國立臺灣師範大學數學系碩士班碩士論文

指導教授：林延輯 博士

Enumeration and Asymptotics on Restricted Growth Functions of Order 2

研 究 生： 陳怡廷

中 華 民 國 一 〇 四 年 六 月

致謝

回顧這一年的生活—重頭打組合學的基礎、修碩班一半以上的課、考過法文B2檢定、寫論文、安排碩班畢業旅行…，覺得自己很幸運，因為很多人的幫忙才能在迷惘中站穩腳步，並順利克服許多困難。

特別要感謝的是指導教授林延輯老師，在去年我臨時拜託指導時，重新陪我建立組合學和相關背景知識的基礎。我是起步要多花點時間的人，老師仍很有耐心地講解基本的概念。其實從大學時期上老師的課就獲益良多，在老師的講解下，很難的理論變得平易近人些，讓我不至於半途而廢。這一年的相處，老師總是不藏私地和我分享他的學習，讓我了解一個研究如何成形，這些對我來說是很重要的經驗。

另外，謝謝口試委員游森棚老師和陳宏賓老師對論文提出的建議，使得論文的內容更加完整。你們對論文的肯定對我而言是很大的鼓勵。

也謝謝組合學課堂的所有老師和同學，從大家的身上學到如何去表達和討論抽象的數學概念，你們對於數學問題的好奇心和做研究的態度是我的楷模。

謝謝惠伶、信傑、嘉惠和張薰文老師口試當天的幫忙及建議，以及同學們幫我聽口試的練習，讓我能專注於口試簡報的準備。

謝謝陳家姍夫婦在我不相信自己時，跟我討論未來的方向；在我退縮時給予支持，你們的熱心付出，讓我十分感動。

還要感謝媽媽在我陷入自己的困境時的陪伴，接納我因為壓力和不確定感而生的情緒，以及這一年我無暇關心其他事務的包容。也謝謝你總是讓我任性地做自己想做的事，並給予無條件的支持。

最後，我想感謝身邊所有勇於追求夢想的人，看著彼此朝著目標前進令我感到激勵，希望這個正面能量可以引領我們克服對未知的恐懼。

怡廷
2015.8

摘要

本篇論文中，我們延伸限制成長函數到更高次，並找到二次限制成長函數和Ｂ型對稱分割的一對一對應關係。為了改善透過傳統方法得到的漸進結果，我們介紹一個類似牛頓法的演算法。假設二次限制成長函數為均勻分佈，我們得到二次限制成長函數最大值的期望值和變異數的漸進公式。最後，我們驗證二次限制成長函數最大值的分佈收斂到常態分佈。

關鍵字：近似常態性、Hayman admissible 函數、機率分佈、限制成長函數、鞍點法

Department of Mathematics
National Taiwan Normal University
Master's Thesis

Enumeration and Asymptotics on Restricted Growth Functions of Order 2

Yi-Ting Chen

Advisor: Dr. Yen-Chi Lin

Taipei, Taiwan

June, 2015

Contents

Enumeration and Asymptotics on Restricted Growth Functions of Order 2

Yi-Ting Chen

Abstract

In this thesis, we extend the restricted growth functions to higher order and find a bijection between restricted growth functions of order 2 and symmetric partitions of type B. To improve the asymptotic results via traditional methods, we introduce an algorithm which is similar to Newton-Raphson method. Assuming that the restricted growth functions of order 2 are uniformly distributed, we obtain the asymptotic formulae for the expectation and variance of the maximum in a random restricted growth function of order 2. Finally, we verify that the distribution of maximum in restricted growth functions of order 2 will converge to a normal distribution.

Keywords. asymptotic normality, Hayman admissible functions, probability distribution, restricted growth functions, saddle-point method

1 Introduction

A partition of $[n]$ can be encoded as a word of length n with the property that for any $i < j$, the first occurrence of i precedes the first occurrence of j. These words are called *restricted growth functions*, see [12] and Definition 1.

We are interested in asymptotic analysis of this combinatorial structure. The basic asymptotic methods can be found in [5, 13]. These two books introduced several approaches to get different approximations of Bell numbers B_n. One of the classic methods is the saddle-point method, which gives an asymptotic expression :

$$\log \frac{B_n}{n!} \sim n \left[-\log\log n + \frac{\log\log n + 1}{\log n} + \mathcal{O}\left(\frac{\log\log^2 n}{\log^2 n} \right) \right].$$

In this article, we extend the restricted growth functions to higher order and give a bijection between restricted growth functions of order 2 and symmetric partitions of type B. A symmetric partition of type B is defined to be a partition π of the set

$$\{1, 2, \ldots, n, \overline{1}, \overline{2}, \ldots, \overline{n}\}$$

such that for any block B of π, $\overline{B}$ is also a block of π. It's a bit different from the ordinary set partitions of type B.

By reinterpreting the partition of $[n]$ as the hyperplane arrangement for the type A root system, Reiner [14] introduced a natural set partition of type B, which was defined to be a symmetric partition π of the set $\{1, 2, \ldots, n, \overline{1}, \overline{2}, \ldots, \overline{n}\}$ such that there is at most one block B_0 containing both i and $\overline{i}$ for some i (sequence A007405 in OEIS). Wang [17] generalized set partitions of type B by coloring its elements. He also found an asymptotic expression of the total number of colored set partitions of type B.

In 1967, Harper [9] first proved the blocks of set partitions (Stirling numbers of second kind) are asymptotically normal. For a general method to establish the asymptotic normality of some limiting distribution, see Bender [1]. The asymptotic behavior and the results of limiting distribution of set partitions can be found in [2, 7, 11, 15]. The expectation and variance of number of blocks in a random partition of $[n]$ are

$$\mu_n \sim \frac{n}{\log n}$$

and

$$\sigma_n^2 \sim \frac{n}{\log^2 n}.$$

Wang [17] proved that the number of non-zero-blocks in colored set partitions of type B is asymptotically normal. In 2005, Drmota, Gittenberger and Klausner [6] extended the Hayman-admissible functions to bivariate generating functions which satisfy the central limit theorem.

The traditional method applied to set partitions gives an asymptotic expansion of the coefficients for the generating functions in terms of n. However, the formula is very inaccurate when an explicit number n is plugged in. Follow Salvy's idea [16], we focus on expansions in a finer scale. We adopt the Lambert W function [3] to get the expansions of the saddle-point roots. The merit of using Lambert W function is that it's quite closer to our saddle-points. We develop an algorithm which is similar to the Newton-Raphson method. The intention of this algorithm is to approximate by tangent line and to extract the leading term of the difference between Lambert W function and saddle-points step by step.

Our main results are based on uniform distribution of restricted growth functions of order 2. The largest letter in a random restricted growth functions of order 2 has the expectation

$$\mu_n \sim \frac{2n}{\log n}$$

and the variance

$$\sigma_n^2 \sim \frac{4n}{\log^2 n}.$$

There are series of studies on geometric distribution of restricted growth functions. In 2013, Fuchs and Prodinger [8] extended the discussion to the generalized restricted growth property which is the same as our definition of restricted growth functions of higher order. They used Mellin transform and depoissonization to get some asymptotic results.

