left\{ e' \in \binom{V}{r-1} : e' \cup \{v\} \in E \right\}.$$

For a hypergraph H and a vertex v we define its *profile* $P_H(v)$ in H as follows

$$P_H(v) = (N_H(v), E(H[N_H(v)]), \mathrm{link}_H(v))$$

and say that two profiles $P_H(v_1)$ and $P_H(v_2)$ are equivalent ($P_H(v_1) \cong P_H(v_2)$) if there is an isomorphism φ that takes $H[N_H(v_1)]$ to $H[N_H(v_2)]$ and $(N_H(v_1), \mathrm{link}_H(v_1))$ to $(N_H(v_2), \mathrm{link}_H(v_2))$. We call $N_H(v)$ the vertices of the profile.

Let $P^{(r)}(\Delta)$ be the set of all possible profiles (Z, E_1, E_2) that we encounter for $F \in \cup_{n\in\mathbb{N}} \mathcal{F}^{(r)}(n, \Delta)$ (up to equivalence). Then, for any $|Z| \leq (r-1)\Delta$, (Z, E_1) is an r-uniform hypergraph with maximum degree $\Delta - 1$ and (Z, E_2) is an $(r-1)$-uniform hypergraph with at most Δ edges and without isolated vertices. It is not difficult to bound $|P^{(r)}(\Delta)|$ by a function exponential in some polynomial in Δ, say, by $A^{\Delta^2 \log \Delta}$ for some absolute constant $A = A(r) > 1$. But since Δ is fixed, all we will care about is that $|P^{(r)}(\Delta)|$ is a constant that depends on r and Δ only.

The following lemma prepares any $F \in \mathcal{F}^{(r)}(n, \Delta)$ for future embedding into $\mathcal{H}^{(r)}(n, p)$.

Lemma 6.1. *Let $r \geq 2$ and $\Delta \geq 1$ be integers. Then for $t = r^3\Delta^3$, any $\varepsilon \leq |P^{(r)}(\Delta)|^{-1}(t-1)^{-1}$ and any $F \in \mathcal{F}^{(r)}(n, \Delta)$ there exists a partition of $V(F)$ in $X_0 \cup \cdots \cup X_t$ (where some of the X_i might be empty) with the following conditions:*

Q1 $|X_t| = \varepsilon n$, $X_0 = N_F(X_t)$,

Q2 *every $x \in X_t$ has the same profile $P_F(x)$ (up to equivalence), and*

Q3 X_i *is 3-independent for $i = 1, \ldots, t$.*

Proof. Let $F \in \mathcal{F}^{(r)}(n, \Delta)$ be given and G be its shadow graph. The third power G^3 of G is the graph which we obtain by connecting any pair of vertices of distance at most 3 by an edge. We estimate the maximum degree of G^3 as follows: $\Delta(G^3) \leq (r-1)^3\Delta^3$. Clearly, G^3 is $(t-1)$-colourable and let $Y_1, \ldots, Y_{t-1}$ be the colour sets in some colouring of $V(G^3)$ such that $|Y_1| \leq |Y_2| \leq \ldots \leq |Y_{t-1}|$. The sets Y_i are 3-independent in F as well because the shadow of a path of length 3 in F contains a path of length 3 in G, which gives an edge in G^3 in contradiction to the colouring above.

We can choose a subset $X_t \subseteq Y_{t-1}$ of size $\varepsilon n \leq n|P^{(r)}(\Delta)|^{-1}(t-1)^{-1}$ of vertices with the same profile in F (up to equivalence). We set $X_0 := N_F(X_t)$ and define $X_i = Y_i \setminus X_0$ for $i = 1, \ldots, t-2$ and $X_{t-1} = Y_{t-1} \setminus (X_0 \cup X_t)$. The partition $V(F) = X_0 \cup \cdots \cup X_t$ satisfies the required conditions. □

Given a partition of $V(F)$ from the above lemma, it follows from **Q1–Q3** that $F[X_0]$ is the disjoint union of εn copies of the same r-uniform hypergraph isomorphic to $F[N_F(x)]$ for all $x \in X_t$. Furthermore, the third condition implies that any edge $e \in E(H)$ intersects each X_i in at most one vertex for $i = 1, \ldots, t$.

Let $H = (V, E)$ be an r-uniform hypergraph. Let $\mathcal{L}$ be a family of pairwise disjoint k-subsets of $\binom{V}{r-1}$ and we write $V(\mathcal{L})$ for $\cup_{e \in L : L \in \mathcal{L}} e$. For a subset $W \subseteq V \setminus V(\mathcal{L})$ we define the auxiliary bipartite graph $B(H, \mathcal{L}, W)$ with the vertex classes $\mathcal{L}$ and W, where $L \in \mathcal{L}$ and $w \in W$ form an edge if and only if $L \subseteq \mathrm{link}_H(w)$. Roughly speaking, for every unembedded $x \in V(F)$ the set $L = L_x \in \mathcal{L}$ will consist of the images of the already fully embedded $(r-1)$-sets from the $\mathrm{link}_F(x)$ and the following definition which resembles the one of good graphs from [56, 73] provides essential properties that will assist us while embedding F into $\mathcal{H}^{(r)}(n, p)$.

Definition 6.2. *We say that an r-uniform hypergraph H is $(n, r, p, t, \varepsilon, \Delta)$-good if there exists a partition $V(H) = V_0 \cup V_1 \cup \cdots \cup V_t$, where $|V_i| = \varepsilon n/(10t)$ for $i = 1, \ldots, t$, and $|V_0| = (1 - \varepsilon/10)n$ that satisfies the following conditions:*

P1 *For any profile $(Z, E_1, E_2) \in P^{(r)}(\Delta)$ there exists a family $\mathcal{F}$ of εn vertex-disjoint copies of the profile (Z, E_1, E_2) with vertices in V_0 and edges E_1 present in H. This family induces a family $\mathcal{F}_2$ of pairwise disjoint copies of E_2 in $\binom{V_0}{r-1}$. Furthermore, for any $W \subseteq V(H) \setminus V(\mathcal{F}_2)$ with $|W| \leq (p/2)^{-\Delta}/2$*

$$|N_{B(H,\mathcal{F}_2,W)}(W)| \geq (p/2)^{\Delta} |W| \varepsilon n/4$$

holds.

P2 *Let $1 \leq k \leq \Delta$ and $\mathcal{L}$ be any collection of disjoint k-subsets of $\binom{V(H)}{r-1}$. If $|\mathcal{L}| \leq (p/2)^{-k}/2$, then for any $i = 1, \ldots, t$ with $V(\mathcal{L}) \cap V_i = \emptyset$ we have*

$$|N_{B(H,\mathcal{L},V_i)}(\mathcal{L})| \geq (p/2)^k |\mathcal{L}|\, |V_i|/4.$$

P3 *Let $1 \leq k \leq \Delta$ and $\mathcal{L}$ be any collection of disjoint k-subsets of $\binom{V(H)}{r-1}$. If $|\mathcal{L}| \geq C'(p/2)^{-k} \log n$, then for any $W \subseteq V(H) \setminus V(\mathcal{L})$ with $|W| \geq C'(p/2)^{-k} \log n$ the graph $B(H, \mathcal{L}, W)$ has at least one edge, where $C' = k(r-1) + 2$.*

6.3 Two lemmas

The proof of Theorem 2.9 follows immediately from the two lemmas we state below, Lemma 6.3 and Lemma 6.4. These lemmas establish the connection between $\mathcal{H}^{(r)}(n, p)$, good hypergraphs and $\mathcal{F}^{(r)}(n, \Delta)$-universality.

Lemma 6.3. *For integers $r \geq 2$, $\Delta \geq 1$, $t \geq 1$ and $\varepsilon \leq 1/(r\Delta)$, there exists a $C > 0$ such that for $p \geq C (\log n/n)^{1/\Delta}$ the random hypergraph $\mathcal{H}^{(r)}(n, p)$ is $(n, r, p, t, \varepsilon, \Delta)$-good a.a.s.*

Lemma 6.4. *For integers $r \geq 2$, $\Delta \geq 1$ and $\varepsilon \leq |P^{(r)}(\Delta)|^{-1} r^{-3} \Delta^{-3}$, there exists a $C > 0$ such that for $p \geq C (\log n/n)^{1/\Delta}$, every $\left(n, r, p, r^3\Delta^3, \varepsilon, \Delta\right)$-good hypergraph is $\mathcal{F}^{(r)}(n, \Delta)$-universal.*

Thus, it remains to prove both lemmas.

6.4 Proofs of auxiliary lemmas

Proof of Lemma 6.3

Let r, Δ, t and $\varepsilon \leq 1/(r\Delta)$ be given, furthermore we assume that $p \geq C\left(\log n/n\right)^{1/\Delta}$, where C is a sufficiently large constant that depends only on ε, r, Δ and t. We will not specify C explicitly but it will be clear from the context how it should be chosen.

We expose $\mathcal{H}^{(r)}(n,p)$ in two rounds and write $\mathcal{H}^{(r)}(n,p) = \mathcal{H}^{(r)}(n,p_1) \cup \mathcal{H}^{(r)}(n,p_2)$, where $p_1 = p_2 \geq p/2$ such that $(1-p) = (1-p_1)(1-p_2)$. In the first round we will find the families $\mathcal{F}$ and in the second round we show that **P1–P3** of Definition 6.2 all hold with probability at least $1-o(1)$. In the beginning we arbitrarily partition V into $V_0 \cup V_1 \cup \cdots \cup V_t$ such that $|V_0| = n - \varepsilon n/10$ and $V_i = \varepsilon n/(10t)$ for $i = 1, \ldots, t$.

1st round.. For a given profile $(Z, E_1, E_2) \in P^{(r)}(\Delta)$ we have that the maximum degree of $G = (Z, E_1)$ is at most $\Delta - 1$. We estimate $m_1(G) \leq \max_{s \geq r} \frac{(\Delta-1)s}{r(s-1)} \leq \Delta - 1$. Theorem 2.19 implies that there exist εn vertex-disjoint copies of G in $\mathcal{H}^{(r)}(n,p_1)$ all of whose vertices are contained inside V_0 a.a.s. Indeed, we apply Theorem 2.19 to $\mathcal{H}^{(r)}(|V_0|,p)$ where $|V_0| \geq (1-\varepsilon/10)n > (r-1)\Delta\varepsilon n + \varepsilon n/10$. We denote this family by $\mathcal{F}_{(Z,E_1)}$.

Since there are constantly many (at most $|P^{(r)}(\Delta)|$) r-uniform hypergraphs G on at most $(r-1)\Delta$ vertices with maximum degree $\Delta - 1$, we will find simultaneously εn vertex-disjoint copies of any such G a.a.s. within V_0. Therefore, with a given profile $(Z, E_1, E_2) \in P^{(r)}(\Delta)$, we associate a family $\mathcal{F}$ of εn vertex-disjoint copies (Y, E', E'') with $(Y, E') \in \mathcal{F}_{(Z,E_1)}$ and such that $(Y, E', E'') \cong (Z, E_1, E_2)$. This gives us a family $\mathcal{F}_2$ of copies of E'' for this kind of profile, thus showing the first part of **P1** from Definition 6.2.

2nd round.. From now on we work in $\mathcal{H} = \mathcal{H}^{(r)}(n,p_2)$. Fix some profile $(Z, E_1, E_2) \in P^{(r)}(\Delta)$ and the corresponding family $\mathcal{F}$ found in the first round. The family $\mathcal{F}$ induces a family $\mathcal{F}_2$ of disjoint copies of E_2 in $\binom{V_0}{r-1}$. Let W be a subset of $V(\mathcal{H}) \setminus V(\mathcal{F}_2)$ with $|W| \leq (p/2)^{-\Delta}/2$. For every $L \in \mathcal{F}_2$ let X_L be the random variable with $X_L = 1$ if and only if $L \subseteq \text{link}_{\mathcal{H}}(w)$ for some $w \in W$. This gives us $|N_{B(\mathcal{H},\mathcal{F}_2,W)}(W)| = \sum_{L \in \mathcal{F}_2} X_L$. The X_L are independent and since $\mathbb{P}[xL \in E(B(\mathcal{H},\mathcal{F}_2,W))] \geq p_2^\Delta \geq (p/2)^\Delta$, we compute using $|W| \leq (p/2)^{-\Delta}/2$

$$\mathbb{P}[X_L = 0] \leq (1-(p/2)^\Delta)^{|W|} \leq 1 - |W|(p/2)^\Delta + |W|^2(p/2)^{2\Delta} \leq 1 - |W|(p/2)^\Delta/2.$$

From this we obtain

$$\mathbb{E}\left[\sum_{L \in \mathcal{F}_2} X_L\right] \geq (p/2)^\Delta |W||\mathcal{F}_2|/2 \overset{|\mathcal{F}_2|=\varepsilon n}{\geq} \varepsilon(C/2)^\Delta |W|(\log n)/2$$

and using Chernoff's inequality, Theorem 2.16, with $\gamma = 1/2$ we get

$$\mathbb{P}\left[\sum_{L \in \mathcal{F}_2} X_L < (p/2)^\Delta |W|\,|\mathcal{F}_2|/4\right] \leq \exp(-\varepsilon(C/2)^\Delta |W|(\log n)/16) = n^{-\varepsilon(C/2)^\Delta |W|/16}. \tag{6.1}$$

Since there are at most n^s choices for a set W of size s we can bound, for C large enough, the probability that there is a set W violating **P1** for $\mathcal{F}_2$ by $o(1)$.

The number of different profiles in $P^{(r)}(\Delta)$ depends only on Δ and thus also the number of families $\mathcal{F}_2$. Thus taking the union bound over the probability that there is a set W violating our condition for some family $\mathcal{F}_2$ is still $o(1)$. This verifies **P1** of Definition 6.2.

To verify **P2** and **P3**, we use the edges of $\mathcal{H}^{(r)}(n, p_2)$. Let $k \in [\Delta]$, $\mathcal{L}$ be a collection of disjoint k-subsets of $\binom{V}{r-1}$ with $|\mathcal{L}| \leq (p/2)^{-k}/2$ and $i \in [t]$ such that $V(\mathcal{L}) \cap V_i = \emptyset$. For $v \in V_i$, let X_v be the random variable with $X_v = 1$ if and only if $L \subseteq \mathrm{link}_{\mathcal{H}}(v)$ for some $L \in \mathcal{L}$. Thus $|N_{B(\mathcal{H},\mathcal{L},V_i))}(\mathcal{L})| = \sum_{v \in V_i} X_v$. As above we obtain

$$\mathbb{P}[X_v = 0] \leq \left(1 - (p/2)^k\right)^{|\mathcal{L}|} \leq 1 - |\mathcal{L}|(p/2)^k + |\mathcal{L}|^2(p/2)^{2k} \leq 1 - |\mathcal{L}|(p/2)^k/2.$$

We have

$$\mathbb{E}\left[\sum_{v \in V_i} X_v\right] \geq (p/2)^k|\mathcal{L}||V_i|/2 \overset{|V_i| = \frac{\varepsilon n}{10t}}{\geq} \varepsilon(C/2)^k|\mathcal{L}|(\log n)/(20t)$$

and from Chernoff's inequality, Theorem 2.16, with $\gamma = 1/2$ we get

$$\mathbb{P}\left[\sum_{v \in V_i} X_v \leq (p/2)^k|\mathcal{L}||V_i|/4\right] \leq \exp(-(p/2)^k|\mathcal{L}||V_i|/16) \leq n^{-\varepsilon(C/2)^k|\mathcal{L}|/(160t)}.$$

There are less than $n^{(r-1)k|\mathcal{L}|}$ possibilities to choose $\mathcal{L}$. Therefore, for C large enough, the probability that there exists $k \in [\Delta]$ and sets $\mathcal{L}$ and V_i that violate **P2** is $o(1)$.

Finally, we verify that **P3** holds a.a.s. in $\mathcal{H}$. For this we set $\ell = C'(p/2)^{-k}\log n$ and let $k \in [\Delta]$. It suffices to consider only sets $\mathcal{L}$ and $W \subseteq V \setminus V(\mathcal{L})$ each of size ℓ. For two such sets $\mathcal{L}$ and W the probability that an edge in $B(\mathcal{H}, \mathcal{L}, W)$ is present equals $p_2^k \geq (p/2)^k$ and therefore the probability that there are no edges is at most $(1 - (p/2)^k)^{\ell^2} \leq \exp(-\ell^2(p/2)^k)$.

There are less than $n^{(r-1)k\ell}$ choices for $\mathcal{L}$ and less than n^ℓ choices for W. Thus, we can bound the probability that there are sets $\mathcal{L}$ and W of size ℓ violating **P3** by

$$\exp\Big(\big((r-1)k\ell + \ell\big)\ln n - \ell^2(p/2)^k\Big) = \exp\Big(\big((r-1)k + 1 - C'\big)\ell \ln n\Big) = o(1).$$

□

Proof of Lemma 6.4

Let r, Δ, $\varepsilon \leq |P^{(r)}(\Delta)|^{-1}r^{-3}\Delta^{-3}$ be given and let $C_{6.3}$ be a constant as asserted by Lemma 6.3 on input r, Δ, $t := r^3\Delta^3$ and ε. Furthermore, we assume that $p \geq C\,(\log n/n)^{1/\Delta}$, where C is a sufficiently large constant that depends only on ε, r, Δ and $C_{6.3}$. We will not specify C explicitly but it will be clear from the context how it should be chosen.

