ABSTRACT. We consider connected components in $k$-uniform hypergraphs for the following notion of connectedness: given integers $k \geq 2$ and $1 \leq j \leq k-1$, two $j$-sets (of vertices) lie in the same $j$-component if there is a sequence of edges from one to the other such that consecutive edges intersect in at least $j$ vertices.

We prove that certain collections of $j$-sets constructed during a breadth-first search process on $j$-components in a random $k$-uniform hypergraph are reasonably regularly distributed with high probability. We use this property to provide a short proof of the asymptotic size of the giant $j$-component shortly after it appears.

Keywords: giant component, phase transition, random hypergraphs, high-order connectedness, degree, branching process

Mathematics Subject Classification: 05C65, 05C80

1. Introduction and main results

One of the most well-known results in the field of random graphs is the phase transition for the emergence of the giant component. Originally observed by Erdős and Rényi and strengthened by various researchers over the years, we may state the result as follows. We say that an event holds with high probability, often abbreviated to whp, if its probability tends to 1 as $n$ tends to infinity.

Theorem 1 ([10, 11, 12]). Let $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$.

(a) If $p = \frac{1-\varepsilon}{n}$, then whp all components of $G(n, p)$ have $O(\varepsilon^{-2} \log(\varepsilon^3 n))$ vertices.
(b) If \( p = \frac{1+\varepsilon}{n} \), then whp the largest component of \( G(n, p) \) has size \((1 \pm o(1))2\varepsilon n\), while all other components have \( O(\varepsilon^{-2} \log(\varepsilon^3 n)) \) vertices.

Our focus in this paper is an extension of this result to \( k \)-uniform hypergraphs\(^1\) for which we need to know what we mean by a connected component in a hypergraph.

Given integers \( k \geq 2 \) and \( 1 \leq j \leq k - 1 \) and a \( k \)-uniform hypergraph \( \mathcal{H} \), we say that two distinct \( j \)-sets (of vertices) \( J_1, J_2 \) are \( j \)-connected if there is a sequence of edges \( E_1, \ldots, E_m \) in \( \mathcal{H} \) such that:

1. \( J_1 \subseteq E_1 \) and \( J_2 \subseteq E_m \);  
2. for each \( 1 \leq i \leq m - 1 \), \(|E_i \cap E_{i+1}| \geq j\).

In other words, we can “walk” from \( J_1 \) to \( J_2 \) using edges which consecutively intersect in at least \( j \) vertices. Additionally we say that a \( j \)-set is \( j \)-connected to itself. Then \( j \)-connectedness is an equivalence relation, and a \( j \)-connected component (or simply \( j \)-component) is an equivalence class of this relation. (Equivalently, a \( j \)-component is a maximal set of pairwise \( j \)-connected \( j \)-sets.) The size of a \( j \)-component is the number of \( j \)-sets it contains.

This provides a whole family of definitions for connectedness. The case \( j = 1 \), also known as vertex-connectedness, is by far the most studied, but larger \( j \), which we refer to as high-order connectedness, provides new and richer challenges.

Given integers \( n, k, j \) and a real number \( p \in [0, 1] \), let \( \mathcal{H}^k(n, p) \) denote the random \( k \)-uniform hypergraph with vertex set \([n] := \{1, \ldots, n\}\) in which each \( k \)-tuple of vertices forms an edge with probability \( p \) independently of one another. Furthermore, we define

\[
\hat{p}_k = \hat{p}_k(n, k, j) := \frac{1}{(\frac{j}{k})!} \frac{1}{(\frac{k}{k})!}.
\]

The following result is a generalisation of the main result in [7].

**Theorem 2.** Let \( 1 \leq j \leq k - 1 \) and let \( \varepsilon = \varepsilon(n) > 0 \) satisfy \( \varepsilon \to 0 \), and \( \varepsilon^2 n^{1-\delta} \to \infty \) for some constant \( \delta > 0 \).

(a) If \( p = (1 - \varepsilon)\hat{p}_k \), then whp all \( j \)-components of \( \mathcal{H}^k(n, p) \) have size at most \( O(\varepsilon^{-2} \log n) \).

(b) If \( p = (1 + \varepsilon)\hat{p}_k \), then whp the size of the largest \( j \)-component of \( \mathcal{H}^k(n, p) \) is \((1 \pm o(1))\frac{1}{(\frac{j}{k})!} \frac{1}{(\frac{k}{k})!} \), while all other \( j \)-components have size at most \( o(\varepsilon cn^j) \).

Note that for \( j = 1 \) the second condition on \( \varepsilon \) becomes \( \varepsilon \gg n^{-1/3} \), which is best possible, while for larger \( j \) the condition \( \varepsilon \gg n^{-\frac{j}{1-j}} \), which is probably not best possible, takes over. We discuss the critical window in more detail in Section 5.

A weaker version of this result appeared in [7] as Theorem 2, where the assumption that \( \varepsilon^3 n^j, \varepsilon^2 n^{1-\delta} \to \infty \) was replaced by the stronger condition \( \varepsilon^3 n^{1-2\delta} \to \infty \), for some constant \( \delta > 0 \). The case \( k = 2 \) and \( j = 1 \) is simply Theorem 1. The case \( j = 1 \) for any \( k \geq 2 \) was proved by Schmidt-Pruzan and Shamir [12].

The proof of Theorem 2 in [7] was based on a short proof of Theorem 1 due to Bollobás and Riordan [6]. The idea is to study an exploration process modelling the growth of components and analyse this process based on a branching process approximation. In the setting of hypergraphs substantial challenges arise when adapting this agenda. In order to overcome these obstacles, the proof of Theorem 2

\(^1A \ k\)-uniform hypergraph is an ordered pair formed by a vertex set and an edge set, where each edge contains precisely \( k \) distinct vertices.
in [7] required significant and lengthy technical details. In this paper we show how much of the technical detail can be avoided and thereby a slightly stronger result can be obtained.

Throughout the paper we fix integers \( k \geq 2 \) and \( 1 \leq j \leq k - 1 \).

The main contribution of [7] to the proof of Theorem 2 was a result stating that certain collections of \( j \)-sets are smooth in the sense that for any \( 0 \leq \ell \leq j - 1 \), no \( \ell \)-set is contained in “too many” \( j \)-sets of the collection (see Corollary 5). Indeed, Corollary 5 followed immediately from a far stronger result in [7]. In this paper we show that we can significantly shorten the proof of Theorem 2 by avoiding this stronger result and proving Corollary 5 more directly.