The organization of this article is as follows. In section 2, we give the definition of restricted growth functions of order 2 and the bijection between the symmetric partitions. At the end of this section, we derive the generating function. In section 3, we apply Hayman's method to get the asymptotic approximation of total number. We also introduced our algorithm to obtain the expansion of the saddle point root in a finer scale. In section 4, we present the results of expected value and variance of the largest letter in a random restricted growth function of order 2. In section 5, we verify that the distribution of restricted growth functions of order 2 is asymptotically normal. We end with some discussions in Section 6.

2 Definition

Definition 1. *A* restricted growth function *or RGF is a word* $w = w_1w_2 \cdots w_n$ *of positive integers such that* $w_1 = 1$*, and* $w_i \leq \max\{w_1, w_2, \cdots, w_{i-1}\} + 1$*, for* $i \geq 2$*. Let* $RG(n)$ *denote the set of all restricted growth functions of length* n*.*

Definition 2. *A partition of a set* S *is a collection* $\pi = B_1/B_2/\ldots/B_k$ *of subsets which are nonempty, pairwise disjoint, and whose union is* S*. The subets* B_i *are called blocks.*

We use the notation Π_n for the partition of $[n] = \{1, 2, \ldots, n\}$. The total number of Π_n is called the Bell number.

There is a one-to-one correspondence between Π_n and $RG(n)$. Let $\pi = B_1/B_2/\ldots/B_k \in \Pi_n$. Put π in standard form with $1 = \min B_1 < \min B_2 < \cdots < \min B_k$. The associated word $w = w_1w_2 \cdots w_n$ is obtained by

$$w_i = j \text{ if and only if } i \in B_j.$$

For example, 1223124 is the RGF corresponing to the partition $\{1, 5\}, \{2, 3, 6\}, \{4\}, \{7\}$. The number of blocks in Π_n is equivalent to the largest letter occuring in $RG(n)$.

Now, we consider the natural extension of RGF.

Definition 3. *A* restricted growth function of order r *is a word* $w = w_1w_2 \cdots w_n$ *of positive integers such that* $w_1 \leq r$*, and* $w_i \leq \max\{w_1, w_2, \cdots, w_{i-1}\} + r$*, for* $i \geq 2$*. Let* $RG_r(n)$ *denote the set of all restricted growth functions of order* r *of length* n*.*

For example, 12315 is RGF of order 2, but 12153 is not RGF of order 2.

Definition 4. *A partition of* $\{1, 2, \ldots, n, \overline{1}, \overline{2}, \ldots, \overline{n}\}$ *into* $B_1/B_2/\ldots/B_k$ *is symmetric if for each* i*,* $\overline{B_i} = B_j$ *for some* j*.*

The number of symmetric partitions of $\{1, 2, \ldots, n, \overline{1}, \overline{2}, \ldots, \overline{n}\}$ is called the Bell number of type B (sequence A002872 in OEIS).

Theorem 5. *There is a bijection between* $RG_2(n)$ *and symmetric partitions of* $\{1, 2, \ldots, n, \overline{1}, \overline{2}, \ldots, \overline{n}\}$*. The number of blocks in symmetric partitions of* $\{1, 2, \ldots, n, \overline{1}, \overline{2}, \ldots, \overline{n}\}$ *equals to the maximum of the corresponding word.*

Proof. We will give a recursive bijection. Let $w = w_1 w_2 \cdots w_n \in RG_2(n)$. If $w_1 = 1$, the partition is $\{1, \overline{1}\}$. If $w_1 = 2$, the partition is $\{1\}, \{\overline{1}\}$.

For $1 < i \leq n$, if $w_i \leq \max\{w_1, w_2, \cdots, w_{i-1}\}$, insert i to the w_i-th block and put $\overline{i}$ in the unique block, so that the partition is symmetric. If $w_i = \max\{w_1, \cdots, w_{i-1}\} + 1$, append a new block and put i and $\overline{i}$ together. If $w_i = \max\{w_1, \cdots, w_{i-1}\} + 2$, append two new blocks and put i and $\overline{i}$ separately.

We always write the partition in standard form with the order $1 < \overline{1} < 2 < \overline{2} < \cdots < n < \overline{n}$. Repeating this process, we obtain the corresponding symmetric partition. □

For instance, $w = 23552 \in RG_2(n)$

$$\begin{aligned}
&\{1\}, \{\overline{1}\} \\
\to\ &\{1\}, \{\overline{1}\}, \{2, \overline{2}\} \\
\to\ &\{1\}, \{\overline{1}\}, \{2, \overline{2}\}, \{3\}, \{\overline{3}\} \\
\to\ &\{1\}, \{\overline{1}\}, \{2, \overline{2}\}, \{3, \overline{4}\}, \{\overline{3}, 4\} \\
\to\ &\{1, \overline{5}\}, \{\overline{1}, 5\}, \{2, \overline{2}\}, \{3, \overline{4}\}, \{\overline{3}, 4\}
\end{aligned}$$

The observation of the previous theorem is the starting point of this paper. We will focus on restricted growth functions of order 2 since the calculation of generating function is more representative. On the other hand, there is NOT a natural bijection between restricted growth functions of order r and extended symmetric partitions, for $r \geq 3$.

Let $a_{n,k}$ be the number of restricted growth functions $w = w_1 w_2 \cdots w_n$ of order 2 with $\max\limits_{1 \leq i \leq n} w_i = k$. If k appears in the first $n-1$ letters, there are k choices for w_n (from 1 to k). If k does not appear in the first $n-1$ letters, w_n must be k and the maximum of the first $n-1$ letters must be $k-1$ or $k-2$. Then we can deduce the recurrence relation :

$$a_{n,k} = k \cdot a_{n-1,k} + a_{n-1,k-1} + a_{n-1,k-2}\ , \quad \text{for } n \geq 1, k \geq 0$$

Now define $A(z,u) = \sum\limits_{n \geq 0} \sum\limits_{k \geq 0} a_{n,k} \frac{z^n}{n!} u^k$ to be the bivariate exponential generating function.

$$\sum_{n \geq 0} \sum_{k \geq 0} a_{n+1,k} \frac{z^n}{n!} u^k = \sum_{n \geq 0} \sum_{k \geq 0} (k \cdot a_{n,k} + a_{n,k-1} + a_{n,k-2}) \frac{z^n}{n!} u^k$$

$$\frac{\partial}{\partial z} A(z,u) = u \cdot \frac{\partial}{\partial u} A(z,u) + u \cdot A(z,u) + u^2 \cdot A(z,u)$$

Solve the partial derivative equation, we get the generating function

$$A(z,u) = \exp\left\{u(e^z - 1) + \frac{1}{2} u^2 (e^{2z} - 1)\right\} \tag{1}$$

3 The Asymptotic Estimation of the Number of Restricted Growth Functions of Order 2

Let $a_n = \sum_k a_{n,k}$ be the total number of restricted growth functions of order 2 of length n.

$$\begin{aligned} a_n &= n![z^n]A(z,1) \\ &= n!e^{-\frac{3}{2}}[z^n]\exp\{e^z + \frac{1}{2}e^{2z}\} \end{aligned} \tag{2}$$

Since $A(z,1)$ is an entire function, we estimate the coefficients a_n by the saddle-point method (or precisely Hayman's method). The saddle-point method is a powerful method for obtaining asymptotic information about fast-growing generating functions. However, it is often cumbersome to apply step by step. In 1956, Hayman [10] defined a class of admissible functions. With this concept, it is easier to verify when the saddle-point method is applicable.

