國立交通大學

應用數學系

博士論文

隨機數位樹上加法性參數之機率分析

Probabilistic Analysis of Additive Shape Parameters in Random Digital Trees

研究生: 李忠逵

指導教授: Prof. Michael Fuchs

中華民國一百零三年六月

隨機數位樹上加法性參數之機率分析
Probabilistic Analysis of Additive Shape Parameters in Random Digital Trees

研究生：李忠逵	Student: Chung-Kuei Lee
指導教授：符麥克	Advisor: Michael Fuchs

國立交通大學
應用數學系
博士論文

A Thesis

Submitted to Department of Applied Mathematics

College of Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Applied Mathematics

June 2014

Hsinchu, Taiwan, Republic of China

中華民國一百零三年六月

隨機數位樹上加法性參數之機率分析

研究生: 李忠逵　　　　指導教授: 符麥克

國立交通大學
應用數學系

摘要

自圖靈獎得主高納德 (D. E. Knuth) 於 1963 年首開先河後，演算法分析成為了一個在數學和理論計算機科學中皆具相當重要性的研究領域。本領域的主要目標之一在於取得對演算法之隨機行為的全面了解。除此之外，資料結構的漸進性質也是演算法分析所關注的重要議題。數位樹家族是在計算機和資訊科學中最常被使用的資料結構之一。在本論文中，我們藉由研究數位樹的加法性構型參數探討了此一資料結構的眾多漸進性質。

本論文的第一部份是對數位樹家族進行一個完整的介紹。內容包括了構造方式、我們將使用的隨機模型、已知的研究成果和相關的數學工具。我們也將利用一些例子來解釋這些數學工具在過去的研究中是如何被使用的。

第二部分則是我們新獲得的研究成果，包括了 Poisson-Laplace-Mellin 方法的新應用以及一套可用來證明隨機數位樹加法性構型參數的中央極限定理的理論框架。過去使用 Poisson-Laplace-Mellin 方法所獲得的結果大多數是有關線性成長的構型參數的，我們在本論文中運用 Poisson-Laplace-Mellin 方法對許多非線性成長的構型參數進行了研究並且推導出了這些參數的期望值和變異數的漸進表示。我們還建立了一套可以近乎”自動”證明構型參數的中央極限定理的理論框架 − 只要此構型參數滿足一定條件，則自動滿足一中央極限定理。

Probabilistic Analysis of Additive Shape Parameters in Random Digital Trees

Student: Chung-Kuei Lee　　Advisor: Michael Fuchs

Department of Applied Mathematics
National Chiao Tung University

Abstract

Established by D. E. Knuth in 1963, analysis of algorithms is an important area which lies in the overlap of mathematics and theoretical computer science. The main aim of this area is to understand the stochastic behavior of algorithms. Moreover, another important issue is to obtain asymptotic properties of data structures. In this thesis, we study one of the most often used data structure, the digital tree family, through additive shape parameters.

The first part of this thesis is an introduction to the digital tree family, including the construction, the random model we are going to use and a survey of known results and related mathematical techniques. We also give some examples to illustrate how these mathematical techniques have been utilized in past researches of random digital trees.

The second part contains the new results we derived, including new applications of the Poisson-Laplace-Mellin method and general frameworks for central limit theorems of additive shape parameters in random digital trees. Most of the known applications of the Poisson-Laplace-Mellin method are for shape parameters of linear order. In this thesis, we study many shape parameters which are of sublinear and superlinear order via the Poisson-Laplace-Mellin method and derive asymptotic expressions for the means and variances of these shape parameters. We also establish frameworks which allows us to prove central limit theorems of the shape parameters in an almost automatic fashion.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The word ”algorithm” is a distortion of the name of the 9th century Persian scientist, astronomer and mathematician Abu Abdallah Muhammad ibn Musa al-Khwarizmi. At al-Khwarizmi’s time, people used the term ”algorism”, which is also a distortion of al-Khwarizmi’s name, to refer to systematic arithmetic methods which used Hindu-Arabic numerals to solve linear and quadratic equations. These methods appeared in a book written by al-Khwarizmi around the year 825 and the book was translated into Latin by the 18th century. The word ”algorithm” came from a Latin translation of al-Khwarizmi’s name.

The modern concept of algorithms was formalized in 1936 by A. Turing’s Turing machines and A. Church’s lambda calculus, which in turn formed the foundation of computer science. Although the formal definition of the term ”algorithm” remains a challenging problem up to date [151], it can be informally defined as a step-by-step computational procedure which takes a set of values as input and produces a set of values as output [25].

Turing machine is the precursor of modern computers. Turing proved that such machines would be capable of performing any conceivable mathematical computation if it can be represented as an algorithm. With the computer rapidly becoming the most important tool in almost every aspects of people’s life, people started seeking more efficient algorithms which make computers solve problems with less resources (time, storage spaces,..., etc.).

There are two different approaches to analyze the efficiency of an algorithm. The first one is to evaluate the performance of an algorithm in the worst-case senario. This kind of studies focus on the growth of resources an algorithm need in the worst-case. One of the major goals of such studies is to find an ”optimal algorithm” whose worst-case performance matches a ”lower bound” in the sense that any other algorithm for the same problem requires at least the same amount of resources in the worst-case scenario.

This approach was popularized by A. V. Aho, J. E. Hopcroft and J. D. Ullman [4] and T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein [25]. P. Flajolet and R. Sedgewick use the term ”theory of algorithms” to refer to this type of studies [193].

The second approach, started by D. E. Knuth, is to study the average-case properties of algorithms. The average-case analysis works by first constructing a suitable mathematical model for the input and then using probabilistic, combinatorial and asymptotic methods to study the expected resource costs of the algorithm given an input drawn from the model. In [193], the authors described the goal of this approach as

> ....to be able to accurately predict the performance characteristics of particular algorithms when run on particular computers, in order to be able to predict resource usage, set parameters, and compare algorithms.

To design efficient algorithms and get a throughout understanding of them, both approaches are necessary. When a new algorithm appears, people need a rough idea of how the algorithm might perform and a rough comparison of this algorithm to other algorithms for the same problem. In such cases, the first approach will be an adequate choice. However, the first approach usually sacrifices many details and hence it provides less precision. Therefore, to acquire more specified information and make more accurate prediction of an algorithm, we need more details of the algorithm and the mathematical properties of the data structures manipulated by the algorithm. In such cases, the second approach plays an important role. Thus, a complete analysis of a specified algorithm should be a study which combines both approaches to get a full understanding of the algorithm. This leads to the emergence of the area ”analysis of algorithms”.

The electronic journal *Discrete Mathematics and Theoretical Computer Science* defines ”analysis of algorithms” as follows:

> Analysis of Algorithms is concerned with accurate estimate of complexity parameters of algorithms and aims at predicting the behavior of a given algorithm run in a given environment. It develops general methods for obtaining closed-form formulae, asymptotic estimates, and probability distributions for combinatorial or probabilistic quantities, that are of interest in the optimization of algorithms. Interest is also placed on the methods themselves, whether combinatorial, probabilistic, or analytic. Combinatorial and statistical properties of discrete structures (strings, trees,

> tries, dags, graphs, and so on) as well as mathematical objects (e.g., continued fractions, polynomials, operators) that are relevant to the design of efficient algorithms are investigated.

According to W. Szpankowski [202], the area of analysis of algorithms was born on July 27, 1963 when D. E. Knuth wrote his ”Notes on Open Addressing” on hashing tables with linear probing. In fact, using the term ”analysis of algorithms” to name the area was proposed by Knuth in his talk ”The Birth of the Giant Component” [63, 105] given during the first *Average Case Analysis of Algorithms Seminar* in 1993. Knuth’s monumental series *The Art of Computer Programming* [125, 126, 127, 128] enriched this area and attracted many researchers devoting their effort to develop it. After 50 years of development and with the contributions from many researchers, analysis of algorithms has became an area with much more abundant content. Nowadays, analysis of algorithms is a field in the overlap of mathematics and theoretical computer science which enjoys close relations to discrete mathematics, combinatorial analysis, probability theory, analytic number theory, asymptotic analysis and complexity theory.

Analysis of algorithms is a fruitful area because of the helps from a wide range of mathematical tools. The development of analysis of algorithms also motivated researches about many different areas in mathematics. The best example might be the emerging of analytic combinatorics.

The major goal of analytic combinatorics is to precisely predict the properties of large structured combinatorial configurations, through an approach based comprehensively on analytic methods. It starts from an exact enumerative characterization of combinatorial structures by means of generating functions. Then, generating functions are viewed as analytical objects which map the complex plane into itself. Finally, techniques from complex analysis are used to derive asymptotic properties of the combinatorial objects from the generating functions.

The basic techniques and ideas of the theory of analytic combinatorics appeared quite early. Since the 18th century, many mathematicians, including L. Euler, A. Cayley, S. Ramanujan, G. Polya, and D. E. Knuth, have made contributions to the theory. However, systematic study of the theory started in 1960s by P. Flajolet because of the need from analysis of algorithms. As we have mentioned, combinatorial and statistical properties of discrete structures are major topics in analysis of algorithms. The need of mathematical tools for such topics fueled the development of analytic combinatorics. In this thesis, we will use techniques from analytic combinatorics to study tree-based data structures.

Trees, defined as acyclic connected graphs, are fundamental objects in graph theory. Moreover, they are also basic objects for data structures and algorithms in computer science. Take binary search trees as an example, which are a variant of the simplest tree structure, the binary trees, in graph theory. Since invented in 1960 in a joint work of A. S. Douglas, P. F. Windley, A. D. Booth, A. J. T. Colin and T. N. Hibbard [13, 45, 88, 216], binary search trees became one of the most widely used data structures in computer science. Consequently, researchers began to study the properties of binary search trees. Started by Knuth [124], many different methods, including classical combinatorics [17], probability theory [34] and analytic combinatorics [46], have been used in related researches.

The purpose of this thesis is to study another widely used tree-based data structure, the digital tree family. There are several outstanding books which introduce some aspects of digital trees [48, 139, 203]. However, many important aspects of digital trees remain unknown and many problems are still unanswered. Our goal is to solve some of these problems and broaden related studies by introducing new ideas and using recently developed methods. For this purpose, we will first review past researches and discuss the advantages and disadvantages of the known mathematical techniques which have been applied in previous studies. In the second half of this thesis, we will state our new results.

The chapters of this thesis are arranged as follows. Chapter 2 is meant to be a survey of known knowledge of digital trees. In this chapter, we first give the formal definition and the construction of digital trees. Then, we are going to explain the random model used in our studies. The final section of Chapter 2 is a collection of the majority of previous researches from the last several decades.

In Chapter 3, we are going to introduce the mathematical techniques which have been used in the study of random digital trees. These methods are from different areas of mathematics, including analytic combinatorics, complex analysis, functional analysis and probability theory. For each method, we will give examples to illustrate how the method was applied to random digital trees.

We are going to show some new applications of the "Poisson-Laplace-Mellin" method and Poissonized variance with correction in Chapter 4. The material of this chapter are from published papers by the author of this thesis and his collaborators. Section 4.1 is based on [78]. In this section, we will derive asymptotic expressions of mean and variance of approximate counting. Section 4.2 gives the expressions of the first and second moment and the limiting distribution of the Wiener index of random digital trees. Results presented in this section are taken from [76, 77]. In Section 4.3, we are going to discuss a generalization of the Wiener index, the total Steiner distance. To analyze this parameter, we will also discuss the k-th total path length. The results in this section are from [130].

The general topic of Chapter 5 is a framework for central limit theorems of certain shape parameters in random digital trees. In the first two sections of this chapter, we state the framework which allows one to prove central limit theorems for shape parameters satisfying certain conditions. In the third section, we introduce a useful lemma which proves lower bounds for the variance of shape parameters and plays a key role in the framework. The rest of this chapter are examples of shape parameters which fit into the framework. These results can be found in [77] and [79].

We conclude the thesis with a summary of the main achievement and some comments in Chapter 6.

Chapter 2

Random Digital Trees

2.1 The Origin and Applications of the Random Digital Trees

Digital trees are an important class of random trees with numerous applications in computer science. In this thesis, we are mainly dealing with four subclasses of digital trees, namely, Tries, PATRICIA Tries, Digital Search Trees (DSTs) and Bucket Digital Search Trees (b-DSTs).

Tries, first introduced by R. de la Briandais [29] and named by E. Fredkin [72], are one of the most widely used data structure in computer science. Tries were created for retrieving data more efficiently. They turn out to be a huge success since they have many advantages over other already-existing data structures such as binary search trees. For example, looking up data in tries is faster in worst case comparing to binary search trees and the collision of different keys is avoided. Nowadays, tries are applied in many areas such as searching, sorting, dynamic hashing coding, polynomial factorizing, regular languages, contention tree algorithms, automatically correcting words in texts, retrieving IP addresses and satellite data, internet routing and molecular biology. For a more detailed introduction and more references, see [167].

Several variants of Tries have been considered. One of them is PATRICIA Tries which were invented by D. R. Morrison [150] in 1986 (PATRICIA is an acronym which stands for Practical Algorithm To Retrieve Information Coded In Alphanumeric). D. R. Morrison proposed this data structure in order to avoid an annoying flaw of tries, namely, one way branching of internal nodes. PATRICIA Tries share many advantages of Tries over balanced trees and binary search trees. Moreover, they require less space for storage. Because of this property, PATRICIA tries find particular application in the area of IP routing, where the ability to contain large ranges of values with a few exceptions is particularly suited to the hierarchical organization of IP addresses. They are also used for inverted indexes of text documents in information retrieval.

Another variant of tries are digital search trees. They were first introduced by Coffman and Eve [24]. They have attracted considerable attention due to their wide applications, especially their close connection to the famous Lempel-Ziv compression scheme [98]. The major difference between DSTs and Tries is that for DSTs, the data are stored in the nodes while the data only appears in the leaves of tries. Bucket digital search trees are a generalization of DSTs, they are sometimes also called generalized digital search tree [90]. In b-DSTs, each node can store b keys. When $b = 1$, they correspond to the original DST. The advantage of such ”bucketing” is multifold such as reducing the root-to-node path lengths and improving the storage utilization. Because bucketing is an important tool in many algorithm designing problems ranging from hashing to computational geometry, b-DSTs are related to many practical algorithms [33, 82, 126, 140, 172]. For example, people use b-DSTs as a tool for memory management in UNIX [90].

General digital trees like tries, PATRICIA tries, DSTs and b-DSTs share a nice property, namely, the average height of the trees is almost optimal. This property ensures that the number of comparing operation during data searching is almost minimal.

There are a great deal of studies about all kinds of properties of digital trees. People study digital trees not only for the practical use we just discussed but also the analysis of digital trees poses mathematical challenges. We will discuss the known properties and studies of digital trees in Section 2.3. However, before we do so, we will make precise the definition of the digital trees discussed above.

2.2 Random Model and Constructions for Random Digital Trees

In this section, we will give the construction and random model for each variant of random digital trees. To make it easier to see the differences, we will always use the same set of infinite binary strings as keys

$$\begin{array}{ll}
S_1 = 0011010\ldots & S_5 = 0000010\ldots \\
S_2 = 0000110\ldots & S_6 = 0110101\ldots \\
S_3 = 1110110\ldots & S_7 = 1000011\ldots \\
S_4 = 1000100\ldots & S_8 = 0010011\ldots
\end{array}$$

to build the examples in the following sections. We will use the same random model, the so-called Bernoulli model, for all the random digital trees we discuss in this thesis. For the Bernoulli model, it is assumed that the letters of each keys are generated independently.

In the binary case, the i-th key will be of the form

$$A_{i,1}, A_{i,2}, \ldots, A_{i,k}, \ldots$$

where for all $1 \leq i \leq n$ and $j \in \mathbb{N}$, $\mathbb{P}(A_{i,j} = 0) = p$ and $\mathbb{P}(A_{i,j} = 1) = q = 1-p$ with some $0 \leq p \leq 1$.

For general m-ary cases, the i-th key is of the same form with $A_{i,j} \in \mathcal{A} = \{a_1, \ldots, a_m\}$ for some alphabet $\mathcal{A}$ of the size m. Moreover, $\mathbb{P}(A_{i,j} = a_k) = p_k$ for all $1 \leq k \leq m$. Of course, we also have that

$$\sum_{i=1}^{m} p_i = 1 \quad \text{and} \quad 0 \leq p_i \leq 1 \text{ for all } 1 \leq i \leq m.$$

The Bernoulli model is the most simple model proposed. More realistic models have been propose by B. Valleé [206] and analyzed by Bourdon [14] and Valleé et. al. [23]. Although the Bernoulli model may seem too idealized for practical applications, typical behaviors under this model often hold under more general models such as Markovian or dynamical sources [96, 134, 203].

2.2.1 Tries

Here we describe how to build a binary trie. For m-ary tries, the procedures are the same only the alphabet is expanded. We start with n data whose keys are infinite 0-1 strings. In a trie, the keys are only stored in the leaves. Whenever a new key is stored, we use it to search in the already existing trie by its prefix until we encounter a leaf which already contains a key. Then, the leaf is replaced by an internal node and the two keys are distributed to the two subtrees. If the two keys go to the same subtree, then the procedure is repeated until both keys go to different subtrees where they are stored in the leaves.

To make this definition more precise, we give a more formal description. Let $\{0,1\}^\infty$ be the set of binary strings of infinite length, we define two operations. The map

$$\underline{\sigma} : \{0,1\}^\infty \longrightarrow \{0,1\}$$

returns the first letter of a string. The shift function

$$\underline{T} : \{0,1\}^\infty \longrightarrow \{0,1\}^\infty$$

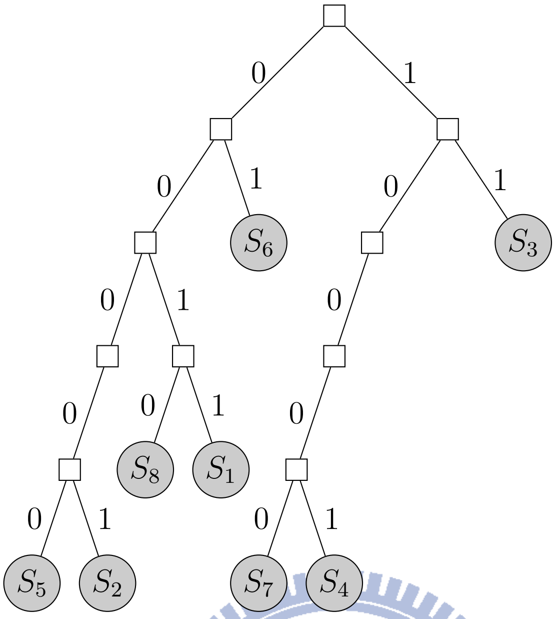

Figure 2.1: A trie built from the keys $S_1, \ldots, S_8$. The rectangle nodes represent the internal nodes while the circle ones are the external nodes which store the keys.

returns the first suffix of a string. Then, the function $\underline{T}_{[a]}$ is the restriction of $\underline{T}$ to the set $\underline{\sigma}^{-1}(a)$ of a word beginning with symbol a.

With any finite set $\mathcal{S}$ of infinite strings, we associate a trie, denoted by $TR(\mathcal{S})$, defined by the following recursive rules:

1. If $\mathcal{S} = \emptyset$, then $TR(\mathcal{S})$ is the empty tree.

2. If $\mathcal{S} = \{s\}$ has cardinality equal to one, then $TR(\mathcal{S})$ consists of a single leaf node containing s.

3. If $|\mathcal{S}| \geq 2$, then $TR(\mathcal{S})$ is an internal node represented generically by $\Box$ to which 2 subtrees are attached

$$TR(\mathcal{S}) =< \Box, TR(\underline{T}_{[0]}\mathcal{S}), TR(\underline{T}_{[1]}\mathcal{S}) > .$$

2.2.2 PATRICIA Tries

PATRICIA tries are tries without one-way branching. As Figure 2.1 shows, tries may contain many internal nodes with only one child. To build a PATRICIA trie, one may build a trie first and ”collapse” all the internal nodes

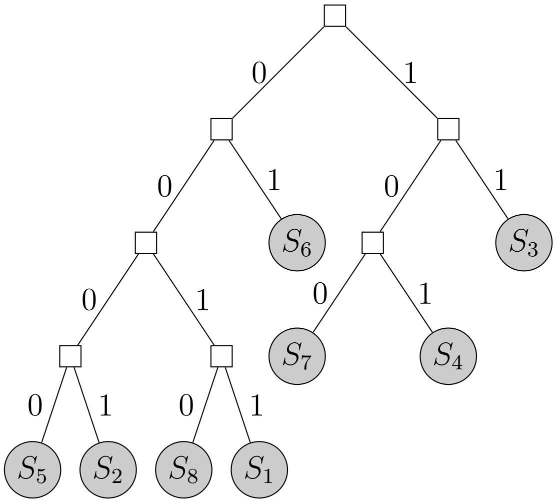

Figure 2.2: A PATRICIA trie built from the keys $S_1, \ldots, S_8$. Note that all the internal nodes have two childs.

with only one child. To see the difference, compare Figure 2.1 and Figure 2.2.

Let $PTR(\mathcal{S})$ be the PATRICIA trie associated to the finite set $\mathcal{S}$ of infinite strings. $PTR(\mathcal{S})$ is constructed by the following recursive rules:

1. If $\mathcal{S} = \emptyset$, then $PTR(\mathcal{S})$ is the empty tree.

2. If $\mathcal{S} = \{s\}$ has cardinality equal to one, then $PTR(\mathcal{S})$ consists of a single leaf node containing s.

3. If $|\mathcal{S}| \geq 2$, by the number of distinct symbols contained in the multiset $\underline{\sigma}(\mathcal{S})$, there are two cases

   (i) If $\underline{\sigma}(\mathcal{S})$ contains only one symbol, then

   $$PTR(\mathcal{S}) = PTR(\underline{T}\mathcal{S}).$$

   (ii) Otherwise, $TR(\mathcal{S})$ is an internal node represented generically by $\square$ to which 2 subtrees are attached

   $$PTR(\mathcal{S}) =< \square, PTR(\underline{T}_0\mathcal{S}), PTR(\underline{T}_1\mathcal{S}) > .$$

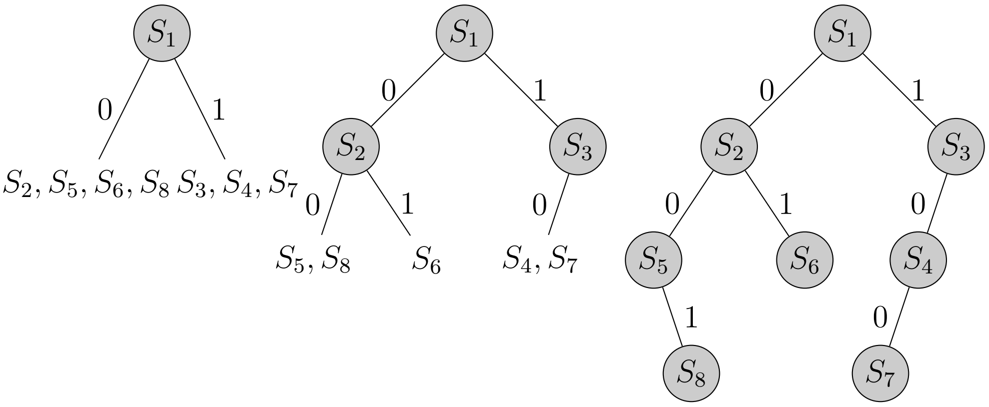

Figure 2.3: The process of building a DST from the keys $S_1, \ldots, S_8$. Note that if the order of the keys is different, the resulting DST would be different.

2.2.3 Digital Search Trees

Here, we demonstrate how to build a binary digital search tree. As for tries and PATRICIA tries, we consider n keys which are infinite $\{0, 1\}$ strings. If $n = 1$, then the only key is put in a node and the building process is finished. Otherwise, we go through the following steps:

1. Store the first key in a node (which will become the root of the tree).

2. Distribute the remaining keys into two sets according to the first letter of the key. If the first letter is:

   0: Put the keys to the left subtree.

   1: Put the keys to the right subtree.

3. Remove all the first letters of the keys in the subtrees. Build the subtrees recursively according to the same rules.

For a bucket digital search tree with bucket size b, the building process is almost the same as for digital search trees. Only in step one of the construction, we put b keys in the node instead of only one node. It is easy to see the differences by comparing Figure 2.3 and Figure 2.4.

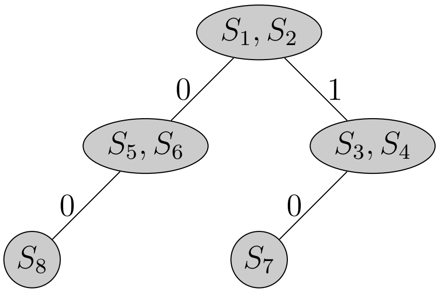

Figure 2.4: A bucket digital search tree built from the keys $S_1, \ldots, S_8$ with bucket size $b = 2$.

2.3 Past Researches about Random Digital Trees

Random digital trees have been studied from many different aspects. In this section, we focus on shape parameters of random digital trees. For each parameter, we will give the definition first and then explain the practical use of the parameter. We will also give a list of references of known results of the parameter.

Size

For random digital trees, the parameter size is the total number of internal nodes. So, this parameter is only relevant for tries and PATRICIA tries. People study this parameter because the number of internal nodes is proportional to the number of pointer needed to store the data structure. The less the number of internal nodes needed, the less the space required. Note that binary PATRICIA tries have a constant size and hence this parameter matters only for m-ary PATRICIA tries with $m \geq 3$. For the study of size of tries, see [23, 39, 93, 103, 107, 109, 112, 116, 139, 145, 155, 95, 194]. For PATRICIA tries, see [14, 15, 38, 39, 146].

Depth

Depth is defined as the distance from the root to a randomly selected node which normally contains data. Some researchers use the name depth of insertion or successful search time. It is one of the most well-studied parameter since it provide a great deal of information for many applications. For example, the depth of a node storing a key represents the search time for the key in searching and sorting algorithms [127]. Depth also gives the length of a conflict resolution session for tree-based communication protocols. For compression algorithms, depth is the length of a substring that may be occupied or compressed [7]. For the study of depth of tries, see [31, 36, 37, 57, 92, 97, 114, 136, 169, 192, 196, 198]. For the study of this parameter in PATRICIA tries, see [38, 39, 181]. Finally, for DSTs, see [36, 135, 169, 176, 200, 139, 37, 30, 100, 102, 127, 113, 197].

Height

Height of a tree is the length of the longest path from the root to a leaf. It can also be understood as the maximum value of the depth. Height in digital trees reflects the the longest common prefix of words stored in a digital tree and is directly related to many important operations such as hashing in computer programming. See [23, 32, 36, 37, 38, 60, 71, 92, 168, 169, 201] for studies of the height tries. See [36, 38, 39, 68, 111, 127, 168, 170, 171, 199, 201, 204] for PATRICIA tries and [122, 35, 37, 47, 139, 169] for digital search trees.

Shortest path length

This parameter is the length of the shortest path from the root to the leaves. It was studied for tries by B. Pittel in [168, 169]. He considered this parameter in order to derive the depth and height of a random trees.

Saturation level

Sometimes also called fill-up level. The levels of a tree are the set of nodes which are of the same distance from the root. The saturation levels are the full levels in a tree. Researcher have investigate the number of saturation levels, the maximum level which is full and so on. This parameter was firstly studied for the purpose of understanding the behavior of the parameter height [35]. However, it found also other applications in improving the efficiency of algorithms for IP address lookup problem [123]. For more details, see [35, 123, 169].

Stack size

In the definition of the height, every edge contributes one when counting the distance. Stack size is a kind of ”biased” height with the edges whose label is the last symbol in the alphabet (for binary case, the edges labeled by 1) make no contribution when counting the distance. For general m-ary tries and PATRICIA tries, if the order of the symbols in the alphabet is given, a preorder traversals of the corresponding trie or PATRICIA trie gives the list of the words stored in the lexicographical order. When the traversal is implemented in a recursive way, the height measures the recursion-depth needed. However, in many cases the recursion is removed or a technique called end-recursion removal is used to save recursion calls. In those cases, the amount of memory space is no longer measured by the height and hence the parameter stack-size is introduced. For stack size of tries, see [16, 157, 158]. For the PATRICIA tries case, L. Devroye gave a method to compute this quantity in [39], however, without giving an explicit result.

Horton-Strahler number

The Horton-Strahler is defined similar to stack size for similar purposes. It specifies the recursion-depth needed for a traversal when pre-recursion removal is applied and the subtree is visited in an order chosen to minimize the recursion depth (this order is fixed for stack-size). Related study for tries can be found in [16, 41, 156, 157, 158]. In [39], the author gave a bound for Horton-Strahler number of PATRICIA tries.

One-sided path length

One-sided path length is length of the path with all edges labeled by the same symbol. This parameter is directly related to a widely used algorithm called leader (or loser) selection. For the algorithm and its applications, see [59]. Because of its recursive nature, the algorithm will generate a tree which has a similar structure as tries during the selecting process. As as result, the one-sided path length of tries will directly reflect the efficiency of the algorithm. Related studies can be found in [59, 106, 174, 212, 214].

Occurrence of certain pattern

People may interested in nodes of the digital trees satisfying certain properties or the number of certain specified subgraphs in a digital tree. For example, the so-called 2-protected nodes, the nodes which are neither the leaves nor parents of a leaf, is a type of nodes which has been studied. For more researches of this kind, see [60, 114, 116, 162, 192, 198, 169, 200, 89, 90, 98, 114, 121, 127]

External path length

This parameter is the sum of distances between leaves and the root. In the case of trie data structures of hashing schemes, this parameter represents the processing time in central unit. However, this parameter has drawn a lot of interests not because of its practical use but for an interesting phenomenon. It is easy to see that the mean of the external path length can be derived directly from the mean of the depth. Therefore, people expected that the variance of the external path length can also be directly derived from the variance of the depth. However, this is not the case [118]. This parameter has been studied for tries [94, 99, 119, 167] and PATRICIA tries [14, 39, 118, 199].

Internal path length

Internal path length is defined as the sum of distances between internal nodes and the root. In contrast to tries and PATRICIA tries which store data in the leaves, digital search trees store data in the internal nodes. Thus, the internal path length can be seen as the counterpart of the external path length in digital search trees. L. Devroye studied this parameter for PATRICIA tries in [38]. Later, Fuchs, Hwang and Vytas gave a general framework for parameters with similar properties of the internal path length in [75]. Researches about internal path length of digital search trees can be found in [200, 89, 173, 90, 117, 121, 98, 192, 160, 74].

Distances

This parameter refers to the distance between two randomly selected nodes in a digital tree. It can be seen as a measure of how "diverse" a tree is. R. Neininger studied this parameter to get better understanding of the Wiener index (which will be discussed later) of recursive trees [159]. After Neininger's work, this parameter was studied for many different classes of trees. For the study of distances of tries, see [1, 3, 22]. The study of distance of digital search trees can be found in [1, 2].

Number of unary nodes in tries

Unary nodes are the nodes with exactly one child. PATRICIA tries are the tries with unary nodes "collapsed". Therefore, studying the number of unary nodes may gives us an estimation how much space is "saved" in PATRICIA tries comparing to tries. Besides comparing the two variations of random digital trees, S. Wagner studied this parameter in [210] in order to get a better understanding of the efficiency of contention tree algorithms. Related studies can be found in [18, 205, 210]

Node profile

The node profile of a random digital tree is a parameter which represents the number of nodes at the same distance from the root. This parameter has drawn a lot of attention because many fundamental parameters, including size, depth, height, shortest path length, internal path length and saturation level can be uniformly analyzed and expressed in terms of node profiles. Although node profile are of great importance, there are not too many researches about it until very recently. See [167] for the node profile of tries, [38, 39] for node profile of PATRICIA tries and [48, 50, 129, 135, 51] for node profile of digital search trees.

Partial match queries

Multidimensional data retrieval is an important issue for the design of data base system. Partial match retrieval is a widely used method of retrieval which found many applications, especially in geographical data and graphic algorithm. For more detailed introduction of partial retrieval operation, see [65] and the references within. Because of the importance of this method, the performance of partial match retrieval on digital trees received a lot of interests. For the analysis of partial match retrieval on tries, PATRICIA tries and bucket-digital search trees, see [39, 42, 65, 73, 120, 190, 191].

Peripheral path length

This parameter was proposed by W. Szpankowski and M. D. Ward in [133, 211, 213] with the name w-parameter. This parameter was originally applied in the study of Lempel-Ziv'77 data compression algorithm on uncompressed suffix trees. In [74], the authors renamed this parameter as the peripheral path length. The fringe-size of a leaf node is defined to be the size of the subtree rooted as its parent node. The peripheral path length is then defined to be the sum of the fringe-size of all leaves of the tree. Peripheral path length has been studied for tries, PATRICIA tries and digital search trees. For related researches, see [49, 74, 133, 211, 213].

Weighted path length

Let l_j be the distance of the j-th node to the root and w_j be the weight attached to the j-th node, the weighted path length is defined to be

$$W_n = \sum_{j=1}^{n} w_j l_j$$

for a tree with n nodes. For many real life applications, people need to assign a weight to each edge or node of a graph. Weighted path length arises for these applications. For more details, see [74].

Differential path length

Also called the colless index, it inspects the internal nodes of trees. We partition the leaves descend end from internal nodes into two groups of sizes L and R. The differential path length is the sum over all absolute values $|L-R|$ for all ancestors. This parameter is investigated in the system biology literature [11]. M. Fuchs et. al. studied this parameter for symmetric random digital search trees [74].

Type of nodes in bucket digital search trees

Because of the construction, a bucket digital search tree with bucket size $b \geq 2$ may have nodes containing different number of keys. The authors of [90] proposed a multivariate frame to study the number of each type of nodes in bucket digital search trees.

Key-wise path length

Since a node of a bucket digital search tree with $b \geq 2$ may contain more than one key, the distance of two keys stored in a b-DST could be 0 in some cases. Therefore, researchers defined two different types of path length to study b-DSTs, key-wise path length and node-wise path length. The key-wise path length is defined as the sum of the distances between keys and the root. Related researches can be found in [74, 89].

Node-wise path length

Node-wise path length of b-DSTS is defined to be the sum of all distances between nodes and the root. See [74, 90] for more details.

Chapter 3

Related Mathematical Techniques

3.1 Rice Method

Rice method, sometimes also called the Nörlund-Rice integral or Rice integral, is named in honor of Niels Erik Nörlund and Stephen O. Rice. It is a fruitful method for finding the asymptotic expansion of sums of the form

$$\sum_{k=n_0}^{n} \binom{n}{k} (-1)^k f(k). \tag{3.1}$$

The formulae of Rice integral is given by

$$\sum_{k=n_0}^{n} \binom{n}{k} (-1)^k f(k) = \frac{(-1)^n}{2\pi i} \oint_C f(z) \frac{n!}{z(z-1)\cdots(z-n)} dz,$$

where C is a positive oriented closed curve encircling the points $n_0, n_1, \ldots, n$ and $f(z)$ is understood to be analytic within C. The integral can also be written as

$$\sum_{k=n_0}^{n} \binom{n}{k} (-1)^k f(k) = \frac{-1}{2\pi i} \oint_C B(n+1, -z) f(z) dz,$$

where C is the contour mentioned before and $B(a, b)$ is the beta function. Now we give a formal statement of the two formulaes.

Lemma 3.1.1. *Let $f(z)$ be analytic in a domain that contains the half-line*

$[n_0, +\infty)$. *Then, we have the representation*

$$\sum_{k=n_0}^{n} \binom{n}{k} (-1)^k f(k) = \frac{(-1)^n}{2\pi i} \oint_C f(z) \frac{n!}{z(z-1)\cdots(z-n)} dz$$
$$= \frac{-1}{2\pi i} \oint_C B(n+1, -z) f(z) dz,$$

where C is a positively oriented closed curve that lies in the domain of analyticity of $f(z)$, encircles $[n_0, n]$, and does not include any of the integers $0, 1, \ldots, n_0 - 1$.

Proof. The proof of these two equalities is omitted here since it is a simple application of residue theorem. For the complete proof, see [69]. □

Suppose we have a finite differences sum of the form in (3.1), the Rice method allows us to compute an asymptotic expansion by the following steps:

Step 1. Extend $f(k)$ which is originally defined on integers to an appropriate meromorphic function $f(z)$.

Step 2. Choose a suitable contour C which encircles the points $n_0, \ldots, n$ and consider the Rice integral

$$\Delta = \frac{(-1)^n}{2\pi i} \oint_C f(z) \frac{n!}{z(z-1)\cdots(z-n)} dz.$$

Step 3. Residue theorem yields that

$$\Delta = \sum_{k=n_0}^{n} \binom{n}{k} (-1)^k f(k) + \{\text{contribution from other poles inside C}\}.$$

Step 4. Estimate Δ.

Remark 1. The most difficult step of the Rice method is usually Step 1. Finding the meromorphic extension of $f(k)$ is usually quite tricky. Also, to carry out Step 4 one needs growth properties of $f(z)$, e.g. $f(z)$ is of polynomial growth.

Rice method has been used in many researches about shape parameters of random digital trees. For example, P. Kirschenhofer, H. Prodinger and W. Szpankowski used it to derive the mean and variance of the internal path length of symmetric DSTs [114, 117, 121]. Here, we use the mean of the internal path length of DSTs as an instance to illustrate how the Rice method works.

We let S_n be the expectation of the internal path length of a symmetric DST built on n strings. Then, under the Bernoulli mode, from the definition of the internal path length, we have the recurrence

$$S_{n+1} = n + 2^{1-n} \sum_{k=0}^{n} \binom{n}{k} S_k, \quad (n \geq 0),$$

where the initial condition is given by $S_0 = 0$. After some manipulations of generating functions, we get that

$$S_n = \sum_{k=2}^{n} \binom{n}{k} (-1)^k Q_{k-2}.$$

Example 3.1.2. *Consider the sum*

$$S_n = \sum_{k=2}^{n} \binom{n}{k} (-1)^k Q_{k-2}, \tag{3.2}$$

where $Q_n = \prod_{j=1}^{n}(1 - 2^{-j})$*.*

Step 1. *We introduce the function*

$$Q(s) = \prod_{n \geq 1} \left(1 - \frac{s}{2^n}\right),$$

then the constant Q_n *can be expressed as* $Q_n = Q(1)/Q(2^{-n})$*. Note that* $Q(1)$ *is finite and numerical results give us* $Q(1) \simeq 0.288788 \cdots$*. Moreover,* $Q(1)/Q(2^{2-z})$ *is analytic on* $[2, \infty)$*. Thus, we get the needed extension.*

Step 2. *We choose the contour* C *to be*

$$\left\{ z = x + iy : \left(x - \frac{1}{2}\right)^2 + y^2 = N^2, x \geq \frac{1}{2} \right\} \bigcup \left\{ z = x + iy : x = \frac{1}{2}, -N \leq y \leq N \right\},$$

and consider the integral

$$\Delta = \frac{(-1)^n}{2\pi i} \oint_C \frac{Q(1)}{Q(2^{2-z})} \frac{n!}{z(z-1)\cdots(z-n)} dz.$$

Step 3. *To compute the residues, we need to find the poles of* $Q(1)/Q(2^{2-z})$*. The poles occur at* $z = 1+2k\pi i/\log 2$ *and hence the integral has a double pole at* $z = 1$ *and simple poles at* $z = 1 + 2k\pi i/\log 2$ *with* $k \in \mathbb{Z} \setminus \{0\}$*. To compute the contribution from the double pole, we derive the series expansion*

$$\begin{aligned}\frac{(-1)^n n!}{z(z-1)\cdots(z-n)} =& \frac{-1}{z(z-1)}\prod_{j=2}^{n}\left(1-\frac{z}{j}\right)^{-1}\\ =& -\frac{n}{z-1} - n(H_{n-1}-1) + \mathcal{O}(z-1)\end{aligned}$$

and

$$\begin{aligned}\frac{Q(1)}{Q(2^{1-z})} =& Q(1)\prod_{j\geq 0}\left(1-2^{-(z+j)}\right)^{-1}\\ =& 1 - \alpha(z-1)\log 2 + \mathcal{O}\left((z-1)^2\right),\end{aligned}$$

where $\alpha = \sum_{n\geq 1}\frac{1}{2^n-1}$*. Combining the expansions, we get*

$$\begin{aligned}\frac{Q(1)}{Q(2^{2-z})}\frac{(-1)^n n!}{z(z-1)\cdots(z-n)} =& \left(\frac{1}{(z-1)\log 2} + \frac{1}{2} + \mathcal{O}(z-1)\right)\\ &\times\left(1-\alpha(z-1)\log 2 + \mathcal{O}\left((z-1)^2\right)\right)\\ &\times\left(-\frac{n}{z-1} - n(H_{n-1}-1) + \mathcal{O}(z-1)\right).\end{aligned}$$

From the above expansion, we get the residue at $z = 1$ *as*

$$-\frac{n}{\log 2}(H_{n-1}-1) + n\left(\alpha - \frac{1}{2}\right) = -n\log_2 n - n\left(\frac{\gamma-1}{\log 2} - \alpha + \frac{1}{2}\right) + \mathcal{O}(1).$$

The poles at $1 + 2k\pi i/\log 2$ *with* $k \in \mathbb{Z} \setminus \{0\}$ *make contributions* $\delta(n)$ *[68], where*

$$\delta(n) = -\frac{1}{\log 2}\sum_{k\in\mathbb{Z}\setminus\{0\}}\Gamma\left(-1-\frac{2k\pi i}{\log 2}\right)e^{2k\pi i\log_2 n}.$$

Step 4. *On the right semi-circle,* Δ *converges to 0 as n grows since*

$$\left|\frac{1}{Q(2^{2-z})}\right| = \mathcal{O}(1) \quad \text{as } |z| \to \infty.$$

On the left line, we have the bound

$$\mathcal{O}\left(\int_{-\infty}^{\infty}\frac{\Gamma(n+1)\Gamma(\frac{1}{2}+iy)}{\Gamma(n+\frac{1}{2}-iy)}dy\right) = \mathcal{O}(n^{1/2}).$$

Combining all the above steps, we have

$$S_n = n\log_2 n + n\left(\frac{\gamma - 1}{\log 2} - \alpha + \frac{1}{2} + \sigma(n)\right) + \mathcal{O}(n^{1/2}). \tag{3.3}$$

The internal path length of DSTs is not the only shape parameter which has been analyzed by the Rice method. For example, the external path length of tries and PATRICIA tries under the Bernoulli model can also be analyzed by the Rice method [118, 119].

However, there are several other mathematical tools which can be used to analyze these shape parameters and they are sometimes more useful than the Rice method. In the following sections, we will introduce these tools. We begin with a standard tool which can transfer a problem under the Bernoulli model into a problem under the so-called Poisson model in which we have many useful tools to conquer the problem.

3.2 Poissonization and Depoissonization

In combinatorics and the analysis of algorithms, a Poisson version of a problem (henceforth called Poisson model or poissonization) is often easier to solve than the original one, which we name here the Bernoulli model. Poissonization is a technique that replaces the original input by a poisson process. This technique was first introduced by Marc Kac [108] in 1949. Recently, this technique flourished in the community of analysis of algorithms and random combinatorial structures. See [99] for a comprehensive survey and many references.

We first introduce the formal definition of analytical poissonization (which is also called Poisson Transform by G. Gonnet and J. Munro [83]).

Definition 3.2.1. *Let $\{g_n\}$ be a sequence, then the Poisson transform $\tilde{G}(z)$ of $\{g_n\}$ is defined as*

$$\tilde{G}(z) = e^{-z}\sum_{n\geq 0} g_n \frac{z^n}{n!}$$

for arbitrary complex z. $\tilde{G}(z)$ is also called the **Poisson generating function**.

If $\tilde{G}(z)$ is known, we can extract the coefficient $g_n = n![z^n]\tilde{G}(z)e^z$. However, in most situations $\tilde{G}(z)$ satisfies a complicated functional/differential equation that is difficult to solve exactly. Fortunately, we have many tools, Mellin transform for example, to find the asymptotic expansion of $\tilde{G}(z)$.

Therefore, the natural next step is to find a method to derive asymptotic expansion of g_n from the asymptotic expansion of $\tilde{G}(z)$. This step is called *analytical depoissonization*. To explain the theory of analytical depoissonization, we give some definitions first.

Definition 3.2.2. *(i) A linear cone is defined as the region in the complex plane satisfying*

$$\mathcal{L}_\theta = \{z : |\arg z| \le \theta\},$$

where $|\theta| \le \pi/2$.

(ii) A polynomial cone $\mathcal{L}(D, \delta)$ *is defined as*

$$\mathcal{L}(D, \delta) = \{z = x + iy : |y| \le Dx^\delta, 0 < \delta \le 1, D > 0\}.$$

We consider now a sequence $\{g_n\}_{n\ge0}$ and its Poisson generating function $\tilde{G}(z) = e^{-z}\sum_{n\ge0} g_n z^n/n!$. Our goal is to derive the asymptotic expansion of g_n from $\tilde{G}(z)$. By Cauchy's formula, we have

$$g_n = \frac{n!}{2\pi i}\oint \frac{\tilde{G}(z)e^z}{z^{n+1}}dz = \frac{n!}{n^n 2\pi}\int_{-\pi}^{\pi} \tilde{G}(ne^{it})\exp(ne^{it})e^{-nit}dt.$$

The depoissonization result will follow from the above by careful estimation of the integral using the saddle point method. Because the complete proof of depoissonization is rather long and technical, we omit it here. For the complete proof, see [99]. Here we only state the results.

Theorem 3.2.3. *(Basic depoissonization lemma) Let* $\tilde{G}(z)$ *be the Poisson generating function of a sequence* $\{g_n\}$ *that is assumed to be an entire function of* z*. Suppose that in a linear cone* $\mathcal{L}_\theta$ *both of the following two conditions hold for some numbers* $A, B, R > 0$, β *and* $\alpha < 1$.

1. *For* $z \in \mathcal{L}_\theta$ *and* $|z| > R$

$$|\tilde{G}(z)| \le B|z|^\beta;$$

2. *For* $z \notin \mathcal{L}_\theta$ *and* $|z| > R$

$$|\tilde{G}(z)e^z| \le Ae^{\alpha|z|}.$$

Then,

$$g_n = \tilde{G}(n) + \mathcal{O}(n^{\beta-1})$$

for large n.

The above theorem can be generalized in several different directions. In [99], the authors gave three generalized versions. Here we give the one which is most often used.

Theorem 3.2.4. *(General depoissonization lemma) Consider a polynomial cone $\mathcal{L}(D, \delta)$ with $1/2 < \delta \leq 1$. Let the following two conditions hold for some numbers $A, B, R > 0$ and $\alpha > 0$, β and γ:*

1. *For $z \in \mathcal{L}(D, \delta)$ and $|z| > R$*

$$|\tilde{G}(z)| \leq B|z|^{\beta}\Psi(|z|),$$

where $\Psi(z)$ is a slowly varying function, that is, a function for which for fixed t, we have that $\lim_{z\to\infty} \Psi(tz)/\Psi(z) = 1$;

2. *For all $z = \rho e^{i\theta}$ with $\theta \leq \pi$ such that $z \notin \mathcal{L}(D, \delta)$ and $\rho = |z| > R$*

$$|\tilde{G}(z)e^{z}| \leq A\rho^{\gamma} e^{\rho(1-\alpha\theta^2)}.$$

Then, for every nonnegative integer m

$$\begin{aligned} g_n &= \sum_{i=0}^{m}\sum_{j=0}^{i+m} b_{ij} n^i \tilde{G}^{(j)}(n) + \mathcal{O}(n^{\beta-(m+1)(2\delta-1)}\Psi(n)) \\ &= \tilde{G}(n) + \sum_{k=1}^{m}\sum_{j=1}^{k} b_{i,k+i} n^i \tilde{G}^{(k+i)}(n) + \mathcal{O}(n^{\beta-(m+1)(2\delta-1)}\Psi(n)), \end{aligned}$$

where b_{ij} are the coefficients of $e^{x\log(1+y)-xy}$ at $x^i y^j$, that is

$$\sum_{i\geq 0}\sum_{j\geq 0} b_{ij} x^i y^j = e^{x\log(1+y)-xy}.$$

Let us take the external path length of tries, which has been mentioned at the end of Section 3.1 as an example. We denote by P_n the random variable of the external path length of random tries built on n records. Then, under the Bernoulli model, we have

$$P_n \overset{d}{=} P_{B_n} + P_{n-B_n} + n, \quad (n \geq 2), \tag{3.4}$$

where $B_n = Binomial(n, p)$ with $p \in (0, 1)$ and $P_0 = P_1 = 0$. Let

$$\tilde{f}(z) = e^{-z}\sum_{n\geq 0} \mathbb{E}(P_n)\frac{z^n}{n!}$$

be the Poisson generating function of the mean of P_n. Then from (3.4), we get the functional equation

$$\tilde{f}(z) = \tilde{f}(pz) + \tilde{f}(qz) + z(1 - e^{-z}). \tag{3.5}$$

One can now check by induction that $\tilde{f}(z)$ satisfies the assumptions of Theorem 3.2.3 (see also Section 3.5.2 for a more systematic method of checking this). Then, this result implies that

$$\mathbb{E}(P_n) = \tilde{f}(n) + \mathcal{O}(n^{\epsilon}), \tag{3.6}$$

where ϵ can be arbitrarily small. Thus, the natural next step is to find more information about the asymptotic behavior of $\tilde{f}(z)$. For this purpose, we introduce one of the most often used tools under such circumstance, the Mellin transform.

3.3 Mellin Transform

Mellin transform is the most popular integral transform in the analysis of algorithms. Its first occurrence is in a memoir of Riemann in which he used it to study the famous Zeta function. However, the transform gets its name from the Finish mathematician Hjalmar Mellin who did the first systematic study of the transform and its inverse. See [132] for a summary of his works. Nowadays, the Mellin transform is used in complex analysis, number theory, applied mathematics and analysis of algorithms. Apart from these applications in mathematics, the Mellin transform has also been applied in many different areas such as physics and engineering.

