The asymptotic number of connected d-uniform hypergraphs*

Michael Behrisch[1 **], Amin Coja-Oghlan[2 ***], and Mihyun Kang[3 †]

[1] Institute of Transportation Systems, German Aerospace Center, Rutherfordstrasse 2, 12489 Berlin, Germany
michael.behrisch@dlr.de
[2] Goethe University, Mathematics Institute, 60054 Frankfurt am Main, Germany
acoghlan@math.uni-frankfurt.de
[3] TU Graz, Institut für Optimierung und Diskrete Mathematik, Steyrergasse 30, 8010 Graz, Austria
kang@math.tugraz.at

Abstract. For $d \geq 2$, let $H_d(n,p)$ denote a random d-uniform hypergraph with n vertices in which each of the $\binom{n}{d}$ possible edges is present with probability $p = p(n)$ independently, and let $H_d(n,m)$ denote a uniformly distributed d-uniform hypergraph with n vertices and m edges. Let either $H = H_d(n,m)$ or $H = H_d(n,p)$, where m/n and $\binom{n-1}{d-1}p$ need to be bounded away from $(d-1)^{-1}$ and 0 respectively. We determine the asymptotic probability that H is connected. This yields the asymptotic number of connected d-uniform hypergraphs with given numbers of vertices and edges. We also derive a local limit theorem for the number of edges in $H_d(n,p)$, conditioned on $H_d(n,p)$ being connected.

2010 *Mathematics subject classification*: Primary 05C80. Secondary 05C65.

1 Introduction and Main Results

1.1 Phase transition and connectivity

A *d-uniform hypergraph* $H = (V,E)$ is a pair of a set $V = V(H)$ of vertices and a set $E = E(H)$ of edges $e \subset V(H)$ with $|e| = d$. The *order* of H is the number of vertices of H, and the *size* of H is the number of edges. A 2-uniform hypergraph is just a graph. We say that a vertex $v \in V(H)$ is *reachable* from $w \in V(H)$ if there exist edges $e_1, \ldots, e_k \in E(H)$ such that $v \in e_1$, $w \in e_k$ and $e_i \cap e_{i+1} \neq \emptyset$ for all $1 \leq i < k$. Reachability is an equivalence relation, and the equivalence classes are called the *components* of H. If H has only a single component, then H is *connected*. We let $\mathcal{N}(H)$ signify the maximum order (i.e., number of vertices) of a component of H. For all hypergraphs H that we deal with the vertex set $V(H)$ will consist of integers. Therefore, the subsets of $V(H)$ can be ordered lexicographically, and we call the lexicographically first component of H that has order $\mathcal{N}(H)$ the *largest component* of H. In addition, we denote by $\mathcal{M}(H)$ the size (i.e., number of edges) of the largest component.

In this paper we consider two models of random d-uniform hypergraphs for $d \geq 2$. The random hypergraph $H_d(n,p)$ has the vertex set $V = \{1, \ldots, n\}$, and each of the $\binom{n}{d}$ possible edges is present with probability p independently. Moreover, $H_d(n,m)$ is a uniformly distributed d-uniform hypergraph with vertex set $V = \{1, \ldots, n\}$ and with exactly m edges. Finally, we say that the random hypergraph $H_d(n,p)$ satisfies a certain property $\mathcal{P}$ *with high probability* ("w.h.p.") if the probability that $\mathcal{P}$ holds in $H_d(n,p)$ tends to 1 as $n \to \infty$; a similar terminology is used for $H_d(n,m)$.

Since the pioneering work of Erdős and Rényi [9, 10] (see also [7, 12]), the component structure of random discrete objects (e.g., graphs, hypergraphs, digraphs, ...) has been among the main subjects of probabilistic combinatorics. Erdős and Rényi [10] studied (among other things) the component structure of *sparse* random graphs with $O(n)$ edges. The main result is that the order $\mathcal{N}(H_2(n,m))$ of the largest component undergoes a *phase transition* as $2m/n \sim 1$. Let us state a more general version from Schmidt-Pruzan and Shamir [17] for $d \geq 2$. Let either $H = H_d(n,m)$ and $c = dm/n$, or $H = H_d(n,p)$ and $c = \binom{n-1}{d-1}p$; we refer to c as the *average degree* of H. Then the result is the following.

* An extended abstract version of this work appeared in the Proceedings of RANDOM 2007, Springer LNCS 4627, 341–352.
** Supported by the DFG research center MATHEON in Berlin.
*** Supported by DFG CO 646
† Supported by the Deutsche Forschungsgemeinschaft (KA 2748/3-1)

(i) If $c < (d-1)^{-1} - \varepsilon$ for an arbitrarily small but fixed $\varepsilon > 0$, then $\mathcal{N}(H) = O(\ln n)$ w.h.p.
(ii) By contrast, if $c > (d-1)^{-1} + \varepsilon$, then H contains a unique component of order $\Omega(n)$ w.h.p., which is called the *giant component*. More precisely, $\mathcal{N}(H) = (1-\rho)n + o(n)$ w.h.p. where ρ is the unique solution to the transcendental equation

$$\rho = \exp(c(\rho^{d-1} - 1)) \tag{1}$$

that lies strictly between 0 and 1. Furthermore, the second largest component has order $O(\ln n)$ w.h.p.

Using probabilistic techniques, we derived in [3] a local limit theorem for $\mathcal{N}(H_d(n,p))$ and in [4] local limit theorems for the joint distribution of $\mathcal{N}(H)$ and $\mathcal{M}(H)$ for $H = H_d(n,m)$, or $H = H_d(n,p)$ in the regime $(d-1)\binom{n-1}{d-1}p > 1+\varepsilon$, resp. $d(d-1)m/n > 1+\varepsilon$, where $\varepsilon > 0$ is arbitrarily small but fixed as $n \to \infty$. Using these results, we determine in this paper the asymptotic probability that H is connected and derive a local limit theorem for the number of edges in $H_d(n,p)$, conditioned on $H_d(n,p)$ being connected.

These problems have been studied by a few authors. For $d = 2$, the asymptotic probability that $H_2(n,p)$ is connected was first computed by Stepanov [18]. Bender, Canfield, and McKay [5] were the first to compute the asymptotic probability that a random graph $H_2(n,m)$ is connected for *any* ratio m/n. Additionally, using their formula for the probability of $H_2(n,m)$ being connected, Bender, Canfield, and McKay [6] inferred the probability that $H_2(n,p)$ is connected as well as a central limit theorem for the number of edges of $H_2(n,p)$ given that $H_2(n,p)$ is connected. Using enumerative arguments, Pittel and Wormald [16] derived an improved version of the main result of [5] and obtained a local limit theorem that in addition to $\mathcal{N}(H)$ and $\mathcal{M}(H)$ also includes the order and size of the 2-core. O'Connell [15] employed the theory of large deviations in order to estimate the probability that $H_2(n,p)$ is connected up to a factor $\exp(o(n))$. While this result is significantly less precise than Stepanov's, O'Connell's proof is simpler. In addition, van der Hofstad and Spencer [11] used a novel perspective on the branching process argument to rederive the formula of Bender, Canfield, and McKay [5] for the number of connected graphs.

In contrast to the case of graphs ($d = 2$), little is known about the connectivity probability of random d-uniform hypergraphs with $d > 2$. Karoński and Łuczak [13] derived an asymptotic formula for the number of connected d-uniform hypergraphs of order n and size $m = \frac{n}{d-1} + o(\ln n / \ln\ln n)$ via combinatorial techniques. Since the minimum number of edges necessary for connectedness is $\frac{n-1}{d-1}$, this formula addresses *sparsely* connected hypergraphs. Furthermore, Andriamampianina and Ravelomanana [1] extended the result from [13] to the regime $m = \frac{n}{d-1} + o(n^{1/3})$ via enumerative techniques. By contrast, the results of this paper concern connected hypergraphs with $m = \frac{n}{d-1} + \Omega(n)$ edges. Thus, our results and those from [1, 13] are complementary.

1.2 Main results

The probability of connectedness. The threshold for $H_d(n,m)$ being connected is $m \sim \frac{n}{d}\ln n$. Hence, for $m = O(n)$ the probability that $H_d(n,m)$ is connected is $o(1)$. In fact, this probability is exponentially small in n. The following theorem gives an asymptotic expression for this exponentially rare event.