Definition 6. *(Hayman-admissibility, [7, p.565])*
Let $f(z)$ be a power series with real coefficients and radius of convergence $\rho \in (0,\infty]$. Assume that $f(r)$ is positive for $r \in (R_0, \rho)$, for some $R_0 \in (0,\rho)$. Let

$$a(r) = r\frac{f'(r)}{f(r)} \quad \text{and} \quad b(r) = r\frac{f'(r)}{f(r)} + r^2\frac{f''(r)}{f(r)} - r^2\left(\frac{f'(r)}{f(r)}\right)^2.$$

The function $f(z)$ is said to be Hayman admissible (or H-admissible) if the following three conditions hold.

$\mathbf{H_1}$. [Capture condition] $\lim_{r\to\rho} a(r) = +\infty$ and $\lim_{r\to\rho} b(r) = +\infty$.

$\mathbf{H_2}$. [Locality condition] For some function $\theta_0(r)$ defined over (R_0,ρ) and satisfying $0 < \theta_0 < \pi$, one has

$$f(re^{i\theta}) \sim f(r)e^{i\theta a(r) - \frac{1}{2}\theta^2 b(r)} \qquad \text{as } r \to \rho,$$

uniformly in $|\theta| \le \theta_0(r)$.

$\mathbf{H_3}$. [Decay condition] Uniformly in $\theta_0(r) \le |\theta| \le \pi$

$$f(re^{i\theta}) = o\left(\frac{f(r)}{\sqrt{b(r)}}\right).$$

Theorem 7. *(Hayman's method, [7, p.565])*
Let $f(z) = \sum a_n z^n$ *be a Hayman-admissible function and* $r \equiv r_n$ *be the positive real root of the equation*

$$a(r) = r\frac{f'(r)}{f(r)} = n.$$

Then the Taylor coefficients of $f(z)$ *satisfy, as* $n \to \infty$ *:*

$$a_n \equiv [z^n]f(z) \sim \frac{f(r)}{r^n\sqrt{2\pi b(r)}}$$

where $b(r) := ra'(r)$.

The H-admissible functions have some useful closure properties (see [7, p. 568]). In equation (2), $f(z) = \exp\{e^z + \frac{1}{2}e^{2z}\}$. Let $P(z) = z + \frac{1}{2}z^2$. Since $P(z)$ is a polynomial with positive coefficients and e^z is H-admissible, $P(e^z) = e^z + \frac{1}{2}e^{2z}$ is H-admissible. Moreover, the exponential of an H-admissible function is H-admissible. We can apply Hayman's method to $f(z)$.

First, solve the equation

$$r(e^{2r} + e^r) = n. \tag{3}$$

We can only get the asymptotic solution of (3).

Consider the Lambert W function $\zeta = \frac{1}{2}W(2n)$ defined by $\zeta e^{2\zeta} = n$ (see [3]). The asymptotic expansion of ζ is

$$\zeta = \frac{1}{2}\log n - \frac{1}{2}\log\log n + \frac{1}{2}\log 2 + \frac{1}{2}\frac{\log\log n}{\log n} - \frac{1}{2}\frac{\log 2}{\log n} + \mathcal{O}\left(\frac{\log^2\log n}{\log^2 n}\right). \tag{4}$$

Using the traditional iteration method, the asymptotic expansion for the solution r to (3) is the same as ζ. Although e^{2r} is much larger than e^r, we cannot neglect the small difference between r and ζ.

Follow Salvy's idea (see [16]), we want to compute expansions of (3) in a finer scale to extract more information in the leading term. Our algorithm is similar to the Newton-Raphson method. The following is the demonstration of Maple script.

```
> restart:
> with(MultiSeries):
> with(gfun):
> f := x -> x*(exp(2*x)+exp(x)) - a*exp(2*a):
> asympt(f(a), a, 10);
> b := a - a*exp(a)/((2*a+1)*exp(2*a));
> asympt(f(b), a, 10);
> L := [-1/4, 0, 1/16, -1/16, 3/64, -1/32, 5/256];
> y := factor(subs(x = 1/a, listtoratpoly(L, x)[1]));
> c := b-(-(a+1)*a/(2*a+1)^2)*(1/((2*a+1)*exp(2*a)));
```

In the script, consider $f(x) = x(e^{2x} + e^x) - n = x(e^{2x} + e^x) - ae^{2a}$ where $a = \frac{1}{2}W(2n)$. Begin with $x_0 = a$ for a root of $f(x) = 0$. We approximate the function by its tangent line. Calculate recursively $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ where $f'(x_n) \sim (2a+1)e^{2a}$. Then we have $x_1 = b = a - \frac{a}{2a+1}e^{-a}$. However, when moving on to the next iteration, the computation load is too big. Hence we do an asymptotic expansion of $f(x_1)$. But the result is still too slow to get important information. With the aid of gfun package in Maple, we write the coefficients of expansion into a list and find a rational approximation of $f(x_1)$. Continuing this process, we can get as many terms of expansion as desired.

According to the algorithm above,

$$r = \zeta - \frac{\zeta}{2\zeta+1}e^{-\zeta} + \frac{\zeta(\zeta+1)}{(2\zeta+1)^3}e^{-2\zeta} + \frac{\zeta(4\zeta^4+8\zeta^3+3\zeta^2-6\zeta-6)}{6(2\zeta+1)^5}e^{-3\zeta} - \frac{\zeta(8\zeta^5+22\zeta^4+24\zeta^3+12\zeta^2-3)}{3(2\zeta+1)^7}e^{-4\zeta} + \mathcal{O}(e^{-5\zeta}).$$

By substituting the expansion of r into the n th coefficient formula, we can obtain the formula of total number of $RG_2(n)$ in terms of ζ.

Theorem 8. *The number a_n of $RG_2(n)$ satisfies*

$$\log\frac{a_n}{n!} = e^{2\zeta}\left(\frac{1}{2} - \zeta\log\zeta\right) + e^{\zeta} - \zeta - \frac{1}{2}\log(2\zeta^2+\zeta) - \frac{7\zeta+3}{4\zeta+2} - \frac{1}{2}\log 2\pi + \mathcal{O}(e^{-\zeta}) \tag{5}$$

where $\zeta = \frac{1}{2}W(2n)$.

