Loose cores and cycles in random hypergraphs

Oliver Cooley*
Institute of Science and Technology Austria (ISTA)
3400 Klosterneuburg, Austria
oliver.cooley@ist.ac.at

Mihyun Kang* Julian Zalla*
Institute of Discrete Mathematics
Graz University of Technology
8010 Graz, Austria
{kang,zalla}@math.tugraz.at

Submitted: Oct 12, 2021; Accepted: Jun 16, 2022; Published: Oct 21, 2022

Abstract

Inspired by the study of loose cycles in hypergraphs, we define the *loose core* in hypergraphs as a structure which mirrors the close relationship between cycles and 2-cores in graphs. We prove that in the r-uniform binomial random hypergraph $H^r(n,p)$, the order of the loose core undergoes a phase transition at a certain critical threshold and determine this order, as well as the number of edges, asymptotically in the subcritical and supercritical regimes.

Our main tool is an algorithm called `CoreConstruct`, which enables us to analyse a peeling process for the loose core. By analysing this algorithm we determine the asymptotic degree distribution of vertices in the loose core and in particular how many vertices and edges the loose core contains. As a corollary we obtain an improved upper bound on the length of the longest loose cycle in $H^r(n,p)$.

Mathematics Subject Classifications: 05C38, 05C65

1 Introduction

1.1 Motivation

One of the first phase transition results for random graphs is the celebrated result of Erdős and Rényi [16] on the emergence of a *giant component* of linear order when the number of edges passes $\frac{n}{2}$, or from the viewpoint that is now more common, when the edge

*Supported by Austrian Science Fund (FWF): I3747, W1230.

probability passes $\frac{1}{n}$. This result has since been strengthened and generalised in a number of directions. In particular in hypergraphs it has been extended to vertex-components (e.g. [4–7, 24, 33, 36]) as well as to high-order components (e.g. [9, 11–13]).

The k-core of a graph G, defined as the maximal subgraph of minimum degree at least k, has been studied extensively in the literature (e.g. [14, 20, 26, 28, 32]). In random graphs, the k-core may be seen as a natural generalisation of the largest component: in the case $k = 2$, whp[1] a linear-sized 2-core emerges at the same time as the giant component, and indeed lies almost entirely within the giant component, while for $k \geqslant 3$, whp the k-core is identical to the largest k-connected subgraph [28, 29]. In [28] Łuczak estimated the order of (i.e. the number of vertices in) the k-core of $G(n, p)$ and the asymptotic probability that the k-core is k-connected. Łuczak also showed in [29] that in the random graph process, in which edges are added to an empty graph one by one in a uniformly random order, whp at the moment the k-core first becomes non-empty, its order is already linear in n. A crucial milestone was achieved by Pittel, Spencer and Wormald [32], who for $k \geqslant 3$ determined the threshold probability at which the non-empty k-core appears whp and determined its asymptotic order and size (i.e. number of edges). This was strengthened by Janson and Luczak [21], who proved a bivariate central limit theorem for the order and size of the k-core. Cain and Wormald [8] determined the asymptotic distribution of vertex degrees within the k-core. Further research has focussed for example on the robustness of the core against edge deletion [35] and how quickly the peeling process arrives at the core [1, 18, 19, 23]. There are many more results in the literature for cores in random graphs, see e.g. [14, 20, 26].

Paths and cycles in random graphs have been investigated at least since 1979 by de la Vega [17] and somewhat later by Ajtai, Komlós, and Szemerédi [2]. Regarding the length of the longest path in the random graph $G(n, p)$, a standard "sprinkling" argument (see Lemma 22 with $r = 2$) shows that in the supercritical regime the length of the longest path and cycle are very similar. Thus it follows from the results of Łuczak [27] on the length of the longest cycle that for $\varepsilon = \varepsilon(n) = o(1)$ and $p = \frac{1+\varepsilon}{n}$ (i.e. shortly after the phase transition), under the assumption $\varepsilon^5 n \to \infty$ whp the longest path has length $\Theta(\varepsilon^2 n)$, where explicit constants can be given. The best-known upper bounds derive from a careful analysis of the 2-core and the simple observation that any cycle must lie within the 2-core.

There are many different ways of generalising the concept of a k-core to hypergraphs; some results for these cores can be found in e.g. Molloy [31] and Kim [26]. However, in the case $k = 2$, all k-cores which have been studied so far do not fully capture the nice connection between the 2-core and cycles in graphs. In [31] Molloy determined the threshold for the appearance of a non-trivial k-core (in that paper defined as a subhypergraph where every vertex has degree at least k) in the r-uniform binomial random hypergraph $H^r(n, p)$ for all $r, k \geqslant 2$ such that $r + k \geqslant 5$. The proof relied on a clever heuristic argument which was first introduced by Pittel, Spencer and Wormald in [32] and has been adapted by many other authors, see e.g. [20, 26, 34, 37]. It turns out that the proofs in [31] can be extended to a wide range of core-type structures. In the case $k = 2$, Dembo and Montanari [15] strengthened this by determining the width of, and examining the behaviour within, the

[1] short for *with high probability*, i.e. with probability tending to one as the number of vertices $n \to \infty$.

critical window.

One of the most natural concepts of paths and cycles in hypergraphs is *loose paths* and *loose cycles* (see Definition 3). A special case of a recent result of Cooley, Garbe, Hng, Kang, Sanhueza-Matamala and Zalla [10] shows that the length of the longest loose path in an r-uniform binomial random hypergraph undergoes a phase transition from logarithmic length to linear, and they also determined the critical threshold, as well as proving upper and lower bounds on the length in the subcritical and supercritical ranges.

Inspired by the substantial body of research on loose cycles, in this paper we introduce the *loose core* (see Definition 2), a structure which does indeed capture the connection between cores and cycles in hypergraphs. Our first main result concerns the degree distribution of vertices in the loose core (see Theorem 4). In fact we prove a stronger result regarding degree distributions of both vertices and edges (see Theorem 21). As a consequence we can deduce both the asymptotic numbers of vertices and edges in the loose core (see Theorem 5) and an improved upper bound on the length of the longest loose cycle in an r-uniform binomial random hypergraph (see Theorem 6).

Before stating our main results, in the next section we introduce some definitions and notations which we will use throughout the paper.

1.2 Setup

Given a natural number $r \geqslant 3$, an *r-uniform hypergraph* consists of a vertex set V and an edge set $E \subset \binom{V}{r}$, where $\binom{V}{r}$ denotes the set of all r-element subsets of V. Let $H^r(n,p)$ denote the *r-uniform binomial random hypergraph* on vertex set $[n]$ in which each set of r distinct vertices forms an edge with probability p independently. For any positive integer k we write $[k] \coloneqq \{1,\dots,k\}$ and $[k]_0 \coloneqq \{0,\dots,k\}$. We also include 0 in the natural numbers, so we write $\mathbb{N} = \{0,1,\dots\}$ and $\mathbb{N}_{\geqslant k} \coloneqq \{k,k+1,\dots\}$. Throughout the paper, unless otherwise stated any asymptotics are taken as $n \to \infty$. In particular, we use the standard Landau notations $o(\cdot), O(\cdot), \Theta(\cdot), \omega(\cdot)$ with respect to these asymptotics.

The loose core will be defined in terms of two parameters, namely the standard notion of (vertex-)degree and a notion we call the connection number.

Definition 1. Let H be an r-uniform hypergraph. Let $d_H(v)$ be the *degree* of a vertex v in H (i.e. the number of edges which contain it) and let $\delta(H)$ denote the *minimum (vertex-)degree* of H, i.e. the smallest degree of any vertex of H. For any edge $e \in E(H)$, define the *connection number* $\kappa(e) \in [r]_0$ of e as

$$\kappa(e) = \kappa_H(e) \coloneqq |\{v \in e : d_H(v) \geqslant 2\}|$$

and let $\kappa(H) \coloneqq \min_{e \in E(H)} \kappa(e)$.

We are now ready to define the loose core.

Definition 2 (Loose core)**.** The *loose core* of an r-uniform hypergraph H is the maximal subhypergraph H' of H such that

(C1) $\delta(H') \geqslant 1$,

(C2) $\kappa(H') \geqslant 2$.

If such a subhypergraph does not exist, then we define the loose core to be the empty hypergraph (i.e. the hypergraph with no vertices and no edges).

Note that the loose core is unique, since the union of two hypergraphs each with properties **(C1)** and **(C2)** again has these properties. The first condition in Definition 2 simply states that the loose core contains no isolated vertices and the second condition specifies how edges are connected to each other in the loose core. Note that for $r \geqslant 3$ the loose core might contain vertices of degree 1, in contrast to the graph case. For $r = 2$, Definition 2 coincides with the 2-core of a graph.

Our motivation to study loose cores arises from the study of loose cycles in hypergraphs which are closely related to loose paths.

Definition 3 (Loose path/cycle)**.** A *loose path of length* ℓ in an r-uniform hypergraph is a sequence of distinct vertices $v_1, \dots, v_{\ell(r-1)+1}$ and a sequence of edges $e_1, \dots, e_\ell$, where $e_i = \{v_{(i-1)(r-1)+1}, \dots, v_{(i-1)(r-1)+r}\}$ for $i \in [\ell]$. A *loose cycle of length* ℓ in an r-uniform hypergraph is defined similarly except that $v_{\ell(r-1)+1} = v_1$ (and otherwise all vertices are distinct).

Note that for $i \in [\ell - 1]$ we have $e_i \cap e_{i+1} = \{v_{i(r-1)+1}\}$ (and in the case of a loose cycle, $e_\ell \cap e_1 = \{v_1\}$), so in particular two consecutive edges intersect in precisely one vertex. Observe that a loose cycle satisfies conditions **(C1)** and **(C2)** of a loose core (Definition 2) and hence it must be contained in the maximal subhypergraph with these properties, i.e. in the loose core.

1.3 Important parameters

We will now define various parameters which will occur often in this paper. Some of these definitions may seem arbitrary and unmotivated initially, but their meaning will become clearer over the course of the paper.

Given $d > 0$, consider a sequence $(d_n)_{n\in\mathbb{N}}$ of real numbers such that $d_n \to d$. Then for $r \in \mathbb{N}_{\geqslant 3}$ and $n \in \mathbb{N}$, set

$$p = p(r, n) \coloneqq \frac{d_n}{\binom{n-1}{r-1}}, \qquad d^* = d^*(r) \coloneqq \frac{1}{r-1}.$$

In addition we define a function $F : [0, \infty) \to \mathbb{R}$ by setting

$$F(x) = F_{r,d}(x) \coloneqq \exp\left(-d\left(1 - x^{r-1}\right)\right) \tag{1}$$

and let $\rho_* = \rho_*(r,d)$ be the largest solution[2] of the fixed-point equation

$$1-\rho = F(1-\rho). \tag{2}$$

Since the function F is dependent on d, so too are the solutions to this equation. It turns out that d^* is a threshold at which the solution set changes its behaviour from containing only the trivial solution 0 to also containing a unique positive solution (see Claim 10).

We define

$$\hat{\rho}_* = \hat{\rho}_*(r,d) \coloneqq 1-(1-\rho_*)^{r-1} \tag{3}$$

and

$$\eta = \eta(r,d) \coloneqq 1 - \frac{(r-1)\rho_*(1-\rho_*)^{r-2}}{\hat{\rho}_*}. \tag{4}$$

Furthermore let

$$\begin{aligned} \alpha = \alpha(r,d) &\coloneqq \rho_*\left(1-d(r-1)(1-\rho_*)^{r-1}\right), \\ \beta = \beta(r,d) &\coloneqq \frac{d}{r}\left(1-(1-\rho_*)^r - r\rho_*(1-\rho_*)^{r-1}\right), \end{aligned} \tag{5}$$

and

$$\gamma = \gamma(r,d) \coloneqq 1-\exp(-d\hat{\rho}_*) - d\hat{\rho}_*\exp(-d\hat{\rho}_*). \tag{6}$$

1.4 Main results: loose cores and cycles in hypergraphs

For any $j \in \mathbb{N}_{\geqslant 1}$, let $v_j(C_H)$ be the number of vertices of $H = H^r(n,p)$ with degree j in the loose core C_H of H and let

$$\mu_j \coloneqq v_j(C_H)\cdot n^{-1}.$$

Let $v(C_H) = \sum_{j\geqslant 1} v_j(C_H)$ denote the number of vertices and $e(C_H)$ the number of edges in the loose core C_H in H. We also define $v_0(C_H)$ to be the number of vertices of H which are not in the loose core of H (so $v_0(C_H) = n - v(C_H)$), and $\mu_0 \coloneqq v_0(C_H)\cdot n^{-1}$. (Observe that this notation is consistent if, with a slight abuse of terminology, we view vertices which are not in the loose core as having degree 0 in the loose core.)

We use the notation $\mathrm{Po}(\lambda), \mathrm{Ber}(q), \mathrm{Bi}(N,q)$ to denote Poisson, Bernoulli and Binomial random variables, respectively, with the given parameters. We interpret a $\mathrm{Po}(0)$ variable as being deterministically 0.

Our first main result describes the asymptotic degree distribution of vertices in the loose core C_H of $H = H^r(n,p)$.

Theorem 4. *Let $r, d, p, \hat{\rho}_*$ and η be as in Section 1.3 and let $H = H^r(n,p)$. Let Y be a random variable with distribution $\mathrm{Po}(d\hat{\rho}_*)$ and define*

$$Z \coloneqq \begin{cases} Y & \text{if } Y \neq 1, \\ \mathrm{Ber}\,(\eta) & \text{if } Y = 1. \end{cases}$$

Then there exists $\varepsilon = \varepsilon(n) = o(1)$ such that whp for any constant $j \in \mathbb{N}$ we have

$$\mu_j = \mathbb{P}(Z=j) \pm \varepsilon.$$

[2] ρ_* is well-defined since 0 is certainly a solution and the set of solutions is closed by continuity.

Our second main result describes the asymptotic numbers of vertices and edges in the loose core C_H of $H = H^r(n,p)$.

Theorem 5. *Let r, p, α and β be as in Section 1.3 and let $H = H^r(n,p)$. Then whp*

$$v(C_H) = (\alpha + o(1))n$$

and

$$e(C_H) = (\beta + o(1))n.$$

By analysing the loose core we obtain an upper bound on the length of the longest loose cycle in $H^r(n,p)$.

Theorem 6. *Let r, p, β and γ be as in Section 1.3 and let $H = H^r(n,p)$. Let L_C be the length of the longest loose cycle in H. Then whp*

$$L_C \leqslant (\min\{\beta, \gamma\} + o(1)) \cdot n.$$

In fact, a standard "sprinkling" argument shows that whp the longest loose path in $H^r(n,p)$ is not significantly longer than the longest loose cycle and therefore we obtain the following corollary.

Corollary 7. *Let r, p, β and γ be as in Section 1.3 and let $H = H^r(n,p)$. Let L_P be the length of the longest loose path in H. Then whp*

$$L_P \leqslant (\min\{\beta, \gamma\} + o(1)) \cdot n.$$

As mentioned previously, Claim 10 will state that $d^* = \frac{1}{r-1}$ is a threshold at which the solution set of (2) changes its behaviour from containing only 0 to also containing a unique positive solution. Together with Theorem 6 and Corollary 7 (and recalling the definitions of β, γ in (5) and (6)) this implies that $d^* = \frac{1}{r-1}$ is a threshold for the existence of a loose path/cycle of linear order, and it is interesting in particular to examine the behaviour shortly after the phase transition.

Corollary 8. *Let $r \in \mathbb{N}_{\geqslant 3}$, let $\varepsilon > 0$ be constant and let $p = \frac{1+\varepsilon}{(r-1)\binom{n-1}{r-1}}$. Let L_C and L_P be the length of the longest loose cycle and the longest loose path in $H^r(n,p)$. Then whp*

$$L_C \leqslant L_P + 1 \leqslant \left(\frac{2\varepsilon^2}{(r-1)^2} + O(\varepsilon^3) \right) \cdot n.$$

In other words, we have an upper bound on L_C and L_P in the barely supercritical regime. For a corresponding lower bound, we will quote (a special case of) a result from [10] (which we later state formally as Theorem 25) which gives a lower bound on L_P. By applying the sprinkling argument again we also obtain a lower bound on L_C, and together with Corollary 8 we obtain the following.

