This is to certify that the

dissertation entitled

Hypertrees in d-uniform Hypergraph

presented by

Wai Cheong Siu

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Mathematics

Edgar M. Palmer
Major professor

Date April 30, 2002

Hypertrees in d-uniform hypergraphs

By

Wai-Cheong Siu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

2002

ABSTRACT

Hypertrees in d-uniform hypergraphs

By

Wai-Cheong Siu

Traditionally a d-uniform hypertree has been defined as a d-uniform hypergraph that is connected and has no cycle. In this dissertation, we study various alternative definitions of d-uniform hypertrees. In particular, we formulate a new kind of hypertree called a (d, k)-tree. We have enumerated rooted and unrooted labeled (d, k)-trees and (d, k)-forests. We also provide some results on (d, k)-trees in random hypergraphs. In particular, we have approximated the number of edges in a random hypergraph so that *with high probability* a spanning hypertree exists.

ACKNOWLEDGMENTS

I would like to thank my thesis advisor Professor Edgar M. Palmer for his patience, guidance and valuable advice. I also would like to thank Professor Bruce E. Sagan for his lectures that equipped me the knowledge of combinatorics and graph theory and for serving in my committee. I would like to thank Professor Abdol-Hossein Esfahanian, Peter Magyar and Eric Torng for serving in my guidance committee. Finally, I would like to thank my parents, Siu, Tsang-Kwai and Chow, How-Ling and my sister Siu, Mei-Lam for their encouragement and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

The study of graph structure has a long history which dates to the early eighteenth century when the Königsberg bridge problem was posed. The problem was solved in 1736 by Euler in the first paper published on graph theory [BiLW76]. For the next two hundred years graph theory underwent steady and important growth and many famous mathematicians, like Cayley, Hamilton, Heawood, Kirchhoff and Tait were involved in formulating and solving interesting new problems. Their efforts formed the footings of the foundation of graph theory. In the last half of the $20^{\underline{th}}$ century, the growth of graph theory has been impressive, as evidenced by the numerous books and new jounals devoted to the subject, as well as by many other important scholarly activities. No doubt the applications of graph theory to other areas such as computer science, operation research, theoretical chemistry, etc, account for some of this growth. But mathematicaians have also been drawn by the beauty of the subject and the promise of the area to continue to be a rich source of attractive and important problems.

Among the many interesting types of graphs that have been studied throughout the history of graph theory, the tree is the simplest and the most useful. Therefore, counting the number of trees (both labeled and unlabeled) attracted the attention of many researchers. For a comprehensive survey of techniques and results on tree enumeration, the reader can consult [M70].

Besides developing the theory of graphs, mathematicans started to generalize the concept of graphs and trees to higher dimensions in the 1960's, when hypergraphs and hypertrees were introduced. Claude Berge was a pioneer in this field and the reader can consult his books [Be70], [Be73] and [Be87] for an introduction to hypergraphs. We will review and discuss verious definitions for hypertrees in the next sections.

At almost the same time, Erdös and Rényi wrote a series of remarkable papers [ErR59], [ErR60], [ErR61a], [ErR61b], [ErR64] and [ErR68] that gave birth to the theory of random graphs. One of their startling discoveries revealed the concept of the probabilistic threshold for monotone graph properties. These properties were found to occur rather abruptly in large random graphs when the number of edges increased only slightly. And Erdös and Rényi discovered many important examples. To illustrate, let $\omega_n \longrightarrow \infty$ as $n \longrightarrow \infty$ and consider a sequence $\{G_n\}$ of random graphs. If each G_n has $\frac{n(\ln n+\omega_n)}{2}$ edges, then *with high probability* G_n has a spanning tree, i.e. the probability that G_n has a spanning tree approaches 1 as $n \longrightarrow \infty$. On the other hand if G_n has only $\frac{n(\ln n-\omega_n)}{2}$ edges, then *with high probability* G_n does not have a spanning tree.

1.1 Hypergraphs

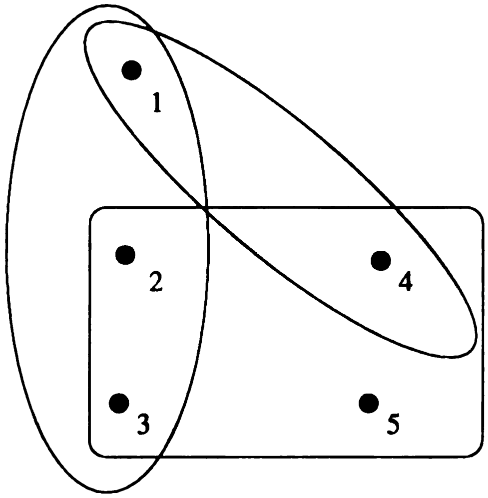
Let $V = \{v_1, v_2, \ldots, v_n\}$ be a finite set and $E = \{e_1, e_2, \ldots, e_m\}$ be a subset of the power set $\mathcal{P}(V)$. The ordered pair $H = (V, E)$ is called a *hypergraph* by [Be87] if

(1) $\quad e_i \neq \varnothing \qquad (i = 1, 2, \ldots, m)$

(2) $\quad \bigcup_{i=1}^{m} e_i = V.$

The finite set V is called the *vertex set* and the elements of V are *vertices*. The set E is the *edge set* of the hypergraph H and the elements of E are *edges* of hypergraph

Figure 1.1: Hypergraph H_1



H. The cardinality of the vertex set V is the *order* of the hypergraph H and is denoted by n. The cardinality of the edge set E is the *size* of the hypergraph H and is denoted by m. For example, $H_1 = (V_1, E_1)$ where $V_1 = \{1,2,3,4,5\}$ and $E_1 = \{\{1,2,3\},\{2,3,4,5\},\{1,4\}\}$ is a hypergraph (see Figure 1.1). The enumeration problem for unlabeled hypergraph was treated by Toru Ishihara [Is01] using standard Pólya theory [HaP73].

A hypergraph H is *d-uniform* if the edges all have the same cardinality d. A one-element edge is a *singleton* or *loop*.

Let $u, v, w_0, w_1, \ldots, w_t$ be vertices of H and $e_1, e_2, \ldots, e_t$ be edges of H where t is a non-negative integer. If $w_0 = u$, $w_t = v$, $w_0 \in e_1$, $w_t \in e_t$, $w_i \in e_i$ and $w_i \in e_{i+1}$ for all $i \in [t-1]$, the sequence $(w_0, e_1, w_1, e_2, \ldots, e_t, w_t)$ is called a $u - v$ *walk* of H . A $u - v$ walk is a $u - v$ *trail* if all the edges e_i are distinct. A $u - v$ walk is called a *path* if all the vertices w_i are distinct. A $u - v$ walk is called a $u - v$ *hyperpath* if all the vertices w_i and all the edges e_i are distinct. A $u - v$ walk is called a *cycle of length* t if $t \geq 2$, $u = v$, w_i are distinct for $1 \leqslant i \leqslant t$ and all the edges e_i are distinct. A hypergraph H is connected if there is a $u-v$ path between any two vertices, u and v in H. A hypergraph H is *acyclic* if it has no cycle.

1.2 Hypertrees

Here are five different definitions of a hypertree that have been used by researchers in the development of the field.

1.2.1 Traditional definition of hypertrees

In the traditional definition, a *hypertree* is defined as a connected hypergraph which contains neither loops nor cycles [SoT00] [Bo84]. So, a d-uniform hypertree with $d = 1$ is an ordinary graph-theoretic tree). For example, H_2 with $V(H_2) = \{1,2,3,4,5,6,7,8,9,10,11\}$ and $E(H_2) = \{\{1,2,3\},\{3,4,5\},\{4,6,7\},\{5,8,9\},\{9,10,11\}\}$ is a 3-uniform hypertree (see Figure 1.2). However H_3 with $V(H_3) = \{1,2,3,4,5,6\}$ and $E(H_3) = \{\{1,2,3\},\{2,3,4\},\{3,4,5\},\{2,4,6\}\}$ (see Figure 1.3) is a 3-uniform hypergraph but not a hypertree, because it contains the following cycle: $\{3\}\{3,4,5\}\{4\}\{2,4,6\}\{2\}\{1,2,3\}\{3\}$.

Table 1.1 and Table 1.2 give the number of 3-uniform and 4-uniform hypertrees with small order n. These numbers were found by constructing the relevant configurations.

1.2.2 NP-hypertrees

The NP-hypertrees were introduced by J. Nieminen and M. Peltola in their paper [NiP99]. A hypergraph H is an *NP-hypertree* if H is trivial or the removal of any edge from it results in a disconnected hypergraph. This family of hypertrees is a superset of

Table 1.1: Number of Unlabeled and Labeled 3-uniform Hypertrees.

n	Unlabeled	Labeled
3	1	1
5	1	15
7	2	735
9	4	76,545
11	8	13,835,745

Table 1.2: Number of Unlabeled and Labeled 4-uniform Hypertrees.

n	Unlabeled	Labeled
4	1	1
7	1	70
10	2	28,000
13	4	33,833,800
16	9	91,842,150,400

the family of traditional hypertrees. For example, H_4 with $V(H_4) = \{1,2,3,4,5,6\}$ and $E(H_4) = \{\{1,2,3\},\{3,4,5\},\{4,5,6\}\}$ is a 3-uniform NP-hypertree(see Figure 1.4). Notice that H_4 is not a traditional hypertree but H_2 (see Figure 1.2) is an NP-hypertree.

Table 1.3 and Table 1.4 give the number of 3-uniform and 4-uniform NP-hypertrees with small n. The data was determined by constructive methods. Notice that the enumeration problem for NP-hypertrees has not been solved.

1.2.3 HP-hypertrees

Harary and Palmer [HaP68] defined certain families of graphs with tree-like structure that correspond to hypergraphs. These are different from the traditional hypertrees and so we will call them *HP-hypertrees*. These trees were characterized by [HaP68]

Figure 1.4: NP-hypertree H_4

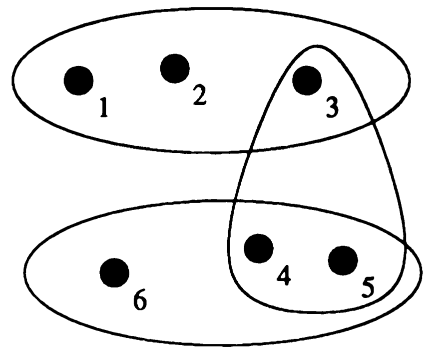


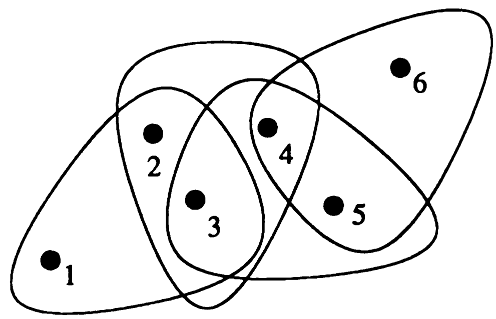
Table 1.3: Number of Unlabeled and Labeled 3-uniform NP-hypertrees

n	Unlabeled	Labeled
1	1	1
3	1	1
4	1	6
5	3	160
6	4	495

Table 1.4: Number of Unlabeled and Labeled 4-uniform NP-hypertrees

n	Unlabeled	Labeled
1	1	1
4	1	1
5	1	10
6	2	110
7	4	2,275

Figure 1.5: HP-hypertree H_5



as:

For any positive integer d greater than 1:

(1) Any d-set V forms a hypertree $H = (V, E)$ with $E = \{V\}$.

(2) A hypertree $H = (V, E)$ of order n can be formed as follows: First consider a hypertree $H' = (V', E')$ of order $n-1$ and a new vertex $v \notin V'$. Then we pick a $(d-1)$-subset e from an edge $e' \in E'$. Next set $V = V' \cup \{v\}$ and $E = E' \cup \{e \cup \{v\}\}$.

For example, H_5 with $V(H_5) = \{1, 2, 3, 4, 5, 6\}$ and $E(H_5) = \{\{1, 2, 3\}, \{2, 3, 4\}, \{3, 4, 5\}, \{4, 5, 6\}\}$ is an 3-uniform HP-hypertree. (see Figure 1.5)

Table 1.5 and Table 1.6 give the number of HP-hypertrees for small n with $d = 3$ and $d = 4$. The enumeration of labeled HP-hypertrees was solved by Beineke and Pippert (see [BeP69]). Harary and Palmer (see [HaP68]) handled the unlabeled case. The numbers for the labeled cases in the tables are calculated by using the formula from [BeP69]. The numbers for the unlabeled cases in the Table 1.5 are from [HaP68] and those in Table 1.6 are found by construction.

Table 1.5: Number of HP-hypertrees with $d = 3$

n	Unlabeled	labeled
3	1	1
4	1	6
5	2	70
6	5	1,215
7	12	2,7951

Table 1.6: Number of HP-hypertrees with $d = 4$

n	Unlabeled	Labeled
4	1	1
5	1	10
6	2	200
7	4	5,915

1.2.4 BD-hypertrees

The BD-hypertrees were studied intensively by Andreas Brandstädt, V. D. Chepoi and Feodor F. Dragan (see [BrCD95], [BrD96] and [BrCD98]). This hypertree structure can be used to construct efficent graph algorithms. Here, a hypergraph $H = (V, E)$ is a *BD-hypertree* if there is an ordinary graph-theoretic tree T with vertex set V such that every edge $e \in E$ induces a subtree in T. For example, H_6 with $V(H_6) = \{1, 2, 3, 4, 5, 6\}$ and $E(H_6) = \{\{1, 2\}, \{2, 3, 4\}, \{3, 5, 6\}\}$ is a BD-hypertree. Notice that the ordinary tree T with $V(T) = \{1, 2, 3, 4, 5, 6\}$ and $E(T) = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{3, 5\}, \{5, 6\}\}$ satisfies the above requirement.(see Figure 1.6) Table 1.7 gives the number of BD-hypertrees with small n.

Notice that the enumeration problem for BD-hypertrees has not been solved.

Figure 1.6: BD-hypertree H_6

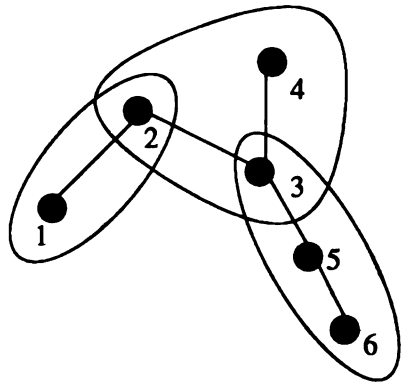


Table 1.7: Number of Labeled BD-hypertrees

n	Labeled
2	1
3	9
4	401

1.2.5 Tomescu definition of hypertrees

The Tomescu hypertrees were introduced to obtain the Bonferronni inequalities by Ioan Tomescu. They are discussed in [To86], [To92], [ToZ94]. At least seven equivalent definitions of these hypertrees have been proposed. The following is the version using recursive definition to define the hypertrees.