The explicit representation of n is derived from substituting the asymptotic approximation of ζ in terms of n (which is (4)) into (5).

$$\log \frac{a_n}{n!} = n \left[\log 2 - \log\log n + \frac{\log\log n + 1 - \log 2}{\log n} + \mathcal{O}\left(\frac{\log\log^2 n}{\log^2 n} \right) \right]. \quad (6)$$

In fact, the asymptotic formula (6) only provides the information of the leading term in (5). For accuracy, it is better to use (5) to estimate. Here is a numerical table comparing the asymptotic estimate $a_n^\star$ provided by $n! \cdot \exp\{(5)\}$ to the exact values a_n.

n	10	100	1000	10000
a_n	36738144	$2.08834 \cdot 10^{140}$	$7.20673 \cdot 10^{2181}$	$3.16277 \cdot 10^{30296}$
$a_n^\star$	34306357	$2.00826 \cdot 10^{140}$	$7.08996 \cdot 10^{2181}$	$3.14351 \cdot 10^{30296}$

The error is about 6.6 % for $n = 10$, 3.8 % for $n = 100$, 1.6 % for $n = 1000$ and 0.6 % for $n = 10000$.

In Figure 1, the blue curve is $\log a_n^\star$, and the black curve is $\log a_n$. Two curves nearly overlap. Figure 2 illustrates the relative errors of $\log a_n$.

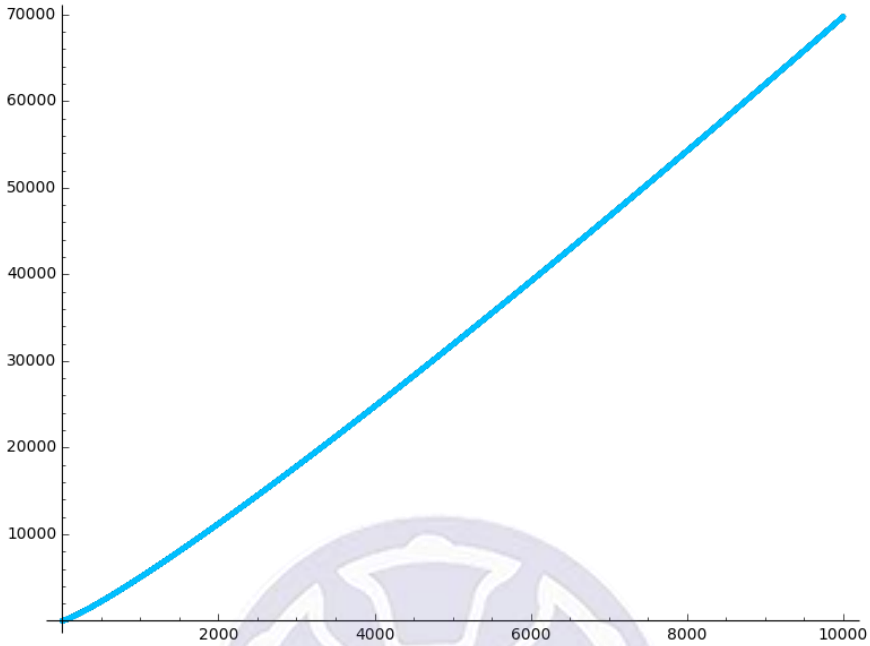

Figure 1: $\log a_n$ and $\log a_n^{\star}$

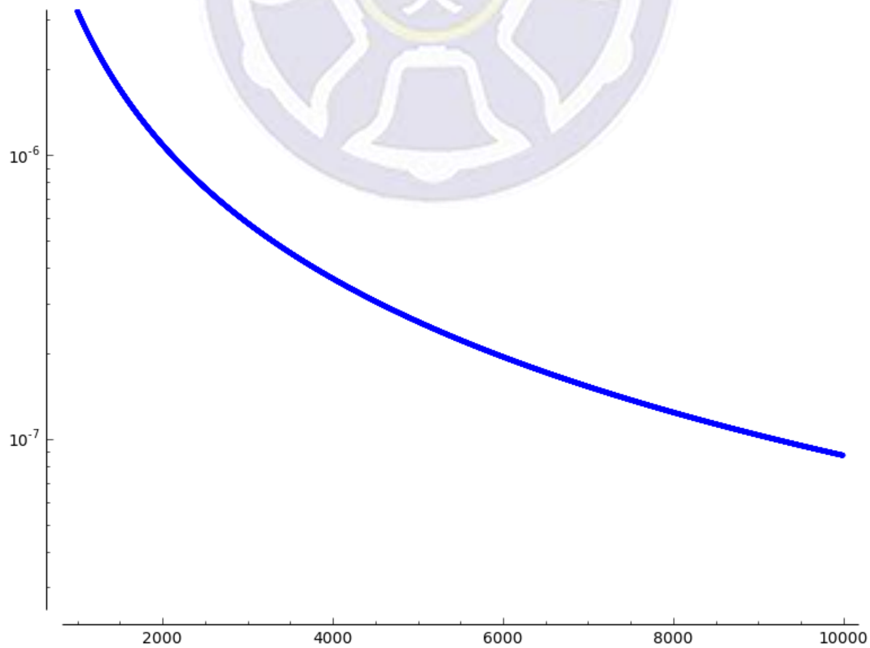

Figure 2: The error of $\frac{\log a_n - \log a_n^{\star}}{\log a_n}$

4 The Asymptotic Estimation of the Expected Value and the Variance of Restricted Growth Functions of Order 2

Assume the restricted growth functions of order 2, $RG_2(n)$, are uniformly distributed. Define the random variable X_n on $RG_2(n)$ by $X_n = k$ if $\max_i w_i = k$, for $w = w_1 w_2 \cdots w_n \in RG_2(n)$. That is, the probability function of random variable X_n is

$$P[X_n = k] = \frac{a_{n,k}}{a_n}$$

where a_n is the number of restricted growth functions of order 2 of length n.