Theorem 9. *Let $r \in \mathbb{N}_{\geqslant 3}$, let $\varepsilon > 0$ be constant and let $p = \frac{1+\varepsilon}{(r-1)\binom{n-1}{r-1}}$. Let L_C and L_P be the length of the longest loose cycle and the longest loose path in $H^r(n,p)$. Then whp*

$$\left(\frac{\varepsilon^2}{4(r-1)^2} + O(\varepsilon^3)\right) \cdot n \leqslant L_C \leqslant L_P + 1 \leqslant \left(\frac{2\varepsilon^2}{(r-1)^2} + O(\varepsilon^3)\right) \cdot n.$$

Theorem 9 provides the best known upper and lower bounds on L_P, L_C in the regime when $p = \frac{1+\varepsilon}{(r-1)\binom{n-1}{r-1}}$, but there is a multiplicative factor of 8 between these two bounds and the correct asymptotic value is still unknown. Indeed it is not even clear that the random variables L_P, L_C are concentrated around a single value.

We note that even for cycles in $G(n,p)$ in the barely supercritical regime the correct asymptotic value is not known despite considerable efforts in this direction. In particular, the best known bounds when $0 < \varepsilon = \varepsilon(n) = o(1)$ satisfies $\varepsilon^3 n \to \infty$ and $p = \frac{1+\varepsilon}{n}$ are due to Łuczak [27] (lower bound) and Kemkes and Wormald [25] (upper bound), and state that whp the length L_C of the longest cycle satisfies

$$\left(\frac{4}{3} + o(1)\right)\varepsilon^2 n \leqslant L_C \leqslant (1.7395 + o(1))\varepsilon^2 n.$$

Very recently, Anastos [3] announced an improvement of the lower bound.

The proofs of all the results of this section appear in Section 4 as a consequence of a single, more general result (Theorem 21).

1.5 Key proof techniques

In order to prove our main results, we transfer the problem from $H^r(n,p)$ to the *factor graph* $G := G(H^r(n,p))$ which will be formally defined in Section 3. In the factor graph we define the *reduced core* R_G, which is closely related to the 2-core of G and from which we can reconstruct the loose core of $H^r(n,p)$, but which is easier to analyse. We use a peeling process (Definition 26) and an auxiliary algorithm called `CoreConstruct` to determine the asymptotic proportion of variable and factor nodes of G with fixed degree in the reduced core (Theorem 21). We also need martingale techniques, in particular an Azuma-Hoeffding inequality and an associated edge-exposure martingale to show concentration of the numbers of vertices and edges of fixed degree around the respective expectations.

1.6 Paper overview

The rest of the paper is structured as follows.

In Section 2 we set basic notation and state some standard probabilistic lemmas which we will use later. In Section 3 we switch our focus to factor graphs, define the reduced core and state Theorem 21 which describes degree distributions in the reduced core and which implies all of our main results, as we prove in Section 4.

Subsequently, Section 5 describes a standard peeling process to obtain the reduced core and contains two main lemmas which together imply Theorem 21. The first of these (Lemma 27) describes the degree distribution after a sufficiently large number of steps of

the peeling process, and will be proved in Section 6. The second main lemma (Lemma 28) states that subsequently, very few further vertices will be deleted in the remainder of the peeling process, and therefore this degree distribution is also a good approximation for the degree distribution in the reduced core. Lemma 28 will be proved in Section 7.

In Section 8, we conclude with some discussion and open questions. We omit from the main body of the paper many proofs which simply involve technical calculations or standard applications of common methods, but include some of them in the appendices for completeness. Appendix A contains an analysis of the fixed-point equation (2), while Appendix B contains the proofs of some basic probabilistic lemmas which are needed throughout the paper. Finally, Appendix C and Appendix D constitute the proofs of Lemma 32 and Lemma 41, respectively.

2 Preliminaries and Notation

For the rest of the paper, $r \in \mathbb{N}_{\geqslant 3}$ and $d > 0$ will be fixed. In particular, we consider these to be constant, so if we say, for example, that $x = O(n)$, we mean that there exists a constant $C = C(r,d)$ such that $x \leqslant Cn$. By the notation $x = a \pm b$ we mean that $a - b \leqslant x \leqslant a + b$. Similarly, the notation $x = (a \pm b)c$ means that $(a-b)c \leqslant x \leqslant (a+b)c$. We will omit floors and ceilings whenever these do not significantly affect the argument.

As mentioned in Section 1.3, the solution set of the fixed-point equation (2) changes its behaviour at $d = d^*$. More precisely we have the following.

Claim 10.

(F1) *If $d < d^*$, then $\rho_* = 0$.*

(F2) *If $d > d^*$, then there is a unique positive solution to (2).*

We defer the (elementary, but rather technical) proof of this claim to Appendix A. Furthermore we will often use the following alternative relation between ρ_* and $\hat{\rho}_*$.

$$1 - \rho_* \overset{(2)}{=} F(1-\rho_*) = \exp\left(-d\left(1-(1-\rho_*)^{r-1}\right)\right) \overset{(3)}{=} \exp(-d\hat{\rho}_*). \tag{7}$$

2.1 Large deviation bounds

In this section, we collect some standard large deviation results which will be needed later. We will use the following Chernoff bound (see e.g. [22, Theorem 2.1]).

Lemma 11 (Chernoff)**.** *If $X \sim \mathrm{Bi}(N,p)$, then for any $s > 0$*

$$\mathbb{P}(|X - Np| \geqslant s) \leqslant 2 \cdot \exp\left(-\frac{s^2}{2\left(Np + \frac{s}{3}\right)}\right).$$

This bound is less precise than the form in [22] since we have combined the upper and lower tail bounds for simplicity.

We will also use a variant of the Azuma-Hoeffding inequality due to Warnke [38], in which the standard Lipschitz condition requirement is weakened by requiring only that it "usually" holds. The following is a simplified form of [38, Theorem 1.3].

Lemma 12. *Suppose $X = (X_1, \dots, X_N)$ is a sequence of independent Bernoulli(p) variables and $\Gamma \subset \{0,1\}^N$ is an event. Suppose that a function $f : \{0,1\}^N \to \mathbb{R}$ satisfies the following condition.*

(∗) There are real numbers $c \leqslant d$ such that for any $x, \tilde{x} \in \{0,1\}^N$ which differ in only one co-ordinate, we have

$$|f(x) - f(\tilde{x})| \leqslant \begin{cases} c & \text{if } x \in \Gamma, \\ d & \text{otherwise.} \end{cases}$$

Then setting $e := \frac{1}{n}(d-c)$, we have

$$\mathbb{P}\Big(\big|f(X) - \mathbb{E}(f(X))\big| \geqslant t\Big) \leqslant 2\exp\left(-\frac{t^2}{2Np(c+e)^2 + 2(c+e)t/3}\right) + nN\mathbb{P}(X \notin \Gamma).$$

Proof. We briefly describe how this lemma is indeed implied by [38, Theorem 1.3]. We have applied that theorem in a symmetric setting where $p_k = p$, where $c_k = c$, where $d_k = d$ and where $\gamma_k = \frac{1}{n}$ for every $k \in [N]$. Then we also obtain $e_k := \gamma_k(d_k - c_k) = e$ for every $k \in [N]$, and $C := \max_{k\in[N]}(c_k + e_k) = c + e$.

While [38, Theorem 1.3] only gives an upper-tail bound (involving the event $f(X) \geqslant \mathbb{E}(f(X)) + t$ rather than $\big|f(X) - \mathbb{E}(f(X))\big| \geqslant t$), as remarked in that paper, applying the same theorem to $-f(X)$ gives the corresponding lower-tail bound; a union bound on the error probabilities leads to the factor of 2 before the exponential above.

Finally, we have also used the fact that

$$\mathbb{P}\Big(\big|f(X) - \mathbb{E}(f(X))\big| \geqslant t\Big) \leqslant \mathbb{P}\Big(\big|f(X) - \mathbb{E}(f(X))\big| \geqslant t \text{ and } \neg\mathcal{B}\Big) + \mathbb{P}(\mathcal{B})$$

for any event $\mathcal{B}$, and [38, Theorem 1.3] guarantees the existence of an appropriate event $\mathcal{B}$ with $\mathbb{P}(\mathcal{B}) \leqslant \left(\sum_{k\in\mathbb{N}} \gamma_k^{-1}\right) \mathbb{P}(X \notin \Gamma) = nN\mathbb{P}(X \notin \Gamma)$. □

3 Factor graphs

There is a natural representation of a hypergraph as a bipartite graph known as a *factor graph*, which is a well-known concept in literature (see e.g. [30]). Although any hypergraph can be represented as a factor graph, for the purposes of this paper we only need and define the notion for r-uniform hypergraphs. In particular, with a slight abuse of terminology, whenever we refer to a "factor graph", we implicitly mean the factor graph of an r-uniform hypergraph.

Definition 13 (Factor graph)**.** Given an r-uniform hypergraph H, the *factor graph* $G = G(H)$ of H is a bipartite graph on vertex classes $\mathcal{V} \coloneqq V(H)$ and $\mathcal{F} \coloneqq E(H)$, where $v \in \mathcal{V}$ and $a \in \mathcal{F}$ are joined by an edge in G if and only if $v \in a$. In other words, the vertices of G are the vertices and edges of H, and the edges of G represent incidences.

To avoid confusion, we refer to the vertices of a factor graph as *nodes*. In particular, the nodes in $\mathcal{V}$ are called *variable nodes* and the nodes in $\mathcal{F}$ are called *factor nodes*.[3] We define

$$G^r(n,p) \coloneqq G(H^r(n,p)),$$

i.e. the factor graph of the r-uniform binomial random hypergraph $H^r(n,p)$.

Note that if H is an r-uniform hypergraph, then the factor nodes of $G(H)$ all have degree r. We will need the following basic fact about the number of factor nodes in $G^r(n,p)$. We omit the proof, which is a simple application of a Chernoff bound (Lemma 11).

Proposition 14. *Let $d > 0$ be a constant and let $p = \frac{(1+o(1))d}{\binom{n}{r-1}}$. Then there exists a function $\omega_0 = \omega_0(n)$ with $\omega_0 \xrightarrow{n\to\infty} \infty$ such that whp the number m of factor nodes in $G^r(n,p)$ satisfies*

$$m = \left(1 \pm \frac{1}{\omega_0}\right)\frac{dn}{r}.$$

It will be more convenient to study the factor graph than the original hypergraph—in order to do this, we need to understand what the structure corresponding to the loose core looks like in the factor graph. We first define the loose core of a factor graph and subsequently observe that it does indeed correspond to the loose core of the hypergraph (Definition 2).

Definition 15 (Loose core)**.** The *loose core* $C = C_G$ of a factor graph G is the maximal subgraph of G such that each factor node of C has degree r in C and furthermore:

(C1') C contains no isolated variable nodes;

(C2') Each factor node in C is adjacent to at least two variable nodes of degree at least two in C.

Proposition 16. *Given an r-uniform hypergraph H, the loose core C_G of the factor graph $G = G(H)$ of H is identical to the factor graph of the loose core C_H of H.*

Proof. The condition that each factor node of $C = C_G$ has degree r in C means that C corresponds to a subhypergraph of H (i.e. no edge of H has a vertex removed from it without itself being removed). Since variable nodes of G correspond to vertices of H, condition **(C1')** in Definition 15 corresponds precisely to **(C1)** in Definition 2. Furthermore, condition **(C2')** in Definition 15 is directly analogous to condition **(C2)** in Definition 2. □

[3] In some contexts in the literature, factor nodes may be called *functional nodes* or *constraint nodes*.

In view of Proposition 16, rather than studying the loose core of the hypergraph, we can study the loose core of the corresponding factor graph instead. In fact, even more convenient than this is a slightly different structure.

Definition 17 (Reduced core)**.** The *reduced core* $R = R_G$ of a factor graph G is the maximal subgraph of G with no nodes of degree 1.

Note that the reduced core is very similar to the 2-core of G—the only difference is that we do not delete isolated nodes, so all original nodes are still present. This will be convenient since it means that all nodes have a well-defined degree within the reduced core (and have degree zero if and only if they are not in the 2-core of G). Similarly we will want to describe degree distributions within the loose core, but also incorporating nodes which are in fact not contained in the loose core. To avoid confusion and abuse of terminology, we define the *padded core*.

Definition 18 (Padded core)**.** The *padded core* $P = P_G$ of a factor graph G is the subgraph of G whose nodes are the nodes of G and whose edges are the edges of C_G.

In other words, the padded core P_G is identical to the loose core C_G except that all nodes of G are still present. Equivalently, P_G is the maximal subgraph of G in which each non-isolated factor node has degree r and is adjacent to at least two variable nodes of degree at least 2. The following observation motivates both our definition of the padded core and the interpretation of μ_0 as the proportion of vertices of $H^r(n,p)$ which do not lie in the loose core.

Remark 19. For each $j \in \mathbb{N}$, the proportion of variable nodes of $G = G^r(n,p)$ which have degree j in the padded core P_G is μ_j.

It is important to observe that, if we have found the reduced core, it is very easy to reconstruct the padded core, and hence also the loose core. Let $\mathcal{F}_R$ be the set of non-isolated factor nodes of the reduced core R_G and let P_G' be the factor graph whose nodes are the nodes of G and whose edges are all edges of G incident to $\mathcal{F}_R$. In other words, P_G' is the factor graph obtained from R_G by adding back in all edges of G attached to non-isolated factor nodes of R_G.

Proposition 20. *Let G be the factor graph of an r-uniform hypergraph. Then $P_G' = P_G$.*

Proof. Let R_1 denote the graph obtained from the padded core P_G of G by removing all edges incident to leaves (which must be variable nodes). Note that, since any non-isolated factor node in P_G has at least two neighbours of degree at least two, the same is still true in R_1. For the sake of intuitive notation, we also denote

$$R_2 \coloneqq R_G, \qquad P_1 \coloneqq P_G, \qquad P_2 \coloneqq P_G'.$$

Our goal is to show that $P_1 = P_2$.

Let us observe that P_1 can be obtained from R_1 by the same operation with which P_2 is obtained from R_2, namely by adding in edges of G incident to non-isolated factor nodes.

We next observe that R_1 is a subgraph of G with no nodes of degree 1, and therefore $R_1 \subseteq R_2 = R_G$, by the maximality of R_G. Since the operation constructing P_i from R_i is inclusion-preserving, $R_1 \subseteq R_2$ implies that $P_1 \subseteq P_2$.

It therefore remains to prove that $P_2 \subseteq P_1$. To this end, we observe that certainly $P_2 = P'_G$ is a subgraph of G in which each non-isolated factor node is in the 2-core of G, and therefore adjacent to at least two variable nodes of degree at least two. Furthermore each non-isolated factor node of P_2 has degree r in C_2, and since $P_1 = P_G$ is the maximal subgraph with these two properties, we have $P_2 \subseteq P_1$, as required. □

Let us observe one further fact about the transformation from the reduced core R_G to the padded core $P_G = P'_G$: although this seemed to be dependent on the initial factor graph G, in fact the operation simply involves connecting non-isolated factor nodes of R_G to (distinct) isolated variable nodes until each factor node has degree precisely r. This means that given R_G, by Proposition 20 we can describe P_G (and therefore also the loose core C_G) entirely, up to the assignment of which nodes are leaves. In other words, R_G already contains all of the "essential" information of both P_G and C_G. It will therefore be enough to study R_G rather than P_G or C_G, and this turns out to be simpler.

Now the main results of this paper are implied by the following theorem about the reduced core R_G of the factor graph $G = G^r(n,p)$ of the r-uniform binomial random hypergraph.

For a non-negative real number λ, let us denote by $\widetilde{\mathrm{Po}}(\lambda)$ the distribution of a random variable X satisfying

$$\mathbb{P}(X=j) = \begin{cases} \mathbb{P}(\mathrm{Po}(\lambda) \leqslant 1) & \text{if } j = 0, \\ 0 & \text{if } j = 1, \\ \mathbb{P}(\mathrm{Po}(\lambda) = j) & \text{if } j \geqslant 2. \end{cases}$$

In other words, the $\widetilde{\mathrm{Po}}$ distribution is identical to the Po distribution except that values of 1 are replaced by 0. We define the $\widetilde{\mathrm{Bi}}$ distribution analogously.

Theorem 21. *Let $r, d, p, \rho_*, \hat{\rho}_*$ be as in Section 1.3 and let $G = G^r(n,p)$, i.e. the factor graph of $H^r(n,p)$. For each $j \in \mathbb{N}$, let ξ_j and $\hat{\xi}_j$ be the proportion of variable nodes and factor nodes of G respectively which have degree j in the reduced core R_G of G. Then there exists a function $\varepsilon = \varepsilon(n) = o(1)$ such that whp for any constant $j \in \mathbb{N}$ we have*

$$\xi_j = \mathbb{P}(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = j) \pm \varepsilon$$

and

$$\hat{\xi}_j = \mathbb{P}(\widetilde{\mathrm{Bi}}(r, \rho_*) = j) \pm \varepsilon.$$

In other words, within R_G, variable nodes and factor nodes have degree distributions which are asymptotically those of a $\widetilde{\mathrm{Po}}(d\hat{\rho}_*)$ and a $\widetilde{\mathrm{Bi}}(r, \rho_*)$ distribution respectively.

