33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms

AofA 2022, June 20–24, 2022, Philadelphia, PA, USA

Edited by

Mark Daniel Ward

Editors

Mark Daniel Ward
Purdue University, West Lafayette, IN, USA
mdw@purdue.edu

ACM Classification 2012
Mathematics of computing; Theory of computation

ISBN 978-3-95977-230-3

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at https://www.dagstuhl.de/dagpub/978-3-95977-230-3.

Publication date
June, 2022

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at https://portal.dnb.de.

Digital Object Identifier: 10.4230/LIPIcs.AofA.2022.0

LIPIcs – Leibniz International Proceedings in Informatics

LIPIcs is a series of high-quality conference proceedings across all fields in informatics. LIPIcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

ISSN 1868-8969

https://www.dagstuhl.de/lipics

Dedicated to the memory of Philippe Flajolet.

Contents

33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022).
Editor: Mark Daniel Ward
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Preface

The 33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022) was held in Philadelphia, PA, USA, at the University of Pennsylvania, during June 20–24, 2022.

Analysis of algorithms is a scientific basis for computation, providing a link between abstract algorithms and the performance characteristics of their implementations in the real world. The general effort to predict precisely the performance of algorithms has come to involve research in analytic combinatorics, the analysis of random discrete structures, asymptotic analysis, exact and limiting distributions, and other fields of inquiry in computer science, probability theory, and enumerative combinatorics.

See `http://aofa.cs.purdue.edu/` for more details.

The Call for Papers invited papers in

- analytic algorithmics and combinatorics,
- probabilistic analysis of algorithms,
- randomized algorithms.

We also welcomed papers addressing problems such as: combinatorial algorithms, string searching and pattern matching, sublinear algorithms on massive data sets, network algorithms, graph algorithms, caching and memory hierarchies, indexing, data mining, data compression, coding and information theory, and computational finance. Papers were also welcomed that address bridges to research in related fields such as statistical physics, computational biology, computational geometry, and simulation.

Authors of selected accepted extended abstracts will be invited to submit full papers for peer review to a special issue of a journal.

Mark Daniel Ward,
on behalf of the Program and Steering Committees

33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022).
Editor: Mark Daniel Ward
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Flajolet Lecture

The Philippe Flajolet Lecture Prize for outstanding contributions to analytic combinatorics and analysis of algorithms is awarded every two years by the Analysis of Algorithms (AofA) community – a community that owes its existence to Philippe Flajolet. The first Flajolet Lecture was presented by Donald E. Knuth at the 25th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms in 2014 in Paris, France. The second Flajolet Lecture was presented by Robert Sedgewick at the 27th AofA Conference in 2016 in Krakow, Poland. The third Flajolet Lecture was presented by Luc Devroye at the 29th AofA Conference in 2018 in Uppsala, Sweden.

The fourth Flajolet Lecture was scheduled to be presented by Wojciech Szpankowski at the 31st AofA Conference in 2020, but the lecture was delayed for two years, due to the COVID-19 pandemic. Therefore, at this year's conference, he presented the fourth Flajolet Lecture, entitled "Analytic Information and Learning Theory: From Compression to Learning." Svante Janson also presented the fifth Flajolet Lecture, entitled "The Sum of Powers of Subtree Sizes for Random Trees."

The prize is named in honor and recognition of the extraordinary accomplishments of the late Philippe Flajolet, who spent most of his scientific life at INRIA, France. Philippe is best known for fundamental advances in mathematical methods for the analysis of algorithms. His research laid the foundation of a subfield of mathematics now known as analytic combinatorics. Analytic combinatorics is a modern basis for the quantitative study of combinatorial structures (such as words, trees, mappings, and graphs), with applications to probabilistic study of algorithms that are based on these structures. It also strongly influences research in other scientific domains, such as statistical physics, computational biology, and information theory. Flajolet's work takes the field forward by introducing original approaches in combinatorics based on two types of methods: symbolic and analytic. The symbolic side is based on the automation of decision procedures in combinatorial enumeration to derive characterizations of generating functions. The analytic side treats those functions as functions in the complex plane and leads to precise characterization of limit distributions. Beyond these foundational contributions, Philippe's research opened new avenues in various domains of applied computer science, including streaming algorithms, communication protocols, database access methods, data mining, symbolic manipulation, text-processing algorithms, and random generation.

33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022).
Editor: Mark Daniel Ward

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Program Committee

- Luc Devroye, McGill Univ., Canada
- Amalia Duch, Univ. Politècnica de Catalunya, Spain
- Cecilia Holmgren, Uppsala Univ., Sweden
- Manuel Lladser, Univ. of Colorado at Boulder, USA
- Cécile Mailler, Univ. of Bath, UK
- Steve Melczer, Univ. of Waterloo, Canada
- Marni Mishna, Simon Fraser Univ., Canada
- Noela Müller, TU Eindhoven, Netherlands
- Ralph Neininger, Goethe-Univ. Frankfurt, Germany
- Cyril Nicaud, Univ. Gustave Eiffel, France
- Robin Pemantle, Univ. of Pennsylvania, USA
- Carine Pivoteau, Univ. Paris-Est, Marne-la-Vallée, France
- Mark Daniel Ward, Purdue Univ., USA (Chair)
- Sebastian Wild, Univ. of Liverpool, UK

33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022).
Editor: Mark Daniel Ward
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Steering Committee

- Frédérique Bassino, Univ. Sorbonne Paris Nord, France
- James Allen Fill, Johns Hopkins Univ., USA
- Clemens Heuberger, Alpen-Adria-Univ. Klagenfurt, Austria
- Hsien-Kuei Hwang, Academia Sinica, Taiwan
- Ralph Neininger, Goethe-Univ. Frankfurt, Germany
- Bruno Salvy (chair), INRIA, Lyon, France

33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022).
Editor: Mark Daniel Ward

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Building Sources of Zero Entropy: Rescaling and Inserting Delays

Ali Akhavi
GREYC, Université de Caen-Normandie, France
LIPN, Université Sorbonne Paris Nord, France

Fréderic Paccaut
LAMFA CNRS UMR 7352, Université de Picardie Jules Verne, France

Brigitte Vallée
GREYC, CNRS and Université de Caen-Normandie, France

Abstract

Most of the natural sources that intervene in Information Theory have a positive entropy. They are well studied. The paper aims in building, in an explicit way, natural instances of sources with zero entropy. Such instances are obtained by slowing down sources of positive entropy, with processes which rescale sources or insert delays. These two processes – rescaling or inserting delays – are essentially the same; they do not change the fundamental intervals of the source, but only the "depth" at which they will be used, or the "speed" at which they are divided. However, they modify the entropy and lead to sources with zero entropy. The paper begins with a "starting" source of positive entropy, and uses a natural class of rescalings of sublinear type. In this way, it builds a class of sources of zero entropy that will be further analysed. As the starting sources possess well understood probabilistic properties, and as the process of rescaling does not change its fundamental intervals, the new sources keep the memory of some important probabilistic features of the initial source. Thus, these new sources may be thoroughly analysed, and their main probabilistic properties precisely described. We focus in particular on two important questions: exhibiting asymptotical normal behaviours à la Shannon-MacMillan-Breiman; analysing the depth of the tries built on the sources. In each case, we obtain a parameterized class of precise behaviours. The paper deals with the analytic combinatorics methodology and makes a great use of generating series.

2012 ACM Subject Classification Theory of computation → Pattern matching; Theory of computation → Randomness, geometry and discrete structures

Keywords and phrases Information Theory, Probabilistic analysis of sources, Sources with zero-entropy, Analytic combinatorics, Dirichlet generating functions, Transfer operator, Trie structure, Continued fraction expansion, Rice method, Quasi-power Theorem

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.1

Category Invited Talk

Funding This work was supported by the Agence Nationale de la Recherche through the project CODYS (ANR-18-CE40-0007).

1 Introduction

General context. A source is one of the main objects of Information Theory. A source is a probabilistic process which emits a digit from a given alphabet Σ, one at each discrete time. Very often, a source $\mathcal{P}$ is defined on the unit interval $\mathcal{I}$ and associates with $x \in \mathcal{I}$ an infinite word $M(x) = (a_1, a_2, \ldots, a_n, \ldots)$ where the successive symbols $a_i = a_i(x)$ belong to Σ. This infinite word is the expansion of x in "base" $\mathcal{P}$. Some sources are directly defined in relation with some concepts of Information Theory (memoryless sources, Markov chains), whereas other ones are related to objects of Number Theory, via numeration systems, for instance.

When both the input x and a depth k are fixed, one considers the reals y for which the word $M(y)$ has the same prefix of length k as $M(x)$. In a quite general setting, this defines an interval denoted as $\mathcal{I}_k(x)$ of length $I_k(x)$. When the unit interval is endowed with some density, and for any fixed k, the random variable $x \mapsto -\log I_k(x)$ plays a central role. The source admits the entropy h (in Shannon's meaning) when $\mathbb{E}[-\log I_k(x)] \sim h \cdot k$. Most of the classical sources (in Information Theory contexts) have a finite entropy.

Here, we focus on sources of zero entropy. Adapting the philosophy described in [3], we say that a source admits a Shannon weight f if $\mathbb{E}[-\log I_k(x)] \sim f(k)$ (when $k \to \infty$). In fact, we deal with a slightly stronger notion, and consider the notion of exponential weight (see Definition 3). Papers [2] and [3] study various sources of zero entropy that arise in Number Theory contexts, and notably two sources, the Stern-Brocot source, and the Sturm source. In [2], the authors analyze the tries built on these two sources; they exhibit in [3] their weight (however not à la Shannon, see definition 2.1 in [3]), and prove that the weight of the Stern-Brocot source is of order $\Theta(k/\log k)$ whereas the weight of the Sturm Source is of order $\Theta(\log k)$. There do not appear, in this Number Theory context, other sublinear weights, for instance of square root type $\sqrt{k}$.

Rescaling or inserting delays. We wish to build a class of sources which appear in a natural way, and admit weights (à la Shannon) of various sublinear growth. We use rescaling processes, that do not really change the source –namely the intervals $\mathcal{I}_k(x)$– but modify the depth where such an interval will be used. Here, rescaling processes, via rescaling functions g, "slow down" an initial source $\mathcal{P}$ of positive entropy and transform it into a source $\mathcal{P}_{\langle g \rangle}$ of zero entropy. Another intuitive process, for slowing down the source, consists in inserting in $\mathcal{P}$ waiting times γ (called here delays) between two symbols of $\mathcal{P}$. One obtains in this way a source $\mathcal{P}^{\langle \gamma \rangle}$ with "delays". The two processes coincide for a convenient choice of the pair (g, γ), and provide a double point of view that will be used in the sequel of the paper.

Our strategy is as follows: We first choose as initial sources $\mathcal{P}$ three types of sources in our favorite set (see Definition 13), and a "natural" class of delays, described with two parameters (a, b) and defined in Eqn (18); we then insert delays $\gamma_{a,b}$ inside each source $\mathcal{P}$ and obtain a class of sources $\mathcal{P}_{a,b}$. The Shannon weights $g_{a,b}$ of $\mathcal{P}_{a,b}$ –except for $(a, b) = (1, 0)$, where we recover the initial source – are all sublinear (and may be of various types: see Proposition 15). As we start from the same initial source $\mathcal{P}$, with nice probabilistic properties, and the rescaling process does not modify the fundamental intervals of the source, (only the depth at which they are used), we expect the source $\mathcal{P}_{a,b}$ to keep the memory of some important features of the source $\mathcal{P}$. Indeed, as the present paper shows it, such a source may be precisely analysed, and its main probabilistic properties exhibited; we focus on two phenomena: the precise behaviour of the function $x \mapsto \log I_k(x)$ (à la Shannon-MacMillan-Breimann), and the probabilistic analysis of the tries built on words emitted from $\mathcal{P}_{a,b}$.

Methods and results. Our methods are inspired by analytic combinatorics described in the book of Flajolet and Sedgewick [15]. In this context of Information Theory, the generating series of a source introduced in [22] are Dirichlet series, here called the Λ series of the source, and may be defined and a priori used for any source: paper [22] exhibits their importance in the analysis of the main probabilistic properties of the source. Later on, various papers [6, 4, 12], using the Rice method, relate the analysis of the shape of a trie built on the source with the $\Lambda(s)$ series. However, until now, except in [2], these series were only used in the context of sources of positive entropy.

Here, we use these Λ series in the wider context of sources with a possible zero entropy. We start with sources of positive entropy (gathered in a favorite set), and we distort the starting sources into sources of zero entropy. This distortion is translated into a relation between the two Λ series –attached respectively to the new source, and to the initial source– (see Proposition 8). The Λ series of sources in our favorite set are precisely described, notably from the tameness point of view (see Definition 12 and Proposition 14). Using a delay $\gamma_{a,b}$ first modifies the nature of the dominant pole of the initial $\Lambda(s)$: the parameter a moves its location, and the parameter b increases its order (See Proposition 17). Theorem 18 then describes some sufficient conditions – both on the initial tameness and parameters (a, b))– under which the tameness of the new source $\mathcal{P}_{a,b}$ may be proven.

We obtain two types of results. A first result (see Theorem 4 and Lemma 16) deals with the lengths $J_k(x)$ of fundamental intervals of the new source and proves phenomena à la Shannon-MacMillan-Breimann, described by the asymptotic normality of $x \mapsto \log J_k(x)$, adjusted with the "speed" $g_{a,b}(k)$ attached to the delay $\gamma_{a,b}$. This result is a straightforward consequence of the Quasi-Power theorem of Hwang [16]. A second result describes the expectation of the trie depth and is more subtle. Theorem 5 first relates, in a very general context, the tameness of a source (of possible zero entropy) and the expectation $\mathbb{E}[D_n]$ of its trie depth. This result (not really new) gathers many various results, some of them being classical, and some of them having not been yet considered. Using this general result and applying it to our sources $\mathcal{P}_{a,b}$ provides in Theorem 19 the analysis of the average depth $\mathbb{E}[D_n]$ for tries built on $\mathcal{P}_{a,b}$. As previously, the parameters (a, b) intervene in the behaviour of the trie, with two regimes, respectively obtained for $a = 1$ and $a > 1$.

Plan of the paper. Section 2 describes the general context. Section 3 explains the two slowing down processes which transform a source with positive entropy into a source of zero entropy. Section 4 describes the favorite set which gathers all the sources that will be used as starting points. Finally, Section 5 describes the class of sources $\mathcal{P}_{a,b}$ and exhibits their main probabilistic properties (asymptotic log-normality à la Shannon-MacMillan-Breimann and estimates for the average depth trie).

2 Sources, weights, generating series

We first describe a general source $\mathcal{Q}$ –of possible zero entropy– related to partitions in Section 2.1 and introduce its generating functions. The following sections then explain how these series intervene in two questions of interest: the asymptotic normality of $x \mapsto \log I_k(x)$ relative to the source $\mathcal{Q}$ (in Section 2.2) and the asymptotic behaviour of the average trie depth, when the trie is built on the words emitted by the source $\mathcal{Q}$ (in Section 2.3).

2.1 General sources associated with partitions

A source $\mathcal{Q}$ is a probabilistic process which emits a digit from a given alphabet Σ of cardinality r (possibly infinite denumerable), one at each discrete time. A source $\mathcal{Q}$ is very often defined on the unit interval $\mathcal{I}$ and associates with $x \in \mathcal{I}$ an infinite word $M(x) = (a_1, a_2, \ldots, a_n, \ldots)$ where the successive digits $a_i = a_i(x)$ belong to Σ. This infinite word is called the expansion of x in "base" $\mathcal{Q}$. Some sources are directly defined in relation with some concepts of Information Theory (memoryless sources, Markov chains), whereas other ones are related to objects of Number Theory, via numeration systems, for instance.
For a fixed depth k, one considers the reals y for which the word $M(y)$ begins with a given prefix w of length k. When both the input x and a depth k are fixed, one also considers the reals y for which the word $M(y)$ has the same prefix of length k as $M(x)$. In a quite general

setting[1], this defines (up to a denumerable subset of $\mathcal{I}$) two families of intervals, respectively denoted as $\mathcal{I}_w$ and $\mathcal{I}_k(x)$. The interval $\mathcal{I}_w$ for $w \in \Sigma^\star$, of length p_w, is the fundamental interval associated with the prefix w; its length p_w is the probability that the word $M(x)$ begins with prefix w; the interval $\mathcal{I}_k(x)$, of length $I_k(x)$, is the k-th coincidence interval of x. When the unit interval is endowed with some density, the mapping $x \mapsto \log I_k(x)$ is, for any fixed k, a random variable of great interest, whose asymptotics (when $k \to \infty$) is widely studied. In particular, the paper [3] defines two notions of weights (related to the asymptotics of this random variable) that extend the notions of entropy (almost everywhere, in probability) introduced in [9].
Here, we consider two weights of different flavour: first, now, the Shannon weight; then, in the next Section, the exponential weight.

▶ **Definition 1.** *A source $\mathcal{Q}$ has a Shannon weight $g(k) \geq 0$, and, resp. a Shannon entropy $h > 0$, if the sequence $\mathbb{E}[-\log I_k]$ satisfies the following (for $k \to \infty$),*

$$(1/g(k))\,\mathbb{E}[-\log I_k] \to 1\,, \qquad (1/k)\,\mathbb{E}[-\log I_k] \to h\,.$$

The notion of Shannon weight thus extends the Shannon entropy, and sources with zero Shannon entropy have a sublinear Shannon weight $g(k)$ for which $g(k)/k \to 0$. The second weight (exponential weight) is described via the generating series of the source. These generating series were first introduced and used in [22].

▶ **Definition 2.** *For $s \in \mathbb{C}$, the Λ generating series of the source involve the fundamental probabilities p_w that a word begins with a prefix w, for a given depth k, or for all depths k,*

$$\Lambda_k(s) = \sum_{w\in\Sigma^k} p_w^s, \qquad \Lambda(s) := \sum_{w\in\Sigma^\star} p_w^s = \sum_{k\geq 0} \Lambda_k(s)\,. \tag{1}$$

These generating series are Dirichlet series that satisfy $|\Lambda_k(s)| \leq \Lambda_k(\Re s)$, $|\Lambda(s)| \leq \Lambda(\Re s)$.

2.2 Exponential weight and asymptotic normality of $\log I_k(x)$.

The exponential weight is defined via the (possible) quasi-power behaviour of the $\Lambda_k(s)$ series. We will relate it later with the Shannon weight (Theorem 4*a*).

▶ **Definition 3.** *The source $\mathcal{Q}$ has an exponential weight f if its Λ_k series satisfies the following: There exist a real number A (with $-\infty \leq A < 0$) and two analytic functions u and v defined on a complex neighborhood $\mathcal{V}$ of the real half-line $s > A$, for which, for any real $c > A$, there exists a complex neighborhhod $\mathcal{V}_c$ of c and a function ϵ_c for which the following estimate for Λ_k,*

$$\Lambda_k(s) = v(s)\cdot u(s)^{f(k)}\cdot[1 + O(\epsilon_c(k))]\,, \qquad \epsilon_c(k) \to 0 \quad (k \to \infty) \tag{2}$$

holds (uniformly) for $s \in \mathcal{V}_c$. The function u is called a base function.

The series $\Lambda_k(s)$ is closely related to the moment generating function $M_k(s)$ of the variable $(-\log I_k(x))$. Indeed, as the random variable $x \mapsto I_k(x)$ is a staircase function that has value p_w on the interval fundamental $\mathcal{I}_w$ ($w \in \mathcal{A}^k$), the moment generating function $M_k(s) = \mathbb{E}[I_k^{-s}(x)]$ is expressed in terms of the family (p_w) for $w \in \mathcal{A}^k$,

$$M_k(s) = \mathbb{E}[I_k^{-s}(x)] = \sum_{w\in\mathcal{A}^k} p_w^{-s}\cdot p_w = \sum_{w\in\mathcal{A}^k} p_w^{1-s} = \Lambda_k(1-s)\,. \tag{3}$$

[1] This is the case when there is an underlying sequence of partitions, in the sense of [9] or [2]. This will be the case here (Section 4).

The importance of exponential weight is related to the Quasi-Power Theorem due to Hwang [16] that deals with the moment generating function, when it has a quasi-power form as in (2). Due to Relation (3), it leads to an asymptotic Gaussian law for the mapping $x \mapsto \log I_k(x)$. The occurrence of a Gaussian law is a refinement of the Shannon-MacMillan-Breimann Property, which usually deals with the almost everywhere behaviour of $\log I_k(x)$.

▶ **Theorem 4.** *Consider a source $\mathcal{Q}$ and the length $I_k(x)$ of its k-th coincidence intervals. If the source $\mathcal{Q}$ admits an exponential weight f with base u, the random variable $x \mapsto \log I_k(x)$ satisfies the following:*

(a) *Its expectation and its variance admit the following asymptotic estimates for $k \to \infty$,*

$$\mathbb{E}[\log I_k(x)] = u'(1) \cdot f(k) + O(1), \quad \mathrm{Var}[\log I_k(x)] = (u''(1) - u'(1)^2) \cdot f(k) + O(1)\,.$$

In particular, the source $\mathcal{Q}$ has a Shannon weight equal to $|u'(1)| \cdot f(k)$.

(b) *If the function $u(s)$ is strictly log-convex, the variable $x \mapsto \log I_k(x)$ asymptotically follows a gaussian law*

$$\Pr\left[x \left| \frac{\log I_k(x) - u'(1) f(k)}{[(u''(1) - u'(1)^2) f(k)]^{1/2}} \le x \right.\right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt\,.$$

Usually, classical applications of the Quasi-Power Theorem exhibit asymptotic gaussian laws for which the expectation and the variance are (most of the time)[2] of order $\Theta(k)$ or $\Theta(\log k)$ (See [16]). The present paper will provide natural instances of applications of the Quasi-Power Theorem with various behaviours of the expectation and the variance (See Section 5.2 and Theorem 16).

2.3 Role of the Λ series in the analysis of tries

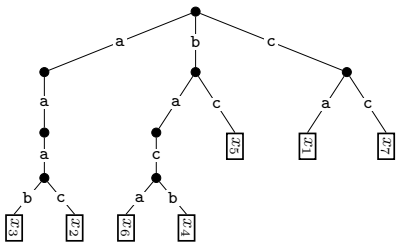
We are also interested in a probabilistic analysis of the shape of a trie $\mathtt{Trie}(\boldsymbol{x})$ built on a sequence $\boldsymbol{x}$ of (infinite) words that are independently drawn from the source $\mathcal{Q}$ of alphabet Σ. $\mathtt{Trie}(\boldsymbol{x})$ is a tree that is recursively defined via the cardinality $N(\boldsymbol{x})$ of the sequence $\boldsymbol{x}$:

(a) If $N(\boldsymbol{x}) = 0$, then $\mathtt{Trie}(\boldsymbol{x}) = \emptyset$;
(b) If $N(\boldsymbol{x}) = 1$, with $\boldsymbol{x} = (x)$, then $\mathtt{Trie}(\boldsymbol{x})$ is a leaf labeled by x;
(c) If $N(\boldsymbol{x}) \ge 2$, then $\mathtt{Trie}(\boldsymbol{x})$ is formed with an internal node and r subtries[a] resp. equal to

$$\mathtt{Trie}(\boldsymbol{x}_{\langle 0\rangle}), \ldots, \mathtt{Trie}(\boldsymbol{x}_{\langle r-1\rangle})\,,$$

where $\boldsymbol{x}_{\langle\sigma\rangle}$ denotes the sequence consisting of words of $\boldsymbol{x}$ which begin with symbol σ, stripped of their initial symbol σ.

[a] r is the cardinality of the alphabet Σ.



Here, we focus on a particular parameter of $\mathtt{Trie}(\boldsymbol{x})$, defined as the length of a *random* branch and called the trie depth. When $N(\boldsymbol{x}) = n$, it is denoted by $D_n(\boldsymbol{x})$.

[2] There is a notable exception in the study of partitions [17] where expectations and variance of order k^d for any $d \in]0, 1[$ occur.

The trie is central in text algorithmics [11]. The study of its geometric parameters, notably its depth, is thus an important subject in analysis of algorithms, where two different strategies have been introduced: the paper [13] uses the Rice method, whereas the book [21] is based on Depoissonization tools. (See [23] for a comparison of these methodologies). Here, we choose the Rice method, which was introduced by Nörlund [18], [19] and widely used in analytic combinatorics since the seminal papers of Flajolet and Sedgewick [13] [14]. The Rice method is indeed dedicated to the study of sequences of the form Eqn (4), which exhibit the main role played by the function $s\Lambda(s)$, provided it *be tame at* $s = c$ for $c < 2$.

Tameness is a central notion in analysis of algorithms which has been only recently introduced in various works (see [4], [7], [5], [2] for instance). It is defined by a position, an order and a shape. Tameness will be precisely described later in Definition 12. For the moment, we give an informal definition:

The function $\Lambda(s)$*, that is analytic on* $\{\Re s > c\}$ *is* tame *at* $s = c$*, with order* $b \geq 0$*, if there exists a region* $\mathcal{R}_c \supset \{\Re s \geq c\}$ *where the function* Λ *is meromorphic, has a sole possible pole of order* $b + 1$ *(*$b \geq 0$*) at* $s = c$ *and is of polynomial growth there as* $|\Im s| \to +\infty$*. Moreover, this region* $\mathcal{R}_c$ *has one of the three following possible shapes, described in Definition 12 and in Figure 1:* a periodic shape (P), an hyperbolic shape (H), or a strip shape (S) .

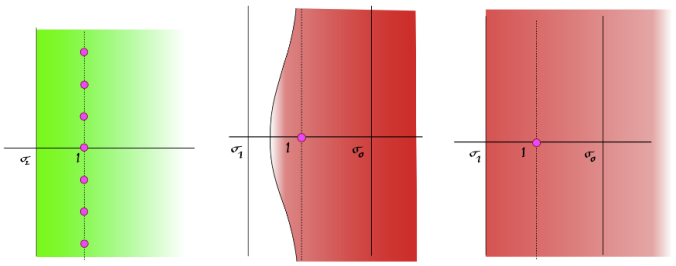


Figure 1 Tameness regions, drawn at $c = 1$, and their possible shapes (from left to right) : Periodic shape (P) – Hyperbolic shape (H) – and Strip shape (S).

The next result relates the analysis of the average depth $\mathbb{E}[D_n]$ and the tameness of the series $\Lambda(s)$. The proof is given in the annex (Section B.1). The result has not been stated before in its full generality, because the study of sources with positive entropy only relies on the case $(c = 1, b = 0)$, and the possible strip shape only recently "discovered". Only two other particular cases $(c = 1, b = 1)$ and $(c = 3/2, b = 0)$ have been already studied in [2].

▶ **Theorem 5.**

(i) If the series $\Lambda(s)$ is well-defined for $s \geq 2$ *(and thus for* $\Re s \geq 2$*), the expectation* $\mathbb{E}[D_n]$ *of the trie depth is expressed as an alternating sum which involves* $\Lambda(\ell)$*, for* $\ell \geq 2$*,*

$$n\mathbb{E}[D_n] = \sum_{\ell=2}^{n} (-1)^{\ell} \binom{n}{\ell} \ell\, \Lambda(\ell)\,. \tag{4}$$

(ii) If the series $s\Lambda(s)$ is tame at $s = c$ for $1 \leq c < 2$ with order b and a tameness region $\mathcal{R}_c$ delimited by a frontier δ_c*, the following holds:*

(a) *the expectation* $n\mathbb{E}[D_n]$ *involves the Rice kernel* $L_n(s)$ *and decomposes as*

$$n\,\mathbb{E}[D_n] = \mathrm{Res}\left[L_n(s) \cdot s\Lambda(s); s = c\right] + \int_{\delta_c} L_n(s) \cdot s\Lambda(s) ds \quad \textit{with} \quad L_n(s) := \frac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)}\,.$$

(b) *The general estimate holds,*

$$n\mathbb{E}[D_n] = \begin{cases} n^c P_b(\log n) & +O(n^{c-\delta})\,, & (S)\text{ shape}\,; \\ n^c P_b(\log n) & +n^c O\left(\exp[-(\log n)^\rho]\right)\,, & (H)\text{ shape}\,; \\ n^c G_b(\tau \log n) & +O(n^{c-\delta})\,, & (P)\text{ shape}\,. \end{cases} \tag{5}$$

The dominant term $\operatorname{Res}\left[L_n(s)\cdot s\Lambda(s); s=c\right]$ *involves a polynomial* P_b*, or a combination* G_b *between polynomials and a periodic function* Π*, together with the boolean* $[\![c=1]\!]$*,*

$$P_b(u) = [\![c=1]\!]\,\alpha_{b+1}\,u^{b+1} + \left[\sum_{i=0}^{b}\alpha_i u^i\right], \quad G_b(u) = [\![c=1]\!]\,\alpha_{b+1}\,u^{b+1} + \left[\sum_{i=0}^{b}\alpha_i u^i\right]\Pi(\{u\})\,.$$

The remainder term $\int_{\delta_c} L_n(s)\cdot s\Lambda(s)ds$ *involves the tameness shapes of Definition 12; in particular, in case of the* (H) *shape with exponent* ν*, the exponent* ρ *satisfies* $\rho < \frac{1}{1+2\nu}$*. For the other two shapes,* δ *is the strip width.*

The two main theorems (Theorems 4 and 5) exhibit the importance of the two generating series $\Lambda_k(s)$ and $\Lambda(s)$ of the source $\mathcal{Q}$: the Λ_k series intervenes via the exponential weight of the source in the study of the map $x \mapsto \log I_k(x)$; the Λ series intervenes via its tameness in the study of the average trie depth. The strategy of the paper is then as follows: we start with a given source $\mathcal{P}$ of positive entropy, with a well-known behaviour of its generating series $\Lambda_k(s)$ and $\Lambda(s)$. Section 4 describes the choice of the "starting" sources. Then, we distort the source $\mathcal{P}$ into another source $\mathcal{Q}$ (of zero entropy), in a way $\mathcal{Q}$ keeps the memory of some important features of $\mathcal{P}$. The next Section describes the distortion process.

3 Change of sources via rescaling or inserting delays

This section describes two processes for slowing down a given source: first, the rescaling process (in Sections 3.1 and 3.2), then the insertion of delays (in Section 3.3). The two processes are essentially the same, and they are compared in Section 3.4. Then we describe in Section 3.5 the particularities of these new sources when using tries. The Section finally establishes (in 3.6) a transfer of fundamental importance, between the Λ series of the initial source and the Λ series of the new source.

3.1 Rescaling sources and change of weights

The following process creates another source $\mathcal{Q}$ from a given source $\mathcal{P}$: It keeps the same fundamental intervals, but it changes the depth where they are used. It thus modifies the "speed" of the subdivision of the related partitions.

▶ **Definition 6.**

(i) *A function* g *is a rescaling function if it is defined on* $\{0\}\cup\mathbb{R}_{\geq 1}$ *and satisfies the following:*

(a) *the equalities* $g(0)=0$*,* $g(1)=1$ *hold ;*

(b) g *is a strictly increasing continuous map* $g:\mathbb{R}_{\geq 1}\to\mathbb{R}_{\geq 1}$*, with* $\lim_{x\to\infty} g(x)=\infty$ *;*

(c) *the inverse function* g^{-1} *has integer values at integer points.*

(ii) *With a source* $\mathcal{P}$*, its coincidence intervals* $\mathcal{I}_k(x)$*, together with a rescaling function* g*, one associates the source* $\mathcal{Q}$ *where the coincidence intervals* $\mathcal{J}_k(x)$ *satisfy*

$$\mathcal{J}_k(x) := \mathcal{I}_{\lfloor g(k)\rfloor}(x)\,. \tag{6}$$

The source $\mathcal{Q}$ *is the rescaled source of* $\mathcal{P}$ *via rescaling* g*. It is denoted as* $\mathcal{P}_{\langle g\rangle}$*.*

Rescaling acts here on the depth of fundamental intervals in a uniform[3] way: it changes the depth in the same way for all the intervals of the same initial depth. Remark that Item (c) entails the equality $\mathcal{I}_k = \mathcal{J}_{g^{-1}(k)}$. Using the notions of the rescaled source (Definition 6) and various weights (given in Definitions 1 and 3), this leads to the following result:

▶ **Proposition 7.** *Consider a rescaling g. If the source $\mathcal{P}$ has weight f (exponential, resp. Shannon), then the source $\mathcal{P}_{\langle g\rangle}$ has a weight (exponential, resp. Shannon) equal to $f \circ \lfloor g \rfloor$.*

Rescaling changes the weight; starting with a source $\mathcal{P}$ of entropy $h > 0$ (i.e., a Shannon weight $f(k) = h \cdot k$), the source $\mathcal{P}_{\langle g\rangle}$ has a Shannon weight $h \cdot \lfloor g(k) \rfloor$.

Then, using rescaling g for which $g(k)/k \to 0$, the source $\mathcal{P}_{\langle g\rangle}$ is of zero entropy.

3.2 Rescaled source

We associate with a rescaling g its sequence of differences,

$$\delta(g)(k) := \lfloor g(k) \rfloor - \lfloor g(k-1) \rfloor \qquad k \geq 1\,. \tag{7}$$

When $g(k) = k$ for any k, the differences $\delta(g)(k)$ are all equal to 1. For other functions g, there may exist integers k for which $\delta(g)(k)$ may be 0, or strictly larger than 1. We now describe the influence of rescaling on the expansion of x in base $\mathcal{P}$,

$$M^{(\mathcal{P})}(x) = a_1\, a_2\, a_3 \dots a_k, \dots\,, \qquad (a_i \in \Sigma)\,; \tag{8}$$

we consider the generalized digits $b_1\, b_2\, \dots, b_k, \dots b_n, \dots$, where each b_k is a block of initial digits of length $\delta(g)(k)$, and the expansion of x in base $\mathcal{P}_{\langle g\rangle}$ is

$$M^{(\mathcal{P}_{\langle g\rangle})}(x) = b_1\, b_2\, b_3 \dots b_k \dots\,, \qquad b_k = \left[a_{\lfloor g(k-1)\rfloor+1} \dots a_{\lfloor g(k)\rfloor}\right], \dots \tag{9}$$

When the k-th block is empty (i.e., $\delta(g)(k) = 0$ or $\lfloor g(k)\rfloor = \lfloor g(k-1)\rfloor$), the equality $\mathcal{J}_k(x) = \mathcal{J}_{k-1}(x)$ holds between two successive coincidence intervals, and the source $\mathcal{P}_{\langle g\rangle}$ does not emit any digit from Σ at time k. In this situation, we decide to emit a fictive symbol $[\cdot]$ at time k; in this way, we get a proper coding for the words emitted by the new source, and remember that it does not emit any digit from Σ at time k. Letting $\Sigma^0 := \{[\cdot]\}$, the generalized digit b_k always belongs to $\Sigma^{\delta(k)}$.