Let us fix a \( j \)-set \( J_1 \) and explore the \( j \)-component containing \( J_1 \) via a breadth-first search process \( \text{BFS} = \text{BFS}(J_1) \). More precisely, given (arbitrary) orderings \( \sigma_j \) and \( \sigma_k \) of the \( j \)-sets and \( k \)-sets respectively, in \( \text{BFS} \) we start with \( J_1 \) being active and all other \( j \)-sets being neutral. The first generation consists only of the \( j \)-set \( J_1 \). We produce generation \( i + 1 \) from generation \( i \) in the following way. For each \( j \)-set \( J \) of generation \( i \) in order according to \( \sigma_j \), we query all previously unqueried \( k \)-sets containing \( J \), in order according to \( \sigma_k \). If such a \( k \)-set forms an edge, any neutral \( j \)-sets within it are added to generation \( i + 1 \) and are called discovered.

We denote the \( i \)-th generation of this process by \( \partial(i) \). Note that \( \partial(i) \) is a set of \( j \)-sets, which we may also view as a \( j \)-uniform hypergraph. Thus for \( 1 \leq \ell \leq j - 1 \) and an \( \ell \)-set \( L \), we may define the degree of \( L \) in \( \partial(i) \), denoted \( d_L(\partial(i)) \), to be the number of \( j \)-sets of \( \partial(i) \) that contain \( L \). Our goal is to prove that the degrees \( d_L(\partial(i)) \) behave “nicely”.

Next, we fix a constant \( \delta \) satisfying \( 0 < \delta < 1/6 \), and think of it as an arbitrarily small constant – in general our results become stronger for smaller \( \delta \) (the bounds on the error probabilities become weaker, but are still exponentially small). Furthermore, we fix a real-valued function \( \lambda = \lambda(n) \) such that

\[
n^{-1/2+\delta/2}, n^{-j/3} \ll \lambda \ll \varepsilon \ll 1.\]

We use \( \text{BFS} \) to grow the component of some \( j \)-set \( J_1 \) until at the beginning of some round \( i \in \mathbb{N} \) one of the following three stopping conditions is reached:

(\text{S1}) the component of \( J_1 \) is fully explored (i.e. \( \partial(i) = \emptyset \));

(\text{S2}) the (partial) component \( \partial(1) \cup \ldots \cup \partial(i) \) has reached size at least \( \lambda n^j \);

(\text{S3}) the \( i \)-th generation \( \partial(i) \) has reached size at least \( \lambda^2 n^j \).

Moreover, we denote the (first) round in which any these stopping conditions is invoked by

\[
i_1 = i_1(\lambda) := \min\left\{ \min_{i \in \mathbb{N}} \{ \text{S1} \lor \text{S2} \lor \text{S3} \} \right\}.
\]

**Theorem 3.** For any \( 0 \leq \ell \leq j - 1 \), there exists a constant \( C_\ell > 0 \) such that for all \( i \leq i_1 \), with probability \( 1 - \exp(-\Theta(n^{j/2})) \) we have

\[
\Delta_\ell(\partial(i)) \leq C_\ell \left( \frac{|\partial(i)|}{n^\ell} + n^\delta \right).
\]

Note that Theorem 3 becomes trivial in the case \( j = 1 \), which is the main reason why vertex-connectedness is so much easier to handle than high-order connectedness.

\footnote{This is well-defined since BFS always terminates in finite time.}
In Section 2, we show how Theorem 3 can be applied to prove the supercritical case of Theorem 2 in a more concise way than in [7]. The proof of Theorem 3 forms the main body (Section 4) of the paper. The strategy is to split the contribution to the degree of an \( \ell \)-set \( L \) into two parts, called “jumps” and “pivots”. The contributions made by each of these parts are bounded separately, in Propositions 11 and 12 respectively.

2. Application of Theorem 3: a simple proof of Theorem 2

In this section we show how Theorem 3 can be used to prove Theorem 2 (i.e. the hard case of Theorem 2) by following the strategy of [6] for the graph case: we aim to determine asymptotically the number \( L \) of \( j \)-sets in large \( j \)-components, i.e. containing at least \( \lambda n^j \) many \( j \)-sets, using the second moment method. Then it is a simple matter to show that in fact, almost all of these \( j \)-sets lie in the same component using a sprinkling argument. Since the argument already appeared in [7], we will give only an outline here.

We first explore the \( j \)-component \( C_{J_1} \) of \( J_1 \) using BFS until one of the three stopping conditions (S1), (S2) and (S3) is fulfilled. We define the partial components \( C_{J_1}(i) := \partial(1) \cup \ldots \cup \partial(i) \) for any \( i \).

We can approximate the search process by a Galton-Watson branching process \( T \) starting with a single vertex (which represents a \( j \)-set), and in which the number of children of any vertex \( v \) is a random variable \( X_v \), where

\[
\frac{X_v}{\binom{k}{j} - 1} \sim \text{Bi} \left( \binom{n}{k-j}, p \right)
\]

and where the \( X_v \) are independent of each other. It is clear that this is an upper coupling for the search process; the fact that the stopping conditions are invoked before the component grows too large ensures that it will always be a good approximation whp (see Lemma 7).

We use this branching process approximation to prove that the probability of the event \( \mathcal{E} \), that one of stopping conditions (S2) and (S3) is invoked, is approximately the same as the survival probability of the associated branching process, i.e. the probability that the branching process survives indefinitely. Standard branching process techniques show that this probability is approximately \( \frac{2^e}{\binom{j}{j}-1} \).

Claim 4. \( \mathbb{P}(\mathcal{E}) = (1 + o(1)) \cdot \frac{2^e}{\binom{j}{j}-1} \).

For completeness, we prove this claim in Appendix B. Furthermore, conditional on (S3) being invoked, \( C_{J_1} \) will be large whp. Also if (S2) is invoked, then clearly \( C_{J_1} \) is large. This already shows that the first moment of \( L \), the number of \( j \)-sets in large components, is \( \frac{2^e}{\binom{j}{j}-1} \).

In order to determine the second moment of \( L \), we consider a second exploration process (of a \( j \)-component \( C_{J_2} \)) starting from another \( j \)-set \( J_2 \) outside \( C_{J_1}(i_1) \). (The contribution to the second moment from \( j \)-sets \( J_2 \) inside \( C_{J_1}(i_1) \) is easily shown to be negligible.) In order to ensure independence, we first delete all the \( j \)-sets of \( C_{J_1}(i_1) \) from the hypergraph, so any \( k \)-set containing such a \( j \)-set may no longer be queried. However, since we have not deleted many \( j \)-sets, \( T \) is still a good approximation for this search process. Again, the probability that this search process becomes large is approximately \( \frac{2^e}{\binom{j}{j}-1} \).
This almost completes the proof of Theorem 2, but there is one more case to consider, namely that the search process from \( J_2 \) remained small, while in fact the component is large, but we did not see this because we had deleted some \( j \)-sets.