Theorem 1. *Let $d \geq 2$ be a fixed integer. For any compact set $\mathcal{J} \subset (d(d-1)^{-1}, \infty)$ and for any $\delta > 0$ there exists $n_0 > 0$ such that the following holds. Let $m = m(n)$ be a sequence of integers such that $\zeta = \zeta(n) = dm/n \in \mathcal{J}$ for all n. There exists a unique number $0 < r = r(n) < 1$ such that*

$$r = \exp\left(-\zeta \cdot \frac{(1-r)(1-r^{d-1})}{1-r^d}\right). \tag{2}$$

Let $\Phi_d(r,\zeta) = r^{\frac{r}{1-r}}(1-r)^{1-\zeta}(1-r^d)^{\frac{\zeta}{d}}$ for $d \geq 2$. Furthermore, define, for $d > 2$,

$$R_d(n,m) = \frac{1 - r^d - (1-r)(d-1)\zeta r^{d-1}}{\sqrt{\left(1 - r^d + \zeta(d-1)(r - r^{d-1})\right)(1-r^d) - d\zeta r(1-r^{d-1})^2}} \cdot \exp\left(\frac{(d-1)\zeta(r - r^2 + r^{d-1} - 2r^d + r^{d+2})}{2(1-r^d)}\right) \cdot \Phi_d(r,\zeta)^n,$$

and for $d = 2$,

$$R_2(n,m) = \frac{1 + r - \zeta r}{\sqrt{(1+r)^2 - 2\zeta r}} \cdot \exp\left(\frac{\zeta r(2 - r - r^2 + \zeta)}{2(1+r)}\right) \cdot \Phi_2(r,\zeta)^n.$$

Finally, let $c_d(n,m)$ denote the probability that $H_d(n,m)$ is connected. Then for all $n > n_0$ we have

$$(1-\delta)R_d(n,m) < c_d(n,m) < (1+\delta)R_d(n,m).$$

Observe that Theorem 1 yields an asymptotic formula for the number $C_d(n,m)$ of connected d-uniform hypergraphs of given order n and size m, because

$$C_d(n,m) = \binom{\binom{n}{d}}{m} c_d(n,m).$$

To prove Theorem 1 we shall consider a "larger" hypergraph $H_d(\nu,p)$ such that the expected order and size of the largest component of $H_d(\nu,p)$ are n and m. Then, we will infer the probability that $H_d(n,m)$ is connected from the local limit theorem for $\mathcal{N}(H_d(\nu,p))$ and $\mathcal{M}(H_d(\nu,p))$, which was proved by the authors in [4] (see below Lemma 6).

We also derive the following theorem on the asymptotic probability that $H_d(n,p)$ is connected, using results from [3, 8] (see below Lemmas 6 and 8).

Theorem 2. *Let $d \geq 2$ be a fixed integer. For any compact set $\mathcal{J} \subset (0,\infty)$, and for any $\delta > 0$ there exists $n_0 > 0$ such that the following holds. Let $p = p(n)$ be a sequence such that $\zeta = \zeta(n) = \binom{n-1}{d-1}p \in \mathcal{J}$ for all n. There exists a unique $0 < \varrho = \varrho(n) < 1$ such that*

$$\varrho = \exp\left(\zeta \cdot \frac{\varrho^{d-1} - 1}{(1-\varrho)^{d-1}}\right). \tag{3}$$

Let $\Psi_d(\varrho,\zeta) = (1-\varrho)\varrho^{\frac{\varrho}{1-\varrho}} \exp\left(\frac{\zeta}{d} \cdot \frac{1-\varrho^d-(1-\varrho)^d}{(1-\varrho)^d}\right)$ for $d \geq 2$. Define, for $d > 2$,

$$S_d(n,p) = \frac{1 - \zeta(d-1)\left(\frac{\varrho}{1-\varrho}\right)^{d-1}}{\sqrt{1 + \zeta(d-1)\frac{\varrho - \varrho^{d-1}}{(1-\varrho)^d}}} \cdot \exp\left(\frac{\zeta(d-1)\varrho(1-\varrho^d-(1-\varrho)^d)}{2(1-\varrho)^d}\right)$$
$$\cdot \exp\left(\frac{\zeta(d-1)\varrho}{2}\left(\left(\frac{\varrho}{1-\varrho}\right)^{d-2} + 1\right)\right) \cdot \Psi_d(\varrho,\zeta)^n,$$

and for $d = 2$,

$$S_2(n,p) = \left(1 - \frac{\zeta}{e^\zeta - 1}\right) \cdot \exp\left(\frac{\zeta(2+\zeta)}{2(e^\zeta - 1)}\right) \cdot (1 - e^{-\zeta})^n.$$

Finally, let $c_d(n,p)$ denote the probability that $H_d(n,p)$ is connected. Then for all $n > n_0$ we have

$$(1-\delta)S_d(n,p) < c_d(n,p) < (1+\delta)S_d(n,p).$$

Remark 3. The formulas for $R_d(n,m)$ and $S_d(n,p)$ for $d \geq 2$ given in an extended abstract version [2] of this work were incorrect.

The distribution of the number of edges in $H_d(n,p)$ given connectedness. Interestingly, if we choose $p = p(n)$ and $m = m(n)$ in such a way that $\binom{n}{d}p = m$ for each n and set $\zeta = \binom{n-1}{d-1}p = dm/n$, then the function $\Psi_d(\varrho,\zeta)$ from Theorem 2 is strictly bigger than $\Phi_d(r,\zeta)$ from Theorem 1. Consequently, the probability that $H_d(n,p)$ is connected exceeds the probability that $H_d(n,m)$ is connected by an exponential factor.

The reason for this is as follows. We can think of generating $H_d(n,p)$ as first choosing a random number m_0 of edges from the binomial distribution $\mathrm{Bin}(\binom{n}{d},p)$, and then generating a random hypergraph $H_d(n,m_0)$. The probability that $H_d(n,m_0)$ is connected increases rapidly as a function of m_0. Hence, $H_d(n,p)$ could "boost" its probability of being connected by including a number of edges that exceeds the expectation $\binom{n}{d}p$ of the binomial distribution considerably. Hence, once we *condition* on $H_d(n,p)$ being connected, the total number of edges in $H_d(n,p)$ will be significantly bigger than $\binom{n}{d}p$. The following local limit theorem quantifies this phenomenon.

Theorem 4. *Let $d \geq 2$ be a fixed integer. For any two compact sets $\mathcal{I} \subset \mathbf{R}$, $\mathcal{J} \subset (0,\infty)$, and for any $\delta > 0$ there exists $n_0 > 0$ such that the following holds. Suppose that $0 < p = p(n) < 1$ is a sequence such that $\zeta = \zeta(n) = \binom{n-1}{d-1}p \in \mathcal{J}$ for all n. Let $0 < \varrho = \varrho(n) < 1$ be the unique solution to (3), and set*

$$\hat{\mu} = \left\lceil \frac{\zeta(1-\varrho^d)}{d(1-\varrho)^d} \cdot n \right\rceil, \quad \hat{\sigma}^2 = \frac{\zeta}{d(1-\varrho)^d}\left(1-\varrho^d - \frac{\zeta d\varrho(1-\varrho^{d-1})^2}{(1-\varrho)^d + \zeta(d-1)(\varrho-\varrho^{d-1})}\right)\cdot n.$$

Finally, let $|E(H_d(n,p))|$ denote the number of edges in $H_d(n,p)$. Then for all $n \geq n_0$ and all integers y such that $n^{-\frac{1}{2}}y \in \mathcal{I}$ we have

$$\begin{aligned}\frac{1-\delta}{\sqrt{2\pi}\hat{\sigma}}\exp\left(-\frac{y^2}{2\hat{\sigma}^2}\right) &\leq \mathbb{P}\left[\,|E(H_d(n,p))| = \hat{\mu}+y \mid H_d(n,p) \text{ is connected}\,\right] \\ &\leq \frac{1+\delta}{\sqrt{2\pi}\hat{\sigma}}\exp\left(-\frac{y^2}{2\hat{\sigma}^2}\right).\end{aligned}$$

In the case $d = 2$ the solution to (3) is $\varrho = \exp(-\zeta)$, whence the formulas from Theorem 4 simplify to

$$\hat{\mu} = \left\lceil \frac{\zeta}{2}\coth(\zeta/2)\cdot n\right\rceil \quad \text{and} \quad \hat{\sigma}^2 = \frac{\zeta}{2}\cdot\frac{1-2\zeta\exp(-\zeta)-\exp(-2\zeta)}{(1-\exp(-\zeta))^2}\cdot n.$$

1.3 Techniques and Outline

In Section 2 we derive Theorem 1 from Lemma 6. The basic reason why this is possible is that *given* that the largest component of $H_d(\nu,p)$ has order n and size m (for suitably chosen $\nu > n$), the largest component is a uniformly distributed connected hypergraph with these parameters. This observation was also exploited by Łuczak [14] to estimate the number of connected graphs up to a polynomial factor, and in [8], where an explicit relation between the numbers $c_d(n,m)$ and $\mathbb{P}\left[\mathcal{N}(H_d(\nu,p)) = n,\ \mathcal{M}(H_d(\nu,p)) = m\right]$ was derived (see Lemma 5 below). Combining this relation with Lemma 6, we obtain Theorem 1. Finally, in Sections 3 and 4 we use similar arguments to establish Theorems 2 and 4.

