Decomposable Combinatorial Structures with Restricted Patterns

by

Li Dong

A thesis submitted to the School of Graduate Studies in partial fulfilment of the
requirements for the degree of

Doctor of Philosophy

Ottawa-Carleton Institute for Mathematics and Statistics

Carleton University

Ottawa, Ontario, Canada

©Copyright

2007, Li Dong

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-40521-5

Our file Notre référence

ISBN: 978-0-494-40521-5

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Abstract

A *decomposable combinatorial structure* consists of simpler objects called *components* which by themselves can not be further decomposed. For example, permutations decompose into cycles, graphs into connected components, and polynomials over finite fields into irreducible factors. The size of an object is the number of its elements. For example, a permutation on n objects, a graph on n vertices and an n -degree polynomial over a finite field have all size n . In this thesis, we study the decomposable combinatorial structures with given *restricted patterns*. A *restricted pattern* for a decomposable object of size n fixes the number of components whose sizes are specified in the restricted pattern. In other words, partial information of this object is known. For instance, the number of permutations with a certain factorization pattern is used in the study of the cycle index of the symmetric group.

Since decomposable combinatorial structures are traditionally divided into labeled and unlabeled cases, we provide results for the two cases throughout this thesis. Applying a combinatorial method, we obtain the generating function of combinatorial structures with a given restricted pattern for both labeled and unlabeled objects. Furthermore, in addition to the restricted pattern, we derive the generating function when we restrict the size of the r th smallest component of decomposable combinatorial structure.

We study two important classes of decomposable combinatorial structures: exp-log class and alg-log class. For the structures in these two classes, using the method of analysis of singularities, we derived an asymptotic expression for the probability that a decomposable structure of size n has a given restricted pattern. In addition to restricted pattern, we estimate the probability that the size of the random decomposable combinatorial structure's r th smallest component is bigger than k , for r, k given integers. Furthermore, we estimate the first moment of the size of the r th smallest component in both exp-log class and alg-log class with restricted patterns. In the exp-log class, we also obtain a result for higher moments.

Acknowledgement

First of all, I would like to thank my supervisors Dr. Zhicheng Gao and Dr. Daniel Panario for their patient guidance, valuable suggestions and enlightening comments during the period of my study at Carleton University. Their profound knowledge has definitely benefited me a lot. Working with them has always been an enjoyable and rewarding experience.

I would like to thank Bruce Richmond, David Sankoff and Michiel Smid for their helpful comments on this thesis .

I really appreciate the school of Mathematics and Statistics, Carleton University, for providing necessary research facilities for preparing this thesis. Many thanks go to members and staffs of the school of Mathematics and Statistics, especially to Valerie Daley, who helped me greatly during my studies at Carleton University.

Finally, I am especially grateful to my wife, Yunyun Song, for her persistent support and irreplaceable help. I would like to thank my parents. I learned a lot from their positive attitude towards life under any circumstances. I dedicate this thesis to them.

Contents

1	Introduction	1
1.1	Introduction	1
1.2	A Brief Survey of Known Results	2
1.2.1	Developed Methods	3
1.2.2	Specific Decomposable Structures	3
1.2.3	Common Properties of Decomposable Combinatorial Structures	4
1.3	Contributions of This Thesis	5
1.4	Brief Description of This Thesis	6
2	Background	9
2.1	Generating Functions	9
2.1.1	Introduction	10
2.1.2	Ordinary and Exponential Generating Functions	10
2.2	Combinatorial Structures	11

2.2.1	Definitions	11
2.2.2	Labeled and Unlabeled Structures	15
2.3	Singularity Analysis of Generating Functions	17
3	Generating Functions of Decomposable Combinatorial Structures	21
3.1	Decomposable Combinatorial Structures with Given Restricted Patterns	22
3.1.1	Labeled Case	23
3.1.2	Unlabeled Case	24
3.2	The r th Smallest Component	26
3.2.1	Labeled Case	26
3.2.2	Unlabeled Case	29
3.3	Examples	32
4	Decomposable Combinatorial Structures in Exp-Log Class	37
4.1	Introduction	38
4.2	Enumeration of Decomposable Combinatorial Structures with Restricted Patterns	39
4.2.1	Labeled Case	40
4.2.2	Unlabeled Case	46
4.3	Distribution of the r th Smallest Component	51
4.3.1	Labeled Case	51

4.3.2	Unlabeled Case	55
4.4	Moments	58
4.5	Examples.	67
5	Decomposable Combinatorial Structures in Alg-Log Class	70
5.1	Introduction	70
5.2	Alg-Log Structures with Restricted Pattern	72
5.2.1	Labeled Case	72
5.2.2	Unlabeled Case	79
5.3	Size of the r th Smallest Component	84
5.3.1	Labeled Case	84
5.3.2	Unlabeled Case	87
5.4	Expectation	89
5.5	Example	91
6	Conclusions and Future Work	93
6.1	Conclusions	93
6.2	Future Work	94
A	Discussion on the Probability of a Combinatorial Structure with Given Restricted Pattern	96

List of Figures

2.1	The cycle decomposition of permutations.	14
2.2	The $\Delta(\nu, \theta)$ region.	18
3.1	The 5 vertices rooted nonplane trees without subtrees formed by 1 vertex.	36
4.1	The contour γ	42
5.1	The Cayley trees with 1 node, 2 nodes and 3 nodes.	92

Chapter 1

Introduction

1.1 Introduction

Decomposable combinatorial structures (decomposable structures) have always been interesting objects. Those structures are widely used in many fields, such as Computer Science, Biostatistics, Cryptology and Analysis of Algorithms.

A decomposable combinatorial structure consists of simpler objects called components which by themselves can not be further decomposed. For example, permutations decompose into cycles, graphs into connected components, polynomials over finite fields into irreducible factors and so on. The decomposition of combinatorial structures into components is analogous to the fact that integer numbers decompose into primes. Some explicit practical problems are: how many different permutations are there of length 20? or of length n , for general n ? These questions are simple, but what if some constraints are imposed, e.g., how many permutations on length n have only three cycles of length 2? In this thesis, we study decomposable combinatorial structures with given constraints. We name the given constraints as restricted pat-

terns. The restricted pattern can be also understood as partial information of a given random decomposable structure. The concept of restricted pattern has been widely used in many academic fields. For example, in the factorization of a polynomial over a finite field, patterns are used to determine the most frequent decomposition of a polynomial into irreducible factors. This gives information about the most common running times of factoring algorithms on random inputs.

Forecasting the occurrence of small components of combinatorial decomposable structures with given restricted pattern is of interest in many branches of combinatorial modeling. For example, given a window of 10000 bases in a DNA sequence consisted of alphabet $\{A, G, C, T\}$ (representing four nucleotides), how significant is it to observe a sequence of 20 ‘independent’ bases which have no identical neighbours? Another mathematical example is, given a permutation of size 2^{50} which has 5 cycles of size 4, what is the possibility that the size of the 50th smallest cycle is bigger than $\log 2^{50}$? Smallest components have been used to analyze a polynomial irreducibility test ([42]). This polynomial irreducibility test is analogous to the basic primality test that stops once we find the smallest prime in the factorization of the number being tested. Because of its significance, the distribution of the smallest components is studied in this thesis.

1.2 A Brief Survey of Known Results

In this section, we give very brief descriptions of the known related results about decomposable combinatorial structures.

1.2.1 Developed Methods

In the study of the decomposable structures, generating function is a useful combinatorial tool to count decomposable structures with same size. But, often, one encounters generating functions where exact extraction of coefficients is extremely hard. In those cases, asymptotic extraction of coefficients is used. To overcome this difficulty, mathematicians brought complex analysis and meromorphic asymptotics in. For example, Flajolet and Odlyzko introduced *singularity analysis of generating function* (see [17]) in 1990, while Flajolet, Grabner, Kirschenhofer, Prodinger and Tichy [14] and Flajolet and Golin [12] developed *mellin transforms*. Some researchers study decomposable structures by alternative methods. For instance, Arratia, Barbour and Tavaré [3] analyze the decomposition of random combinatorial structures from a probabilistic standpoint. All the methods above have contributed to the advancement of the research of decomposable combinatorial structures in the last two decades.

1.2.2 Specific Decomposable Structures

In the early days, mathematicians only study some specific decomposable structures with restricted patterns, such as the cycle distribution of permutation [25, 26, 34, 48] and the factorization of a polynomial into irreducible factors [7, 33, 38]. Additionally, the specific decomposable structures with restricted patterns have been approached in the analysis of some particular algorithms [13, 23, 24, 41, 42].

1.2.3 Common Properties of Decomposable Combinatorial Structures

As decomposable structures were being studied more and more frequently, the study of the common properties of such objects became an interesting research project. In the early 1990's, some important classes of decomposable structures were defined. Flajolet and Soria introduced the *exp-log* class in [20] and the *alg-log* (*algebraic-logarithmic* type) in [21]. They studied the behavior of the number of components of structures in these classes.

In 1993, Flajolet and Soria [21] show that, under general conditions, the number of components in some random classical combinatorial structures had a Gaussian limit distribution with an exponential tail. Some related research can be found in [4, 6, 8, 47]. The combinatorial structures studied by Flajolet and Soria are formed using the Sequence, Cycle and MultiSet constructions, such that each object of the composed class is uniquely decomposed into elements of the component class.

Moreover, Gourdon [28] considered the distribution of the r th largest components for objects of the algebraic-logarithmic type. It involved the study of restricted patterns where the sizes of all components are smaller than or equal to a value k . More details of Gourdon's work can be found in his Ph.D thesis [29]. In 2001, Panario and Richmond gave the distribution of the r th smallest component of a decomposable combinatorial structure of size n [43]. They studied decomposable combinatorial structures of size n in the *exp-log* class as $n \rightarrow \infty$.

1.3 Contributions of This Thesis

Up to now, most research have focused on some special restricted patterns. For instance, the authors of [34] give estimates for the number of permutations on n elements that have no more than k cycles, and the equivalent result for polynomials over finite fields. In our study, we take more general conditions of the restricted patterns into consideration. In other words, we analyze the properties of decomposable structures with arbitrary given patterns.

In this thesis, we aim to achieve the following goals

- To obtain the generating functions of random decomposable combinatorial structures with given partial information (restricted pattern). In addition to the restricted pattern, we provide the generating functions when, at the same time, we restrict the size of the r th smallest component of the decomposable structure.
- To estimate the distribution of the decomposable structures in exp-log class with given restricted pattern, and discuss the joint probability that the size of the r th smallest component is bigger than some integer k and that the object has a given restricted pattern.
- To estimate the distribution of the decomposable structures of algebraic-logarithmic type with given restricted pattern, and provide the joint distribution of the r th smallest component of decomposable structures with a given restricted pattern.

Throughout this thesis, a random object is chosen uniformly from a set under consideration.

1.4 Brief Description of This Thesis

In this section, we describe briefly the structure of this thesis.

In Chapter 2, we introduce the mathematical background of this thesis. In Section 2.1, we define generating functions ([50]) which is a classical tool in combinatorics. In Section 2.2, we introduce some basic concepts, such as combinatorial class, restricted factorization pattern and generating functions of decomposable combinatorial structures. In Section 2.3, we refer to a powerful tool from analytic combinatorics [22]: singularity analysis.

Chapter 3 is about generating functions of decomposable combinatorial structures. In Section 3.1, we consider decomposable structures with a given restricted pattern. In Section 3.2, in addition to the restricted pattern, we provide the generating functions when we restrict the size of the r th smallest component of decomposable structure. In Section 3.3, some explicit examples are provided.

In Chapter 4, we focus on decomposable combinatorial structures in the Exp-Log class. We introduce in Section 4.1 the definition of the Exp-Log Type which contains many interesting decomposable combinatorial structures. In Section 4.2, applying the generating functions obtained in Chapter 3, we estimate the probability that a random decomposable combinatorial structure of exp-log type has a given restricted pattern. Section 4.3 further deals with the r th smallest component condition. In Section 4.4, we discuss the moments of the size of the r th component of a decomposable structure in the exp-log class with a restricted pattern. Section 4.5 exemplifies our results for some decomposable combinatorial structures, including labeled and unlabeled cases.

Chapter 5 deals with parallel problems on the alg-log type which is another im-

portant class of decomposable combinatorial structures. In Section 5.2, we derived an asymptotic expression for the probability that a random decomposable combinatorial structure of the alg-log type has a given restricted pattern. The r th smallest component is discussed in Section 5.3. Section 5.4 gives the first moment of the size of the r th smallest component of a decomposable structure of the alg-log type with a restricted pattern. Section 5.5 provides one example of object in the alg-log class.

In Chapter 6, we summarize the results in this thesis and list some open problems. Among these problems, for some of them, we have some ideas for attacking them, but the results are far from satisfactory. Some problems are too far removed from the current research, and we need more efficient tools and more advanced technique to solve them.

Notations:

We give some widely used notations which will be discussed later, but are gathered here for easy reference.

- Let $\mathbb{N}$ be the set of natural numbers $\{0, 1, 2, \dots\}$.
- $A := B$ means that A is defined to be equal to expression B .
- For a real number x , $\lfloor x \rfloor$ denotes the largest integer which is smaller than or equal to x .
- $|A|$ denotes the size of the set A (namely, the number of elements in set A).
- Let $I(A)$ be a characteristic function of an event A , that is, $I(A) = 1$ if event A happens and $I(A) = 0$ otherwise.
- N_k represents a number set $\{1, 2, \dots, k\}$ where k is a positive integer.

- Given a pattern S , $C_L(z; S)$ and $C_U(z; S)$ denote the generating functions of the labeled and unlabeled pattern, respectively.
- $X_n^{[r]}(S)$ is the size of the r th smallest component of a random decomposable combinatorial structure of size n having a restricted pattern S .

Chapter 2

Background

The definitions of generating functions and combinatorial class are given in Sections 2.1 and 2.2. Some useful notations which will be used in this thesis are also introduced in those two sections. In Section 2.3, we introduce singularity analysis which is the main asymptotic technique in this thesis. The purpose of this chapter is to provide a brief introduction of this field and of the important mathematical tools that we use in later chapters.

2.1 Generating Functions

Generating functions (see [35, 50, 52]) are a bridge between discrete mathematics, on the one hand, and continuous analysis (particularly complex variable theory) on the other. As we shall see, generating functions are the central mathematical object of combinatorial analysis. An excellent general reference on generating functions is [9], which contains numerous beautiful examples.

2.1.1 Introduction

A generating function is a framework in which we display a sequence of numbers. Suppose we have a sequence of numbers, $a_0, a_1, a_2, \dots$. We want to express them in a simple form. The best results that could be hoped for is a simple recurrence formula of a_n . Unfortunately, under many conditions, we can not get such a formula, or the formula is really huge and complicated. Generating function helps us to compress the sequence of numbers into a single formula. When we want to get the numbers back, we only need to extract the corresponding coefficients out.

We use the following symbol to represent coefficients of generating functions.

Definition 2.1.1. *Let $f(z)$ be a power series of z . Then by the symbol $[z^n]f(z)$ we mean the coefficient of z^n in the series of $f(z)$.*

A perfectly obvious property of this symbol, that we will use repeatedly, is

$$[z^n]f(\rho z) = \rho^n [z^n]f(z).$$

2.1.2 Ordinary and Exponential Generating Functions

Generating functions are normally divided into two categories, *ordinary* and *exponential* generating functions. Given a sequence $\{a_n\}$, when we want to distinguish the two kinds of generating functions, we write the ordinary generating function (OGF) of the sequence $\{a_n\}$ as $\sum_n a_n z^n$ while the exponential generating function (EGF) of the sequence $\{a_n\}$ is written as $\sum_n a_n z^n / n!$.

Although the only difference between OGF and EGF is the coefficient $1/n!$, OGF and EGF count two important types of combinatorial structures. We will introduce them in Section 2.2.2.

2.2 Combinatorial Structures

A combinatorial structure can be represented by a specific generating function. The problem of enumerating a class of combinatorial structures then simply reduces to finding the coefficients of the associated generating function. In this section, we introduce Combinatorial Structures and relationships between generating functions and combinatorial structures.

2.2.1 Definitions

As an object being discussed in this thesis, a combinatorial class is a set of discrete elements which follow some construction rules. Familiar examples are trees, graphs, permutations, polynomials over finite fields, topological configurations and etc. In studying combinatorial classes, one of the always interesting questions is to enumerate such objects according to some characteristic parameter(s).

Definition 2.2.1. (*Flajolet and Sedgewick [22]*) A **combinatorial class** is a finite or denumerable set on which a size function is defined, satisfying the following conditions:

1. the size of an element is a nonnegative integer;
2. the number of elements of any given size is finite.

An element of a combinatorial class is also called a **combinatorial structure**.

We use size to measure a combinatorial structure. For example, a permutation on n objects, a graph on n vertices and an n -degree polynomial over a finite field all have size n . If C is a combinatorial class, the size of an element $\alpha \in C$ is denoted by $|\alpha|$.

Given a combinatorial class C , we consistently let C_n be the number of structures in C that have size n .

Definition 2.2.2. (Wilf [52]) *The counting sequence of a combinatorial class C is the sequence of integers $(C_n)_{n \geq 0}$ where C_n is the number of combinatorial structures in class C that have size n .*

From symbolic analysis, the counting sequence of a combinatorial class C is the sequence of coefficients in the generating function of class C . Therefore, we can study the distribution of combinatorial structures by analyzing their generating functions.

Many combinatorial structures are built directly in terms of simpler objects by means of a collection of elementary combinatorial **constructions**. For example, permutations decompose into cycles, graphs into connected components, and polynomials over finite fields into irreducible factors.

Definition 2.2.3. *Let C be a combinatorial class such that its elements can not be further decomposed. A class F is a **decomposable combinatorial class** over C if any of its elements can be uniquely decomposed into a set of structures of C . Those structures of C are called the components of F . The elements of F are called **decomposable combinatorial structures***

Every decomposable combinatorial structure has a ‘pattern’, that is, an object of size n has S_j components of size j such that $\sum_{j=1}^n jS_j = n$. For a decomposable combinatorial structure, we have a restricted pattern if the numbers of the components with certain sizes are fixed. For instance, an n -degree polynomial over a finite field can be uniquely decomposed into a product of irreducible polynomials over the finite field. We get a restricted pattern, for example, if we know that there are only 3 irreducible polynomials of degree 2 among their factors. We have the following

1	2	.	.	.	n
$\downarrow$	$\downarrow$		$\downarrow$		$\downarrow$
α_1	α_2	.	.	.	α_n

Table 2.1: Permutations.

definition.

Definition 2.2.4. A restricted factorization pattern S (or simply pattern) for a combinatorial structure of size n is defined as a mapping $S : J \mapsto \mathbb{N}$, where J is the set of the components' sizes which are known and $\mathbb{N}$ is the set of nonnegative integers. Let $S(j)$ be the number of components of size j in the given restricted pattern S . If we do not have a restricted pattern, J is the empty set.

According to the definition of restricted factorization pattern, we emphasize that we want to count objects such that the components with sizes not in J may appear any number of times but there are exactly $S(j)$ components of size $j \in J$.

Example 2.2.1. Permutations. A permutation of $N_n := \{1, 2, \dots, n\}$ is a bijective mapping that is representable by an array $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ as in Table 2.2.1. From elementary mathematics, it is known that a permutation admits a unique decomposition into cycles. The length is defined as the size of a permutation or of a cycle. Given a set $\mathcal{P}$ of permutations, the number $\mathcal{P}_n$ of permutations of size n is well known to be $\mathcal{P}_n = n!$. Now, we want to know how many permutations with size n have only 5 cycles with size 3 and 2 cycles with size $\lfloor \log n \rfloor$. In this question, we have a pattern S with $J = \{3, \lfloor \log n \rfloor\}$ and $S(3) = 5$, $S(\lfloor \log n \rfloor) = 2$.

Besides enumeration, probabilistic properties of the combinatorial structures are also interesting objects to mathematicians. For example, what is the probability that the size of smallest component of a random Catalan tree with size n is bigger

A permutations may be viewed as a set of cycles that are labeled circular digraphs. The diagram shows the decomposition of the permutation

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
7	6	15	12	13	5	1	2	8	14	3	4	9	10	11

Figure 2.1: The cycle decomposition of permutations.

than $\log n$? The probabilistic properties can help us to understand combinatorial structures better. As one of the most important probabilistic properties, moments will be discussed in this thesis. The first moment is also called expected value (or expectation). It describes the average size of a random variable.

Definition 2.2.5. *The n th moment of a probabilistic distribution is given by the Riemann-Stieltjes integral*

$$E^{(n)}(X) = \int_{-\infty}^{\infty} x^n dF(x).$$

where X is a random variable that has distribution $F(x)$.

Normally, in deriving the n th moments of a discrete random variable X , we have

a popular expression that will be used many times in this thesis

$$E^{(n)}(X) = \sum_{k \geq 0} ((k+1)^n - k^n) P(X > k).$$

2.2.2 Labeled and Unlabeled Structures

Traditionally, decomposable combinatorial structures are divided into two categories depending on whether the structures under consideration are labeled or unlabeled.

Many combinatorial objects present themselves naturally as labeled structures where atoms of an object (typically nodes in a graph or a tree) are distinguishable from one another by the fact that they bear distinct labels. We name those objects as labeled combinatorial structures. Permutations, 2-regular graphs, Cayley tree, random mappings and simple family of trees [36] are all labeled combinatorial structures. For instance, a permutation can be viewed as a linear arrangement of distinct labels; its cycle decomposition represents it as an unordered collection of circular directed graphs whose node are labeled by integers. The labeled objects are represented by *exponential generating function*.

For the unlabeled combinatorial structures, as told by the name, we do not distinguish the atoms of the object. Some typical unlabeled combinatorial structures are polynomials over finite fields, partitions of integers, random mapping patterns, some types of trees etc. The unlabeled objects are represented by *ordinary generating function*.

Definition 2.2.6. *A combinatorial class C is labeled (unlabeled) if every element of C is labeled (unlabeled). Given a decomposable combinatorial class F over C , we say F is labeled (unlabeled) if C is a labeled (unlabeled) combinatorial class.*

As mentioned in Section 2.2.1, there are some basic constructions of decomposable combinatorial structures. One of the most important constructions is **multiset construction** [22]. Multisets are like finite sets (the order between elements does not count), but arbitrary repetitions of elements are allowed. The notation is $F = MSet(C)$ when F is obtained by forming all finite multisets of elements from C . In this thesis, all decomposable combinatorial structures studied by us are assumed to have the multiset construction.