As seen earlier in Section 3, the bivariate generating function of the restricted growth functions of order 2 counted according to the largest letter is

$$A(z,u) = \exp\{u(e^z - 1) + \frac{1}{2}u^2(e^{2z} - 1)\}.$$

Define

$$f_i(z) = \exp\{e^z + \frac{1}{2}e^{2z} + iz\}, \quad \text{for } i = 1, 2, 3, 4.$$

Then the expected value and the variance of X_n are

$$\begin{aligned} \mu_n &= \frac{[z^n]\frac{\partial}{\partial u}A(z,u)|_{u=1}}{[z^n]A(z,1)} \\ &= \frac{[z^n]\exp\{e^z + \frac{1}{2}e^{2z} - \frac{3}{2}\} \cdot (e^{2z} + e^z - 2)}{[z^n]\exp\{e^z + \frac{1}{2}e^{2z} - \frac{3}{2}\}} \\ &= \frac{[z^n]\, f_2(z) + [z^n]\, f_1(z)}{[z^n]\exp\{e^z + \frac{1}{2}e^{2z}\}} - 2 \end{aligned} \tag{7}$$

and

$$\begin{aligned} \sigma_n^2 &= \frac{[z^n]\frac{\partial^2}{\partial u^2}A(z,u)|_{u=1}}{[z^n]A(z,1)} + \frac{[z^n]\frac{\partial}{\partial u}A(z,u)|_{u=1}}{[z^n]A(z,1)} - \left(\frac{[z^n]\frac{\partial}{\partial u}A(z,u)|_{u=1}}{[z^n]A(z,1)}\right)^2 \\ &= \frac{[z^n]\exp\{e^z + \frac{1}{2}e^{2z} - \frac{3}{2}\} \cdot (e^{4z} + 2e^{3z} - 2e^{2z} - 4e^z + 3)}{[z^n]\exp\{e^z + \frac{1}{2}e^{2z} - \frac{3}{2}\}} + \mu_n - \mu_n^2 \\ &= \frac{[z^n]\, f_4(z) + 2[z^n]\, f_3(z) - 2[z^n]\, f_2(z) - 4[z^n]\, f_1(z)}{[z^n]\exp\{e^z + \frac{1}{2}e^{2z}\}} + 3 + \mu_n - \mu_n^2. \end{aligned} \tag{8}$$

To calculate μ_n and σ_n^2, we still need the nth coefficient of $f_i(z)$. Because $\exp\{e^z + \frac{1}{2}e^{2z}\}$ and $\exp\{iz\}$ are H-admissible, $f_i(z)$ are all H-admissible. Apply Theorem 7 again, we have

$$r_i(e^{2r_i} + e^{r_i} + i) = n \ , \quad \text{for } i = 1, 2, 3, 4.$$

When we compute the expected value and variance, the small difference between r_i and ζ which is hidden behind an indefinite cancellation indeed has an impact. In order to distinguish them, we write r_i as the expansion of ζ. Using the same algorithm as Section 3 yields the following results.

$$\begin{aligned} r_1 = \zeta &- \frac{\zeta}{2\zeta+1}e^{-\zeta} - \frac{\zeta^2(4\zeta+3)}{(2\zeta+1)^3}e^{-2\zeta} \\ &- \frac{\zeta(44\zeta^4+16\zeta^3-63\zeta^2-48\zeta+6)}{6(2\zeta+1)^5}e^{-3\zeta} \\ &- \frac{\zeta(96\zeta^6+104\zeta^5-98\zeta^4-168\zeta^3-48\zeta^2+12\zeta+3}{3(2\zeta+1)^7}e^{-4\zeta} + \mathcal{O}(e^{-5\zeta}) \end{aligned}$$

$$\begin{aligned} r_2 = \zeta &- \frac{\zeta}{2\zeta+1}e^{-\zeta} - \frac{\zeta(8\zeta^2+7\zeta+1)}{(2\zeta+1)^3}e^{-2\zeta} \\ &- \frac{\zeta(92\zeta^4+40\zeta^3-123\zeta^2-102\zeta-18)}{6(2\zeta+1)^5}e^{-3\zeta} \\ &- \frac{\zeta(384\zeta^6+584\zeta^5-26\zeta^4-456\zeta^3-240\zeta^2-24\zeta+3)}{3(2\zeta+1)^7}e^{-4\zeta} + \mathcal{O}(e^{-5\zeta}) \end{aligned}$$

$$\begin{aligned} r_3 = \zeta &- \frac{\zeta}{2\zeta+1}e^{-\zeta} - \frac{\zeta(12\zeta^2+11\zeta+2)}{(2\zeta+1)^3}e^{-2\zeta} \\ &- \frac{\zeta(140\zeta^4+64\zeta^3-183\zeta^2-156\zeta-30)}{6(2\zeta+1)^5}e^{-3\zeta} \\ &- \frac{\zeta(864\zeta^6+1448\zeta^5+238\zeta^4-840\zeta^3-564\zeta^2-108\zeta-3)}{3(2\zeta+1)^7}e^{-4\zeta} + \mathcal{O}(e^{-5\zeta}) \end{aligned}$$

$$\begin{aligned} r_4 = \zeta &- \frac{\zeta}{2\zeta+1}e^{-\zeta} - \frac{\zeta(16\zeta^2+15\zeta+3)}{(2\zeta+1)^3}e^{-2\zeta} \\ &- \frac{\zeta(188\zeta^4+88\zeta^3-243\zeta^2-210\zeta-42)}{6(2\zeta+1)^5}e^{-3\zeta} \\ &- \frac{\zeta(96\zeta^6+104\zeta^5-98\zeta^4-168\zeta^3-48\zeta^2+12\zeta+3)}{3(2\zeta+1)^7}e^{-4\zeta} + \mathcal{O}(e^{-5\zeta}) \end{aligned}$$

By substituting the expansions of r and r_i into the n th coefficient formulae (7) and (8), we get

Theorem 9. *The expected value and variance of X_n satisfy:*

$$\mu_n = e^{2\zeta} + \frac{e^{\zeta}}{2\zeta+1} - \frac{16\zeta^3+20\zeta^2+9\zeta+2}{(2\zeta+1)^3} + \mathcal{O}(e^{-\zeta}) \tag{9}$$

$$\sigma_n^2 = \frac{2e^{2\zeta}}{2\zeta+1} + \frac{(6\zeta+1)e^{\zeta}}{(2\zeta+1)^3} - \frac{96\zeta^5+240\zeta^4+224\zeta^3+112\zeta^2+36\zeta+3}{(2\zeta+1)^5} + \mathcal{O}(e^{-\zeta}) \tag{10}$$

where $\zeta = \frac{1}{2}W(2n)$.

The asymptotic formula of μ_n and σ_n^2 in terms of n are obtained by substituting (4) into the leading term of (9) and (10). Recall that the results of ordinary set partitions are $\mu_n \sim \frac{n}{\log n}$ and $\sigma_n^2 \sim \frac{n}{\log^2 n}$. We have the corresponding results of these formulae.

Theorem 10. *The expected value and variance of X_n satisfy:*

$$\mu_n = \frac{2n}{\log n}\left[1 + \frac{\log\log n - \log 2}{\log n} + \mathcal{O}\left(\frac{\log^2\log n}{\log^2 n}\right)\right]$$

$$\sigma_n^2 = \frac{4n}{\log^2 n}\left[1 + \frac{2\log\log n - 2\log 2 - 1}{\log n} + \mathcal{O}\left(\frac{\log^2\log n}{\log^2 n}\right)\right]$$

Here is a numerical table comparing the asymptotic estimate $\mu_n^\star$ and $\sigma_n^{2\star}$provided by (9) and (10) to the exact values μ_n and σ_n^2, respectively.

n	10	100	1000	10000
μ_n	8.23783	50.51227	343.49702	2553.60995
$\mu_n^\star$	8.25811	50.51915	343.49912	2553.61057
σ_n^2	3.41788	18.45577	98.34212	575.44135
$\sigma_n^{2\star}$	3.42929	18.45750	98.34242	575.44141

The error of μ_n is about 0.25 % for $n = 10$ and less than 10^{-4} for $n > 100$. The error of σ_n^2 is about 0.33 % for $n = 10$ and less than 10^{-4} for $n > 100$.