1.4 Notation

We use the "O-notation" to express asymptotic estimates as $n \to \infty$. Occasionally we will apply this notation to expressions that do not only depend on n, but also on further parameters. Suppose that $f(x_1,\ldots,x_k,n)$, $g(x_1,\ldots,x_k,n)$ are functions of n and further parameters x_i are from domains $D_i \subset \mathbf{R}$ ($1 \leq i \leq k$), and that $g \geq 0$. Then we say that the estimate $f(x_1,\ldots,x_k,n) = O(g(x_1,\ldots,x_k,n))$ holds *uniformly in* $x_1,\ldots,x_k$ if the following is true: there exist numbers C and n_0 such that

$$|f(x_1,\ldots,x_k,n)| \leq Cg(x_1,\ldots,x_k,n) \text{ for all } n \geq n_0 \text{ and } (x_1,\ldots,x_k) \in \textstyle\prod_{j=1}^k D_j.$$

Similarly, we say that $f(x_1,\ldots,x_k,n) \sim g(x_1,\ldots,x_k,n)$ holds *uniformly in* $x_1,\ldots,x_k$ if for any $\varepsilon > 0$ there exists $n_0 > 0$ such that for all $n > n_0$

$$\sup_{(x_1,\ldots,x_k)\in D_1\times\cdots\times D_k}\left|\frac{f(x_1,\ldots,x_k,n)}{g(x_1,\ldots,x_k,n)} - 1\right| < \varepsilon.$$

We define uniformity analogously for the other Landau symbols Ω, Θ, etc.

2 The Probability that $H_d(n, m)$ is Connected: Proof of Theorem 1

We will derive the probability that $H_d(n, m)$ is connected (Theorem 1) from the local limit theorem for the joint distribution of the order and size of the largest component in $H_d(\nu, p)$, for suitably chosen $\nu > n$. The latter was proved by us in [3] and restated below in Lemma 6.

Let $\mathcal{J} \subset (d(d-1)^{-1}, \infty)$ be a compact interval, and let $m(n)$ be a sequence of integers such that $\zeta = \zeta(n) = dm/n \in \mathcal{J}$ for all n. The basic idea is to choose ν and p in such a way that $|n - \mathbb{E}(\mathcal{N}(H_d(\nu, p)))|$ and $|m - \mathbb{E}(\mathcal{M}(H_d(\nu, p)))|$ are "small", i.e., n and m will be "probable" outcomes of $\mathcal{N}(H_d(\nu, p))$ and $\mathcal{M}(H_d(\nu, p))$. Since given that $\mathcal{N}(H_d(\nu, p)) = n$ and $\mathcal{M}(H_d(\nu, p)) = m$, the largest component of $H_d(\nu, p)$ is a uniformly distributed connected graph of order n and size m, we can then express the probability that $H_d(n, m)$ is connected in terms of the probability

$$\chi = \mathbb{P}\left[\mathcal{N}(H_d(\nu, p)) = n,\ \mathcal{M}(H_d(\nu, p)) = m\right].$$

The (somewhat technical) details of this approach were carried out in [8], where the following lemma was established.

Lemma 5. *Suppose that $n > n_0$ for some large enough number $n_0 = n_0(\mathcal{J})$. Then there exist an integer $\nu = \nu(n) = \Theta(n)$ and a number $0 < p = p(n) < 1$ such that the following is true.*

(i) Let $c = \binom{\nu-1}{d-1} p$. Then $(d-1)^{-1} < c = O(1)$, and letting $0 < \rho = \rho(c) < 1$ signify the solution to (1), we have

$$n = (1-\rho)\nu, \quad \left|m - (1-\rho^d)\binom{\nu}{d} p\right| = O(1).$$

(ii) The solution r to (2) satisfies $|r - \rho| = o(1)$ and $|c - \frac{1-r}{1-r^d}\zeta| = o(1)$.

(iii) Furthermore,

$$c_d(n, m) \sim \nu \cdot \chi \cdot uvw \cdot \Phi_d(r, \zeta)^n \tag{4}$$

uniformly for $\zeta \in \mathcal{J}$, where

$$\Phi_d(r, \zeta) = (1-r)^{1-\zeta}\, r^{r/(1-r)}\, \left(1 - r^d\right)^{\zeta/d}, \tag{5}$$

$$u = 2\pi\sqrt{r(1-r)(1-r^d)c/d}, \tag{6}$$

$$v = \exp\left(\frac{(d-1)rc}{2}\left(1 - r^d + (1-r)r^{d-2}\right)\right), \qquad \text{and} \tag{7}$$

$$w = \begin{cases} 1 & \text{if } d > 2, \\ \exp\left(\frac{c^2 r(1+r)}{2}\right) & \text{if } d = 2. \end{cases} \tag{8}$$

The formulas (4)–(8) are reformulated from the corresponding ones in [8] by translating the notations as follows. We exchange the roles of ν and n and those of μ and m respectively; r and ρ play the same role as $1 - a_1$ and $1 - a_5$ respectively. The formula (5) follows from the term $(a_5(1-a_5)^{(1-a_5)/a_5})^\nu (a_5^{-d} b_5)^\mu = (a_5^{1-\zeta}(1-a_5)^{(1-a_5)/a_5}(1-(1-a_5)^d)^{\zeta/d})^\nu$ in (15) of [8]. Letting $\Phi_d(x, \zeta) := (1-x)^{1-\zeta} x^{\frac{x}{1-x}} (1-x^d)^{\frac{\zeta}{d}}$, we have from Lemma 12 of [8] that $\Phi_d(1-a_5, \zeta)^\nu \sim \Phi_d(1-a_1, \zeta)^\nu$, so we have in the current setting that $\Phi_d(\rho, \zeta)^n \sim \Phi_d(r, \zeta)^n$. Furthermore, (6) follows from the term $\frac{2\pi}{n}\sqrt{a_5(1-a_5)b_5 nm} \sim u$ in (15) of [8]; (7) from the term $\exp\left[\frac{1}{2}(d-1)(1-a_5)c(b_5 + a_5(1-a_5)^{d-2})\right] \sim v$; and (8) from the term $\exp\left[\frac{b_5 mp(1-a_5^d-(1-a_5)^d)}{2a_5^d}\right] \sim w$.

Thus, once we know the explicit expression for

$$\chi = \mathbb{P}\left[\mathcal{N}(H_d(\nu, p)) = n,\ \mathcal{M}(H_d(\nu, p)) = m\right],$$

we can derive the exact asymptotic expression for $c_d(n, m)$ from (4). We can in fact compute χ explicitly using the following local limit theorem for the joint distribution of $\mathcal{N}(H_d(\nu, p))$ and $\mathcal{M}(H_d(\nu, p))$ from [4].

Lemma 6. *Let $d \geq 2$ be a fixed integer. For any two compact sets $\mathcal{I} \subset \mathbf{R}^2$, $\mathcal{J} \subset ((d-1)^{-1}, \infty)$, and for any $\delta > 0$ there exists $\nu_0 > 0$ such that the following holds. Let $p = p(\nu)$ be a sequence such that $c = c(\nu) = \binom{\nu-1}{d-1} p \in \mathcal{J}$ for all ν and let $0 < \rho = \rho(\nu) < 1$ be the unique solution to (1). Further, let*

$$\sigma_{\mathcal{N}}^2 = \frac{\rho\left(1-\rho+c(d-1)(\rho-\rho^{d-1})\right)}{(1-c(d-1)\rho^{d-1})^2} \cdot \nu, \tag{9}$$

$$\sigma_{\mathcal{M}}^2 = c^2\rho^d \cdot \frac{2+c(d-1)(\rho^{2d-2}-2\rho^{d-1}+\rho^d)-\rho^{d-1}-\rho^d}{(1-c(d-1)\rho^{d-1})^2} \cdot \nu + (1-\rho^d)\frac{c}{d} \cdot \nu, \tag{10}$$

$$\sigma_{\mathcal{N}\mathcal{M}} = c\rho \cdot \frac{1-\rho^d-c(d-1)\rho^{d-1}(1-\rho)}{(1-c(d-1)\rho^{d-1})^2} \cdot \nu. \tag{11}$$