In the analysis of algorithms, the Mellin transform is mostly used to derive asymptotic expansions. The transform is defined in the following:

Definition 3.3.1. *Let $f(z)$ be a function which is locally Lebesque integrable over $(0, +\infty)$. The Mellin transform of $f(z)$ is defined by*

$$\mathscr{M}[f(z); s] = f^*(s) = \int_0^{\infty} f(z) z^{s-1} dz.$$

The largest open strip $\langle \alpha, \beta \rangle$ in which the integral converges is called the *fundamental strip*. To determine the fundamental strip, we have the following lemma.

Lemma 3.3.2. *If the function $f(z)$ satisfies*

$$f(z) = \begin{cases} \mathcal{O}(z^u), & z \to 0^+; \\ \mathcal{O}(z^v), & z \to +\infty, \end{cases}$$

then $\mathscr{M}[f(z); s]$ exists in the strip $\langle -u, -v \rangle$ and is analytic there.

	$f(z)$	$f^*(s)$	$< \alpha, \beta >$	
F1	$z^\nu f(z)$	$f^*(s+\nu)$	$< \alpha-\nu, \beta-\nu >$	
F2	$f(z^\rho)$	$\frac{1}{\rho} f^*(s/\rho)$	$< \rho\alpha, \rho\beta >$	$\rho > 0$
	$f(1/z)$	$-f^*(-s)$	$< -\beta, -\alpha >$	
F3	$f(\mu z)$	$\frac{1}{\mu^s} f^*(s)$	$< \alpha, \beta >$	$\mu > 0$
	$\sum_k \lambda_k f(\mu_k z)$	$(\sum_k \lambda_k \mu_k^{-s}) f^*(s)$	$< \alpha, \beta >$	
F4	$f(z) \log z$	$\frac{d}{ds} f^*(s)$	$< \alpha, \beta >$	
F5	$\Theta f(z)$	$-s f^*(s)$	$< \alpha', \beta' >$	$\Theta = z\frac{d}{dz}$
	$\frac{d}{dz} f(z)$	$-(s-1) f^*(s-1)$	$< \alpha'-1, \beta'-1 >$	
	$\int_0^z f(t)dt$	$-\frac{1}{s} f^*(s+1)$		

Table 3.1: Functional properties of Mellin transform

Before we start to develop a systematic theory of the Mellin transform, we give an easy example to show how the transform works.

Example 3.3.3. *For $f(z) = e^{-z}$, it is obvious that*

$$f(z) = \begin{cases} \mathcal{O}(z^0), & z \to 0^+; \\ \mathcal{O}(z^{-b}), & z \to +\infty, \end{cases}$$

where b is an positive real number which can be arbitrarily large. As a result,

$$\mathscr{M}[f(z); s] = \Gamma(s)$$

is defined in $\langle 0, +\infty \rangle$ and analytic there.

Simple changes of variables in the definition of Mellin transforms yields many useful functional properties summarized in Table 3.1.

Similar to other integral transformations, the Mellin transform has an inverse. For a given function $f(z)$ and its Mellin transform $f^*(s)$, we let $s = \sigma + 2\pi i t$ and $z = e^{-y}$. Then, the Mellin transform becomes a Fourier transform

$$f^*(s) = \int_0^\infty f(z) z^{s-1} dz = \int_{-\infty}^\infty f(e^{-y}) e^{-\sigma y} e^{-2\pi i t y} dy = \mathscr{F}[f(e^{-y}) e^{-\sigma y}; t],$$

where $\mathscr{F}[f(z); t]$ denotes the Fourier transform of $f(z)$. As a result, the inversion theorem for the Mellin transform follows from the corresponding one for the Fourier transform. If $\hat{f}(t) = \mathscr{F}[f; t]$ is the Fourier transform, then the original function is recovered by

$$f(z) = \int_{-\infty}^\infty \hat{f}(t) e^{2\pi i t z} dt.$$

Thus,

$$f(e^{-y})e^{-\sigma y} = \int_{-\infty}^{\infty} f^*(\sigma + 2\pi i t)e^{2\pi i t y}dt.$$

Changing the variables y back to z, we get

$$f(z) = z^{-\sigma}\int_{-\infty}^{\infty} f^*(\sigma + 2\pi i t)z^{-2\pi i t}dt.$$

Finally, by replacing $\sigma + 2\pi i t$ by s, we have

$$f(z) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f^*(s)z^{-s}ds.$$

Theorem 3.3.4. *Let $f(z)$ be integrable with fundamental strip $\langle\alpha, \beta\rangle$. If c is such that $\alpha < c < \beta$ and $f^*(c + it)$ is integrable, then the equality*

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f^*(s)z^{-s}ds = f(z)$$

holds almost everywhere. Moreover, if $f(z)$ is continuous, then the equality holds everywhere on $(0, +\infty)$.

As we mentioned before, the major application of the Mellin transform in the analysis of algorithms is to derive asymptotic expansions. This application comes from the correspondence between the asymptotic expansion of a function at 0 (and ∞), and poles of its Mellin transform in a left (resp. right) half-plane. Before we explain the correspondence, we introduce the singular expansion of a function.

Definition 3.3.5. *Let $\phi(s)$ be meromorphic with poles in Ω. The singular expansion of $\phi(s)$ is*

$$\phi(s) \asymp \sum_{s_0\in\Omega} \Delta(s; s_0),$$

where $\Delta(s, s_0)$ is the principal part of the Laurent expansion of ϕ around the pole $s = s_0$.

Example 3.3.6. *Given $\phi(s) = \frac{1}{s^2(s+1)}$, since*

$$\frac{1}{s^2(s+1)} = \frac{1}{s+1} + \mathcal{O}(1) \quad \text{as } s \to -1$$

and

$$\frac{1}{s^2(s+1)} = \frac{1}{s^2} - \frac{1}{s} + \mathcal{O}(1),$$

we have

$$\phi(s) \asymp \left[\frac{1}{s+1}\right]_{s=-1} + \left[\frac{1}{s^2} - \frac{1}{s}\right]_{s=0}.$$

With the help of the singular expansion of a function, we now give the correspondence.

Theorem 3.3.7. *(Direct mapping) Let $f(z)$ be a function with transform $f^*(s)$ in the fundamental strip $\langle \alpha, \beta \rangle$.*

(i) Assume that, as $z \to 0^+$, $f(z)$ admits a finite asymptotic expansion of the form

$$f(z) = \sum_{(\xi,k)\in\mathcal{A}} c_{\xi,k} z^{\xi} (\log z)^k + \mathcal{O}(z^{\gamma}),$$

where $-\gamma < -\xi \le \alpha$ and the $k's$ are nonnegative. Then $f^*(s)$ is continuable to a meromorphic function in the strip $\langle -\gamma, \beta \rangle$ where it admits a singular expansion

$$f^*(s) \asymp \sum_{(\xi,k)\in\mathcal{A}} c_{\xi,k} \frac{(-1)^k k!}{(s+\xi)^{k+1}}, \quad s \in \langle -\gamma, \beta \rangle.$$

(ii) Similarly, assume that as $z \to +\infty$, $f(z)$ admits a finite asymptotic expansion of the same form where now $\beta \le -\xi < -\gamma$. then $f^*(s)$ is continuable to a meromorphic function in the strip $\langle \alpha, -\gamma \rangle$ where

$$f^*(s) \asymp -\sum_{(\xi,k)\in\mathcal{A}} c_{\xi,k} \frac{(-1)^k k!}{(s+\xi)^{k+1}}, \quad s \in \langle \alpha, -\gamma \rangle.$$

The direct mapping theorem gives a connection from the asymptotic expansion of the function to the singular expansion of the Mellin transform. Thus, the natural question is whether or not there is a way to get the asymptotic expansion from the singular expansion. The answer is yes, when the Mellin transform is small enough at $\pm i\infty$. More precisely, we have the following result.

Theorem 3.3.8. *(Converse mapping) Let $f(z)$ be continuous in $[0, +\infty)$ with Mellin transform $f^*(s)$ having a nonempty fundamental strip $\langle \alpha, \beta \rangle$.*

(i) Assume that $f^*(s)$ admits a meromorphic continuation to the strip $\langle \gamma, \beta \rangle$ for some $\gamma < \alpha$ with a finite number of poles there, and is analytic on $\Re(s) = \gamma$. Assume also that there exists a real number $\eta \in (\alpha, \beta)$ such that

$$f^*(s) = \mathcal{O}(|s|^{-r}) \quad \textit{with } r > 1,$$

when $|s| \to \infty$ *in* $\gamma \le \Re(s) \le \eta$*. If* $f^*(s)$ *admits the singular expansion for* $\Re(s) \in \langle \gamma, \alpha \rangle$*,*

$$f^*(s) \asymp \sum_{(\xi,k)\in\mathcal{A}} d_{\xi,k} \frac{1}{(s-\xi)^k},$$

then an asymptotic expansion of $f(z)$ *at* 0 *is*

$$f(z) = \sum_{(\xi,k)\in\mathcal{A}} d_{\xi,k} \frac{(-1)^{k-1}}{(k-1)!} z^{-\xi} (\log z)^{k-1} + \mathcal{O}(z^{-\gamma}).$$

(ii) Similarly assume that $f^*(s)$ *admits a meromorphic continuation to* $\langle \alpha, \gamma \rangle$ *for some* $\gamma > \beta$ *and is analytic on* $\Re(s) = \gamma$*. Assume also that*

$$f^*(s) = \mathcal{O}(|s|^{-r}) \quad \text{with } r > 1,$$

for $\eta \le \Re(s) \le \gamma$ *with* $\eta \in (\alpha, \beta)$*. If* $f^*(s)$ *admits the singular expansion of the same form for* $\Re(s) \in \langle \eta, \gamma \rangle$*, then an asymptotic expansion of* $f(z)$ *at* ∞ *is*

$$f(z) = -\sum_{(\xi,k)\in\mathcal{A}} d_{\xi,k} \frac{(-1)^{k-1}}{(k-1)!} z^{-\xi} (\log z)^{k-1} + \mathcal{O}(z^{-\gamma}).$$

Note that to apply Theorem 3.3.8, we need the condition

$$f^*(s) = \mathcal{O}(|s|^{-r}) \quad \text{with } r > 1,$$

for s in a suitable strip. In other words, we need the Mellin transforms to be sufficiently small along an vertical line. It is well-known that the smallness of a Mellin transform is directly related to the degree of "smoothness" (differentiability, analyticity) of the original function. Here we introduce two theorems which are useful in determining the "smallness" of the Mellin transforms.

Theorem 3.3.9. *Let* $f(x) \in \mathcal{C}^r$ *with fundamental strip* $\langle \alpha, \beta \rangle$*. Assume that* $f(x)$ *admits an asymptotic expansion as* $x \to 0^+$ *(*$x \to \infty$*) of the form*

$$f(x) = \sum_{(\xi,k)\in\mathcal{A}} c_{\xi,k} x^{\xi} (\log x)^k + \mathcal{O}(x^{\gamma}) \tag{3.7}$$

where the ξ *satisfy* $-\alpha \le \xi < \gamma$ *(*$\gamma < \xi \le -\beta$*). Assume also that each derivative* $f^{(j)}(x)$ *for* $j = 1, \ldots, r$ *satisfies an asymptotic expansion obtained by termwise differentiation of (3.7). Then the continuation of* $f^*(s)$ *satisfies*

$$f^*(\sigma + it) = o(|t|^{-r}) \quad \text{as } |t| \to \infty$$

uniformly for σ *in any closed subinterval of* $(-\gamma, \beta)$ *(*$(\alpha, -\gamma)$*).*

Theorem 3.3.9 shows that smoothness implies smallness. The strongest possible form of smoothness for a function is analyticity. The following theorem shows that the Mellin transform of an analytic function will decay exponentially in a quantifiable way.

Theorem 3.3.10. *Let $f(z)$ be analytic in a sector S_θ which is defined as*

$$S_\theta = \{z \in \mathbb{C} | 0 < |z| < \infty \text{ and } |\arg(z)| \leq \theta\} \quad \text{with } 0 < \theta < \pi.$$

Assume that for $f(z) = \mathcal{O}(|z|^{-\alpha})$ as $|z| \to 0$ in S_θ, and $f(z) = \mathcal{O}(|z|^{-\beta})$ as $|z| \to \infty$ in S_θ. Then

$$f^*(\sigma + it) = \mathcal{O}(e^{-\theta|t|}) \tag{3.8}$$

uniformly for σ in every closed subinterval of (α, β).

Note that a polynomial bound with positive power is enough for Theorem 3.3.8 while (3.8) provides an exponential bound. In fact, if the assumption

$$f^*(s) = \mathcal{O}(|s|^{-r}) \quad \text{with } r > 1,$$

in Theorem 3.3.8 is replaced by

$$f^*(\sigma + it) = \mathcal{O}(e^{-\theta|t|}),$$

we will get a stronger version of the converse mapping theorem as follows.

Theorem 3.3.11. *Let $f(z)$ be continuous in $[0, +\infty)$ with Mellin transform $f^*(s)$ having a nonempty fundamental strip $\langle \alpha, \beta \rangle$.*

(i) Assume that $f^*(s)$ admits a meromorphic continuation to the strip $\langle \gamma, \beta \rangle$ for some $\gamma < \alpha$ with a finite number of poles there, and is analytic on $\Re(s) = \gamma$. Assume also that there exists a real number $\eta \in (\alpha, \beta)$ such that

$$f^*(\sigma + it) = \mathcal{O}(e^{-\theta|t|}), \quad 0 < \theta < \pi$$

when $|t| \to \infty$ in $\gamma \leq \Re(s) \leq \eta$. If $f^*(s)$ admits the singular expansion for $\Re(s) \in \langle \gamma, \alpha \rangle$,

$$f^*(s) \asymp \sum_{(\xi,k)\in\mathcal{A}} d_{\xi,k} \frac{1}{(s-\xi)^k},$$

then an asymptotic expansion of $f(z)$ at 0 is

$$f(z) = \sum_{(\xi,k)\in\mathcal{A}} d_{\xi,k} \frac{(-1)^{k-1}}{(k-1)!} z^{-\xi} (\log z)^{k-1} + \mathcal{O}(z^{-\gamma})$$

and the asymptotic expansion holds in the cone

$$S_\theta = \{z \in \mathbb{C} | 0 < |z| < \infty \text{ and } |\arg(z)| \leq \theta\} \quad \text{with } 0 < \theta < \pi.$$

(ii) Similarly assume that $f^*(s)$ *admits a meromorphic continuation to* $\langle \alpha, \gamma \rangle$ *for some* $\gamma > \beta$ *and is analytic on* $\Re(s) = \gamma$. *Assume also that*

$$f^*(s) = \mathcal{O}(|s|^{-r}) \quad \textit{with } r > 1,$$

for $\eta \leq \Re(s) \leq \gamma$ *with* $\eta \in (\alpha, \beta)$. *If* $f^*(s)$ *admits the singular expansion of the same form for* $\Re(s) \in \langle \eta, \gamma \rangle$, *then an asymptotic expansion of* $f(z)$ *at* ∞ *is*

$$f(z) = - \sum_{(\xi,k) \in \mathcal{A}} d_{\xi,k} \frac{(-1)^{k-1}}{(k-1)!} z^{-\xi} (\log z)^{k-1} + \mathcal{O}(z^{-\gamma})$$

and the asymptotic expansion holds in the cone

$$S_\theta = \{z \in \mathbb{C} | 0 < |z| < \infty \textit{ and } |\arg(z)| \leq \theta\} \quad \textit{with } 0 < \theta < \pi.$$

The advantage of asymptotic expansions holding in a cone in the complex plane is that asymptotic expressions of the derivative are obtained by term-by-term differentiation (the same is not true for asymptotic expansions which just hold on the real line). The justification of this follows from a useful theorem due to J. Ritt [163].

Theorem 3.3.12. *Let* $f(z)$ *be analytic in an annular sector* $\mathcal{S}_R$ *which is defined as*

$$S_R = \{z \in \mathbb{C} | R < |z| < \infty \textit{ and } \theta_1 < \arg(z) \leq \theta_2\} \quad \textit{for some } \theta_1, \theta_2 \textit{ and } 0 \leq R.$$

If for some fixed real number p, *we have*

$$f(z) = \mathcal{O}(z^p) \quad (\textit{or } f(z) = o(x^p)) \quad \textit{as } z \to \infty \textit{ in } S_R,$$

then

$$f^{(m)}(z) = \mathcal{O}(z^{p-m}) \quad (\textit{ or } f^{(m)}(z) = 0(x^{p-m})) \quad \textit{as } z \to \infty$$

in any closed annular sector properly interior to S_R.

Now, with the knowledge of the Mellin transform, we can handle the functional equation (3.5). From the assumptions of Theorem 3.2.3 and the fact that $P_0 = P_1 = 0$, we have that

$$\tilde{f}(z) = \begin{cases} \mathcal{O}(z^2), & \text{as } z \to 0^+, \\ \mathcal{O}(z^{1+\epsilon}), & \text{as } z \to \infty. \end{cases}$$

Now, applying the Mellin transform on (3.5), we obtain for $\Re(s) \in \langle -2, -1 - \epsilon \rangle$

$$\tilde{f}^*(s) = \frac{-\Gamma(s+1)}{1 - p^{-s} - q^{-s}}. \tag{3.9}$$

For the sake of simplicity, we assume that $p = q = 1/2$, then (3.9) become

$$\tilde{f}^*(s) = \frac{-\Gamma(s+1)}{1-2^{s+1}}.$$

We let $\chi_k = 2k\pi i/\log 2$ for $k \in \mathbb{Z}$. By a simple computation, we get

$$\tilde{f}^*(s) \asymp \frac{1}{(s+1)^2}\frac{1}{\log 2} - \frac{1}{s+1}\left(\frac{\gamma}{\log 2} + \frac{1}{2}\right) - \frac{1}{\log 2}\sum_{k\in\mathbb{Z}\backslash\{0\}} \frac{\Gamma(\chi_k)}{s+1-\chi_k}.$$

From [54], we have that the gamma function admits a bound

$$|\Gamma(\sigma + it)| = \mathcal{O}\left(|t|^{\sigma-1/2}e^{-\pi|t|/2}\right), \qquad \text{as } |t| \to \infty.$$

Thus, we can apply Theorem 3.3.8. This plus the result from depoissonization in (3.6) yields

$$\mathbb{E}(P_n) = \tilde{f}(n) = n\log_2 n + n\left(\frac{\gamma}{\log 2} + \frac{1}{2} + P(\log_2 n)\right) + o(n),$$

where $P(t)$ is a 1-periodic function with the Fourier expansion given by

$$P(t) = \sum_{k\in\mathbb{Z}\backslash\{0\}} \frac{\Gamma(-\chi_k)}{\log 2} e^{2k\pi it}.$$

With the mean of the external path length of symmetric tries solved, let us turn our attention back to the internal path length of symmetric DSTs. Apart from the Rice method, P. Flajolet and B. Richmond proposed another method to handle such problems in [66]. The Flajolet-Richmond method is a combination of the Euler transform, the Mellin transform and the singularity analysis. Before we explain their approach, we introduce singularity analysis.

3.4 Singularity Analysis

It has been recognized for a long time that generating function's dominant singularities (the ones with smallest modulus) contains a great deal of information on the coefficients. Therefore, studying the singularities of generating function may give us how the number of objects which the generating function is counting will behave in the long term. Although the idea has been recognized a long time ago, there was no systematical research about this subject until P. Flajolet and A. M. Odlyzko constructed the theory of singularity analysis [64]. Nowadays, singularity analysis is one the the most often used techniques in the analysis of algorithms. Here, we will briefly explain how to apply this method. For a more comprehensive introduction to the whole theory, see [70].

Many combinatorial counting problems with a solution a_n depending on n and satisfying certain recursion relation can be solved by introducing the generating function

$$f(z) = \sum_{n\geq 0} a_n z^n.$$

Then, the desired result can be retrieved by

$$a_n = [z^n]f(z).$$

There are many methods to retrieve the coefficient. A method which is much more productive than elementary real analysis techniques is to use the Cauchy's coefficient formula:

$$[z^n]f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z^{n+1}}dz.$$

As an example, we use $f(z) = (1-z)^{-\alpha}$ with $\alpha > 0$ to illustrate the idea. We choose the contour γ at a distance $1/n$ from the singularity $z = 1$. Then, by using the change of variables $z = 1 + t/n$, we get that $dz = \frac{dt}{n}$, $(1-z)^{-\alpha} = n^\alpha(-t)^{-\alpha}$ and

$$\frac{1}{z^{n+1}} \xrightarrow{n\to\infty} e^{-t}.$$

This gives us (for a rigorous proof, see [70])

$$[z^n](1-z)^{-\alpha} \sim g_a n^{\alpha-1}, \quad \text{where} \quad g_a := \frac{1}{2\pi i}\int_{\mathcal{H}} e^{-t}(-t)^{-\alpha}dt,$$

with $\mathcal{H}$ being the Hankel contour. We recall the Hankel's integral representation of $\Gamma(\alpha)$:

$$\frac{1}{\Gamma(\alpha)} = \frac{1}{2\pi i}\int_{\mathcal{H}} e^{-t}(-t)^{-\alpha}dt.$$

Thus,

$$[z^n](1-z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}.$$

Utilizing the same idea for logarithmic factors with singularities at 1, we get the following theorem.

Theorem 3.4.1. *Let α be an arbitrary complex number in $\mathbb{C}\backslash\mathbb{Z}_{\leq 0}$ and $\beta \in \mathbb{R}$. The coefficient of z^n in the function of the form*

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

admits for large n a full asymptotic expansion in descending power of $\log n$,

$$f_n \equiv [z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}(\log n)^{\beta}\left[1 + \frac{C_1}{\log n} + \frac{C_2}{\log^2 n} + \dots\right],$$

where

$$C_k = \binom{\beta}{k}\Gamma(\alpha)\frac{d^k}{ds^k}\left.\frac{1}{\Gamma(s)}\right|_{s=\alpha}.$$

Remark 2. In many situations, the location of the dominating singularity will not be at 1. However, we can easily shift the location of the dominating singularity to 1 and then apply Theorem 3.4.1. Suppose that

$$f(z) = \left(1 - \frac{z}{\xi}\right)^{-\alpha}\left(\frac{\xi}{z}\log\frac{1}{1-\frac{z}{\xi}}\right)^{\beta},$$

then we have

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

and hence

$$[z^n]f(z) = \xi^{-n}[z^n]f(\xi z) = \xi^{-n}[z^n](1-z)^{-\alpha}\left(\frac{1}{z}\log\frac{1}{1-z}\right)^{\beta}.$$

Example 3.4.2. *Planted trees, sometimes also called Catalan trees, are rooted trees where each node has an arbitrary number of children and subtrees have a natural left-to-right-order. Let f_n be the number of planted trees with n nodes and $f(z) = \sum_{n\geq 0} f_n z^n$. It is well known that*

$$f(z) = \frac{1-\sqrt{1-4z}}{2}.$$

By Theorem 3.4.1, we get

$$[z^n]f(z) = 4^n[z^n]f\left(\frac{z}{4}\right) = 4^n[z^n]\frac{-\sqrt{1-z}}{2} \sim 4^{n-1}\frac{n^{-3/2}}{\sqrt{\pi}}.$$

Theorem 3.4.1 gives us a way to derive asymptotic expansions of the coefficients of generating functions satisfying a certain form. However, generating functions do not always admit such an elegant expression in practical cases. For general use, we usually expand the generating function $f(z)$ near the dominant singularity in the form

$$f(z) = g(z) + \mathcal{O}(h(z)) \quad \text{or} \quad f(z) = g(z) + o(h(z)),$$

where $\tilde{h}(z)$ is a function of the above form and $g(z)$ is written as a linear combination of functions of the above form. What is required at this stage is a way to extract coefficients of error terms. For this purpose, assumptions on $h(z)$ are necessary.

One such assumption is to assume that $h(z)$ is analytic in the complex plane slit at the half line $\mathbb{R}_{\geq 1}$. In fact, weaker conditions suffice: any domain whose boundary makes an acute angle with the half line appears to be suitable.

Definition 3.4.3. *Given two number ϕ, R with $R > 1$ and $0 < \phi < \frac{\pi}{2}$, the open domain $\Delta(\phi, R)$ is defined as*

$$\Delta(\phi, R) = \{z : |z| < R, z \neq 1, |\arg(z - 1)| > \phi\}.$$

A domain is a Δ-domain at 1 if it is a $\Delta(\phi, R)$ for some R and ϕ. For a complex number $\zeta \neq 0$, a Δ-domain at ζ is the image by the mapping $z \mapsto \zeta z$ of a Δ-domain at 1. A function is Δ-analytic if it is analytic in some Δ-domain.

With the definitions of Δ-domain and Δ-analytic, we may now introduce the transfer theorem for the error terms.

Theorem 3.4.4. *Let α, β be arbitrary real numbers, $\alpha, \beta \in \mathbb{R}$ and let $f(z)$ be a function that is Δ-analytic.*

(i) Assume that $f(z)$ satisfies in the intersection of a neighborhood of 1 with its Δ-domain the condition

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

Then, one has

$$[z^n]f(z) = \mathcal{O}(n^{\alpha-1}(\log n)^{\beta}).$$

(ii) Assume that $f(z)$ satisfies in the intersection of a neighborhood of 1 with its Δ-domain the condition

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

Then, one has

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

Example 3.4.5. *We let f_n be the number of labeled 2-regular graphs with n vertices and $f(z) = \sum_{n\geq 0}\frac{f_n}{n!}z^n$ be the exponential generating function of f_n. Then, by symbolic combinatorics (for more details, see [70]), we get*

$$f(z) = \frac{\exp(-\frac{2z-z^2}{4})}{\sqrt{1-z}}.$$

Expanding the numerator around $z = 1$, we have

$$f(z) = e^{-3/4}(1-z)^{-1/2} + \mathcal{O}\left((1-z)^{1/2}\right).$$

Now, an application of Theorem 3.4.1 and Theorem 3.4.4 yields

$$\frac{f_n}{n!} = [z^n]f(z) = \frac{e^{-3/4}}{\sqrt{n\pi}} + \mathcal{O}(n^{-3/2}).$$

In Example 3.1.2, we derived the asymptotic expression of the total path length of symmetric DSTs via the Rice method. Now, as promised in the previous sections, we display how the Flajolet-Richmond approach works by deriving the asymptotic expression of the total path length of symmetric DSTs again.

Example 3.4.6. *(*Flajolet-Richmond approach for the total path length of DSTs*)*
As in Example 3.1.2, we let S_n be the mean of the total path length of symmetric DSTs built on n strings. We also let $A(z) := \sum_n S_n z^n$. Now, we apply the Flajolet-Richmond approach to derive the asymptotic expression of S_n by the following steps:

(1) Euler Transform. *We apply the Euler transform on $A(z)$ by letting*

$$\hat{A}(s) = \frac{1}{s+1}A\left(\frac{1}{s+1}\right).$$

Then, from (3.2), we get

$$(s+1)\hat{A}(s) = 4\hat{A}(2s) + s^{-2}. \tag{3.10}$$

(2) Normalization. *We denote by* $\bar{A}(s) = \hat{A}/Q(-s)$*, where* $Q(-s)$ *is defined in Step 2 of Example 3.1.2. Dividing both sides of (3.10) by* $Q(-2s)$*, we get*

$$\bar{A}(s) = 4\bar{A}(2s) + \frac{1}{s^2 Q(-2s)}. \tag{3.11}$$

(3) Mellin Transform. *By applying Mellin transform on (3.11) and the results in [68], for* $\Re(\omega) > 2$*, we have*

$$\mathscr{M}[\bar{A};\omega] = \frac{G_E(\omega)}{1-2^{2-\omega}},$$

where

$$G_E(\omega) = \frac{Q(2^{\omega-2})}{Q(1)}\Gamma(\omega)\Gamma(1-\omega).$$

Then, the inverse Mellin transform yields as $s \to 0$

$$\begin{aligned}\bar{A}(s) =& s^{-2}\log_2\frac{1}{s} + \frac{1}{s^2}\left(\frac{1}{2} - \alpha\right)\\ &+ \frac{1}{\log 2}\sum_{k\in\mathbb{Z}\setminus\{0\}} G_E(2+\chi_k)s^{-2-\chi_k} + \mathcal{O}(|s|^{-1}),\end{aligned} \tag{3.12}$$

where α *is defined as in Example 3.1.2 and* $\chi_k = 2k\pi i/\log 2$*.*

(4) Asymptotic for the Ordinary Generating Function. *We multiply both sides of (3.12) by* $Q(-2s)$ *and reverse the Euler transform by*

$$A(z) = \frac{1}{z}\hat{A}\left(\frac{1-z}{z}\right).$$

From the fact that $Q(-2s) = 1 + \mathcal{O}(|s|)$*, we get*

$$\begin{aligned}A(z) =& \frac{z}{(1-z)^2}\log_2\frac{z}{1-z} + \frac{z}{(1-z)^2}\left(\frac{1}{2} - \alpha\right)\\ &+ \sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{G_E(2+\chi_k)}{\log 2}\frac{z^{1+\chi_k}}{(1-z)^{2+\chi_k}} + \mathcal{O}(|1-z|^{-1}).\end{aligned} \tag{3.13}$$

(5) Singularity Analysis. *Now, we handle the terms in (3.13) individu-*

ally. First,

$$\begin{aligned}\frac{z}{(1-z)^2}\log_2\frac{z}{1-z} =&\frac{z}{(1-z)^2}\left(\log_2 z+\log_2\frac{1}{1-z}\right)\\ =&\frac{1-(1-z)}{(1-z)^2}\left(\frac{1-(1-z)}{\log 2}\left(\frac{1}{z}\log\frac{1}{1-z}\right)\right.\\ &\left.+\mathcal{O}((1-z)^{-1})\right)\\ =&\frac{1}{\log 2}(1-z)^{-2}\left(\frac{1}{z}\log\frac{1}{1-z}\right)\\ &-\frac{2}{\log 2}(1-z)^{-1}\left(\frac{1}{z}\log\frac{1}{1-z}\right)\\ &+\mathcal{O}((1-z)^{-1}).\end{aligned}$$

By Theorem 3.4.1, we get that

$$[z^n]\frac{z}{(1-z)^2}\log_2\frac{z}{1-z}=n\log_2 n+n\frac{\gamma-1}{\log 2}+\mathcal{O}(n^{1-\epsilon}).$$

Similarly, we get

$$\begin{aligned}[z^n]\frac{z}{(1-z)^2}=&[z^n](1-z)^{-2}(1-(1-z))=n+\mathcal{O}(n^{1-\epsilon})\\ [z^n]\frac{z^{1+\chi_k}}{(1-z)^{2+\chi_k}}=&\frac{n^{1+\chi_k}}{\Gamma(2+\chi_k)}+\mathcal{O}(n^{1-\epsilon})\end{aligned}$$

and

$$\mathcal{O}([z^n]|1-z|^{-1})=\mathcal{O}(1).$$

By substituting them back into (3.13), we get

$$\begin{aligned}S_n=&n\log_2 n+n\left(\frac{\gamma-1}{\log 2}+\frac{1}{2}-\alpha\right)\\ &+\frac{n}{\log 2}\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{G_E(2+\chi_k)}{\Gamma(2+\chi_k)}n^{\chi_k}+\mathcal{O}(n^{1-\epsilon}).\end{aligned}$$

Note that the asymptotic expression coincides with (3.3) which was derived by the Rice method.

3.5 Recently developed Methods

In the last several sections, we have introduced four often used methods. We also demonstrated how the methods work by working out asymptotic expressions of the mean for two shape parameters of digital trees. So, it is natural to ask whether or not the variance and higher moments can be derived by the methods as well. The answer is yes. However, the computation can be extremely complicated and tricky.

For example, assume that we try to derive the asymptotic expression of the variance for the total path length of symmetric DSTs. Let P_n be the total path length of symmetric DSTs built on n strings. Then we can use the introduced methods to derive asymptotic expressions for $\mathbb{E}(P_n)$ and $\mathbb{E}(P_n^2)$ and then compute $\mathbb{E}\left(P_n^2\right) - \mathbb{E}(P_n)^2$. However, the order of $\mathbb{E}(P_n)^2$ is $n^2(\log n)^2$ while the order of the variance is n (see [121]). This implies that the first several terms of $\mathbb{E}(P_n)^2$ and $\mathbb{E}\left(P_n^2\right)$ will be canceled. If we do not know the right order of the variance (which normally is the case), we will have to derive very long asymptotic expressions and face many cancelations. This would be extremely difficult not only because deriving long asymptotic expressions can be complicated, but also because the cancelation part often needs some deep knowledge on all kinds of constants, Fourier series and q-analysis.

To avoid such disadvantages and derive the variance more efficiently, a new set of mathematical tools has been proposed by M. Fuchs, H.-K. Hwang and V. Zacharovas in [74]. In this section, we will introduce these tools.

3.5.1 Poissonized Variance with Correction

For a given random variable X_n, we let

$$\tilde{f}_1(z) = e^{-z} \sum_{n \geq 0} \mathbb{E}(X_n) \frac{z^n}{n!}$$

and

$$\tilde{f}_2(z) = e^{-z} \sum_{n \geq 0} \mathbb{E}(X_n^2) \frac{z^n}{n!}.$$

Then, from the definition of the variance and analytical depoissonization, one may guess that if $\tilde{f}_1(z)$ and $\tilde{f}_2(z)$ are smooth enough, then

$$\mathbb{V}(X_n) = \mathbb{E}(X_n^2) - (\mathbb{E}(X_n))^2 \sim \tilde{f}_2(n) - \tilde{f}_1(n)^2. \tag{3.14}$$

This asymptotic equivalent holds for many cases. However, for a large class of problems, $\tilde{f}_2(n) - \tilde{f}_1(n)^2$ is not asymptotically equivalent to the variance. One class of such problem are those with mean and variance satisfying

$$\lim_{n\to\infty} \frac{\log \mathbb{E}(X_n)}{\log n} = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{\log \mathbb{V}(X_n)}{\log n} = 1. \tag{3.15}$$

Examples of these problems are the two shape parameters we discussed before.

To solve this problem, H.-K. Hwang, M. Fuchs and V. Zacharovas proposed the poissonized variance with correction

$$\tilde{V}(z) := \tilde{f}_2(z) - \tilde{f}_2(z)^2 - z\tilde{f}'(z)^2 \tag{3.16}$$

in [74]. For the problems satisfying (3.15), we have that

$$\mathbb{V}(X_n) = \tilde{V}(n) + \mathcal{O}\left((\log n)^c\right)$$

for some $c \geq 0$ under suitable assumptions. Comparing (3.14) and (3.16), the difference is the appearance of the term $z\tilde{f}'(z)^2$. To see why this term is necessary, we introduce the following lemma :

Lemma 3.5.1. *Let $\tilde{f}(z) = e^{-z}\sum_{n\geq 0} a_n \frac{z^n}{n!}$. If $\tilde{f}(z)$ is an entire function, then*

$$a_n = \sum_{j\geq 0} \frac{\tilde{f}^{(j)}(n)}{j!}\tau_j(n), \tag{3.17}$$

where

$$\tau_j(n) := n![z^n](z-n)^j e^z = \sum_{k=0}^{j} \binom{j}{k} (-1)^{j-k} \frac{n! n^{j-k}}{(n-k)!}.$$

Now we let $\tilde{D}(z) := \tilde{f}_2(z) - \tilde{f}_1(z)^2$. By the above lemma we get that

$$\begin{aligned}
\mathbb{V}(X_n) &= \mathbb{E}(X_n^2) - (\mathbb{E}(X_n))^2 \\
&= \sum_{j\geq 0} \frac{\tilde{f}_2^{(j)}(n)}{j!}\tau_j(n) - \left(\sum_{j\geq 0} \frac{\tilde{f}_1^{(j)}(n)}{j!}\tau_j(n)\right)^2 \\
&= \tilde{D}(n) - n\tilde{f}'(n)^2 - \frac{n}{2}\tilde{D}''(n) + \text{smaller order terms.}
\end{aligned}$$

For the case $\tilde{f}_1(n) \asymp n\log n$, we find that $n\tilde{f}_1'(n)^2 \asymp n(\log n)^2$ is of larger order than $\tilde{D}(n)$.

Comparing to other methods like the second moment approach (using $\tilde{f}_2(z)$) or usage of $\tilde{D}(z)$, one of the major advantages of using the poissonized variance with correction is that the computation is largely simplified. Analytical poissonization transfers a problem from Bernoulli model to the Poisson model. While the other approaches apply the analytical depoissonization first and then dealt with many cancelations, the poissonized variance with corrections has already incorporated all the cancelations in the poisson model. See [74] for more comparison between poissonized variance with corrections and other methods.

3.5.2 Poisson-Laplace-Mellin Method

The Poisson-Laplace-Mellin method was developed together with poissonized variance with correction in [74] to deal with shape parameters of symmetric random digital search trees. As its name implies, this approach combines poissonization, Laplace transform and Mellin transform. Before we explain this method step by step, we first need an important lemma.

Lemma 3.5.2. *Let $\tilde{f}(z)$ be a function whose Laplace transform exists and is analytic in $\mathbb{C}\setminus(-\infty,0]$. Assume that*

$$\mathscr{L}[\tilde{f}(z);s]=\begin{cases}\mathcal{O}\left(|s|^{-\alpha}|\log s|^{m}\right),\\ cs^{-\beta}\left(\log\left(\frac{1}{s}\right)\right)^{m},\end{cases}$$

uniformly for $s\to\infty$ with $|\arg(s)|\le\pi-\epsilon$, where $\alpha\in\mathbb{R}$, $\beta\in\mathbb{C}$ and $m\ge 0$ is an integer. Moreover, assume that

$$\mathscr{L}[\tilde{f}(z);s]=\mathcal{O}(|s|^{-1-\epsilon})$$

uniformly for $s\to\infty$ with $|\arg(s)|\le\pi-\epsilon$. Then,

$$\tilde{f}(z)=\begin{cases}\mathcal{O}(|z|^{\alpha-1}|\log z|^{m}),\\[2ex] cz^{\beta-1}\displaystyle\sum_{j=0}^{m}\binom{m}{j}(\log z)^{m-j}\frac{\partial^{j}}{\partial\omega^{j}}\frac{1}{\Gamma(\omega)}\bigg|_{\omega=\beta},\end{cases}$$

uniformly as $z\to\infty$ and $|\arg(z)|\le\frac{\pi}{2}-\epsilon$.

Now, we give a step by step description of how the Poisson-Laplace-Mellin method work:

1. Use the Poisson generating functions of the first and second moment. The Poisson generating function of both the mean and the variance will satisfy a differential-functional equation of the form
$$\tilde{f}(z)+\tilde{f}'(z)=2\tilde{f}(z/2)+\tilde{t}(z),\tag{3.18}$$
where $\tilde{t}(z)$ is some suitable function.

2. Substitute the Poisson generating function of the first and second moment into the formula of poissonized variance with correction. The poissonized variance with corrections will also satisfy a differential-functional equation of the type (3.18).

3. Now, we have to asymptotically solve two differential-functional equations of the form in (3.18). We apply Laplace transform to (3.18) to get rid of the differential operator. The resulting functional equation will be
$$(1+s)\mathscr{L}[\tilde{f}(z);s] = 4\mathscr{L}[\tilde{f}(z);2s] + \mathscr{L}[\tilde{t}(z);s]. \tag{3.19}$$

4. We let
$$Q(s) = \prod_{k\geq 1}\left(1-\frac{s}{2^k}\right), \qquad \bar{\mathscr{L}}[\tilde{f}(z);s] = \frac{\mathscr{L}[\tilde{f}(z);s]}{Q(-s)},$$
and
$$\bar{\mathscr{L}}[\tilde{t}(z);s] = \frac{\mathscr{L}[\tilde{t}(z);s]}{Q(-2s)}.$$
Dividing both sides of (3.19) by $Q(-2s)$, we will get a simplified functional equation
$$\bar{\mathscr{L}}[\tilde{f}(z);s] = 4\bar{\mathscr{L}}[\tilde{f}(z);2s] + \bar{\mathscr{L}}[\tilde{t}(z);s]. \tag{3.20}$$

5. Apply the Mellin transform to (3.20), which yields
$$\mathscr{M}[\bar{\mathscr{L}}[\tilde{f}(z);s];\omega] = \frac{\mathscr{M}[\bar{\mathscr{L}}[\tilde{t}(z);s];\omega]}{1-2^{2-\omega}}.$$
By the standard theory of inverse Mellin transform which we have introduced in Section 3.3, we derive an asymptotic expansion of $\bar{\mathscr{L}}[\tilde{f}(z);s]$ as $s\to 0$.

6. Applying Lemma 3.5.2 to the asymptotic expansion of $\bar{\mathscr{L}}[\tilde{f}(z);s]$ will give an asymptotic expansion of $\tilde{f}(z)$ as $z\to\infty$.

7. The last step is to apply analytical depoissonization in order to get the desired results from the asymptotic expansions of $\tilde{f}(z)$. For this, one may use the standard method from [99] or the theory of JS-admissiblility which will be explained below.

3.5.3 JS-admissibility

For analytical depoissonization, the standard theory by P. Jacquet and W. Szpankowski was introduced in Section 3.2. However, there are many complicated conditions to be checked before using the depoissonization lemmas from Section 3.2 and similar results in [99]. Thus, the authors of [74] proposed a systematic method for this which they called the JS-admissibility. We start with the following definition which arises from Theorem 3.2.4.

Definition 3.5.3. *We let $\epsilon, \epsilon' \in (0,1)$ be arbitrarily small numbers. An entire function $\tilde{f}$ is said to be JS-admissible, denoted by $\tilde{f} \in \mathscr{J}\mathscr{S}$, if the following two conditions hold for $|z| \geq 1$.*

(I) There exists $\alpha, \beta \in \mathbb{R}$ such that uniformly for $|\arg(z)| \leq \epsilon$,

$$\tilde{f}(z) = \mathcal{O}\left(|z|^{\alpha}(\log_{+}|z|)^{\beta}\right),$$

where $\log_{+} x := \log(1+x)$.

(O) Uniformly for $\epsilon \leq |\arg(z)| \leq \pi$,

$$f(z) := e^{z}\tilde{f}(z) = \mathcal{O}\left(e^{(1-\epsilon')|z|}\right).$$

Then Theorem 3.2.4 can be reformulated as follows.

Lemma 3.5.4. *Assume $\tilde{f} \in \mathscr{J}\mathscr{S}$. Let $f(z) = e^{z}\tilde{f}(z)$, then we have that*

$$\begin{aligned} a_n := f^{(n)}(0) = n![z^n]f(z) &= n![z^n]e^{z}\tilde{f}(z) \\ &= \sum_{j=0}^{2k} \frac{\tilde{f}^{(j)}(n)}{j!}\tau_j(n) + \mathcal{O}\left(n^{\alpha-k}(\log n)^{\beta}\right) \end{aligned}$$

for $k = 1, 2, \ldots$.

The real advantage of introducing admissibility is that it opens the possibility of developing closure properties as we discuss here.

Lemma 3.5.5. *Let m be a nonnegative integer and $\alpha \in (0,1)$, we have the following properties.*

1. *z^m, $e^{-\alpha z} \in \mathscr{J}\mathscr{S}$.*
2. *If $\tilde{f} \in \mathscr{J}\mathscr{S}$, then $\tilde{f}(\alpha z)$, $z^m\tilde{f} \in \mathscr{J}\mathscr{S}$.*
3. *If $\tilde{f}, \tilde{g} \in \mathscr{J}\mathscr{S}$, then $\tilde{f} + \tilde{g} \in \mathscr{J}\mathscr{S}$.*
4. *If $\tilde{f} \in \mathscr{J}\mathscr{S}$ and $\tilde{P}$ is a polynomial, then the product $\tilde{P}\tilde{f} \in \mathscr{J}\mathscr{S}$.*
5. *If $\tilde{f}, \tilde{g} \in \mathscr{J}\mathscr{S}$, then $\tilde{h}(z) = \tilde{f}(\alpha z)\tilde{g}\left((1-\alpha)z\right) \in \mathscr{J}\mathscr{S}$.*
6. *If $\tilde{f} \in \mathscr{J}\mathscr{S}$, then $\tilde{f}' \in \mathscr{J}\mathscr{S}$ and thus $\tilde{f}^{(m)} \in \mathscr{J}\mathscr{S}$.*

In the last step of the Poisson-Laplace-Mellin method we have mentioned that the depoissonization step can be finished by JS-admissibility. Here we use the closure properties together with the following Proposition.

Proposition 3.5.6. *Let $\tilde{f}$ and $\tilde{g}$ be entire functions satisfying*

$$\tilde{f}(z) + \tilde{f}'(z) = 2\tilde{f}(z/2) + \tilde{g}(z),$$

with $\tilde{f}(0) = 0$, then

$$\tilde{g} \in \mathscr{J}\mathscr{S} \quad \text{if and only if} \quad \tilde{f} \in \mathscr{J}\mathscr{S}.$$

With this proposition, for all differential functional equation of the form of (3.18), we only need to check whether $\tilde{t}(z)$ is JS-admissible to finish the depoissonization step.

3.6 Contraction Method

Since introduced in the 1960s by Knuth [125, 126, 127], probabilistic analysis of algorithms has been mostly depending on analytic techniques for generating functions. However, over the last decade of the 20th century, among other probabilistic techniques, the so called contraction method has been developed. The contraction method was first proposed by Rösler to analyze Quicksort [185]. It was then further developed independently by Rösler [186] and Rachev and Rüschendorf [180], and later on in Rösler [187] and Neininger and Rüschendorf [160, 161]. See also the survey article by Rösler and Rüschendorf [183].

The contraction method is used to prove that certain sequences of random variables which satisfies a distributional recurrences will converge to a fixed point which is then shown to be the limiting distribution. The method was first used for univariate cases and then generalized to multivariate cases. Here we will introduce the multivariate version with the univariate version as a special case. Before we discuss the settings, we need some definitions.

Definition 3.6.1. *We denote by $\mathcal{M}^d$ the space of all probability measures on $\mathbb{R}^d$. For $\mu, \nu \in \mathcal{M}^d$ with $X \sim \mu$ and $Y \sim \nu$, the Zolotarev metric ζ_s with $s > 0$ is defined by*

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}(f(X) - f(Y))|,$$

where $s = m + \alpha$, $0 < \alpha \leq 1$, $m \in \mathbb{N}_0$, and

$$\mathcal{F}_s := \left\{ f \in C^m(\mathbb{R}^d, \mathbb{R}) : |f^{(m)}(x) - f^{(m)}(y)| \leq \|x - y\|^\alpha \right\}.$$

Here, $C^m(\mathbb{R}^d, \mathbb{R})$ denotes the space of m times differentiable functions.

Let us now consider a sequence of d-dimensional random vectors $\{Y_n\}_{n\geq 0}$ which satisfy the distributional recursion

$$Y_n \overset{d}{=} \sum_{r=1}^{k} A_r(n) Y^{(r)}_{I_r^{(n)}} + b_n, \quad n \geq n_0,$$

where

1. $I_r^{(n)}$ is a vector of random cardinalities with $I_r^{(n)} \in \{0, \dots, n\}$
2. $\left(A_1(n), \dots, A_k(n), b_n, I_1^{(n)}, \dots, I_k^{(n)}\right), (Y_n^{(1)}), \dots, (Y_n^{(k)}), (Y_n)$ are independent,
3. $A_1(n), \dots, A_k(n)$ are random $d \times d$ matrices,
4. b_n is a random d-dimensional vector,
5. $(Y_n^{(1)}), \dots, (Y_n^{(k)})$ are identically distributed as (Y_n).

The symbol $\overset{d}{=}$ denotes equality in distribution and we have $n_0 \geq 0$.

Next, we normalize the Y_n by

$$X_n := C_n^{-1/2}\left(Y_n - M_n\right), \quad n \geq n_0,$$

where $M_n \in \mathbb{R}^d$ and C_n is a positive-definite square matrix. In the case $2 < s \leq 3$, we assume that $\mathrm{Cov}(Y_n)$ is positive definite for $n \geq n_1 \geq n_0$. Our eventual goal is to prove that X_n converges to a fixed point in $\mathcal{M}_s^d(0, \mathrm{Id}_d)$. If the first and second moments of Y_n are finite, we choose M_n and C_n for different s as follows.

$$M_n := \begin{cases} \mathbb{E}(Y_n), \quad C_n := \begin{cases} \mathrm{Id}_d, & \text{for } 0 \leq n < n_1, \\ \mathrm{Cov}(Y_n), & \text{for } n \geq n_1, \end{cases} & \text{if } 2 < s \leq 3, \\ \mathbb{E}(Y_n), \quad C_n \text{ is any positive definite matrix}, & \text{if } 1 < s \leq 2, \end{cases}$$

The normalized quantities X_n then satisfy the modified recurrence

$$X_n \overset{d}{=} \sum_{r=1}^{k} A_r(n) X^{(r)}_{I_r^{(n)}} + b^{(n)}, \quad n \geq n_0,$$

with

$$A_r^{(n)} := C_n^{-1/2} A_r(n) C^{1/2}_{I_r^{(n)}}, \quad b^{(n)} := C_n^{-1/2}\left(b_n - M_n + \sum_{r=1}^{k} A_r(n) M_{I_r^{(n)}}\right)$$

and the independence relations are as for Y_n. The normalized quantities will converge in ζ_s under suitable conditions.

Theorem 3.6.2. *Let (X_n) be normalized as before and s-integrable and $0 < s \le 3$. Assume that as $n \to \infty$,*

1. $\left(A_1^{(n)}, \dots, A_k^{(n)}, b_n\right) \xrightarrow{\mathcal{L}_s} (A_1^*, \dots, A_k^*, b^*)$,
2. $\mathbb{E}\sum_{r=1}^{k} \|A_r^*\|_{op}^s < 1$, *and*
3. $\mathbb{E}\left[\mathbf{1}_{\{I_r^{(n)} \le l\} \cup \{I_r^{(n)} = n\}} \|A_r^{(n)}\|_{op}^s\right] \to 0$ *for all* $l \in \mathbb{N}$ *and* $r = 1, \dots, k$.