The proof of this theorem will form the main body of the paper. In Section 5 we will prove how Theorem 21 follows from two auxiliary statements, stating that for some large integer ℓ the proportions of variable and factor nodes of degree j in the graph obtained after ℓ rounds of a peeling process are approximately the values given in Theorem 21 (Lemma 27), and furthermore not many nodes are deleted after round ℓ (Lemma 28).

4 Back to hypergraphs: Proofs of main results

We now show how all of the results of Section 1.4 follow from Theorem 21. First we deduce our result on the asymptotic degree distribution of vertices in the loose core of $H^r(n,p)$.

Proof of Theorem 4. We will apply Theorem 21 to provide us with a function ε, and we will prove Theorem 4 with $\varepsilon' := \sqrt{\varepsilon} + \frac{1}{\sqrt{\omega_0}}$, where $\omega_0 = \omega_0(n)$ is the function given by Proposition 14.

For convenience, for any $j \in \mathbb{N}$, let us define

$$\mu'_j := \begin{cases} \mathbb{P}(\mathrm{Po}(d\hat{\rho}_*) = j) & \text{if } j \geqslant 2; \\ \eta \cdot \mathbb{P}(\mathrm{Po}(d\hat{\rho}_*) = j) & \text{if } j = 1; \\ \mathbb{P}(\mathrm{Po}(d\hat{\rho}_*) = 0) + (1-\eta) \cdot \mathbb{P}(\mathrm{Po}(d\hat{\rho}_*) = 1) & \text{if } j = 0. \end{cases}$$

In other words, μ'_j is the "idealised version" of μ_j, and our goal is simply to prove that whp, for each $j \in \mathbb{N}$ we have $\mu_j = \mu'_j \pm \varepsilon$. Similarly we also define

$$\xi'_j := \mathbb{P}(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = j),$$

so by Theorem 21 we have $\xi_j = \xi'_j \pm \varepsilon$ whp for each $j \in \mathbb{N}$. The proof of Theorem 4 now simply consists of relating the μ_j to the ξ_j, relating the μ'_j to the ξ'_j and applying Theorem 21. Note that it follows instantly from the definitions that $\mu'_j = \xi'_j$ for $j \in \mathbb{N}_{\geqslant 2}$. We will split the proof into three cases.

Case 1: $j \geqslant 2$.
We start by showing that $\mu_j = \xi_j$. Observe that by Remark 19, μ_j is simply the proportion of variable nodes with degree j in the padded core P_G of $G = G^r(n,p)$. Theorem 21 tells us the degrees of variable and factor nodes in the reduced core R_G of G. By Proposition 20, moving from R_G to P_G means that we connect all non-isolated factor nodes of R_G to their original neighbours in G, and any variable nodes which receive additional incident edges in this process have their degrees changed from 0 to 1. It follows that for $j \geqslant 2$, the proportion μ_j of variable nodes in G with degree j in the padded core P_G is precisely equal to ξ_j, the proportion of variable nodes in G with degree j in the reduced core R_G. Therefore

$$\mu_j = \xi_j \overset{\text{Th. 21}}{=} \xi'_j \pm \varepsilon = \mu'_j \pm \varepsilon,$$

and the statement of Theorem 4 is certainly true for $j \geqslant 2$ (indeed, we have proved something stronger since $\varepsilon < \varepsilon'$).

Case 2: $j = 1$.
To prove the case $j = 1$, we need to consider how many isolated variable nodes become leaves when moving from R_G to P_G. Since by Proposition 20 every factor node of R_G with degree $j \geqslant 2$ has $r - j$ leaves connected to it, and since whp the number m of factor nodes in total is $m = \left(1 \pm \frac{1}{\omega_0}\right)\frac{dn}{r}$ for some growing function $\omega_0 \xrightarrow{n\to\infty} \infty$ by Proposition 14, whp

the number of leaves added, which is simply $\mu_1 n$, satisfies

$$\mu_1 n = \sum_{j=2}^{r}(r-j)\hat{\xi}_j m = \left(1 \pm \frac{1}{\omega_0}\right)\frac{dn}{r}\sum_{j=2}^{r}(r-j)\left(\mathbb{P}(\widetilde{\mathrm{Bi}}(r,\rho_*)=j) \pm \varepsilon\right)$$
$$= \frac{dn}{r}\sum_{j=2}^{r}(r-j)\mathbb{P}(\widetilde{\mathrm{Bi}}(r,\rho_*)=j) \pm \frac{\varepsilon' n}{2}, \tag{8}$$

where the last line follows since $\frac{1}{\omega_0}, \varepsilon = o\left(\sqrt{\varepsilon} + \frac{1}{\sqrt{\omega_0}}\right) = o(\varepsilon')$. The sum can be estimated using the definition of the $\widetilde{\mathrm{Bi}}$ distribution and equations (3) and (7):

$$\begin{aligned}
&\textstyle\sum_{j=2}^{r}(r-j)\mathbb{P}(\widetilde{\mathrm{Bi}}(r,\rho_*)=j)\\
&\quad= \textstyle\sum_{j=0}^{r}(r-j)\mathbb{P}(\mathrm{Bi}(r,\rho_*)=j) - r(1-\rho_*)^r - (r-1)r\rho_*(1-\rho_*)^{r-1}\\
&\quad= r(1-\rho_*)\left(1-(1-\rho_*)^{r-1} - (r-1)\rho_*(1-\rho_*)^{r-2}\right)\\
&\quad\overset{(3),(7)}{=} r\exp(-d\hat{\rho}_*)\left(\hat{\rho}_* - (r-1)\rho_*(1-\rho_*)^{r-2}\right).
\end{aligned}$$

Substituting this into (8) gives

$$\mu_1 = d\exp(-d\hat{\rho}_*)\left(\hat{\rho}_* - (r-1)\rho_*(1-\rho_*)^{r-2}\right) \pm \varepsilon'/2. \tag{9}$$

On the other hand, we have

$$\mu_1' = \eta\cdot\mathbb{P}(\mathrm{Po}(d\hat{\rho}_*)=1) = \left(1 - \frac{(r-1)\rho_*(1-\rho_*)^{r-2}}{\hat{\rho}_*}\right)d\hat{\rho}_*\exp(-d\hat{\rho}_*)$$
$$= d\exp(-d\hat{\rho}_*)\left(\hat{\rho}_* - (r-1)\rho_*(1-\rho_*)^{r-2}\right),$$

which combined with (9) tells us that

$$\mu_1 = \mu_1' \pm \varepsilon'/2, \tag{10}$$

which is in fact slightly stronger than required.

Case 3: $j = 0$.
Finally to prove the statement for $j = 0$, note that $\mu_0 = \xi_0 - \mu_1$ (deterministically). Furthermore, we have $\sum_{j=0}^{\infty}\mu_j' = \sum_{j=0}^{\infty}\xi_j' = 1$, and we have already observed that $\mu_j' = \xi_j'$ if $j \geqslant 2$, and therefore $\mu_0' + \mu_1' = \xi_0' + \xi_1'$. Observing also that $\xi_1' = 0$, we deduce that $\mu_0' = \xi_0' - \mu_1'$. Therefore, applying Theorem 21 (for $j = 0$) and (10), we obtain

$$\mu_0 = \xi_0 - \mu_1 = \xi_0' \pm \varepsilon - (\mu_1' \pm \varepsilon'/2) = \mu_0' \pm \varepsilon'$$

as required. □

With a little more calculation we can also determine the number of vertices and edges in the loose core, and therefore also prove Theorem 5.

Proof of Theorem 5. Observe that the number of vertices in the loose core of $H = H^r(n,p)$ is simply the number of variable nodes of $G = G(H)$ which have degree at least one in the padded core of G, and thus the proportion of such vertices is $1 - \mu_0$ (see Remark 19). By Theorem 4, whp

$$\begin{aligned} 1 - \mu_0 &= 1 - \exp(-d\hat{\rho}_*) - d\hat{\rho}_* \exp(-d\hat{\rho}_*)(1-\eta) + o(1) \\ &\overset{(3),(4),(7)}{=} \rho_* - d(1-\rho_*)(r-1)\rho_*(1-\rho_*)^{r-2} + o(1) \\ &= \rho_*(1 - d(r-1)(1-\rho_*)^{r-1}) + o(1) = \alpha + o(1), \end{aligned}$$

precisely as stated in Theorem 5.

The number of edges in the loose core of H is the number of factor nodes with degree at least 1 in R_G, which is $(1-\hat{\xi}_0)m$, where recall that m denotes the total number of factor nodes of G. Applying Theorem 21 to estimate $\hat{\xi}_0$ and Proposition 14 to estimate m, we deduce that whp the number of edges in the loose core is

$$\left(1-\hat{\xi}_0\right) m = \left(1 - (1-\rho_*)^r - r\rho_*(1-\rho_*)^{r-1} \pm o(1)\right) \frac{(1+o(1))dn}{r} \overset{(5)}{=} (\beta + o(1))\, n,$$

as claimed. □

Now we can also prove the bound on the length of the longest loose cycle in Theorem 6.

Proof of Theorem 6. Let us observe that for any loose cycle in $H = H^r(n,p)$, the edges and the vertices which lie in two edges form a cycle in the factor graph $G = G(H)$, which must clearly lie within the reduced core R_G of G. Thus the length of the loose cycle is bounded both by the number of variable nodes and the number of factor nodes which are not isolated in R_G. In other words, the length L_C of the longest loose cycle (deterministically) satisfies

$$L_C \leqslant \min\left\{(1-\xi_0)n\ ,\ (1-\hat{\xi}_0)m\right\}. \tag{11}$$

By Proposition 14 we have that whp $m = (1+o(1))\,\frac{dn}{r}$. Observe also that by Theorem 21, whp ξ_0 is asymptotically

$$\mathbb{P}(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = 0) = \mathbb{P}(\mathrm{Po}(d\hat{\rho}_*) \leqslant 1) = \exp(-d\hat{\rho}_*)(1+d\hat{\rho}_*) = 1-\gamma,$$

while whp $\hat{\xi}_0$ is asymptotically

$$\mathbb{P}(\widetilde{\mathrm{Bi}}(r,\rho_*) = 0) = \mathbb{P}(\mathrm{Bi}(r,\rho_*) \leqslant 1) = (1-\rho_*)^r + r\rho_*(1-\rho_*)^{r-1} = 1 - \frac{\beta r}{d}.$$

Substituting these values into (11) gives the bound in Theorem 6. □

Our next goal is to prove the remaining results of Section 1.4, for which we will need to relate L_P and L_C. To do this, we use a standard sprinkling argument.

Lemma 22. *Let $\omega = \omega(n)$ be any function and $p_1 = p_1(n)$ and $p_2 = p_2(n)$ be any probabilities satisfying*

1. $p_1 \leqslant \left(1 + \frac{1}{\omega}\right) p_1 \leqslant p_2$;
2. $p_1 n^r/\omega \to \infty$.

Suppose that $H_1 \sim H^r(n, p_1)$ and $H_2 \sim H^r(n, p_2)$ are coupled in such a way that $H_1 \subset H_2$. For $i = 1, 2$, let $L_P^{(i)}, L_C^{(i)}$ denote the length of the longest loose path and loose cycle, respectively, in H_i. Then whp

$$L_C^{(2)} \geqslant L_P^{(1)} + o(n).$$

We defer the proof of this lemma to Appendix B.1. The following slightly different form will be a little more convenient to apply. We omit the proof, which is elementary given Lemma 22.

Corollary 23. *Given the setup of Lemma 22, the following hold.*

1. *If there exists a constant ζ_1 such that whp $L_P^{(1)} \geqslant (\zeta_1 + o(1))n$, then whp $L_C^{(2)} \geqslant (\zeta_1 + o(1))n$.*
2. *If there exists a constant ζ_2 such that whp $L_C^{(2)} \leqslant (\zeta_2 + o(1))n$, then whp $L_P^{(1)} \leqslant (\zeta_2 + o(1))n$.*

We also need a further technical result which states that the parameters β, γ, with which we bound L_C in Theorem 6, are continuous in p (except at the threshold $p = d^*/\binom{n-1}{r-1}$). Let r, d, p be as in Section 1.3. Let $p' = (1 + 1/\omega)p$ for a function $\omega = \omega(n) \to \infty$ but $\omega = o(\log n)$. The following lemma shows that if we replace p by p', the parameters β, γ remain essentially the same. The (technical) proof can be found in Appendix A.

Lemma 24. *Let β, γ be defined as in (5) and (6), and let β', γ' be defined similarly but with p' in place of p. If $d \neq d^*$, then*

$$\min\{\beta', \gamma'\} = \min\{\beta, \gamma\} + o(1).$$

We can now bound the length of the longest loose path in $H^r(n, p)$.

Proof of Corollary 7. Let us set $\omega = 1/(\log n)$, set $p_1 = p$ and set $p_2 = \left(1 + \frac{1}{\omega}\right) p_1$. It is easy to check that these parameters satisfy the assumptions of Lemma 22, and therefore also of Corollary 23. Theorem 6 applied to $H_2 \sim H^r(n, p_2)$ implies that whp $L_C^{(2)} \leqslant (\min\{\beta_2, \gamma_2\} + o(1))n$, where β_2, γ_2 are defined analogously to β, γ, but with $p_2 = (1+1/\omega)p$ in place of p. Furthermore, Lemma 24 implies that $\min\{\beta_2, \gamma_2\} = \min\{\beta, \gamma\} + o(1)$, so we deduce that whp $L_C^{(2)} \leqslant (\min\{\beta, \gamma\} + o(1))n$. Finally, Corollary 23 then implies that whp $L_P = L_P^{(1)} \leqslant (\min\{\beta, \gamma\} + o(1))n$, as required. □

By applying Corollary 7 shortly beyond the phase transition threshold, we are able to prove Corollary 8.

Proof of Corollary 8. Since $\min\{\beta,\gamma\} \leqslant \gamma$ it suffices to show that

$$\gamma = \frac{2\varepsilon^2}{(r-1)^2} + O(\varepsilon^3).$$

(In fact a similar computation for β gives exactly the same result.) By definition

$$\begin{aligned}\gamma &\overset{(6)}{=} 1 - \exp(-d\hat{\rho}_*) - d\hat{\rho}_* \exp(-d\hat{\rho}_*) \\ &= 1 - \left(1 - d\hat{\rho}_* + \frac{d^2\hat{\rho}_*^2}{2} + O\left(\hat{\rho}_*^3\right)\right) - d\hat{\rho}_*\left(1 - d\hat{\rho}_* + O\left(\hat{\rho}_*^2\right)\right) \\ &= \frac{d^2\hat{\rho}_*^2}{2} + O\left(\hat{\rho}_*^3\right) \qquad (12)\end{aligned}$$

Recall from (3) that $\hat{\rho}_*$ was defined as a function of ρ_*, which itself was defined as the largest solution of the fixed-point equation (2). We therefore need to estimate ρ_*. From (2) we obtain

$$\rho = \frac{-d(r-1)+1}{(-\frac{1}{2} - \frac{d}{2}(r-1)(r-2))} + O\left(\rho^2\right).$$

Substituting $d = \frac{1+\varepsilon}{r-1}$ gives

$$\rho = \frac{2\varepsilon}{1+(1+\varepsilon)(r-2)} + O\left(\rho^2\right) = \frac{2\varepsilon}{r-1+O(\varepsilon)} + O\left(\rho^2\right) = \frac{2\varepsilon}{r-1} + O\left(\rho^2\right).$$

In particular this implies that there exists a solution $\rho = \frac{2\varepsilon}{r-1} + O(\varepsilon^2)$ of the fixed point equation (2), and by Claim 10 this is the unique positive solution and therefore $\rho_* = \frac{2\varepsilon}{r-1} + O(\varepsilon^2)$. Substituting this into (3) we obtain

$$\hat{\rho}_* = 1 - \left(1 - \frac{2\varepsilon}{r-1} + O(\varepsilon^2)\right)^{r-1} = 2\varepsilon + O(\varepsilon^2).$$

Substituting this into (12), we obtain

$$\gamma = \frac{2\varepsilon^2}{(r-1)^2} + O\left(\varepsilon^3\right). \qquad \square$$

In order to prove Theorem 9, we also need a lower bound on L_C. We will use a result of [10], which provides a lower bound on L_P together with Lemma 22 to relate L_P and L_C. More precisely, one special case (the supercritical regime for $j = 1$) of [10, Theorem 4] can be reformulated (in a slightly weakened but much simplified way) as follows.