In the end of this section, we provide the generating functions of labeled and unlabeled combinatorial structures. We have a combinatorial class C and a class F of decomposable combinatorial structures over C . Let C_k be the number of different elements of size k in the class C . In the case of labeled structures, the exponential generating function of the elements of C is

$$C_L(z) = \sum_{k \geq 0} C_k \frac{z^k}{k!}.$$

An element of F is formed by taking a multiset of labeled elements of C and performing all consistent relabellings. The corresponding generating function of the elements in F ([20]) is

$$L(z) = \exp(C_L(z)). \quad (2.1)$$

In the case of unlabeled structures, the ordinary generating function of the elements of C is

$$C_U(z) = \sum_{k \geq 0} C_k z^k.$$

Since the objects can be regarded as a product of components, it follows that, the corresponding generating function [20] of F satisfies

$$U(z) = \prod_{k=1}^{\infty} (1 + z^k + z^{2k} + \dots)^{C_k}$$

$$\begin{aligned}
&= \prod_{k=1}^{\infty} \left(\frac{1}{1-z^k} \right)^{C_k} \\
&= \exp \left(\sum_k \log(1-z^k)^{-C_k} \right) \\
&= \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \frac{C_U(z^3)}{3} + \dots \right). \tag{2.2}
\end{aligned}$$

2.3 Singularity Analysis of Generating Functions

In many enumerative problems in computer science and combinatorics one encounters the difficulty that no closed form formulae exist for the quantities of interest and only recurrences for generating functions are available. For solving those problems, complex analysis were brought into combinatorics (e.g. [15, 16, 37]). For some other excellent surveys of combinatorial asymptotics, see [3, 4, 5, 39].

In 1990, Flajolet and Odlyzko published an important paper ([17]) which introduced *singularity analysis* of generating functions. They restricted their attention to functions with a unique dominant singularity. This work presents a method by which one can translate, on a term-by-term basis, an asymptotic expansion of a function around the dominant singularity into a corresponding asymptotic expansion for the Taylor coefficients of the function. This approach is based on contour integration using Cauchy's formula and Hankel-like contours [51]. We extensively use similar contours in this thesis. In the following, we shall give a brief description of their results.

Definition 2.3.1. Let $\Delta(\nu, \theta)$ be the region $|z| \leq 1 + \nu$ minus the region $\theta < \arg(z - 1) < 2\pi - \theta$, with $\nu > 0$ and $0 < \theta < \pi/2$ in the complex plane.

This domain has the form of an indented disk as shown in Figure 2.2. In Flajolet

and Odlyzko's paper, the object function $f(z)$ is asked to be analytic in $\Delta(\nu, \theta)$ except at the singularity $z = 1$. This domain is also used to identify the *exp-log class* and the *alg-log* type which will be discussed in Chapter 4 and 5, respectively.

Figure 2.2: The $\Delta(\nu, \theta)$ region.

Let f_n be the coefficient of z^n in the function $f(z)$. Flajolet and Odlyzko proved the following two transfer theorems.

Proposition 2.3.1. (*O-transfer theorem*) Assume that, with the sole exception of the singularity $z = 1$, $f(z)$ is analytic in the domain $\Delta(\nu, \theta)$, where $\nu > 0$ and $0 < \theta < \pi/2$. Assume further that as z tends to 1 in $\Delta(\nu, \theta)$,

$$f(z) = O\left((1-z)^\alpha L\left(\frac{1}{1-z}\right)\right)$$

where $L(u) = (\log u)^\gamma (\log \log u)^\delta$, for some real numbers α, γ, δ . Then the n th Taylor coefficient of $f(z)$ satisfies

$$f_n = [z^n]f(z) = O(n^{-\alpha-1}L(n)).$$

Proposition 2.3.2. (*o-transfer theorem*) Assume that $f(z)$ is analytic in the domain $\Delta(\nu, \theta) \setminus \{1\}$, and that as z tends to 1 in $\Delta(\nu, \theta)$,

$$f(z) = o\left((1-z)^\alpha L\left(\frac{1}{1-z}\right)\right)$$

where $L(u) = (\log u)^\gamma$. Then, as $n \rightarrow \infty$,

$$f_n = [z^n]f(z) = o(n^{-\alpha-1}L(n)).$$

By normalization, the generating functions of many decomposable structures satisfy the analytic condition in the domain $\Delta(\nu, \theta) \setminus \{1\}$ and have the form $(1-z)^\alpha L\left(\frac{1}{1-z}\right)$. Hence, we can get the asymptotic value of the coefficient which represents the number of decomposable structures with certain size using singularity analysis. For instance, 2-regular graph has the generating function

$$f(z) = \exp\left(\frac{1}{2} \log \frac{1}{1-z} - \frac{z}{2} - \frac{z^2}{4}\right) = (1-z)^{-1/2} \exp\left(-\frac{z}{2} - \frac{z^2}{4}\right)$$

with the coefficient

$$f_n \sim \frac{e^{-3/4}}{\sqrt{\pi}} n^{-1/2}.$$

Flajolet and Odlyzko also provided detailed asymptotic expansions for the coefficient of a set of functions which satisfy the analytic conditions in the domain $\Delta(\nu, \theta)$.

Proposition 2.3.3. Let α and γ be real or complex numbers, $\alpha \notin \{0, 1, 2, \dots\}$.

Define the function $f(z)$ by

$$f(z) = (1-z)^\alpha \left(\frac{1}{z} \log \frac{1}{1-z}\right)^\gamma.$$

Then, the Taylor coefficients of $f(z)$ satisfy

$$f_n \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)} (\log n)^\gamma \left(1 + \sum_{k \geq 1} \frac{e_k^{(\alpha, \gamma)}}{\log^k n}\right)$$

with

$$e_k^{(\alpha, \gamma)} = (-1)^k \binom{\gamma}{k} \Gamma(-\alpha) \frac{d^k}{ds^k} \left(\frac{1}{\Gamma(-s)} \right) \Big|_{s=\alpha},$$

where

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt,$$

when the complex number $z \notin \{0, -1, -2, \dots\}$, is the well known Gamma function.

Chapter 3

Generating Functions of Decomposable Combinatorial Structures

As fundamental mathematical object of combinatorial analysis, generating functions of decomposable combinatorial structures are our first goal in this thesis. As we will see throughout this thesis, we divide every problem into two categories: labeled and unlabeled cases.

In Section 3.1, applying well known generating function of multiset construction, we deduce the generating function of a decomposable combinatorial structure with a given restricted pattern. In Section 3.2, we provide the generating function of a random decomposable combinatorial structure with, in addition to the given restricted pattern, the size of r th smallest component bigger than or equal to some integer k .

Without any assumption or condition, the objects that we study in this thesis can

be any decomposable combinatorial structures. But we notice that, since we keep the coefficients C_j (number of different components that have size j) in our final generating functions, we require knowledge of C_j to apply our results.

The results of this chapter are part of [10].

3.1 Decomposable Combinatorial Structures with Given Restricted Patterns

In this section, we apply the combinatorial methodology to obtain generating function of decomposable combinatorial structure with a given restricted pattern. The results in this section will be repeatedly applied in later chapters.

Before beginning with our first theorem, we make the following assumptions. In both labeled and unlabeled cases, let C be a combinatorial class. According to the introduction in Section 2.2.2, elements of C have different generating functions depending on the labeled or unlabeled case. Let F be a decomposable combinatorial class over C . As we claimed in Chapter 2, F has the *multiset* construction. We use $L(z)$ to denote the generating function of F in the labeled case and $U(z)$ to denote the generating function of F in the unlabeled case. Those are default assumptions for each chapter in this thesis and we will not repeat them again.

3.1.1 Labeled Case

In the labeled case, as we referred to in Section 2.2.2, the exponential generating function of C is

$$C_L(z) = \sum_{j \geq 1} C_j \frac{z^j}{j!},$$

and the exponential generating function of F is

$$L(z) = \exp(C_L(z)).$$

We have a restricted pattern $S : J \mapsto \mathbb{N}$. Let us write $C_L(z; J) = \sum_{j \in J} C_j \frac{z^j}{j!}$.

As one of the most repeatedly used results in this thesis, the following theorem provides the generating function of labeled decomposable combinatorial structures with a given restricted pattern:

Theorem 3.1.1. *Let $S : J \mapsto \mathbb{N}$ be a given restricted pattern. The exponential generating function of labeled combinatorial structures which have size n and restricted pattern S is*

$$L(z; S) = \exp(C_L(z) - C_L(z; J)) \prod_{j \in J} \frac{C_j^{S(j)} z^{jS(j)}}{(j!)^{S(j)} S(j)!}. \quad (3.1)$$

Proof. Obviously, with a given restricted pattern S , the number of components is at least $\sum_{j \in J} S(j)$. Components with sizes not included in S can be combined in any possible way. The generating function of these components can be expressed as $C_L(z) - C_L(z; J)$ and so, we have generating function

$$\sum_{k \geq 0} \frac{\left(C_L(z) - \sum_{j \in J} C_j \frac{z^j}{j!}\right)^k}{k!} = \exp(C_L(z) - C_L(z; J)).$$

For each $j \in J$, we choose $S(j)$ components of size j with repetitions allowed, that is, we have generating function $\frac{C_j^{S(j)} z^{jS(j)}}{(j!)^{S(j)}}$. Hence, the exponential generating function of labeled combinatorial structures with restricted pattern S is

$$L(z; S) = \exp(C_L(z) - C_L(z; J)) \prod_{j \in J} \frac{C_j^{S(j)} z^{jS(j)}}{(j!)^{S(j)} S(j)!}.$$

□

3.1.2 Unlabeled Case

Let $C_U(z)$ and $U(z)$ be the ordinary generating functions for components and decomposable combinatorial structures, respectively. As introduced in Chapter 2, we know that

$$U(z) = \exp\left(C_U(z) + \frac{1}{2}C_U(z^2) + \frac{1}{3}C_U(z^3) + \cdots\right),$$

with

$$C_U(z) = \sum_{k \geq 1} C_k z^k.$$

The following theorem extends the above relation to combinatorial structures with a restricted pattern.

Theorem 3.1.2. *Let $S : J \mapsto \mathbb{N}$ be a given restricted pattern. The generating function of unlabeled combinatorial structures which have size n and restricted pattern S is*

$$U(z; S) = \left(\prod_{j \in J} (1 - z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} \right) \exp\left(C_U(z) + \frac{C_U(z^2)}{2} + \frac{1}{3}C_U(z^3) + \cdots\right). \quad (3.2)$$

Proof. Essentially, an unlabeled combinatorial structure is formed by components. If we do not have a restricted pattern, a component of size k can appear i times

where i is any non-negative integer. The contribution of i components of size k to the generating function is z^{ik} . There are C_k possible components with size k that could appear any number of times. Therefore, generating function of all the components of size k is $(1 + z^k + z^{2k} + \dots)^{C_k}$. Hence, generating function of unlabeled combinatorial structures is

$$\prod_{k \geq 1} (1 + z^k + z^{2k} + \dots)^{C_k} = \prod_{k \geq 1} \left(\frac{1}{1 - z^k} \right)^{C_k}.$$

This generating function has another expression which has been repeatedly applied in combinatorics (for example see [43])

$$\begin{aligned} \prod_{k \geq 1} \left(\frac{1}{1 - z^k} \right)^{C_k} &= \exp \left(\sum_{k \geq 1} \log(1 - z^k)^{-C_k} \right) \\ &= \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right). \end{aligned}$$

Now we fix the number of components with sizes corresponding to the restricted pattern S . If we know that there are $S(i)$ components of size i , there is a small change in the above generating function. Indeed, we have to remove the original factor $(1 + z^i + z^{2i} + \dots)^{C_i}$ of components of size i and put instead $\binom{C_i + S(i) - 1}{S(i)} z^{iS(i)}$ since we are counting with repetition. Repeating this process for every $i \in J$, we have

$$\begin{aligned} U(z; S) &= \prod_{k \geq 1} \left(\frac{1}{1 - z^k} \right)^{C_k} \prod_{j \in J} (1 - z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} \\ &= \left(\prod_{j \in J} (1 - z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} \right) \\ &\quad \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right). \end{aligned}$$

□

3.2 The r th Smallest Component

In this section, we consider combinatorial structures with a given restricted pattern such that the size of the r th smallest component is bigger than or equal to k , $k \geq 1$. We provide generating functions of such decomposable combinatorial structures in both labeled and unlabeled case. Some examples are given in Section 3.3. An analysis on these generating functions could provide the distribution of the r th smallest component of a random decomposable combinatorial structure with a given pattern.

In the following, we consider the generating function of components in two parts $r_k(z)$ and $s_k(z)$. In the unlabeled case, $r_k(z) = \sum_{j \geq k} C_j z^j$, and $s_k(z) = \sum_{j < k} C_j z^j$. In the labeled case, $r_k(z) = \sum_{j \geq k} C_j \frac{z^j}{j!}$, and $s_k(z) = \sum_{j < k} C_j \frac{z^j}{j!}$. We need to give new notations. Let $X_n^{[r]}(S)$ denote the size of the r th smallest component of a decomposable combinatorial structure of size n having a restricted pattern S . We also define the characteristic function

$$I_j = I(j < k) + I(S(j) = 0) - I(j < k)I(S(j) = 0),$$

with

$$I(j < k) = \begin{cases} 1 & \text{if } j < k \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad I(S(j) = 0) = \begin{cases} 1 & \text{if } S(j) = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Finally, we denote by $f^{[j]}(z)$ and $g^{[j]}(z)$ the coefficient of u^j in the function $f(z, u)$ and $g(z, u)$, respectively.

3.2.1 Labeled Case

The generating function for labeled combinatorial structures are given by the following theorem.

Theorem 3.2.1. *Let $S : J \mapsto \mathbb{N}$ be a given restricted pattern and let $L_k^{[r]}(z)$ be the exponential generating function for labeled combinatorial structures of size n having $X_n^{[r]}(S) \geq k$, $1 \leq k \leq n$. Let $f(z, u)$ be*

$$f(z, u) = \exp \left(us_k(z) - \sum_{j \in J} u I(j < k) C_j \frac{z^j}{j!} \right) \left(\exp \left(r_k(z) - \sum_{j \in J} I(j \geq k) C_j \frac{z^j}{j!} \right) - \prod_{j \in J} I_j \right) \prod_{j \in J} \frac{\left(u^{I(j < k)} C_j \frac{z^j}{j!} \right)^{S(j)}}{S(j)!}.$$

Then we have

$$L_k^{[r]}(z) = f^{[0]}(z) + f^{[1]}(z) + \cdots + f^{[r-1]}(z).$$

Proof. We divide the generating function into three parts: those components which come from the given pattern, other components whose size is smaller than k and other components whose size is bigger than or equal to k . Let us begin from the case $|J| = 1$, which means only one element (say j) in J . We consider two conditions: $j \geq k$ and $j < k$.

Case (i): $j \geq k$.

In this case, the given size is bigger than or equal to k . The generating function can be expressed as the product of the generating function of those components whose size is bigger than or equal to k and the generating function of those components whose size is smaller than k . The generating function of the smaller part is $\exp \left(\sum_{i < k} C_i \frac{z^i}{i!} \right)$. In this term, the size of the components is up to $k - 1$. Moreover we use u to count the number of components of size up to $k - 1$. For the bigger part, we have the pattern term which is $\frac{\left(C_j \frac{z^j}{j!} \right)^{S(j)}}{S(j)!}$ and the undefined term which is

obtained by eliminating the pattern from $\exp\left(\sum_{i \geq k} C_i \frac{z^i}{i!}(z)\right)$. Hence we have

$$\begin{aligned} f(z, u) &= \left(\sum_{i \geq S(j)} \frac{\left(\sum_{l \geq k} C_l \frac{z^l}{l!} - C_j \frac{z^j}{j!}\right)^{i-S(j)}}{(i-S(j))!} - I(S(j)=0) \right) \\ &\quad \frac{\left(C_j \frac{z^j}{j!}\right)^{S(j)}}{S(j)!} \exp\left(u \sum_{i < k} C_i \frac{z^i}{i!}(z)\right) \\ &= e^{us_k(z)} \left(\exp\left(r_k(z) - C_j \frac{z^j}{j!}\right) - I(S(j)=0) \right) \frac{\left(C_j \frac{z^j}{j!}\right)^{S(j)}}{S(j)!}. \end{aligned}$$

Case (ii): $j < k$.

Similar to the Case (i), the generating function is divided into two parts, but in this case we need to separate the pattern from the smaller part. We have

$$\begin{aligned} f(z, u) &= \sum_{i \geq S(j)} \frac{\left(u \left(s_k(z) - C_j \frac{z^j}{j!}\right)\right)^{i-S(j)}}{(i-S(j))!} \frac{\left(u C_j \frac{z^j}{j!}\right)^{S(j)}}{S(j)!} (e^{r_k(z)} - 1) \\ &= (e^{r_k(z)} - 1) \exp\left(u \left(s_k(z) - C_j \frac{z^j}{j!}\right)\right) \frac{\left(u C_j \frac{z^j}{j!}\right)^{S(j)}}{S(j)!}. \end{aligned} \quad (3.3)$$

Then we combine these two cases using the characteristic function I_j

$$\begin{aligned} f(z, u) &= \left(\exp\left(r_k(z) - I(j \geq k) C_j \frac{z^j}{j!}\right) - I_j \right) \\ &\quad \exp\left(us_k(z) - u I(j < k) C_j \frac{z^j}{j!}\right) \frac{\left(u^{I(j < k)} C_j \frac{z^j}{j!}\right)^{S(j)}}{S(j)!}. \end{aligned} \quad (3.4)$$

Because the size of the r th smallest component is bigger than or equal to k , we can have at most $r-1$ components whose size is smaller than k . In the previous equation, u is used to count the number of components whose size is smaller than k . Hence the generating function $L_k^{[r]}(z)$ is the sum of the coefficients of u^i , $0 \leq i \leq r-1$.

So the generating function we are looking for is

$$L_k^{[r]}(z) = f^{[0]}(z) + f^{[1]}(z) + \cdots + f^{[r-1]}(z).$$

For the general case, as an arbitrary pattern $S : J \mapsto \mathbb{N}$ is given, we apply the same method and get

$$L_k^{[r]}(z) = f^{[0]}(z) + f^{[1]}(z) + \cdots + f^{[r-1]}(z), \quad (3.5)$$

with

$$\begin{aligned} f(z, u) = & \exp \left(us_k(z) - \sum_{j \in J} u I(j < k) C_j \frac{z^j}{j!} \right) \\ & \left(\exp \left(r_k(z) - \sum_{j \in J} I(j \geq k) C_j \frac{z^j}{j!} \right) - \prod_{j \in J} I_j \right) \\ & \prod_{j \in J} \frac{\left(u^{I(j < k)} C_j \frac{z^j}{j!} \right)^{S(j)}}{S(j)!}. \end{aligned}$$

□

3.2.2 Unlabeled Case

The unlabeled case is similar to the labeled case with ordinary generating functions instead of exponential generating functions.

Theorem 3.2.2. *Given a restricted given pattern $S : J \mapsto \mathbb{N}$, let $U_k^{[r]}(z)$ be the ordinary generating function for unlabeled combinatorial structures of size n having pattern S and $X_n^{[r]}(S) \geq k$, $1 \leq k \leq n$. Let $g(z, u)$ be the function*

$$\begin{aligned} g(z, u) = & \left(\prod_{j \in J} (1 - u^{I(j < k)} z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} u^{S(j)I(j < k)} \right) \\ & \left(\exp \left(r_k(z) + \frac{r_k(z^2)}{2} + \cdots \right) - \prod_{j \in J} I_j \right) \end{aligned}$$

$$\exp \left(u s_k(z) + u^2 \frac{s_k(z^2)}{2} + \dots \right). \quad (3.6)$$

Then generating function of the decomposable combinatorial structure is

$$U_k^{[r]}(z) = g^{[0]}(z) + g^{[1]}(z) + \dots + g^{[r-1]}(z). \quad (3.7)$$

Proof:

The proof is similar to the one of Theorem 3.2.1. We consider the generating function in three parts. Let us begin with the case $|J| = 1$ which means only one size component in the pattern, say $J = \{j\}$.

Case (i): $j \geq k$.

For the smaller part, similar to the proof of Theorem 3.1.2, the generating function is $\prod_{i < k} \left(\frac{1}{1 - uz^i} \right)^{C_i}$. In this term, the size of a component is up to $k - 1$. Moreover we use u to count the number of components of size up to $k - 1$. Later we will impose $r - 1$ components of size at most $k - 1$. For the bigger part, also similar to Theorem 3.1.2, we use $\binom{C_i + \alpha_i - 1}{\alpha_i} z^{i\alpha_i}$ to replace $\left(\frac{1}{1 - z^i} \right)^{C_i}$ as the generating function of components whose size is i . Then we have

$$\begin{aligned} g(z, u) &= \left(\prod_{i \geq k} \left(\frac{1}{1 - z^i} \right)^{C_i} - I(S(j) = 0) \right) (1 - z^j)^{C_j} \\ &\quad \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} \prod_{i < k} \left(\frac{1}{1 - uz^i} \right)^{C_i} \\ &= (1 - z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} \exp \left(\sum_{i < k} C_i u z^i + \sum_{i < k} C_i u^2 \frac{z^{2i}}{2} + \dots \right) \\ &\quad \left(\exp \left(\sum_{i \geq k} C_i z^i + \sum_{i \geq k} C_i \frac{z^{2i}}{2} + \dots \right) - I(S(j) = 0) \right) \\ &= (1 - z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} \exp \left(u s_k(z) + u^2 \frac{s_k(z^2)}{2} + \dots \right) \end{aligned}$$

$$\left(\exp \left(r_k(z) + \frac{r_k(z^2)}{2} + \dots \right) - I(S(j) = 0) \right).$$

Case (ii): $j < k$.