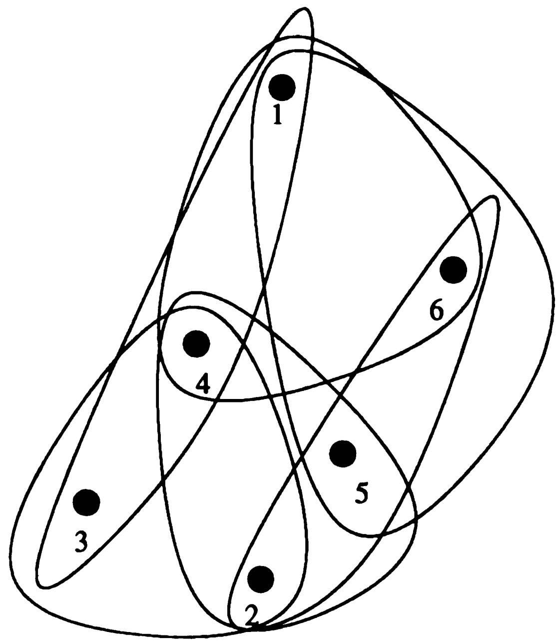
Let $T = (V, E)$ be an h-uniform hypergraph. When $h = 2$, T is a graph and T is called an *h-hypertree* if T is a tree. When $h \geq 3$, T is an *h-hypertree* if and only if

(1) $\mid V \mid= h$ and $E = \{V\}$, i.e. T has exactly one edge consisting of all h vertices of V.

OR

(2) $\mid V \mid\geq h + 1$ then there is a vertex $v_i \in V$ such that if $e_1, e_2, \ldots, e_q$ denote

Figure 1.7: Tomescu 3-hypertree H_7



all edges containing v_i then $e_1\backslash\{v_i\},\ldots,e_q\backslash\{v_i\}$ induce an $(h-1)$-hypertree with vertex set $V\backslash\{v_i\}$ and the remaining edges of T induce an h-hypertree with vertex set $V\backslash\{v_i\}$.

For example, H_7 with $V(H_7) = \{1,2,3,4,5,6\}$, $E(H_7) = \{\{1,3,4\},\{1,4,6\},\{1,5,6\},\{2,3,4\},\{2,4,5\},\{2,4,6\}\}$ (see Figure 1.7) is a Tomescu 3-hypertree.

The enumeration of Tomescu hypertrees remains an unsolved problem.

CHAPTER 2

The definition and enumeration of (d, k)-hypertrees

In this chapter, we define and enumerate two tree-like hypergraph structures which we call them (d, k)-trees and d-trees, where $d \geq 2$ and $k > 0$ are integers. These new definitions generalize traditional and HP-hypertrees.

2.1 Definition of a (d, k)-tree

For k fixed, with $1 \leq k \leq d - 1$, a (d, k)-*tree* is defined inductively as follows :

(1) A single edge is a (d, k)-tree.

(2) Suppose T is a (d, k)-tree with m vertices, then the hypergraph formed by adding a new edge consisting of $d - k$ new vertices and any k vertices of any edge of T is also a (d, k)-tree.

Note that when $k = d - 1$, these are the d-dimensional HP-hypertrees that first appeared in [HaP68] and were subsequently studied in [M70] [BeM69] and [BeP69]. Of course, when $d = 2$, then $k = 1$, we have the ordinary trees of graph theory with at least one edge. When $d = 3$ and $k = 1$, they are pure Husimi trees [Hu50].

If we relax condition (2) and allow k to vary $(1 \leq k \leq d - 1)$, we obtain a different kind of hypertree called a *d-tree*.

2.2 Enumeration of (d,k)-trees

We begin by determining egf's for various types of rooted (d,k)-trees. The relationships between these egf's are determined in the following lemmas, concluding with a specific generalization of Cayly's famous n^{n-2} formula. Note that the number of vertices $|V(T)|$ and the number of edges $|E(T)|$ of a (d,k)-tree satisfy the equation

$$|V|-k=|E|(d-k) \tag{2.1}$$

which reduces to the familiar

$$|V|-1=|E| \tag{2.2}$$

for $(2,1)$-trees.

A *simply rooted* (d,k)-tree has as its root a linearly ordered k-subset of vertices which belongs to exactly one edge.

Let y_m be the number of labeled simply rooted (d,k)-trees with m edges, whose vertices are labeled except for the k linearly ordered vertices of the root and let y be the exponential generating function for these labeled simply rooted (d,k)-trees. Then it follows from (2.1) that y has the form :

$$y=\sum_{m=1}^{\infty} y_m \frac{x^{m(d-k)}}{(m(d-k))!} \quad . \tag{2.3}$$

Note that $y_1=1$ and $y_2=\binom{2(d-k)}{(d-k)}\left(\binom{d}{k}-1\right)$.

Lemma 2.1 *The egf y for simply rooted (d,k)-trees satisfies the functional equation*

$$y=\frac{(e^y)^{\binom{d}{k}-1}x^{d-k}}{(d-k)!} \quad . \tag{2.4}$$

Proof : Since y is the egf for the labeled simply rooted (d,k)-trees, e^y is the egf for the labeled (d,k)-trees which are rooted at an unlabeled, linearly ordered k-subset of an edge (the k-subset may belong to many edges). Then, $(e^y)^{\binom{d}{k}-1}$ is the egf of $\binom{d}{k}-1$ ordered copies of this kind of (d,k)-tree. Now, if we start with an ordered d-set of vertices, then the order of the vertices imposes a natural order on all the k-subsets of it. Next we use the orders to match up the $\binom{d}{k}-1$ copies of special (d,k)-trees with the k-subsets, and then, identify the vertices in the (d,k)-trees with the vertices in the corresponding k-subsets using the orders. The result is a tree like structure whose egf is $(e^y)^{\binom{d}{k}-1}$. Now, if we label the unlabeled vertices and remove the order on the d-set except for those vertices in the last k-subset (which has no special (d,k)-tree assigned to it), we get a labeled, simply rooted (d,k)-tree. But its egf is $\frac{(e^y)^{\binom{d}{k}-1}x^{d-k}}{(d-k)!}$ as in (2.4). □

A *rooted* (d,k)-tree has as its root a k-subset of vertices which belongs to at least one edge. We denote by Y the egf for rooted (d,k)-trees whose vertices are all labeled, even those which belong to the root. We define the coefficients of Y as follows :

$$Y = \sum_{m=1}^{\infty} Y_m \frac{x^{m(d-k)+k}}{(m(d-k)+k)!} \quad . \tag{2.5}$$

So Y_m is the number of these trees with m edges, and we have $Y_1 = \binom{d}{k}$ and $Y_2 = \binom{d}{k}\binom{2(d-k)+k}{d-k,k,d-k}$.

Lemma 2.2 *The egf Y for rooted (d,k)-trees can be expressed in terms of the egf y for simply rooted trees as follows :*

$$Y = \frac{(e^y-1)x^k}{k!} \quad . \tag{2.6}$$

Proof : From the proof of Lemma 2.1, e^y-1 is the egf for the labeled (d,k)-trees rooted at an unlabeled, linearly ordered k-subset of an edge (the k-subset may belong to many edges), with at least one edge. If we label the k vertices of the root and remove the order on them, we have a rooted (d,k)-tree. The egf of the rooted (d,k)-trees resulting from the above operations is $\frac{(e^y-1)x^k}{k!}$. Notice that when $d=2$ and

$k = 1$ (i.e. ordinary trees), we have to use $Y = e^y x$ instead of (2.6). It's because unlike $d \geq 3$ cases, a single rooted vertex is considered as a rooted tree with no edge in $d = 2$ and $k = 1$. □

An *edge rooted* (d, k)-tree has as its root a single edge. Let z be the egf for these. Then the coefficients of z are defined by

$$z = \sum_{m=1}^{\infty} z_m \frac{x^{m(d-k)+k}}{(m(d-k)+k)!} \quad . \tag{2.7}$$

So we have $z_1 = 1$ and $z_2 = \binom{2(d-k)+k}{d-k,k,d-k}$.

Lemma 2.3 *The egf for edge rooted (d,k)-trees can be expressed in terms of the egf for simply rooted trees as follows :*

$$z = \frac{(e^y)^{\binom{d}{k}} x^d}{d!} \quad . \tag{2.8}$$

Proof : We follow the proof of lemma 1, but use ordered $\binom{d}{k}$ copies of the special kind of (d, k)-trees and match them to all the k-subsets of the ordered d-set. Then we fill in labels for all d unlabeled vertices and remove the order on them to obtain we have an edge rooted (d, k)-tree. The egf of the edge rooted (d, k)-tree resulting from the above operations is $\frac{(e^y)^{\binom{d}{k}} x^d}{d!}$. □

We denote the egf for (d, k)-trees by Z and define its coefficients by

$$Z = \sum_{m=1}^{\infty} Z_m \frac{x^{m(d-k)+k}}{(m(d-k)+k)!} \quad . \tag{2.9}$$

Then $Z_1 = 1$, $Z_2 = \frac{z_2}{2} = \frac{\binom{2(d-k)+k}{d-k,k,d-k}}{2}$.

Lemma 2.4 *The egf for (d,k)-trees is expressed in terms of Y and z as follows :*

$$Z = Y - \left(\binom{d}{k} - 1\right) z. \tag{2.10}$$

Proof : Consider a (d, k)-tree T. By the inductive definition of (d, k)-trees, there is a way to construct T by adding edges one by one. Following the order of construction, we can order the edge set of T. By using the order of the edge set, we can construct a many-to-one mapping from the set of labeled rooted (d,k)-tree obtained from T to the set of labeled edge rooted (d,k)-trees obtained from T as follows : For a labeled rooted (d,k)-tree of T, we map it to a labeled edge rooted (d,k)-tree of T whose rooted edge contains the k-set which form the root of the labeled rooted (d,k)-tree of T. If the k-set is contained in more than one edge of T, we choose the labeled edge rooted (d,k)-tree of T whose rooted edge has the highest priority among the edges which contain the rooted k-set. It's easy to see that the above construction provides a many to one mapping. Futhermore, there are exactly $\binom{d}{k}-1$ many labeled rooted (d,k)-trees of T mapped to every labeled edge rooted (d,k)-tree of T, except the one which is rooted at the edge of T with highest priority among all the edges of T. For the only exception here, it is mapped by exactly $\binom{d}{k}$ rooted (d,k)-trees. Therefore, we have :

$$\begin{aligned} 1 &= (\text{number of ways to root the } (d,k)\text{-tree } T) \\ &\quad -(\tbinom{d}{k}-1)(\text{number of ways to root the } (d,k)\text{-tree } T \text{ at an edge}) \end{aligned}$$

or

$$Z_m = Y_m - \left(\binom{d}{k}-1\right) z_m \tag{2.11}$$

where m is the number of edges of a (d,k)-tree.

Multiplying both sides of (2.11) by $\frac{x^{m(d-k)+k}}{(m(d-k)+k)!}$ and summing over $m \geq 1$, we arrive the generating function equation (2.10). □

From (2.4), (2.6), (2.8) and (2.11), we obtain the following theorm:

Theorem 2.1 *The number of (d,k)-trees of order $n = m(d-k)+k$ is*

$$Z_m = Y_m - \left(\binom{d}{k}-1\right) z_m \tag{2.12}$$

where

$$Y_m = \frac{((m(d-k)+k)!)\left(m\left(\binom{d}{k}-1\right)+1\right)^{m-1}}{m!k!\left((d-k)!\right)^m} \tag{2.13}$$

$$z_m = \frac{\binom{d}{k}\left((m\,(d-k)+k)!\right)\left(\left(\binom{d}{k}-1\right)m+1\right)^{m-2}}{d!\,((d-k)!)^{m-1}\,(m-1)!} \tag{2.14}$$

Proof : Let $n = m(d-k)+k$ and $\alpha = x^{d-k}$. If we write y in terms of α, then from (2.4), we find

$$y = \frac{(e^y)^{\binom{d}{k}-1}x^{d-k}}{(d-k)!} = \frac{(e^y)^{\binom{d}{k}-1}\alpha}{(d-k)!} = \phi(y)\alpha \tag{2.15}$$

where $\phi(y) = \frac{(e^y)^{\binom{d}{k}-1}}{(d-k)!}$.

The coefficients of Y are extracted from (2.6) by the following steps:

$$\begin{aligned}
\left[\frac{x^n}{n!}\right]Y &= \left[\frac{x^n}{n!}\right]\frac{(e^y-1)x^k}{k!} \\
&= \frac{n!}{(n-k)!}\left[\frac{x^{n-k}}{(n-k)!}\right]\frac{e^y-1}{k!} \\
&= \frac{n!}{(n-k)!k!}\frac{(n-k)!}{m!}\left[\frac{\alpha^m}{m!}\right]e^y - 1 \\
&= \frac{(m(d-k)+k)!}{m!k!}\left[\frac{\alpha^m}{m!}\right]e^y - 1.
\end{aligned} \tag{2.16}$$

On applying Lagrange's inversion formula to (2.15), we have

$$\begin{aligned}
\left[\frac{\alpha^m}{m!}\right]e^y &= \left(\frac{d^{m-1}}{dy^{m-1}}\right)_{y=0}(e^y)'\,\alpha^m(y) \\
&= \frac{\left(m\left(\binom{d}{k}-1\right)+1\right)^{m-1}}{((d-k)!)^m}.
\end{aligned}$$

Therefore, $Y_m = \frac{((m(d-k)+k)!)\left(m\left(\binom{d}{k}-1\right)+1\right)^{m-1}}{m!k!((d-k)!)^m}$. We can obtain (2.14) in a similar way. Note that the formula works for $d \geq 3$ with $m \geq 1$ and $d = 2$ with $m \geq 0$. For $d \geq 3$ with $m = 0$, the number should be 0. □

To obtain Cayley's Theorem as a corollary, we take $d = 2$, $k = 1$ and $m = n-1$. Then Z_m is the number of labeled trees of order n, and the formulas in the Theorem yield

$$z_m = (n-1)n^{n-2} \tag{2.17}$$

$$Y_m = n^{n-1} \tag{2.18}$$

$$Z_m = n^{n-2}. \tag{2.19}$$

When $k = d-1$, The Theorem provides the formulas for the number of $(d-1)$-dimensional trees found by Beineke and Pippert [BeP69]. In this case $m = n-d+1$ and formulas (2.12), (2.13) and (2.14) give

$$z_m = \binom{n}{d}\left(d\left((n-d)(d-1)+d\right)^{n-d-1}\right) \tag{2.20}$$

$$Y_m = \binom{n}{d-1}\left(\left((n-d)(d-1)+d\right)^{n-d}\right) \tag{2.21}$$

$$Z_m = \binom{n}{d-1}\left((d-1)(n-(d-1))+1\right)^{n-(d-1)-2} = \binom{n}{k}\left(k(n-k)+1\right)^{n-k-2} \tag{2.22}$$

Note that

$$mZ_m = z_m \tag{2.23}$$

provides the preferred formula

$$Z_m = \frac{\binom{d}{k}\left((m\,(d-k)+k)!\right)\left(\left(\binom{d}{k}-1\right)m+1\right)^{m-2}}{d!m!\left((d-k)!\right)^{m-1}} \tag{2.24}$$

for Z_m over (2.12).