3.3 Inserting delays in a source $\mathcal{P}$.

The rescaled source $\mathcal{P}_{\langle g\rangle}$ may emit generalized digits as soon as $\delta(g)(k) \geq 2$. We now decide to always deal with a rescaling g for which all the differences defined in (7) satisfy $\delta(g)(k) \in \{0, 1\}$. In this case, the source $\mathcal{P}_{\langle g\rangle}$ only emits digits from the alphabet $\overline{\Sigma} = \{[\cdot]\} \cup \Sigma$. In such a source $\mathcal{P}_{\langle g\rangle}$, we view the symbol $[\cdot]$ as a waiting symbol, describing the situation where we wait for emitting the next symbol and do not emit any symbol from Σ. In this way, the length $\gamma(\ell) \geq 1$ of the plateau between the two successive indices (where symbols $a_{\ell-1} \in \Sigma$ and $a_\ell \in \Sigma$ are emitted) measures the ℓ-th waiting time between two times when symbols from Σ are emitted. In computer science contexts, the waiting time $\gamma(\ell)$ is called the ℓ-th "delay".

We are then led to another process for slowing down a source $\mathcal{P}$ defined on Σ: starting with source $\mathcal{P}$, where the expansion of x in base $\mathcal{P}$ is given in (8), we insert, for each ℓ, a delay $\gamma(\ell)$ between $a_{\ell-1}$ and a_ℓ, and obtain a new expansion, on the alphabet $\overline{\Sigma} = \{[\cdot]\} \cup \Sigma$,

$$M^{(\mathcal{P}^{\langle\gamma\rangle})}(x) = a_1\, [\cdot]^{\gamma(2)-1}\, a_2\, \dots [\cdot]^{\gamma(\ell)-1}\, a_\ell\, \dots\,. \tag{10}$$

[3] See the conclusion for a possible non-uniform definition.

3.4 Comparing the two points of view

Starting from a source $\mathcal{P}$ on Σ, we have thus introduced two new sources on $\overline{\Sigma} = \{[\cdot]\} \cup \Sigma$,
(a) $\mathcal{P}_{\langle g \rangle}$ is rescaled from $\mathcal{P}$ via a rescaling g with differences $\delta(g) \in \{0,1\}$ (described in (9)).
(b) $\mathcal{P}^{\langle \gamma \rangle}$ is obtained from $\mathcal{P}$ via inserting delays $\gamma \geq 1$ (described in (10)).

First, the equivalence holds: $\delta(g)(k) \in \{0,1\} \Longleftrightarrow \gamma(\ell) \geq 1$
Second, in this case, comparing the two expansions described in (9) and (10), and using Item (c) of Definition 6 proves the coincidence of the two sources $\mathcal{P}_{\langle g \rangle}$ and $\mathcal{P}^{\langle \gamma \rangle}$ when the pair (g, γ) satisfies

$$\text{Relation } (R) \qquad \sum_{k=1}^{\ell} \gamma(k) = g^{-1}(\ell) \qquad \text{for any integer } \ell \geq 1\,. \tag{11}$$

3.5 Sources $\mathcal{P}^{\langle \gamma \rangle}$ in Information Theory contexts.

We deal with infinite words emitted by $\mathcal{P}^{\langle \gamma \rangle}$, described in (10). As the inserted delays γ are the same for any word of the source, the words of $\mathcal{P}^{\langle \gamma \rangle}$ are very particular, and easy recognizable: they are obtained with the insertion of a deterministic process (the delays) at deterministic indices into a probabilistic one (the source $\mathcal{P}$).

We are interested in sorting a sequence $\boldsymbol{x}$ of words emitted by $\mathcal{P}^{\langle \gamma \rangle}$. A first idea is to use the trie $\mathtt{Trie}(\boldsymbol{x})$. In fact, there is a close connection between the two tries: $\mathtt{Trie}(\boldsymbol{x})$ and $\mathtt{Trie}(\boldsymbol{y})$ built on the compressed sequence $\boldsymbol{y}$, formed with the compressed words of $\boldsymbol{x}$, where the delays are removed. In the annex, (see Section D.1), we compare the branches of the two tries, and their lengths, and prove the following important property which will be the starting point for proving Proposition 20:

When the branches of the compressed $\mathtt{Trie}(\boldsymbol{y})$ *have length* $k_i(\boldsymbol{y})$*, the branches of* $\mathtt{Trie}(\boldsymbol{x})$ *have length* $g^{-1}(k_i(\boldsymbol{y}))$*, where* g *is the rescaling associated with* γ *via Relation* (R)*.*

When we use $\mathtt{Trie}(\boldsymbol{x})$ for sorting the sequence $\boldsymbol{x}$, there are three cases: the first case arises when the "trie-woman" knows that she deals with a source of type $\mathcal{P}^{\langle \gamma \rangle}$, with a precise knowledge of the delays. In this case, as she is not stupid, she probably does not build the trie $\mathtt{Trie}(\boldsymbol{x})$: she first compresses the sequence $\boldsymbol{x}$ (i.e. removes the delays), obtains a sequence $\boldsymbol{y}$, and then uses the trie $\mathtt{Trie}(\boldsymbol{y})$ –built on the source $\mathcal{P}$– to sort the sequence $\boldsymbol{x}$. However, there are other two cases, that are called "blind cases" : [(i) the "trie-woman" knows that she deals with a source of type $\mathcal{P}^{\langle \gamma \rangle}$, without knowing the form of the delay or (ii) she does not know that she deals with a $\mathcal{P}^{\langle \gamma \rangle}$ source]. In these blind cases, she builds the trie $\mathcal{T}(\boldsymbol{x})$ and uses it. We assume here that the "trie woman" is always blind.

3.6 Influence of rescaling or inserting delays on the Λ series.

The following result describes the influence of rescaling –or inserting delays– on the Λ series.

▶ **Proposition 8.** *The* Λ_k *series of the source* $\mathcal{P}_{\langle g \rangle}$ *obtained by rescaling* $\mathcal{P}$ *with* g *is*

$$\Lambda_k^{(\mathcal{P}_{\langle g \rangle})}(s) = \Lambda_{\lfloor g(k) \rfloor}^{(\mathcal{P})}(s)\,. \tag{12}$$

The Λ *series of the source* $\mathcal{P}^{\langle \gamma \rangle}$ *obtained by inserting delays* γ *in* $\mathcal{P}$ *is*

$$\Lambda^{(\mathcal{P}^{\langle \gamma \rangle})}(s) = \sum_{k \geq 0} \Lambda_k^{(\mathcal{P}^{\langle \gamma \rangle})}(s) = \sum_{\ell \geq 0} \gamma(\ell)\, \Lambda_\ell^{(\mathcal{P})}(s)\,. \tag{13}$$

This easy result, of fundamental importance, shows that each point of view –rescaling or inserting delays– is of interest: The Λ_k series of a rescaled source $\mathcal{P}_{\langle g \rangle}$, via a rescaling g, will be studied via g. The Λ series of a source $\mathcal{P}^{\langle \gamma \rangle}$, with delays γ, will be studied via γ.

4 The favorite set

The paper starts with a well understood source of positive entropy, and transforms it with rescaling or inserting delays. This section is devoted to the "starting" sources. We will use dynamical sources (see Section 4.1), and we need these sources to be good (see Section 4.2) and tame (see Section 4.3). We then describe in Section 4.4 our favorite set which gathers the "starting" sources we choose: they are all good and tame, with various tameness shapes.

4.1 Dynamical sources

Dynamical sources, introduced by Vallée in [22], are related to dynamical systems of the interval $\mathcal{I} := [0, 1]$. One starts with a topological partition $\{\mathcal{I}_\sigma\}$ of $\mathcal{I}$ indexed by symbols $\sigma \in \Sigma$, a coding map $\tau : \mathcal{I} \to \Sigma$ which equals σ on $\mathcal{I}_\sigma$, and a shift map $T : \mathcal{I} \to \mathcal{I}$. The mapping T is defined by its branches $T_\sigma : \mathcal{I}_\sigma \to \mathcal{I}$ that are assumed to be *surjective*, of class $\mathcal{C}^2$, and strictly monotonic. The source produces on the real x the word $M(x)$ that encodes the trajectory $(x, Tx, T^2x \dots)$ via the coding map τ, namely, $M(x) = (\tau(x), \tau(Tx), \tau(T^2x), \dots)$ ..
When the input x is randomly drawn from $\mathcal{I}$, this becomes a probabilistic process.
Each (local) inverse of T^k is associated with a prefix $w \in \Sigma^k$ and denoted as h_w. As each branch of T is surjective, each inverse branch h_w is defined on $\mathcal{I}$, and the image of h_w is the interval $\mathcal{I}_w = [h_w(0), h_w(1)]$ which gathers all the reals x for which $M(x)$ begins with w. There is thus an underlying sequence $(\mathcal{P}_k)_k$ of partitions, with $\mathcal{P}_k := \{\mathcal{I}_w \mid w \in \Sigma^k\}$ which will be inherited for rescaled sources.

All memoryless sources are dynamical sources, associated with a increasing piecewise linear shift. A main instance is the standard binary system, obtained by $T(x) = \{2x\}$ (where $\{\cdot\}$ is the fractional part). Another dynamical source plays a central role here: the source `CF`, related to continued fraction, and associated with the non-linear shift $T(x) = \{1/x\}$.

In the context of analytic combinatorics, the importance of dynamical sources relies on the following fact: their Λ series are themselves generated by the secant transfer operator:

▶ **Lemma 9.** *The secant transfer operator $\mathbf{G}_s$ of the source involves the inverses h_σ of the branch T_σ and acts on functions $F : \mathcal{I}^2 \to \mathbb{C}$, as follows:*

$$\mathbf{G}_s[F](x, y) = \sum_{\sigma \in \Sigma} \left| \frac{h_\sigma(x) - h_\sigma(y)}{x - y} \right|^s F(h_\sigma(x), h_\sigma(y)) \,. \tag{14}$$

The k-th iterate $\mathbf{G}_s^k$ involves the inverse branches h_w of the shift T^k

$$\mathbf{G}_s^k[F](x, y) = \sum_{w \in \Sigma^k} \left| \frac{h_w(x) - h_w(y)}{x - y} \right|^s F(h_w(x), h_w(y)) \,.$$

The fundamental relations hold: $\quad \Lambda_k(s) = \mathbf{G}_s^k[1](0, 1), \quad \Lambda(s) = (I - \mathbf{G}_s)^{-1}[1](0, 1) \,.$

4.2 Good sources

The Good Class gathers the dynamical sources for which there is an iterate T^n of the shift T that is strictly expansive. This class contains, together with all the memoryless sources, many other sources, as the Continued Fraction Source.

▶ **Definition 10.** *A dynamical source is good if it satisfies the following:*

(i) *The constant ρ defined as follows satisfies $\rho < 1$,*

$$\rho = \limsup_{k\to\infty} (\sup_{w\in\Sigma^k} p_w)^{1/k} .$$

(ii) *There is a constant $B > 0$, for which, for any $w \in \Sigma$, one has $|h''_w| \le B|h'_w|$.*

(iii) *There exists $A < 1$ for which the series $\sum_{w\in\Sigma} p_w^s$ converges on $\Re s > A$.*

Item (i) is the most important; items (ii) and (iii) are only useful for sources on infinite alphabets Σ, and the real A in (iii) is the convergence abscissa of the series $\Lambda_1(s)$.
When the source is good, the secant transfer operator, defined in (14), when acting on the functional space $\mathcal{C}^1([0,1]^2)$, has, for any real $s > A$, a unique dominant eigenvalue $\lambda(s)$, separated from the remainder of the spectrum by a spectral gap. The function $\lambda(s)$ (which depends analytically on s) is called the dominant eigenvalue of the source. It satisfies $\lambda(1) = 1$, $\lambda'(1) = -h$ where $h > 0$ is the entropy of the source. Moreover, there exists a (complex) neighborhood $\mathcal{V}$ of the real axis ($s > A$) on which the series $\Lambda_k(s)$ has a quasi-power behaviour that involves functions $v(s)$ and $c_k(s)$, that are analytic on $\mathcal{V}$, under the form,

$$\Lambda_k(s) = v(s)\lambda(s)^k + c_k(s) = v(s)\lambda(s)^k \left[1 + \frac{1}{v(s)}\frac{c_k(s)}{\lambda(s)^k}\right] = v(s)\lambda(s)^k \left[1 + O(\rho^k)\right] . \quad (15)$$

The last estimate indeed holds, due to the Spectral Radius Theorem and the existence of a spectral gap. Comparing with Eqn (2) leads to the following result:

▶ **Proposition 11.** *A good source has a "dominant eigenvalue" denoted as $\lambda(s)$. It has an exponential weight equal to k, with a base $u(s) = \lambda(s)$. Its Shannon entropy is $|\lambda'(1)|$.*

Moreover the pressure function $L(s) = \log \lambda(s)$ admits convexity properties : it is always convex and the two conditions are equivalent (see for instance [1]):
(a) the source is conjugated[4] to an unbiased memoryless source;
(b) the function L is affine.
For a memoryless source (p, q), the function $\lambda(s)$ is defined on the whole plane $\mathbb{C}$, and

$$\lambda(s) = (p^s + q^s) , \qquad \Lambda(s) = \frac{1}{1-\lambda(s)} . \quad (16)$$

The three conditions are equivalent (See Section C.1 in the annex).
(a) There exists $\tau > 0$ for which $\lambda(1 + 2i\pi\tau) = 1$;
(a) The function λ is periodic with period $2i\pi\tau$;
(a) The ratio $(\log q)/(\log p)$ is rational.
In this case, the memoryless source is periodic with period τ. Otherwise, it is aperiodic.

4.3 Tameness and tameness shapes

With Eqn (15), the Λ series satisfies

$$\Lambda(s) = v(s)\frac{1}{1-\lambda(s)} + c(s), \qquad c(s) = \sum_k c_k(s) , \quad (17)$$

[4] Two sources are conjugated if their shifts are conjugated in $\mathcal{C}^2([0,1])$.

where $v(s)$ and $c(s)$ are analytic for $s \in \mathcal{W} := \mathcal{V} \cap \{\Re s > \tau\}$ for some $\tau < 1$. As the inequality $|\Lambda(s)| \leq \Lambda(\Re s)$ holds on the hyperplane $\Re s > 1$, the function $\Lambda(s)$ is meromorphic on the domain $\{\Re s > 1\} \cup \mathcal{W}$ with an only simple pole at $s = 1$. However, in order to apply the Rice formula and to be able to study tries (see Theorem 5), we need to know a region on the left of the vertical line $\Re s = 1$ where the Λ series is *tame*, i.e, meromorphic and of polynomial growth when $|\Im s| \to \infty$. We now give a general definition of tameness, where the point of tameness may be any $c \geq 1$. This definition is already used in the statement of Theorem 5.

▶ **Definition 12.**

(i) *A function $M(s)$ analytic on $\{\Re s > c\}$ (c real) is* tame *at c of order $b \geq 0$, if there exists a region $\mathcal{R}_c \supset \{\Re s \geq c\}$ where the function M is meromorphic, has a sole possible pole of order $b+1$ at $s = c$ and is of polynomial growth as $|\Im s| \to +\infty$. Moreover, this region $\mathcal{R}_c$ has one of the three following possible[5] shapes (See Fig. 1):*

(S) Strip shape: $\mathcal{R}_c = \{s \mid \Re s > c - \delta\}$ *for some* $\delta > 0$.

(H) Hyperbolic shape (of exponent ν): $\mathcal{R}_c = \mathcal{R}_c^+ \cup \mathcal{R}_c^-$ *(with positive constants A, B)*

$$\mathcal{R}_c^+ := \{s = \sigma + it \mid |t| \geq B,\ \ \sigma > c - \tfrac{A}{|t|^\nu}\ ,\ \ \mathcal{R}_c^- := \{s = \sigma + it \mid |t| \leq B\ \ \sigma > c - \tfrac{A}{B^\nu}, \}\,.$$

(P) Periodic[6] shape (of period τ): $\mathcal{R}_c = \{s \mid \Re s > c - \delta\} \setminus \bigcup\{s_k\}$ *where the points s_k, of the form $s_k = c + 2ki\pi\tau$ (for $k \in \mathbb{Z}$, $k \neq 0$) are poles of Γ of order at most $b+1$.*

(ii) *A source $\mathcal{Q}$ is tame at c of order b with a tameness shape in $\{S, H, P\}$ if its Λ series is tame at c of order b with the given tameness shape.*

(iii) *A good source is tame if its Λ series is tame at $s = 1$ with order 0.*

4.4 Instances of tame sources. Our favorite set of starting sources

These various prescribed tameness shapes (that may appear at a first glance somewhat artificial) indeed intervene in possible behaviors of classical sources. Here, we choose particular sources $\mathcal{P}$ of positive entropy as starting points: we wish good sources, proven to be tame, with various tameness shapes in $\{S, H, P\}$. This leads us to our favorite set of starting sources:

▶ **Definition 13.** *The favorite set of starting sources contains three types of sources:*

(a) *the class $\mathcal{MP}$ of all the periodic memoryless sources ;*

(b) *the class $\mathcal{MA}$ of all the aperiodic memoryless sources related to a pair $(p, q = 1-p)$ for which the ratio $\beta = (\log q)/(\log p)$ is irrational with a finite irrationality exponent;*

(c) *the Continued Fraction source, denoted as* `CF`*, associated with the shift $T(x) = \{1/x\}$.*

The following theorem first shows that our favorite set is well-chosen. Then, using Theorem 14 inside Theorem 5 leads to the analysis of the trie depth for each "favorite" source.

▶ **Theorem 14.** *All the starting sources of the favorite set are tame. Their shape is*

(P) *for any source of the class $\mathcal{MP}$*

(H) *for any source of the class $\mathcal{MA}$: the exponent ν of the tameness region is related to the irrationality exponent μ of $\beta = \log p / \log q$ via the equality $\nu = 2(\mu + 1)$.*

(S) *for the* `CF` *source.*

---

[5] To the best of our knowledge, these shapes are the only ones to occur in "classical" analyses.

[6] This means here that $M(s)$ is of polynomial growth on a family of horizontal lines $t = t_k$ with $t_k \to \infty$, and on vertical lines $\Re(s) = c - \delta'$ with some $\delta' < \delta$.

The proof of the previous result is not trivial (see a sketch of proof in the annex). The first assertion is classical, but the second and the third assertions are based on subtle properties. For aperiodic memoryless sources, the geometry of the tameness region has not been precisely described before the works of [12] and [20], which relate it to the irrationality exponent. This is why the associated remainder term in previous trie analyses was not precise, too. The third assertion is based on a seminal work of Dologopyat [10], extended by Baladi and Vallée [1] to the case of the CF dynamical system, then to the secant operator by Cesaratto and Vallée [4].

Sketch of the proof. See Section C in the annex. ◀

5 Sources $\mathcal{P}_{a,b}$ associated with a source $\mathcal{P}$ from the favorite set.

The sources to be analyzed are defined in Section 5.1 by the insertion of a delay $\gamma_{a,b}$ described in (18). Then, after computing the rescaling $g_{a,b}$ which is associated with the delay $\gamma_{a,b}$, we obtain in Section 5.2 our first result about asymptotic normality à la Shannon-MacMillan-Breiman (Theorem 16). Then Section 5.3 is devoted to the description of Λ series of the sources $\mathcal{P}_{a,b}$. Section 5.4 "inserts" the results of Section 5.3 inside Theorem 5 and provides the final result (Theorem 19). Section 5.5 provides a direct comparison between the tries of the two sources (the initial one, and the one with delays). Finally, Section 5.6 focuses on a particular class of delays, the computational delays, of main algorithmic interest.

5.1 The sources $\mathcal{P}_{a,b}$.

We start with a source $\mathcal{P}$ of our favorite set and consider the source $\mathcal{P}_{a,b}$ obtained from $\mathcal{P}$ by inserting in $\mathcal{P}$ delays $\gamma_{a,b}$ that satisfy, for any integer $\ell \geq 1$,

$$\gamma_{a,b}(\ell) = \lfloor a^\ell \rfloor \cdot \ell^b\,, \quad (a>1, \quad b \text{ integer}, \quad b \geq 0)\,. \tag{18}$$

The particular form of these delays is interesting from four points of view

(a) The initial source $\mathcal{P}$ is the source $\mathcal{P}_{1,0}$. The parameters (a,b) describe the distortion the delays $\gamma_{a,b}$ impose to the initial source $\mathcal{P}$;

(b) This class contains the computational delays, defined in Section 5.6 ;

(c) Proposition 8 provides an expression for the $\Lambda(s)$ series of source $\mathcal{P}_{a,b}$, from which we exhibit its main singularity (in Proposition 17) and then describe its possible tameness (in Proposition 18). This will be central for applying Theorem 5 and obtain estimates for the expected depth $\mathbb{E}[D_n^{\langle a,b\rangle}]$ of the trie built on the source $\mathcal{P}_{a,b}$;

(d) Proposition 8 provides an expression for the $\Lambda_k(s)$ series of source $\mathcal{P}_{a,b}$ that will be used for applying Theorem 4 and obtain gaussian laws for $\log I_k^{\langle a,b\rangle}(x)$, *provided that* the source $\mathcal{P}_{a,b}$ be viewed as a rescaled source via an explicit rescaling $g_{a,b}$ associated with delay $\gamma_{a,b}$ via Relation (R). We study this last point in the next Section.

5.2 Rescaling $g_{a,b}$ and asymptotic gaussian laws for $\log I_k(x)$.

Relation (R) described in (11) does not lead (generally speaking) to an explicit expression of $g_{a,b}$; but, as we deal with exponential weights, we only need the principal[7] part $g_{a,b}^{(0)}$ of $g_{a,b}$, for which the following lemma provides an explicit expression.

[7] We recall that the principal part $g^{(0)}$ of g is given by the decomposition

$$g(x) = g^{(0)}(x) + A + \epsilon(x), \qquad \epsilon(x) \to 0, \quad g^{(0)}(1/x) \to 0 \quad (x \to \infty)\,. \tag{19}$$

▶ **Lemma 15.** *Consider a delay $\gamma_{a,b}$ defined in* (18)*. The principal part $g^{(0)}_{a,b}$ of the rescaling $g_{a,b}$ associated with $\gamma_{a,b}$ via Relation* (R) *satisfies*

$$g^{(0)}_{1,b}(x) = (b+1)^{1/(b+1)}\, x^{1/(b+1)}\,, \quad (b \geq 0)\,; \quad g^{(0)}_{a,b}(x) = \log_a \left[\frac{x}{(\log_a(x))^b}\right], \quad (a > 1,\, b \geq 0)\,.$$

Proof. Given in the annex. (Section B.2). ◀

Using Proposition 8 and Definition 3 leads to the first important result of the paper:

▶ **Theorem 16.** *The source $\mathcal{P}_{a,b}$ has an exponential weight equal to $\lfloor g^{(0)}_{a,b} \rfloor$ defined in Lemma 15, with the eigenvalue function $\lambda(s)$ of the initial source $\mathcal{P}$ as a base function. Applying Theorem 4 to the source $\mathcal{P}_{a,b}$ with $g := \lfloor g^{(0)}_{a,b} \rfloor$ provides asymptotical normality à la Shannon-MacMillan-Breimann for the source $\mathcal{P}_{a,b}$.*

This result is a straightforward application of the Quasi-Power Theorem of Hwang [16]. It leads to asymptotical normality phenomena related to a quite general behaviour of expectations and variances. For more effective results, –in particular, in analyses related to tries–, we have limited ourselves to rescalings $g = g_{a,b}$ that are associated with delays $\gamma = \gamma_{a,b}$ defined in (18). However, with Proposition 7, the previous result holds for any rescaling g that fulfills Definition 6 and satisfies $g(k)/k \to 0$.

We thus provide "natural" instances of asymptotic normality phenomena, with expectations and variances of order $\Theta(g(k))$ where g is any increasing function with $g(k)/k \to 0$.

5.3 Study of the $\Lambda(s)$ series of the $\mathcal{P}_{a,b}$ source.

The delays $\gamma_{a,b}$ involve the integer part $\lfloor a^\ell \rfloor$. Using Proposition 8 and decomposition $\lfloor a^\ell \rfloor = a^\ell - \{a^\ell\}$, the Λ series of the source $\mathcal{P}_{a,b}$ decomposes as a sum,

$$\left[\sum_\ell a^\ell \ell^b \Lambda_\ell(s)\right] - \left[\sum_\ell \{a^\ell\} \ell^b \Lambda_\ell(s)\right]. \tag{20}$$

We then mainly study the first series, which provides the "main" behaviour of the total $\Lambda(s)$ series (its dominant singularity, its possible tameness), and finally, for non integer values of a, we have to restrict ourselves to the half plane $\Re s > 1$, due to the occurrence of the second series that is (only) analytic on $\Re s > 1$ and bounded on each half-plane $\Re s \geq 1 + \rho > 1$.

We first prove that the dominant singularity of the Λ series attached to the source $\mathcal{P}_{a,b}$ is a pole of order $(b+1)$ located at σ_a, where $\sigma_a \geq 1$ is defined by the equation $a\lambda(s) = 1$.

▶ **Proposition 17** (Dominant singularity)**.** *Consider any source $\mathcal{P}$ from the favorite set, with a dominant eigenvalue $\lambda(s)$, and the modified source $\mathcal{P}_{a,b}$ obtained from $\mathcal{P}$ by inserting delays of the form $\gamma_{a,b}$ with $a \geq 1$. Consider the real σ_a defined by the relation $a\lambda(\sigma_a) = 1$.*
The Λ series of the source $\mathcal{P}_{a,b}$ has a pole of order $b+1$ located at σ_a, and is meromorphic on the domain $\{\Re s > \sigma_a\} \cup \mathcal{W}_a$; with $\mathcal{W}$ defined in Section 4.3, there are two cases for $\mathcal{W}_a$: $\mathcal{W}_a = \mathcal{W}$ for $a = 1$. If $a > 1$, there exists $\tau_a < \sigma_a$ for which $\mathcal{W}_a = \mathcal{W} \cap \{\Re s > \max(1, \tau_a)\}$.

Proof. Given in the annex. (Section B.3). ◀

We now describe conditions under which the source $\mathcal{P}_{a,b}$ is proven to be tame at $s = \sigma_a$.

▶ **Proposition 18** (Tameness). *Consider any source $\mathcal{P}$ from the favorite set. The source $\mathcal{P}_{a,b}$ is tame at σ_a in the following cases:*

- **(i)** [$a = 1$, any $\mathcal{P}$]: *The source $\mathcal{P}_{a,b}$ is tame at $s = 1$, of order b, with the same shape as $\mathcal{P}$;*
- **(ii)** [$a > 1$ and $\mathcal{P}$ is memoryless and tame]: *The source $\mathcal{P}_{a,b}$ is tame at $s = \sigma_a$ of order b with a tameness region $\mathcal{R}_{\sigma_a}$ of the same shape as $\mathcal{R}_1$, but limited to the half plane $\Re s > 1$;*
- **(iii)** [$\mathcal{P} = \texttt{CF}$ and $a \neq 1$ is close enough to 1]: *The source $\mathcal{P}_{a,b}$ is tame at $s = \sigma_a$ with a tameness strip defined from the tameness strip $\mathcal{S}$ of $\mathcal{P}$ as $\mathcal{S} \cap \{\Re s > 1\}$.*

▶ Remark. For $\mathcal{P} = \texttt{CF}$ and a not close enough to 1, we do not know the tameness of the source $\mathcal{P}_{a,b}$ at $s = \sigma_a$. The case σ_a close enough to 1 is particular from this point of view.

Proof. It is done in Section C of the annex. ◀

5.4 Average depth for tries built on the source $\mathcal{P}_{a,b}$

We now obtain our second main result. We first limit the real σ_a to be less than 2 (i.e., $a < 1/\lambda(2)$, where $\lambda(s)$ is the eigenvalue of $\mathcal{P}$), so that Condition (a) of Theorem 5 is fullfilled.

▶ **Theorem 19.** *Consider any source $\mathcal{P}$ from the favorite set (with an eigenvalue λ), a delay $\gamma_{a,b}$, the source $\mathcal{P}_{a,b}$ obtained with inserting a delay $\gamma_{a,b}$ in $\mathcal{P}$, and the blind*[8] *trie built on words of the source $\mathcal{P}_{a,b}$. Define σ_a by the relation $a\lambda(\sigma_a) = 1$ and assume the parameter a be less than $1/\lambda(2)$.*
With notations of Theorem 5, the estimate holds for the average depth of the blind trie,

$$\mathbb{E}[D_n^{\langle a,b\rangle}] = \begin{cases} n^{\sigma_a - 1} P_b(\log n) & [1 + O(n^{-\delta})]\,, & \mathcal{P} = \texttt{CF},\ a \text{ close to } 1\,; \\ n^{\sigma_a - 1} P_b(\log n) & [1 + O\left(\exp[-(\log n)^{\rho}]\right)]\,, & \mathcal{P} \in \mathcal{MA},\ 1 \le a \le 1/\lambda(2)\,; \\ n^{\sigma_a - 1} G_b(\tau \log n) & [1 + O(n^{-\delta})\,, & \mathcal{P} \in \mathcal{MP},\ 1 \le a \le 1/\lambda(2)\,. \end{cases}$$

This theorem exhibits a class of various behaviours that may occur in the average trie depth $\mathbb{E}[D_n^{\langle a,b\rangle}]$, both in the dominant term and the remainder term. In particular, a term n^d with $d \in]0,1[$ arises in the dominant term as soon as parameters d and a are related by the equality $d = \lambda^{[-1]}(1/a) - 1$, where $\lambda^{[-1]}$ is the inverse of the eigenvalue mapping. For the unbiased memoryless source $(1/2, 1/2)$, the relation between a and d is $d = \log_2 a$.

5.5 A direct estimate for the average depth $\mathbb{E}[D_n^{\langle a,b\rangle}]$.

We have mentioned in Section 3.5 a direct connection between the two tries

- **(i)** $\texttt{Trie}(\boldsymbol{x})$ built on a sequence $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ of words emitted by the $\mathcal{P}^{\langle\gamma\rangle}$ source;
- **(ii)** $\texttt{Trie}(\boldsymbol{y})$ built on the compressed sequence $\boldsymbol{y}$ obtained by removing all the delays in the words x_i of the sequence $\boldsymbol{x}$. $\texttt{Trie}(\boldsymbol{y})$ is thus built on the source $\mathcal{P}$.

This relation leads to the following result (proven in Section D.1 of the annex.)

▶ **Proposition 20.** *Consider any source $\mathcal{P}$ of the favorite set, and the source $\mathcal{P}_{a,b}$ built with insertion of delays $\gamma_{a,b}$. The following inequality holds between the average trie depth of the $\mathcal{P}_{a,b}$ source and the average trie depth of the initial source $\mathcal{P} = \mathcal{P}_{1,0}$:*

$$\mathbb{E}[D_n^{\langle a,b\rangle}] \ge g_{a,b}^{-1}\left(\mathbb{E}[D_n^{\langle 1,0\rangle}]\right)\,.$$

The estimate is sharp for $a = 1$. For $a > 1$, the right member involves a term $n^{(1/h)\log a}$ whereas the left member involves a term $n^{\sigma_a - 1}$ (see Theorem 19). The equality between the two exponents holds for some $a > 1$ only if the starting source $\mathcal{P}$ is memoryless unbiased.

[8] This notion is defined in Section 3.5.

This is an interesting result which applies to sources $\mathcal{P}_{a,b}$ (even if they are not proven to be tame) provided they be associated with a tame source $\mathcal{P}$; this is the case for the CF source when the parameter a is not close to 1 and varies in the whole interval $]1, 1/\lambda(2)[$. We obtain

$$\mathbb{E}[D_n^{\langle a,b\rangle}] \geq n^{(1/h)\log a}\, P_b(\log n)\,[1 + O(n^{-\delta})]\,, \qquad \text{with} \qquad 1/h = (6\log 2)/\pi^2\,.$$

5.6 Computational version of a source from the favorite set

The *exact* computation of the word $M(x)$ given in (8) is an important algorithmic issue for the initial source $\mathcal{P}$. See for instance [8] for a description in the case when when $\mathcal{P} = \mathtt{CF}$. This leads to two important questions: – the first one about the *needed precision* $\pi(\ell)$ for computing *exactly* the ℓ-th digit of the source $\mathcal{P}$ when the first $(\ell-1)$ digits have been already computed, – the second one, about the *cost* (in arithmetical complexity) for computing exactly the ℓ-th digit of the source $\mathcal{P}$ when the first $(\ell-1)$ digits have been already computed. As all the favorite sources have a positive entropy, the needed precision $\pi(\ell)$ is $\Theta(\ell)$ and the main operations that are performed are multiplications (on integers with $\Theta(\ell)$ digits). Then, the computational cost is $\Theta(\ell^2)$. This defines a (particular) delay $\widehat{\gamma}$, called the computational delay, that is equal here to $\gamma_{1,2}$ for any reference source. The source $\widehat{\mathcal{P}}$ attached to the delay $\widehat{\gamma}$ "replaces" the theoretical source $\mathcal{P}$ in any complexity issues (related in particular to tries), and is called the computational source. (See Section D.2 in the annex for precisions).