Let H be an $(n, r, p, t, \varepsilon, \Delta)$-good hypergraph and fix the partition $V_0 \cup V_1 \cup \cdots \cup V_t$ of $V(G)$ as specified by Definition 6.2. Fix any $F \in \mathcal{F}^{(r)}(n, \Delta)$ and apply Lemma 6.1 with r, Δ, $t = r^3\Delta^3$ and ε to obtain a partition of $V(F)$ in $X_0 \cup \cdots \cup X_t$ with **Q1–Q3** from Lemma 6.1.

An embedding of F into G is an injective map $\phi\colon V(F) \to V(H)$, where edges are mapped onto edges. We start with constructing an embedding ϕ_0 that maps X_0 into $V_0 \subset V(H)$. From **Q2**, we know that every $x \in X_t$ has the same profile in F. Therefore, let (Z, E_1, E_2) be the profile of any $x \in X_t$. By **P1**, there is a family $\mathcal{F}$ of copies of (Z, E_1, E_2) with vertices in V_0. Furthermore, $F[X_0]$ is the disjoint union of εn copies of (Z, E_1) which holds because of the 3-independence of X_t (refprop:indep of Lemma 6.1). Therefore, we can construct ϕ_0 by mapping bijectively every copy $(N_F(x), F[N_F(x)], \mathrm{link}_F(x))$ to one member (Y, E', E'') of $\mathcal{F}$. This is for sure a valid embedding of $F[X_0]$ into H.

Now we construct ϕ_i from ϕ_{i-1} for $i = 1, \dots, t$ by embedding X_i such that $\phi_i\left(F[\cup_{j=0}^{i} X_j]\right) \subseteq H$. The available vertices for this step are $V_i^* = (V_0 \cup \dots \cup V_i) \setminus \mathrm{Im}(\phi_{i-1})$. For $x \in X_i$ we collect the images of the already fully embedded $(r-1)$-sets from the $\mathrm{link}_F(x)$ in

$$L(x) := \left\{ \phi_{i-1}(e) : e \in \mathrm{link}_F(x) \cap \binom{\bigcup_{j=0}^{i-1} X_j}{r-1} \right\}.$$

Since X_i is 3-independent we have $L(x_1) \cap L(x_2) = \emptyset$ for $x_1, x_2 \in X_i$ and we set $\mathcal{L}_i = \{L(x) : x \in X_i\}$ which is a collection of vertex-disjoint sets in $\binom{V(H)}{r-1}$. A possible image for $x \in X_i$ is any $v \in V_i^*$, for which $L(x) \subseteq \mathrm{link}_H(v)$. It remains to find an $\mathcal{L}_i$-matching in $B_i = B(H, \mathcal{L}_i, V_i^*)$ since then we set $\phi_i(x) := v$ for every edge $vL(x)$ in this matching and, since any edge $e \in E(F)$ intersects X_i in at most one vertex, we obtain $\phi_i\left(F[\cup_{j=0}^{i} X_j]\right) \subseteq H$.

To guarantee an $\mathcal{L}_i$-matching in B_i we will verify Hall's condition. Let $U \subseteq \mathcal{L}_i$ and one needs to show that $|N_{B_i}(U)| \geq |U|$ holds. We assume $\emptyset \notin U$, because otherwise $N_{B_i}(U) = V_i^*$ and $|V_i^*| \geq |\mathcal{L}_i|$.

First, we verify Hall's condition for all sets U with $|U| \leq |V_i^*| - \varepsilon n/10$. Notice that there exists a $k \in [\Delta]$ and a subset $U' \subseteq U$ of size at least $|U|/\Delta$ and $|L| = k$ for every $L \in U'$. If $|U'| \leq (p/2)^{-k}/2$, then by **P2** we have for C large enough

$$|N_{B_i}(U)| \geq |N_{B_i}(U')| \geq (p/2)^k |U'||V_i|/4 \geq \varepsilon (C/2)^k |U| (\log n)/(40 t \Delta) \geq |U|.$$

If $(p/2)^{-k}/2 < |U'| < C'(p/2)^{-k} \log n$, then we take any subset $U'' \subseteq U'$ of size $(p/2)^{-k}/2$ and use again **P2** and $|U''| \geq 2|U|/(C'\Delta \log n)$ to obtain for C large enough

$$|N_{B_i}(U)| \geq |N_{B_i}(U'')| \geq (p/2)^k |U''||V_i|/4 \geq \varepsilon (C/2)^k |U|/(20 C' t \Delta) \geq |U|.$$

If $|U'| > C'(p/2)^{-k} \log n$, then $|U| > C'(p/2)^{-k} \log n$ and there are no edges between U and $V_i^* \setminus N_{B_i}(U)$ in B_i. Therefore, **P3** yields for C large enough

$$|V_i^* \setminus N_{B_i}(U)| < C'(p/2)^{-k} \log n \leq C'(C/2)^{-k} (n/\log n)^{k/\Delta} \log n \leq \varepsilon n/10,$$

and thus $|N_{B_i}(U')| > |V_i^*| - \varepsilon n/10$ which verifies Hall's condition in B_i for $|U| \leq |V_i^*| - \varepsilon n$.

For $i = 1, \dots, t-1$ it follows from $|\bigcup_{i=1}^{t} V_i| = \varepsilon n/10$ and $|X_t| = \varepsilon n$, that

$$|V_i^*| - |X_i| \geq (n - |\mathrm{Im}\, \phi_{i-1}| - \varepsilon n/10) - (n - |\mathrm{Im}\, \phi_{i-1}| - \varepsilon n) \geq 9/10 \varepsilon n$$

and therefore $|\mathcal{L}_i| = |X_i| \leq |V_i^*| - \varepsilon n/10$.

Therefore, we find $\mathcal{L}_i$-matchings in B_i for $i \in [t-1]$ one after each other extending at each step our embedding.

In the last step, we have $|V_t^*| = |X_t| = \varepsilon n$ and, by the partitioning of $V(F)$ with $X_0 = N_F(X_t)$ we have $\mathcal{L}_t = \mathcal{F}_2$, where $\mathcal{F}_2$ is the family guaranteed by **P1**. Since we already saw that $|N_{B_t}(U)| \geq |U|$ for all $U \subseteq \mathcal{L}_t$ with $|U| \leq |\mathcal{L}_t| - \varepsilon n/10$ in $B_t = B(H, \mathcal{L}_t, V_t^*)$, it suffices to verify $|N_{B_t}(W)| \geq |W|$ for all $W \subseteq V_t^*$ with $|W| \leq \varepsilon n/10$. If $|W| > (p/2)^{-\Delta}/2$ then we take an arbitrary subset $W' \subseteq W$ of size exactly $(p/2)^{-\Delta}/2$ and otherwise we set $W' := W$. By **P1**, we have

$$|N_{B_t}(W')| \geq (p/2)^{\Delta}|W'|\varepsilon n/4,$$

which is at least $\varepsilon n/8 > \varepsilon n/10 \geq |W|$ if $W' \subsetneq W$ and is at least $\varepsilon(C/2)^{\Delta}(\log n)|W'|/4 > |W|$ if $W = W'$. Therefore, $N_{B_t}(U) \geq |U|$ for all $|U| \geq |\mathcal{L}_t| - \varepsilon n/10$ as well and there exists a (perfect) $\mathcal{L}_t$-matching in B_t that allows us to finish embedding F into H. □

In the proofs of Lemmas 6.3 and 6.4 (and thus of Theorem 2.9) we only considered the case of constant Δ. Similarly to the arguments in [56] this also works in the range where Δ is some function of n but then the C in the bound on the probability is no longer a constant and rather growing exponentially with Δ. Furthermore, the proof yields a randomised polynomial time algorithm that on input $\mathcal{H}^{(r)}(n,p)$ embeds a.a.s. any given $F \in \mathcal{F}^{(r)}(n,\Delta)$ into $\mathcal{H}^{(r)}(n,p)$. All steps of the proof can be performed in polynomial time and the only place where we need to use additional random bits is to split $\mathcal{H}^{(r)}(n,p)$ into $\mathcal{H}^{(r)}(n,p_1) \cup \mathcal{H}^{(r)}(n,p_2)$ and this can be done similarly as was done in [4].

Chapter 7

Constructions of universal hypergraphs

In this last chapter we finally give the proofs[34] of Theorem 2.11, 2.13 and 2.14 aquired with Hetterich and Person [64]. The third theorem is the hardest of these three and we need to prove a new decomposition result for hypergraphs with maximum degree 2 (cf. Lemma 7.5 and 7.7). For all proofs we need the concept of hitting graphs first introduced in [94], which we will refine in the next section following [64]. But before we come to that let us briefly justify our lower bound on the number of edges in an $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraph as shown in [94][35].

As claimed above we first observe that *any* $\mathcal{F}^{(r)}(n,\Delta)$-universal r-uniform hypergraph must possess $\Omega(n^{r-r/\Delta})$ edges. Indeed, it follows e.g. from a result of Dudek, Frieze, Ruciński, and Šileikis [46] that for any $\Delta \geq 1$ and $r \geq 3$ the number of labelled r-uniform Δ-regular hypergraphs on n vertices (whenever $r|n\Delta$) is $\Theta\left(\frac{(\Delta n)!}{(\Delta n/r)!(r!)^{\Delta n/r}(\Delta!)^n}\right)$. Thus, the number of non-isomorphic r-uniform Δ-regular hypergraphs on n vertices is $\Omega\left(\frac{(\Delta n)!}{(\Delta n/r)!(r!)^{\Delta n/r}(\Delta!)^n n!}\right)$ and a similar bound holds for the cardinality of $\mathcal{F}^{(r)}(n,\Delta)$. On the other hand an r-uniform hypergraph with m edges contains exactly $\binom{m}{n\Delta/r}$ hypergraphs with $n\Delta/r$ edges. Thus, it holds $\binom{m}{n\Delta/r} = \Omega\left(\frac{(\Delta n)!}{(\Delta n/r)!(r!)^{\Delta n/r}(\Delta!)^n n!}\right)$ and solving for m yields $m = \Omega\left(n^{r-r/\Delta}\right)$.

The random hypergraph $\mathcal{H}^{(r)}(n,p)$ with $p = C(\log n/n)^{1/\Delta}$ is $\mathcal{F}^{(r)}(n,\Delta)$-universal (by Theorem 2.9) and has $\Theta(n^{r-1/\Delta}(\log n)^{1/\Delta})$ edges and thus the exponent in the density of $\mathcal{H}^{(r)}(n,p)$ is off by roughly a factor of r from the lower bound $\Omega\left(n^{-r/\Delta}\right)$ on the density for any $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraph. In the following we show how one can construct sparser $\mathcal{F}^{(r)}(n,\Delta)$-universal r-uniform hypergraphs out of the universal graphs from [9, 10].

7.1 Hitting graphs

Here we define a concept of hitting graphs first introduced in [94]. This will allow us later to obtain r-uniform universal hypergraphs out of universal hypergraphs of smaller uniformity.

Let $r \geq 3$ and $2 \leq s < r$ be integers. Given two s-uniform hypergraphs G and F and an r-graph H on the same vertex set as G, we say that *G hits H on F* if for all edges $f \in E(H)$ there is a copy of F in G induced on f, i.e. in $G[f]$. A family of s-uniform hypergraphs $\mathcal{G}$ *hits* a family of r-uniform hypergraphs $\mathcal{F}$ *on F* if for every $H \in \mathcal{F}$ there is a $G \in \mathcal{G}$ such that G hits H on F.

This concept allows us to reduce the uniformity from r to s keeping at the same time much of the information about H. This motivates a definition that allows us to recover all the edges of the hypergraph H which is being hit by G on F. For given s-uniform hypergraphs G and F let $\mathcal{H}_{(F,r)}(G)$

[34] The proofs that we give in this chapter are a close adaption of [64].
[35] The second and third paragraph are almost verbatim copies from [94].

be the r-graph on the vertex set $V(G)$ whose edges $f \in \binom{V(G)}{r}$ are such that a copy of F is contained in $G[f]$. Then G hits H on F if and only if $H \subseteq \mathcal{H}_{(F,r)}(G)$.

The following lemma establishes the connection between hitting hypergraphs and $\mathcal{H}_{(F,r)}(G)$. It is an extension of Lemma 5.2 from [94]. For completeness, we include its easy proof.

Lemma 7.1. *Let $r > s \geq 2$, $\Delta \geq 1$ be integers and F be an s-graph on at most r vertices. Further, let $\mathcal{F}$ be a family of r-uniform hypergraphs and $\mathcal{G}$ a family of s-uniform hypergraphs hitting $\mathcal{F}$ on F. If G' is a $\mathcal{G}$-universal s-graph, then $\mathcal{H}_{(F,r)}(G')$ is $\mathcal{F}$-universal.*

Proof. Let $H \in \mathcal{F}$ be an r-graph together with the s-graph $G \in \mathcal{G}$ that hits H on F. Since G' is $\mathcal{G}$-universal, there exists an embedding $\varphi\colon V(G) \to V(G')$ of G into G'.

It is now easy to see that φ is an embedding of H into $\mathcal{H}_{F,r}(G')$, and thus, $\mathcal{H}_{F,r}(G')$ is $\mathcal{F}$-universal. This can be seen as follows. For any edge $f \in E(H)$ there is a copy of F in $G[f]$. Since φ is an embedding of G into G', there is a copy of F in $G'[\varphi(f)]$. By the definition of $\mathcal{H}_{F,r}(G')$, $\varphi(f)$ is a hyperedge in $\mathcal{H}_{F,r}(G')$. Thus, φ is an embedding of H into $\mathcal{H}_{F,r}(G')$. □

The lemma above suggests a way of obtaining r-uniform universal hypergraphs out of hypergraphs of smaller uniformity. This will be exploited for particular choices of F in the following sections.

7.2 Proofs for general Δ

In this section we provide proofs of Theorem 2.11, Corollary 2.12 and Theorem 2.13, which are valid for all $\Delta \geq 2$.

Proof of Theorem 2.11

Let $r > r' \geq 2$ and $\Delta \geq 1$ be integers such that $r' \mid r$. We take F to be the r'-uniform perfect matching on r vertices (and thus with r/r' edges). Let $H \in \mathcal{F}^{(r)}(n, \Delta)$. Since every vertex lies in at most Δ edges there is an r'-graph $H' \in \mathcal{F}^{(r')}(n, \Delta)$ hitting H on F. Such an H' can be obtained from H by replacing every edge f of H with an arbitrary perfect r'-uniform matching on f. Therefore, $\mathcal{F}^{(r')}(n, \Delta)$ hits $\mathcal{F}^{(r)}(n, \Delta)$ on F.

Now if G' is $\mathcal{F}^{(r')}(n, \Delta)$-universal then, by Lemma 7.1, $\mathcal{H}_{(F,r)}(G')$ is $\mathcal{F}^{(r)}(n, \Delta)$-universal. Moreover, since any collection of r/r' independent edges from G' forms an r-edge in $\mathcal{H}_{(F,r)}(G')$, we have $e(\mathcal{H}_{(F,r)}(G')) \leq e(G)^{r/r'}$.

□

Proof of Corollary 2.12

If $r' \mid r$ and there exists an $\mathcal{F}^{(r')}(n, \Delta)$-universal hypergraph H' with $O(n^{r'-r'/\Delta})$ edges, then we immediately obtain an $\mathcal{F}^{(r)}(n, \Delta)$-universal hypergraph with the vertex set $V(H')$ and with

$$O\left((n^{r'-r'/\Delta})^{r/r'}\right) = O(n^{r-r/\Delta})$$

edges.

By Theorem 2.10 there exist optimal explicitly constructible $\mathcal{F}(n,\Delta)$-universal graphs on $O(n)$ vertices with $O(n^{2-2/\Delta})$ edges. This yields for even r an explicitly constructible optimal $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraph with $O(n^{r-r/\Delta})$ edges. A similar argument applies also for the case of explicitly constructible $\mathcal{F}(n,\Delta)$-universal graphs on n vertices with $O(n^{2-2/\Delta}\log^{4/\Delta} n)$ edges, giving $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraphs on n vertices with $O(n^{r-r/\Delta}\log^{2r/\Delta}(n))$ edges. □

We remark, that obtaining $\mathcal{F}^{(r')}(n,\Delta)$-universal hypergraphs on $O(n)$ vertices with $O(n^{r'-r'/\Delta})$ edges for r' being prime would provide then the conjectured optimal upper bound $O(n^{r-r/\Delta})$ for all r and Δ.