Suppose that $\nu \geq \nu_0$ and that n, m are integers such that

$$x = n-(1-\rho)\nu \quad \text{and} \quad y = m-(1-\rho^d)\binom{\nu}{d}p \tag{12}$$

satisfy $\nu^{-\frac{1}{2}}(x,y) \in \mathcal{I}$. Define

$$P(x,y) = \frac{1}{2\pi\sqrt{\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{N}\mathcal{M}}^2}} \cdot \exp\left(-\frac{\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2}{2(\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{N}\mathcal{M}}^2)}\left(\frac{x^2}{\sigma_{\mathcal{N}}^2}-\frac{2\sigma_{\mathcal{N}\mathcal{M}}xy}{\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2}+\frac{y^2}{\sigma_{\mathcal{M}}^2}\right)\right). \tag{13}$$

Then we have

$$(1-\delta)P(x,y) \leq \mathbb{P}\left[\mathcal{N}(H_d(\nu,p)) = n,\ \mathcal{M}(H_d(\nu,p)) = m\right] \leq (1+\delta)P(x,y). \tag{14}$$

Note that from (9)–(11) we have

$$\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{N}\mathcal{M}}^2 = \frac{\frac{c\rho}{d}\left(1-\rho+c(d-1)(\rho-\rho^{d-1})\right)(1-\rho^d)-c^2\rho^2(1-\rho^{d-1})^2}{(1-c(d-1)\rho^{d-1})^2} \cdot \nu^2. \tag{15}$$

From Lemma 5 (i) and (12), $x = 0, y = O(1)$, and from (10) $\sigma_{\mathcal{M}} = \Theta(\nu)$. Thus (13)–(15) yield

$$\begin{aligned}\chi &= \mathbb{P}\left[\mathcal{N}(H_d(\nu,p)) = n,\ \mathcal{M}(H_d(\nu,p)) = m\right] \\ &\sim \frac{1}{2\pi\sqrt{\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{N}\mathcal{M}}^2}} \\ &= \frac{1-c(d-1)\rho^{d-1}}{2\pi\nu\sqrt{\frac{c\rho}{d}\left(1-\rho+c(d-1)(\rho-\rho^{d-1})\right)(1-\rho^d)-c^2\rho^2(1-\rho^{d-1})^2}}.\end{aligned} \tag{16}$$

Since $r \sim \rho$ and $c \sim \frac{1-r}{1-r^d}\zeta$ by Lemma 5 (ii), we can express $\nu \cdot \chi, u, v, w$ in (16) and (6)–(8) solely in terms of r and ζ:

$$\begin{aligned}\nu \cdot \chi &\sim \frac{1-\frac{1-r}{1-r^d}\zeta(d-1)r^{d-1}}{2\pi\sqrt{\frac{1-r}{1-r^d}\zeta\frac{r}{d}\left(1-r+\frac{1-r}{1-r^d}\zeta(d-1)(r-r^{d-1})\right)(1-r^d)-\left(\frac{1-r}{1-r^d}\zeta\right)^2 r^2(1-r^{d-1})^2}} \\ &= \frac{1-\frac{1-r}{1-r^d}\zeta(d-1)r^{d-1}}{2\pi\sqrt{\frac{(1-r)^2}{1-r^d}\frac{\zeta r}{d}\left(1-r^d+\zeta(d-1)(r-r^{d-1})\right)-\left(\frac{1-r}{1-r^d}\right)^2\zeta^2r^2(1-r^{d-1})^2}}\end{aligned}$$

$$= \frac{1 - \frac{1-r}{1-r^d}\zeta(d-1)r^{d-1}}{2\pi\sqrt{\frac{\zeta r}{d}\left(\frac{1-r}{1-r^d}\right)^2\left(\left(1-r^d+\zeta(d-1)(r-r^{d-1})\right)(1-r^d) - d\zeta r(1-r^{d-1})^2\right)}}$$

$$= \frac{1-r^d-(1-r)\zeta(d-1)r^{d-1}}{2\pi\sqrt{\frac{\zeta r}{d}(1-r)^2\left(\left(1-r^d+\zeta(d-1)(r-r^{d-1})\right)(1-r^d) - d\zeta r(1-r^{d-1})^2\right)}},$$

$$u \sim 2\pi\sqrt{r(1-r)(1-r^d)\frac{1-r}{1-r^d}\zeta\frac{1}{d}} \;=\; 2\pi\sqrt{\frac{\zeta r}{d}}\cdot(1-r),$$

$$v \sim \exp\left(\frac{(d-1)r}{2}\frac{1-r}{1-r^d}\zeta\left(1-r^d+(1-r)r^{d-2}\right)\right)$$

$$= \exp\left(\frac{\zeta(d-1)(r-r^2+r^{d-1}-2r^d+r^{d+2})}{2(1-r^d)}\right), \qquad \text{and}$$

$$w \sim \begin{cases} 1 & \text{if } d>2,\\ \exp\left(\frac{(1-r)^2\zeta^2 r(1+r)}{2(1-r^2)^2}\right) \;=\; \exp\left(\frac{\zeta^2 r}{2(1+r)}\right) & \text{if } d=2.\end{cases}$$

Putting these together, we obtain for $d>2$,

$$\nu\cdot\chi\cdot uvw \sim \frac{1-r^d-(1-r)\zeta(d-1)r^{d-1}}{\sqrt{\left(1-r^d+\zeta(d-1)(r-r^{d-1})\right)(1-r^d)-d\zeta r(1-r^{d-1})^2}} \cdot\exp\left(\frac{\zeta(d-1)(r-r^2+r^{d-1}-2r^d+r^{d+2})}{2(1-r^d)}\right), \tag{17}$$

and for $d=2$,

$$\nu\cdot\chi\cdot uvw \sim \frac{1+r-\zeta r}{\sqrt{(1+r)^2-2\zeta r}}\cdot\exp\left(\frac{\zeta r(2-r-r^2+\zeta)}{2(1+r)}\right). \tag{18}$$

Thus, (4), (17) and (18) imply the desired result.

Remark 7. While Lemma 5 was established in Coja-Oghlan, Moore, and Sanwalani [8], the exact joint limiting distribution of $\mathcal{N}(H_d(\nu,p))$ and $\mathcal{M}(H_d(\nu,p))$ (i.e. Lemma 6) was not known at that point. Therefore, Coja-Oghlan, Moore, and Sanwalani could only compute the $c_d(n,m)$ up to a constant factor. By contrast, combining Lemma 6 with Lemma 5, here we have obtained *tight* asymptotics for $c_d(n,m)$.

3 The Probability that $H_d(\nu,p)$ is Connected: Proof of Theorem 2

Let $\mathcal{J}\subset(0,\infty)$ be a compact set, and let $0<p=p(n)<1$ be a sequence such that $\zeta=\zeta(n)=\binom{n-1}{d-1}p\in\mathcal{J}$ for all n. All asymptotics in this section are uniform in ζ.

To compute the probability $c_d(n,p)$ that a random hypergraph $H_d(n,p)$ is connected, we will establish that

$$\mathbb{P}\left[\mathcal{N}(H_d(\nu,p))=n\right] \sim \binom{\nu}{n}c_d(n,p)(1-p)^{\binom{\nu}{d}-\binom{\nu-n}{d}-\binom{n}{d}} \tag{19}$$

for a suitably chosen integer $\nu>n$. Then, we will employ the local limit theorem for $\mathcal{N}(H_d(\nu,p))$, which is implied by Lemma 6 and as well as our previous result [3] on the local limit theorem for $\mathcal{N}(H_d(n,p))$, to compute the l.h.s. of (19), so that we can just solve (19) for $c_d(n,p)$.

In order to carry this out, we use the following lemma on the component structure of $H_d(\nu,p)$, which is a slight variant of Theorem 5 of [8]. To obtain it, we can easily adapt the argements of the proof of Theorem 5 of [8]. We may skip here the details, as the computations become quite technical and tedious without providing useful new insights.