In this case we have

$$\begin{aligned} g(z, u) &= \prod_{i < k} \left(\frac{1}{1 - uz^i} \right)^{C_i} (1 - uz^k)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} u^{S(j)} \\ &\quad \left(\prod_{i \geq k} \left(\frac{1}{1 - z^i} \right)^{C_i} - 1 \right) \\ &= (1 - uz^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} u^{S(j)} \\ &\quad \exp \left(\sum_{i < k} C_i uz^i + \sum_{i < k} C_i u^2 \frac{z^{2i}}{2} + \dots \right) \\ &\quad \left(\exp \left(\sum_{i \geq k} C_i z^i + \sum_{i \geq k} C_i \frac{z^{2i}}{2} + \dots \right) - 1 \right) \\ &= (1 - uz^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} u^{S(j)} \\ &\quad \left(\exp \left(r_k(z) + \frac{r_k(z^2)}{2} + \dots \right) - 1 \right) \exp \left(us_k(z) + u^2 \frac{s_k(z^2)}{2} + \dots \right). \end{aligned}$$

Then we combine these two cases, using the characteristic function I_j

$$\begin{aligned} g(z, u) &= (1 - u^{I(j < k)} z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} u^{S(j)I(j < k)} \\ &\quad \left(\exp \left(r_k(z) + \frac{r_k(z^2)}{2} + \dots \right) - I_j \right) \\ &\quad \exp \left(us_k(z) + u^2 \frac{s_k(z^2)}{2} + \dots \right). \end{aligned}$$

Using the same reasoning as the one in the proof of Theorem 3.2.1, the generating function we are looking for is

$$U_k^{[r]}(z) = g^{[0]}(z) + g^{[1]}(z) + \dots + g^{[r-1]}(z).$$

In general, the number of elements in J is arbitrary. By the same method, we get

$$g(z, u) = \left(\prod_{j \in J} (1 - u^{I(j < k)} z^j)^{C_j} \binom{C_j + S(j) - 1}{S(j)} z^{jS(j)} u^{S(j)I(j < k)} \right) \\ \left(\exp \left(r_k(z) + \frac{r_k(z^2)}{2} + \dots \right) - \prod_{j \in J} I_j \right) \\ \exp \left(us_k(z) + u^2 \frac{s_k(z^2)}{2} + \dots \right),$$

and

$$U_k^{[r]}(z) = g^{[0]}(z) + g^{[1]}(z) + \dots + g^{[r-1]}(z).$$

□

3.3 Examples

In this section, we provide examples which cover some important decomposable combinatorial structures, such as permutations, polynomials over finite fields and rooted nonplane trees. We apply our theorems to calculate the number of decomposable combinatorial structures of a given size with a restricted pattern. Some of these applications may be of interest for practitioners in cryptography (see [13] and [45]).

Example 3.3.1. (*Cycles in Permutations*). The simplest example for the labeled case comes from permutations viewed as a product of cycles. The generating function for the cycles of permutations ([52]) is

$$C_L(z) = \ln \frac{1}{1-z} = z + \frac{z^2}{2!} + 2! \frac{z^3}{3!} + 3! \frac{z^4}{4!} + \dots.$$

Let us consider the restricted pattern $\alpha_2 = 2$ and $\alpha_4 = 1$. Thus, from Theorem 3.1.1

the generating function is

$$L(z; S) = \left(1 + \left(z + 2 \frac{z^3}{3!} + \dots \right) + \frac{\left(z + 2 \frac{z^3}{3!} + \dots \right)^2}{2!} + \dots \right) \frac{1^2 z^4}{2^2 2!} \frac{6 z^4}{4!}.$$

If we want the number of permutations of size 11 having 2 cycles of length 2 and 1 cycle of length 4, we have

$$\left[\frac{z^{11}}{11!} \right] L(z; S) = 11! \frac{1}{64} = 623700.$$

Now let us take another restricted pattern $S(\lfloor \log n \rfloor) = 1$. The generating function of the permutations whose smallest cycle size is bigger than or equal to v with $v < \lfloor \log n \rfloor$ can be obtained by applying Theorem 3.2.1

$$L_v^{[1]}(z) = \exp \left(\sum_{i \geq v} \frac{z^i}{i} - \frac{z^{\lfloor \log n \rfloor}}{\lfloor \log n \rfloor} \right) \frac{z^{\lfloor \log n \rfloor}}{\lfloor \log n \rfloor}.$$

Example 3.3.2. (*Irreducible Factors in Polynomials over Finite Fields*). A typical unlabeled structure is polynomials over finite fields. The generating function for polynomials over $\mathbb{F}_q$ ($[20]$) is

$$U(z) = \frac{1}{1 - qz} = \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right),$$

with

$$C_U(z) = \log \frac{1}{1 - qz} + F(z),$$

where $F(z)$ is an analytic function in $|z| < q^{-1/2}$. As an example, let us consider the binary field $\mathbb{F}_2$ and assume $n = 6$ and $S(2) = 1$, that is, we look for the number of polynomials of degree 6 over $\mathbb{F}_2$ with exactly one irreducible factor of degree 2. The generating function of irreducible polynomials over $\mathbb{F}_2$ is

$$C_U(z) = 2z + z^2 + 2z^3 + 3z^4 + \dots$$

$x(x^2 + x + 1)(x^3 + x + 1)$	$x(x^2 + x + 1)(x^3 + x^2 + 1)$
$(x + 1)(x^2 + x + 1)(x^3 + x + 1)$	$(x + 1)(x^2 + x + 1)(x^3 + x^2 + 1)$
$(x^2 + x + 1)(x^4 + x + 1)$	$(x^2 + x + 1)(x^4 + x^3 + 1)$
$(x^2 + x + 1)(x^4 + x^3 + x^2 + x + 1)$	$x^4(x^2 + x + 1)$
$x^3(x + 1)(x^2 + x + 1)$	$x^2(x + 1)^2(x^2 + x + 1)$
$x(x + 1)^3(x^2 + x + 1)$	$(x + 1)^4(x^2 + x + 1)$

Table 3.1: Polynomials of degree 6 with the restricted pattern in Example 3.3.2.

From Theorem 3.1.2, we have the generating function

$$\begin{aligned}
 U(z; S) &= (1 - z^2) \binom{1 + 1 - 1}{1} z^2 \left(1 + \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right) \right. \\
 &\quad \left. + \frac{\left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right)^2}{2!} + \frac{\left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right)^3}{3!} + \dots \right) \\
 &= z^2 + 2z^3 + 3z^4 + 6z^5 + 12z^6 + \dots
 \end{aligned}$$

Hence, $[z^6]U(z; S) = 12$. The 12 polynomials over $\mathbb{F}_2$ of degree 6 with exactly one irreducible factor of degree 2 are given in Table 3.1.

Given a polynomial over $\mathbb{F}_q$ with the restricted pattern $S(2) = 1$, if we know the size of the second small component is bigger than or equal to v where $2 < v < n$, what is the generating function as $n \rightarrow \infty$? By Theorem 3.2.2, we have

$$\begin{aligned}
 f(z, u) &= uz^2(1 - uz^2) \left(\exp \left(r_v(z) + \frac{r_v(z^2)}{2} + \dots \right) - 1 \right) \\
 &\quad \exp \left(us_v(z) + u^2 \frac{s_v(z^2)}{2} + \dots \right).
 \end{aligned}$$

Therefore, the generating function is

$$\begin{aligned}
 L_v^{[2]}(z) &= f^{[0]}(z) + f^{[1]}(z) \\
 &= z^2 \left(\exp \left(r_v(z) + \frac{r_v(z^2)}{2} + \dots \right) - 1 \right).
 \end{aligned}$$

$(x^2 + x + 1)(x^4 + x + 1)$
$(x^2 + x + 1)(x^4 + x^3 + 1)$
$(x^2 + x + 1)(x^4 + x^3 + x^2 + x + 1)$

Table 3.2: Polynomials of degree 6 with the restricted pattern $S(2) = 1$ and the second smallest component has size bigger than or equal to 3.

Let $v = 3$, we have the generating function

$$L_3^{[2]}(z) = z^2 \left(\exp \left((2z^3 + 3z^4 + \dots) + \frac{2z^6 + 3z^8 + \dots}{2} + \dots \right) - 1 \right).$$

Then, the number of polynomials of degree 6 when the second smallest component has size bigger than or equal to 3 and there is only one factor of degree 2 is

$$[z^6]L_3^{[2]}(z) = 3.$$

We list the corresponding polynomials in Table 3.2

Example 3.3.3. (*Rooted Nonplane Trees*). Another unlabeled example is rooted nonplane tree. A rooted nonplane tree, also called rooted unordered tree, is a root node linked to a multiset of trees. The generating function of the components ([52]) is

$$C_U(z) = z + z^2 + 2z^3 + 4z^4 + \dots$$

Let the restricted pattern be $S = \{\alpha_1 = 0\}$, that is, rooted nonplane trees with no subtree of the root with 1 vertex. From Theorem 3.1.2 the generating function of rooted nonplane trees with restricted pattern S is

$$U(z; S) = 1 + z^2 + 2z^3 + 5z^4 + \dots$$

So, if we want to know the number of rooted nonplane trees of size n which have no subtree formed by 1 vertex, we can use this generating function. We observe that a rooted nonplane tree consists of the root node and subtrees. Hence, if we want to

Figure 3.1: The 5 vertices rooted nonplane trees without subtrees formed by 1 vertex.

know the number of rooted nonplane trees on 5 vertices with no subtree formed by 1 vertex, we want the coefficient of z^4 in $U(z; S)$. As Figure 3.1 demonstrates, the answer is

$$[z^4]U(z; S) = 5.$$

Chapter 4

Decomposable Combinatorial Structures in Exp-Log Class

From the discussion in Chapters 2 and 3, we can understand the relationship between a decomposable combinatorial structure and its generating function. A very natural classification of combinatorial structures is based on their generating functions. The exp-log class introduced by Flajolet and Soria (see [20, 21]) is a class of decomposable combinatorial structures which has multiset construction and generating functions of the components are of the so-called *logarithmic* type.

In this chapter, using the method of analysis of singularities, we provide the probability that a decomposable structure of size n in exp-log class has a given restricted pattern. Furthermore, we estimate the probability that a random decomposable combinatorial structure has a given restricted pattern and the size of its r th smallest component is bigger than k , for r, k given integers. Our studies apply to labeled and unlabeled structures. We also give several concrete examples in Section 4.5. The results in this chapter are published in [10, 11].

4.1 Introduction

We start by defining generating functions of the logarithmic type.

Definition 4.1.1. *Let $\Delta(\nu, \theta)$ be the region defined in Definition 2.3.1. We say that $C(z)$ is of logarithmic type with multiplicity constant $a > 0$ if*

$$C(\rho z) = a \log \left(\frac{1}{1-z} \right) + R(z),$$

where ρ is a positive constant, $R(z)$ is analytic in $\Delta(\nu, \theta)$, and as $z \rightarrow 1$ in $\Delta(\nu, \theta)$,

$$R(z) = K + O((1-z)^\alpha)$$

with $0 < \alpha < 1$ and K is a complex constant.

We name the combinatorial class which is a set of elements of logarithmic type as exp-log class. We assume that generating functions of the structures we study in this chapter to be of this type.

Definition 4.1.2. *We say that $e^{C(z)}$ is in exp-log class if $C(z)$ is of logarithmic type.*

Some examples of decomposable combinatorial structures which belong to exp-log class are: permutations, 2-regular graph, polynomials over finite fields, rooted nonplane trees, children's yards, etc. More examples can be found in [21].

We note that the theorems of singularity analysis work in the region $\Delta(\nu, \theta)$ (see Section 2.3). Therefore, applying Propositions 2.3.1 and 2.3.3, we obtain

$$[z^n]C(z) = \frac{1}{\rho^n} \frac{a}{n} (1 + O(n^{-\alpha})). \quad (4.1)$$

In the labeled case, it follows from (2.1) and Flajolet-Odlyzko's transfer theorem that

$$[z^n]L(z) = \frac{1}{\rho^n} \frac{e^K}{\Gamma(a)} n^{a-1} (1 + O(n^{-\alpha})). \quad (4.2)$$

In the unlabeled case, we require the singular point $0 < \rho < 1$ such that the radius of convergence of

$$\frac{C(z^2)}{2} + \frac{C(z^3)}{3} + \dots$$

is bigger than ρ . This requirement holds for all unlabeled decomposable combinatorial structures in this thesis. Let

$$r_0 = \frac{C(\rho^2)}{2} + \frac{C(\rho^3)}{3} + \dots.$$

Then the transfer theorem and (2.2) entails

$$[z^n]U(z) = \frac{1}{\rho^n} \frac{e^{K+r_0}}{\Gamma(a)} n^{a-1} (1 + O(n^{-\alpha})). \quad (4.3)$$

In this chapter, to simplify the notation, we use $C(z)$ and $C(z; J)$ instead of $C_L(z)$ ($C_U(z)$) and $C_L(z; J)$ ($C_U(z; J)$) when they do not lead to misunderstanding.

4.2 Enumeration of Decomposable Combinatorial Structures with Restricted Patterns

Let us assume we have a restricted pattern $S : J \mapsto \mathbb{N}$ as defined in Chapter 2. We are interested in the probability that a structure of size n has this restricted pattern S . We consider labeled and unlabeled cases as in former sections.

In the rest of this thesis we use the following notations (we recall that C_j is the number of components of size j):

- $|A|$ denotes the number of elements in a set A ;
- $d(k) = \sum_{j \in J \cap N_k} S(j)$;

- $C_L(z; J) = \sum_{j \in J} C_j \frac{z^j}{j!}$ and $\mathcal{S}_L(S) = \prod_{j \in J} \frac{C_j^{S(j)}}{(j!)^{S(j)} S(j)!}$;
- $C_U(z; J) = \sum_{j \in J} C_j z^j$ and $\mathcal{S}_U(S) = \prod_{j \in J} \binom{C_j + S(j) - 1}{S(j)}$;
- $m = n - \sum_{j \in J} jS(j)$ denotes the size of the unrestricted part of the structure of size n (one may think of m as the degree of freedom of a structure of size n with a given restricted pattern S). we always assume $m \rightarrow \infty$ in our theorems.

4.2.1 Labeled Case

We apply singularity analysis of generating functions to obtain the probability that a random labeled structure has a given restricted pattern. From Theorem 3.1.1 we have

$$[z^n]L(z; S) = \mathcal{S}_L(S)[z^m] \exp(C(z) - C(z; J)).$$

The following theorem gives the asymptotic probability that a random labeled combinatorial structure has a restricted pattern for the case $m \rightarrow \infty$.

Theorem 4.2.1. *Let $P_L(S, n)$ be the probability that a random decomposable labeled structure with size n has a restricted pattern S . Let $J^{(1)} = \{j \in J : j \leq m\}$ and $J^{(2)} = \{j \in J : j > m\}$. Suppose $|J^{(1)}| = o(m/\log m)$, $j = O(m/\log m)$ when $j \in J^{(1)}$, and $m \rightarrow \infty$ as $n \rightarrow \infty$. Then, as $n \rightarrow \infty$,*

$$P_L(S, n) \sim \left(\frac{m}{n}\right)^{a-1} \mathcal{S}_L(S) \rho^{n-m} \exp(-C_L(\rho; J^{(1)})).$$

Proof. We know that

$$P_L(S, n) = \frac{[z^n]L(z; S)}{[z^n]L(z)}.$$

Define restricted patterns $S^{(1)}$ and $S^{(2)}$ such that $S^{(1)}(j) = S(j)$ for $j \in J^{(1)}$ and $S^{(2)}(j) = S(j)$ for $j \in J^{(2)}$. Let $n^* = n - \sum_{j \in J^{(2)}} j S^{(2)}(j)$. Applying Theorem 3.1.1, we obtain

$$\begin{aligned} P_L(S, n) &= \frac{[z^n] \exp(C_L(z) - C_L(z; J)) \mathcal{S}_L(S) z^{n-m}}{[z^n] L(z)} \\ &= \frac{\mathcal{S}_L(S^{(1)} \cup S^{(2)}) [z^m] \exp(C_L(z) - C_L(z; J^{(1)}) - C_L(z; J^{(2)}))}{[z^n] L(z)}. \end{aligned}$$

Since $m < \min\{j : j \in J^{(2)}\}$, we can remove the term $C_L(z; J^{(2)})$ without changing the result, namely,

$$[z^m] \exp(C_L(z) - C_L(z; J^{(1)}) - C_L(z; J^{(2)})) = [z^m] \exp(C_L(z) - C_L(z; J^{(1)})).$$

Therefore, we obtain

$$P_L(S, n) \sim \mathcal{S}_L(S^{(2)}) \frac{[z^{n^*}] L(z; S^{(1)})}{[z^n] L(z)}. \quad (4.4)$$

For the term $[z^{n^*}] L(z; S^{(1)})$, we shall use the Cauchy integration formula. We choose a contour similar to the one in [17]. Let $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$ be positively oriented, where (see Figure 4.1)

$$\begin{aligned} \gamma_1 &= \{z : |z - 1| = 1/m, \theta < \arg(z - 1) < 2\pi - \theta\}, \\ \gamma_2 &= \{z : 1/m \leq |z - 1| \leq E \log m/m, \arg(z - 1) = \theta\}, \\ \gamma_3 &= \{z : |z| = 1 + (3a + 1) \log m/m, \theta < \arg(z - 1) < 2\pi - \theta\}, \\ \gamma_4 &= \{z : 1/m \leq |z - 1| \leq E \log m/m, \arg(z - 1) = 2\pi - \theta\}, \end{aligned}$$

where $E = O(1)$ satisfies

$$\left| 1 + E \frac{\log m}{m} e^{i\theta} \right| = 1 + (3a + 1) \log m/m.$$

Let

$$f(z) = \frac{\exp(C(\rho z) - C(\rho z; J^{(1)}))}{z^{m+1}}.$$

Figure 4.1: The contour γ .

Then we have

$$[z^{n^*}]L(z; S^{(1)}) = \rho^{-n^*} [z^{n^*}]L(\rho z; S^{(1)}) = \rho^{-m} \mathcal{S}_L(S^{(1)}) \frac{1}{2\pi i} \int_{\gamma} f(z) dz. \quad (4.5)$$

We first analyze the integral on contour $\gamma' = \gamma_1 + \gamma_2 + \gamma_4$. Later we will show that the integral over γ_3 is negligible. Let $z = 1 + w/m$. In the following analysis, we note $|w| = O(\log m)$ when $z \in \gamma'$ and $j = O(m/\log m)$ for $j \in J^{(1)}$. We have

$$C(\rho z) = a \log \left(-\frac{w}{m} \right)^{-1} + K + O \left(\left(\frac{\log m}{m} \right)^{\alpha} \right),$$

and

$$\exp(C(\rho z)) = \left(-\frac{m}{w} \right)^a e^K (1 + o(1)).$$

Since

$$z^j = \left(1 + \frac{w}{m} \right)^j = \exp \left(j \log \left(1 + \frac{w}{m} \right) \right)$$

$$= \exp(O(jw/m)) = 1 + O(jw/m),$$

we also have

$$\begin{aligned} \sum_{j \in J^{(1)}} C_j \frac{(\rho z)^j}{j!} &= \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} \left(1 + O\left(\frac{jw}{m}\right)\right) \\ &= C(\rho; J^{(1)}) + O\left(\frac{w}{m}\right) \sum_{j \in J^{(1)}} j C_j \frac{\rho^j}{j!}. \end{aligned}$$

Now since by (4.1)

$$C_j \frac{\rho^j}{j!} \sim \frac{a}{j},$$

it follows that

$$\sum_{j \in J^{(1)}} C_j \frac{(\rho z)^j}{j!} = C(\rho; J^{(1)}) + O\left(|J^{(1)}| \frac{w}{m}\right) = C(\rho; J^{(1)}) + o(1).$$

We comment that the term $o(1)$ above is uniform for w and $J^{(1)}$ in the ranges under consideration. This also holds for all the $O(\cdot)$ and $o(\cdot)$ terms below, without explicitly saying so.

Putting all pieces together, we have

$$\begin{aligned} &\frac{1}{2\pi i} \int_{\gamma'} f(z) dz \\ &= e^K m^{a-1} \frac{1}{2\pi i} \int_{\gamma''} \left(-\frac{1}{w}\right)^a (1 + o(1)) \exp(-C(\rho; J^{(1)}) + o(1)) \\ &\quad \left(1 + \frac{w}{m}\right)^{-(m+1)} dw \\ &= e^K m^{a-1} \exp(-C(\rho; J^{(1)})) \frac{1}{2\pi i} \int_{\gamma''} \left(-\frac{1}{w}\right)^a e^{-w} (1 + o(1)) dw, \end{aligned}$$

where $\gamma'' = \gamma'_1 + \gamma'_2 + \gamma'_3$,

$$\gamma'_1 = \{w : |w| = 1, \theta < \text{Arg}(w) < 2\pi - \theta\},$$

$$\gamma'_2 = \{w : 1 \leq |w| \leq E \log m, \text{Arg}(w) = \theta\},$$

$$\gamma'_3 = \{w : 1 \leq |w| \leq E \log m, \text{Arg}(w) = 2\pi - \theta\}.$$

We require the following property of the Gamma function [30, 51]

$$\frac{1}{\Gamma(a)} = \frac{1}{2\pi i} \int_{\mathcal{C}} (-w)^{-a} \exp(-w) dw,$$

where $\mathcal{C}$ denotes a contour starting at $+\infty$ in the lower plane, winding clockwise around the origin, and proceeding towards $+\infty$ in the upper half plane. Hence we have

$$\frac{1}{2\pi i} \int_{\gamma''} (-w)^{-a} \exp(-w) dw \sim \frac{1}{\Gamma(a)}.$$

It follows that

$$\begin{aligned} \frac{1}{2\pi i} \int_{\gamma'} f(z) dz &= \frac{1}{\Gamma(a)} e^K m^{a-1} \exp(-C(\rho; J^{(1)}) + o(1)) \\ &\quad + e^K m^{a-1} \exp(-C(\rho; J^{(1)})) \frac{1}{2\pi i} \int_{\gamma''} \left(-\frac{1}{w}\right)^a e^{-w} o(1) dw. \end{aligned}$$

It is not difficult to see that

$$\int_{\gamma''} \left| \left(-\frac{1}{w}\right)^a e^{-w} \right| |dw| = O(1).$$

Hence we have

$$\frac{1}{2\pi i} \int_{\gamma'} f(z) dz = \frac{1}{\Gamma(a)} e^K m^{a-1} \exp(-C(\rho; J^{(1)})) (1 + o(1)). \quad (4.6)$$

Now, we show that the integral over γ_3 is negligible compared with that over γ' . We have $z = (1 + \log m/m) e^{it}$ where $\theta < t < 2\pi - \theta$. Therefore, for $z \in \gamma_3$,

$$\begin{aligned} \exp(C(\rho z)) &= \exp(-a \log(1 - (1 + (3a + 1) \log m/m) e^{it}) + K + O(1)) \\ &= O(\exp(-a \log((3a + 1) \log m/m))) \\ &= O((m/\log m)^a). \end{aligned} \quad (4.7)$$

Using arguments similar to those used to derive (4.6), we obtain, for $z \in \gamma_3$,

$$\begin{aligned} |C(\rho z; J^{(1)})| &\leq \sum_{j \in J^{(1)}} \frac{C_j \rho^j}{j!} \left(1 + \frac{(3a+1) \log m}{m}\right)^j \\ &\leq \sum_{j \in J^{(1)}} \frac{a + o(1)}{j} \left(1 + O\left(\frac{j \log m}{m}\right)\right) \\ &\leq 1.5a \log |J^{(1)}| + O(1). \end{aligned}$$

Hence

$$\begin{aligned} \exp(-C(\rho z; J^{(1)})) &= \exp(-C(\rho; J^{(1)})) \exp(C(\rho; J^{(1)}) - C(\rho z; J^{(1)})) \\ &= \exp(-C(\rho; J^{(1)})) O(\exp(3a \log |J^{(1)}|)) \\ &= O\left(\left(\frac{m}{\log m}\right)^{3a} \exp(-C(\rho; J^{(1)}))\right). \end{aligned}$$

We observe that

$$|z^{-(m+1)}| = |(1 + (3a+1) \log m/m)^{-(m+1)}| = O\left(\frac{1}{m^{3a+1}}\right),$$

and thus we obtain

$$\frac{1}{2\pi i} \int_{\gamma_3} f(z) dz = O((\log m)^{-4a} m^{a-1} \exp(-C(\rho; J^{(1)}))),$$

which is negligible compared with (4.6).