Proof of Theorem 2.13

In the case when r is odd, our hitting r'-uniform hypergraphs will be simply graphs, i.e. $r' = 2$. Moreover, the graph F can no longer be perfect matching, and thus we take F as the disjoint union of a matching on $r-3$ vertices and a path P_3 of length 2, i.e. a path with 2 edges. We remark, that the cases when $F = K_2$ (a single edge) and $F = K_r$ were considered in [94]. We use the following lemma which asserts that one can find a family of graphs with not too large maximum degree which hits $\mathcal{F}^{(r)}(n,\Delta)$ on F.

Lemma 7.2. *Let $r \geq 3$ be odd and $\Delta \geq 1$ be integers. Let F be the disjoint union of a matching on $r-3$ vertices and a path P_3. Then $\mathcal{F}(n,\lceil (r+1)\Delta/r\rceil)$ hits $\mathcal{F}^{(r)}(n,\Delta)$ on F.*

Proof. Let $H \in \mathcal{F}^{(r)}(n,\Delta)$. One defines an auxiliary bipartite incidence graph B as follows. The first class V_1 consists of $\lceil \Delta/r\rceil$ copies of $V(H)$ and the second class V_2 is equal to $E(H)$, while an edge of B corresponds to a pair (v,f), where v is some copy of a vertex from $V(H)$ and $f \in E(H)$ is such that $v \in f$. The vertices in V_1 have degree at most Δ and every hyperedge is connected to all $\lceil \Delta/r\rceil$ copies of its r vertices, i.e. the vertices from V_2 have degree $r\lceil \Delta/r\rceil \geq \Delta$. By Hall's condition, there is then a matching M covering V_2 and thus of size $e(H)$.

We build the hitting graph H' on the vertex set $V(H)$ by replacing edges $f \in E(H)$ through copies of F as follows. For every edge f in $E(H)$ we use the edge (v,f) of the matching M and place a copy of F on f such that the vertex v is the degree 2 vertex of the path P_3 from F while the other vertices are placed on $f \setminus \{v\}$ arbitrary. We see that each *placed* copy of F that contains v contributes 1 (in case $(v,f) \notin M$) or 2 (in case $(v,f) \in M$) to $\deg_{H'}(v)$. Since there are $\lceil \Delta/r\rceil$ copies of every vertex v and every vertex v lies in at most Δ edges of H, the maximum degree in H' is at most $\Delta + \lceil \frac{\Delta}{r}\rceil$ and therefore $\Delta(H') \leq \lceil (r+1)\Delta/r\rceil$. This implies $H' \in \mathcal{F}(n,\lceil (r+1)\Delta/r\rceil)$. □

For any $\mathcal{F}(n,\lceil (r+1)\Delta/r\rceil)$-universal graph G we use Lemma 7.1 to get an $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraph $H = \mathcal{H}_{(F,r)}(G)$ on the same number of vertices with at most $2|E(G)|^{(r-1)/2}\Delta(G)$ edges, where the bound comes from first choosing a matching on $r-1$ vertices and then one of the two possible endpoints enlarging one edge to a P_3. The maximum degree of universal graphs G in the constructions of Alon and Capalbo from Theorem 2.10 is $O(|E(G)|/|V(G)|)$, and thus we obtain Theorem 2.13 with $\mathcal{F}(n,\lceil (r+1)\Delta/r\rceil)$-universal graph G on $O(n)$ vertices with $O(n^{2-2/\lceil (r+1)\Delta/r\rceil})$ edges since

$$O\left((n^{2-2/\lceil (r+1)\Delta/r\rceil})^{(r-1)/2}\cdot n^{1-2/\lceil (r+1)\Delta/r\rceil}\right) = O\left(n^{r-(r+1)/\lceil (r+1)\Delta/r\rceil}\right).$$

A similar calculation yields $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraphs on n vertices with

$$O(n^{r-(r+1)/\lceil (r+1)\Delta/r\rceil} \log^{2(r+1)/\lceil (r+1)\Delta/r\rceil} n)$$

edges, which we obtain from $\mathcal{F}(n, \lceil (r+1)\Delta/r\rceil)$-universal graphs G on n vertices with

$$O(n^{2-2/\lceil (r+1)\Delta/r\rceil} \log^{4/\lceil (r+1)\Delta/r\rceil} n)$$

edges. □

In contrary to the F chosen as a matching plus P_3 we could work with any forest F. To find hitting graphs of small maximum degree we can use similar matching techniques and counting arguments, but in general it is not clear how low we can get. For example, if F is the path P_r on r vertices one can show that $\mathcal{F}(n, \lceil 2(r-1)\Delta/r\rceil)$ hits $\mathcal{F}^{(r)}(n,\Delta)$ on F. This leads to an $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraph on $O(n)$ vertices with $O(n^{r-2(r-1)/\lceil 2(r-1)\Delta/r\rceil})$ edges. It depends on the values of r and Δ, which bound is better, but one does not get anything significantly better than $O\left(n^{r-(r+1)/\lceil (r+1)\Delta/r\rceil}\right)$ edges and therefore we do not further pursue this here.

Reducing the number of vertices

Note that it is possible to reduce the number of vertices from $O(n)$ to $(1+\varepsilon)n$ in Theorems 2.10, 2.13, and 2.14 and in Corollary 2.12, for any fixed $\varepsilon > 0$, by using a *concentrator* as was done in [12]. Consider the $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraph H on $O(n)$ vertices and with m edges. A *concentrator* is a bipartite graph C on the vertex sets $V(H)$ and Q, where $|Q| = (1+\varepsilon)n$ such that for every $S \subseteq V(H)$ with $|S| \leq n$ we have $|N(S)| \geq |S|$ and every vertex from $V(H)$ has $O_\varepsilon(1)$ neighbours in C. We define a new hypergraph H' on Q by taking all sets $f' \in \binom{Q}{r}$ as edges for which there exists a perfect matching in C from an edge $f \in E(H)$ to f'. Since every vertex from $V(H)$ has $O_\varepsilon(1)$ degree in C, the hypergraph H' has $O_\varepsilon(m)$ edges. It is also not difficult to see that H' is $\mathcal{F}^{(r)}(n,\Delta)$-universal. Indeed, let $F \in \mathcal{F}^{(r)}(n,\Delta)$ and let $\varphi\colon V(F) \to V(H)$ be its embedding into H. By the property of the concentrator C, there is a matching of $\varphi(V(F))$ in C which we can describe by an injection $\psi\colon \varphi(V(F)) \to V(H')$. But now, by construction of H', $\psi \circ \varphi$ is an embedding of F into H'.

7.3 Proof for $\Delta = 2$

At this point in all cases where r is not even and r does not divide Δ we do not have constructions of $\mathcal{F}^{(r)}(n,\Delta)$-universal hypergraphs that match the lower bound $\Omega(n^{r-r/\Delta})$ on the number of edges. In this section we will deal with the *smallest* open case $\Delta = 2$ by constructing optimal $\mathcal{F}^{(r)}(n,2)$-universal hypergraphs on $O(n)$ vertices with $O(n^{r/2})$ edges. So, for example, if $r = 3$ then Theorem 2.13 yields $\mathcal{F}^{(3)}(n,2)$-universal hypergraphs on $O(n)$ vertices with $O(n^{3-4/\lceil 8/3\rceil}) = O(n^{5/3})$ edges, while the lower bound is $\Omega(n^{3/2})$.

We will first deal with the case $r = 3$ and $\Delta = 2$ and then reduce the case of general r and $\Delta = 2$ to this one. Let us say a few words how an improvement from $O(n^{5/3})$ to $O(n^{3/2})$ can be accomplished. We will use the concept of a graph G that hits some hypergraph H on P_3 (the path on 3 vertices). If we would follow the arguments in the previous section, then we see that taking a hypergraph

$H \in \mathcal{F}^{(3)}(n,2)$ and replacing every hyperedge by P_3 we can obtain a hitting graph G of maximum degree 3 and of average degree $8/3$. Thus, if we would like to use Theorem 2.10 we need to consider $\mathcal{F}(n,3)$-universal graphs, which results in the loss of some $n^{1/6}$-factor in the edge density. Instead, we will seek to decompose the hitting graph G into appropriate subgraphs G_1, G_2, G_3 and G_4 such that every edge of G lies in *exactly* three of the graphs G_i. A decomposition result of Alon and Capalbo from [9] will assist us in this. Finally, following closely the arguments again due to Alon and Capalbo but now from [10] will allow us to construct a universal graph $\mathcal{G}$ on $O(n)$ vertices and with maximum degree $O(n^{1/4})$ for a carefully chosen family $\mathcal{F}'$ of graphs allowing a decomposition as above, which hits $\mathcal{F}^{(3)}(n,2)$ on P_3. Lemma 7.1 implies then that $\mathcal{H}_{P_3,3}(G)$ is $\mathcal{F}^{(3)}(n,2)$-universal and has $O(n^{3/2})$ edges.

A graph decomposition result

The following notation is from [9]. Let G be a graph and $S \subseteq V(G)$ be a subset of its vertices. A graph G' which is obtained from G by adding additionally $|S|$ new vertices to G and placing an (arbitrary) matching between these new vertices and the vertices from S is called an *augmentation* of G. We call a graph *thin* if every of its components is an augmentation of a path or a cycle, or if they contain at most two vertices of degree 3. We also call any subgraph of a thin graph thin.

The following decomposition theorem may be seen as a generalisation of Petersen's Theorem to graphs of odd degree.

Theorem 7.3 (Theorem 3.1 from [9])**.** *Let Δ be an integer and G a graph with maximum degree Δ. Then there are Δ spanning subgraphs $G_1, \ldots, G_\Delta$ such that each G_i is thin and every edge of G appears in precisely two graphs G_i.*

Its proof is built on the Gallai-Edmonds decomposition theorem and is implied by the following lemma.

Lemma 7.4 (Lemma 3.3 from [9])**.** *Let $\Delta \geq 3$ be an odd integer and G a Δ-regular graph. Then G contains a spanning subgraph in which every vertex has degree 2 or 3 and every connected component has at most 2 vertices of degree 3.*

We will use the two results above to prove the existence of a hitting graph G with nice properties so that we can later take advantage of them when constructing a universal graph for the family of such *nice* hitting graphs.

Lemma 7.5. *Let $H \in \mathcal{F}^{(3)}(n,2)$. Then there exists a graph G that hits H on P_3 with the following properties:*

(i) there are spanning subgraphs G_1, G_2, G_3 and G_4 of G such that every G_i is an augmentation of a thin graph, and

(ii) every edge lies in exactly *three of the G_i.*

Proof. Let $H \in \mathcal{F}^{(3)}(n,2)$. We assume first that H is linear, i.e. edges are always intersecting in at most one vertex. Further, we assume that H is 2-regular[36].

[36]Otherwise we add *dummy* vertices and edges and obtain a 2-regular hypergraph, and, once the desired graph G is constructed, we delete these dummy vertices from G.

The rough outline of the proof is to find a graph G that hits H on P_3 and such that G contains a matching M so that $G \setminus M$ is an augmentation of a thin graph and if we contract the matching edges from M in G we obtain a graph of maximum degree at most 3. Decomposing such contracted graph via Theorem 7.3 into thin graphs G_1', G_2' and G_3' and then *recontracting* edges yields the desired family $G_1, \ldots, G_4$ (where $G_4 = G \setminus M$).

Let H^* be the line graph of H, that is $V(H^*) = E(H)$ and $e \neq f \in E(H)$ form an edge ef in H^* if $e \cap f \neq \emptyset$. Thus, H^* is a 3-regular graph on $2n/3$ vertices. Lemma 7.4 asserts then the existence of a matching M^* in H^* such that in $H^* \setminus M^*$ every component has at most 2 vertices of degree 3 and all other vertices have degree 2. Such a decomposition implies thus that every component of $H^* \setminus M^*$ is either a cycle, or has exactly two vertices, say a and b, of degree 3, so that either there are 3 internally vertex-disjoint paths between a and b or there is one path between a and b and, additionally, a and b lie on vertex-disjoint cycles (which also do not contain inner vertices from the path between a and b). We assume that a and b are not adjacent, because otherwise we could add the edge ab to M^*, splitting this component into two cycles.

From the matching M^* we define a subset $D := \{v \colon e \cap f = \{v\} \text{ where } ef \in E(M^*)\}$. Since M^* is a matching in the line graph of H it follows that no two vertices from D lie in an edge from H.

We denote by H_D the hypergraph which we obtain from H if we delete from the edges of H the vertices in D but we keep the edges, obtaining thus a hypergraph on the vertex set $V(H) \setminus D$, whose edges have cardinality 2 or 3. Thus, if ef is an edge in H^* and $e \cap f = \{v\}$ then the deletion of v from e and f implies that the edges $e \setminus \{v\}$ and $f \setminus \{v\}$ are no longer adjacent in the line graph $(H_D)^*$, which corresponds to the deletion of the edge ef in H^*. This implies that every component of $H^* \setminus M^*$ corresponds to a component of H_D, and therefore in every component of H_D there are at most two edges of cardinality 3 and all other edges have cardinality exactly 2. Again, the structure of every component of H_D is thus either a (graph) cycle, or there are exactly two edges, say g and h, of cardinality 3, with $g \cap h = \emptyset$ and there are three vertex-disjoint (graph) paths that connect the vertices from $g \cup h$.

Finally, we come to the definition of the hitting graph G. For every component C of H_D, let D_C be the vertices that have been deleted from the hyperedges in H that lie now in H_D. Thus, there is a (natural) map ψ_C between the edges from C of cardinality 2 and D_C: $\psi_C(f) = v$ if $\{v\} \cup f \in E(H)$. Note that this map is not necessarily injective. Since every vertex from D lies in exactly two edges of H, it will suffice to explain how we replace the 3-uniform edges of H_D and the edges of H incident with D by paths P_3. If C is the (graph) cycle, then we replace every edge of the form $\{v\} \cup f$, where $\psi_C(f) = v$, by P_3 so that the graph G_C obtained contains all the edges from $E(C)$ and is such that $\Delta(G_C) \leq 3$ and the vertices from D_C have degree at most 2 in G_C. If C contains exactly two 3-uniform edges (say g and h), then it is possible to replace the edges g, h and every edge of the form $\{v\} \cup f$, where $\psi_C(f) = v$, by P_3 such that the graph G_C satisfies the following: It contains all 2-uniform edges of C, is such that $\Delta(G_C) \leq 3$, the vertices from D_C have degree at most 2 in G_C and $G_C \setminus D_C$ is connected and has exactly two vertices of degree 3[37]. The graph G is then the union of all G_C and observe that G_C and $G_{C'}$ intersect in $D_C \cap D_{C'}$ for $C \neq C'$ and in particular have no common edges. Furthermore, every vertex from D has degree 2 in G, since it is an image of ψ_C precisely twice.

Let M be a matching in G that saturates D. Such a matching exists since D is independent in G (no

[37]This is easily done by considering the structure of the components C from H_D described in the previous paragraph.

two vertices from D lie in an edge from H), every vertex of D is connected to a vertex of degree 2 in $G \setminus D$ and $\deg(G) \leq 3$. By the definition of G above, every component in $G \setminus M$ is an augmentation of a graph with at most two vertices of degree 3, and thus an augmentation of a thin graph. We set $G_4 := G \setminus M$. Next, we contract the edges of M in G obtaining the graph G/M. Since M saturates D, which are vertices of degree 2 in G, it follows that G/M has maximum degree at most 3. Theorem 7.3 yields a decomposition of G/M into thin graphs G_1', G_2', G_3' such that every edge of G/M appears in precisely two of the graphs. Now we reverse the recontraction procedure. This leads to three graphs G_1, G_2 and G_3 where every edge of $G \setminus M$ appears in exactly two of the graphs, every edge from M appears in all three of them, and each of the G_1, G_2 and G_3 is an augmentation of a thin graph. Together with the graph $G_4 = G \setminus M$ we thus constructed the desired decomposition of a hitting graph G.

If H is not linear, then things get in some sense even easier, so we shall be brief. We proceed essentially in the same way. That is, we define the line graph H^* of H, which is now not necessarily 3-regular, but whose maximum degree is at most 3. Again, Lemma 7.4 asserts then the existence of a matching M^* in H^* such that in $H^* \setminus M^*$ every component has at most 2 vertices of degree 3 and all other vertices have degree at most 2. We then define the set D as before but in the case that the edge $ef \in M^*$ with, say, $e = \{a, b, c\}$ and $f = \{b, c, d\}$ we simply replace the edge e by $\{a, b\}$ and f by $\{c, d\}$ without putting anything into D. Once the components of H_D are identified and the graphs G_C are defined we add the edge bc (which we call nonlinear) to those graphs G_C, which contain either b or c (or both). Then we choose edges into the matching M as before and add all nonlinear edges such as bc to M. The rest of the argument remains the same. □

Recall that the ℓ-th power of a graph G, denoted by G^ℓ, is the graph on $V(G)$, whose vertices at distance at most ℓ in G are connected. It is not difficult to see that a thin graph on n vertices can be embedded into P_n^4, and thus, an augmentation of a thin graph into P_n^8. This motivates the following general definition.