Theorem 25 ([10])**.** *Let L_P denote the length of the longest loose path in $H^r(n,p)$. For all $r \in \mathbb{N}_{\geqslant 3}$ there exists $\varepsilon_0 \in (0,1]$ such that for any function $\varepsilon = \varepsilon(n) < \varepsilon_0$ which satisfies $\varepsilon^5 n \xrightarrow{n\to\infty} \infty$, setting $\delta = \varepsilon/\sqrt{\varepsilon_0}$ the following holds. If $p = \frac{1+\varepsilon}{(r-1)\binom{n-1}{r-1}}$, then whp*

$$(1-\delta)\frac{\varepsilon^2 n}{4(r-1)^2} \leqslant L_P \leqslant (1+\delta)\frac{2\varepsilon n}{(r-1)^2}.$$

Note that Theorem 25 allows for a wider range of ε than we consider in this paper, in particular allowing ε to tend to zero sufficiently slowly. However, there is a $\Theta(1/\varepsilon)$ gap between the upper and lower bounds. Theorem 9 improves the upper bound and thus narrows the gap to just a constant factor.

Proof of Theorem 9. The second and third inequalities are simply the statement of Corollary 7, so it remains to show that whp

$$\left(\frac{\varepsilon^2}{4(r-1)^2} + O(\varepsilon^3)\right) \cdot n \leqslant L_C.$$

Note that we may assume that $\varepsilon < \varepsilon_0$, where ε_0 is the parameter from Theorem 25, since otherwise the $O(\varepsilon^3)$ error term may in fact be the dominant term, and the result is trivial.

Let us set $p_2 = p$ and $p_1 = \left(1 - \frac{1}{\log n}\right) p$. It is easy to check that these parameters satsify the assumptions of Lemma 22, and therefore also of Corollary 23.

It is also clear that $p_1 = \frac{1+\varepsilon_1}{(r-1)\binom{n-1}{r-1}}$, where $\varepsilon_1 = \varepsilon - \frac{1}{\log n} - \frac{\varepsilon}{\log n} = \varepsilon + O(\varepsilon^2)$, and therefore the lower bound in Theorem 25 (together with the observation that $\varepsilon_1/\sqrt{\varepsilon_0} = O(\varepsilon_1)$) states that whp

$$L_P^{(1)} \geqslant \left(\frac{\varepsilon_1^2}{4(r-1)^2} + O(\varepsilon_1^3)\right) \cdot n = \left(\frac{\varepsilon^2}{4(r-1)^2} + O(\varepsilon^3)\right) \cdot n,$$

and an application of Corollary 23 completes the proof. □

5 Peeling process

Recall that for a given hypergraph H the reduced core of the factor graph $G = G(H)$ is defined as the maximum subgraph with no nodes of degree one, which is similar to the 2-core of G except that isolated nodes are not deleted. There is a simple peeling process to obtain the 2-core of G which is a standard procedure and has been used and analysed extensively in the literature. We will consider the obvious adaptation of this process which obtains the reduced core rather than the 2-core.

Definition 26 (Peeling Process)**.** In every round we check whether the factor graph has any nodes of degree one and delete edges incident to such nodes. More precisely, we recursively define a sequence of graphs $(G_i)_{i\in\mathbb{N}}$ where G_0 is the input graph and for $i \in \mathbb{N}_{\geqslant 1}$, G_i is the graph obtained from G_{i-1} by removing all edges incident to nodes of degree one. We say that we *disable* a node if we delete its incident edges.

In the remainder of the paper, whenever we consider G_ℓ the associated input graph G_0 will be $G^r(n,p)$.

Note that deterministically there exists an i_0 such that $G_{i_0} = G_{i_0+k} = R_{G_0}$ for any $k \in \mathbb{N}$. We recall the definition of ξ_j and $\hat{\xi}_j$ in Theorem 21 and observe that

$$\xi_j \coloneqq \lim_{\ell\to\infty} \xi_j^{(\ell)} \quad \text{and} \quad \hat{\xi}_j \coloneqq \lim_{\ell\to\infty} \hat{\xi}_j^{(\ell)},$$

where $\xi_j^{(\ell)}, \hat{\xi}_j^{(\ell)}$ are the proportions of variable nodes and factor nodes respectively which have degree j in G_ℓ for $\ell \in \mathbb{N}$. These limits exist since both $\left(\xi_j^{(\ell)}\right)_\ell$ and $\left(\hat{\xi}_j^{(\ell)}\right)_\ell$ remain constant after a finite number of steps. We will prove Theorem 21 with the help of two lemmas. The first describes the asymptotic distribution of $\xi_j^{(\ell)}$ and $\hat{\xi}_j^{(\ell)}$ for large ℓ.

Lemma 27. *Let $r, d, \rho_*, \hat{\rho}_*$ be as in Section 1.3. There exist an integer $\ell = \ell(n) \in \mathbb{N}$ and a real number $\varepsilon_1 = \varepsilon_1(n) = o(1)$ such that whp, for any constant $j \in \mathbb{N}$*

$$\xi_j^{(\ell)} = \mathbb{P}(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = j) \pm \varepsilon_1$$

and

$$\hat{\xi}_j^{(\ell)} = \mathbb{P}(\widetilde{\mathrm{Bi}}(r, \rho_*) = j) \pm \varepsilon_1.$$

The second lemma states that $\xi_j^{(\ell)}$ and $\hat{\xi}_j^{(\ell)}$ approximate ξ_j and $\hat{\xi}_j$, respectively.

Lemma 28. *Let r, d be as in Section 1.3. For each $j \in \mathbb{N}$, let $\xi_j, \hat{\xi}_j$ be as defined in Theorem 21, let ℓ, ε_1 be as in Lemma 27 and set $\varepsilon_2 := \sqrt{\varepsilon_1}$. Then whp the peeling process will disable at most $\varepsilon_2 n$ nodes after round ℓ. In particular whp, for any constant $j \in \mathbb{N}$*

$$\xi_j = \xi_j^{(\ell)} \pm \varepsilon_2$$

and

$$\hat{\xi}_j = \hat{\xi}_j^{(\ell)} \pm \frac{2\varepsilon_2 r}{d}.$$

Before proving these two lemmas, we show how together they imply our main result.

Proof of Theorem 21. Let $\ell, \varepsilon_1, \varepsilon_2$ be as in Lemmas 27 and 28. Applying these two lemmas, whp we have

$$\xi_j \overset{\text{L.28}}{=} \xi_j^{(\ell)} \pm \varepsilon_2 \overset{\text{L.27}}{=} \mathbb{P}(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = j) \pm (\varepsilon_1 + \varepsilon_2).$$

Similarly, whp we have

$$\hat{\xi}_j \overset{\text{L.28}}{=} \hat{\xi}_j^{(\ell)} \pm \frac{2\varepsilon_2 r}{d} \overset{\text{L.27}}{=} \mathbb{P}(\widetilde{\mathrm{Bi}}(r, \rho_*) = j) \pm \left(\varepsilon_1 + \frac{2\varepsilon_2 r}{d}\right).$$

The statement of Theorem 21 follows by setting $\varepsilon = \varepsilon_1 + \varepsilon_2 \max\{1, 2r/d\}$. □

6 `CoreConstruct` Algorithm: Proof of Lemma 27

6.1 Main algorithm

In this section we will introduce the `CoreConstruct` algorithm, which is related to the peeling process. To do so, we need to define some notation—this notation could apply to any graph, but since we will need it specifically for factor graphs, we introduce it in this (slightly restrictive) setting for clarity.

Definition 29. Let G be a factor graph with variable node set $\mathcal{V}$ and factor node set $\mathcal{F}$. We denote by $d_G(u,v)$ the distance between two nodes $u,v \in \mathcal{V} \cup \mathcal{F}$, i.e. the number of edges in a shortest path between them. For each $\ell \in \mathbb{N}$ and each $w \in \mathcal{V} \cup \mathcal{F}$, we define

$$D_\ell(w) \coloneqq \{u \in \mathcal{V} \cup \mathcal{F} : d_G(w,u) = \ell\}$$

and

$$d_\ell(w) \coloneqq |D_\ell(w)|.$$

Let

$$D_{\leqslant \ell}(w) = \bigcup_{i=0}^{\ell} D_i(w)$$

and

$$N_{\leqslant \ell}(w) \coloneqq G[D_{\leqslant \ell}(w)],$$

i.e. the subgraph of G induced on $D_{\leqslant \ell}(w)$.

We consider a procedure called `CoreConstruct`. Given a factor graph G on node set $\mathcal{V} \cup \mathcal{F}$ and a node $w \in \mathcal{V} \cup \mathcal{F}$, we consider the factor graph as being rooted at w. In particular, neighbours of a node v which are at distance $d_G(v,w)+1$ from w are called *children* of v. Starting at distance $\ell \in \mathbb{N}$ and moving up towards the root w, we recursively delete any node with no (remaining) children; Algorithm 1 gives a formal description of this procedure. We will denote by $D^*_{\ell-i}(w)$ the set of nodes in $D_{\ell-i}(w)$ which survive round i and let $d^*_i(w) \coloneqq |D^*_i(w)|$.

Algorithm 1: `CoreConstruct`

Input: Integer $\ell \in \mathbb{N}$, node $w \in \mathcal{V} \cup \mathcal{F}$, factor graph $N_{\leqslant \ell+1}(w)$
Output: $d^*_1(w)$

$D^*_{\ell+1}(w) = D_{\ell+1}(w)$
for $1 \leqslant i \leqslant \ell$ **do**
  $D^*_{\ell-i+1}(w) \leftarrow D_{\ell-i+1}(w) \setminus \left\{v : N(v) \cap D^*_{\ell-i+2}(w) = \emptyset\right\}$
  $d^*_{\ell-i+1}(w) \leftarrow |D^*_{\ell-i+1}(w)|$

It is rather difficult to analyse the peeling process directly and it turns out that `CoreConstruct` is easier to analyse while also being closely related. `CoreConstruct` is intended to model the effect of the peeling process on the degree of w after ℓ steps (although note that `CoreConstruct` does delete nodes rather than merely disabling them). Note, however, that it does not mirror the peeling process precisely; some nodes may be disabled in the peeling process much earlier than they are deleted in `CoreConstruct`, and some nodes may be deleted in `CoreConstruct` even though they are actually in the reduced core, and are therefore never disabled in the peeling process. Nevertheless, we obtain the following important relation. Recall that G_ℓ is the graph obtained after the ℓ-th round of the peeling process (see Definition 26) and that $d_{G_\ell}(w)$ is the degree of the node w in G_ℓ.

Lemma 30. *Let $\ell \in \mathbb{N}_{\geqslant 1}$ and $w \in \mathcal{V} \cup \mathcal{F}$. If there are no cycles in $N_{\leqslant \ell+1}(w)$, then the output $d_1^*(w)$ of* `CoreConstruct` *with input ℓ, w and $N_{\leqslant \ell+1}(w)$ satisfies*

$$d_{G_\ell}(w) \begin{cases} = d_1^*(w) & \text{if } d_1^*(w) \neq 1, \\ \leqslant d_1^*(w) & \text{if } d_1^*(w) = 1. \end{cases}$$

Proof. For an upper bound, we will show that if a given node $v \in \mathcal{V} \cup \mathcal{F}$ is deleted in round i of `CoreConstruct`, it must have been disabled at some round $i' \leqslant i$ in the peeling process for the reduced core. In particular, by setting $i = \ell$ we immediately obtain $d_{G_\ell}(w) \leqslant d_1^*(w)$. We prove the statement by induction on i.

For $i = 1$, if a node v is deleted in round one of `CoreConstruct`, then v had no children, and therefore it has only one neighbour in $G = G_0$ (its unique parent in $N_{\leqslant \ell+1}(w)$). Thus v will be disabled in round one of the peeling process. Now suppose v is deleted in round $i \geqslant 2$ of `CoreConstruct`, which must mean that all its children (if it had any) are deleted in step $i-1$ of `CoreConstruct`. By the induction hypothesis, all its children are disabled by step at most $i-1$ of the peeling process and so have degree 0 in G_{i-1}. Therefore v itself has degree at most one in G_{i-1} (from its unique parent) and so will be disabled in round i of the peeling process if it has not been disabled already.

It remains to prove that $d_{G_\ell}(w) \geqslant d_1^*(w)$ if $d_1^*(w) \geqslant 2$. Let $j \coloneqq d_1^*(w) \geqslant 2$ be the number of children of the root w which survive `CoreConstruct`. Each such child must have a descendant in $D_{\ell+1}(w)$, otherwise it would not survive `CoreConstruct`. Thus we have j paths of length $\ell+1$ which all meet at w, but are otherwise disjoint (since $N_{\leqslant \ell+1}(w)$ contains no cycles). By induction on i, we deduce that after i rounds of the peeling process, there are j paths of length $\ell+1-i$ which meet only in w, and in particular after ℓ rounds of the peeling process, $d_{G_\ell}(w) \geqslant j$, as required. □

From now on for the rest of this section, we will always have $G = G^r(n,p)$. Observe that if $d_{G_\ell}(w) \in \{0,1\}$, then $d_{R_G}(w) = 0$. However, if ℓ is sufficiently large we can even say that in G_ℓ the degree will almost always be 0.

Proposition 31. *For any integer-valued function $\ell = \ell(n) \xrightarrow{n\to\infty} \infty$ and node w we have $\mathbb{P}\left(d_{G_\ell}(w) = 1\right) = o(1)$.*

Proof. Let us first assume that w is a variable node. For any integer $i \geqslant 1$, let $\mathcal{V}_i$ and $\mathcal{F}_i$ be the set of variable nodes and factor nodes respectively which are disabled in round i of the peeling process. It is an elementary fact about the peeling process that for any integer $i \geqslant 2$ we have $|\mathcal{V}_i| \leqslant |\mathcal{F}_{i-1}|$ and $|\mathcal{F}_i| \leqslant |\mathcal{V}_{i-1}|$ (deterministically), from which it follows that $|\mathcal{V}_i| \leqslant |\mathcal{V}_{i-2}|$ for $i \geqslant 3$. Therefore we have

$$|\mathcal{V}_\ell| \leqslant |\mathcal{V}_{\ell-2}| \leqslant \ldots \leqslant |\mathcal{V}_{\ell-2\lfloor \frac{\ell-1}{2} \rfloor}|.$$

Furthermore, the $\mathcal{V}_i$ are all disjoint, and so we have

$$|\mathcal{V}_\ell| \leqslant \frac{n}{1 + \lfloor \frac{\ell-1}{2} \rfloor} = O(n\ell^{-1}) = o(n)$$

deterministically. Therefore for large enough n we have $|\mathcal{V}_\ell| \in K := \left[n/\sqrt{\ell}\right]_0$ and so by symmetry we have

$$\mathbb{P}(w \in \mathcal{V}_\ell) = \sum_{k \in K} \left(\mathbb{P}(|\mathcal{V}_\ell| = k) \cdot \frac{k}{n} \right) \leqslant \frac{1}{\sqrt{\ell}} \cdot \sum_{k \in K} \mathbb{P}(|\mathcal{V}_\ell| = k) = o(1),$$

as required. The proof when w is a factor node is essentially identical. □

6.2 Analysis of `CoreConstruct`

We proceed with the analysis of `CoreConstruct` and we will choose the parity of ℓ such that $D_{\ell+1}(w)$ consists of variable nodes, i.e. if $w \in \mathcal{V}$, then we will choose ℓ odd, and if $w \in \mathcal{F}$ we will choose ℓ to be even. This convention is merely for technical convenience since it ensures that we know which type of nodes are being considered in round i of `CoreConstruct` and thus avoid a case distinction.

Let $w \in \mathcal{V} \cup \mathcal{F}$ and $\ell \in \mathbb{N}_{\geqslant 1}$ be given. We say the event $E_w(\ell)$ holds if *neither* of the following two events occur.

(E1) $|D_{\leqslant \ell+1}(w)| \geqslant (\log n)^2$;

(E2) $D_{\leqslant \ell+1}(w)$ contains a node which lies on a cycle of length at most 2ℓ.

We will later condition on the event $E_w(\ell)$ holding, and therefore need to know that it is very likely.

Lemma 32. *For any function $\ell = o(\log\log n)$ and node $w \in \mathcal{V} \cup \mathcal{F}$,*

$$\mathbb{P}(E_w(\ell)) \geqslant 1 - \exp\left(-\Theta\left(\sqrt{\log n}\right)\right).$$

Furthermore, whp all but $o(n)$ nodes $w \in \mathcal{V} \cup \mathcal{F}$ satisfy $E_w(\ell)$.