With Lemma 15, the principal part of the scaling $g_{1,2}$ is equal to $(3x)^{1/3}$. Applying the main Theorems (Theorems 16 and 19) in this particular case leads to the following result:

▶ **Theorem 21.** *Consider the computational source $\widehat{\mathcal{P}}$ associated with a reference source $\mathcal{P}$.*

(a) *Except in the case when the reference source is memoryless unbiased, the random variable $x \mapsto \log \widehat{I}_k(x)$ asymptotically follows a gaussian law*

$$\Pr\left[x \,\middle|\, \frac{\log \widehat{I}_k(x) - |\lambda'(1)| \cdot \lfloor (3k)^{1/3} \rfloor}{\left[|\lambda''(1) - \lambda'(1)^2| \cdot \lfloor (3k)^{1/3} \rfloor\right]^{1/2}} \leq u\right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt\,.$$

(b) *The general estimate holds for the average depth of the blind trie, with the notations of Theorem 5 and $b = 2$, for which P_2 and G_2 are of degree 3,*

$$\mathbb{E}[\widehat{D}_n] = \begin{cases} P_2(\log n) & [1 + O(n^{-\delta})]\,, & \mathcal{P} = \mathtt{CF}\,; \\ P_2(\log n) & [1 + O\left(\exp[-(\log n)^{\rho}]\right)]\,, & \mathcal{P} \in \mathcal{MA}\,; \\ G_2(\tau \log n) & [1 + O(n^{-\delta})\,, & \mathcal{P} \in \mathcal{MP}\,. \end{cases}$$

6 Conclusion and open problems

The paper introduces sources with delays that provide "natural" instances of sources with zero entropy: here, the ℓ-th delay (that is inserted between the symbols $a_{\ell-1} = a_{\ell-1}(x)$ and $a_\ell = a_\ell(x)$) only depends on the depth ℓ where it is inserted, and is uniform (i.e., the same for any input x). In this sense, this is a "toy-model", that however "shows the path" for future research. We may indeed think about more realistic models where the (non uniform) ℓ-th delay may depend on the prefix $M_\ell(x)$ of length ℓ of the word $M(x)$, that is itself defined in terms of interval $\mathcal{I}_\ell(x)$. Such delays intervene in the modelling of two classical sources in Number Theory: the Stern-Brocot source is viewed as a source with delays from the CF source, and the Sturm source as a source with delays from the Stern-Brocot source. Such generalized delays also intervene in the analysis of VLMC sources (VLMC = Variable Length Markov Chains). In this sense, using generalized (i.e., non uniform) delays may be of great interest in the modelling and analysis of a large variety of sources of zero entropy.

References

1 Viviane Baladi and Brigitte Vallée. Euclidean Algorithms are gaussian. *Journal of Number Theory, Volume 110, Issue*, pages 331–386, 2005.

2 Valérie Berthé, Eda Cesaratto, Frédéric Paccaut, Pablo Rotondo, Martín D. Safe, and Brigitte Vallée. Two Arithmetical Sources and their associated Tries. In Michael Drmota and Clemens Heuberger, editors, *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, AofA 2020, June 15-19, 2020, Klagenfurt, Austria (Virtual Conference)*, volume 159 of *LIPIcs*, pages 4:1–4:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.AofA.2020.4.

3 Valérie Berthé, Eda Cesaratto, Pablo Rotondo, and Martin Dario Safe. Lochs-type Theorems beyond positive entropy. arXiv preprint, 2022. doi:10.48550/arXiv.2202.04008v1.

4 Eda Cesaratto and Brigitte Vallée. Gaussian Distribution of Trie depth for Strongly Tame sources. *Comb. Probab. Comput.*, 24(1):54–103, 2015. doi:10.1017/S0963548314000741.

5 Julien Clément, James Allen Fill, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a realistic analysis of the Quickselect algorithm. *Theory Comput. Syst.*, 58(4):528–578, 2016. doi:10.1007/s00224-015-9633-5.

6 Julien Clément, Philippe Flajolet, and Brigitte Vallée. Dynamical Sources in Information Theory: A General Analysis of Trie structures. *Algorithmica*, 29(1):307–369, 2001. doi:10.1007/BF02679623.

7 Julien Clément, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a realistic analysis of some popular sorting algorithms. *Comb. Probab. Comput.*, 24(1):104–144, 2015. doi:10.1017/S0963548314000649.

8 Robert M Corless. Continued fractions and chaos. *American Mathematical Monthly*, 99:203–215, 1992.

9 Karma Dajani and A. Howard Fieldsteel. Equipartition of interval partitions and an application to number theory. In *Proceedings of the American Mathematical Society*, volume 129, pages 3453–3460, 2001.

10 Dmitry Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math.*, 147(2):357–390, 1998. doi:10.2307/121012.

11 Philippe Flajolet. The Ubiquitous Digital Tree. In *STACS 2006, 23rd Annual Symposium on Theoretical Aspects of Computer Science, Marseille, France, February 23-25, 2006, Proceedings*, volume 3884 of *Lecture Notes in Computer Science*, pages 1–22. Springer, 2006. doi:10.1007/11672142_1.

12 Philippe Flajolet, Mathieu Roux, and Brigitte Vallée. Digital Trees and Memoryless sources: from Arithmetics to Analysis. *Proceedings of AofA'10, DMTCS, proc AM*, pages 231–258, 2010. URL: http://www.dmtcs.org/dmtcs-ojs/index.php/proceedings/article/view/dmAM0117/3289.

13 Philippe Flajolet and Robert Sedgewick. Digital Search Trees revisited. *SIAM J. Comput.*, 15(3):748–767, 1986. doi:10.1137/0215054.

14 Philippe Flajolet and Robert Sedgewick. Mellin Transforms and Asymptotics: Finite differences and Rice's integrals. *Theor. Comput. Sci.*, 144(1&2):101–124, 1995. doi:10.1016/0304-3975(94)00281-M.

15 Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. URL: http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=9780521898065.

16 Hsien-Kuei Hwang. On Convergence Rates in the central limit theorems for Combinatorial Structures. *Eur. J. Comb.*, 19(3):329–343, 1998. doi:10.1006/eujc.1997.0179.

17 Hsien-Kuei Hwang. Limit Theorems for the Number of Summands in Integer Partitions. *J. Comb. Theory, Ser. A*, 96(1):89–126, 2001. doi:10.1006/jcta.2000.3170.

18 N. E. Nörlund. Leçons sur les équations linéaires aux différences finies. In *Collection de monographies sur la théorie des fonctions*. Gauthier-Villars, Paris, 1929.

19 Niels Erik Nörlund. *Vorlesungen über Differenzenrechnung*. Chelsea Publishing Company, New York, 1954.

20 Mathieu Roux and Brigitte Vallée. Information theory: Sources, Dirichlet series, and realistic analyses of data structures. In Petr Ambroz, Stepan Holub, and Zuzana Masáková, editors, *Proceedings 8th International Conference Words 2011, Prague, Czech Republic, 12-16th September 2011*, volume 63 of *EPTCS*, pages 199–214, 2011. doi:10.4204/EPTCS.63.26.

21 Wojtek Szpankowski. *Average case analysis of algorithms on sequences*. Interscience series in Discrete Mathematics and Optimization. Wiley, 2001.

22 Brigitte Vallée. Dynamical Sources in Information Theory: Fundamental Intervals and Word Prefixes. *Algorithmica*, 29(1):262–306, 2001. doi:10.1007/BF02679622.

23 Brigitte Vallée. The Depoissonisation Quintet: Rice-Poisson-Mellin-Newton-Laplace. In James Allen Fill and Mark Daniel Ward, editors, *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, AofA 2018, June 25-29, 2018, Uppsala, Sweden*, volume 110 of *LIPIcs*, pages 35:1–35:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.AofA.2018.35.

A Annex

The annex is organized into three Sections. The first Section (Section B) gathers proofs of Theorem 5, Lemma 15, Proposition 17. The second Section (Section C) gathers the material which is related to tameness. It is first used in the proof of Theorem 14 but also in its extension in Theorem 18. Finally the third Section (Section D) is devoted to the proofs of the last two results that are stated in Section 5: direct comparison of the tries (stated in Section 5.5) and computational version of the source (stated in Section 5.6).

B Proofs of Theorem 5, Lemma 15, Proposition 17

B.1 Proof of Theorem 5

We summarize (Part I) the general estimates that are needed. Then, we describe the dominant parts (Part II) and finally the remainder terms (Part III)

Part I. Needed estimates. Along the proof, we compare the two functions, the Rice kernel $L_n(s)$ and the function $n^s\Gamma(-s)$ along a vertical (or an hyperbolic) line. The comparisons are based on the estimates provided in the annex of [4], and summarized as follows. Note that the exponential decreasing of the Γ function along vertical (or hyperbolic) lines plays a central role.

▶ **Lemma 22.** Consider a curve γ, that is a vertical line $\gamma = \{\Re s = c\}$ or an hyperbolic curve, and, for some $\tau > 0$, the infinite set of points $c_k = c + 2i\pi k\tau$ for $k \in \mathbb{Z}$. The curve γ is decomposed into two parts : the central part $\gamma_n^- := \{s \in \gamma \mid |\Im s| \leq \sqrt{n}\}$ and the exterior part $\gamma_n^+ := \{s \in \gamma \mid |\Im s| > \sqrt{n}\}$. In the same way, the set $\mathbb{Z}$ is decomposed into two parts, the central part $\mathbb{Z}_n^- := \{k \mid 2\pi|k|\tau \leq \sqrt{n}\}$ and the exterior part $\mathbb{Z}_n^+ := \{k \mid 2\pi|k|\tau > \sqrt{n}\}$. The following holds:

(i) The series $\sum_{k\in\mathbb{Z}} c_k\Gamma(-c_k)$ and the integral $\int_\gamma s\Gamma(-s)$ are absolutely convergent.

(ii) The following integrals or sums are $O(1/n)$ for $n \to \infty$,

$$\int_{\gamma_n^+} n^s\Gamma(-s)ds\,, \quad \int_{\gamma_n^+} L_n(s)ds\,, \quad \sum_{k\in\mathbb{Z}_n^+} n^{c_k}c_k\Gamma(-c_k)\,, \quad \sum_{k\in\mathbb{Z}_n^+} L_n(c_k)\,,$$

(iii) The following estimates holds between the residues $A_n^{(1)}(c) = \mathrm{Res}[sL_n(s)) \cdot \Lambda(s); s = c]$ and $A_n^{(0)}(c) = \mathrm{Res}[n^s s\Gamma(-s)\Lambda(s); s = c]$,

$$A_n^{(1)}(c) = A_n^{(0)}(c) \left[1 + O\left(\frac{1}{n}\right)\right]. \tag{21}$$

(iv) The following estimates hold between the series of residues,

$$\Sigma_n^{(1)} := \sum_{k\in\mathbb{Z}} A_n^{(1)}(c_k), \qquad \Sigma_n^{(0)} := \sum_{k\in\mathbb{Z}} A_n^{(0)}(c_k) \qquad \Sigma_n^{(1)} = \Sigma_n^{(0)} \cdot \left[1 + O\left(\frac{1}{n}\right)\right]. \tag{22}$$

(v) The two integrals satisfy

$$\int_\gamma L_n(s)ds = \left[\int_\gamma n^s\Gamma(-s)ds\right]\left[1 + O\left(\frac{1}{n}\right)\right].$$

Part II. Study of the dominant parts. The dominant terms are brought by residues. For a meromorphic function F with a pole of order $b+1$ at $s = c$, the term $\mathrm{Res}[n^s sF(s); s = c]$ is computed from the singular expansion of $n^s \cdot F(s)$ at $s = c$, using the analytic expansion of n^s at $s = c$,

$$n^s = e^{s\log n} = n^c \sum_{i\geq 0} \frac{1}{i!}(s-c)^i \log^i n.$$

Then, the residue $\mathrm{Res}[n^s sF(s); s = c]$ is written as $n^c \cdot Q(\log n)$ with a polynomial Q of degree b. There are four cases, according to the shape: in cases (S) and (H), there is only one pole at c, whereas in case (P), all the points $c_k = c + 2i\pi k\tau$ are also poles. Moreover, the boolean $[[c = 1]]$ intervenes due to a supplementary pole brought by $\Gamma(-s)$ at $s = 1$.

Case (S) or (H). The two residues $A_n^{(1)}(c)$ and $A_n^{(0)}(c)$ are compared with (21) and:

(i) $c \neq 1$. As c is a pole of order $b+1$ of $\Lambda(s)$ at $s = c$, the function $sL_n(s)) \cdot \Lambda(s)$ has a pole c of order $b+1$, and $A_n^{(0)}(c) = n^c c\,\Gamma(-c)\,Q(\log n)$, with a polynomial Q of degree b.

(ii) $c = 1$. As $\Gamma(-s)$ has a pole at $s = 1$, the function $\Gamma(-s) \cdot \Lambda(s)$ has now a pole at $s = 1$ of order $b+2$, and $A_n^{(0)}(c) = n^c c\,Q(\log n)$, with a polynomial Q of degree $b+1$.

Case (P). The function Λ admits poles $c_k = c + 2i\pi k\tau$ (for $k \in \mathbb{Z}$) of order $b+1$. Then, there arise the two series $\Sigma_n^{(1)}$ and $\Sigma_n^{(0)}$ that are compared in (22).

(i) Case $c \neq 1$. In the series $\Sigma_n^{(0)}$, each pole c_k brings a term $A_n^{(0)}(c_k)$, where the polynomial Q is the same for each c_k (because $\Lambda(s)$ is periodic), and

$$\Sigma_n^{(0)} = \sum_{k\in\mathbb{Z}} n^{c_k} c_k\,\Gamma(-c_k)\,Q(\log n) = n^c Q(\log n)\Pi(n), \quad \Pi(n) := \left[\sum_{k\in\mathbb{Z}} n^{2i\pi k\tau} c_k\,\Gamma(-c_k)\right].$$

The function Π is absolutely convergent, due to exponential decreasing of the Γ function. Moreover, with the equality $n^{2i\pi k\tau} = e^{2i\pi k\tau\log n}$, the function Π is a periodic function (of period 1) of the variable $\tau\log n$, and finally the series of residues is

$$\Sigma_n^{(0)} = n^c\,Q(\log n)\,\Pi(\{\tau\log n\}), \qquad \Pi \text{ periodic of period 1, } Q \text{ polynomial of degree } b.$$

(ii) The pole $c = 1$ is now a pole of order $b+2$, whereas the other poles c_k (for $k \neq 0$) are of order $b+1$. We consider the pole $c = 1$ separately, together the expansion it brings, then the other poles for $k \neq 0$, and we obtain

$$\mathrm{Res}[n^s s\Gamma(-s)\Lambda(s); s = c] + \Sigma_n^{(0)} = n\left[R(\log n) + Q(\log n)\,\Pi(\{\tau\log n\})\right],$$

where R and Q are polynomials of resp. degree $b+1$ and b.

Finally, we check, in each of the four cases, the form of the dominant terms that is stated in Theorem 5.

Part III. Study of the remainder terms. The remainder terms are easily obtained in the cases (S) or (P) where the curve δ_c is a vertical line. In case (H), the remainder term is not precisely studied in the litterature except in the case of a memoryless aperiodic source, where the study is done in [12] and [4]. We recall the study here. On the hyperbolic curve δ, and letting $s = \sigma + it$ and $L := \log n$, the following estimates hold

$$|n^s| = n^\sigma = n^c \exp[-ALt^{-\nu}], \qquad |\Lambda(s)\Gamma(-s)| \leq \exp[-Kt]\,,$$

$$\text{and entails the bound} \qquad |n^s \Lambda(s)\Gamma(-s)| \leq n^c \exp[-Kt - ALt^{-\nu}]\,.$$

When n (and then L) is fixed, the minimum of the function $t \mapsto Kt + ALt^{-\nu}$ is reached for $t^{\nu+1} = \nu L/K$ and the maximum of the function is of order $n^c \exp[-(\log n)^\beta]$ with $\beta < 1/(1+\nu)$. Using the same principles as in Laplace's method, this entails the estimate

$$\int_\delta L_n(s)\Lambda(s) = n^c O\left(\exp[-(\log n)^\beta]\right)\,.$$

B.2 Proof of Lemma 15

Consider a delay $\gamma_{a,b}$ defined in (18)*. The associated rescaling $g_{a,b}$ via Relation (R) and its inverse $g_{a,b}^{-1}$ satisfy the following:*

(a) *the inverse g^{-1} satisfies*

$$g_{1,b}^{-1}(x) = \frac{1}{b+1} x^{b+1}\left[1 + O\left(\frac{1}{\ell}\right)\right]\,, \quad (a=1)\,; \quad g_{a,b}^{-1}(x) = \frac{a}{a-1} a^\ell \ell^b \left[1 + O\left(\frac{1}{\ell}\right)\right]\,, \quad (a>1)\,.$$

(b) *the function $g_{a,b}$ and the principal part $g_{a,b}^{(0)}$ satisfy*

$$g_{1,b}^{(0)}(x) = (b+1)^{1/(b+1)}\, x^{1/(b+1)}\,, \quad (a=1)\,; \qquad g_{a,b}^{(0)}(x) = \log_a\left[\frac{x}{(\log_a(x))^b}\right]\,, \quad (a>1)\,;$$

$$\textit{with} \quad g_{1,b}(x) = g_{1,b}^{(0)}(x) + L_b + O\left(\frac{1}{x}\right)\,; \qquad g_{a,b}(x) = g_{a,b}^{(0)}(x) + K_a + O\left(\frac{\log\log x}{\log x}\right)\,.$$

Proof. The proof has two steps.

Proof of Item (a). Expression de $\Gamma_{a,b}(\ell) = \sum_{k\leq \ell} \gamma_{a,b}(k)\,.$

$$\text{When } a = 1, \text{ one has} \quad \Gamma_{1,b}(\ell) := \sum_{k=1}^{\ell} k^b = \frac{1}{b+1}\ell^{b+1}\left[1 + O\left(\frac{1}{\ell}\right)\right]\,. \tag{23}$$

For the general case $a > 1$, we consider the sum, first without integer parts, then with integer parts. First, without integer parts,

$$\widehat{\Gamma}_{a,b}(\ell) := \sum_{k=1}^{\ell} a^k k^b = a^\ell \ell^b \cdot \Theta_{a,b}(\ell) \quad \text{with} \quad \Theta_{a,b}(\ell) = \sum_{k=0}^{\ell-1} a^{-k}\left(1 - \frac{k}{\ell}\right)^b\,.$$

Now, $\Theta_{a,b}(\ell)$ has a limit (when $\ell \to \infty$) equal to the geometric sum $\sum_{k\geq 0} a^{-k} = a/(1-a)$. The difference indeed decomposes as

$$\sum_{k\geq 0} a^{-k} - \Theta_{a,b}(\ell) = \sum_{k=0}^{\ell-1} a^{-k}\left[1 - \left(1 - \frac{k}{\ell}\right)^b\right] + \sum_{k\geq \ell} a^{-k}\,.$$

The second term tends to 0 for $\ell \to \infty$ (with exponential speed). As each term of the first term is less than $(k/\ell) \cdot b$, the first term is itself less than

$$\frac{b}{\ell} \sum_{k=0}^{\ell-1} k a^{-k} \leq \frac{b}{\ell} \sum_{k=0}^{\infty} k a^{-k}, \qquad \sum_{k=0}^{\infty} k a^{-k} < \infty\,,$$

and is $O(1/\ell)$ (for $\ell \to \infty$). Finally,

$$\sum_{k\geq 0} a^{-k} - \Theta_{a,b}(\ell) = O\left(\frac{1}{\ell}\right), \qquad \widehat{\Gamma}_{a,b}(\ell) = \frac{a}{a-1} a^\ell \ell^b \left[1 + O\left(\frac{1}{\ell}\right)\right].$$

We now consider the sums with integer parts. The difference with the previous case is

$$\sum_{k=1}^{\ell} \{a^k\} k^b = O(\ell^{b+1}) = a^\ell \ell^b O\left(\frac{\ell}{a^\ell}\right) = a^\ell \ell^b O\left(\frac{1}{\ell}\right),$$

$$\text{and finally for } a > 1 \qquad \Gamma_{a,b}(\ell) := \sum_{k=1}^{\ell} \lfloor a^k \rfloor k^b = \frac{a}{a-1} a^\ell \ell^b \left[1 + O\left(\frac{1}{\ell}\right)\right]. \tag{24}$$

Proof of Item (b). Expression for $g^{(0)}_{a,b}$.
With Relation (R) described in (11) and the estimates (24) and (23), we have to determine the principal part $g^{(0)}_{a,b}(x)$ of the function $g_{a,b}(x)$ for which the inverse $g^{-1}_{a,b}(x)$ admits the estimate

$$g^{-1}_{a,b}(x) = \frac{a}{a-1} a^x x^b \left[1 + O_b\left(\frac{1}{x}\right)\right] \quad (a > 1), \qquad g^{-1}_{1,b}(x) = \frac{1}{b+1} x^{b+1} \left[1 + O_b\left(\frac{1}{x}\right)\right].$$

We begin with the easy case $a = 1$, and obtain, for some constant L_b,

$$g_{1,b}(x) = (b+1)^{1/(b+1)} x^{1/(b+1)} + L_b + O\left(x^{-1/(b+1)}\right).$$

When $a > 1$, we check the estimate $\quad g_{a,b}(x) = \log_a(x) - b \log_a(\log_a x)) - \log_a\left(\frac{a}{a-1}\right),$

with the following remainder terms $\quad O\left(\frac{\log\log x}{\log x}\right) \quad (b \neq 0), \quad O\left(\frac{1}{\log x}\right) \quad (b = 0)\,.$ ◀

B.3 Proof of Proposition 17

Consider a source $\mathcal{P}$ of the Good Class, and the modified source $\mathcal{P}_{a,b}$ obtained from $\mathcal{P}$ by inserting delays of the form $\gamma_{a,b}$ with $a \geq 1$. Consider the real $\sigma_a \geq 1$ defined by the equation $a\lambda(s) = 1$ so that $\sigma_1 = 1$.
The Λ series of the source $\mathcal{P}_{a,b}$ has a pole of order $b+1$ located at σ_a, and is meromorphic on a neighborhood $\mathcal{W}_a$ of the real axis. With $\mathcal{W}$ defined in Section 4.3, there are two cases for $\mathcal{W}_a$: for $a = 1$, $\mathcal{W}_a = \mathcal{W}$. If $a > 1$, there exists $\tau_a < \sigma_a$ for which $\mathcal{W}_a = \mathcal{W} \cap \{\Re s > \max(1, \tau_a)\}$.

Proof. We will deal in the paper with the delay $\gamma_{a,b}$ already defined in (18), but we first study a delay $\pi_{a,b}$ closely related to $\gamma_{a,b}$ and defined as

$$\pi_{a,b}(\ell) = \lfloor a^\ell \rfloor \cdot \pi_b(\ell), \qquad \text{with} \quad \pi_b(\ell) := (\ell+1)(\ell+2)\dots(\ell+b)\,. \tag{25}$$

We will return at the end of the proof to the delay $\gamma_{a,b}$ with the relation

$$\ell^b = \sum_{c \leq b} \alpha_c \pi_c(\ell) \qquad \gamma_{a,b}(\ell) = \lfloor a^\ell \rfloor \ell^b = \sum_{c \leq b} \alpha_c \pi_{a,c}(\ell), \qquad \alpha_b = 1\,. \tag{26}$$

We first deal with the modified series (associated with π instead of γ),

$$\Lambda^{\langle a,b\rangle}(s) = \sum_{\ell} \pi_{a,b}(\ell) \Lambda_\ell(s) = \widehat{\Lambda}^{\langle a,b\rangle}(s) + \sum_{\ell} \{a^\ell\} \pi_b(\ell) \Lambda_\ell(s), \quad \widehat{\Lambda}^{\langle a,b\rangle}(s) = \sum_{\ell} \lfloor a^\ell \rfloor \pi_b(\ell) \Lambda_\ell(s)\,.$$

Case $a = 1$. With the definition of delay π, the equality holds, with (17), and the real neighborhood $\mathcal{W}$ defined there,

$$b! \left[\frac{1}{1-\lambda(s)}\right]^{b+1} = \sum_{\ell\geq 0} \pi_b(\ell)\lambda(s)^\ell, \qquad \sum_\ell \pi_b(\ell)\Lambda_\ell(s) = b!v(s)\left[\frac{1}{1-\lambda(s)}\right]^{b+1} + d(s)\,, \quad (27)$$

where $d(s) := \sum_\ell \pi_b(\ell)c_\ell(s)$ is analytic on $\mathcal{W}$. The series $\Lambda^{\langle 1,b\rangle}(s)$ is then meromorphic on $\{\Re s > 1\} \cup \mathcal{W}$ and has a pole of order $b+1$ at $s = 1$.

Case $a > 1$. The series $\widehat{\Lambda}^{\langle a,b\rangle}(s)$ is written as

$$\sum_\ell a^\ell \pi_b(\ell)\Lambda_\ell(s) = b!v(s)\left[\frac{1}{1-a\lambda(s)}\right]^{b+1} + d_a(s), \qquad \text{with} \quad d_a(s) = \sum_\ell a^\ell \pi_b(\ell)c_\ell(s)\,.$$

It has a dominant singularity (a pole of order $b+1$) at $s = \sigma_a$ where σ_a is the real s for which $a\lambda(s) = 1$. Near the real axis, using the neighborhood $\mathcal{W}$, there exists $\tau_a < \sigma_a$ for which $d_a(s)$ is analytic and $\widehat{\Lambda}^{\langle a,b\rangle}(s)$ is meromorphic on $\mathcal{W} \cap \{\Re s > \ \tau_a\}$. As the second series (with fractional parts) is analytic for $\Re s > 1$, the series $\Lambda^{\langle a,b\rangle}(s)$ is meromorphic on the domain $\{\Re s > \sigma_a\} \cup \mathcal{W}_a$ with $\mathcal{W}_a := \mathcal{W} \cap \{\Re s > \max(1, \tau_a)\}$ with a unique pole of order $b+1$ at $s = \sigma_a$.

We now return to the Λ series defined with delays $\gamma_{a,b}$ with (26). ◀

C Tameness

This section is devoted to tameness studies. First, in Section C.1, we focus on the memoryless case and recall the dichotomy between periodic memoryless sources and aperiodic ones. Then, Section C.2 describes the aperiodic memoryless case in the general case $a \geq 1$. The first case $a = 1$ was studied in works [12] and [20] and used in the proof of Theorem 14. The second case $a > 1$ is new (however in the same spirit as the case $a = 1$), and used in the proof of Theorem 18. Finally, Section C.3 studies the tameness of the source $\mathcal{P} = \texttt{CF}$, first for the parameters $(a = 1,\ b = 0)$ (where the Dolgopyat-Baladi-Cesaratto-Vallée result is recalled and used in the proof of Theorem 14) then for a close enough to 1 (where the previous study is extended and used in the proof of Theorem 18).

C.1 Generalities for memoryless sources

When the source $\mathcal{P}$ is memoryless, the Λ series of the source $\mathcal{P}_{a,b}$ decomposes as in (20). The second series (that does not appear for integer values of a) is analytic on $\Re s > 1$ and bounded on each vertical line $\Re s = 1 + \rho$ with $\rho > 0$. We focus on the first series (without integer parts), here in the memoryless case. We use the coefficients $\pi_{a,b}(\ell)$ defined in (25), the polynomial decomposition (26) and the identity (27). This gives in the memoryless case

$$\widehat{\Lambda}^{(\mathcal{P}_{a,b})}(s) = \sum_\ell a^\ell \ell^b \Lambda_\ell(s) = \sum_\ell a^\ell \sum_{c\leq b} \alpha_c \pi_c(\ell)\lambda(s)^\ell = \sum_{c\leq b} \alpha_c\, c! \left[\frac{1}{1-a\lambda(s)}\right]^{c+1}. \quad (28)$$

For a real $a \geq 1$, we consider the value σ_a defined by the implicit equation $a\lambda(s) = 1$.

Periodicity condition. We now recall conditions under which a memoryless source (p, q) is periodic. Assume that there is $\tau > 0$ for which the two equations hold

$$[p + q = 1, \quad p^{1+2i\pi\tau} + q^{1+2i\pi\tau} = 1] .$$

Then the triangular inequality holds

$$1 = \left|p^{1+2i\pi\tau} + q^{1+2i\pi\tau}\right| \leq \left|p^{1+2i\pi\tau}\right| + \left|q^{1+2i\pi\tau}\right| = p + q = 1 ,$$

and becomes an equality. This may occur only if there exists a real θ for which the equality holds, $q^{1+2i\pi\tau} = \theta \cdot p^{1+2i\pi\tau}$. This leads to the sequence of equalities

$$1 = p^{1+2i\pi\tau} + q^{1+2i\pi\tau} = (1 + \theta) \cdot p^{1+2i\pi\tau} = (1 + \theta) \cdot p = 1 ,$$

which entails $\quad p^{2i\pi\tau} = q^{2i\pi\tau} = 1 \quad$ and then $\quad \Big[(2i\pi\tau) \log p = 2i\pi k, \quad (2i\pi\tau) \log q = 2i\pi\ell\Big]$.
Finally, $(\log q/ \log p) = (k/\ell)$ for some pair of integers (k, ℓ) for which $\gcd(k, \ell) = 1$. Now, the equality satisfied by $e^{1/\tau}$, namely $1 = p + q = (e^{1/\tau})^k + (e^{1/\tau})^\ell$ proves that $e^{1/\tau}$ is an algebraic integer.

Periodic memoryless sources. If the source $\mathcal{P}$ is periodic, then the function $s \mapsto \lambda(s)$ is periodic of period $2i\pi\tau$ with some $\tau > 0$. This entails that the function $s \mapsto a\lambda(s)$ is periodic too with the same period, and the poles of $\widehat{\Lambda}^{\mathcal{P}_{a,b}}$ are located at points $\sigma_a + 2i\pi k\tau$. The second series (with fractional parts) has no poles on the hyperplane $\Re s > 1$. Then, the poles of $\Lambda^{\mathcal{P}_{a,b}}$ are located at points $\sigma_a + 2i\pi k\tau$.

C.2 Aperiodic memoryless sources

The pair (p, q) is here denoted as (p_1, p_2). We consider together the cases $a = 1$ and $a > 1$. In the case $a = 1$, the proof of [12] studies the set $\mathcal{Z} := \{s \mid \lambda(s) = 1, \quad s \neq 1\}$, and exhibits its hyperbolic shape defined via the irrationality exponent μ of $\beta = (\log p_2)/(\log p_1)$. We consider here, for any $a \geq 1$ the set

$$\mathcal{Z}_a := \{s \mid \lambda(s) = 1/a, \quad s \neq \sigma_a\} , \qquad \lambda(\sigma_a) = p_1^{\sigma_a} + p_2^{\sigma_a} = 1/a , \tag{29}$$

which gathers the poles of $\widehat{\Lambda}^{(\mathcal{P}_{a,b})}$ defined in (28). As previously, these poles will be the only poles of $\Lambda^{\mathcal{P}_{a,b}}$ on $\Re s > 1$. We wish to determine a region on the left of the vertical line $\Re s = \sigma_a$ but close to the vertical line $\Re s = \sigma_a$ where $\Lambda^{\mathcal{P}_{a,b}}$ is analytic: we look for a (non empty) region $\mathcal{R}_{\sigma_a}$ located "between" $\mathcal{Z}_a$ and the vertical line $\Re s = \sigma_a$, and defined by an equation of hyperbolic type as in Definition 12. We follow the general scheme of [12] given for $a = 1$ that we easily extend to a general $a \geq 1$.

An implicit equation. We let $\log p_1 := w_1$, $\log p_2 = w_2$, and consider rational approximations of the ratio w_2/w_1 (that do not depend –of course– on the real a), and thus integers q for which there exists an integer q_2 such that

$$v = q\frac{w_2}{w_1} - q_2 \quad \text{is small} \qquad \text{and thus} \quad w_2\left(2i\pi\frac{q}{w_1}\right) = 2i\pi v + 2i\pi q_2 \qquad (v \ \text{ small}) . \tag{30}$$

We focus on the points of $\mathcal{Z}_a$ close enough to the vertical line $\Re s = \sigma_a$ with an imaginary part close to a multiple of $(2\pi)/w_1$ of the form $q(2\pi)/w_1$. We then let :

$$s \in \mathcal{Z}_a, \qquad s = \sigma_a + \Delta_a + 2i\pi\frac{q}{w_1} ,$$

and focus on complex numbers s for which the complex number Δ_a has both a small real part and a small imaginary part. With (30), the two equalities hold,

$$\left\{ \begin{array}{lll} p_1^s & = \exp\left[w_1(\sigma_a+\Delta_a+2i\pi\frac{q}{w_1})\right] & = p_1^{\sigma_a+\Delta_a} \\ p_2^s & = \exp\left[w_2(\sigma_a+\Delta_a+2i\pi\frac{q}{w_1})\right] & = p_2^{\sigma_a+\Delta_a}\cdot\exp[2i\pi v] \end{array} \right. \tag{31}$$

$$\text{and entail the equality} \quad p_1^{\sigma_a+\Delta_a}+p_2^{\sigma_a+\Delta_a}\cdot\exp[2i\pi v]=\frac{1}{a}\,. \tag{32}$$

Then $\Delta_a=\Delta_a(v)$ is implicitely determined (via Eqn (32)) as a function of v, and Eqn (29) entails the equality $\Delta_a(0)=0$. With the complex version of the implicit theorem, the first two derivatives of Δ_a at $v=0$ involve two real numbers $\delta_a^{(1)}$ and $\delta_a^{(2)}>0$ under the form,

$$\Delta_a'(0)=i\delta_a^{(1)}, \qquad \Delta_a''(0)=-\delta_a^{(2)}\,,$$

and there exists, for any $q\in\mathbb{Z}$, for each (small) real number v, a point $s=\sigma+it$ of $\mathcal{Z}_a$, with

$$\sigma=\sigma_a-\delta_a^{(2)}v^2+O(v^4), \qquad t=2i\pi\frac{q}{w_1}+i\delta_a^{(1)}v+O(v^3)\,. \tag{33}$$

Diophantine approximations. We consider now the reals v which are associated with a pair (q,q_2) as in (30), and focus on rationals q_2/q which are a best diophantine approximation of $\beta=w_2/w_1$: this means that, for every rational number q_2'/q' different from q_2/q with $0<q'\le q$, one has

$$|q\beta-q_2|<|q'\beta-q_2'|\,.$$

We recall that ν is an irrationality exponent for β if, for any $\epsilon>0$, there is a finite number of rationals q_2/q for which

$$|q\beta-q_2|\le q^{1-\nu-\epsilon}\,.$$

Continued fractions theory proves the existence of an infinite number of rationals q_2/q for which $|q\beta-q_2|\le q^{-1}$. Then an irrationality exponent is at least equal to 2.
The irrationality exponent of β is the smallest possible irrationality exponent of β; it is denoted as $\mu(\beta)$. The inequality $\mu(\beta)\ge 2$ always holds.