Definition 7.6 ((k, r, ℓ)-decomposable graphs)**.** *Let k, r and ℓ be integers. A graph G on n vertices is called* (k, r, ℓ)-decomposable *if there exist k graphs G_i with the following properties. Every edge of G appears in exactly r of the G_i and there are maps $g_i \colon G_i \to [n]$, which are injective homomorphisms from G_i into P_n^ℓ. Then we denote by $\mathcal{F}_{k,r,\ell}(n)$ the family of (k, r, ℓ)-decomposable graphs on n vertices.*

We can restate our Lemma 7.5 in the following slightly weaker form.

Lemma 7.7. *The family $\mathcal{F}_{4,3,8}(n)$ hits $\mathcal{F}^{(3)}(n, 2)$ on a path P_3.*

This lemma implies that it is the family $\mathcal{F}_{4,3,8}(n)$ for which a universal graph is needed. This graph will be constructed in the section below and briefly explained why a desired embedding works, which will follow from the results of Alon and Capalbo from [10].

Constructions of universal graphs

First we briefly describe the construction from [10] of $\mathcal{F}(n, k)$-universal graphs on $O(n)$ vertices with $O(n^{2-2/k})$ edges. One chooses $m = 20n^{1/k}$, a fixed $d > 720$ and a graph R to be a d-regular graph on m vertices with the absolute value of all but the largest eigenvalues at most λ (such graphs are called

(n, d, λ)-graphs). One can assume that $\lambda \leq 2\sqrt{d-1}$ (then R is called Ramanujan) and $\text{girth}(R) \geq \frac{2}{3}\log m/\log(d-1)$. Explicit constructions of such Ramanujan graphs have been found first for $d-1$ being a prime congruent to 1 mod 4 in [85, 88]. Finally, the graph $G_{k,n}$ is defined on the vertex set $V(R)^k$ where two vertices $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ are adjacent if and only if there are at least two indices i such that x_i and y_i are within distance 4 in R. It is easily seen that such a graph $G_{k,n}$ has $O(n)$ vertices, $O(n^{2-2/k})$ edges and maximum degree $O(n^{1-2/k})$.

The first step in the proof of $\mathcal{F}(n,k)$-universality of $G_{k,n}$ is Theorem 7.3 implying that any graph F with $\Delta(F) \leq k$ is $(k, 2, 4)$-decomposable. In what follows we summarise a straightforward generalisation of the central claim from [10] (which is inequality (3.1) there), from which an existence of embedding of any graph $G \in \mathcal{F}(n,k)$ into $G_{k,n}$ follows. Its proof can be taken almost verbatim from [10].

Lemma 7.8. *Let $k \geq 3$, r and ℓ be natural numbers. For any choice of k permutations $g_i \colon [n] \to [n]$ there are k homomorphisms $f_i \colon [n] \to V(R)$ from the path P_n to the Ramanujan graph R introduced above such that the map $f \colon [n] \to V(G_{k,r,\ell}(n))$ defined by $f(v) = (f_1(g_1(v)), \ldots, f_k(g_k(v)))$ is injective.*

More precisely, the f_i's are inductively constructed as non-returning walks preserving the property that for any i vertices $v_1, \ldots, v_i \in V(G)$, $i \leq k$, one has

$$|\{v \in [n] : f_1(g_1(v)) = v_1, \ldots, f_i(g_i(v)) = v_i\}| \leq n^{(k-i)/k}.$$

For the last step $i = k$ this is equivalent to injectivity.

Finally, we explain, how we obtain $\mathcal{F}_{k,r,\ell}(n)$-universal graphs. The choice of the Ramanujan graph R along with the parameters m and d remains the same. The graph $G_{k,r,\ell}(n)$ is defined on the vertex set $V(R)^k$ and two vertices $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ are adjacent if and only if there are at least r indices i such that x_i and y_i are within distance ℓ in R. It is then an easy calculation to show that $G_{k,r,\ell}(n)$ has $O(n)$ vertices, at most $n\binom{k}{r}d^{r\ell}m^{k-r} = O(n^{2-r/k})$ edges and maximum degree $O(n^{1-r/k})$, where the constants in the O-notation depend on k, r, ℓ and d. Lemma 7.8 implies then the following.

Theorem 7.9. *Let $k \geq 3$, r and ℓ be natural numbers. The graph $G_{k,r,\ell}(n)$ is $\mathcal{F}_{k,r,\ell}(n)$-universal.*

Proof. Let G be a (k, r, ℓ)-decomposable graph on n vertices together with a decomposition $G_1, \ldots, G_k$ and an injective homomorphisms $g_i \colon V(G_i) \to [n]$ from G_i into P_n^ℓ. Lemma 7.8 asserts the existence of the homomorphisms $f_i \colon [n] \to V(R)$ from P_n to R for every $i \in [k]$, so that the map $f \colon V(G) \to V(G_{k,r,\ell}(n))$ given by $f(v) = (f_1(g_1(v)), \ldots, f_k(g_k(v)))$ is injective.

It is clear that the composition of f_i with g_i is a homomorphism from G_i to R^ℓ. Furthermore, every edge $\{u, v\}$ from G lies in r graphs G_i. Thus, there are r indices i such that $g_i(u)$ and $g_i(v)$ are distinct and within distance ℓ in P_n. This implies that $f_i(g_i(u))$ and $f_i(g_i(v))$ are also distinct and within distance ℓ in G. By the definition of $G_{k,r,\ell}(n)$ this implies that $f(u)$ and $f(v)$ are adjacent in $G_{k,r,\ell}(n)$ and f is the desired embedding of G into $G_{k,r,\ell}(n)$. $\square$

From this, Theorem 2.14 follows immediately for $r = 3$.

Proof of Theorem 2.14, case $r = 3$. Note, that the graph $G_{4,3,8}(n)$ has $m^4 = O(n)$ vertices and $O(nm) = O(n^{5/4})$ edges. By Theorem 7.9 $G_{4,3,8}(n)$ is $\mathcal{F}_{4,3,8}(n)$-universal, and since $\mathcal{F}_{4,3,8}(n)$ hits $\mathcal{F}^{(3)}(n, 2)$ on

P_3, Lemma 7.1 implies that $\mathcal{H}_{P_3,3}(G_{4,3,8}(n))$ is $\mathcal{F}^{(3)}(n,2)$-universal, has $O(n)$ vertices and $O(n^{3/2})$ edges. This proves the case $r = 3$. □

We believe that the constructions from [9] can also be adapted to work with (k, r, ℓ)-decomposable graphs. For the cases discussed here this would lead to universal graphs on n vertices, where the number of edges is some polylog factor larger.

$\mathcal{F}^{(r)}(n,2)$-universal hypergraphs of uniformity $r \geq 5$

Proof of Theorem 2.14 for odd $r \geq 5$. First we define the hypergraph $\mathcal{H}$ which will then turn out to be $\mathcal{F}^{(r)}(n,2)$-universal. Let $t = (r-3)/2$. Let $G_1, \ldots, G_{t+1}$ be vertex-disjoint graphs, where $G_1, \ldots, G_t$ are copies of C_n^4 (the fourth power of the cycle C_n) and G_{t+1} is a copy of the graph $G_{4,3,8}(n)$, introduced in the previous section. Furthermore, we add on top of G_{t+1} another graph G^*_{t+1} containing as edges all pairs of vertices which have a common neighbour in G_{t+1}. We define $\mathcal{H}$ to be the r-graph on the vertex set $\dot\bigcup_{i=1}^{t+1} V(G_i)$, and the edges are r-element subsets f such that, with $f_i := f \cap V(G_i)$, we have $|f_i| \leq 3$ and each $G_i[f_i]$ contains a copy of $P_{|f_i|}$, a path on $|f_i|$ vertices (thus, P_0 is the empty graph, $P_1 = K_1$ and $P_2 = K_2$). Additionally, in the case $|f_{t+1}| = 2$, we allow f_{t+1} to be an edge (i.e. P_2) in G^*_{t+1} instead of G_{t+1}.

Certainly, $\mathcal{H}$ has $O(n)$ vertices. How many edges does the hypergraph $\mathcal{H}$ contain? For this we need to choose paths P_{ℓ_i} from every G_i (respectively G^*_{t+1}) such that $\ell_i \in \{0,1,2,3\}$ and $\sum_{i=1}^{t+1} \ell_i = r$. Because $G_1, \ldots, G_t$ have maximum degree 8, G_{t+1} has maximum degree $O(n^{1/4})$, and G^*_{t+1} has maximum degree $n^{1/2}$, we compute the number of edges of $\mathcal{H}$ to be $O(n^{t+1}n^{2/4}) = O(n^{r/2})$, as desired.

Given a hypergraph H and a subset of vertices $X \subseteq V$, we denote through $H(X)$ the (not necessarily uniform) hypergraph on the vertex set X, whose edges are restrictions to X, i.e. $E(H(X_i)) = \{f \cap X_i \colon f \in E(H)\}$.

The rest of the proof hinges on the following auxiliary lemma (whose proof can be found below) and the case $r = 3$ of Theorem 2.14 shown in the previous section.

Lemma 7.10. *Let $H \in \mathcal{F}^{(r)}(n,2)$ and $t = (r-3)/2$. Then there exists a partition of the vertex set of H into disjoint subsets $X_1, \ldots, X_{t+1}$, such that $H(X_1), \ldots, H(X_{t+1})$ have maximum vertex degree 2 and contain hyperedges of cardinality at most 3. Moreover in $H(X_1), \ldots, H(X_t)$ every component contains at most 2 hyperedges of size 3.*

Let us see how then H can be embedded into the hypergraph $\mathcal{H}$. Owing to the special structure of $H(X_1), \ldots, H(X_t)$, one can easily find injective maps $g_i \colon X_i \to V(G_i)$, such that every hyperedge $f \in E(H(X_i))$ is such that $G_i[g_i(f)]$ contains a path $P_{|f|}$. This can be seen by replacing f in $H(X_i)$ through an arbitrary path $P_{|f|}$ obtaining thus the graph G'_i on the vertex set X_i. Then, since in every component of $H(X_i)$ there are at most two edges of size 3, it is easy to find an injective graph homomorphism from G'_i into G_i.

For $H(X_{t+1})$ we can assume first that it is 3-uniform and lies in $\mathcal{F}^{(3)}(n,2)$ by adding some *dummy* vertices appropriately (but still using the notation $H(X_{t+1})$). The $\mathcal{F}_{4,3,8}(n)$-universality of $G_{t+1} = G_{4,3,8}(n)$ and the fact that $\mathcal{F}_{4,3,8}(n)$ hits $H(X_{t+1})$ on P_3 yields an injective map $g_{t+1} \colon X_{t+1} \to V(G_{t+1})$ such that $G_{t+1}[g_{t+1}(f)]$ contains P_3 for every $f \in E(H(X_{t+1}))$. Deleting the dummy vertices (but keeping the edges) we see that g_{t+1} remains injective and $G_{t+1}[g_{t+1}(f)]$ contains $P_{|f|}$ for every $f \in$

$E(H(X_{t+1}))$ except possibly for the case, when the center vertex of some P_3 was deleted (being a dummy vertex). But in this case we observe that $G^*_{t+1}[g_{t+1}(f)]$ induces P_2 instead, because both vertices of $g_{t+1}(f)$ were incident to the deleted vertex in G_{t+1}.

It should be clear that $g\colon V(H) \to V(\mathcal{H})$ with $g|_{X_i} = g_i$, for all $i \in [t+1]$, is injective. It remains to show that g is a homomorphism into $\mathcal{H}$. Given an edge e of H, by the definition of $H(X_i)$ and the choices of g_i's, we see that $e \cap X_i \in E(H(X_i))$ and $G_i\,[g_i(e \cap X_i)]$ contains a path $P_{|e \cap X_i|}$ for all i, except possibly for the case when $|g_{t+1}(e \cap X_{t+1})| = 2$. But in this case one must necessarily have $g_{t+1}(e \cap X_{t+1}) \in E(G^*_{t+1})$. These conditions fulfill exactly the requirement for $g(e)$ to be the edge in $\mathcal{H}$. Thus, g embeds H into $\mathcal{H}$. □

Finally, we provide the proof for the auxiliary lemma above, Lemma 7.10.

Proof of Lemma 7.10. Let $H \in \mathcal{F}^{(r)}(n, 2)$. Again we assume first that H is linear and 2-regular. We consider, as in the case $r = 3$, the line graph H^*, which is r-regular now. Hence Lemma 7.4 yields a spanning subgraph H^*_1, in which every vertex has degree 2 or 3 and every component has at most 2 vertices of degree 3.

If C is a component of H^*_1, then we define the set V_C as all vertices v such that $\{v\} = e \cap f$ for some $ef \in E(C)$ (recall that H is assumed to be a linear hypergraph). We set $X_1 = \cup V_C$ where the union is over all components C of H^*_1 and then the set $\{v\colon \{v\} = e \cap f \text{ for some } ef \in E(C)\}$ is an edge of $H(X_1)$ for every edge $f \in E(H)$. Observe, that these edges have cardinality either 2 or 3. Indeed, a vertex of degree j in some component C is the edge of H that intersects j other edges of H in different vertices, which give rise to a j-uniform edge in $H(X_1)$. By construction, $H(X_1)$ is linear and 2-regular. Crucially, the components of $H(X_1)$ have simple structure, since these are *inherited* from the components C. More precisely, each component of $H(X_1)$ has at most two 3-uniform edges and all other edges have cardinality 2.

We denote by $\tilde{H}_1 = H(V(H) \setminus X_1)$ the hypergraph obtained from H by deleting from its edges all vertices from X_1 (we call this procedure as *reducing uniformity*). It should be clear that, in this way every edge of H can be written uniquely as the union of one edge of $H(X_1)$ and the other from $\tilde{H}_1$. Since $H(X_1)$ is not necessarily uniform, the hypergraph $\tilde{H}_1$ is now a not necessarily uniform hypergraph as well, but its edges have cardinalities either $r-3$ or $r-2$.

The next step calls for an inductive procedure with a blemish, that $\tilde{H}_1$ is not necessarily uniform. But this can be remedied by adding *dummy* vertices and edges to $\tilde{H}_1$ and obtaining an $(r-2)$-uniform linear hypergraph still denoted by $\tilde{H}_1$ which is 2-regular[38]. We keep doing this reduction until we arrive at the hypergraph $\tilde{H}_t$ where $t = (r-3)/2$, thereby generating $X_2, \ldots, X_t$ and $\tilde{H}_2(X_2) \ldots, \tilde{H}_{t-1}(X_t)$. Finally we get $X_{t+1} := V(H) \setminus \cup_{i=1}^{t} X_i$ and a 3-uniform linear hypergraph $\tilde{H}_t$ on X_{t+1}, which is 2-regular.

Before we proceed, let us summarise what we achieved so far. We have found hypergraphs $H(X_1)$, $\tilde{H}_2(X_2), \ldots, \tilde{H}_{t-1}(X_t)$, so that each of them is linear, 2-regular and its edge uniformities are either 2 or 3 and each of its components has simple structure (recall: each component has at most two 3-uniform edges and all other edges have cardinality 2). Furthermore $\tilde{H}_t$ is a 3-uniform linear hypergraph, which is 2-regular, and the vertex sets $X_1, \ldots, X_{t+1}$ are a partition of $V(H)$.

[38]Once we are finished with the decomposition, we will reduce the uniformity by deleting these dummy vertices from edges, but keeping the altered edges.

We finally obtain the promised family $H(X_1), \dots, H(X_{t+1})$. This can be seen as reducing uniformities of the hypergraphs $H(X_1), \tilde{H}_2(X_2), \dots, \tilde{H}_{t-1}(X_t)$ and $\tilde{H}_t$ by deleting dummy edges and dummy vertices from the edges. In this way it may happen, that the uniformity of some edges of the hypergraph family will be reduced to 0 (in which case they disappear from that particular hypergraph), while some others will be reduced to 1, in which case we get edges of the type $\{v\}$, which we will use.

The case when H is not a linear hypergraph can be treated similarly. We slightly extend the definition of the line graph H^* such that it contains multiple edges, i.e. for $e, f \in E(H)$ there are $|e \cap f|$ edges between e and f in H^* and we label each of them with a distinct vertex from $e \cap f$. Then H^* is again r-regular and we can again apply Lemma 7.4, because the proof from [9] extends verbatim to multigraphs. In this way we obtain a multigraph H_1^* and for every component C we define the vertex set V_C as follows: for a given edge $g \in E(C)$, the set V_C contains $e \cap f$ where the edge g connects e and f and it holds $|e \cap f| = 1$, and otherwise (i.e. there are parallel edges to g) the vertex set V_C contains precisely the vertex of the label that the edge g carries. The set X_1 is then the union of the V_C over all components C from H_1^*. The construction of $X_2, \dots, X_t$ is similar to above. The rest of the proof proceeds along the lines of the linear case and we omit further details. □

A general problem

To prove the embedding for other parameters of r and Δ we would need the analogue of Lemma 7.7, that is, a solution to the following problem.