The (standard) proof appears in Appendix C.

Now let us define

$$\tilde{d}_1(w) := \begin{cases} d_1^*(w) & \text{if } d_1^*(w) \neq 1 \\ 0 & \text{if } d_1^*(w) = 1. \end{cases}$$

In other words, $\tilde{d}_1(w)$ is identical to $d_1^*(w)$ except that values of 1 are replaced by 0 (similar to the $\widetilde{\text{Po}}$ and $\widetilde{\text{Bi}}$ distributions compared to the Po and Bi distributions). We can combine Lemmas 32 and 30 and Proposition 31 to obtain the following.

Corollary 33. *For any integer-valued function $\ell = \ell(n) \xrightarrow{n\to\infty} \infty$ which also satisfies $\ell = o(\log\log n)$ and any node w we have*

$$\mathbb{P}\left(d_{G_\ell}(w) \neq \tilde{d}_1(w)\right) = o(1).$$

Proof. By Lemma 30, the only cases in which $d_{G_\ell}(w)$ and $\tilde{d}_1(w)$ can differ are if $E_w(\ell)$ does not hold or if $d_{G_\ell}(w) = 1$. Thus by applying Lemma 32 and Proposition 31, we obtain

$$\begin{aligned}\mathbb{P}\left(d_{G_\ell}(w) \neq \tilde{d}_1(w)\right) &\leqslant \mathbb{P}\left(\bar{E}_w(\ell)\right) + \mathbb{P}\left(d_{G_\ell}(w) = 1\right) \\ &\leqslant \exp\left(-\Theta\left(\sqrt{\log n}\right)\right) + o(1) = o(1).\end{aligned}$$

□

We next describe the survival probabilities of internal (i.e. non-root) variable and factor nodes in each round of `CoreConstruct`. Recall that for any $i \in [\ell]$ the set $D^*_{\ell+1-i}(w)$ consists of nodes within $D_{\ell+1-i}(w)$ which survive the i-th round of `CoreConstruct`. We define the recursions

$$\begin{aligned}\rho_0 &= 1, \\ \hat{\rho}_t &= \mathbb{P}(\mathrm{Bi}(r-1, \rho_{t-1}) \geqslant 1), \qquad (13) \\ \rho_t &= \mathbb{P}(\mathrm{Po}(d\hat{\rho}_t) \geqslant 1). \qquad (14)\end{aligned}$$

Lemma 34. *Let $w \in \mathcal{V} \cup \mathcal{F}$ and ℓ be odd if $w \in \mathcal{V}$ and even if $w \in \mathcal{F}$. Let $t \in \mathbb{N}$ with $0 \leqslant t \leqslant \frac{\ell+1}{2}$ be given. Conditioned on the event $E_w(\ell)$:*

1. *For each $u \in D_{\ell+1-2t}(w)$ independently of each other we have*

$$\mathbb{P}[u \in D^*_{\ell+1-2t}(w)] = \rho_t + o(1);$$

2. *For each $a \in D_{\ell-2t}(w)$ independently of each other we have*

$$\mathbb{P}[a \in D^*_{\ell-2t}(w)] = \hat{\rho}_{t+1} + o(1).$$

In particular:

(i) If $w \in \mathcal{F}$ and $t_1 = \ell/2$, then for each $u \in D_1(w)$ independently of each other,

$$\mathbb{P}[u \in D^*_1(w)] = \rho_{t_1} + o(1);$$

(ii) If $w \in \mathcal{V}$ and $t_2 = (\ell+1)/2$, then for each $a \in D_1(w)$ independently of each other,

$$\mathbb{P}[a \in D^*_1(w)] = \hat{\rho}_{t_2} + o(1).$$

To prove this lemma, we will need the asymptotic degree distribution of a variable node in $N_{\leqslant \ell}(w)$, which is a standard result.

Proposition 35. *Let $w \in \mathcal{V} \cup \mathcal{F}$ and an integer $\ell = o(\log\log n)$ be given. Conditioned on the event $E_w(\ell)$, for each $u \in D_{\leqslant \ell}(w) \cap \mathcal{V}$ independently, the number of children of u in $N_{\leqslant \ell+1}(w)$ is asymptotically distributed as $\mathrm{Po}(d)$.*

We defer the proof to Appendix B.2. We can now prove Lemma 34.

Proof of Lemma 34. We will prove both statements (1) and (2) by a common induction on t. For $t = 0$ the statements are clear since we have $\rho_0 = 1$, which corresponds to the fact that nothing ever gets deleted from $D_{\ell+1}(w)$, while $\hat{\rho}_1 = \mathbb{P}(\mathrm{Bi}(r-1,1) \geqslant 1) = 1$, corresponding to the fact that internal factor nodes have $r - 1 \geqslant 1$ children in the input graph and therefore also no nodes of $D_\ell(w)$ will be deleted.

We assume both statements are true for $t-1$ and aim to prove that they also hold for t. We first consider $u \in D_{\ell+1-2t}$ and let X_u be the number of children of u in $D^*_{\ell+2-2t}$. Observe that the probability that u survives in `CoreConstruct` is simply $\mathbb{P}(X_u \geqslant 1)$.

By Proposition 35, the number of children of u is asymptotically $\mathrm{Po}(d)$, and by the induction hypothesis each child survives with probability $\hat{\rho}_t + o(1)$ independently of each other. Therefore the asymptotic survival probability of u is given by

$$\mathbb{P}\left(\mathrm{Bi}(\mathrm{Po}(d), \hat{\rho}_t + o(1)) \geqslant 1\right) = \mathbb{P}\left(\mathrm{Po}(d(\hat{\rho}_t + o(1))) \geqslant 1\right) = \rho_t + o(1),$$

by definition of ρ_t, as required for statement (1). Independence simply follows from the conditioning on $E_w(\ell)$, which in particular means that $N_{\leqslant \ell+1}(w)$ is a tree.

Similarly for $a \in D_{\ell-2t}(w)$ we define X_a to be the number of children of a which survive. Since a is an internal factor node it has precisely $r-1$ children, and by statement 1 which we have just proved, each child survives with probability $\rho_t + o(1)$ independently. Therefore the probability that a survives the `CoreConstruct` is

$$\mathbb{P}(\mathrm{Bi}(r-1, \rho_t + o(1)) \geqslant 1) = \hat{\rho}_{t+1} + o(1),$$

as required for statement (2). Again, independence simply follows from the conditioning on $E_w(\ell)$. □

A consequence of Lemma 34 is that, if ℓ is large, the distribution of the number of children of the root w which survive `CoreConstruct` is almost identical to the claimed distributions in Theorem 21. In order to quantify this, for two discrete random variables X, Y taking values in $\mathbb{N}$ we use the standard notion of the *total variation distance*, defined as

$$d_{\mathrm{TV}}(X,Y) \coloneqq \sum_{m \in \mathbb{N}} \Big| \mathbb{P}(X = m) - \mathbb{P}(Y = m) \Big|.$$

Corollary 36. *There exist $\varepsilon = o(1)$ and $\ell = \ell(\varepsilon)$ such that if we run* `CoreConstruct` *with input ℓ and root $w \in \mathcal{V} \cup \mathcal{F}$, then:*

1. *If $w \in \mathcal{V}$, then*

$$d_{\mathrm{TV}}\left(\tilde{d}_1(w), \widetilde{\mathrm{Po}}(d\hat{\rho}_*)\right) \leqslant d_{\mathrm{TV}}\left(d_1^*(w), \mathrm{Po}(d\hat{\rho}_*)\right) < \varepsilon;$$

2. *If $w \in \mathcal{F}$, then*

$$d_{\mathrm{TV}}\left(\tilde{d}_1(w), \widetilde{\mathrm{Bi}}(r, \rho_*)\right) \leqslant d_{\mathrm{TV}}\left(d_1^*(w), \mathrm{Bi}(r, \rho_*)\right) < \varepsilon.$$

Let us note that the second statement involves the $\mathrm{Bi}(r, \rho_*)$ distribution rather than $\mathrm{Bi}(r-1, \rho_*)$ since w is the root, and therefore all r of its neighbours are children rather than just $r-1$ children and one parent.

Proof. We will prove only the first statement; the proof of the second is very similar.

Note that the first inequality follows directly from the definitions of $d_1^*(w)$ and the $\widetilde{\mathrm{Po}}$ distributions. There are 2 reasons why the second inequality is not quite immediate from Lemma 34:

(R1) Lemma 34 has the conditioning on the event $E_w(\ell)$;

(R2) $\hat{\rho}_{(\ell+1)/2}$ in Lemma 34 has been replaced by $\hat{\rho}_*$.

To overcome **(R1)**, for ease of notation for each $j \in \mathbb{N}$ we set $q_j := \mathbb{P}\left(d_1^*(w) = j\right)$ and $q'_j := \mathbb{P}\left(d_1^*(w) = j | E_w(\ell)\right)$, and define

$$J^+ := \left\{j : q_j > q'_j\right\} \qquad \text{and} \qquad J^- := \left\{j : q_j < q'_j\right\}.$$

Since for any $w \in \mathcal{V} \cup \mathcal{F}$ we have

$$\sum_{j=0}^{\infty}(q_j - q'_j) = \sum_{j=0}^{\infty} q_j - \sum_{j=0}^{\infty} q'_j = 1 - 1 = 0,$$

we deduce that

$$\sum_{j=0}^{\infty}\left|q_j - q'_j\right| = \sum_{j\in J^+}\left(q_j - q'_j\right) - \sum_{j\in J^-}\left(q_j - q'_j\right) = 2\sum_{j\in J^+}\left(q_j - q'_j\right).$$

Therefore we have

$$\begin{aligned} d_{\mathrm{TV}}\left(d_1^*(w), d_1^*(w)|_{E_w(\ell)}\right) &= 2\sum_{j\in J^+}\left(q_j - q'_j\right) \\ &\leqslant 2\cdot\sum_{j\in J^+}\mathbb{P}\left(d_1^*(w) = j\right) - \mathbb{P}\left(\left(d_1^*(w) = j\right) \cap E_w(\ell)\right) \\ &\leqslant 2\cdot\mathbb{P}\left(\overline{E_w(\ell)}\right) \\ &\leqslant 2\exp\left(-\Theta\left(\sqrt{\log n}\right)\right) \\ &\leqslant \varepsilon/3, \end{aligned} \tag{15}$$

where we applied Lemma 32 in the penultimate line, and where the last line holds for ε tending to 0 sufficiently slowly.

To address **(R2)** we first claim that

$$\hat{\rho}_t \xrightarrow{t\to\infty} \hat{\rho}_* \tag{16}$$

holds. To see this, observe that

$$1-\rho_t \overset{(14)}{=} \exp(-d\hat{\rho}_t) \overset{(13)}{=} \exp\left(-d\left(1-(1-\rho_{t-1})^{r-1}\right)\right) \overset{(1)}{=} F(1-\rho_{t-1})$$

for all $t \in \mathbb{N}_{\geqslant 1}$. Since $1-\rho_0 = 0$ it follows by elementary analytic arguments that as t tends to infinity, $1-\rho_t$ converges to the smallest fixed point of the function F and thus ρ_t converges to the largest solution of the fixed-point equation (2), which we defined as ρ_* (see (1)). Then (16) follows immediately from (3), (13) by continuity.

As a consequence of (16), if $\ell = \ell(\varepsilon)$ is sufficiently large, then

$$d_{\mathrm{TV}}\left(\mathrm{Po}(d\hat{\rho}_{(\ell+1)/2}), \mathrm{Po}(d\hat{\rho}_*)\right) < \varepsilon/3. \tag{17}$$

To complete the proof, observe that Lemma 34 (ii) implies that

$$d_{\mathrm{TV}}\left(d_1^*(w)|_{E_w(\ell)}\,,\, \mathrm{Po}(d\hat{\rho}_{(\ell+1)/2})\right) = d_{\mathrm{TV}}\left(d_1^*(w)|_{E_w(\ell)}\,,\, \mathrm{Bi}\left(\mathrm{Po}(d), \hat{\rho}_{(\ell+1)/2}\right)\right) \leqslant \varepsilon/3,$$

and combining this with (15) and (17), we obtain

$$\begin{aligned} d_{\mathrm{TV}}\left(d_1^*(w), \mathrm{Po}(d\hat{\rho}_*)\right) &\leqslant d_{\mathrm{TV}}\left(d_1^*(w), d_1^*(w)|_{E_w(\ell)}\right) + d_{\mathrm{TV}}\left(d_1^*(w)|_{E_w(\ell)}, \mathrm{Po}(d\hat{\rho}_{(\ell+1)/2})\right) \\ &\qquad + d_{\mathrm{TV}}\left(\mathrm{Po}(d\hat{\rho}_{(\ell+1)/2}, \mathrm{Po}(d\hat{\rho}_*)\right) \\ &\leqslant \varepsilon \end{aligned}$$

as required. □

As a further consequence of Corollary 36, we can asymptotically determine the expected degree distribution in G_ℓ.

Corollary 37. *There exist $\varepsilon = o(1)$ and $\ell = \ell(\varepsilon)$ such that for all $j \in \mathbb{N}$,*

$$\mathbb{E}\left(\xi_j^{(\ell)}\right) = \mathbb{P}\left(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = j\right) \pm \varepsilon,$$

and

$$\mathbb{E}\left(\hat{\xi}_j^{(\ell)}\right) = \mathbb{P}\left(\widetilde{\mathrm{Bi}}(r, \rho_*) = j\right) \pm \varepsilon.$$

Proof. Observe that for any $j \in \mathbb{N}$ and any variable node w, by linearity of expectation and Corollary 33, for some $\varepsilon_1 = o(1)$ we have

$$\mathbb{E}\left(\xi_j^{(\ell)}\right) = \mathbb{E}\left(\frac{1}{n}\sum_{w\in\mathcal{V}} \mathbb{1}_{\{d_{G_\ell}(w)=j\}}\right) = \mathbb{P}(d_{G_\ell}(w) = j) = \mathbb{P}\left(\tilde{d}_1(w) = j\right) \pm \varepsilon_1$$

Furthermore, Corollary 36 implies that for some $\varepsilon_2 = o(1)$ we have

$$\mathbb{P}\left(\tilde{d}_1(w) = j\right) = \mathbb{P}\left(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = j\right) \pm \varepsilon_2.$$

Combining these two approximations and setting $\varepsilon = \varepsilon_1 + \varepsilon_2 = o(1)$ completes the proof of the first statement. The second statement is proven similarly. □

Now for $j, \ell \in \mathbb{N}$, let us define the events $B_j = B_j(\ell) := \left\{\left|\xi_j^{(\ell)} - \mathbb{E}\left(\xi_j^{(\ell)}\right)\right| \geqslant n^{-1/3}\right\}$ and $\hat{B}_j = \hat{B}_j(\ell) := \left\{\left|\hat{\xi}_j^{(\ell)} - \mathbb{E}\left(\hat{\xi}_j^{(\ell)}\right)\right| \geqslant n^{-1/3}\right\}$. We can apply Lemma 12 to prove the following.

Lemma 38. *Let $G = G^r(n,p)$ be a random factor graph and $\ell = o(\log\log n)$. Then*

$$\mathbb{P}\left(\bigcup_{j\in\mathbb{N}}\left(B_j \cup \hat{B}_j\right)\right) = o(1).$$

Proof. We will model the factor graph $G = G^r(n,p)$ as a subgraph of a larger factor graph $\tilde{G} = \tilde{G}^r(n,p)$, which is obtained from G by adding some additional isolated factor nodes, which represent the *non-edges* of $H^r(n,p)$, such that the total number of factor nodes is precisely $\binom{n}{r}$. Thus each factor node comes with a set of r associated variable nodes, these sets all being distinct, and a factor node is adjacent to its r associated variable nodes (and no others) whenever the corresponding edge is present in $H^r(n,p)$, and otherwise it is isolated. (Note, however, that these extra factor nodes are *not* considered for the calculation of $\hat{\xi}_j^{(\ell)}$.)

Now set $N := \binom{n}{r}$ and for some arbitrary order of the factor nodes, let X_k denote the indicator function of the event that the k-th factor node is not isolated, i.e. that the corresponding edge is present in $H^r(n,p)$. (This simply describes the standard edge-exposure process, but rephrased in the language of factor graphs.) Clearly there is a one-to-one correspondence between sequences in $\{0,1\}^N$ and possible instances of the graph $\tilde{G}$, and in what follows we will identify a sequence with its associated factor graph.

We let f describe the function on $\{0,1\}^N$ corresponding to $\xi_j^{(\ell)}$ or $\hat{\xi}_j^{(\ell)}$ as appropriate. We need to check that f satisfies $(*)$, the *typical Lipschitz condition* (to borrow the terminology of [38]) required in Lemma 12, for an appropriate choice of the parameters c, d and event Γ.