Lemma 8. *Let $c = c(\nu)$ be a sequence of non-negative reals and let $p = c\binom{\nu-1}{d-1}^{-1}$ and $m = \binom{\nu}{d}p = c\nu/d$. Then for both $H = H_d(\nu, p)$ and $H = H_d(\nu, \mu)$ the following holds.*

(i) For any $c_0 < (d-1)^{-1}$ there is a number ν_0 such that for all $\nu > \nu_0$ for which $c = c(\nu) \le c_0$ we have

$$\mathbb{P}\left[\mathcal{N}(H) \le 300(d-1)^2(1-(d-1)c_0)^{-2}\ln\nu\right] \ge 1 - \nu^{-100}.$$

(ii) For any $c_0 > (d-1)^{-1}$ there are numbers $\nu_0 > 0$, $0 < c_0' < (d-1)^{-1}$ such that for all $\nu > \nu_0$ for which $c_0 \le c = c(\nu) < \ln\nu/\ln\ln\nu$ the following holds. The transcendental equation (1) has a unique solution $0 < \rho = \rho(\nu) < 1$, which satisfies

$$\rho^{d-1}c < c_0'.$$

Furthermore, with probability $\ge 1-\nu^{-100}$ there exists precisely one component of order $(1-\rho)\nu + o(\nu)$ in H, while all other components have order $\le \ln^2\nu$. In addition,

$$\mathbb{E}\left[\mathcal{N}(H)\right] = (1-\rho)\nu + o(\sqrt{\nu}).$$

We pick ν as follows. By Lemma 8 for each integer k such that $c(k) = \binom{k-1}{d-1}p > (d-1)^{-1}$ the transcendental equation $\rho(k) = \exp(c(k)(\rho(k)^{d-1} - 1))$ has a unique solution $\rho(k)$ that lies strictly between 0 and 1. We let $\nu = \max\{k \in \mathbb{N} : (1-\rho(k))k < n\}$. Moreover, set $\rho = \rho(\nu)$ and $c = c(\nu) = \binom{\nu-1}{d-1}p$, and let $0 < s < 1$ be such that $(1-s)\nu = n$. We claim

$$|n - (1-\rho)\nu| < O(1). \tag{20}$$

To see this, we observe that $(1-\rho(\nu))\nu < n = (1-s)\nu \le (1-\rho(\nu+1))(\nu+1)$. In order to establish (20), it suffices to show that $|\rho(\nu+1) - \rho(\nu)| = O(1/\nu)$, because $n - (1-\rho(\nu))\nu < (1-\rho(\nu+1))(\nu+1) - (1-\rho(\nu))\nu < 1 + \nu(\rho(\nu) - \rho(\nu+1))$. To prove this, we note that since $\zeta = \binom{n-1}{d-1}p = \binom{(1-s)\nu-1}{d-1}p$,

$$\begin{aligned} c(\nu+1) - c(\nu) &= \binom{\nu}{d-1}p - \binom{\nu-1}{d-1}p = p\binom{\nu-1}{d-1}\frac{d-1}{\nu-d+1} \\ &= \frac{\zeta\binom{\nu-1}{d-1}}{\binom{(1-s)\nu-1}{d-1}} \cdot \frac{(d-1)}{\nu-d+1} = O(1/\nu). \end{aligned}$$

This, together with Taylor series expansion, implies that $|\rho(\nu+1) - \rho(\nu)| = O(1/\nu)$, because $\rho(k) = \exp(c(k)(\rho(k)^{d-1} - 1))$ and $\rho(k)$ is differentiable due to the implicit function theorem.

To establish (19), note that the r.h.s. is just the expected number of components of order n in $H_d(\nu, p)$. For there are $\binom{\nu}{n}$ ways to choose the vertex set $\mathcal{C}$ of such a component, and the probability that $\mathcal{C}$ spans a connected hypergraph is $c_d(n, p)$. Moreover, if $\mathcal{C}$ is a component, then $H_d(\nu, p)$ features no edge that connects $\mathcal{C}$ with $V \setminus \mathcal{C}$, and there are $\binom{\nu}{d} - \binom{\nu-n}{d} - \binom{n}{d}$ possible edges of this type, each being present with probability p independently. Hence, we conclude that

$$\mathbb{P}\left[\mathcal{N}(H_d(\nu, p)) = n\right] \le \binom{\nu}{n}c_d(n, p)(1-p)^{\binom{\nu}{d} - \binom{\nu-n}{d} - \binom{n}{d}}. \tag{21}$$

On the other hand,

$$\mathbb{P}\left[\mathcal{N}(H_d(\nu, p)) = n\right] \ge \binom{\nu}{n}c_d(n, p)(1-p)^{\binom{\nu}{d} - \binom{\nu-n}{d} - \binom{n}{d}}\mathbb{P}\left[\mathcal{N}(H_d(\nu-n, p)) < n\right], \tag{22}$$

because the r.h.s. equals the probability that $H_d(\nu, p)$ has *exactly* one component of order n. Furthermore, as $|n - (1-\rho)\nu| < O(1)$ by (20), Lemma 8 entails that

$$\mathbb{P}\left[\mathcal{N}(H_d(\nu-n, p)) < n\right] \sim 1.$$

Hence, combining (21) and (22), we obtain (19).

To derive an explicit formula for $c_d(n, p)$ from (19), we need the following lemma.

Lemma 9. *(i) We have* $c = \zeta(1-s)^{1-d}\left(1+\binom{d}{2}\frac{s}{(1-s)\nu}+O(\nu^{-2})\right)$.

(ii) The transcendental equation (3) has a unique solution $0<\varrho<1$*, which satisfies* $|s-\varrho| = O(\nu^{-1})$.

(iii) Letting

$$\Psi(x) = \Psi_d(x,\zeta) := (1-x)x^{\frac{x}{1-x}}\exp\left(\frac{\zeta}{d}\cdot\frac{1-x^d-(1-x)^d}{(1-x)^d}\right),$$

we have $\Psi(\varrho)^n \sim \Psi(s)^n$.

Proof of Lemma 9. Regarding the first assertion, we note that

$$\frac{(1-s)^{d-1}\binom{\nu-1}{d-1}}{\binom{(1-s)\nu-1}{d-1}} = \prod_{j=1}^{d-1}\left(1+\frac{sj}{(1-s)\nu-j}\right) = 1+\binom{d}{2}\frac{s}{(1-s)\nu}+O(\nu^{-2}). \tag{23}$$

Since $c = \binom{\nu-1}{d-1}p = \zeta\frac{\binom{\nu-1}{d-1}}{\binom{n-1}{d-1}}$ and $n = (1-s)\nu$, (23) implies the first assertion.

In order to show the second assertion, we set

$$\varphi_z : (0,1)\to\mathbf{R},\ t\mapsto \exp\left(z\frac{t^{d-1}-1}{(1-t)^{d-1}}\right) \text{ for } z>0.$$

Then $\lim_{t\searrow 0}\varphi_z(t) = \exp(-z)>0$, while $\lim_{t\nearrow 1}\varphi_z(t)=0$. In addition, φ_z is convex for any $z>0$. Therefore, for each $z>0$ there is a unique $0<t_z<1$ such that $t_z=\varphi_z(t_z)$, whence (3) in Theorem 2 has the unique solution $0<\varrho=t_\zeta<1$. Moreover, letting $\zeta' = (1-\rho)^{d-1}c$, we have $\rho = t_{\zeta'}$. Thus, since $t\mapsto t_z$ is differentiable by the implicit function theorem and $|\zeta-\zeta'| = O(\nu^{-1})$ by the first assertion, we conclude that $|\varrho-\rho| = O(\nu^{-1})$. In addition, $|s-\rho| = O(\nu^{-1})$ by (20). Hence, $|s-\varrho| = O(\nu^{-1})$, as desired.

To establish the third assertion, we compute

$$\begin{aligned}\frac{\partial}{\partial x}\Psi(x) = (1-x)^{-d-1}x^{\frac{2x-1}{1-x}}\exp\left(\frac{\zeta}{d}\frac{1-x^d-(1-x)^d}{(1-x)^d}\right)&\\ \times\left(\zeta(1-x)(x-x^d)+(1-x)^d x\ln x\right).&\end{aligned} \tag{24}$$

As $\varrho = \exp\left(\zeta\frac{\varrho^{d-1}-1}{(1-\varrho)^{d-1}}\right)$, (24) entails that $\frac{\partial}{\partial x}\Psi(\varrho)=0$. Therefore, Taylor's formula yields that $\Psi(s)-\Psi(\varrho) = O(s-\varrho)^2 = O(\nu^{-2})$, because $s-\varrho = O(\nu^{-1})$ by the second assertion. Consequently, we obtain

$$\left(\frac{\Psi(s)}{\Psi(\varrho)}\right)^\nu = \left(1+\frac{\Psi(s)-\Psi(\varrho)}{\Psi(\varrho)}\right)^\nu \sim \exp\left(\nu\cdot\frac{\Psi(s)-\Psi(\varrho)}{\Psi(\varrho)}\right) = \exp(O(\nu^{-1}))\sim 1,$$

thereby completing the proof of the third assertion. □

Let us continue with the proof of Theorem 2. Note that Lemma 6 implies

$$\mathbb{P}\left[\mathcal{N}(H_d(\nu,p)) = n\right] \sim \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{N}}}\exp\left(-\frac{(n-(1-\rho)\nu)^2}{2\sigma_{\mathcal{N}}^2}\right). \tag{25}$$