Note that Beineke and Pippert found the formula for Z_m with $k = d-1$ by using a special case of (2.6) and did not extend the egfs to include our equations (2.8) and (2.10). We have done this to enable us to enumerate forests in the next section.

2.3 Forests of (d, k)-trees

Let $f_l(m, d, k)$ be the number of forests of (d, k)-trees with l components and m edges, then $n = m(d-k)+kl$ and the egf for forests of (d, k)-trees with l components is

$$\frac{Z^l}{l!} = \frac{\left(Y - \left(\binom{d}{k}-1\right)z\right)^l}{l!}$$

$$= \sum_{i=0}^{l} \frac{(-1)^i \binom{l}{i} (Y)^{l-i} \left(\binom{d}{k} - 1\right)^i z^i}{l!}$$

On extracting the coefficients of this egf, we have the following theorem.

Theorem 2.2 *If $f_l(m,d,k)$ be the number of forests of (d,k)-trees with l components and m edges, then*

$$f_l(m,d,k) = \sum_{i=0}^{l} \binom{l}{i} \frac{(-1)^i \left(\binom{d}{k} - 1\right)^i (m(d-k)+kl)!}{l!(k!)^{l-i}(d!)^i(m-i)!\,((d-k)!)^{m-i}} B_i$$

where $m = \frac{n-kl}{d-k}$ and

$$B_i = \begin{cases} \sum_{j=0}^{l-i-1} \binom{l-i-1}{j} (-1)^{l-i-1-j} \left(\left(i\left(\binom{d}{k}-1\right)+l\right)\left(m\left(\binom{d}{k}-1\right) \right.\right. & \\ \left.\left. +i+j+1\right)^{m-i-1} - i\binom{d}{k}\left(m\left(\binom{d}{k}-1\right)+i+j\right)^{m-i-1}\right) & 0 \le i < l \\ l\binom{d}{k}\left(m\left(\binom{d}{k}-1\right)+l\right)^{m-l-1} & i = l \end{cases}$$

Proof : To find $f_l(m,d,k)$, we use (2.4), (2.6), (2.8) and Lagrange's inversion formula to extract coefficients. Here are some steps:

$$\begin{aligned}
\left[\frac{x^n}{n!}\right] \frac{Z^l}{l!} &= \left[\frac{x^n}{n!}\right] \sum_{i=0}^{l} \frac{(-1)^i \binom{l}{i} (Y)^{l-i} \left(\binom{d}{k}-1\right)^i z^i}{l!} \\
&= \left[\frac{x^n}{n!}\right] \sum_{i=0}^{l} \binom{l}{i} \frac{(-1)^i \left(\binom{d}{k}-1\right)^i}{l!} Y^{l-i} z^i \\
&= \sum_{i=0}^{l} \binom{l}{i} \frac{(-1)^i \left(\binom{d}{k}-1\right)^i}{l!} \left[\frac{x^n}{n!}\right] Y^{l-i} z^i \\
&= \sum_{i=0}^{l} \binom{l}{i} \frac{(-1)^i \left(\binom{d}{k}-1\right)^i}{l!} \left[\frac{x^n}{n!}\right] \left(\frac{(e^y-1)x^k}{k!}\right)^{l-i} \left(\frac{(e^y)^{\binom{d}{k}} x^d}{d!}\right)^i \\
&= \sum_{i=0}^{l} \binom{l}{i} \frac{(-1)^i \left(\binom{d}{k}-1\right)^i}{l!(k!)^{l-i}(d!)^i} \left[\frac{x^n}{n!}\right] (e^y-1)^{l-1} (e^y)^{i\binom{d}{k}} x^{k(l-i)+id} \\
&= \sum_{i=0}^{l} \binom{l}{i} \frac{(-1)^i \left(\binom{d}{k}-1\right)^i n!}{l!(k!)^{l-i}(d!)^i(m-i)!} \left[\frac{\alpha^{m-i}}{(m-i)!}\right] (e^y-1)^{l-i} (e^y)^{i\binom{d}{k}}
\end{aligned}$$

If we apply Lagrange's inversion formula and expand $(e^y - 1)^{l-i}$ when necessary, then we can continue in a similar way as in the proof of Theorem 2.1 to obtain the formula of the theorem. □

The formula in Theorem 2.2 opened the way for us to estimate the average number of (d,k)-trees in a labeled (d,k)-forest with large n. The ordinary labeled tree case of this problem was treated in [M70] and the ordinary unlabeled tree case was solved by E.M. Palmer and A.J. Schewnk in [PaS79].

Let us consider the special case $d = (b+1)a$ and $k = ba$ where a, b are any natural numbers not not both equal to 1. Note that when $a = 1$, we get d-dimensional HP-hypertrees. We also let $l \geq 2$ and fix the number of vertices n such that $n = m_1(d-k) + k = m(d-k) + kl$. Therefore m_1 is the number of edges in a (d,k)-tree with n vertices and m is the number of edges in a (d,k)-forest with n vertices and l components. We can write m_1 in terms of m and then we have $m_1 = m+h$ where $h = b(l-1)$ is an integer greater than 0. Consider the function $\beta(n) = \frac{n!\left(\left(\binom{d}{k}-1\right)m_1\right)^{m_1-1}}{m_1!((d-k)!)^{m_1}}$. Note that $\beta(n) \sim m_1 Z_{m_1}$. If we divide $f_l(m,d,k)$ by $\beta(n)$, we have

$$\begin{aligned}
\frac{f_l(m,d,k)}{\beta(n)} &= \sum_{i=0}^{l} \binom{l}{i} \left(\frac{(-1)^i\left(\binom{d}{k}\right)^i}{l!(k!)^{l-i}(d!)^i}\right)\left(\frac{m_1!}{(m-i)!}\right)\left(\frac{((d-k)!)^{m_1}}{((d-k)!)^{m-i}}\right)\left(\frac{B_i}{\left(\binom{d}{k}-1\right)^{m_1-1}(m_1)^{m_1-1}}\right) \\
&= \sum_{i=0}^{l} \binom{l}{i} \left(\frac{(-1)\left(\binom{d}{k}-1\right)k!}{d!}\right)^i \left(\frac{((d-k)!)^{h+1}}{l!(k!)^l}\right)\left(\frac{(m+h)!}{(m-i)!}\right)\left(\frac{B_i}{\left(\binom{d}{k}-1\right)^{m+h-1}(m+h)^{m+h-1}}\right)
\end{aligned} \tag{2.25}$$

where B_i is defined in Theorem 2.2.

Continuing to simplfy formula (2.25), we have

$$\begin{aligned}
&\left(\frac{(m+h)!}{(m-i)!}\right)\left(\frac{B_i}{\left(\binom{d}{k}-1\right)^{m+h-1}(m+h)^{m+h-1}}\right) = \\
&(m+h)(m+h-1)\cdots(m-(i-1)) \\
&\times \left(\sum_{j=0}^{l-i-1}\binom{l-i-1}{j}(-1)^{l-i-1-j}\left(\left(i\left(\binom{d}{k}-1\right)+l\right)\right.\right. \\
&\times \left.\left.\left(m\left(\binom{d}{k}-1\right)+i+j+1\right)^{m-i-1} - i\binom{d}{k}\left(m\left(\binom{d}{k}-1\right)+i+j\right)^{m-i-1}\right)\right) \\
&/\left(\left(\binom{d}{k}-1\right)^{m+h-1}(m+h)^{m+h-1}\right) = \\
&\frac{1}{\left(\binom{d}{k}-1\right)^{h+i}}(1+h/m)(1+(h-1)/m)\cdots(1-(i-1)/m) \\
&\times \left(\sum_{j=0}^{l-i-1}\binom{l-i-1}{j}(-1)^{l-i-1-j}\left(\left(i\left(\binom{d}{k}-1\right)+l\right)\right.\right. \\
&\times \left.\left.\left(1+\left((i+j+1)/\left(\binom{d}{k}-1\right)\right)/m\right)^{m-i-1} - i\binom{d}{k}\left(1+\left((i+j)/\left(\binom{d}{k}-1\right)\right)/m\right)^{m-i-1}\right)\right) \\
&/(1+h/m)^{m+h-1}
\end{aligned}$$

when $0 \le i < l$ and

$$\left(\frac{(m+h)!}{(m-i)!}\right)\left(\frac{B_1}{\left(\binom{d}{k}-1\right)^{m+h-1}(m+h)^{m+h-1}}\right) =$$

$$\frac{1}{\left(\binom{d}{k}-1\right)^{h+l}} l\binom{d}{k}(1+h/m)(1+(h-1)/m)\cdots(1-(l-1)/m)$$
$$\times\left(1+\left(l/\left(\binom{d}{k}-1\right)\right)/m\right)^{m-l-1}/\left(1+h/m\right)^{m+h-1}$$

when $i = l$. If we let $n \longrightarrow \infty$, we have $\frac{f_l(m,d,k)}{\beta(n)} \longrightarrow$

$$\sum_{i=0}^{l}\binom{l}{i}\left(\frac{(-1)\left(\binom{d}{k}-1\right)k!}{d!}\right)^{i}\left(\frac{1}{l!(k!)^l}\right)\left(\frac{(d-k)!}{\binom{d}{k}-1}\right)^{h+i}\left(\frac{1}{e^h}\right)e^{\frac{i}{\binom{d}{k}-1}}\times$$
$$\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)^{l-i-1}\left(\left(i\left(\binom{d}{k}-1\right)+l\right)e^{\frac{1}{\binom{d}{k}-1}}-i\binom{d}{k}\right)$$

or

$$\left(\frac{1}{l!(k!)^l}\right)\left(\frac{(d-k)!}{e\left(\binom{d}{k}-1\right)}\right)^{h}\left(\sum_{i=0}^{l}\binom{l}{i}\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)^{i}\left(e^{1/\left(\binom{d}{k}-1\right)}-1\right)^{l-i-1}\right.$$
$$\left.\left(\left(i\left(\binom{d}{k}-1\right)+l\right)e^{1/\left(\binom{d}{k}-1\right)}-i\binom{d}{k}\right)\right)$$

A little more work on the sum and use the fact that $h = b(l-1)$ shows that $\frac{f_l(m,d,k)}{\beta(n)} \longrightarrow$

$$\left(\frac{1}{\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)}\right)\left(\frac{1}{l!(k!)^l}\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/\left(\binom{d}{k}-1\right)}A\right)lC^{l-1}$$

where

$$A = \left(e^{\frac{1}{\binom{d}{k}-1}}-1-\frac{e^{1/\left(\binom{d}{k}-1\right)}}{\binom{d}{k}}\right)$$

and

$$C = \left(e^{\frac{1}{\binom{d}{k}-1}}-1-\frac{e^{1/\left(\binom{d}{k}-1\right)}}{\binom{d}{k}}\right)\left(\frac{(d-k)!}{e\left(\binom{d}{k}-1\right)}\right)^{b}.$$

If we sum all the $\frac{f_l(m,d,k)}{\beta(n)}$ from $l = 2$ up to ∞, we have the following limit:

$$\lim_{n\to\infty}\frac{\sum_{l=2}^{\infty}f_l(m,d,k)}{\beta(n)} = \sum_{l=2}^{\infty}\lim_{n\to\infty}\frac{f_l(m,d,k)}{\beta(n)} =$$
$$\sum_{l=2}^{\infty}\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\frac{1}{l!(k!)^l}\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/\left(\binom{d}{k}-1\right)}A\right)lC^{l-1}$$
$$=$$
$$\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/\left(\binom{d}{k}-1\right)}A\right)\sum_{l=2}^{\infty}\left(\frac{1}{l!(k!)^l}\right)lC^{l-1} \tag{2.26}$$

Since

$$\sum_{l=2}^{\infty}\left(\frac{1}{l!(k!)^l}\right)lC^{l-1}=\frac{1}{k!}\sum_{l=2}^{\infty}\left(\frac{1}{(l-1)!}\right)\left(\frac{C}{k!}\right)^{l-1}=\frac{1}{k!}\left(e^{\frac{C}{k!}}-1\right),$$

We have the limit of the sum is

$$\lim_{n\to\infty}\frac{\sum_{l=2}^{\infty}f_l(m,d,k)}{\beta(n)}=$$
$$\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/(\binom{d}{k}-1)}A\right)\frac{1}{k!}\left(e^{\frac{C}{k!}}-1\right). \tag{2.27}$$

Similarly, if we use (2.26), we can find the limit of $\sum_{l=2}^{\infty}\frac{lf_l(m,d,k)}{\beta(n)}$ and the limit is:

$$\lim_{n\to\infty}\frac{\sum_{l=2}^{\infty}lf_l(m,d,k)}{\beta(n)}=\sum_{l=2}^{\infty}\lim_{n\to\infty}\frac{lf_l(m,d,k)}{\beta(n)}=$$
$$\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/(\binom{d}{k}-1)}A\right)\sum_{l=2}^{\infty}\left(\frac{l}{l!(k!)^l}\right)lC^{l-1}.$$

Since

$$\sum_{l=2}^{\infty}\left(\frac{l}{l!(k!)^l}\right)lC^{l-1}=\frac{1}{k!}\sum_{l=2}^{\infty}\left(\frac{l}{(l-1)!}\right)\left(\frac{C}{k!}\right)^{l-1}=\frac{1}{k!}\left(\left(\frac{C}{k!}+1\right)e^{\frac{C}{k!}}-1\right),$$

therefore

$$\lim_{n\to\infty}\frac{\sum_{l=2}^{\infty}lf_l(m,d,k)}{\beta(n)}=$$
$$\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/(\binom{d}{k}-1)}A\right)\frac{1}{k!}\left(\left(\frac{C}{k!}+1\right)e^{\frac{C}{k!}}-1\right). \tag{2.28}$$

Notice that $f_1(m,d,k)\sim Z_{m_1}$ only, therefore the average number of (d,k)-trees in a (d,k)-forest, when n is big, $d=(b+1)a$ and $k=ba$, is $\frac{\sum_{l=1}^{\infty}lf_l(m,d,k)}{\sum_{l=1}\infty f_l(m,d,k)}=\frac{\sum_{l=2}^{\infty}lf_l(m,d,k)}{\sum_{l=2}^{\infty}f_l(m,d,k)}$. By using (2.27) and (2.28), we find the limit is:

$$\frac{\sum_{l=2}^{\infty}lf_1(m,d,k)}{\sum_{l=2}^{\infty}f_l(m,d,k)}=$$
$$\frac{\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/(\binom{d}{k}-1)}A\right)\frac{1}{k!}\left(\left(\frac{C}{k!}+1\right)e^{\frac{C}{k!}}-1\right)}{\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/(\binom{d}{k}-1)}A\right)\frac{1}{k!}\left(e^{\frac{C}{k!}}-1\right)} \tag{2.29}$$
$$=\frac{\left(\left(\frac{C}{k!}+1\right)e^{\frac{C}{k!}}-1\right)}{\left(e^{\frac{C}{k!}}-1\right)}.$$

We summarize these results in the theorem and corollary below.