The first observation is that since most of the \( j \)-sets of \( C_{J_1}(i_1) \) were fully explored, many potential such queries had already been made while exploring \( C_{J_1} \). The only ones we might have missed are the ones containing a \( j \)-set of \( \partial(i_1) \) and \( j \)-set of \( C_{J_2} \).

However, given two \( j \)-sets \( J, J' \), one from each of these components, how many \( k \)-sets contain both of them? For \( j = 1 \) (so in particular for graphs) this is simply \( (n - 2k - 2) \), but for \( j \geq 2 \) the answer is fundamentally dependent on the size of the intersection \( J \cap J' \). We therefore need the following corollary of Theorem 3.

**Corollary 5.** There exists a real-valued function \( \xi = \xi(n) \) satisfying \( \xi/\log n \to \infty \) and \( \xi = o(\lambda^2 n) \) and constants \( \kappa_1, \ldots, \kappa_{j-1} > 0 \) (independent of \( n \)) such that, conditional on \( E \), with probability at least \( 1 - \exp(-\Theta(\xi^{1/4})) \), for all \( 1 \leq \ell \leq j - 1 \) and \( \ell \)-sets \( L \) we have

\[
d_L(\partial(i_1)) \leq \kappa_\ell (|\partial(i_1)|n^{-\ell} + \xi).
\]

Now given \( J \in C_{J_2} \), we can consider all subsets \( L \) of \( J \) and bound from above the number of \( j \)-sets of \( \partial(i_1) \) that intersect \( J \) in \( L \).

Thus we obtain an upper bound on the number of \( k \)-sets that we have not queried because of deleting \( C_{J_1}(i_1) \). With some careful calculation, we observe that the expected number of edges among these \( k \)-sets is small, and applying Markov’s inequality, the probability that there is at least one such edge is small enough that it is negligible.

It remains to show how Corollary 5 can be deduced from Theorem 3.

**Proof of Corollary 5.** Note that Theorem 3 is a generalisation of Corollary 5 in that it applies to any generation, and furthermore, is not conditional on the event \( E \).

Let us denote by \( A \) the event that (3) (i.e. the conclusion of Corollary 5) holds and by \( B \) the event that (2) (i.e. the conclusion of Theorem 3) holds for all \( i \leq i_1 \). Then we have \( B \subset A \) and therefore

\[
P(A|E) \geq 1 - \frac{\Pr(B) + \Pr(E)}{\Pr(E)} \geq 1 - \frac{\Pr(B)}{\Pr(E)}.
\]

Hence, using Theorem 3 and Claim 4 we obtain

\[
P(A|E) \geq 1 - \frac{\exp(-\Theta(n^{\delta/2}))}{\Omega(\varepsilon)} = 1 - \exp(-\Theta(n^{\delta/2}))
\]

where the final equality we used the fact that \( \log(1/\varepsilon) = o(n^{\delta/2}) \).

This also completes the proof of Theorem 2.

**Remark 6.** As mentioned earlier, Corollary 5 already appeared as Lemma 4 in [7]. The proof was long and complicated, and in fact a much stronger result was proved (in which the degrees of \( \ell \)-sets are asymptotically determined). Using Theorem 3 allows us to avoid much of this work.
3. Preliminaries

We first provide some basic properties of the search process BFS. Most of these
results were already stated in [7, 8] in a similar fashion. Because the proofs are
short, we include them for local reference.

We define two abstract branching processes, which will form upper and lower
couplings on the search process BFS. For these branching processes, the j-sets
will be represented by vertices.

Let \( T^* \) denote a Galton-Watson branching process starting with a single vertex,
and in which the number of children of any vertex is a random variable \( X \), where
\[
\frac{X}{\binom{k}{j} - 1} \sim \text{Bi} \left( \binom{n}{k - j}, p \right)
\]
It is clear that \( T^* \) forms an upper coupling on BFS, since from any j-set we may
query at most \( \binom{n}{k - j} \) many k-sets, each such k-set forms an edge with probability
\( p \), and if it forms an edge, we discover at most \( \binom{k}{j} - 1 \) new j-sets. If we actually
discover fewer j-sets in BFS, we can artificially add in some dummy j-sets, and
equally make some additional dummy queries, to “fill up” to \( T^* \). We denote this
coupling by BFS \( \prec T^* \).

For the lower coupling we define a similar branching process \( T_\epsilon \), where in this
case the number of children of any vertex is distributed as a random variable \( Y \),
where
\[
\frac{Y}{\binom{k}{j} - 1} \sim \text{Bi} \left( (1 - \epsilon_\star) \binom{n}{k - j}, p \right)
\]
with some \( \epsilon_\star = \epsilon_\star(n) > 0 \) satisfying \( \lambda \ll \epsilon_\star \ll \epsilon \), say \( \epsilon_\star := \sqrt{\lambda \epsilon} \). Note that \( T_\epsilon \)
does not always form a lower coupling for BFS, but early on in the search process,
it is very likely to. More specifically, as long as it is still true that from each j-set
we make at least \( (1 - \epsilon_\star) \binom{n}{k - j} \) queries to k-sets which contain \( \binom{k}{j} - 1 \) undiscovered
j-sets, \( T_\epsilon \) forms a lower coupling, which we denote by \( T_\epsilon \prec \text{BFS} \).

We first show that whp \( T_\epsilon \) will remain a lower coupling during the period of the
search process that we are interested in. We introduce a new stopping time: let \( i_2 \)
be the first round \( i \) for which either (S1) or (S2) is satisfied. Note that we have
\( i_2 \geq i_1 \).

**Lemma 7.** With probability at least \( 1 - \exp(-\Theta(n^{3/2})) \) for each \( 0 \leq \ell \leq j - 1 \) and
\( 1 \leq i \leq i_2 \) we have
\[
\Delta_\ell(C_{J_1}(i)) = O(\lambda n^{2 - \ell}) = o(\epsilon_\star n^{2 - \ell}).
\]
In particular, with probability at least \( 1 - \exp(-\Theta(n^{3/2})) \) for any \( 1 \leq i \leq i_2 \) we have
\( T_\epsilon \prec \text{BFS} \).