Problem 7.11. *Let $r \geq 3$ and $\Delta \geq 3$ be integers. Find ℓ such that $\mathcal{F}_{(r-1)\Delta,r,\ell}(n)$ hits $\mathcal{F}^{(r)}(n, \Delta)$ on P_r.*

It is immediate that, with $k = (r-1)\Delta$, Theorem 7.9 yields $\mathcal{F}_{(r-1)\Delta,r,\ell}(n)$-universal graphs $G = G_{(r-1)\Delta,r,\ell}$ on $O(n)$ vertices with $O(n^{2-r/((r-1)\Delta)})$ edges and maximum degree $O(n^{1-r/((r-1)\Delta)})$. From this the solution to Problem 7.11 would yield optimal universal hypergraphs on $O(n)$ vertices with $|V(G)|(|E(G)|/|V(G)|)^{r-1} = O(n^{r-r/\Delta})$ edges. Clearly, the interesting cases are $\Delta \geq 3$, $r \nmid \Delta$ and r odd.

An alternative to our approach is to extend the constructions for universal graphs from [9, 10, 12] to hypergraphs. To follow a similar embedding scheme one would ask for appropriate decomposition results for hypergraphs. For example, for $H \in \mathcal{F}^{(3)}(n, 2)$ the task is to find subhypergraphs $H_1, \dots, H_4$ which are *thin* and such that every hyperedge appears in exactly three of them.

7.4 Proof for $\mathcal{E}^{(r)}(m)$-universal hypergraphs

Proof of Theorem 2.15. To prove the existence of optimal $\mathcal{E}^{(r)}(m)$-universal hypergraphs we exploit the proof of Alon and Asodi [8].

Take any $H \in \mathcal{E}^{(r)}(m)$ and replace all edges of H by cliques of size r. This gives a graph with at most $\binom{r}{2}m$ edges and thus there exists a graph G with $O(m^2/\log^2 m)$ edges which is $\mathcal{E}(\binom{r}{2}m)$-universal. We define the r-graph $\mathcal{K}_r(G)$ on the vertex set $V(G)$ with edges being the vertex sets of the copies of K_r in G. It is straightforward to see that $\mathcal{K}_r(G)$ is $\mathcal{E}^{(r)}(m)$-universal and thus it remains to estimate the number of edges in $\mathcal{K}_r(G)$.

The $\mathcal{E}(m)$-universal graph G of Alon and Asodi [8] is defined on the vertex set $V = V_0 \cup V_1 \cup \dots \cup V_k$ where $k = \lceil \log_2 \log_2 m \rceil$, $|V_0| = 4m/\log_2^2 m$ and $|V_i| = 4m2^i/\log_2 m$ for $i \in [k]$. A vertex in V_0 is connected to any other vertex and the graph induced on V_1 is a clique. For any $u \in V_i$, $i \geq 2$, and $v \in V_1 \cup V_2 \cup \cdots \cup V_i$ with $u \neq v$ the edge uv is present independently with probability $\min\left(1, 8^{3-i}\right)$. It is shown in [8] that with probability at least $1/4$ the graph G has $O(m^2/\log^2 m)$ edges and is $\mathcal{E}(m)$-universal. We count the expected number of copies of K_r in G, i.e. $\mathbb{E}\left[|E(\mathcal{K}_r(G))|\right]$.

There are several possible types of cliques K_r in G. Indeed, we need to choose r vertices from $V_0,\ldots,V_k$, and a particular *type* of a possible r-clique K in G is specified by α, which is the number of its vertices in V_0 and by numbers $t_1 \leq \ldots \leq t_\gamma$ (all from $[k]$), which specify to which sets V_i the remaining $\gamma = r - \alpha$ vertices belong to. There are at most $|V_0|^\alpha \prod_{j=1}^{\gamma} |V_{t_j}|$ cliques of a particular type, and each such clique occurs with probability $\prod_{j=1}^{\gamma}\left[\min\left(1, 8^{3-t_j}\right)\right]^{j-1}$. It is clear that there are at most $|V_0|^{r-1}|V(G)| \leq \frac{(4m)^{r-1}\cdot(32m)}{(\log_2 m)^{2(r-1)}} = o\left(\frac{m^r}{\log_2^r m}\right)$ cliques K_r in G that intersect V_0 in at least $r-1$ vertices.

Next we upper bound the expected number of edges in $\mathcal{K}_r(G)$ as follows:

$$\begin{aligned}
\mathbb{E}\left[|E(\mathcal{K}_r(G))|\right] &\leq |V_0|^{r-1}|V(G)| + \sum_{\substack{\alpha+\gamma=r\\ \gamma\geq 2}} \sum_{1\leq t_1\leq\cdots\leq t_\gamma\leq k} |V_0|^\alpha \prod_{j=1}^{\gamma}|V_{t_j}| \cdot \prod_{j=1}^{\gamma}\left[\min\left(1, 8^{3-t_j}\right)\right]^{j-1} \\
&\leq o\left(\frac{m^r}{\log_2^r m}\right) + \sum_{\gamma\geq 2}^{r}\left(\frac{4m}{\log_2 m}\right)^r \frac{1}{\log_2^{r-\gamma} m} \sum_{1\leq t_1\leq\cdots\leq t_\gamma\leq k} 2^{\sum_{j=1}^{\gamma} t_j} \cdot 2^{\sum_{j=1}^{\gamma}\min\{0,(9-3t_j)(j-1)\}}, \qquad (7.1)
\end{aligned}$$

and in order to simplify it further we first estimate the inner sum of the second summand by splitting it according to t_1 as follows:

$$\begin{aligned}
&\sum_{1\leq t_1\leq\cdots\leq t_\gamma\leq k} 2^{\sum_{j=1}^{\gamma} t_j} \cdot 2^{\sum_{j=1}^{\gamma}\min\{0,(9-3t_j)(j-1)\}} \\
&\leq \sum_{t_1\leq 19} \sum_{\substack{t_j\geq 1\\ j=2,\ldots,\gamma}} 2^{\sum_{j=1}^{\gamma} t_j + \sum_{j=1}^{\gamma}\min\{0,(9-3t_j)(j-1)\}} + \sum_{t_1\geq 20} \sum_{\substack{t_j\geq t_1\\ j=2,\ldots,\gamma}} 2^{\sum_{j=1}^{\gamma} t_j + \sum_{j=1}^{\gamma}\min\{0,(9-3t_j)(j-1)\}} \\
&\leq 2^{20} \sum_{\substack{t_j\geq 1\\ j=2,\ldots,\gamma}} 2^{\sum_{j=2}^{\gamma}(t_j+\min\{0,(9-3t_j)(j-1)\})} + \sum_{t_1\geq 20} 2^{t_1} \sum_{\substack{t_j\geq t_1\\ j=2,\ldots,\gamma}} 2^{\sum_{j=2}^{\gamma}(t_j+(9-3t_j)(j-1))} \\
&\leq 2^{20}\left(\sum_{t\geq 1} 2^{t+\min\{0,(9-3t)\}}\right)^{\gamma-1} + \sum_{t_1\geq 20} 2^{t_1}\left(\sum_{t\geq t_1} 2^{t+(9-3t)}\right)^{\gamma-1} \\
&\leq 2^{20}\left(6+\sum_{t\geq 3} 2^{9-2t}\right)^{\gamma-1} + \sum_{t_1\geq 20} 2^{t_1}\left(\sum_{t\geq t_1} 2^{-3t/2}\right)^{\gamma-1} \leq 2^{20+5\gamma} + \sum_{t_1\geq 20} 2^{t_1-\frac{3t_1(\gamma-1)}{2}+2(\gamma-1)} \\
&\leq 2^{20+5\gamma} + 2^{2(\gamma-1)}\sum_{t_1\geq 20} 2^{-t_1/2} \leq 2^{21+5\gamma} \leq 2^{21+5r}.
\end{aligned}$$

This allows us to further upper bound (7.1) by

$$\mathbb{E}\left[|E(\mathcal{K}_r(G))|\right] \leq r2^{21+5r}\left(\frac{4m}{\log_2 m}\right)^r.$$

By Markov's inequality, the probability that $|E(\mathcal{K}_r(G))|$ is at least $5r2^{21+5r}\left(\frac{4m}{\log_2 m}\right)^r$ is at most $1/5$. Thus, taking $\hat{m} = \binom{r}{2}m$, there exists an $\mathcal{E}(\hat{m})$-universal graph with $O\left(\frac{\hat{m}^r}{\log_2^r \hat{m}}\right)$ copies of K_r. This implies that there exists an $\mathcal{E}^{(r)}(m)$-universal hypergraph H with $O(m^r/\log^r m)$ edges. $\square$

It is possible to prove that there exist such hypergraphs H with rm vertices which is optimal. However, no explicit construction is known.

Chapter 8

Conclusion and open problems

In this thesis we have seen different perspectives on embedding spanning structures in random graphs and hypergraphs. We first proved a generalisation of a theorem by Riordan [97], which gives the right thresholds for several classes of graphs and opens up new possibilities for extending results to hypergraphs. Then we improved upon the best-known algorithms for finding a tight Hamilton cycle in $\mathcal{H}^{(r)}(n,p)$. Next we obtained the threshold for embedding spanning bounded degree graphs into randomly perturbed graphs. And finally we worked on universality in random hypergraphs as well as the construction of optimal universal hypergraphs.

While all these results contribute to our knowledge and bring us a step forward in understanding random graphs, there are still many open problems and regimes that are not well understood. We will now discuss some related problems, some of which were mentioned earlier, and suggest ideas for future work. In the topic of embedding spanning structures, among the central objectives are Conjecture 2.3, its generalisation Conjecture 2.8, and the Kahn-Kalai Conjecture [68]. Of course, an ultimate solution of any of these three would be a great achievement, but there are more accessible problems on the way to these conjectures, which are compelling on their own.

Embeddings into $\mathcal{G}(n,p)$

A very concrete problem towards Conjecture 2.3 is the case $\Delta = 4$. Even in the almost spanning case, this is still open. The only obstruction is a triangle with two pending edges on each vertex, which can not be embedded using Janson's inequality directly because $p_4^9 n^3 = o(1)$. For any Δ it would be interesting to get rid of the $\log$-term in the almost spanning version. As discussed previously this is plausible, as for example the almost $K_{\Delta+1}$-factor already appears after $n^{-2/(\Delta+1)}$, cf. Theorem 2.19. In Chapter 5 we manged to do this, for the case where we have a larger graph and do not care which small part is left over.

If we want to make Theorem 2.4 spanning, new ideas are required. In fact if there is a linear number of dense spots from one type, then using the result of Johannson, Kahn, and Vu [67], it is possible to find an almost spanning embedding for the rest and then use a new round of randomness to apply Theorem 2.2 and embed the whole graph. If there are no dense spots than it is also easy, but in between it is much harder. One approach could be to use absorbers, but these would have to be specifically set up for each dense spot.

Turning to universality questions we have to leave the second moment method behind as used in Theorem 2.5 and with more detailed analysis for Hamilton cycles [43, 44]. The results by Johannson, Kahn, and Vu [67] on factors in random graphs were obtained using martingales, which results in a probability that is also too small for a union bound. Nevertheless as shown in [53] this can still be used for universality with some extra caution. The variant of Gerke and McDowell [61] gives some more flexibility for applying these theorems. A better understanding of these results, might play a crucial role for developing further techniques.

Advancing to universality, it would be interesting to improve on the almost spanning case. That would be either getting rid of the log-terms for $\Delta = 3$ or improving significantly for larger Δ. A more careful deletion of some structures inside the graph might help to reduce the probability needed. For spanning universality the case $\Delta = 3$ is the next to approach. But in contrast to disjoint unions of cycles in $\mathcal{F}(n,2)$, the graphs in $\mathcal{F}(n,3)$ are expanders and thus much harder to embed into a pseudo-random environment obtained at this probability. It might be necessary to split $\mathcal{F}(n,3)$ into several classes depending on their properties and deal with each of them separately in a different way.

Randomly perturbed graphs

Regarding universality it would also be worthwhile to extend our Theorem 2.7 in the model $G_\alpha \cup \mathcal{G}(n,p)$ such that we can embed all graphs from $\mathcal{F}(n,\Delta)$ simultaneously. However, our use of Riordan's result, which was proved by second moment calculations, and the multi-round exposure make it unlikely that our techniques can be used to obtain such a result. Even though we believe, that $\mathcal{F}(n,\Delta)$-universality holds for $p = \omega(n^{-2/(\Delta+1)})$, new ideas are needed to show this.

But this model also provides several other interesting questions. The result of Balogh, Treglown, and Wagner [19] mentioned before shows that there are nontrivial spanning structures for which starting with G_α confers no benefit. That is, there are structures whose appearance threshold in $G(n,p)$ is not larger than in $G_\alpha \cup G(n,p)$. On the other hand, in the hypergraph setting McDowell and Mycroft [89] showed that the thresholds can differ by some factor n^ε. The question when (and why) the thresholds in $G(n,p)$ and $G_\alpha \cup G(n,p)$ are different and by how much they can differ still merits more systematic study.

A first question in this direction is whether in the graph case there is some spanning structure where we can benefit a polynomial n^ε compared to the threshold in $\mathcal{G}(n,p)$. A natural candidate for this is the square of the Hamilton cycle, because powers of Hamilton cycles resemble this property in the hypergraph case [89]. As discussed before the threshold for the appearance should be $n^{-1/2}$, even though the currenlty best known upper bound is a polylog-factor off. Together with Böttcher, Montgomery, and Person [32] we are able to extend the result of McDowell and Mycroft [89] to graphs.

Theorem 8.1. *For every $\alpha > 0$ there exists an $\varepsilon > 0$ such that $G_\alpha \cup \mathcal{G}(n, n^{-1/2-\varepsilon})$ a.a.s. contains the square of a Hamilton cycle.*

The proof is again based on our method, but is not included in this thesis. The optimal dependence between α and ε is unclear. There is another result for the square of the Hamilton cycle by Bennett, Dudek, and Frieze [21], which requires $\alpha > 1/2$ and $p \geq Cn^{-2/3}\log^{1/3} n$, and thus, a lower probability but higher minimum degree. At this range G_α already contains many Hamilton cycles by Dirac's Theorem. Together with our result, this raises the question if some sort of interpolation is possible inbetween. Note that it was proved by Komlós, Sárközy, and Szemerédi [76] that G_α on its own contains the square of a Hamilton cycle, provided that $\alpha \geq 2/3$ and $v(G_\alpha)$ is large enough. Another problem where we feel that the comparison of these thresholds would be interesting is the d-dimensional cube, which appears in $G(n,p)$ shortly after $p = 1/4$ [97].

Hypergraphs

As demonstrated on many examples throughout this thesis, some phenomenons do generalise to hypergraphs in a straight-forward way, whereas others behave differently. It would be very interesting to know which of the previously discussed improvements from the graph case can be easily extended, and which require substantial additional work.

The result by Ferber, Luh, and Nguyen [54] (Theorem 2.4) and our result on embedding graphs from $\mathcal{F}(n, \Delta)$ in $G_\alpha \cup \mathcal{G}(n, p)$ (Theorem 2.5), both use Riordan's [97] result, which already is generalised to hypergraphs in Theorem 2.5. Apart from this the decomposition into sparse and dense parts and the application of Janson's inequality need to be checked. For the result in $G_\alpha \cup \mathcal{G}(n, p)$ it is unclear how to extend the switching idea to hypergraphs, which is essential for our approach. Furthermore, extensions of the almost spanning universality result by Conlon, Ferber, Nenadov, and Škorić [36], where the question is what kind of cycle we want to remove, and the spanning universality by Ferber and Nenadov [55] would be nice.

Algorithmic questions

In this thesis we also provided a deterministic algorithm for finding tight Hamilton cycles in $\mathcal{H}^{(r)}(n, p)$ with runtime $O(n^r)$. This gives an affirmative answer to a question of Dudek and Frieze [44] in this regime, but the question remains open for $e/n \leq p < C(\log n)^3 n^{-1}$, where for $r = 3$ the precise threshold is not clear. Furthermore there are various other structures, in particular ℓ-overlapping Hamilton cycles for $1 \leq \ell \leq r - 2$, for which no efficient algorithms are known.

A closely related problem is finding the k-th power of a Hamilton cycle in $\mathcal{G}(n, p)$, which is the shadow graph of a tight k-uniform Hamilton cycle. As discussed the threshold for the appearance is given by $n^{-1/k}$ for $k \geq 3$ [84, 97]. This results is based on the second moment method and thus inherently non-constructive. However, the proof by Nenadov and Škorić [92] gives a quasi-polynomial time algorithm to find the k-th power for $k \geq 2$ a.a.s. provided that $p \geq C(\log n)^{8/k} n^{-1/k}$. This algorithm is very similar to their algorithm for finding tight Hamilton cycles in $\mathcal{H}^{(r)}(n, p)$. The main difference between the problems is that in the graph case two overlapping K_t's are not independent in contrast to two overlapping hyperedges. We think that our ideas are also applicable in this context and would provide an improved algorithm for finding k-th powers of Hamilton cycles in $\mathcal{G}(n, p)$, though we did not check any details.