Figure 3 illustrates the expected value and variance of X_n from $n = 1$ to $n = 10000$. The blue curve is $\mu_n^\star$, and the black curve is μ_n. The yellow curve is $\sigma_n^{2\star}$, and the red curve is σ_n^2. Each pairs of curves nearly overlap. In Figure 4, the blue dots are the relative errors of μ_n, and the red dots are the relative errors of σ_n^2.

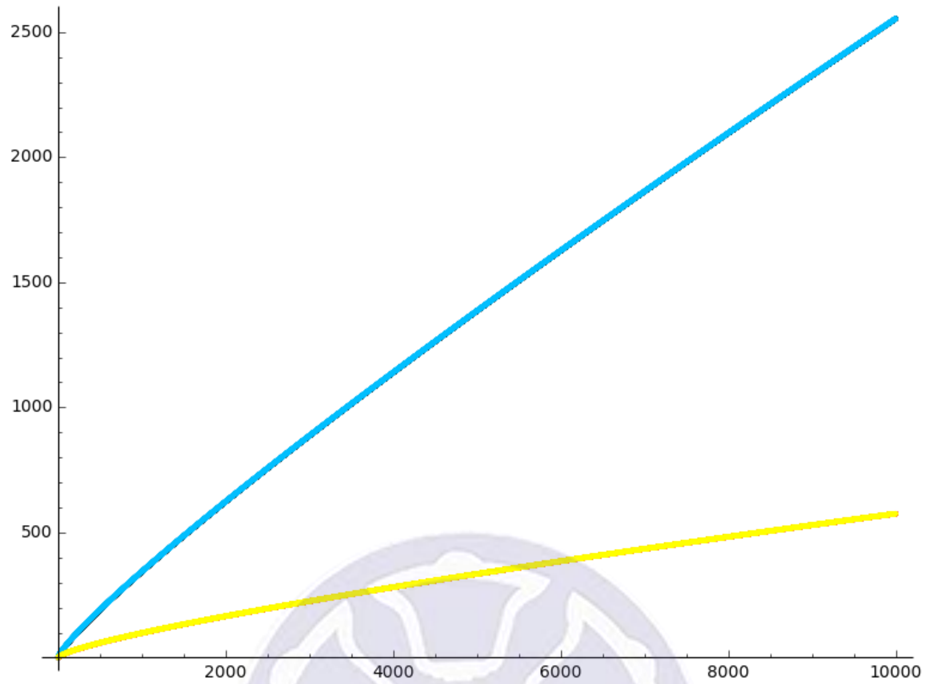

Figure 3: The exact value and estimate of expected value and variance of X_n

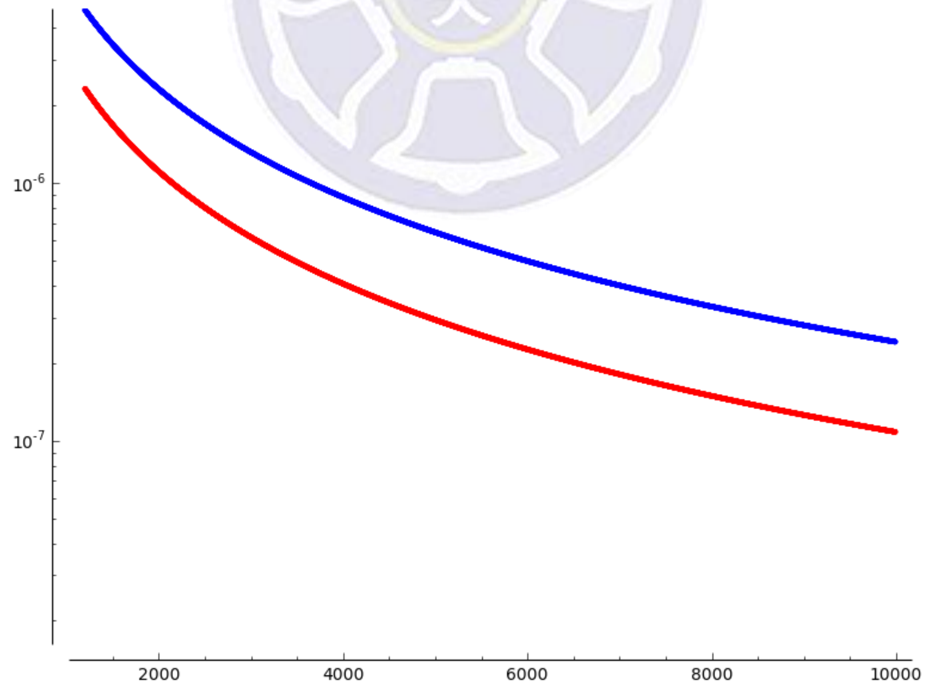

Figure 4: The error of the expected value and variance of X_n

5 The Asymptotic Normality of Restricted Growth Functions of Order 2

In order to prove the central limit theorem directly by looking at the bivariate generating function, Drmota et al. [6] extended Hayman admissibility to Definition 11 which also has strong closure properties.

Definition 11. *(Extended-admissibility, [6])*
A function $f(z,u) = \sum_{n,k\geq 0} a_{n,k}\, z^n u^k$ *is called extended admissible or e-admissible if there exists* $0 < \rho \leq \infty$ *such that the following conditions are satisfied:*

(1) $f(z,u)$ *is analytic in* $\Delta_{\rho,\xi} := \{(z,u) : |z| < R, |u| < 1+\xi\}$*, where* $\xi > 0$*, and for some* $R_0 < \rho$ *we have*

$$f(r,1) > 0, \quad R_0 < r < \rho.$$

(2) Set

$$a(z,u) = z\frac{f_z(z,u)}{f(z,u)}, \qquad \overline{a}(z,u) = u\frac{f_u(z,u)}{f(z,u)},$$
$$b(z,u) = za_z(z,u) = z\frac{f_z(z,u)}{f(z,u)} + z^2\frac{f_{zz}(z,u)}{f(z,u)} - z^2\left(\frac{f_z(z,u)}{f(z,u)}\right)^2,$$
$$\overline{b}(z,u) = u\overline{a}_u(z,u), \qquad c(z,u) = ua_u(z,u),$$

and

$$\varepsilon(r) = K\left(\overline{b}(r,1) - \frac{c(r,1)^2}{b(r,1)}\right)^{-\frac{1}{2}},$$

where $K > 0$ *is an arbitrary constant. Then, for each choice of* $K > 0$ *there exists a function* $\delta(r) : (R_0,\rho) \to (0,\pi)$ *such that for* $R_0 < r < \rho$ *we have*

$$f(re^{i\theta},u) \sim f(r,u)\exp\left(i\theta a(r,u) - \frac{\theta^2}{2}b(r,u)\right), \text{ as } r \to \rho,$$

uniformly for $|\theta| \leq \delta(r)$ *and* $u \in [1-\varepsilon(r), 1+\varepsilon(r)]$.