Note that this result has a broadly similar flavour to Corollary 5 and Theorem 6
with the crucial difference that the degree bound applies to the whole component,
not to a single generation. While Lemma 7 has a strong resemblance to Lemma 12
in [8] and Lemmas 14 and 17 in [7], it is not immediately apparent that these
directly imply Lemma 7. Since a formal proof of this implication would involve
checking tedious technicalities, we instead provide a stand-alone proof for Lemma 7
in Appendix A for completeness.
We use Lemma 7 twice: firstly, we use it in the proof of Claim 8 in Appendix B (with \( i > i_2 \)); secondly, we use it to deduce that after seeing a sufficiently large generation in BFS, the next one will typically not be smaller.

**Lemma 8.** With probability at least \( 1 - \exp(-\Theta(n^{\delta/2})) \), for all \( 1 \leq i \leq i_1 \) such that \( |\partial(i)| \geq n \) we have

\[
|\partial(i + 1)| \geq |\partial(i)|.
\]

This lemma will be applied in Section 4 during the proofs of Propositions 11 and 12.

**Proof.** By the coupling \( T_s \prec \text{BFS} \) provided by Lemma 7, conditional on \( |\partial(i)| \), we may couple \( |\partial(i + 1)| \) with a random variable \( Z_\ast \) such that \( |\partial(i + 1)| \geq \left( \binom{k}{j} - 1 \right) Z_\ast \), where

\[
Z_\ast \sim \text{Bi} \left( 1 - \varepsilon_\ast \left( \frac{n}{k - j} \right) |\partial(i)|, p \right).
\]

Note that \( E(Z_\ast) = (1 - \varepsilon_\ast)(1 + \varepsilon) \left( \binom{k}{j} - 1 \right)^{-1} |\partial(i)| \), and thus we obtain

\[
\Pr(|\partial(i + 1)| < |\partial(i)|) \leq \Pr \left( Z_\ast \leq \left( \frac{k}{j} - 1 \right)^{-1} |\partial(i)| \right)
\]

\[
\leq \exp \left( \frac{((1 - \varepsilon_\ast)(1 + \varepsilon) - 1)^2 \left( \binom{k}{j} - 1 \right)^{-1} |\partial(i)|}{2(1 - \varepsilon_\ast)(1 + \varepsilon)} \right)
\]

\[
\leq \exp \left( -\Theta(\varepsilon^2 n) \right) \leq \exp(-\Theta(n^{\delta/2})),
\]

where the penultimate inequality follows since \( |\partial(i)| \geq n \). \( \square \)

At various points in the proof of Theorem 3 the following result, which is a direct consequence of a Chernoff bound, will be more convenient than the standard Chernoff bound formulations.

**Lemma 9.** For any \( m = m(n) \in \mathbb{N} \) and \( p = p(n) \in [0,1] \) and constant \( a > 0 \) the following holds. Let \( X \) be a random variable distributed as \( \text{Bi}(m, p) \). Then with probability at least \( 1 - \exp(-\Theta(n^{\delta/2})) \),

\[
X \leq (1 + a)mp + 2n^{\delta}.
\]

**Proof.** We split the proof into two cases. If \( mp \leq n^{\delta} \), then \( X \) is dominated by a random variable \( Y \sim \text{Bi}(m, p') \), where \( p' := n^{\delta}/m \) (if \( p' > 1 \), then \( m < n^{\delta} \) and the upper bound is trivial). Let \( \mu := E(Y) = n^{\delta} \). Now using a multiplicative Chernoff bound, which states that

\[
\Pr(Y \geq (1 + \eta)\mu) \leq \left( \frac{e^\eta}{(1 + \eta)^{1+\eta}} \right)^\mu
\]

we have

\[
\Pr(X \geq 2n^{\delta}) \leq \Pr(Y \geq 2n^{\delta}) = \Pr(Y \geq 2\mu) \leq \left( \frac{e^\delta}{2^\delta} \right)^n = \exp \left( -\Theta(2n^\delta) \right).
\]

On the other hand, if \( mp > n^{\delta} \), then by an additive Chernoff bound, we obtain

\[
\Pr(X \geq (1 + a)mp) \leq \exp \left( -\frac{a^2 mp}{2(a + 1/3)} \right) = \exp \left( -\Omega(n^\delta) \right),
\]
and the statement follows. □

4. Proof of main result

We prove Theorem 3 for a set of recursively defined constants $C_\ell$. (Note that we have made no attempt to optimise these constants.)

We first define $c_\ell := \binom{k-\ell}{j-\ell} - 1$ for each $0 \leq \ell \leq j - 1$, and set $u_0 = w_0(\ell) := \max\{0, j + \ell - k\}$ and $r_\ell := \frac{c_\ell}{c_0}$. Note that $r_\ell < 1$ for $\ell > 0$. Therefore we can fix some constant $\alpha > 0$ such that $r_\ell := (1 + \alpha)(1 + \varepsilon)r_\ell < 1$ for all $1 \leq \ell \leq j - 1$.

We set

$$C_0 := 1$$

and iteratively for $1 \leq \ell \leq j - 1$ we let

$$C'_\ell := \left\lceil \frac{k-\ell}{j-\ell} \right\rceil \max \left\{ \frac{(1+\alpha)(1+\varepsilon)(k-j)!}{c_0} \sum_{w=\max(0,\ell)}^{\ell-1} \frac{C_w}{(k-j-\ell+w)!^3} \right\},$$

$$C_\ell := \frac{C'_\ell + 2\alpha + 1}{1 - r_\ell}. \tag{5}$$

Let us observe that $C'_\ell = f(C_0, \ldots, C_{\ell-1}; k, j, \ell)$, while $C_\ell = g(C'_\ell; k, j, \ell)$, so these constants are recursively well-defined.

We first note that Theorem 3 is trivial in the case $\ell = 0$, since then $\Delta_{\ell}(\partial(i)) = |\partial(i)| = C_0 \frac{|\partial(i)|}{n}$ for any $i$. Therefore in the remainder of the proof we will assume that $\ell \geq 1$.

The proof of Theorem 3 relies on distinguishing two types of contribution to the degrees of $\ell$-sets.

**Definition 10.** Let $L$ be an $\ell$-set with $1 \leq \ell \leq j - 1$ and let $i \geq 1$.

(i) A jump to $L$ (in step $i$) occurs when we query a $k$-set containing $L$ from a $j$-set in $\partial(i - 1)$ which did not contain $L$ and the $k$-set forms an edge of $H^k(n, p)$. Such an edge contributes at most $\binom{k-\ell}{j-\ell}$ to $d_L(\partial(i))$.

(ii) A pivot at $L$ (in step $i$) occurs when we query any $k$-set from a $j$-set in $\partial(i - 1)$ containing $L$ and it forms an edge of $H^k(n, p)$. Such an edge contributes at most $\binom{k-\ell}{j-\ell} - 1$ to $d_L(\partial(i))$.