Then (X_n) converges to a limit X,

$$\zeta_s(X_n, X) \to 0, \quad n \to \infty.$$

Proof. The proof is quite long and technical and hence we omit it here. See [160] for the complete proof. □

There are many variants of the contraction method for different cases, here we give a specialized version which is very useful for proving central limit theorems.

Corollary 3.6.3. *(Central Limit Theorem) Let (Y_n) be s-integrable, $s > 2$ and satisfies the recurrence*

$$Y_n \overset{d}{=} \sum_{r=1}^{k} Y_{I_r^{(n)}}^{(r)} + b_n, \quad n \ge n_0$$

with

$$\mathbb{E}(Y_n) = f(n) + o(g^{1/2}(n)) \quad \textit{and} \quad \mathrm{Var}(Y_n) = g(n) + o(g(n)),$$

where $g(n) > 0$ for all n huge enough. Assume for all $r = 1, \dots, k$ and some $2 < s \le 3$,

1. $\left(\dfrac{g(I_r^{(n)})}{g(n)}\right)^{1/2} \xrightarrow{\mathcal{L}_s} A_r^*$,
2. $\dfrac{1}{g^{1/2}(n)}\left(b_n - f(n) + \sum_{r=1}^{k} f(I_r^{(n)})\right) \xrightarrow{\mathcal{L}_s} 0$,
3. $\sum_{r=1}^{k} (A_r^*)^2 = 1, \quad \mathbb{P}(\exists r : A_r^* = 1) < 1$.

Then

$$\frac{Y_n - f(n)}{g^{1/2}(n)} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ *denotes the standard normal distribution.*

Example 3.6.4. *We let* Y_n *be the size of random* m*-ary search tree (see [139]) containing* n *data. Then* Y_n *satisfies the recursion*

$$Y_n \overset{d}{=} \sum_{r=1}^{m} Y^{(r)}_{I_r^{(n)}} + 1, \quad n \geq m$$

with initial conditions $Y_0 = 0$ *and* $Y_1 = \cdots = Y_{m-1} = 1$*. We let* $V = (U_{(1)}, U_{(2)} - U_{(1)}, \ldots, 1 - U_{(m-1)})$ *denote the vector of spacings of independent Unif[0,1] random variables* $U_{(1)}, \ldots, U_{(m-1)}$*. Then we have*

$$\frac{1}{n}(I_1^{(n)}, \ldots, I_m^{(n)}) \xrightarrow{\mathcal{L}_{1+\epsilon}} V.$$

From [9, 20, 127, 141], we have that for $3 \leq m \leq 26$*,*

$$\mathbb{E}(Y_n) = \frac{n}{2(H_m - 1)} + \mathcal{O}(1 + n^{\alpha - 1}), \quad \mathrm{Var}(Y_n) = n\gamma_m + o(n),$$

where H_m *is the* m*-th harmonic number,* $\gamma_m > 0$ *and* $\alpha < 3/2$ *are constants depending on* m*. To apply Corollary 3.6.3, we choose*

$$f(n) = \frac{n}{2(H_m - 1)} \quad \text{and} \quad g(n) = \gamma_m n.$$

Then,

$$\frac{g(I_r^{(n)})}{g(n)} = \frac{I_r^{(n)}}{n} \xrightarrow{\mathcal{L}_s} U_{(r)} - U_{(r-1)} = (A_r^*)^2$$

and

$$\begin{aligned}
&\frac{1}{\sqrt{g(n)}}\left(b_n - f(n) + \sum_{r=1}^{k} f(I_r^{(n)})\right) \\
=&\frac{1}{\sqrt{n\gamma_m}}\left(1 - \frac{n}{2(H_m - 1)} + \frac{n-1-m}{2(H_m - 1)}\right) \to 0 \qquad \text{as } n \to \infty.
\end{aligned}$$

Finally,

$$\sum_{r=1}^{m} (A_r^*)^2 = U_{(1)} + \sum_{r=2}^{m-1} \left(U_{(r)} - U_{(r-1)}\right) + 1 - U_{(m-1)} = 1$$

and

$$\mathbb{P}(A_r^* = 1) < 1 \quad \text{for all } r.$$

Thus, all conditions of Corollary 3.6.3 are satisfied and hence we rederived the limit law (see [9, 131, 141]) for Y_n*.*

Chapter 4

New Applications of the Poisson-Laplace-Mellin Method

4.1 Approximate Counting

4.1.1 Introduction

Approximate counting, an algorithm proposed by Morris [149] in 1978, is used for counting within a certain error tolerance a huge amount of objects with very limited space. The algorithm has found many applications such as in the analysis of the Webgraph, monitoring network traffic, finding patterns in protein and DNA sequencing, computing frequency moments of data streams, data storage in flash memory, and many variants and improvements have been proposed; see Csűrös [27], Mitchell and Day [147], Gronemeier and Sauerhoff [84], Aspnes and Censor [8], Cichoń and Macyna [21] and references therein.

Here, we are going to revisit the analysis of the classical algorithm which is described as follows: a counter C_n is maintained with initial value $C_0 = 0$. After "counting n objects", a random decision based only on the current content of the counter determines whether or not the counter should be increased when "counting the $n+1$-st object". More precisely, the counter obeys the following rule

$$C_{n+1} = \begin{cases} C_n + 1, & \text{with probability } q^{C_n}; \\ C_n, & \text{with probability } 1 - q^{C_n}, \end{cases} , \tag{4.1}$$

where $0 < q < 1$ is fixed. Hence, $(C_n)_{n\geq 0}$ is a Markov chain describing a pure birth process. The same chain was also encountered in a couple of other problems: width of greedy decomposition of random acyclic digraphs into node-disjoint paths (see Simon [195]), size of greedy independent set and greedy clique in random graphs (see Simon [195]) and length of the leftmost path in digital search trees (see below).

We mention in passing that many variants of the above Markov chain have been investigated as well; e.g. see Crippa and Simon [26], Louchard and Prodinger [138], Bertoin, Biane and Yor [10] and Guillemin, Robert and Zwart [86]. Applications range from Computer Science over Particle Physics to Molecular Biology; see the detailed discussion in [26].

As for the classical chain C_n, the first detailed analysis was given by Flajolet in [61] who used Mellin transform (see also Prodinger [173] for a similar analysis). Other approaches have been given by Kirschenhofer and Prodinger [115] via Rice method, Prodinger [175] via Euler transform, Louchard and Prodinger [137] via analysis of extreme value distributions, Rosenkrantz [184] via martingale theory and Robert [182] via probabilistic tools. Here, we are going to use the "Poisson-Laplace-Mellin" method described in Section 3.5.2 to analyze approximate counting via the connection between the algorithm and shape parameters of digital search trees.

We now explain the connection between approximate counting and digital search trees equipped with the Bernoulli model (introduced in Chapter 2) in more details. The parameter which is related to approximate counting is the number of vertices on the leftmost path from the root to the leftmost leaf. We denote this length in a random digital search tree of size n by X_n. Obviously, X_n satisfies the following distributional recurrence

$$X_{n+1} \stackrel{d}{=} X_{B_n} + 1, \qquad (n \geq 0) \tag{4.2}$$

with $X_0 = 0$ and $B_n \stackrel{d}{=} \text{Binom}(n, q)$. This recurrence just reflects the trivial fact that X_n can be computed by starting from the root (which counts as 1) and then moving on to the left subtree (which has size B_n) where the same procedure is repeated. Now, a moment's reflection reveals that C_n is related to X_n as

$$C_n \stackrel{d}{=} X_n.$$

This relation will be the starting point of our analysis. We will use it to derive asymptotic expansions for mean and variance of C_n.

Apart from the original approximate counting algorithm, we will also discuss extensions and variations of approximate counting. One such extension was proposed in [21] where instead of one counter, m counters $C_n^{(1)}, \ldots, C_n^{(m)}$ were used (m fixed). Then, when "counting the $n+1$-st object", one of the counters is chosen uniformly at random and increased according to the stochastic rule (4.1).

The m-counter problem has been analyzed in [177], where mean and variance of $D_n := C_n^{(1)} + \cdots + C_n^{(m)}$ were derived. For this variant, we will show that the "Poisson-Laplace-Mellin" method will greatly simplify the analysis since the case of m counters can be reduced to the case of one counter. Moreover, similar simplifications can also be achieved for shape parameters in m-DST trees recently introduced in [178].

Remark 3. Before stating our new result, we explain what is known about C_n. Flajolet in [61] showed that, as $n \to \infty$,

$$\mathbb{E}(C_n) \sim \log_{1/q} n + F_C(\log_{1/q} n),$$

where $F_C(z) = \sum_k f_k e^{2k\pi i}$ is a 1-periodic function with Fourier coefficients

$$f_0 = \frac{\gamma}{\log(1/q)} + \frac{1}{2} - \alpha, \qquad f_k = \frac{\Gamma(-\chi_k)}{\log(1/q)} \quad (k \neq 0),$$

where γ is Euler's constant, $\alpha = \sum_{l\geq 1} q^l/(1-q^l)$ and $\chi_k = 2k\pi i/\log(1/q)$. As for the variance, he showed that, as $n \to \infty$,

$$\mathrm{Var}(C_n) \sim G_C(\log_{1/q} n),$$

where $G_C(z) = \sum_k g_k e^{2k\pi i}$ is again a 1-periodic function with computable Fourier coefficients. Moreover, he gave the following expression for the average value of $G_C(z)$

$$g_0 = \frac{\pi^2}{6\log^2(1/q)} - \alpha - \beta + \frac{1}{12} - \frac{1}{\log(1/q)} \sum_{l\geq 1} \frac{1}{l\sinh(2l\pi^2/\log(1/q))},$$

where $\beta = \sum_{l\geq 1} q^{2l}/(1-q^l)^2$.

4.1.2 Analysis of Approximate Counting

The starting point of our analysis will be (4.2).

Poissonization and depoissonization. First define

$$\tilde{P}(y,z) = e^{-z} \sum_{n\geq 0} \mathbb{E}(e^{X_n y}) \frac{z^n}{n!}.$$

Then, from (4.2), we obtain

$$\tilde{P}(y,z) + \frac{\partial}{\partial z}\tilde{P}(y,z) = e^y \tilde{P}(y,qz)$$

with $\tilde{P}(y,0) = \exp(-z)$.

From this, by differentiation with respect to y and setting $y = 0$, we obtain for the Poisson generating functions of the first and second moment of X_n (denoted by $\tilde{f}_1(z)$ and $\tilde{f}_2(z)$, respectively)

$$\begin{aligned} \tilde{f}_1(z) + \tilde{f}_1'(z) &= \tilde{f}_1(qz) + 1, \\ \tilde{f}_2(z) + \tilde{f}_2'(z) &= \tilde{f}_2(qz) + 2\tilde{f}_1(qz) + 1, \end{aligned} \tag{4.3}$$

with $\tilde{f}_1(0) = \tilde{f}_2(0) = 0$. Moreover, as *poissonized variance with corrections* we introduced in Section 3.5.1 $\tilde{V}(z) := \tilde{f}_2(z) - \tilde{f}_1^2(z)$. Then, the above two relations in turn yield

$$\tilde{V}(z) + \tilde{V}'(z) = \tilde{V}(qz) + \tilde{f}_1'(z)^2 \tag{4.4}$$

with $\tilde{V}(0) = 0$.

By Proposition 3.5.6, we have that $\tilde{f}_1(z)$ and $\tilde{f}_2(z)$ are both JS-admissible. Depoissonization then yields, as $n \to \infty$,

$$\mathbb{E}(X_n) \sim \tilde{f}_1(n) \qquad \text{and} \qquad \mathrm{Var}(X_n) \sim \tilde{V}(n).$$

Thus, we only have to find asymptotics of $\tilde{f}_1(z)$ and $\tilde{V}(z)$.

Analysis of the Mean. Here, we analyze the mean, where we start from (4.3). Since $\tilde{f}_1(z)$ is JS-admissible, we may apply Laplace transform on it to get rid of the differential operator and obtain

$$(s+1)\mathscr{L}[\tilde{f}_1; s] = \frac{1}{q}\mathscr{L}[\tilde{f}_1; -s/q] + 1/s. \tag{4.5}$$

Next, we derive an exact expression for the mean. By iterating the above functional equation, we get

$$\begin{aligned} \mathscr{L}[\tilde{f}_1; s] &= \frac{1}{s}\sum_{j\ge 0} \frac{1}{(s+1)(sq^{-1}+1)\cdots(q^{-j}s+1)} \\ &= \frac{1}{s}\sum_{j\ge 0}\sum_{0\le l\le j} \frac{(-1)^{j-l} q^{\binom{j-l+1}{2}}}{(q^{-l}s+1)Q_l Q_{j-l}} \\ &= \frac{1}{s}\sum_{l\ge 0} \frac{1}{Q_l(q^{-l}s+1)} \sum_{j\ge 0} \frac{(-1)^j q^{\binom{j+1}{2}}}{Q_j} \\ &= \frac{Q_\infty}{s}\sum_{l\ge 0}\frac{1}{Q_l(q^{-l}s+1)}, \end{aligned}$$

where Q_j and Q_∞ have been defined in the introduction. So, by inverse Laplace transform,

$$\tilde{f}_1(z) = Q_\infty \sum_{l\geq 0} \frac{1}{Q_l}(1-e^{-q^l z})$$

and hence

$$\mathbb{E}(X_n) = Q_\infty \sum_{l\geq 0} \frac{1}{Q_l}(1-(1-q^l)^n).$$

We record this result for future reference.

Proposition 4.1.1. *We have,*

$$\mathbb{E}(X_n) = Q_\infty \sum_{l\geq 0} \frac{1}{Q_l}(1-(1-q^l)^n).$$

Next, we derive an asymptotic expansion. Set $\bar{\mathscr{L}}[\tilde{f}_1; s] = \mathscr{L}[\tilde{f}_1; s]/Q(-s)$, where

$$Q(-s) = \prod_{i=1}^{\infty}\left(1+sq^i\right).$$

Then, by dividing (4.5) by $Q(-s/q)$,

$$\bar{\mathscr{L}}[\tilde{f}_1; s] = \frac{1}{q}\bar{\mathscr{L}}[\tilde{f}_1; s/q] + \frac{1}{sQ(-s/q)}.$$

From the fact that $\tilde{f}_1(z)$ is JS-admissible and properties of the well-studied function $Q(-s/q)$ [5], $\bar{\mathscr{L}}[\tilde{f}_1; s]$ admits a polynomial bound as s tends to both zero and ∞. Therefore, we may apply Mellin transform and obtain

$$\mathscr{M}[\bar{\mathscr{L}}; \omega] = \frac{M_1(\omega)}{1-q^{\omega-1}}, \qquad (\Re(\omega) > 1),$$

where

$$M_1(\omega) = \int_0^\infty \frac{s^{\omega-2}}{Q(-s/q)}\mathrm{d}s = \frac{Q(q^{1-\omega})}{Q_\infty}\Gamma(\omega+1)\Gamma(-\omega).$$

Next, from the exponential decay of the Gamma function along vertical lines, we have $M_1(c+it) = O(e^{-(\pi-\epsilon)|t|})$ for $c > 1$ and $|t|$ large. Thus, $M_1(\omega)$ is integrable and hence the inverse Mellin transform exists for $|\arg(s)| \leq \pi - \epsilon$. Using inverse Mellin transform, we obtain that for $|\arg(s)| \leq \pi - \epsilon$ and $|s| \to 0$,

$$\bar{\mathscr{L}}[\tilde{f}_1; s] \sim \frac{1}{s}\log_{1/q}\frac{1}{s} + \frac{1}{s}\left(\frac{1}{2} - \alpha + \frac{1}{L}\sum_{k\neq 0} M_1(1+\chi_k)s^{-\chi_k}\right),$$

where notations are as in the introduction. Thus, for $|\arg(s)| \leq \pi - \epsilon$ and $|s| \to 0$,

$$\mathscr{L}[\tilde{f}_1; s] \sim \frac{1}{s}\log_{1/q}\frac{1}{s} + \frac{1}{s}\left(\frac{1}{2} - \alpha + \frac{1}{L}\sum_{k\neq 0} M_1(1+\chi_k)s^{-\chi_k}\right),$$

By Lemma 3.5.2, we may apply inverse Laplace transform and we obtain that for $|\arg(z)| \leq \frac{\pi}{2} - \epsilon$, as $|z| \to \infty$,

$$\begin{aligned}\tilde{f}_1(z) &\sim \log_{1/q} z + \frac{\gamma}{L} + \frac{1}{2} - \alpha + \frac{1}{L}\sum_{k\neq 0}\frac{M_1(1+\chi_k)}{\Gamma(1+\chi_k)}z^{\chi_k}\\ &= \log_{1/q} z + \frac{\gamma}{L} + \frac{1}{2} - \alpha - \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)z^{\chi_k}.\end{aligned}$$

The same asymptotic expansion also holds for $\mathbb{E}(X_n)$ by depoissonization.

Analysis of the Variance. For an asymptotic expansion of the variance, we start from (4.4) and proceed by a similar method as above. We already know that $\tilde{f}_1(z)$ is JS-admissible and $\tilde{V}(z)$ satisfies the functional equation (4.4). Because $\tilde{f}_1(z)$ is entire and JS-admissible, $\tilde{f}_1'(z)^2$ admits a polynomial bound by Ritt's Theorem (Theorem 4.2 of [163]). Therefore, we have the following rough bounds for $\tilde{V}(z)$:

$$\tilde{V}(z) = \begin{cases} O(z), & |z| \to 0;\\ O(z^{\epsilon}), & |z| \to \infty.\end{cases}$$

Consequently, we may apply Laplace transform to (4.4) and it yields

$$(s+1)\mathscr{L}[\tilde{V}; s] = \frac{1}{q}\mathscr{L}[\tilde{V}; s/q] + \tilde{g}(s), \qquad (4.6)$$

where $\tilde{g}(s) = \mathscr{L}[\tilde{f}_1'^2; s]$. Again, from the polynomial bound for $\tilde{f_1}'(z)^2$, we get bounds for $\tilde{g}(z)$:

$$\tilde{g}(s) = \begin{cases} \mathcal{O}(s), & s \to 0;\\ \mathcal{O}(s^{-1}), & s \to \infty.\end{cases}$$

Next, set $\bar{\mathscr{L}}[\tilde{V}; s] = \mathscr{L}[\tilde{V}; s]/Q(-s)$. Dividing both sides of (4.6) by $Q(-s/q)$ yields

$$\bar{\mathscr{L}}[\tilde{V}; s] = \frac{1}{q}\bar{\mathscr{L}}[\tilde{V}; s/q] + \tilde{g}(s)/Q(-s/q).$$

Again, by the same reason as for the mean, we may apply Mellin transform and obtain

$$\mathscr{M}[\bar{\mathscr{L}};\omega] = \frac{M_2(\omega)}{1-q^{\omega-1}}, \qquad (\Re(\omega) > 1),$$

where

$$M_2(\omega) = \int_0^\infty \frac{s^{\omega-1}}{Q(-s/q)} \int_0^\infty e^{-zs} \tilde{f}_1'(z)^2 \mathrm{d}z\mathrm{d}s.$$

By the bounds for $\tilde{g}(s)$ above, we have that $M_2(\omega)$ is analytic in the half plane $\Re(\omega) \in (-1,\infty)$. Because $\bar{\mathscr{L}}[\tilde{V};s]$ admits a polynomial bound, the inverse Mellin transform of $\mathscr{M}[\bar{\mathscr{L}};\omega]$ exists by Proposition 5 of [62]. The rest of the analysis is as for the mean and we obtain, as $z \to \infty$,

$$\tilde{V}(z) \sim \frac{1}{L} \sum_{k\in\mathbb{Z}} \frac{M_2(1+\chi_k)}{\Gamma(1+\chi_k)} z^{\chi_k}.$$

By depoissonization, the same holds for $\mathrm{Var}(X_n)$ as well.

We conclude by simplifying $M_2(1+\chi_k)$. Therefore, we use that

$$\tilde{f}_1'(z) = Q_\infty \sum_{l\geq 0} \frac{q^l}{Q_l} e^{-zq^l}$$

and

$$\frac{1}{Q(-s/q)} = \frac{1}{Q_\infty} \sum_{j\geq 0} \frac{(-1)^j q^{\binom{j}{2}}}{Q_j(s+q^{-j})}.$$

Plugging this into the above integral yields

$$M_2(1+\chi_k) = Q_\infty \sum_{h,l,j\geq 0} \frac{(-1)^j q^{\binom{j}{2}+l+h}}{Q_h Q_l Q_j} \int_0^\infty \frac{s^{\chi_k}}{(s+q^{-j})(s+q^h+q^l)} \mathrm{d}s.$$

Denote by

$$\varphi(\chi;x) := \begin{cases} \pi(x^\chi - 1)/(\sin(\pi\chi)(x-1)), & \text{if } x \neq 1, \\ \pi\chi/\sin(\pi\chi), & \text{if } x = 1. \end{cases}$$

Then,

$$M_2(1+\chi_k) = Q_\infty \sum_{h,l,j\geq 0} \frac{(-1)^j q^{\binom{j}{2}+(\chi_k-1)j+l+h}}{Q_h Q_l Q_j} \varphi(\chi_k, q^{h+j} + q^{l+j}).$$

4.1.3 Average Value of $G_C(z)$

We will use the abbreviations $Q = 1/q$ and $L = \log Q$. Furthermore, in order to be closer to the q-hypergeometric world and the identities of relevance (see the book of Andrews-Askey-Roy [6]), we use the classical notation $(q)_n$ instead of Q_n.

In [115], the alternative expression

$$\mathcal{P} := \frac{\log 2}{L} - \alpha - \beta + \frac{2}{L}\tau \quad \text{with} \quad \tau := \sum_{k\ge 1} \frac{(-1)^{k-1}}{k(Q^k-1)}$$

was given for the constant in the variance, and we will show now the equality of this and

$$\mathcal{F} := \frac{(q)_\infty}{L} \sum_{j,l,h\ge 0} \frac{(-1)^j q^{\binom{j+1}{2}+l+h}}{(q)_j(q)_l(q)_h} \frac{\log(q^{h+j}+q^{l+j})}{q^{h+j}+q^{l+j}-1}.$$

In this expression, we have replaced the ψ function by what it is; in some exceptional cases a limit has to be taken.

We use the symmetry in l and h and set $l = h + d$ with $d \ge 0$; then we have to take the sum over $h, d \ge 0$ twice, and subtract the sum for $h \ge 0$ and $d = 0$. Therefore

$$\mathcal{F} = 2\sum_{j,h,d\ge 0} \cdots - \sum_{j,h\ge 0,\ d=0} \cdots.$$

We think about d as being fixed, set $h = N - j$ and fix N as well: This leads to

$$\frac{(q)_\infty[-LN + \log(1+q^d)]}{L} \sum_{j=0}^{N} \frac{(-1)^j q^{\binom{j+1}{2}+2(N-j)+d}}{(q)_j(q)_{N-j+d}(q)_{N-j}} \frac{1}{q^N+q^{N+d}-1}.$$

By automatic summation (q-Zeilberger's algorithm) we have the simplification

$$\sum_{j=0}^{N} \frac{(-1)^j q^{\binom{j+1}{2}+2(N-j)+d}}{(q)_j(q)_{N-j}(q)_{N+d-j}} \frac{1}{q^N+q^{N+d}-1} = \frac{q^{N^2+dN}}{(q)_N(q)_{N+d}}.$$

Consequently,

$$\begin{aligned}\mathcal{F} =& 2(q)_\infty \sum_{N,d\ge 0} \frac{-NL+\log(1+q^d)}{L} \frac{q^{N^2+dN}}{(q)_N(q)_{N+d}} \\ &+ (q)_\infty \sum_{N\ge 0} \frac{NL-\log 2}{L} \frac{q^{N^2}}{(q)_N(q)_N}.\end{aligned}$$

We will soon show that

$$(q)_\infty \sum_{N,d\geq 0} \frac{\log(1+q^d)}{L} \frac{q^{N^2+dN}}{(q)_N(q)_{N+d}} = \frac{\tau}{L} + \frac{\log 2}{L}, \tag{4.7}$$

which leaves us to prove that

$$2(q)_\infty \sum_{N,d\geq 0} \frac{Nq^{N^2+dN}}{(q)_N(q)_{N+d}} - (q)_\infty \sum_{N\geq 0} \frac{(N-\frac{\log 2}{L})q^{N^2}}{(q)_N(q)_N} = \frac{\log 2}{L} + \alpha + \beta.$$

Because of the identity [6]

$$\sum_{N\geq 0} \frac{q^{N^2}}{(q)_N^2} = \frac{1}{(q)_\infty},$$

this leaves us with

$$2(q)_\infty \sum_{N,d\geq 0} \frac{Nq^{N^2+dN}}{(q)_N(q)_{N+d}} - (q)_\infty \sum_{N\geq 0} \frac{Nq^{N^2}}{(q)_N(q)_N} = \alpha + \beta. \tag{4.8}$$

Now, expanding $\log(1+q^d)$, (4.7) is proved once we can prove that

$$(q)_\infty \sum_{N\geq 0,\ d\geq 1} \frac{q^{N^2+dN+dk}}{(q)_N(q)_{N+d}} = \frac{1}{Q^k-1}.$$

But this follows from

$$\sum_{N\geq 0,\ d\geq 1} \frac{1}{(q)_d} \frac{q^{N^2+dN+dk}}{(q)_N(q^{d+1})_N} = \sum_{d\geq 1} \frac{q^{dk}}{(q)_d} \frac{1}{(q^{d+1})_\infty} = \frac{1}{(q)_\infty} \frac{1}{Q^k-1}.$$

We have used here the classical identity (Cauchy's identity) [6]

$$\sum_{n\geq 0} \frac{x^n q^{n^2}}{(q)_n(xq)_n} = \frac{1}{(xq)_\infty}.$$

In order to prove (4.8), we will show that

$$-(q)_\infty \sum_{N\geq 0} \frac{Nq^{N^2}}{(q)_N(q)_N} = \sum_{r\geq 1} \frac{(-1)^r q^{\binom{r+1}{2}}}{1-q^r}, \tag{4.9}$$

$$(q)_\infty \sum_{N,d\geq 0} \frac{Nq^{N^2+dN}}{(q)_N(q)_{N+d}} = -\sum_{r\geq 1} \frac{(-1)^r q^{\binom{r+1}{2}}}{(1-q^r)^2}. \tag{4.10}$$

Since in [121, (3.16)], it was proved that

$$\sum_{r\geq 1}\frac{(-1)^r q^{\binom{r+1}{2}}}{1-q^r}-2\sum_{r\geq 1}\frac{(-1)^r q^{\binom{r+1}{2}}}{(1-q^r)^2}=\alpha+\beta,$$

that would finish the proof. We start from

$$\sum_{n\geq 0}\frac{x^n q^{n^2}}{(q)_n(xq)_n}=\sum_{n\geq 0}\frac{x^n q^{n^2}}{(q)_n(xq)_\infty}(xq^{n+1})_\infty=\frac{1}{(xq)_\infty},$$

which is equivalent to

$$\sum_{n\geq 0}\frac{x^n q^{n^2}}{(q)_n}\sum_{k\geq 0}\frac{(-1)^k q^{\binom{k}{2}}x^k q^{(n+1)k}}{(q)_k}=1.$$

Now differentiate this, and then set $x=1$:

$$\sum_{n\geq 0}\frac{nq^{n^2}}{(q)_n^2}+\frac{1}{(q)_\infty}\sum_{n\geq 0}\frac{q^{n^2}}{(q)_n}\sum_{k\geq 0}\frac{(-1)^k q^{\binom{k}{2}}kq^{(n+1)k}}{(q)_k}=0.$$

Rearranging,

$$\sum_{n\geq 0}\frac{nq^{n^2}}{(q)_n^2}+\frac{1}{(q)_\infty}\sum_{N\geq 1}\sum_{n=0}^{N}\frac{q^{n^2}}{(q)_n}\frac{(-1)^{N-n}q^{\binom{N-n}{2}}(N-n)q^{(n+1)(N-n)}}{(q)_{N-n}}=0,$$

and again by a mechanical proof,

$$\sum_{n\geq 0}\frac{nq^{n^2}}{(q)_n^2}+\frac{1}{(q)_\infty}\sum_{N\geq 1}\frac{(-1)^N q^{\binom{N+1}{2}}}{1-q^N}=0.$$

This is (4.9). Now let us plug in $x=q^d$ after differentation (instead of $x=1$, as before):

$$\sum_{n\geq 0}\frac{nq^{d(n-1)}q^{n^2}}{(q)_n}\sum_{k\geq 0}\frac{(-1)^k q^{\binom{k}{2}}q^{kd}q^{(n+1)k}}{(q)_k}$$
$$+\sum_{n\geq 0}\frac{q^{dn}q^{n^2}}{(q)_n}\sum_{k\geq 0}\frac{(-1)^k q^{\binom{k}{2}}kq^{(k-1)d}q^{(n+1)k}}{(q)_k}=0.$$

After some simplifications (using Rothe's identity [6]), this leads to

$$(q)_\infty\sum_{n\geq 0}\frac{nq^{dn}q^{n^2}}{(q)_n(q)_{n+d}}+\sum_{N\geq 1}\frac{(-1)^N q^{\binom{N+1}{2}+dN}}{1-q^N}=0.$$

Now sum this on d:

$$(q)_\infty \sum_{n,d\geq 0} \frac{nq^{dn}q^{n^2}}{(q)_n(q)_{n+d}} + \sum_{N\geq 1} \frac{(-1)^N q^{\binom{N+1}{2}}}{(1-q^N)^2} = 0,$$

which is (4.10).

4.1.4 Approximate Counting with m Counters and m-DSTs

Approximate Counting with m Counters. I. Here, we consider approximate counting with m counters as discussed in the introduction. Recall that D_n denotes the sum of the counters after counting n objects. Then, we have

$$D_n \overset{d}{=} C_{I_1}^{(1)} + \cdots + C_{I_m}^{(m)},$$

where $C_n^{(1)}, \dots, C_n^{(m)}$ are independent copies of C_n and

$$P(I_1 = n_1, \dots, I_m = n_m) = \binom{n}{n_1, \dots, n_m} \frac{1}{m^n}$$

with $n_1 + \cdots + n_m = n$. Now, set

$$\tilde{Q}(y,z) = e^{-z} \sum_{n\geq 0} \mathbb{E}(e^{D_n y}) \frac{z^n}{n!}, \quad \tilde{P}(y,z) = e^{-z} \sum_{n\geq 0} \mathbb{E}(e^{C_n y}) \frac{z^n}{n!}.$$

Then, by a straightforward computation

$$\tilde{Q}(y,z) = \tilde{P}(y, z/m)^m.$$

From this, we can derive the following relations for the Poisson generating functions of the first and second moment of D_n and C_n (denoted by $\tilde{g}_1(z), \tilde{g}_2(z)$ for the former and as above for the latter)

$$\begin{aligned} \tilde{g}_1(z) &= m\tilde{f}_1(z/m), \\ \tilde{g}_2(z) &= m(m-1)\tilde{f}_1(z/m)^2 + m\tilde{f}_2(z/m). \end{aligned}$$

Moreover, again consider the poisson variance $\tilde{W}(z) := \tilde{g}_2(z) - \tilde{g}_1(z)^2$. Then,

$$\tilde{W}(z) = m\tilde{V}(z/m).$$

Now, it follows from the closure properties of JS-admissibility that both $\tilde{g}_1(z)$ and $\tilde{g}_2(z)$ are JS-admissible. Hence, we only have to concentrate on $\tilde{g}_1(z)$ and $\tilde{W}(z)$ whose asymptotic expansions, due to the above formulas, follow from the case $m = 1$.

***m*-DSTs.** m-DSTs have been introduced in [178]. They are defined as follows: again we start with n keys, but they are now stored in m DSTs. For every key, one of the m DSTs is chosen uniformly and at random and the key is then stored in the chosen tree.

Clearly, the previous analysis also gives the sum of the lengths of the leftmost path in m-DSTs. Similarly, one can consider other shape parameters in DSTs and extend them linearly to m-DSTs. Our method above can then be applied to such parameters as well and again the analysis will be reduced to the case $m = 1$.

We give two examples. The first example is the depth of a random node which was discussed in [178]. As a second example, consider the total path length T_n in a random digital search tree of size n which is the sum over all distances of nodes to the root. The mean of this parameter has already been computed in Example 3.4.6. The result is

$$\mathbb{E}(T_n) \sim n\log_2 n + n\left(\frac{\gamma-1}{\log 2} + \frac{1}{2} - \alpha\right) + \frac{n}{\log 2}F_T(\log_2 n),$$

where $F_T(z)$ is a 1-periodic function with its Fourier coefficients given in Example 3.4.6. For the variance, it was proved that for $q = 1/2$ (see Kirschenhofer, Prodinger and Szpankowski [121] and [74]), as $n \to \infty$,

$$\mathrm{Var}(T_n) \sim nG_T(\log_2 n),$$

where $G_T(z)$ are 1-periodic functions with computable Fourier coefficients. Similar results are known for the case $q \neq 1/2$ as well. Now, denote by U_n the sum of all total path lengths in an m-DST. Then, with the same approach as above, we have the following result.

Theorem 4.1.2. *For the total path length in m-DSTs, we have, as $n \to \infty$,*

$$\begin{aligned}\mathbb{E}(U_n) &\sim (n/m)\log_2(n/m) + (n/m)F_T(\log_2(n/m)),\\ \mathrm{Var}(U_n) &\sim (n/m)G_T(\log_2(n/m)).\end{aligned}$$

Approximate Counting with m Counters. II. Here, we again consider approximate counting with m counters, but this time we label them from 1 to m. First, we use the first counter until it will be increased, then we use the second one until it will be increased, etc. until the last counter is increased then we return to the first one and repeat this procedure.

Let again D_n denote the sum of the m counters after counting n objects. This clearly corresponds to the length of the leftmost path in random digital search trees, where every node can hold up to m keys, namely, the bucket digital search trees discussed in Section 2.2.3 (here, the length is the sum of all nodes on the leftmost path weighted by the number of keys the nodes contain). Consequently, $D_n \overset{d}{=} X_{n+1}$, where X_n satisfies

$$X_{n+m} \overset{d}{=} X_{B_n} + m, \qquad (n \geq 0)$$

with $X_i = i, 0 \leq i \leq m-1$. The Poisson-Laplace-Mellin approach can be applied to this sequence as well. We only sketch some details.

First, for the poisson generating functions of the first and second moment and the poissonized variance (again denoted by $\tilde{f}_1(z)$ and $\tilde{V}(z)$) respectively, we have

$$\sum_{i=0}^{m} \binom{m}{i} \tilde{f}_1^{(i)}(z) = \tilde{f}_1(qz) + m$$

and

$$\sum_{i=0}^{m} \binom{m}{i} \tilde{V}^{(i)}(z) = \tilde{V}(qz) + \tilde{g}(z).$$

Where $\tilde{g}(z)$ is of the form

$$\begin{aligned}\tilde{g}(z) &= 2m \sum_{i=0}^{m} \binom{m}{i} \tilde{f_1}^{(i)}(z) + \left(\sum_{i=0}^{m} \binom{m}{i} \tilde{f_1}^{(i)}(z) - m \right)^2 \\ &\quad + \frac{z}{q} \left(\sum_{i=0}^{m} \binom{m}{i} \tilde{f_1}^{(i+1)}(z) - m \right)^2 + m^2 - \sum_{i=0}^{m} \binom{m}{i} \left(\tilde{f}_1(z)^2 \right)^{(i)}\end{aligned}$$

Applying the Poisson-Laplace-Mellin method then yields asymptotic expansion of mean and variance. We content ourself with stating the result for the variance. As usual, we check JS-admissibility of $\tilde{f}_1(z)$ first. By similar methods as in the proof of Proposition 3.2 in [74], we can easily verify the JS-admissibility of $\tilde{f}_1(z)$ and hence the existence of the Laplace transform and Mellin transform below are ensured by the same argument as in Section 4.1.2. We apply the Laplace transform and divide it by $Q(-s/q) = \prod_{l \geq 1} (1 + q^l s)^m$:

$$\bar{\mathscr{L}}[\tilde{V}; s] = \frac{1}{q} \bar{\mathscr{L}}[\tilde{V}; s/q] + \frac{\mathscr{L}[\tilde{g}; s] - p(s)}{Q(-s/q)}$$

where

$$p(s) = \sum_{n=0}^{m} \sum_{k=0}^{n-1} \binom{m}{n} s^{n-1-k} \left(2^k \binom{k}{2} - k^2 \right)$$

Then the Mellin transform gives us

$$\mathscr{M}[\bar{\mathscr{L}}; \omega] = \frac{M(\omega)}{1 - q^{\omega - 1}}, \qquad (\Re(\omega) > 1),$$

where

$$M(\omega)=\int_0^\infty s^{\omega-1}\int_0^\infty \frac{e^{-sz}}{Q(s/q)}\left(\tilde{g}(z)dz-p(s)\right)ds=\mathscr{M}[\frac{\mathscr{L}[\tilde{g};s]-p(s)}{Q(-s/q)};\omega].$$

Finally, by reversing the above process, we obtain

$$\tilde{V}(z)\sim\frac{1}{L}\sum_k\frac{M(\omega_k)}{\Gamma(\omega_k)}z^{\omega_k-1}.$$

where $\omega_k=1+\frac{2k\pi i}{\log(\frac{1}{q})}$, $k\in\mathbb{Z}$.

Theorem 4.1.3. *For approximate counting with m-counters, where counters are chosen cyclically, we have, as $n\to\infty$,*

$$\mathrm{Var}(D_n)\sim G_D(\log_{1/q}n),$$

where $G_D(z)=\sum_k g_k e^{2k\pi i}$ is a 1-periodic function with Fourier coefficients

$$g_k=\frac{M(\omega_k)}{L\Gamma(\omega_k)}.$$

4.2 Wiener Index

4.2.1 Introduction

Topological indices of molecular graphs are of great importance in combinatorial chemistry and many papers have been dedicated to them. One of the most well-known index is the so-called *Wiener index* which is defined as the sum of distances of all unordered pairs of nodes of a graph. This index was proposed by Wiener in [215] in order to investigate the boiling point of alkanes. It has been intensively studied, in particular for trees since trees arise as molecular graphs of acyclic organic molecules; see the survey paper of Dobrynin, Entringer and Gutman [43] for many results and references.

Here, we are interested in the Wiener index of random trees. The first class of random trees for which the Wiener index was studied were simple generated random trees. In [53], Entringer, Meir, Moon and Székely showed that the mean of the Wiener index in a simple generated random tree of size n is of order $n^{5/2}$. The mean for families of random trees more relevant in chemistry has been investigated by Dobrynin and Gutman in [44] and Wagner in [207], [208].

As for deeper stochastic properties, Neininger in [159] was the first who considered variance and limit laws. More precisely, he showed for random binary search trees and random recursive trees that the mean of the Wiener index is of order $n^2 \log n$ and the variance is of order n^4. Moreover, he also proved a bivariate limit law of the Wiener index and the total path length. Janson in [101] then carried out a similar study for simple generated random trees whose Wiener index has variance of order n^5 and again satisfies a bivariate limit law with the total path length (however, the limiting distribution is quite different from the one found by Neininger for random binary search trees and random recursive trees). The same results were very recently also proved to hold for non-plane unlabeled trees by Wagner [209] (he considered both the rooted and unrooted case).

Finally, also very recently, Munsonius in [153] extended the above results of Neininger to the class of random split trees which was introduced by Devroye in [37]. The class of split trees is a very large class of random trees containing many important types of random trees as special cases, e.g., binary search trees, m-ary search trees, median-of-$(2k+1)$ search trees, quadtrees, simplex trees, digital trees, etc. Munsonius proved in [153] that for a huge subclass of the class of random split trees, the variance of the Wiener index has order n^4 and a bivariate limit law with the total path length holds. The subclass he considered includes most of the classes of random trees mentioned above but not the important class of random digital trees. We will derive the stochastic properties of Wiener index for random digital trees in this section. The result will answer two questions of Neininger from [159] in affirmative who asked whether or not periodic oscillations are present in the moments of the Wiener index for digital trees and whether or not the Wiener index is asymptotically normal distributed.

Before discussing our results in more details, we want to mention that apart from limit laws, results about tail probabilities of the distribution of the Wiener index have been proved as well; see Janson and Chassaing [104], Ali Khan and Neininger [110], Fill and Janson [58] and Munsonius [154].

Now, fix a symmetric random digital search tree of size n and denote by T_n its total path length and by W_n its Wiener index. Then, we have the following result for first and second moments.

Theorem 4.2.1. *We have for the mean of the total path length and the Wiener index of digital search trees,*

$$\mathbb{E}(T_n) = n\log_2 n + nP_1(\log_2 n) + \mathcal{O}(\log n),$$
$$\mathbb{E}(W_n) = n^2\log_2 n + n^2 P_1(\log_2 n) - n^2 + \mathcal{O}(n\log n),$$

where $P_1(z)$ is a one-periodic function given in Remark 4 below. Moreover, variances and covariances of the total path length and the Wiener index of *digital search trees are given by*

$$\begin{aligned}\mathrm{Var}(T_n) &= nP_2(\log_2 n) + \mathcal{O}(1),\\ \mathrm{Cov}(T_n, W_n) &= n^2 P_2(\log_2 n) + \mathcal{O}(n \log n),\\ \mathrm{Var}(W_n) &= n^3 P_2(\log_2 n) + \mathcal{O}(n^2 \log n),\end{aligned}$$

where $P_2(z)$ is again a one-periodic function given in Remark 5 below.

Remark 4. The result for the mean of the total path length has already been introduced in Example 3.4.6. The periodic function is given by

$$P_1(z) = \frac{\gamma - 1}{\log 2} + \frac{1}{2} - \sum_{k\ge 1} \frac{1}{2^k - 1} + \frac{1}{\log 2} \sum_{k \neq 0} \Gamma(-1 - \chi_k) e^{2k\pi i z},$$

where γ is Euler's constant and $\chi_k = 2k\pi i / \log 2$.

Note that the result for the mean of the Wiener index is also not new since it can be derived from the result in [2].

Finally, we want to remark that with our method of proof it is straightforward to compute longer asymptotic expansions.

Remark 5. Similar to the mean, the result about the variance of the total path length is also not new; see Kirschenhofer, Prodinger and Szpankowski [121]. In [74] the following explicit expression was given for the periodic function

$$P_2(z) = \frac{1}{\log 2} \sum_k \frac{G_2(2 + \chi_k)}{\Gamma(2 + \chi_k)} e^{2k\pi i z},$$

where

$$G_2(2 + \chi_k) = Q_\infty \sum_{j,h,l \ge 0} \frac{(-1)^j 2^{-\binom{j+1}{2}}}{Q_j Q_h Q_l 2^{h+l}} \varphi(2 + \chi_k; 2^{-j-h} + 2^{-j-l}).$$

Here, $Q_j = \prod_{1 \le l \le j}(1 - 2^{-l}), Q_\infty = \lim_{j\to\infty} Q_j$ and

$$\varphi(\omega; x) = \begin{cases} \dfrac{\pi(1 + x^{\omega-2}((\omega - 2)x + 1 - \omega))}{(x-1)^2 \sin(\pi\omega)}, & \text{if } x \neq 1;\\ \dfrac{\pi(\omega - 1)(\omega - 2)}{2\sin(\pi\omega)}, & \text{if } x = 1.\end{cases}$$

Moreover, it was proved in [121] that $P_2(\log_2 n) > 0$ for all n; see also Schachinger [189] for a more elementary proof of this fact.

As for the covariance between total path length and Wiener index and the variance of the Wiener index, these results are new. In particular, note that the variance is of order n^3 which is different from the order obtained for other random split trees; see [153]. This is actually not surprising since it is well-known that random digital search trees are "less random" than other random split trees.

Again it is straightforward to obtain more terms in the asymptotic expansion.

As a corollary of Theorem 4.2.1, we obtain the following result.

Corollary 4.2.2. *For the correlation coefficient of the total path length and the Wiener index of digital search trees, denoted by $\rho(T_n, W_n)$, we obtain that*

$$\lim_{n\to\infty} \rho(T_n, W_n) = 1.$$

This will allow us to prove the following result.

Theorem 4.2.3. *We have,*

$$\left(\frac{T_n - \mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}, \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}}\right) \xrightarrow{d} (X, X),$$

where X is a standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

Remark 6. Again the central limit theorem for the total path length is not new; see Jacquet and Szpankowski [98] and the discussion in Section 5 in [74]. In fact, our result will follow from Jacquet and Szpankowski's result and Corollary 4.2.2.

Next, we give a brief description of the method we will use in order to prove our results. First, note that from the definition of the total path length and the Wiener index, we immediately get the following distributional recurrences: for $n \geq 0$, we have

$$T_{n+1} \overset{d}{=} T_{B_n} + T^*_{n-B_n} + n, \tag{4.11}$$

$$\begin{aligned} W_{n+1} \overset{d}{=} & W_{B_n} + W^*_{n-B_n} + (B_n + 1)(T^*_{n-B_n} + n - B_n) \\ & + (n - B_n + 1)(T_{B_n} + B_n), \end{aligned} \tag{4.12}$$

where $B_n = \mathrm{Binomial}(n, 1/2)$, (T^*_n, W^*_n) denotes an independent copy of (T_n, W_n), and (T_n, W_n) and (B_n) are independent. Also, note that initial conditions are given by $T_0 = W_0 = 0$.

Remark 7. It is interesting to point out that Schachinger in [192] studied a general distributional recurrence which is very similar to the two recurrences above. More precisely, he investigated the distributional recurrence

$$X_n \overset{d}{=} X_{B_n} + X^*_{n-B_n} + T_n,$$

where notation is as above and T_n is a general random variable called *toll function*. For the case $T_n = n^\alpha, \alpha > 0$, he proved that the limit law is normal if and only if $\alpha \leq 3/2$. In view of this result, it might come as a surprise that the Wiener index is asymptotically normal distributed since the toll sequence in (4.12) should be roughly of order n^2. However, note that in Schachinger's result T_n is deterministic and hence independent of X_n whereas in our situation we have strong dependence.

4.2.2 Wiener Index for Digital Search Trees

In order to obtain the moments, we will use the Poisson-Laplace-Mellin method. Here, we will prove Theorem 4.2.1 and Theorem 4.2.3. Note that the total path length is already analyzed in [74]. In fact, we will heavily use results from this analysis in our derivation below (for the relevant results see Section 2.5 and Section 2.6 in [74]).

Now, we will start with our analysis. Therefore, set

$$\tilde{f}_{1,0}(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(T_n)\frac{z^n}{n!} \qquad \text{and} \qquad \tilde{f}_{0,1}(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(W_n)\frac{z^n}{n!}.$$

Then, from (4.11), (4.12) and a straightforward computation, one obtains

$$\begin{aligned}\tilde{f}_{1,0}(z) + \tilde{f}'_{1,0}(z) &= 2\tilde{f}_{1,0}(z/2) + z,\\ \tilde{f}_{0,1}(z) + \tilde{f}'_{0,1}(z) &= 2\tilde{f}_{0,1}(z/2) + (z+2)\tilde{f}_{1,0}(z/2) + \frac{z^2}{2} + z \end{aligned} \qquad (4.13)$$

with $\tilde{f}_{1,0}(0) = \tilde{f}_{0,1}(0) = 0$. Similarly, set

$$\tilde{f}_{2,0}(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(T_n^2)\frac{z^n}{n!}, \quad \tilde{f}_{1,1}(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(T_nW_n)\frac{z^n}{n!},$$

and

$$\tilde{f}_{0,2}(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(W_n^2)\frac{z^n}{n!}.$$

Then, again from (4.11), (4.12) with a slightly more involved computation,

$$\begin{aligned}\tilde{f}_{2,0}(z)+\tilde{f}'_{2,0}(z)=&2\tilde{f}_{2,0}(z/2)+2\tilde{f}^2_{1,0}(z/2)+4z\tilde{f}_{1,0}(z/2)+2z\tilde{f}'_{1,0}(z/2)\\&+z^2+z\\\tilde{f}_{1,1}(z)+\tilde{f}'_{1,1}(z)=&2\tilde{f}_{1,1}(z/2)+2\tilde{f}_{1,0}(z/2)\tilde{f}_{0,1}(z/2)+z\tilde{f}_{1,0}(z/2)\tilde{f}'_{1,0}(z/2)\\&+(z+2)\tilde{f}_{2,0}(z/2)+(z+2)\tilde{f}^2_{1,0}(z/2)\\&+(2z^2+5z)\tilde{f}_{1,0}(z/2)+\frac{3z^2+4z}{2}\tilde{f}'_{1,0}(z/2)+2z\tilde{f}_{0,1}(z/2)\\&+z\tilde{f}'_{0,1}(z/2)+\frac{z^3+4z^2+2z}{2}\end{aligned}$$

$$\begin{aligned}\tilde{f}_{0,2}(z)+\tilde{f}'_{0,2}(z)=&2\tilde{f}_{0,2}\left(z/2\right)+\left(\frac{z^3}{2}+3z+2\right)\tilde{f}_{2,0}\left(z/2\right)\\&+(2z+4)\tilde{f}_{1,1}\left(z/2\right)+(2z+4)\tilde{f}_{1,0}\left(z/2\right)\\&+\tilde{f}_{0,1}\left(z/2\right)+2z\tilde{f}_{1,0}\left(z/2\right)\tilde{f}'_{0,1}\left(z/2\right)+2\tilde{f}_{0,1}\left(z/2\right)^2\\&+(2z^2+4z)\tilde{f}_{0,1}\left(z/2\right)+(2z^2+2z)\tilde{f}'_{0,1}\left(z/2\right)\\&+\left(\frac{z^2}{2}+2z+2\right)\tilde{f}_{1,0}\left(z/2\right)^2+(z^2+2z)\tilde{f}_{1,0}\left(z/2\right)\tilde{f}'_{1,0}\left(z/2\right)\\&+\frac{z^2}{2}\tilde{f}'_{1,0}\left(z/2\right)^2+(z^3+6z^2+6z)\tilde{f}_{1,0}\left(z/2\right)\\&+(z^3+5z^2+2z)\tilde{f}'_{1,0}\left(z/2\right)+\frac{z^4}{4}+2z^3+4z^2+z,\end{aligned}$$

where $\tilde{f}_{2,0}(0)=\tilde{f}_{1,1}(0)=\tilde{f}_{0,2}(0)=0$.