Two curves. As in [12], we consider a ratio $\beta=w_2/w_1$ with an irrationality exponent $\mu(\beta)$, and we deal with v as in (30), now associated with a rational q_2/q which is a best diophantine approximation of $\beta=w_2/w_1$. Then the point $s=s(q)\in\mathcal{Z}_a$ in (33) is "close" to a curve

$$s=\sigma_a-C^{(a)}t^{-2\mu-2}\,, \qquad \text{for some constant } C_a\,.$$

More precisely, as in [12], the following holds, with the irrationality exponent $\mu=\mu(\beta)$:

- **(i)** *For any $\nu>\mu$, there exists $B_\nu^{(a)}>0$, for which all elements $s=\sigma+it$ of the set $\mathcal{Z}_a\cap\{\Im s\ge 1\}$ satisfy $\sigma\le\sigma_a-B_\nu^{(a)}t^{-2\nu-2}$.*
- **(ii)** *For any $\theta<\mu$, there exist $A_\theta^{(a)}>0$, and an infinite set of elements $s=\sigma+it$ of $\mathcal{Z}_a$ such that $\sigma\ge\sigma_a-A_\theta^{(a)}t^{-2\theta-2}$.*

Part (i) of the result provides the free of poles region $\mathcal{R}_{\sigma_a}$ we look for, with its frontier. Part (ii) says that this region is in a sense optimal.

Tameness. Consider $\nu' = \nu(1+\epsilon)$ with $\nu > \mu$ and the curve $\mathcal{C}$ defined by the equation $s = \sigma_a - B_\nu^{(a)} t^{-2\nu'-2}$. Any point $s = \sigma + it$ of this curve is at a distance δ of $\mathcal{Z}_a$, with

$$\delta > B_\nu^{(a)} \left[\frac{1}{t^{2\nu+2}} - \frac{1}{t^{2\nu'+2}} \right] = B_\nu^{(a)} \frac{t^{2(\nu'-\nu)} - 1}{t^{2\nu'+2}} \geq \frac{1}{t^{2\nu'+2}}$$

for $|t|$ large enough. Moreover, for any $s = \sigma + it$ that is at a distance δ from a point of $\mathcal{Z}_a$, the following inequality holds and proves that Λ is of polynomial growth on the curve $\mathcal{C}$.

$$|\Lambda(s)| \leq \frac{1}{|\lambda'(\sigma_a)|} \frac{1}{\delta} = \Theta(t^{2\nu'+2})$$

C.3 Tameness in the CF case

We begin with the decomposition (20). We first focus on the first series and use the Dolgopyat-Baladi-Cesaratto-Vallée results:

For $\mathcal{P} =$ CF, there is a "truncated" vertical strip

$$\mathcal{S} := \{s = \sigma + it \mid |\sigma - 1| \leq \theta, \quad |t| \geq t_0, \quad t_0 > 0\},$$

where the $\Lambda_\ell(s)$ series satisfy for any $\ell \geq 1$, the bound

$$|\Lambda_\ell(s)| \leq \rho^\ell \cdot |t|^\xi \text{ for some } \rho < 1 \text{ and some } \xi > 0.$$

We first give sufficient conditions on a under which σ_a belongs to the strip S. As the relation holds,

$$(a-1)/a = |\lambda(1) - \lambda(\sigma_a)| \sim |\sigma_a - 1||\lambda'(1)|\,,$$

a sufficient condition for a under which σ_a belong to S is $a < 1/(1-\theta|\lambda'(1)|)$.
When s belongs to the vertical strip S, and as soon as $a\rho \leq \delta < 1$, one has

$$|\widehat{\Lambda}(s)| \leq \sum_\ell a^\ell \ell^b |\Lambda_\ell(s)| \leq \left(\sum_\ell \ell^b a^\ell \rho^\ell \right) \cdot |t|^\xi \leq \left(\sum_\ell \ell^b \delta^\ell \right) \cdot |t|^\xi,$$

and using a similar decomposition as in (28) one obtains

$$\widehat{\Lambda}(s) \leq A_\delta \cdot |t|^\xi \qquad \text{with} \quad A_\delta := \left(\sum_\ell \ell^b \delta^\ell \right) = \sum_{c \leq b} |\alpha_c| \cdot c! \left(\frac{1}{1-\delta} \right)^{c+1}.$$

Finally, when a satisfies the two following conditions, for some $\delta < 1$, that involve the geometric characteristics on the "free of poles" strip of the source $\mathcal{P}$,

$$a \leq \min\left(\frac{\delta}{\rho}, \frac{1}{1-\theta|\lambda'(1)|} \right),$$

the real σ_a belongs to S, the series $\widehat{\Lambda}(s)$ is tame at σ_a, with a tameness strip

$$\{s = \sigma + it \mid 1 - \theta < \sigma < 1 + \theta\}\,.$$

However, as $a \neq 1$, the second series (with fractional parts) appears: it is analytic on $\Re s > 1$ and bounded on each vertical line $\Re s = 1 + \kappa$ with $\kappa > 0$. Finally, a tameness strip for the total series $\Lambda(s)$ is $\{s \mid 1 + \kappa \leq \sigma < 1 + \theta\}\,,$ with $1 + \kappa < \sigma_a < 1 + \theta\,.$
We have to compare the two terms, the principal term n^{σ_a} and the remainder term $n^{1+\kappa} = n^{\sigma_a} \cdot n^{1-\sigma_a+\kappa}$ where κ is any positive number such that $1 + \kappa \in]1, \sigma_a[$. With barycentric coordinates, we have $1 + \kappa = \mu + (1-\mu)\sigma_a$ with $\mu \in]0,1[$ so that

$$1 - \sigma_a + \kappa = \mu + (1+\mu)\sigma_a - \sigma_a = \mu(1-\sigma_a) = -\mu(\sigma_a - 1)\,.$$

Then, the multiplicative remainder term is $n^{1-\sigma_a+\kappa} = n^{-\mu(\sigma_a-1)}$, for any $\mu \in]0,1[$.

D Comparisons between tries. Computational version of a source

D.1 Direct comparison between tries of the two sources

We first prove an inequality related to a conjecture which appears at the end of [3]. This conjecture states the equality in (34). We will exhibit in the following cases when the equality does not hold.

▶ **Proposition 23.** *Consider a delay γ, the associated rescaling g via Relation (R), (see Eqn (11)) and its inverse g^{-1}. If the function g^{-1} is convex, the following inequality holds between the average trie depth $\mathbb{E}[D_n^{\langle\gamma\rangle}]$ of a trie built on the source $\mathcal{P}_{\langle\gamma\rangle}$ and the average trie depth $\mathbb{E}[D_n]$ of a trie built on the source $\mathcal{P}$,*

$$\mathbb{E}[D_n^{\langle\gamma\rangle}] \geq g^{-1}\left(\mathbb{E}[D_n]\right) . \tag{34}$$

Proof. For a sequence $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathcal{I}^n$, we consider the two sequences of expansions

$$M(\boldsymbol{x}) = (M(x_1), M(x_2) \ldots M(x_n)), \quad M^{\langle\gamma\rangle}(\boldsymbol{x}) = (M^{\langle\gamma\rangle}(x_1), M^{\langle\gamma\rangle}(x_2) \ldots M^{\langle\gamma\rangle}(x_n)) .$$

We wish to compare the two tries $\texttt{Trie}\left(M(\boldsymbol{x})\right)$ and $\texttt{Trie}\left(M^{\langle\gamma\rangle}(\boldsymbol{x})\right)$, and notably their depths

$$D_n(\boldsymbol{x}) = D_n\left[\texttt{Trie}(M(\boldsymbol{x}))\right], \qquad D_n^{\langle\gamma\rangle}(\boldsymbol{x}) = D_n\left[\texttt{Trie}\left(M^{\langle\gamma\rangle}(\boldsymbol{x})\right)\right] .$$

Due to the definition of the source $\mathcal{P}^{\langle\gamma\rangle}$ in terms of the source $\mathcal{P}$, and definition of $\texttt{Trie}\left(M^{\langle\gamma\rangle}(\boldsymbol{x})\right)$, the branches of the two tries are related, as we now see: we denote by $M_k(x)$ the prefix of the expansion $M(x)$ of length k, and by $M_k^{\langle\gamma\rangle}(x)$ the prefix of the expansion $M^{\langle\gamma\rangle}(x)$ of length k. Remark first that a comparison between two words of $\mathcal{P}^{\langle\gamma\rangle}$ always terminates on a symbol in Σ. Remark also that the prefixes of $M^{\langle\gamma\rangle}(x)$ that end with a symbol in Σ are exactly those of length equal to $g^{-1}(k)$ for some k. Then, for $i \in [1..n]$, the i-th branch of the second trie $\texttt{Trie}\left(M^{\langle\gamma\rangle}(\boldsymbol{x})\right)$ is thus written for some $k_i = k_i(\boldsymbol{x})$ as $M^{\langle\gamma\rangle}_{g^{-1}(k_i)}(x_i)$ (and has a length $g^{-1}(k_i(\boldsymbol{x}))$) whereas the corresponding i-th branch of the first trie $\texttt{Trie}\left(M(\boldsymbol{x})\right)$ is $M_{k_i}(x_i)$ (and has a length $k_i(\boldsymbol{x})$). Then the depths of the two tries are respectively

$$D_n(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} k_i(\boldsymbol{x}), \qquad D_n^{\langle\gamma\rangle}(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} g^{-1}(k_i(\boldsymbol{x})) .$$

The depth D_n of a trie for which the lengths of branches are $D_n^{(i)}$ is defined via its law

$$\Pr[D_n \geq k] = \frac{1}{n}\sum_{i=1}^{n} \Pr[D_n^{(i)} \geq k], \qquad \text{so that} \quad \mathbb{E}[D_n] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[D_n^{(i)}] .$$

The expectations of the two depths are thus

$$\mathbb{E}[D_n] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[k_i(\boldsymbol{x})] \qquad \mathbb{E}[D_n^{\langle\gamma\rangle}] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[g^{-1}(k_i(\boldsymbol{x})] .$$

When g^{-1} is a convex function, the following inequalities hold and entail the stated bound,

$$g^{-1}\left(\mathbb{E}[D_n]\right) = g^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[k_i(\boldsymbol{x})]\right) \leq \frac{1}{n}\left(\sum_{i=1}^{n} g^{-1}\left(\mathbb{E}[k_i(\boldsymbol{x})]\right)\right)$$

$$g^{-1}\left(\mathbb{E}[k_i(x)]\right) \leq \mathbb{E}[g^{-1}(k_i(\boldsymbol{x}))] . \qquad \blacktriangleleft$$

The present paper exhibits a class of delays of the form $\gamma_{a,b}(\ell) = a^\ell \ell^b$ for which the inverse of the weight is

$$g^{-1}(x) = \begin{cases} g_{1,b}^{-1}(x) & = \Theta\left(x^{b+1}\right) \quad (a=1)\,, \\ g_{a,b}^{-1}(x) & = \Theta\left(a^x x^b\right) \quad (a>1)\,, \end{cases} \tag{35}$$

and is convex for $x \geq 1$.
As our starting sources are tame sources $\mathcal{P}$ for which Theorem 5 applies with ($c=1$, $b=0$), the depth $\mathbb{E}[D_n]$ satisfies, with $h = |\lambda'(1)|$,

$$\mathbb{E}[D_n] = A_n\,[1+O(\epsilon_n)] \qquad A_n := \frac{1}{h}\log n\,, \quad \epsilon_n = \frac{1}{\log n}\,;$$

For $n \to \infty$ and, for any $g := g_{a,b}$, the estimates $\mathbb{E}[D_n] \sim A_n$ lead to the estimates $g^{-1}(\mathbb{E}[D_n]) = \Theta\left(g^{-1}(A_n)\right)$. Then, using the equality $a^{(1/h)\log n} = n^{(1/h)\log a}$), one obtains

$$g^{-1}(\mathbb{E}[D_n]) = \Theta\left(g^{-1}(A_n)\right) = \begin{cases} \Theta\left((\log n)^{b+1}\right) & (a=1)\,, \\ \Theta\left(n^{(1/h)\log a}\,(\log n)^b\right) & (a>1)\,. \end{cases} \tag{36}$$

Theorem 19 of the present paper exhibits the following estimates for $\mathbb{E}[D_n^{\langle\gamma\rangle}]$ that involve the real σ_a given by the equation $a\lambda(\sigma_a) = 1$,

$$\mathbb{E}[D_n^{\langle\gamma\rangle}] = \begin{cases} \Theta\left((\log n)^{b+1}\right) & (a=1)\,, \\ \Theta\left(n^{\sigma_a - 1}\right)(\log n)^b) & (a>1)\,. \end{cases} \tag{37}$$

Then, in the case $a > 1$, there are two exponents (a priori) distinct:

- **(i)** an exponent equal to $(1/h)\log a$ in the estimate of $g^{-1}(\mathbb{E}[D_n])$
- **(ii)** an exponent $\sigma_a - 1$ in the estimate of $E[D_n^{\langle\gamma\rangle}]$.

With Proposition 23, we prove (in an indirect way) the inequality $\sigma_a - 1 \geq \frac{1}{h}\log a$. For an unbiased memoryless source of cardinality r with entropy $\log r$, the equality holds, $[\sigma_a - 1 = \log a/\log r]$, and the two exponents are the same. We now directly compare the two exponents :

▶ **Lemma 24.** *The following inequality holds between the two exponents,*

$$\sigma_a - 1 \geq \frac{1}{h}\log a\,,$$

and the equality holds only in the case of a source conjugated to an unbiased memoryless source.

Proof. Inside the proof, we use the notation $\sigma_a = \sigma(a)$. The function $x \mapsto \sigma(x)$ is defined through the implicit equation $x\lambda(\sigma(x)) = 1$ that involves the dominant eigenvalue $s \mapsto \lambda(s)$. The derivative $\sigma'(x)$ satisfies $\lambda(\sigma(x)) + x\sigma'(x)\lambda'(\sigma(x)) = 0$ and thus

$$\sigma'(x) = \frac{1}{x}\left[-\frac{\lambda(\sigma(x))}{\lambda'(\sigma(x))}\right] = \frac{1}{x}\left[\frac{-1}{L'(\sigma(x))}\right]\,,$$

where $L(s)$ is the pressure function equal to $L(s) = \log\lambda(s)$. The function L is always convex, so that $u \mapsto L'(u)$ is increasing. Then, the function $u \mapsto -1/L'(u)$ is thus increasing, too, and as the function $\sigma(x)$ is increasing, this entails that the function $\tau : x \mapsto \frac{-1}{L'(\sigma(x))}$ is also increasing, and satisfies, for $u \geq 1$ the inequality $\tau(u) \geq \tau(1)$ with $\tau(1) = \frac{-1}{L'(1)} = \frac{1}{h}$. Thus the inequality holds,

$$\sigma(a) - 1 = \int_1^a \frac{1}{u}\,\tau(u)du \geq \frac{1}{h}\int_1^a \frac{1}{u}du = \frac{1}{h}\log a\,.$$

The inequalities become equalities only in the case when L' is linear, which occurs only when the source is conjugated to an unbiased memoryless source. ◀

D.2 Computational version of a good source

The computation of the successive symbols of the word $M(x)$ indeed involves the computation of the intervals $(\mathcal{I}_1(x), \mathcal{I}_2(x) \ldots, \mathcal{I}_\ell(x), \ldots)$. The computation of the ℓ-th digit of the word $M(x)$, when the first $\ell - 1$ digits (which form the prefix w) have been already computed, continues inside the interval $\mathcal{I}_{\ell-1}(x) = \mathcal{I}_w$. We already know the prefix w of length $\ell - 1$, and, as this interval $\mathcal{I}_w$ is the union of intervals $\mathcal{I}_{w\cdot\sigma}$ for $\sigma \in \Sigma$, we have to compare x and the end points of the possible intervals $\mathcal{I}_{w\cdot\sigma}$, in order to output the new digit σ for which x belongs to $\mathcal{I}_{w\cdot\sigma}$.

We have then to deal with the convenient precision $\pi(\ell)$ (on the real x and on the end points $a_\ell(x)$ and $b_\ell(x)$ of $\mathcal{I}_\ell(x)$) that is needed for an exact comparison between the end points and the input x. We deal in the classical model where there is an oracle which *freely* gives, for each real y and a given precision π, a rational approximation of y whose numerator p and denominator q are integers, with a number π of binary digits, for which $|y - p/q| \leq 2^{-\pi}$. We freely ask the oracle, but, when the rational approximations are given by the oracle, we have to compute with the integers p, q we are given; now, the computation is not free, and depends on the binary size of the rationals, and thus on π.

The main questions are: first, how is the integer $\pi(\ell)$ compared to the depth ℓ? second, what are the main operations on the integers (numerators and denominators) to be done? For our reference sources, of positive entropy, the needed precision is $\pi(\ell) = \Theta(\ell)$ and the main operations used in the computation of the shift T are additions and multiplications. Then, for our reference sources, the arithmetic complexity for computing a_ℓ when $(a_1, a_2, \ldots, a_{\ell-1})$ is already computed is of the form $\Theta(\pi(\ell)^b) = \Theta(\ell^b)$ with $b = 2$.

On the Independence Number of Random Trees via Tricolourations

Etienne Bellin
Ecole Polytechnique, Palaiseau, France

Abstract

We are interested in the independence number of large random simply generated trees and related parameters, such as their matching number or the kernel dimension of their adjacency matrix. We express these quantities using a canonical tricolouration, which is a way to colour the vertices of a tree with three colours. As an application we obtain limit theorems in L^p for the renormalised independence number in large simply generated trees (including large size-conditioned Bienaymé-Galton-Watson trees).

2012 ACM Subject Classification Mathematics of computing → Random graphs

Keywords and phrases Independence number, simply generated tree, Galton-Watson tree, tricolouration

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.2

Acknowledgements I am grateful to Igor Kortchemski for his careful reading of the manuscript and for telling me Frederic Chapoton's suggestion to consider canonical tricolourations of random trees. I am also grateful to the anonymous referees and their useful remarks.

1 Introduction

A subset S of vertices of a finite graph G is called an *independent set* if there is no pair of connected vertices in S. The *independence number* of G, denoted by $I(G)$, is the biggest cardinal of an independent set of G. The independence number is a well studied quantity in computational complexity theory. It is known that computing the independence number is NP-hard in general (see e.g. [10, Sec. 3.1.3]). A lot of work has been carried out to describe algorithms computing the independence number in general graphs [19, 22] and in special classes of graphs where the computational time can be decreased (e.g. cubic graphs [21], claw-free graphs [17], P_5-free graphs [15]). The independence number has also received interest in combinatorics and in probability. Upper bounds have been found using probabilistic methods for cubic graphs [3]. Asymptotics have been found in certain classes of random trees (e.g. conditioned Bienaymé-Galton-Watson trees [8], simply generated trees [2], random recursive trees and binary search trees [9], and a wider class of random trees constructed from a Crump–Mode–Jagers branching process [12]). Finally we mention three articles giving applications of the independence number in scheduling theory [13], coding theory [4] and collusion detection in voting pools [1].

The goal of this article is to study the independence number of large simply generated trees, generalising some results of [7] and [2]. Simply generated trees are a wide class of random plane trees (i.e. rooted and ordered trees) introduced in [16] and encompass Bienaymé-Galton-Watson trees (BGW trees for short) conditioned to have a fixed number of vertices. Informally, a BGW tree with offspring distribution μ is a plane tree where vertices have an i.i.d. number of children with law μ. Various natural models of random trees are obtained with appropriate choices of the offspring distribution: e.g., uniform plane trees, uniform plane d-ary trees and uniform Cayley trees (see [11, Sec. 10]). In order to study the independence number, we will use a particular tricolouration of trees introduced in [23] and

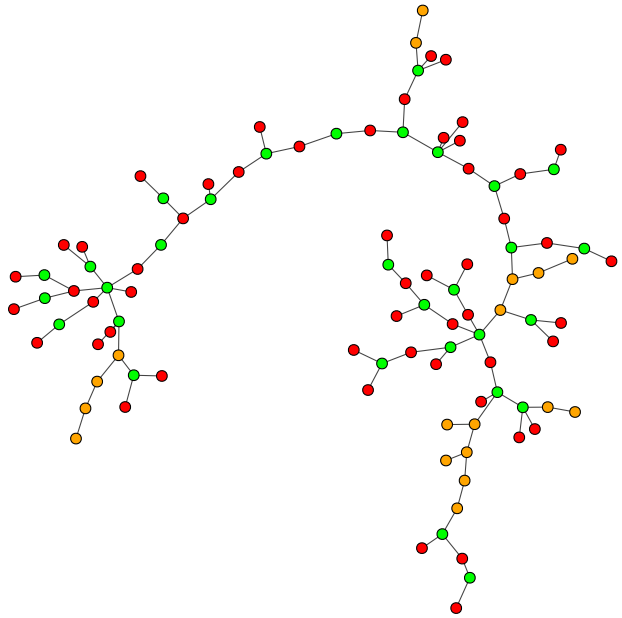


Figure 1 Tricolouration of a BGW tree with 100 vertices and a Poisson offspring distribution of parameter 1. The algorithm used to tricolour this tree can be found in [5, Appendix A].

later studied in [6], [7] and [5]. This colouring is based on the notion of covering. A *covering* of a finite tree T is a subset of vertices S of T such that every edge of T is adjacent to a vertex of S. A *smallest covering* of T is a covering with minimal cardinality. In general, a tree has more than one smallest covering. For every vertex v of T we colour v in the following way:

- If v belongs to every smallest covering, we colour v in green.
- If v belongs to no smallest covering, we colour it in red.
- If v belongs to some smallest coverings but not all, we colour it in orange.

For a tree T, denote by $n_g(T)$, $n_o(T)$ and $n_r(T)$, respectively, the number of green, orange and red vertices in T. It has been noticed in [7] that the size of a smallest covering of a tree T is equal to $n_g(T) + n_o(T)/2$. Since the complementary of a smallest covering is an independent set of maximal size, the independence number of T is $n_r(T) + n_o(T)/2$. Actually, other statistics of the tree T can be expressed as a linear combination of $n_g(T)$, $n_o(T)$ and $n_r(T)$. For instance the *matching number* $M(T)$ (i.e. the maximum size of a partial vertex matching) is equal to the size of a smallest covering which is $n_g(T) + n_o(T)/2$. The *nullity* $N(T)$ (i.e. the kernel dimension of the adjacency matrix) is $n_r(T) - n_g(T)$. The *edge cover number* and the *clique cover number* also coincide with the independence number on trees (see [9]). Our main result (Theorem 4) concerns simply generated trees, but, to keep this introduction short, let us state here a particular case for critical BGW trees.

▶ **Theorem 1.** *Let T_n be a Bienaymé-Galton-Watson tree with reproduction law μ, conditioned on having n vertices. Denote by $G(t) := \sum_{k=0}^{\infty} \mu_k t^k$ the generating function of μ and let q be the unique solution of $G(1-q) = q$ in $[0,1]$. Suppose that μ has mean 1, then, the following convergences hold in L^p for every $p > 0$:*

$$\frac{n_g(T_n)}{n} \xrightarrow[n\to\infty]{L^p} \frac{1-q+(1-2q)G'(1-q)}{1+G'(1-q)}, \qquad \frac{n_o(T_n)}{n} \xrightarrow[n\to\infty]{L^p} \frac{2\,q\,G'(1-q)}{1+G'(1-q)},$$

$$\frac{n_r(T_n)}{n} \xrightarrow[n\to\infty]{L^p} \frac{q}{1+G'(1-q)}.$$

Explicit computations of the expected number of green, orange and red vertices have been carried out in the case of a uniform Cayley tree with fixed size in [7] using generating functions. This confirms the convergence, in mean value, of Theorem 1. Indeed, it is well known that a BGW tree with Poisson distribution of parameter 1 conditioned to have n vertices has the same law as a uniform Cayley tree with n vertices. To our knowledge, there is no estimates or asymptotics, other than [7], for the expected number of green, orange and red vertices in simply generated trees. As we said before, different quantities such as the independence number $I(T)$, the matching number $M(T)$ and the nullity $N(T)$ of a tree T can be expressed in terms of $n_g(T)$, $n_o(T)$ and $n_r(T)$, so Theorem 1 yields limit theorems for these in the case of critical BGW trees.

▶ **Corollary 2.** *With the same notation and hypothesis as in Theorem 1, the following convergences hold in L^p for every $p > 0$:*

$$\frac{I(T_n)}{n} \xrightarrow[n\to\infty]{L^p} q, \qquad \frac{M(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 1-q, \qquad \frac{N(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 2q-1.$$

With the same hypothesis, the authors of [8] show the convergence of $I(T_n)/n$ in probability towards q. Moreover, in [2] the convergence of the first and second moment of $I(T_n)/n$ is studied. More precisely, it is shown that $\mathbb{E}\left[I(T_n)\right] = nq + O(1)$ and $V(I(T_n)) = \nu n + O(1)$ for some constant ν. Therefore, the convergence of the independence number, in the settings of Corollary 2 is not new, but Corollary 5 generalises this convergence for simply generated trees that are not equivalent to conditioned critical BGW trees. The main tool to prove Theorem 4 is the use of limit theorems for uniformly pointed simply generated trees found in [20] (see Section 4). In the first section we introduce simply generated trees. In the next section we state our main result in its most general form. To prove our main result, we first explain the limit theorems of [20] in the third section. The next two sections give properties of the tricolouration and describe how to colour the limiting trees. Finally, in the last section we prove Theorem 4.

2 Simply generated trees

Let $\mathbf{w} := (w_i)_{i\geq 0}$ be a sequence of nonnegative weights. A *simply generated tree* having n vertices with weight sequence $\mathbf{w}$ is a random plane tree T_n such that for every finite plane tree T,

$$\mathbb{P}\left(T_n = T\right) = \frac{1}{Z_n}\left(\prod_{v\in T} w_{k_v}\right)\mathbb{1}_{|T|=n}$$

where k_v is the outdegree of the vertex v in T, $|T|$ is the number of vertices in T and Z_n is the normalising constant defined by

$$Z_n := \sum_{|T|=n}\prod_{v\in T} w_{k_v}.$$

Notice that, when the weight sequence $\mathbf{w}$ is actually a probability sequence (i.e. the sum of the weights is equal to 1) then we recover the class of BGW trees. For T_n to be well defined, one needs Z_n to be nonzero. First of all, suppose that $w_0 > 0$ and $w_k > 0$ for some $k \geq 1$ otherwise $Z_n = 0$ for all $n \geq 1$. Let $\mathrm{span}(\mathbf{w}) := \gcd\{i \geq 0 \,|\, w_i > 0\}$ (since $w_k > 0$ this quantity is well defined). The following result, found for instance in [11, Cor. 15.6], characterises the n's such that $Z_n > 0$:

▶ **Lemma 3** (Janson 2012). *If $Z_n > 0$ then $n \equiv 1 \mod \mathrm{span}(\mathbf{w})$. Conversely, there exists n_0 such that for all $n \geq n_0$ satisfying $n \equiv 1 \mod \mathrm{span}(\mathbf{w})$, $Z_n > 0$.*

Throughout this document and in Theorem 4, we suppose that $w_0 > 0$, $w_k > 0$ for some $k \geq 1$ and that all the n's appearing satisfy $n \geq n_0$ and $n \equiv 1 \mod \mathrm{span}(\mathbf{w})$

Let $\rho \in [0, +\infty]$ be the radius of convergence of the generating series

$$\phi(x) := \sum_{i \geq 0} w_i x^i.$$

It is shown in [11, Lemma 3.1] that, if $\rho > 0$, then the function defined by

$$\psi(x) := \frac{x\phi'(x)}{\phi(x)}$$

is increasing on $[0, \rho)$ and we can define $\nu := \lim_{x \to \rho} \psi(x) \in (0, +\infty]$. We distinguish three different regimes:

- Regime 1 when $\rho > 0$ and $\nu \geq 1$. In this case there is a unique $\tau \in [0, \rho]$ such that $\tau < +\infty$ and $\psi(\tau) = 1$.
- Regime 2 when $\rho > 0$ and $0 < \nu < 1$. In this case $\rho < +\infty$ and we set $\tau := \rho$.
- Regime 3 when $\rho = 0$.

In regime 1 and 2 we can define a probability function given by

$$\pi_k := \frac{\tau^k w_k}{\phi(\tau)}. \tag{1}$$

The associated mean and generating function are respectively given by

$$m := \min(1, \nu) \qquad \text{and} \qquad G(x) := \frac{\phi(\tau x)}{\phi(x)}. \tag{2}$$

An important result of [11] is that, T_n, in regime 1 or 2, has the same law as a BGW tree with reproduction law π conditioned to have n vertices. In regime 3, T_n is not distributed like a conditioned BGW tree. Note that a critical or super-critical BGW tree or a BGW tree with a reproduction law with infinite mean, conditioned to have n vertices, always lays in regime 1. Moreover for a critical BGW tree with reproduction law μ, the probability π is the same as μ. A sub-critical BGW tree, conditioned to have n vertices, is either in regime 1 or in regime 2. We define *complete condensation* to be the condition:

$$\Delta(T_n) = (1 - m)n + n\mathcal{E}_n \tag{3}$$

where $\Delta(T_n)$ is the maximum degree of a vertex of T_n and $\mathcal{E}_n$ is a random variable converging in probability towards 0. For instance, complete condensation happens in regime 2 when there exists $\theta > 1$ and a slowly varying function ℓ such that $\pi_k = \ell(k)k^{-(1+\theta)}$ (see [14]). Complete condensation also happens in regime 3 for the weight sequence $w_k = k!^\alpha$ for $\alpha > 0$ (see [11, Ex. 19.36]).

3 Main results

In this section we state our main results. We keep all the notations and assumptions of Section 2.

▶ **Theorem 4.** *Let T_n be a simply generated tree with n vertices according to the weight sequence $\mathbf{w} = (w_i)_{i\geq 0}$. Recall that, in regime 1 and 2, G denotes the generating function of π defined in (1). Let q be the unique solution of $q = G(1-q)$ in $[0,1]$.*

1. *In regime 1 and in regime 2 with complete condensation (meaning that (3) is satisfied), the following convergences hold in L^p for every $p > 0$:*

$$\frac{n_g(T_n)}{n} \xrightarrow[n\to\infty]{L^p} \frac{1-q+(1-2q)G'(1-q)}{1+G'(1-q)}, \qquad \frac{n_o(T_n)}{n} \xrightarrow[n\to\infty]{L^p} \frac{2\,q\,G'(1-q)}{1+G'(1-q)},$$

$$\frac{n_r(T_n)}{n} \xrightarrow[n\to\infty]{L^p} \frac{q}{1+G'(1-q)}.$$

2. *In regime 3 with complete condensation, the following convergences hold in L^p for every $p > 0$:*

$$\frac{n_g(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 0, \qquad \frac{n_o(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 0, \qquad \frac{n_r(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 1.$$

▶ **Corollary 5.** *We keep the same notation and hypothesis as in Theorem 4. Recall that $I(T_n)$, $M(T_n)$ and $N(T_n)$ are, respectively, the independence number, the matching number and the nullity of T_n.*

1. *In regime 1 and in regime 2 with complete condensation, the following convergences hold in L^p for every $p > 0$:*

$$\frac{I(T_n)}{n} \xrightarrow[n\to\infty]{L^p} q, \qquad \frac{M(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 1-q, \qquad \frac{N(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 2q-1.$$

2. *In regime 3 with complete condensation, the following convergences hold in L^p for every $p > 0$:*

$$\frac{I(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 1, \qquad \frac{M(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 0, \qquad \frac{N(T_n)}{n} \xrightarrow[n\to\infty]{L^p} 1.$$

4 Limit theorems for uniformly pointed simply generated trees

In this section we explain the results proved in [20] which will be our basic tool to prove Theorem 4. All the proofs and details of this section can be found in the above mentioned article. As said in the introduction, these results are limit theorems for uniformly pointed simply generated trees. A *pointed tree* is simply a couple (T, v) with a plane tree (i.e. rooted and ordered tree) T and a distinguished vertex v of T. A uniformly pointed simply generated tree is a couple (T_n, v_n) where T_n is a simply generated tree with n vertices and v_n is a distinguished vertex chosen uniformly at random among the n vertices of T_n. Basically, in regime 1 and in regime 2 and 3 with complete condensation, the local tree structure around v_n converges towards an infinite random tree which depends only on the regime. To formally define the notion of convergence used here, one needs to consider T_n to be a subtree of a big ambient tree denoted by $\mathcal{U}_\infty^\bullet$. Every plane tree (e.g. T_n) is considered, by definition, to be rooted, however we will encounter some infinite trees without any root which is unusual in the classic framework of plane trees. Let

$$\mathcal{V}_\infty := \{\emptyset\} \cup \bigcup_{n\geq 1} (\mathbb{N}^*)^n$$

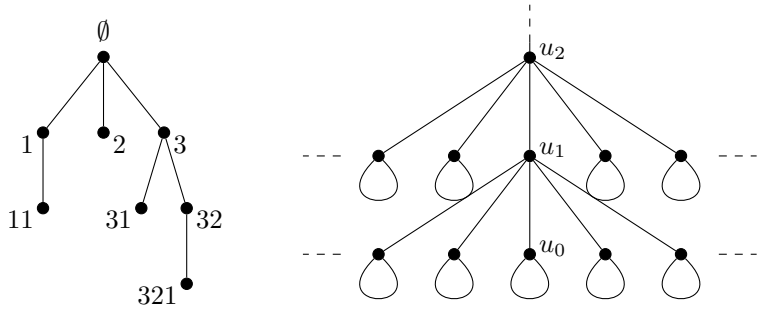


Figure 2 On the left, a subtree of the Ulam-Harris tree $\mathcal{U}_\infty$. On the right, the tree $\mathcal{U}_\infty^\bullet$. Each loop represents a copy of the Ulam-Harris tree.

be the set of words (empty word included) formed in the alphabet $\mathbb{N}^* = \{1, 2, 3, \dots\}$. Usually, plane trees are defined as subtrees of the so-called *Ulam-Haris tree*, denoted here by $\mathcal{U}_\infty$, which is the tree with vertex set $\mathcal{V}_\infty$ and edge set $\{(a_1 \dots a_{n-1}; a_1 \dots a_n) \,|\, \forall n, a_1, \dots, a_n \in \mathbb{N}^*\}$. With this definition, all the plane trees have a root which is a common ancestor to every vertex of the tree (it is the vertex designated by the empty word $\emptyset$). However, in regime 1, the root of T_n is, in a local point of view, at infinite distance from the distinguished vertex v_n. It suggests that the local limit of T_n around v_n has an infinite spine of ancestors and therefore, has no root. This is why we need a more general framework than the usual one for plane trees. Here we explain informally the construction of $\mathcal{U}_\infty^\bullet$. Let $u_0, u_1, \dots$ be vertices, in the plane, lined up to form an infinite connected spine. Each vertex u_i with $i > 0$ gets an infinite countable number of children on the left and on the right of its child u_{i-1}. Then, all the leaves of the current tree (u_0 included) give birth to the Ulam-Harris tree $\mathcal{U}_\infty$. The tree we obtain from this construction is denoted by $\mathcal{U}_\infty^\bullet$ and its set of vertices is denoted by $\mathcal{V}_\infty^\bullet$ (see Figure 2). To formally define this tree one could start by creating the vertex set $\mathcal{V}_\infty^\bullet$ as a subset of $\mathbb{N} \times \mathbb{Z} \times \mathcal{V}_\infty$ and then describing the edge set. However we think that the above informal construction is enough for our purpose. As we said, we want to represent T_n as a subtree of $\mathcal{U}_\infty^\bullet$. First we make clear what we call a subtree of $\mathcal{U}_\infty$ and $\mathcal{U}_\infty^\bullet$. Denote by $\mathcal{E}_\infty$ the edge set of $\mathcal{U}_\infty$ and $\mathcal{E}_\infty^\bullet$ the edge set of $\mathcal{U}_\infty^\bullet$.