The following two propositions bound the contribution to the degree $d_L(\partial(i))$ made by jumps and pivots, respectively. We first state these propositions and then show how they combine to prove Theorem 3 before providing their proofs.

**Proposition 11.** For any $0 \leq \ell \leq j - 1$, any $\ell$-set $L$ and any $2 \leq i \leq i_1$, with probability at least $1 - \exp(-\Theta(n^{\delta/2}))$ the contribution made to $d_L(\partial(i))$ by jumps is

$$d^{(j)}_L(\partial(i)) \leq C'_\ell \left( \frac{|\partial(i)|}{n^\ell} + n^\delta \right).$$

**Proposition 12.** For any $0 \leq \ell \leq j - 1$, any $\ell$-set $L$ and any $2 \leq i \leq i_1$, with probability at least $1 - \exp(-\Theta(n^{\delta/2}))$ the contribution made to $d_L(\partial(i))$ by pivots is

$$d^{(p)}_L(\partial(i)) \leq (r'_\ell C_\ell + 2\alpha + 1) \left( \frac{|\partial(i)|}{n^\ell} + n^\delta \right).$$
Proof of Theorem 3. Now assuming Propositions 11 and 12 hold, we note that for any $0 \leq \ell \leq j - 1$, any $\ell$-set $L$ and any $2 \leq i \leq i_1$, assuming the conclusions of Propositions 11 and 12 hold, we have

$$d_L(\partial(i)) = d_L^{(j \ell)}(\partial(i)) + d_L^{(j \ell \ell)}(\partial(i)) \leq (r'_\ell C\ell + 2c\ell + 1 + C'_\ell) \left( \frac{||\partial(i)||}{n^\ell} + n^\delta \right)$$

and furthermore

$$r'_\ell C\ell + 2c\ell + 1 + C'_\ell = r'_\ell C\ell + (1 - r'_\ell)C\ell = C\ell.$$

Taking a union bound over all choices of $\ell$, $L$ and $i$ (of which there are certainly at most $jn^2\delta$, and observing that $jn^2\delta \exp(-\Theta(n^\delta/2)) = \exp(-\Theta(n^\delta/2))$, this completes the proof of Theorem 3. □

4.1. Jumps: proof of Proposition 11. Fix some $\ell$, $L$ and $i$ and let $m'_2 = m'_2(L, i)$ denote the number of queries to be made in generation $i$ that would result in jumps to $L$ if an edge is found. Given a $j$-set $J$, the number of $k$-sets containing $L$ and $J$ is 0 if $|J \cap L| < w_0$, and at most $(k - j - \ell + w)$ if $|J \cap L| = w \geq w_0$. Thus we consider the number of $j$-sets in $\partial(i - 1)$ which intersect $L$ in $w$ vertices with $w_0 \leq w \leq \ell - 1$. This is at most $(\ell - w)\Delta_w(\partial(i - 1))$. Consequently, we have

$$m'_2 \leq \sum_{w = w_0}^{\ell - 1} \left( \begin{array}{c} \ell - w \cr w \end{array} \right) \Delta_w(\partial(i - 1)) \left( \begin{array}{c} n \cr k - j - \ell + w \end{array} \right) \leq \sum_{w = w_0}^{\ell - 1} \left( \begin{array}{c} \ell \cr w \end{array} \right) C_w \left( \begin{array}{c} ||\partial(i - 1)|| \cr n^w \end{array} \right) n^{k - j - \ell + w} \left( \begin{array}{c} n^{k - j - \ell + w} \cr (k - j - \ell + w)! \end{array} \right) = m_2,$$

with probability at least $1 - \exp(-\Theta(n^\delta/2))$, where we have used the induction hypothesis for the second inequality.

Thus we consider the number of such edges we discover is dominated by $Bi(m_2, p)$, and so by Lemma 3 with probability $1 - \exp(-\Theta(n^\delta/2))$ is at most

$$(1 + \alpha)m_2p + 2n^\delta.$$

Furthermore, each such edge results in at most $(\ell - j - \ell)$ new $j$-sets containing $L$ becoming active. Thus we have

$$d_L^{(j \ell \ell)}(\partial(i)) \leq \left( \begin{array}{c} \ell \cr k - \ell \end{array} \right) \left( \begin{array}{c} 1 + \varepsilon \cr c_0 \end{array} \right) \frac{1}{C_0(1 - O(1/n))} \sum_{w = w_0}^{\ell - 1} \left( \begin{array}{c} \ell \cr w \end{array} \right) C_w \left( \begin{array}{c} ||\partial(i - 1)|| \cr n^w \end{array} \right) n^{k - j - \ell + w} \left( \begin{array}{c} n^{k - j - \ell + w} \cr (k - j - \ell + w)! \end{array} \right) = \left( \begin{array}{c} \ell \cr k - \ell \end{array} \right) \left( \begin{array}{c} 1 + \varepsilon \cr c_0 \end{array} \right) \left( \begin{array}{c} \ell - w \cr w \end{array} \right) C_w \left( \begin{array}{c} ||\partial(i - 1)|| \cr n^w \end{array} \right) n^{k - j - \ell + w} \left( \begin{array}{c} n^{k - j - \ell + w} \cr (k - j - \ell + w)! \end{array} \right) + 2n^\delta.$$

We now observe that since $w \leq \ell - 1$, the term in $n^{-\ell + w + \delta}$ is always $O(n^{-1 + \delta})$, and therefore we may absorb all such terms into the $n^\delta$ term by increasing the constant slightly.

We would also like to replace $||\partial(i - 1)||$ by $||\partial(i)||$. This is certainly possible for an upper bound if $||\partial(i - 1)|| \geq n$ by Lemma 3. However, if $||\partial(i - 1)|| \leq n$, we simply observe that $d_L^{(j \ell \ell)}(\partial(i)) \leq n^{1 - \ell} = o(n^\delta)$ (because we have $\ell \geq 1$), and in this case we have $d_L^{(j \ell \ell)}(\partial(i)) \leq 3n^\delta$. 
Thus in either case we have
\[
\frac{d_L^{(pj)}(\partial(i))}{\binom{k-\ell}{j-\ell}} \leq (1 + \alpha)(1 + \varepsilon)(k - j)! \sum_{w=w_0}^{\ell-1} \binom{\ell}{w} \frac{C_w}{(k - j - \ell + 1)!} \left( \frac{\partial(i)}{n^\ell} \right) + 3n^\delta
\]
and
\[
\leq \binom{C'}{\ell} \left( \frac{\partial(i)}{n^\ell} + n^\delta \right)
\]
as claimed. \qed

4.2. Pivots: proof of Proposition 12. Fix some \(\ell, L\) and \(i\) and let \(m_3' = m_3'(L, i)\) denote the number of queries to be made in generation \(i\) that would result in pivots at \(L\) if an edge is found. The number of \(j\)-sets in \(\partial(i - 1)\) which could lead to a pivot is at most \(\Delta(\partial(i - 1))\) and causes at most \(\binom{n}{k-j}\) many queries. Thus, with probability at least \(1 - \exp(-\Theta(n^{\delta/2}))\), we obtain
\[
m_3' \leq \Delta(\partial(i - 1)) \binom{n}{k-j} \leq C_L \left( \frac{\partial(i-1)}{n^\ell} + n^\delta \right) \binom{n}{k-j} =: m_3,
\]
where we have used the induction hypothesis for the second inequality.