Next, we define poissonized variances and covariances by using the poissonized variance with corrections which was introduced in Section 3.5.1.

$$\tilde{V}(z)+\tilde{V}'(z)=2\tilde{V}(z/2)+z\tilde{f}''_{1,0}(z)^2,$$

$$\tilde{C}(z)+\tilde{C}'(z)=2\tilde{C}(z/2)+(z+2)\tilde{V}(z/2)+z\tilde{f}''_{1,0}(z)\tilde{f}''_{0,1}(z), \qquad (4.14)$$

$$\begin{aligned}\tilde{W}(z)+\tilde{W}'(z)=&2\tilde{W}(z/2)+(2z+4)\tilde{C}(z/2)+\left(\frac{z^2}{2}+3z+2\right)\tilde{V}(z/2)\\&+z^2\tilde{f}'_{1,0}(z/2)^2+2z^2\tilde{f}'_{1,0}(z/2)+z\tilde{f}''_{0,1}(z)^2+z^2\end{aligned} \qquad (4.15)$$

with $\tilde{V}(0)=\tilde{C}(0)=\tilde{W}(0)=0$.

We will now apply the "Poisson-Laplace-Mellin" method to these differential-functional equations. We will start with the mean value.

Mean Value of Wiener Index. We will start from (4.13). We first apply Laplace transform which yields

$$(1+s)\mathscr{L}[\tilde{f}_{0,1}(z);s]=4\mathscr{L}[\tilde{f}_{0,1}(z);2s]-2\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{L}[\tilde{f}_{1,0}(z);2s]+4\mathscr{L}[\tilde{f}_{1,0}(z);2s]+\frac{1+s}{s^3}.$$

Next, dividing by $Q(-2s)$ and setting

$$\bar{\mathscr{L}}[\tilde{f}_{0,1}(z);s]=\frac{\mathscr{L}[\tilde{f}_{0,1}(z);s]}{Q(-s)},\qquad \bar{\mathscr{L}}[\tilde{f}_{1,0}(z);s]=\frac{\mathscr{L}[\tilde{f}_{1,0}(z);s]}{Q(-s)}$$

gives

$$\begin{aligned}\bar{\mathscr{L}}[\tilde{f}_{0,1}(z);s]=&4\bar{\mathscr{L}}[\tilde{f}_{0,1}(z);2s]-\frac{2}{Q(-2s)}\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{L}[\tilde{f}_{1,0}(z);2s]+4\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);2s]\\&+\frac{1+s}{s^3Q(-2s)}.\end{aligned}\tag{4.16}$$

Observe that

$$\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);2s]=\mathscr{L}[\tilde{f}_{1,0}(z);2s]\frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{Q(-2s)}+\frac{1}{Q(-2s)}\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{L}[\tilde{f}_{1,0}(z);2s].\tag{4.17}$$

Moreover, logarithmic differentiation yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}Q(-2s)&=\frac{\mathrm{d}}{\mathrm{d}s}\exp\{\log(Q(-2s))\}=Q(-2s)\frac{\mathrm{d}}{\mathrm{d}s}\sum_{j\ge0}\log\left(1+\frac{s}{2^j}\right)\\&=Q(-2s)\sum_{j\ge0}\frac{1}{2^j+s}.\end{aligned}$$

Set $A(s)=\sum_{j\ge0}\frac{1}{2^j+s}$ whose Maclaurin series is given by

$$A(s)=\sum_{j\ge0}\sum_{k\ge0}\frac{(-s)^k}{2^{(k+1)j}}=\sum_{k\ge0}\frac{2^{k+1}}{2^{k+1}-1}(-s)^k.$$

Next,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{Q(-2s)}&=-\frac{1}{Q(-2s)^2}\frac{\mathrm{d}}{\mathrm{d}s}Q(-2s)=-\frac{A(s)}{Q(-2s)}\\&=-\frac{2}{Q(-2s)}-\frac{\bar{A}(s)}{Q(-2s)},\end{aligned}\tag{4.18}$$

where $\bar{A}(s) = \sum_{k\geq 1} 2^{k+1}(-s)^k/(2^{k+1}-1)$. Plugging (4.18) into (4.17) and (4.17) in turn into (4.16) gives

$$\begin{aligned}\bar{\mathscr{L}}[\tilde{f}_{0,1}(z);s] =& 4\bar{\mathscr{L}}[\tilde{f}_{0,1}(z);2s] - 2\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);2s] - 2\bar{A}(s)\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);2s]\\ &+ \frac{1+s}{s^3Q(-2s)}.\end{aligned} \tag{4.19}$$

The next step is to apply Mellin transform. Therefore, note that from [74], we know that

$$\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);s] = \begin{cases}\mathcal{O}\left(|s|^{-2}|\log s|\right), & \text{as } s \to 0;\\ \mathcal{O}\left(|s|^{-b}\right), & \text{as } s \to \infty\end{cases}$$

uniformly for s with $|\arg(s)| \leq \pi - \epsilon$, where $b > 0$ is an arbitrary large constant. Moreover, again from [74], for $Q(-2s)$ (and consequently also for $\bar{A}(s)$), we have the bounds

$$Q(-2s) = \begin{cases}1+\mathcal{O}(|s|), & \text{as } s \to 0;\\ \mathcal{O}(|s|^{-b}), & \text{as } s \to \infty\end{cases}, \qquad \bar{A}(s) = \begin{cases}\mathcal{O}(|s|), & \text{as } s \to 0;\\ \mathcal{O}(|s|^{-b}), & \text{as } s \to \infty\end{cases} \tag{4.20}$$

again uniformly for s with $|\arg(s)| \leq \pi - \epsilon$, where $b > 0$ is an arbitrary large constant. As a consequence of this and Ritt's theorem (see Chapter 1, Section 4.3 in Olver [163]), the Mellin transform of

$$\tilde{s}_{0,1}(s) = -2\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);2s] + \frac{1+s}{s^3Q(-2s)},$$

which we denote by $S_{0,1}(\omega)$, exists for $\Re(\omega) > 3$ and the Mellin transform of

$$\tilde{t}_{0,1}(s) = -2\bar{A}(s)\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);2s],$$

which we denote by $T_{0,1}(\omega)$, exists for $\Re(\omega) > 1$. Moreover, by Proposition 5 in [62], we have, as $|t| \to \infty$,

$$S_{0,1}(c+it) = \mathcal{O}\left(e^{-(\pi-\epsilon)|t|}\right), \qquad T_{0,1}(c+it) = \mathcal{O}\left(e^{-(\pi-\epsilon)|t|}\right) \tag{4.21}$$

for all $c \in \mathbb{R}$ contained in the fundamental strip. In fact, using the expression for the Mellin transform for $\mathscr{L}[\tilde{f}_{1,0}(z);s]$ from [74], we obtain for $S_{0,1}(\omega)$ the expression

$$\begin{aligned}S_{0,1}(\omega) =& \frac{Q(2^{\omega-3})\Gamma(\omega)\Gamma(2-\omega)}{2Q_\infty(2^{\omega-3}-1)} + \frac{Q(2^{\omega-3})\Gamma(\omega-1)\Gamma(2-\omega)}{Q_\infty}\\ &+ \frac{Q(2^{\omega-2})\Gamma(\omega)\Gamma(1-\omega)}{Q_\infty}.\end{aligned}$$

Note that from this, it follows that (4.21) holds for all $c \in \mathbb{R}$. Finally, by applying Mellin transform to (4.19), we have

$$\mathscr{M}[\bar{\mathscr{L}}[\tilde{f}_{1,0}];\omega] = \frac{S_{0,1}(\omega)+T_{0,1}(\omega)}{1-2^{2-\omega}}.$$

From this and the above explicit expression for $S_{0,1}(\omega)$, we obtain by inverse Mellin transform

$$\begin{aligned}\bar{\mathscr{L}}[\tilde{f}_{1,0}(z);s] = {}& 2s^{-3}\log_2\frac{1}{s} + \left(\frac{1}{\log 2} - 1 - 2c\right)s^{-3}\\ &+ \frac{1}{\log 2}\sum_{k\neq 0}\Gamma(3+\chi_k)\Gamma(-1-\chi_k)s^{-3-\chi_k} + \mathcal{O}\left(|s|^{-2}|\log s|\right)\end{aligned}$$

where $c = \sum_{k\geq 1} 1/(2^k-1)$, χ_k was defined in Remark 4 and the above asymptotic expansion holds uniformly as $s \to 0$ with $|\arg(s)| \leq \pi - \epsilon$. Moreover, due to (4.20), the same asymptotic expansion holds for $\mathscr{L}[\tilde{f}_{1,0}(z);s]$ as well.

Next, we apply inverse Laplace transform and obtain

$$\tilde{f}_{0,1}(z) = z^2\log_2 z + z^2 P_1(\log_2 z) - z^2 + \mathcal{O}(|z\log z|) \tag{4.22}$$

uniformly as $z \to \infty$ with $|\arg(z)| \leq \pi/2 - \epsilon$, where $P_1(z)$ was introduced in Remark 4.

The final step is depoissonization which is done by the closure properties of JS-admisibility. Hence,

$$\mathbb{E}(W_n) = \tilde{f}_{0,1}(n) - \frac{n}{2}\tilde{f}''_{0,1}(n) + \text{lower order terms.}$$

Note that from (4.22) and Ritt's theorem, we obtain that the second term on the right-hand side above is of order $\mathcal{O}(n\log n)$. Consequently, the above gives the claimed expansion for the mean.

Covariance of Total Path Length and Wiener Index. Here, we start from (4.14) and use the same method as for the mean. First, from [74], we have that

$$\tilde{f}_{1,0}(z) = z\log_2 z + zP_1(\log_2 z) + \mathcal{O}(|\log z|) \tag{4.23}$$

uniformly as $z \to \infty$ with $|\arg(z)| \leq \pi/2 - \epsilon$. From this, (4.22) and Ritt's theorem, we obtain the bounds

$$z\tilde{f}''_{1,0}(z)\tilde{f}''_{0,1}(z) = \begin{cases}\mathcal{O}(|z|), & \text{as } z \to 0;\\ \mathcal{O}(|\log z|), & \text{as } z \to \infty\end{cases} \tag{4.24}$$

uniformly for z with $|\arg(z)| \leq \pi/2 - \epsilon$.

Next, we apply Laplace transform to (4.14) and divide it by $Q(-2s)$. Then, by similar manipulations as for the mean, we obtain

$$\bar{\mathscr{L}}[\tilde{C}(z); s] = 4\bar{\mathscr{L}}[\tilde{C}(z); 2s] - 2\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{V}(z); 2s] - 2\bar{A}(s)\bar{\mathscr{L}}[\tilde{V}(z); 2s] + \bar{g}_{1,1}(s), \tag{4.25}$$

where

$$\bar{g}_{1,1}(s) = \frac{\mathscr{L}[z\tilde{f}''_{1,0}(z)\tilde{f}''_{0,1}(z); s]}{Q(-2s)}.$$

Before applying Mellin transform, we note that from [74], we have

$$\bar{\mathscr{L}}[\tilde{V}(z); s] = \begin{cases} \mathcal{O}\left(|s|^{-2}\right), & \text{as } s \to 0; \\ \mathcal{O}\left(|s|^{-b}\right), & \text{as } s \to \infty \end{cases}$$

uniformly for s with $|\arg(s)| \leq \pi - \epsilon$, where $b > 0$ is an arbitrary large constant. Moreover, from (4.24) and (4.20), we obtain

$$\bar{g}_{1,1}(s) = \begin{cases} \mathcal{O}\left(|s|^{-1}|\log s|\right), & \text{as } s \to 0; \\ \mathcal{O}\left(|s|^{-b}\right), & \text{as } s \to \infty \end{cases}$$

again uniformly for s with $|\arg(s)| \leq \pi - \epsilon$, where $b > 0$ is an arbitrary large constant. Hence, the Mellin transform of

$$\tilde{s}_{1,1}(s) = -2\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{V}(z); 2s],$$

which we denote by $S_{1,1}(\omega)$, exists for $\Re(\omega) > 3$ and the Mellin transform of

$$\tilde{t}_{1,1}(s) = -2\bar{A}(s)\bar{\mathscr{L}}[\tilde{V}(z); 2s] + \bar{g}_{1,1}(s),$$

which we denote by $T_{1,1}(\omega)$, exists for $\Re(\omega) > 1$. Also, both Mellin transforms satisfy a bound of the form (4.21) inside their fundamental strips. Moreover, in [74], we showed that

$$\mathscr{M}[\bar{\mathscr{L}}[\tilde{V}]; \omega] = \frac{G_2(\omega)}{1 - 2^{2-\omega}},$$

where $G_2(\omega)$ is analytic for $\Re(\omega) > 0$ and satisfies a bound of the form (4.21) in this half-plane. Consequently, by applying Mellin transform to (4.25),

$$\mathscr{M}[\bar{\mathscr{L}}[\tilde{C}]; \omega] = \frac{S_{1,1}(\omega) + T_{1,1}(\omega)}{1 - 2^{2-\omega}} = \frac{2^{2-\omega}(\omega - 1)G_2(\omega - 1)}{(1 - 2^{3-\omega})(1 - 2^{2-\omega})} + \frac{T_{1,1}(\omega)}{1 - 2^{2-\omega}}.$$

From this by inverse Mellin transform

$$\bar{\mathscr{L}}[\tilde{C}(z);s] = \frac{1}{\log 2}\sum_k (2+\chi_k)G_2(2+\chi_k)s^{-3-\chi_k} + \mathcal{O}\left(|s|^{-2}\right)$$

uniformly as $s \to 0$ with $|\arg(s)| \leq \pi - \epsilon$. (For $G_2(\omega)$, the expressions given in Remark 5 was proved in [74]) From (4.20), we get the same asymptotic for $\mathscr{L}[\tilde{C}(z);s]$.

Inverse Laplace transform yields

$$\tilde{C}(z) = z^2 P_2(\log_2 z) + \mathcal{O}(|z|) \tag{4.26}$$

uniformly as $z \to \infty$ and $|\arg(z)| \leq \pi/2 - \epsilon$, where $P_2(z)$ is given in Remark 5.

The final step is depoissonization. Therefore, observe that $\tilde{f}_{1,0}(z)$, $\tilde{f}_{0,1}(z)$ and $\tilde{f}_{1,1}(z)$ are all JS-admissible. Hence,

$$\mathrm{Cov}(T_n, W_n) = \tilde{C}(n) - \frac{n}{2}\tilde{C}''(n) - \frac{n^2}{2}\tilde{f}''_{1,0}(n)\tilde{f}''_{0,1}(n) + \text{lower order terms.}$$

Note that due to Ritt's theorem, the second term on the right hand side is $\mathcal{O}(n)$ and the third term is $\mathcal{O}(n\log n)$. Hence, our claimed result for the covariance is proved.

Variance of Wiener Index. Next, we turn to the variance of the Wiener index. We start from (4.15) which we rewrite as

$$\tilde{W}(z) + \tilde{W}'(z) = 2\tilde{W}(z/2) + 2z\tilde{C}(z/2) + \frac{z^2}{2}\tilde{V}(z/2) + \tilde{g}_{0,2}(z)$$

with

$$\tilde{g}_{0,2}(z) = 4\tilde{C}(z/2)+(3z+2)\tilde{V}(z/2)+z^2\tilde{f}'_{1,0}(z/2)^2+2z^2\tilde{f}'_{1,0}(z/2)+z\tilde{f}''_{0,1}(z)^2+z^2.$$

From [74], we have that

$$\tilde{V}(z) = zP_2(\log_2 z) + \mathcal{O}(1)$$

uniformly as $z \to \infty$ with $|\arg(z)| \leq \pi/2 - \epsilon$. From this, (4.26), (4.23), (4.22) and Ritt's theorem it follows that

$$\tilde{g}_{0,2}(z) = \begin{cases} \mathcal{O}(|z|), & \text{as } z \to 0; \\ \mathcal{O}(|z|^2|\log z|^2), & \text{as } z \to \infty \end{cases} \tag{4.27}$$

uniformly for z with $|\arg(z)| \leq \pi/2 - \epsilon$.

Next, applying Laplace transform to the above differential-functional equation and dividing by $Q(-2s)$ yields

$$\begin{aligned}\bar{\mathscr{L}}[\tilde{W}(z);s] =& 4\bar{\mathscr{L}}[\tilde{W}(z);2s] - \frac{4}{Q(-2s)}\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{L}[\tilde{C}(z);2s] \\ &+ \frac{1}{Q(-2s)}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathscr{L}[\tilde{V}(z);2s] + \bar{g}_{0,2}(s),\end{aligned} \tag{4.28}$$

where

$$\bar{g}_{0,2}(s) = \frac{\mathscr{L}[\tilde{g}_{0,2}(z);s]}{Q(-2s)}.$$

Using the same manipulations as for mean and covariance

$$-\frac{4}{Q(-2s)}\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{L}[\tilde{C}(z);2s] = -4\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{C}(z);2s] - 4A(s)\bar{\mathscr{L}}[\tilde{C}(z);2s]. \tag{4.29}$$

Moreover, observe that

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\bar{\mathscr{L}}[\tilde{V}(z);2s] =& \frac{1}{Q(-2s)}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathscr{L}[\tilde{V}(z);2s] - 2\frac{A(s)}{Q(-2s)}\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{L}[\tilde{V}(z);2s] \\ &+ \mathscr{L}[\tilde{V}(z);2s]\frac{\mathrm{d}^2}{\mathrm{d}s^2}\frac{1}{Q(-2s)}\end{aligned}$$

and note that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\frac{1}{Q(-2s)} = -\frac{\mathrm{d}}{\mathrm{d}s}\frac{A(s)}{Q(-2s)} = \frac{A(s)^2}{Q(-2s)} - \frac{B(s)}{Q(-2s)},$$

where

$$B(s) = -\sum_{k\geq 0}\frac{k2^{k+2}}{2^{k+2}-1}(-s)^k.$$

This implies that

$$\begin{aligned}\frac{1}{Q(-2s)}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathscr{L}[\tilde{V}(z);2s] =& \frac{\mathrm{d}^2}{\mathrm{d}s^2}\bar{\mathscr{L}}[\tilde{V}(z);2s] + 2A(s)\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{V}(z);2s] \\ &+ (A(s)^2 + B(s))\bar{\mathscr{L}}[\tilde{V}(z);2s]\end{aligned} \tag{4.30}$$

and plugging (4.30) and (4.29) into (4.28) yields

$$\bar{\mathscr{L}}[\tilde{W}(z);s] = 4\bar{\mathscr{L}}[\tilde{W}(z);2s] - 4\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{C}(z);2s] + \frac{\mathrm{d}^2}{\mathrm{d}s^2}\bar{\mathscr{L}}[\tilde{V}(z);2s] + \tilde{t}_{0,2}(s)$$

with

$$\begin{aligned}\tilde{t}_{0,2}(s) =& -4A(s)\bar{\mathscr{L}}[\tilde{C}(z);2s] + 2A(s)\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{V}(z);2s]\\ &+ (A(s)^2+B(s))\bar{\mathscr{L}}[\tilde{V}(z);2s] + \bar{g}_{0,2}(s).\end{aligned}$$

Before we apply Mellin transform, note that from (4.27) and (4.20),

$$\bar{g}_{0,2}(s) = \begin{cases}\mathcal{O}(|s|^{-3}|\log s|^2), & \text{as } s\to 0;\\ \mathcal{O}(|s|^{-b}), & \text{as } s\to\infty\end{cases}$$

uniformly for s with $|\arg(s)| \le \pi-\epsilon$, where $b>0$ is an arbitrary large constant. Moreover,

$$B(s) = \begin{cases}\mathcal{O}(1), & \text{as } s\to 0;\\ \mathcal{O}(|s|^{-b}), & \text{as } s\to\infty\end{cases}$$

again uniformly for s with $|\arg(s)| \le \pi-\epsilon$, where $b>0$ is an arbitrary large constant. From this and corresponding bounds for $A(s), \bar{\mathscr{L}}[\tilde{C}(z);s]$ and $\bar{\mathscr{L}}[\tilde{V}(z);s]$ obtained in the analysis of the mean and covariance, we see that the Mellin transform of $\tilde{t}_{0,2}(s)$, which we denote by $T_{0,2}(\omega)$, exists for $\Re(\omega)>3$. Similarly, the Mellin transform of

$$\tilde{s}_{0,2}(s) = -4\frac{\mathrm{d}}{\mathrm{d}s}\bar{\mathscr{L}}[\tilde{C}(z);2s] + \frac{\mathrm{d}^2}{\mathrm{d}s^2}\bar{\mathscr{L}}[\tilde{V}(z);2s],$$

which we denote by $S_{0,2}(\omega)$, exists for $\Re(\omega)>4$. Both of these Mellin transforms satisfy a bound of the form (4.21) inside their fundamental strip. Moreover, observe that using the expressions from the analysis of the covariance, $S_{0,2}(\omega)$ is given by

$$S_{0,2}(\omega) = \frac{2^{2-\omega}(2^{3-\omega}+1)(\omega-1)(\omega-2)G_2(\omega-2)}{(1-2^{3-\omega})(1-2^{4-\omega})} + \frac{2^{3-\omega}(\omega-1)T_{1,1}(\omega-1)}{1-2^{3-\omega}},$$

where $G_2(\omega)$ is an analytic function for $\Re(\omega)>0$, $T_{1,1}(\omega)$ is an analytic function for $\Re(\omega)>1$ and both satisfy a bound of the form (4.21) in their half-plane of analyticity. Overall, we obtain for the Mellin transform of $\bar{\mathscr{L}}[\tilde{W}(z);s]$

$$\begin{aligned}\mathscr{M}[\bar{\mathscr{L}}[\tilde{W}];\omega] =& \frac{S_{0,2}(\omega)+T_{0,2}(\omega)}{1-2^{2-\omega}}\\ =& \frac{2^{2-\omega}(2^{3-\omega}+1)(\omega-1)(\omega-2)G_2(\omega-2)}{(1-2^{2-\omega})(1-2^{3-\omega})(1-2^{4-\omega})}\\ &+ \frac{2^{3-\omega}(\omega-1)T_{1,1}(\omega-1)}{1-2^{3-\omega}} + \frac{T_{0,2}(\omega)}{1-2^{2-\omega}}.\end{aligned}$$

From this, by applying inverse Mellin transform

$$\bar{\mathscr{L}}[\tilde{W}(z);s]=\frac{1}{\log 2}\sum_{k}(3+\chi_k)(2+\chi_k)G_2(2+\chi_k)s^{-4-\chi_k}+\mathcal{O}(|s|^{-3-\epsilon})$$

uniformly as $s\to 0$ with $|\arg(s)|\le\pi-\epsilon$. Moreover, due to (4.20), the same is also true for $\mathscr{L}[\tilde{W}(z);s]$.

Again, we apply inverse Laplace transform and obtain

$$\tilde{W}(z)=z^3P_2(\log_2 z)+\mathcal{O}(|z|^{2+\epsilon})$$

uniformly as $z\to\infty$ with $|\arg(z)|\le\pi/2-\epsilon$.

The final step is the depoissonization step where as above we use the closure properties of JS-admissiblity. By these results, $\tilde{f}_{0,2}(z)$ and $\tilde{f}_{0,1}(z)$ are both JS-admissible. Consequently,

$$\mathrm{Var}(W_n)=\tilde{W}(n)-\frac{n}{2}\tilde{W}''(n)-\frac{n^2}{2}\tilde{f}''_{0,1}(n)^2+\text{smaller order terms.}$$

By Ritt's theorem, the second term on the right-hand side is $\mathcal{O}(n^2)$ and the third term is $\mathcal{O}(n^2\log^2 n)$. From this our result follows (the claimed error term in Theorem 4.2.1 is obtained by a slightly refined analysis which we leave as an exercise to the reader).

This concludes our proof of Theorem 4.2.1 and consequently also Corollary 4.2.2. We will use now the latter to give a proof of Theorem 4.2.3. As a second ingredient, we need the following central limit theorem for the total path length.

Theorem 4.2.4 (Jacquet and Szpankowski; [98]). *We have,*

$$\frac{T_n-\mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}\xrightarrow{d}X,$$

where X has a standard normal distribution.

Proof of Theorem 4.2.3. First set

$$X_n=\frac{T_n-\mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}.$$

Then, by the above result

$$X_n\xrightarrow{d}X,$$

where X has a standard normal distribution. Consequently,

$$(X_n,X_n)\xrightarrow{d}(X,X).$$

Next, define

$$Y_n = \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}} - \frac{T_n - \mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}.$$

Note that

$$\begin{aligned}\mathbb{E}(Y_n^2) &= \frac{\mathbb{E}(W_n - \mathbb{E}(W_n))^2}{\mathrm{Var}(W_n)} + \frac{\mathbb{E}(T_n - \mathbb{E}(T_n))^2}{\mathrm{Var}(T_n)} - 2\frac{\mathbb{E}(W_n - \mathbb{E}(W_n))(T_n - \mathbb{E}(T_n))}{\sqrt{\mathrm{Var}(W_n)\mathrm{Var}(T_n)}}\\ &= 2 - 2\rho(T_n, W_n).\end{aligned}$$

Hence, by Markov's inequality

$$P(|Y_n| \geq \epsilon) \leq \frac{\mathbb{E}(Y_n^2)}{\epsilon} \longrightarrow 0, \qquad \text{as } n \to \infty.$$

Thus, $Y_n \xrightarrow{P} 0$ and consequently $(0, Y_n) \xrightarrow{P} (0,0)$ (here, $\xrightarrow{P}$ denotes convergence in probability). Using Slutsky's theorem (also called Cramér's theorem; see Theorem 11.4 in Gut [87]) now implies

$$(X_n, X_n) + (0, Y_n) \xrightarrow{d} (X, X).$$

Since

$$(X_n, X_n) + (0, Y_n) = \left(\frac{T_n - \mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}, \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}} \right)$$

this proves our claim. □

4.2.3 Wiener Index for Variants of Digital Search Trees

In this section, we are going to discuss similar results as in Section 4.2.2 for other digital trees. Proofs of these results follow along the same lines (or are even easier since in some cases Laplace transform is not needed) and will not be given). For the reader's convenience, we will list the (differential-)functional equations for poissonized mean, variances and covariances which are crucial to the proofs in the Appendix A. Our results can be deduced from them with a similar approach as used in Section 4.2.2.

The first member of the digital tree family we are going to discuss is the bucket digital search trees. Note that there are two types of total path length in bucket digital trees: the sum of distances of all keys to the root and the sum of distances of all nodes to the root; the former is called *key-wise path length* and the latter *node-wise path length* (see [74] for more details). Accordingly, we also have a *key-wise Wiener index* and a *node-wise Wiener* *index*. Results for both Wiener indices in random bucket digital search trees will be presented below.

Another member of the digital tree family are tries. Note that for tries, the number of leaves is n whereas the number of internal nodes is random. Hence, there are again two different types of Wiener indices, namely, the *external Wiener index* which only uses external nodes and the *internal Wiener index* where internal nodes are used. Again both of these Wiener indices will be discussed below.

As a final member of the digital tree family, we consider PATRICIA tries. For binary PATRICIA tries, the number of internal nodes is not random and hence there is only external Wiener index which make sense. However, for m-ary PATRICIA tries with $m > 2$, the number of internal nodes is no longer definite and hence the internal Wiener index is well-defined. We will give the results of internal Wiener index for m-ary PATRICIA tries in the end of this section.

As in Section 4.2.2, we will denote by T_n the total path length (either key-wise or node-wise or external or internal depending on the context) and by W_n the Wiener index (again either key-wise or node-wise or external or internal). Moreover, for the node-wise Wiener index and the internal Wiener index, we also need the number of nodes (internal in case of the internal Wiener index) which will be denoted by N_n.

Key-wise Wiener Index of Bucket Digital Search Trees. Here, we have the following distributional recurrences for T_n and W_n: for $n \geq 0$,

$$\begin{aligned} T_{n+b} &\overset{d}{=} T_{B_n} + T^*_{n-B_n} + n, \\ W_{n+b} &\overset{d}{=} W_{B_n} + W^*_{n-B_n} + (B_n + 1)(T^*_{n-B_n} + n - B_n) \\ &\quad + (n - B_n + 1)(T_{B_n} + B_n), \end{aligned}$$

where notation is as in Section 1 and initial conditions are given by $T_0 = \cdots = T_{b-1} = W_0 = \cdots = W_{b-1} = 0$.

From these recurrences, we obtain the following results for mean and variance.

Theorem 4.2.5. *We have for the mean of the key-wise path length and key-wise Wiener index of bucket digital search trees,*

$$\begin{aligned} \mathbb{E}(T_n) &= n \log_2 n + nP_1(\log_2 n) + \mathcal{O}(\log n), \\ \mathbb{E}(W_n) &= n^2 \log_2 n + n^2 P_1(\log_2 n) - n^2 + \mathcal{O}(n \log n), \end{aligned}$$

where $P_1(z)$ is a one-periodic function given in the remark below. Moreover, variances and covariances of the key-wise path length and key-wise Wiener index of bucket digital search trees are given by

$$\begin{aligned}
\mathrm{Var}(T_n) &= nP_2(\log_2 n) + \mathcal{O}(1),\\
\mathrm{Cov}(T_n, W_n) &= n^2 P_2(\log_2 n) + \mathcal{O}(n \log n),\\
\mathrm{Var}(W_n) &= n^3 P_2(\log_2 n) + \mathcal{O}(n^2 \log n),
\end{aligned}$$

where $P_2(z)$ is again a one-periodic function given in the remark below.

Remark 8. The result for the mean and variance of the key-wise path length were first obtained by Hubalek in [89]. In [74], we gave the following expressions for the periodic functions

$$P_1(z) = \frac{\gamma - 1}{\log 2} + \frac{1}{2} + \frac{c}{\log 2} + \frac{1}{\log 2} \sum_{k \neq 0} \frac{G_1(2+\chi_k)}{\Gamma(2+\chi_k)} e^{2k\pi i z},$$

where

$$G_1(\omega) = \int_0^\infty \frac{s^{\omega-3}}{Q(-2s)^b} \mathrm{d}s, \qquad c = \lim_{\omega \to 2} (G_1(\omega) - 1/(\omega - 2))$$

and

$$P_2(z) = \frac{1}{\log 2} \sum_{k} \frac{G_2(2+\chi_k)}{\Gamma(2+\chi_k)} e^{2k\pi i z},$$

where

$$G_2(\omega) = \int_0^\infty \frac{s^{\omega-1}}{Q(-2s)^b} \int_0^\infty e^{-zs} \tilde{g}(z) \mathrm{d}z \mathrm{d}s$$

with

$$\begin{aligned}
\tilde{g}(z) = &\left(\sum_{0 \le j \le b} \binom{b}{j} \tilde{f}_{1,0}^{(j)}(z) \right)^2 + z \left(\sum_{0 \le j \le b} \binom{b}{j} \tilde{f}_{1,0}^{(j+1)}(z) \right)^2\\
&- \sum_{0 \le j \le b} \binom{b}{j} \left(\tilde{f}_{1,0}^2(z) + z \tilde{f}'_{1,0}(z)^2 \right)^{(j)}
\end{aligned}$$

and $\tilde{f}_{1,0}(z)$ denotes the Poisson generating function of $\mathbb{E}(T_n)$.

Note that the result for the mean of the Wiener index also follows from [22].

Moreover, we have the following bivariate central limit theorem.

Theorem 4.2.6. *We have,*

$$\left(\frac{T_n-\mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}},\frac{W_n-\mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}}\right)\xrightarrow{d}(X,X),$$

where X is a standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

Remark 9. The central limit theorem for the key-wise path length was first proved in [90].

Node-wise Wiener Index of Bucket Digital Search Trees. Here, the distributional recurrences for N_n, T_n and W_n are given by: for $n \geq 0$,

$$\begin{aligned}
N_{n+b} &\overset{d}{=} N_{B_n}+N^*_{n-B_n}+1,\\
T_{n+b} &\overset{d}{=} T_{B_n}+T^*_{n-B_n}+N_{B_n}+N^*_{n-B_n},\\
W_{n+b} &\overset{d}{=} W_{B_n}+W^*_{n-B_n}+(N_{B_n}+1)(T^*_{n-B_n}+N^*_{n-B_n})\\
&\quad+(N^*_{n-B_n}+1)(T_{B_n}+N_{B_n}),
\end{aligned}$$

where B_n is as in Section 1, the triplet (N^*_n, T^*_n, W^*_n) denotes an independent copy of (N_n, T_n, W_n) and (N_n, T_n, W_n) is independent of (B_n). Initial conditions are given by $T_0 = \cdots = T_{b-1} = W_0 = \cdots = W_{b-1} = N_0 = 0$ and $N_1 = \cdots = N_{b-1} = 1$.

From this, we obtain the following result.

Theorem 4.2.7. *We have for the mean of the number of nodes, node-wise path length and node-wise Wiener index of bucket digital search trees,*

$$\begin{aligned}
\mathbb{E}(N_n) &= nP_1(\log_2 n)+\mathcal{O}(1),\\
\mathbb{E}(T_n) &= n(\log_2 n)P_1(\log_2 n)+\mathcal{O}(n),\\
\mathbb{E}(W_n) &= n^2(\log_2 n)P_1(\log_2 n)^2+\mathcal{O}(n^2),
\end{aligned}$$

where $P_1(z)$ is a one-periodic function given in the remark below. Moreover, variances and covariances of the number of nodes, node-wise path length and node-wise Wiener index of bucket digital search trees are given by

$$\begin{aligned}
\mathrm{Var}(N_n) &= nP_2(\log_2 n)+\mathcal{O}(1),\\
\mathrm{Cov}(N_n,T_n) &= n(\log_2 n)P_2(\log_2 n)+\mathcal{O}(n),\\
\mathrm{Var}(T_n) &= n(\log_2 n)^2P_2(\log_2 n)+\mathcal{O}(n\log n),\\
\mathrm{Cov}(N_n,W_n) &= 2n^2(\log_2 n)P_1(\log_2 n)P_2(\log_2 n)+\mathcal{O}(n^2),\\
\mathrm{Cov}(T_n,W_n) &= 2n^2(\log_2 n)^2P_1(\log_2 n)P_2(\log_2 n)+\mathcal{O}(n^2\log n),\\
\mathrm{Var}(W_n) &= 4n^3(\log_2 n)^2P_1(\log_2 n)^2P_2(\log_2 n)+\mathcal{O}(n^3\log n),
\end{aligned}$$

where $P_2(z)$ is again a one-periodic function given in the remark below.

Remark 10. The results for the number of nodes were first proved in [90]. Moreover, the results were reproved in [74] where in addition we also proved the results for the node-wise path length and gave the following expressions for $P_1(z)$ and $P_2(z)$

$$P_1(z) = \frac{1}{\log 2} \sum_k \frac{G_1(2+\chi_k)}{\Gamma(2+\chi_k)} e^{2k\pi i z},$$

where

$$G_1(\omega) = \int_0^\infty \frac{s^{\omega-2}}{Q(-2s)^b}(s+1)^{b-1}\mathrm{d}s$$

and

$$P_2(z) = \frac{1}{\log 2} \sum_k \frac{G_2(2+\chi_k)}{\Gamma(2+\chi_k)} e^{2k\pi i z},$$

where

$$G_2(\omega) = \int_0^\infty \frac{s^{\omega-1}}{Q(-2s)^b}\left(\int_0^\infty e^{-zs}\tilde{g}(z)\mathrm{d}z + H(s)\right)\mathrm{d}s$$

with

$$H(s) = \frac{(s+1)^{b-1} - (-1)^b(2b-3+(b-1)s)}{(s+2)^2}$$

and

$$\tilde{g}(z) = \left(\sum_{0\le j\le b}\binom{b}{j}\tilde{f}_{1,0}^{(j)}(z)\right)^2 + z\left(\sum_{0\le j\le b}\binom{b}{j}\tilde{f}_{1,0}^{(j+1)}(z)\right)^2 - \sum_{0\le j\le b}\binom{b}{j}\left(\tilde{f}_{1,0}^2(z) + z\tilde{f}'_{1,0}(z)^2\right)^{(j)}$$

and $\tilde{f}_{1,0}(z)$ denotes the Poisson generating function of $\mathbb{E}(T_n)$.

Theorem 4.2.7 yields the following trivariate central limit theorem.

Theorem 4.2.8. *We have,*

$$\left(\frac{N_n - \mathbb{E}(N_n)}{\sqrt{\mathrm{Var}(N_n)}}, \frac{T_n - \mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}, \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}}\right) \xrightarrow{d} (X, X, X),$$

where X is a standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

Remark 11. The central limit theorem for the number of nodes was first proved in [90]. Also note that we posed the problem of proving a bivariate central limit law of number of nodes and node-wise path length in Section 5 of [74].

External Wiener Index of Tries. Here, the distributional recurrences for T_n and W_n are as follows: for $n \geq 2$,

$$\begin{aligned} T_n &\stackrel{d}{=} T_{B_n} + T^*_{n-B_n} + n, \\ W_n &\stackrel{d}{=} W_{B_n} + W^*_{n-B_n} + B_n(T^*_{n-B_n} + n - B_n) + (n - B_n)(T_{B_n} + B_n), \end{aligned}$$

where notation is as in Section 1 and initial conditions are given by $T_0 = T_1 = W_0 = W_1 = 0$.

From this, we obtain the following theorem.

Theorem 4.2.9. *We have for the mean of external path length and external Wiener index of tries,*

$$\begin{aligned} \mathbb{E}(T_n) &= n\log_2 n + nP_1(\log_2 n) + \mathcal{O}(\log n), \\ \mathbb{E}(W_n) &= n^2\log_2 n + n^2P_1(\log_2 n) - n^2 + \mathcal{O}(n\log n), \end{aligned}$$

where $P_1(z)$ is a one-periodic function given in the remark below. Moreover, variances and covariances of the external path length and external Wiener index of tries are given by

$$\begin{aligned} \mathrm{Var}(T_n) &= nP_2(\log_2 n) + \mathcal{O}(1), \\ \mathrm{Cov}(T_n, W_n) &= n^2P_2(\log_2 n) + \mathcal{O}(n\log n), \\ \mathrm{Var}(W_n) &= n^3P_2(\log_2 n) + \mathcal{O}(n^2\log n), \end{aligned}$$

where $P_2(z)$ is again a one-periodic function given in the remark below.

Remark 12. The result about the mean of the total path length was first obtained in [127]. A detailed analysis of the variance of the total path length was first undertaken by Kirschenhofer, Prodinger and Szpankowski [119] (see also Jacquet and Régnier [94] for preliminary results). In Hwang, Fuchs and Zacharovas [75], we obtained the following expressions for the periodic functions

$$P_1(z) = \frac{\gamma}{\log 2} + \frac{1}{2} - \frac{1}{\log 2}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi iz}$$

and

$$P_2(z) = \frac{1}{\log 2}\sum_{k} G_2(-1-\chi_k)e^{2k\pi iz},$$

where

$$\begin{aligned} G_2(\omega) =& \Gamma(\omega+1)\left(1 - \frac{\omega^2+\omega+4}{2^{\omega+3}}\right) \\ &+ 2\sum_{l\geq 1}\frac{(-1)^l\Gamma(\omega+l+1)}{l!(2^l-1)}(l(\omega+l)-1). \end{aligned}$$

Note that the result about the mean of the Wiener index also follows from [22].

From the previous result, we again obtain the following theorem.

Theorem 4.2.10. *We have,*

$$\left(\frac{T_n - \mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}, \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}}\right) \xrightarrow{d} (X, X),$$

where X is a standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

Remark 13. The central limit theorem for the key-wise path length was first proved in [94].

Internal Wiener Index of Tries. Here, the distributional recurrences for N_n, T_n and W_n are as follows: for $n \geq 2$,

$$\begin{aligned} N_n &\overset{d}{=} N_{B_n} + N^*_{n-B_n} + 1, \\ T_n &\overset{d}{=} T_{B_n} + T^*_{n-B_n} + N_{B_n} + N^*_{n-B_n}, \\ W_n &\overset{d}{=} W_{B_n} + W^*_{n-B_n} + (N_{B_n} + 1)(T^*_{n-B_n} + N^*_{n-B_n}) \\ &\quad + (N^*_{n-B_n} + 1)(T_{B_n} + N_{B_n}), \end{aligned}$$

where notation is as for the node-wise Wiener index and initial conditions are given by $N_0 = N_1 = T_0 = T_1 = W_0 = W_1 = 0$.

Then, we have the following result for mean values, variances and covariances.

Theorem 4.2.11. *We have for the mean of the number of internal nodes, internal path length and internal Wiener index of tries,*

$$\begin{aligned} \mathbb{E}(N_n) &= nP_1(\log_2 n) + \mathcal{O}(1), \\ \mathbb{E}(T_n) &= n(\log_2 n)P_1(\log_2 n) + \mathcal{O}(n), \\ \mathbb{E}(W_n) &= n^2(\log_2 n)P_1(\log_2 n)^2 + \mathcal{O}(n^2), \end{aligned}$$

where $P_1(z)$ is a one-periodic function given in the remark below. Moreover, variances and covariances of the number of internal nodes, internal path *length and internal Wiener index of tries are given by*

$$\begin{aligned}
\mathrm{Var}(N_n) &= nP_2(\log_2 n) + \mathcal{O}(1),\\
\mathrm{Cov}(N_n, T_n) &= n(\log_2 n)P_2(\log_2 n) + \mathcal{O}(n),\\
\mathrm{Var}(T_n) &= n(\log_2 n)^2 P_2(\log_2 n) + \mathcal{O}(n \log n),\\
\mathrm{Cov}(N_n, W_n) &= 2n^2(\log_2 n)P_1(\log_2 n)P_2(\log_2 n) + \mathcal{O}(n^2),\\
\mathrm{Cov}(T_n, W_n) &= 2n^2(\log_2 n)^2 P_1(\log_2 n)P_2(\log_2 n) + \mathcal{O}(n^2 \log n),\\
\mathrm{Var}(W_n) &= 4n^3(\log_2 n)^2 P_1(\log_2 n)^2 P_2(\log_2 n) + \mathcal{O}(n^3 \log n),
\end{aligned}$$

where $P_2(z)$ is again a one-periodic function given in the remark below.

Remark 14. The result for the mean of the number of internal nodes was first proved in [127]. The variance of the number of internal nodes was first derived by Régnier and Jacquet [95] (see also [94], [93]). In [75], we gave the following expression for the periodic functions

$$P_1(z) = \frac{1}{\log 2} + \frac{1}{\log 2}\sum_{k\neq 0} \chi_k \Gamma(-1-\chi_k)e^{2k\pi i z}.$$

and

$$P_2(z) = \frac{1}{\log 2}\sum_{k} G_2(-1-\chi_k)e^{2k\pi i z},$$

where

$$\begin{aligned}
G_2(\omega) =&(\omega+1)\Gamma(\omega)\left(1 - \frac{\omega^2+4\omega+8}{2^{\omega+3}}\right)\\
&+ 2\sum_{l\geq 1} \frac{(-1)^l l\Gamma(\omega+l+1)}{(l+1)!(2^l-1)}(l(\omega+l+1)-1).
\end{aligned}$$

The results for mean and variance of internal path length and covariance with the number of internal nodes are due to Nguyen-The [162].

As before, we have a central limit theorem which now reads as follows.

Theorem 4.2.12. *We have,*

$$\left(\frac{N_n - \mathbb{E}(N_n)}{\sqrt{\mathrm{Var}(N_n)}}, \frac{T_n - \mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}, \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}}\right) \xrightarrow{d} (X, X, X),$$

where X is a standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

Remark 15. The central limit theorem for the number of internal nodes was first proved in [93] and [94]. The bivariate central limit theorem for the number of internal nodes and the internal path length was wrongly stated in [162] (the author of this work did not observe that the covariance matrix is singular leading to a wrong proof).

External Wiener Index of Binary PATRICIA tries. Here, we have for T_n and W_n: for $n \geq 2$,

$$T_n \stackrel{d}{=} \begin{cases} T_{B_n} + T^*_{n-B_n} + n, & \text{if } B_n \neq 0 \text{ or } B_n \neq n; \\ T_n, & \text{otherwise,} \end{cases}$$

$$W_n \stackrel{d}{=} \begin{cases} W_{B_n} + W^*_{n-B_n} + B_n(T^*_{n-B_n} + n - B_n) & \\ \quad +(n - B_n)(T_{B_n} + B_n), & \text{if } B_n \neq 0 \text{ or } B_n \neq n; \\ W_n, & \text{otherwise,} \end{cases}$$

where notations is as in Section 1 and $T_0 = T_1 = W_0 = W_1 = 0$.

Then, we have the following result.

Theorem 4.2.13. *We have for the mean of the total path length and Wiener index of PATRICIA tries,*

$$\begin{aligned} \mathbb{E}(T_n) &= n \log_2 n + nP_1(\log_2 n) + \mathcal{O}(\log n), \\ \mathbb{E}(W_n) &= n^2 \log_2 n + n^2 P_1(\log_2 n) - n^2 + \mathcal{O}(n \log n), \end{aligned}$$

where $P_1(z)$ is a one-periodic function given in the remark below. Moreover, variances and covariances of the total path length and Wiener index of PATRICIA tries are given by

$$\begin{aligned} \mathrm{Var}(T_n) &= nP_2(\log_2 n) + \mathcal{O}(1), \\ \mathrm{Cov}(T_n, W_n) &= n^2 P_2(\log_2 n) + \mathcal{O}(n \log n), \\ \mathrm{Var}(W_n) &= n^3 P_2(\log_2 n) + \mathcal{O}(n^2 \log n), \end{aligned}$$

where $P_2(z)$ is again a one-periodic function given in the remark below.

Remark 16. The result for the mean of the external path length was first derived in [127]. The result for the variance of the total path length is due to Kirschenhofer, Prodinger and Szpankowski [118]. In [75], we obtained the expressions for the period functions

$$P_1(z) = \frac{\gamma - 1}{\log 2} + \frac{1}{\log 2} \sum_{k \neq 0} \Gamma(-\chi_k) e^{2k\pi i z}$$

and

$$P_2(z) = \frac{1}{\log 2} \sum_k G_2(-1-\chi_k) e^{2k\pi i z},$$

where

$$\begin{aligned} G_2(\omega) =& \Gamma(\omega+1)\left(2^{\omega+1}(\omega+2) - \frac{\omega^2+3\omega+6}{4}\right) \\ &+ 2^{\omega+2} \sum_{l\geq 1} \frac{(-1)^l \Gamma(\omega+l+2)}{(l-1)!(2^l-1)}. \end{aligned}$$

The latter result again implies the following bivariate central limit theorem.

Theorem 4.2.14. *We have,*

$$\left(\frac{T_n - \mathbb{E}(T_n)}{\sqrt{\mathrm{Var}(T_n)}}, \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathrm{Var}(W_n)}}\right) \xrightarrow{d} (X, X),$$

where X is a standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

Remark 17. Up to our knowledge, this result was first obtained by Neininger and Rüschendorf in [160].

Internal Wiener Index of m-ary PATRICIA tries. First, observe that the internal path length and internal Wiener index satisfy the following distribution recurrences for $n \geq 2$

$$N_n \overset{d}{=} \begin{cases} \sum_{i=1}^m N^{(i)}_{I_n^{(i)}} + 1, & \text{if } I_n^{(i)} \neq n \text{ for all } i, \\ N_n, & \text{otherise}, \end{cases} \tag{4.31}$$

$$T_n \overset{d}{=} \begin{cases} \sum_{i=1}^m \left(T^{(i)}_{I_n^{(i)}} + N^{(i)}_{I_n^{(i)}}\right), & \text{if } I_n^{(i)} \neq n \text{ for all } i; \\ T_n, & \text{otherwise} \end{cases} \tag{4.32}$$

and

$$W_n \overset{d}{=} \begin{cases} \begin{matrix} \sum_{i=1}^m \left(W^{(i)}_{I_n^{(i)}} + T^{(i)}_{I_n^{(i)}} + N^{(i)}_{I_n^{(i)}}\right) \\ + \sum_{(i,j)\in S_2} N^{(i)}_{I_n^{(i)}} \left(T^{(j)}_{I_n^{(j)}} + N^{(j)}_{I_n^{(j)}}\right) \end{matrix}, & \text{if } I_n^{(i)} \neq n \text{ for all } i, \\ W_n, & \text{otherwise}, \end{cases} \tag{4.33}$$

where notation is as in Section 4.2.1, $S_2 = \{(i,j) \ : \ 1 \leq i, \leq m, i \neq j\}$ and $T_0 = T_1 = W_0 = W_1 = 0$.

Theorem 4.2.15. *Consider m-ary PATRICIA tries built on strings with digits from alphabet $\mathcal{S} = \{a_1, \dots, a_m\}$. Suppose that the probability for a digit of the random string being a_i is p_i for all $1 \le i \le m$. Set $h = -\sum_{i=1}^{m} p_i \log p_i$, then we have that for the mean of internal nodes, internal path length and internal Wiener index of m-ary PATRICIA tries, as $n \to \infty$,*

$$\begin{aligned}
\mathbb{E}(N_n) &\sim nP_1(\log_{1/a} n),\\
\mathbb{E}(T_n) &\sim h^{-1} n \log n P_1(\log_{1/a} n),\\
\mathbb{E}(W_n) &\sim h^{-1} n^2 \log n P_1(\log_{1/a} n)^2,
\end{aligned}$$

where $P_1(z)$ is a one-periodic function. Moreover, variances and covariances of the number of internal nodes, internal path length and internal Wiener index of m-ary PATRICIA tries are given by

$$\begin{aligned}
\operatorname{Var}(N_n) &\sim nP_2(\log_{1/a} n),\\
\operatorname{Cov}(N_n, T_n) &\sim h^{-1} n \log n P_2(\log_{1/a} n),\\
\operatorname{Var}(T_n) &\sim h^{-2} n \log^2 n P_2(\log_{1/a} n),\\
\operatorname{Cov}(N_n, W_n) &\sim 2h^{-1} n^2 \log n P_1(\log_{1/a} n) P_2(\log_{1/a} n),\\
\operatorname{Cov}(T_n, W_n) &\sim 2h^{-2} n^2 \log^2 n P_1(\log_{1/a} n) P_2(\log_{1/a} n),\\
\operatorname{Var}(W_n) &\sim 4h^{-2} n^3 \log^2 n P_1(\log_{1/a} n)^2 P_2(\log_{1/a} n),
\end{aligned}$$

where $Q(z)$ is again a one-periodic function. In particular,

$$\rho(N_n, T_n) \longrightarrow 0, \quad \rho(N_n, W_n) \longrightarrow 0, \quad \rho(T_n, W_n) \longrightarrow 0,$$

where $\rho(\cdot, \cdot)$ denotes the correlation coefficient.