▶ **Definition 6.** *A subtree t of $\mathcal{U}_\infty$ is a tree with vertex set included in $\mathcal{V}_\infty$ and edge set included in $\mathcal{E}_\infty$, such that the vertex $\emptyset$ belongs to t and such that there are no holes in t, meaning that: if $v = (a_1 \dots a_n)$ is a vertex of t with $a_n > 0$ then $(a_1 \dots a_n - 1)$ is also a vertex of t. Similarly a subtree t of $\mathcal{U}_\infty^\bullet$ is a tree with vertex set included in $\mathcal{V}_\infty^\bullet$ and edge set included in $\mathcal{E}_\infty^\bullet$, such that the vertex u_0 belongs to t and such that there are no holes in t (see Figure 5). A subtree of $\mathcal{U}_\infty^\bullet$ is rooted if the set $\{k \geq 0 \,|\, u_k \in t\}$ is finite. In this case the vertex u_k with maximal k in t is called the root of t.*

The representation of T_n as a subtree of $\mathcal{U}_\infty^\bullet$ will obviously depend on the distinguished vertex v_n since we want to look at the local structure around this vertex. More precisely, let T be a plane tree and v be a vertex of T. We identify the distinguished vertex v with the element u_0 and the root of T is identified with the element u_h where h is the graph distance between the root and v in T (it is the height of v). All the other vertices of T are identified such that the plane order is preserved (See Figure 3). We denote by (T, v) the subtree of $\mathcal{U}_\infty^\bullet$ obtained this way. We can now formally define the notion of convergence we use.

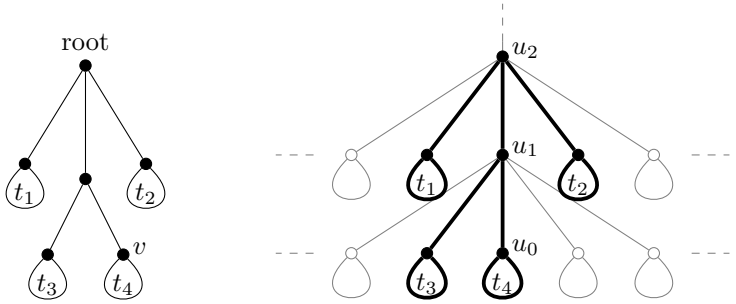


Figure 3 On the left, a finite tree T with a distinguished vertex v at distance 2 from the root. The loops t_1, t_2, t_3 and t_4 represent subtrees of T which can be seen as subtrees of the Ulam-Harris tree $\mathcal{U}_\infty$. On the right, the representation of T as a subtree of $\mathcal{U}_\infty^\bullet$.

▶ **Definition 7.** *Let t_n and t be subtrees of $\mathcal{U}_\infty^\bullet$ for all n. We say that (t_n) converges towards t and write $t_n \to t$ if for all $v \in \mathcal{V}_\infty^\bullet$,*

$$\mathbb{1}_{v \in t_n} \xrightarrow[n\to\infty]{} \mathbb{1}_{v \in t}.$$

This notion of convergence induces a topology that is metrizable and compact over the set of subtrees of $\mathcal{U}_\infty^\bullet$. Before stating the limit theorems, one needs to define the limiting trees T_1^*, T_2^* and T_3^* (seen as subtrees of $\mathcal{U}_\infty^\bullet$) that correspond, respectively, to regime 1, 2 and 3. Let T be a BGW tree with reproduction law π, given by (1), in regime 1 or 2. Let $\hat{\pi}$ be the probability measure on $\mathbb{N} \cup \{\infty\}$ given by

$$\hat{\pi}_k = k\pi_k \quad \forall k \in \mathbb{N} \quad \text{and} \quad \hat{\pi}_\infty = 1 - m$$

where $m := \min(1, \nu)$ is the mean of π.

- First we define T_1^* in regime 1. Notice that in this case $\hat{\pi}_\infty = 0$. We attach to u_0 an independent copy of T. For $k \geq 1$, u_k receives offspring according to an independent copy of $\hat{\pi}$. Then u_{k-1} is identified with a child of u_k chosen uniformly at random. Finally, we attach an independent copy of T to all the children of u_k, except u_{k-1} (see Figure 4).
- Now we define T_2^* in regime 2. In this case $\hat{\pi}_\infty > 0$. We attach to u_0 an independent copy of T. For $k \geq 1$, u_k receives offspring according to an independent copy of $\hat{\pi}$. Notice that almost surely, there exist $1 \leq i < j$ two integers such that $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{j-1}$ have a finite number of children and u_i and u_j have an infinite number of children. For every $k \in \{1, \ldots, i-1, i+1, \ldots, j-1\}$, u_{k-1} is identified with a child of u_k chosen uniformly at random, while u_i gets infinitely many children on the left and the right of its child u_{i-1}. Finally, for all $k \geq 1$, we attach an independent copy of T to all the children of u_k, except u_{k-1}. The tree T_2^* is the tree obtained by keeping all the descendants of u_{j-1} (see Figure 4).
- Finally, T_3^* is simply composed of the vertex u_1 having infinitely many children on the left and on the right of u_0, all of them, including u_0, being leaves (see Figure 4).

▶ **Theorem 8** (Stufler 2018). *Let (T_n, v_n) be a uniformly pointed simply generated tree with n vertices. Suppose that we are in regime $i = 1$ or in regime $i \in \{2, 3\}$ with complete condensation (meaning that (3) is satisfied). Then the convergence*

$$(T_n, v_n) \xrightarrow[n\to\infty]{(d)} T_i^*$$

holds in distribution for the topology induces by the convergence of Definition 7.

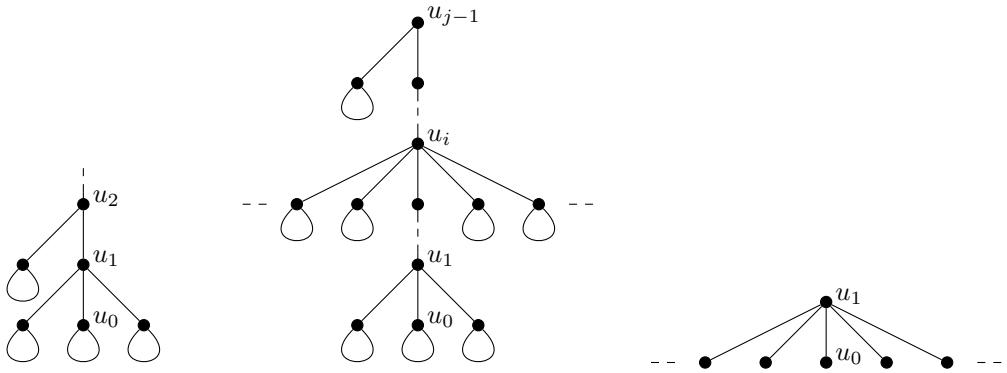


Figure 4 Representation of the trees T_1^*, T_2^* and T_3^*, respectively, from left to right. Each loop represents a copy of a Bienaymé-Galton-Watson tree of reproduction law π.

5 Properties of the tricolouration

In this section, we look at some general properties of the tricolouration defined in the introduction that will be essential to prove our main result. Let $T_1, \dots, T_n$ be n rooted finite trees. We define $T_1 * \cdots * T_n$ the rooted tree obtained by creating an edge between each root of $T_1, \dots, T_n$ and a new vertex which will be the root of $T_1 * \cdots * T_n$. In particular the number of vertices $\#V(T_1 * \cdots * T_n)$ equals $1 + \#V(T_1) + \cdots + \#V(T_n)$. And the number of edges $\#E(T_1 * \cdots * T_n)$ equals $n + \#E(T_1) + \cdots + \#E(T_n)$. We say that a rooted tree has colour c if the root has colour c.

▶ **Lemma 9.** *Let $T_1, \dots, T_n$ be n rooted finite trees. Set $T := T_1 * \cdots * T_n$, then*

1. *T is red if $T_1, \dots, T_n$ are all non-red.*
2. *T is orange if exactly one tree among $T_1, \dots, T_n$ is red.*
3. *T is green if two or more trees among $T_1, \dots, T_n$ are red.*

Proof. Denote by $C(T)$ the size of a smallest covering of T. Notice that $C(T)$ is equal to $C(T_1) + \cdots + C(T_n) + \Delta$ with $\Delta \in \{0, 1\}$. More precisely, $C(T) = C(T_1) + \cdots + C(T_n)$ if and only if all the T_i's are non-red.

1. Suppose that $T_1, \dots, T_n$ are all non-red. We can take a smallest covering for each T_i such that the root of T_i is included in the covering. Then the union of these coverings gives a smallest covering of T. Moreover we can see that all the smallest coverings of T are obtained this way. Thus T is red.
2. Suppose than T_1 is red and $T_2, \dots, T_n$ are all non-red. We can take a smallest covering for each T_i in addition to the root of T. This gives a smallest covering of T, so T is either green or orange. We can also take a smallest covering for each T_i, $i > 1$, such that the root of T_i is included in the covering, a smallest covering of T_1 and the root of T_1. This also gives a smallest covering of T. Thus T is orange.
3. Suppose that T_1 and T_2 are red. As for the previous case, we can take a smallest covering for each T_i in addition to the root of T. This gives a smallest covering for T. But, as opposed to the previous case, all the smallest coverings of T are obtained this way. Thus T is green. ◀

If T_1, T_2 are finite trees and v_1, v_2 are vertices of, respectively, T_1 and T_2, then we denote by $(T_1, v_1) * (T_2, v_2)$ the tree obtained from T_1 and T_2 by drawing an edge between v_1 and v_2.

▶ **Lemma 10.** *With the same notation as above, set $T := (T_1, v_1) * (T_2, v_2)$. If v_1 is green in the tricolouration of T_1 then, the colour of every vertex v in the tricolouration of T is just the same as its colour in the tricolouration of T_1 (if v is a vertex of T_1) or T_2 (if v is a vertex of T_2). In other words, the tricoloured tree T is simply obtained by drawing an edge between v_1 and v_2 and keeping the colours of T_1 and T_2.*

Proof. Notice that, since v_1 is green in the tricolouration of T_1, a smallest covering of T_1 combined with a smallest covering of T_2 gives a smallest covering of T. Conversely a smallest covering of T is necessarily obtained by combining a smallest covering of T_1 and T_2. ◀

6 Tricolouration of the infinite limiting trees

In this section, we extend the definition of the tricolouration given in the introduction to the random infinite limiting trees T_1^*, T_2^* and T_3^* defined in Section 4. The initial definition applies only to finite trees since a covering of smallest size only makes sense in this context. Even though it seems not obvious to find a satisfactory definition of "smallest covering" for an infinite tree, it is still possible to describe a canonical way to tricolour the trees T_1^*, T_2^* and T_3^* using the properties found in Section 5.

Let t be a subtree of $\mathcal{U}_\infty^\bullet$ (finite or not) such that for every $k \geq 0$ and for every child v of u_k, distinct from u_{k-1}, v has a finite number of descendants. In other words, for all $k \geq 0$, all the children of u_k, distinct from u_{k-1}, are roots of finite trees. Since those trees are finite, it makes sense to consider their tricolouration. A *good* vertex of t is a vertex u_k with $k \geq 0$ such that, at least two of its children, distinct from u_{k-1}, are red in the tricolouration of the finite subtree they produce. If t is finite, then, from Lemma 9, a good vertex is a green vertex for the tricolouration of t. Notice that T_1^*, T_2^* and T_3^* satisfy the same hypothesis as t almost surely.

- We begin with the definition of the tricolouration of T_1^*. Almost surely, there exists an increasing sequence $(k_i)_i$ such that for all i, u_{k_i} is good. For all i, all the vertices below u_{k_i} (u_{k_i} included) get the same colour in T_1^* as their colour in the tricolouration of the subtree rooted at u_{k_i} (which is a finite tree). Notice that, from Lemma 9, u_{k_i} gets necessarily the colour green. Lemma 10 ensures that this way of colouring is consistent when taking larger i.
- For T_2^* we colour the unique vertex with infinite degree in green. Then, by cutting this vertex from T_2^* we obtain a (infinite) forest of finite trees who gets their induced tricolouration. Notice that, almost surely, the vertex with infinite degree is good.
- Lastly, all the leaves of T_3^* are coloured in red and the root u_1 is coloured in green.

We finish this section with the following lemma which explicitly gives the colour distribution of the vertex u_0 in T_i^*. This lemma will be useful when proving Theorem 4.

▶ **Lemma 11.** *Let $p_i(c)$ be the probability that u_0 has colour c in T_i^*.*

1. *In regime $i = 1$ and $i = 2$ with complete condensation we have*

$$p_i(green) = \frac{1 - q + (1 - 2q)G'(1 - q)}{1 + G'(1 - q)}, \qquad p_i(orange) = \frac{2\,q\,G'(1 - q)}{1 + G'(1 - q)},$$

$$p_i(red) = \frac{q}{1 + G'(1 - q)}.$$

2. *In regime 3 with complete condensation we have that $p_3(red) = 1$.*

Proof. The case of regime 3 is obvious, let us focus on regime 1 and 2. Let T be a BGW tree with reproduction law π. Denote by q the probability that the root of T is red. From Lemma 9 we deduce that

$$q = \sum_{k\geq 0} \pi_k (1-q)^k = G(1-q).$$

Let $\widetilde{T}$ be the tree obtained from T_i^* by cutting the edge between u_0 and u_1 and keeping the component containing u_1. Let $\widetilde{q}$ be the probability that u_1 is red in $\widetilde{T}$. Then, from Lemma 9 again,

$$\widetilde{q} = \sum_{k\geq 1} k\pi_k (1-q)^{k-1}(1-\widetilde{q}) = (1-\widetilde{q})G'(1-q).$$

Finally,

$$p_i(\text{red}) = \sum_{k\geq 0} \pi_k (1-q)^k (1-\widetilde{q}) = \frac{q}{1+G'(1-q)}.$$

And

$$p_i(\text{orange}) = \sum_{k\geq 0} \pi_k (1-q)^k \widetilde{q} + \sum_{k\geq 1} k\pi_k (1-q)^{k-1} q(1-\widetilde{q}) = \frac{2qG'(1-q)}{1+G'(1-q)}.$$

Finally we deduce the value $p_i(\text{green})$ by the law of total probability. ◀

7 Proof of Theorem 4

All this section is devoted to the proof of Theorem 4. We keep all the notation of Theorem 4 and suppose that we are in regime $i = 1$ or in regime $i \in \{2, 3\}$ with complete condensation. Let c be a colour in {red, green, orange}. Recall that $p_i(c)$ is the probability that the vertex u_0 has colour c in the tree T_i^* in regime i. The idea is to prove the convergence of the first two moments of $n_c(T_n)/n$, namely

$$\frac{1}{n}\mathbb{E}\left[n_c(T_n)\right] \xrightarrow[n\to\infty]{} p_i(c) \quad \text{and} \quad \frac{1}{n^2}\mathbb{E}\left[n_c(T_n)^2\right] \xrightarrow[n\to\infty]{} p_i(c)^2.$$

Then, using Lemma 12, we will conclude that $n_c(T_n)/n$ converges in L^p towards $p_i(c)$ for all $p > 0$. Actually, the convergence of the second moment won't be required in regime 3. Recall that the explicit computation of $p_i(c)$ can be found in Lemma 11.

▶ **Lemma 12.** *Let (X_n) be a sequence of random variables with values in $[0,1]$, and $\alpha \in [0,1]$. Suppose that one of the following condition is satisfied.*

- *The convergences $\mathbb{E}[X_n] \to \alpha$ and $\mathbb{E}\left[X_n^2\right] \to \alpha^2$ hold when $n \to \infty$.*
- *The convergence $\mathbb{E}[X_n] \to \alpha$ holds when $n \to \infty$ and $\alpha = 1$.*

Then for all $p > 0$, (X_n) converges towards α in L^p.

Proof of Lemma 12. First, we show that (X_n) converges towards α in probability. Let $\varepsilon > 0$. In the first case we use Markov's inequality which gives

$$\mathbb{P}\left(|X_n - \alpha| \geq \varepsilon\right) \leq \frac{\mathbb{E}\left[(X_n-\alpha)^2\right]}{\varepsilon^2} \xrightarrow[n\to\infty]{} 0.$$

In the second case we notice that

$$\mathbb{E}[X_n] \leq (1-\varepsilon)\mathbb{P}\left(X_n \leq 1-\varepsilon\right) + \mathbb{P}\left(X_n > 1-\varepsilon\right) = 1 - \varepsilon\mathbb{P}\left(X_n \leq 1-\varepsilon\right).$$

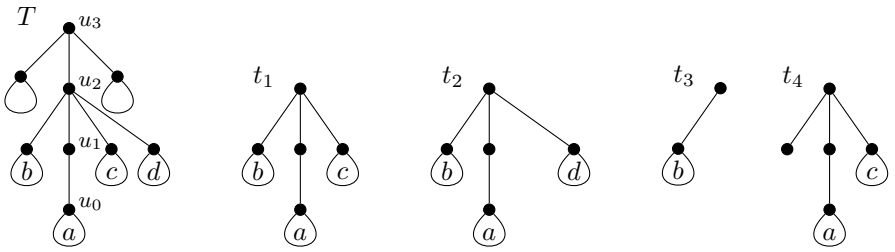


Figure 5 Illustration of Definitions 6 and 13. Only the tree t_1 satisfies $t_1 \preceq T$. The tree t_3 is not even a subtree of $\mathcal{U}_\infty^\bullet$ since it doesn't contain u_0. The tree t_2 is not a subtree of $\mathcal{U}_\infty^\bullet$ either since it has a hole between u_1 and the rightmost child of u_2. Finally t_4 is a subtree of $\mathcal{U}_\infty^\bullet$ but doesn't satisfy $t_4 \preceq T$ because the descendants of the left child of u_2 are missing.

Consequently $\lim_{n\to\infty} \mathbb{P}(X_n \leq 1-\varepsilon) = 0$ and the convergence in probability is shown in both cases. Second, let ℓ be an accumulation point of the sequence $(\mathbb{E}[|X_n - \alpha|^p])_n$ and $(n_k)_k$ be an extraction such that the convergence to ℓ occurs. From $(X_{n_k})_k$ we can extract a subsequence that converges almost surely to α. From the dominated convergence theorem we deduce that $\ell = 0$. ◀

Let v_n, v_n' be vertices chosen independently and uniformly in T_n. Notice that

$$\frac{1}{n}\mathbb{E}[n_c(T_n)] = \mathbb{P}(v_n \text{ has colour } c \text{ in } T_n) \quad \text{and}$$

$$\frac{1}{n^2}\mathbb{E}\left[n_c(T_n)^2\right] = \mathbb{P}(v_n \text{ and } v_n' \text{ have colour } c \text{ in } T_n).$$

Convergence of the first moment

First we prove the convergence of the first moment. Recall the notation of Section 6 when defining the tricolouration of the infinite trees T_1^*, T_2^* and T_3^*. Denote by $k \geq 0$ the first positive integer such that u_k is good ($k = 1$ almost surely in regime 3). Let τ_i^* be the subtree of $\mathcal{U}_\infty^\bullet$ obtained from T_i^* by cutting the edge between u_k and u_{k+1} and keeping the component containing u_0 ($\tau_3^* = T_3^*$ in regime 3). Note that, by construction, the tricolouration of τ_i^* is the restriction of its tricolouration in T_i^* and that u_k is green. The following definition introduces a useful order relation between trees.

▶ **Definition 13.** *Let T and t be subtrees of $\mathcal{U}_\infty^\bullet$ such that t is rooted at u_j for some $j \geq 0$. Suppose that u_j is also a vertex of T and denote by E_j the set of edges of T adjacent to u_j. We write $t \preceq T$ if there exists a subset $e_j \subset E_j$ such that t is the tree obtained from T by cutting all the edges from e_j and keeping the component containing u_j (see Figure 5).*

Let $\mathcal{F}$ be the set of rooted subtrees t of $\mathcal{U}_\infty^\bullet$ such that the root $u_j \in t$ is the only good vertex of t. For $t \in \mathcal{F}$ such that the root u_j of t has finite degree, denote by $v_\ell(t)$ (resp. $v_r(t)$) the leftmost (resp. rightmost) child of the root u_j of t. Let $\mathcal{F}_0$ be the set of elements $t \in \mathcal{F}$ such that: the root u_j of t has finite degree ; u_j has exactly two red neighbors distinct from u_{j-1} ; and ($v_\ell(t) = u_{j-1}$ or $v_\ell(t)$ is red) and ($v_r(t) = u_{j-1}$ or $v_r(t)$ is red). Notice that if t_1 and t_2 are distinct elements of $\mathcal{F}_0$, then we can't have $t_1 \preceq t_2$ nor $t_2 \preceq t_1$. Moreover for every $t_1 \in \mathcal{F}$ there exists a unique $t_2 \in \mathcal{F}_0$ such that $t_2 \preceq t_1$. In other words $\mathcal{F}_0$ is the set of equivalence classes for the equivalence relation $t_1 \sim t_2$ iff $t_1 \preceq t_2$ or $t_2 \preceq t_1$. Notice that almost surely $\tau_i^* \in \mathcal{F}$.

Fix $\varepsilon > 0$. Let $\mathcal{T}$ be a finite subset of $\mathcal{F}_0$ such that the event $\{\mathcal{T} \preceq \tau_i^*\} := \{\exists t \in \mathcal{T},\, t \preceq \tau_i^*\}$ happens with probability at least $1-\varepsilon$. Let $\mathcal{T}(c)$ be the set of trees $t \in \mathcal{T}$ such that u_0 has colour c in t. Remember that we see (T_n, v_n) as a subtree of $\mathcal{U}_\infty^\bullet$. For all $t \in \mathcal{T}$, define the event $A_n(t) := \{t \preceq (T_n, v_n)\}$. Using Theorem 8, we have that for all $t \in \mathcal{T}$

$$\mathbb{P}(A_n(t)) \xrightarrow[n\to\infty]{} \mathbb{P}(t \preceq T_i^*). \tag{4}$$

The properties of τ_i^* and $t \in \mathcal{T}$ imply that $t \preceq T_i^*$ if and only if $t \preceq \tau_i^*$. Thus

$$\mathbb{P}(t \preceq T_i^*) = \mathbb{P}(t \preceq \tau_i^*).$$

Denote by E_n the event $\cup_{t\in\mathcal{T}} A_n(t)$. Notice that the event $E_n \cap \{v_n \text{ has colour } c \text{ in } T_n\}$ is equal to the event $\cup_{t\in\mathcal{T}(c)} A_n(t)$. It is a consequence of Lemma 9 and 10. Notice also that for t_1, t_2 distinct trees of $\mathcal{T}$, $A_n(t_1)$ and $A_n(t_2)$ are disjoint for all n. Consequently,

$$\mathbb{P}(\{v_n \text{ has colour } c \text{ in } T_n\} \cap E_n) = \sum_{t\in\mathcal{T}(c)} \mathbb{P}(A_n(t)) \xrightarrow[n\to\infty]{} \mathbb{P}(\mathcal{T}(c) \preceq \tau_i^*).$$

Since u_0 has colour c in τ_i^* if and only if u_0 has colour c in T_i^*,

$$\mathbb{P}(\mathcal{T}(c) \preceq \tau_i^*) = \mathbb{P}(u_0 \text{ has colour } c \text{ in } T_i^* \text{ and } \mathcal{T} \preceq \tau_i^*) \in [p_1(c) - \varepsilon, p_1(c) + \varepsilon].$$

Finally, using Theorem 8 again, one has

$$\mathbb{P}(E_n) \xrightarrow[n\to\infty]{} \mathbb{P}(\mathcal{T} \preceq \tau_i^*) \geq 1 - \varepsilon.$$

The convergence $\mathbb{E}[n_c(T_n)/n] \to p_1(c)$ readily follows.

Convergence of the second moment in regime 1 and 2

The next step is to show convergence for the second moment in regime $i = 1$ or $i = 2$ with complete condensation. Let c' be another colour in {red, green, orange}. We will actually show that

$$\frac{1}{n^2}\mathbb{E}[n_c(T_n) n_{c'}(T_n)] \xrightarrow[n\to\infty]{} p_i(c) p_i(c').$$

We keep the notation of the previous part which shows the convergence of the first moment. For all $t \in \mathcal{T}$, let $A'_n(t) := \{t \preceq (T_n, v'_n)\}$ and $E'_n := \cup_{t\in\mathcal{T}} A'_n(t)$. We have,

$$\mathbb{P}(\{v_n \text{ has colour } c \text{ in } T_n\} \cap E_n \cap E'_n) = \sum_{t\in\mathcal{T}(c)} \sum_{t'\in\mathcal{T}(c')} \mathbb{P}(A_n(t) \cap A'_n(t')). \tag{5}$$

Fix $t, t' \in \mathcal{T}$. Recall that the trees t, t' and T_n are rooted plane trees, thus we can consider their so-called Łukasiewicz walk. More precisely, let T be a rooted plane tree with n vertices and w be a vertex of T. Denote by $\ell(w, T)$ the rank of w in T for the lexicographic order. Equivalently, w is the $\ell(w,T)$-th vertex of T explored by the depth first search starting from the root of T. Let $w_1, \ldots, w_n$ be the vertices of T ordered according to the lexicographic order (so $\ell(w_i, T) = i$ for all i). The *Łukasiewicz walk* associated with T is the sequence $(s_k)_{1\leq k\leq n}$ such that $s_0 = 0$ and $s_k - s_{k-1} + 1$ is the out-degree of w_k for all $k \in \{1, \ldots, n\}$. An important property of the Łukasiewicz walk is that it uniquely encodes its tree, meaning that the tree T can be retrieved from $(s_k)_{1\leq k\leq n}$. Let $(S_k^{(n)})_{0\leq k\leq n}$ be the Łukasiewicz walk associated with the tree T_n. Let $X_1, \ldots, X_n, \ldots$ be i.i.d random variables such that $\mathbb{P}(X_1 = m) = \pi_{m+1}$ for all integer $m \geq -1$ and set $S_k := \sum_{i=1}^k X_i$ for all $k \geq 0$. It is well known that, the random

walk $(S_k)_{0\le k\le n}$, starting at 0 and conditioned on reaching -1 for the first time at time n, has the same law as $(S_k^{(n)})_{0\le k\le n}$. Let $m := |t|$ be the number of vertices of t and $m' := |t'|$. Let $(s_k)_{0\le k\le m}$ and $(s'_k)_{0\le k\le m'}$ be, respectively, the Łukasiewicz walks associated with t and t' and denote by $x_k := s_{k+1} - s_k$ and $x'_k := s'_{k+1} - s'_k$ the associated steps. Write $k_0 := \ell(u_0, t)$, $k'_0 := \ell(u_0, t')$, $i_n := \ell(v_n, T_n)$ and $i'_n := \ell(v'_n, T_n)$. The indices i_n and i'_n are independent random elements of $\{1, \dots, n\}$ with uniform distribution. The event $A_n(t)$ happens if and only if the Łukasiewicz walk $(S_k^{(n)})_{0\le k\le n}$ coincides with $(s_k)_{0\le k\le m}$, up to a vertical shifting, on the interval $[\![i_n - k_0, i_n + m - k_0]\!]$. The same goes for $A'_n(t')$. More precisely

$$A_n(t) \cap A'_n(t') = \{X^{(n)}_{i+i_n-k_0} = x_i \ \forall i \in [\![1, m]\!] \text{ and } X^{(n)}_{i+i'_n-k'_0} = x'_i \ \forall i \in [\![1, m']\!]\}.$$

Applying the reverse Vervaat transform, we can change the initial excursion type conditioning into a bridge type conditioning (see e.g. [18, Sec. 6.1]). Namely

$$\begin{aligned}&\mathbb{P}\left(A_n(t) \cap A'_n(t')\right)\\ &\qquad = \mathbb{P}\left(X_{i+i_n-k_0} = x_i \ \forall i \in [\![1, m]\!] \text{ and } X_{i+i'_n-k'_0} = x'_i \ \forall i \in [\![1, m']\!] \,\middle|\, X_n = -1\right).\end{aligned}$$

Denote by D_n the event $\{[\![i_n - k_0 + 1, i_n - k_0 + m]\!] \cap [\![i'_n - k'_0 + 1, i'_n - k'_0 + m']\!] = \emptyset\}$. This event has a probability tending to 1 and one can see that

$$\begin{aligned}&\mathbb{P}\left(A_n(t) \cap A'_n(t') \cap D_n\right)\\ &\qquad = \mathbb{P}\left(X_i = x_i \ \forall i \in [\![1, m]\!] \text{ and } X_{i+m} = x'_i \ \forall i \in [\![1, m']\!] \,\middle|\, X_n = -1\right) \mathbb{P}\left(D_n\right).\end{aligned}$$

According to [11, Thm. 11.7] the steps $X_1, \dots, X_{m+m'}$ conditioned on $\{X_n = -1\}$ are asymptotically independent and the conditioning fades for large values of n, consequently

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left(A_n(t) \cap A'_n(t')\right) &= \mathbb{P}\left(X_i = x_i \ \forall i \in [\![1, m]\!]\right) \mathbb{P}\left(X_i = x'_i \ \forall i \in [\![1, m']\!]\right)\\ &= \lim_{n\to\infty} \mathbb{P}\left(A_n(t)\right) \mathbb{P}\left(A'_n(t')\right).\end{aligned}$$

Finally, using (4) and (5), we have that

$$\lim_{n\to\infty} \mathbb{P}\left(\{v_n \text{ has colour } c \text{ in } T_n\} \cap E_n \cap E'_n\right) = \mathbb{P}\left(\mathcal{T}(c) \preceq \tau_i^*\right) \mathbb{P}\left(\mathcal{T}(c) \preceq \tau_i^*\right).$$

and the result follows.

References

1. F. Araujo, J. Farinha, P. Domingues, G.C. Silaghi, and D. Kondo. A Maximum Independent Set Approach for Collusion Detection in Voting Pools. *Journal of Parallel and Distributed Computing*, 71(10):1356–1366, 2011.
2. C. Banderier, M. Kuba, and A. Panholzer. Analysis of three graph parameters for random trees. *Random Structures & Algorithms*, 35(1):42–69, 2009.
3. B. Bollobás. The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, 83(2):433–436, 1981.
4. S. Butenko, P. Pardalos, I. Sergienko, V. Shylo, and P. Stetsyuk. Finding maximum independent sets in graphs arising from coding theory. In *Proceedings of the 2002 ACM Symposium on Applied Computing*, SAC '02, pages 542–546, New York, NY, USA, 2002. Association for Computing Machinery.
5. F. Chapoton. On some varieties associated with trees. *Michigan Mathematical Journal*, 64(4):721–758, 2015.

6 S. Coulomb. Minimal vertex covers of random trees. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(06):P06007, June 2005.

7 S. Coulomb and M. Bauer. On vertex covers, matchings and random trees, 2004. `arXiv:math/0407456`.

8 L. Devroye, M.K. Goh, and R.Y. Zhao. The independence number of a Bienaymé-Galton-Watson tree and related parameters, 2021. `arXiv:2106.14389`.

9 M. Fuchs, C. Holmgren, D. Mitsche, and R. Neininger. A note on the independence number, domination number and related parameters of random binary search trees and random recursive trees. *Discrete Applied Mathematics*, 292:64–71, 2021.