Finally, applying (4.4), (4.5) and (4.6), we obtain

$$P_L(S, n) \sim \mathcal{S}_L(S) \rho^{n-m} \left(\frac{m}{n}\right)^{a-1} \exp(-C_L(\rho; J^{(1)}))$$

This completes the proof of Theorem 4.2.1.

□

Remark. Theorem 4.2.1 restricts the number of components and the maximum component size in $J^{(1)}$ in the restricted pattern. Relaxing the restriction seems to be

hard and may require using a different contour.

4.2.2 Unlabeled Case

In terms of the unlabeled case, by Theorem 3.1.2, the generating function of unlabeled objects with a given restricted pattern is

$$U(z; S) = \mathcal{S}_U(S) z^{n-m} \prod_{j \in J} (1 - z^j)^{C_j} \exp \left(C(z) + \frac{C(z^2)}{2} + \cdots \right).$$

It follows that

$$[z^n]U(z; S) = \mathcal{S}_U(S) [z^m] \prod_{j \in J} (1 - z^j)^{C_j} \exp \left(C(z) + \frac{C(z^2)}{2} + \cdots \right). \quad (4.8)$$

Theorem 4.2.2 gives the asymptotic probability that a random unlabeled object has a given restricted pattern in the case $m \rightarrow \infty$.

Theorem 4.2.2. *Let $P_U(S, n)$ be the probability that a random decomposable unlabeled structure with size n has a restricted pattern S . Suppose $0 < \rho < 1$. Let $J^{(1)} = \{j \in J : j \leq m\}$ and $J^{(2)} = \{j \in J : j > m\}$. Suppose $|J^{(1)}| = o(m/\log m)$, $j = O(m/\log m)$ when $j \in J^{(1)}$, and $m \rightarrow \infty$ as $n \rightarrow \infty$. Then, as $n \rightarrow \infty$,*

$$P_U(S, n) \sim \left(\frac{m}{n} \right)^{a-1} \mathcal{S}_U(S) \rho^{n-m} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j}.$$

Proof. The proof is similar to the one of Theorem 4.2.1. We apply

$$P_U(S, n) = \frac{[z^n]U(z; S)}{[z^n]U(z)}.$$

Let us define restricted patterns $S^{(1)}$ and $S^{(2)}$ as $S^{(1)}(j) = S(j)$ for $j \in J^{(1)}$ and $S^{(2)}(j) = S(j)$ for $j \in J^{(2)}$. Let $n^* = n - \sum_{j \in J^{(2)}} j S^{(2)}(j)$. Applying Theorem 4.3, we obtain

$$\begin{aligned}
P_U(S, n) &= \frac{[z^n] \prod_{j \in J} (1 - z^j)^{C_j} \mathcal{S}_U(S) z^{n-m} \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right)}{[z^n] U(z)} \\
&= \frac{\mathcal{S}_U(S^{(1)} \cup S^{(2)}) [z^m] \prod_{j \in J^{(1)} \cup J^{(2)}} (1 - z^j)^{C_j} \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right)}{[z^n] U(z)}.
\end{aligned}$$

Since $m < \min\{j : j \in J^{(2)}\}$, it is clear that

$$\begin{aligned}
&[z^m] \prod_{j \in J^{(1)} \cup J^{(2)}} (1 - z^j)^{C_j} \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right) \\
&= [z^m] \prod_{j \in J^{(1)}} (1 - z^j)^{C_j} \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \dots \right).
\end{aligned}$$

Therefore, we obtain

$$P_U(S, n) \sim \mathcal{S}_U(S^{(2)}) \frac{[z^{n^*}] U(z; S^{(1)})}{[z^n] U(z)}. \quad (4.9)$$

For the term $[z^{n^*}] U(z; S^{(1)})$, we use the Cauchy integration formula. We choose the same contour as the one in the labeled case (see Figure 4.1). Let $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$ be positively oriented, where

$$\begin{aligned}
\gamma_1 &= \{z : |z - 1| = 1/m, \theta < \text{Arg}(z - 1) < 2\pi - \theta\}, \\
\gamma_2 &= \{z : 1/m \leq |z - 1| \leq 1 + E \log m/m, \text{Arg}(z - 1) = \theta\}, \\
\gamma_3 &= \{z : |z| = 1 + (3a + 1) \log m/m, \theta < \text{Arg}(z - 1) < 2\pi - \theta\}, \\
\gamma_4 &= \{z : 1/m \leq |z - 1| \leq 1 + E \log m/m, \text{Arg}(z - 1) = 2\pi - \theta\},
\end{aligned}$$

where E satisfies $|1 + E \frac{\log m}{m} e^{i\theta}| = 1 + \log m/m$.

Let

$$f(z) = \prod_{j \in J} (1 - (\rho z)^j)^{C_j} \exp \left(C(\rho z) + \frac{C((\rho z)^2)}{2} + \dots \right) z^{-(m+1)}.$$

Then we have

$$[z^{n^*}]U(z; S^{(1)}) = \rho^{-n^*} [z^{n^*}]U(\rho z; S^{(1)}) = \rho^{-m} \mathcal{S}_U(S^{(1)}) \frac{1}{2\pi i} \int_{\gamma} f(z) dz.$$

Applying the same method as in the proof of Theorem 4.2.1, we begin with the integral on contour $\gamma' = \gamma_1 + \gamma_2 + \gamma_4$. Let $z = 1 + w/m$ and we observe that $|w| = O(\log m)$ when $z \in \gamma'$.

Since $0 < \rho < 1$ in the unlabeled case,

$$\frac{C((\rho z)^2)}{2} + \frac{C((\rho z)^3)}{3} + \dots$$

has a radius of convergence bigger than 1. Hence $\exp\left(\frac{C((\rho z)^2)}{2} + \dots\right)$ is bounded for $z \in \gamma$. Let

$$r_0 = \frac{C(\rho^2)}{2} + \frac{C(\rho^3)}{3} + \dots.$$

We then have, for $z \in \gamma'$,

$$\frac{C((\rho z)^2)}{2} + \frac{C((\rho z)^3)}{3} + \dots = r_0(1 + o(1)),$$

and using (4.6), we get

$$\begin{aligned} \exp\left(C(\rho z) + \frac{C((\rho z)^2)}{2} + \frac{C((\rho z)^3)}{3} + \dots\right) &= e^{r_0} \exp(C(\rho z))(1 + o(1)) \\ &= e^{r_0} (-m/w)^a e^K (1 + o(1)). \end{aligned}$$

Using $0 < \rho < 1$ and $C_j \rho^j = O(1/j)$ as $j \rightarrow \infty$, we have

$$\begin{aligned} \prod_{j \in J^{(1)}} (1 - (\rho z)^j)^{C_j} &= \prod_{j \in J^{(1)}} \left(1 - \rho^j + O\left(\frac{\rho^j j w}{m}\right)\right)^{C_j} \\ &= \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \left(1 + O\left(\frac{w}{m C_j}\right)\right)^{C_j} \\ &= \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \left(1 + O\left(\frac{w}{m}\right)\right) \\ &= \left(1 + O\left(\frac{|J^{(1)}| w}{m}\right)\right) \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j}. \end{aligned}$$

Since $w = O(\log m)$ and $|J^{(1)}| = o(m/\log m)$, it follows that

$$\prod_{j \in J^{(1)}} (1 - (\rho z)^j)^{C_j} = \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} (1 + o(1)). \quad (4.10)$$

Therefore we have

$$\begin{aligned} & \frac{1}{2\pi i} \int_{\gamma'} f(z) dz \\ &= \frac{e^{r_0} e^K}{2\pi i} \int_{\gamma''} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} (-m/w)^a (1 + o(1)) \left(1 + \frac{w}{m}\right)^{-(m+1)} \frac{dw}{m} \\ &= \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} m^{a-1} \frac{e^{r_0} e^K}{2\pi i} \int_{\gamma''} \left(-\frac{1}{w}\right)^a e^{-w} (1 + o(1)) dw, \end{aligned}$$

where the contour γ'' is defined in (4.6).

As in the proof of Theorem 4.2.1, we use the Gamma function to obtain

$$\frac{1}{2\pi i} \int_{\gamma'} f(z) dz \sim \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} e^{K+r_0} \frac{m^{a-1}}{\Gamma(a)}. \quad (4.11)$$

Next, we estimate the integral on contour γ_3 . If it is negligible compared with (4.11), our proof is done. We have $z = (1 + (3a+1)\log m/m)e^{it}$ where $\theta < t < 2\pi - \theta$. Similarly to (4.7) and (4.8), we get

$$\exp(C(\rho z)) = O((m/\log m)^a),$$

and

$$|z^{-(m+1)}| = O\left(\frac{1}{m^{3a+1}}\right).$$

Using an argument similar to the one used to derive (4.10), we obtain

$$\left| \prod_{j \in J^{(1)}} (1 - (\rho z)^j)^{C_j} \right| = O\left(\prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \right).$$

Hence, we have

$$\begin{aligned}
\left| \prod_{j \in J^{(1)}} (1 - (\rho z)^j)^{C_j} \right| &= O \left(\prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \left(\frac{1 + \rho^j}{1 - \rho^j} \right)^{C_j} \right) \\
&= O \left(\prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \exp(C_j(2\rho^j + O(\rho^{2j}))) \right) \\
&= O \left(\prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \exp(3a/j) \right) \\
&= O \left(\exp(3a \log |J^{(1)}|) \prod_{j \in J} (1 - \rho^j)^{C_j} \right) \\
&= O \left(\left(\frac{m}{\log m} \right)^{3a} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \right).
\end{aligned}$$

Therefore we have

$$\frac{1}{2\pi i} \int_{\gamma_3} f(z) dz = O \left(\frac{m^{a-1}}{(\log m)^{4a}} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \right),$$

which is negligible compared with (4.11).

Finally, we apply (4.11) in (4.9). This completes the proof of Theorem 4.2.2. $\square$

As basic theorems of this chapter, the required conditions in Theorems 4.2.1 and 4.2.2 will be satisfied by all later theorems. If we can weaken the required conditions, our result can be applied under more general conditions. In the Appendix, we extend this discussion about the required conditions in our theorems.

4.3 Distribution of the r th Smallest Component

In this section, we discuss the distribution of r th smallest component of a random decomposable combinatorial structure with a given restricted pattern. If we apply the generating functions obtained in Section 3.2, it is difficult to estimate the distribution since the characteristic function depends on the pattern. That is why we apply the results in Section 4.2. Hence, theorems in this section inherit the conditions of theorems in Section 4.2.

As defined in Section 3.2, $X_n^{[r]}(S)$ is the size of the r th smallest component of a random decomposable combinatorial structure of size n having a restricted pattern S . We derive an asymptotic expression for $P(X_n^{[r]}(S) > k)$. For convenience, we define the number set $N_k = \{1, 2, \dots, k\}$ where k is a positive integer. Let $d(k) = \sum_{j \in J \cap N_k} S(j)$ be the total number of components in the pattern which have sizes smaller than or equal to k . We observe that $P(X_n^{[r]}(S) > k) = 0$ if $d(k) \geq r$.

As usual, we divide this section into two subsections: Subsection 4.3.1 presents the theorem and the corollaries for the labeled case, and Subsection 4.3.2 provides the analogous results for the unlabeled case.

4.3.1 Labeled Case

In the following, we adopt the convention $0^0 = 1$ and use the notations $J^{(1)}$ and $J^{(2)}$ which are defined in Theorems 4.2.1 and 4.2.2. The probability that the size of the r th smallest component of a random decomposable combinatorial structure of size n having a restricted pattern is bigger than constant k is given by the following theorem.

Theorem 4.3.1. *Let S be a restricted pattern with $|J^{(1)}| = o(m/\log m)$. Assume $j = O(m/\log m)$ when $j \in J^{(1)}$, $r - d(k) = O(\log m)$, $k = o(m/\log m)$ and $m \rightarrow \infty$.*

We have

$$P(X_n^{[r]}(S) > k) \sim P_L(S, n) \exp(-C_L(\rho; N_k \setminus J)) \sum_{j=0}^{r-1-d(k)} \frac{(C_L(\rho; N_k \setminus J))^j}{j!}.$$

If $k \rightarrow \infty$ (as $n \rightarrow \infty$), letting $G(k) = a \log(k) + a\gamma + K - C(\rho; J \cap N_k)$, we have

$$P(X_n^{[r]}(S) > k) \sim P_L(S, n) \frac{\exp(C_L(\rho; J \cap N_k))}{k^a \exp(a\gamma + K)} \sum_{j=0}^{r-1-d(k)} \frac{(G(k))^j}{j!}.$$

Proof:

To keep track of the size of the r th smallest component, it is convenient to consider augmented patterns $S^*: J^* \mapsto N$ where $J^* = J \cup N_k$, $S^*(j) = S(j)$ for $j \in J$, and, for $j \in N_k \setminus J$, $S^*(j)$ is specified below. We note that each structure with the restricted pattern S such that its r th smallest component has size greater than k corresponds to structures with an augmented pattern S^* such that $\sum_{i \in N_k \setminus J} S^*(i) \leq r - 1 - d(k)$.

Hence we have

$$P(X_n^{[r]}(S) > k) = \sum_{\sum_{i \in N_k \setminus J} S^*(i) \leq r-1-d(k)} P_L(S^*).$$

We observe that, if we divide J^* into two parts $J^{*(1)}$ and $J^{*(2)}$ as we did to J , we have $J^{*(1)} = J^{(1)} \cup N_k$ and $J^{*(2)} = J^{(2)}$ since $k = o(m/\log m)$.

Let $m^* = n - \sum_{j \in J \cup N_k} j S^*(j)$. Since $k = o(m/\log m)$ and $\sum_{j \in N_k \setminus J} S^*(j) \leq r - 1 - d(k) = O(\log m)$, we have $\frac{m^*}{n} \sim \frac{m}{n}$. Applying Theorem 4.2.1 and noticing that $\sum_{j \in J \cap N_k} j S(j)$ is given by the pattern S , we get

$$\begin{aligned} & P(X_n^{[r]}(S) > k) \\ & \sim \sum_{\sum_{j \in N_k \setminus J} S^*(j) \leq r-1-d(k)} \left(\frac{m^*}{n}\right)^{a-1} \prod_{i \in J \cup N_k} \frac{C_i^{S^*(i)}}{(i!)^{S^*(i)} S^*(i)!} \rho^{n-m^*} \exp(-C_L(\rho; J^{(1)} \cup N_k)) \end{aligned}$$

$$\begin{aligned}
& \sim \left(\frac{m}{n}\right)^{a-1} \mathcal{S}_L(S) \rho^{n-m} \exp(-C_L(\rho; J^{(1)}) - C_L(\rho; N_k \setminus J^{(1)})) \\
& \quad \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \left(\prod_{i \in N_k \setminus J} \frac{C_i^{S^*(i)} \rho^{i S^*(i)}}{(i!)^{S^*(i)} S^*(i)!} \right) \\
& \sim P_L(S, n) \exp(-C_L(\rho; N_k \setminus J^{(1)})) \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \left(\prod_{i \in N_k \setminus J} \frac{C_i^{S^*(i)} \rho^{i S^*(i)}}{(i!)^{S^*(i)} S^*(i)!} \right).
\end{aligned} \tag{4.12}$$

Since $k = o(m/\log m)$, we have

$$N_k \setminus J = N_k \setminus J^{(1)}.$$

It follows from (4.12) and the multinomial formula

$$\begin{aligned}
(Y_1 + Y_2 + \dots + Y_k)^j &= \sum_{a_1 + a_2 + \dots + a_k = j} \binom{j}{a_1, a_2, \dots, a_k} Y_1^{a_1} Y_2^{a_2} \dots Y_k^{a_k} \\
&= j! \sum_{a_1 + a_2 + \dots + a_k = j} \prod_{i=1}^k \frac{Y_i^{a_i}}{a_i!}
\end{aligned}$$

that

$$P(X_n^{[r]}(S) > k) \sim P_L(S, n) \exp(-C_L(\rho; N_k \setminus J)) \sum_{j=0}^{r-1-d(k)} \frac{(C_L(\rho; N_k \setminus J))^j}{j!}.$$

If $k \rightarrow \infty$, we have, using the transfer theorem again,

$$\begin{aligned}
C_L(\rho; N_k) &= \sum_{i=1}^k C_i \rho^i / i! = [z^k] \frac{1}{1-z} C(\rho z) \\
&= [z^k] \frac{a}{1-z} \log \frac{1}{1-z} + K + O(k^{-\alpha}) \\
&= a \log(k) + a\gamma + K + O(k^{-\alpha}),
\end{aligned} \tag{4.13}$$

where γ is the Euler-Mascheroni constant.

Let $G(k) = a \log(k) + a\gamma + K - C_L(\rho; J \cap N_k)$. Then

$$P(X_n^{[r]}(S) > k) \sim P_L(S, n) \frac{\exp(C_L(\rho; J \cap N_k))}{k^a \exp(a\gamma + K)} \sum_{j=0}^{r-1-d(k)} \frac{(G(k))^j}{j!}.$$

□

When there is no restricted pattern, let $X_n^{[r]}$ denote the size of the r th smallest component of a random labeled decomposable structure of size n . From Theorem 4.3.1, we get the following result. We observe that there is another method to prove this result by using singularity analysis obtained by Panario and Richmond [43].

Corollary 4.3.1. *Let $r = O(\log n)$, $k = o(n/\log n)$ and $n \rightarrow \infty$. We have*

$$P(X_n^{[r]} > k) \sim \exp(-C_L(\rho; N_k)) \sum_{j=0}^{r-1} \frac{C_L(\rho; N_k)^j}{j!}.$$

If $k \rightarrow \infty$, we have

$$P(X_n^{[r]} > k) \sim k^{-a} \exp(-a\gamma - K) \left(\sum_{j=0}^{r-1} \frac{(a \log k)^j}{j!} \right).$$

The expression in Theorem 4.3.1 can be simplified when $r - d$ and $|J \cap N_k|$ are small compared with k , as the following corollary shows.

Corollary 4.3.2. *Using the same notations and conditions as those in Theorem 4.3.1, if we further limit $\log |J \cap N_k| = o(\log k)$ and $r - d(k) = o(\log k)$ as $k \rightarrow \infty$, we have, as $m \rightarrow \infty$,*

$$P(X_n^{[r]}(S) > k) \sim P_L(S, n) \frac{\exp(C_L(\rho; J \cap N_k))}{k^a \exp(a\gamma + K)} \frac{(a \log(k))^{r-1-d(k)}}{(r-1-d(k))!}.$$

Proof: Since $\log |J \cap N_k| = o(\log k)$ and $C_L(\rho; J \cap N_k) = O(C_L(\rho; N_{|J \cap N_k|}))$, applying (4.13), we get

$$C_L(\rho; J \cap N_k) = o(\log k).$$

Hence, $a \log(k) + a\gamma + K - C_L(\rho; J \cap N_k) = a \log k(1 + o(1))$.

From $r - d(k) = o(\log k)$, we have

$$\sum_{j=0}^{r-1-d(k)} \frac{(C_L(\rho; N_k \setminus J))^j}{j!} \sim \frac{(a \log k)^{r-1-d(k)}}{(r-1-d(k))!}.$$

Now the corollary follows immediately from Theorem 4.3.1. $\square$

4.3.2 Unlabeled Case

In the unlabeled case, we have a parallel theorem to the one in the labeled case. However, observe that we require the singularity point $0 < \rho < 1$ which is not a condition in labeled case. The reason has been given in Section 4.1.

Define

$$U(z; J) = \prod_{i \in J} (1 - z^i)^{C_i}.$$

Theorem 4.3.2 provides the probability that the size of the r th smallest component of a random decomposable combinatorial structure of size n has a restricted pattern bigger than a constant k .

Theorem 4.3.2. *Let S be a restricted pattern with $|J^{(1)}| = o(m/\log m)$. Let us assume that $j = O(m/\log m)$ when $j \in J^{(1)}$, $r = O(\log m)$ and $k = o(m/\log m)$. As $m \rightarrow \infty$, we have,*

$$P(X_n^{[r]}(S) > k) \sim P_U(S, n) U(\rho; N_k \setminus J) \sum_{j=0}^{r-1-d(k)} [z^j] \exp \left(z C_U(\rho; N_k \setminus J) + z^2 \frac{C_U(\rho^2; N_k \setminus J)}{2} + \dots \right).$$

If $k \rightarrow \infty$, we have

$$P(X_n^{[r]}(S) > k) \sim P_U(S, n) \frac{e^{-a\gamma - K - r_0}}{k^a U(\rho; J \cap N_k)} \sum_{j=0}^{r-1-d(k)} [z^j] \exp \left(z C_U(\rho; N_k \setminus J) + z^2 \frac{C_U(\rho^2; N_k \setminus J)}{2} + \dots \right).$$

Proof:

We define the same augmented pattern S^* as the one in the proof of Theorem 4.3.1. Let $m^* = n - \sum_{j \in J \cup N_k} j S^*(j)$. We still have $\frac{m^*}{n} \sim \frac{m}{n}$ for the same reason as in the labeled case. Applying Theorem 4.2.2, we get

$$\begin{aligned}
& P(X_n^{[r]}(S) > k) \\
& \sim \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \left(\frac{m^*}{n}\right)^{a-1} \prod_{j \in J \cup N_k} \binom{C_j + S^*(j) - 1}{S^*(j)} \prod_{j \in J^{(1)} \cup N_k} (1 - \rho^j)^{C_j} \rho^{n-m^*} \\
& \sim P_U(S, n) U(\rho; N_k \setminus J) \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \prod_{j \in N_k \setminus J} \binom{C_j + S^*(j) - 1}{S^*(j)} \rho^{j S^*(j)}, \quad (4.14)
\end{aligned}$$

where we apply $N_k \setminus J = N_k \setminus J^{(1)}$ in the last line. The negative binomial theorem states that

$$(1 - \rho^i z)^{-C_i} = \sum_{X_i \geq 0} \binom{C_i + X_i - 1}{X_i} \rho^{i X_i} z^{X_i}.$$

Hence, we get

$$\begin{aligned}
& \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \prod_{j \in N_k \setminus J} \binom{C_j + S^*(j) - 1}{S^*(j)} \rho^{j S^*(j)} = \sum_{u=0}^{r-1-d(k)} [z^u] \prod_{j \in N_k \setminus J} (1 - \rho^j z)^{-C_j} \\
& = \sum_{u=0}^{r-1-d(k)} [z^u] \exp \left(- \sum_{j \in N_k \setminus J} C_j \log(1 - \rho^j z) \right) \\
& = \sum_{u=0}^{r-1-d(k)} [z^u] \exp \left(\sum_{i \geq 1} \sum_{j \in N_k \setminus J} C_j \rho^{ij} z^i / i \right) \\
& = \sum_{u=0}^{r-1-d(k)} [z^u] \exp \left(z(C_U(\rho; N_k) - C_U(\rho; J \cap N_k)) \right. \\
& \quad \left. + z^2 \frac{C_U(\rho^2; N_k) - C_U(\rho^2; J \cap N_k)}{2} + \dots \right). \quad (4.15)
\end{aligned}$$

When $k \rightarrow \infty$, by (4.13), we have

$$C_U(\rho; N_k) = a \log(k) + a\gamma + K + O(k^{-\alpha}).$$

Since ρ is the radius of convergence of $C_U(z)$, and in the unlabeled case $0 < \rho < 1$, $C_U(\rho^t)$ is bounded for an integer $t \geq 2$. We have that $C_U(\rho^t; N_k) \sim C_U(\rho^t)$ as $k \rightarrow \infty$.