The number of such edges we discover is dominated by \(Bi(m_3, p)\), and so by Lemma 9 with probability \(1 - \exp(-\Theta(n^{\delta/2}))\) is at most
\[(1 + \alpha)m_3p + 2n^\delta.
\]
Furthermore, each such edge results in at most \((k-\ell)-1 = c_\ell\) new \(j\)-sets containing \(L\) becoming active (the one from which we are currently querying is already active). Thus we have
\[
d_L^{(pv)}(\partial(i)) \leq \frac{1 + \varepsilon}{c_\ell} \frac{n^\delta}{\binom{n}{k-j}} C_L \left( \frac{\partial(i-1)}{n^\ell} + n^\delta \right) \binom{n}{k-j} + 2n^\delta
\]
and
\[
d_L^{(pv)}(\partial(i)) \leq (1 + O(1/n))r'_L \left( \frac{\partial(i-1)}{n^\ell} \right) + (r'_L C_L + 2c_\ell)n^\delta,
\]
where we recall that \(r'_L = (1 + \alpha)(1 + \varepsilon)c_\ell/c_0\).

As before, if \(|\partial(i-1)| \geq n\), we have \(|\partial(i-1)| \leq |\partial(i)|\) by Lemma 8 while if \(|\partial(i-1)| \leq n\), we have \(|\partial(i-1)| = o(n^\delta)\). In either case we have
\[
d_L^{(pv)}(\partial(i)) \leq (1 + O(1/n))r'_L \left( \frac{\partial(i)}{n^\ell} \right) + (r'_L C_L + 2c_\ell + 1)n^\delta
\]
as claimed. \qed

5. Concluding remarks

There are several topics which have been studied extensively for random graphs but remain open questions in random hypergraphs, particularly for \(j \geq 2\).
5.1. **Critical window.** With our improvements on Theorem 2 compared to [7], we have come one step closer to determining the width of the critical window for the emergence of a unique largest \( j \)-component in \( \mathcal{H}^k(n, p) \). However, the lower bound on \( \varepsilon \) given by \( \varepsilon^2 n^{1-\delta} \to \infty \) is probably still not best possible for \( j \geq 2 \). We conjecture the following:

**Conjecture 13.** Theorem 2 holds for all \( \varepsilon = \varepsilon(n) \) satisfying \( \varepsilon \to 0 \) and \( \varepsilon^3 n^j \to \infty \). Furthermore, if \( p = (1 + cn^{-j/3})\hat{p}_n \) for some fixed \( c \in \mathbb{R} \), then whp all \( j \)-components are of size \( O(n^{2j/3}) \) and there is more than one \( j \)-component of \( \mathcal{H}^k(n, p) \) of size \( \Theta(n^{2j/3}) \).

Note that if \( \varepsilon^3 n^j \) is constant, the bounds on the size of the largest component from the super-critical case (\( \Theta(\varepsilon n^j) \)) and the sub-critical case (\( O(\varepsilon^{-2} \log n) \)) match up to the \( \log n \) term, suggesting that we have a smooth transition. In particular, this condition is also sufficient for the sprinkling argument in Section 2 to work.

5.2. **Asymptotic normality of the giant.** In the supercritical case, it would be interesting to determine the asymptotic distribution of the size of the giant component more precisely, as was done for the case \( j = 1 \) in [3, 5], where the size of the giant component was shown to tend to a normal distribution.

5.3. **Nullity.** One can also consider the structure of the components in the subcritical, critical, or supercritical regimes. For graphs it is well-known that whp all components in the subcritical regime, and whp all except the giant in the supercritical regime, are either trees or contain at most one cycle. Thus we aim to generalise the notion of a tree to \( j \)-connectedness. One possibility is via the nullity (with respect to \( j \)-connectedness) of a \( k \)-uniform hypergraph \( \mathcal{H} = (V, E) \), which we define to be

\[
\nu = \nu(\mathcal{H}; j, k) := |C| + c_0 |E| - \left( \binom{|V|}{j} \right),
\]

where \( C \) denotes the set of \( j \)-components of \( \mathcal{H} \), and we recall that \( c_0 = \binom{k}{j} - 1 \). For a collection \( J \subset \binom{V}{j} \) and an edge set \( E \subset \binom{V}{k} \) we define the pair \( (J, E) \) to be a (hyper-)tree if \( J \) is a \( j \)-connected component in \( \mathcal{H} \) and \( \nu = 0 \).

In contrast to graphs, a \( j \)-component does not necessarily contain a spanning tree, but in [7] it was asked what the minimal nullity of a spanning structure in the giant \( j \)-component after the phase transition is. We might also ask about the nullities of other components in the subcritical and supercritical regimes, and whether these are small whp, as is the case for graphs.

An approach based on nullities may also help to analyse components within the critical regime, as was done for graphs in e.g. [1, 2, 10, 11].

References

1. L. Addario-Berry, N. Broutin, and C. Goldschmidt, *The continuum limit of critical random graphs*, Probability Theory and Related Fields 152 (2012), no. 3, 367–406.
2. D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, The Annals of Probability 25 (1997), no. 2, 812–854.
3. M. Behrisch, A. Coja-Oghlan, and M. Kang, *Local limit theorems for the giant component of random hypergraphs*, Combin. Probab. Comput. 23 (2014), no. 3, 331–366. MR 3189416

---

Note that this is not the only reasonable definition of a tree in hypergraphs.
Our aim in this section is to prove Lemma 4. We first prove that with very high probability
\[ |C_{J_1}(i_2)| \leq 3\lambda n^j. \] (6)
For by the definition of \(i_2\) we have \(|\partial(i_2 - 1)| \leq |C_{J_1}(i_2 - 1)| < \lambda n^j\), and therefore the number of queries we make while exploring generation \(i_2\) is at most
\[ \lambda n^j \binom{n}{k-j}. \]
Therefore the expected number of edges we discover while exploring generation \(i_2\) is at most \(3\lambda n^j\), as claimed.