We define Γ to be the event that $\Delta(G) \leqslant \log n$ and that G contains $\Theta(n)$ factor nodes. We claim that $\mathbb{P}(\Gamma) \geqslant 1 - n^{-\omega(1)}$. This can be proved by a union bound on the two bad events. The probability that the number of factor nodes is *not* $\Theta(n)$ is exponentially small in n by a Chernoff bound. (This is similar to Proposition 14, but with a larger deviation and a correspondingly smaller error probability.) Meanwhile, the probability that a variable node has degree at least $\log n$ can be approximated by

$$\mathbb{P}\left(\mathrm{Bi}\left(\binom{n-1}{r-1}, p\right) \geqslant \log n\right) \leqslant \binom{\binom{n-1}{r-1}}{\log n} p^{\log n} = \left(\frac{\Theta(1)}{\log n}\right)^{\log n} = n^{-\omega(1)}.$$

It follows that the expected number of vertices of degree at least $\log n$ is at most $n \cdot n^{-\omega(1)} = n^{-\omega(1)}$, and Markov's inequality implies that the probability that there is at least one such vertex is at most $n^{-\omega(1)}$, as required.

Now if $x \in \Gamma$ and $\tilde{x} \in \{0,1\}^N$ is a sequence which differs from x in only one co-ordinate, this corresponds to adding or removing r edges incident to a single factor node in a factor graph of maximum degree at most $\log n$. Let R be the set of r variable nodes incident to these edges. In the peeling process up to step ℓ, this change can only affect those variable

nodes at distance at most ℓ from R, of which, due to the maximum degree condition, there are at most $r\sum_{i=0}^{\ell}(\log n)^i \leqslant (\log n)^{\ell+1}$, and therefore can change $\xi_j^{(\ell)}$ by at most $\frac{(\log n)^{\ell+1}}{n}$. When considering $\hat{\xi}_j^{(\ell)}$ the argument is similar, but we use the fact that since $x \in \Gamma$ there are $\Theta(n)$ factor nodes in the corresponding graphs $G, \tilde{G}$, and so we can change $\hat{\xi}_j^{(\ell)}$ by at most $\frac{(\log n)^{\ell+1}}{\Theta(n)} \leqslant \frac{(\log n)^{\ell+2}}{n}$. In either case we have $|f(x) - f(\tilde{x})| \leqslant \frac{(\log n)^{\ell+2}}{n} =: c$. For the worst case, we simply use the trivial upper bound $|f(x) - f(\tilde{x})| \leqslant 1 =: d$. Then the typical Lipschitz condition $(*)$ is indeed satisfied.

Now Lemma 12 tells us (using the simplification $c + e \leqslant 2c$) that

$$\begin{aligned}\mathbb{P}\Big(\Big|f(X) - \mathbb{E}(f(X))\Big| \geqslant t\Big) &\leqslant 2\exp\left(-\frac{t^2}{2\binom{n}{r}p \cdot 4c^2 + 4ct/3}\right) + n\binom{n}{r}n^{-\omega(1)} \\ &\leqslant 2\exp\left(-\frac{t^2}{\Theta\left(\frac{(\log n)^{2\ell+4}}{n}\right) + \Theta\left(\frac{t(\log n)^{\ell+2}}{n}\right)}\right) + o(1).\end{aligned}$$

Setting $t = n^{-1/3}$, this last expression is certainly $o(1)$, so we deduce that whp

$$\Big|f(X) - \mathbb{E}(f(X))\Big| \leqslant n^{-1/3},$$

as required. □

We are now able to give the proof of Lemma 27.

Proof of Lemma 27. By Corollary 37 (with $\varepsilon/2$ in place of ε), if $\ell = \ell(\varepsilon)$ is sufficiently large we have

$$\mathbb{E}\left(\xi_j^{(\ell)}\right) = \mathbb{P}\left(\widetilde{\mathrm{Po}}\left(d\hat{\rho}_*\right) = j\right) \pm \varepsilon/2$$

for any $j \in \mathbb{N}$. Furthermore, by Lemma 38, we have that whp for all $j \in \mathbb{N}$

$$\xi_j^{(\ell)} = \mathbb{E}\left(\xi_j^{(\ell)}\right) + o(1) = \mathbb{E}\left(\xi_j^{(\ell)}\right) \pm \varepsilon/2$$

for ε tending to 0 sufficiently slowly, and combining these two facts proves the lemma for variable nodes. The proof for factor nodes is essentially identical. □

7 Subcriticality: Proof of Lemma 28

Our goal in this section is to show that after some large number ℓ rounds of the peeling process on $G^r(n,p)$ have been completed, whp very few nodes will be disabled in subsequent rounds (at most εn for some $\varepsilon = \varepsilon(n) = o(1)$), thus proving Lemma 28.

Let us fix ℓ, ε_1 as in Lemma 27, and for the rest of this section we will assume that the high probability events of Lemma 27 and Proposition 14 both hold.

To help intuitively describe the argument, let us suppose for simplicity that in round ℓ exactly *one* node x_0 is disabled and we consider the future effects of such a disabling. Since x_0 was disabled, it had degree at most one, and therefore there is at most one neighbour

x_1 whose degree is decreased as a result. If x_1 originally had degree two, it now has degree one and will therefore be disabled in round $\ell + 1$. Continuing in this way, we observe that we will never be disabling more than one node in any subsequent round. Furthermore, if we reach a node x_i whose original degree was not exactly two, the peeling process stops (either without disabling this node, or once it has been disabled and no further nodes' degrees are decreased). Heuristically, it will not take long before we reach a node whose original degree was not exactly two—this is because Lemma 27 implies in particular that a constant proportion of the nodes have degree at least three.

Of course, in reality there may be more than one node disabled in round ℓ. This slightly complicates matters because some node may receive the knock-on effects of more than one disabling in round ℓ, and therefore have its degree decreased by more than one. However, this will turn out to be no more than a technical nuisance.

We first need a result which states that almost any graph with a fixed (reasonable) degree sequence is approximately equally likely to be G_ℓ, the graph obtained from $G = G^r(n,p)$ after ℓ rounds of the peeling process. To introduce this result, we need some definitions.

Definition 39. An *r-duplicate* in a factor graph consists of two factor nodes and r variable nodes which together form a copy of $K_{2,r}$.

Observe that an r-duplicate would correspond to a double-edge in an r-uniform hypergraph, which in our model cannot occur since the hypergraph must be simple. Therefore our factor graphs may not contain any r-duplicates. On the other hand, a "loop", in the sense of an edge which contains the same vertex more than once, must involve a double-edge in the corresponding factor graph. This motivates the following definition, which (roughly) describes when a factor graph corresponds to a simple hypergraph.

Definition 40. We say that a factor graph is *r-plain* if:

1. it contains no double-edge, i.e. two edges between the same variable node and factor node;

2. it contains no r-duplicates.

Claim 41. *Suppose H_1, H_2 are two r-plain factor graphs with common variable node set $\mathcal{V} = \mathcal{V}(H_1) = \mathcal{V}(H_2) = [n]$ and with factor node sets $\mathcal{F}_1 = \mathcal{F}(H_1)$ and $\mathcal{F}_2 = \mathcal{F}(H_2)$. Suppose further that there is a bijection $\phi : \mathcal{V} \cup \mathcal{F}(H_1) \to \mathcal{V} \cup \mathcal{F}(H_2)$ such that*

- *$\phi(\mathcal{V}) = \mathcal{V}$;*
- *$d_{H_2}(\phi(v)) = d_{H_1}(v)$ for all $v \in \mathcal{V} \cup \mathcal{F}(H_1)$.*

Let $G = G^r(n,p)$. Then

$$\mathbb{P}\left(G_\ell = H_1\right) = \mathbb{P}\left(G_\ell = H_2\right).$$

This claim is very similar to standard results for simple graphs or hypergraphs (see e.g. [31]) and we defer the proof to Appendix D. However, we note that in our setting there is one subtle technical difficulty to overcome which does not appear in many other cases,

namely that given a factor graph G such that $G_\ell = H_1$, if we transform G by changing H_1 to H_2 but otherwise leaving G unchanged, we need to show that the resulting graph is indeed the factor graph of an r-uniform hypergraph, and in particular is r-plain.

Claim 41 tells us that the factor graph G_ℓ after ℓ rounds of the peeling process is uniformly random conditioned on its degree sequence and being r-plain. Since Lemma 27 tells us the degree sequence quite precisely, this is very helpful. We can change our point of view by saying that we first reveal the degree sequence of G_ℓ without revealing any of its edges, and subsequently we reveal edges only as required. More precisely, we consider the *configuration model*, in which each node is given half-edges based on its degree, and we generate a uniformly random perfect matching between the two classes of half-edges (at variable and factor nodes) conditioned on the resulting factor graph being r-plain. We need to know that this conditioning is not too restrictive, i.e. that the probability that the resulting factor graph is r-plain is not too small. This will be stated in Proposition 44, for which we first need some preliminaries. We begin by observing that $G^r(n,p)$ does not have too many nodes of high degree.

Definition 42. Given a function $\omega = \omega(n) \to \infty$, we say that a factor graph H has property $\widetilde{\mathcal{D}} = \widetilde{\mathcal{D}}(\omega, n)$ if

$$\sum_{v \in \mathcal{V}(H):\ d(v) > \omega} d(v)^2 = o(n).$$

Furthermore given $\varepsilon > 0$ we say that H has property $\mathcal{D} = \mathcal{D}(\varepsilon, \omega, n)$ if it satisfies property $\widetilde{\mathcal{D}}(\omega, n)$ and also satisfies the conclusion of Lemma 27 (with this ε).

Claim 43. *For any $\omega \xrightarrow{n\to\infty} \infty$, with high probability $G^r(n,p)$ has property $\widetilde{\mathcal{D}}(\omega, n)$.*

Proof. For $k \in \mathbb{N}$, let us define X_k to be the number of variable nodes of degree k and $X_{\geqslant k} \coloneqq \sum_{j \in \mathbb{N}_{\geqslant k}} X_j$. Observe that the expected degree of a vertex is $\binom{n-1}{r-1}p = (1+o(1))d$. It is a standard fact that the degrees of vertices are approximately distributed as independent $\mathrm{Po}(d)$ variables. More formally (though much weaker), it is an easy exercise in the second-moment method to prove that whp, for any $k \in \mathbb{N}$ we have $X_{\geqslant k} \leqslant n \cdot \mathbb{P}(\mathrm{Po}(2d) \geqslant k)$ (we omit the details). We therefore have

$$\sum_{v \in \mathcal{V}(H): d(v) \geqslant \omega} d(v)^2 = \sum_{k \geqslant \omega} k^2 X_k \leqslant n \sum_{k \geqslant \omega} k^2 \frac{e^{-2d}(2d)^k}{k!} = n \cdot (1 + o(1))\omega^2 \frac{e^{-2d}(2d)^\omega}{\omega!} = o(n),$$

as required. □

Note that if $\widetilde{\mathcal{D}}$ holds in a factor graph G, then it also holds in any subgraph of G, and in particular in G_ℓ, the factor graph obtained after ℓ steps of the peeling process. Together with Lemma 27, this shows that, with ℓ and ε as in given in that lemma and any $\omega \to \infty$, setting $G = G^r(n,p)$, with high probability G_ℓ satisfies $\mathcal{D}(\varepsilon, \omega, n)$.

Let us observe further that $\mathcal{D}$ is a property that depends only on the degree sequences of variable and factor nodes of the graph, and therefore with a slight abuse of terminology we may also say that it holds in a factor graph with half-edges, where we have not yet determined which half-edges will be matched together.

Proposition 44. *Let $\varepsilon = o(1)$ and suppose that n variable nodes and $m = (1+o(1))\frac{dn}{r}$ factor nodes are given half-edges in such a way that property $\mathcal{D}(\varepsilon, \varepsilon^{-1/4}, n)$ holds. Suppose also that the total numbers of half-edges at factor nodes and at variable nodes are equal, and that we construct a uniformly random perfect matching between these two sets of half-edges. Then there exists a constant $c_0 > 0$ (independent of ε, n) such that for sufficiently large n the probability that the resulting factor graph is r-plain is at least c_0.*

The proof of Proposition 44 is a standard exercise in applying the method of moments to determine the asymptotic distribution of the number of double edges and r-duplicates – we omit the details. The proposition states in particular that we may condition on the resulting graph being r-plain without skewing the distribution of the matching too much. More precisely, any statements that are true with high probability for the uniformly random perfect matching are also true with high probability under the condition that the resulting graph is simple. Therefore in what follows, for simplicity we will suppress this conditioning.

Definition 45 (Change process)**.** We will track the changes that the peeling process makes after reaching round ℓ by revealing information a little at a time as follows.

- Reveal the degrees of all nodes.
- While there are still nodes of degree one, pick one such node x_0.
  - Reveal its neighbour x_1, delete the edge x_0x_1 and update the degrees of x_0, x_1.
  - If x_1 now has degree one, continue from x_1; otherwise find a new x_0 (if there is one).

In other words, we track the changes in a depth-first search manner (rather than the breadth-first view of considering rounds of the peeling process). We call this the *change process*.

Observe that we only reveal edges one at a time (just before deleting them). The following claim is simple given Lemma 27, but is the essential heart of our proof. Recall that $\varepsilon_2 := \sqrt{\varepsilon_1}$ as defined in Lemma 28.

Claim 46. *Let G' be any graph obtained from G_ℓ by deleting at most $\varepsilon_2 n$ edges. Then when revealing the second endpoint of any half-edge, the probability of revealing a node of degree at least three is at least*

$$\min\left\{\frac{(d\hat{\rho}_*)^2\exp(-d\hat{\rho}_*)}{2}, \frac{(r-1)(r-2)\rho_*^2(1-\rho_*)^{r-3}}{2}\right\} - 20\varepsilon_2.$$

In particular there exists a constant $c = c(r,d) > 0$ such that this probability is at least c.

We defer the proof to Appendix B.3.
This claim tells us that, provided we have not deleted too many edges so far, there is a reasonable probability of revealing a node of degree at least three, which blocks the continued propagation of any deletions.

Let $c = c(r,d)$ be as in Claim 46 and let us set $\delta_1 := \varepsilon_1^{3/4}$. We now define an abstract branching process which will provide an upper coupling on the change process starting from G_ℓ.

Definition 47. Let $\mathcal{T}$ be a branching process which begins with $\delta_1 n$ vertices in generation 0, and in which each vertex independently has a child with probability $1-c$, and otherwise has no children.

Proposition 48. *The process $\mathcal{T}$ can be coupled with the change process in such a way that, if both processes are run until one of the stopping conditions*

- *$\mathcal{T}$ has reached size at least $\varepsilon_2 n$;*
- *$\mathcal{T}$ has died out,*

is satisfied, then $\mathcal{T}$ forms an upper coupling on the change process.

Proof. We first need to show that whp we make at most $\delta_1 n$ changes in round $\ell+1$ of the peeling process. Since we have assumed that the high probability statement of Lemma 27 holds, we have $\xi_j^{(\ell)} = \mathbb{P}(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = j) \pm \varepsilon_1$ and $\hat{\xi}_j^{(\ell)} = \mathbb{P}(\widetilde{\mathrm{Bi}}(r,\rho_*) = j) \pm \varepsilon_1$. By the definition of the peeling process (Definition 26) the only change we make when moving from G_ℓ to $G_{\ell+1}$ is that we disable all nodes of degree one. The proportion of such variable and factor nodes in G_ℓ is $\xi_1^{(\ell)}$ and $\hat{\xi}_1^{(\ell)}$ respectively. Recalling that $\mathbb{P}(\widetilde{\mathrm{Po}}(d\hat{\rho}_*) = 1) = \mathbb{P}(\widetilde{\mathrm{Bi}}(r,\rho_*) = 1) = 0$, this immediately implies that at most $\varepsilon_1(m+n) \leqslant \delta_1 n$ nodes are disabled in round $\ell+1$ of the peeling process, and these disablings represent the first nodes of the change process.

Now the proposition follows directly from the observation that in the change process, a node only has at most one incident edge deleted (if it has degree one), and therefore at most one neighbour is revealed, along with Claim 46, which implies that the probability of not causing any further changes is at least c. The first stopping condition ensures that we have deleted at most $\varepsilon_2 n$ edges, and therefore the assumptions of Claim 46 are indeed satisfied. □

In view of Proposition 48, it is enough to prove that whp the branching process $\mathcal{T}$ dies out (i.e. fulfills the second stopping condition) before reaching size $\varepsilon_2 n$.