Theorem 4.2.16. *We have,*

$$\left(\frac{N_n - \mathbb{E}(N_n)}{\sqrt{\operatorname{Var}(N_n)}}, \frac{T_n - \mathbb{E}(T_n)}{\sqrt{\operatorname{Var}(T_n)}}, \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\operatorname{Var}(W_n)}} \right) \xrightarrow{d} (X, X, X).$$

4.3 Steiner Distance

4.3.1 Introduction

In this section, we are interested in two parameters, the k-th total path length and the total Steiner k-distance, which have not been analyzed for digital trees. We start with the definition of these two parameters.

For a given tree T with vertex set V and a subset $M \subset V$, the smallest spanning tree containing M is called the *Steiner tree* for M in T while the smallest subtree containing M and the root is the so-called *ancestor-tree* for M in T. The size of the Steiner tree for M in T (denoted by $S_M(T)$) and the size of the ancestor-tree for M in T (denoted by $D_M(T)$) are called the Steiner distance and the $|M|$-th path length, respectively. Furthermore, for the given tree T and integer $k \in \mathbb{N}$, the k-th total path length $P_k(T)$ and the Steiner k-distance $W_k(T)$ are defined as

$$P_k(T) = \sum_{|M|=k} D_M(T) \quad \text{and} \quad W_k(T) = \sum_{|M|=k} S_M(T).$$

Steiner trees and ancestor trees have many real-life applications, e.g. in transportation and multiprocessor networks [166], circuit layouts, internet communication [179] and many others. Consequently, the Steiner distance and the $|M|$-th path length are useful statistics. For example, when comparing the efficiency of communication potential of different networks, the Steiner distance can be used [28]. Moreover, the Steiner distance and k-th total path length have also applications to Multiple Quickselect algorithm [166] and the efficiency of certain traceroute algorithms [85].

In the last decade, several papers dedicated to the analysis of the two parameters in various random trees, including random increasing tree [166], random binary search tree [148], generalized random m-ary search tree [165], recursive trees [152, 164] and random simply generated trees [152, 164] have been published. As mentioned in [152], the size of a Steiner tree is related to the communication potential of its nodes. Thus, it is of interest to study the Steiner k-distance of different data structures, such as DSTs.

In this section, we again use the ”Poisson-Laplace-Mellin Method” to obtain the means, variances and covariances of the k-th total path length and the total Steiner k-distance for symmetric DSTs under the Bernoulli model. Limit laws for the two parameter are derived as well. In the remainder of this section, we use $P_n^{[k]}$ and $S_n^{[k]}$ to denote the k-th total path length and total k-th Steiner distance of symmetric random digital search trees built on n strings, respectively. Also, we use the common notation for the constant $Q_m = \prod_{j=1}^{m} (1 - 2^{-j})$ and $Q_\infty = \lim_{m \to \infty} Q_m$. The main results are:

Theorem 4.3.1. *We have that for $k \geq 2$,*

$$\mathbb{E}\left(P_n^{[k]}\right) \sim \mathbb{E}\left(S_n^{[k]}\right) \sim \frac{n^k \log_2 n}{(k-1)!}.$$

Moreover, the variance and covariance of $P_n^{[k]}$ and $S_n^{[k]}$ are given by

$$\begin{aligned}\mathrm{Var}\left(P_n^{[k]}\right) \sim \mathrm{Var}\left(S_n^{[k]}\right) &\sim n^{2k-1}\frac{2^{2-2k}}{Q_{k-1}^2}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right),\\ \mathrm{Cov}\left(P_n^{[k_1]},P_n^{[k_2]}\right) &\sim \mathrm{Cov}\left(S_n^{[k_1]},P_n^{[k_2]}\right)\\ &\sim n^{k_1+k_2-1}\frac{2^{2-k_1-k_2}}{Q_{k_1-1}Q_{k_2-1}}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right).\end{aligned}$$

where the expressions of C_{kps} and ϖ_{kps} can be derived from the results in Remark 5.

Theorem 4.3.2. *Let*

$$X_n^{[k]} = \frac{P_n^{[k]} - \mathbb{E}(P_n^{[k]})}{\sqrt{\mathrm{Var}(P_n^{[k]})}} \quad \text{and} \quad Y_n^{[k]} = \frac{S_n^{[k]} - \mathbb{E}(S_n^{[k]})}{\sqrt{\mathrm{Var}(S_n^{[k]})}}.$$

We have that for any $k \geq 2$,

$$\left(X_n^{[1]},\dots,X_n^{[k-1]},Y_n^{[k]}\right) \xrightarrow{d} (X,\dots,X),$$

where X is the standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

Remark 18. The asymptotics of $S_n^{[k]}$ can be explained intuitively. It is well-known that the expected value of the depth of a node is of order $\log_2 n$. For a Steiner tree, the size will be more or less the sum of the depth of the k chosen nodes. Thus, for k chosen nodes, the expected size of the Steiner tree will be of order $k\log_2 n$. Since there are $\binom{n}{k}$ ways to choose the k nodes, the mean of the total Steiner k-distance will be roughly $\binom{n}{k}k\log_2 n \sim \frac{n^k\log_2 n}{(k-1)!}$.

Remark 19. As we have seen, the leading terms for the asymptotics of k-th total path length and Steiner k-distance are the same. This is not surprising, intuitively speaking, because the k-subsets which are most relevant are those contain vertices from both subtrees for which the ancestor tree and the Steiner tree will be the same. This is similar to the distance between two random nodes (see [1, 3]) which is also twice the depth, because the most relevant cases are again those include the root.

Remark 20. In fact, we can find more terms in the asymptotic of the means, variances and covariances for $P_n^{[k]}$ and $S_n^{[k]}$ by the same method applied in the following sections. For example, let $\chi_m = 2m\pi i/\log 2$, we have that

$$\begin{aligned}&\mathbb{E}\left(P_n^{[k]}\right) \sim \mathbb{E}\left(S_n^{[k]}\right) + D^{[k]}n^k\\ &\qquad\sim \frac{n^k\log n}{(k-1)!} + \frac{n^k}{(k-1)!}\left(c_k + \frac{e_k}{k} + \frac{1}{\log 2}\sum_{m\in\mathbb{Z}\backslash\{0\}}\frac{G_k(\chi_m)n^{\chi_m}}{\Gamma(k+1+\chi_m)}\right).\end{aligned}$$

where

$$G_k(\chi_m) = \Gamma(k+1-\chi_m)\Gamma(-1-\chi_m), \quad D^{[k]} = \frac{1}{k!(2^{k-1}-1)}$$

and the constant c_k is given by

$$c_k = \frac{\gamma - 1}{\log 2} + \frac{1}{2} - \sum_{j\geq 1} \frac{(k-1)!}{2^j - 1} + \frac{(k-1)! d_k}{\log 2}.$$

In the expression, d_k is defined recursively as $d_1 = 0$ and

$$d_k = \frac{1}{2^{k-1}-1} \sum_{r=1}^{k-1} \frac{d_{k-r}}{r!} - \frac{2^{k-1}}{2^{k-1}-1} \frac{\log 2}{(k-1)!}.$$

Also, the sequence $\{e_k\}_{k\geq 1}$ is defined recursively as $e_1 = 0$ and

$$e_k = \frac{1}{2^{k-1}-1} \sum_{r=1}^{k-2} \frac{k!}{r!} e_{k-r} + \frac{2^k - 1}{2^{k-1}-1}, \quad \text{for } k \geq 2.$$

We state the main result in the form of Theorem 1 because the leading term is the most interesting part and it would be enough for proving the central limit theorem. Also, computing more terms can be extremely complicated. As we see from the above statements, the difference between the asymptotics of the two shape parameters is $\frac{n^k}{k!(2^{k-1}-1)}$. This can be explained heuristically. Let $d_n^{[k]}$ be the difference of the two shape parameters, then $d_n^{[k]} \sim 2d_{n/2}^{[k]} + 2\binom{\frac{n}{2}}{k}$ since the size of both subtrees will be roughly $n/2$ under the Bernoulli model. Iterating it, we get $d_n^{[k]} = \Theta(n^k)$, which matches the difference above.

Remark 21. Note that the Steiner k-distance is a generalization of the Wiener index, namely, for $k = 2$ we obtain the Wiener index. Thus, Theorem 4.2.1 is actually a special case of Theorem 4.3.1 with $k = 2$.

4.3.2 k-th Total Path Length

In this section, we start with the recurrence under the Bernoulli model and then use it to get the differential-functional equation of the Poisson model. The rest of the analysis will focus on the Poisson model, since the depoissonization is standard with the language of JS-admissible.

Mean of the k-th Total Path Length of DSTs

First, we start with deriving a distributional recurrence relation for the k-th total path length. Recall the notation $P_n^{[k]}$ for the k-th total path length from the introduction. Moreover, we will use the notation $B_n \stackrel{d}{=} \mathrm{Binom}(n, \frac{1}{2})$. Let a DST with $n+1$ nodes given. Depending on how the k nodes are chosen, there are 4 cases:

1. All k nodes are from one subtree.

The contribution to the k-th total path length will be

$$P_{B_n}^{[k]} + P_{n-B_n}^{[k]*} + \binom{B_n}{k} + \binom{n-B_n}{k},$$

where $P_{B_n}^{[k]}$ is independent of $P_{n-B_n}^{[k]*}$ and $P_{B_n}^{[k]} \stackrel{d}{=} P_{n-B_n}^{[k]*}$.

2. The k nodes are chosen from both subtrees and the root is not chosen.

We will have the contribution

$$\sum_{r=1}^{k-1}\left(\binom{n-B_n}{k-r}P_{B_n}^{[r]} + \binom{B_n}{r}P_{n-B_n}^{[k-r]*} + 2\binom{B_n}{r}\binom{n-B_n}{k-r}\right).$$

3. The root is chosen, the other $k-1$ nodes are all from one subtree.

It will contribute

$$P_{B_n}^{[k-1]} + P_{n-B_n}^{[k-1]*} + \binom{n-B_n}{k-1} + \binom{B_n}{k-1}.$$

4. The root is chosen, the other $k-1$ nodes are from both subtrees.

The contribution will be

$$\sum_{r=1}^{k-2}\left(\binom{n-B_n}{k-r-1}P_{B_n}^{[r]} + \binom{B_n}{r}P_{n-B_n}^{[k-r-1]*} + 2\binom{B_n}{r}\binom{n-B_n}{k-r-1}\right).$$

Combining all four cases, we get that for $n+1 \geq k \geq 1$:

$$\begin{aligned}P_{n+1}^{[k]} \stackrel{d}{=} P_{B_n}^{[k]} &+ P_{n-B_n}^{[k]*} + P_{B_n}^{[k-1]} + P_{n-B_n}^{[k-1]*} + \sum_{r=1}^{k-1}\left(\binom{n-B_n}{k-r}P_{B_n}^{[r]} + \binom{B_n}{r}P_{n-B_n}^{[k-r]*}\right)\\ &+ 2\binom{n}{k} + 2\binom{n}{k-1} + \sum_{r=1}^{k-2}\left(\binom{n-B_n}{k-r-1}P_{B_n}^{[r]} + \binom{B_n}{r}P_{n-B_n}^{[k-r-1]*}\right)\\ &- \binom{n-B_n}{k} - \binom{B_n}{k} - \binom{n-B_n}{k-1} - \binom{B_n}{k-1}.\end{aligned}$$

Note that from the above equation, we see that the k-th total path length depends on the 1-st, 2-nd,$\ldots$, $(k-1)$-th total path length. Thus, we actually have a system of recurrences. The initial conditions are $P_n^{[0]} = 0$ for all n and $P_n^{[k]} = 0$ for all $k > n$.

Let $\tilde{f}^{[k]}(z) = e^{-z} \sum_{n \geq 0} \mathbb{E}(P_n^{[k]}) \frac{z^n}{n!}$ which is the mean in the Poisson model. Then, from the recurrence relation above, we get

$$\begin{aligned}\tilde{f}^{[k]}(z) + \tilde{f}^{[k]\prime}(z) =& 2\tilde{f}^{[k]}\left(\frac{z}{2}\right) + 2\tilde{f}^{[k-1]}\left(\frac{z}{2}\right) + 2\sum_{r=1}^{k-1} \frac{\left(\frac{z}{2}\right)^r}{r!} \tilde{f}^{[k-r]}\left(\frac{z}{2}\right) \\ &+ 2\sum_{r=1}^{k-2} \frac{\left(\frac{z}{2}\right)^r}{r!} \tilde{f}^{[k-r-1]}\left(\frac{z}{2}\right) + 2\left(\frac{z^k - \left(\frac{z}{2}\right)^k}{k!} + \frac{z^{k-1} - \left(\frac{z}{2}\right)^{k-1}}{(k-1)!}\right).\end{aligned}$$

Note that when $k = 1$, the above equation will be exactly the same as the one derived in [74] and hence the order of $\tilde{f}^{[1]}(z)$ is known. Thus, by induction and the closure properties of JS-admissibility from [74], we get that

$$\tilde{f}^{[k]}(z) = \begin{cases} \mathcal{O}(z^{k+\epsilon}), & \text{as } z \to \infty; \\ \mathcal{O}(z^k), & \text{as } z \to 0^+ \end{cases}$$

uniformly for z with $|\arg z| \leq \frac{\pi}{2} - \epsilon$, where $\epsilon > 0$ is an arbitrary small constant. Applying Laplace transform, we get the differential-functional equation

$$\begin{aligned}(1+s)\mathscr{L}[\tilde{f}^{[k]}; s] =& 4\mathscr{L}[\tilde{f}^{[k]}; 2s] + 4\mathscr{L}[\tilde{f}^{[k-1]}; 2s] + 4\sum_{l=1}^{k-1} \frac{(-1)^l}{l!} \mathscr{L}^{(l)}[\tilde{f}^{[k-l]}; 2s] \\ &+ 4\sum_{l=1}^{k-2} \frac{(-1)^l}{l!} \mathscr{L}^{(l)}[\tilde{f}^{[k-l-1]}; 2s] + 2\left(\frac{1+s}{s^{k+1}} - \frac{1+2s}{2^k s^{k+1}}\right),\end{aligned}$$

where $\mathscr{L}^{(l)}[\tilde{f}^{[k-l]}; s]$ is the l-th differentiation of $\mathscr{L}[\tilde{f}^{[k-l]}; s]$. Let

$$Q(-s) = \prod_{j \geq 1} \left(1 - \frac{s}{2^j}\right) \quad \text{and} \quad \bar{\mathscr{L}}[\tilde{f}^{[k]}; s] = \frac{\mathscr{L}[\tilde{f}^{[k]}; s]}{Q(-s)}$$

and divide both sides of above equation by $Q(-2s)$. This yields

$$\begin{aligned}\bar{\mathscr{L}}[\tilde{f}^{[k]};s] =& 4\bar{\mathscr{L}}[\tilde{f}^{[k]};2s] + 4\bar{\mathscr{L}}[\tilde{f}^{[k-1]};2s] + 4\sum_{l=1}^{k-1}\frac{(-1)^l}{l!}\bar{\mathscr{L}}^{(l)}[\tilde{f}^{[k-l]};2s]\\ &+ 4\sum_{l=1}^{k-2}\frac{(-1)^l}{l!}\bar{\mathscr{L}}^{(l)}[\tilde{f}^{[k-l-1]};2s] + 2\left(\frac{1}{s^{k+1}Q(-s)} - \frac{1+2s}{2^k s^{k+1}Q(-2s)}\right)\\ &- 4\sum_{l=1}^{k-1}\sum_{r=0}^{l-1}\frac{(-1)^l}{r!(l-r)!}2^{r-l}\mathscr{L}^{(r)}[\tilde{f}^{[k-l]};2s]h^{(l-r)}(s)\\ &- 4\sum_{l=1}^{k-2}\sum_{r=0}^{l-1}\frac{(-1)^l}{r!(l-r)!}2^{r-l}\mathscr{L}^{(r)}[\tilde{f}^{[k-l-1]};2s]h^{(l-r)}(s),\end{aligned}$$

where $h(s) = \frac{1}{Q(-2s)}$ and $h^{(n)}(s)$ is the n-th derivative of $h(s)$. From the bound for $1/Q(-2s)$ obtained in [74]

$$\frac{1}{Q(-2s)} = \begin{cases} \mathcal{O}(s^{-b}), & \text{as } s \to \infty;\\ \mathcal{O}(1), & \text{as } s \to 0,\end{cases}$$

where b can be arbitrarily large, we obtain the bounds

$$\bar{\mathscr{L}}[\tilde{f}^{[k]};s] = \begin{cases} \mathcal{O}(|s|^{-b}), & \text{as } s \to \infty;\\ \mathcal{O}(|s|^{-(k+1+\epsilon)}), & \text{as } s \to 0^+,\end{cases}$$

and

$$h^{(n)}(s) = \begin{cases} \mathcal{O}(|s|^{-b}), & \text{as } s \to \infty;\\ \mathcal{O}(1), & \text{as } s \to 0^+,\end{cases}$$

uniformly for s with $|\arg(s)| \leq \pi - \epsilon$. We let

$$\begin{aligned}R^{[k]}(s) =& -4\sum_{l=1}^{k-1}\sum_{r=0}^{l-1}\frac{(-1)^l}{r!(l-r)!}2^{r-l}\mathscr{L}^{(r)}[\tilde{f}^{[k-l]};2s]h^{(l-r)}(s)\\ &- 4\sum_{l=1}^{k-2}\sum_{r=0}^{l-1}\frac{(-1)^l}{r!(l-r)!}2^{r-l}\mathscr{L}^{(r)}[\tilde{f}^{[k-l-1]};2s]h^{(l-r)}(s).\end{aligned}$$

Then, by Ritt's Theorem (Theorem 4.2 of [163]), we derive the bounds

$$R^{[k]}(s) = \begin{cases} \mathcal{O}(|s|^{-b}), & \text{as } s \to \infty;\\ \mathcal{O}(|s|^{-(k+\epsilon)}), & \text{as } s \to 0^+\end{cases}$$

uniformly for s with $|\arg z| \leq \pi - \epsilon$. Thus, we may apply the Mellin transform:

$$\begin{aligned}
\mathscr{M}[\bar{\mathscr{L}}^{[k]};\omega] =& \frac{2^{2-\omega}}{1-2^{2-\omega}}\mathscr{M}[\bar{\mathscr{L}}^{[k-1]};\omega] \\
&+ \frac{2^{2-\omega}}{1-2^{2-\omega}}\sum_{l=1}^{k-1}\frac{\prod_{i=1}^{l}(\omega-i)}{l!}\mathscr{M}[\bar{\mathscr{L}}^{[k-l]};\omega-l] \\
&+ \frac{2^{2-\omega}}{1-2^{2-\omega}}\sum_{l=1}^{k-2}\frac{\prod_{i=1}^{l}(\omega-i)}{l!}\mathscr{M}[\bar{\mathscr{L}}^{[k-l-1]};\omega-l] \\
&+ \frac{2}{1-2^{2-\omega}}\frac{Q(2^{\omega-k-1})}{Q(1)}\Gamma(k-\omega)\Gamma(\omega-k+1)(1-2^{-k}) \\
&+ \frac{2}{1-2^{2-\omega}}\frac{Q(2^{\omega-k})}{Q(1)}\Gamma(k+1-\omega)\Gamma(\omega-k)(1-2^{1-k}) \\
&+ \frac{\mathscr{M}[R^{[k]};\omega]}{1-2^{2-\omega}},
\end{aligned}$$

where for convenience, we use the notation $\mathscr{M}[\bar{\mathscr{L}}^{[k]};\omega]$ for $\mathscr{M}[\bar{\mathscr{L}}[\tilde{f}^{[k]};s];\omega]$. The fundamental strip of the above expression will be the half plane $\Re(\omega) > k+1$. To apply the inverse Mellin transform, we need to figure out all the singularities of the above expression. Since the case $k=1$ is already solved in [74] and the general case k will be determined by $1,\dots,k-1$, we get that for $k \geq 2$ the expression can be simplified as

$$\begin{aligned}
\mathscr{M}[\bar{\mathscr{L}}^{[k]};\omega] =& \frac{2^{2-\omega}}{1-2^{2-\omega}}\sum_{r=1}^{k-1}\frac{\prod_{i=1}^{r}(\omega-i)}{r!}\mathscr{M}[\bar{\mathscr{L}}^{[k-r]};\omega-r] \\
&+ \frac{1}{1-2^{2-\omega}}\frac{Q(2^{\omega-k-1})}{Q(1)}\Gamma(k-\omega)\Gamma(\omega-k+1)(2-2^{1-k}) + \bar{g}_k(\omega)
\end{aligned}$$

where $\bar{g}_k(\omega)$ is the sum of all the remaining terms in the expression. From the bound we derived for $R^{[k]}(s)$ and $\mathscr{L}[\tilde{f}^{[k]};s]$ and the properties of the Mellin transform [62], we get that if α is a singularity of $\bar{g}_k(\omega)$, then $\Re(\alpha) \leq k$. From [74], we have that

$$\mathscr{M}[\bar{\mathscr{L}}^{[1]};\omega] = \frac{G_1(\omega)}{1-2^{2-\omega}},$$

where

$$G_1(\omega) = \frac{Q(2^{\omega-2})}{Q(1)}\Gamma(\omega)\Gamma(1-\omega).$$

Plugging this into the recurrence and iterating, we get that for $k \geq 2$

$$\mathscr{M}[\bar{\mathscr{L}}^{[k]};\omega] = \frac{\prod_{i=1}^{k-1}(\omega-i)}{1-2^{k+1-\omega}} G_1(\omega-k+1)A_k(\omega) + T_k(\omega)G_1(\omega-k+1) + g_k(\omega)$$

where $g_k(\omega)$ is defined recursively by $g_1(\omega)=0$, $g_2(\omega)=\bar{g}_2(\omega)$ and

$$g_k(\omega) = \frac{2^{2-\omega}}{1-2^{2-\omega}} \sum_{r=1}^{k-1} \frac{\prod_{i=1}^{r}(\omega-i)}{r!} g_{k-r}(\omega-r) + \bar{g}_k(\omega).$$

Again, by similar argument as above, we have that if α is a singularity of $g_k(\omega)$, then $\Re(\alpha) \leq k$. The function $A_k(\omega)$ is defined recursively as $A_1(\omega) = 1$, $A_2(\omega) = \dfrac{1}{2^{\omega-2}-1}$ and

$$A_k(\omega) = \frac{2^{2-\omega}}{1-2^{2-\omega}} \sum_{r=1}^{k-1} \frac{A_{k-r}(\omega-r)}{r!}.$$

Also, $T_k(\omega)$ is defined recursively as $T_1(\omega)=0$, $T_2(\omega) = \frac{6}{4(1-2^{2-\omega})}$ and

$$T_k(\omega) = \frac{2^{2-\omega}}{1-2^{2-\omega}} \sum_{r=1}^{k-1} \frac{\prod_{i=1}^{k-1}(\omega-i)}{r!} T_{k-r}(\omega-r) + \frac{2(1-2^{-k})}{1-2^{2-\omega}}.$$

Note that one can easily prove that

$$A_k(k+1+\chi_m) = A_k(k+1) = \frac{1}{(k-1)!}$$

for $\chi_m = \dfrac{2i\pi m}{\log 2}$, $m \in \mathbb{Z}$ by induction. Moreover, the Laurent series of $A_k(\omega)$ at $\omega = k+1+\chi_r$ is given as

$$A_k(\omega) = \frac{1}{(k-1)!} + d_k(\omega-k-1) + \mathcal{O}((\omega-k-1)^2),$$

where $\{d_k\}_{k\geq 1}$ is a sequence which is defined recursively as $d_1 = 0$ and

$$d_k = \frac{1}{2^{k-1}-1} \sum_{r=1}^{k-1} \frac{d_{k-r}}{r!} - \frac{2^{k-1}}{2^{k-1}-1} \frac{\log 2}{(k-1)!}.$$

Because we have the explicit form of $G_1(\omega)$, we rewrite the expression as

$$\mathscr{M}[\bar{\mathscr{L}}^{[k]};\omega] = \frac{Q(2^{\omega-k-1})}{(1-2^{k+1-\omega})Q(1)} \Gamma(\omega)\Gamma(k-\omega)A_k(\omega) + g_k(\omega).$$

Finally, applying the inverse Mellin transform and collecting residues, we get that

$$\begin{aligned}\bar{\mathscr{L}}[\tilde{f}^{[k]}; s] =& k s^{-(k+1)} \log_2 \frac{1}{s} + s^{-(k+1)} \left(c'_k + e_k + \frac{1}{\log 2} \sum_{m\in\mathbb{Z}\setminus\{0\}} \frac{G_k(\chi_m)}{(k-1)!} s^{-\chi_m} \right) \\ &+ \mathcal{O}(|s|^{-k-\epsilon})\end{aligned}$$

where $G_k(\chi_m)$ is introduced in previous section, $e_k = T_k(k+1)$ and

$$c'_k = k \left(\frac{H_k - 1}{\log 2} + \frac{1}{2} - \sum_{j\geq 1} \frac{(k-1)!}{2^j - 1} + \frac{(k-1)! d_k}{\log 2} \right).$$

Note that the asymptotic hold uniformly as $|s| \to 0$ with $|\arg(s)| \leq \pi - \epsilon$. Finally, we apply Proposition 1 of [74] and obtain that, as $z \to \infty$,

$$\begin{aligned}\tilde{f}^{[k]}(z) =& \frac{z^k \log z}{(k-1)!} + \frac{z^k}{(k-1)!} \left(c_k + \frac{e_k}{k} + \frac{1}{\log 2} \sum_{m\in\mathbb{Z}\setminus\{0\}} \frac{G_k(\chi_m) z^{\chi_m}}{\Gamma(k+1+\chi_m)} \right) \\ &+ \mathcal{O}(z^{k-1+\epsilon}).\end{aligned}$$

Variance and Covariance of the k-th Total Path Length

Next, let us consider the variance. Here we introduce the poissonized variance and covariance as

$$\begin{aligned}\tilde{V}^{[k]}(z) =& \tilde{f}_2^{[k]}(z) - \tilde{f}^{[k]}(z)^2 - z\tilde{f}^{[k]\prime}(z)^2, \\ \tilde{C}^{[k_1,k_2]}(z) =& \tilde{f}_2^{[k_1,k_2]}(z) - \tilde{f}^{[k_1]}(z)\tilde{f}^{[k_2]}(z) - z\tilde{f}^{[k_1]\prime}(z)\tilde{f}^{[k_2]\prime}(z),\end{aligned}$$

where

$$\tilde{f}_2^{[k]}(z) = e^{-z} \sum_{n\geq 0} \mathbb{E}\left(P_n^{[k]2}\right) \frac{z^n}{n!} \quad \text{and} \quad \tilde{f}_2^{[k_1,k_2]}(z) = e^{-z} \sum_{n\geq 0} \mathbb{E}\left(P_n^{[k_1]} P_n^{[k_2]}\right) \frac{z^n}{n!}.$$

For detailed explanation of why we choose them this way, see [74]. Note that when $k_1 = k_2 = k$, $\tilde{V}^{[k]}(z) = \tilde{C}^{[k_1,k_2]}(z)$. Thus, we will consider only $\tilde{C}^{[k_1,k_2]}(z)$ in this section.

From the given definition, we derive that

$$\begin{aligned}\tilde{C}^{[k_1,k_2]}(z) + \tilde{C}^{[k_1,k_2]\prime}(z) =& \tilde{f}_2^{[k_1,k_2]}(z) + \tilde{f}_2^{[k_1,k_2]\prime}(z) - \tilde{f}^{[k_1]}(z)\tilde{f}^{[k_2]}(z) \\ & - z\tilde{f}^{[k_1]\prime}(z)\tilde{f}^{[k_2]\prime}(z) - \tilde{f}^{[k_1]\prime}(z)\tilde{f}^{[k_2]}(z) \\ & - \tilde{f}^{[k_1]}(z)\tilde{f}^{[k_2]\prime}(z) - \tilde{f}^{[k_1]\prime}(z)\tilde{f}^{[k_2]\prime}(z) \\ & - z\tilde{f}^{[k_1]\prime\prime}(z)\tilde{f}^{[k_2]\prime}(z) - z\tilde{f}^{[k_1]\prime}(z)\tilde{f}^{[k_2]\prime\prime}(z).\end{aligned}$$

From the recurrence of $P_{n+1}^{[k]}$, we derive the differential-functional equations of $\tilde{f}_2^{[k]}$ and $\tilde{f}_2^{[k_1,k_2]}$ and plug them into the above equation. Thus, by the same argument we used in the mean case, we find the bounds

$$\tilde{C}^{[k_1,k_2]}(z) = \begin{cases} \mathcal{O}(z^{k_1+k_2-1+\epsilon}), & \text{as } z \to \infty; \\ \mathcal{O}(z^{\max\{k_1,k_2\}}), & \text{as } z \to 0^+ \end{cases}$$

uniformly for z with $|\arg z| \le \frac{\pi}{2} - \epsilon$. With the help of computer algebra systems, we get that

$$\tilde{C}^{[k_1,k_2]}(z) + \tilde{C}^{[k_1,k_2]\prime}(z) = 2\sum_{r_1=1}^{k_1}\sum_{r_2=1}^{k_2}\left(\frac{z}{2}\right)^{k_1+k_2-r_1-r_2}\tilde{C}^{[k_1,k_2]}\left(\frac{z}{2}\right) + \tilde{g}^{[k_1,k_2]}(z).$$

Because the exact expression of $\tilde{g}_2^{[k_1,k_2]}(z)$ is way too complicated, we do not list the whole expression here. For the later computation, we only need the property that $\tilde{g}_2^{[k_1,k_2]}(z) = \mathcal{O}(z^{k_1+k_2-2})$ as $z \to \infty$. Similar to our analysis of the mean, we apply Laplace transform to the differential-functional equations and divide both sides by $Q(-2s)$. Let $k' = k_1 + k_2$, then

$$\bar{\mathscr{L}}[\tilde{C}^{[k_1,k_2]}; s] = 4\sum_{r_1=1}^{k_1}\sum_{r_2=1}^{k_2}(-1)^{k'-r_1-r_2}\bar{\mathscr{L}}^{(k'-r_1-r_2)}[\tilde{C}^{[r_1,r_2]}; 2s] + R_2^{[k_1,k_2]}(s),$$

where

$$\begin{aligned} R_2^{[k_1,k_2]}(s) =& (-1)^{k'}\frac{\mathscr{L}[\tilde{g}_2^{[m_1,m_2]}; s]}{Q(-2s)} \\ & - 4\sum_{r_1=1}^{k_1}\sum_{r_2=1}^{k_2}\sum_{j=1}^{k'-r_1-r_2}(-1)^{k'-r_1-r_2}\binom{k'-r_1-r_2}{j}\frac{h^{(j)}(s)}{2^j}L(s) \end{aligned}$$

with the function $L(s)$ defined as

$$L(s) = \mathscr{L}^{(k'-r_1-r_2-k)}[\tilde{C}^{[r_1,r_2]}; 2s].$$

Before we proceed to apply the Mellin transform, we derived similar bounds as in the analysis of the mean:

$$\bar{\mathscr{L}}[\tilde{C}^{[k_1,k_2]}; s] = \begin{cases} \mathcal{O}(|s|^{-b}), & \text{as } s \to \infty; \\ \mathcal{O}(|s|^{-(k'+\epsilon)}), & \text{as } s \to 0^+, \end{cases}$$

where b is a constant which can be arbitrarily large. Note that the bounds hold uniformly for $|\arg s| \le \pi - \epsilon$. Now, we apply the Mellin transform on both sides of the above equalities.

Again, we use the simplified notation $\mathscr{M}[\bar{\mathscr{L}}^{[k_1,k_2]};\omega] = \mathscr{M}[\bar{\mathscr{L}}[\tilde{C}^{[k_1,k_2]}];\omega]$. Then, the equation becomes

$$\begin{aligned}\mathscr{M}[\bar{\mathscr{L}}^{[k_1,k_2]};\omega] =& 2^{2-\omega}\sum_{r_1=1}^{k_1}\sum_{r_2=1}^{k_2}\mathscr{M}[\bar{\mathscr{L}}^{[r_1,r_2]};\omega+r_1+r_2-k']\prod_{i=1}^{k'-r_1-r_2}(\omega-i)\\ &+\mathscr{M}[R_2^{[k_1,k_2]};\omega]\end{aligned}$$

for $\Re(\omega) > k'$. From [74], we already have that

$$\mathscr{M}[\bar{\mathscr{L}}^{[1]};s];\omega] = \frac{H_1(\omega)}{1-2^{2-\omega}},$$

where

$$H_1(\omega) = Q_\infty \sum_{j,h,l\geq 0}\frac{(-1)^j 2^{-\binom{j+1}{2}+j(\omega-2)}}{Q_jQ_hQ_l2^{h+l}}\varphi(\omega;2^{-j-h}+2^{-j-l})$$

with

$$\begin{aligned}\varphi(\omega;x) &= \int_0^\infty \frac{s^{\omega-1}}{(s+1)(s+x)^2}ds\\ &= \begin{cases}\dfrac{\pi(1+x^{\omega-2}((\omega-2)\zeta+1-\omega))}{(x-1)^2\sin(\pi\omega)}, & \text{if } x\neq 1;\\[2ex] \dfrac{\pi(\omega-1)(\omega-2)}{2\sin(\pi\omega)}, & \text{if } x=1.\end{cases}\end{aligned}$$

Consequently, we can express $\mathscr{M}[\bar{\mathscr{L}}^{[k_1,k_2]};\omega]$ in terms of $H_1(\omega)$

$$\mathscr{M}[\bar{\mathscr{L}}^{[k_1,k_2]};\omega] = A_{k_1,k_2}(\omega)\frac{H_1(\omega+2-k')}{1-2^{k'-\omega}}\prod_{i=1}^{k'-2}(\omega-i)+\bar{g}_2^{[k_1,k_2]}(\omega),$$

where $A_{r_1,r_2}(\omega)$ satisfies the recurrence

$$A_{k_1,k_2}(\omega) = \frac{2^{2-\omega}}{1-2^{2-\omega}}\left(\sum_{r_1=1}^{k_1}\sum_{r_2=1}^{k_2}A_{r_1,r_2}(\omega+r_1+r_2-k')\right)$$

with the initial condition $A_{1,1}(\omega)=1$. Note that $\bar{g}_2^{[k_1,k_2]}(\omega)$ has no singularities with real part larger than $k'-1$. From above recurrence, we can easily prove that for all $k\in\mathbb{Z}$

$$A_{k_1,k_2}(k_1+k_2+\chi_k) = \frac{2^{(k_1-1)(k_1-2)/2}}{\prod_{j=1}^{k_1-1}(2^j-1)}\frac{2^{(k_2-1)(k_2-2)/2}}{\prod_{i=1}^{k_2-1}(2^i-1)} = \frac{2^{2-k'}}{Q_{k_1-1}Q_{k_2-1}}$$

by induction. For convenience, we set $C_{k_1,k_2}=A_{k_1,k_2}(k_1+k_2)$. Applying the inverse Mellin transform and collecting residues, we get

$$\mathscr{L}[\tilde{C}^{[k_1,k_2]};\omega]=\frac{s^{-k'}}{\log 2}\sum_{r\in\mathbb{Z}}C_{k_1,k_2}H_1(2+\chi_r)s^{-\chi_r}\prod_{i=2}^{k_1+k_2-1}(i+\chi_r)+\mathcal{O}(|s|^{1-k'})$$

uniformly as $|s|\to 0$ with $|\arg s|\leq \pi-\epsilon$. Finally, we apply inverse Laplace transform and Proposition 1 of [98] and obtain that, as $z\to\infty$,

$$\tilde{C}^{[k_1,k_2]}(z)=z^{k_1+k_2-1}C_{k_1,k_2}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right)+\mathcal{O}(|z|^{k_1+k_2-2+\epsilon}).$$

In particular,

$$\tilde{V}^{[k]}(z)=\frac{z^{2k-1}}{\log 2}C_{k,k}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right)+\mathcal{O}(|z|^{2k-2+\epsilon})$$

as $z\to\infty$.

Remark 22. Note that from the expression of C_{k_1,k_2}, we have $C^2_{k_1,k_2}=C_{k_1,k_1}C_{k_2,k_2}$. Thus,

$$\begin{aligned}\rho(P_n^{[k_1]},P_n^{[k_2]})=&\frac{\mathrm{Cov}(P_n^{[k_1]},P_n^{[k_2]})}{\sqrt{\mathrm{Var}(P_n^{[k_1]})\mathrm{Var}(P_n^{[k_2]})}}\\ \sim&\sqrt{\frac{n^{2k_1+2k_2-2}C^2_{m,m-1}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right)^2}{n^{2k_1+2k_2-2}C_{m,m}C_{m-1,m-1}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right)^2}}=1.\end{aligned}$$

Remark 23. Since we already know that $P_n^{[1]}$ satisfies a central limit theorem [98], together with the result in the above remark and applying similar argument as of [76], we obtain that

$$\left(\frac{P_n^{[1]}-\mathbb{E}(P_n^{[1]})}{\sqrt{\mathrm{Var}(P_n^{[1]})}},\ldots,\frac{P_n^{[k]}-\mathbb{E}(P_n^{[k]})}{\sqrt{\mathrm{Var}(P_n^{[k]})}}\right)\xrightarrow{d}(X,\ldots,X),$$

where X is a standard normal distributed random variable and $\xrightarrow{d}$ denotes weak convergence.

4.3.3 Total Steiner k-distance

Let $S_n^{[k]}$ be the Steiner k-distance. Then, using the same idea as for the k-th total path length, we consider four cases:

1. **All k nodes are from one subtree.**
   $S^{[k]}_{B_n} + S^{[k]*}_{n-B_n}$.

2. **The k nodes are chosen from both subtrees and the root is not chosen.**

$$\sum_{l=1}^{k-1} \left(\binom{n-B_n}{k-l} P^{[l]}_{B_n} + \binom{B_n}{l} P^{[k-l]*}_{n-B_n} + 2\binom{B_n}{l}\binom{n-B_n}{k-l} \right).$$

3. **The root is chosen, the other $k-1$ nodes are all from one subtree.**

$$P^{[k-1]}_{B_n} + P^{[k-1]*}_{n-B_n} + \binom{n-B_n}{k-1} + \binom{B_n}{k-1}.$$

4. **The root is chosen, the other $k-1$ nodes are from both subtrees.**

$$\sum_{l=1}^{k-2} \left(\binom{n-B_n}{k-l-1} P^{[l]}_{B_n} + \binom{B_n}{l} P^{[k-l-1]*}_{n-B_n} + 2\binom{B_n}{l}\binom{n-B_n}{k-l-1} \right).$$

Note that as for the k-th total path length, here we have a system of recurrences for the Steiner k-distance. Similar to the analysis of the k-th total path length, we let $\tilde{g}^{[k]}(z)$ be the Poisson generating function of the mean of the total Steiner k-distance, $\tilde{W}^{[k_1,k_2]}(z)$ be the Poissonized covariance of the total k_1-th Steiner distance and the total k_2-th total path length and $\tilde{V}^{[k]}_S(z)$ be the variance of the k-th Steiner distance. With the help from computer algebra systems, we get the differential-functional equations

$$\begin{aligned}
\tilde{g}^{[k]}(z) + \tilde{g}^{[k]\prime}(z) =& 2\tilde{g}^{[k]}\left(\frac{z}{2}\right) + 2\tilde{f}^{[k-1]}\left(\frac{z}{2}\right) + 2\sum_{r=1}^{k-1} \frac{(\frac{z}{2})^r}{r!} \tilde{f}^{[k-r]}\left(\frac{z}{2}\right) \\
&+ 2\sum_{r=1}^{k-2} \frac{(\frac{z}{2})^r}{r!} \tilde{f}^{[k-r-1]}\left(\frac{z}{2}\right) + \frac{2z^k - 4(\frac{z}{2})^k}{k!} \\
&+ \frac{2z^{k-1} - 2(\frac{z}{2})^{k-1}}{(k-1)!}, \\
\tilde{W}^{[k_1,k_2]}(z) + \tilde{W}^{[k_1,k_2]\prime}(z) =& 2\sum_{r=1}^{k_2} \left(\frac{z}{2}\right)^{k_2-r} \tilde{W}^{[k_1,r]}\left(\frac{z}{2}\right) \\
&+ 2\sum_{r_1=1}^{k_1-1}\sum_{r_2=1}^{k_2} \left(\frac{z}{2}\right)^{k'-r_1-r_2} \tilde{C}^{[r_1,r_2]}\left(\frac{z}{2}\right) \\
&+ \tilde{h}^{[k_1,k_2]}_2(z)
\end{aligned}$$

and

$$\begin{aligned}\tilde{V}_S^{[k]}(z)+\tilde{V}_S^{[k]\prime}(z)=&2\tilde{V}_S^{[k]}\left(\frac{z}{2}\right)+4\sum_{r=1}^{k-1}\left(\frac{z}{2}\right)^{k-r}\tilde{W}^{[k,r]}\left(\frac{z}{2}\right)\\&+4\sum_{r=1}^{k-1}\left(\frac{z}{2}\right)^{k-r}\tilde{W}^{[k,r]}\left(\frac{z}{2}\right)\\&+2\sum_{r_1=1}^{k-1}\sum_{r_2=1}^{k-1}\left(\frac{z}{2}\right)^{2k-r_1-r_2}\tilde{C}^{[r_1,r_2]}\left(\frac{z}{2}\right)+\tilde{h}_S^{[k]}(z).\end{aligned}$$

We use $\tilde{h}_2^{[k_1,k_2]}(z)$ and $\tilde{h}_S^{[k]}(z)$ to denote the lower order terms. Because the rest of the analysis will be very similar to the one with the k-th total path length, we skip the details and list only the results

$$\begin{aligned}\mathbb{E}\left(S_n^{[k]}\right)=&\frac{n^k\log n}{(k-1)!}+\frac{n^k}{(k-1)!}\left(c_k+\frac{e_k}{k}-D^{[k]}\right)\\&+\frac{n^k}{(k-1)!}\frac{1}{\log 2}\sum_{r\in\mathbb{Z}\setminus\{0\}}\frac{G_k(\chi_r)n^{\chi_r}}{\Gamma(k+1+\chi_r)}\\&+\mathcal{O}(n^{k-1+\epsilon}),\\\operatorname{Cov}\left(S_n^{[k_1]},P_n^{[k_2]}\right)=&\frac{n^{k_1+k_2-1}}{\log 2}C_{k_1,k_2}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right)+\mathcal{O}(n^{k_1+k_2-2}),\\\operatorname{Var}\left(S_n^{[k]}\right)=&\frac{n^{2k-1}}{\log 2}C_{k,k}\left(C_{kps}+\varpi_{kps}(\log_2 n)\right)+\mathcal{O}(n^{2k-2}).\end{aligned}$$

Since the leading terms are exactly the same as for the k-th total path length, the same arguments as for $P_n^{[k]}$ gives us the results stated in Theorem 4.3.1.

Chapter 5

A General Framework for Central Limit Theorems

5.1 Framework for m-ary Tries

In this section, we will discuss a general framework for the limiting distribution of additive shape parameters in random digital trees. For m-ary tries and PATRICIA tries, an additive shape parameter is defined as follows: X_n is a sequence of random variables satisfying the distributional recurrence

$$X_n \overset{d}{=} \sum_{r=1}^{m} X^{(r)}_{I_n^{(r)}} + T_n, \qquad (n \geq n_0), \tag{5.1}$$

where $n_0 \geq 0$ is an integer, $X_n, X_n^{(1)}, \ldots, X_n^{(m)}, (I_n^{(1)}, \ldots, I_n^{(m)})$, T_n are independent and $X_n^{(i)}$ has the same distribution as X_n. The random model we are using is the Bernoulli model which is introduced in Chapter 2. For digital search trees and bucket digital search trees, the distributional recurrence will be

$$X_{n+b} \overset{d}{=} \sum_{r=1}^{m} X^{(r)}_{I_n^{(r)}} + T_{n+b}, \qquad (n \geq n_0), \qquad \text{where } b \geq 1 \text{ is an integer.} \tag{5.2}$$

The remaining notations are as in the trie case.

Because of the development of related mathematical techniques, including poissonization, poissonized variance with correction, Mellin transform and contraction method, we have many tools to characterize the asymptotics of additive shape parameters under the Bernoulli model. The authors of [77] and [75] proposed a systematical way to derive the asymptotics for mean and variance and the limit laws of additive shape parameters of random tries. It turns out that the same method works for random digital search trees as well.

Definition 5.1.1. *If a set* $\mathbf{P} = \{p_1, \ldots, p_m\}$ *satisfies that* $p_i \in (0,1)$ *for all* $1 \leq i \leq m$ *and* $\sum_i p_i = 1$*, then we say* $\mathbf{P}$ *is a* **probability family***.*

For a probability family $\mathbf{P} = \{p_1, \ldots, p_m\}$*, if there exists a constant* $a \in \mathbb{R}$ *and a sequence* $\{k_i\}_{i=1}^m$*,* $k_i \in \mathbb{N}$ *for all* $1 \leq i \leq m$ *such that* $p_i = a^{k_i}$ *for all* i*, then we say* $\mathbf{P}$ *is* **periodic***. Otherwise,* $\mathbf{P}$ *is said to be* **aperiodic***.*

For a probability family $\mathbf{P} = \{p_1, \ldots, p_m\}$, we define a function

$$\Lambda(s) = 1 - p_1^{-s} - \cdots - p_m^{-s}.$$

We let $\mathscr{Z}$ be the set of roots of $\Lambda(s) = 0$ and define the following notations

$$\mathscr{Z}_{<\alpha} = \mathscr{Z} \cap \{\Re(z) < \alpha\} \quad \text{and} \quad \mathscr{Z}_{=\alpha} = \mathscr{Z} \cap \{\Re(z) = \alpha\}.$$

Then from [55] and [67], we have the following properties

Theorem 5.1.2. *Depending on the real part of the solutions of* $\Lambda(s)$*, we have three cases:*

(i) If $\Re(s) < -1$*, then* $\Lambda(s)$ *has no solutions. In other words,* $\mathscr{Z}_{<-1} = \emptyset$*.*

(ii) If $\Re(s) = -1$*, then* $\mathscr{Z}_{=-1} = \{-1\} \cup S$ *where*

$$S = \begin{cases} \{-1 + \chi_k | \chi_k = 2k\pi i / \log a, k \in \mathbb{Z} \setminus \{0\}\}, & \mathbf{P} \textit{is periodic;} \\ \emptyset, & \mathbf{P} \textit{ is aperiodic.} \end{cases}$$

(iii) If $\Re(s) > -1$*, then there exists a positive constant* η *such that for any solutions* ω_1, ω_2*, we have* $|\omega_1 - \omega_2| > \eta$*.*

Lemma 5.1.3. *Let* $\tilde{f}(z)$ *and* $\tilde{h}(z)$ *be entire functions satisfying a functional equation of the form*

$$\tilde{f}(z) = \sum_{r=1}^{m} \tilde{f}(p_r z) + \tilde{h}(z) \tag{5.3}$$

where $\{p_1, \ldots, p_m\}$ *forms a probability family. We denote by* $h = -\sum_{r=1}^{m} p_r \log p_r$*. If* $\tilde{h}(z) \in \mathscr{J}\mathscr{S}_{\alpha,\gamma}$ *with* $0 \leq \alpha < 1$ *and* $\tilde{f}(0) = \tilde{f}'(0) = 0$*, then*

$$\tilde{f}(z) = \frac{1}{h} \sum_{\omega_k \in \mathscr{Z}_{<-\alpha-\epsilon}} G(\omega_k) z^{-\omega_k} + \mathcal{O}(z^{\alpha+\epsilon}),$$

where the sum expression is infinitely differentiable and

$$G(\omega) = \int_0^\infty z^{\omega-1} \tilde{h}(z) dz = \mathscr{M}[\tilde{h}; \omega].$$

Proof. Since $\tilde{h}(z) \in \mathscr{J}\mathscr{S}_{\alpha,\gamma}$ with $0 \leq \alpha < 1$, by a similar proof as of Proposition 3.3 in [75], we get that $\tilde{f}(z) = \mathcal{O}(z)$ as $z \to \infty$. At the same time, the assumptions that $\tilde{f}(0) = \tilde{f}'(0) = 0$ imply that $\tilde{f}(z) = \mathcal{O}(z^2)$ as $z \to 0$. Thus, the Mellin transform of $\tilde{f}(z)$ exists in the strip $\langle -2, -1\rangle$ and from (5.3), we get

$$\mathscr{M}[\tilde{f};\omega] = \frac{G(\omega)}{\Lambda(\omega)}, \quad \text{for } -2 < \Re(\omega) < -1.$$

By the converse mapping theorem, Theorem 3.3.10 and Theorem 5.1.2, we get the desired result. □

Now, we consider the moment generating function of X_n

$$M_n(y) := \mathbb{E}\left(e^{X_n y}\right).$$

Then, by (5.1), we get that

$$M_n(y) = \mathbb{E}\left(e^{T_n y}\right) \sum_{j_1+\cdots+j_m=n} \pi_{j_1,\ldots,j_m} M_{j_1}(y)\cdots M_{j_m}(y), \quad (n \geq n_0),$$

where

$$\pi_{j_1,\ldots,j_m} = \binom{n}{j_1 \cdots j_m} p_1^{j_1} \cdots p_m^{j_m}.$$

Let $\mu_n = \mathbb{E}(X_n)$ and $s_n = \mathbb{E}\left(X_n^2\right)$, then from the definition of the moment generating function, we get that

$$\begin{aligned}
\mu_n = M_n'(0) &= \sum_{j_1+\cdots+j_m=n} \pi_{j_1,\ldots,j_m} \sum_{r=1}^{m} \mu_{j_r} + \mathbb{E}(T_n),\\
s_n = M_n''(0) &= \sum_{j_1+\cdots+j_m=n} \pi_{j_1,\ldots,j_m} \sum_{r=1}^{m} \left(s_{j_r} + 2\mathbb{E}(T_n)\mu_{j_r}\right)\\
&\quad + \sum_{r\neq s} \sum_{j_1+\cdots+j_m=n} \pi_{j_1,\ldots,j_m} \mu_{j_r}\mu_{j_s} + \mathbb{E}(T_n^2). \qquad (5.4)
\end{aligned}$$

For the sake of simplicity, from now on, we assume that $X_0 = X_1 = 0$ and $n_0 = 2$. For more general cases, our method will also apply with slight modifications.