Finally, it would be interesting to know the average case complexity of determining whether an n-vertex r-uniform hypergraph with m edges contains a tight Hamilton cycle. Our results imply that if $m = \omega(n^{r-1} \log^3 n)$ then a typical such hypergraph will contain a Hamilton cycle, but the failure probability of our algorithm is not good enough to show that the average case complexity is polynomial time. For this one would need a more robust algorithm which can tolerate some *errors* at the cost of doing extra computation to determine whether the *error* causes Hamiltonicity to fail or not.

Bibliography

[1] D. Achlioptas and A. Naor, *The two possible values of the chromatic number of a random graph*, Annals of Mathematics (2) **162** (2005), no. 3, 1335–1351.

[2] R. Aharoni and P. Haxell, *Hall's theorem for hypergraphs*, Journal of Graph Theory **35** (2000), no. 2, 83–88.

[3] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*, arXiv:1612.00622 (2016), 122 pages.

[4] P. Allen, J. Böttcher, Y. Kohayakawa, and Y. Person, *Tight Hamilton cycles in random hypergraphs*, Random Structures & Algorithms **46** (2015), no. 3, 446–465.

[5] P. Allen, C. Koch, O. Parczyk, and Y. Person, *Finding tight Hamilton cycles in random hypergraphs faster*, extended abstract, to appear in LATIN 2018: Theoretical Informatics - 13th Latin American Symposium, Buenos Aires, Argentina, April 16-19, 2018, Proceedings.

[6] ———, *Finding tight Hamilton cycles in random hypergraphs faster*, arXiv:1710.08988 (2017), 17 pages.

[7] N. Alon, *Universality, tolerance, chaos and order*, An Irregular Mind: Szemerédi is 70, Springer, 2010, 21–37.

[8] N. Alon and V. Asodi, *Sparse universal graphs*, Journal of Computational and Applied Mathematics **142** (2002), no. 1, 1–11.

[9] N. Alon and M. R. Capalbo, *Sparse universal graphs for bounded-degree graphs*, Random Structures & Algorithms **31** (2007), no. 2, 123–133.

[10] ———, *Optimal universal graphs with deterministic embedding*, Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008 , SIAM, 2008, 373–378.

[11] N. Alon, M. R. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi, *Universality and tolerance*, 41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA, IEEE Computer Society, 2000, 14–21.

[12] ———, *Near-optimum universal graphs for graphs with bounded degrees*, Approximation, Randomization and Combinatorial Optimization: Algorithms and Techniques, 4th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2001 and 5th International Workshop on Randomization and Approximation Techniques in Computer Science, RANDOM 2001 Berkeley, CA, USA, August 18-20, 2001 , Lecture Notes in Computer Science, vol. 2129, Springer, 2001, 170–180.

[13] N. Alon and Z. Füredi, *Spanning subgraphs of random graphs*, Graphs and Combinatorics **8** (1992), no. 1, 91–94.

[14] N. Alon, M. Krivelevich, and B. Sudakov, *Embedding nearly-spanning bounded degree trees*, Combinatorica **27** (2007), no. 6, 629–644.

[15] N. Alon and J. H. Spencer, *The probabilistic method*, 4th ed., John Wiley & Sons, 2016.

[16] D. Angluin and L. G. Valiant, *Fast probabilistic algorithms for Hamiltonian circuits and matchings*, Journal of Computer and System Sciences **18** (1979), no. 2, 155–193.

[17] L. Babai, F. R. Chung, P. Erdös, R. L. Graham, and J. H. Spencer, *On graphs which contain all sparse graphs*, North-Holland Mathematics Studies **60** (1982), 21–26.

[18] J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, Journal of the American Mathematical Society **28** (2015), no. 3, 669–709.

[19] J. Balogh, A. Treglown, and A. Z. Wagner, *Tilings in randomly perturbed dense graphs*, arXiv:1708.09243 (2017), 18 pages.

[20] W. Bedenknecht, J. Han, Y. Kohayakawa, and G. O. Mota, *Powers of tight Hamilton cycles in random perturbed hypergraphs*, in preparation.

[21] P. Bennett, A. Dudek, and A. M. Frieze, *Adding random edges to create the square of a Hamilton cycle*, arXiv:1710.02716 (2017), 7 pages.

[22] A. Björklund, *Determinant sums for undirected Hamiltonicity*, SIAM Journal on Computing **43** (2014), no. 1, 280–299.

[23] T. Bohman, A. M. Frieze, M. Krivelevich, and R. R. Martin, *Adding random edges to dense graphs*, Random Structures & Algorithms **24** (2004), no. 2, 105–117.

[24] T. Bohman, A. M. Frieze, and R. R. Martin, *How many random edges make a dense graph Hamiltonian?*, Random Structures & Algorithms **22** (2003), no. 1, 33–42.

[25] B. Bollobás, *Threshold functions for small subgraphs*, Mathematical Proceedings of the Cambridge Philosophical Society **90** (1981), no. 2, 197–206.

[26] ______, *The evolution of sparse graphs*, Graph theory and combinatorics (Cambridge, 1983), Academic Press, London, 1984, 35–57.

[27] ______, *The chromatic number of random graphs*, Combinatorica **8** (1988), no. 1, 49–55.

[28] B. Bollobás, T. I. Fenner, and A. M. Frieze, *An algorithm for finding Hamilton cycles in a random graph*, Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA , ACM, 1985, 430–439.

[29] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica **7** (1987), no. 1, 35–38.

[30] B. Bollobás, *Random graphs*, 2nd ed., Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2001.

[31] J. Böttcher, R. Montgomery, O. Parczyk, and Y. Person, *Embedding spanning bounded degree subgraphs in randomly perturbed graphs*, Electronic Notes in Discrete Mathematics **61** (2017), 155–161.

[32] ______, *Embedding spanning bounded degree graphs in randomly perturbed graphs*, arXiv:1802.04603 (2018), 25 pages.

[33] G. Burosch and P. V. Ceccherini, *Isometric embeddings into cube-hypergraphs*, Discrete Mathematics **137** (1995), no. 1-3, 77–85.

[34] D. Clemens, J. Ehrenmüller, and Y. Person, *A Dirac-type theorem for Hamilton berge cycles in random hypergraphs*, Electronic Notes in Discrete Mathematics **54** (2016), 181–186.

[35] D. Conlon and W. Gowers, *Combinatorial theorems in sparse random sets*, Annals of Mathematics **184** (2016), no. 2, 367–454.

[36] D. Conlon, A. Ferber, R. Nenadov, and N. Skoric, *Almost-spanning universality in random graphs*, Random Structures & Algorithms **50** (2017), no. 3, 380–393.

[37] D. Conlon, W. T. Gowers, W. Samotij, and M. Schacht, *On the KŁR conjecture in random graphs*, Israel Journal of Mathematics **203** (2014), no. 1, 535–580.

[38] D. Conlon and R. Nenadov, *Size ramsey numbers of graphs with bounded degree*, in preparation.

[39] C. Cooper, A. M. Frieze, M. Molloy, and B. A. Reed, *Perfect matchings in random r-regular, s-uniform hypergraphs*, Combinatorics, Probability & Computing **5** (1996), 1–14.

[40] D. Dellamonica Jr., Y. Kohayakawa, V. Rödl, and A. Ruciński, *Universality of random graphs*, SIAM Journal on Discrete Mathematics **26** (2012), no. 1, 353–374.

[41] ______, *An improved upper bound on the density of universal random graphs*, Random Structures & Algorithms **46** (2015), no. 2, 274–299.

[42] G. A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society **3** (1952), no. 1, 69–81.

[43] A. Dudek and A. M. Frieze, *Loose Hamilton cycles in random uniform hypergraphs*, The Electronic Journal of Combinatorics **18** (2011), no. 1, P48.

[44] ______, *Tight Hamilton cycles in random uniform hypergraphs*, Random Structures & Algorithms **42** (2013), no. 3, 374–385.

[45] A. Dudek, A. M. Frieze, P. Loh, and S. Speiss, *Optimal divisibility conditions for loose Hamilton cycles in random hypergraphs*, The Electronic Journal of Combinatorics **19** (2012), no. 4, P44.

[46] A. Dudek, A. M. Frieze, A. Ruciński, and M. Sileikis, *Approximate counting of regular hypergraphs*, Information Processing Letters **113** (2013), no. 19-21, 785–788.

[47] A. Dudek and L. Helenius, *On offset Hamilton cycles in random hypergraphs*, arXiv:1702.01834 (2017), 13 pages.

[48] P. Erdős and A. Rényi, *On random graphs I*, Publicationes Mathematicae (Debrecen) **6** (1959), 290–297.

[49] ______, *On the evolution of random graphs*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences, Series A **5** (1960), 17–61.

[50] ______, *On the existence of a factor of degree one of a connected random graph*, Acta Mathematica Academiae Scientiarum Hungaricae **17** (1966), 359–368.

[51] P. Erdös, *Some remarks on the theory of graphs*, Bulletin of the American Mathematical Society **53** (1947), no. 4, 292–294.

[52] A. Ferber, *Closing gaps in problems related to Hamilton cycles in random graphs and hypergraphs*, The Electronic Journal of Combinatorics **22** (2015), no. 1, P1.61.

[53] A. Ferber, G. Kronenberg, and K. Luh, *Optimal threshold for a random graph to be 2 universal*, arXiv:1612.06026 (2016), 23 pages.

[54] A. Ferber, K. Luh, and O. Nguyen, *Embedding large graphs into a random graph*, Bulletin of the London Mathematical Society **49** (2017), no. 5, 784–797.

[55] A. Ferber and R. Nenadov, *Spanning universality in random graphs*, arXiv:1707.07914 (2017), 14 pages.

[56] A. Ferber, R. Nenadov, and U. Peter, *Universality of random graphs and rainbow embedding*, Random Structures & Algorithms **48** (2016), no. 3, 546–564.

[57] E. Friedgut, *Sharp thresholds of graph properties, and the k-sat problem*, Journal of the American Mathematical Society **12** (1999), no. 4, 1017–1054, With an appendix by Jean Bourgain.

[58] ______, *Hunting for sharp thresholds*, Random Structures & Algorithms **26** (2005), no. 1-2, 37–51.

[59] A. M. Frieze, *Loose Hamilton cycles in random 3-uniform hypergraphs*, The Electronic Journal of Combinatorics **17** (2010), no. 1, N28.

[60] A. M. Frieze and M. Karoński, *Introduction to random graphs*, Cambridge University Press, 2015.

[61] S. Gerke and A. McDowell, *Nonvertex-balanced factors in random graphs*, Journal of Graph Theory **78** (2015), no. 4, 269–286.

[62] E. N. Gilbert, *Random graphs*, The Annals of Mathematical Statistics **30** (1959), no. 4, 1141–1144.

[63] D. Hefetz, M. Krivelevich, and T. Szabó, *Sharp threshold for the appearance of certain spanning trees in random graphs*, Random Structures & Algorithms **41** (2012), no. 4, 391–412.

[64] S. Hetterich, O. Parczyk, and Y. Person, *On universal hypergraphs*, The Electronic Journal of Combinatorics **23** (2016), no. 4, P4.28.

[65] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, John Wiley & Sons, 2000.

[66] D. Johannsen, M. Krivelevich, and W. Samotij, *Expanders are universal for the class of all spanning trees*, Combinatorics, Probability & Computing **22** (2013), no. 2, 253–281.

[67] A. Johansson, J. Kahn, and V. H. Vu, *Factors in random graphs*, Random Structures & Algorithms **33** (2008), no. 1, 1–28.

[68] J. Kahn and G. Kalai, *Thresholds and expectation thresholds*, Combinatorics, Probability & Computing **16** (2007), no. 3, 495–502.

[69] J. Kahn, E. Lubetzky, and N. C. Wormald, *The threshold for combs in random graphs*, Random Structures & Algorithms **48** (2016), no. 4, 794–802.

[70] R. M. Karp, *Reducibility among combinatorial problems*, Proceedings of a symposium on the Complexity of Computer Computations, March 20-22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York , The IBM Research Symposia Series, Plenum Press, New York, 1972, 85–103.

[71] P. Keevash, D. Kühn, R. Mycroft, and D. Osthus, *Loose Hamilton cycles in hypergraphs*, Discrete Mathematics **311** (2011), no. 7, 544–559.

[72] J. H. Kim, *Perfect matchings in random uniform hypergraphs*, Random Structures & Algorithms **23** (2003), no. 2, 111–132.

[73] J. H. Kim and S. Lee, *Universality of random graphs for graphs of maximum degree two*, SIAM Journal Discrete Mathematics **28** (2014), no. 3, 1467–1478.

[74] Y. Kohayakawa, T. Łuczak, and V. Roedl, *On k_4-free subgraphs of random graphs*, Combinatorica **17** (1997), no. 2, 173–213.

[75] J. Komlós, *Tiling turán theorems*, Combinatorica **20** (2000), no. 2, 203–218.

[76] J. Komlós, G. N. Sárközy, and E. Szemerédi, *On the square of a Hamiltonian cycle in dense graphs*, Random Structures & Algorithms **9** (1996), no. 1-2, 193–211.

[77] J. Komlós and E. Szemerédi, *Limit distribution for the existence of Hamiltonian cycles in a random graph*, Discrete Mathematics **43** (1983), no. 1, 55–63.

[78] A. D. Koršunov, *Solution of a problem of P. Erdős and A. Rényi on Hamiltonian cycles in undirected graphs*, Doklady Akademii Nauk SSSR **228** (1976), no. 3, 529–532.

[79] M. Krivelevich, *Triangle factors in random graphs*, Combinatorics, Probability & Computing **6** (1997), no. 3, 337–347.

[80] ———, *Embedding spanning trees in random graphs*, SIAM Journal on Discrete Mathematics **24** (2010), no. 4, 1495–1500.

[81] M. Krivelevich, M. Kwan, and B. Sudakov, *Cycles and matchings in randomly perturbed digraphs and hypergraphs*, Combinatorics, Probability & Computing **25** (2016), no. 6, 909–927.

[82] ______, *Bounded-degree spanning trees in randomly perturbed graphs*, SIAM Journal on Discrete Mathematics **31** (2017), no. 1, 155–171.

[83] M. Krivelevich, C. Lee, and B. Sudakov, *Long paths and cycles in random subgraphs of graphs with large minimum degree*, Random Structures & Algorithms **46** (2015), no. 2, 320–345.

[84] D. Kühn and D. Osthus, *On Pósa's conjecture for random graphs*, SIAM Journal on Discrete Mathematics **26** (2012), no. 3, 1440–1457.

[85] A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.

[86] T. Łuczak, *On the equivalence of two basic models of random graphs*, Proceedings of Random Graphs'87 , Wiley, Chichester, 1990, 171–174.

[87] T. Łuczak and A. Rucinski, *Tree-matchings in graph processes*, SIAM Journal on Discrete Mathematics **4** (1991), no. 1, 107–120.

[88] G. Margulis, *Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators*, Problems of Information Transmission **24** (1988), no. 1, 39–46.

[89] A. McDowell and R. Mycroft, *Hamilton ℓ-cycles in randomly-perturbed hypergraphs*, arXiv:1802.04242 (2018), 24 pages.

[90] R. Montgomery, *Embedding bounded degree spanning trees in random graphs*, arXiv:1405.6559v2 (2014), 14 pages.

[91] ______, *Sharp threshold for embedding combs and other spanning trees in random graphs*, arXiv:1405.6560 (2014), 20 pages.

[92] R. Nenadov and N. Škorić, *Powers of Hamilton cycles in random graphs and tight Hamilton cycles in random hypergraphs*, arXiv:1601.04034 (2017), 23 pages.

[93] O. Parczyk and Y. Person, *On spanning structures in random hypergraphs*, Electronic Notes in Discrete Mathematics **49** (2015), 611–619.

[94] ______, *Spanning structures and universality in sparse hypergraphs*, Random Structures & Algorithms **49** (2016), no. 4, 819–844.

[95] D. Poole, *On weak Hamiltonicity of a random hypergraph*, arXiv:1410.7446 (2014), 25 pages.

[96] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics **14** (1976), no. 4, 359–364.

[97] O. Riordan, *Spanning subgraphs of random graphs*, Combinatorics, Probability & Computing **9** (2000), no. 2, 125–148.

[98] V. Rödl and A. Ruciński, *Threshold functions for Ramsey properties*, Journal of the American Mathematical Society **8** (1995), no. 4, 917–942.

[99] ______, *Dirac-type questions for hypergraphs—a survey (or more problems for endre to solve)*, An Irregular Mind: Szemerédi is 70, Springer, 2010, 561–590.

[100] V. Rödl, A. Ruciński, and E. Szemerédi, *A Dirac-type theorem for 3-uniform hypergraphs*, Combinatorics, Probability & Computing **15** (2006), no. 1-2, 229–251.

[101] D. Saxton and A. Thomason, *Hypergraph containers*, Inventiones mathematicae **201** (2015), no. 3, 925–992.

[102] M. Schacht, *Extremal results for random discrete structures*, Annals of Mathematics **184** (2016), no. 2, 333–365.

[103] E. Shamir, *How many random edges make a graph Hamiltonian?*, Combinatorica **3** (1983), no. 1, 123–131.

[104] B. Sudakov and J. Vondrák, *How many random edges make a dense hypergraph non-2-colorable?*, Random Structures & Algorithms **32** (2008), no. 3, 290–306.