It follows also from our previous result [3] on the local limit theorem for $\mathcal{N}(H_d(n,p))$. Since $|s-\rho| = O(\nu^{-1})$ by (20), we can express $\sigma_{\mathcal{N}}^2$ (in (9)) in terms of s:

$$\begin{aligned}\sigma_{\mathcal{N}}^2 &= \frac{\rho\left(1-\rho+c(d-1)(\rho-\rho^{d-1})\right)}{(1-c(d-1)\rho^{d-1})^2}\cdot\nu\\ &\sim \frac{s\left(1-s+c(d-1)(s-s^{d-1})\right)}{(1-c(d-1)s^{d-1})^2}\cdot\nu.\end{aligned} \tag{26}$$

Further, since $|n-(1-\rho)\nu| < O(1)$ by (20), we have from (25) and (26)

$$\mathbb{P}\left[\mathcal{N}(H_d(\nu,p)) = n\right] \sim (2\pi)^{-\frac{1}{2}}\left(\frac{s\left(1-s+c(d-1)(s-s^{d-1})\right)}{(1-c(d-1)s^{d-1})^2}\cdot\nu\right)^{-1/2}. \tag{27}$$

Via Stirling's formula and $n=(1-s)\nu$ we can estimate the binomial coefficient

$$\binom{\nu}{n} \sim \left(s^{s\nu}(1-s)^{(1-s)\nu}\sqrt{2\pi s(1-s)\nu}\right)^{-1}. \tag{28}$$

Plugging (27) and (28) into (19), we obtain

$$\begin{aligned} c_d(n,p) &\sim \binom{\nu}{n}^{-1} \cdot \mathbb{P}\left[\mathcal{N}(H_d(\nu,p))=n\right] \cdot (1-p)^{\binom{\nu-n}{d}+\binom{n}{d}-\binom{\nu}{d}} \\ &\sim s^{s\nu}(1-s)^{(1-s)\nu} \cdot \eta \cdot (1-p)^{\binom{\nu-n}{d}+\binom{n}{d}-\binom{\nu}{d}}, \end{aligned} \tag{29}$$

where

$$\eta = \left(\frac{(1-s)(1-c(d-1)s^{d-1})^2}{1-s+c(d-1)(s-s^{d-1})}\right)^{1/2}. \tag{30}$$

Let us consider the cases $d=2$ and $d>2$ separately, because $\binom{\nu}{d}p^2 = o(1)$ for $d>2$, while $\binom{\nu}{2}p^2 = \Theta(1)$ and therefore the asymptotics for $(1-p)^{\binom{\nu-n}{d}+\binom{n}{d}-\binom{\nu}{d}}$ behave quite differently.

1st case: $d=2$**.** Note first that $\binom{\nu-n}{2}+\binom{n}{2}-\binom{\nu}{2} = s(s-1)\nu^2$, because $n=(1-s)\nu$. Using $p=\frac{c}{\nu-1}$, we get

$$\begin{aligned} (1-p)^{\binom{\nu-n}{2}+\binom{n}{2}-\binom{\nu}{2}} &= (1-p)^{s(s-1)\nu^2} \\ &\sim \exp\left(-\left(p+\frac{p^2}{2}\right)s(s-1)\nu^2\right) \\ &\sim \exp\left(-\frac{c}{\nu-1}s(s-1)\left((\nu-1)(\nu+1)+1\right)-\frac{1}{2}\left(\frac{c}{\nu-1}\right)^2 s(s-1)\nu^2\right) \\ &\sim \exp\left(cs(1-s)(\nu+1)+\frac{c^2}{2}s(1-s)\right). \end{aligned} \tag{31}$$

Moreover, (30) simplifies to $\eta = 1-cs$. Hence, recalling that $\nu=(1-s)^{-1}n$ and using Lemma 9 (i)-(iii), i.e. $c=\frac{\zeta}{1-s}\left(1+\frac{s}{(1-s)\nu}+O(\nu^{-2})\right)$, $|s-\varrho| = O(\nu^{-1})$ and $\left((1-s)s^{\frac{s}{1-s}}\exp\left(\frac{\zeta s}{1-s}\right)\right)^n \sim \left((1-\varrho)\varrho^{\frac{\varrho}{1-\varrho}}\exp\left(\frac{\zeta\varrho}{1-\varrho}\right)\right)^n$, we can estimate (29) as

$$\begin{aligned} c_2(n,p) &\sim s^{s\nu}(1-s)^{(1-s)\nu} \cdot (1-cs) \cdot \exp\left(cs(1-s)\nu + cs(1-s) + \frac{c^2}{2}s(1-s)\right) \\ &\sim s^{\frac{sn}{1-s}}(1-s)^n\left(1-\frac{\zeta s}{1-s}\right)\exp\left(\frac{\zeta sn}{1-s}+\frac{\zeta s^2}{1-s}+\zeta s+\frac{\zeta^2 s}{2(1-s)}\right) \\ &= \left(s^{\frac{s}{1-s}}(1-s)\exp\left(\frac{\zeta s}{1-s}\right)\right)^n\left(1-\frac{\zeta s}{1-s}\right)\exp\left(\frac{\zeta s^2}{1-s}+\zeta s+\frac{\zeta^2 s}{2(1-s)}\right) \\ &\sim \left(\varrho^{\frac{\varrho}{1-\varrho}}(1-\varrho)\exp\left(\frac{\zeta\varrho}{1-\varrho}\right)\right)^n\left(1-\frac{\zeta\varrho}{1-\varrho}\right)\exp\left(\frac{\zeta\varrho^2}{1-\varrho}+\zeta\varrho+\frac{\zeta^2\varrho}{2(1-\varrho)}\right) \\ &= \left(\varrho\exp(\zeta)\right)^{\frac{\varrho n}{1-\varrho}}(1-\varrho)^n\left(1-\frac{\zeta\varrho}{1-\varrho}\right)\exp\left(\frac{\zeta(2+\zeta)\varrho}{2(1-\varrho)}\right). \end{aligned} \tag{32}$$

Finally, for $d=2$ the unique solution to (3) is just $\varrho=\exp(-\zeta)$, so we have $\frac{\varrho}{1-\varrho}=\frac{1}{e^\zeta-1}$. Plugging these into (32), we obtain

$$c_2(n,p) \sim (1-e^{-\zeta})^n\left(1-\frac{\zeta}{e^\zeta-1}\right)\exp\left(\frac{\zeta(2+\zeta)}{2(e^\zeta-1)}\right), \tag{33}$$

as desired.

2nd case: $d > 2$. For $0 < \alpha < 1$, using

$$\alpha^d \binom{\alpha\nu}{d}^{-1} \binom{\nu}{d} = \prod_{i=0}^{d-1} \frac{\alpha(\nu - i)}{\alpha\nu - i} = \prod_{i=0}^{d-1} \left(1 + \frac{(1-\alpha)i}{\alpha\nu - i}\right) = 1 + \frac{1-\alpha}{\alpha\nu}\binom{d}{2} + O(\nu^{-2}),$$

and $n = (1-s)\nu$, we estimate

$$\begin{aligned}
&\binom{n}{d}\binom{\nu}{d}^{-1} + \binom{\nu - n}{d}\binom{\nu}{d}^{-1} \\
&= \binom{(1-s)\nu}{d}\binom{\nu}{d}^{-1} + \binom{s\nu}{d}\binom{\nu}{d}^{-1} \\
&= (1-s)^d\left(1 - \frac{s}{(1-s)\nu}\binom{d}{2} + O(\nu^{-2})\right) + s^d\left(1 - \frac{1-s}{s\nu}\binom{d}{2} + O(\nu^{-2})\right) \\
&= (1-s)^d + s^d - \frac{1}{\nu}\binom{d}{2}\left(s(1-s)^{d-1} + (1-s)s^{d-1}\right) + O(\nu^{-2})
\end{aligned}$$

and thus we have

$$\begin{aligned}
&\binom{n}{d} + \binom{\nu - n}{d} - \binom{\nu}{d} \\
&= \binom{\nu}{d}\left((1-s)^d + s^d - 1\right) \\
&\quad - \binom{\nu}{d}\frac{1}{\nu}\binom{d}{2}\left(s(1-s)^{d-1} + (1-s)s^{d-1}\right) + O(\nu^{d-2}).
\end{aligned} \tag{34}$$