Similarly to the derivation of (4.13) and the above discussion, we get

$$\begin{aligned} U(\rho; N_k) &= \exp \left(\sum_{i=1}^k C_i \log(1 - \rho^i) \right) \sim \exp \left(- \sum_{i=1}^k C_i \rho^i \right) e^{-r_0} \\ &\sim k^{-a} \exp(-a\gamma - K - r_0), \end{aligned} \quad (4.16)$$

where $r_0 = \frac{C_U(\rho^2)}{2} + \frac{C_U(\rho^3)}{3} + \dots$. Finally, in Equation (4.14), we do the following substitution

$$U(\rho; N_k \setminus J) = \frac{U(\rho; N_k)}{U(\rho; J \cap N_k)},$$

and the proof is complete. $\square$

As in the labeled case, we provide the following two corollaries. The first one is for the special case $J = \emptyset$.

Corollary 4.3.3. *Let $r = o(\log n)$ and $k = o(n/\log n)$. Using the same notations as in Theorem 4.3.2, we have*

$$P(X_n^{[r]} > k) \sim U(\rho; N_k) \sum_{j=0}^{r-1-d} [z^j] \exp \left(z C_U(\rho; N_k) + z^2 \frac{C_U(\rho^2; N_k)}{2} + \dots \right).$$

If $k \rightarrow \infty$, we have

$$P(X_n^{[r]} > k) \sim \frac{e^{-a\gamma - K - r_0}}{k^a} \sum_{j=0}^{r-1-d} [z^j] \exp \left((a \log k) z + z^2 \frac{C_U(\rho^2)}{2} + \dots \right).$$

Corollary 4.3.4. *Assume the same notations and conditions as in Theorem 4.3.2. Assume further $\log |J \cap N_k| = o(\log k)$ and $r - d = o(\log k)$ as $k \rightarrow \infty$. We then have*

$$P(X_n^{[r]}(S) > k) \sim P_U(S) \frac{(a \log k)^{r-1-d}}{(r-1-d)! k^a e^{a\gamma + K + r_0} U(\rho; J \cap N_k)}.$$

Proof:

Since $\log |J \cap N_k| = o(\log k)$ and $C(\rho; J \cap N_k) = O(C(\rho; N_{J \cap N_k}))$, applying (4.13), we have

$$C(\rho; J \cap N_k) = o(\log k).$$

Hence,

$$C(\rho; N_k \setminus J) = a \log k (1 + o(1)).$$

From (4.15), noticing $r - d = o(\log k)$, we have

$$\begin{aligned} \sum_{j=0}^{r-1-d(k)} [z^j] \exp \left(z(C(\rho; N_k \setminus J)) + z^2 \frac{C(\rho^2; N_k \setminus J)}{2} + \dots \right) &\sim \sum_{j=0}^{r-1-d} [z^j] \exp(az \log k) \\ &\sim \frac{(a \log k)^{r-1-d}}{(r-1-d)!}. \end{aligned}$$

The corollary is proved using these simplifications in the result of Theorem 4.3.2. $\square$

4.4 Moments

In this section, given a restricted pattern $S : J \mapsto \mathbb{N}$, we consider the asymptotic behavior of the average size of the r th smallest component. We only discuss the case when $J^{(2)} = \emptyset$ where $J^{(2)}$, defined in Section 4.2, is the set of the sizes of large components in the pattern S .

Let $E(X_n^{[r]}(S))$ stand for the first moment of $X_n^{[r]}(S)$. To estimate the first moment, we use the well-known expression

$$E(X_n^{[r]}(S)) = \sum_{k \geq 0} P(X_n^{[r]}(S) > k).$$

Let $i^{[j]}$ denote the size of the j th component in the given restricted pattern. If $r \leq \sum_{j \in J} S(j)$, it is clear that $P(X_n^{[r]}(S) > k) = 0$ when $\sum_{j \in J \cap N_k} S(j) \geq r$. Hence we have

$$E(X_n^{[r]}(S)) = \sum_{k=0}^{i^{[r]}-1} P(X_n^{[r]}(S) > k).$$

If $r > \sum_{j \in J} S(j)$, we need to estimate $P(X_n^{[r]}(S) > k)$ when k exceeds the range given in Theorem 4.3.1.

Let $\lfloor x \rfloor$ be the biggest integer which is smaller than or equal to x . We shall estimate $E(X_n^{[r]}(S))$ under either of the following two conditions

(C1) $a > 1$, $r = o(\frac{\log n}{\log \log n})$ and $\log |J| = o(\log n)$,

(C2) $a = 1$, $r = O(1)$ and $|J| = O(\log n)$.

To prove later theorems, we introduce the following lemma.

Lemma 4.4.1. *Let a , b , and c be positive integers. Suppose $a \rightarrow \infty$, $\frac{\log b}{\log a} \rightarrow \infty$, and $c = O(1)$. Then,*

$$\sum_{k=a}^b \frac{(\log k)^c}{k} \sim \frac{1}{c+1} (\log b)^{c+1}.$$

Proof: Let us consider the continuous real function $f(x) = \frac{1}{x}(\log x)^c$ where $x \in [a, b]$.

We have

$$f'(x) = \frac{1}{x^2}(c - \log x)(\log x)^{c-1} < 0,$$

since c is a constant and $a \rightarrow \infty$. Hence, $f(x)$ is monotonic decreasing and we have

$$\int_a^b f(k) dk < \sum_{k=a}^b \frac{(\log k)^c}{k} < \int_a^b f(k-1) dk,$$

and the proof is done since

$$\int_a^b f(k)dk \sim \int_a^b f(k-1)dk \sim \frac{1}{c+1}(\log b)^{c+1}.$$

□

For a random decomposable combinatorial structure with a given restricted pattern, the first moment of the r th smallest component is given by the following theorem.

Theorem 4.4.1. *Let $M = \max\{j : j \in J\}$. Assume $r > \sum_{j \in J} S(j)$ and $M = o(n/\log n)$. Define $w = \max\left\{\left\lfloor \frac{n}{\log n \log \log n} \right\rfloor, M\right\}$. Under either (C1) or (C2) conditions, and as $n \rightarrow \infty$, for both labeled and unlabeled cases, we have*

$$E(X_n^{[r]}(S)) \sim \sum_{k=0}^w P(X_n^{[r]}(S) > k).$$

Proof:

We only prove this theorem for the labeled case since the proof of the unlabeled case is very similar. We observe that $m \sim n$.

We start with the condition (C1), that is, $a > 1$, $r = o(\log n / \log \log n)$ and $\log |J| = o(\log n)$. From the definition, we have $P(X_n^{[r]}(S) > k) > 0$, for any positive k . According to Theorem 4.3.1, we have

$$\sum_{k=0}^w P(X_n^{[r]}(S) > k) > P(X_n^{[r]}(S) > 0) = P_L(S, n).$$

Let $\hat{d} = \sum_{j \in J} S(j)$. From Corollary 4.3.2, we have

$$\begin{aligned} & \sum_{k=w+1}^n P(X_n^{[r]}(S) > k) \\ & \leq nP(X_n^{[r]}(S) > w) \end{aligned}$$

$$\begin{aligned}
& \sim nP_L(S, n)O\left(\frac{\exp(C_L(\rho; J \cap N_w))(a \log w)^{r-1-\hat{d}}}{(r-1-\hat{d})!w^a}\right) \\
& \leq nP_L(S, n)O\left(\frac{\exp(C_L(\rho; J \cap N_w))a^{r-1-\hat{d}}(\log \log n)^a(\log n)^{r+a-1-\hat{d}}}{(r-1-\hat{d})!n^a}\right).
\end{aligned}$$

In this case, the constant a is bigger than 1. According to condition (C1), we have

$$\exp(C_L(\rho; J \cap N_w)) = O(\exp(C_L(\rho; N_{|J|}))) = O(\exp(o(\log n))) = O(n^{o(1)}),$$

and

$$(\log n)^{r+a-1-\hat{d}} = (\log n)^{o(\frac{\log n}{\log \log n})} = n^{o(1)}.$$

Hence,

$$\sum_{k=w+1}^n P(X_n^{[r]}(S) > k) = P_L(S, n)o(1),$$

which is negligible compared with $\sum_{k=0}^w P(X_n^{[r]}(S) > k)$.

We consider now condition (C2), that is, $a = 1$, $r = O(1)$ and $|J| = O(\log n)$. We divide this proof into three parts. In the first part, we prove that

$$\begin{aligned}
& \sum_{k=0}^w P(X_n^{[r]}(S) > k) \\
& > \exp(-\gamma - K) \frac{P_L(S, n) \exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor}))}{2(r-1)!(r-\hat{d})} (\log n)^{r-\hat{d}}. \quad (4.17)
\end{aligned}$$

In the second part, we prove

$$\begin{aligned}
& \sum_{k=w}^{\lfloor n/\log n \rfloor} P(X_n^{[r]}(S) > k) \\
& = o(P_L(S, n) \exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor})) (\log n)^{r-\hat{d}}). \quad (4.18)
\end{aligned}$$

In the third part, we explain that $\sum_{k=n/\log n}^n P(X_n^{[r]}(S) > k)$ is negligible compared with $P_L(S, n) \exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor})) (\log n)^{r-\hat{d}}$.

Before we start the proof, let us first estimate $C_L(\rho; J \cap N_k)$ when $k > (\log n)^{\log \log n}$.

Under the condition (C2), $|J| = O(\log n)$. Then we have

$$\begin{aligned}
 C_L(\rho; J \cap N_k) &= C_L(\rho; J \cap N_{(\log n)^{\log \log n}}) + C_L(\rho; J \cap (N_k \setminus N_{(\log n)^{\log \log n}})) \\
 &\leq C_L(\rho; J \cap N_{(\log n)^{\log \log n}}) + C_L(\rho; (N_{(\log n)^{\log \log n} + |J|} \setminus N_{(\log n)^{\log \log n}})) \\
 &= C_L(\rho; J \cap N_{(\log n)^{\log \log n}}) + o(1).
 \end{aligned} \tag{4.19}$$

Part I: proof of (4.17).

Applying Lemma 4.4.1 and Corollary 4.3.2, we get

$$\begin{aligned}
 &\sum_{k=0}^w P(X_n^{[r]}(S) > k) \\
 &> \sum_{k=\lfloor (\log n)^{\log \log n} \rfloor}^w P(X_n^{[r]}(S) > k) \\
 &> \frac{\exp(-\gamma - K)}{2(r-1)!} \exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor})) P_L(S, n) \sum_{k=\lfloor (\log n)^{\log \log n} \rfloor}^w \frac{(\log k)^{r-1-\hat{d}}}{k} \\
 &\sim \exp(-\gamma - K) \frac{P_L(S, n) \exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor}))}{2(r-1)!(r-\hat{d})} (\log n)^{r-\hat{d}}.
 \end{aligned} \tag{4.20}$$

Part II: proof of (4.18).

Applying $r = O(1)$ and Corollary 4.3.2, we obtain

$$\begin{aligned}
 &\sum_{k=w+1}^{\lfloor \frac{n}{\log n} \rfloor} P(X_n^{[r]}(S) > k) \\
 &< \frac{n}{\log n} P(X_n^{[r]}(S) > w) \\
 &\leq \frac{n}{\log n} P_L(S, n) O\left(\frac{\exp(C_L(\rho; J \cap N_w)) \log \log n (\log n)^{r-\hat{d}}}{n}\right).
 \end{aligned} \tag{4.21}$$

Now applying (4.19), we get

$$C(\rho; J \cap (N_w \setminus N_{\lfloor (\log n)^{\log \log n} \rfloor})) = o(1).$$

Putting this back to (4.21), we have

$$\begin{aligned} & \sum_{k=w+1}^{\lfloor \frac{n}{\log n} \rfloor} P(X_n^{[r]}(S) > k) \\ &= P_L(S, n) o(\exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor})) (\log n)^{r-\hat{d}}). \end{aligned} \quad (4.22)$$

Part III: we prove that $\sum_{k=n/\log n}^n P(X_n^{[r]}(S) > k)$ is negligible.

Let $\varepsilon = \frac{1}{\log \log \log n}$ and $w' = \frac{\sqrt{\varepsilon n}}{\log n}$. We have

$$\varepsilon n \frac{(\log w')^{r-\hat{d}-1}}{w'} = o((\log n)^{r-\hat{d}}).$$

Similar to (4.22), we get

$$\begin{aligned} \sum_{k=w'}^{\lfloor \varepsilon n \rfloor} P(X_n^{[r]}(S) > k) &= O\left(\varepsilon n P_L(S, n) \exp(C_L(\rho; J \cap N_{w'})) \frac{(\log w')^{r-\hat{d}-1}}{w'}\right) \\ &= P_L(S, n) o(C_L(\rho; N_{\lfloor (\log n)^{\log \log n} \rfloor})) (\log n)^{r-\hat{d}}. \end{aligned} \quad (4.23)$$

In order to estimate $P(X_n^{[r]}(S) > k)$ when $k = \lfloor \varepsilon n \rfloor$, we have to go back to the proof of Theorem 4.3.1. Let $\mathbb{S}$ be the set of augmented patterns of S as defined in the proof of Theorem 4.3.1. For any $S^* \in \mathbb{S}$, the corresponding J^* satisfies $J^* = N_{\lfloor \varepsilon n \rfloor}$ and $0 \leq \sum_{j \in J^* \setminus J} S^*(j) \leq r - \hat{d} - 1$. We need to estimate $P_L(S^*, n)$ for every $S^* \in \mathbb{S}$. We notice that $J \subset N_{\lfloor \varepsilon n \rfloor}$. We base our analysis on (C2). Given $a = 1$, we get the expression of $P_L(S^*, n)$

$$\begin{aligned} P_L(S^*, n) &\sim \mathcal{J}_L(S) \rho^{n-m^*} e^{-K} \prod_{j \in N_{\lfloor \varepsilon n \rfloor} \setminus J} \frac{C_j^{S^*(j)}}{(j!)^{S^*(j)} S^*(j)!} [z^{m^*}] \exp\left(\sum_{j=\lfloor \varepsilon n \rfloor+1}^n \frac{z^j}{j}\right) \\ &\sim \mathcal{J}_L(S) \rho^{n-m} e^{-K} \prod_{j \in N_{\lfloor \varepsilon n \rfloor} \setminus J} \frac{(C_j \rho^j)^{S^*(j)}}{(j!)^{S^*(j)} S^*(j)!} [z^{m^*}] \prod_{j=\lfloor \varepsilon n \rfloor+1}^n \left(\sum_{i \geq 0} \frac{z^{ij}}{i! j^i}\right). \end{aligned}$$

Recall that $m = n - \sum_{j \in J} j S(j)$ for a given pattern S . For any $S^* \in \mathbb{S}$, we have $n/m^* \rightarrow 1$ since $r > \hat{d}$ and $r = O(1)$. Let $S' : J' \mapsto N$ be an augmented pattern

belonging to $\mathbb{S}$ and the corresponding $m' = n - \sum_{j \in J'} j S'(j)$ such that, for any $S^* \in \mathbb{S}$, we have

$$[z^{m^*}] \prod_{j=\lfloor \varepsilon n \rfloor + 1}^n \left(\sum_{i \geq 0} \frac{z^{ij}}{i! j^i} \right) \leq [z^{m'}] \prod_{j=\lfloor \varepsilon n \rfloor + 1}^n \left(\sum_{i \geq 0} \frac{z^{ij}}{i! j^i} \right).$$

Following a similar reasoning as the one we used in the proof of Theorem 4.2.1 to derive $P(X_n^{[r]}(S) > k)$, we obtain

$$\begin{aligned} & P(X_n^{[r]}(S) > \lfloor \varepsilon n \rfloor) \\ &= \sum_{\substack{r-1-\hat{d} \\ \sum_{i \in N_{\lfloor \varepsilon n \rfloor} \setminus J} S^*(i)=0}} P_L(S^*, n) \\ &\leq \mathcal{S}_L(S) \rho^{n-m} e^{-K} [z^{m'}] \prod_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} \left(\sum_{i \geq 0} \frac{z^{ij}}{i! j^i} \right) \sum_{\substack{r-1-\hat{d} \\ \sum_{i \in N_{\lfloor \varepsilon n \rfloor} \setminus J} S^*(i)=0}} \prod_{j \in N_{\lfloor \varepsilon n \rfloor} \setminus J} \frac{(C_j \rho^j)^{S^*(j)}}{(j!)^{S^*(j)} S^*(j)!}. \end{aligned}$$

Again, using the same method as in the proof of Theorem 4.3.1, we have

$$\begin{aligned} \sum_{\substack{r-1-\hat{d} \\ \sum_{i \in N_{\lfloor \varepsilon n \rfloor} \setminus J} S^*(i)=0}} \prod_{j \in N_{\lfloor \varepsilon n \rfloor} \setminus J} \frac{(C_j \rho^j)^{S^*(j)}}{(j!)^{S^*(j)} S^*(j)!} &= \sum_{i=0}^{r-1-\hat{d}} \frac{(C(\rho; N_{\lfloor \varepsilon n \rfloor} \setminus J))^i}{i!} \\ &= O((\log n)^{r-1-\hat{d}}). \end{aligned}$$

Hence, we obtain

$$\begin{aligned} P(X_n^{[r]}(S) > \lfloor \varepsilon n \rfloor) &= O \left(\mathcal{S}_L(S) \rho^{n-m} (\log n)^{r-1-\hat{d}} [z^{m'}] \prod_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} \left(\sum_{i \geq 0} \frac{z^{ij}}{i! j^i} \right) \right) \\ &= O \left(\mathcal{S}_L(S) \rho^{n-m} (\log n)^{r-1-\hat{d}} \sum_{\substack{m' \\ \sum_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} i_j j = m'}} \prod_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} \frac{1}{i_j! j^{i_j}} \right) \quad (4.24) \\ &= O \left(\mathcal{S}_L(S) \rho^{n-m} (\log n)^{r-1-\hat{d}} \sum_{h=1}^{\lfloor 1/\varepsilon \rfloor} \sum_{\substack{m' \\ \sum_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} i_j = h}} \prod_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} \frac{1}{i_j! j^{i_j}} \right). \end{aligned}$$

Since $\varepsilon = 1/\log \log \log n$, we have

$$\left(\frac{1}{\varepsilon}\right)^{\lfloor \frac{1}{\varepsilon} \rfloor + 1} = o(\log n). \quad (4.25)$$

If $\sum_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} i_j = h$,

$$\prod_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} \frac{1}{i_j! j^{i_j}} = O\left(\frac{1}{(\varepsilon n)^h}\right). \quad (4.26)$$

For integer $0 < k \leq n$, let $p_n(k)$ denote the number of solutions, in positive integers, to $x_1 + x_2 + \cdots + x_k = n$. It is well-known that $p_n(k) = \binom{n-1}{k-1} = O(n^{k-1})$.

Applying (4.26) in (4.24), since $p_{m'}(h)$ is the upper bound of

$$\sum_{\substack{\sum_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} i_j = h \\ \sum_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} i_j j = m'}} 1,$$

we obtain

$$\sum_{\substack{\sum_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} i_j = h \\ \sum_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} i_j j = m'}} \prod_{j=\lfloor \varepsilon n \rfloor + 1}^{m'} \frac{1}{i_j! j^{i_j}} \leq O\left(\frac{1}{(\varepsilon n)^h}\right) p_{m'}(h) = O\left(\frac{1}{(\varepsilon n)^h} n^{h-1}\right).$$

In the last line, we use n^{h-1} instead of $(m')^{h-1}$ since $m' \leq n$. Hence, applying (4.25), we have

$$\begin{aligned} P(X_n^{[r]}(S) > \lfloor \varepsilon n \rfloor) &= O\left(\mathcal{S}_L(S) \rho^{n-m} (\log n)^{r-1-d} \sum_{h=1}^{\lfloor \frac{1}{\varepsilon} \rfloor} O\left(\frac{1}{(\varepsilon n)^h} n^{h-1}\right)\right) \\ &= o\left(\mathcal{S}_L(S) \rho^{n-m} \frac{(\log n)^{r-d}}{n}\right). \end{aligned}$$

From Theorem 4.2.1, we know $P_L(S, n) \sim \left(\frac{m}{n}\right)^{a-1} \mathcal{S}_L(S) \rho^{n-m} \exp(-C_L(\rho; J))$. Using (4.19), we get,

$$C_L(\rho; J \cap (N_n \setminus N_{\lfloor (\log n) \log \log n \rfloor})) = o(1),$$

namely

$$P_L(S, n) \exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor})) \sim \mathcal{S}_L(S) \rho^{n-m}.$$

We have

$$\begin{aligned} & \sum_{k=\lfloor \varepsilon n \rfloor}^{n-1} P(X_n^{[r]}(S) > k) \\ & \leq n P(X_n^{[r]}(S) > \lfloor \varepsilon n \rfloor) \\ & = o \left(P_L(S, n) \exp(C_L(\rho; J \cap N_{\lfloor (\log n)^{\log \log n} \rfloor})) (\log n)^{r-d} \right). \end{aligned} \quad (4.27)$$

This part is completed.

Finally, combining (4.27) with (4.20), (4.22) and (4.23), we finish our proof.