**Proof of Lemma 4** We now prove the first assertion of Lemma 4 by induction on \(\ell\). Note that by monotonicity we may assume \(i = i_2\). The base case \(\ell = 0\) follows immediately from (6), so assume \(\ell \geq 1\) and that \(\Delta_v(C_{J_1}(i_2)) \leq S_v\lambda n^{j-\ell}\) for each \(0 \leq \ell' \leq \ell - 1\) and some constants \(S_0, \ldots, S_{\ell-1}\).

Fix an \(\ell\)-set \(L\). Let us consider how the degree of \(L\) in \(C_{J_1}(i_2)\) might grow. We bound the contribution from jumps and pivots separately.

We first consider how many queries we may make from a \(j\)-set \(J \in C_{J_1}(i_2)\) not containing \(L\) to a \(k\)-set containing \(L\), i.e. the number of queries which might lead to a jump to \(L\). Given \(J\), the number of \(k\)-sets containing \(L\) and \(J\) is at most \(\binom{n}{k-j-w} \), where \(w = |J \cap L|\). Thus we consider the number of \(j\)-sets in \(C_{J_1}(i_2)\) which intersect \(L\) in \(w\) vertices. This is at most \(\binom{\ell}{w}\Delta_w(C_{J_1}(i_2))\). Thus the number
of queries which might result in a jump to $L$ is

\[ \sum_{w=w_0}^{\ell-1} \binom{\ell}{w} D_w(C_{j_1}(i_2)) \left( \frac{n}{k-j-\ell+w} \right)^{n-w} \leq \sum_{w=w_0}^{\ell-1} \binom{\ell}{w} S_w \lambda n^{k-\ell} \]

\[ \leq 2^\ell \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \lambda n^{k-\ell}. \]

By Lemma 9 and the fact that $\lambda p = \omega(n^{j-k-1+\delta})$, the number of edges we find in this way is at most $2p\lambda n^{k-\ell} 2^\ell \max_{w=w_0,\ldots,\ell-1} S_w$ with probability at least $1 - \exp(-\Theta(n^{\delta/2}))$. Finally, each such edge contributes at most $(k-\ell)^{-1}$ to the degree of $L$, and so the contribution made by jumps is at most

\[ 2^{k+\ell+1} \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \lambda p n^{k-\ell} \leq 2^{k+\ell+1} \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \lambda n^{j-\ell} \]

with probability at least $1 - \exp(-\Theta(n^{\delta/2}))$.

On the other hand, let us consider the pivots at $L$ as forming a set of pivot processes – if we discover a $j$-set $J'$ from a $j$-set $J$ via a pivot at $L$, then $J$ and $J'$ both contain $L$ and are part of the same pivot process. Each $j$-set arising from a jump to $L$ (and possibly also $J_1$ if this contains $L$) gives rise to such a pivot process at $L$. A pivot process is a search process on $j$-sets containing $L$ in a $k$-uniform hypergraph. By removing $L$ from each of the $j$-sets, it becomes a search process on $(j-\ell)$-sets in a $(k-\ell)$-uniform hypergraph. We note that the number of children in such a process is dominated by a random variable $X^{(pv)}$ with

\[ \frac{X^{(pv)}}{\binom{k-\ell}{j-\ell}} - 1 \sim \text{Bi} \left( \binom{n}{k-j}, p \right) \]

and we therefore define $\mathcal{T}^{(pv)}$ to be an abstract branching process on vertices (which represent $j$-sets containing $L$) in which the number of children of each vertex has this distribution. The expected number of children is

\[ (1 + \varepsilon) \frac{\binom{k-\ell}{j-\ell} - 1}{\binom{k}{j} - 1} < 1, \]

where the inequality follows since $\ell \geq 1$ and $\varepsilon = o(1)$. In other words, $\mathcal{T}^{(pv)}$ is subcritical. We can then show that with probability $1 - \exp(-\Theta(n^{\delta/2}))$, if we start $x \geq n^3$ such processes, then their combined total size is $O(x)$.

More precisely, consider a set of $x \geq n^3$ independent copies of $\mathcal{T}^{(pv)}$. If $x-1$ is the maximum of $\binom{k-\ell}{j-\ell}$ times the number of jumps to $L$ and $n^3$, then the total number of vertices in these $x$ processes dominates the number of pivots at $L$.

We imagine generating children via a sequence of Bernoulli queries with success probability $p$, each success giving rise to $\binom{k-\ell}{j-\ell} - 1$ children. In order for the processes to reach total size $C_x$ for some constant $C$, we would need at least \( \frac{(C-1)x}{\binom{k-\ell}{j-\ell}} \) of the first $C_x \binom{n}{k-j}$ queries to be successful. But the probability of this can be bounded
by a Chernoff bound:

\[ \mathbb{P}\left( \text{Bi}\left( Cx\left( \frac{n}{k-j}\right), p \right) \geq \frac{(C-1)x}{(k-j-\varepsilon)} - 1 \right) \leq \exp \left( -\frac{x^2}{2} \left( \frac{C}{c_\ell} - Cp(\frac{n}{k-j}) \right)^2 \right) \leq \exp \left( -\frac{x}{2} \left( \frac{C}{c_\ell} - \frac{1+(\varepsilon)}{c_\ell} \right)^2 \right) \]

where the last line follows for sufficiently large \( C \) because \( c_\ell < c_0 \) and \( \varepsilon = o(1) \) and because \( x \geq n^{\delta} \).

Now recall that with probability at least \( 1 - \exp(-\Theta(n^{\delta/2})) \) the number of jumps to \( L \) is at most \( 2^{k+\ell+1} \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \lambda n^{j-\ell} \), therefore for sufficiently large \( C \), the total contribution to the degree of \( L \) made by pivots is at most

\[ C \left( 2^{k+\ell+1} \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \lambda n^{j-\ell} + 1 \right), \]

since \( \lambda n^{j-\ell} \geq \lambda n = \omega(n^{\delta}) \).

Consequently, the total number of \( j \)-sets containing \( L \) is at most

\[ (C+1) \left( 2^{k+\ell+1} \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \lambda n^{j-\ell} + 1 \right) \leq (C+2)2^{k+\ell+1} \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \lambda n^{j-\ell} \]

with probability at least \( 1 - \exp(-\Theta(n^{\delta/2})) \). We obtain the inductive step by setting \( S_\ell := (C+2)2^{k+\ell+1} \max_{w=w_0,\ldots,\ell-1} \{ S_w \} \) and taking a union bound over all \( \ell \)-sets \( L \) (since \( \binom{n}{\ell} \exp(-\Theta(n^{\delta/2})) = \exp(-\Theta(n^{\delta/2})) \)).