Because $\binom{\nu-1}{d-1}p = c = \Theta(1)$, we have $\binom{\nu}{d}p^2 = o(1)$ for $d > 2$, and hence

$$\begin{aligned}
(1-p)^{\binom{\nu}{d}\left((1-s)^d + s^d - 1\right)} &\sim \exp\left(-p\binom{\nu}{d}\left((1-s)^d + s^d - 1\right)\right) \\
&= \exp\left(\frac{c\nu}{d}\left(1 - s^d - (1-s)^d\right)\right)
\end{aligned} \tag{35}$$

and

$$\begin{aligned}
&(1-p)^{-\binom{\nu}{d}\frac{1}{\nu}\binom{d}{2}\left(s(1-s)^{d-1} + (1-s)s^{d-1}\right)} \\
&\sim \exp\left(p\binom{\nu}{d}\frac{1}{\nu}\binom{d}{2}\left(s(1-s)^{d-1} + (1-s)s^{d-1}\right)\right) \\
&= \exp\left(p\binom{\nu-1}{d-1}\frac{d-1}{2}\left(s(1-s)^{d-1} + (1-s)s^{d-1}\right)\right) \\
&\sim \exp\left(\frac{c(d-1)}{2}\left(s(1-s)^{d-1} + (1-s)s^{d-1}\right)\right).
\end{aligned} \tag{36}$$

Putting (34)–(36) together, we get

$$\begin{aligned}
&(1-p)^{\binom{n}{d} + \binom{\nu-n}{d} - \binom{\nu}{d}} \\
&\sim \exp\left(\frac{c\nu}{d}(1 - s^d - (1-s)^d) + \frac{c(d-1)}{2}((1-s)s^{d-1} + s(1-s)^{d-1})\right).
\end{aligned} \tag{37}$$

Before proceeding further computations toward the asymptotic estimation of $c_d(n,p)$, we note that taking $d = 2$ in the estimate (37) yields $(1-p)^{\binom{n}{2} + \binom{\nu-n}{2} - \binom{\nu}{2}} \sim \exp\left(cs(1-s)(\nu+1)\right)$, which differs by a factor $\exp(\frac{c^2}{2}s(1-s))$ from the estimate (31), the reason being that $\binom{\nu}{d}p^2 = o(1)$ for $d > 2$, while $\binom{\nu}{2}p^2 = \Theta(1)$. This in turn results in an extra factor $\exp(\frac{c^2}{2}\varrho(1-\varrho))$ in the estimate (32) of $c_2(n,p)$, in comparison to the estimate of $c_d(n,p)$ when taking $d = 2$ in (41).

We now return to the computation of (37). Using

$$c=\zeta(1-s)^{1-d}\left(1+\binom{d}{2}\frac{s}{(1-s)\nu}+O(\nu^{-2})\right)$$

by Lemma 9 (i) and recalling that $\nu=(1-s)^{-1}n$,

$$\frac{c\nu}{d}=\frac{\zeta n}{d(1-s)^d}+\frac{\zeta(d-1)s}{2(1-s)^d}+O(n^{-1}),$$

and thus

$$\begin{aligned}
&\frac{c\nu}{d}(1-s^d-(1-s)^d)+\frac{c(d-1)}{2}((1-s)s^{d-1}+s(1-s)^{d-1})\\
&=\frac{\zeta n}{d(1-s)^d}(1-s^d-(1-s)^d)+\frac{\zeta(d-1)s}{2(1-s)^d}(1-s^d-(1-s)^d)\\
&\quad+\frac{\zeta(1-s)^{1-d}(d-1)}{2}((1-s)s^{d-1}+s(1-s)^{d-1})+O(n^{-1})\\
&=\frac{\zeta n}{d(1-s)^d}(1-s^d-(1-s)^d)+\frac{\zeta(d-1)s}{2(1-s)^d}(1-s^d-(1-s)^d)\\
&\quad+\frac{\zeta(d-1)s}{2}\left(\left(\frac{s}{1-s}\right)^{d-2}+1\right)+O(n^{-1}).
\end{aligned}\tag{38}$$

Using this, we can restate (37) as

$$\begin{aligned}
&(1-p)^{\binom{n}{d}+\binom{\nu-n}{d}-\binom{\nu}{d}}\\
&\quad\sim\exp\left(\frac{\zeta\left(1-s^d-(1-s)^d\right)n}{d(1-s)^d}+\frac{\zeta(d-1)s(1-s^d-(1-s)^d)}{2(1-s)^d}\right)\\
&\qquad\cdot\exp\left(\frac{\zeta(d-1)s}{2}\left(\left(\frac{s}{1-s}\right)^{d-2}+1\right)\right).
\end{aligned}\tag{39}$$

Due to the same reasons, we estimate (30) as

$$\begin{aligned}
\eta&=\left(\frac{(1-s)(1-c(d-1)s^{d-1})^2}{1-s+c(d-1)(s-s^{d-1})}\right)^{1/2}\\
&=(1-c(d-1)s^{d-1})\left(1+c(d-1)(1-s)^{-1}(s-s^{d-1})\right)^{-1/2}\\
&=\left(1-\zeta(d-1)\left(\frac{s}{1-s}\right)^{d-1}+O(n^{-1})\right)\\
&\qquad\cdot\left(1+\frac{\zeta(d-1)(s-s^{d-1})}{(1-s)^d}+O(n^{-1})\right)^{-1/2}\\
&=\left(1-\zeta(d-1)\left(\frac{s}{1-s}\right)^{d-1}\right)\\
&\qquad\cdot\left(1+\frac{\zeta(d-1)(s-s^{d-1})}{(1-s)^d}\right)^{-1/2}+O(n^{-1}).
\end{aligned}\tag{40}$$

Plugging (39) and (40) into (29) and recalling $\nu=(1-s)^{-1}n$, we obtain

$$\begin{aligned}
c_d(n,p)&\sim s^{s\nu}(1-s)^{(1-s)\nu}(1-p)^{\binom{\nu-n}{d}+\binom{n}{d}-\binom{\nu}{d}}\cdot\eta\\
&\sim s^{\frac{sn}{1-s}}(1-s)^n\exp\left(\frac{\zeta\left(1-s^d-(1-s)^d\right)n}{d(1-s)^d}\right)
\end{aligned}$$

$$\cdot \exp\left[\frac{\zeta(d-1)s(1-s^d-(1-s)^d)}{2(1-s)^d}+\frac{\zeta(d-1)s}{2}\left(\left(\frac{s}{1-s}\right)^{d-2}+1\right)\right]$$
$$\cdot\left(1-\zeta(d-1)\left(\frac{s}{1-s}\right)^{d-1}\right)\left(1+\frac{\zeta(d-1)(s-s^{d-1})}{(1-s)^d}\right)^{-1/2}.$$

Finally, using Lemma 9 (ii)–(iii), i.e. $|s-\varrho|=O(\nu^{-1})$ and

$$\left(s^{\frac{s}{1-s}}(1-s)\exp\left(\frac{\zeta\left(1-s^d-(1-s)^d\right)}{d(1-s)^d}\right)\right)^n$$
$$\sim\left(\varrho^{\frac{\varrho}{1-\varrho}}(1-\varrho)\exp\left(\frac{\zeta\left(1-\varrho^d-(1-\varrho)^d\right)}{d(1-\varrho)^d}\right)\right)^n,$$

we estimate (41) as

$$\begin{aligned} c_d(n,p)\sim&\left((1-\varrho)\varrho^{\frac{\varrho}{1-\varrho}}\exp\left(\frac{\zeta(1-\varrho^d-(1-\varrho)^d)}{d(1-\varrho)^d}\right)\right)^n\\ &\cdot\exp\left(\frac{\zeta(d-1)\varrho(1-\varrho^d-(1-\varrho)^d)}{2(1-\varrho)^d}+\frac{\zeta(d-1)\varrho}{2}\left(\left(\frac{\varrho}{1-\varrho}\right)^{d-2}+1\right)\right)\\ &\times\left(1-\zeta(d-1)\left(\frac{\varrho}{1-\varrho}\right)^{d-1}\right)\left(1+\frac{\zeta(d-1)(\varrho-\varrho^{d-1})}{(1-\varrho)^d}\right)^{-1/2}, \end{aligned}\tag{41}$$

which is exactly the formula stated in Theorem 2.

□

4 The Conditional Edge Distribution: Proof of Theorem 4

Let $\mathcal{J}\subset(0,\infty)$ and $\mathcal{I}\subset\mathbf{R}$ be compact sets, and let $0<p=p(n)<1$ be a sequence such that $\zeta=\zeta(n)=\binom{n-1}{d-1}p\in\mathcal{J}$ for all n. All asymptotics in this section are uniform in ζ.