Now, we utilize the idea of Poissonization which was already used in previous sections. We let

$$\tilde{f}_1(z) = e^{-z} \sum_{n\geq 0} \mathbb{E}(X_n) \frac{z^n}{n!}, \qquad \tilde{f}_2(z) = e^{-z} \sum_{n\geq 0} \mathbb{E}(X_n^2) \frac{z^n}{n!}$$

$$\tilde{h}_1(z) = e^{-z} \sum_{n\geq 0} \mathbb{E}(T_n) \frac{z^n}{n!}, \qquad \tilde{h}_2(z) = e^{-z} \sum_{n\geq 0} \mathbb{E}(T_n^2) \frac{z^n}{n!},$$

then (5.4) yields that

$$\tilde{f}_1(z) = \sum_{r=1}^{m} \tilde{f}_1(p_r z) + \tilde{h}_1(z),$$
$$\tilde{f}_2(z) = \sum_{r=1}^{m} \tilde{f}_2(p_r z) + \sum_{r\neq s} \tilde{f}_1(p_r z)\tilde{f}_1(p_s z) + \tilde{h}_2(z) + \tilde{g}(z), \qquad (5.5)$$

where

$$\tilde{g}(z) = 2e^{-z} \sum_{n\geq 0} \sum_{j_1+\cdots+j_m=n} \pi_{j_1,\ldots,j_m} \mathbb{E}(T_n) \left(\sum_{r=1}^{m} \mu_{j_r} \right) \frac{z^n}{n!}.$$

Next, we utilize the idea of Poissonized variance with correction and let

$$\tilde{V}_X(z) = \tilde{f}_2(z) - \tilde{f}_1(z)^2 - z\tilde{f}_1'(z)^2,$$
$$\tilde{V}_T(z) = \tilde{h}_2(z) - \tilde{h}_1(z)^2 - z\tilde{h}_1'(z)^2.$$

From (5.5), we derive that

$$\tilde{V}_X(z) = \sum_{r=1}^{m} \tilde{V}_X(p_r z) + \tilde{V}_T(z) + \tilde{\phi}_1(z) + \tilde{\phi}_2(z), \qquad (5.6)$$

where

$$\tilde{\phi}_1(z) = \tilde{g}(z) - 2\tilde{h}_1(z) \sum_{r=1}^{m} \tilde{f}_1(p_r z) - 2z\tilde{h}_1'(z) \sum_{r=1}^{m} p_r \tilde{f}_1'(p_r z),$$
$$\tilde{\phi}_2(z) = z \sum_{r<s} p_r p_s \left(\tilde{f}_1'(p_r z) - \tilde{f}_1'(p_s z) \right)^2.$$

Before we go on to derive asymptotic expressions, we introduce the Hadamard product of Poisson generating functions.

Definition 5.1.4. *Given two Poisson generating functions*

$$\tilde{F}_1(z) = e^{-z} \sum_{n\geq 0} \frac{a_n}{n!} z^n \quad \textit{and} \quad \tilde{F}_2(z) = e^{-z} \sum_{n\geq 0} \frac{b_n}{n!} z^n,$$

we define the Hadamard product of these two functions as

$$\tilde{F}_3(z) := \tilde{F}_1(z) \odot \tilde{F}_2(z) = e^{-z} \sum_{n\geq 0} \frac{a_n b_n}{n!} z^n.$$

Note that the definition is different from the usual one since we consider the exponential generating function.

Subsequently, we will use Hadamard products to handle the function $\tilde{\phi}_1$ in (5.6). For this, we will need the following theorem which shows that JS-admissibility is closed under the Hadamard product.

Theorem 5.1.5. *If $\tilde{F}_1 \in \mathscr{J}\mathscr{S}_{\alpha_1,\beta_1}$ and $\tilde{F}_2 \in \mathscr{J}\mathscr{S}_{\alpha_2,\beta_2}$, then $\tilde{F}_3 \in \mathscr{J}\mathscr{S}_{\alpha_1+\alpha_2,\beta_1+\beta_2}$. More precisely, we have*

$$\tilde{F}_3(z) = \tilde{F}_1(z)\tilde{F}_2(z) + z\tilde{F}_1'(z)\tilde{F}_2'(z) + \mathcal{O}\left(|z|^{\alpha_1+\alpha_2-2}(\log_+ |z|)^{\beta_1+\beta_2}\right),$$

uniformly as $|z| \to \infty$ and $|\arg(z)| \le \theta$, where $0 < \theta < \pi/2$.

Proof. See the proof of Proposition 3.5 of [75]. □

We now can state the result on asymptotic expressions of mean and variance. (This result was first obtained by Fuchs et al. in [75].)

Proposition 5.1.6. *If $\tilde{h}_1(z) \in \mathscr{J}\mathscr{S}_{\alpha_1,\gamma_1}$ with $0 \le \alpha_1 < 1$, then*

$$\mathbb{E}(X_n) = \frac{1}{h} \sum_{\omega_k \in \mathcal{Z}_{<-\alpha_1-\epsilon}} G_E(\omega_k) n^{-\omega_k} + \mathcal{O}(n^{\alpha_1+\epsilon}),$$

where the sum expression is infinitely differentiable and

$$G_E(\omega) = \mathscr{M}[\tilde{h}_1; \omega] = \int_0^\infty \tilde{h}(z) z^{\omega-1} dz.$$

Moreover, if $\tilde{V}_T(z) \in \mathscr{J}\mathscr{S}_{\alpha_2,\gamma_2}$ with $0 \le \alpha_2 < 1$ and $\tilde{h}_2(z) \in \mathscr{J}\mathscr{S}$, then

$$\mathrm{Var}(X_n) \sim \frac{1}{h} \sum_{\omega_k \in \mathcal{Z}_{=-1}} G_V(\omega_k) n^{-\omega_k},$$

where the sum expression is infinitely differentiable and

$$G_V(\omega) = \mathscr{M}[\tilde{V}_T + \tilde{\phi}_1 + \tilde{\phi}_2; \omega] = \int_0^\infty \left(\tilde{V}_T(z) + \tilde{\phi}_1(z) + \tilde{\phi}_2(z)\right) z^{\omega-1} dz.$$

Proof. The expression of the mean follows directly from (5.5), Lemma 5.1.3 and depoissonization.

For the variance, we start from (5.6). We apply Theorem 5.1.5 to $\tilde{g}(z)$, which yields

$$\begin{aligned}\tilde{g}(z) =&2e^{-z}\sum_{n\geq 0}\sum_{j_1+\cdots+j_m=n}\pi_{j_1,\dots,j_m}\mathbb{E}(T_n)\left(\sum_{r=1}^{m}\mu_{j_r}\right)\frac{z^n}{n!}\\ =&2e^{-z}\sum_{n\geq 0}\sum_{j_1+\cdots+j_m=n}\pi_{j_1,\dots,j_m}\mathbb{E}(T_n)\left(\mu_n-\mathbb{E}(T_n)\right)\frac{z^n}{n!}\\ =&2\tilde{f}_1(z)\odot\tilde{h}_1(z)-2\tilde{h}_1(z)\odot\tilde{h}_1(z)\\ =&2\tilde{h}_1(z)\left(\sum_{r=1}^{m}\tilde{f}_1(p_rz)+\tilde{h}_1(z)\right)+2z\tilde{h}_1'(z)\left(\sum_{r=1}^{m}p_r\tilde{f}_1'(p_rz)+\tilde{h}_1'(z)\right)\\ &-2\tilde{h}_1(z)^2-2z\tilde{h}_1'(z)^2+\mathcal{O}(1)\\ =&2\tilde{h}_1(z)\sum_{r=1}^{m}\tilde{f}_1(p_rz)+2z\tilde{h}_1'(z)\sum_{r=1}^{m}p_r\tilde{f}_1'(p_rz)+\mathcal{O}(1).\end{aligned}$$

Plugging the result into the expression of $\tilde{\phi}_1(z)$, we get that $\tilde{\phi}_1(z)=\mathcal{O}(1)$.

Now, we turn to $\tilde{\phi}_2(z)$. First, by applying the Mellin transform to (5.5), we get that for $-2<\Re(\omega)<-1$,

$$\mathscr{M}[\tilde{f}_1;\omega]=\frac{G_E(\omega)}{\Lambda(\omega)}.$$

Thus, from inverse Mellin transform,

$$\begin{aligned}\tilde{f}_1'(z)=\frac{d}{dz}\tilde{f}_1(z)=&\frac{1}{2i\pi}\int_{(-1-\epsilon)}\frac{G_E(\omega)}{\Lambda(\omega)}\left(\frac{d}{dz}z^{-\omega}\right)dz\\ =&\frac{-1}{2i\pi}\int_{(-1-\epsilon)}\frac{G_E(\omega)}{\Lambda(\omega)}\omega z^{-\omega-1}d\omega.\end{aligned}$$

Therefore, we get

$$\begin{aligned}\tilde{f}_1'(p_rz)-\tilde{f}_1'(p_sz)=&\frac{-1}{2i\pi}\int_{(-1-\epsilon)}\frac{G_E(\omega)\omega}{\Lambda(\omega)}\left(p_r^{-\omega-1}-p_s^{-\omega-1}\right)z^{-\omega-1}d\omega\\ =&o(1),\end{aligned}$$

where the latter follows from the fact that the integral has no poles at $\Re(\omega)=-1$. As a result, $\tilde{\phi}_2(z)=o(|z|)$ as $z\to\infty$ which in turn shows that $\mathscr{M}[\tilde{\phi}_2;\omega]$ has no poles at $\Re(\omega)=-1$. Now, the converse mapping theorem proves the claimed expansion for $\tilde{V}(z)$. Moreover, by JS-admissibility, the expansion holds for $\mathrm{Var}(X_n)$, as well.

□

Now, we can state the general central limit theorem.

Theorem 5.1.7. *Suppose that $\tilde{h}_1(z) \in \mathscr{J}\mathscr{S}_{\alpha_1,\gamma_1}$ with $0 \le \alpha_1 < 1/2$, $\tilde{h}_2(z) \in \mathscr{J}\mathscr{S}$ and $\tilde{V}_T(z) \in \mathscr{J}\mathscr{S}_{\alpha_2,\gamma_2}$ with $0 \le \alpha_2 < 1$. Moreover, we assume that $\|T_n\|_s = o(\sqrt{n})$ with $2 < s \le 3$ and $\mathbb{V}(X_n) \ge cn$ for all n large enough and some $c > 0$. Then, as $n \to \infty$,*

$$\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \xrightarrow{d} \mathcal{N}(0,1).$$

Proof. From Proposition 5.1.6, we get that

$$\begin{aligned}\mathbb{E}(X_n) =& \frac{1}{h} \sum_{\omega_k \in \mathcal{Z}_{<-\alpha_1-\epsilon}} G_E(\omega_k) n^{-\omega_k} + \mathcal{O}(n^{\alpha_1+\epsilon}),\\ \mathbb{V}(X_n) \sim& \frac{1}{h} \sum_{\omega_k \in \mathcal{Z}_{=-1}} G_V(\omega_k) n^{-\omega_k}.\end{aligned}$$

From the assumption, we can choose ϵ such that $\alpha_1 + \epsilon < 1/2$. Next, we set

$$\begin{aligned}\varpi_1(x) &= \sum_{\omega_k \in \mathcal{Z}_{<-\alpha_1-\epsilon}} \frac{G_1(\omega_k)}{h} x^{-\omega_k},\\ \varpi_2(x) &= \sum_{\omega_k \in \mathcal{Z}_{=-1}} \frac{G_2(\omega_k)}{h} x^{-\omega_k-1}.\end{aligned}$$

To apply the contraction method, we need to verify the following conditions:

(a)

$$\left(\frac{I_n^{(r)} \varpi_2(I_n^{(r)})}{n\varpi_2(n)}\right)^{1/2} \xrightarrow{L_s} A_r, \quad \sum_{r=1}^{m} A_r^2 = 1 \quad \text{and} \quad \mathbb{P}(\exists r : A_r = 1) < 1.$$

(b)

$$(n\varpi_2(n))^{-1/2} \left(T_n - \varpi_1(n) + \sum_{r=1}^{m} \varpi_1(I_n^{(r)})\right) \xrightarrow{\mathcal{L}_s} 0.$$

We begin with the verification of (a). By the strong law of large number and the dominating converge theorem,

$$I_n^{(r)} \xrightarrow{L_p} p_r, \quad 1 \le r \le m. \tag{5.7}$$

Moreover, by the definition of $\varpi_2(x)$, we have that

$$\varpi_2(p_r n) = \varpi_2(n) \text{ for all } 1 \le r \le m \quad \text{and} \quad \varpi_2'(n) = \mathcal{O}(n^{-1}).$$

By the Taylor series expansion of ϖ_2:

$$\varpi_2(I_n^{(r)}) = \varpi_2(n) + \mathcal{O}\left(\left|\frac{I_n^{(r)}}{n} - p_r\right|\right) \quad \text{for } 1 \le r \le m.$$

This implies that

$$\frac{\varpi_2(I_n^{(r)})}{\varpi_2(n)} - 1 = \frac{1}{\varpi_2(n)}\mathcal{O}\left(\left|\frac{I_n^{(r)}}{n} - p_r\right|\right) \xrightarrow{a.s} 0 \tag{5.8}$$

Combining (5.7) and (5.8), we get

$$\left(\frac{I_n^{(r)}}{n}\frac{\varpi_2(I_n^{(r)})}{\varpi_2(n)}\right)^{1/2} \xrightarrow{L_s} p_r^{1/2} = A_r.$$

Moreover,

$$\sum_{r=1}^{m} A_r^2 = \sum_{r=1}^{m} p_r = 1 \quad \text{and} \quad \mathbb{P}(\exists r : A_r = 1) < 1$$

and hence the condition (a) is verified.

Now, we turn to the verification of condition (b). Note that from the assumption on $\|T_n\|_s$ and $\mathbb{V}(X_n)$, the term T_n can be dropped from (b). Therefore, we only need to check that

$$(n\varpi_2(n))^{-1/2}\left(\sum_{r=1}^{m} \varpi_1(I_n^{(r)}) - \varpi_1(n)\right) \xrightarrow{L_s} 0. \tag{5.9}$$

We let

$$A_n = \bigcap_{r=1}^{m} \left\{\left|I_n^{(r)} - p_r n\right| \le p_r n^{2/3}\right\}$$

and χ_{A_n} be the indicator function of A_n. We also let $p' = \min_{1\le r\le m} p_r$. Then Chernoff's bound yields that

$$\mathbb{P}(A_n^c) = \mathcal{O}\left(\exp\left(\frac{-p'n^{1/3}}{3}\right)\right).$$

Thus, it suffices to show (5.9) on A_n. From the way ϖ_1 was chosen, we have that $\varpi_1(n) = \mathcal{O}(n)$ and $\varpi_1''(n) = \mathcal{O}(n^{-1})$.

Again, we compute the Taylor expansion of ϖ_1 (on A_n):

$$\varpi_1(I_n^{(r)}) = \varpi_1(p_r n) + \varpi_1'(p_r n)\left(I_n^{(r)} - p_r n\right) + \mathcal{O}\left(\frac{(I_n^{(r)} - p_r n)^2}{n}\right)$$

for all $1 \le r \le m$. Consequently,

$$\begin{aligned}
&\left\| (n\varpi_2(n))^{-1/2} \left(\sum_{r=1}^{m} \varpi_1(I_n^{(r)}) - \varpi_1(n) \right) \chi_{A_n} \right\|_s \\
&\qquad \le \left\| (n\varpi_2(n))^{-1/2} \sum_{r=1}^{m} \varpi_1'(p_r n)\left(I_n^{(r)} - p_r n\right) \right\|_s \\
&\qquad\quad + \left\| (n\varpi_2(n))^{-1/2} \sum_{r=1}^{m} \mathcal{O}\left(\frac{(I_n^{(r)} - p_r n)^2}{n}\right) \right\|_s .
\end{aligned} \tag{5.10}$$

We estimate the terms in (5.10) individually. First, we consider

$$\varpi_1'(p_r n) - \varpi_1'(p_s n) = \sum_{\omega_k \in \mathcal{Z}_{<-\alpha_1-\epsilon}} \frac{G_1(\omega_k)}{h} (-\omega_k) n^{-\omega_k-1} \left(p_r^{-\omega_k-1} - p_s^{-\omega_k-1}\right) = o(n).$$

Together with the assumption on $\mathbb{V}(X_n)$, we get

$$\begin{aligned}
&\left\| (n\varpi_2(n))^{-1/2} \sum_{r=1}^{m} \varpi_1'(p_r n)\left(I_n^{(r)} - p_r n\right) \right\|_s \\
&\qquad = \left\| (n\varpi_2(n))^{-1/2} \sum_{r=1}^{m-1} \left(\varpi_1'(p_r n) - \varpi_1'(p_m n)\right)\left(I_n^{(r)} - p_r n\right) \right\|_s \\
&\qquad \le o(1) \sum_{r=1}^{m-1} \left\| \frac{I_n^{(r)} - p_r n}{\sqrt{n}} \right\|_s \\
&\qquad = o\left(\|\mathcal{N}(0,1)\|_s \right) = o(1).
\end{aligned} \tag{5.11}$$

Similarly, we also have

$$\left\| (n\varpi_2(n))^{-1/2} \sum_{r=1}^{m} \mathcal{O}\left(\frac{(I_n^{(r)} - p_r n)^2}{n}\right) \right\|_s = \mathcal{O}\left(\frac{\|\mathcal{N}(0,1)^2\|_s}{\sqrt{n}}\right) = o(1). \tag{5.12}$$

Substituting (5.11) and (5.12) back into (5.9) shows that (b) holds. □

5.2 Framework for Symmetric DSTs

In the previous section, we have established a framework for central limit theorems of shape parameter of m-ary tries satisfying recurrence (5.1). In this section, we are going to establish a similar framework for shape parameters of symmetric DSTs satisfying (5.2). For the sake of simplicity, we will only consider the case $b = 1$ and $m = 2$. However, more general cases can be obtained by similar methods. The structure of both frameworks are quite alike and the proofs are more or less the same. Thus, we will only display the proofs which are different. First, we start with an analogue of Lemma 5.1.3.

Lemma 5.2.1. *Let $\tilde{f}(z)$ and $\tilde{h}(z)$ be entire functions satisfying a differential functional equation of the form*

$$\tilde{f}(z) + \tilde{f}'(z) = 2\tilde{f}\left(\frac{z}{2}\right) + \tilde{h}(z),$$

where $\tilde{f}(0) = 0$ and $\tilde{h}(z) \in \mathscr{J}\mathscr{S}_{\alpha,\gamma}$ with $0 \le \alpha < 1$, then

$$\tilde{f}(z) = \frac{z}{\log 2} \sum_{k\in\mathbb{Z}} \frac{G(2+\chi_k)}{\Gamma(2+\chi_k)} n^{\chi_k} + \mathcal{O}(z^{\alpha+\epsilon}),$$

where $\chi_k = 2k\pi i/\log 2$ and

$$G(\omega) = \int_0^\infty \frac{s^{\omega-1}}{Q(-2s)} \left(\int_0^\infty e^{-sz}\tilde{h}(z)dz\right) ds.$$

Proof. First, we apply Laplace transform to the differential functional equation. This yields

$$(1+s)\mathscr{L}[\tilde{f}; s] = 4\mathscr{L}[\tilde{f}; 2s] + \mathscr{L}[\tilde{h}; s].$$

Because $\tilde{h}(z) \in \mathscr{J}\mathscr{S}_{\alpha,\gamma}$, from the proof of Proposition 2.4 of [74], we get that

$$\tilde{f}(z) = \begin{cases} \mathcal{O}(z^1), & \text{if } z \to 0^+; \\ \mathcal{O}(z^{1+\epsilon}), & \text{if } z \to \infty, \end{cases}$$

where $\epsilon > 0$ can be arbitrarily small. Therefore,

$$\mathscr{L}[\tilde{f}; s] = \begin{cases} \mathcal{O}(s^{-2}), & \text{as } s \to \infty; \\ \mathcal{O}(s^{-2-\epsilon}), & \text{as } s \to 0^+. \end{cases}$$

Now, divide both sides by $Q(-2s)$ and denote $\mathscr{L}[\tilde{f}; s]/Q(-2s)$ by $\bar{\mathscr{L}}[\tilde{f}; s]$. This gives

$$\bar{\mathscr{L}}[\tilde{f}; s] = 4\bar{\mathscr{L}}[\tilde{f}; 2s] + \frac{\mathscr{L}[\tilde{h}; s]}{Q(-2s)}. \tag{5.13}$$

From [74], we have

$$\log Q(-2s) = \frac{(\log s)^2}{\log 2} + \frac{\log s}{2} + \sum_{k\in\mathbb{Z}} q_k s^{-\chi_k} + \mathcal{O}(|s|^{-1}) \tag{5.14}$$

uniformly for $|s| \to \infty$ and $|\arg(s)| \leq \pi - \epsilon$, where $\chi_k = 2k\pi i/\log 2$,

$$q_0 = \frac{\log 2}{12} + \frac{\pi^2}{6\log 2} \quad \text{and} \quad q_k = \frac{1}{2k\sinh(2k\pi/\log 2)} \quad (\text{for } k \neq 0).$$

From (5.14) and the Taylor series expansion

$$Q(-2s) = 1 + \mathcal{O}(|s|), \quad (|s| \to 0),$$

we get

$$\bar{\mathscr{L}}[\tilde{f}; s] = \begin{cases} \mathcal{O}(s^{-M}), & \text{as } s \to \infty \\ \mathcal{O}(s^{-2-\epsilon}), & \text{as } s \to 0^+, \end{cases}$$

where M can be arbitrarily large. Thus, we may apply Mellin transform on both sides of (5.13). Then, we get that for $\Re(\omega) > 2$

$$\mathscr{M}[\bar{\mathscr{L}}[\tilde{f}; s]; \omega] = \frac{G(\omega)}{1 - 2^{2-\omega}},$$

where

$$G(\omega) = \mathscr{M}\left[\frac{\mathscr{L}[\tilde{h}; s]}{Q(-2s)}; \omega\right] = \int_0^\infty \frac{s^{\omega-1}}{Q(-2s)} \left(\int_0^\infty s^{-sz}\tilde{h}(z)dz\right) ds.$$

Since $\tilde{h}(z) \in \mathscr{J}\mathscr{S}_{\alpha,\gamma}$ with $0 \leq \alpha < 1$, $G(\omega)$ is analytic on the half plane $\Re(\omega) > \alpha + 1$. As a result, inverse Mellin transform gives

$$\bar{\mathscr{L}}[\tilde{f}; s] = \frac{s^{-2}}{\log 2} \sum_{k\in\mathbb{Z}} G(2 + \chi_k) s^{-\chi_k} + \mathcal{O}(|s|^{-\alpha-1-\epsilon}).$$

Finally, by Theorem 3.5.2, we get that

$$\tilde{f}_1(z) = \frac{z}{\log 2} \sum_{k\in\mathbb{Z}} \frac{G(2+\chi_k)}{\Gamma(2+\chi_k)} z^{\chi_k} + \mathcal{O}(|z|^{\alpha+\epsilon}).$$

This proves the claimed result. □

As in the previous section, we use the notations

$$\tilde{f}_1(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(X_n)\frac{z^n}{n!}, \quad \tilde{f}_2(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(X_n^2)\frac{z^n}{n!},$$

$$\tilde{\tau}_1(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(T_n)\frac{z^n}{n!}, \quad \tilde{\tau}_2(z) = e^{-z}\sum_{n\geq 0}\mathbb{E}(T_n)\frac{z^n}{n!}.$$

Moreover, again similar to the previous section, we assume that $X_0 = X_1 = 0$ and $n_0 = 2$ to simplify the computation. More general cases can be handled by the same method with slight modifications.

By a similar computation, we get the following differential functional equations

$$\begin{aligned} \tilde{f}_1(z) + \tilde{f}_1'(z) =& 2\tilde{f}_1\left(\frac{z}{2}\right) + \tilde{\tau}_1(z), \\ \tilde{f}_2(z) + \tilde{f}_2'(z) =& 2\tilde{f}_2\left(\frac{z}{2}\right) + 2\tilde{f}_1\left(\frac{z}{2}\right)^2 + \tilde{\tau}_2(z) + \tilde{\lambda}(z), \end{aligned} \tag{5.15}$$

where

$$\begin{aligned} \tilde{\lambda}(z) =& 2e^{-z}\sum_{n\geq 0}\mathbb{E}(T_n)2^{-n}\sum_{0\leq k\leq n}\binom{n}{k}(\mathbb{E}(X_k)+\mathbb{E}(X_{n-k}))\frac{z^n}{n!} \\ =& 2\tilde{\tau}_1(z)\odot\tilde{f}_1(z) + 2\tilde{\tau}_1(z)\odot\tilde{f}_1'(z) - 2\tilde{\tau}_1(z)\odot\tilde{\tau}_1(z). \end{aligned}$$

Moreover, we again use the Poissonized variance with correction and obtain that

$$\begin{aligned} \tilde{V}(z) + \tilde{V}'(z) =& 2\tilde{V}\left(\frac{z}{2}\right) + \tilde{V}_T(z) + \tilde{\lambda}(z) - 4\tilde{\tau}_1(z)\tilde{f}_1\left(\frac{z}{2}\right) \\ & - 2z\tilde{\tau}_1'(z)\tilde{f}_1'\left(\frac{z}{2}\right) + z\tilde{f}_1''(z)^2 \end{aligned} \tag{5.16}$$

where $\tilde{V}_T(z) = \tilde{\tau}_2(z) - \tilde{\tau}_1(z)^2 - z\tilde{\tau}_1'(z)^2$.

We apply Theorem 5.1.5 to (5.16). By similar arguments as in the proof of Proposition 5.1.6, we get the following analogue.

Proposition 5.2.2. *Let all the functions be defined as above. If $\tilde{\tau}_1(z) \in \mathscr{J}\mathscr{S}_{\alpha_1,\gamma_1}$ with $0 \leq \alpha_1 < 1$, then*

$$\mathbb{E}(X_n) = \frac{n}{\log 2}\sum_{k\in\mathbb{Z}}\frac{G_E(2+\chi_k)}{\Gamma(2+\chi_k)}n^{\chi_k} + \mathcal{O}(n^{\alpha+\epsilon}),$$

where the sum expression is infinitely differentiable and

$$G_E(\omega) = \int_0^\infty \frac{1}{Q(-2s)}\left(\int_0^\infty e^{-sz}\tilde{\tau}_1(z)dz\right)ds.$$

In addition, if $\tilde{V}_T(z) \in \mathscr{J}\mathscr{S}_{\alpha_2,\gamma_2}$ with $0 \le \alpha_2 < 1$ and $\tilde{\tau}_2(z) \in \mathscr{J}\mathscr{S}$, then

$$\mathbb{V}(X_n) \sim \frac{n}{\log 2} \sum_{k\in\mathbb{Z}} \frac{G_V(2+\chi_k)}{\Gamma(2+\chi_k)} n^{\chi_k},$$

where the sum expression is infinitely differentiable and

$$G_V(\omega) = \int_0^\infty \frac{1}{Q(-2s)} \left(\int_0^\infty e^{-sz} \tilde{R}(z) dz \right) ds$$

with

$$\tilde{R}(z) = \tilde{V}_T(z) + \tilde{\lambda}(z) - 4\tilde{\tau}_1(z)\tilde{f}_1\left(\frac{z}{2}\right) - 2z\tilde{\tau}_1'(z)\tilde{f}_1'\left(\frac{z}{2}\right) + z\tilde{f}_1''(z)^2.$$

Finally, we give the general central limit theorem of shape parameters for DSTs. The proof of the following Theorem is similar to Theorem 5.1.7 (even easier since DSTs are binary and we only consider the symmetric case here) and hence skipped.

Theorem 5.2.3. *Suppose $\tilde{\tau}_1(z) \in \mathscr{J}\mathscr{S}_{\alpha_1,\gamma_1}$ with $0 \le \alpha_1 < 1/2$, $\tilde{\tau}_2(z) \in \mathscr{J}\mathscr{S}$ and $\tilde{V}_T(z) \in \mathscr{J}\mathscr{S}_{\alpha_2,\gamma_2}$ with $0 \le \alpha_2 < 1$. Moreover, we assume that $\|T_n\|_s = o(\sqrt{n})$ with $2 < s \le 3$ and $\mathbb{V}(X_n) \ge cn$ for all n large enough and some $c > 0$. Then, as $n \to \infty$*

$$\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \xrightarrow{d} \mathcal{N}(0,1).$$

5.3 Lower Bounds for the Variance

Note that in the statement of Theorem 5.1.7 and Theorem 5.2.3, we require that $\mathbb{V}(X_n) = \Omega(n)$. Here we introduce a useful lemma which helps us to establish lower bounds for recurrences of some specific form. This lemma can be used to check the assumption $\mathbb{V}(X_n) = \Omega(n)$ in Theorem 5.1.7 and Theorem 5.2.3. The proof of the following lemma is largely based on ideas of Schachinger in [189].

Lemma 5.3.1. *Consider two nonnegative sequence $\{\alpha_i\}$ and $\{\beta_i\}$ satisfying a recurrence of the form*

$$\alpha_{n+1} = \sum_{i=1}^{m} a_i \sum_{j=0}^{n} f(n,j,p_i)\alpha_j + \beta_n, \quad (n \ge n_0),$$

where $a_1, \ldots, a_m$ are positive real numbers, $p_i \in (0,1)$ for all $1 \le i \le m$ and $f(n,j,p)$ is a nonnegative-valued function. We assume that there exists some $j' \ge n_0$ such that $\beta_{j'} > 0$. We also assume that $f(n,j,p)$ satisfies that $\sum_{j=0}^{n} f(n,j,p) = 1$ and there exists $n_1 \ge n_0$ such that for all $n > n_1$ and $p < 1$,

$$\sum_{|j-pn|>pn^{\tau}} f(n,j,p) = \mathcal{O}(n^{\tau-1})$$

for some constant $1 > \tau > 0$, then $\alpha_n = \Omega(n^{\lambda})$ with λ being the unique real root of $F(z) = 1 - \sum_{i=1}^{m} a_i p_i^z$.

Proof. Let $C_n = \alpha_n / n^{\lambda}$. We may rewrite the recurrence as

$$C_{n+1} = \sum_{i=1}^{m} a_i \sum_{j=0}^{n} f(n,j,p_i) \left(\frac{j}{n+1}\right)^{\lambda} C_j + \frac{\beta_n}{(n+1)^{\lambda}}.$$

We set $\underline{C}_n = \min_{j'+1 \le \hat{n} \le n} C_{\hat{n}}$ and $\mathscr{N} = \{n \in \mathbb{N} | \underline{C}_n < \underline{C}_{n-1}\}$.

If $|\mathscr{N}| < \infty$, we get the desired result. Otherwise, we let $n' = n + n^{\tau}$ and $n'' = n - \operatorname{sgn}(\lambda) n^{\tau}$. For all $1 \le i \le m$, we can find $n_2 \ge n_1$ such that for all $n \ge n_2$

$$\begin{aligned}
\sum_{j=0}^{n} f(n,j,p_i) \left(\frac{j}{n+1}\right)^{\lambda} C_j &\ge \sum_{|j-p_i n| \le p_i n^{\tau}} f(n,j,p_i) \left(\frac{j}{n+1}\right)^{\lambda} C_j \\
&\ge \sum_{|j-p_i n| \le p_i n^{\tau}} f(n,j,p_i) \left(\frac{p_i n''}{n+1}\right)^{\lambda} \underline{C}_{\lceil p_i n' \rceil} \\
&\ge \underline{C}_{\lceil p_i n' \rceil} \left(\frac{p_i n''}{n+1}\right)^{\lambda} \left(1 - \sum_{|j-p_i n| > p_i n^{\tau}} f(n,j,p_i)\right) \\
&\ge \underline{C}_{\lceil p_i n' \rceil} \left(\frac{p_i n''}{n+1}\right)^{\lambda} \left(1 - F' n^{\tau-1}\right)
\end{aligned}$$

for some constant F'. We choose $p' = \max_{1 \le i \le m} p_i$ and $p'' = \frac{1+p'}{2}$, then there exists $n_3 \ge n_2$ and a constant c such that $\lceil p_i n' \rceil \le \lfloor p'' n \rfloor$ for all $n \ge n_3$ and

$$\left(\frac{p_i n''}{n+1}\right)^{\lambda} \ge p_i^{\lambda} \left(1 - |c| n^{\tau-1}\right) \quad \text{for all } 1 \le i \le m.$$

Let $F = |c| + F'$. Then, we get that

$$\sum_{j=0}^{n} f(n,j,p_i) \left(\frac{j}{n+1}\right)^{\lambda} C_j \ge \underline{C}_{\lfloor p'' n \rfloor} p_i^{\lambda} (1 - F n^{\tau-1})$$

and hence $C_{n+1} \geq \underline{C}_{\lfloor p''n \rfloor}\left(1 - Fn^{\tau-1}\right)$.

Now we construct an increasing sequence $\{N_i\}_{i\geq 0}$ by letting $N_0 = n_3$ and

$$N_{i+1} = \max\{(n+1) \in \mathbb{N} | \underline{C}_n > C_{n+1}, \underline{C}_{\lfloor p''n \rfloor} \geq C_{N_i}\}.$$

Note that $N_{i+2} > \frac{N_i}{p''}$. As a result, we get that $\prod_{j\geq 0}\left(1 - FN_j^{\tau-1}\right)$ is convergent and hence we can find j_0 big enough such that

$$\prod_{j\geq j_0}\left(1 - FN_j^{\tau-1}\right) \geq \frac{1}{2}.$$

Finally,

$$\underline{C}_{N_m} \geq \underline{C}_{j_0} \prod_{j=j_0+1}^{m}\left(1 - FN_j^{\tau-1}\right) \geq \frac{1}{2}C_{j_0}.$$

This implies that $C_n = \Omega(1)$ and the result follows. □

Now, we explain how to use this lemma. First, we let $\mu_n = \mathbb{E}(X_n)$ and

$$M_n(y) = \mathbb{E}\left(e^{(X_n-\mu_n)y}\right). \tag{5.17}$$

Note that $M_n''(0) = \mathbb{E}\left((X_n - \mu_n)^2\right) = \mathbb{V}(X_n)$.

For the trie case, we substitute (5.1) into (5.17), then for $n \geq n_0$

$$\begin{aligned} M_n(y) = \sum_{j_1+\cdots+j_m=n} & \pi_{j_1,\ldots,j_m} M_{j_1}(y)\cdots M_{j_m}(y) \\ & \mathbb{E}\left(e^{(T_n-\mu_n+\sum_i \mu_{j_i})y} | I_n^{(1)} = j_1, \ldots, I_n^{(m)} = j_m\right), \end{aligned}$$

where

$$\pi_{j_1,\ldots,j_m} = \binom{n}{j_1,\cdots,j_m} p_1^{j_1}\cdots p_m^{j_m}.$$

Let $\sigma_n^2 = M_n''(0)$ and differentiate the equation, we get

$$\sigma_n^2 = \sum_{i=1}^{m}\sum_{k=0}^{n-1} \frac{\binom{n}{k}p_i^k(1-p_i)^{n-k}}{1 - \sum_{r=1}^{m} p_r^n}\sigma_k^2 + \eta_n, \tag{5.18}$$

where

$$\eta_n = \sum_{j_1+\cdots+j_m=n} \pi_{j_1,\ldots,j_m}\mathbb{E}\left((T_n - \mu_n + \sum_i \mu_{j_i})^2 | I_n^{(1)} = j_1, \ldots, I_n^{(m)} = j_m\right).$$

Choose

$$f(n,j,p) = \frac{\binom{n}{j} p^j (1-p)^{n-j}}{1 - \sum_{r=1}^{m} p_r^n}.$$

Then, by Lemma 5.3.1, $\mathbb{V}(X_n) = \Omega(n)$ if $\eta_{n'} > 0$ for some $n' \geq n_0$.

For symmetric DSTs, we substitute (5.2) into (5.17). By similar computations, we get that for $n \geq n_0$

$$\sigma_n^2 = 2\sum_{k=0}^{n-1} 2^{1-n}\binom{n-1}{k}\sigma_k^2 + \vartheta_n, \tag{5.19}$$

where

$$\vartheta_n = \sum_{k=0}^{n-1} 2^{1-n}\binom{n-1}{k}\mathbb{E}\left((T_{n-1} + \mu_k + \mu_{n-k-1} - \mu_n)^2 | I_{n-1} = k\right). \tag{5.20}$$

To apply Lemma 5.3.1, we choose $f(n,j,p) = 2^{1-n}\binom{n-1}{j}$ and check whether $\vartheta_{n'} > 0$ for some $n' \geq n_0$.

5.4 Internal Nodes of m-ary Tries with Specified Outdegree

As an application of the framework introduced in Section 5.1, we will consider the number of internal nodes of outdegree k in a random trie of size n which will be denoted by $N_n^{(k)}$. (This is a refinement of the size of tries and PATRICIA tries; see Corollary 5.4.4 below.) We will give a multivariate study of these parameters by considering

$$Z_n = \sum_{k=1}^{m} a_k N_n^{(k)},$$

where $a_1, \dots, a_m$ are arbitrary real number with $a_i \neq (i-1)a_2$ for some i (this is to make sure that Z_n is not deterministic; see Lemma 5.4.1 and the remark succeeding it). Note that a similar multivariate framework was considered in Hubalek et al. [90] for shape parameters in digital search trees. However, our analysis will take into account many tools developed after [90]. Before we state the main result about Z_n, we need some preparation.

Lemma 5.4.1. *Z_n is not deterministic for n large enough.*

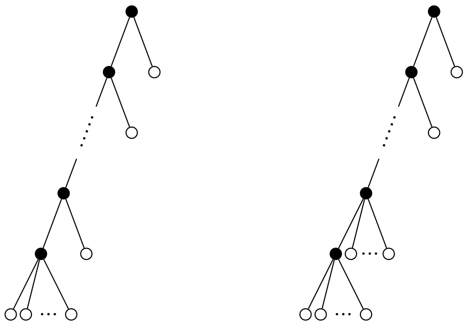

Figure 5.1: Two tries with internal nodes black and external nodes white. The trie on the left has all internal nodes of outdegree 2 except the last which is of outdegree i; the trie on the right has all internal nodes of outdegree 2 expect the last two which are of outdegree i.

Proof. First, observe that the claim is trivial if $a_1 \neq 0$. Thus, we may assume that $m \geq 3$ and $a_i \neq (i-1)a_2$ for $i \geq 3$. For this case, consider the two tries from Figure 5.1. For the first trie, we have $Z_n = (n-i)a_2 + a_i$; for the second, we have $Z_n = (n-2i+1)a_2 + 2a_i$. From our assumption, these two values are different. This concludes the proof. □

Remark 24. If $a_i = (i-1)a_2$ for all i, then it is easy to see that $Z_n = a_2(n-1)$ for all n.

Proposition 5.4.2. *We have,* $\mathrm{Var}(Z_n) \geq cn$ *with* $c > 0$ *for all* n *large enough.*

Proof. First, observe that Z_n is an additive shape parameter satisfying a recurrence of type (5.1). In order to see this, note that

$$N_n^{(k)} \overset{d}{=} \sum_{i=1}^{m} (N_{I_n^{(i)}}^{(k)})^{(i)} + T_n^{(k)}, \qquad (n \geq 2),$$

with the initial conditions $N_0^{(k)} = N_1^{(k)} = 0$ and

$$T_n^{(k)} = \begin{cases} 1, & \text{if } \#\{1 \leq i \leq m \ : \ I_n^{(i)} \neq 0\} = k; \\ 0, & \text{otherwise.} \end{cases}$$

Consequently,

$$Z_n \overset{d}{=} \sum_{i=1}^{m} Z_{I_n^{(i)}}^{(i)} + T_n, \qquad (n \geq 2), \tag{5.21}$$

where $Z_0 = Z_1 = 0$ and

$$T_n = \sum_{k=1}^{m} a_k T_n^{(k)}. \tag{5.22}$$

Now, to apply Lemma 5.3.1, we derive the recurrence for $\mathrm{Var}(Z_n)$. Set $\mu_n = \mathbb{E}(Z_n)$ and

$$M_n(y) = \mathbb{E}\left(e^{(Z_n-\mu_n)y}\right).$$

Then, from (5.21), similar computation as in Section 5.3 yields that

$$\mathrm{Var}(Z_n) = \sigma_n^2 = \sum_{i=1}^{m}\sum_{j=0}^{n}\binom{n}{j}p_i^j(1-p_i)^{n-j}\sigma_j^2 + \eta_n, \qquad (n \geq 2),$$

where $\sigma_0^2 = \sigma_1^2 = 0$ and

$$\eta_n = \sum_{j_1+\cdots+j_m=n} \pi_{j_1,\ldots,j_m}\mathbb{E}\Big(\big(T_n - \mu_n + \sum_i \mu_{j_i}\big)^2 | I_n^{(1)} = j_1, \ldots, I_n^{(m)} = j_m\Big).$$

From the last expression, we see that $\eta_n \geq 0$. Consequently, by Lemma 5.3.1, either $\mathrm{Var}(X_n)$ grows at least linearly or equals zero for all n. The latter, however, is impossible by Lemma 5.4.1. □

With the preparation work done, we now state the main result about Z_n.

Theorem 5.4.3. *We have, as $n \to \infty$,*

$$\mathbb{E}(Z_n) \sim nP(\log_{1/a} n), \qquad \mathrm{Var}(Z_n) \sim nQ(\log_{1/a} n),$$

where $a > 0$ is a suitable constant and $P(z), Q(z)$ are infinitely differentiable, 1-periodic functions (possibly constant). Moreover, $\mathrm{Var}(Z_n) > 0$ *for all n large enough and*

$$\frac{Z_n - \mathbb{E}(Z_n)}{\sqrt{\mathrm{Var}(Z_n)}} \xrightarrow{d} N(0,1).$$

Proof. We are going to use the results of Proposition 5.1.6 and Theorem 5.1.7 to prove this theorem. Note, however, that T_n is not independent of $(I_n^{(1)}, \ldots, I_n^{(m)})$ and hence, strictly speaking, the two propositions do not apply. However, it is easily checked that the proofs of the propositions still work for the current situation under the same assumptions; see Section 5.4 in [75] for a similar example.

Now, we will check that the assumptions of the propositions hold. This is not complicated since $\tilde{h}_1(z)$ and $\tilde{h}_2(z)$ are easily computed. For instance, to compute $\tilde{h}_1(z)$, note that

$$\mathbb{E}\left(T_n^{(k)}\right) = \sum_{\{i_1,\ldots,i_k\}\subseteq S}\ \sum_{\substack{j_{i_1}+\cdots+j_{i_k}=n \\ j_{i_1},\ldots,j_{i_k}\geq 1}} \binom{n}{j_{i_1},\ldots,j_{i_k}} p_{i_1}^{j_{i_1}}\cdots p_{i_k}^{j_{i_k}}.$$

Consequently, for $k \geq 2$,

$$e^{-z}\sum_{n\geq 2}\mathbb{E}\left(T_n^{(k)}\right)\frac{z^n}{n!} = \sum_{\{i_1,\dots,i_k\}\subseteq S} e^{-z}\left(e^{p_{j_{i_1}}z}-1\right)\cdots\left(e^{p_{j_{i_k}}z}-1\right)$$

and similar for $k = 1$. From this, we obtain $\tilde{h}_1(z)$ by (5.22) and linearity of the mean.

In particular, we see that $\tilde{h}_1(z)$ is a linear combination of functions of the form e^{-az} with $a \geq 0$. Hence, from the closure properties from Section 3.5.3, we have that $\tilde{h}_1(z) \in \mathscr{JS}_{0,0}$. The same result is also easily verified to hold for $\tilde{h}_2(z)$. Thus, the claims about mean and variance in Theorem 5.4.3 follow from Proposition 5.1.6.

Next, we turn to the limit law. We are going to apply Theorem 5.1.7. The only assumption of this theorem which needs further explanation is the assumption on the positiveness of the variance (or more precisely, the assumption of the at least linear growth of the variance). By Proposition 5.4.2, this assumption is verified and the result follows. □

As a consequence, we consider the size of tries and PATRICIA tries

$$N_n^{(T)} = \sum_{k=1}^{m} N_n^{(k)}, \qquad N_n^{(P)} = \sum_{k=2}^{m} N_n^{(k)}.$$

Note that $N_n^{(P)}$ equals $n-1$ if $m = 2$ and this case was excluded from our definition of Z_n. We have the following consequence of Theorem 5.4.3.

Corollary 5.4.4. *For $m \geq 2$, as $n \to \infty$,*

$$\frac{N_n^{(T)} - \mathbb{E}(N_n^{(T)})}{\sqrt{\operatorname{Var}(N_n^{(T)})}} \xrightarrow{d} N(0,1)$$

and for $m \geq 3$, as $n \to \infty$,

$$\frac{N_n^{(P)} - \mathbb{E}(N_n^{(P)})}{\sqrt{\operatorname{Var}(N_n^{(P)})}} \xrightarrow{d} N(0,1).$$

The result for the size of tries with $m = 2$ is classical; see [93] for an analytic proof and Neininger and Rüschendorf [160] for a proof using the contraction method.

Note that the covariance of $N_n^{(k_1)}$ and $N_n^{(k_2)}$ can be obtained from Theorem 5.4.3 via the relation

$$2\mathrm{Cov}\left(N_n^{(k_1)}, N_n^{(k_2)}\right) = \operatorname{Var}\left(N_n^{(k_1)} + N_n^{(k_2)}\right) - \operatorname{Var}\left(N_n^{(k_1)}\right) - \operatorname{Var}\left(N_n^{(k_2)}\right).$$

By this relation and Theorem 5.4.3, we obtain

$$\mathrm{Cov}(N_n^{(k_1)}, N_n^{(k_2)}) \sim nQ^{(k_1,k_2)}(\log_{1/a} n) \tag{5.23}$$

for all $1 \leq k_1, k_2 \leq m$, where $Q^{(k_1,k_2)}(z)$ is an infinitely differentiable, 1-periodic function (possibly constant). Set

$$\mathrm{Var}(N_n^{(k_1)}) \sim nQ^{(k_1)}(\log_{1/a} n), \quad \mathrm{Var}(N_n^{(k_2)}) \sim nQ^{(k_2)}(\log_{1/a} n)$$

and

$$\Sigma_n = \begin{pmatrix} nQ^{(k_1)}(\log_{1/a} n) & nQ^{(k_1,k_2)}(\log_{1/a} n) \\ nQ^{(k_1,k_2)}(\log_{1/a} n) & nQ^{(k_2)}(\log_{1/a} n) \end{pmatrix}.$$

Then, we are going to show a bivariate limit law

$$\Sigma_n^{-1/2} \begin{pmatrix} N_n^{(k_1)} - \mathbb{E}(N_n^{(k_1)}) \\ N_n^{(k_2)} - \mathbb{E}(N_n^{(k_2)}) \end{pmatrix} \xrightarrow{d} N(0, I_2).$$

Similar to Theorem 5.4.3, we need some preparation before proving the limit law.

First, we have to show for normalization purposes that Σ_n is positive definite. We will do this again in two steps. From now on, we will assume that $(k_1, k_2, m) \notin \{(1,2,2),(2,3,3)\}$. (Otherwise, Σ_n is not positive definite, see Remark 25 below.)

Lemma 5.4.5. *The correlation coefficient $\rho(N_n^{(k_1)}, N_n^{(k_2)})$ is not -1 or 1 for all n large enough.*

Proof. We use proof by contradiction. Thus, assume that $\rho(N_n^{(k_1)}, N_n^{(k_2)}) \in \{-1, 1\}$ which implies that for some $a_n, b_n \in \mathbb{R}$ with $a_n \neq 0$, we have that

$$N_n^{(k_1)} = a_n N_n^{(k_2)} + b_n.$$

Obviously this cannot hold if $k_1 = 1$. Thus, we may assume that $k_1 \geq 2$. First, consider $k_1 = 2$ and set $i \neq k_2$ (this is possible due to the assumption on (k_1, k_2, m)). Then, we get a contradiction from the two tries in Figure 1 (since $N_n^{(2)}$ decreases, whereas $N_n^{(k_2)}$ remains constant). Next, consider $k_1 > 2$ and set $i = k_2$. Then, again a contradiction is obtained from Figure 1 (now, $N_n^{(k_1)}$ remains constant, whereas $N_n^{(i)}$ increases). □

Remark 25. $(k_1, k_2, m) = (1, 2, 2)$ is the only case where the correlation coefficient is not defined ($N_n^{(2)}$ is deterministic in this case; see Remark 24). If $(k_1, k_2, m) = (2, 3, 3)$, then $N_n^{(2)} = n - 1 - 2N_n^{(3)}$ (again by Remark 24). Hence, in this case $\rho(N_n^{(2)}, N_n^{(3)}) = -1$.