(3) For $R_0 < r < \rho$*, we have*

$$f(re^{i\theta},u) = o\left(\frac{f(r,u)}{\sqrt{b(r,u)}}\right), \quad \text{as } r \to \rho,$$

uniformly for $\delta(r) \leq |\theta| \leq \pi$ *and* $u \in [1-\varepsilon(r), 1+\varepsilon(r)]$.

(4) For $r \to \rho$, we have $b(r,1) \to \infty$.

(5) $b(r,u) \sim b(r,1)$ for $r \to \rho$, uniformly for $u \in [1-\varepsilon(r), 1+\varepsilon(r)]$.

(6) $a(r,u) = a(r,1) + c(r,1)(u-1) + \mathcal{O}(c(r,1)(u-1)^2)$ for $r \in (R_0, \rho)$ and $u \in [1-\varepsilon(r), 1+\varepsilon(r)]$.

(7) $\overline{a}(r,u) = \mathcal{O}(\overline{a}(r,1))$ and $\overline{b}(r,u) = \mathcal{O}(\overline{b}(r,1))$ for all u in an arbitrary but fixed complex neighborhood of 1 and all r.

(8) $\overline{b}(r,1) - \frac{c(r,1)^2}{b(r,1)} \to \infty$ as $r \to \rho$.

(9) $\varepsilon(r)^3 \overline{b}(r,1) \to 0$ for $r \to \rho$.

(10) $\overline{b}(r,1) = \mathcal{O}(\overline{a}(r,1)^2)$ and $\overline{a}(r,1) = \mathcal{O}(f(r,1)^\eta)$ for every $\eta > 0$

Theorem 12. *([6]) Let $f(z,u)$ be an e-admissible function such that for sufficiently large n all coefficients $a_{n,k}$ are nonnegative. X_n denotes a sequence of random variables satisfying*

$$P[X_n = k] = \frac{a_{n,k}}{a_n}.$$

Then the following central limit theorem holds:

$$\frac{X_n - \overline{a}(r_n,1)}{\sqrt{|B(r_n,1)|/b(r_n,1)}} \to \mathcal{N}(0,1),$$

where r_n is the positive solution of $a(r,1) = n$ and $|B|$ is the determinant of B with

$$B(r,u) = \begin{pmatrix} b(r,u) & c(r,u) \\ c(r,u) & \overline{b}(r,u) \end{pmatrix}.$$

Furthermore we have, as $n \to \infty$,

$$\mu_n = \overline{a}(r_n,1) + o(|B(r_n,1)|/b(r_n,1))$$

and

$$\sigma_n^2 \sim \frac{|B(r_n,1)|}{b(r_n,1)}.$$

First, we claim our generating function is e-admissible. We need the next theorem.

Theorem 13. *Let $f(z)$ be an H-admissible function and satisfy*

$$f(r) - \frac{(rf'(r))^2}{rf'(r) + r^2 f''(r)} \to \infty,$$
$$f(r) \cdot \left(f(r) - \frac{(rf'(r))^2}{rf'(r) + r^2 f''(r)} \right)^{-\frac{3}{2}} \to 0, \quad \text{as } r \to \rho,$$

then $e^{uf(z)}$ and $e^{\frac{1}{2}u^2 f(z)}$ are extended-admissible.

Proof. Let $f(z,u) = e^{g(u)f(z)}$. We have

$$\begin{aligned} a(r,u) &= rg(u)f'(r) & \overline{a}(r,u) &= ug'(u)f(r) \\ b(r,u) &= (rf'(r) + r^2 f''(r))g(u) & \overline{b}(r,u) &= (ug'(u) + ug''(u))f(r) \\ c(r,u) &= rug'(u)f'(r) \end{aligned}$$

$$\overline{b}(r,1) - \frac{c(r,1)^2}{b(r,1)} = (g'(1) + g''(1))f(r) - \frac{r^2 g'(1)^2 f'(r)^2}{(rf'(r) + r^2 f''(r))g(1)}.$$

Conditions (1)−(7) and (10) of e-admissibility follow from the proof of Theorem 3 in [6]. We only check conditions (8) and (9).

If $g(u) = u$, then $\overline{b}(r,1) - \frac{c(r,1)^2}{b(r,1)} = f(r) - \frac{(rf'(r))^2}{rf'(r)+r^2 f''(r)} \to \infty$ and

$$\varepsilon(r)^3 \overline{b}(r,1) = K^3 \left(f(r) - \frac{(rf'(r))^2}{rf'(r)+r^2 f''(r)} \right)^{-\frac{3}{2}} \cdot f(r) \to 0$$

where $K > 0$ is an arbitrary constant.
If $g(u) = \frac{1}{2}u^2$, then $\overline{b}(r,1) - \frac{c(r,1)^2}{b(r,1)} = 2 \left(f(r) - \frac{(rf'(r))^2}{rf'(r)+r^2 f''(r)} \right) \to \infty$ and

$$\varepsilon(r)^3 \overline{b}(r,1) = K^3 \left(f(r) - \frac{(rf'(r))^2}{rf'(r)+r^2 f''(r)} \right)^{-\frac{3}{2}} \cdot \frac{1}{\sqrt{2}} f(r) \to 0$$

where $K > 0$ is an arbitrary constant. □

In our example, we have to verify $f(z) = e^z - 1$ and $f(z) = e^{2z} - 1$. As $r \to \rho$,

$$\begin{aligned} f(r) - \frac{(rf'(r))^2}{rf'(r) + r^2 f''(r)} &= e^r - 1 - \frac{r^2 e^{2r}}{re^r + r^2 e^r} \\ &= \frac{e^r}{r+1} - 1 \to \infty \end{aligned}$$

$$f(r) \cdot \left(f(r) - \frac{(rf'(r))^2}{rf'(r) + r^2 f''(r)} \right)^{-\frac{3}{2}} = (e^r - 1) \cdot \left(\frac{e^r}{r+1} - 1 \right)^{-\frac{3}{2}} \to 0$$

and

$$f(r) - \frac{(rf'(r))^2}{rf'(r) + r^2 f''(r)} = e^{2r} - 1 - \frac{4r^2 e^{4r}}{2re^{2r} + 4r^2 e^{2r}}$$

$$= \frac{e^{2r}}{2r+1} - 1 \ \to \infty$$

$$f(r) \cdot \left(f(r) - \frac{(rf'(r))^2}{rf'(r) + r^2 f''(r)} \right)^{-\frac{3}{2}} = (e^{2r} - 1) \cdot \left(\frac{e^{2r}}{2r+1} - 1 \right)^{-\frac{3}{2}} \to 0.$$

By Theorem 13, $\exp\{u(e^z - 1)\}$ and $\exp\{\frac{1}{2}u^2(e^{2z} - 1)\}$ are e-admissible. Since the product of two e-admissible functions is e-admissible (see [6]), $\exp\{u(e^z - 1) + \frac{1}{2}u^2(e^{2z} - 1)\}$ is e-admissible. Applying Theorem 12, we have proven

Theorem 14. *The distribution of restricted growth functions of order 2 is asymptotically normal, with mean and variance that satisfy* $\mu_n \sim \frac{2n}{\log n}$ *and* $\sigma_n^2 \sim \frac{4n}{\log^2 n}$.