To compute the limiting distribution of the number of edges of $H_d(n,p)$ given that this random hypergraph is connected, we choose $\nu>n$ as in Section 3. Thus, letting $c=\binom{\nu-1}{d-1}p$, we know from Section 3 that $c>(d-1)^{-1}$, and that the solution $0<\rho<1$ to (1) satisfies $(1-\rho)\nu\le n<(1-\rho)\nu+O(1)$. Now, we investigate the random hypergraph $H_d(\nu,p)$ *given that* $\mathcal{N}(H_d(\nu,p))=n$. Then the largest component of $H_d(\nu,p)$ is a random hypergraph $H_d(n,p)$ *given that* $H_d(n,p)$ *is connected*. Therefore,

$$\begin{aligned}\mathbb{P}&\left[|E(H_d(n,p))|=m\mid H_d(n,p)\text{ is connected}\right]\\ &=\mathbb{P}\left[\mathcal{M}(H_d(\nu,p))=m\mid\mathcal{N}(H_d(\nu,p))=n\right]\\ &=\frac{\mathbb{P}\left[\mathcal{M}(H_d(\nu,p))=m,\ \mathcal{N}(H_d(\nu,p))=n\right]}{\mathbb{P}\left[\mathcal{N}(H_d(\nu,p))=n\right]}.\end{aligned}\tag{42}$$

Furthermore, as $|n-(1-\rho)\nu|<O(1)$ by (20), we can apply Lemma 6 to get an explicit expression for the r.h.s. of (42). Namely, using (13) with $x=O(1)$, for any integer m such that $\nu^{-\frac12}y\in\mathcal{I}$ and $y=m-(1-\rho^d)\binom{\nu}{d}p$ satisfying $\nu^{-\frac12}y\in\mathcal{I}$ we obtain

$$\begin{aligned}&\mathbb{P}\left[|E(H_d(n,p))|=m\mid H_d(n,p)\text{ is connected}\right]\\ \sim\ &\frac{1}{\sqrt{2\pi}}\cdot\left(\frac{\sigma_{\mathcal{N}}^2}{(\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{NM}}^2)}\right)^{\frac12}\exp\left(-\frac{\sigma_{\mathcal{N}}^2}{2(\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{NM}}^2)}\cdot y^2\right).\end{aligned}\tag{43}$$

From (9) and (15) we have

$$\sigma_{\mathcal{N}}^2=\frac{\rho\left(1-\rho+c(d-1)(\rho-\rho^{d-1})\right)}{(1-c(d-1)\rho^{d-1})^2}\cdot\nu,$$
$$\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{NM}}^2=\frac{c\rho\left(\left(1-\rho+c(d-1)(\rho-\rho^{d-1})\right)(1-\rho^d)-dc\rho(1-\rho^{d-1})^2\right)}{d\left(1-c(d-1)\rho^{d-1}\right)^2}\cdot\nu^2.$$

Thus we have

$$\frac{\sigma_{\mathcal{N}}^2}{\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{N}\mathcal{M}}^2} = \frac{d\big(1-\rho+c(d-1)(\rho-\rho^{d-1})\big)}{c\left(\left(1-\rho+c(d-1)(\rho-\rho^{d-1})\right)(1-\rho^d)-dc\rho(1-\rho^{d-1})^2\right)}\cdot\frac{1}{\nu}$$
$$= \frac{d}{c\,\nu}\left(1-\rho^d-\frac{dc\rho(1-\rho^{d-1})^2}{1-\rho+c(d-1)(\rho-\rho^{d-1})}\right)^{-1}. \quad (44)$$

In order to reformulate (44) in terms of n, ζ, and the solution ϱ to (3), we just observe that $|c-\zeta(1-\rho)^{1-d}| = O(\nu^{-1})$ and $|\rho-\varrho| = O(\nu^{-1})$ by Lemma 9, and that $|\nu-(1-\rho)^{-1}n| = O(\nu^{-1})$. Using these we obtain

$$\left(\frac{\sigma_{\mathcal{N}}^2}{\sigma_{\mathcal{N}}^2\sigma_{\mathcal{M}}^2-\sigma_{\mathcal{N}\mathcal{M}}^2}\right)^{-1} = \frac{c\nu}{d}\left(1-\rho^d-\frac{dc\rho(1-\rho^{d-1})^2}{1-\rho+c(d-1)(\rho-\rho^{d-1})}\right)$$
$$\sim \frac{\zeta n}{d(1-\rho)^d}\left(1-\rho^d-\frac{d\zeta(1-\rho)^{1-d}\rho(1-\rho^{d-1})^2}{1-\rho+\zeta(1-\rho)^{1-d}(d-1)(\rho-\rho^{d-1})}\right)^{-1}$$
$$= \frac{\zeta}{d(1-\rho)^d}\left(1-\rho^d-\frac{d\zeta\rho(1-\rho^{d-1})^2}{(1-\rho)^d+\zeta(d-1)(\rho-\rho^{d-1})}\right)\cdot n$$
$$\sim \frac{\zeta}{d(1-\varrho)^d}\left(1-\varrho^d-\frac{d\zeta\varrho(1-\varrho^{d-1})^2}{(1-\varrho)^d+(d-1)\zeta(\varrho-\varrho^{d-1})}\right)\cdot n$$
$$= \hat{\sigma}^2, \quad (45)$$

and

$$(1-\rho^d)\binom{\nu}{d}p = (1-\rho^d)\frac{\nu}{d}c \;\sim\; (1-\rho^d)\frac{n}{d(1-\rho)}\zeta(1-\rho)^{1-d} = \frac{\zeta(1-\varrho^d)}{d(1-\varrho)^d}\cdot n.$$

Plugging (45) into (43) we have

$$\mathbb{P}\left[|E(H_d(n,p))| = m \mid H_d(n,p) \text{ is connected}\right] \sim \frac{1}{\sqrt{2\pi}\hat{\sigma}}\exp\left(-\frac{y^2}{2\hat{\sigma}^2}\right),$$

as desired.

Acknowledgment. We thank the anonymous referee for useful comments and suggestions on earlier versions of this paper.

References

1. Andriamampianina, T., Ravelomanana, V.: Enumeration of connected uniform hypergraphs. Proceedings of FPSAC 2005.
2. Behrisch, M., Coja-Oghlan, A., Kang, M.: in the Proceedings of RANDOM 2007, Springer LNCS 4627, 341–352.
3. Behrisch, M., Coja-Oghlan, A., Kang, M.: The order of the giant component of random hypergraphs. Random Structures and Algorithms **36** (2010), 149–184.
4. Behrisch, M., Coja-Oghlan, A., Kang, M.: Local limit theorems for the giant component of random hypergraphs. to appear in Combinatorics, Probability and Computing.
5. Bender, E.A., Canfield, E.R., McKay, B.D.: The asymptotic number of labeled connected graphs with a given number of vertices and edges. Random Structures and Algorithms **1** (1990) 127–169.
6. Bender, E.A., Canfield, E.R., McKay, B.D.: Asymptotic properties of labeled connected graphs. Random Structures and Algorithms **3** (1992) 183–202.
7. Bollobás, B.: Random graphs. 2nd edition. Cambridge University Press (2001).
8. Coja-Oghlan, A., Moore, C., Sanwalani, V.: Counting connected graphs and hypergraphs via the probabilistic method. Random Structures and Algorithms **31** (2007) 288–329.
9. Erdős, P., Rényi, A.: On random graphs I. Publicationes Mathematicae Debrecen **5** (1959) 290–297.
10. Erdős, P., Rényi, A.: On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. **5** (1960) 17–61.

11. van der Hofstad, R., Spencer, J.: Counting connected graphs asymptotically. European Journal on Combinatorics **27** (2006) 1294–1320.
12. Janson, S., Łuczak, T, Ruciński, A.: Random Graphs, Wiley 2000.
13. Karoński, M., Łuczak, T.: The number of connected sparsely edged uniform hypergraphs. Discrete Math. **171** (1997) 153–168.
14. Łuczak, T.: On the number of sparse connected graphs. Random Structures and Algorithms **1** (1990) 171–173.
15. O'Connell, N.: Some large deviation results for sparse random graphs. Prob. Th. Relat. Fields **110** (1998) 277–285.
16. Pittel, B., Wormald, N.C.: Counting connected graphs inside out. J. Combin. Theory, Series B **93** (2005) 127–172.
17. Schmidt-Pruzan, J., Shamir, E.: Component structure in the evolution of random hypergraphs. Combinatorica **5** (1985) 81–94.
18. Stepanov, V. E.: On the probability of connectedness of a random graph $g_m(t)$. Theory Prob. Appl. **15** (1970) 55–67.