Finally, we prove that \( T_* \) is a lower coupling. For given a \( j \)-set \( J \in \partial(i) \), we bound the number of \( k \)-sets \( K \) containing another \( j \)-set \( J' \in C_{J_1}(i) \), and which therefore would give fewer than \( c_0 = \binom{\ell}{2} - 1 \) new \( j \)-sets. We distinguish cases based on \( \ell = |J \cap J'| \) and observe that the number of such \( k \)-sets is at most

\[ \sum_{\ell=\max\{0,2j-k\}}^{j-1} \Delta_{\ell}(C_{J_1}(i)) \left( \binom{n}{k-2j+\ell} \right) \leq \sum_{\ell=0}^{j-1} o(\varepsilon_* n^{j-\ell}) n^{k-2j+\ell} = o(\varepsilon_* n^{k-j}). \]

Thus the number of \( k \)-sets that can be queried from \( J \) and which contain no further discovered \( j \)-sets is

\[ \binom{n-j}{k-j} - o(\varepsilon_* n^{k-j}) = (1 - O(\frac{1}{n}) - o(\varepsilon_*)) \left( \binom{n}{k-j} \right) \geq (1 - \varepsilon_*) \left( \binom{n}{k-j} \right) \]

as required. \( \square \)
Appendix B. Survival Probability

Proof of Claim \[4\]. Let \( \mathcal{A} \) denote the event that \( \mathcal{T}_* \prec \text{BFS} \) for all \( i \leq i_2 \). We observe that \( \mathbb{P}(\mathcal{A}) \geq 1 - \exp(-\Theta(n^{3/2})) \) by Lemma \[7\]. Thus we have

\[
\mathbb{P}(\mathcal{E}) \geq \mathbb{P}(|C_j(i_2)| \geq \lambda n^j) \geq \mathbb{P}(|C_j(i_2)| \geq \lambda n^j \land \mathcal{A}) \\
\geq \mathbb{P}(|\mathcal{T}_*| \geq \lambda n^j \land \mathcal{A}) \\
\geq \mathbb{P}(\mathcal{T}_* \text{ survives}) - \mathbb{P}(\neg \mathcal{A}) \\
\geq \mathbb{P}(\mathcal{T}_* \text{ survives}) - \exp(-\Theta(n^{3/2})).
\]

On the other hand, note that by some elementary calculation (see Appendix A in \[7\]) the following holds: if we condition on the process \( \mathcal{T}^* \) dying out, we obtain a subcritical Galton-Watson branching process \( \bar{T} \) where the number of children of each individual is distributed as a random variable \( \bar{X} \), where

\[
\frac{\bar{X}}{\epsilon} \sim \text{Bi} \left( \binom{n}{k-j}, \hat{p} \right)
\]

for some \( \hat{p} = \hat{p}(n) \) satisfying \( \left( \binom{n}{k-j} - 1 \right) \binom{n}{k-j} \hat{p} = 1 - \varepsilon \pm o(\varepsilon) \). Furthermore, we have \( \mathbb{E}(|\bar{T}|) = (1 \pm o(1))\varepsilon^{-1} \) and thus we obtain the upper bound

\[
\mathbb{P}(\mathcal{E}) \leq \mathbb{P}(\mathcal{T}^* \text{ survives}) + \mathbb{P} \left( |\mathcal{T}^*| \geq \lambda n^j \mid \mathcal{T}^* \text{ dies out} \right) \\
\leq \mathbb{P}(\mathcal{T}^* \text{ survives}) + (1 \pm o(1))(\varepsilon \lambda n^j)^{-1},
\]

by Markov’s Inequality. Note that the last term is \( o(\varepsilon) \).

We therefore need to calculate the survival probabilities of \( \mathcal{T}_* \) and \( \mathcal{T}^* \). We treat both cases in parallel by setting \( \mathcal{T} := \mathcal{T}_* \text{ or } \mathcal{T} := \mathcal{T}^* \), \( \zeta := 0 \) or \( \zeta := \varepsilon \), and \( \varepsilon' := \varepsilon - \zeta - \varepsilon \zeta = (1 - \zeta)(1 + \varepsilon) - 1 \).

It is slightly more convenient to consider the event \( \mathcal{D} \) of the process \( \mathcal{T} \) dying out and calculate its probability \( \mathbb{P}(\mathcal{D}) \). The process dies out if every subprocess starting at a child of the root also dies out. Recall that \( c_0 := \binom{n}{k} - 1 \). Because of the recursive nature of the tree, we have

\[
\mathbb{P}(\mathcal{D}) = \sum_{i=0}^{\infty} \mathbb{P} \left( \text{Bi} \left( (1 - \zeta) \binom{n}{k-j}, p \right) = i \right) \mathbb{P}(\mathcal{D})^{(i)-1} \}
\]

\[
= \sum_{i=0}^{\infty} \left( (1 - \zeta) \binom{n}{k-j} \right) \mathbb{P}(\mathcal{D})^{c_0 + 1 - p (1 - (1 - \zeta) \binom{n}{k-j})} \\
= (p \mathbb{P}(\mathcal{D})^{c_0} + 1 - p) (1 - \zeta) \binom{n}{k-j} = (1 - p (1 - \mathbb{P}(\mathcal{D})^{c_0}) (1 - \zeta) \binom{n}{k-j}).
\]

We set \( x := 1 - \mathbb{P}(\mathcal{D})^{c_0} \) and \( y := (1 - \zeta) \binom{n}{k-j} \) and note that \( p = (1 + \varepsilon) \hat{p}_g = (1 + \varepsilon) c_0^{-1} (1 - \zeta) \binom{n}{k-j} = (1 + \varepsilon') y \). Hence, we obtain

\[
1 - x = (1 - (1 + \varepsilon') y) x^{1/y},
\]

and furthermore solving for \( \varepsilon' \) yields

\[
\varepsilon' = \frac{1 - xy - (1 - x)^y}{xy} = \frac{y(1 - y) x^2 + (1 - y) 2y x^3 + \ldots}{xy}.
\]
implying
\[ \varepsilon' = \frac{x}{2} + O(x^2). \]

In other words, since \( \varepsilon' = (1 \pm o(1))\varepsilon \), we have
\[ 1 - P(D) = 1 - (1 - x)^{1/c_0} = (1 \pm o(1)) \frac{2\varepsilon}{c_0}, \]
as claimed. \qed