Theorem 2.3 *When $d = (b+1)a$ and $k = ba$ where a,b are natural numbers, the average number of (d,k)-trees in a (d,k)-forest is asymptotic to*

$$\frac{\left(\left(\frac{C}{k!}+1\right)e^{\frac{C}{k!}}-1\right)}{\left(e^{\frac{C}{k!}}-1\right)} \tag{2.30}$$

where

$$C = \left(e^{\frac{1}{\binom{d}{k}-1}} - 1 - \frac{e^{1/(\binom{d}{k}-1)}}{\binom{d}{k}}\right)\left(\frac{(d-k)!}{e\left(\binom{d}{k}-1\right)}\right)^{b}.$$

Corollary 2.4 *The average number of d-dimensional HP-hypertrees in a d-dimensional HP-forest is asymptotic to*

$$\frac{\left(\left(\frac{C}{(d-1)!}+1\right)e^{\frac{C}{(d-1)!}}-1\right)}{\left(e^{\frac{C}{(d-1)!}}-1\right)} \tag{2.31}$$

where

$$C = \left(e^{\frac{1}{d-1}} - 1 - \frac{e^{1/(d-1)}}{d}\right)\left(\frac{1}{e\,(d-1)}\right)^{d-1}.$$

In particular, the average number of 3-dimensional HP-hyptrees in a 3-dimensional HP-forest is 2.0008

Now, let us consider the general case. When d, k and l fixed, we observe that (d,k)-forest with l components does not exists at all natural number $n > (d-k)+kl$. When forest with l components exists at a particular n, then it doesn't exist at $n+1$, $n+2, \ldots, n+(d-k)-1$ and then it exists at $n+(d-k)$. Also, forest with $l+1$ components doesn't exist at $n, n+1, \ldots, n+k-1$ and it exists at $n+k$. From this observation, we can group the number of components that the forests exist at the same size of the hypergraph into a single group. A little calculation shows that we have $\frac{d-k}{gcd(d-k,k)}$ many groups. Note that the special case we discussed above is the case when $\frac{d-k}{gcd(d-k,k)} = 1$. If we pick the smallest number of components within each group to represent that group and write it as l_q, $q = 1, 2, \ldots, \frac{d-k}{gcd(d-k,k)}$, then $l_q = q$. Now, we can find the average number of components of (d,k)-forests within each group one by one.

For a particular group l_q, we define $\beta(n,q) = \frac{n!\left(\left(\binom{d}{k}-1\right)m_1\right)^{m_1-1}}{m_1!((d-k)!)^{m_1}}$ where n is a natural number such that $(d,k)-forest$ with l_q components exists and $n = m_1(d-k)+kl_q$. Consider a forest with l components, m edges and size n that belongs to the same group. We have $n = m(d-k)+kl = m_1(d-k)+kl_q$ and so $m_1 = m+h$ where $h = \frac{k}{gcd(d-k,k)}r$ and r is certain positive integer. With this setup, we can follow in the same way as in the special case above and we found that, when $n \to \infty$,

$$\sum_{r=0 or r=1}^{\infty} \lim_{n\to\infty} \frac{f_l}{\beta n} \to$$
$$\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/(\binom{d}{k}-1)}A\right)\left(\frac{A}{k!}\right)^{l_q-1}\left(\frac{1}{k!}\right)^{l_q}\times$$
$$\sum_{r=0 or r=1}^{\infty}\left(\frac{1}{\left(\frac{d-k}{gcd(d-k,k)}r+l_q-1\right)_{(l_q-1)}}\right)\left(\frac{1}{\left(\left(\frac{d-k}{gcd(d-k,k)}\right)r\right)!}\right)\left(\left(\frac{A}{k!}\right)\left(\frac{(d-k)!}{e\left(\binom{d}{k}-1\right)}\right)^{\frac{k}{d-k}}\right)^{\frac{d-k}{gcd(d-k,k)}r}$$

and

$$\sum_{r=0 or r=1}^{\infty} \lim_{n\to\infty} \frac{lf_l}{\beta n} \to$$
$$\left(1/\left(e^{\frac{1}{\binom{d}{k}-1}}-1\right)\right)\left(\left(e^{\frac{1}{\binom{d}{k}-1}}-\binom{d}{k}\right)\left(\frac{-e^{\frac{1}{\binom{d}{k}-1}}}{\binom{d}{k}}\right)+e^{1/(\binom{d}{k}-1)}A\right)\left(\frac{A}{k!}\right)^{l_q-1}\left(\frac{1}{k!}\right)^{l_q}\times$$
$$\sum_{r=0 or r=1}^{\infty}\left(\frac{\frac{d-k}{gcd(d-k,k)}r+l_q}{\left(\frac{d-k}{gcd(d-k,k)}r+l_q-1\right)_{(l_q-1)}}\right)\left(\frac{1}{\left(\left(\frac{d-k}{gcd(d-k,k)}\right)r\right)!}\right)\left(\left(\frac{A}{k!}\right)\left(\frac{(d-k)!}{e\left(\binom{d}{k}-1\right)}\right)^{\frac{k}{d-k}}\right)^{\frac{d-k}{gcd(d-k,k)}r}$$

where A is defined in the special case and the sums start at $r=1$ when $q=1$ and at $r=0$ when $q>1$.

With the help of the above formulas, Theorem 2.3 and Theorem 2.2, we get the Table 2.1.

The average number of trees in a edge rooted (d,k)-forest can also be obtained by a similar approach.

Table 2.1: Average Number of Components in a (d,k)-forest

(d,k)	q	limit value	$n=12$	$n=13$
$(3,1)$	1	3.000602687	$\varnothing$	1.580972649
	2	2.001205156	2.001203732	$\varnothing$
$(3,2)$	1	2.000838869	1.004255875	1.005977298
$(4,1)$	1	4.000000964	$\varnothing$	1
	2	2.000004686	$\varnothing$	$\varnothing$
	3	3.000001874	3	$\varnothing$
$(4,2)$	1	2.000656280	1.414027545	$\varnothing$
$(4,3)$	1	2.000007175	1.000034210	1.000052513

2.4 Enumeration of 3-trees

In the first section of this chapter we defined d-trees. Here we will use egfs of various rooted cases to enumerate 3-trees.

A *1-rooted* 3-tree has as its root a single unlabeled, non-cutting vertex. Let a_n be the number of 1-rooted 3-tree with n labeled vertices. Then $a_0 = 1$, $a_1 = 0$ and $a_2 = 1$. Let $y_1 = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$ be the egf of 1-rooted 3-trees and $t(x)$ be the egf of unrooted labeled trees. Note that $t(x) = \sum_{i=1}^{\infty} n^{n-2} \frac{x^n}{n!}$. Let $T(x,y) = \frac{t(xy)-xy}{y} = \sum_{n=2}^{\infty} a_n \frac{x^n y^{n-1}}{n!}$. Then $T(x,y)$ is a kind of egf of unrooted labeled trees where the degree of y gives the number of edges in a tree with order n. Notice that $T(x,y)$ counts only trees with order at least 2.

A *2-rooted* 3-tree has as its root 2 ordered, unlabeled vertices of an edge and there is exactly one edge that contains both the root vertices. Let b_n be the number of 2-rooted 3-trees with n labeled vertices and let $y_2 = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!}$ be the egf of 2-rooted 3-trees. Then we have the following lemma.

Lemma 2.5 *The egfs y_1 and y_2 for 1-rooted 3-trees and 2-rooted 3-trees satisfy the following two equations.*

$$y_1 = T(xe^{y_1}, e^{y_2}) = \frac{t(xe^{y_1}e^{y_2})}{e^{y_2}} - xe^{y_1} \tag{2.32}$$

$$y_2 = xe^{y_1}e^{2y_2} = xe^{y_1+2y_2} \tag{2.33}$$

Proof : Consider the case of 1-rooted 3-trees. If we take all the edges incident with the root vertex together and remove from them the root vertex, we get an edge set of an ordinary unrooted labeled tree. Therefore, we can construct a 1-rooted 3-tree by first adding an ordinary labeled tree to the root vertex (i.e. add the root vertex to all the edges of an ordinary tree), and then we add other 1-rooted 3-trees, one by one to the vertices of the tree (we can order the vertices and edges of the tree by using the labels of the vertices). Then we add 2-rooted 3-trees to the edges of the tree one by one (the edges and their two end vertices are ordered). The end result is a 1-rooted 3-tree. This implies that, to get the egf y_1, we just have to subsitute xe^{y_1} to x and e^{y_2} to y into the egf $T(x,y)$. Therefore, we have $y_1 = T(xe^{y_1}, e^{y_2}) = \frac{t(xe^{y_1}e^{y_2})}{e^{y_2}} - xe^{y_1}$. For 2-rooted 3-trees, we can construct them by first starting with a special hyperedge. This edge has one labeled vertex and two unlabeled vertices. The two unlabeled vertices are ordered. The order of the unlabeled vertices can be used to induce an order to the sides of the edge, in particular, the two sides with the labeled vertex. Now we add 2-rooted 3-trees to the two sides with the labeled vertex, one by one. Then we add 1-rooted 3-trees to the labeled verex, identify the roots with the labeled vertex. The end result is a 2-rooted 3-tree. This process shows us that the egf y_2 can be obtained by multiplying x with e^{y_1} and e^{2y_2}. Therefore, we have $y_2 = xe^{y_1}e^{2y_2} = xe^{y_1+2y_2}$. □

Let F be the egf of 3-trees rooted at a vertex. Then we can write the egf $F = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$ in terms of y_1 and y_2.

Lemma 2.6 *Let F be the egf of rooted 3-trees, then we have*

$$F = xe^{y_1} \tag{2.34}$$

Proof : If we arrange many 1-rooted trees together and identify the roots as a single root vertex, the egf of this structure is e^{y_1}. Now we fill the unlabeled root vertex with a label. We get a rooted 3-tree. It's egf is xe^{y_1}. $\Box$

Now, we solve the equations (2.32), (2.33) and extract the coefficients of (2.32), (2.33) and (2.34), we have the following formulas:

$$\left(\frac{d^n y_1}{dx^n}\right)_{x=0} = \sum_{r=0}^{n} \binom{n}{r} A_{n,r} B_{n,r} - nC_n \tag{2.35}$$

$$\left(\frac{d^n y_2}{x^n}\right)_{x=0} = nD_n \tag{2.36}$$

$$\left(\frac{d^n F}{dx^n}\right)_{x=0} = C_n \tag{2.37}$$

where the $A_{n,r}$, $B_{n,r}$, C_n and D_n are given below.

First we have

$$A_{n,r} = \begin{cases} \sum \left(\frac{(n-r)!}{\prod_{i=1}^{n-r} (i!)^{k_i} k_i!} (-1)^{\sum_{i=1}^{n-r} k_i} \prod_{j=1,k_j\neq 0}^{n-r} \left(\frac{d^j y_2}{dx^j}\right)^{k_j} \right) & n > r \geq 0 \\ 1 & r = n. \end{cases} \tag{2.38}$$

The summation is over all partitions $(k_1, k_2, \ldots, k_{n-r})$ of $n-r$ such that each k_i for $i = 1, 2, \ldots, n-r$ is a nonnegative integer and $\sum_{i=1}^{n-r} ik_i = n-r$.

Second comes

$$B_{n,r} = \begin{cases} \sum \left(\frac{r!}{\prod_{i=1}^{r} (i!)^{k_i} k_i!} \left(\sum_{i=1}^{r} k_i\right)^{\left(\sum_{i=1}^{r} k_i\right)-2} \prod_{j=1,k_j\neq 0}^{r} (b_j)^{k_j} \right) & n \geq r > 0 \\ 0 & \mathrm{r} = 0 \end{cases}$$

$$\textit{with} \tag{2.39}$$

$$b_j = \begin{cases} j \left(\sum \left(\frac{(j-1)!}{\prod_{i=1}^{j-1} (i!)^{h_i} h_i!} \prod_{m=1,h_m^{j-1}\neq 0} j-1 \left(\frac{d^m y_1+y_2}{dx^m}\right)^{h_m} \right) \right) & r \geq j > 1 \\ 1 & j = 1. \end{cases}$$

The summation in $B_{n,r}$ is over all partitions $(k_1, k_2, \ldots, k_r)$ of r such that each k_i for $i = 1, 2, \ldots, r$ is a nonnegative integer and $\sum_{i=1}^{r} ik_i = r$. The summation in b_j is over

Table 2.2: Number of 1-rooted, 2-rooted, rooted 3-trees and unrooted 3-trees.

n	1-rooted 3-trees	2-rooted 3-trees	rooted 3-trees	unrooted 3-trees
1	0	1	1	1
2	1	4	0	0
3	6	39	3	1
4	82	584	24	6
5	1515	11985	425	85
6	36951	312522	9450	1575
7	1112083	9898168	269892	38556

all partitions $(h_1, h_2, \ldots, h_{j-1})$ of $j-1$ such that each h_i for $i = 1, 2, \ldots, j-1$ is a nonnegative integer and $\sum_{i=1}^{j-1} ih_i = j-1$.

Next we have

$$C_n = \begin{cases} \sum \left(\frac{(n-1)!}{\prod_{l=1}^{n-1} (l!)^{k_l} k_l!} \prod_{j=1, k_j \neq 0}^{n-1} \left(\frac{d^j y_1}{dx^j} \right)^{k_j} \right) & n > 1 \\ 1 & n = 1. \end{cases} \tag{2.40}$$

The summation is over all partitions $(k_1, k_2, \ldots, k_{n-1})$ of $n-1$ such that each k_i for $i = 1, 2, \ldots, n-1$ is a nonnegative integer and $\sum_{i=1}^{n-1} ik_i = n-1$.

And finally

$$D_n = \begin{cases} \sum \left(\frac{(n-1)!}{\prod_{l=1}^{n-1} (l!)^{k_l} k_l!} \prod_{j=1, k_j \neq 0}^{n-1} \left(\frac{d^j y_1 + 2y_2}{dx^j} \right)^{k_j} \right) & n > 1 \\ 1 & n = 1. \end{cases} \tag{2.41}$$

The summation is over all partitions $(k_1, k_2, \ldots, k_{n-1})$ of $n-r$ such that each k_i for $i = 1, 2, \ldots, n-1$ is a nonnegative integers and $\sum_{i=1}^{n-1} ik_i = n-1$.