Proposition 5.4.6. *Σ_n is positive definite for all n large enough.*

Proof. It is sufficient to show that $\det(\Sigma_n) > 0$ for all n large enough. For the proof of this, we will need some notation. First,

$$\mu_n^{(k_1)} = \mathbb{E}(N_n^{(k_1)}), \qquad \mu_n^{(k_2)} = \mathbb{E}(N_n^{(k_2)}).$$

Moreover,

$$\xi_n = \mathrm{Var}(N_n^{(k_1)}), \qquad \nu_n = \mathrm{Cov}(N_n^{(k_1)}, N_n^{(k_2)}), \qquad \kappa_n = \mathrm{Var}(N_n^{(k_2)}).$$

Then, by setting

$$F_n(u,v) = \mathbb{E}\left(e^{(N_n^{(k_1)} - \mu_n^{(k_1)})u + (N_n^{(k_2)} - \mu_n^{(k_2)})v}\right)$$

and arguing as in Section 5.3, we obtain (after a lengthy computation)

$$\begin{aligned}
&\xi_{n_1}\kappa_{n_2} + \xi_{n_2}\kappa_{n_1} - 2\nu_{n_1}\nu_{n_2} \\
&= \sum_{j_1+\cdots+j_m=n_1} \sum_{l_1+\cdots+l_m=n_2} \pi_{j_1,\ldots,j_m}\pi_{l_1,\ldots,l_m} \sum_{i=1}^{m}\sum_{u=1}^{m} (\xi_{j_i}\kappa_{l_u} + \xi_{l_u}\kappa_{j_i} - 2\nu_{j_i}\nu_{l_u}) \\
&\quad + \tau_{n_1,n_2} \qquad (5.24)
\end{aligned}$$

for $n_1, n_2 \geq 2$ and all initial conditions equal to 0. In order to describe τ_{n_1,n_2} set

$$\alpha_{j_1,\ldots,j_m} = \mathbb{E}\Big(\big(T_n^{(k_1)} - \mu_n^{(k_1)} + \sum_i \mu_{j_i}^{(k_1)}\big)^2 | I_n^{(1)} = j_1, \ldots, I_n^{(m)} = j_m\Big),$$

$$\beta_{j_1,\ldots,j_m} = \mathbb{E}\Big(\big(T_n^{(k_2)} - \mu_n^{(k_2)} + \sum_i \mu_{j_i}^{(k_2)}\big)^2 | I_n^{(1)} = j_1, \ldots, I_n^{(m)} = j_m\Big).$$

Then,

$$\begin{aligned}
&\tau_{n_1,n_2} \\
&= \sum_{j_1+\cdots+j_m=n_1} \sum_{l_1+\cdots+l_m=n_2} \pi_{j_1,\ldots,j_m}\pi_{l_1,\ldots,l_m} \left(\Theta_{j_1,\ldots,j_m,l_1,\ldots,l_m} + \Xi_{j_1,\ldots,j_m,l_1,\ldots,l_m}\right),
\end{aligned}$$

where

$$\Theta_{j_1,\ldots,j_m,l_1,\ldots,l_m} = (\alpha_{j_1,\ldots,j_m}\beta_{l_1,\ldots,l_m} - \alpha_{l_1,\ldots,l_m}\beta_{j_1,\ldots,j_m})^2$$

and

$$\begin{aligned}
\Xi_{j_1,\ldots,j_m,l_1,\ldots,l_m} &= \sum_{i=1}^{m} \mathbb{E}\left(\alpha_{l_1,\ldots,l_m}(N_{j_i}^{(k_1)} - \mu_{j_i}^{(k_1)}) - \beta_{l_1,\ldots,l_m}(N_{j_i}^{(k_1)} - \mu_{j_i}^{(k_2)})\right)^2 \\
&\quad + \sum_{u=1}^{m} \mathbb{E}\left(\alpha_{j_1,\ldots,j_m}(N_{l_u}^{(k_2)} - \mu_{j_u}^{(k_2)}) - \beta_{j_1,\ldots,j_m}(N_{l_u}^{(k_1)} - \mu_{l_u}^{(k_1)})\right)^2.
\end{aligned}$$

Now, note that $\tau_{n_1,n_2} \geq 0$ for all n_1, n_2. Using a similar argument as in the proof of Lemma 5.3.1 for (5.24) twice, one obtains that

$$\xi_{n_1}\kappa_{n_2} + \xi_{n_2}\kappa_{n_1} - 2\nu_{n_1}\nu_{n_2}$$

is either identical zero for all n_1, n_2 or $\geq cn_1n_2$ with $c > 0$. The former is however impossible due to Lemma 5.4.5. Finally, by setting $n_1 = n_2$, we obtain that $\det(\Sigma_n) \geq cn^2$ with $c > 0$. □

As a consequence of Proposition 5.4.6, $\Sigma_n^{1/2}$ exists for n large enough. For the proof of the bivariate limit law, we need some notation

$$\begin{pmatrix} b_n^{(1)} \\ b_n^{(2)} \end{pmatrix} = \Sigma_n^{-1/2}\left(\begin{pmatrix} T_n^{(k_1)} \\ T_n^{(k_2)} \end{pmatrix} - \begin{pmatrix} \mu_n^{(k_1)} \\ \mu_n^{(k_2)} \end{pmatrix} + \sum_{i=1}^{k} \begin{pmatrix} \mu_{I_n^{(i)}}^{(k_1)} \\ \mu_{I_n^{(i)}}^{(k_2)} \end{pmatrix}\right),$$

where $\mu_n^{(k_1)}$ and $\mu_n^{(k_2)}$ are as in the proof of the above proposition and

$$A_n^{(i)} = \Sigma_n^{-1/2}\Sigma_{I_n^{(i)}}^{1/2}, \qquad 1 \leq i \leq m.$$

Explicit expressions for these vectors and matrices can be derived by Maple and are given in Appendix B (which the reader should consult before reading the proof of Theorem 5.4.7.)

Theorem 5.4.7. *Assume that $(k_1, k_2, m) \notin \{(1, 2, 2), (2, 3, 3)\}$. Then, Σ_n is positive definite for n large enough and, as $n \to \infty$,*

$$\Sigma_n^{-1/2}\begin{pmatrix} N_n^{(k_1)} - \mathbb{E}(N_n^{(k_1)}) \\ N_n^{(k_2)} - \mathbb{E}(N_n^{(k_2)}) \end{pmatrix} \xrightarrow{d} N(0, I_2),$$

where I_2 denotes the 2×2 identity matrix.

Proof. We use the multivariate version of the contraction method; see Neininger and Rüschendorf [160]. We have to verify the following assumptions for $2 < s \leq 3$:

$$\begin{pmatrix} b_n^{(1)} \\ b_n^{(2)} \end{pmatrix} \xrightarrow{\mathcal{L}_s} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad A_n^{(i)} \xrightarrow{\mathcal{L}_s} A_i, \tag{5.25}$$

$$\mathbb{E}\sum_{i=1}^{m} \|A_i\|_{\text{op}}^s < 1, \qquad \mathbb{E}\left(\|A_i\|_{\text{op}}^s \chi_{\{I_n^{(i)} \leq j\} \cup \{I_n^{(i)} = n\}}\right) \to 0 \tag{5.26}$$

for all $1 \leq i \leq m$ and $j \in \mathbb{N}$, where $\|\cdot\|_{\text{op}}$ denotes the operator norm of a matrix.

First, from the proof of Theorem 5.4.3, we have

$$\frac{T_n^{(\star)} - \mu_n^{(\star)} + \sum_{i=1}^{m} \mu_i^{(\star)}}{\sqrt{n}} \xrightarrow{L_s} 0$$

for $\star \in \{k_1, k_2\}$. This together with the boundedness of $\Omega_1(n)$and $\Omega_2(n)$ (from Proposition 5.4.2) and $D(n)$ (from the proof of Proposition 5.4.6) shows the claimed result for $b_n^{(1)}$ and $b_n^{(2)}$ in (5.25).

Next, to show the second claim (5.26), we argue as in the proof of condition (a) in Theorem 5.1.7. For instance, for the expressions in $A_n^{(i)}(1,1)$, we obtain

$$\frac{\Omega_1(I_n^{(i)}) + \Omega_2(I_n^{(i)}) + 2\sqrt{D(I_n^{(i)})}}{\Omega_1(n) + \Omega_2(n) + 2\sqrt{D(n)}} \xrightarrow{\text{a.s.}} 1$$

and

$$\frac{\left(\Omega_1(I_n^{(i)}) + \sqrt{D(I_n^{(i)})}\right)\left(\Omega_2(n) + \sqrt{D(n)}\right) - \Omega_3(n)\Omega_3(I_n^{(i)})}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}} \xrightarrow{\text{a.s.}} 1$$

which are proved similar as (5.8). Plugging this into the expression for $A_n^{(i)}(1,1)$ and using (5.7) then gives

$$A_n^{(i)}(1,1) \xrightarrow{\text{a.s.}} \sqrt{p_i}.$$

By dominated convergence, the same holds in L_s. Similarly, the other entries of $A_n^{(i)}$ are treated. Overall, we obtain

$$A_n^{(i)} \xrightarrow{\mathcal{L}_s} \sqrt{p_i} I_2 \tag{5.27}$$

which shows the second claim in (5.25).

From (5.27) it follows immediately that $\|A_i\|_{\text{op}} = \sqrt{p_i}$. Using this, the two conditions in (5.26) are easily checked.

Finally, by applying the multivariate contraction method, we obtain convergence in distribution of the random vector from Theorem 5.4.7 to a random variable whose distribution is the unique solution of a distributional fixed point equation (see [188]). It is easily verified that the only solution of this fixed point equation is the (2-dimensional) standard normal distribution. This completes the proof of Theorem 5.4.7. □

A similar result could be also given for $(N_n^{(k_1)}, N_n^{(k_2)}, N_n^{(k_3)})$, however, proving that the corresponding covariance matrix is positive definite would be technically complicated (and the problem becomes even more intractable when considering stochastic vectors of higher dimension).

5.5 2-Protected Nodes in Symmetric Digital Search Trees

5.5.1 Introduction

A node in a tree is said to be k-protected if the distance from the node to each leaf is at least k. For example, every node which is not a leaf is 1-protected and 2-protected nodes are those nodes which are neither a leaf nor a parent of a leaf.

2-protected nodes have drawn some attention in recent years due to some practical relevance. For instance, in a network model with the hackers represented by leaves, the distance between the computers (represented by internal nodes) and hackers would be an important measure of how ”vulnerable” a computer is in the network. This is why the parameter is named ”protected”. The parameter also found real life applications in social networks, computer security models and so on; see [81] for more details.

The concept of protected nodes was first proposed by Cheon and Shapiro in [19]. They considered the number of 2-protected nodes in some families of ordered trees, including planar trees, Motzkin trees and ternary trees, and showed that the portion of 2-protected nodes in the trees they considered will converge to some constants. Since then, 2-protected nodes have been considered for various tree models. In [144], the author computed the portion of 2-protected nodes in k-ary trees. Mahmoud and Ward showed that the number of 2-protected nodes in binary search trees satisfies a central limit theorem in [142]. Bóna also considered the probabilistic properties of k-protected nodes in binary search trees in [12]. The authors of [80] and [81] derived the asymptotic expression of the mean and variance of 2-protected nodes in tries. Du and Prodinger computed an asymptotic expression of the mean of 2-protected nodes in DSTs [52]. Recently, Devroye and Janson proposed a parameter called protected fringe subtree which generalizes k-protected nodes and studied it for simply generated trees, binary search trees and random recursive trees [40].

In this section, we are going to rederive the asymptotic expression of the mean for the number of 2-protected nodes in DSTs. Moreover, we also obtain an asymptotic expression for the variance and prove a central limit theorem by the framework introduced in Section 5.2.

5.5.2 Limit Laws

Let L_n be the random variable which counts the number of 2-protected nodes in a DST constructed by n strings. Then, we have the following recurrence under the Bernoulli model:

$$L_{n+1} \overset{d}{=} L_{B_n} + L^*_{n-B_n} + T_n, \quad (n \geq 3),$$

where $B_n = \text{Binomial}(n, 1/2)$ and

$$T_n = \begin{cases} 0, & B_n = 1 \vee n-1; \\ 1, & \text{Otherwise.} \end{cases} \tag{5.28}$$

The initial conditions are $L_0 = L_1 = L_2 = 0$ and $L_3 = 1/2$.

Theorem 5.5.1. *We have that*

$$\mathbb{E}(L_n) = \frac{n}{\log 2} \sum_{k \in \mathbb{Z}} \frac{G_E(2+\chi_k)}{\Gamma(2+\chi_k)} n^{\chi_k} + \mathcal{O}(n^{\epsilon}) \quad \textit{as } n \to \infty,$$

and

$$\mathbb{V}(L_n) \sim \frac{n}{\log 2} \sum_{k \in \mathbb{Z}} \frac{G_V(2+\chi_k)}{\Gamma(2+\chi_k)} n^{\chi_k} \qquad \textit{as } n \to \infty.$$

where $G_E(\omega)$ and $G_V(\omega)$ are infinitely differentiable, 1-periodic functions (possibly constant).

Proof. To apply Proposition 5.2.2, we need to compute $\tilde{\tau}_1(z)$, $\tilde{\tau}_2(z)$ and $\tilde{V}_T(z)$. (Note that again the independence assumption of B_n and T_n is not certified. However, the proof of Proposition 5.2.2 still holds for the current assumption.) From (5.28), we get that

$$\tilde{\tau}_1(z) = \tilde{\tau}_2(z) = 1 - e^{-z} - ze^{-z/2}.$$

Thus,

$$\tilde{V}_T(z) = \left(1 - z - \frac{z^3}{4}\right) e^{-z} - (1+z)e^{-2z} + z^{-z/2} - z^2 e^{-3z/2}.$$

We can easily verify that $\tilde{\tau}_1(z)$, $\tilde{\tau}_2(z)$ as well as $\tilde{V}_T(z)$ fulfill the requirements. □

Next, we derive an explicit expression of the periodic function in the asymptotic expression of $\mathbb{E}(L_n)$. Let $\tilde{f}_1(z) = e^{-z} \sum_{n \geq 0} \mathbb{E}(L_n) \frac{z^n}{n!}$, we get

$$\tilde{f}_1(z) + \tilde{f}_1'(z) = 2\tilde{f}_1\left(\frac{z}{2}\right) + \left(\frac{z^2}{4} - 1\right) e^{-z} - ze^{-z/2} + 1.$$

Now, set $\tilde{p}(z) = \left(\frac{z^2}{4} - 1\right) e^{-z} - ze^{-z/2} + 1$. We need to find an explicit expression of

$$G_E(\omega) = \int_0^\infty \frac{s^{\omega-1}}{Q(-2s)} \left(\int_0^\infty \tilde{p}(z)e^{-sz}dz\right) ds.$$

We begin with the Laplace transform of $\tilde{p}(z)$

$$\mathscr{L}[\tilde{p}; s] = \frac{1}{2(s+1)^3} - \frac{1}{(s+1)} - \frac{1}{(s+1/2)^2} + \frac{1}{s}.$$

Set

$$g(s) = \frac{1}{2(s+1)^3} - \frac{1}{s+1} - \frac{1}{(s+1/2)^2},$$

we first compute the Mellin transform of $\frac{g(s)}{Q(-2s)}$. Equation 2.2.5 of [5] gives us that

$$\frac{1}{Q(-2s)} = \sum_{n\geq 0} \frac{(-1)^n s^n}{Q_n}, \quad \text{when } |s| < 1.$$

Thus,

$$\begin{aligned}\frac{g(s)}{Q(-2s)} =& \sum_{r\geq 0} \left(\frac{(r+2)(r+1)}{4} - 1 - (r+1)2^{r+2}\right)(-1)^r s^r \sum_{n\geq 0} \frac{(-1)^n s^n}{Q_n} \\ =& \sum_{n\geq 0} (-1)^n s^n \sum_{r\geq 0} \frac{1}{Q_{n-r}} \left(\frac{(r+2)(r+1)}{4} - 1 - (r+1)2^{r+2}\right).\end{aligned}$$

By the direction mapping theorem from Section 3.3, the singular expansion will be

$$\mathscr{M}\left[\frac{g(s)}{Q(-2s)}; \omega\right] \asymp \sum_{n\geq 0} \left(\sum_{r=0}^{n} \frac{1}{Q_{n-r}} \left(\frac{(r+2)(r+1)}{4} - 1 - (r+1)2^{r+2}\right)\right) \frac{(-1)^n}{\omega+n}.$$

From equation 2.2.6 of [5], we get that

$$\frac{1}{Q_n} = \frac{Q(2^n)}{Q(1)} = \frac{1}{Q(1)} \sum_{l\geq 0} a_{l+1} 2^{-nl},$$

where

$$a_{l+1} = \frac{(-1)^l 2^{-\binom{l+1}{2}}}{Q_l}.$$

Next, we compute the meromorphic extension of

$$\begin{aligned}
&\sum_{r=0}^{n}\frac{1}{Q_{n-r}}\left(\frac{(r+2)(r+1)}{4}-1-(r+1)2^{r+2}\right)\\
=&\frac{1}{Q(1)}\sum_{l\geq 0}a_{l+1}\sum_{r=0}^{n}\left(\frac{(n-r+2)(n-r+1)}{4}2^{-rl}-2^{-rl}-(n-r+1)2^{n+2-r(l+1)}\right)\\
=&\frac{8\cdot 2^{4l}-32\cdot 2^{3l}+46\cdot 2^{2l}-32\cdot 2^{l}+9}{2^{1-ln}(2\cdot 2^{l}-1)^{2}(2^{l}-1)^{3}}-\frac{2^{n+l+3}\left(n(2^{l+1}-1)+2^{l+1}-2\right)}{(2\cdot 2^{l}-1)^{2}}\\
&+\frac{2^{l}\left(2^{l}(n^{2}+3n-2)-2^{l+1}(n^{2}+4n-2)+n^{2}+5n+2\right)}{4(2^{l}-1)^{3}}
\end{aligned}$$

We let

$$\begin{aligned}
\kappa(\omega)=&\frac{8\cdot 2^{4l}-32\cdot 2^{3l}+46\cdot 2^{2l}-32\cdot 2^{l}+9}{2^{1-l\omega}(2\cdot 2^{l}-1)^{2}(2^{l}-1)^{3}}-\frac{2^{\omega+l+3}\left(\omega(2^{l+1}-1)+2^{l+1}-2\right)}{(2\cdot 2^{l}-1)^{2}}\\
&+\frac{2^{l}\left(2^{l}(\omega^{2}+3\omega-2)-2^{l+1}(\omega^{2}+4\omega-2)+\omega^{2}+5\omega+2\right)}{4(2^{l}-1)^{3}},
\end{aligned}$$

then

$$\mathscr{M}\left[\frac{g(s)}{Q(-2s)};\omega\right]\asymp\sum_{n\geq 0}\kappa(n)\frac{(-1)^{n}}{\omega+n}.$$

By the same argument as in Example 5 of [62], we get that

$$\mathscr{M}\left[\frac{g(s)}{Q(-2s)};\omega\right]=\kappa(-\omega)\Gamma(\omega)\Gamma(1-\omega).$$

Moreover, from [66], we have

$$\int_{0}^{\infty}\frac{s^{\omega-2}}{Q(-2s)}ds=\frac{Q(2^{\omega-1})}{Q(1)}\Gamma(-\omega)\Gamma(\omega+1)\quad\text{for}\quad\Re(\omega)>2.$$

As a result,

$$G_{E}(\omega)=\kappa(-\omega)\Gamma(\omega)\Gamma(1-\omega)+\frac{Q(2^{\omega-1})}{Q(1)}\Gamma(-\omega)\Gamma(\omega+1).$$

Note that the asymptotic expression of L_n in DSTs has already been derived by Du and Prodinger with the Rice integral method in [52]. Their result is given as

$$\mathbb{E}(L_{n})=\frac{n}{Q(1)}\sum_{m\geq 0}a_{m+1}b_{m}+n\delta(\log_{2}n)+\mathcal{O}(1),$$

where

$$b_m := \frac{1}{4\log 2}\frac{B(2^{-m})}{(2^{-m}-1)^3(2^{-m}-2)^2}$$

with

$$\begin{aligned} B(x) =& 16(1-x)^3\log 2-(x-1)(x-2)(7x^2-15x+10) \\ &-2x(4+6x-5x^2+2x^3)\log x. \end{aligned}$$

Du and Prodinger computed the numerical value of $\frac{1}{Q(1)}\sum_{m\geq 0} a_{m+1}b_m$ as 0.3070798... and claimed that $\delta(x)$ is a periodic function with tiny amplitude. However, they did not compute the Fourier coefficient of the periodic function $\delta(x)$.

Comparing with our result, we have that

$$\lim_{\omega\to 2}\frac{G_E(\omega)}{\log 2} \approx 0.3070798...$$

which coincide with Du and Prodinger's result. Our method can also derive all the Fourier coefficients of the periodic function at once, which is obviously an advantage over the Rice integral method.

Now, we turn to the limiting distribution.

Theorem 5.5.2.
$$\frac{L_n-\mathbb{E}(L_n)}{\sqrt{\operatorname{Var}(L_n)}} \xrightarrow{d} \mathscr{N}(0,1).$$

Proof. To apply Theorem 5.2.3, we need to verify that

1. $\|T_n\|_s = o(\sqrt{n})$ with $2 < s \leq 3$.

2. $\tau_1(z) \in \mathscr{J}\mathscr{S}_{\alpha_1,\gamma_1}$ with $0 \leq \alpha_1 < 1/2$ and $\tilde{\tau}_2(z) \in \mathscr{J}\mathscr{S}$.

3. $\tilde{V}_T(z) \in \mathscr{J}\mathscr{S}_{\alpha_2,\gamma_2}$ with $0 \leq \alpha_2 < 1$.

4. $\mathbb{V}(L_n) \geq cn$ for some $c > 0$.

The first condition is trivial since $\|T_n\|_s = \mathcal{O}(1)$ for all $s \geq 1$. The second, third and fourth conditions are already verified in the previous theorem. Thus, we only need the show that $\mathbb{V}(L_n) \geq cn$ for some $c > 0$.

From the discussion in Section 5.3, we only need to find some $n' > 3$ such that $\vartheta_{n'} > 0$ where ϑ_n is defined as (5.20). From (5.28) and the initial conditions $\mu_0 = \mu_1 = \mu_2 = 0$ and $\mu_3 = 1/2$, we can easily compute that $\vartheta_4 = 27/16$. Therefore, the fourth condition is satisfied and the result follows. □

5.6 k-Cousins in Digital Trees

5.6.1 Introduction

In computer science, the similarity of strings is an important area with many subareas such as string metric, string matching algorithms and fuzzy string searching. This issue is of great important because it has numerous applications, including DNA analysis, data mining, image analysis and many others. In this article, we will study the similarity of strings stored in digital trees via the parameter k-**cousins**.

In [143], H. Mahmoud and M. D. Ward proposed a new parameter, the so-called k-cousins, to study the similarity of strings stored in digital trees. In a digital tree, a subtree with k keys is said to be ”on the fringe” if there is no proper subtree which also contain k keys. Those k keys contained in a subtree on the fringe form a k-cousin. In other word, k-cousins can be seen as prefixes contained by exactly k keys. For example, in binary tries, 2-cousins are exactly pairs of two strings sharing a common parent node.

The mean of number of k-cousins in m-ary tries is studied in [143] by Poissonization and Mellin transform. In this section, we will use the framework introduced in Section 5.1 and Section 5.2 to study the mean, variance and limit laws of k-cousins in digital trees.

5.6.2 k-cousin in m-ary Tries

Let $X_{n,k}^{(T)}$ be the random variable which count the number of k-cousins in a m-ary Trie containing n keys. From the construction of Tries and the definition of k-cousins, we get the following distributional recurrence

$$X_{n,k}^{(T)} = \sum_{r=1}^{m} X_{I_n^{(r)},k}^{(T)}, \qquad n > k, \tag{5.29}$$

with the initial conditions $X_{n,k}^{(T)} = 0$ whenever $n < k$ and $X_{n,k}^{(T)} = 1$ when $n = k$. We can easily seen that (5.29) satisfies (5.3) with the toll function $T_n = 0$ for all n. Applying Proposition 5.1.6 and Theorem 5.1.7, we get the following results.

Lemma 5.6.1. *$\mathbb{V}(X_{n,k}^{(T)}) \geq cn$ for some positive constant c.*

Proof. Let $\mu_{n,k} = \mathbb{E}\left(X_{n,k}^{(T)}\right)$, then from the recurrence (5.29) and the initial conditions, we get that

$$\mu_{k+1,k}=\sum_{r=1}^{m}\sum_{t=0}^{k+1}\binom{k+1}{t}p_r^t\mu_{t,k}=(k+1)\sum_{r=1}^{m}p_r^k+\mu_{k+1,k}\sum_{r=1}p_r^{k+1}.$$

Thus,

$$\mu_{k+1,k}=(k+1)\frac{\sum_{r=1}^{m}p_r^k}{1-\sum_{r=1}^{m}p_r^{k+1}}.$$

If $\mu_{k+1,k}\neq 1$, then by the same argument as the proof of Theorem 5.5.2, we get the desired result. Otherwise, we assume that

$$(k+1)\frac{\sum_{r=1}^{m}p_r^k}{1-\sum_{r=1}^{m}p_r^{k+1}}=1,$$

then

$$\begin{aligned}\mu_{k+2,k}=&\sum_{r=1}^{m}\sum_{j_r=0}^{k+2}\binom{k+2}{j_r}p_r^{j_r}\mu_{j_r,k}\\=&\sum_{r=1}^{m}\left(\binom{k+2}{2}p_r^k+(k+2)p_r^{k+1}+p_r^{k+2}\mu_{k+2,k}\right)\\=&\frac{1}{1-\sum_{r=1}^{m}p_r^{k+2}}\left(m\binom{k+2}{2}\sum_{r=1}^{m}p_r^k+(k+2)\sum_{r=1}^{m}p_r^{k+1}\right)\\=&\frac{m(k+2)}{2\left(1-\sum_{r=1}^{m}p_r^{k+2}\right)}.\end{aligned}$$

It implies that

$$1-\sum_{r=1}^{m}p_r^{k+2}=\frac{m(k+2)}{2}.$$

However, from our assumption,

$$1-\sum_{r=1}^{m}p_r^{k+2}<1\quad\text{while}\quad\frac{m(k+2)}{2}>1.$$

A contradiction. Therefore, $\mu_{k+1,k}\neq 1$ and the result follows. □

Theorem 5.6.2. *We have, as $n\to\infty$,*

$$\mathbb{E}(X_{n,k}^{(T)})\sim nP^{[k]}(\log_{1/a}n),\qquad \mathrm{Var}(X_{n,k}^{(T)})\sim nQ^{[k]}(\log_{1/a}n),$$

where $a > 0$ is a suitable constant and $P(z), Q(z)$ are infinitely differentiable, 1-periodic functions (possibly constant) for all $k \in \mathbb{N}$. Moreover,

$$\frac{X_{n,k}^{(T)} - \mathbb{E}(X_{n,k}^{(T)})}{\sqrt{\operatorname{Var}(X_{n,k}^{(T)})}} \xrightarrow{d} N(0,1).$$

Proof. Similar to the proof of Theorem 5.5.1 and Theorem 5.5.2, we only need to check that the conditions of Proposition 5.1.6 and Theorem 5.1.7 are satisfied. From 5.29, we get that for given k

$$\tilde{f}_1^{[k]}(z) = e^{-z} \sum_{n \geq 0} \mathbb{E}\left(X_{n,k}^{(T)}\right) \frac{z^n}{n!} = \sum_{r=1}^{m} \tilde{f}_1^{[k]}(p_r z) + \left(1 - \sum_{i=1}^{m} p_i^k\right) \frac{z^k e^{-z}}{k!}$$

and

$$\begin{aligned} \tilde{f}_2^{[k]}(z) =& e^{-z} \sum_{n \geq 0} \mathbb{E} X_{n,k}^{(T)2} \frac{z^n}{n!} \\ =& \sum_{r=1}^{m} \tilde{f}_2^{[k]}(p_r z) + 2 \sum_{r=1}^{m-1} \sum_{s=r+1}^{m} \tilde{f}_1^{[k]}(p_r z) \tilde{f}_1^{[k]}(p_s z) + \left(1 - \sum_{r=1}^{m} p_r^k\right) \frac{z^k e^{-z}}{k!}. \end{aligned}$$

Since $1 - \sum_{i=1}^{m} p_i^k$ is a constant and $z^k e^{-z}/k! \in \mathscr{J}\mathscr{S}_{\alpha,1}$ for all $\alpha < 1$, all the conditions of JS-Admissibility in Proposition 5.1.6 and Theorem 5.1.7 are satisfied. Moreover, we can easily see that $T_n = 0$ for all n from (5.29) and hence $\|T_n\|_s = o(\sqrt{n})$ for all $s \in \mathbb{R}_+$. Plus Lemma 5.6.1, all the conditions are checked and the result follows. □

5.6.3 k-cousins in Digital Search Trees

Similar to the Trie case, we use $X_{n,k}^{(P)}$ to denote the random variable of number of k-cousin in a symmetric DSTs built on n keys. The random variable satisfies the following distributional recurrence

$$X_{n+1,k}^{(P)} = \sum_{r=1}^{m} X_{I_n^{(r)},k}^{(P)}, \quad n \geq k,$$

with the initial conditions $X_{k,k}^{(P)} = 1$ and $X_{n,k}^{(P)} = 0$ for all $n < k$. To apply the framework, we need to establish the lower bound for the variance as we have seen before.

Lemma 5.6.3. *$\mathbb{V}(X_{n,k}^{(P)}) \geq cn$ for some positive constant c.*

Proof. Similar to Lemma 5.6.1. □

Proposition 5.6.4. *For all $k \in \mathbb{N}$, we use $G^{(k)}(\omega)$ to denote the Mellin transform of*

$$\frac{1}{Q(-2s)(s+1)^k}.$$

Then,

$$G^{(k)}(\omega) = \Psi^{(k)}(-\omega)\Gamma(\omega)\Gamma(1-\omega),$$

where

$$\Psi^{(k)}(\omega) = \frac{1}{Q(1)}\sum_{l\geq 0} a_{l+1}\left(\frac{2^{-\omega l}}{(1-2^l)^k} + \frac{\prod_{s=1}^{k}(\omega+s)}{(k-1)!(1-2^l)^k}\sum_{j=1}^{k}\frac{(-1)^j(2^l)^j}{\omega+j}\binom{k-1}{j-1}\right)$$

with

$$a_{l+1} = \frac{(-1)^l 2^{-\binom{l+1}{2}}}{Q_l}.$$

Proof. By Theorem 3.3.7, we get

$$G^{(k)}(\omega) \asymp \sum_{n\geq 0}\left(\sum_{r=0}^{n}\binom{r+k-1}{r}\frac{1}{Q_{n-r}}\right)\frac{(-1)^n}{\omega+n}.$$

Since $\dfrac{1}{Q_n} = \dfrac{Q(2^{-n})}{Q(1)}$, apply equation 2.2.6 of [5], we derive that

$$\sum_{r=0}^{n}\frac{\binom{r+k-1}{k-1}}{Q_{n-r}} = \frac{1}{Q(1)}\sum_{l\geq 0} a_{l+1}\sum_{r=0}^{n}\binom{n-r+k-1}{k-1}2^{-rl}.$$

With the help of Maple, we get the following meromorphic extension

$$\sum_{r=0}^{n}\binom{n-r+k-1}{k-1}2^{-rl} = \frac{2^{-ln}}{(1-2^l)^k} + \frac{\prod_{s=1}^{k}(n+s)}{(k-1)!(1-2^l)^k}\sum_{j=1}^{k}\frac{(-1)^j(2^l)^j}{n+j}\binom{k-1}{j-1}.$$

We let

$$\Psi^{(k)}(\omega) = \frac{1}{Q(1)}\sum_{l\geq 0} a_{l+1}\left(\frac{2^{-\omega l}}{(1-2^l)^k} + \frac{\prod_{s=1}^{k}(\omega+s)}{(k-1)!(1-2^l)^k}\sum_{j=1}^{k}\frac{(-1)^j(2^l)^j}{\omega+j}\binom{k-1}{j-1}\right),$$

then

$$G^{(k)}(\omega) \asymp \sum_{n\geq 0} \Psi^{(k)}(n)\frac{(-1)^n}{\omega+n}.$$

By the same argument as in Example 5 of [62], we get that

$$G^{(k)}(\omega) = \Psi^{(k)}(-\omega)\Gamma(\omega)\Gamma(1-\omega).$$

□

Theorem 5.6.5. *We have, as $n \to \infty$,*

$$\mathbb{E}(X_{n,k}^{(P)}) \sim n\varpi_E^{[k]}(\log_{1/a} n), \qquad \mathrm{Var}(X_{n,k}^{(P)}) \sim n\varpi_V^{[k]}(\log_{1/a} n),$$

where $a > 0$ is a suitable constant and $\varpi_E^{[k]}(z), \varpi_V^{[k]}(z)$ are infinitely differentiable, 1-periodic functions (possibly constant) for all $k \in \mathbb{N}$. The explicit expression of $\varpi_E^{[k]}(z)$ is given by

$$\varpi_E^{[k]}(\omega) = \Psi^{(k)}(-\omega)\Gamma(\omega)\Gamma(1-\omega)$$

where $\Psi^{(k)}(\omega)$ is defined as Proposition 5.6.4. Moreover,

$$\frac{X_{n,k}^{(P)} - \mathbb{E}(X_{n,k}^{(P)})}{\sqrt{\mathrm{Var}(X_{n,k}^{(P)})}} \xrightarrow{d} N(0,1).$$

Proof. Similar to Theorem 5.6.2, Theorem 5.5.1 and Theorem 5.5.2, we use the results from Lemma 5.6.3 and Proposition 5.6.4 to show that all the assumptions of Proposition 5.2.2 and Theorem 5.2.3 are satisfied. The details are emitted here. □

Chapter 6

Conclusion

The main purpose of this thesis was to contribute to the analysis of additive shape parameters in random digital trees. The results in this thesis can be divided into two topic areas.

The first topic area was concerned with new applications of the recently proposed Poisson-Laplace-Mellin method. In [74], all the applications of the Poisson-Laplace-Mellin method were for shape parameters of linear order (up to a power of logarithms). In Chapter 4, we collected many examples of shape parameters which are not of linear order, including the leftmost path length, the Wiener index and the total Steiner distance. We derived asymptotic expansions of the mean and variance for these parameters. Moreover, we proved limit laws as well.

The second topic area was concerened with general framworks for central limit theorems of additive shape parameters in random digital trees. In Chapter 5, we first introduced our framework from [77] for proving central limit theorems for shape paramters in m-ary tries. Then, we extended this framework to shape parameters in symmetric digital search trees. We also gave two examples to illustrate how our frameworks work.

As for open problems, the most straightforward one is the extension of our study of the total Steiner distance to other digital trees. In fact, such a study can be performed by the methods we introduced in this thesis. However, the computations are cumbersome. Another obvious question is whether our results for symmetric DSTs can be extended to asymmetric DSTs? For parameters satisfying one-sided distributional recurrences, such as the leftmost path length, we saw that the Poisson-Laplace-Mellin method still works in the asymmetric case. On the other hand, for parameters satisfying two-sided distributional recurrences, this is no longer true. Netherless, with similar tools as in our thesis, deriving asymptotic expansions of mean, variance and obtaining the limit law is still possible. However, asymptotic expressions are not explicit. Thus, finding a general method for deriving explicit asymptotic expressions of the mean and variance of shape parameters satisfying a two-sided distributional recurrence in asymmetric DSTs is an important open question. As a final open problem, note that our frameworks in Chapter 5 are for proving central limit laws. So, a natural question is whether or not similar frameworks can be given for local limit laws and rates of convergence?

We end this thesis by placing our research in a larger context. Therefore, we point out that research of random digital trees is part of the more general study of binomial splitting processes (BSPs) in which the binomial distribution and some of its extensions play an important role. For an extensive introduction into BSPs, see [75]. In this thesis, we mainly dealt with functional equations of the form

$$\tilde{f}(z)+\tilde{f}'(z)=2\tilde{f}\left(\frac{z}{2}\right)+\tilde{g}(z)$$

or

$$\tilde{f}(z)=\sum_{r=1}^{m}\tilde{f}(p_r z)+\tilde{h}(z).$$

Such (differential-)functional equations are special cases of the more general form

$$\sum_{j=0}^{b}\binom{b}{j}\tilde{f}^{(j)}(z)=\sum_{r=1}^{m}a_r\tilde{f}(p_r z+\lambda)+\tilde{g}(z)$$

which underlies the study of BSPs. Most of the recent research has focused on the case $\lambda = 0$. Very little is known about the case $\lambda > 0$ which is also important in applications; see [55, 56, 91]. Thus, there is still a lot of research to be done and the we have still a long way ahead of us before having a complete understanding of the stochastic properties of BSPs.

Appendices

Appendix A

We use the same notation for poissonized means, variances and covariances as in Section 2. In addition, for the node-wise Wiener index of bucket digital search trees and the internal Wiener index for tries, we denote by $\tilde{h}_1(z)$ the Poisson generating function of $\mathbb{E}(N_n)$ and

$$\begin{aligned}\tilde{H}_N(z) &= \tilde{g}_N(z) - \tilde{h}_1(z)^2 - z\tilde{h}_1'(z)^2,\\ \tilde{H}_T(z) &= \tilde{g}_T(z) - \tilde{h}_1(z)\tilde{f}_{1,0}(z) - z\tilde{h}_1'(z)\tilde{f}_{1,0}'(z),\\ \tilde{H}_W(z) &= \tilde{g}_W(z) - \tilde{h}_1(z)\tilde{f}_{0,1}(z) - z\tilde{h}_1'(z)\tilde{f}_{0,1}'(z),\end{aligned}$$

where $\tilde{g}_N(z), \tilde{g}_T(z)$ and $\tilde{g}_W(z)$ denote the Poisson generating function of $\mathbb{E}(N_n^2)$, $\mathbb{E}(N_nT_n)$ and $\mathbb{E}(N_nW_n)$, respectively.

Key-wise Wiener Index of Bucket Digital Search Trees. We have,

$$\begin{aligned}\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j)}(z) &= 2\tilde{f}_{1,0}(z/2) + z,\\ \sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j)}(z) &= 2\tilde{f}_{0,1}(z/2) + (z+2)\tilde{f}_{1,0}(z/2) + \frac{z^2}{2} + z\end{aligned}$$

and

$$\begin{aligned}\sum_{j=0}^{b}\binom{b}{j}\tilde{V}^{(j)}(z) = 2\tilde{V}(z/2) &+ \left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j)}(z)\right)^2 + z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j+1)}(z)\right)^2\\ &- \sum_{j=0}^{b}\binom{b}{j}\left(\tilde{f}_{1,0}(z)^2 + z\tilde{f}_{1,0}'(z)^2\right)^{(j)},\end{aligned}$$

$$
\begin{aligned}
\sum_{j=0}^{b}\binom{b}{j}\tilde{C}^{(j)}(z) =& 2\tilde{C}(z/2)+(z+2)\tilde{V}(z/2)\\
&+\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j)}(z)\right)\\
&+z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j+1)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j+1)}(z)\right)\\
&-\sum_{j=0}^{b}\binom{b}{j}\left(\tilde{f}_{1,0}(z)\tilde{f}_{0,1}(z)+z\tilde{f}'_{1,0}(z)\tilde{f}'_{0,1}(z)\right)^{(j)},\\
\sum_{j=0}^{b}\binom{b}{j}\tilde{W}^{(j)}(z) =& 2\tilde{W}(z/2)+(2z+4)\tilde{C}(z/2)+\left(\frac{z^2}{2}+3z+2\right)\tilde{V}(z/2)\\
&+z^2\tilde{f}'_{1,0}(z/2)^2+2z^2\tilde{f}'_{1,0}(z/2)+z^2\\
&+\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j)}(z)\right)^2+z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j+1)}(z)\right)^2\\
&-\sum_{j=0}^{b}\binom{b}{j}\left(\tilde{f}_{0,1}(z)^2+z\tilde{f}'_{0,1}(z)^2\right)^{(j)}.
\end{aligned}
$$

Node-wise Wiener Index of Bucket Digital Search Trees. We have,

$$
\begin{aligned}
\sum_{j=0}^{b}\binom{b}{j}\tilde{h}_1^{(j)}(z) =& 2\tilde{h}_1(z/2)+1,\\
\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(b)}(z) =& 2\tilde{f}_{1,0}(z/2)+2\tilde{h}_1(z/2),\\
\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(b)}(z) =& 2\tilde{f}_{0,1}(z/2)+2\tilde{f}_{1,0}(z/2)\tilde{h}_1(z/2)+2\tilde{h}_1(z/2)^2\\
&+2\tilde{f}_{1,0}(z/2)+2\tilde{h}_1(z/2)
\end{aligned}
$$

and

$$
\sum_{j=0}^{b}\binom{b}{j}\tilde{H}_N^{(j)}(z)=2\tilde{H}_N(z/2)+\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{h}_1^{(j)}(z)\right)^2
$$

$$
\begin{aligned}
&+ z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{h}_1^{(j+1)}(z)\right)^2 - \sum_{j=0}^{b}\binom{b}{j}\left(\tilde{h}_1(z)^2 + z\tilde{h}_1'(z)^2\right)^{(j)},\\
\sum_{j=0}^{b}\binom{b}{j}\tilde{H}_T^{(j)}(z) =&2\tilde{H}_T(z/2) + 2\tilde{H}_N(z/2)\\
&+ \left(\sum_{j=0}^{b}\binom{b}{j}\tilde{h}_1^{(j)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j)}(z)\right)\\
&+ z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{h}_1^{(j+1)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j+1)}(z)\right)\\
&- \sum_{j=0}^{b}\binom{b}{j}\left(\tilde{h}_1(z)\tilde{f}_{1,0}(z) + z\tilde{h}_1'(z)\tilde{f}_{1,0}'(z)\right)^{(j)},\\
\sum_{j=0}^{b}\binom{b}{j}\tilde{V}^{(j)}(z) =&2\tilde{V}(z/2) + 4\tilde{H}_T(z/2) + 2\tilde{H}_N(z/2) + \left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j)}(z)\right)^2\\
&+ z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j+1)}(z)\right)^2 - \sum_{j=0}^{b}\binom{b}{j}\left(\tilde{f}_{1,0}(z)^2 + z\tilde{f}_{1,0}'(z)^2\right)^{(j)},\\
\sum_{j=0}^{b}\binom{b}{j}\tilde{H}_W^{(j)}(z) =&2\tilde{H}_W(z/2) + 2\tilde{H}_T(z/2)(\tilde{h}_1(z/2) + 1) + 2\tilde{H}_N(z/2)(2\tilde{h}_1(z/2)\\
&+ \tilde{f}_{1,0}(z/2) + 1) + \left(\sum_{j=0}^{b}\binom{b}{j}\tilde{h}_1^{(j)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j)}(z)\right)\\
&+ z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{h}_1^{(j+1)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j+1)}(z)\right)\\
&- \sum_{j=0}^{b}\binom{b}{j}\left(\tilde{h}_1(z)\tilde{f}_{0,1}(z) + z\tilde{h}_1'(z)\tilde{f}_{0,1}'(z)\right)^{(j)},\\
\sum_{j=0}^{b}\binom{b}{j}\tilde{C}^{(j)}(z) =&2\tilde{C}(z/2) + 2\tilde{H}_W(z/2) + 2\tilde{V}(z/2)(\tilde{h}_1(z/2) + 1)\\
&+ 2\tilde{H}_T(z/2)(3\tilde{h}_1(z/2) + \tilde{f}_{1,0}(z/2) + 2) + 2\tilde{H}_N(z/2)(2\tilde{h}_1(z/2)\\
&+ \tilde{f}_{1,0}(z/2) + 1) + \left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j)}(z)\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{1,0}^{(j+1)}(z)\right)\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j+1)}(z)\right)\\
&- \sum_{j=0}^{b}\binom{b}{j}\left(\tilde{f}_{1,0}(z)\tilde{f}_{0,1}(z) + z\tilde{f}'_{1,0}(z)\tilde{f}'_{0,1}(z)\right)^{(j)},
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=0}^{b}\binom{b}{j}\tilde{W}^{(j)}(z) =& 2\tilde{W}(z/2) + 4\tilde{C}(z/2)(\tilde{h}_1(z/2)+1) + 4\tilde{H}_W(z/2)(2\tilde{h}_1(z/2)\\
&+ \tilde{f}_{1,0}(z/2) + 1) + 2\tilde{V}(z/2)\tilde{H}_N(z/2)\\
&+ \tilde{V}(z/2)((2+z)\tilde{h}_1(z/2)^2 + 4\tilde{h}_1(z/2) + 2) + 2\tilde{H}_T(z/2)^2\\
&+ \tilde{H}_T(z/2)(8\tilde{h}_1(z/2)^2 + 16\tilde{h}_1(z/2)\\
&+ 4z\tilde{h}'_1(z/2)^2 + 4\tilde{h}_1(z/2)\tilde{f}_{1,0}(z/2) + 2z\tilde{h}'_1(z/2)\tilde{f}'_{1,0}(z/2) + 4)\\
&+ 4\tilde{H}_N(z/2)^2 + 8\tilde{H}_N(z/2)^2\tilde{h}_1(z/2)^2\\
&+ 8\tilde{H}_N(z/2)\tilde{H}_T(z/2) + \tilde{H}_N(z/2)(8\tilde{h}_1(z/2)\\
&+ 4z\tilde{h}'_1(z/2)^2 + 8\tilde{h}_1(z/2)\tilde{f}_{1,0}(z/2)\\
&+ 4z\tilde{h}'_1(z/2)\tilde{f}'_{1,0}(z/2) + 2\tilde{f}_{1,0}(z/2)^2 + 4\tilde{f}_{1,0}(z/2)\\
&+ z\tilde{f}'_{1,0}(z/2)^2 + 2) + z^2\tilde{h}_1(z/2)^4 + 2z^2\tilde{h}_1(z/2)^3\tilde{f}_{1,0}(z/2)\\
&+ z^2\tilde{h}'_1(z/2)^2\tilde{f}'_{1,0}(z/2)^2\\
&+ \left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j)}(z)\right)^2 + z\left(\sum_{j=0}^{b}\binom{b}{j}\tilde{f}_{0,1}^{(j+1)}(z)\right)^2\\
&- \sum_{j=0}^{b}\binom{b}{j}\left(\tilde{f}_{0,1}(z)^2 + z\tilde{f}'_{0,1}(z)^2\right)^{(j)}.
\end{aligned}
$$

External Wiener Index of Tries. We have,

$$
\begin{aligned}
\tilde{f}_{1,0}(z) &= 2\tilde{f}_{1,0}(z/2) + z - ze^{-z},\\
\tilde{f}_{0,1}(z) &= 2\tilde{f}_{0,1}(z/2) + z\tilde{f}_{1,0}(z/2) + \frac{z^2}{2}
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{V}(z) = 2\tilde{V}(z/2) &+ e^{-z}(4z\tilde{f}_{1,0}(z/2) + 2z\tilde{f}'_{1,0}(z/2) - 2z^2\tilde{f}'_{1,0}(z/2))\\
&+ e^{-z}(z - ze^{-z} + z^2e^{-z} - z^3e^{-z}),
\end{aligned}
$$

$$
\begin{aligned}
\tilde{C}(z) &= 2\tilde{C}(z/2) + z\tilde{V}(z/2) + e^{-z}\Big(z\tilde{f}_{1,0}(z/2) + \frac{z^2}{2}\tilde{f}'_{1,0}(z/2) - \frac{z^3}{2}\tilde{f}'_{1,0}(z/2) \\
&\qquad + 2z\tilde{f}_{0,1}(z/2) + z\tilde{f}'_{0,1}(z/2) - z^2\tilde{f}'_{0,1}(z/2)\Big) + e^{-z}\left(z^2 - \frac{z^3}{2}\right), \\
\tilde{W}(z) &= 2\tilde{W}(z/2) + 2z\tilde{C}(z/2) + \left(\frac{z^2}{2} + z\right)\tilde{V}(z/2) + z^2\tilde{f}'_{1,0}(z/2)^2 \\
&\qquad + 2z^2\tilde{f}'_{1,0}(z/2) + z^2.
\end{aligned}
$$