10 M.R. Garey and D.S. Johnson. *Computers and intractability*, volume 174. freeman San Francisco, 1979.

11 S. Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probability Surveys*, 9(none):103–252, 2012.

12 S. Janson. On the independence number of some random trees. *Electronic Communications in Probability*, 25(none):1–14, 2020.

13 C. Joo, X. Lin, J.Ryu, and N.B. Shroff. Distributed greedy approximation to maximum weighted independent set for scheduling with fading channels. In *Proceedings of the Fourteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, MobiHoc '13, pages 89–98, New York, NY, USA, 2013. Association for Computing Machinery.

14 I. Kortchemski. Limit theorems for conditioned non-generic Galton–Watson trees. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 51(2):489–511, 2015.

15 D. Lokshtanov, M. Vatshelle, and Y. Villanger. Independent set in p5-free graphs in polynomial time. *Proceedings of SODA*, pages 570–581, January 2014.

16 A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 30:997–1015, 1978.

17 G.J. Minty. On maximal independent sets of vertices in claw-free graphs. *Journal of Combinatorial Theory, Series B*, 28(3):284–304, 1980.

18 J Pitman. Combinatorial stochastic processes. *Lecture Notes in Mathematics*, 1875, 2006. Springer, New York.

19 J.M. Robson. Algorithms for maximum independent sets. *Journal of Algorithms*, 7(3):425–440, 1986.

20 B. Stufler. Local limits of large Galton–Watson trees rerooted at a random vertex. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(1):155–183, 2019.

21 M. Xiao and H. Nagamochi. Confining sets and avoiding bottleneck cases: A simple maximum independent set algorithm in degree-3 graphs. *Theor. Comput. Sci.*, 469:92–104, 2013.

22 M. Xiao and H. Nagamochi. Exact algorithms for maximum independent set. *Information and Computation*, 255:126–146, 2017.

23 J. Zito. The structure and maximum number of maximum independent sets in trees. *Journal of Graph Theory*, 15(2):207–221, 1991.

Fragmentation Processes Derived from Conditioned Stable Galton-Watson Trees

Gabriel Berzunza Ojeda
Department of Mathematical Sciences, University of Liverpool, United Kingdom

Cecilia Holmgren
Department of Mathematics, Uppsala University, Sweden

Abstract

We study the fragmentation process obtained by deleting randomly chosen edges from a critical Galton-Watson tree $\mathbf{t}_n$ conditioned on having n vertices, whose offspring distribution belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$. This fragmentation process is analogous to that introduced in the works of Aldous, Evans and Pitman (1998), who considered the case of Cayley trees. Our main result establishes that, after rescaling, the fragmentation process of $\mathbf{t}_n$ converges as $n \to \infty$ to the fragmentation process obtained by cutting-down proportional to the length on the skeleton of an α-stable Lévy tree of index $\alpha \in (1, 2]$. We further establish that the latter can be constructed by considering the partitions of the unit interval induced by the normalized α-stable Lévy excursion with a deterministic drift studied by Miermont (2001). In particular, this extends the result of Bertoin (2000) on the fragmentation process of the Brownian CRT.

2012 ACM Subject Classification Mathematics of computing → Probabilistic algorithms

Keywords and phrases Additive coalescent, fragmentation, Galton-Watson trees, spectrally positive stable Lévy processes, stable Lévy tree, Prim's algorithm

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.3

Related Version *Full Version*: `https://arxiv.org/pdf/2010.07880.pdf` [11]

Funding This work was supported by the Knut and Alice Wallenberg Foundation, a grant from the Swedish Research Council and The Swedish Foundations' starting grant from Ragnar Söderbergs Foundation.

1 Introduction and main results

Aldous, Evans and Pitman [3, 19, 28] (see also [13, 24]) considered a fragmentation process of a uniform random tree $\mathbf{t}_n$ on $n \in \mathbb{N}$ labelled vertices (or Cayley tree with n vertices) by deleting the edges of $\mathbf{t}_n$ one by one in uniform random order. More precisely, as time passes, the deletion of edges creates more and more subtrees of $\mathbf{t}_n$ (connected components) such that the evolution of the ranked vector of sizes (number of vertices) of these subtrees (in decreasing order) evolves as a fragmentation process. It turns out that the asymptotic behavior of this fragmentation process, in reverse time, is related to the so-called *standard additive coalescent* [3, 19]. Moreover, this leads to a continuous representation of the standard additive coalescent in terms of the time-reversal of an analogue fragmentation process of the Brownian continuum random tree (Brownian CRT). Evans and Pitman [19, Theorem 2] showed that an additive coalescent is a Feller Markov process with values in the infinite ordered set

$$\mathbb{S} := \Big\{ \mathbf{x} = (x_1, x_2, \dots) : x_1 \geq x_2 \geq \cdots \geq 0 \ \text{ and } \ \sum_{i=1}^{\infty} x_i < \infty \Big\},$$

endowed with the ℓ^1-norm, $\|\mathbf{x}\|_1 = \sum_{i=1}^{\infty} |x_i|$ for $\mathbf{x} \in \mathbb{S}$, whose evolution is described formally by: given that the current state is $\mathbf{x}$, two terms x_i and x_j, $i < j$, of $\mathbf{x}$ are chosen and merged into a single term $x_i + x_j$ (which implies some reordering of the resulting sequence) at rate equal to $x_i + x_j$.

In this extended abstract, we study the situation where one wants to cut-down critical Galton–Watson trees conditioned on having a fixed number of vertices, but whose offspring distribution belongs to the domain of attraction of a stable law. More precisely, consider a critical offspring distribution $\mu = (\mu(k), k \geq 0)$, i.e., a probability distribution on the nonnegative integers satisfying $\sum_{k=0}^{\infty} k\mu(k) = 1$. In addition, we always implicitly assume that $\mu(0) > 0$ and $\mu(0) + \mu(1) < 1$ to avoid degenerate cases, and that μ is aperiodic[1]. We say that μ belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$ if either the variance of μ is finite, or if $\mu([k, \infty)) = k^{-\alpha} L(k)$ as $k \to \infty$, where $L : \mathbb{R}_+ \to \mathbb{R}_+$ is a function such that $L(x) > 0$ for $x \in \mathbb{R}_+$ large enough and $\lim_{x\to\infty} L(tx)/L(x) = 1$ for all $t > 0$ (such function is called slowly varying function). In other terms, if $(Y_i)_{i\geq 1}$ is a sequence of i.i.d. random variables with distribution μ, then there exists a sequence of positive real numbers $(B_n)_{n\geq 1}$ such that

$$B_n \to \infty \quad \text{and} \quad \frac{Y_1 + Y_2 + \cdots + Y_n - n}{B_n} \xrightarrow{d} Y_\alpha, \quad \text{in distribution as } n \to \infty \tag{1}$$

where the Laplace exponent of Y_α is given by $\mathbb{E}[\exp(-\lambda Y_\alpha)] = \exp(-\lambda^\alpha)$ whenever $\alpha \in (1, 2)$, and $\mathbb{E}[\exp(-\lambda Y_2)] = \exp(-\lambda^2/2)$ if $\alpha = 2$, for every $\lambda > 0$ ([20, Section XVII.5] guarantees its existence). In particular, for $\alpha = 2$, we have that Y_2 is distributed as a standard Gaussian random variable. The factor B_n is of order $n^{1/\alpha}$ (more precisely, $B_n/n^{1/\alpha}$ is a slowly varying function), and one may take $B_n = \sigma n^{1/2}$ when μ has finite variance σ^2.

We henceforth let $\mathbf{t}_n$ denote a critical Galton-Watson tree whose offspring distribution μ belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$ and refer to it as an α-stable GW-tree, for simplicity. Following Aldous, Evans and Pitman [3, 19], we are interested in the evolution of the ranked vector of sizes (in decreasing order) of the subtrees created by deleting randomly chosen edges from $\mathbf{t}_n$. Indeed, we will consider a continuous-time version of this cutting-down process. Equip each of the edges of $\mathbf{t}_n$ with i.i.d. uniform random variables (or weights) on $[0, 1]$ and independently of the tree $\mathbf{t}_n$. For $u \in [0, 1]$, we then keep the edges of $\mathbf{t}_n$ with weight smaller than u and discard the others. Therefore, one obtains a (fragmentation) forest $\mathbf{f}_n(u)$ conformed by the connected components (or subtrees of $\mathbf{t}_n$) created by the above procedure. In particular, the forest $\mathbf{f}_n(u)$ has the same set of vertices as $\mathbf{t}_n$ but clearly it has a different set of edges. Let $\mathbf{F}_n = (\mathbf{F}_n(u), u \in [0, 1])$ be the process given by

$$\mathbf{F}_n(u) = (F_{n,1}(1 - u), F_{n,2}(1 - u), \dots), \quad \text{for } u \in [0, 1],$$

the sequence of sizes (number of vertices) of the connected components of the forest $\mathbf{f}_n(1-u)$, ranked in decreasing order. We have strategically viewed the sequence of sizes of the components of $\mathbf{f}_n(1 - u)$ as an infinite sequence, by completing with an infinite number of zero terms. Plainly as time passes more and more subtrees are created, and thus, the process $\mathbf{F}_n$ evolves as a fragmentation process. Note also that $\mathbf{F}_n(0) = (n, 0, 0, \dots)$ and that $\mathbf{F}_n(1) = (1, 1, \dots, 1, 0, 0, \dots)$ where the first n terms $\mathbf{F}_n(1)$ are ones and the remaining terms are zeros. Since we are interested in studying the asymptotic behaviour of $\mathbf{F}_n$, we consider the (rescaled in time and space) fragmentation process $\mathbf{F}_n^{(\alpha)} = (\mathbf{F}_n^{(\alpha)}(t), t \geq 0)$ given by

$$\mathbf{F}_n^{(\alpha)}(t) = \frac{1}{n}\mathbf{F}_n\left(\frac{B_n}{n}t\right), \quad \text{for } 0 \leq t \leq n/B_n, \quad \text{and } \mathbf{F}_n^{(\alpha)}(t) = \frac{1}{n}\mathbf{F}_n(1) \quad \text{for } t > n/B_n, \tag{2}$$

[1] In the sense that the additive subgroup of the integers $\mathbb{Z}$ spanned by $\{i \geq 0 : \mu(i) \neq 0\}$ is $\mathbb{Z}$.

where $(B_n)_{n\geq 1}$ is a sequence satisfying (1). The process $\mathbf{F}_n^{(\alpha)}$ takes values on the set $\mathbb{S}$. The aim of this extended abstract is to establish a convergence result for the fragmentation process $\mathbf{F}_n^{(\alpha)}$. To state the precise statement (Theorem 1), it will be convenient to first introduce the limiting object.

Bertoin [6] showed that the fragmentation process of the Brownian CRT in [3] can be constructed by considering the partitions of the unit interval induced by a standard Brownian excursion with drift. This latter is sometimes called the *Brownian fragmentation*. In a similar vein, Miermont [25] built other fragmentation processes from Lévy processes with no positive jumps. Specifically, let $X_\alpha^{\mathrm{exc}} = (X_\alpha^{\mathrm{exc}}(s), s \in [0,1])$ be the normalized excursion (with unit length) of an α-stable spectrally positive Lévy process of index $\alpha \in (1,2]$; see e.g., [5]. In particular, X_2^{exc} is the normalized standard Brownian excursion. For every $t \geq 0$, define the processes $Y_\alpha^{(t)} = (Y_\alpha^{(t)}(s), s \in [0,1])$ and $I_\alpha^{(t)} = (I_\alpha^{(t)}(s), s \in [0,1])$ by letting

$$Y_\alpha^{(t)}(s) = X_\alpha^{\mathrm{exc}}(s) - ts \quad \text{and} \quad I_\alpha^{(t)}(s) = \inf_{u\in[0,s]} Y_\alpha^{(t)}(u), \quad \text{for } s \in [0,1]. \tag{3}$$

For $t \geq 0$, we introduce

$$\mathbf{F}^{(\alpha)}(t) = (F_1^{(\alpha)}(t), F_2^{(\alpha)}(t), \dots) \tag{4}$$

as the random element of $\mathbb{S}$ defined by the ranked sequence (in decreasing order) of the lengths of the intervals components of the complement of the support of the Stieltjes measure $\mathrm{d}(-I_\alpha^{(t)})$; note that $s \mapsto -I_\alpha^{(t)}(s) = \sup_{u\in[0,s]} -Y_\alpha^{(t)}(u)$ is an increasing process. More precisely, the support of $\mathrm{d}(-I_\alpha^{(t)})$ is defined as the set of times when the process $Y_\alpha^{(t)}$ reaches a new infimum. On the other hand, it can be shown that the support of $\mathrm{d}(-I_\alpha^{(t)})$ coincides with the so-called ladder time set of $-Y_\alpha^{(t)}$ which is given by the closure of the set of times when $Y_\alpha^{(t)}$ is equal to its infimum, i.e.,

$$\mathcal{L}^\alpha(t) := \overline{\left\{s \in [0,1] : Y_\alpha^{(t)}(s) = I_\alpha^{(t)}(s)\right\}};$$

see e.g., [5, Proposition 1, Chapter VI] and the discussion after that. Then $\mathbf{F}^{(\alpha)}(t)$ is the lengths of the open intervals in the canonical decomposition of $[0,1] \setminus \mathcal{L}^\alpha(t)$ arranged in the decreasing order. It is well-known that $\mathcal{L}^\alpha(t)$ is a.s. a random closed set with zero Lebesgue measure which implies that $\mathbf{F}^{(\alpha)}(t) \in \mathbb{S}_1$ a.s., where $\mathbb{S}_1 \subset \mathbb{S}$ is the space of the elements of $\mathbb{S}$ with sum 1; see [5, Corollary 5, Chapter VII]. Observe that for every fixed $0 \leq t < t'$, the process $s \mapsto Y_\alpha^{(t)}(s) - Y_\alpha^{(t')}(s) = (t'-t)s$ is monotone increasing which entails that $\mathcal{L}^\alpha(t) \subseteq \mathcal{L}^\alpha(t')$. Then the partition of $[0,1]$ induced by $\mathcal{L}^\alpha(t')$ is finer than that induced by $\mathcal{L}^\alpha(t)$. As a consequence, it has been shown by Miermont [25, Proposition 2] (see also [6, Theorem 1] for $\alpha = 2$) that $\mathbf{F}^{(\alpha)} = (\mathbf{F}^{(\alpha)}(t), t \geq 0)$ is a fragmentation process issued from $\mathbf{F}^{(\alpha)}(0) = (1,0,0,\dots)$. A precise description of its transition kernel is given in [25, Definition 4]. From now on, we will refer to $\mathbf{F}^{(\alpha)}$ as the α-stable fragmentation of index $\alpha \in (1,2]$.

We are now able to state our first main result. Let $\mathbb{D}(I, \mathbb{M})$ be the space of càdlàg functions from an interval $I \subseteq \mathbb{R}$ to the separable, complete metric space $(\mathbb{M}, d)$ equipped with the Skorohod topology; (see e.g. [12, Chapter 3] or [21, Chapter VI] for details on this space). We write $\xrightarrow{d}$ to denote convergence in distribution.

▶ **Theorem 1.** *Let $\mathbf{t}_n$ be an α-stable* GW*-tree of index $\alpha \in (1,2]$. Then, we have that*

$$(\mathbf{F}_n^{(\alpha)}(t), t \geq 0) \xrightarrow{d} (\mathbf{F}^{(\alpha)}(t), t \geq 0), \quad \text{as } n \to \infty, \quad \text{in the space } \mathbb{D}(\mathbb{R}_+, \mathbb{S}).$$

As mentioned earlier, $\mathbf{F}^{(2)}$ is exactly the Brownian fragmentation studied by Bertoin [6], that is to say, it corresponds to the fragmentation process derived from the Brownian CRT of Aldous and Pitman [3]; see also [1]. In view of this, the second goal of this paper is to show that indeed $\mathbf{F}^{(\alpha)}$ is the fragmentation process obtained by cutting-down the "edges" of the α-stable Lévy tree.

The α-stable Lévy tree of index $\alpha \in (1, 2]$ is the continuum random tree analogue of (discrete) α-stable GW-trees. They were introduced by Duquesne and Le Gall [18], and in particular, they also appear as scaling limits of α-stable GW-trees. In brief, the α-stable Lévy tree $\mathcal{T}_\alpha = (\mathcal{T}_\alpha, d_\alpha, \rho_\alpha)$ is a random compact metric space $(\mathcal{T}_\alpha, d_\alpha)$ with one distinguished element $\rho \in \mathcal{T}_\alpha$ called the root such that $(\mathcal{T}_\alpha, d_\alpha)$ is a tree-like space in that for $v, w \in \mathcal{T}_\alpha$, there is a unique non-self-crossing path $[v, w]$ from v to w in $\mathcal{T}_\alpha$, whose length equals $d_\alpha(v, w)$. The leaves $\mathrm{Lf}(\mathcal{T}_\alpha)$ of $\mathcal{T}_\alpha$ are those points that do not belong to the interior of any path leading from one point to another, and the skeleton of the tree is the set $\mathrm{Sk}(\mathcal{T}_\alpha) = \mathcal{T}_\alpha \setminus \mathrm{Lf}(\mathcal{T}_\alpha)$ of non-leaf points. The α-stable Lévy tree $\mathcal{T}_\alpha$ is naturally endowed with a uniform probability measure μ_α (*the mass measure*) that is supported on $\mathrm{Lf}(\mathcal{T}_\alpha)$, and a unique σ-finite measure λ_α (*the length measure*) carried by $\mathrm{Sk}(\mathcal{T}_\alpha)$ that assigns measure $d(v, w)$ to the geodesic path between v and w in $\mathcal{T}_\alpha$.

Following Aldous-Pitman's fragmentation [3] of the Brownian CRT, the analogue of deleting randomly chosen edges in $\mathbf{t}_n$ is to cut the skeleton of $\mathcal{T}_\alpha$ by a Poisson point process of cuts with intensity $\mathrm{d}t \otimes \lambda_\alpha(\mathrm{d}v)$ on $[0, \infty) \times \mathcal{T}_\alpha$. For all $t \geq 0$, define an equivalence relation $\sim_t$ on $\mathcal{T}_\alpha$ by saying that $v \sim_t w$, for $v, w \in \mathcal{T}_\alpha$, if and only if, no atom of the Poisson process that has appeared before time t belongs to the path $[v, w]$. These cuts split the α-stable Lévy tree into a (continuum) forest, that is a countably infinite set of smaller subtrees (connected components) of $\mathcal{T}_\alpha$. Let $\mathcal{T}_{\alpha,1}^{(t)}, \mathcal{T}_{\alpha,2}^{(t)}, \dots$ be the distinct equivalence classes for $\sim_t$ (connected components of $\mathcal{T}_\alpha$), ranked according to the decreasing order of their μ_α-masses. The subtrees $(\mathcal{T}_{\alpha,i}^{(t)}, i \geq 1)$ are nested as t varies, that is, for every $0 \leq t < t'$ and $i \geq 1$, there exits $j \geq 1$ such that $\mathcal{T}_{\alpha,i}^{(t')} \subset \mathcal{T}_{\alpha,j}^{(t)}$. Let $\mathbf{F}_{\mathcal{T}_\alpha} = (\mathbf{F}_{\mathcal{T}_\alpha}(t), t \geq 0)$ be the process given by

$$\mathbf{F}_{\mathcal{T}_\alpha}(t) = (\mu_\alpha(\mathcal{T}_{\alpha,1}^{(t)}), \mu_\alpha(\mathcal{T}_{\alpha,2}^{(t)}), \dots), \quad t \geq 0,$$

where $\mathbf{F}_{\mathcal{T}_\alpha}(0) = (1, 0, 0, \dots)$. Indeed, $\mathbf{F}_{\mathcal{T}_\alpha}$ is a fragmentation process in the sense that $\mathbf{F}_{\mathcal{T}_\alpha}(t')$ is obtained by splitting at random the elements of $\mathbf{F}_{\mathcal{T}_\alpha}(t)$, for $0 \leq t < t'$. We call $\mathbf{F}_{\mathcal{T}_\alpha}$ the fragmentation process of the α-stable Lévy tree. In particular, $\mathbf{F}_{\mathcal{T}_2}$ is the fragmentation process of the Brownian CRT introduced in [3, Section 2.2]. Note that $\mathbf{F}_{\mathcal{T}_\alpha}$ takes values in $\mathbb{S}$, and that [11, Lemma 7] shows that $\mathbf{F}_{\mathcal{T}_\alpha}(t) \in \mathbb{S}_1$ a.s., for every $t \geq 0$. We can now state our second main result.

▶ **Proposition 2.** *We have that* $(\mathbf{F}^{(\alpha)}(t), t \geq 0) \overset{d}{=} (\mathbf{F}_{\mathcal{T}_\alpha}(t), t \geq 0)$*, where* $\overset{d}{=}$ *means equal in distribution (in the sense of finite-dimensional distributions).*

Theorem 3 in [3] shows that the time-reversed fragmentation process of the Brownian CRT, i.e. $(\mathbf{F}_{\mathcal{T}_2}(e^{-t}), t \in \mathbb{R})$, is a version of the standard additive coalescent providing an explicit construction of this last process. In general, Miermont [25, Section 6] has shown that the time-reversed α-stable fragmentation process, i.e. $(\mathbf{F}^{(\alpha)}(e^{-t}), t \in \mathbb{R})$, is an eternal additive coalescent as described by Evans and Pitman [19]. More precisely, it is a mixing of so-called *extremal coalescents* of Aldous and Pitman [4] (see also [7]) which exact law is given in [25, Proposition 3]. Thus, Proposition 2 implies that this eternal additive coalescent can also be constructed from the α-stable Lévy tree by Poisson splitting along its skeleton. On the other hand, Theorem 1 and Proposition 2 clearly generalize Bertoin's work [6] and moreover,

complete Miermont's [25] one by identifying the distribution of the α-stable fragmentation with that of the fragmentation process of the α-stable Lévy tree. In particular, Bertoin [8] proved that $\mathbf{F}^{(2)}$ (or equivalently, $\mathbf{F}_{\mathcal{T}_2}$) is a so-called *self-simlar fragmentation process* of index 1/2. However, Miermont [26] has already pointed out that $\mathbf{F}^{(\alpha)}$ (and therefore $\mathbf{F}_{\mathcal{T}_\alpha}$), for $\alpha \in (1,2)$, is not a self-similar fragmentation due to the existence of points in $\mathcal{T}_\alpha$ with infinite degree.

The proof of Theorem 1 uses some of the ideas developed in [13] where only the case of Cayley tree was treated. However, in our more general framework, there are technical challenges that do not appear in [13], mostly due to the lack of some properties that only the Cayley tree satisfies. Informally, we use the so-called *Prim's algorithm* [29] to obtain a consistent ordering on the vertices of the forest created by deleting randomly chosen edges from $\mathbf{t}_n$ that we refer to as the *Prim order*; see Section 3. This will allow us to precisely encode this forest (and in particular, the sizes of connected components) using a discrete analogue of the process $Y_\alpha^{(t)}$ defined in (3) that we refer to as the *Prim path*. We then show that this (properly rescaled) Prim path indeed converges to its continuous version. Finally, we use a general approach developed in the complete version [11, Section 6] for the convergence of fragmentation processes encoded by functions in $\mathbb{D}([0,1],\mathbb{R})$ to conclude our proof.

The proof of Proposition 2 follows along the lines of that of Theorem 3 in [3] for the Brownian CRT (see also the proof of Proposition 13 in [4]). Informally, one uses the convergence of rescaled α-stable GW-trees toward the α-stable Lévy tree $\mathcal{T}_\alpha$ in order to approximate the fragmentation process of $\mathcal{T}_\alpha$. The detailed proof of Proposition 2 is given in the complete version [11].

The rest of the manuscript is organized as follows. In Section 2, we discuss some connections with some combinatorial and probabilistic models: additive coalescents, parking schemes, laminations and Bernoulli bond-percolation. Section 3 is devoted to the introduction of Galton-Watson trees as well as the formal definition of the exploration process (the Prim path) associated with the fragmentation forest. Finally, in Section 4 and 5, we provide a fair enough guideline of the proof of Theorem 1.

2 Further remarks

In this section, we highlight some connections with previous works.

Additive coalescents

A Cayley tree of size n can be viewed as a Galton-Watson tree with Poissonian offspring distribution of parameter 1 and conditioned to have n vertices, where the labels are assigned to the vertices uniformly at random. In particular, Aldous, Evans and Pitman fragmentation process [3, 19, 28], say $\mathbf{F}_n^+ = (\mathbf{F}_n^+(t), t \geq 0)$, corresponds precisely to $\mathbf{F}_n^{(\alpha)}$ in (2), with $\alpha = 2$ and $B_n = n^{1/2}$. The fragmentation process $\mathbf{F}_n^+$ leads to a representation of an additive coalescent by an appropriate time reversal, that is, the exponential time-change $t \to e^{-t}$. Specifically, $(\mathbf{F}_n^+(e^{-t}), t \geq -(1/2)\ln n)$ is an additive coalescent starting at time $-(1/2)\ln n$ from the state $(1/n, 1/n, \dots, 1/n, 0, 0, \dots) \in \mathbb{S}$. Evans and Pitman [19] (see also [3, Proposition 2]) showed that this time-reversed version of $\mathbf{F}_n^+$ converges in distribution to the standard additive coalescent, i.e., $(\mathbf{F}_{\mathcal{T}_2}(e^{-t}), t \in \mathbb{R})$.

Aldous and Pitman [4] (see also [19, Construction 5]) also studied the fragmentation process derived by cutting-down *birthday trees*. They are a family of trees that generalizes the Cayley tree in allowing "weights" on the vertices. Aldous and Pitman showed that this

fragmentation process, suitable rescaled, converges to the fragmentation process associated to the continuum counterpart of birthday trees, the *inhomogeneous continuum random trees* (ICRT). Moreover, the time-reversed version of the fragmentation process of the ICRT can be viewed as version of an eternal additive coalescent. On the other hand, Bertoin [7] has proved that the fragmentation process of the ICRT can also be constructed by considering the partitions of the unit interval induced by certain bridges with exchangeable increments.

Parking schemes

Chassaing and Louchard [14] provided another representation of the standard additive coalescent as parking schemes related to the *Knuth's parking problem*; see also [15, 24]. Bertoin and Miermont [10] extended the work [14] and relate the Knuth's parking problem for caravans to different versions of eternal additive coalescent. On the other hand, the Knuth's parking problem bear some similarities with the dynamics of an aggregating server studied by Bertoin [7].

Lamination process

In [30], Thévenin has provided a geometric representation of the fragmentation process $\mathbf{F}_{\mathcal{T}_\alpha}$ by introducing a new lamination-valued process. In particular, Theorem 1.1 in [30] combined with Proposition 1 allows us to deduce the exact distribution of the ranked sequence (in decreasing order) of the masses of the faces of this lamination-valued process.

Bernoulli bond-percolation

Bernoulli bond-percolation on finite connected graphs is perhaps the simplest example of a percolation model. In this model, each edge in the connected graph is removed with probability $1-p \in (0,1)$, and it is kept with probability p, independently of the other edges. This induces a partition of the set of vertices of the graph into connected components usually referred to as clusters. It should be intuitively clear that there is a link between Bernoulli bond-percolation on α-stable GW-trees and their associated fragmentation processes. More precisely, let $\mathbf{t}_n$ be an α-stable GW-tree. For $u \in [0,1]$, recall that the continuous-time cutting-down procedure of $\mathbf{t}_n$ described in the introduction results in a random forest of connected components. Indeed, the probability that a given edge of $\mathbf{t}_n$ has not yet been removed at time u is exactly u. Thus, the configuration of the connected components at time u is precisely that resulting from Bernoulli bond-percolation on $\mathbf{t}_n$ with parameter u. A natural problem in this setting is then to investigate the asymptotic behavior of the sizes of the largest clusters for appropriate percolation regimes. In this direction, let $(B_n)_{n\geq 1}$ be a sequence of positive real numbers satisfying (1). An application of Theorem 1 shows that for the percolation parameter $1-(B_n/n)t$ with a fixed $t \geq 0$, the sequence of sizes of the clusters ranked in decreasing order and renormalized by a factor of $1/n$ (i.e. $\mathbf{F}_n^{(\alpha)}(t)$) converges in distribution, as $n \to \infty$, to $\mathbf{F}^{(\alpha)}(t)$. Theorem 2 in [25] allows us to describe explicitly the distribution of $\mathbf{F}^{(\alpha)}$ at fixed times. Let $(p_s(z), z \in \mathbb{R}, s \geq 0)$ be the family of densities of the distribution of a strictly stable spectrally positive Lévy process with index $\alpha \in (1,2]$.

▶ **Corollary 3.** *For $t > 0$, let $\mathrm{a}_1^{(\alpha)}(t) > \mathrm{a}_2^{(\alpha)}(t) > \cdots$ be the atoms of a Poisson measure on $(0,\infty)$ with intensity $\Lambda_\alpha^{(t)}(\mathrm{d}z) := z^{-1}p_z(-tz)\mathbb{1}_{\{z>0\}}\mathrm{d}z$, ranked in decreasing order. Then*

$$\mathbf{F}^{(\alpha)}(t) \stackrel{d}{=} \Big((\mathrm{a}_1^{(\alpha)}(t), \mathrm{a}_2^{(\alpha)}(t), \dots) \,\Big|\, \sum_{i=1}^{\infty} \mathrm{a}_i^{(\alpha)}(t) = 1\Big).$$

Following Bertoin's [9] work about Bernoulli bond-percolation on random trees. The percolation regime $1-(B_n/n)t$ on $\mathbf{t}_n$ corresponds to the so-called supercritical regime. Indeed, the result in Corollary 3 has already been proved by Pitman [28] for Cayley trees.

3 The coding of Galton-Watson trees and their fragmentation

In this section, we formally introduce the family of critical Galton-Watson trees and explain how they can be coded by different functions, namely the so-called Łukasiewicz path and a similar path derived by the Prim's algorithm.

Plane trees

We follow the formalism of Neveu [27]. Let $\mathbb{N} = \{1, 2, \dots\}$ be the set of positive integers, set $\mathbb{N}^0 = \{\varnothing\}$ and consider the set of labels $\mathbb{U} = \bigcup_{n\geq 0} \mathbb{N}^n$. For $u = (u_1, \dots, u_n) \in \mathbb{U}$, we denote by $|u| = n$ the length (or generation, or height) of u; if $v = (v_1, \dots, v_m) \in \mathbb{U}$, we let $uv = (u_1, \dots, u_n, v_1, \dots, v_m) \in \mathbb{U}$ be the concatenation of u and v. A plane tree is a nonempty, finite subset $\tau \subset \mathbb{U}$ such that: (i) $\varnothing \in \tau$; (ii) if $v \in \tau$ and $v = uj$ for some $j \in \mathbb{N}$, then $u \in \tau$; (iii) if $u \in \tau$, then there exists an integer $c(u) \geq 0$ such that $ui \in \tau$ if and only if $1 \leq i \leq c(u)$. We will view each vertex u of a tree τ as an individual of a population whose τ is the genealogical tree. The vertex $\varnothing$ is called the root of the tree and for every $u \in \tau$, $c(u)$ is the number of children of u (if $c(u) = 0$, then u is called a leaf, otherwise, u is called an internal vertex). The total progeny (or size) of τ will be denoted by $\zeta(\tau) = \mathrm{Card}(\tau)$ (i.e., the number of vertices of τ). We denote by $\mathbb{T}$ the set of plane trees and for each $n \in \mathbb{N}$, by $\mathbb{T}_n$ the set of plane trees with n vertices, or equivalently $n-1$ edges.

Galton-Watson trees

Let μ be a probability measure on $\mathbb{Z}_+$ which satisfies $\mu(0) > 0$, expectation $\sum_{k=0}^{\infty} k\mu(k) = 1$ and such that $\mu(0) + \mu(1) < 1$. The law of a critical Galton–Watson tree with offspring distribution μ is the unique probability measure $\mathbb{P}_\mu$ on $\mathbb{T}$ satisfying: (i) $\mathbb{P}_\mu(c(\varnothing) = k) = \mu(k)$ for every $k \geq 0$; (ii) For every $k \geq 1$ such that $\mu(k) > 0$, conditioned on the event $\{c(\varnothing) = k\}$, the subtrees that stem from the children of the root $\{u \in \mathbb{U} : 1u \in \tau\}, \dots, \{u \in \mathbb{U} : ku \in \tau\}$ are independent and distributed as $\mathbb{P}_\mu$. A random tree whose distribution is $\mathbb{P}_\mu$ will be called a Galton–Watson tree with offspring distribution μ. We also denote by $\mathbb{P}_\mu^{(n)}$ the law on $\mathbb{T}_n$ of a Galton-Watson tree with offspring distribution μ conditioned to have n vertices, providing that this conditioning makes sense.

Coding planar trees by discrete paths

We will use two different orderings of the vertices of a tree $\tau \in \mathbb{T}$:

- **(i) Lexicographical ordering.** Given $v, w \in \tau$, we write $v \prec_{\mathrm{lex}} w$ if there exits $z \in \tau$ such that $v = z(v_1, \dots, v_n)$, $w = z(w_1, \dots, w_m)$ and $v_1 < w_1$.
- **(ii) Prim ordering.** Let $\mathbf{edge}(\tau)$ be the set of edges of τ and consider a sequence of distinct and positive weights $\mathbf{w} = (w_e : e \in \mathbf{edge}(\tau))$ (i.e., each edge e of τ is marked with a different and positive weight w_e). Given two distinct vertices $u, v \in \tau$, we write $\{u, v\}$ for the edge connecting u and v in τ. Let us describe the Prim order $\prec_{\mathrm{prim}}$ of the vertices in τ, that is, $\varnothing = u(0) \prec_{\mathrm{prim}} u(1) \prec_{\mathrm{prim}} \cdots \prec_{\mathrm{prim}} u(\zeta(\tau) - 1)$. We will use the notation V_i for the set $\{u(0), \dots, u(i-1)\}$, for $1 \leq i \leq \zeta(\tau)$. First set $u(0) = \varnothing$ and $V_0 = \{u(0)\}$. Suppose that for some $1 \leq i \leq \zeta(\tau) - 1$, the vertices $u(0), \dots, u(i-1)$ have been defined. Consider the weights $\{w_{\{u,v\}} : u \in V_i, v \notin V_i\}$ of edges between a vertex

of V_i and another outside of V_i. Since all the weights are distinct, the minimum weight in $\{w_{\{u,v\}} : u \in V_i, v \notin V_i\}$ is reached at an edge $\{\tilde{u}, \tilde{v}\}$ where $\tilde{u} \in V_i$ and $\tilde{v} \notin V_i$. Then set $u(i) = \tilde{v}$. This iterative procedure completely determines the Prim order $\prec_{\mathrm{prim}}$.