We used these formulas to calculate the entries in Table 2.2.

To extend this approach to enumerate 4-trees would clearly require substantially more effort.

CHAPTER 3

Threshold problems for random hypergraphs

There are several important problems that have been solved for graphs but not for hypergraphs. Two of these ask for probabilistic thresholds for monotone properties of random hypergraphs and both involve hypertrees.

The first seeks an approximation to the number of edges in a random hypergraph sufficient to insure that, with high probability, it has a spanning hypertree. We will focus in the simplest unsolved cases, which deal with spanning $(3,1)$-trees and $(3,2)$-trees. We provide upper and lower bounds for these approximations.

The second problem concerns maximum matchings in hypergraphs. We want to estimate the number of edges in a random hypergraph sufficient to insure that, with high probability, it has a spanning set of independent edges. As above, the simplest case involves 3-uniform hypergraphs. We have also determined bounds for this threshold. We have two methods for determining upper bounds. First we used an algorithmic approach. But a better bound can be found following Krivelevich [Kr97] and this makes use of a special class of hypertrees.

There is another asymptotic problem for hypergraphs that should be mentioned here. The Turan number, $t_3(n,4)$, is the maximum number of edges in a 3-uniform hypergraph of order n that has no complete hypergraph of order 4. Turan conjectured that

$$t_3(n,4) \sim \frac{5}{9}\binom{n}{3}. \tag{3.1}$$

A simple construction shows that the right side of 3.1 is a lower bound for $t_3(n,4)$. Professor Erdös liked this problem so much that he posted a bounty of \$1000 for its solution. We have no progress to report on this question.

3.1 Random hypergraphs

We define $H(n,d,p)$ as the probability space of all labelled d-uniform hypergraphs with vertex set $V = \{1,\ldots,n\}$. A hypergraph $H = (V,E) \in H(n,d,p)$ has probability $P(H) = p^{|E|}(1-p)^{\binom{n}{d}-|E|}$. We called p the *edge probability* and, of course, $0 < p < 1$. Thus each subset $\{v_1, v_2, \ldots, v_d\} \in \binom{V}{d}$ is chosen independently to be an edge of a random hypergraph H with probability p.

Suppose $Q \subseteq H(n,d,p)$ is a hypergraph property. We say that a random hypergraph $H \in H(n,d,p)$ has Q *with high probability (whp)* or *almost surely (a.s.)* if the probability of Q tends to 1 as n approaches infinity.

It is well-known that if Q is a monotone property, then there exists a threshold function p_t such that

$$P(Q) \to \begin{cases} 0 & \text{if } \frac{p}{p_t} \to 0 \\ 1 & \text{if } \frac{p}{p_t} \to \infty \end{cases}$$

Note that the limits indicted by the arrows are all taken as n approaches infinity.

We illustrate with a threshold for connectedness in random d-uniform hypergraphs found in [PaR85]. Recall from section 1.1 that a hypergraph H is connected if there is a $u-v$ path between any two vertices u and v. Clearly connectedness is a monotone property. Therefore a threshold function exists. It was determined in [PaR85] that

$p_t = \frac{(d-1)!}{n^{d-1}} log\, n$ is a threshold for connectness. In fact, the following much stronger result is proved in [PaR85]. Let $C \subseteq H(n, d, p)$ be the property of connectedness.

Theorem 3.1 *If the edge probability is* $p = \frac{(d-1)!}{n^{d-1}}(\ln n + x)$, *where x is fixed, then*

$$P(C) \to e^{-e^{-x}} \tag{3.2}$$

For definitions and random graph methods not explained here, the reader can consult an array of books, such as [Bo85], [Pa85], [AS92] and [JaLR90].

3.2 Spanning trees in random hypergraphs

Suppose $Q \subseteq H(n, d, p)$ is the hypergraph property that $H \in Q$ has a spanning (d, k)-tree. Then clearly Q is monotone and so there is a threshold function. In this section, we will establish upper and lower bounds for these thresholds when $d = 3$.

First observe that it follows from theorem 3.1 above that if the edge probability is

$$p = \frac{2!}{n^2}(\ln n - \omega_n), \quad where\ \omega_n \to +\infty, \tag{3.3}$$

then $P(Q) \to 0$. That is, almost surely the random hypergraph is not connected and hence has no spanning $(3, 1)$-tree, $(3, 2)$-tree or mixed 3-tree. Hence $\frac{2!}{n^2} \ln n$ is a lower bound on the spanning tree threshold for all of these types of trees.

Now we focus on the upper bound for spanning $(3, 1)$-trees. Our approach is algorithmic. We assume that $p >> \frac{2!}{n^2}$ and so whp there are many edges. Here is how the algorithm works. We pick an edge, say $\{u, v, w\}$, at random and form a $(3, 1)$-tree T to start with. Then we try to find another edge $\{u, x, y\}$ where $\{x, y\} \cap \{u, v, w\} = \emptyset$. We extend the $(3, 1)$-tree T with the new edge and get a new, bigger $(3, 1)$-tree T. We repeat this step until this $(3, 1)$-tree reaches a size that depends on $0 < \epsilon < 1$. Then we add new edges made from a pair of the remaining vertices and *any* vertex of the $(3, 1)$-tree under construction. Here is a more concrete description of the algorithm.

Greedy Algorithm for a Spanning $(3,1)$-tree

INPUT a uniform hypergraph H and an edge $\{u, v, w\}$ of H

Choose $0 < \epsilon < 1$

$T \longleftarrow \{\{u, v, w\}\}$

$S \longleftarrow \{w\}$

WHILE $|V(T)| < \epsilon |V(H)|$ DO

$S_1 \longleftarrow V(H) - V(T)$

    Pick $w \in S$

    REPEAT

        Pick $x_1 \in S_1$

        $S_2 \longleftarrow S_1 - \{x_1\}$

        REPEAT

            pick $x_2 \in S_2$

            $S_2 \longleftarrow S_2 - \{x_2\}$

        UNTIL $\{w, x_1, x_2\} \in E(H)$ OR $S_2 = \emptyset$

        IF $\{w, x_1, x_2\} \in E(H)$ THEN $T \longleftarrow T \cup \{\{w, x_1, x_2\}\}$

            $S \longleftarrow \{x_1\}$

        $S_1 \longleftarrow S_1 - \{x_1\}$

    UNTIL $\{w, x_1, x_2\} \in E(H)$

$R \longleftarrow V(H) - V(T)$

WHILE $R \neq \emptyset$ DO

    Pick $y_1 \in R$

    $R_1 \longleftarrow R - \{y_1\}$

    $T_1 \longleftarrow V(T)$

    REPEAT

        pick $y_2 \in R_1$

$R_1 \longleftarrow R_1 - \{y_2\}$

pick $v \in T_1$

$T_1 \longleftarrow T_1 - \{v\}$

UNTIL $\{y_1, y_2, v\} \in E(H)$

$T \longleftarrow T \cup \{\{y_1, y_2, v\}\}$

$R \longleftarrow V(H) - V(T)$

OUTPUT T

This algorithm terminates and outputs a spanning $(3,1)$-tree T of the random hypergraph H if it can find one. Notice that after we expended the first input edge to a $(3,1)$-tree of size $\epsilon|V(H)|$ in the first half of the algorithm, the input hypergraph H loses randomness and therefore the probability of successful adding edges in the second half of the algorithm is not easy to compute. To overcome the difficulty and analyze the algorithm, we need to use a trick that is well-known to probabilists. We will use two colors, blue and green, for the edges of our random graphs. Thus we express the edge probability p as $p = p_1 + p_2 - p_1 p_2$, where p_1 is the probability of having a blue hyperedge between three vertices of H and p_2 is the probability of having a green hyperedge. In fact, this setup enables us to decompose the 3-uniform random hypergraph H into the union of two 3-uniform random hypergraphs H_1 and H_2, where $H_1 \in H(n, 3, p_1)$ has all the random blue hyperedges of H and $H_2 \in H(n, 3, p_2)$ has all the random green hyperedges. If we use only H_1 in the first half and use only H_2 in the second half of our algorithm, we can find a spanning $(3,1)$-tree of H without losing randomness.

Let $H \in H(n, p, 3)$ with $p = p_1 + p_2 - p_1 p_2$. Let $H_1 \in H(n, 3, p_1)$ be the random blue hypergraph, $H_2 \in H(n, 3, p_2)$ be the random green hypergraph where H is the union of H_1 and H_2. Let $r = \epsilon|V(H)|$, r odd. In many of the proofs we deal with quantites such as $\epsilon|V(H)|$ where $\epsilon > 0$ is very small but $\epsilon|V(H)|$ is very large. Rather than using the round-off notation, we treat $\epsilon|V(H)|$ as an integer itself. The validity of the proofs still holds. Let $T = \{\{u, v, w\}\} \subset E(H_1)$, that is, T contains the blue edge we used to start the algorithm.

The probability of adding a new blue edge $\{w, x_1, x_2\}$ to T is

$$1 - (1 - p_1)^{\binom{n-3}{2}}. \tag{3.4}$$

Next we search for another blue edge $\{x, y_1, y_2\}$ with $\{y_1, y_2\} \cap \{u, v, w, x_1, x_2\} = \emptyset$. The probability for success in our second attempt is

$$1 - (1 - p)^{\binom{n-5}{2}}. \tag{3.5}$$

We repeat this step until $|V(T)| = r$. The probability of successfully constructing a blue $(3,1)$-tree T of odd order r in this way is

$$\prod_{i=1}^{\frac{(r-3)}{2}} \left(1 - (1 - p_1)^{\binom{n-(2i+1)}{2}}\right) > \left(1 - (1 - p_1)^{\binom{n-(r-2)}{2}}\right)^{\frac{r-3}{2}}. \tag{3.6}$$

Note that the lower bound above does not approach 1 as $n \to \infty$ when $r = n$. So we terminate this part of the algorithm for $r = \epsilon n$, with ϵ to be determined later. And now we continue to extend our tree T in a slightly different way.

In the second half of the algorithm, we switch to H_2. First we pick a vertex z_1 in $R = V(Q) - V(T)$. The probability that z_1 belongs to an green edge of the form $\{z_1, z_2, z_3\}$ with $z_2 \in V(T)$ and $z_3 \in R$ is $1 - (1 - p_2)^{r(n-r-1)}$. The probability that this step can be repeated until the tree spans all vertices is

$$\prod_{i=0}^{\frac{(n-r-2)}{2}} \left(1 - (1 - p_2)^{(r+2i)(n-r-(2i+1))}\right) > \left(1 - (1 - p_2)^{(n-2)2}\right)^{\frac{n-r}{2}}. \tag{3.7}$$

From 3.6 we have

$$\begin{aligned} &\left(1 - (1 - p_1)^{\binom{n-(r-2)}{2}}\right)^{\frac{r-3}{2}} \\ &> \left(1 - e^{-p_1\binom{n-(r-2)}{2}}\right)^{\frac{r-3}{2}} \\ &> 1 - \left(\tfrac{r-3}{2}\right)\left(e^{-p_1\binom{n-(r-2)}{2}}\right). \end{aligned} \tag{3.8}$$

In order for

$$1-\left(\frac{r-3}{2}\right)\left(e^{-p_1\left(n-\frac{(r-2)}{2}\right)}\right)\longrightarrow 1 \tag{3.9}$$

when $n \longrightarrow \infty$, we must have $\frac{r-3}{2}e^{-p_1\left(n-\frac{(r-2)}{2}\right)} \longrightarrow 0$. Let $\frac{r-1}{2}e^{-p_1\left(\frac{n-r}{2}\right)} = \frac{1}{\omega_n}$ where $\omega_n \longrightarrow \infty$. Now $p_1 = \frac{2\ln\left(\frac{(r-3)\omega_n}{2}\right)}{(n-(r-2))^2} \geq \frac{2\ln(r-3)}{(n-(r-2))^2} \sim O\left(\frac{\ln n}{n^2}\right)$.

From 3.7 we have

$$\begin{aligned} &\left(1-(1-p_2)^{(n-2)2}\right)^{\frac{n-r}{2}} \\ &> \left(1-e^{-p_2 2(n-2)}\right)^{\frac{n-r}{2}} \\ &> 1-\left(\tfrac{n-r}{2}\right)\left(e^{-p_2 2(n-2)}\right). \end{aligned} \tag{3.10}$$

In order for

$$1-\left(\frac{n-r}{2}\right)\left(e^{-p_2 2(n-2)}\right)\longrightarrow 1 \tag{3.11}$$

when $n \longrightarrow \infty$, we need $\left(\frac{n-r}{2}\right)e^{-p_2(2n-4)} \longrightarrow 0$. If we let $\left(\frac{n-r}{2}\right)e^{-p_2(2n-4)} = \frac{1}{\omega_n}$ where $\omega_n \longrightarrow \infty$, then $p_2 = \frac{\ln\left(\frac{(n-r)\omega_n}{2}\right)}{2n-4} \geq \frac{\ln n}{2n-4} \sim O\left(\frac{\ln n}{n}\right)$. Now, if we combine all the blue edges and green edges we collected, we get a spanning $(3,1)$-tree of H and we have $p = p_1 + p_2 - p_1p_2 \sim O\left(\frac{\ln n}{n}\right)$. This result is summarized in the following theorem.

Theorem 3.2 *If the edge probability of a random hypergraph H in $H(n,3,p)$ is $p \geq \left(\frac{1}{2}+\epsilon\right)\frac{\ln n}{n}$, then* ***whp*** *$H$ has a spanning $(3,1)$-tree. Hence the threshold for a spanning $(3,1)$-tree is $O\left(\frac{\ln n}{n}\right)$.*

By modifying the algorithm above, we obtain an upper bound for the spanning $(3,2)$-tree threshold. Let $H \in H(n,3,p)$ be the 3-uniform random hypergraph. We start with an edge in H, then we extend it as a $(3,2)$-tree by adding a new edge with exactly one new vertex x_1 outside the $(3,2)$-tree. Suppose the edge we just added to the $(3,2)$-tree is $\{u,v,x_1\}$, then what we do next is to find another edge of the form $\{u,x_1,x_2\}$ or $\{v,x_1,x_2\}$ where x_2 is a new vertex outside our current $(3,2)$-tree. We keep on expanding our $(3,2)$-tree till our hypertree's size reaches $r = \epsilon|V(H)|$, where $0 < \epsilon < 1$. Then we expand our $(3,2)$-tree in a different way. We add the remaining vertices, one by one, to the hypertree by finding an edge with the choosen vertex and two vertices from an edge of our current $(3,2)$-tree. If this process terminates, we get a spanning $(3,2)$-tree of H. Slightly modify this, we have the following algorithm.