Proposition 49. *Whp $\mathcal{T}$ contains at most $\varepsilon_2 n$ vertices.*

Proof. In order to reach size $\varepsilon_2 n$, the first (at most) $\varepsilon_2 n$ vertices of the process would have to have a total of at least $(\varepsilon_2 - \delta_1)n$ children, which, by Lemma 11, occurs with probability at most

$$\begin{aligned}\mathbb{P}\left(\mathrm{Bi}(\varepsilon_2 n, 1-c) \geqslant (\varepsilon_2-\delta_1)n\right) &\leqslant 2\cdot\exp\left(-\frac{(c\varepsilon_2-\delta_1)^2 n^2}{2\varepsilon_2 n + 2(c\varepsilon_2-\delta_1)n/3}\right)\\ &\leqslant 2\cdot\exp\left(-\frac{(c\varepsilon_2/2)^2 n}{3\varepsilon_2}\right) = o(1)\end{aligned}$$

(since $\varepsilon_2 n = \sqrt{\varepsilon_1} n \to \infty$) as required. □

We can now complete the proof of Lemma 28.

Proof of Lemma 28. Proposition 48 implies that the probability that at least $\varepsilon_2 n$ further nodes are disabled after round ℓ in the peeling process is at most the probability that $\mathcal{T}$ reaches size $\varepsilon_2 n$. But this event has probability $o(1)$ by Proposition 49.

This proves that whp at most $\varepsilon_2 n$ further nodes are disabled after round ℓ in the peeling process. To show the remaining two statements, observe that since disabling a node deletes at most one edge, and therefore changes the degree of at most one variable node and at most one factor node, at most $\varepsilon_2 n$ nodes of each type will have their degree changed. It follows immediately that for any $j \in \mathbb{N}$ we have $\xi_j = \xi_j^{(\ell)} \pm \varepsilon_2$, while similarly

$$\hat{\xi}_j = \hat{\xi}_j^{(\ell)} \pm \frac{\varepsilon_2 n}{m} = \hat{\xi}_j^{(\ell)} \pm \frac{2\varepsilon_2 r}{d},$$

where we have used the fact that whp $\frac{n}{m} \leqslant \frac{2r}{d}$ by Proposition 14. □

8 Concluding remarks

8.1 Upper bound on L_P, L_C

Theorem 6 and Corollary 7 state that whp L_P and L_C, the length of the longest loose path and longest loose cycle respectively in $H^r(n,p)$, satisfy $L_P, L_C \leqslant (\min\{\beta,\gamma\} + o(1)) \cdot n$, but which of β, γ is smaller? Recall that both β and γ are functions of d and r, with $\beta = \gamma = 0$ for $d < d^*$. Numerical approximations and plots with Mathematica suggest that for $r = 2, 3$, we have $\beta \leqslant \gamma$ for all d, but that for $r \geqslant 4$, for values of d not too much larger than d^* we have $\gamma < \beta$, and there is a crossing point after which (for larger d) we have $\beta < \gamma$. It would be interesting to investigate this behaviour more closely to determine whether this is indeed true, and to determine the crossing point precisely as a function of r.

8.2 High-order cores

The methods used in this paper are amenable to more general definitions of cores in hypergraphs. More precisely, vertex degrees are far from the only type of degrees that have been extensively studied in hypergraphs—one can consider the degrees of j-sets of vertices for each $1 \leqslant j \leqslant r-1$, and for each j there is a natural associated definition of a core. So far only the case $j = 1$ has been studied, but it would also be interesting to consider "high-order cores", i.e. $j \geqslant 2$.

9 Acknowledgement

We would like to thank Matthew Kwan for bringing to our attention the variants of the Azuma-Hoeffding inequality presented in [38].

References

[1] Dimitris Achlioptas and Michael Molloy. The solution space geometry of random linear equations. *Random Structures Algorithms*, 46(2):197–231, 2015.

[2] Miklós Ajtai, János Komlós, and Endre Szemerédi. The longest path in a random graph. *Combinatorica*, 1(1):1–12, 1981.

[3] Michael Anastos. Talk at the 20th International Conference on Random Structures and Algorithms. 1-5 August 2022.

[4] Michael Behrisch, Amin Coja-Oghlan, and Mihyun Kang. The order of the giant component of random hypergraphs. *Random Structures Algorithms*, 36(2):149–184, 2010.

[5] Michael Behrisch, Amin Coja-Oghlan, and Mihyun Kang. Local limit theorems for the giant component of random hypergraphs. *Combin. Probab. Comput.*, 23(3):331–366, 2014.

[6] Béla Bollobás and Oliver Riordan. Asymptotic normality of the size of the giant component in a random hypergraph. *Random Structures Algorithms*, 41(4):441–450, 2012.

[7] Béla Bollobás and Oliver Riordan. Exploring hypergraphs with martingales. *Random Structures Algorithms*, 50(3):325–352, 2017.

[8] Julie Cain and Nicholas Wormald. Encores on cores. *Electron. J. Combin.*, 13(1):Research Paper 81, 13, 2006.

[9] Oliver Cooley, Wenjie Fang, Nicola Del Giudice, and Mihyun Kang. Subcritical random hypergraphs, high-order components, and hypertrees. *SIAM J. Discrete Math.*, 34(4):2033–2062, 2020.

[10] Oliver Cooley, Frederik Garbe, Eng Keat Hng, Mihyun Kang, Nicolás Sanhueza-Matamala, and Julian Zalla. Longest paths in random hypergraphs. *SIAM J. Discrete Math.*, 35(4):2430–2458, 2021.

[11] Oliver Cooley, Mihyun Kang, and Christoph Koch. The size of the giant high-order component in random hypergraphs. *Random Structures Algorithms*, 53(2):238–288, 2018.

[12] Oliver Cooley, Mihyun Kang, and Christoph Koch. The size of the giant component in random hypergraphs: a short proof. *Electron. J. Combin.*, 26(3):Paper No. 3.6, 17, 2019.

[13] Oliver Cooley, Mihyun Kang, and Yury Person. Largest components in random hypergraphs. *Combin. Probab. Comput.*, 27(5):741–762, 2018.

[14] Colin Cooper. The cores of random hypergraphs with a given degree sequence. *Random Structures Algorithms*, 25(4):353–375, 2004.

[15] Amir Dembo and Andrea Montanari. Finite size scaling for the core of large random hypergraphs. *Ann. Appl. Probab.*, 18(5):1993–2040, 2008.

[16] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[17] Wenceslas Fernandez de la Vega. Long paths in random graphs. *Studia Sci. Math. Hungar.*, 14(4):335–340, 1979.

[18] Pu Gao. The stripping process can be slow: Part II. *SIAM J. Discrete Math.*, 32(2):1159–1188, 2018.

[19] Pu Gao and Michael Molloy. The stripping process can be slow: Part I. *Random Structures Algorithms*, 53(1):76–139, 2018.

[20] Svante Janson and Malwina J. Luczak. A simple solution to the k-core problem. *Random Structures Algorithms*, 30(1-2):50–62, 2007.

[21] Svante Janson and Malwina J. Luczak. Asymptotic normality of the k-core in random graphs. *Ann. Appl. Probab.*, 18(3):1085–1137, 2008.

[22] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[23] Jiayang Jiang, Michael Mitzenmacher, and Justin Thaler. Parallel peeling algorithms. In *Proceedings of the 26th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA'14)*, pages 319–330, 2014.

[24] Michał Karoński and Tomasz Łuczak. The phase transition in a random hypergraph. In *Probabilistic methods in combinatorics and combinatorial optimization*, volume 142, pages 125–135. 2002.

[25] Graeme Kemkes and Nicholas Wormald. An improved upper bound on the length of the longest cycle of a supercritical random graph. *SIAM J. Discrete Math.*, 27(1):342–362, 2013.

[26] Jeong Han Kim. Poisson cloning model for random graphs. In *International Congress of Mathematicians. Vol. III*, pages 873–897. Eur. Math. Soc., Zürich, 2006.

[27] Tomasz Łuczak. Cycles in a random graph near the critical point. *Random Structures Algorithms*, 2(4):421–439, 1991.

[28] Tomasz Łuczak. Size and connectivity of the k-core of a random graph. *Discrete Math.*, 91(1):61–68, 1991.

[29] Tomasz Łuczak. Sparse random graphs with a given degree sequence. In *Random graphs, Vol. 2 (Poznań, 1989)*, Wiley-Intersci. Publ., pages 165–182. Wiley, New York, 1992.

[30] Marc Mézard and Andrea Montanari. *Information, physics, and computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009.

[31] Michael Molloy. Cores in random hypergraphs and Boolean formulas. *Random Structures Algorithms*, 27(1):124–135, 2005.

[32] Boris Pittel, Joel Spencer, and Nicholas Wormald. Sudden emergence of a giant k-core in a random graph. *J. Combin. Theory Ser. B*, 67(1):111–151, 1996.

[33] Daniel Poole. On the strength of connectedness of a random hypergraph. *Electron. J. Combin.*, 22(1):Paper 1.69, 16, 2015.

[34] Oliver Riordan. The k-core and branching processes. *Combin. Probab. Comput.*, 17(1):111–136, 2008.

[35] Cristiane M. Sato. On the robustness of random k-cores. *European J. Combin.*, 41:163–182, 2014.

[36] Jeanette Schmidt-Pruzan and Eli Shamir. Component structure in the evolution of random hypergraphs. *Combinatorica*, 5(1):81–94, 1985.

[37] Kathrin Skubch. The core in random hypergraphs and local weak convergence. arXiv:1511.02048, [v2], 2015.

[38] Lutz Warnke. On the method of typical bounded differences. *Combin. Probab. Comput.*, 25(2):269–299, 2016.

A Analysis of fixed-point equation

Recall the fixed-point equation (2)

$$1 - \rho = F(1 - \rho)$$

with largest solution ρ_*. Note that this equation has no solutions for $\rho \geqslant 1$ and that 0 is a solution, hence $0 \leqslant \rho_* < 1$. Our goal in this section is to prove Claim 10 and Lemma 24. It will be more convenient to use a transformed equation: by substituting $x = 1 - \rho$ and taking logarithms on both sides (which is permissible since both sides are positive in any solution) we get the equivalent equation $\log x = -d(1 - x^{r-1})$, or equivalently

$$f(x) \coloneqq \log x + d(1 - x^{r-1}) = 0, \tag{18}$$

for $0 < x \leqslant 1$. Furthermore we define τ_* to be the smallest solution of (18) and note that $\tau_* = 1 - \rho_*$ holds. We now restate Claim 10 in this new setting.

Claim 50.

(F1)' *If $d < d^*$, then $\tau_* = 1$.*

(F2)' *If $d > d^*$, then $f(1) = 0$ and there is a unique solution to (18) in $(0, 1)$.*

Proof of Claim 10. It is clear that Claim 10 follows directly from Claim 50 since $\rho_* = 1 - \tau_*$. □

To prove Claim 50, we define

$$f_n(x) \coloneqq \log x + d\left(1 + \frac{1}{\omega}\right)(1 - x^{r-1}),$$

where $\omega = \omega(n)$ is a function with $\omega \xrightarrow{n\to\infty} \infty$. Observe that $f(x) = \lim_{n\to\infty} f_n(x)$ for each x. We would like to treat f_n and f simultaneously, therefore with a slight abuse of notation we define $f_\infty = f$, and set $1/\omega = 0$ for $n = \infty$. Furthermore, whenever we use statements such as "for n sufficiently large", this also includes $n = \infty$.

Now observe that

$$f_n'(x) = \frac{1}{x} - d\left(1 + \frac{1}{\omega}\right)(r-1)x^{r-2},$$
$$f_n''(x) = -\frac{1}{x^2} - d\left(1 + \frac{1}{\omega}\right)(r-1)(r-2)x^{r-3}.$$

The following fact collects properties of f_n which are trivial to check and that will be used later in this section.

Fact 51. *For sufficiently large n we have*

(P1) $f_n(1) = 0$.

(P2) $f_n(x) \xrightarrow{x\to 0} -\infty$.

(P3) *If $d < d^*$, then $f_n'(x) > 0$ for all $0 < x < 1$.*

(P4) *If $d > d^*$, then $f_n'(1) < 0$.*

(P5) $f_n''(x) < 0$ *for all* $0 < x \leqslant 1$.

Proof of Claim 50. First we show **(F1)'**. Let $d < d^*$. By **(P1)** and **(P3)** we have $f(x) < 0$ for all $0 < x < 1$ and hence that $x = 1$ is the unique solution of $f(x) = 0$, i.e. $\tau_* = 1$.

To see why **(F2)'** holds we observe that **(P1)** and **(P4)** imply that there exists x_0 with $0 < x_0 < 1$ such that $f(x_0) > 0$. Since **(P2)** holds by the intermediate value theorem there is y_0 with $0 < y_0 < x_0$ such that $f(y_0) = 0$, which implies that $\tau_* < 1$.

To prove uniqueness, suppose that there are two solutions x_1, x_2 of (18) with $0 < x_1 < x_2 < 1$. Since 1 is also a solution, by the mean value theorem there exist y_1, y_2 with $x_1 < y_1 < x_2 < y_2 < 1$ such that $f'(y_1) = f'(y_2) = 0$. Then by the mean value theorem applied to f', there exists a z_1 with $y_1 < z_1 < y_2$ such that $f''(z_1) = 0$. However, this contradicts **(P5)**. □

We need one last fact before we can give the proof of Lemma 24.

Claim 52. *If $d > d^*$, then there exists $\delta > 0$ such that for all sufficiently large n and all $0 < x \leqslant \tau_*$, we have*

$$f_n'(x) > \delta. \tag{19}$$

Proof. Observe that $f'(\tau_*) > 0$, which follows from the fact that $\tau_* < 1$, the fact that $f(1) = 0$ and **(P5)**. Now $f_n'(\tau_*) \xrightarrow{n\to\infty} f'(\tau_*)$, and therefore there exists $\delta > 0$ and $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $f_n'(\tau_*) > \delta$. Together with **(P5)**, the claim follows since $f_n'(x) > f_n'(\tau_*) > \delta$ holds for all $0 < x < \tau_*$ and $n \geqslant N$. □

With these considerations we are able to give the proof of Lemma 24.

Proof of Lemma 24. Let Q_n denote the set of zeros of $f_n(x)$ and let $q_n := \tau_*\left(r, d\left(1 + \frac{1}{\omega}\right)\right)$. Observe that $q_n = \min_n Q_n$ and let $q_* := \lim_{n\to\infty} q_n$. Our goal is to show that

$$q_* = \tau_*(r, d) =: \tau_*,$$

or in other words,

$$\rho_*\left(r, d\left(1 + \frac{1}{\omega}\right)\right) = \rho_* + o(1). \tag{20}$$

For all $n \in \mathbb{N}$ we have $q_n \leqslant 1$ since $x = 1$ is a solution to $f_n(x) = 0$, and $q_n \geqslant 0$ by **(P2)**. By Claim 50 there exists some $N \in \mathbb{N}$ such that for all $n \geqslant N$ the function f_n has exactly one solution ν_n, and since $f_n(\tau_*) > 0$ we have that $\nu_n < \tau_*$. By the mean value theorem there exists y_n with $\nu_n < y_n < \tau_*$ such that

$$f_n'(y_n) = \frac{f_n(\tau_*) - f_n(\nu_n)}{\tau_* - \nu_n}.$$

By (19) we deduce that

$$\tau_* - \nu_n = \frac{f_n(\tau_*) - f_n(\nu_n)}{f_n'(y_n)} \overset{(19)}{\leqslant} \frac{f_n(\tau_*)}{\delta}$$

and hence $\nu_n \xrightarrow{n\to\infty} \tau_*$ (since $f_n(\tau_*) \to 0$).

We have now proved that (20) holds, and (3) shows that then we also have

$$\hat{\rho}_*\left(r, d\left(1 + \frac{1}{\omega}\right)\right) = \hat{\rho}_*(r, d) + o(1).$$

Furthermore β and γ as defined in (5) and (6) are continuous functions in ρ_* and $\hat{\rho}_*$, respectively, both of which are themselves continuous functions in d and the statement of Lemma 24 follows. □

B Probabilistic Lemmas

In this appendix we include various proofs of probabilistic results which were omitted from the main text. These proofs are all standard applications of common techniques.

B.1 Sprinkling

Proof of Lemma 22. First let ω' be a function which tends to infinity arbitrarily slowly, and in particular such that $\frac{p_1 n^r}{(\omega')^r \omega} \to \infty$. Observe that we may assume that $L_P^{(1)} \geqslant n/\omega'$, since otherwise the trivial bound $L_C^{(2)} \geqslant 0$ is sufficient.