For $* \in \{\mathrm{lex}, \mathrm{prim}\}$, we associate to every ordering $\varnothing = u(0) \prec_* u(1) \prec_* \cdots \prec_* u(\zeta(\tau)-1)$ of the vertices of τ a path $W^* = (W^*(k), 0 \le k \le \zeta(\tau))$, by letting $W^*(0) = 0$ and for $0 \le k \le \zeta(\tau)-1$, $W^*(k+1) = W^*(k) + c(u(k)) - 1$, where we recall that $c(u(k))$ denotes the number of children of the vertex $u(k) \in \tau$. Observe that $W^*(k+1) - W^*(k) = c(u(k)) - 1 \ge -1$ for every $0 \le k \le \zeta(\tau) - 1$, with equality if and only if $u(k)$ is a leaf of τ. Note also that $W^*(k) \ge 0$, for every $0 \le k \le \zeta(\tau) - 1$, but $W^*(\zeta(\tau)) = -1$. We shall think of such a path as the step function on $[0, \zeta(\tau)]$ given $s \mapsto W^*(\lfloor s \rfloor)$. The path W^{lex} is commonly called Łukasiewicz path of τ, and from now on we refer to W^{prim} as the Prim path; see [23] for more details and properties on the Łukasiewicz path.

The procedure just described to obtain the Prim ordering is known as Prim's algorithm (or Prim-Jarník algorithm); see [29]. This algorithm associates to any properly weighted graph its unique minimum spanning tree. In practice, one could also consider that $\mathbf{w}$ is a sequence of i.i.d. positive random variables such that they are all distinct a.s. and independent of the tree.

Define the probability measure $\hat{\mu}$ on $\{-1, 0, 1, \dots\}$ by $\hat{\mu}(k) = \mu(k+1)$ for every $k \ge -1$. Let $X = (X(k), k \ge 0)$ be a random walk which starts at 0 with jump distribution $\hat{\mu}$ and define also the time $\zeta_1 = \inf\{k \ge 0 : X(k) = -1\}$. In the Prim ordering, consider that the weights $\mathbf{w}$ is a sequence of i.i.d. positive random variables such that they are distinct a.s. and independent of the tree.

▶ **Proposition 4.** *For every $* \in \{\mathrm{lex}, \mathrm{prim}\}$, if we sample a plane tree $\mathbf{t}$ according to $\mathbb{P}_\mu$, then W^* is distributed as $(X(0), X(1), \dots, X(\zeta_1))$. In particular, the total progeny of the sample plane tree has the same distribution as ζ_1.*

Proof. The proof for the Łukasiewicz path can be found in [23, Proposition 1.5]. For the Prim path the proof follows from a simple adaptation of that of [23, Proposition 1.5]; see also [13, Lemmas 15 and 16] for an alternative approach. ◀

Fragmentation of a plane tree

Consider $\tau \in \mathbb{T}$ and let $\mathbf{edge}(\tau)$ denote its set of edges. Equip the edges of τ with i.i.d. uniform random variables (or weights) $\mathbf{w} = (w_e : e \in \mathbf{edge}(\tau))$ on $[0,1]$ and independently of the tree τ. In particular, for a vertex $v \in \tau$ with $c(v) \ge 1$ children, we write $(w_{v,k}, 1 \le k \le c(v))$ for the weights of the edges connecting v with its children. For $t \in [0,1]$, we then keep the edges of τ with weight smaller than t and discard the others. This gives rise to a forest $\mathbf{f}_\tau(t)$ with the same set of vertices as τ but with set of edges given by $\mathbf{edge}(\mathbf{f}_\tau(t)) = \{e \in \mathbf{edge}(\tau) : w_e \le t\}$. Furthermore, each vertex $v \in \mathbf{f}_\tau(t)$ has $c_t(v) = \sum_{k=1}^{c(v)} \mathbb{1}_{\{w_{v,k} \le t\}}$ children if $c(v) \ge 1$; otherwise, $c_t(v) = 0$ whenever $c(v) = 0$. In what follows, we refer to the forest $\mathbf{f}_\tau(t)$ associated to a plane tree τ and uniform weights $\mathbf{w}$ as the *fragmentation forest*. In this manuscript we restrict ourselves to the case uniform i.i.d. weights, but certainly some of the forthcoming results can be extended for more general sequences of weights.

Prim exploration of the fragmentation forest

For a plane tree $\tau \in \mathbb{T}$ and sequence of i.i.d. uniform random weights $\mathbf{w}$ on $[0,1]$, let $\mathbf{f}_\tau(t)$ be the fragmentation forest of τ at time $t \in [0,1]$. Let us now explain how to explore the subtree components of the forest $\mathbf{f}_\tau(t)$ by using the approach outlined in [13, page 532] (see also [2]).

For $t \in [0,1]$, denote by $\mathbf{Neigh}_t(v) \coloneqq \{u \in \mathbf{f}_\tau(t) : \{u,v\} \in \mathbf{edge}(\mathbf{f}_\tau(t))\}$ the set of neighbours of $v \in \mathbf{f}_\tau(t)$. For a set of vertices V of $\mathbf{f}_\tau(t)$, let also $\mathbf{Neigh}_t(V) \coloneqq (\bigcup_{v\in V} \mathbf{Neigh}_t(v)) \setminus V$, the set of neighbours of vertices in V but not in V. We associate to the prim ordering $\varnothing = u(0) \prec_{\text{prim}} u(1) \prec_{\text{prim}} \cdots \prec_{\text{prim}} u(\zeta(\tau)-1)$ of the vertices of τ the following exploration process of $\mathbf{f}_\tau(t)$ (recall that $\mathbf{f}_\tau(t)$ and τ have the same set of vertices). The first visited vertex is $v_t(0) = u(0)$. Suppose that we have explored the vertices $V_k = \{v_t(0), \dots, v_t(k-1)\}$ at some time $1 \le k \le \zeta(\tau)$. If $k = \zeta(\tau)$, we have finished the exploration, and otherwise, one has two possibilities:

- **(i)** if $\mathbf{Neigh}_t(V_k) \neq \varnothing$, then $v_t(k)$ is the next vertex according to the order $\prec_{\text{prim}}$ that belongs to $\mathbf{Neigh}_t(V_k)$, or
- **(ii)** if $\mathbf{Neigh}_t(V_k) = \varnothing$, then $v_t(k)$ is the next vertex according to the order $\prec_{\text{prim}}$ that belongs to $\tau \setminus V_k$.

This exploration process results in an order for the vertices of $\mathbf{f}_\tau(t)$ (equivalently, to the vertices of τ) that we denote by $<_{\text{prim}}$ (i.e. $\varnothing = v_t(0) <_{\text{prim}} v_t(1) <_{\text{prim}} \cdots <_{\text{prim}} v_t(\zeta(\tau)-1)$) and call *Prim exploration*. An important feature of the Prim exploration of $\mathbf{f}_\tau(t)$ is that the Prim ordering $<_{\text{prim}}$ of its vertices is preserved for all values of $t \in [0,1]$. More precisely, for $t_1, t_2 \in [0,1]$, $v_{t_1}(k) = v_{t_2}(k)$, for all $0 \le k \le \zeta(\tau)-1$. This is a consequence of the algorithm to obtain the Prim ordering of the vertices in τ which associates to any properly weighted graph its unique minimum spanning tree. We henceforth write $\prec_{\text{prim}}$ instead of $<_{\text{prim}}$ and remove the subindex t from our notation, i.e., we write $\varnothing = v(0) \prec_{\text{prim}} v(1) \prec_{\text{prim}} \cdots \prec_{\text{prim}} v(\zeta(\tau)-1)$ for the vertices of $\mathbf{f}_\tau(t)$ in Prim order, which is the same as the Prim ordering of the vertices of the tree τ, $\varnothing = u(0) \prec_{\text{prim}} u(1) \prec_{\text{prim}} \cdots \prec_{\text{prim}} u(\zeta(\tau)-1)$ presented earlier.

Following the presentation of [13, pages 532-533], one can associate to the Prim ordering of the vertices of $\mathbf{f}_\tau(t)$, an *exploration path* $Z_t = (Z_t(k), 0 \le k \le \zeta(\tau)+1)$ by letting $Z_t(0) = Z(\zeta(\tau)+1) = 0$, and for $1 \le k \le \zeta(\tau)$, $Z_t(k) = \mathrm{Card}(\mathbf{Neigh}_t(V_k))$. Furthermore, let $\mathbf{CC}(\mathbf{f}_\tau(t))$ be the set of connected components of $\mathbf{f}_\tau(t)$. Then [13, Lemma 14] shows that

$$\mathrm{Card}(\{k \in \{1,\dots,\zeta(\tau)\} : Z_t(k) = 0\}) = \mathrm{Card}(\mathbf{CC}(\mathbf{f}_\tau(t))),$$

and that the successive sizes of the connected components ordered by the exploration coincide with the distances between successive 0's in the sequence $Z_t = (Z_t(k), 0 \le k \le \zeta(\tau)+1)$.

In this manuscript, and in analogy with the coding paths of τ introduced earlier, we will consider a slight modification of the exploration path Z_t. More precisely, define the Prim path $W_t^{\text{prim}} = (W_t^{\text{prim}}(k), 0 \le k \le \zeta(\tau))$ by letting $W_t^{\text{prim}}(0) = 0$, and for $0 \le k \le \zeta(\tau)-1$, $W_t^{\text{prim}}(k+1) = W_t^{\text{prim}}(k) + c_t(v_t(k)) - 1$, where $c_t(v)$ denotes the number of children of $v \in \mathbf{f}_\tau(t)$. We shall also think of such a path as the step function on $[0, \zeta(\tau)]$ given by $s \mapsto W_t^{\text{prim}}(\lfloor s \rfloor)$.

▶ **Lemma 5.** *Let $\tau \in \mathbb{T}$ and $\mathbf{w}$ be a sequence of i.i.d. uniform random weights on $[0,1]$ which is also independent of τ. For any time $t \in [0,1]$,*

$$\mathrm{Card}\left(\left\{k \in \{1,\dots,\zeta(\tau)\} : W_t^{\text{prim}}(k) = \min_{0\le m\le k} W_t^{\text{prim}}(m)\right\}\right) = \mathrm{Card}(\mathbf{CC}(\mathbf{f}_\tau(t))),$$

Moreover, the successive sizes of the connected components of $\mathbf{f}_\tau(t)$ ordered by the exploration process coincide with the distances between successive new minimums in the sequence $(W_t^{\text{prim}}(k), 0 \le k \le \zeta(\tau))$.

Proof. The result is an immediate consequence of the previous discussion. ◀

Indeed, the sizes of the connected components of $\mathbf{f}_\tau(t)$ coincides with the length of the excursions of the walk W_t^{prim} above its minimum.

Following Proposition 4, the Prim path of the fragmentation forest associated to a critical Galton-Watson tree with offspring distribution μ can also be related to a random walk. Recall that $X = (X(k), k \geq 0)$ denotes a random walk that starts at 0 and has jump distribution $\hat{\mu}$ on $\{-1, 0, 1, \dots\}$. Recall also that $\zeta_1 = \inf\{k \geq 0 : X(k) = -1\}$. Denote by $(\xi(k), k \geq 1)$ the increments of X, i.e. $\xi(k) = X(k) - X(k-1)$, for $k \geq 1$. Let $(U_k(j))_{k,j\geq 1}$ be a sequence of i.i.d. uniform random variables on $[0, 1]$. For $t \in [0, 1]$, define $(\xi_t(k), k \geq 1)$ by letting

$$\xi_t(k) = \sum_{j=1}^{\xi(k)+1} \mathbb{1}_{\{U_k(j)\leq t\}}, \quad \text{for } t \in [0,1], \ k \geq 1,$$

with the convention $\sum_{j=1}^{0} \mathbb{1}_{\{U_k(j)\leq t\}} = 0$. Hence, $\xi_0(k) = 0$, $\xi_1(k) = \xi(k) + 1$ and for any $k \geq 1$, the mapping $t \mapsto \xi_t(k)$ is non-decreasing. Let $X_t = (X_t(k), k \geq 0)$ be the process defined by

$$X_t(0) = 0 \quad \text{and} \quad X_t(k) = \sum_{i=1}^{k} (\xi_t(i) - 1), \quad \text{for } t \in [0,1], \ k \geq 1. \tag{5}$$

▶ **Proposition 6.** *Sample a plane tree* $\mathbf{t}$ *according to* $\mathbb{P}_\mu$*, i.e., consider a critical Galton-Watson tree* $\mathbf{t}$ *with offspring* μ*. Let* $\mathbf{w} = (w_e : e \in \mathbf{edge}(\mathbf{t}))$ *be a sequence of i.i.d. uniform random weights on* $[0, 1]$ *which is also independent of* $\mathbf{t}$*. Then, the Prim path* W_t^{prim} *satisfies*

$$(W_t^{\text{prim}}(0), W_t^{\text{prim}}(1), \dots, W_t^{\text{prim}}(\zeta(\mathbf{t})))_{t\in[0,1]} \overset{d}{=} (X_t(0), X_t(1), \dots, X_t(\zeta_1))_{t\in[0,1]},$$

where $\overset{d}{=}$ *means equal in distribution (in the sense of finite-dimensional distributions).*

Proof. The proof of can be found in the complete version [11]. ◀

4 Convergence of the exploration processes

Recall that $\mathbb{P}_\mu^{(n)}$ denotes the law of a critical Galton-Watson tree with offspring distribution μ conditioned to have $n \in \mathbb{N}$ vertices. For every $n \in \mathbb{N}$, for which $\mathbb{P}_\mu^{(n)}$ is well-defined, sample a plane tree on $\mathbb{T}_n$, say $\mathbf{t}_n$, according to $\mathbb{P}_\mu^{(n)}$, i.e., $\mathbf{t}_n$ is a critical Galton-Watson tree conditioned to have n vertices. Through this section we assume that μ belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$, and refer to $\mathbf{t}_n$ as an α-stable GW-tree. We will always let $\mathbf{w} = (w_e : e \in \mathbf{edge}(\mathbf{t}_n))$ be a sequence of i.i.d. uniform random weights on $[0, 1]$ which is also independent of $\mathbf{t}_n$. We write $W_n^{\text{lex}} = (W_n^{\text{lex}}(\lfloor nu \rfloor), u \in [0, 1])$ for the associated time-scaled Łukasiewicz path of $\mathbf{t}_n$. We also write $W_n^{\text{prim}} = (W_n^{\text{prim}}(\lfloor nu \rfloor), u \in [0, 1])$ for the time-scaled Prim path of $\mathbf{t}_n$ with respect to the weights $\mathbf{w}$.

The asymptotic behavior of large α-stable GW-trees is well understood, in particular through scaling limits of their associated Łukasiewicz paths; see, e.g., [17]. In this section, we first show that the Prim path of $\mathbf{t}_n$ has the same asymptotic behavior as its Łukasiewicz path. This will serve as a stepping stone to study the Prim path of the fragmentation forest of $\mathbf{t}_n$ associated to the weights $\mathbf{w}$. Recall that $X_\alpha^{\text{exc}} = (X_\alpha^{\text{exc}}(u), u \in [0, 1])$ denotes the α-stable excursion of index α.

▶ **Theorem 7.** *Let* $\mathbf{t}_n$ *be an* α*-stable* GW*-tree, and let* $(B_n)_{n\geq 1}$ *be a sequence of positive real numbers satisfying (1). For* $* \in \{\text{lex}, \text{prim}\}$*, we have that*

$$\left(\frac{1}{B_n} W_n^*(\lfloor nu \rfloor), u \in [0,1]\right) \xrightarrow{d} (X_\alpha^{\text{exc}}(u), u \in [0,1]), \quad \text{as } n \to \infty, \text{ in the space } \mathbb{D}([0,1], \mathbb{R}).$$

Proof. It follows from [17, Theorem 3.1] and Proposition 4. ◀

For $s \in [0,1]$, let $\mathbf{f}_n(s)$ be the fragmentation forest of $\mathbf{t}_n$ at time s. Denote by $W_{n,s}^{\mathrm{prim}} = (W_{n,s}^{\mathrm{prim}}(\lfloor nu \rfloor), u \in [0,1])$ the time-scaled Prim path of $\mathbf{f}_n(s)$. In particular, $W_{n,1}^{\mathrm{prim}}$ is exactly W_n^{prim}. For fixed $t \geq 0$, consider the sequence $(s_n(t))_{n\geq 1}$ of positive times given by $s_n(t) = 1 - (B_n/n)t$, where $(B_n)_{n\geq 1}$ is a sequence of positive real numbers satisfying (1). Define the process $W_n^{(t)} = (W_n^{(t)}(u), u \in [0,1])$ by letting

$$W_n^{(t)}(u) = \frac{1}{B_n} W_{n,s_n(t)}^{\mathrm{prim}}(\lfloor nu \rfloor), \quad \text{for } u \in [0,1]. \tag{6}$$

Later, we refer to $W_n^{(t)}$ as the (normalized and time-scaled) Prim path of the fragmentation forest at time $s_n(t)$, i.e., $\mathbf{f}(s_n(t))$. We set $W_n = (W_n^{(t)}, t \geq 0)$. From the previous section, the mapping $t \mapsto W_n^{(t)}(u)$ is non-increasing in t which implies that W_n has càdlàg paths. Thus, we will view $(t,u) \mapsto W_n^{(t)}(u)$ as a random variable taking values in the space $\mathbb{D}(\mathbb{R}_+, \mathbb{D}([0,1],\mathbb{R}))$ of $\mathbb{D}([0,1],\mathbb{R})$-valued càdlàg functions on $\mathbb{R}$ equipped with the Skorokhod topology. In other words, for fixed $t \geq 0$, $W_n^{(t)}$ is a random variable in $\mathbb{D}([0,1],\mathbb{R})$.

We introduce the continuous counterpart of the process W_n. For every $t \geq 0$, let $Y_\alpha^{(t)} = (Y_\alpha^{(t)}(u), u \in [0,1])$ be defined by $Y_\alpha^{(t)}(u) = X_\alpha^{\mathrm{exc}}(u) - tu$, for $u \in [0,1]$. In particular, for $t = 0$, $Y_\alpha^{(0)} = X_\alpha^{\mathrm{exc}}$. Then, define the process $Y_\alpha = (Y_\alpha^{(t)}, t \geq 0)$.

The following theorem is the main result of this section.

▶ **Theorem 8.** *We have the convergence*

$$(W_n^{(t)}, t \geq 0) \xrightarrow{d} (Y_\alpha^{(t)}, t \geq 0), \quad \textit{as } n \to \infty, \quad \textit{in the space } \mathbb{D}(\mathbb{R}_+, \mathbb{D}([0,1],\mathbb{R})).$$

Theorem 8 generalizes [13, Theorem 10]. Specifically, in [13], the authors only consider the case when $\mathbf{t}_n$ is a GW-tree with μ being the law of a Poisson random variable of parameter 1 (i.e., $\mathbf{t}_n$ is a Cayley tree) while our setting is clearly more general.

As in most proofs for convergence of stochastic processes, the proof of Theorem 8 consists in two steps: convergence of the finite-dimensional distributions and tightness of the sequence of processes $(W_n)_{n\geq 1}$. To accomplish the above, one uses the random walk connected to the Prim path of the fragmentation forest of the α-stable GW-tree $\mathbf{t}_n$ in Proposition 6. More precisely, for $s \in [0,1]$, let $X_s = (X_s(k), k \geq 0)$ be the stochastic process defined in (5). For $n \in \mathbb{N}$ and $t \geq 0$, define the process $Y_n^{(t)} = (Y_n^{(t)}(u), u \in [0,1])$ by letting

$$Y_n^{(t)}(u) = \frac{1}{B_n} X_{s_n(t)}(\lfloor nu \rfloor), \quad \text{for } u \in [0,1],$$

and set $Y_n = (Y_n^{(t)}, t \geq 0)$. From Proposition 6, we see that W_n has the same finite-dimensional distribution as Y_n under the conditional probability distribution $\mathbb{P}_n(\cdot) := \mathbb{P}(\cdot|\zeta_1 = n)$. Therefore, the proof of Theorem 8 boils down to establishing the convergence for Y_n instead of W_n. Although it is simpler to work with Y_n than with W_n, the proof of the convergence is rather technical and it is given in the complete version [11, Section 5].

5 Proof of Theorem 1

In this section, we prove Theorem 1. The final ingredient is the general approach developed in the complete version for the convergence of fragmentation processes encoded by functions in $\mathbb{D}([0,1],\mathbb{R})$; see [11, Section 6]. Before that, we need to introduce some notation. For an increasing function $h = (h(s), s \in [0,1]) \in \mathbb{D}([0,1],\mathbb{R})$, write $\mathbf{F}(h) := (F_1(h), F_2(h), \dots) \in \mathbb{S}$, for the sequence of the lengths of the intervals components of the complement of the support of the Stieltjes measure $\mathrm{d}h$, arranged in decreasing order; we tacitly understand $\mathbf{F}(h)$ as an infinite sequence, by completing with an infinite number of zero terms.

Proof of Theorem 1. Let $\mathbf{t}_n$ be an α-stable GW-tree of index $\alpha \in (1,2]$. Recall that $(B_n)_{n\geq 1}$ denotes a sequence of positive real numbers satisfying (1). For $t \geq 0$, let $W_n^{(t)}$ be the (normalized and time-scaled) Prim path defined in (6) of the fragmentation forest at time $s_n(t) = 1 - (B_n/n)t$, i.e. $\mathbf{f}(s_n(t))$, associated to $\mathbf{t}_n$ and the i.i.d. uniform random weights $\mathbf{w}$. Define the process $I_n^{(t)} = (I_n^{(t)}(u), u \in [0,1])$ by letting

$$I_n^{(t)}(u) = \inf_{s\in[0,u]} W_n^{(t)}(s), \quad \text{for } s \in [0,1].$$

Recall that $\mathbf{F}_n^{(\alpha)} = (\mathbf{F}_n^{(\alpha)}(t), t \geq 0)$ stands for the fragmentation process of $\mathbf{t}_n$ defined in (2). From Lemma 5 and the preceding discussion, it is clear that $\mathbf{F}_n^{(\alpha)}(t) = \mathbf{F}(-I_n^{(t)})$, for $t \geq 0$. Let $Y_\alpha^{(t)}$ and $I_\alpha^{(t)}$ be the processes defined in (3), and recall that the α-stable fragmentation process, $\mathbf{F}^{(\alpha)} = (\mathbf{F}^{(\alpha)}(t), t \geq 0)$, is given by $\mathbf{F}^{(\alpha)}(t) = \mathbf{F}(-I_\alpha^{(t)})$, for $t \geq 0$. Note that for all $t \geq 0$, $W_n^{(t)}(0) = Y_\alpha^{(t)}(0) = 0$. Then, to prove Theorem 1, one only needs to check that the processes $W_n = (W_n^{(t)}, t \geq 0)$ and $Y_\alpha = (Y_\alpha^{(t)}, t \geq 0)$ satisfy the conditions of [11, Lemma 5].

We start by verifying that Y_α fulfills (i), (ii) and (iii) of [11, Lemma 5]. Indeed, (i) has been proven in Theorem 8. Recall that X_α^{exc} can be defined as the Vervaat transform of the so-called stable Lévy bridge; see [16]. Since the stable Lévy bridge has exchangeable increments (see e.g., [22, Chapters 11 and 16]), (ii) follows along the lines of the proof of Lemma 7 (i) in [7] thanks to [16, Theorem 4]. To prove that $Y_\alpha^{(t)}$ fulfills condition (iii) for every $t \geq 0$, recall that the support of the Stieltjes measure $\mathrm{d}(-I_\alpha^{(t)})$ coincides with the ladder time set $\mathcal{L}^\alpha(t)$ of $Y_\alpha^{(t)}$, which is a random closed set with zero Lebesgue measure. The latter follows from [5, Corollary 5, Chapter VII] but alternatively, it can be deduced from [16, Theorem 4] by following the same argument as in [7, Proof of Lemma 7]. Since $\mathbf{F}(-I_\alpha^{(t)})$ is defined as the ranked sequence of the lengths of the open intervals in the canonical decomposition of $[0,1]/\mathcal{L}^\alpha(t)$, condition (iii) follows.

We now check that the sequence $(W_n)_{n\geq 1}$ fulfills [11, (17)]. Note that, for every $t \geq 0$, $\|\mathbf{F}(-I_n^{(t)})\|_1 = 1$. Fix $t_\star, t^\star$ such that $0 \leq t_\star \leq t^\star < \infty$. For every $t \in [t_\star, t^\star]$ and $m \in \mathbb{N}$,

$$\|\mathbf{F}(-I_n^{(t)})\|_1 - \sum_{i=1}^m \mathbf{F}_i(-I_n^{(t)}) = \sum_{i>m} \mathbf{F}_i(-I_n^{(t)})$$

reaches its maximum at $t = t_\star$. Then for [11, (17)] to be satisfied, it suffices that for any $\varepsilon > 0$, there exists $m \in \mathbb{N}$ and $n \in \mathbb{N}$ such that

$$\sum_{i=1}^m \mathbf{F}_i(-I_n^{(t_\star)}) \geq \lim_{r\to\infty} \sum_{i=1}^r \mathbf{F}_i(-I_n^{(t_\star)}) - \varepsilon = 1 - \varepsilon. \tag{7}$$

This would imply that for any $t \in [t_\star, t^\star]$, we have that $\sum_{i=1}^m \mathbf{F}_i(-I_n^{(t)}) \geq 1 - \varepsilon$, which shows that $(W_n)_{n\geq 1}$ satisfies [11, (17)].

Theorem 8 implies that $(W_n^{(t)}, t \in [t_\star, t^\star]) \to (Y_\alpha^{(t)}, t \in [t_\star, t^\star])$, in distribution, as $n \to \infty$, in the space $\mathbb{D}([t_\star, t^\star], \mathbb{D}([0,1], \mathbb{R}))$. By the Skorokhod representation theorem, we can and we will work on a probability space on which this convergence holds almost surely. Since we have proven that the process $Y_\alpha^{(t_\star)}$ fulfills (iii) of [11, Lemma 5], for any $\varepsilon > 0$, there exists an $m \in \mathbb{N}$ such that $\sum_{i=1}^m \mathbf{F}_i(-I_\alpha^{(t_\star)}) \geq 1 - \varepsilon/2$. On the other hand, recall that $Y_\alpha^{(t_\star)}$ fulfills (ii) of [11, Lemma 5]. Then [7, Lemma 4] implies that a.s., $\mathbf{F}(-I_n^{(t_\star)}) \to \mathbf{F}(-I_\alpha^{(t_\star)})$, as $n \to \infty$ in the space $\mathbb{S}$ with the ℓ^1-norm. Hence, a.s. for all n large enough, $\sum_{i=1}^m \mathbf{F}_i(-I_n^{(t_\star)}) \geq 1 - \varepsilon$, which proves (7). ◀

References

1 Romain Abraham and Laurent Serlet. Poisson snake and fragmentation. *Electron. J. Probab.*, 7:no. 17, 15, 2002. doi:10.1214/EJP.v7-116.

2 David Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997. doi:10.1214/aop/1024404421.

3 David Aldous and Jim Pitman. The standard additive coalescent. *Ann. Probab.*, 26(4):1703–1726, 1998. doi:10.1214/aop/1022855879.

4 David Aldous and Jim Pitman. Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Related Fields*, 118(4):455–482, 2000. doi:10.1007/PL00008751.

5 Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

6 Jean Bertoin. A fragmentation process connected to Brownian motion. *Probab. Theory Related Fields*, 117(2):289–301, 2000. doi:10.1007/s004400050008.

7 Jean Bertoin. Eternal additive coalescents and certain bridges with exchangeable increments. *Ann. Probab.*, 29(1):344–360, 2001. doi:10.1214/aop/1008956333.

8 Jean Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002. doi:10.1016/S0246-0203(00)01073-6.

9 Jean Bertoin. Almost giant clusters for percolation on large trees with logarithmic heights. *J. Appl. Probab.*, 50(3):603–611, 2013. doi:10.1239/jap/1378401225.

10 Jean Bertoin and Grégory Miermont. Asymptotics in Knuth's parking problem for caravans. *Random Structures Algorithms*, 29(1):38–55, 2006. doi:10.1002/rsa.20092.

11 Gabriel Berzunza Ojeda and Cecilia Holmgren. Invariance principle for fragmentation processes derived from conditioned stable Galton-Watson trees. *arXiv e-prints*, page arXiv:2010.07880, October 2020. arXiv:2010.07880.

12 Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication. doi:10.1002/9780470316962.

13 Nicolas Broutin and Jean-François Marckert. A new encoding of coalescent processes: applications to the additive and multiplicative cases. *Probab. Theory Related Fields*, 166(1-2):515–552, 2016. doi:10.1007/s00440-015-0665-1.

14 P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms*, 21(1):76–119, 2002. doi:10.1002/rsa.10039.

15 Philippe Chassaing and Svante Janson. A Vervaat-like path transformation for the reflected Brownian bridge conditioned on its local time at 0. *Ann. Probab.*, 29(4):1755–1779, 2001. doi:10.1214/aop/1015345771.

16 L. Chaumont. Excursion normalisée, méandre et pont pour les processus de Lévy stables. *Bull. Sci. Math.*, 121(5):377–403, 1997.

17 Thomas Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.*, 31(2):996–1027, 2003. doi:10.1214/aop/1048516543.

18 Thomas Duquesne and Jean-François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, 281:vi+147, 2002.

19 Steven N. Evans and Jim Pitman. Construction of Markovian coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(3):339–383, 1998. doi:10.1016/S0246-0203(98)80015-0.

20 William Feller. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

21 Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003. doi:10.1007/978-3-662-05265-5.

22 Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002. doi:10.1007/978-1-4757-4015-8.

23 Jean-François Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005. doi:10.1214/154957805100000140.

24 Jean-François Marckert and Minmin Wang. A new combinatorial representation of the additive coalescent. *Random Structures Algorithms*, 54(2):340–370, 2019. doi:10.1002/rsa.20775.

25 Grégory Miermont. Ordered additive coalescent and fragmentations associated to Levy processes with no positive jumps. *Electron. J. Probab.*, 6:no. 14, 33, 2001. doi:10.1214/EJP.v6-87.

26 Grégory Miermont. Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. *Probab. Theory Related Fields*, 131(3):341–375, 2005. doi:10.1007/s00440-004-0373-8.

27 J. Neveu. Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(2):199–207, 1986. URL: http://www.numdam.org/item?id=AIHPB_1986__22_2_199_0.

28 Jim Pitman. Coalescent random forests. *J. Combin. Theory Ser. A*, 85(2):165–193, 1999. doi:10.1006/jcta.1998.2919.

29 R. C. Prim. Shortest connection networks and some generalizations. *The Bell System Technical Journal*, 36(6):1389–1401, 1957.

30 Paul Thévenin. A geometric representation of fragmentation processes on stable trees. *The Annals of Probability*, 49(5):2416–2476, 2021. doi:10.1214/21-AOP1512.

A Modification of the Random Cutting Model

Fabian Burghart ✉
Department of Mathematics, Uppsala University, Sweden

Abstract

We propose a modification to the random destruction of graphs: Given a finite network with a distinguished set of sources and targets, remove (cut) vertices at random, discarding components that do not contain a source node. We investigate the number of cuts required until all targets are removed, and the size of the remaining graph. This model interpolates between the random cutting model going back to Meir and Moon [28] and site percolation. We prove several general results, including that the size of the remaining graph is a tight family of random variables for compatible sequences of expander-type graphs, and determine limiting distributions complete binary trees.

2012 ACM Subject Classification Mathematics of computing → Probabilistic algorithms

Keywords and phrases Random cutting model, Random separation of graphs, Percolation

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.4

Related Version *Full Version*: `https://arxiv.org/abs/2111.02968` [9]

Funding Partially supported by grants from the Knut and Alice Wallenberg Foundation, the Ragnar Söderberg Foundation, and the Swedish Research Council.

Acknowledgements The author wishes to thank his academic advisors, Cecilia Holmgren and Svante Janson, for their generous support and many helpful remarks and discussions. The author also wishes to thank the anonymous referees for their helpful comments.

1 Introduction and structure of the paper

Investigating the behaviour of trees when randomly removing vertices was first done by Meir and Moon in [28]. The process starts with a rooted tree, where at every step a uniformly chosen vertex is deleted, and all remaining components that do not include the root vertex are discarded. Since the process naturally stops once the root node has been cut, the question of interest is the random number of cuts needed to reach this state. In [28], the expected value and the variance of this random variable for a random labelled tree are found.

Since this initial paper, the model has been considered for different types of random and deterministic graphs, such as in the works of Panholzer [29, 30], Janson [20, 21] (see also [2] for an alternative approach by Addario-Berry, Broutin and Holmgren), or Holmgren [18, 19].

The random cutting model relates to both the record problem (as was observed in [21]) and to fragmentation processes and their Cut-trees, see e.g. [4], [5], or [8]. Moreover, it connects to Union-Find algorithms, see [31] and the further work in [13, 15, 22, 23].

In recent years, modifications of the original cutting model have been discussed: Kuba and Panholzer regarded the case of isolating a leaf, a general node, or multiple nodes simultaneously (instead of isolating the root) in [24, 25, 26], and Cai, Holmgren et al. proposed and investigated the k-cut model in [6, 10, 11], where a node is only removed after it has been cut for the k-th time.