□

For some special restricted patterns, the above result can be simplified using Theorems 4.2.1 and 4.2.2.

Corollary 4.4.1. *Suppose $r = O(1)$, $a = 1$ and $J = \emptyset$. Then, for the labeled case,*

$$E(X_n^{[r]}) \sim \frac{1}{r! \exp(\gamma + K)} (\log n)^r,$$

and for the unlabeled case,

$$E(X_n^{[r]}) \sim \frac{1}{r! \exp(\gamma + K + r_0)} (\log n)^r.$$

Proof: We can easily prove this corollary by applying Theorem 4.4.1, Corollary 4.3.2 and Corollary 4.3.4. □

For higher moments, we can get the i th moment when $2 \leq i < a$ by applying the approach in the proof of Theorem 4.4.1 under condition (C1). The results are given by the following corollary:

Corollary 4.4.2. *Let $\max\{j : j \in J\} = o(n/\log n)$ and $w = \left\lfloor \frac{n}{\log n \log \log n} \right\rfloor$. Assume $2 \leq i < a$. As $m \rightarrow \infty$, under conditions (C1), we have, for both labeled and unlabeled cases,*

$$E^{(i)}(X_n^{[r]}(S)) \sim \sum_{k=1}^w ((k+1)^i - k^i) P(X_n^{[r]}(S) > k).$$

4.5 Examples.

In this section, we exemplify our results for several decomposable structures. The first two examples are for the labeled case and the last one is for the unlabeled case. We apply our theorems to estimate the asymptotic probability that a random decomposable structure of size n has a restricted pattern. Moreover, the case where the smallest component size is bigger than or equal to any integer v with $v = o(n/\log n)$ follows from Theorems 4.3.1 and 4.3.2 immediately. These theorems apply to many other combinatorial structures, for example, to unlabeled structures such as random mappings patterns, arithmetical semigroups, and to labeled structures such as functional digraphs, children's yards [43].

Example 4.5.1. (*Cycles in Permutations*). *As an example of labeled decomposable combinatorial structure, let us consider a random permutation. The generating function for the cycles of permutations is*

$$C_L(z) = \log \frac{1}{1-z}.$$

Thus, the cycles are of logarithmic type. We have that $a = 1$ and the singular point is 1. Let us say, for example, that we want the probability that a random permutation has restricted pattern $S(2) = 2$ and $S(4) = 1$, and the length of its 10th smallest

cycle is bigger than $\lfloor \log n \rfloor$, as $n \rightarrow \infty$. Applying Corollary 4.3.2, we have

$$P(X_n^{[10]}(S) > \lfloor \log n \rfloor) \sim \frac{(\log \log n)^6}{23040e^\gamma \log n}.$$

Applying now Theorem 4.4.1 and Corollary 4.3.2, we have

$$E_L(X_n^{[10]}(S)) \sim \frac{e^{-\gamma}}{161280} (\log n)^7.$$

In the non-restricted pattern case, applying Shepp and Lloyd's result (see [48]), we obtain that the expectation of the size of the 10th cycle is $e^{-\gamma}(\log n)^{10}$. Compared with our previous result, the $(\log)^3$ disappearing makes sense because we fixed 3 of the smallest cycles.

Example 4.5.2. (2-Regular Graphs). Another familiar labeled structure is 2-regular graph. A 2-regular graph, whose vertices all have degree 2, consist of cycles. The size of a 2-regular graph is the number of vertices. The generating function of 2-regular graphs is (see [52])

$$L(z) = \exp \left(\frac{1}{2} \log \frac{1}{1-z} - \frac{z}{2} - \frac{z^2}{4} \right).$$

Thus, $L(z)$ is in the exp-log class with $a = \frac{1}{2}$. Given a random 2 regular graph of size n , let us say we want to estimate the probability that there is no cycle of length smaller than $\lfloor \log n \rfloor$ and the size of the $\lfloor \log \log n \rfloor$ th smallest cycle be bigger than $(\log n)^{\log \log n}$. According to Corollary 4.3.2, we get

$$P(X_n^{\lfloor \log \log n \rfloor}(S) > (\log n)^{\log \log n}) \sim \frac{e^{0.75-0.5\gamma} (\log \log n)^{2\lfloor \log \log n \rfloor - 2}}{2^{\lfloor \log \log n \rfloor - 1} (\log n)^{0.5 \log \log n} (\lfloor \log \log n \rfloor - 1)!}.$$

Example 4.5.3. (Irreducible Factors in Polynomials over Finite Fields). A typical unlabeled structure is the set of polynomials over finite fields. The generating function for polynomials over $\mathbb{F}_q$ is

$$U(z) = \exp \left(C_U(z) + \frac{C_U(z^2)}{2} + \cdots \right),$$

with

$$C_U(z) = \log \frac{1}{1 - qz} + F(z),$$

where $F(z)$ is an analytic function in $|z| < q^{-1/2}$. Of course, there is a more precise expression for $U(z)$ but this form is enough for our interests here; see [13].

Let us consider, for example, the restricted pattern $S(2) = 1$ and $S(\lfloor \log n \rfloor) = 1$. Then, what is the probability that the 10th smallest irreducible factor has degree bigger than, say, $\lfloor \sqrt{n} \rfloor$? Applying Corollary 4.3.4, we have for each fixed q

$$P(X_n^{[10]}(S) > \lfloor \sqrt{n} \rfloor) \sim \frac{\left(1 - \frac{1}{q}\right) (\log n)^6}{2^8 7! \exp(\gamma + F(1/q) + r_0) \sqrt{n}}.$$

Chapter 5

Decomposable Combinatorial Structures in Alg-Log Class

In this chapter, we study another class of decomposable combinatorial structures which is named alg-log class. Applying the generating functions obtained in Chapter 3, we estimate the probability that a decomposable structure of size n in alg-log class has a given restricted pattern. In addition, we consider the distribution of the r th smallest component. The first moment of the r th smallest component with given restricted pattern is discussed in Section 5.4. One example of labeled case is provided in Section 5.5.

5.1 Introduction

Like exp-log class, the algebraic-logarithmic type is another important class of decomposable combinatorial structures defined by Flajolet and Soria (see [21]).

Definition 5.1.1. A complex function $C(z)$ is said to be algebraic-logarithmic at $\rho > 0$ if

- $C(z)$ is analytic in $\Delta(\nu, \theta)$ as given in the Definition 2.3.1;
- as z tends to ρ in $\Delta(\nu, \theta)$

$$C(z) = c + \left(1 - \frac{z}{\rho}\right)^\alpha \left(\log \frac{1}{1 - z/\rho}\right)^\beta (d + o(1)), \quad (5.1)$$

with $c, d \neq 0$, α and β real numbers. The constant α is called the algebraic exponent, and β is called the logarithmic exponent of $C(z)$.

Many elementary combinatorial structures have generating functions of this type, for example simple families of trees [36], families of random mappings [1] and many classes of permutations defined by cycle constraints, etc. More interesting examples can be found in [19].

From the definition of alg-log type, we can see that the generating function of alg-log type has good analytic properties. A very natural idea is to apply singularity analysis to it.

In this chapter, we study the distribution of alg-log type structures with given restricted patterns. Furthermore, we estimate the distribution of the r th smallest component as we did in Chapter 4 for the exp-log class. A related work was done by Gourdon in [28]. An example is given in the last section.

Throughout this chapter, we assume the algebraic exponent $0 < \alpha < 1$ and inherit all notations from Chapter 4 without declaration.

5.2 Alg-Log Structures with Restricted Pattern

As usual, we divide this section into two subsections: labeled case and unlabeled case.

5.2.1 Labeled Case

In the labeled case, let the exponential generating function $L(z)$ satisfy $L(z) = \exp(C_L(z))$ [20]. Therefore, $L(z)$ is also of alg-log type since, by Taylor expansion, as $z \rightarrow \rho$

$$\begin{aligned} L(z) &= e^c \exp(C_L(z) - c) \\ &= e^c (1 + (C_L(z) - c)(1 + o(1))). \end{aligned}$$

Propositions 2.3.1 and 2.3.3 entails

$$[z^n]L(z) = \frac{e^c}{\rho^n \Gamma(-\alpha)} \frac{(\log n)^\beta}{n^{1+\alpha}} (d + o(1)), \quad (5.2)$$

$$\frac{C_n}{n!} = \frac{1}{\rho^n \Gamma(-\alpha)} \frac{(\log n)^\beta}{n^{1+\alpha}} (d + o(1)), \quad (5.3)$$

whenever $\alpha \notin \{0, 1, 2, \dots\}$.

The following lemma will be repeatedly applied later.

Lemma 5.2.1. *Let real numbers $0 < a < 1$ and $b = O(1)$. Then*

$$\sum_{j=1}^c \frac{(\log j)^b}{j^a} \sim \frac{1}{1-a} c^{1-a} (\log c)^b, \quad \text{as } c \rightarrow \infty.$$

Proof: It is easy to see that, as $x > 1$, $\frac{(\log x)^b}{x^a}$ achieves the maximum value at $x = e^{\frac{b}{a}}$.

Let $c' = \lfloor e^{\frac{b}{a}} \rfloor + 1$. We have

$$\sum_{j=1}^c \frac{(\log j)^b}{j^a} = \sum_{j=1}^{c'} \frac{(\log j)^b}{j^a} + \sum_{j=c'+1}^c \frac{(\log j)^b}{j^a},$$

and

$$\sum_{j=1}^{c'} \frac{(\log j)^b}{j^a} = O(1).$$

Let $f(x) = \frac{(\log x)^b}{x^a}$. Since $f(x)$ is monotonic decreasing as $x \geq c'$, applying the Riemann integral, we obtain

$$\int_{c'}^c f(x) dx \leq \sum_{j=c'}^c \frac{(\log j)^b}{j^a} \leq \int_{c'}^c f(x-1) dx.$$

Let $x = e^t$. We have

$$\begin{aligned} \int_{c'}^c f(x) dx &= \int_{\log c'}^{\log c} \frac{t^b}{e^{(a-1)t}} dt \\ &= \frac{1}{1-a} e^{(1-a)t} t^b \Big|_{\log c'}^{\log c} - \frac{b}{1-a} \int_{\log c'}^{\log c} \frac{t^{b-1}}{e^{(a-1)t}} dt \\ &= \frac{1}{1-a} c^{1-a} (\log c)^b + O(c^{1-a} (\log c)^{b-1}). \end{aligned}$$

We have the same result for $\int_{c'}^c f(x-1) dx$. Hence, the proof is complete. $\square$

Our first result in this chapter is given below. Let us recall that, given a structure with a restricted pattern $S : J \mapsto \mathbb{N}$, $J^{(1)} = \{j \in J : j \leq m\}$ and $J^{(2)} = \{j \in J : j > m\}$ where m is the size of the unrestricted part of the structure.

Theorem 5.2.1. *Let $P_L(S, n)$ be the probability that a random labeled decomposable combinatorial structure of size n , in the alg-log class with an algebraic exponent $0 < \alpha < 1$, has a restricted pattern S . For $\hat{j} = \max\{j; j \in J^{(1)}\}$, suppose $\hat{j} = O(m/\log m)$ and $\sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha} (\log m)^{\beta-1})$. We have, as $m \rightarrow \infty$,*

$$P_L(S, n) \sim \left(\frac{n}{m}\right)^{1+\alpha} \left(\frac{\log m}{\log n}\right)^\beta \mathcal{S}_L(S) \rho^{n-m} \exp(-C_L(\rho; J^{(1)})).$$

Proof: We have

$$P_L(S, n) = \frac{[z^n] L(z; S)}{[z^n] L(z)}.$$

Define restricted patterns $S^{(1)}$ and $S^{(2)}$ such that $S^{(1)}(j) = S(j)$ for $j \in J^{(1)}$ and $S^{(2)}(j) = S(j)$ for $j \in J^{(2)}$. Let $n^* = n - \sum_{j \in J^{(2)}} j S^{(2)}(j)$. Applying Theorem 3.1.1, as we did in the proof of Theorem 4.2.1, we obtain

$$P_L(S, n) = \frac{\mathcal{S}_L(S^{(1)} \cup S^{(2)})[z^m] \exp(C_L(z) - C_L(z; J^{(1)}) - C_L(z; J^{(2)}))}{[z^n]L(z)}.$$

Since $m < \min\{j : j \in J^{(2)}\}$, we can remove the term $C_L(z; J^{(2)})$, namely,

$$[z^m] \exp(C_L(z) - C_L(z; J^{(1)}) - C_L(z; J^{(2)})) = [z^m] \exp(C_L(z) - C_L(z; J^{(1)})).$$

Therefore,

$$P_L(S, n) \sim \mathcal{S}_L(S^{(2)}) \frac{[z^{n^*}]L(z; S^{(1)})}{[z^n]L(z)}. \quad (5.4)$$

For the term $[z^{n^*}]L(z; S^{(1)})$, we shall use the Cauchy integration formula. We choose a contour similar to the one in [17]. Let $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$ be positively oriented, where

$$\begin{aligned} \gamma_1 &= \{z : |z - 1| = 1/m, \theta < \text{Arg}(z - 1) < 2\pi - \theta\}, \\ \gamma_2 &= \{z : 1/m \leq |z - 1| \leq E \log m/m, \text{Arg}(z - 1) = \theta\}, \\ \gamma_3 &= \{z : |z| = 1 + 3 \log m/m, \theta < \text{Arg}(z - 1) < 2\pi - \theta\}, \\ \gamma_4 &= \{z : 1/m \leq |z - 1| \leq E \log m/m, \text{Arg}(z - 1) = 2\pi - \theta\}, \end{aligned}$$

where real number E satisfies

$$\left| 1 + E \frac{\log m}{m} e^{i\theta} \right| = 1 + 3 \log m/m.$$

Let

$$f(z) = \frac{\exp(C(\rho z) - C_L(\rho z; J^{(1)}))}{z^{m+1}}.$$

Then we have

$$\begin{aligned} [z^{n^*}]L(z; S^{(1)}) &= \rho^{-n^*} [z^{n^*}]L(\rho z; S^{(1)}) \\ &= \rho^{-m} \mathcal{S}_L(S^{(1)}) \frac{1}{2\pi i} \int_{\gamma} f(z) dz. \end{aligned} \quad (5.5)$$

Since $|z| \leq 1 + 3 \log m/m$ and $\hat{j} = O(m/\log m)$, we have

$$|z^j| \leq \left(1 + 3 \frac{\log m}{m}\right)^j = 1 + O(j \log m/m) = O(1).$$

From (5.3), we have

$$\sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} = O(1). \quad (5.6)$$

Hence, when $|z| \leq 1 + 3 \log m/m$,

$$\sum_{j \in J^{(1)}} C_j \frac{(\rho z)^j}{j!} = O(1). \quad (5.7)$$

Applying (5.1) and Taylor expansion, on the contour γ , we obtain

$$\begin{aligned} f(z) &= z^{-(m+1)} \exp \left(c + (1-z)^\alpha \left(\log \frac{1}{1-z} \right)^\beta (d + o(1)) - \sum_{j \in J^{(1)}} C_j \frac{(\rho z)^j}{j!} \right) \\ &= \frac{\exp \left(c - \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} \right)}{z^{m+1}} \\ &\quad \exp \left((1-z)^\alpha \left(\log \frac{1}{1-z} \right)^\beta (d + o(1)) + \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} (1-z^j) \right) \\ &= \frac{\exp \left(c - \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} \right)}{z^{m+1}} (1 + g(z, J^{(1)}) + O(g^2(z, J^{(1)}))), \end{aligned} \quad (5.8)$$

where $g(z, J^{(1)}) = (1-z)^\alpha \left(\log \frac{1}{1-z} \right)^\beta (d + o(1)) + \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} (1-z^j)$. In the last line, we can expand the exponential function because of (5.6), (5.7) and $0 < \alpha < 1$.

Substituting (5.8) for $f(z)$ in (5.5), we obtain

$$\begin{aligned}
[z^{n^*}]L(z; S^{(1)}) &= \rho^{-m} \mathcal{S}_L(S^{(1)}) \frac{1}{2\pi i} \int_{\gamma} \frac{\exp\left(c - \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!}\right)}{z^{m+1}} (1 + g(z, J^{(1)}) + O(g^2(z, J^{(1)}))) dz \\
&= \rho^{-m} \mathcal{S}_L(S^{(1)}) \frac{\exp\left(c - \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!}\right)}{2\pi i} \left(\int_{\gamma} z^{-(m+1)} dz + \int_{\gamma} \frac{g(z, J^{(1)})}{z^{m+1}} dz + \int_{\gamma} \frac{O(g^2(z, J^{(1)}))}{z^{m+1}} dz \right) \\
&= \rho^{-m} \mathcal{S}_L(S^{(1)}) \frac{\exp\left(c - \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!}\right)}{2\pi i} \left(\int_{\gamma} \frac{g(z, J^{(1)})}{z^{m+1}} dz + \int_{\gamma} \frac{O(g^2(z, J^{(1)}))}{z^{m+1}} dz \right),
\end{aligned}$$

where the term $\int_{\gamma} z^{-(m+1)} dz$ in the second line is 0.

In the following, we first discuss the integral $\int_{\gamma} \frac{g(z, J^{(1)})}{z^{m+1}} dz$. Then, we show that the second integral $\int_{\gamma} \frac{O(g^2(z, J^{(1)}))}{z^{m+1}} dz$ is negligible compared with $\int_{\gamma} \frac{g(z, J^{(1)})}{z^{m+1}} dz$.

I. First integral.

Estimate on $\gamma_1 \cup \gamma_2 \cup \gamma_4$:

Let $z = 1 + \frac{w}{m}$. From the definition of γ_1 , γ_2 , and γ_4 , we notice that $w = O(\log m)$.

Since

$$z^j = \left(1 + \frac{w}{m}\right)^j = 1 + O(jw/m),$$

applying (5.3) and $\sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha}(\log m)^{\beta-1})$, we get

$$\begin{aligned}
\sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} (1 - z^j) &= \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} O\left(\frac{jw}{m}\right) = O\left(\frac{w}{m}\right) \sum_{j \in J^{(1)}} j C_j \frac{\rho^j}{j!} \\
&= O\left(\frac{w}{m} \sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^\alpha}\right) = o\left(\frac{(\log m)^\beta}{m^\alpha}\right).
\end{aligned}$$

Hence,

$$\begin{aligned} g(z, J^{(1)}) &= \left(-\frac{w}{m}\right)^\alpha \left(\log\left(-\frac{m}{w}\right)\right)^\beta (d + o(1)) + \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!} (1 - z^j) \\ &= \frac{(\log m)^\beta}{m^\alpha} (-w)^\alpha (d + o(1)), \end{aligned} \quad (5.9)$$

where the $o(1)$ is uniform as $m \rightarrow \infty$ for w on the contour $\gamma_1 \cup \gamma_2 \cup \gamma_4$.

Let $\gamma' = \gamma'_1 + \gamma'_2 + \gamma'_4$ where

$$\begin{aligned} \gamma'_1 &= \{w : |w| = 1, \theta < \text{Arg}(w) < 2\pi - \theta\}, \\ \gamma'_2 &= \{w : 1 \leq |w| \leq E \log m, \text{Arg}(w) = \theta\}, \\ \gamma'_4 &= \{w : 1 \leq |w| \leq E \log m, \text{Arg}(w) = 2\pi - \theta\}. \end{aligned}$$

Then

$$\begin{aligned} &\frac{1}{2\pi i} \int_{\gamma_1 \cup \gamma_2 \cup \gamma_4} \frac{g(z, J^{(1)})}{z^{m+1}} dz \\ &= \frac{1}{2\pi i m} \int_{\gamma'} \frac{(\log m)^\beta}{m^\alpha} (-w)^\alpha (d + o(1)) \left(1 + \frac{w}{m}\right)^{-(m+1)} dw \\ &= \frac{1}{2\pi i m} \int_{\gamma'} \frac{(\log m)^\beta}{m^\alpha} (-w)^\alpha e^{-w} (d + o(1)) dw \\ &= \frac{d(\log m)^\beta}{2\pi i m^{1+\alpha}} \int_{\gamma'} (-w)^\alpha e^{-w} dw + o\left(\frac{(\log m)^\beta}{m^{1+\alpha}} \int_{\gamma'} (-w)^\alpha e^{-w} dw\right). \end{aligned}$$

We apply the Gamma function as we did in the proof of Theorem 4.2.1. It follows that

$$\frac{1}{2\pi i} \int_{\gamma_1 \cup \gamma_2 \cup \gamma_4} \frac{g(z, J^{(1)})}{z^{m+1}} dz \sim \frac{(\log m)^\beta}{m^{1+\alpha}} \frac{d}{\Gamma(-\alpha)}. \quad (5.10)$$

Estimate on γ_3 :

Let $z = (1 + 3 \log m / m) e^{it}$ where $\theta < t < 2\pi - \theta$. Using (5.7) and the definition

of alg-log type, we obtain

$$|g(z, J^{(1)})| = O(1).$$

We also have

$$|z^{-(m+1)}| = |(1 + 3 \log m/m)^{-(m+1)}| = O(1/m^3). \quad (5.11)$$

Therefore,

$$\frac{1}{2\pi i} \int_{\gamma_3} \frac{g(z, J^{(1)})}{z^{m+1}} dz = O(m^{-3}). \quad (5.12)$$

Since we have $0 < \alpha < 1$, the integral on γ_3 is negligible compared with the integral on $\gamma_1 \cup \gamma_2 \cup \gamma_4$.

II. Second integral.

On $\gamma_1 \cup \gamma_2 \cup \gamma_4$, let $z = 1 + \frac{w}{m}$. From (5.9),

$$g^2(z, J^{(1)}) = \frac{(\log m)^{2\beta}}{m^{2\alpha}} (-w)^{2\alpha} (d^2 + o(1)).$$

Using arguments similar to those used to derive (5.10), we have

$$\int_{\gamma_1 \cup \gamma_2 \cup \gamma_4} \frac{g^2(z, J^{(1)})}{z^{m+1}} dz = O\left(\frac{(\log m)^{2\beta}}{m^{1+2\alpha}}\right).$$

On γ_3 , as in the derivation of (5.26), we get

$$\int_{\gamma_3} \frac{g^2(z, J^{(1)})}{z^{m+1}} dz = O(m^{-3}),$$

which is negligible compared with the integral on $\gamma_1 \cup \gamma_2 \cup \gamma_4$.

It is clear that the integral

$$\rho^{-m} \mathcal{S}_L(S^{(1)}) \frac{\exp\left(c - \sum_{j \in J^{(1)}} C_j \frac{\rho^j}{j!}\right)}{2\pi i} \int_{\gamma} \frac{g(z, J^{(1)})}{z^{m+1}} dz$$

is the dominant part in $[z^{n^*}]L(z; S^{(1)})$ since $0 < \alpha < 1$. □

According to Lemma 5.2.1, we can simplify the requirements as in the following corollary although this makes the applicability of our theorem a little bit weaker.