Greedy Algorithm for a Spanning $(3,2)$-tree

INPUT a uniform hypergraph H and an edge $\{u,v,w\}$ of H

Choose $0 < \epsilon < 1$

$T \longleftarrow \{\{u,v,w\}\}$

$T_1 \longleftarrow \{u,v\}$

$S \longleftarrow \{w\}$

WHILE $|V(T)| < \epsilon|V(H)|$ DO

  $S_1 \longleftarrow V(H) - V(T)$

  Pick the vertex $w \in S$

  Pick a vertex $u \in T_1$

  Pick the remaining vertex $v \in T_1 - \{u\}$

  REPEAT

    Pick $x_1 \in S_1$

    $S_1 \longleftarrow S_1 - \{x_1\}$

  UNTIL $\{w,u,x_1\} \in E(H)$ OR $\{w,v,x_1\} \in E(H)$

  IF $\{w,u,x_1\} \in E(H)$ THEN $T \longleftarrow T \cup \{\{w,u,x_1\}\}$

    $T_1 \longleftarrow \{w,u\}$

    ELSE IF $\{w,v,x_1\} \in E(H)$ THEN $T \longleftarrow T \cup \{\{w,v,x_1\}\}$

      $T_1 \longleftarrow \{w,v\}$

      $S \longleftarrow \{x_1\}$

$R \longleftarrow V(H) - V(T)$

WHILE $R \neq \emptyset$ DO

  Pick $y_1 \in R$

$T_1 \longleftarrow T$

REPEAT Pick $\{u,v,w\} \in T_1$

$T_1 \longleftarrow T_1 - \{\{u,v,w\}\}$

UNTIL $\{u,v,y_1\} \in E(H)$ OR $\{u,w,y_1\} \in E(H)$ OR $\{v,w,y_1\} \in E(H)$

IF $\{u,v,y_1\} \in E(H)$ THEN $T \longleftarrow T \cup \{\{u,v,y_1\}\}$

ELSE IF $\{u,w,y_1\} \in E(H)$ THEN $T \longleftarrow T \cup \{\{u,w,y_1\}\}$

ELSE IF $\{v,w,y_1\} \in E(H)$ THEN $T \longleftarrow T \cup \{\{v,w,y_1\}\}$

$R \longleftarrow V(H) - V(T)$

OUTPUT T

We follow the same trick to analyze the algorithm and decompose $H \in H(n,3,p)$ into the union of two random hypergraphs $H_1 \in H(n,3,p_1)$ and $H_2 \in H(n,3,p_2)$. Where $p = p_1 + p_2 - p_1p_2$. H_1 is the blue random hypergraph and H_2 is the green random hypergraph.

The probability of success in the first part with blue edge probability p_1 is

$$\begin{aligned}
&\left(1-(1-p_1)^{3(n-3)}\right)\prod_{i=2}^{r-3}\left(1-(1-p_1)^{2(n-(2+i))}\right)\\
&\geq \prod_{i=1}^{r-3}\left(1-(1-p_1)^{2(n-(2+i))}\right)\\
&\geq \prod_{i=1}^{r-3}\left(1-(1-p_1)^{2(n-r+1)}\right)\\
&\geq \left(1-(1-p_1)^{2(n-r+1)}\right)^{r-3}\\
&\geq 1-(r-3)(1-p_1)^{2(n-r+1)}\\
&\geq 1-\frac{r-3}{e^{p_1 2(n-r+1)}}.
\end{aligned}$$

We need $\frac{r-3}{e^{p_1 2(n-r+1)}} = \frac{1}{\omega_n} \longrightarrow 0$ and we find that this occurs when $p_1 \sim O(\frac{\ln n}{2n})$. Therefore, if $p_1 \sim O(\frac{\ln n}{2n})$, **whp** the first part will succeed.

The probability for second part to succeed with green edge probability p_2 is

$$\begin{aligned}
&\prod_{i=1}^{n-r}\left(1-(1-p_2)^{(2(r+i-1)-3)}\right)\\
&\geq \prod_{i=1}^{n-r}\left(1-(1-p_2)^{(2r-3)}\right)\\
&\geq \left(1-(1-p_2)^{(2r-3)}\right)^{(n-r)}\\
&\geq \left(1-e^{-p_2(2r-3)}\right)^{(n-r)}\\
&\geq 1-(n-r)e^{-p_2(2r-3)}.
\end{aligned}$$

Now we need $(n-r)e^{-p_2(2r-3)} \longrightarrow 0$. As before, we let $(n-r)e^{-p_2(2r-3)} = \frac{1}{\omega_n}$ where $\omega_n \longrightarrow \infty$ and solve for p_2. We find $p_2 \sim O\left(\frac{\ln n}{2n}\right)$. Therefore, if $p_2 \sim O\left(\frac{\ln n}{2n}\right)$, **whp** the second part succeeds.

Since $H \in H(n,3,p)$ is the union of the blue random hypergraph and green random hypergraph with $p = p_1 + p_2 \sim O\left(\frac{\ln n}{n}\right)$, we have the following theorem.

Theorem 3.3 *If the edge probability of a random hypergraph H in $H(n,3,p)$ is $p \geq \left(\frac{1}{2}+\epsilon\right)\frac{\ln n}{n}$, then **whp** H has a spanning $(3,2)$-tree. Hence the threshold for a spanning $(3,2)$-tree is $O\left(\frac{\ln n}{n}\right)$.*

Since a $(3,2)$-spanning tree is also a spanning mixed 3-tree, the upper bound of the threshold for the former also serves as a bound for the latter.

Theorem 3.4 *If the edge probability of a random hypergraph H in $H(n,3,p)$ is $p \geq \left(\frac{1}{2}+\epsilon\right)\frac{\ln n}{n}$, then **whp** H has a spanning mixed 3-tree. Hence the threshold for a spanning mixed 3-tree is $O\left(\frac{\ln n}{n}\right)$.*

Our approach would also work on (d,k)-trees, $d \geq 4$, but much more effort would be required. Furthermore, we would still have only crude uppper and lower bounds for the threshold.

3.3 Maximum matchings in hypergraphs

A subgraph of a hypergraph which is regular of degree 1 is called a *matching.* A *complete matching* is a matching that spans all vertices. We also call a complete matching a *1-factor.* Having a 1-factor is a monotone property. Erdös and Rényi established the corresponding random graph threshold [ErR66]. Perhaps it is not a surprise that this threshold coincides with the event that the minimum degree is at least 1, i.e. $\delta \geq 1$. For all even n let $F_n \subseteq H(n, 2, p)$ be the set of graphs of order n with a 1-factor. Next define the edge probability:

$$pn = \ln n + c_n. \tag{3.12}$$

Then, if $c_n \longrightarrow -\infty$, the second moment method shows that almost surely there are vertices of degree 0, i.e. $P(\delta \geq 1) \longrightarrow 0$. But $P(F_n) \leq P(\delta \geq 1)$, hence also $P(F_n) \longrightarrow 0$. So almost surely a random graph does not have a 1-factor when $c_n \longrightarrow -\infty$. So much for the trivial portion of the next theorem.

Theorem 3.5 (Erdös and Rényi) *With the edge probability defined by* $pn = \ln n + c_n$*, the probability that the random graph in* $H(n, 2, p)$ *has a 1-factor has the same limiting value as* $P(\delta \geq 1)$*, namely:*

$$P(\delta \geq 1) \rightarrow \begin{cases} 0 & \textit{if } c_n \rightarrow -\infty \\ e^{-e^{-c}} & \textit{if } c_n \rightarrow c \\ 1 & \textit{if } c_n \rightarrow +\infty \end{cases}$$

The hard part of the proof requires serious graph theory, namely Tutte's famous 1-factor theorem [Tu47], which characterizes graphs with complete matchings. Paul Catlin used to prove Tutte's theorem using Hall's matching theorem, so perhaps the latter is the basis for matching theorems in graphs. At any rate, graph theory is availiable when called for in the proof of Theorem 3.5 or for proving strengthened versions involving hitting times [Bo85].

Now an interesting thing happens when this problem is generalized to random 3-uniform hypergraphs in $H(n,3,p)$. The new problem is far more difficult! This phenomennon often happens when moving from graphs to hypergraphs. For example, recall that finding a maximum matching in a graph can be done in polynomial time, whereas the corresponding problem for 3-uniform hypergraphs is in NP [GaJ79].

Let us consider the matching problem in terms of uniform random hypergraphs. Let $Q_n \subseteq H(n,3,p)$. A *complete matching* M of $H \in Q_n$ is a collection of isolated edges of H that span all n vertices of H. An easy calculation shows that if

$$pn^2 = 2\ln n + \omega_n, \qquad \omega_n \to +\infty, \tag{3.13}$$

then almost surely every vertex belongs to at least one edge. This provides a lower bound for the threshold for a complete matching.

Schmidt and Shamir [ScS83] found an upper bound for this threshold (for r-uniform hypergraphs) with an insightful application of the second moment method. On translating from their probability model to ours and taking $r = 3$ we found that if

$$pn^{\frac{3}{2}} = \omega_n \to \infty \tag{3.14}$$

then almost all 3-uniform hypergraphs have a complete matching.

Erdös was preaching about this problem as early as 1985. Evidently it originated with Schmidt and Shamir who passed it along to Uncle Paul.

There are other versions of it but they all lead only to upper and lower bounds for the threshold. The root of the problem seems to be the absence of a suitable description or characterization of factorizations in triangles or tripartite matchings.

There is another important, relevant result that must be mentioned. It shows that a large matching can still be found when the edge probability is slightly smaller than that of (3.14). Using a closely related probability model for r-uniform hypergraphs, de la Vega [Ve82] found that a necessary and sufficient condition for a random hypergraph to have a matching that spans all but $o(n)$ vertices is

$$pn^{r-1} = \omega_n \to \infty. \tag{3.15}$$

The proof of the necessity is a straightforward argument using Chebyshev's inequality. But the sufficiency makes nice use of Markov chains to show that a greedy algorithm will produce a big matching if given enough edges. Note that the edge probability in (3.15) is indeed smaller than that of (3.14)

There is a simple way to get a upperbound of the threshold. Although the upperbound is not as good as the above upperbounds, this method is easy to follow and it relates the problem to the threshold of complete matching of random bipartite graphs. Here is how we construct the complete matching. Given an order n 3-uniform random hypergraph H with edge probability p. We divide the vertex set $V(H)$ into 3 equal size vertex set V_1, V_2 and V_3. Therefore, we have $|V_1| = |V_2| = |V_3| = \frac{n}{3}$. Next, we arbitrary pair up vertices in V_1 with vertices in V_2. Now, we have a set of vertex pairs. Let us call this set V_4. Note that $|V_3| = |V_4| = \frac{n}{3}$. If we treat the elements in V_4 as vertices, then we can construct a new random bipartite graph G with vertex sets V_3 and V_4 and edge probability p. An edge in G that joins a vertex $\{u, v\} \in V_4$ and a vertex $w \in V_3$ can be viewed as an edge $\{u, v, w\}$ in H. Therefore the probability that G has a complete matching is an upperbound of the probability that H has a complete matching. It's already known that the threshold of complete matching in random bipartite graph is of $o\left(\frac{\ln n}{n}\right)$ (See [Bo85]). Therefore, we concludes that the threshold of complete matching in 3-uniform hypergraph is bounded above by $\frac{\ln n}{n} + \omega_n$ where $\omega_n \longrightarrow \infty$. We have the following theorem.

Theorem 3.6 *A 3-uniform hypergraph $H \in H(n, 3, \frac{\ln n}{n})$ **whp** contains a complete matching if 3 divides n.*

In his paper [Kr97], Michael Krivelevich introduced a new random graph algorithm to find a new and improved upper bound for the threshold of a triangle factor in a random graph. We find that this new algorithm for random graphs can be modified and be used to improve the upper bound in theorem 3.6. In the following sections, we shall provide the modified version of his algorithm and use it to find a better upper bound of the threshold of complete matching in 3-uniform random hypergraphs. Eventually, the algorithm leads to the following theorem.

Theorem 3.7 *A 3-uniform random hypergraph* $H \in H(n, 3, 57867n^{-\frac{3}{2}})$ ***whp*** *contains a complete matching, assuming 3 divides n.*

In order to describe our random hypergraph algorithm, we need some basic hypergraph structures. These hypergraph structures play the central role in constructing a complete matching of a random 3-uniform hypergraph.

3.3.1 Hypergraph H_0 and H_0-hypertree

The *hypergraph* H_0 has 4 vertices v_0, v_1, v_2, v_3 and 2 edges (v_0, v_1, v_2) and (v_1, v_2, v_3). The vertices v_0, v_3 are called *removable*, while the vertices v_1, v_2 are called the *kernel* of H_0.

You can see here that the hypergraph H_0 is not a matching over it's vertices. However, if we remove any one of the removable vertices, we get an edge (complete matching over the three vertices). A hypergraph H_0 can be extended naturally to a tree like structure together with a set of vertices. We call this tree like structure *H_0-hypertree* and the set of vertices *set of removable vertices.* The following is the recursive definition of H_0-hypertrees and its set of removable vertices:

(1) A hypergraph H_0 is an H_0-hypertree with the set of removable vertices $R = \{v_0, v_3\}$;

(2) If $T = (V, E)$ is an H_0-hypertree with the set of removable vertices R and H is a copy of H_0 with the set of removable vertices $\{u_0, u_3\}$ and kernel $\{u_1, u_2\}$ so that $V(H) \bigcap V(T) = V(H) \bigcap R = \{u_0\}$, then the graph $T' = (V', E')$, with $V' = V(T) \bigcup V(H)$ and $E' = E(T) \bigcup E(H)$, is an H_0-hypertree with the set of removable vertices $R' = R(T) \bigcup \{u_3\}$;

(3) Every H_0-hypertree can be obtained from hypergraph H_0 by applying (2).

Notice that an H_0-hypertree has the following properties.

Proposition 3.8 *If $T = (V, E)$ is an H_0-hypertree with its set of removable vertices R, we have*

(1) $|V(T)| \equiv 1 \ (mod\ 3)$.

(2) $|R| \geq \frac{|V(T)|}{3}$.

(3) For every $v \in R$, the hypergraph $T - \{v\}$ contains a complete matching.

Proof : We prove the above properties by induction. Let T be an H_0-hypertree and m be the number of H_0 hypergraphs in T. When $m = 1$, T is an H_0 hypergraph. So $|V(T)| = 4$, $|R| = 2$ and T satisfies the three properties. Suppose for any T with $m = k$, $k \geq 1$ satisfies the above properties. For any T with $m = k + 1$, T is obtained by identifing a removable vertex of a new H_0 hypergraph H_1 as one of the removable vertices of an H_0-hypertree T_1 with $m = k$. Therefore $|V(T)| = |V(T_1)| + 3$ and $|R(T)| = |R(T_1)| + 1$. Since T_1 satisfies properties (1) and (2), therefore $|V(T)| \equiv 1 \ (mod\ 3)$ and $|R(T)| \geq \frac{|V(T_1|}{3} + 1 \geq \frac{|V(T_1)|+3}{3} \geq \frac{|V(T)|}{3}$. Suppose v_0 is the removable vertex of H_1 that is identified as one of the removable vertices of T_1. Let $v \in R(T)$. If $v \in R(T_1)$, then $T - \{v\}$ contains a matching that covers T_1. If that matching contains v_0, than $H_1 - v_0$ is an hyperedge that covers all the vertices of H_1 other than v_0 and is independent of the matching of T_1. So we found a complete matching of

$T - \{v\}$. If the matching of T_1 in $T - \{v\}$ does not contain v_0, then $v = v_0$ and we add the hyperedge $H_1 - v_0$ to the matching to get a complete matching of $T - \{v\}$. If $v \notin R(T_1)$, then the hyperedge $H_1 - \{v\}$ and a matching in $T_1 - \{v_0\}$ form a complete matching of $T - \{v\}$. Therefore, T satisfies all three properties. By induction, we proved proposition 3.8. □

Like hypergraph H_0, if we take away a removable a vertex from a H_0-hypertree, the resultant subgraph has a complete matching.