Internal Wiener Index of Tries. We have,

$$
\begin{aligned}
\tilde{h}_1(z) =& 2\tilde{h}_1(z/2) + 1 - e^{-z}(1+z), \\
\tilde{f}_{1,0}(z) =& 2\tilde{f}_{1,0}(z/2) + 2\tilde{h}_1(z/2), \\
\tilde{f}_{0,1}(z) =& 2\tilde{f}_{0,1}(z/2) + 2\tilde{f}_{1,0}(z/2)\tilde{h}_1(z/2) + 2\tilde{h}_1(z/2)^2 + 2\tilde{f}_{1,0}(z/2) + 2\tilde{h}_1(z/2)
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{H}_N(z) =& 2\tilde{H}_N(z/2) + e^{-z}(4\tilde{h}_1(z/2) + 4z\tilde{h}_1(z/2) - 2z^2\tilde{h}'_1(z/2)) \\
&\quad + e^{-z}(1 + z - e^{-z} - 2ze^{-z} - z^2e^{-z} - z^3e^{-z}), \\
\tilde{H}_T(z) =& 2\tilde{H}_T(z/2) + 2\tilde{H}_N(z/2) + e^{-z}(2\tilde{h}_1(z/2) + 2z\tilde{h}_1(z/2) - z^2\tilde{h}'_1(z/2) \\
&\quad + 2\tilde{f}_{1,0}(z/2) + 2z\tilde{f}_{1,0}(z/2) - z^2\tilde{f}'_{1,0}(z/2)), \\
\tilde{V}(z) =& 2\tilde{V}(z/2) + 4\tilde{H}_T(z/2) + 2\tilde{H}_N(z/2), \\
\tilde{H}_W(z) =& 2\tilde{H}_W(z/2) + 2\tilde{H}_T(z/2)(\tilde{h}_1(z/2) + 1) + 2\tilde{H}_N(z/2)(2\tilde{h}_1(z/2) \\
&\quad + \tilde{f}_{1,0}(z/2) + 1) + e^{-z}(2\tilde{h}_1(z/2)^2 + 2z\tilde{h}_1(z/2) + 2\tilde{h}_1(z/2) + 2z\tilde{h}_1(z/2) \\
&\quad - z^2\tilde{h}_1(z/2)\tilde{h}'_1(z/2) - z^2\tilde{h}'_1(z/2) + 2\tilde{h}_1(z/2)\tilde{f}_{1,0}(z/2) + 2z\tilde{h}_1(z/2)\tilde{f}_{1,0}(z/2) \\
&\quad - z^2\tilde{h}_1(z/2)\tilde{f}'_{1,0}(z/2) - z^2\tilde{h}'_1(z/2)\tilde{f}_{1,0}(z/2) + 2\tilde{f}_{1,0}(z/2) + 2z\tilde{f}_{1,0}(z/2) \\
&\quad - z^2\tilde{f}'_{1,0}(z/2) + 2\tilde{f}_{0,1}(z/2) + 2z\tilde{f}_{0,1}(z/2) - z^2\tilde{f}'_{0,1}(z/2)), \\
\tilde{C}(z) =& 2\tilde{C}(z/2) + 2\tilde{H}_W(z/2) + 2\tilde{V}(z/2)(\tilde{h}_1(z/2) + 1) \\
&\quad + 2\tilde{H}_T(z/2)(3\tilde{h}_1(z/2) + \tilde{f}_{1,0}(z/2) + 2) + 2\tilde{H}_N(z/2)(2\tilde{h}_1(z/2) \\
&\quad + \tilde{f}_{1,0}(z/2) + 1), \\
\tilde{W}(z) =& 2\tilde{W}(z/2) + 4\tilde{C}(z/2)(\tilde{h}_1(z/2) + 1) + 4\tilde{H}_W(z/2)(2\tilde{h}_1(z/2) + \tilde{f}_{1,0}(z/2) + 1) \\
&\quad + 2\tilde{V}(z/2)\tilde{H}_N(z/2) + \tilde{V}(z/2)((2+z)\tilde{h}_1(z/2)^2 + 4\tilde{h}_1(z/2) + 2) \\
&\quad + 2\tilde{H}_T(z/2)^2 + \tilde{H}_T(z/2)(8\tilde{h}_1(z/2)^2 + 16\tilde{h}_1(z/2) + 4z\tilde{h}'_1(z/2)^2 \\
&\quad + 4\tilde{h}_1(z/2)\tilde{f}_{1,0}(z/2) + 2z\tilde{h}'_1(z/2)\tilde{f}'_{1,0}(z/2) + 4) + 4\tilde{H}_N(z/2)^2
\end{aligned}
$$

$$+ 8\tilde{H}_N(z/2)^2\tilde{h}_1(z/2)^2 + 8\tilde{H}_N(z/2)\tilde{H}_T(z/2) + \tilde{H}_N(z/2)(8\tilde{h}_1(z/2)$$
$$+ 4z\tilde{h}_1'(z/2)^2 + 8\tilde{h}_1(z/2)\tilde{f}_{1,0}(z/2) + 4z\tilde{h}_1'(z/2)\tilde{f}_{1,0}'(z/2)$$
$$+ 2\tilde{f}_{1,0}(z/2)^2 + 4\tilde{f}_{1,0}(z/2) + z\tilde{f}_{1,0}'(z/2)^2 + 2) + z^2\tilde{h}_1(z/2)^4$$
$$+ 2z^2\tilde{h}_1(z/2)^3\tilde{f}_{1,0}(z/2) + z^2\tilde{h}_1'(z/2)^2\tilde{f}_{1,0}'(z/2)^2.$$

External Wiener Index of PATRICIA Tries. We have,

$$\tilde{f}_{1,0}(z) = 2\tilde{f}_{1,0}(z/2) + z - ze^{-z/2},$$
$$\tilde{f}_{0,1}(z) = 2\tilde{f}_{0,1}(z/2) + z\tilde{f}_{1,0}(z/2) + \frac{z^2}{2}$$

and

$$\begin{aligned}\tilde{V}(z) =& 2\tilde{V}(z/2) + e^{-z/2}(2z\tilde{f}_{1,0}(z/2) - z^2\tilde{f}_{1,0}'(z/2)) + e^{-z/2}\left(z + \frac{z^2}{2}\right)\\ &- e^{-z}\left(z + \frac{z^3}{4}\right),\\ \tilde{C}(z) =& 2\tilde{C}(z/2) + z\tilde{V}(z/2) + e^{-z/2}\Big(z\tilde{f}_{1,0}(z/2) + \frac{z^2}{2}\tilde{f}_{1,0}(z/2) + \frac{z^2}{2}\tilde{f}_{1,0}'(z/2)\\ &- \frac{z^3}{4}\tilde{f}_{1,0}'(z/2) + z\tilde{f}_{0,1}(z/2) - \frac{z^2}{2}\tilde{f}_{0,1}'(z/2)\Big) + z^2e^{-z},\\ \tilde{W}(z) =& 2\tilde{W}(z/2) + 2z\tilde{C}(z/2) + \left(\frac{z^2}{2} + z\right)\tilde{V}(z/2)\\ &+ z^2\tilde{f}_{1,0}'(z/2)^2 + 2z^2\tilde{f}_{1,0}'(z/2) + z^2.\end{aligned}$$

Internal Wiener Index of m-ary PATRICIA Tries. We have,

$$\begin{aligned}\tilde{h}(z) =& \sum_{r=1}^{m}\tilde{h}(p_rz) + 1 + (m-1)e^{-z} - \sum_{r=1}^{m}e^{(p_r-1)z},\\ \tilde{f}_{1,0}(z) =& \sum_{r=1}^{m}\tilde{f}_{1,0}(p_rz) + \sum_{r=1}^{m}\tilde{h}(p_rz) - \sum_{r=1}^{m}e^{(p_r-1)z}\tilde{h}(p_rz),\\ \tilde{f}_{0,1}(z) =& \sum_{r=1}^{m}\tilde{f}_{0,1}(p_rz) + \sum_{(r,s)\in S_2}\tilde{h}(p_rz)\tilde{f}_{1,0}(p_sz) + \sum_{(r,s)\in S_2}\tilde{h}(p_rz)\tilde{h}(p_sz)\\ &+ \sum_{r=1}^{m}\left(\tilde{f}_{1,0}(p_rz) + \tilde{h}(p_rz)\right) - \sum_{r=1}^{m}e^{(p_r-1z)}\left(\tilde{f}_{1,0}(p_rz) + \tilde{h}(p_rz)\right),\end{aligned}$$

where $S_2 = \{(r,s) : 1 \le r, s \le m, r \ne s\}$ and

$$\begin{aligned}
\tilde{H}_T(z) &= \sum_{i=1}^{m} \tilde{H}_T(p_i z) + \sum_{i=1}^{m} \tilde{H}_N(p_i z) + \tilde{g}_{N,T}(z), \\
\tilde{V}(z) &= \sum_{i=1}^{m} \tilde{V}(p_i z) + 2\sum_{i=1}^{m} \tilde{H}_T(p_i z) + \tilde{g}_T(z), \\
\tilde{H}_W(z) &= \sum_{i=1}^{m} \tilde{H}_W(p_i z) + \sum_{(i,j)\in S_2} \Big(\tilde{H}_N(p_i z)\tilde{f}_{1,0}(p_j z) + \tilde{H}_T(p_i z)\tilde{h}(p_j z)\Big) \\
&\quad + \tilde{g}_{N,W}(z), \\
\tilde{C}(z) &= \sum_{i=1}^{m} \tilde{C}(p_i z) + \sum_{(i,j)\in S_2} \Big(\tilde{H}_T(p_i z)\tilde{f}_{1,0}(p_j z) + \tilde{V}(p_i z)\tilde{h}(p_j z)\Big) \\
&\quad + \tilde{g}_{T,W}(z), \\
\tilde{W}(z) &= \sum_{i=1}^{m} \tilde{W}(p_i z) + \sum_{(i,j)\in S_2} \Big(\tilde{H}_N(p_i z)\tilde{f}_{1,0}(p_j z)^2 + 2\tilde{H}_T(p_i z)\tilde{h}(p_j z)\tilde{f}_{1,0}(p_j z) \\
&\quad + \tilde{V}(z)\tilde{f}_{1,0}(p_j z)^2 + 2\tilde{H}_W \tilde{f}_{1,0}(p_j z) + 2\tilde{C}(p_i z)\tilde{f}_{1,0}(p_j z)\Big) \\
&\quad + \sum_{(i,j,k)\in S_3} \Big(\tilde{H}_N(p_i z)\tilde{f}_{1,0}(p_j z)\tilde{f}_{1,0}(p_k z) + 2\tilde{H}_T(p_i z)\tilde{h}(p_j z)\tilde{f}_{1,0}(p_k z) \\
&\quad + \tilde{V}(p_i z)\tilde{h}(p_j z)\tilde{h}(p_k z)\Big),
\end{aligned}$$

where $S_3 = \{(i,j,k)\ : 1 \le i, j, k \le m, i \ne j, j \ne k, i \ne k\}$ and

$$\begin{aligned}
\tilde{g}_{N,T}(z) &= o(z), \\
\tilde{g}_T(z) &= \mathcal{O}(z), \\
\tilde{g}_{N,W}(z) &= \mathcal{O}(z^2), \\
\tilde{g}_{T,W}(z) &= \mathcal{O}(z^2 \log z), \\
\tilde{g}_W(z) &= \mathcal{O}(z^3 \log z)
\end{aligned}$$

uniformly in z with $|\arg(z)| \le \phi$ and $0 < \phi < \pi/2$. The expression of $\tilde{g}_{N,T}(z), \tilde{g}_T(z), \tilde{g}_{N,W}(z), \tilde{g}_{T,W}(z)$ and $\tilde{g}_W(z)$ can be computed by computer algebra such as maple. However, the explicit expressions of them are too complicated and hence we do not list them here. We give only the bounds since it is already enough for our purpose.

Appendix B

We will use the following notations

$$\Omega_1(n) = Q^{(k_1)}(\log_{1/a} n), \quad \Omega_2(n) = Q^{(k_2)}(\log_{1/a} n), \quad \Omega_3(n) = Q^{(k_1,k_2)}(\log_{1/a} n)$$

and

$$D(n) = \Omega_1(n)\Omega_2(n) - \Omega_3(n)^2.$$

Then,

$$\begin{aligned} b_n^{(1)} =& \frac{T_n^{(k_1)} - \mu_n^{(k_1)} + \sum_{i=1}^{m} \mu_{I_i}^{(k_1)}}{\sqrt{n}} \\ & \cdot \frac{\left(\Omega_1(n) + \sqrt{D(n)}\right)\left(\sqrt{\Omega_1(n) + \Omega_2(n) + 2\sqrt{D(n)}}\right)}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}} \\ & - \frac{T_n^{(k_2)} - \mu_n^{(k_2)} + \sum_{i=1}^{m} \mu_{I_i}^{(k_2)}}{\sqrt{n}} \cdot \frac{\Omega_3(n)\sqrt{\Omega_1(n) + \Omega_2(n) + 2\sqrt{D(n)}}}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}}, \\ b_n^{(2)} =& \frac{T_n^{(k_2)} - \mu_n^{(k_2)} + \sum_{i=1}^{m} \mu_{I_i}^{(k_2)}}{\sqrt{n}} \\ & \cdot \frac{\left(\Omega_2(n) + \sqrt{D(n)}\right)\left(\sqrt{\Omega_1(n) + \Omega_2(n) + 2\sqrt{D(n)}}\right)}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}} \\ & - \frac{T_n^{(k_1)} - \mu_n^{(k_1)} + \sum_{i=1}^{m} \mu_{I_i}^{(k_1)}}{\sqrt{n}} \cdot \frac{\Omega_3(n)\sqrt{\Omega_1(n) + \Omega_2(n) + 2\sqrt{D(n)}}}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}} \end{aligned}$$

and

$$A_n^{(i)}(1,1) = B_n^{(i)} \cdot \frac{\left(\Omega_1(I_n^{(i)}) + \sqrt{D(I_n^{(i)})}\right)\left(\Omega_2(n) + \sqrt{D(n)}\right) - \Omega_3(n)\Omega_3(I_r^{(n)})}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}},$$

$$A_n^{(i)}(1,2) = B_n^{(i)} \cdot \frac{\Omega_3(I_n^{(i)}) \left(\Omega_2(n) + \sqrt{D(n)}\right) - \Omega_3(n) \left(\Omega_2(I_n^{(i)}) + \sqrt{D(I_n^{(i)})}\right)}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}},$$

$$A_n^{(i)}(2,1) = B_n^{(i)} \cdot \frac{\Omega_3(I_n^{(i)}) \left(\Omega_1(n) + \sqrt{D(n)}\right) - \Omega_3(n) \left(\Omega_1(I_n^{(i)}) + \sqrt{D(I_n^{(i)})}\right)}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}},$$

$$A_n^{(i)}(2,2) = B_n^{(i)} \cdot \frac{\left(\Omega_1(n) + \sqrt{D(n)}\right) \left(\Omega_2(I_n^{(i)}) + \sqrt{D(I_n^{(i)})}\right) - \Omega_3(n)\Omega_3(I_r^{(n)})}{2D(n) + (\Omega_1(n) + \Omega_2(n))\sqrt{D(n)}},$$

where

$$B_n^{(i)} = \sqrt{\frac{I_n^{(i)}}{n}} \cdot \sqrt{\frac{\Omega_1(n) + \Omega_2(n) + 2\sqrt{D(n)}}{\Omega_1(I_n^{(i)}) + \Omega_2(I_n^{(i)}) + 2\sqrt{D(I_n^{(i)})}}}.$$

Bibliography

[1] Rafik Aguech, Nabil Lasmar, and Hosam M. Mahmoud. Distribution of inter-nodes distance in digital trees. In *Discrete Mathematics and Theoretical Computer Science Proceedings, Proceedings of 2005 International Conference on Analysis of Algorithms*, pages 1–10, 2005.

[2] Rafik Aguech, Nabil Lasmar, and Hosam M. Mahmoud. Distances in random digital search trees. *Acta Informatica*, 43(4):243–264, 2006.

[3] Rafik Aguech, Nabil Lasmar, and Hosam M. Mahmoud. Limit distribution of distances in biased random tries. *Journal of Applied Probability*, 43:1–14, 2006.

[4] Alfred V. Aho, Jeffrey D. Ullman, and John E. Hopcroft. *Data Structures and Algorithms*. Addison-Wesley, 1983.

[5] George E. Andrews. *The Theory of Partitions*, volume 2 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 1998.

[6] George E. Andrews, Richard Askey, and Ranjan Roy. *Special Functions*. Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1999.

[7] Alberto Apostolico. *The Myriad Virtue of Suffix Trees*, volume F12 of *NATO ASI*, pages 85–96. Springer, 1985.

[8] James Aspnes and Keren Censor. Approximate shared-memory counting despite a strong adversary. *ACM Transactions on Algorithms (TALG)*, 6(2):25, 2010.

[9] Ricardo A. Baeza-Yates. Some average measures in m-ary search trees. *Information Processing Letters*, 25(6):375–381, 1987.

[10] Jean Bertoin, Philippe Biane, and Marc Yor. Poissonian exponential functionals, q-series, q-integrals, and the moment problem for log-normal distributions. In *Seminar on Stochastic Analysis, Random Fields and Applications IV*, pages 45–56. Springer, 2004.

[11] Michael G. B. Blum, Olivier François, and Svante Janson. The mean, variance and limiting distribution of two statistics sensitive to phylogenetic tree balance. *Annals of Applied Probability*, 16(4):2195–2214, 2006.

[12] Miklós Bóna. k-Protected vertices in binary search trees. *Advances in Applied Mathematics*, 53:1–11, 2014.

[13] Andrew D. Booth and Andrew J. T. Colin. On the efficiency of a new method of dictionary construction. *Information and Control*, 3(4):327–334, 1960.

[14] Jérémie Bourdon. Size and path length of Patricia tries: dynamical sources context. *Random Structures & Algorithms*, 19(3-4):289–315, 2001.

[15] Jérémie Bourdon. *Analyze dynamique d'algorithmes: examples en arithmétique et en théorie de l'information*. PhD thesis, Université de Caen Basse-Normandie, 2002.

[16] Jérémie Bourdon, Markus Nebel, and Brigitte Vallée. On the stack-size of general tries. *RAIRO-Theoretical Informatics and Applications*, 35(02):163–185, 2001.

[17] William H. Burge. An analysis of binary search trees formed from sequences of nondistinct keys. *Journal of the ACM*, 23(3):451–454, 1976.

[18] John Capetanakis. Tree algorithms for packet broadcast channels. *IEEE Transactions on Information Theory*, 25(5):505–515, 1979.

[19] Gi-Sang Cheon and Louis W. Shapiro. Protected points in ordered trees. *Applied Mathematics Letters*, 21(5):516–520, 2008.

[20] Hua-Huai Chern and Hsien-Kuei Hwang. Phase changes in random m-ary search trees and generalized quicksort. *Random Structures & Algorithms*, 19(3-4):316–358, 2001.

[21] Jacek Chichoń and Wojciech Macyna. Approximate counters for flash memory. In *Proceedings of the seventeenth IEEE International Conference on Embedded and Real-time Computing Systems and Applications*, pages 185–189, 2011.

[22] Costas A. Christophi and Hosam M. Mahmoud. The oscillatory distribution of distances in random tries. *Annals of Applied Probability*, 15(2):1536–1564, 2005.

[23] Julien Clément, Philippe Flajolet, and Brigitte Vallée. Dynamical source in information theory: a general analysis of trie structures. *Algorithmica*, 29(1-2):307–369, 2001.

[24] Edward G. Coffman Jr and J. Eve. File structures using hashing functions. *Communications of the ACM*, 13(7):427–432, 1970.

[25] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, et al. *Introduction to Algorithms*, volume 2. MIT press Cambridge, 2001.

[26] Davide Crippa and Klaus Simon. q-distributions and Markov processes. *Discrete Mathematics*, 170(1):81–98, 1997.

[27] Miklós Csűrös. Approximate counting with a floating-point counter. In *Computing and Combinatorics*, pages 358–367. Springer, 2010.

[28] Peter Dankelmann, Ortrud R Oellermann, and Henda C Swart. The average steiner distance of a graph. *Journal of Graph Theory*, 22(1):15–22, 1996.

[29] Rene de la Briandais. File searching using variable length keys. In *Papers Presented at the March 3-5, 1959, Western joint Computer Conference*, pages 295–298. ACM, 1959.

[30] Florian Dennert and Rudolf Grübel. Renewals for exponentially increasing lifetimes, with an application to digital search trees. *Annals of Applied Probability*, 17(2):676–687, 2007.

[31] Luc Devroye. A note on the average depth of tries. *Computing*, 28(4):367–371, 1982.

[32] Luc Devroye. A probabilistic analysis of the height of tries and of the complexity of triesort. *Acta Informatica*, 21(3):229–237, 1984.

[33] Luc Devroye. *Lecture Notes on Bucket Algorithms*. Birkhauser Boston, 1986.

[34] Luc Devroye. A note on the height of binary search trees. *Journal of the ACM*, 33(3):489–498, 1986.

[35] Luc Devroye. A Note on the Probabilistic Analysis of Patricia Trees. *Random Structures & Algorithms*, 3(2):203–214, 1992.

[36] Luc Devroye. A study of tree-like structures under the density model. *Annals of Applied Probability*, 2:402–434, 1992.

[37] Luc Devroye. Universal limit laws for depths in random trees. *SIAM Journal on Computing*, 28(2):409–432, 1998.

[38] Luc Devroye. Laws of large numbers and tail inequalities for random tries and PATRICIA trees. *Journal of Computational and Applied Mathematics*, 142(1):27–37, 2002.

[39] Luc Devroye. Universal asymptotics for random tries and PATRICIA trees. *Algorithmica*, 42(1):11–29, 2005.

[40] Luc Devroye and Svante Janson. Protected nodes and fringe subtrees in some random trees. *arXiv preprint arXiv:1310.0665*, 2013.

[41] Luc Devroye and Paul Kruszewski. On the Horton-Strahler number for random tries. *Informatique Théorique et Applications*, 30(5):443–456, 1996.

[42] Luc Devroye and Carlos Zamora-Cura. Expected worst-case partial match in random quadtries. *Discrete Applied Mathematics*, 141(1):103–117, 2004.

[43] Andrey A. Dobrynin, Roger Entringer, and Ivan Gutman. Wiener index of trees: Theory and applications. *Acta Applicandae Mathematica*, 66(3):211–249, 2001.

[44] Andrey A. Dobrynin and Ivan Gutman. The average Wiener index of trees and chemical trees. *Journal of Chemical Information and Computer Sciences*, 39(4):679–683, 1999.

[45] Alexander S. Douglas. Techniques for the recording of, and reference to data in a computer. *The Computer Journal*, 2(1):1–9, 1959.

[46] Michael Drmota. An analytic approach to the height of binary search trees. *Algorithmica*, 29(1-2):89–119, 2001.

[47] Michael Drmota. The variance of the height of digital search trees. *Acta Informatica*, 38(4):261–276, 2002.

[48] Michael Drmota. *Random Trees: An Interplay between Combinatorics and Probability.* Springer, 2009.

[49] Michael Drmota, Bernhard Gittenberger, Alois Panholzer, Helmut Prodinger, and Mark D. Ward. On the shape of the fringe of various types of random trees. *Mathematical Methods in the Applied Sciences*, 32(10):1207–1245, 2009.

[50] Michael Drmota and Wojciech Szpankowski. (un) expected behavior of digital search tree profile. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 130–138. Society for Industrial and Applied Mathematics, 2009.

[51] Michael Drmota and Wojciech Szpankowski. The expected profile of digital search trees. *Journal of Combinatorial Theory, Series A*, 118(7):1939–1965, 2011.

[52] Rosena R.-X. Du and Helmut Prodinger. Notes on protected nodes in digital search trees. *Applied Mathematics Letters*, 25(6):1025–1028, 2012.

[53] Roger C. Entringer, Amram Meir, John W. Moon, and László A. Székely. The Wiener index of trees from certain families. *Australasian Journal of Combinatorics*, 10:211–224, 1994.

[54] Arthur Erdélyi, Wilhelm Magnus, Fritz Oberhettinger, Francesco G. Tricomi, and Harry Bateman. *Higher Transcendental Functions*, volume 1. New York McGraw-Hill, 1953.

[55] Guy Fayolle, Philippe Flajolet, and Micha Hofri. On a functional equation arising in the analysis of a protocol for a multi-access broadcast channel. *Advances in Applied Probability*, pages 441–472, 1986.

[56] Guy Fayolle, Philippe Flajolet, Micha Hofri, and Philippe Jacquet. Analysis of a stack algorithm for random multiple-access communication. *IEEE Transactions on Information Theory*, 31(2):244–254, 1985.

[57] Julien Fayolle and Mark D. Ward. Analysis of the average depth in a suffix tree under a markov model. In *Discrete Mathematics and Theoretical Computer Science Proceedings, Proceedings of 2005 International Conference on Analysis of Algorithms*, pages 95–104, 2005.

[58] James A. Fill and Svante Janson. Precise logarithmic asymptotics for the right tails of some limit random variables for random trees. *Annals of Combinatorics*, 12(4):403–416, 2009.

[59] James Allen Fill, Hosam M. Mahmoud, and Wojciech Szpankowski. On the distribution for the duration of a randomized leader election algorithm. *Annals of Applied Probability*, 6:1260–1283, 1996.

[60] Philippe Flajolet. On the performance evaluation of extendible hashing and trie searching. *Acta Informatica*, 20(4):345–369, 1983.

[61] Philippe Flajolet. Approximate counting: a detailed analysis. *BIT Numerical Mathematics*, 25(1):113–134, 1985.

[62] Philippe Flajolet, Xavier Gourdon, and Philippe Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144(1):3–58, 1995.

[63] Philippe Flajolet, Rainer Kemp, and Helmut Prodinger. Average-case analysis of algorithms. *Dagstuhl Seminar Report*, 68, 1993.

[64] Philippe Flajolet and Andrew Odlyzko. Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, 3(2):216–240, 1990.

[65] Philippe Flajolet and Claude Puech. Tree structures for partial match retrieval. In *Proceedings of the 24th Annual Symposium on Foundations of Computer Science*, pages 282–288, 1983.

[66] Philippe Flajolet and Bruce Richmond. Generalized digital trees and their difference¡xdifferential equations. *Random Structures & Algorithms*, 3(3):305–320, 1992.

[67] Philippe Flajolet, Mathieu Roux, and Brigitte Vallée. Digital trees and memoryless sources: from arithmetics to analysis. In *21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA’10), Discrete Mathematics Theoretical Computer Science Proceedings*, number 01, 2010.

[68] Philippe Flajolet and Robert Sedgewick. Digital search trees revisited. *SIAM Journal on Computing*, 15(3):748–767, 1986.

[69] Philippe Flajolet and Robert Sedgewick. Mellin transforms and asymptotics: finite differences and rice's integrals. *Theoretical Computer Science*, 144(1):101–124, 1995.

[70] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[71] Philippe Flajolet and Jean-Marc Steyaert. *A branching process arising in dynamic hashing, trie searching and polynomial factorization*, volume 140 of *Lecture Notes in Computer Science*, pages 101–124. Springer-Verlag, 1982.

[72] Edward Fredkin. Trie memory. *Communications of the ACM*, 3(9):490–499, 1960.

[73] Michael Fuchs. The variance for partial match retrievals in k-dimensional bucket digital trees,. In *Discrete Mathematics and Theoretical Computer Science Proceedings, Proceedings of the 21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10)*, pages 261–275, 2010.

[74] Michael Fuchs, Hsien-Kuei Hwang, and Vytas Zacharovas. Asymptotic variance of random digital search trees. *Discrete Mathematics & Theoretical Computer Science (Special Issue in Honor of Phillipe Flajolet)*, 12(2):103–166, 2010.

[75] Michael Fuchs, Hsien-Kuei Hwang, and Vytas Zacharovas. An analytic approach to the asymptotic variance of trie statistics and related structures. *Theoretical Computer Science*, 2014.

[76] Michael Fuchs and Chung-Kuei Lee. The Wiener index of random digital trees. *Submitted for publication.*

[77] Michael Fuchs and Chung-Kuei Lee. A General Central Limit Theorem for Shape Parameters of m-ary Tries and PATRICIA Tries. *Electronic Journal of Combinatorics*, 21(1), 2014.

[78] Michael Fuchs, Chung-Kuei Lee, and Helmut Prodinger. Approximate counting via the poisson-laplace-mellin method. *Disrete Mathematics and Theoretical Computer Science Proceedings*, (01):13–28, 2012.

[79] Michael Fuchs, Chung-Kuei Lee, and Guan-Ru Yu. 2-protected nodes in random digital trees. *in preparation*.

[80] Jeffrey Gaither, Yushi Homma, Mark Sellke, and Mark D. Ward. On the number of 2-protected nodes in tries and suffix trees. *Disrete Mathematics and Theoretical Computer Science Proceedings*, (01):381–398, 2012.

[81] Jeffrey Gaither and Mark D. Ward. The variance of the number of 2-protected nodes in a trie. In *ANALCO*, pages 43–51. SIAM, 2013.

[82] Gaston H. Gonnet and Ricardo Baeza-Yates. *Handbook of Algorithms and Data Structures: in Pascal and C*. Addison-Wesley Longman Publishing Co., Inc., 1991.

[83] Gaston H. Gonnet and J. Ian Munro. The analysis of linear probing sort by the use of a new mathematical transform. *Journal of Algorithms*, 5(4):451–470, 1984.

[84] André Gronemeier and Martin Sauerhoff. Applying approximate counting for computing the frequency moments of long data streams. *Theory of Computing Systems*, 44(3):332–348, 2009.

[85] Fabrice Guillemin and Philippe Robert. Analysis of steiner subtrees of random trees for traceroute algorithms. *Random Structures & Algorithms*, 35(2):194–215, 2009.

[86] Fabrice Guillemin, Philippe Robert, Bert Zwart, et al. Aimd algorithms and exponential functionals. *Annals of Applied Probability*, 14(1):90–117, 2004.

[87] Allan Gut. *Probability: A Graduate Course*, volume 200 of *Springer Texts in Statistics*. Springer Verlag, 2005.

[88] Thomas N. Hibbard. Some combinatorial properties of certain trees with applications to searching and sorting. *Journal of the ACM*, 9(1):13–28, 1962.

[89] Friedrich Hubalek. On the variance of the internal path length of generalized digital trees–the Mellin convolution approach. *Theoretical Computer Science*, 242(1):143–168, 2000.

[90] Friedrich Hubalek, Hsien-Kuei Hwang, William Lew, Hosam M. Mahmoud, and Helmut Prodinger. A multivariate view of random bucket digital search trees. *Journal of Algorithms*, 44(1):121–158, 2002.

[91] Philippe Jacquet, Paul Muhlethaler, et al. Marginal throughtput of a stack algorithm for CSMA/CD random length packet communication when the load is over the channel efficiency. *Technical Report RR-0436, INRIA, Rocquencourt*, 1990.

[92] Philippe Jacquet and Mireille Régnier. Trie partitioning process: limiting distributions. In *CAAP'86*, pages 196–210. Springer, 1986.

[93] Philippe Jacquet and Mireille Régnier. Normal limiting distribution of the size of tries. In *Proceedings of the 12th IFIP WG 7.3 International Symposium on Computer Performance Modelling, Measurement and Evaluation*, pages 209–223. North-Holland Publishing Co., 1987.

[94] Philippe Jacquet and Mirelle Régnier. Normal limiting distribution for the size and the external path length of tries. Technical Report RR-0827, INRIA, 1988.

[95] Philippe Jacquet and Mirelle Régnier. New results on the size of tries. *IEEE Transactions on Information Theory*, 35(1):203–205, 1989.

[96] Philippe Jacquet and Wojciech Szpankowski. Analysis of digital tries with Markovian dependency. *IEEE Transactions on Information Theory*, 37(5):1470–1475, 1991.

[97] Philippe Jacquet and Wojciech Szpankowski. Autocorrelation on words and its applications: analysis of suffix trees by string-ruler approach. *Journal of Combinatorial Theory, Series A*, 66(2):237–269, 1994.

[98] Philippe Jacquet and Wojciech Szpankowski. Asymptotic behavior of the Lempel-Ziv parsing scheme and digital search trees. *Theoretical Computer Science*, 144(1):161–197, 1995.

[99] Philippe Jacquet and Wojciech Szpankowski. Analytical depoissonization and its applications. *Theoretical Computer Science*, 201(1):1–62, 1998.

[100] Philippe Jacquet, Wojciech Szpankowski, and Jing Tang. Average profile of the Lempel-Ziv parsing scheme for a Markovian source. *Algorithmica*, 31(3):318–360, 2001.

[101] Svante Janson. The Wiener index of simply generated random trees. *Random Structures & Algorithms*, 22(4):337–358, 2003.

[102] Svante Janson. Rounding of continuous random variables and oscillatory asymptotics. *Annals of Probability*, 34:1807–1826, 2006.

[103] Svante Janson. Renewal theory in the analysis of tries and strings. *Theoretical Computer Science*, 416:33–54, 2012.

[104] Svante Janson, Philippe Chassaing, et al. The center of mass of the ISE and the Wiener index of trees. *Electronic Communication in Probability*, 9:178–187, 2004.

[105] Svante Janson, Donald E. Knuth, Tomasz Łuczak, and Boris Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3):233–358, 1993.

[106] Svante Janson and Wojciech Szpankowski. Analysis of asymmetric leader election algorithm. *Electronic Journal of Combinatorics*, 64(R17):1–62, 1997.

[107] Augustus J. E. M. Janssen and Marc J. M. de Jong. Analysis of contention tree algorithms. *IEEE Transactions on Information Theory*, 46(6):2163–2172, 2000.

[108] Marc Kac. On deviations between theoretical and empirical distributions. *Proceedings of the National Academy of Sciences of the United States of America*, 35(5):252, 1949.

[109] Michael Kaplan and Eugene Gulko. Analytic properties of multiple-access trees. *IEEE Transactions on Information Theory*, 31(2):255–263, 1985.

[110] Tämur Ali Khan and Ralph Neininger. Tail bound for the Wiener index of random trees. In *Discrete Mathematics and Theoretical Computer Science Proceedings, Proceedings of the 2007 Conference on the Analysis of Algorithms*, number 01, pages 279–289, 2007.

[111] Peter Kirschenhofer and Helmut Prodinger. *Some further results on digital trees*, volume 214 of *Lecture Notes in Computer Science*, pages 177–185. Springer-Verlag, 1986.

[112] Peter Kirschenhofer and Helmut Prodinger. b-tries: a paradigm for the use of number-theoretic methods in the analysis of algorithms. *Contributions to General Algebra*, 6:141–154, 1988.

[113] Peter Kirschenhofer and Helmut Prodinger. Eine Anwendung der Theorie der Modulfunktionen in der Informatik. *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II*, 197(4-7):339–366, 1988.

[114] Peter Kirschenhofer and Helmut Prodinger. Further results on digital search trees. *Theoretical Computer Science*, 58(1):143–154, 1988.

[115] Peter Kirschenhofer and Helmut Prodinger. Approximate counting: an alternative approach. *Informatique Théorique et Applications*, 25(1):43–48, 1991.

[116] Peter Kirschenhofer and Helmut Prodinger. On some applications of formulae of ramanujan in the analysis of algorithms. *Mathematika*, 38(1):14–33, 1991.

[117] Peter Kirschenhofer, Helmut Prodinger, and Wojciech Szpankowski. Digital search trees-further results on a fundamental data structure. In *IFIP Congress*, pages 443–447, 1989.

[118] Peter Kirschenhofer, Helmut Prodinger, and Wojciech Szpankowski. On the balance property of Patricia tries: external path length viewpoint. *Theoretical Computer Science*, 68(1):1–17, 1989.

[119] Peter Kirschenhofer, Helmut Prodinger, and Wojciech Szpankowski. On the variance of the external path length in a symmetric digital trie. *Discrete Applied Mathematics*, 25(1):129–143, 1989.

[120] Peter Kirschenhofer, Helmut Prodinger, and Wojciech Szpankowski. Multidimensional digit searching and some new parameters in tries. *International Journal of Foundations of Computer Science (IJFCS)*, 4:69–84, 1993.

[121] Peter Kirschenhofer, Helmut Prodinger, and Wojciech Szpankowski. Digital search trees again revisited: The internal path length perspective. *SIAM Journal on Computing*, 23(3):598–616, 1994.

[122] Charles Knessl and Wojciech Szpankowski. Asymptotic Behavior of the Height in a Digital Search Tree and the Longest Phrase of the Lempel-Ziv Scheme. *SIAM Journal on Computing*, 30(3):923–964, 2000.

[123] Charles Knessl and Wojciech Szpankowski. On the number of full levels in tries. *Random Structures & Algorithms*, 25(3):247–276, 2004.

[124] Donald E. Knuth. Optimum binary search trees. *Acta Informatica*, 1(1):14–25, 1971.

[125] Donald E. Knuth. *The Art of Computer Programming. Volume 1: Fundamental Algorithms*. Addison Wiley Publishing Co., Third edition, 1997.

[126] Donald E. Knuth. *The Art of Computer Programming. Volume 2: Seminumerical Algorithms.* Addison Wiley Publishing Co., Third edition, 1997.

[127] Donald E. Knuth. *The Art of Computer Programming. Volume 3: Sorting and Searching.* Addison Wiley Publishing Co., Second edition, 1998.

[128] Donald E. Knuth. *The Art of Computer Programming, Volume 4A: Combinatorial Algorithms, Part 1.* Pearson Education India, 2011.

[129] Alan G. Konheim and Donald J. Newman. A note on growing binary trees. *Discrete Mathematics*, 4(1):57–63, 1973.

[130] Chung-Kuei Lee. The k-th total path length and total steiner k-distance for digital search trees. *in press.*

[131] William Lew and Hosam M. Mahmoud. The joint distribution of elastic buckets in multiway search trees. *SIAM Journal on Computing*, 23(5):1050–1074, 1994.

[132] Ernst Lindelöf. Robert Hjalmar Mellin. *Acta Mathematica*, 61(1):I–VI, 1933.

[133] Stefano Lonardi, Wojciech Szpankowski, and Mark D. Ward. Error resilient LZ'77 data compression: algorithm, analysis, and experiments. *Information Theory, IEEE Transactions on*, 53(5):1799–1813, 2007.

[134] M. Lothaire. *Applied Combinatorics on Words*, volume 105. Cambridge University Press, 2005.

[135] Guy Louchard. Exact and asymptotic distributions in digital and binary search trees. *Informatique Théorique et Applications*, 21(4):479–495, 1987.

[136] Guy Louchard. Trie size in a dynamic list structure. *Random Structures & Algorithms*, 5(5):665–702, 1994.

[137] Guy Louchard and Helmut Prodinger. Asymptotics of the moments of extreme-value related distribution functions. *Algorithmica*, 46(3-4):431–467, 2006.

[138] Guy Louchard and Helmut Prodinger. Generalized approximate counting revisited. *Theoretical Computer Science*, 391(1):109–125, 2008.

[139] Hosam M. Mahmoud. *Evolution of Random Search Tree.* Wiley-Interscience, 1992.

[140] Hosam M. Mahmoud, Philippe Flajolet, Philippe Jacquet, and Mireille Régnier. Analytic variations on bucket selection and sorting. *Acta Informatica*, 36(9-10):735–760, 2000.

[141] Hosam M. Mahmoud and Boris Pittel. Analysis of the space of search trees under the random insertion algorithm. *Journal of Algorithms*, 10(1):52–75, 1989.

[142] Hosam M. Mahmoud and Mark D. Ward. Asymptotic distribution of two-protected nodes in random binary search trees. *Applied Mathematics Letters*, 25(12):2218–2222, 2012.

[143] Hosam M. Mahmoud, Mark D. Ward, et al. Average-case analysis of cousins in m-ary tries. *Journal of Applied Probability*, 45(3):888–900, 2008.

[144] Toufik Mansour. Protected points in k-ary trees. *Applied Mathematics Letters*, 24(4):478–480, 2011.

[145] James L. Massey. Collision-resolution algorithms amd random-access communications. In *Multi-User Communication Systems*, pages 73–137, 1981.

[146] Collin McDiarmid. *On the method of bounded differences*, volume 141 of *London Mathematical Society Lecture Notes Series - Survey in Combinatorics*, pages 148–188. Cambridge University Press, 1989.

[147] Scott A. Mitchell and David M. Day. Flexible approximate counting. In *Proceedings of the 15th Symposium on International Database Engineering & Applications*, pages 233–239, 2011.

[148] Kate Morris, Alois Panholzer, and Helmut Prodinger. On some parameters in heap ordered trees. *Combinatorics Probability & Computing*, 13(4-5):677–696, 2004.

[149] Robert Morris. Counting large numbers of events in small registers. *Communications of the ACM*, 21(10):840–842, 1978.

[150] Donald R. Morrison. PATRICIA - Practical Algorithm To Retrieve Information Coded In Alphanumeric. *Journal of the ACM (JACM)*, 15(4):514–534, 1968.

[151] Yiannis N. Moschovakis. What is an algorithm. *Mathematics Unlimited–2001 and Beyond*, pages 919–936, 2001.

[152] Götz O. Munsonius. The total Steiner k-distance for b-ary recursive trees and linear recursive trees. In *Discrete Mathematics and Theoretical Computer Science Proceedings, Proceedings of the 21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10)*, pages 529–550, 2010.

[153] Götz O. Munsonius. On the asymptotic internal path length and the asymptotic Wiener index of random split trees. *Electronic Journal of Probability*, 16:1020–1047, 2011.

[154] Götz O. Munsonius et al. On tail bounds for random recursive trees. *Journal of Applied Probability*, 49(2):566–581, 2012.

[155] Jean-Fréderic Myoupo, Loys Thimonier, and Vlady Ravelomanana. Average case analysis-based protocols to initialize packet radio networks. *Wireless Communications and Mobile Computing*, 3(4):539–548, 2003.

[156] Markus E. Nebel. On the horton-strahler number for combinatorial tries. *RAIRO-Theoretical Informatics and Applications*, 34(04):279–296, 2000.

[157] Markus E. Nebel. The Stack-Size of Combinatorial Tries Revisited. *Discrete Mathematics & Theoretical Computer Science*, 5(1):1–16, 2002.

[158] Markus E. Nebel. The stack-size of tries: a combinatorial study. *Theoretical Computer Science*, 270(1):441–461, 2002.

[159] Ralph Neininger. The Wiener index of random trees. *Combinatorics, Probability & Computing*, 11(6):587–597, 2002.

[160] Ralph Neininger and Ludger Rüschendorf. A general limit law for recursive algorithms and combinatorial structures. *Annals of Applied Probability*, 14(1):378–418, 2004.

[161] Ralph Neininger and Ludger Rüschendorf. On the contraction method with degenerate limit equation. *Annals of Probability*, 32(3B):2838–2856, 2004.

[162] Michel Nguyên-Thê. *Distribution de valuations sur les arbres.* PhD thesis, LIX, Ecole polytechnique, 2003.

[163] Frank W. J. Olver. *Asymptotics and Special Functions.* Akademic Press., 1974.

[164] Alois Panholzer. The distribution of the size of the ancestor-tree and of the induced spanning subtree for random trees. *Random Structures & Algorithms*, 25(2):179–207, 2004.

[165] Alois Panholzer. Distribution of the steiner distance in generalized M-ary search trees. *Combinatorics Probability and Computing*, 13(4-5):717–733, 2004.

[166] Alois Panholzer and Helmut Prodinger. Analysis of some statistics for increasing tree families. *Discrete Mathematics & Theoretical Computer Science*, 6(2):437–460, 2004.

[167] Gahyun Park, Hsien-Kuei Hwang, Pierre Nicodeme, and Wojciech Szpankowski. Profiles of tries. *SIAM Journal on Computing*, 38(5):1821–1880, 2009.

[168] Boris Pittel. Asymptotical growth of a class of random trees. *Annals of Probability*, 18:414–427, 1985.

[169] Boris Pittel. Paths in a random digital tree: Limiting distributions. *Advances in Applied Probability*, pages 139–155, 1986.

[170] Boris Pittel. On the height of patricia search tree. In *ORSA/TIMS Special Interest Conference on Applied Probability in the Engineering*, Monterey, CA., 1991.

[171] Boris Pittel and Herman Rubin. How many random questions are necessary to identify n distinct objects? *Journal of Combinatorial Theory, Series A*, 55(2):292–312, 1990.

[172] Franco P. Preparatat and Michael I. Shamos. *Computational Geometry: An Introduction.* Springer-Verlag, 1985.

[173] Helmut Prodinger. Hypothetical analyses: approximate counting in the style of Knuth, path length in the style of Flajolet. *Theoretical Computer Science*, 100(1):243–251, 1992.

[174] Helmut Prodinger. How to select a loser. *Discrete Mathematics*, 120(1):149–159, 1993.

[175] Helmut Prodinger. Approximate counting via euler transform. *Mathematica Slovaca*, 44(5):569–574, 1994.

[176] Helmut Prodinger. Digital search trees and basic hypergeometric functions. *Bulletin-European Association for Theoretical Computer Science*, 56:112–112, 1995.

[177] Helmut Prodinger. Digital search trees with m trees: level polynomials and insertion costs. *Discrete Mathematics & Theoretical Computer Science*, 13(3):1–8, 2011.

[178] Helmut Prodinger. Approximate counting with m counters: a detailed analysis. *Theoretical Computer Science*, 439:58–68, 2012.

[179] Hans Jürgen Prömel and Angelika Steger. *The Steiner Tree Problem. A Tour Through Graphs, Algorithms and Complexity.* Vieweg Verlag, Wiesbaden, 2002.

[180] Svetlozar T. Rachev and Ludger Rüschendorf. Probability metrics and recursive algorithms. *Advances in Applied Probability*, 27(3):770–799, 1995.

[181] Bonita Rais, Philippe Jacquet, and Wojciech Szpankowski. A limiting distribution for the depth in Patricia tries. *SIAM Journal on Discrete Mathematics*, 6(2):197–213, 1993.

[182] Philippe Robert. On the asymptotic behavior of some algorithms. *Random Structures & Algorithms*, 27(2):235–250, 2005.

[183] Uwe Roesler and Ludger Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29(1-2):3–33, 2001.

[184] Walter A. Rosenkrantz. Approximate counting: a martingale approach. *Stochastics: An International Journal of Probability and Stochastic Processes*, 20(2):111–120, 1987.

[185] Uwe Rösler. A limit theorem for Quicksort. *RAIRO Theoretical Informatics and Applications*, 25:85–100, 1991.

[186] Uwe Rösler. A fixed point theorem for distributions. *Stochastic Processes and their Applications*, 42(2):195–214, 1992.

[187] Uwe Rösler. On the analysis of stochastic divide and conquer algorithms. *Algorithmica*, 29(1-2):238–261, 2001.

[188] Ludger Rüschendorf and Ralf Neininger. Survey of multivariate aspects of the contraction method. *Discrete Mathematics & Theoretical Computer Science*, 8(1), 2006.

[189] Werner Schachinger. On the variance of a class of inductive valuations of data structures for digital search. *Theoretical Computer Science*, 144(1):251–275, 1995.

[190] Werner Schachinger. The variance of a partial match retrieval in a multidimensional symmetric trie. *Random Structures & Algorithms*, 7(1):81–95, 1995.

[191] Werner Schachinger. Limiting distributions for the costs of partial match retrievals in multidimensional tries. *Random Structures & Algorithms*, 17(3-4):428–459, 2000.

[192] Werner Schachinger. Asymptotic normality of recursive algorithms via martingale difference arrays. *Discrete Mathematics & Theoretical Computer Science*, 4(2):363–398, 2001.

[193] Robert Sedgewick and Philippe Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, 2013.

[194] Shyue-Horng Shiau and Chang-Biau Yang. A fast initialization algorithm for single-hop wireless networks. *IEICE Transactions on Communications*, 88(11):4285–4292, 2005.

[195] Klaus Simon. An improved algorithm for transitive closure on acyclic digraphs. *Theoretical Computer Science*, 58(1):325–346, 1988.

[196] Wojciech Szpankowski. *Average complexity of additive properties for multiway tries: a unified approach*, volume 249 of *Lecture Notes in Computer Science*, pages 13–25. Springer-Verlag, 1987.

[197] Wojciech Szpankowski. The evaluation of an alternating sum with applications to the analysis of some data structure. *Information Processing Letter*, 28(1):13–19, 1988.

[198] Wojciech Szpankowski. Some results on v-ary asymmertic tries. *Journal of Algorithms*, 9(2):224–244, 1988.

[199] Wojciech Szpankowski. Patricia tries again revisited. *Journal of the ACM*, 37(4):691–711, 1990.

[200] Wojciech Szpankowski. A characterization of digital search trees from the successful search viewpoint. *Theoretical Computer Science*, 85(1):117–134, 1991.

[201] Wojciech Szpankowski. On the height of digital trees and related problems. *Algorithmica*, 6(1-6):256–277, 1991.

[202] Wojciech Szpankowski. Analysis of Algorithms (AofA) part I: 1993–1998 ("Dagstuhl Period"). *Current Trends in Theoretical Computer Science: Algorithms and Complexity*, 1:39, 2004.

[203] Wojciech Szpankowski. *Average Case Analysis of Algorithms*. Chapman & Hall/CRC, 2010.

[204] Wojciech Szpankowski and Charles Knessl. *Height in generalized tries and PATRICIA tries*, volume 1776 of *Lecture Notes in Computer Science*, pages 298–307. 2000.

[205] Boris Solomonovich Tsybakov and Viktor Alexandrovich Mikhaĭlov. Free synchronous packet access in a broadcast channel with feedback. *Problemy Peredachi Informatsii*, 14(4):32–59, 1978.

[206] Brigitte Vallée. Dynamical sources in information theory: Fundamental intervals and word prefixes. *Algorithmica*, 29(1-2):262–306, 2001.

[207] Stephan G. Wagner. A class of trees and its Wiener index. *Acta Applicandae Mathematica*, 91(2):119–132, 2006.

[208] Stephan G. Wagner. On the average Wiener index of degree-restricted trees. *Australasian Journal of Combinatorics*, 37:187–203, 2007.

[209] Stephan G. Wagner. On the Wiener index of random trees. *Discrete Mathematics*, 312(9):1502–1511, 2012.

[210] Stephen G. Wagner. On unary nodes in tries. In *Discrete Mathematics and Theoretical Computer Science Proceedings, Proceedings of the 21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10)*, pages 577–589, 2010.

[211] Mark D. Ward. *Analysis of the multiplicity matching parameter in suffix trees*. PhD thesis, Purdue University, 2005.

[212] Mark D. Ward and Wojciech Szpankowski. Analysis of randomized selection algorithm motivated by the LZ'77 scheme. In *The First Workshop on Analytic Algorithmics and Combinatorics (ANALCO 04)*, New Orlean, 2004.

[213] Mark D. Ward and Wojciech Szpankowski. Analysis of the multiplicity matching parameter in suffix trees. In *International Conference on Analysis of Algorithms, Disrete Mathematics and Theoretical Computer Science Proceedings AD*, volume 307, page 322, 2005.

[214] Mark D. Ward and Wojciech Szpankowski. Analysis of the multiplicity matching parameters in suffix trees. In *Discrete Mathematics and Theoretical Computer Science Proceedings, Proceedings of 2005 International Conference on Analysis of Algorithms*, pages 307–322, 2005.

[215] Harry Wiener. Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69(1):17–20, 1947.

[216] Peter F. Windley. Trees, forests and rearranging. *The Computer Journal*, 3(2):84–88, 1960.