[105] E. Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes, Colloque International CNRS, Université Orsay, Orsay, 1976, vol. 260, CNRS, Paris, 1978, 399–401.

Deutsche Zusammenfassung

Das Studium von zufälligen Graphen ist ein faszinierendes Gebiet innerhalb der diskreten Mathematik. Beginnend mit frühen, bahnbrechenden Arbeiten von Erdős und Rényi [48, 49, 50] hat sich dieses Gebiet in den letzen 60 Jahren sehr stark entwickelt. Einen guten Einblick in diesen Prozess gewähren die Bücher von Bollobás [30], Janson, Łuczak und Ruciński [65] und Karonski und Frieze [60]. Typische Fragestellungen in zufälligen Graphen beschäftigen sich mit verschiedenen Graphenparametern (z.B. die chromatische Zahl), der Struktur (z.B. die größte Zusammenhangskomponente) und dem Finden von bestimmten Teilgraphen (z.B. ein Dreieck). Die aufspannende Version des letzten Punktes ist das Hauptthema dieser Arbeit.

Aufspannende Strukturen, die intensiv studiert wurden, sind perfekte Matchings, Hamiltonkreise, Bäume und allgemeine gradbeschränkte Graphen. Trotz Jahren intensiver Forschung bleiben allerdings noch immer viele Fragen offen. Erweitert man die Fragestellungen auf Hypergraphen ist noch weniger bekannt. In dieser Arbeit betrachten wir diese Probleme von verschiedenen Standpunkten aus. Zunächst bestimmen wir Schranken für den Schwellenwert für die Existenz eines beliebigen Hypergraphen und das Einbetten einer ganzen Familie von Hypergraphen. Desweiteren beschäftigen wir uns mit einem algorithmischen Zugang und dem Finden von bestimmten Hamiltonkreisen in Hypergraphen. Zuletzt betrachten wir noch gradbeschränkte Strukturen in einem besonderen Modell von zufällig manipulierten dichten Graphen. In der nächsten Sektion führen wir kurz die wichtigsten Begriffe und Notationen ein, um dann im weiteren Verlauf dieser Übersicht die Ergebnisse einzeln vorzustellen, einzuordnen und deren Beweise zu umreißen.

Schwellenwerte

Das am weitesten verbreite Graphenmodell ist der *zufällige binomiale Graph* $\mathcal{G}(n,p)$, in dem auf n Knoten jede mögliche Kante unabhängig von allen anderen mit Wahrscheinlichkeit p existiert. Eine *Grapheigenschaft* $\mathcal{F}$ ist eine Teilmenge von Graphen und wir sagen, dass $G(n,p)$ die Eigenschaft *asymptotisch fast sicher* (a.f.s.) besitzt, falls $\mathbb{P}[G(n,p) \in \mathcal{F}]$ gegen 1 konvergiert für n gegen unendlich. Ein *Schwellenwert* für eine Eigenschaft $\mathcal{F}$ ist nun eine Funktion $p : \mathbb{N} \to [0,1]$ für die gilt

$$\mathbb{P}\big[\mathcal{G}(n,p) \in \mathcal{F}\big] \begin{cases} \to 0 & \text{falls } p = o(\hat{p}) \\ \to 1 & \text{falls } p = \omega(\hat{p}). \end{cases}$$

Genügt $p \geq (1+\varepsilon)\hat{p}$ und $p \leq (1-\varepsilon)\hat{p}$ für die Konvergenz so spricht man sogar von einem *scharfen* Schwellenwert. Bollobás und Thomason [29] haben gezeigt, dass alle nicht trivialen *monotonen* Grapheneingenschaften, dass heißt das Hinzufügen von Kanten kann die Eigenschaft nicht zerstören, eine Schwellenwertfunktion besitzen. Die Jagd nach den Schwellenwerten für monotone Grapheneigenschaften ist ein zentrales Thema in der Theorie der zufälligen Graphen. Die Existenz eines bestimmten Teilgraphen ist eine monotone Eigenschaft und besitzt damit einen Schwellenwert. Für zwei Graphen G und H ist G ein *Teilgraph* von H, falls es eine Funktion $\phi : V(G) \to V(H)$ gibt, mit $\{\phi(u), \phi(v)\} \in E(H)$ für alle $\{u, v\} \in E(G)$.

Erdős und Rényi [50] bestimmten den Schwellenwert für perfekte Matchings auf $\log n/n$. Etwas später zeigten Pósa [96] und Korŝunov [78] unabhängig voneinander, dass ein Hamiltonkreis den selben Schwellenwert besitzt. Neben vielen anderen Ergebnissen befürworten diese beiden eine Vermutung von Kahn und Kalai [68]. Diese besagt, dass der Schwellenwert für eine Eigenschaft immer innerhalb eines Faktors von $O(\log n)$ von p_E ist, dem sogenannten *Erwartungsschwellenwert*, bei dem die erwartete Anzahl von Kopien von jedem Teilgraphen G' von G in $\mathcal{G}(n, p)$ mindestens 1 ist. Für die beiden Beispiele, perfekte Matchings und Hamiltonkreise, ist p_E in der Nähe von $1/n$ und der extra $\log n$-Faktor ist notwendig, da sonst a.f.s. isolierte Knoten in $\mathcal{G}(n, p)$ auftreten. Sucht man allerdings nur nach der fast aufspannenden Variante in $\mathcal{G}(n, p)$, also auf $(1 - \varepsilon)n$ vielen Knoten für ein beliebiges $\varepsilon > 0$, dann genügt in der Regel p_E.

Einzelne aufspannende Strukturen

Neben perfekten Matchings und Hamiltonkreisen wurden gradbeschränkte Bäume intensiver studiert, wobei das momentan beste Resultat von Montgomery [90, 91] eine Wahrscheinlichkeit $p \geq \Delta \log^5 n/n$ benötig. Ein sehr allgemeines Resultat zum Einbetten von aufspannenden Strukturen wurde von Riordan [97] bewiesen, welches mit Hilfe der zweiten Moment Methode insbesondere den Schwellenwert für Würfel und Gitter bestimmte. Desweiteren wurde in [84] bemerkt, dass daraus auch der Schwellenwert $n^{-1/k}$ für die k-te Potenz des Hamiltonkreises für $k \geq 3$ folgt. Für $k = 2$ ist die beste obere Schranke einen polylog-Faktor von der unteren Schranke $n^{-1/2}$ entfernt [92]. Wir bemerken an dieser Stelle, dass im Gegensatz zu den zuvor erwähnten Schwellenwerten, hier vermutlich keine weiteren $\log n$-Faktoren benötigt werden, was sich mit der Abwesenheit von lokalen Ursachen erklären lässt. Außerdem lässt sich beobachten das die zweite Moment Methode hier für $k \geq 3$ gut funktioniert, im Gegensatz zu perfekten Matchings und Hamiltonkreisen.

Eine natürliche Verallgemeinerung von Matchings ist der Faktor eines Graphen G, dies sind $n/v(G)$ knotendisjunkte Kopien von G. Das bestimmen der Schwellenwerte war lange Zeit ein offenes Problem (vergleiche vorherige Ergebnisse für den Dreiecksfaktor [72, 79]). Die vollständige Lösung gelang Johansson, Kahn und Vu [67]. Sie zeigten insbesondere, dass der Schwellenwert für den $K_{\Delta+1}$-Faktor durch

$$p_\Delta := \left(n^{-1} \log^{1/\Delta} n\right)^{\frac{2}{\Delta+1}}$$

gegeben ist und somit erneut extra $\log$-Terme benötigt werden um sicherzustellen, dass jeder Knoten in einer Kopie von $K_{\Delta+1}$ liegt.

Von hier aus wenden wir uns einer viel allgemeineren Klasse von Graphen zu, die alle bisher behandelten umfasst. Wir bezeichnen mit $\mathcal{F}(n, \Delta)$ die Familie der Graphen auf n Knoten mit Maximalgrad höchstens Δ. Alon und Füredi [13] studierten diese Klasse in $\mathcal{G}(n, p)$ und zeigten, dass $p \geq C(\log n/n)^{1/\Delta}$ immer genügt, für eine absolute Konstante $C > 0$. Bei dieser Wahrscheinlichkeit hat jede Menge von Δ Knoten viele gemeinsame Nachbarn und daher ist eine gierige Einbettungsstrategie mit einem Matchingtrick ausreichend. Da der oben beschriebene $K_{\Delta+1}$-Faktor in $\mathcal{F}(n, \Delta)$ liegt und allgemein als am schwierigsten einzubetten angesehen wird, hat sich die folgende Vermutung verbreitet.

Vermutung. *Sei* $\Delta > 0$*,* $F \in \mathcal{F}^{(r)}(n, \Delta)$ *und* $p = \omega(p_\Delta)$*. Dann enthält* $\mathcal{G}(n,p)$ *a.f.s. eine Kopie von* F*.*

Für $\Delta = 2$ wurde diese Vermutung von Ferber, Kronenberg und Luh [53] bewiesen, die sogar Universalität zeigten. Für größere Δ gibt Riordans Ergebnis [97] eine Wahrscheinlichkeit innerhalb eines Faktors $n^{\Theta(1/\Delta^2)}$ von p_Δ. Für den fast aufspannende Fall gelang es Ferber, Luh und Nguyen [54] zu zeigen, dass p_Δ genügt falls $\Delta \geq 5$, wobei der $\log n$-Term überflüssig sein sollte. Für ihren Beweis unterteilten sie den Graphen in einen großen dünnen Teil, der mit Hilfe von Riordans Ergebnis [97] eingebettet wird, und viele kleine dichte Graphen, die mit Jansons Ungleichung [65] und einem Hypergraphmatchingergebnis von Aharoni und Haxell [2] nachträglich hinzugefügt werden. Wir werden uns diesen Beweisansatz in einem späteren Ergebnis zunutze machen.

Einzelne aufspannende Strukturen in Hypergraphen

Wenn wir uns Hypergraphen zuwenden, dann war abgesehen von Faktoren und Hamiltonkreisen nicht viel bekannt. Der *zufällige* r*-uniform Hypergraph* $\mathcal{H}^{(r)}(n,p)$ ist die natürlich Erweiterung von $\mathcal{G}(n,p)$, es wird jede r-elemetinge Menge mit Wahrscheinlichkeit p undbhängig von allen anderen als Kante gewählt. Das Ergebnis von Johansson, Kahn und Vu [67] gilt auch für Hypergraphen und zeigte unter anderem, dass der Schwellenwert für perfekte Matchings $\log n/n^{r-1}$ ist (zuvor als Shamirs Problem bekannt).

Zusammen mit Person [94] haben wir das bereits mehrmals erwähnte Ergebnis von Riordan [97] auf Hyerpragphen verallgemeinert. Mit $e_H(v) = \max\{e(F) : F \subseteq H, v(F) = v\}$ definieren wir die folgenden Dichte

$$\gamma(H) := \max_{r+1 \leq v \leq n} \left\{ \frac{e_H(v)}{v-2} \right\}.$$

Theorem. *Sei* $r \geq 2$ *eine gazne Zahl und* H *ein* r*-uniformer Hypergraph auf* n *Knoten mit* $e(H) = \alpha\binom{n}{r} = \alpha(n)\binom{n}{r}$ *Kanten und* $\Delta = \Delta(H)$*. Sei weiter* $p\colon \mathbb{N} \to [0,1)$*. Falls* H *einen Knoten von Grad* 2 *hat und*

$$np^{\gamma(H)}\Delta^{-4} \to \infty$$

gilt, dann enthält $\mathcal{H}^{(r)}(n,p)$ *a.f.s. eine Kopie von* H*.*

Für $r = 2$ ist dies Riordans Ergebnis [97, Theorem 2.1]. Unser Beweis für Hypergraphen folgt in weiten Teilen dem Vorgehen von Riordan, muss an einigen Stellen allerdings angepasst werden. Aus dem Theorem lassen sich die Schwellenwerte für einige Hamiltonkreise, Würfel, Gitter und Potenzen von Hamiltonkreisen in Hypergraphen folgern. Desweiteren bereitet es die Möglichkeit die Ergebnisse aus [54] auf Hypergraphen zu erweitern.

Es gibt verschiedene Möglichkeiten Hamiltonkreise auf Hypergraphen zu verallgemeinern. Wir betrachten hier ℓ*-überlappende* Hamiltonkreise, dass sind n Knoten in einer zyklischen Anordnung, wobei die Kanten Segemente von r Knoten sind und benachbarte Kanten sich jeweils in ℓ Knoten überlappen. Für $\ell = 1$ oder $\ell = r-1$ nennen wir dies einen *losen* bzw. *engen* Hamiltonkreis. Die

Schwellenwerte für ℓ-überlappenden Hamiltonkreisen in Hypergraphen wurden hauptsächlich von Dudek und Frieze [43, 44] studiert. Im Allgemeinen sind diese $n^{\ell-r}$ für $\ell \geq 2$ und $n^{1-r}\log n$ für $\ell = 1$, aber es sind noch genauere Ergebnisse bekannt, z.B. ist e/n ein scharfer Schwellenwert für enge Hamiltonkreise, falls $r \geq 4$. Für $\ell \geq 2$ können wir den Schwellenwert auch direkt aus unserem ersten Theorem folgern.

Algorithmen für Hamiltonkreise

Die meisten der bisher vorgestellen Resultate liefern uns zwar Informationen darüber wann mit hoher Wahrscheinlichkeit eine Struktur in einem zufälligen Graphen existiert, geben uns aber keinerlei Möglichkeit eine solche zu finden. Vor allem die Ergebnisse, die die zweite Moment Methode benutzen oder das Resultat von Johannson, Kahn und Vu [67], liefern keine brauchbaren Algorithmen. Wir sind interessiert an Algorithmen mit einer Laufzeit, die polynomiell in der Anzahl der Knoten n ist. Im Folgenden diskutieren wir dieses Problem anhand von Hamiltonkreisen.

Zu entscheiden ob ein gegebener Graph einen Hamiltonkreis enthält, ist eins der 21 klassischen NP-volsständigen Probleme von Karp [70]. Der beste bekannte Algorithmus ist ein Monte-Carlo Algorithmus von Björklund [22] mit Laufzeit $O^*(1.657^n)$. Aber wie sieht es mit *typischen* Instanzen aus, also zum Beispiel $\mathcal{G}(n,p)$? Wir wissen, dass für $p = \omega(\log n/n)$ es a.f.s. einen Hamiltonkreis in $\mathcal{G}(n,p)$ gibt, aber können wir diesen dann auch finden?

Die ersten Polynomialzeit-Algorithmen zum Finden von Hamiltonkreisen in $\mathcal{G}(n,p)$ von Angluin und Valiant [16] und Shamir [103] waren randomisiert. Später gelang es Bollobás, Fenner und Frieze [28] einen deterministischen Algorithmus zu entwickeln, dessen Erfolgswahrscheinlichkeit auf $\mathcal{G}(n,p)$ zu der Wahrscheinlichkeit für Hamiltonizität in $\mathcal{G}(n,p)$ passt für n gegen unendlich.

Für Hypergraphen haben Dudek und Frieze am Ende von [44] das Problem gestellt die verschiedenen Hamiltonkreise in $\mathcal{H}^{(r)}(n,p)$ an den Schwellenwerten zu finden. Für enge Hamiltonkreise wurde ein erster randomisierter Algorithmus von Allen, Böttcher, Kohayakawa und Person [4] beschrieben, der $p \geq n^{-1+\varepsilon}$ benötigt für ein festes $\varepsilon \in (0, 1/6r)$ und Laufzeit n^{20/ε^2} besitzt. Daraufhin gelang es Nenadov und Škorić [92] einen randomisierten quasipolynomiellen Algorithmus zu finden, der mit $p \geq C(\log n)^8 n^{-1}$ auskommt. Zusammen mit Allen, Koch und Person [6] konnten wir dies weiter verbessern.

Theorem. *Für jedes $r \geq 3$ existiert ein $C > 0$ und ein deterministischer Algorithmus mit Laufzeit $O(n^r)$, der für $p \geq C(\log n)^3 n^{-1}$ a.f.s. einen engen Hamiltonkreis im zufälligen r-uniformen Hypergraphen $\mathcal{H}^{(r)}(n,p)$ findet.*

Unser Ergebnis benutzt eine Variante der Absorbiertechnik von Rödl, Ruciński und Szemerédi [100]. Diese Technik wurde bereits von Krivelevich [79] in zufälligen Graphen angewendet, aber die ersten nahezu optimalen Ergebnisse damit wurden in [4] und unabhängig davon von Kühn und Osthus [84] erzielt, welche den Schwellenwert für Potenzen von Hamiltonkreisen in $\mathcal{G}(n,p)$ studierten. Bei einer Wahrscheinlichkeit $p \geq C(\log n)^3 n^{-1}$ benötigen wir eine Reservoirstruktur von polylogarithmischer Größe, wie sie zuvor von Montgomery [91] für das Finden von aufspannenden Bäumen benutzt wurde, und später auch in [92].