3.3.2 A matching that covers all vertices

As a base to construct a proof for Theorem 3.7, the following proposition guarantees us to have a matching to start with, provided that the edge probability is big enough.

Proposition 3.9 ***Whp*** *every set of at least $n^{0.95}$ vertices of a random 3-uniform hypergraph $H \in H(n, 3, p)$, where $p = Cn^{-\frac{3}{2}}$ for any absolute constant $C > 0$, contains an edge.*

Proof: Given a subset V_0 of the vertex set $V(H)$ with size $|V_0| = n^{0.95}$. The Probability that the subgraph of H spanned by V_0 contains no edge, is

$$P\left[G \text{ has no edge}\right] = (1-p)^{\binom{|V_0|}{3}} < e^{-p\binom{|V_0|}{3}} < e^{-\frac{C}{n^{1.5}}\binom{|V_0|}{3}} = e^{-\Theta\left(\frac{Cn^{1.35}}{3!}\right)} = e^{-\Theta(n^{1.35})}. \tag{3.16}$$

Therefore, the probability of the existance of a size $n^{0.95}$ vertex set that contains no edge is bounded above by

$$\binom{n}{n^{0.95}} P\left[G \text{ has no edge}\right] \leq 2^n e^{-\Theta(n^{1.35})} = o(1). \tag{3.17}$$

Hence, **whp** every subgraph of H with size $n^{0.95}$ contains an edge. □

Corollary 3.10 *For any constant $C > 0$, a random 3-uniform hypergraph $H \in H(n, 3, p)$ with $p = Cn^{-\frac{3}{2}}$* ***whp*** *contains a matching, covering all but at most $n^{0.95}$ vertices.*

Proof: The matching can be constructed by picking edges one by one from subgraphs of H with size $n^{0.95}$. We can continue this process till we have less than $n^{0.95}$ not covered by the picked edges. □

After we find a matching to start with, our algorithm tries to find a H_0-forest. This forest can be used to expand the size of our matching. The follwing proposition and lemmas show us the forest exists.

Proposition 3.11 *Let $p_0 = 77n^{-\frac{3}{2}}$. Then* **whp** *for every triple of disjoint subsets $U', U", W$ of the vertex set of a random hypergraph $H \in H(n, d, p_0)$, satisfying $|U'| \geq \frac{n}{18}, |U"| \geq \frac{n}{6}$, $|W| \geq \frac{n}{3}$, there exists in H a copy of the hypergraph H_0, having it's kernel vertices in W, one of its removable vertices in U' and the other one in $U"$.*

Proof: Given U', $U"$ and W satisfy the above conditions. WLOG, we may assume that $|U'| = \frac{n}{18}$, $|U"| = \frac{n}{6}$ and $|W| = \frac{n}{3}$. The probability that no such H_0 exist is

$$P\left[no\ such\ H_0\right] = \left(1 - p^2\right)^{|U'||U"|\binom{|W|}{2}} < e^{-\frac{n}{18}\frac{n}{6}\binom{\frac{n}{3}}{2}p^2} < e^{-3n}. \tag{3.18}$$

Hence the probability that there exists U', $U"$ and W with no such kind of H_0 is bounded above by

$$\binom{n}{|U'|}\binom{n}{|U"|}\binom{n}{|W|}P\left[no\ such\ H_0\right] \leq \left(2^n\right)^3 e^{-3n} = o(1). \tag{3.19}$$

Therefore **whp** every triple U', $U"$, W has a copy of H_0 with it's kernel vertices in W, one removable vertices in U' and the other one in $U"$. □

Note that the constant 77 guarantees us that (3.18) holds.

Lemma 3.1 *If $p_0 = 77n^{-\frac{3}{2}}$, then for every integer k, satifying $4 \leq k \leq \frac{n}{6}$ and $k \equiv 1 \pmod 3$, a random hypergraph $H \in H(n, 3, p_0)$* **whp** *contains $\lfloor \frac{n}{6k} \rfloor$ vertex disjoint copies of H_0-hypertrees, each having k vertices.*

Proof. Let $p_0 = 77n^{-\frac{3}{2}}$, k fixed with $4 \leq k \leq \frac{n}{6}$ and $k \equiv 1 \pmod 3$. Suppose we already have t many H_0-hypertrees of size k where $0 \leq t < \frac{n}{6}$. It's sufficient for us to find one H_0-hypertree of size k from the remaining vertices. Consider the following algorithm. At step i, it generates a family of H_0-trees $\mathcal{T}_i = \{T_1, \ldots, T_m\}$ where T_l, $1 \leq l \leq m$, are vertex disjoint H_0-hypertrees. Associated with each member T_l of $\mathcal{T}_i$ is a vertex subset $U(T_l) \subseteq V(T_l)$. Let V_0 be the set of vertices of the t already found H_0-hypertrees. Let $V_1 = V(H) - V_0$ be the set of the remaining vertices.

<u>Algorithm to find an H_0-hypertree</u>

INPUT V_1 and any $\frac{n}{6}$ vertices $\{v_1, \ldots, v_{\frac{n}{6}}\}$ from V_1

$T_l \longleftarrow \{v_l\}$, $1 \leq l \leq \frac{n}{6}$

$V \longleftarrow V_1 - \{v_1, \ldots, v_{\frac{n}{6}}\}$

$\mathcal{T}_1 \longleftarrow \{T_1, \ldots, T_{\frac{n}{6}}\}$

$U(T_l) \longleftarrow \{v_l\}$, $1 \leq l \leq \frac{n}{6}$

$i \longleftarrow 1$

WHILE NOT EXISTS $T_l \in \mathcal{T}_i$ with $|T_l| = k$ DO

  Find $G = \{u_0, u_1, u_2, u_3\}$, a copy of H_0 with u_0 (one of the removable vertex) in $U' = \cup_{T \in \mathcal{T}_i} U(T)$, the other removable vertex u_3 in an arbitrary subset U" of $V_1 - \cup_{T \in \mathcal{T}_i} V(T)$ of size $\frac{n}{6}$ and $\{u_1, u_2\} \subset V_1 - \cup_{T \in \mathcal{T}_i} V(T) \cup U$"

  $T_j \longleftarrow T_j \cup G$ and $U(T_j) \longleftarrow U(T_j) \cup \{u_0, u_3\}$ where $T_j \in \mathcal{T}_i$ such that $u_0 \in U(T_j)$

  IF $|\cup_{T \in \mathcal{T}_i} V(T)| > \frac{n}{3}$ THEN $\mathcal{T}_i \longleftarrow \mathcal{T}_i - \{T_l\}$ where T_l is of minimum size in $\mathcal{T}_i$

  $\mathcal{T}_{i+1} \longleftarrow \mathcal{T}_i$

  $i = i + 1$

OUTPUT T_l where $|T_l| = k$

Notice that the average size of the H_0-trees increases when i increase. Within the WHILE loop, the size of U' is at least $\frac{n}{18}$ because of proposition 3.8 and at any time $|\cup_{T \in \mathcal{T}_i} V(T)| \geq \frac{n}{6}$. Also, $|V_1 - \cup_{T \in \mathcal{T}_i} V(T) \cup U"| \geq \frac{n}{3}$ because $|\cup_{T \in \mathcal{T}_i} V(T)|$

is always being kept no more than $\frac{n}{3}$ by the IF statement. Therefore, all conditions of proposition 3.11 are satisfied and so **whp**, this algorithm increases the size of a H_0-hypertree in the family $\mathcal{T}_i$. Each time, the increment of the size is 3. Therefore, we cannot miss the size k and the WHILE loop terminates within finite number of iteration and output a desire H_0-hypertree. □

With the help of above propositions and lemmas, we are able to prove the following lemma. This lemma leads to Theorem 3.7.

Lemma 3.2 *Define sequences* $\{p_l\}_{l=0}^{\infty}$ *and* $\{\epsilon\}_{l=0}^{\infty}$ *by* $p_0 = 77n^{-\frac{3}{2}}$, $p_l = p_0 + \left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1} - \left(\frac{6}{5}\right)^{\frac{3}{2}} p_0 p_{l-1}$ *for* $l \geq 1$; *and* $\epsilon_l = 0.95 - 0.05l$ *for* $l \geq 0$. *Then for every integer* l *satisfying* $0 \leq l \leq 18$, *a random hypergraph* $H \in H(n, 3, p_l)$ ***whp*** *contains a family of vertex disjoint hyperedges covering all but at most* n^{ϵ_l} *vertices.*

Proof. We are going to prove the lemma by induction on l. For the base $l = 0$, we have $\epsilon_l = 0.95$ and the lemma is true by corollary 3.10. Suppose it's true up to $l-1$ where $l \geq 1$. Note that $1 - p_l = (1 - p_0)(1 - \left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}$. This allows us to treat the random hypergraph H as a union of two random hypergraphs $H_1 \in H(n, 3, p_0)$ and $H_2 \in H(n, 3, \left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1})$. Also, $p_l = \Theta(n^{-\frac{3}{2}})$ for every $1 \leq l \leq 18$. Consider the hypergraph H_1. Let $m_1 = \lfloor n^{\epsilon_{l-1}} \rfloor$ and m_2 be the smallest integer satisfying $m_2 \geq 2m_1 + n^{\epsilon_l}$ and $m_1 + m_2 \equiv 0$ (mod 3). Rather than using the round-off notation, we use $m_1 = n^{\epsilon_{l-1}}$, $m_2 = 2n^{\epsilon l-1} + n^{\epsilon_l}$. The validity of the proof still holds. Now we let k be the largest integer such that $km_2 \leq \frac{n}{6}$ and $k \equiv 1$ (mod 3). Since $k \sim \frac{n^{1-\epsilon_{l-1}}}{2*6}$, all the conditions of lemma 3.1 are satisfied. According to lemma 3.1 H_1 **whp** contains $\lfloor \frac{n}{6k} \rfloor \geq m_2$ vertex disjoint H_0-hypertrees, each having k vertices. We pick only m_2 of these subgraphs and denote this family by $\mathcal{T}_0 = \{T_1, \ldots, T_{m_2}\}$. Let $V_0 = \cup_{j=1}^{m_2} V(T_j)$ be the set of vertices of $\mathcal{T}_0$'s members and $V_1 = V - V_0$ be the set of vertices not covered by the H_0-hypertrees. Since $|V_0| = \lfloor \frac{5n}{6} \rfloor$, we have $|V_1| \geq \frac{5n}{6}$.

Now we shift our attention to H_2. Consider the subgraph of H_2 that is spanned by the vertex set V_1. This subgraph has size at least $\frac{5n}{6}$ and edge probability $\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}$. Therefore, this subgraph satisfies the conditons of this lemma (lemma 3.2). By the induction hypothesis, this subgraph **whp** contains a matching that covers all but at most $|V_1|^{\epsilon_{l-1}} < m_1$ vertices. Let us pick just enough edges from the matching to cover exactly m_1 vertices of V_1 and call this edge set $\mathcal{F}_0$. It can be done because $km_2 \equiv m_2 \ (mod\ 3)$ (as $k \equiv 1 \ (mod\ 3)$) and so $m_1 + km_2 \equiv m_1 + m_2 \equiv 0 \ (mod\ 3)$. We denote the set of vertices of V_1, not yet covered by $\mathcal{F}_0$, by $W = \{v_1, \dots, v_{m_1}\}$.

Now, we want to find edges in H_2 that has one vertex in W, one vertex in the removable vertex set of an H_0-hypertree in $\mathcal{T}$ and the last vertex in the removable vertex set of another H_0-hypertree in $\mathcal{T}$. If we can find, for each vertex in W, an edge like that and the H_0-hypertrees involved are different for each edge, then we can expand the covering $\mathcal{F}_0$ with edges from H_1 and H_2. Therefore we define the following process. At the beginning of step i, where $1 \leq i \leq m_1$, we have a subfamily $\mathcal{T}_{i-1} \subseteq \mathcal{T}_0, |\mathcal{T}_{i-1}| = m_2 - 2(i-1)$, and a family of edges $\mathcal{F}_{i-1} \supseteq \mathcal{F}_0$. We try to find an edge $\{v_i, u_1, u_2\}$, where $u_1 \in R(T_{j_1})$ and $u_2 \in R(T_{j_2})$, $T_{j_1}, T_{j_2} \in \mathcal{T}_{i-1}$, $j_1 \neq j_2$. If such an edge exists, then we add it to $\mathcal{F}_{i-1}$, obtaining $\mathcal{F}'_{i-1}$. Since u_1 and u_2 are removable vertices of T_{j_1} and T_{j_2}, both subgraphs $T_{j_1} - \{u_1\}$ and $T_{j_2} - \{u_2\}$ contain complete matchings on their vertices. We add these matchings to $\mathcal{F}'_{i-1}$, thus obtaining $\mathcal{F}_i$. The subgraph T_{j_1} and T_{j_2} are then removed from $\mathcal{T}_{i-1}$, resulting in a new family $\mathcal{T}_i$.

We claim that **whp** this process can be performed successfully for m_1 steps. To prove this, we consider the edges of H_2. At step i of the above process the family $\mathcal{T}_{i-1}$ contains at least $m_2 - 2m_1 \geq n^{\epsilon_l}$ subgraphs. Choose $m_2 - 2m_1$ subgraphs from $\mathcal{F}_{i-1}$ arbitrarily and denote them by $T_1, \dots, T_{m_2-2m_1}$. Note that $\sum_{j=1}^{m_2-2m_1} |V(T_j)| = k(m_2 - 2m_1) = \Theta(n^{0.95})$. By proposition 3.8, we have $|R(T_j)| \geq \frac{|V(T_j)|}{3}$, so $|R(T_j)| \geq \frac{k}{3}$, and thus $|R(T_j)| = \Theta(n^{1-\epsilon_{l-1}})$ for every $1 \leq j \leq m_2 - 2m_1$.