Corollary 5.2.1. *With the same notations as the ones in Theorem 5.2.1, suppose $\hat{j} = O(m/\log m)$ and $|J^{(1)}| = o\left(\frac{m}{(\log m)^{\frac{1}{1-\alpha}}}\right)$. We have, as $m \rightarrow \infty$,*

$$P_L(S, n) \sim \left(\frac{n}{m}\right)^{1+\alpha} \left(\frac{\log m}{\log n}\right)^\beta \mathcal{S}_L(S) \rho^{n-m} \exp(-C_L(\rho; J^{(1)})).$$

Proof: Since $|J^{(1)}| = o\left(\frac{m}{(\log m)^{\frac{1}{1-\alpha}}}\right)$, applying Lemma 5.2.1, we have

$$\begin{aligned} \sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^\alpha} &\leq \sum_{j \leq |J^{(1)}|} \frac{(\log j)^\beta}{j^\alpha} \\ &\sim \frac{1}{1-\alpha} |J^{(1)}|^{1-\alpha} (\log |J^{(1)}|)^\beta \\ &= o(m^{1-\alpha} (\log m)^{\beta-1}). \end{aligned}$$

Hence, the requirements of Theorem 5.2.1 are satisfied. $\square$

5.2.2 Unlabeled Case

In the unlabeled case, the generating function of decomposable combinatorial structures is

$$U(z) = \exp\left(C(z) + \frac{C(z^2)}{2} + \frac{C(z^3)}{3} + \dots\right).$$

Let $0 < \rho < 1$. Therefore, $\frac{C(z^2)}{2} + \frac{C(z^3)}{3} + \dots$ has a radius of convergence bigger than ρ . As $z \rightarrow \rho$, we have

$$U(z) = \exp(C(z) + r_0(1 + o(1))). \quad (5.13)$$

where $r_0 = \frac{C(\rho^2)}{2} + \frac{C(\rho^3)}{3} + \dots$. It is easy to see that $U(z)$ is also of alg-log type.

Propositions 2.3.1 and 2.3.3 implies

$$[z^n]U(z) = \frac{e^{c+r_0}}{\rho^n \Gamma(-\alpha)} \frac{(\log n)^\beta}{n^{1+\alpha}} (d + o(1)), \quad (5.14)$$

$$C_n = \frac{1}{\rho^n \Gamma(-\alpha)} \frac{(\log n)^\beta}{n^{1+\alpha}} (d + o(1)), \quad (5.15)$$

whenever $\alpha \notin \{0, 1, 2, \dots\}$.

Theorem 5.2.2 gives the asymptotic probability that a random unlabeled object has a given restricted pattern in the case $m \rightarrow \infty$.

Theorem 5.2.2. *Let $P_U(S, n)$ be the probability that a random decomposable unlabeled object of size n , which is in the alg-log class with an algebraic exponent $0 < \alpha < 1$, has a restricted pattern S . For $\hat{j} = \max\{j; j \in J^{(1)}\}$, suppose $\hat{j} = O(m/\log m)$ and $\sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha}(\log m)^{\beta-1})$. We have, as $m \rightarrow \infty$,*

$$P_U(S, n) \sim \left(\frac{n}{m}\right)^{1+\alpha} \left(\frac{\log m}{\log n}\right)^\beta \mathcal{S}_U(S) \rho^{n-m} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j}.$$

Proof: We have

$$P_U(S, n) = \frac{[z^n]U(z; S)}{[z^n]U(z)}.$$

Similar to the discussion in the proof of Theorem 5.2.1, we can prove

$$P_U(S, n) \sim \mathcal{S}_U(S^{(2)}) \frac{[z^{n^*}]U(z; S^{(1)})}{[z^n]U(z)}. \quad (5.16)$$

The only term unknown in the above expression is $[z^{n^*}]U(z; S^{(1)})$. Then we apply singularity analysis to estimate $[z^{n^*}]U(z; S^{(1)})$. As always, we choose a contour similar to the one in [17]. Let $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$ be positively oriented, where

$$\begin{aligned} \gamma_1 &= \{z : |z - 1| = 1/m, \theta < \text{Arg}(z - 1) < 2\pi - \theta\}, \\ \gamma_2 &= \{z : 1/m \leq |z - 1| \leq E \log m/m, \text{Arg}(z - 1) = \theta\}, \\ \gamma_3 &= \{z : |z| = 1 + 3 \log m/m, \theta < \text{Arg}(z - 1) < 2\pi - \theta\}, \\ \gamma_4 &= \{z : 1/m \leq |z - 1| \leq E \log m/m, \text{Arg}(z - 1) = 2\pi - \theta\}, \end{aligned}$$

where real number $E = O(1)$ satisfies

$$\left| 1 + E \frac{\log m}{m} e^{i\theta} \right| = 1 + 3 \log m/m.$$

Let

$$f(z) = \frac{\prod_{j \in J^{(1)}} (1 - (\rho z)^j)^{C_j} \exp \left(C(\rho z) + \frac{C((\rho z)^2)}{2} + \dots \right)}{z^{m+1}}. \quad (5.17)$$

Then, applying (4.8), we have

$$[z^{n^*}]U(z; S^{(1)}) = \rho^{-n^*} [z^{n^*}]U(\rho z; S^{(1)}) = \rho^{-m} \mathcal{S}_U(S^{(1)}) \frac{1}{2\pi i} \int_{\gamma} f(z) dz.$$

From (5.13) and (5.8), for $z \in \gamma$, we can get

$$\begin{aligned} & \exp \left(C(\rho z) + \frac{C((\rho z)^2)}{2} + \frac{C((\rho z)^3)}{3} + \dots \right) \\ &= e^{c+r_0} \exp(C(\rho z)(1 + o(1))) \\ &= e^{c+r_0} \exp \left((1-z)^\alpha \left(\log \frac{1}{1-z} \right)^\beta (d + o(1)) \right) (1 + o(1)). \end{aligned} \quad (5.18)$$

For $z \in \gamma$, $|z| \leq 1 + 3 \log m/m$ as $m \rightarrow \infty$. Since $\hat{j} = O(m/\log m)$ and $0 < \rho < 1$, for $j \in J^{(1)}$, it is easy to see that

$$|z^j| = O(1), \quad (5.19)$$

and

$$\left| \sum_{i \geq 0} \frac{(\rho z)^{ij}}{i+1} \right| = O(1). \quad (5.20)$$

Using (5.19), (5.20) and $C_j \rho^j \sim \frac{d(\log j)^\beta}{\Gamma(-\alpha)j^{1+\alpha}}$ as $j \rightarrow \infty$, we have

$$\begin{aligned} \left| \sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{(\rho z)^{ij}}{i} \right| &= \left| \sum_{j \in J^{(1)}} \frac{d(\log j)^\beta}{\Gamma(-\alpha)j^{1+\alpha}} z^j (1 + o(1)) \sum_{i \geq 0} \frac{(\rho z)^{ij}}{i+1} \right| \\ &= \left| O \left(\sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^{1+\alpha}} \right) \right| \\ &= O(1), \end{aligned} \quad (5.21)$$

where the last line results from $0 < \alpha < 1$. Let $z = 1$, we get

$$\left| \sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{\rho^{ij}}{i} \right| = O(1). \quad (5.22)$$

We have

$$\begin{aligned} \prod_{j \in J^{(1)}} (1 - (\rho z)^j)^{C_j} &= \exp \left(\sum_{j \in J^{(1)}} C_j \log(1 - (\rho z)^j) \right) \\ &= \exp \left(- \sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{(\rho z)^{ji}}{i} \right) \\ &= \exp \left(- \sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{\rho^{ij}}{i} + \sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{\rho^{ij}}{i} (1 - z^{ij}) \right) \\ &= \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \exp \left(\sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{\rho^{ij}}{i} (1 - z^{ij}) \right). \end{aligned} \quad (5.23)$$

Therefore, using (5.18), (5.23), (5.21) and (5.22),

$$\begin{aligned} &\frac{1}{2\pi i} \int_{\gamma} f(z) dz \\ &= \frac{e^{c+r_0}}{2\pi i} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \int_{\gamma} z^{-m-1} \exp \left((1-z)^{\alpha} \left(\log \frac{1}{1-z} \right)^{\beta} \right. \\ &\quad \left. (d + o(1)) + \sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{\rho^{ij}}{i} (1 - z^{ij}) \right) dz \\ &= \frac{e^{c+r_0}}{2\pi i} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \int_{\gamma} (1 + \hat{g}(z, J^{(1)}) + O(\hat{g}^2(z, J^{(1)}))) \frac{(1 + o(1))}{z^{m+1}} dz \\ &= \frac{e^{c+r_0}}{2\pi i} \prod_{j \in J^{(1)}} (1 - \rho^j)^{C_j} \left(\int_{\gamma} \frac{(1 + o(1))}{z^{m+1}} dz + \int_{\gamma} \frac{\hat{g}(z, J^{(1)})(1 + o(1))}{z^{m+1}} dz \right. \\ &\quad \left. + \int_{\gamma} \frac{O(\hat{g}^2(z, J^{(1)}))(1 + o(1))}{z^{m+1}} dz \right), \end{aligned} \quad (5.24)$$

where $\hat{g}(z, J^{(1)}) = (1-z)^{\alpha} \left(\log \frac{1}{1-z} \right)^{\beta} (d + o(1)) + \sum_{j \in J^{(1)}} C_j \sum_{i \geq 1} \frac{\rho^{ij}}{i} (1 - z^{ij})$. It is easy to see that the first integral $\int_{\gamma} \frac{(1+o(1))}{z^{m+1}} dz$ is 0.

Following the proof of Theorem 5.2.1, we begin with the integral on contour $\gamma' = \gamma_1 + \gamma_2 + \gamma_4$. Let $z = 1 + w/m$ and note $|w| = O(\log m)$ when $z \in \gamma'$.

Using the condition $\sum_{j \in J(1)} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha}(\log m)^{\beta-1})$, we obtain

$$\begin{aligned} \sum_{j \in J(1)} C_j \sum_{i \geq 1} \frac{\rho^{ij}}{i} (1 - z^{ij}) &= \sum_{j \in J(1)} C_j \rho^j (1 - z^j) \sum_{i \geq 0} \frac{\rho^{ij}}{i+1} (1 + z^j + z^{2j} + \dots + z^{ij}) \\ &= \sum_{j \in J(1)} O\left(\frac{(\log j)^\beta}{j^{1+\alpha}}\right) O\left(\frac{jw}{m}\right) O(1) \\ &= o\left(\frac{(\log m)^\beta}{m^\alpha}\right). \end{aligned}$$

In the third line, the last term $O(1)$ comes from $\sum_{i \geq 0} \frac{\rho^{ij}}{i+1} (1 + z^j + z^{2j} + \dots + z^{ij})$. Indeed, when $z = 1 + w/m$ as $|w| \leq 3 \log m$, the condition $0 < \rho < 1$ derives $0 < |(\rho z)^j| < 1$. Hence,

$$\left| \sum_{i \geq 0} \frac{\rho^{ij}}{i+1} (1 + z^j + z^{2j} + \dots + z^{ij}) \right| \leq \left| \sum_{i \geq 0} (\rho z)^{ij} \right| = O(1).$$

Replacing z by $1 + \frac{w}{m}$, we obtain

$$\begin{aligned} \hat{g}(z, J(1)) &= \left(-\frac{w}{m}\right)^\alpha \left(\log\left(-\frac{m}{w}\right)\right)^\beta (d + o(1)) + o\left(\frac{(\log m)^\beta}{m^\alpha}\right) \\ &= \frac{(\log m)^\beta}{m^\alpha} (-w)^\alpha (d + o(1)), \end{aligned}$$

where the $o(1)$ is uniform as $m \rightarrow \infty$ for w on the contour γ' .

Now, the problem turns out to be the one we already solved in the proof of Theorem 5.2.1. We apply properties of the Γ function and a change of variable to get

$$\frac{1}{2\pi i} \int_{\gamma'} \frac{\hat{g}(z, J(1))}{z^{m+1}} dz \sim \frac{(\log m)^\beta}{m^{1+\alpha}} \frac{d}{\Gamma(-\alpha)}. \quad (5.25)$$

On γ_3 , let $z = (1 + 3 \log m/m)e^{it}$ where $\theta < t < 2\pi - \theta$. From (5.21) and (5.22), it

is obvious that $|\hat{g}(z, J^{(1)})| = O(1)$. Considering (5.11), we have

$$\frac{1}{2\pi i} \int_{\gamma_3} \frac{g(z, J^{(1)})}{z^{m+1}} dz = O(m^{-3}). \quad (5.26)$$

Since the constant α satisfies $0 < \alpha < 1$, the integral on γ_3 is negligible compared with the integral on γ' .

The last integral $\int_{\gamma} \frac{O(\hat{g}^2(z, J^{(1)}))(1+o(1))}{z^{m+1}} dz$ can also be proved to be $O\left(\frac{(\log m)^{2\beta}}{m^{1+2\alpha}}\right)$ as we did in the proof of Theorem 5.2.1. Applying (5.25) and (5.16), we complete the proof. $\square$

5.3 Size of the r th Smallest Component

In this section, we discuss the distribution of the r th smallest component with a given restricted pattern. We continue to use the notations defined in Section 4.3. We derive an asymptotic expression for $P(X_n^{[r]}(S) > k)$ where $X_n^{[r]}(S)$ is the size of the r th smallest component of a decomposable combinatorial structure of size n having a restricted pattern S . In the following, we take $P(X_n^{[r]}(S) > k) = 0$ when $d(k) < r$ without declaring it.

5.3.1 Labeled Case

Theorem 5.3.1. *Let $S: J \mapsto \mathbb{N}$ be a restricted pattern with*

$$\sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha}(\log m)^{\beta-1}).$$

Assume $\hat{j} = O(m/\log m)$. Let $r - d(k) = O(\log m)$, $k = o\left(\frac{m}{(\log m)^{\frac{1}{1-\alpha}}}\right)$. As $m \rightarrow \infty$, we have

$$P(X_n^{[r]}(S) > k) \sim P_L(S, n) \exp(-C_L(\rho; N_k \setminus J)) \sum_{j=0}^{r-1-d(k)} \frac{(C_L(\rho; N_k \setminus J))^j}{j!}.$$

Proof: We use a similar approach as the one of the corresponding theorem for the exp-log class. Define the augmented pattern $S^*: J^* \mapsto \mathbb{N}$ where $J^* = J \cup N_k$, and $S^*(j) = S(j)$ as $j \in J$, and, for $j \in N_k \setminus J$, $S^*(j)$ is specified below. We note that each structure with restricted pattern S and r th smallest component with size greater than k corresponds to structures with an augmented pattern S^* such that $\sum_{i \in N_k \setminus J} S^*(i) \leq r - 1 - d(k)$. Hence we have

$$P(X_n^{[r]}(S) > k) = \sum_{\sum_{i \in N_k \setminus J} S^*(i) \leq r-1-d(k)} P_L(S^*, n).$$

Let $m^* = n - \sum_{j \in J \cup N_k} j S^*(j)$. Since $k = o(m/\log m)$ and $\sum_{j \in N_k \setminus J} S^*(j) \leq r - 1 - d(k) = O(\log m)$, we have $m^* \sim m$. Hence,

$$\left(\frac{n}{m^*}\right)^{1+\alpha} \left(\frac{\log m^*}{\log n}\right)^\beta \sim \left(\frac{n}{m}\right)^{1+\alpha} \left(\frac{\log m}{\log n}\right)^\beta.$$

Applying $k = o\left(\frac{m}{(\log m)^{\frac{1}{1-\alpha}}}\right)$ and Lemma 5.2.1, we get

$$\sum_{j \in N_k} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha}(\log m)^{\beta-1}).$$

Therefore, we have that the augmented pattern S^* satisfies the requirements of Theorem 5.2.1. Noticing that $\sum_{j \in J \cap N_k} j S(j)$ is given by the pattern S , we get

$$P(X_n^{[r]}(S) > k) \sim \sum_{\sum_{j \in N_k \setminus J} S^*(j) \leq r-1-d(k)} \left(\frac{n}{m^*}\right)^{1+\alpha} \left(\frac{\log m^*}{\log n}\right)^\beta \prod_{i \in J \cup N_k} \frac{C_i^{S^*(i)}}{(i!)^{S^*(i)} S^*(i)!}$$

$$\begin{aligned}
& \rho^{n-m^*} \exp(-C_L(\rho; J^{(1)} \cup N_k)) \\
& \sim \left(\frac{n}{m}\right)^{1+\alpha} \left(\frac{\log m}{\log n}\right)^\beta \mathcal{S}_L(S) \rho^{n-m} \exp(-C_L(\rho; J^{(1)}) - C_L(\rho; N_k \setminus J)) \\
& \quad \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \left(\prod_{i \in N_k \setminus J} \frac{C_i^{S^*(i)} \rho^{i S^*(i)}}{(i!)^{S^*(i)} S^*(i)!} \right) \\
& \sim P_L(S, n) \exp(-C_L(\rho; N_k \setminus J)) \\
& \quad \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \left(\prod_{i \in N_k \setminus J} \frac{C_i^{S^*(i)} \rho^{i S^*(i)}}{(i!)^{S^*(i)} S^*(i)!} \right). \tag{5.27}
\end{aligned}$$

Applying the same transformation as we did in Theorem 4.3.1, we obtain

$$P(X_n^{[r]}(S) > k) \sim P_L(S, n) \exp(-C_L(\rho; N_k \setminus J)) \sum_{j=0}^{r-1-d(k)} \frac{(C_L(\rho; N_k \setminus J))^j}{j!}.$$

□

In this theorem, we can not give a better expression in the case $k \rightarrow \infty$ as we did in Theorem 4.3.1. In the alg-log type, we have

$$\begin{aligned}
C_L(\rho; N_k \setminus J) &= \sum_{i=1}^k C_i \rho^i / i! - C_L(\rho; J \cap N_k) \\
&= [z^k] \frac{1}{1-z} \left(c + (1-z)^\alpha \left(\log \frac{1}{1-z} \right)^\beta (d + o(1)) \right) - C_L(\rho; J \cap N_k) \\
&= c + \frac{(\log k)^\beta}{k^\alpha \Gamma(1-\alpha)} + O(k^{-\alpha} (\log k)^{\beta-1}) - C_L(\rho; J \cap N_k).
\end{aligned}$$

Moreover, since $C_L(\rho; J \cap N_k)$ depends on the pattern, we can not give the dominant part of $C_L(\rho; N_k \setminus J)$ as before.

When there is no restricted pattern, applying Theorem 4.3.1, we obtain the following corollary.

Corollary 5.3.1. *Let $r = O(\log n)$ and $k = o\left(n/(\log n)^{\frac{1}{1-\alpha}}\right)$. As $n \rightarrow \infty$, we have*

$$P(X_n^{[r]} > k) \sim \exp(-C_L(\rho; N_k)) \sum_{j=0}^{r-1} \frac{C_L(\rho; N_k)^j}{j!},$$

and if k is unbounded,

$$P(X_n^{[r]} > k) \sim e^{-c} \left(\sum_{j=0}^{r-1} \frac{c^j}{j!} \right).$$

5.3.2 Unlabeled Case

In the unlabeled case, we require the singularity point $0 < \rho < 1$ which is not a required condition in the labeled case. The reason has been given in Section 5.2.2.

Theorem 5.3.2. *Let $S: J \mapsto \mathbb{N}$ be a restricted pattern with*

$$\sum_{j \in J^{(1)}} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha}(\log m)^{\beta-1}).$$

Assume $\hat{j} = O(m/\log m)$. Let $r = O(\log m)$, $k = o\left(\frac{m}{(\log m)^{\frac{1}{1-\alpha}}}\right)$. As $m \rightarrow \infty$, we have

$$P(X_n^{[r]}(S) > k) \sim P_U(S, n) U(\rho; N_k \setminus J) \sum_{j=0}^{r-1-d(k)} [z^j] \exp \left(z C_U(\rho; N_k \setminus J) + z^2 \frac{C_U(\rho^2; N_k \setminus J)}{2} + \dots \right), \quad (5.28)$$

where $U(\rho; N_k \setminus J)$ is defined in Section 4.3.2.

Proof: In the unlabeled case, we still have

$$P(X_n^{[r]}(S) > k) = \sum_{\substack{i \in N_k \setminus J \\ S^*(i)=0}}^{r-1-d(k)} P_U(S^*, n),$$

which is explained in the proof of Theorem 5.3.1. We define the same augmented pattern S^* as the one in the proof of Theorem 5.3.1. Let $m^* = n - \sum_{j \in J \cup N_k} j S^*(j)$.

For the same reason as in the labeled case, we obtain

$$\left(\frac{n}{m^*} \right)^{1+\alpha} \left(\frac{\log m^*}{\log n} \right)^\beta \sim \left(\frac{n}{m} \right)^{1+\alpha} \left(\frac{\log m}{\log n} \right)^\beta,$$

and the augmented pattern S^* satisfies the requirements of Theorem 5.2.2.

Applying Theorem 5.2.2, we get

$$\begin{aligned}
& P(X_n^{[r]}(S) > k) \\
& \sim \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \left(\frac{n}{m^*}\right)^{1+\alpha} \left(\frac{\log m^*}{\log n}\right)^\beta \prod_{j \in J \cup N_k} \binom{C_j + S^*(j) - 1}{S^*(j)} \\
& \quad \prod_{j \in J^{(1)} \cup N_k} (1 - \rho^j)^{C_j} \rho^{n-m^*} \\
& \sim P_U(S, n) U(\rho; N_k \setminus J) \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \prod_{j \in N_k \setminus J} \binom{C_j + S^*(j) - 1}{S^*(j)} \rho^{j S^*(j)},
\end{aligned} \tag{5.29}$$

where $U(z; J) = \prod_{i \in J} (1 - z^i)^{C_i}$.