In einem Hypergraphen $H = (V, E)$ findet unser Algorithmus zuerst einen langen engen Pfad mit der besonderen Eigenschaft, dass er aus einem Reservoir $R \subseteq V$ von polylogarithmisch vielen Knoten jede beliebige Teilmenge $R' \subseteq R$ absorbieren kann. Dieser Pfad wird dann zu einem engen Pfad erweitert, der $V \setminus R$ enthält, wobei am Ende auch einige Knoten aus R verwendet werden dürfen. Mit Hilfe der restlichen Reservoirknoten kann der Pfad zu einem engen Hamiltonkreis geschlossen werden und die übrig gebliebenen Knoten R' aus R können durch die oben beschriebene Eigenschaft in den Pfad absorbiert werden. Dieses Vorgehen wurde auch in ähnlicher Weise in [4, 92] genutzt. Durch einen simpleren Algorithmus und ein präziseres Vorgehen beim Verbinden von zwei Endpunkten gelingt es uns aber diese Ergebnisse wie beschrieben zu verbessern.

Trotz unser Verbesserung bleibt Platz zu der unteren Schranke von e/n und es ist offen wie weit ein algorithmischer Beweis gehen kann. Der Algorithmus von Nenadov und Škorić [92] funktioniert auch bei Potenzen von Hamiltonkreisen im Graphenfall, welche engen Hamiltonkreisen in Hypergraphen sehr ähnlich sind. Unsere Ideen sollten auch auf diesen Fall erweiterbar sein und bessere Algorithmen dafür liefern.

Zufällig veränderte dichte Graphen

Wir betrachten nun ein etwas anderes Modell, das von Bohman, Frieze und Martin [24] iniziiert wurde. Für $\alpha \in (0, 1)$ sei G_α ein beliebiger Graph mit Minimalgrad αn. Nun fügen wir weitere Kanten zufällig mit Wahrscheinlichkeit p hinzu. Wir studieren also Eigenschaften und insbesondere Schwellenwerte für aufspannende Strukturen in dem Modell $G_\alpha \cup \mathcal{G}(n, p)$.

Für $\alpha \in (0, 1/2)$ haben Bohman, Frieze und Martin [24] gezeigt, dass $1/n$ der Schwellenwert für einen Hamiltonkreis in $G_\alpha \cup \mathcal{G}(n, p)$ ist. Wir sparen also einen $\log n$-Faktor gegenüber dem Schwellenwert in $\mathcal{G}(n, p)$, da wir durch G_α bereits einen hohen Minimalgrad garantiert haben und somit in diesem Fall p_E ausreicht. Für $\alpha \geq 1/2$ brauchen wir keine zusätzlichen Kanten, da G_α bereits selber einen Hamiltonkreis enthält (Diracs Theorem). Ein ähnliches Phänomen, wie im Hamiltonkreisfall, lässt sich auch bei vielen weiteren aufspannende Strukturen beobachten, bei denen sich der Schwellenwert in $\mathcal{G}(n, p)$ und der Erwartungsschwellenwert p_E unterscheiden.

Krivelevich, Kwan und Sudakov [82] studierten das Problem für gradbeschränkte Bäume und zeigten, dass $1/n$ auch hier der Schwellenwert ist. Es ist bekannt, dass $1/n$ für den fast aufspannenden Fall genügt [14] und damit wird G_α nur gebraucht, um die Einbettung zu vollenden und insbesondere den nötigen Minimalgrad sicherzustellen. Erst kürzlich gelang es Balogh, Treglown und Wagner [19] für Faktoren zu zeigen, dass die $\log n$-Terme in diesem Modell nicht gebraucht werden und p_E genügt. Für eine bestimmte Klasse von Faktoren zeigten sie damit auch, dass das hinzufügen von G_α keinen Vorteil gegenüber $\mathcal{G}(n, p)$ allein ergibt. Weitere Ergebnisse in diesem Modell beschäftigen sich mit kleinen Cliquen, dem Durchmesser, k-Zusammenhang [23] und Nicht-2-Färbbarkeit [104]. Unser Resulat mit Böttcher, Montgomery und Person [32] in diesem Modell ist für gradbeschränkte Graphen.

Theorem. *Sei $\alpha > 0$ eine Konstante, $\Delta \geq 5$ eine natürliche Zahl und G_α ein Graph mit Minimalgrad mindestens αn. Dann gilt für jedes $F \in \mathcal{F}(n, \Delta)$ und mit $p = \omega\left(n^{-\frac{2}{\Delta+1}}\right)$, dass $G_\alpha \cup G(n, p)$ a.f.s. eine Kopie von F enthält.*

Unser Beweis nutzt das oben beschriebene Vorgehen von Ferber, Luh und Nguyen [54] zusammen mit einer neuen Art von Reservoirstruktur. Ohne größere Schwierigkeiten lassen sich daraus auch die vorheringen Ergebnisse über gradbeschränkte Bäume [82] und Faktoren [19] zurückgewinnen und somit erhalten wir Beweise dieser Ergebnisse ohne das dort verwendete Regularitätslemma von Szemerédi [105].

Wie bei den anderen Resultaten finden wir zuerst eine fast aufspannende Einbettung für die wir nur Kanten von $\mathcal{G}(n,p)$ benutzen. Die entscheidende Beobachtung ist, dass diese Einbettung zufällig und uniform verteilt auf G_α liegt. Dadurch gelingt es uns für jeden der übrigen Knoten v eine große Menge von Knoten $B(v)$ zu finden, die durch v ersetzt werden könnten, ohne das die Einbettung verletzt wird. Nun verfolgen wir eine ähnliche Einbettungsstrategie wie vorher, nur das wir jetzt in die Mengen $B(v)$ einbetten. Aufgrund dieser großen Auswahl an Zielknoten und durch weitere Kanten von G_α gelingt es uns auch mit dem niedrigeren p die Einbettung zu vervollständigen und dann die Austauscheigenschaft der Knoten anzuwenden.

Das Modell lässt sich einfach auf r-uniforme Hypergraphen verallgemeinern, es muss aber jeweils festgelegt werden, welche Gradbedingung an G_α gefordert wird. In einer weiteren Arbeit studierten Krivelevich, Kwan und Sudakov [82] perfekte Matchings und lose Hamiltonkreise in Hypergraphen und konnten in beiden Fällen den $\log n$-Faktor gegenüber den vorher besprochenen Schwellenwerten in $\mathcal{H}^{(r)}(n,p)$ einsparen, falls G_α einen minimalen $(r-1)$-Grad von mindestens αn hat. Interessanterweise gelang es McDowell und Mycroft [89] zu zeigen, dass es für ℓ-überlappende Hamiltonkreise mit $\ell \geq 2$ möglich ist sogar einen polynomiellen Faktor n^ε gegenüber dem Schwellenwert in $\mathcal{H}^{(r)}(n,p)$ einzusparen unter der Annahme von hohem ℓ und $r-\ell$ Grad in G_α. Dieses Ergebnis wurde von Bedenknecht, Han, Kohayakawa und Mota [20] auf Potenzen von engen Hamiltonkreisen erweitert, wofür sie einen noch höheren Minimalgrad benötigen.

Zusammen mit Böttcher, Montgomery und Person [32] ist es uns gelungen das analoge Ergebnis für das Quadrat des Hamiltonkreises zu beweisen.

Theorem. *Für jedes $\alpha > 0$ existiert ein $\varepsilon > 0$, so dass $G_\alpha \cup \mathcal{G}(n, n^{-1/2-\varepsilon})$ a.f.s. das Quadrat eines Hamiltonkreises enthält.*

Der Beweis nutzt die von uns vorgestellte Methode, ist aber nicht in dieser Arbeit enthalten. Es gibt noch ein weiteres Ergebnis für das Quadrat des Hamiltonkreises von Benett, Dudek und Frieze [21], welches $\alpha > 1/2$ und $p \geq Cn^{-2/3}\log^{1/3} n$ benötigt und somit bei höherem Minimalgrad eine geringer Wahrscheinlichkeit vorraussetzt. Zusammen mit unserem Ergebnis stellt sich die spannenden Frage, ob eine Art Interpolation dazwischen möglich ist. Desweiteren bedarf die Frage, wann und warum der Schwellenwert in $\mathcal{G}(n,p)$ und in $G_\alpha \cup \mathcal{G}(n,p)$ sich wie weit voneinander unterscheiden, einer tiefgehenderen systematischen Untersuchung.

Universalität

Alles bisher diskutierten Ergebnisse beschäftigen sich mit dem Problem eine einzelne Struktur in einem zufälligen Graphenmodell zu finden. Aber wie verhält es sich, wenn wir eine ganze Familie von Graphen gleichzeitig finden möchten? Wir nennen einen Graphen $\mathcal{F}$-universal für eine Familie $\mathcal{F}$, falls er jeden Graphen $F \in \mathcal{F}$ als Teilgraph enthält. Wir sind hauptsächlich interessiert an Schranken für den Schwellenwert für $\mathcal{F}(n,\Delta)$-Universalität in $\mathcal{G}(n,p)$. Die meisten der bisher diskutierten Ergebnisse lassen sich nicht ohne weiteres auf Universalität verallgemeinern, da die erlangten Wahrscheinlichkeiten zu groß für eine Abschätzung gegen die Anzahl der Graphen sind.

Universalität in zufälligen Graphen

Das Studium der Universalität wurde von Alon, Capalbo, Kohayakawa, Rödl, Ruciński und Szemerédi [11] begonnen. Nach einigen Zwischenergebnissen zeigten Dellamonica, Kohayakawa, Rödl und Ruciński [41], dass $p \geq C\,(\log n/n)^{1/\Delta}$ genügt damit $\mathcal{G}(n,p)$ a.f.s. $\mathcal{F}(n,\Delta)$-universal ist, wobei der Fall $\Delta = 2$ von Kim und Lee [73] bewiesen wurde. Wie zuvor besprochen, bildet diese Wahrscheinlichkeit eine natürliche Grenze für Einbettungsalgorithmen. Es gelang zuerst Conlon, Ferber, Nenadov und Škorić [36] diese zu durchbrechen mit einem fast aufspannenden Ergebnis. Dieses konnte kürzlich von Ferber und Nenadov [55] zu einem aufspannende Ergebnis mit Wahrscheinlichkeit $p \geq (n^{-1}\log^3 n)^{\frac{1}{\Delta-1/2}}$ verbessert werden, wofür sie eine Einbettungstechnik von Conlon und Nenadov [38], Ideen von [36] und Absorbierer verwendeten.

Die untere Schranke für den Schwellenwert kommt erneut vom $K_{\Delta+1}$-Faktor und daher lässt sich die erste Vermutung wie folgt verallgemeinern.

Vermutung. *Sei $\Delta > 0$ und $p = \omega(p_\Delta)$. Dann ist $\mathcal{G}(n,p)$ a.f.s. $\mathcal{F}(n,\Delta)$-universal.*

Für $\Delta = 2$ wurde diese Vermutung bereits von Ferber, Kronenberg und Luh [53] bewiesen.

Universalität in zufälligen Hypergraphen

Zusammen mit Person haben wir uns diese Fragestellung im Hypergraphenfall angeschaut. Für welche p ist $\mathcal{H}^{(r)}(n,p)$ universal für $\mathcal{F}^{(r)}(n,\Delta)$, die Familie der r-uniformen Hypergraphen mit maximalem Knotengrad Δ. Uns gelang es, das Ergebnis von Dellamonica, Kohayakawa, Rödl und Ruciński [41] zu erweitern und zu zeigen, dass wir bis zu der natürlichen Schranke gehen können.

Theorem. *Für jedes $r \geq 2$ und jede natürliche Zahl $\Delta \geq 1$, existiert eine Konstante $C > 0$, so dass mit $p \geq C(\log n/n)^{1/\Delta}$ der zufällige r-uniforme Hypergraph $\mathcal{H}^{(r)}(n,p)$ a.f.s. $\mathcal{F}^{(r)}(n,\Delta)$-universal ist.*

Für den Beweis bedienen wir uns der Ansätze aus [41, 56, 73] und führen diesen in zwei Schritten durch. Zuerst beweisen wir, dass $\mathcal{H}^{(r)}(n,p)$ a.f.s. einige *gute* deterministische, pseudozufällige Eigenschaften erfüllt. Im zweiten Schritt zeigen wir, dass wir jeden Graphen aus $\mathcal{F}^{(r)}(n,\Delta)$ in einen Graphen mit solchen Eigenschaften einbetten können. Dieser Umweg über den deterministischen Teilgraphen erlaubt es uns Universalität zu beweisen und nur im ersten Schritt ist der Zufall involviert.

Wie auch im Graphenfall ist diese Schranke vermutlich nicht optimal und ähnliche Verbesserungen, wie in [36, 55], könnten möglich sein. Die beste untere Schranke für den Schwellenwert kommt wiederum vom Cliquen-Faktor.

Explizite Kontruktionen von universalen Hypergraphen

Ein eng verwandtes Problem ist die Existenz und die explizite Konstruktion von universalen Hypergraphen. Eine einfache Rechnung zeigt, dass jeder $\mathcal{F}(n,\Delta)$-universale Hypergraph mindestens

$\Omega(n^{2-2/\Delta})$ viele Kanten besitzt. Nach anfänglichen Konstruktionen in der bereits erwähnten Arbeit von Alon, Capalbo, Kohayakawa, Rödl, Ruciński und Szemerédi [11], verfolgten Alon und Capalbo weiter diese Schiene der Universalität. Sie konstruierten universale Graphen auf $O(n)$ Knoten mit $O(n^{2-2/\Delta})$ Kanten [10] und weitere auf n Knoten mit $O(n^{2-2/\Delta}\log^{4/\Delta} n)$ Kanten [9]. Dies ist also nahezu optimal.

Beginnend mit Person [94] und darauffolgend mit Hetterich und Person [64] haben wir uns der Konstruktionen von $\mathcal{F}^{(r)}(n,\Delta)$-universalen r-uniformen Hypergraphen zugewand. Mit ähnlichen Rechnungen wie zuvor erhielten wir eine untere Schranke von $\Omega(n^{r-r/\Delta})$ Kanten. Indem wir die Konstruktionen für Graphen von Alon und Capalbo [9, 10] ausnutzten, gelangen uns ebenso optimale Ergebnisse in Hyergraphen für viele Kombinationen von r und Δ zu erhalten.

Theorem. *Falls $r \mid \Delta$ oder $2 \mid r$, dann existieren explizit kontruierbare $\mathcal{F}^{(r)}(n,\Delta)$-universale Hypergraphen auf $O(n)$ Knoten mit $O(n^{r-r/\Delta})$ Kanten und auf n Knoten mit $O(n^{r-r/\Delta}\log^{2r/\Delta} n)$ Knoten.*

Desweiteren existieren explizit kontruierbare $\mathcal{F}^{(r)}(n,2)$-universale Hypergraphen auf $O(n)$ Knoten mit $O(n^{r/2})$ Kanten.

Für alle weiteren Fälle erhalten wir Konstruktionen, die höchstens um einen Faktor n^{r/Δ^2} von der unteren Schranke abweichen.

Eine einfache Möglichkeit eine Konstruktionen für einen $\mathcal{F}(n,\Delta')$-universalen Graphen G auszunutzen, ist daraus einen r-uniformen Hypergraphen H zu definieren, in dem wir jede Clique auf r Knoten durch eine Hyperkante ersetzen. Wollen wir jetzt einen Graphen $F \in \mathcal{F}^{(r)}(n,\Delta)$ in H einbetten, dann genügt es jede Kante durch einen K_r zu ersetzten und somit einen Graphen F' daraus zu erhalten. Dieser Graph F' liegt in $\mathcal{F}(n,\Delta')$ für ein geeingetes Δ' und die Einbettung von F' in G definiert uns auch eine Einbettung von F in H. Somit is H also $\mathcal{F}^{(r)}(n,\Delta)$-universal.

Eine Verbesserung dieser simplen Grundidee in verschiedene Richtungen ermöglicht uns den ersten Teil des Theorems zu beweisen. Für den zweiten Teil müssen wir allerdings tiefer in die Konstruktionen einsteigen und eine bestimmte Unterteilung für alle $F \in \mathcal{F}^{(r)}(n,2)$ erzeugen. In der Tat stehen uns die optimalen Konstruktionen sogar für alle Parameter zur Verfügung, aber um die Universalität zu beweisen fehlt uns eine geeignete Unterteilung.

Eine weitere Familie von Hypergraphen ist $\mathcal{E}^{(r)}(m)$, alle r-uniformen Hypergraphen mit m Kanten und ohne isolierte Knoten. Ein $\mathcal{E}^{(2)}(m)$-universaler Graph muss $\Omega(m^2/\log^2 m)$ Kanten besitzen und Alon und Asodi [8] zeigten die Existenz von solchen Graphen mit $O(m^2/\log^2 m)$ Kanten. Mit einer ähnlichen Technik wie zuvor erweitern wir dieses Ergebnis auf Hypergraphen und zeigen die Existenz von $\mathcal{E}^{(r)}(m)$-universalen Hypergraphen passend zur unteren Schranke von $\Omega(m^r/\log^r m)$.