Given T_{j_1} and T_{j_2}, the probability that an edge $\{v_i, u_1, u_2\}$ doesn't exists is $\left(1-\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}\right)^{\left(\frac{k}{3}\right)^2}$. Therefore the probability that the process fails at step i where $1 \le i \le m_1$ is $\left(1-\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}\right)^{\binom{m_2-2m_1}{2}\left(\frac{k}{3}\right)^2}$. So, the probability that the process gives us a matching is $\left(1-\left(1-\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}\right)^{\binom{m_2-2m_1}{2}\left(\frac{k}{3}\right)^2}\right)^{m_1}$. We have

$$\begin{aligned}
&\left(1-\left(1-\left(\tfrac{6}{5}\right)^{\frac{3}{2}} p_{l-1}\right)^{\binom{m_2-2m_1}{2}\left(\frac{k}{3}\right)^2}\right)^{m_1} \\
&> 1-m_1\left(1-\left(\tfrac{6}{5}\right)^{\frac{3}{2}} p_{l-1}\right)^{\binom{m_2-2m_1}{2}\left(\frac{k}{3}\right)^2} \\
&> 1-m_1 e^{-\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}\binom{m_2-2m_1}{2}\left(\frac{k}{3}\right)^2}.
\end{aligned}$$

Because $\binom{m_2-2m_1}{2}\left(\frac{k}{3}\right)^2 = \Theta(\frac{k^2(m_2-2m_1)^2}{18}) = \Theta\left((n^{0.95})^2\right) = \Theta\left(n^{1.9}\right)$, $m_1 = n^{\epsilon_{l-1}}$ and $p_{l-1} = \Theta\left(n^{-\frac{3}{2}}\right)$, we get

$$\begin{aligned}
&1-m_1 e^{-\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}\binom{m_2-2m_1}{2}\left(\frac{k}{3}\right)^2} \\
&\sim 1-\Theta\left(\frac{n^{\epsilon_{l-1}}}{e^{\left(\frac{6}{5}\right)^{\frac{3}{2}} n^{0.4}}}\right) \\
&\ge 1-\Theta\left(\frac{n^{0.95}}{e^{\left(\frac{6}{5}\right)^{\frac{3}{2}} n^{0.4}}}\right) \longrightarrow 1.
\end{aligned}$$

Therefore, **whp** the process finds the desired edges.

After the process terminated, we have a family $\mathcal{T}_{m_1} = \{T_1, \ldots, T_{m_2-2m_1}\}$ of H_0-hypertrees. Deleting a removable vertex $u_j \in R(T_j)$ from every $T_j \in \mathcal{T}_{m_1}$, we have a matching in each T_j. We add all these edges to $\mathcal{F}_{m_1}$. Now $\mathcal{F}_{m_1}$ covers all vertices of V except the deleted removable vertices $\{u_1, \ldots, u_{m_2-2m_1}\}$ from $\mathcal{T}_{m_1}$. Since $2m_2-2m_1 = n^{\epsilon_l}$, $\mathcal{F}_{m_1}$ is the matching we want to get. □

Proof of Theorem 3.7. Refer to lemma 3.2 and consider a random graph $H \in H(n,3,p_{19})$ and represent it as a union of $H_1 \in H(n,3,p_0)$ and $H_2 \in H(n,3,\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{18})$. Follow lemma 3.2's construction, we have, **whp**, H_2 contains a matching that covers all but at most $n^{0.05}$ vertices. Define m_1 to be the largest integer not exceeding $n^{0.05}$ and satisfying $m_1 \equiv 0 \ (mod\, 3)$, and define $m_2 = 2m_1$. Now, we follow a similar process as the one in lemma 3.2, but in this time, instead of choosing 2 H_0-hypertrees from a set of $m_2 - 2m_1$ H_0-hypertrees as in lemma 3.2, we arbitrary assign 2 H_0-hypertrees to every uncovered vertices and then we follow the old process. At this time, we have $|R(T_j)| \geq \frac{k}{3} = \Theta(n^{1-\epsilon_{l-1}}) = \Theta(n^{0.95})$ for every $T_j \in \mathcal{F}$. Now, the probability that we fail to find a desired edge in one of the step is given by $\left(1 - \left(\frac{6}{5}\right)^{\frac{3}{2}} p_{18}\right)^{\left(\frac{k}{3}\right)^2}$. Therefore, the probability we succeed is

$$
\begin{aligned}
&\left(1-\left(1-\left(\tfrac{6}{5}\right)^{\frac{3}{2}} p_{18}\right)^{\left(\frac{k}{3}\right)^2}\right)^{m_1} \\
\geq &\left(1-\left(1-\left(\tfrac{6}{5}\right)^{\frac{3}{2}} p_{18}\right)^{\left(\frac{k}{3}\right)^2}\right)^{n^{0.05}} \\
\geq &\left(1-e^{-\frac{\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{18}k^2}{9}}\right)^{n^{0.05}} \\
\geq &\, 1-n^{0.05}e^{-\frac{\left(\frac{6}{5}\right)^{\frac{3}{2}} p_{18}k^2}{9}} \\
\sim &\, 1-\Theta\left(\frac{n^{0.05}}{e^{\left(\frac{6}{5}\right)^{\frac{3}{2}} n^{1.9-1.5}}}\right) \\
\sim &\, 1-\Theta\left(\frac{n^{0.05}}{e^{\left(\frac{6}{5}\right)^{\frac{3}{2}} n^{0.4}}}\right) \longrightarrow 1.
\end{aligned}
$$

So, **whp** the process finds the desired edges and then we get, at this time, a complete matching of H.

In order to estimate p_{19}, we consider the fact that $p_l = p_0 + \left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1} - \left(\frac{6}{5}\right)^{\frac{3}{2}} p_0 p_{l-1} \leq p_0 + \left(\frac{6}{5}\right)^{\frac{3}{2}} p_{l-1}$ for every $l \geq 1$. Therefore $p_l \leq \left(\frac{\left(\frac{6}{5}\right)^{\frac{3}{2}(l+1)}-1}{\left(\frac{6}{5}\right)^{\frac{3}{2}}-1}\right) p_0$ or $p_{19} \leq \left(\frac{\left(\frac{6}{5}\right)^{\frac{3}{2}(20)}-1}{\left(\frac{6}{5}\right)^{\frac{3}{2}}-1}\right)\left(77n^{-\frac{3}{2}}\right) \leq 57867n^{-\frac{3}{2}}$. □

BIBLIOGRAPHY

BIBLIOGRAPHY

[AcI93] R. Acharya and D. Ibaroudene, Linear hypertree for multidimensional image representation, *Inform. Sci.* **68** (1993) 123-154.

[An82] S. Antonuicci, Sur les colorations généralisées des hyperarbres et sur les multicolorations des graphes, *Riv. Mat. Univ. Parma(4)* **8** (1982) 235-242.

[AS92] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley (1992).

[Be70] C. Berge, *Graphs et Hypergraphes*, Dunod (1970).

[Be73] C. Berge, *Introduction Á La Théorie Des Hypergraphes*, Les Presses De L'Université De Montréal (1973).

[Be87] C. Berge, *Hypergraphs : Combinatorics of Finite Sets*, North-Holland (1989).

[BeM69] L.W. Beineke and J.W. Moon, Several Proofs of the Number of Labeled 2-Dimensional Trees, *Proof Techniques in Graph Theory*, Academic New York, (1969).

[BeP69] L.W. Beineke and R.E. Pippert, The number of labelled k-dimensional trees, *Journal of Combinatorial Theory* **6** (1969) 200-205.

[BiLW76] N.L. Biggs, E.K. Lloyd and R.J. Wilson, *Graph Theory 1736-1936*, Oxford University Press (1976).

[Bo84] V. Boonyasombat, Degree sequences of connected hypergraphs and hypertrees, *Graph theory, Singapore 1983* 236-247, Lecture Notes in Math., 1073, *Springer, Berlin-New York,* 1984.

[Bo85] B. Bollobás, *Random Graphs*, Academic Press (1985).

[BrCD95] A. Brandstädt, V. D. Chepoi and F. F. Dragan, The algorithmic use of hypertree structure and maximum neighbourhood orderings, *Graph-theoretic concepts in computer science (Herrsching, 1994)* 65-80, Lecture Notes in Compt. Sci., 903, *Springer, Berlin,* 1995.

[BrCD98] A. Brandstädt, V. D. Chepoi and F. F. Dragan, The algorithmic use of hypertree structure and maximum neighbourhood orderings, *Discrete Appl. Math.* **82** (1998) 43-77.

[BrD94] A. Brandstädt, Andreas and F. F. Dragan, Dominating cliques in graphs with hypertree structure, 735-746, Lecture Notes in Comput. Sci., 775, *Springer, Berlin* (1994) 735-746.

[BrD96] A. Brandstädt and F. F. Dragan, r-dominating cliques in graphs with hypertree structure, *Discrete Math.* **162** (1996) 93-108.

[Dr99] F. Drewes, A characterization of the sets of hypertrees generated by hyperedge-replacement graph grammars, *Theory Comput. Syst.* **32** (1999) 159-208.

[ErR59] P. Erdös and A. Rényi, On random graphs I, *Publ. Math. Debrecen* **6** (1959) 290-297.

[ErR60] P. Erdös and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* **5** (1960) 17-61.

[ErR61a] P. Erdös and A. Rényi, On the evolution of random graphs, *Bull. Inst. Internat. Statist.* **38, 4** (1961) 343-347.

[ErR61b] P. Erdös and A. Rényi, On the strength of connectedness of a random graph, *Acta Math. Acad. Sci. Hungar.* **12** (1961) 261-267.

[ErR64] P. Erdös and A. Rényi, On random matrices, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8** (1964) 455-461.

[ErR66] P. Erdös and A. Rényi, On the existance of a fractor of degree one of a connected random hypergraph, *Acta Math. Acad. Sci. Hungar.* **17** (1966) 359-368.

[ErR68] P. Erdös and A. Rényi, On random matrices II, *Studia Sci. Math. Hungar.* **3** (1968) 459-464.

[GaJ79] M. R. Garey and D. S. Johnson, *Computers and Intractability: a guide to the theory of NP-completeness*, W. H. Freeman, New York (1979)

[Gi78] M. Gionfriddo, Hypergraphs lacking significant r-cycles and K-hypertrees, *Riv. Mat. Univ. Parma (4)* **4** (1978) 259-268.

[Gr91] D. Grieser, Counting complements in the partition lattice, and hypertrees, *J. Combin. Theory Ser. A* **57** (1991) 144-150.

[GrH93] D. A. Grable and F. M. Hoppe, Permutation hypertrees in probability, *Ars Combin.* **35** (1993) 135-142.

[HaP68] F. Harary and E. M. Palmer, On acyclic simplicial complexes, *Mathematika* **15** (1968) 115-122.

[HaP73] F, Harary and E. M. Palmer, *Graphical Enumeration*, Academic, New York (1973).

[Hu50] K. Husimi, Note on Mayer's Theory of Cluster Integrals, *J. Chem. Phys.* **18** (1950) 682-684.

[Is01] T. Ishihara, Enumeration of hypergraphs, *Europe J. Combinatorics* **22** (2001) 503-509.

[JaLR90] S. Janson, T. Luczak and A. Ruciński, An exponential bound for the probability of nonexistence of a specified subgraph in a random graph, *Random Graphs '87* (1990) 73-87.

[Ka82] M. Karoński, Random hypertrees, allocations and partitions, *Graphs and other combinatorial topics (Pargue, 1982)* 159-162, Teubner-Texte Math. 59, *Teubner, Leipzig,* 1983.

[KaN91] M. Karoński and K. Nowicki, On the entropy of vertex degree distributions in a random hypertree, *Advances in graph theory* 215-226, *Vishwa, Gulbarga,* 1991.

[Ko99] V. F. Kolchin, Branching processes and random hypertrees, *Diskret. Mat.* **11** (1999), 8-23 *translation in Discrete Appl. Math.* **9** (1999) 7-23.

[Kr97] M. Krivelevich, Triangle Factors in Random Graphs, *Combinatorics, Probability and Computing* **6** (1997) 337-347.

[Li86] G. Z. Liu, A theorem on the factors of r-hypertrees, *Adv. in Math. (Beijing)* **15** (1986) 381-383.

[LiZ88] B. L. Liu and X. K. Zhang, The enumeration theory of hypertrees. I, *Chinese Quart. J. Math.* **3** (1988) 78-82.

[LjS74] V. N. Ljamin and B. I. Selivanov, Hypertrees with a given number of end vertices and edges, *Kombinatornyĭ Anal.* Byp. **3** (1974) 68-71.

[LjS74b] V. N. Ljamin and B. I. Selivanov, Random hypertrees and hyperforests, *Mat. Zametki* **15** (1974) 641-650.

[M70] J. W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs (1970).

[NiP99] J. Nieminen and M. Peltola, Hypertrees, *Appl. Math. Lett.* **12** (1999) 35-38.

[NiV00] A. Niculitsa and V. Voloshin, About uniquely colorable mixed hypertrees, *Discuss. Math. Graph Theory* **20** (2000) 81-91.

[Pa73] E. M. Palmer, On the number of n-plexes, *discrete Math.* **6** (1973) 377-390.

[Pa85] E. M. Palmer, *Graphical Evolution*, John Wiley & Sons, 1985.

[PaR85] E. M. Palmer and R. W. Robinson, Connectivity of a random m-graph, *Revue Roumaine De Mathématiques Pures Et Appliquées* , **XXX N.3** (1985) 229-232.

[PaS79] E. M. Palmer and A. J. Schwenk, On the number of trees in a random forest, *Journal of Combinatorial Theory, Series B* **27** (1979) 109-121.

[ScS83] J. Schmidt and E. Shamir, A threshold for perfect matchings in random d-pure hypergraphs, *Discrete Math.* **45** (1983) 287-295.

[Se75] B. I. Selivanov, Additional remarks on homogeneous hypertrees and hyperforests, *Combinatorial and Asymptotic Analysis, Krasnojarsk. Gos. Univ., Krasnoyarsk* (1975) 137-146.

[SoT00] M. Sonntag and H. Teichert, Sum numbers of hypertrees, *Discrete Math.* **214** (2000) 285-290.

[To86] I. Tomescu, Hypertrees and Bonferroni inequalities, *J. Combin. Theory Ser. B* **41** (1986) 209-217.

[To92] I. Tomescu, Ordered h-hypertrees, *Discrete Math.* **105** (1992) 241-248.

[ToZ94] I. Tomescu and M. Zimand, Minimum spanning hypertrees, *Discrete Appl. Math.* **54** (1994) 67-76.

[Tu47] W. T. Tutte, The factorizations of linear graphs, *J. London Math. Soc.* **22** (1947) 107-111.

[Ve82] W. Fernandez de la Vega, Sur la cardinalité maximum des couplages d'hypergraphes aléatorires uniformes, *Discrete Math.* **40** (1982) 315- 318.

[Yu00] R. Yuster, Decomposing hypergraphs into simple hypertrees, *Combinatorica* **20** (2000) 119-140.