The negative binomial theorem states that

$$(1 - \rho^i z)^{-C_i} = \sum_{X_i \geq 0} \binom{C_i + X_i - 1}{X_i} \rho^{i X_i} z^{X_i}.$$

Hence, we get

$$\begin{aligned}
& \sum_{\sum_{j \in N_k \setminus J} S^*(j)=0}^{r-1-d(k)} \prod_{j \in N_k \setminus J} \binom{C_j + S^*(j) - 1}{S^*(j)} \rho^{j S^*(j)} \\
& = \sum_{u=0}^{r-1-d(k)} [z^u] \prod_{j \in N_k \setminus J} (1 - \rho^j z)^{-C_j} \\
& = \sum_{u=0}^{r-1-d(k)} [z^u] \exp \left(- \sum_{j \in N_k \setminus J} C_j \log(1 - \rho^j z) \right) \\
& = \sum_{u=0}^{r-1-d(k)} [z^u] \exp \left(\sum_{i \geq 1} \sum_{j \in N_k \setminus J} C_j \rho^{ij} z^i / i \right) \\
& = \sum_{u=0}^{r-1-d(k)} [z^u] \exp \left(z(C_U(\rho; N_k) - C_U(\rho; J \cap N_k)) \right. \\
& \quad \left. + z^2 \frac{C_U(\rho^2; N_k) - C_U(\rho^2; J \cap N_k)}{2} + \dots \right).
\end{aligned} \tag{5.30}$$

This completes the proof. $\square$

5.4 Expectation

In this section, we discuss the expectation of the r th smallest component with a given restricted pattern. We assume $J^{(2)} = \emptyset$. With all notations defined in Section 4.4, we still use $\hat{d}$ to denote the total number of components in the given restricted pattern, namely $\hat{d} = \sum_{j \in J} S(j)$. If $r \leq \hat{d}$, it is clear that $P(X_n^{[r]}(S) > k) = 0$ when $\sum_{j \in J \cap N_k} S(j) \geq r$. Hence we have

$$E(X_n^{[r]}(S)) = \sum_{k=0}^{i^{[r]}-1} P(X_n^{[r]}(S) > k).$$

In the following, we only discuss the case $r = \hat{d} + 1$. From Theorem 5.3.1, we have the following theorem.

Theorem 5.4.1. *Let $S: J \mapsto \mathbb{N}$ be a restricted pattern with*

$$\sum_{j \in J} \frac{(\log j)^\beta}{j^\alpha} = o(m^{1-\alpha}(\log m)^{\beta-1}).$$

Assume $\hat{j} = o\left(\frac{m}{(\log m)^{\frac{1}{1-\alpha}}}\right)$. For both labeled and unlabeled cases, when $r = \hat{d} + 1$,

(i) in labeled case,

$$E(X_n^{[r]}(S)) \sim P_L(S, n)m \exp(-c + C_L(\rho; J));$$

(ii) in unlabeled case,

$$E(X_n^{[r]}(S)) \sim P_U(S, n)m \exp(-c - r_0) \prod_{j \in J} (1 - \rho^j)^{-C_j}.$$

Proof: To estimate the first moment, we use the well known expression

$$E(X_n^{[r]}(S)) = \sum_{k \geq 0} P(X_n^{[r]}(S) > k). \quad (5.31)$$

We only prove our theorem in the labeled case since the proof is very similar in the unlabeled case.

Let us begin from a special case: there is only one component except the components of the pattern. Namely, we have $\hat{d} + 1$ components in total. Therefore, the size of the only component not in the pattern is m . The probability that an object has exact $\hat{d} + 1$ components is

$$\begin{aligned}
 P(X_n^{[\hat{d}+1]}(S) = m) &= \frac{\mathcal{S}_L(S) \frac{C_m}{m!}}{[z^n]L(z)} \\
 &= \frac{\mathcal{S}_L(S) \frac{1}{\rho^m \Gamma(-\alpha)} \frac{(\log m)^\beta}{m^{1+\alpha}} (d + o(1))}{\frac{e^c}{\rho^n \Gamma(-\alpha)} \frac{(\log n)^\beta}{n^{1+\alpha}} (d + o(1))} \\
 &= \mathcal{S}_L(S) \rho^{n-m} e^{-c} \left(\frac{n}{m}\right)^{1+\alpha} \left(\frac{\log m}{\log n}\right)^\beta (1 + o(1)) \\
 &\sim P_L(S, n) e^{-c} \exp(C_L(\rho; J)).
 \end{aligned}$$

Then, let us see what happens when k is small. Let $t = \max\left(\hat{j} + 1, \left\lfloor \frac{m}{(\log m)^{\frac{2}{1-\alpha}}} \right\rfloor\right)$.

Applying Theorem 5.3.1, we obtain

$$P(X_n^{[\hat{d}+1]}(S) > t) \sim P_L(S, n) e^{-c} \exp(C_L(\rho; J)).$$

Clearly, for any $k < m$, we always have

$$P(X_n^{[\hat{d}+1]}(S) > k) \geq P(X_n^{[\hat{d}+1]}(S) = m) = P_L(S, n) e^{-c} \exp(C_L(\rho; J)).$$

Uniformly for $t \leq k \leq m$, we have

$$P(X_n^{[\hat{d}+1]}(S) > k) \sim P_L(S, n) e^{-c} \exp(C_L(\rho; J)).$$

Hence,

$$\begin{aligned}
 E(X_n^{[\hat{d}+1]}(S)) &= \sum_{k \geq 0} P(X_n^{[\hat{d}+1]}(S) > k) \\
 &= \sum_{0 \leq k < t} P(X_n^{[\hat{d}+1]}(S) > k) + \sum_{t \leq k \leq m} P(X_n^{[\hat{d}+1]}(S) > k) \\
 &\sim P_L(S, n) e^{-c} m \exp(C_L(\rho; J)).
 \end{aligned}$$

We get the last line since $\hat{j} = o\left(\frac{m}{(\log m)^{\frac{1}{1-\alpha}}}\right)$. $\square$

In the non-pattern case, let $X_n^{[r]}$ be the size of the r th smallest component of a random decomposable combinatorial structure of size n . The expectation of $X_n^{[1]}$ immediately follows from Theorem 5.4.1.

Corollary 5.4.1. *For both labeled and unlabeled cases,*

1. *in labeled case,*

$$E(X_n^{[1]}) \sim \frac{n}{e^c};$$

2. *in unlabeled case,*

$$E(X_n^{[1]}) \sim \frac{n}{e^{c+\tau_0}}.$$

5.5 Example

In this section, we provide one example for labeled case.

Example 5.5.1. *(Cayley Trees). A Cayley tree is a non-planar rooted labeled tree.*

Figure 5.1 shows all Cayley trees with 1, 2 and 3 nodes.

Let $T(z)$ be the generating function of trees. It is well known [22] that

$$T(z) = z \exp(T(z)) = 1 - \sqrt{2(1 - ez)} + O(1 - ez).$$

The tree function $L(z) = \exp(T(z))$ is in the alg-log class with the singularity point $\rho = 1/e$. By Corollary 5.4.1, the expected size of the r th smallest subtree is, when $r = 1$,

$$E(X_n^{[1]}) \sim \frac{n}{e}.$$

Figure 5.1: The Cayley trees with 1 node, 2 nodes and 3 nodes.

Let us consider a restricted pattern $S(\lfloor \log n \rfloor) = 3$ and $r = 5$. Then, Theorems 5.2.1 and 5.3.1 give

$$P(X_n^{[5]}(S) > k) \sim P_L(S) \exp(-C_L(\rho; N_k \setminus J)) \sum_{j=0}^{r-1-d} \frac{(C_L(\rho; N_k \setminus J))^j}{j!}.$$

If $k \rightarrow \infty$,

$$P(X_n^{[5]}(S) > k) \sim \frac{1+e}{6e\sqrt{2\pi}} (\log n)^{-1.5}.$$

If $k = O(1)$,

$$P(X_n^{[5]}(S) > k) \sim \frac{1}{6\sqrt{2\pi}} (\log n)^{-1.5} \exp(-C_L(\rho; N_k \setminus J)) \sum_{j=0}^{r-1-d} \frac{(C_L(\rho; N_k \setminus J))^j}{j!}.$$

Chapter 6

Conclusions and Future Work

In the previous chapters, we detailed our contributions of this thesis. In this chapter, we summarize the contributions and provide some problems for future work.

6.1 Conclusions

In this thesis, we study decomposable combinatorial structures with given restricted patterns. By using combinatorial methods, we obtain the generating functions of the decomposable combinatorial structures with given restricted patterns. We provide the generating functions when, in addition to the restricted pattern, we restrict the size of the r th smallest component of the decomposable structure.

For objects in the exp-log class, based on their generating functions, we provide the probability that random decomposable combinatorial structures have a given restricted pattern. Moreover, in addition to the restricted pattern, we discuss the probability that the size of the r th smallest component be bigger than some integer

k . We also provide the moments of the r th smallest component.

For other important objects, alg-log structures, we do a similar study as we did for exp-log class. We provide the probability of the appearance of a given restricted pattern and study the r th smallest component. About the first moment (expectation), we obtain the estimate for the expectation of the size of r th smallest component only if r equal to one plus the number of components in the given pattern.

6.2 Future Work

We now give some natural extensions of this work.

First, one could do further studies on the moments of the r th smallest component, especially for the decomposable combinatorial objects of alg-log type. The expectation of r th smallest component in an alg-log structure are more likely to grow very fast as r increasing. We plan to continue working on this problem.

According to the discussion in Appendix A, we can weaken the restrictive conditions in most theorems. But we have to be very careful on the different cases depending on the sizes of components of the pattern.

Furthermore, the distribution of the r th largest component is another potential subject of study. In the largest component case, new methods will have to be provided. Indeed, the largest size among the components in the restricted pattern would exceed the range allowed in the theorems given in this thesis. Gourdon's fundamental work [28] could be helpful in that study. Panario and Richmond also did some work [44] about the exact size of the largest component. Perhaps it is possible to extend [44] in the case of restricted patterns.

Finally, we intend to study the alg-log structures when the algebraic exponent α is not in $(0, 1)$. This problem may be more challenging than what it looks like. The condition $0 < \alpha < 1$ makes the problem much easier when we apply analytic method.

Appendix A

Discussion on the Probability of a Combinatorial Structure with Given Restricted Pattern

Given a restricted pattern S , recall that $J^{(1)} = \{j \in J : j \leq m\}$ and $J^{(2)} = \{j \in J : j > m\}$ where $m = n - \sum_{j \in J} jS(j)$. In Chapter 4 and 5, all theorems which involve restricted patterns require that $j = O(m/\log m)$ when $j \in J^{(1)}$. Namely, our theorems hold only if there is no component whose size is between $m/\log m$ and m in the given pattern. Then, a natural question is: do we have a method to deal with the ‘troublesome’ objects which have components with sizes between $m/\log m$ and m in the pattern? In this appendix, we try to weaken the restrictive conditions in our theorems. We put forward some ideas to attack it, instead of providing concrete result. This will be material for future studies.

Let us begin with the simplest pattern. A random decomposable combinatorial

structure with size n contains a restricted pattern S which has one component of size $n/\log n$ and one component of size $\frac{n+a}{4}$ where $n \rightarrow \infty$, $0 \leq a < n - \frac{2n}{\log n}$ and $\frac{n+a}{4}$ is a positive integer. We have $m = \frac{3n-a}{4} - n/\log n$ which is bigger than or equal to the size of the component. If we ask for the probability that a random decomposable combinatorial structure with size n contains such a pattern, our theorems can not deal with this question. Then, is there any method to transform this question into a new one for which our theorems apply?

Since the number of combinations without pattern is easy to know, we only need to seek for the number of combinations with the pattern. For a random combinatorial structure with size n , let $T^n(i)$ denote the number of combinations when each combination has one component of size i and $T_k^n(i, j)$ denote the number of different combinations when each combination has one component of size i and k components of size j . In this question, we are looking for the value of $T_1^n(n/\log n, \frac{n+a}{4})$. By applying a combinatorial method, it is clear that we can obtain $T_1^n(n/\log n, \frac{n+a}{4})$ only if we know the value of $T_0^{m+n/\log n}(n/\log n, \frac{n+a}{4})$. The values of the other two terms depend on a . In the next, we discuss the value of a .

(I) If $\frac{n}{3} - \frac{4n}{3\log n} \leq a < n - \frac{2n}{\log n}$.

In this case, $\frac{m}{2} \leq \frac{n+a}{4} < m$. We solve the question by using

$$T_0^{m+n/\log n}\left(\frac{n}{\log n}, \frac{n+a}{4}\right) = T^{m+n/\log n}(n/\log n) - T_1^{m+n/\log n}\left(\frac{n}{\log n}, \frac{n+a}{4}\right),$$

where $T^{m+n/\log n}(n/\log n)$ can be obtained by applying our theorems in Chapter 4 and 5.

Applying a combinatorial approach again, we can transform the question into asking for $T_0^{m+\frac{n}{\log n}-\frac{n+a}{4}}(n/\log n, \frac{n+a}{4})$. And, since $m + \frac{n}{\log n} - \frac{n+a}{4} < \frac{n+a}{4}$, it is possible to apply our theorems in Chapter 4 and 5. We have the following three conditions

depending on a :

1. $\frac{n}{\log n} = O\left(\frac{\frac{1}{2}n - \frac{1}{2}a - \frac{n}{\log n}}{\log\left(\frac{1}{2}n - \frac{1}{2}a - \frac{n}{\log n}\right)}\right)$. In this case, the requirements of our theorems are satisfied and $T_0^{m + \frac{n}{\log n} - \frac{n+a}{4}}(n/\log n, \frac{n+a}{4})$ can be obtained. In other words, if a is not very close to n , we can solve the question by applying our theorems.
2. $m + \frac{n}{\log n} - \frac{n+a}{4}$ is bounded by a constant. In this case, we can solve the question by applying a combinatorial method directly.
3. Above two conditions do not hold. This is the difficult part in this question. We have to try other methods which are beyond this thesis and is left for future work.

(II) If $0 \leq a < \frac{n}{3} - \frac{4n}{3\log n}$.

Similar to the Condition (I), we recurrently apply a combinatorial approach to decrease the size of the structure until it is smaller than $\frac{n+a}{4}$. Then, to solve the question, for some values of a , we can apply our theorems in this thesis or the combinatorial method we have commented.

In summary, we have provided an example about how to attack the questions not concerned by our theorems. The idea is to apply a combinatorial approach repeatedly, to reduce the unknown sizes until only those desired components in the given pattern remain in the structure. In practical questions, the pattern normally are not so simple. When the number of components with ‘troublesome’ sizes in the pattern grows, the complexity of the problem will increase rapidly. And, even if the size of the non covered components are very small, the problem will be very hard to deal with.

Bibliography

- [1] ARNEY, J. AND BENDER, E.D. *Random mappings with constraints on coalescence and number of origins*. Pacific J. Math, 103 (1982), 269-294.
- [2] ARRATIA, R. BARBOUR, A.D. AND TAVARÉ, S. *Random combinatorial structures and prime factorizations*. Notices of the American Mathematical Society, 44 (1997), no.8, 903-910.
- [3] ARRATIA, R. BARBOUR, A.D. AND TAVARÉ, S. *Logarithmic Combinatorial Structures: a Probabilistic Approach*. EMS Monographs in Mathematics, European Mathematical Society(EMS), Zürich. (2003).
- [4] BENDER, E. *Central and local limit theorems applied to asymptotic enumeration*. J. Combinatorial Theory Series. A, 15 (1973), 91-111.
- [5] BENDER, E. *Asymptotic methods in enumeration*. SIAM Reviews, 16 (1974), 485-515.
- [6] CANFIELD, E.R. *Central and local limit theorems for coefficients of binomial type*. J. Combinatorial Theory Series. A, 23 (1977), 275-290.
- [7] CAR, M. *Théorèmes de densité dans $\mathbb{F}_q[x]$* . Acta Arith., 48 (1987), 145-165.

- [8] COMPTON, K.J. *Some methods for computing component distribution probabilities in relational structures*. Discrete Mathematics, 66 (1987), 59-77.
- [9] COMTET, L. *Advance Combinatorics; The Art of Finite and Infinite Expansions*. Boston, MA, D. Reidel Publ. Co. (1974).
- [10] DONG, L., GAO, Z. AND PANARIO, D. *Enumeration of combinatorial decomposable structures with restricted patterns*. To appear in Annals of Combinatorics.
- [11] DONG, L., GAO, Z. AND PANARIO, D. *The size of the r th smallest component in decomposable structures with a restricted pattern*. To appear in Discrete Mathematics and Theoretical Computer Science.
- [12] FLAJOLET, P. AND GOLIN, M. *Mellin transforms and asymptotics: the merge-sort recurrence*. Acta Informatica, 31 (1994), 673-696.
- [13] FLAJOLET, P., GOURDON, X. AND PANARIO, D. *The complete analysis of a polynomial factorization algorithm over finite fields*. Journal of Algorithms, 40 (2001), 37-81.
- [14] FLAJOLET, P., GRABNER, P., KIRSCHENHOFER, P., PRODINGER, H. AND TICHY, R. *Mellin transforms and asymptotics: digital sums*. Theoretical Computer Science, 123 (1994), 291-314.
- [15] FLAJOLET, P. AND ODLYZKO, A.M. *The average height of binary trees and other simple trees*. J. Comput. System. Sci., 25 (1982), 171-213.
- [16] FLAJOLET, P. AND ODLYZKO, A.M. *Limit distributions for coefficients of iterates of polynomials with applications to combinatorial enumeration*. Math. Proc. Cambridge Phil. Soc., 96 (1984), 237-253.

- [17] FLAJOLET, P. AND ODLYZKO, A.M. *Singularity analysis of generating functions*. SIAM J. Discrete Math., 3 (1990), 216-240.
- [18] FLAJOLET, P. AND PRODINGER, H. *Stirling Numbers for Complex Arguments and Hankel Contours*. SIAM Journal on Discrete Mathematics, 12 (1999), 155-159.
- [19] FLAJOLET, P., SALVY, B. AND ZIMMERMANN, P. *Automatic average-case analysis of algorithms*. Theoretical Computer Science, 79 (1991), 37-109.
- [20] FLAJOLET, P. AND SORIA, M. *Gaussian limiting distribution for the number of components in combinatorial structures*. J. Combin. Theory Ser. A, 53 (1990), 165-182.
- [21] FLAJOLET, P. AND SORIA, M. *General combinatorial schemas: Gaussian limit distributions and exponential tails*. Discrete Mathematics, 114 (1993), 159-180.
- [22] FLAJOLET, P. AND SEDGEWICK, R. *Analytic Combinatorics*. To appear.
- [23] GAREFALAKIS, T. AND PANARIO, D. *Polynomials over finite fields free from large and small degree irreducible factors*. Journal of Algorithms, 44 (2002), 98-120.
- [24] GAREFALAKIS, T. AND PANARIO, D. *The index calculus method using non-smooth polynomials*. Mathematics of Computation, 70 (2001), 1253-1264.
- [25] GONCHAROV, V. *Sur la distribution des cycles dans une permutation*. Acad. Sci. URSS(N.S.), 35 (1942), 267-269.
- [26] GONCHAROV, V. *On the field of combinatorial analysis*. Amer. Math. Soc. Transl., 2 (1962), 1-46.

- [27] GOULDEN, I. AND JACKSON, D. *Combinatorial Enumerations*. Wiley, New York, (1983).
- [28] GOURDON, X. *Largest component in random combinatorial structures*. Discrete Mathematics, 180 (1998), 185-209.
- [29] GOURDON, X. *Combinatoire, algorithmique et géométrie des polynômes*. Ph.D thesis. École Polytechnique (1996).
- [30] HENRICI, P. *Applied and Computational Complex Analysis*. Vol.2, John Wiley, New York (1974).
- [31] HENRICI, P. *Applied and Computational Complex Analysis*. Vol.3, John Wiley, New York (1977).
- [32] JUNGEN, R. *Sur les séries de Taylor n'ayant que des singularités algebrico-logarithmiques sur leur circle de convergence*. Commentarii Mathematici Helvetici, 3 (1931), 266-306.
- [33] KNOPFMACHER, A. AND WARLIMONT, R. *Distinct degree factorizations for polynomials over a finite field*. Trans. Amer. Math. Soc, 347 (1995), 2235-2243.
- [34] KNOPFMACHER, A. AND WARLIMONT, R. *Counting permutations and polynomials with a restricted factorization pattern*. Australasian Journal of Combinatorics, 13 (1996), 151-162.
- [35] LANDO, S.K. *Lectures on Generating Functions*. American Mathematical Society (2003).
- [36] MEIR, A. AND MOON, J.W. *On the altitude of modes in random trees*. Canad. J. Math., 30 (1978), 997-1015.

- [37] ODLYZKO, A. *Periodic oscillations of coefficients of power series that satisfy functional equations*. Adv. in Math., 44 (1982), 180-205.
- [38] ODLYZKO, A. *Discrete logarithms in finite fields and their cryptographic significance*. Lecture Notes in Computer Science, 209 (1985), 224-314.
- [39] ODLYZKO, A. *Asymptotic enumeration methods*. In Handbook of Combinatorics. vol. 2 (1996).
- [40] OLVER, F.W.J. *Asymptotics and Special Functions*. Academic Press. New York (1974).
- [41] PANARIO, D., PITTEL, B., RICHMOND, B. AND VIOLA, A. *Analysis of Rabin's irreducibility test for polynomials over finite fields*. Random Structures and Algorithms, 19 (2001), 525-551.
- [42] PANARIO, D. AND RICHMOND, B. *Analysis of Ben-Or's polynomial irreducibility test*. Random Structures and Algorithms, 13 (1998), 439-456.
- [43] PANARIO, D. AND RICHMOND, B. *Smallest components in decomposable structures: exp-log class*. Algorithmica, 29 (2001), 205-226.
- [44] PANARIO, D. AND RICHMOND, B. *Exact largest and smallest size of components*. Algorithmica, 31 (2001), 413-432.
- [45] QUISQUATER, J.J. AND DELESCAILLE, J.P. *How easy is collision search! Application to DES*. Lecture Notes in Computer Science, 434 (1989), 429-433.
- [46] SACHKOV, V.N. *Combinatorial methods in discrete mathematics*. Encyclopedia of Mathematics and its Applications, 55 (1996).
- [47] SACHKOV, V.N. *Verojatnosnie metody v kombinatornom analize*. Nauka, Moscow (1978).

- [48] SHEPP, L. AND LLOYD, S. *Ordered cycle lengths in a random permutation*. Trans. Amer. Math. Soc., 121 (1966), 340-357.
- [49] SORIA-COUSINEAU, M. *Méthodes d'analyse pour les constructions combinatoires et les algorithmes*. Doctorat ès Sciences, Université de Paris-Sud, Orsay, July (1990).
- [50] STANLEY, R.P. '*Generating functions*' in *Studies in Combinatorics*. The Mathematical Association of America (1978), 100-141.
- [51] WHITTAKER, E.T AND WATSON, G.N. *A Course of Modern Analysis*. Cambridge University Press. 4th ed. (1927).
- [52] WILF, H. *Generatingfunctionology*. A K Peters, Ltd. (2005).
- [53] WONG, R. *Asymptotic Approximations of Integrals*. Academic Press (1989).
- [54] ZAVE, D.A. *A series expansion involving the harmonic numbers*. Inform. Process. Lett., 5 (1976), 75-77.