Figure 5 supports our result. The blue points are $P(X_n = k)$, and the red curve is normal distribution with $n = 10000$.

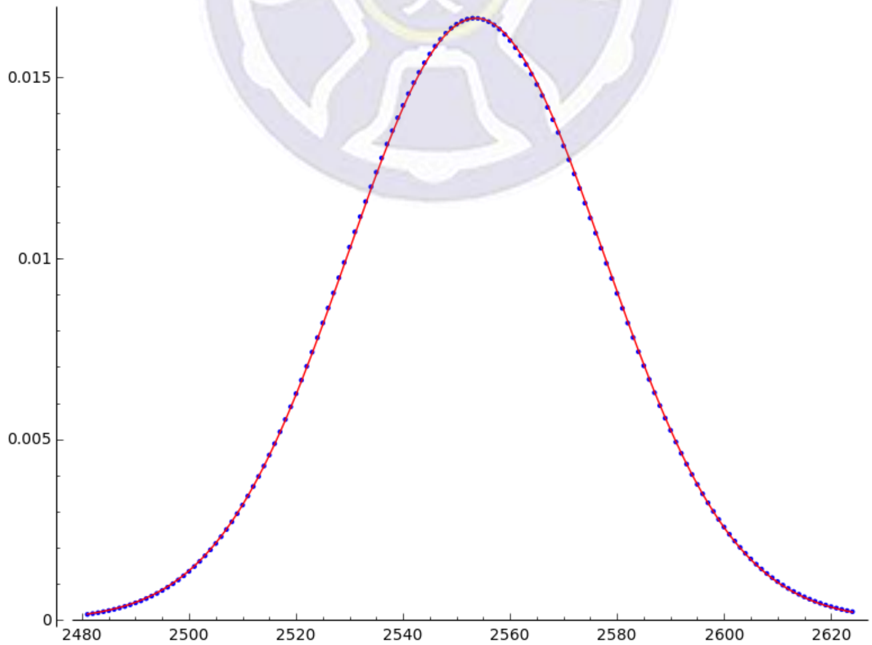

Figure 5: The distribution of X_n with $n = 10000$

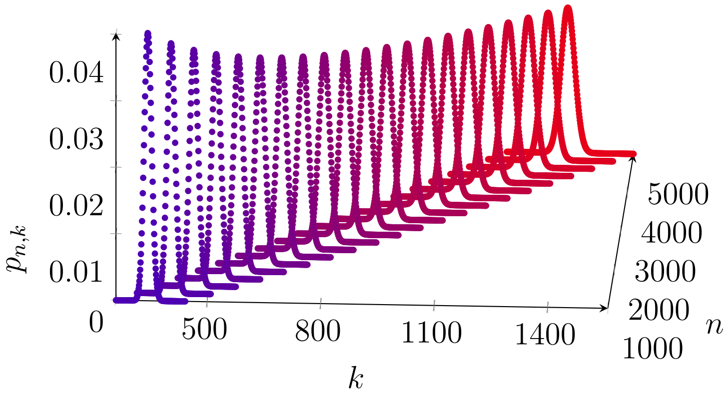

Figure 6: The distribution of X_n from $n = 1000$ to $n = 5000$

6 Conclusion and Future Work

In this thesis, we connected symmetric partitions of type B to restricted growth functions of order 2. For ordinary set partitions, one of natural definitions of pattern avoidance is derived from considering them as restricted growth functions (see [4, 11]). Hence, it's a good direction to explore the patterns and statistics for symmetric partitions of type B.

Furthermore, we developed an algorithm to obtain the expansions of the saddle point root in terms of Lambert W function. With this algorithm, we are able to easily get the asymptotic approximations of the expected value and variance of symmetric partitions of type B. We believe that it's applicable for set partitions and (colored) set partitions of type B. Also, future research could use our algorithm to solve the saddle point equation which has the relative small and non-negligible term and then improve the asymptotic estimates.

References

[1] E. A. Bender. Central and local limit theorems applied to asymptotic enumeration. *J. Combinatorial Theory Ser. A*, 15:91–111, 1973.

[2] E. A. Bender and L. B. Richmond. Admissible functions and asymptotics for labelled structures by number of components. *Electron. J. Combin.*, 3(1):Research Paper 34, approx. 23 pp. (electronic), 1996.

[3] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Adv. Comput. Math.*, 5(4):329–359, 1996.

[4] Samantha Dahlberg, Robert Dorward, Jonathan Gerhard, Thomas Grubb, Carlin Purcell, Lindsey Reppuhn, and Bruce E. Sagan. Set partition patterns and statistics, 2015.

[5] N. G. de Bruijn. *Asymptotic methods in analysis.* Dover Publications, Inc., New York, third edition, 1981.

[6] M. Drmota, B. Gittenberger, and T. Klausner. Extended admissible functions and gaussian limiting distributions. *Math. Comp.*, 74(252):1953–1966, 2005.

[7] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.

[8] Michael Fuchs and Helmut Prodinger. Words with a generalized restricted growth property. *Indag. Math. (N.S.)*, 24(4):1024–1033, 2013.

[9] L. H. Harper. Stirling behavior is asymptotically normal. *Ann. Math. Statist.*, 38:410–414, 1967.

[10] W. K. Hayman. A generalisation of Stirling's formula. *J. Reine Angew. Math.*, 196:67–95, 1956.

[11] T. Mansour. *Combinatorics of set partitions.* Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, 2013.

[12] S. Milne. Restricted growth functions and incidence relations of the lattice of partitions of an n-set. *Advances in Math.*, 26(3):290–305, 1977.

[13] A. M. Odlyzko. Asymptotic enumeration methods. In *Handbook of combinatorics, Vol. 1, 2*, pages 1063–1229. Elsevier, Amsterdam, 1995.

[14] V. Reiner. Non-crossing partitions for classical reflection groups. *Discrete Math.*, 177(1-3):195–222, 1997.

[15] V. N. Sachkov. *Probabilistic methods in combinatorial analysis*, volume 56 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1997.

[16] B. Salvy and J. Shackell. Symbolic asymptotics: Multiseries of inverse functions. *J. Symbolic Computation*, 27(6):543–563, 1999.

[17] David G. L. Wang. On colored set partitions of type B_n. *Cent. Eur. J. Math.*, 12(9):1372–1381, 2014.