Since H_1, H_2 are coupled such that $H_1 \subset H_2$, we have $H_2 \sim H_1 \cup H_0$ for a random hypergraph $H_0 \sim H^r(n, p_0)$ which is independent of H_1 and where $p_0 = \frac{p_2 - p_1}{1 - p_1} \geqslant p_1/\omega$.

Let P_0' be a longest loose path in H_1. Let V_1 be the set of the first $\delta n/(4\omega')$ vertices of P_0'. Recall that we omit floors and ceilings, and in particular we assume that $\delta n/(4(r-1)\omega') \in \mathbb{N}$. Let I_1 be defined similarly for the following $\delta n/(4\omega')$ vertices and I_2 for the last $\delta n/(4\omega')$ vertices of P_0'. Let P_0 be the loose path which results after deleting V_1, I_1, I_2 and all incident edges. Denote by V_0 the set of vertices contained in P_0.

Now let $\mathcal{A}$ be the set of r-sets such that one vertex lies in I_1, one vertex lies in I_2 and $r-2$ vertices lie in V_1. If some $x \in \mathcal{A}$ forms an edge in H_0, then we would obtain a loose cycle in H_2 containing P_0 and thus of length at least $L_P - 3n/(4(r-1)\omega') = L_P - o(n)$, as required. We have

$$|\mathcal{A}| = |I_1||I_2|\binom{|V_1|}{r-2} = \left(\frac{n}{\omega'}\right)^2 \binom{n/(4\omega')}{r-2} = \Theta\left(\left(\frac{n}{\omega'}\right)^r\right).$$

Therefore the probability that no $x \in \mathcal{A}$ is an edge in $H_0 \sim H^r(n, p_0)$ is at most

$$(1-p_0)^{\Theta((n/\omega')^r)} \leqslant \exp\left(-p_0\Theta((n/\omega')^r)\right) \leqslant \exp\left(-\frac{p_1}{\omega}\Theta\left(\frac{n^r}{(\omega')^r}\right)\right) = o(1),$$

and therefore with high probability there is at least one such edge, as required. □

B.2 Offspring distribution

Proof of Proposition 35. We will consider an auxiliary breadth-first search algorithm starting from w, which goes through all $\binom{n-1}{r-1}$ many r-sets of variable nodes containing the currently active node x and queries an r-set R to determine whether it is an edge of $H^r(n,p)$ (i.e. if there exists a factor node whose neighbourhood is this r-set) if x is the only node of R which lies in the current tree, and otherwise the query is skipped. Initially the current tree consists simply of w, and the tree is updated if a query turns out to be an edge—in this case the corresponding factor node and all its neighbours are added to the current tree and we proceed. The event $E_w(\ell)$ implies that $|D_{\leqslant \ell+1}(w)| \leqslant (\log n)^2$ and thus the number q_x of queries made from any node x satisfies $\binom{n-(\log n)^2}{r-1} \leqslant q_x \leqslant \binom{n}{r-1}$. The number of edges found is asymptotically distributed as $\mathrm{Bi}\,(q_x, p)$—this is not exact since we still have the conditioning on the event $E_w(\ell)$, but since by Lemma 32 this is a very likely event, it does not affect the distribution significantly. Now we can couple the number Z_x of edges discovered from x from below and above by random variables with distributions $\mathrm{Bi}\left(\binom{n-(\log n)^2}{r-1}, p\right)$ and $\mathrm{Bi}\left(\binom{n}{r-1}, p\right)$, respectively. Since each of these random variables tends asymptotically to the $\mathrm{Po}(\binom{n}{r-1}p)$ distribution, which in turn converges to the $\mathrm{Po}(d)$ distribution, so does Z_x.

Finally, to prove independence observe that the upper and lower couplings in the previous step were independent of the number of edges found from any previous vertex (the lower coupling only used the conditioning on $E_w(\ell)$). Furthermore no r-set is queried more than once, and therefore the upper and lower couplings can be considered independent for each vertex. Thus the Z_x are also asymptotically independent of each other. □

B.3 Proof of Lemma 46

Proof. Defining $\xi_j(G')$ to be the proportion of variable nodes of degree j in G' for each $j \in \mathbb{N}$, observe that when revealing the neighbour of a factor node, the probability of

revealing a variable node of degree at least three is

$$\begin{aligned}\frac{\sum_{j\geqslant 3} j\xi_j(G')}{\sum_{j\in\mathbb{N}} j\xi_j(G')} &\geqslant \frac{3\xi_3(G')}{\sum_{j=1}^{3} j\xi_j(G')}\\ &\geqslant \frac{3(\xi_3^{(\ell)}-2\varepsilon_2)}{\sum_{j=1}^{3} j(\xi_j^{(\ell)}+2\varepsilon_2)}\\ &\overset{\text{L.27}}{\geqslant} \frac{3\mathbb{P}\left(\widetilde{\mathrm{Po}}(d\hat{\rho}_*)=3\right)-3\varepsilon-6\varepsilon_2}{\sum_{j=1}^{3} j\mathbb{P}\left(\widetilde{\mathrm{Po}}(d\hat{\rho}_*)=j\right)+6\varepsilon+12\varepsilon_2}\\ &\geqslant \frac{(d\hat{\rho}_*)^3\exp(-d\hat{\rho}_*)/2}{\mathbb{E}\left(\widetilde{\mathrm{Po}}(d\hat{\rho}_*)\right)}-9\varepsilon-18\varepsilon_2\\ &\geqslant \frac{(d\hat{\rho}_*)^2\exp(-d\hat{\rho}_*)}{2}-20\varepsilon_2,\end{aligned}$$

where in the last line we used that $\varepsilon_2=\sqrt{\varepsilon}$. Similarly, the probability of revealing a factor node of degree at least three is at least

$$\frac{3\mathbb{P}\left(\widetilde{\mathrm{Bi}}(r,\rho_*)=3\right)-3\varepsilon-6\varepsilon_2}{\sum_{j=1}^{3} j\mathbb{P}\left(\widetilde{\mathrm{Bi}}(r,\rho_*)=j\right)+6\varepsilon+12\varepsilon_2}\geqslant\frac{(r-1)(r-2)\rho_*^2(1-\rho_*)^{r-3}}{2}-20\varepsilon_2,$$

which proves our claim. □

C Proof of Lemma 32

Recall that for a node $w\in\mathcal{V}\cup\mathcal{F}$ the event $E_w(\ell)$ holds if and only if $|D_{\leqslant\ell+1}(w)|\leqslant(\log n)^2$ and furthermore $D_{\leqslant\ell+1}(w)$ contains no node which lies in a cycle of length at most 2ℓ. We will prove Lemma 32 under the assumption that $w\in\mathcal{V}$. The case when $w\in\mathcal{F}$ follows since clearly $E_w(\ell)$ is implied by $E_{v_1}(\ell)\cap\ldots\cap E_{v_r}(\ell)$, where $v_1,\ldots,v_r$ are the neighbours of w. Furthermore, applying the statement of the lemma to the variable nodes $v_1,\ldots,v_r$ we have

$$\begin{aligned}\mathbb{P}(E_{v_1}(\ell)\cap\ldots\cap E_{v_r}(\ell))&\geqslant 1-\sum_{i=1}^{r}(1-\mathbb{P}(E_{v_i}(\ell)))\\ &\geqslant 1-r\exp\left(-\Theta\left(\sqrt{\log n}\right)\right)\\ &=1-\exp\left(-\Theta\left(\sqrt{\log n}\right)\right),\end{aligned}$$

as required. We therefore assume in the following that $w\in\mathcal{V}$.

We begin a breadth-first search from w, and index the generations by a *time* t. We run this process until a stopping time T_{stop} which is defined as follows.

Definition 53. Let $w \in \mathcal{V}$ be given and $\ell = \ell(n) = o(\log\log n)$. Let T_{stop} be defined as the smallest time t such that one of the following stopping conditions is invoked:

(S1) The BFS tree has size $(\log n)^2$;

(S2) The BFS tree contains a node which lies in a cycle of length at most 2ℓ;

(S3) $t = \ell + 1$.

We aim to show that whp **(S3)** is applied first.

Proposition 54. *With probability at least* $1 - \exp(-\Theta(\sqrt{\log n}))$, **(S1)** *is not applied.*

Proof. We consider a branching process starting at one node and in which each variable node has offspring distribution $\mathrm{Bi}\left(\binom{n-1}{r-1}, p\right)$ independently, and each factor node has $r-1$ offspring independently—this is an upper coupling on the BFS process started at w. (Note in particular that each factor node has precisely $r-1$ children since the parent has already been counted and the root, which is the only node which has no parent, is a variable node.)

Let us consider the first generation t_0 in which the number of nodes is at least $\sqrt{\log n}$. If no such t_0 exists, or if $t_0 \geqslant \ell + 1$, then the total size of the first $(\ell+2)$ generations is at most $(\ell+2)\sqrt{\log n} \leqslant (\log n)^2$ as required. On the other hand, if $t_0 \leqslant \ell + 1$, let us set x_t to be the size of the t-th generation, for each $t \in \mathbb{N}$. We will assume for technical convenience that $x_t \geqslant \sqrt{\log n}$ for all $t \geqslant t_0$, i.e. the size of a generation does not decrease back below $\sqrt{\log n}$—if this does occur, we simply add in some fictitious nodes to reach this size threshold, which is clearly permissible for an upper bound.

By a standard Chernoff bound (Lemma 11) we have for sufficiently large n that with probability at least $1 - n^{-1}$ the maximum degree of any of the at most $\sqrt{\log n}$ nodes in generation $t_0 - 1$ is $\log n$ and thus we have $x_{t_0} \leqslant (\log n)^{3/2}$. Furthermore, again by Lemma 11, for each even $t \geqslant t_0$ (meaning that the t-th generation consists of variable nodes), the probability that $x_{t+1} \geqslant 2\binom{n-1}{r-1} p x_t$ is at most $\exp\left(-\Theta(\sqrt{\log n})\right)$. Taking a union bound over all at most $\lceil \frac{\ell+3-t_0}{2} \rceil \leqslant \ell + 1$ even generations, and recalling that each variable node has $r-1$ children, we deduce that whp the total size of the process up to generation $\ell + 1$ is at most

$$\begin{aligned} t_0\sqrt{\log n} + (1+(r-1))(\log n)^{3/2} \sum_{i=0}^{\lceil \frac{\ell+3-t_0}{2} \rceil} \left(2\binom{n-1}{r-1}p\right)^i &\leqslant \log n + r(\log n)^{3/2} \sum_{i=0}^{\ell+1} (3d)^i \\ &= (\log n)^{3/2}\Theta\left((3d)^{\ell+1}\right) \\ &\leqslant (\log n)^2, \end{aligned}$$

where the last line follows since $\ell = o(\log\log n)$. In total, the error probability is at most $n^{-1/3} + \exp(-\Theta(\sqrt{\log n})) = \exp(-\Theta(\sqrt{\log n}))$. □

Proposition 55. *With probability at least* $1 - 2n^{-1/3}$, **(S2)** *is not applied.*

Proof. Let $X_{\leqslant 2\ell}$ be the random variable counting the number of variable nodes in $D_{\leqslant \ell+1}(w)$ which lie on cycles of length at most 2ℓ in $G^r(n,p)$. Then

$$\mathbb{E}(X_{\leqslant 2\ell}) \leqslant \sum_{i=2}^{2\ell} i \frac{n^i}{2i} \binom{n}{r-2}^i p^i \leqslant \sum_{i=3}^{2\ell} (n^{r-1}p)^i = o(\log n),$$

where the last line follows since $\ell = o(\log n)$ and $n^{r-1}p = O(1)$. By Markov's inequality, with probability at least $1 - n^{-1/3}$, at most $n^{1/3}\log n$ variable nodes lie on a cycle of length at most 2ℓ. Assuming that this is indeed the case, since **(S1)** was not invoked we infer that the probability that there exists a node in $D_{\leqslant \ell+1}(w)$ that lies in a cycle of length at most 2ℓ is at most $\frac{(\log n)^3 n^{1/3}}{n} \leqslant n^{-1/3}$, and so with probability at least $1 - 2n^{-1/3}$ **(S2)** is not invoked first. □

Proof of Lemma 32. By Propositions 54 and 55, the probability that **(S3)** is invoked first for some particular node w is at least $1 - \exp(-\Theta(\sqrt{\log n})) - 2n^{-1/3} \geqslant 1 - \exp(-\Theta(\sqrt{\log n}))$. In particular, since $\exp(-\Theta(\sqrt{\log n})) = o(1)$, by Markov's inequality whp $(1 - o(1))n$ nodes satisfy $E_w(\ell)$. □

D Proof of Lemma 41

Proof. First observe that any permutation of $\mathcal{V}$ does not affect the probability that G_ℓ is equal to a factor graph H with variable node set $\mathcal{V}$, and therefore we assume without loss of generality that $\phi|_{\mathcal{V}}$ is the identity map.

We will further simplify the proof by simply identifying each $a \in \mathcal{F}_1$ with $\phi(a) \in \mathcal{F}_2$. Note that this is an abuse of terminology: if it were in fact true that $\phi(a) = a$, then the two factor nodes would represent exactly the same edge in the original r-uniform hypergraph, and therefore have exactly the same neighbours. However, although we identify the two factor nodes with one another, we do not carry over these restrictions (indeed, otherwise we would necessarily have $H_1 = H_2$). An alternative way of considering this is to say that we regard the factor nodes no longer as edges of a hypergraph but as abstract nodes stripped of all information.

Now for $i = 1, 2$, let $\mathcal{G}_i$ be the set of factor graphs G of r-uniform hypergraphs such that $G_\ell = H_i$. Observe that any $G \in \mathcal{G}_i$ has precisely the same node set as H_i, and in particular has $|\mathcal{F}_i|$ factor nodes, and therefore the probability that $G^r(n,p) = G$ is simply $p^{|\mathcal{F}_i|}(1-p)^{\binom{n}{r}-|\mathcal{F}_i|}$. Since this value is identical for $i = 1, 2$, what remains to prove is simply that

$$|\mathcal{G}_1| = |\mathcal{G}_2|.$$

We now observe that, as follows from the definition of the peeling process, if $G_\ell = H_i$ then any edge of G which runs between nodes which are non-isolated in H_i is also in H_i. (In other words, the edge set of H_i is induced by the set of non-isolated nodes.) Since the sets of isolated nodes in H_1, H_2 are identical, let $\mathcal{W}$ be this set, so for any $G \in \mathcal{G}_i$, the subgraph of G with all nodes but only those edges of G which lie within $\mathcal{W}$ is precisely $G_\ell = H_i$. Let us define a graph function $f : \mathcal{G}_1 \to \mathcal{G}_2$ which, for each $G \in \mathcal{G}_1$, changes the

edges within $\mathcal{W}$ to those of H_2 instead of H_1, but otherwise leaves all nodes and edges unchanged. We aim to prove that f is a bijection from $\mathcal{G}_1$ to $\mathcal{G}_2$, which implies that these classes have the same size, as required.

The critical part of the proof is to observe that the range of this function does indeed lie within $\mathcal{G}_2$, i.e. that for every $G \in \mathcal{G}_1$ we have $f(G) \in \mathcal{G}_2$. To see this, observe that it is a simple exercise to prove by induction on t that $f(G_t) = (f(G))_t$ for all $G \in \hat{\mathcal{G}}_1$ and all $t \in [\ell]_0$, and in particular $H_2 = f(H_1) = f(G_\ell) = (f(G))_\ell$.

However, slightly more subtly, we have to observe that $f(G)$ is indeed a factor graph arising from an r-uniform hypergraph. It is clear from the construction that $f(G)$ respects the bipartition of variable and factor nodes, and also that the degrees of all nodes are identical in G and $f(G)$, so in particular every factor node has degree r in $f(G)$. Note also that G contained no double edges which means that $f(G)$ cannot contain double edges which do not lie entirely within $\mathcal{W}$, while the fact that H_2 contains no double edges also means that $f(G)$ contains no double edges which lie entirely within $\mathcal{W}$. It therefore remains to prove that $f(G)$ contains no r-duplicates. Observe that in an r-duplicate already all nodes have degree at least two, and therefore none of these nodes will ever be disabled in the peeling process. Thus if $f(G)$ contains an r-duplicate, it is also present in $f(G)_\ell = H_2$, which contradicts our assumption.

We now further observe that f is clearly an injection, since any two distinct graphs $G, G' \in \mathcal{G}_1$ must differ in edges which do not lie completely within $\mathcal{W}$, and therefore $f(G), f(G')$ also differ in those edges. Finally, the function f has an obvious inverse, and therefore is a bijection. It follows that $|\hat{\mathcal{G}}_1| = |\hat{\mathcal{G}}_2|$. □