In this paper, a different modification of the cutting model is introduced, where, additionally to one or several root vertices (which we will call sources) a second set of vertices (targets) are given. This allows for defining a stopping time for the cutting procedure on the graph by looking at the first moment when all of the targets have been removed (i.e. the

sources have been *separated* from the targets) – see Section 2 for the detailed definitions. We can then ask several natural questions, such as about the number of cuts necessary to separate the two sets and about the size of the remaining graph.

Section 3 will contain several basic estimates, and we will formalise the imprecise notion that separation interpolates between the cutting model and site percolation (Propositions 1 and 3). This requires the right definition that enables us to approximate such a graph by finite graphs, all while respecting the choice of sources and targets, see Definition 2.

In Section 4 we obtain the probability for a fixed subgraph to be the remaining graph at the time of separation. This leads to Theorem 9, which could be regarded as the main result of the paper and gives sufficient conditions for the size of the graph at separation to be a tight sequence of random variables when the graph approximates a locally finite infinite graph in the sense of Definition 2.

For the final section, our scope will focus on rooted trees, since their recursive nature can be used to simplify many of the arguments and calculations. The arguments here relate to earlier work by Devroye et al., [12, 14]. We consider the separation sizes and separation times for complete binary trees as an illustrating example, and finish by indicating potential future research directions.

In order to comply with the requirements for this extended abstract, proofs of auxiliary statements have been omitted, but can be found in the full version of this paper, [9]. Moreover, the full version contains several examples that were not included here.

2 Cutting procedures

2.1 Some notation

We will always use $G = (V(G), E(G))$ to denote a graph, consisting of its vertex and edge set, but will shorten the notation to $V = V(G)$ and $E = E(G)$ if there is no ambiguity from the context. Since most subgraphs we will consider are induced and therefore uniquely determined by their vertex set, we will not distinguish between an induced subgraph and its vertex (sub-)set.

If two vertices v, w are neighbours, we also use the notation $v \sim w$. More generally, we write $\mathrm{dist}(v, w)$ for the graph distance between vertices v, w. In the case where $A, B \subseteq V(G)$ are subsets, $\mathrm{dist}(A, B)$ is to be understood as $\min\{\mathrm{dist}(v, w) : v \in A, w \in B\}$.

Given any set $A \subseteq V(G)$ and a fixed set S of source nodes, we define the closure of A to be $\mathrm{clos}_S(A) := A \cup S \cup \{v \in V(G) : v \sim w \text{ for some } w \in A\}$. The (exterior) boundary of A is defined as $\partial_S A := \mathrm{clos}_S(A) \setminus A$. In other words, the vertices in $\partial_S A$ are precisely the vertices not in A that are in S or neighbour some vertex in A. Note that this implies e.g. $\partial_S \emptyset := S$.

2.2 Cutting and separation

Consider a finite simple connected graph $G = (V, E)$ with a distinguished subset $S \subseteq V$ whose vertices are referred to as sources. Now, proceed as follows:

1. Choose a vertex $v \in V$ uniformly at random, and remove it – together with all edges incident to v – from the graph. This will potentially split the graph into connected components, in which case we only keep the components containing sources, regarding them jointly as a new (potentially disconnected graph).
2. Iterate step 1, where the randomness in choosing the node is assumed to be independent from everything that happened previously.

3. The process terminates once the graph contains no more vertices. Equivalently, this happens as soon as the last source node has been removed.

This defines a finite sequence $G =: G_0 \supsetneq G_1 \supsetneq ... \supsetneq G_{r-1} \supsetneq G_r = \emptyset$ of random induced subgraphs, where we denoted the empty subgraph consisting of no vertices with $\emptyset$. We will denote this process by $\mathsf{Cut}(G)$.

Introducing a second set of distinguished vertices, T, whose vertices we refer to as targets, we can now consider the following functionals of the cutting process:

- The cutting number $\mathfrak{C}(G)$. This is merely the number of cuts until the last source node is cut, or equivalently, until the remaining graph is empty, i.e. $\mathfrak{C}(G) = \inf\{i \geq 0 : G_i = \emptyset\}$. Note that this does not rely on T.
- The separation number $\mathfrak{S}(G)$, defined to be the number of cuts until the remaining graph does not contain target nodes anymore (independently of how many sources are still present). In other words, $\mathfrak{S}(G) = \inf\{i \geq 0 : V(G_i) \cap T = \emptyset\}$. We say that at this time, separation (of S and T) occurs.
- The separation subgraph $G_{\mathfrak{S}} := G_{\mathfrak{S}(G)}$, defined to be the random subgraph of G at separation.

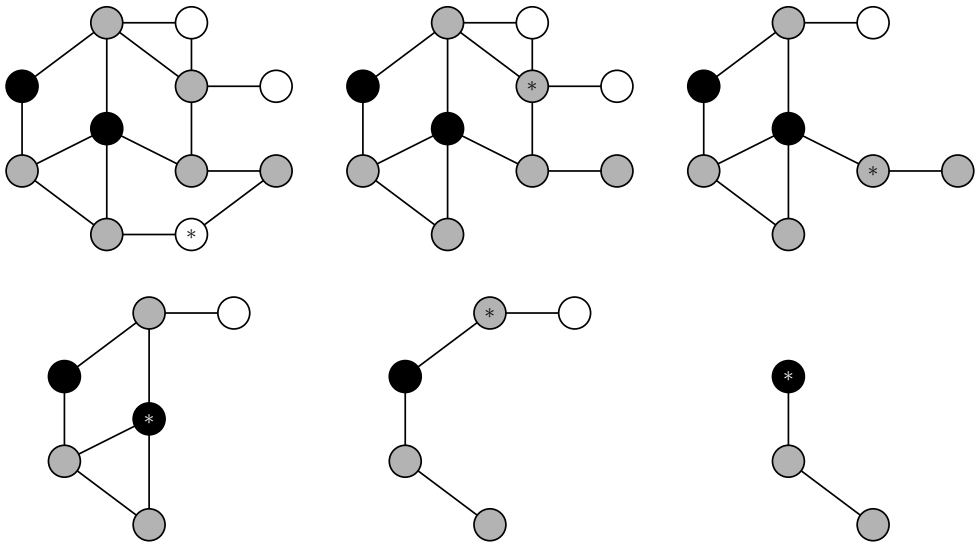


Figure 1 The cutting procedure on a graph G from top left to bottom right, with source nodes in black and target nodes in white. The symbol $*$ in a vertex indicates that this is the vertex about to be cut in the next step. In this example, $\mathfrak{S}(G) = 5$, $\mathfrak{C}(G) = 6$ (after the last source node has been cut) and $G_{\mathfrak{S}}$ is a path on three vertices.

2.3 The continuous-time model

As has been observed previously by [21] and since been brought to effective use, the above cutting model is equivalent to a model where each node is equipped with a random alarm clock whose alarm triggers after time X_v, $v \in V$. Whenever an alarm rings, and the corresponding node is still in the graph at this time, that node will be removed together with any new components that do not contain source nodes. Here, to ensure equivalence to the discrete-time cutting model above, we assume that $(X_v)_{v \in V}$ is an i.i.d. family of $\mathsf{Exp}(1)$-distributed random variables – while any i.i.d. family of continuous random variables would suffice, the memoryless property will be useful later.

This once again yields a monotone stochastic process of subgraphs of G, but now parametrised by continuous time, $(G_t^c)_{t\in[0,\infty)}$. We will denote this process by $\mathsf{Cut}^c(G)$. However, G_t will only attain finitely many different graphs, and we will still denote those graphs by $G_0, G_1, G_2, ...$ in order of occurrence, as before. Hence, we can denote by G_{t^-} the graph that was attained by $\mathsf{Cut}^c(G)$ immediately before time t; so, $G_{t^-} = G_t$ iff no cut happened at time t.

Note that there are two ways of generalising the random variables $\mathfrak{C}$ and $\mathfrak{S}$ to the continuous-time setting: By default, $\mathfrak{C}$ and $\mathfrak{S}$ respectively denote the quantities

$$\mathfrak{C} = \inf\{i \in \mathbb{N} : G_i = \emptyset\} \quad \text{and} \quad \mathfrak{S} = \inf\{i \in \mathbb{N} : V(G_i) \cap T = \emptyset\},$$

exactly as before, while $\mathfrak{C}_c$ and $\mathfrak{S}_c$ denote

$$\mathfrak{C}_c = \inf\{t \geq 0 : G_t = \emptyset\} \quad \text{and} \quad \mathfrak{S}_c = \inf\{t \geq 0 : V(G_t) \cap T = \emptyset\},$$

respectively.

3 Cutting and site percolation

The following proposition asserts that the additional freedom of choosing target nodes for the separation number can be used to obtain the original cutting number. In other words, $\mathfrak{S}(G)$ can be understood as a generalisation of $\mathfrak{C}(G)$.

▶ **Proposition 1.** *Let $G = (V, E)$ be a finite connected graph, and let $S, T \subseteq V$ be the sets of source and target nodes, respectively. Then, we have $\mathfrak{S}(G) \leq \mathfrak{C}(G)$ deterministically, with equality if $S \subseteq T$. Conversely, if $\mathfrak{S}(G) = \mathfrak{C}(G)$ in distribution, then $S \subseteq T$. Moreover, all of those statements also hold true for $\mathfrak{S}_c(G)$ and $\mathfrak{C}_c(G)$ in the continuous-time model.*

We remark that therefore, $\mathfrak{S}(G) = \mathfrak{C}(G)$ holds in distribution if and only if it holds deterministically.

Proof. At time $\mathfrak{C}(G)$, the remaining graph is empty, so separation must have occurred already. Thus $\mathfrak{S}(G) \leq \mathfrak{C}(G)$.

If $S \subseteq T$, then separation will occur as soon as the last source node has been removed, at which time the remaining graph will be empty. Thus $\mathfrak{S}(G) = \mathfrak{C}(G)$. Conversely, if there exists $v \in S \setminus T$ then $G_{\mathfrak{S}}$ contains v with some positive probability p_0. If this happens, $\mathfrak{S}(G) \leq \mathfrak{C}(G) - 1$, so $\mathbf{E}[\mathfrak{C}(G)] - \mathbf{E}[\mathfrak{S}(G)] \geq p_0$, and equality in distribution cannot hold.

For the continuous-time model, only the last argument requires modification: Once again, if $v \in S \setminus T$ then $G_{\mathfrak{S}}$ contains v with probability $p_0 > 0$. In this case, $\mathfrak{S}_c(G) < X_v \leq \mathfrak{C}_c(G)$, and we have

$$\mathbf{E}[\mathfrak{C}_c(G)] - \mathbf{E}[\mathfrak{S}_c(G)] \geq p_0 \mathbf{E}[X_v - \mathfrak{S}_c(G) \mid v \in V(G_{\mathfrak{S}})] = p_0$$

since $X_v \sim \mathsf{Exp}(1)$ is memoryless. ◀

We now show that in a certain sense, the continuous-time separation model on an infinite graph G with infinite distance $\mathrm{dist}(S, T)$ contains the site percolation model on G.

More precisely, recall that for Bernoulli site percolation in an infinite graph G, every node is independently kept with some probability $p \in [0, 1]$ and otherwise rejected, thus giving a random subgraph of G. We denote by $\mathrm{perc}_S(p)$ the probability that the $\mathsf{Ber}(p)$-site percolation on G exhibits an infinite cluster containing at least one vertex of S.

▶ **Definition 2.** *Let G be a locally finite, infinite connected graph, containing two subsets $S, T \subseteq V(G)$. We say that the sequence $(G^{(n)})_{n\geq 1}$ of finite induced subgraphs of G* exhausts *G if the following conditions are satisfied:*

(i) *The $G^{(n)}$ are connected subgraphs satisfying*

$$G^{(1)} \subseteq G^{(2)} \subseteq \ldots \subseteq G \quad \text{and} \quad \bigcup_{n\geq 1} V\left(G^{(n)}\right) = V(G).$$

(ii) *The set S is entirely contained in $G^{(n)}$ for all n (and understood to be the set of source nodes of $G^{(n)}$), and each subgraph $G^{(n)}$ is endowed with the target set $T^{(n)} := \left(T \cap V\left(G^{(n)}\right)\right) \cup \partial_\emptyset\left(G \setminus G^{(n)}\right)$.*

Observe that the target nodes $T^{(n)}$ are indeed a subset of $V\left(G^{(n)}\right)$ and will be non-empty even if $T = \emptyset$. Moreover, condition (ii) necessitates that S is finite.

▶ **Proposition 3.** *Let G be a locally finite, infinite graph with a finite set of source nodes, S, and let $T = \emptyset$. Assume that $(G^{(n)})_{n\geq 1}$ exhausts G.*

Then,

$$\lim_{n\to\infty} \mathbf{P}\left[\mathfrak{S}_c\left(G^{(n)}\right) \geq x\right] = \mathrm{perc}_S(e^{-x}). \tag{1}$$

Proof. Note first that if $T = \emptyset$, then (using the notation of Definition 2) $\mathrm{dist}\left(S, T^{(n)}\right) \to \infty$ as $n \to \infty$. Indeed, there would otherwise be a bound $C \in \mathbb{N}$ with $\mathrm{dist}\left(S, T^{(n)}\right) < C$. Consider the neighbourhood $B_C(S) := \{v \in V(G) : \mathrm{dist}(v, S) \leq C\}$ of S. Since $B_C(S)$ is finite, it will eventually be contained in all $G^{(n)}$, contradicting the notion that $T^{(n)} = \partial_\emptyset\left(G \setminus G^{(n)}\right)$ contains vertices in $B_{C-1}(S)$.

Recall next that independently removing each vertex v of G at a random time $X_v \sim \mathsf{Exp}(1)$ gives rise to the monotonous coupling of Bernoulli site percolation for all parameters $p \in [0, 1]$ (cf. [27, p.138]). Indeed, at time $x \in [0, \infty]$, the graph we observe that way is a sample of $\mathsf{Ber}(e^{-x})$-site percolation on G. We can couple the process obtained in this way to the continuous-time cutting model by restricting our attention to the intersection of $G^{(n)}$ with those percolation clusters that intersect S.

To show $\geq$ in (1), assume that $\mathsf{Ber}(e^{-x})$-site percolation exhibits an infinite cluster which intersects S, such cluster necessarily intersects $T^{(n)}$ as well and hence, for each n, contains a path connecting S with $T^{(n)}$. By the coupling indicated above, this path must then also be present in the sample of the continuous-time cutting model on $G^{(n)}$ at time x. Therefore, $\mathrm{perc}_S(e^{-x}) \leq \mathbf{P}\left[\mathfrak{S}_c\left(G^{(n)}\right) \geq x\right]$, and letting n tend to ∞ yields $\liminf_{n\to\infty} \mathbf{P}\left[\mathfrak{S}_c\left(G^{(n)}\right) \geq x\right] \geq \mathrm{perc}_S(e^{-x})$.

For the other inequality, suppose now that $\mathsf{Ber}(e^{-x})$-site percolation does not exhibit an infinite cluster intersecting S, so that the total mass of clusters intersecting S is bounded by some finite integer, say k. By the second assumption, we have $\mathrm{dist}(S, T^{(n)}) > k$ for all but finitely many n. However, this implies that eventually, the clusters intersecting S cannot intersect $T^{(n)}$, which, for the coupled cutting procedure, means that separation in $G^{(n)}$ must have occurred before time x. So,

$$\mathbf{P}\left[|V(\text{Clusters intersecting } S)| \leq k\right] \leq \mathbf{P}\left[\mathfrak{S}_c\left(G^{(n)}\right) < x\right].$$

Taking the limit superior for $n \to \infty$ yields

$$\limsup_{n\to\infty} \mathbf{P}\left[\mathfrak{S}_c\left(G^{(n)}\right) \geq x\right] \leq 1 - \mathbf{P}\left[|V(\text{Clusters intersecting } S)| \leq k\right],$$

which implies the existence of the limit and $\leq$ in (1) after passing to the limit for $k \to \infty$ as well. ◀

4 Visiting probability of subgraphs and size of the separation graph

Consider a finite simple connected graph G with $S, T \subseteq V$ as usual. The aim of this section is to determine the probability that at some time $i \geq 1$, the cutting procedure $\mathsf{Cut}(G)$ will produce a specific subgraph G_*.

▶ **Lemma 4.** *Fix an induced subgraph G_* of G with every component of G_* containing at least one source node. Then, for all times $t \geq 0$ in the continuous-time cutting model, we have*

$$\mathbf{P}\left[G_t = G_*\right] = e^{-|G_*|t}\left(1 - e^{-t}\right)^{|\partial_S G_*|}. \tag{2}$$

Moreover, consider $v_* \in \partial_S G_*$. Then

$$\mathbf{P}\left[G_{X_{v_*}} = G_* \mid X_{v_*}\right] = e^{-|G_*|X_{v_*}}\left(1 - e^{-X_{v_*}}\right)^{|\partial_S G_*|-1}. \tag{3}$$

▶ **Corollary 5.** *Fix an induced subgraph G_* of G with every component of G_* containing at least one source node. Let $v_* \in \partial_S G_*$. Denote the i-th graph obtained in the cutting process by G_i and the i-th cut node by v_i. Then*

$$\mathbf{P}\left[\exists i \geq 1 : G_i = G_* \text{ and } v_i = v_*\right] = \frac{1}{|\partial_S G_*|}\binom{|G_*| + |\partial_S G_*|}{|\partial_S G_*|}^{-1} \tag{4}$$

and therefore

$$\mathbf{P}\left[\exists i \geq 1 : G_i = G_*\right] = \binom{|G_*| + |\partial_S G_*|}{|\partial_S G_*|}^{-1}. \tag{5}$$

▶ **Definition 6.** *Let G be a finite connected graph equipped with sources S and targets $T \neq \emptyset$. An induced subgraph G_* is called* admissible *if the probability $\mathbf{P}[G_{\mathfrak{S}} = G_*]$ is positive. We denote by $\mathscr{A}_m(G)$ the set of all admissible subgraphs G_* of G of size $|V(G_*)| = m$.*

Assuming that $T \neq \emptyset$, it can be shown (Lemma 12 in [9]) that $G_* \subseteq G$ is admissible if and only if G_* contains no target nodes and every component of G_* contains at least one source node.

Relying on the preceding results, we can establish the following connection between the graph $G_{\mathfrak{S}}$ at separation and the continuous-time separation number $\mathfrak{S}_c$:

▶ **Proposition 7.** *Let G_* be an admissible subgraph of G. Then,*

$$\mathbf{P}[G_{\mathfrak{S}} = G_*] = \sum_{v_* \in \partial_S G_*} \int_0^1 u^{|G_*|-1}(1-u)^{|\partial_S G_*|-1}\mathbf{P}\left[\mathfrak{S}_c(H[v_*]) \geq -\ln u\right] \mathrm{d}u, \tag{6}$$

where $H[v_*]$ denotes the graph obtained from G in the following way: Remove all vertices in G_* and in $\partial_S G_* \setminus \{v_*\}$ from G, and from what remains, let $H[v_*]$ be the connected component containing v_*. We endow $H[v_*]$ with source node v_* and target nodes $T \cap V(H[v_*])$.

Proof. Fix a vertex $v_* \in \partial_S G_*$, and assume that this is the last node to be removed for separation to occur. We observe first that by definition of the separation number, any graphs obtained by $\mathsf{Cut}(G)$ before separation must have contained a path from v_* to T. In particular, the last graph before separation occurred contained such a path, which additionally was not

passing through any other nodes in G_* or $\partial_S G_*$ and must have therefore been contained in $H[v_*]$. The existence of such a path means, however, that the graph $H[v_*]$ is not yet separated. Thus, by transitioning from the discrete to the continuous-time model, we obtain

$$\begin{aligned}
&\mathbf{P}[G_{\mathfrak{S}} = G_*, v_{\mathfrak{S}} = v_* \mid X_{v_*}] \\
&\quad = \mathbf{P}\left[G_{X_{v_*}} = G_* \text{ and } v_* \text{ connects to } T \text{ in } G_{X_{v_*}^-} \mid X_{v_*}\right] \\
&\quad = \mathbf{P}\left[G_{X_{v_*}} = G_* \mid X_{v_*}\right] \mathbf{P}\left[\mathfrak{S}_c(H[v_*]) \geq X_{v_*} \mid X_{v_*}\right], \qquad (7)
\end{aligned}$$

where conditional independence holds true because

$$\{G_{X_{v_*}} = G_*\} \text{ and } \{\mathfrak{S}_c(H[v_*]) \geq X_{v_*}\}$$

are events on vertex sets which only share v_*. Conditioned on X_{v_*} being x, the event $\{\mathfrak{S}_c(H[v_*]) \geq X_{v_*}\}$ amounts to the existence of a path from v_* to the set of target nodes in $H[v_*]$, none of whose clocks have rung yet at time X_{v_*}. On the other hand, without the conditioning, the same event is equivalent to the existence of a path from a neighbour of v_* to the set of target nodes in $H[v_*]$. Hence,

$$\mathbf{P}\left[\mathfrak{S}_c(H[v_*]) \geq X_{v_*} \mid X_{v_*} = x\right] = \frac{\mathbf{P}\left[\mathfrak{S}_c(H[v_*]) \geq x\right]}{e^{-x}}.$$

In light of (3) from Lemma 4, we can now rewrite equation (7) as

$$\mathbf{P}[G_{\mathfrak{S}} = G_*, v_{\mathfrak{S}} = v_* \mid X_{v_*} = x] = e^{-|G_*|x} \left(1 - e^{-x}\right)^{|\partial_S G_*| - 1} \frac{\mathbf{P}\left[\mathfrak{S}_c(H[v_*]) \geq x\right]}{e^{-x}}. \qquad (8)$$

Finally, observe that, with μ_X denoting the distribution of X_{v_*},

$$\mathbf{P}[G_{\mathfrak{S}} = G_*] = \sum_{v_* \in \partial_S G_*} \int_0^\infty \mathbf{P}[G_{\mathfrak{S}} = G_*, v_{\mathfrak{S}} = v_* \mid X_{v_*} = x] \, \mathrm{d}\mu_X(x),$$

so that, after plugging in the expression from (8) and using $X_{v_*} \sim \mathsf{Exp}(1)$, we obtain

$$\mathbf{P}[G_{\mathfrak{S}} = G_*] = \sum_{v_* \in \partial_S G_*} \int_0^\infty (e^{-x})^{|G_*|} \left(1 - e^{-x}\right)^{|\partial_S G_*| - 1} \frac{\mathbf{P}\left[\mathfrak{S}_c(H[v_*]) \geq x\right]}{e^{-x}} \cdot e^{-x} \, \mathrm{d}x,$$

which only differs from (6) by the substitution $e^{-x} = u$. ◀

▶ Remark 8. The (unconditioned) probability $\mathbf{P}[\mathfrak{S}_c(H[v_*]) \geq X_{v_*}]$, has a number of equivalent versions. Indeed, if we consider any G with $S = \{v_*\}$, then we have the following equalities:

$$\mathbf{P}[\mathfrak{S}_c(G) \geq X_{v_*}] = \mathbf{P}[\mathfrak{S}_c(G) \geq \mathfrak{C}_c(G)] = \mathbf{P}[\mathfrak{S}(G) \geq \mathfrak{C}(G)] = \mathbf{P}[G_{\mathfrak{S}} = \emptyset] = \mathbf{P}[|G_{\mathfrak{S}}| = 0].$$

Moreover, in the first three formulations, the strict inequality "$>$" is impossible, so one could just as well write "$=$" there. Additionally, since $\mathfrak{S}_c(H[v_*]) \leq X_{v_*}$, we have the estimate

$$\mathbf{P}\left[\mathfrak{S}_c(H[v_*]) \geq -\ln u\right] \leq \mathbf{P}\left[X_{v_*} \geq -\ln u\right] = u \qquad \forall u \in [0, 1]. \qquad (9)$$

Recall that a family of real-valued random variables $X_i, i \in I$, is *tight* if for all $\varepsilon > 0$, there exists a constant M such that $\mathbf{P}[|X_i| \geq M] < \varepsilon$ for all $i \in I$, cf. [7, p.37]. We conclude this section by showing the following tightness result about the size of $G_{\mathfrak{S}}$:

▶ **Theorem 9.** *Let $(G^{(n)})_{n\geq 1}$ be a sequence of finite induced subgraphs exhausting the locally finite, infinite graph G equipped with subsets $S, T \subseteq V(G)$. Define $a_{m,n} = \sum_{A\in\mathscr{A}_m(G^{(n)})} |\partial_S A|$ and let $a_m := \lim_{n\to\infty} a_{m,n}$. Assume that there are constants $b > 0$ and $L \geq 0$ such that*

(i) *$|\partial_S A| \geq bm + 1$ for all $A \in \mathscr{A}_m(G^{(n)})$ and all $m \geq L$, and*

(ii) *the function $f(x) = \sum_{m=L}^{\infty} a_m x^m$ has radius of convergence at least $\frac{b^b}{(b+1)^{b+1}}$ and satisfies*

$$\int_0^1 f(x(1-x)^b)\,\mathrm{d}x < \infty \tag{10}$$

Then the sizes of the separation graphs, $\left|G_{\mathfrak{S}}^{(n)}\right|$, form a tight family of random variables.

Observe that, since $c := \max_{x\in[0,1]} x(1-x)^b = \frac{b^b}{(b+1)^{(b+1)}}$, condition (ii) requires that the radius of convergence, r, of the power series f is at least c. In case $r > c$, this condition is trivially satisfied as the integrand is bounded. However, in case of equality $r = c$, f has a singularity at c by virtue of Pringsheim's theorem, cf. [16, Theorem IV.6], and (10) is a non-trivial requirement. We also remark that condition (i) above is a variant of the notion of expander graphs, which play a crucial role in the theory of percolation, see e.g. [3].

Proof. We first show that for fixed m, the sequence $a_{m,n}$ is nondecreasing and eventually constant. Recall from Definition 2 that the sets $T^{(n)}$ consist of two parts, namely $T \cap V\left(G^{(n)}\right)$ and vertices that are incident to edges leaving $G^{(n)}$. By copying the respective argument from the proof of Proposition 3, we can show again that $\operatorname{dist}\left(S, T^{(n)} \setminus T\right) \to \infty$ as $n \to \infty$. Hence, for n sufficiently large, the closed neighbourhood of S of radius m will become independent of n, and so will $a_{m,n}$. Moreover, since the $G^{(n)}$ are monotonously growing, any admissible subgraph in $G^{(n)}$ will also be admissible in $G^{(n+1)}$ and the number of boundary vertices will not decrease when changing from $G^{(n)}$ to $G^{(n+1)}$. Hence $a_{m,n}$ is nondecreasing in n. In particular, the limit a_m exists, is finite, and an upper bound to $a_{m,n}$ for all n.

Let $M \geq L$. We now apply Proposition 7, where we set $p_{v_*}(x) := \mathbf{P}[\mathfrak{S}_c(H[v_*]) \geq -\ln x]$ (note that this still depends on $G_{\mathfrak{S}}$ as well!) for brevity. Summing over all $m \geq M$ and all $A \in \mathscr{A}_m(G^{(n)})$ yields

$$\begin{aligned}\mathbf{P}\left[\left|G_{\mathfrak{S}}^{(n)}\right| \geq M\right] &= \sum_{m=M}^{\infty} \sum_{A\in\mathscr{A}_m(G^{(n)})} \int_0^1 x^{m-1}(1-x)^{|\partial_S A|-1} \sum_{v_*\in\partial_S A} p_{v_*}(x)\,\mathrm{d}x \\ &\leq \sum_{m=M}^{\infty} \sum_{A\in\mathscr{A}_m(G^{(n)})} |\partial_S A| \int_0^1 x^m (1-x)^{|\partial_S A|-1}\,\mathrm{d}x,\end{aligned}$$

where we applied the estimate (9). Using assumption (i), we get

$$\begin{aligned}\mathbf{P}\left[\left|G_{\mathfrak{S}}^{(n)}\right| \geq M\right] &\leq \sum_{m=M}^{\infty} \sum_{A\in\mathscr{A}_m(G^{(n)})} |\partial_S A| \int_0^1 (x(1-x)^b)^m\,\mathrm{d}x \\ &= \sum_{m=M}^{\infty} \int_0^1 a_{m,n}(x(1-x)^b)^m\,\mathrm{d}x.\end{aligned}$$

Then, by the above argument on the monotonicity of $a_{m,n}$, we obtain

$$\mathbf{P}\left[\left|G_{\mathfrak{S}}^{(n)}\right| \geq M\right] \leq \sum_{m=M}^{\infty} \int_0^1 a_m(x(1-x)^b)^m\,\mathrm{d}x.$$

By monotone convergence, we moreover obtain for all $M \geq L$ that

$$\sum_{m=M}^{\infty} \int_0^1 a_m (x(1-x)^b)^m \,\mathrm{d}x = \int_0^1 \sum_{m=M}^{\infty} a_m (x(1-x)^b)^m \,\mathrm{d}x$$
$$\leq \int_0^1 \sum_{m=L}^{\infty} a_m (x(1-x)^b)^m \,\mathrm{d}x = \int_0^1 f(x(1-x)^b) \,\mathrm{d}x$$

which is finite by assumption (ii). Hence the tail of the series on the left-hand side converges to zero, and therefore $\mathbf{P}\left[\left|G_{\mathfrak{S}}^{(n)}\right| \geq M\right] \to 0$ uniformly in n as $M \to \infty$ as well. ◀

▶ Remark 10. Tightness of the sizes of the separation graphs $\left|G_{\mathfrak{S}}^{(n)}\right|$ is a helpful property in order to translate limit laws from the cutting times $\mathfrak{C}\left(G^{(n)}\right)$ to the separation times $\mathfrak{S}\left(G^{(n)}\right)$: Assume that there exist sequences α_n and $\beta_n > 0$ such that

$$\frac{\mathfrak{C}\left(G^{(n)}\right) - \alpha_n}{\beta_n} \xrightarrow{\mathscr{D}} X \qquad \text{as } n \to \infty$$

for a random variable X with positive variance. If $\beta_n \to \infty$ and $\left|G_{\mathfrak{S}}^{(n)}\right|$ forms a tight family of random variables, then the deterministic estimate $\mathfrak{S}\left(G^{(n)}\right) \leq \mathfrak{C}\left(G^{(n)}\right) \leq \mathfrak{S}\left(G^{(n)}\right) + \left|G_{\mathfrak{S}}^{(n)}\right|$ implies

$$\frac{\mathfrak{S}\left(G^{(n)}\right) - \alpha_n}{\beta_n} \xrightarrow{\mathscr{D}} X \qquad \text{as } n \to \infty.$$

5 Separating trees

In this section, let $G = \mathcal{T}$ be a rooted tree, where we will always interpret the root node as the (unique) source vertex, and the leaves as targets.

To each node $w \in V(\mathcal{T})$, we assign a polynomial $p[w]$ from $\mathbb{Z}[x]$ recursively as follows: If w is a leaf, define $p[w](x) = x$. Otherwise, denote the children of w by $w_1, ..., w_r$ for $r \geq 1$. Then, define

$$p[w](x) := x\left(1 - \prod_{i=1}^{r} (1 - p[w_i](x))\right). \tag{11}$$

Observe that in the case where w only has a single child w_1, this simplifies to $p[w](x) = xp[w_1](x)$. Furthermore, if $\mathcal{T}_*$ is the fringe subtree of $\mathcal{T}$ rooted at w (i.e. the subtree consisting of w together with all its descendants), we also write $p[\mathcal{T}_*] := p[w]$. In particular, $p[\mathcal{T}] := p[\text{root}]$.

▶ **Proposition 11.** *For the continuous-time cutting model on a rooted tree $\mathcal{T}$, we have*

$$\mathbf{P}[\mathfrak{S}_c(\mathcal{T}) \geq x] = p[\mathcal{T}](e^{-x}) \tag{12}$$

for all $x \geq 0$. Equivalently, one can interpret $p[\mathcal{T}](q)$ for $q \in [0,1]$ as the probability that $\mathsf{Ber}(q)$*-site percolation on $\mathcal{T}$ contains a path from the root to a leaf.*

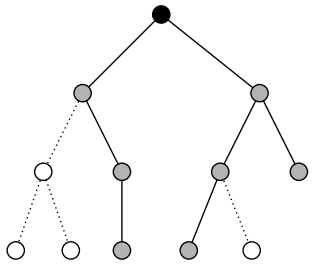


Figure 2 A faithful subtree of an underlying rooted tree $\mathcal{T}$. The root is shown in black, with vertices belonging to the subtrees being shaded grey. Dotted edges and white vertices belong to $\mathcal{T}$, but not to the subgraph.

▶ **Corollary 12.** *We have*

$$\mathbf{P}[G_{\mathfrak{S}} = \emptyset] = \int_0^1 \frac{p[\mathcal{T}](u)}{u}\,\mathrm{d}u. \tag{13}$$

Moreover, we have for any subtree $\mathcal{T}_* \subseteq \mathcal{T}$ containing the root node but none of the leaves:

$$\mathbf{P}[G_{\mathfrak{S}} = \mathcal{T}_*] = \sum_{v_* \in \partial\mathcal{T}_*} \int_0^1 u^{|\mathcal{T}_*|-1}(1-u)^{|\partial\mathcal{T}_*|-1} p[v_*](u)\,\mathrm{d}u. \tag{14}$$

A *transversal* in a rooted tree $\mathcal{T}$ is defined to be a subset of vertices that intersect every path from the root to a leaf. It then follows from the proof of Proposition 11 that $1-p[\mathcal{T}](1-q)$ yields the probability that a random set of vertices, containing each vertex independently with probability q, is a transversal of $\mathcal{T}$. It is this expression that was investigated in [14, 12].

▶ Remark 13. There are some straightforward conclusions to be drawn from the previous proposition: The polynomials $p[\mathcal{T}](x)$ map the interval $[0,1]$ to itself, with $p[\mathcal{T}](0) = 0$ and $p[\mathcal{T}](1) = 1$. Furthermore, they are monotonically increasing on this interval, and by (11), $p[\mathcal{T}](x) \leq x$ for $x \in [0,1]$.

▶ **Definition 14.** *Let $\mathcal{T}$ be a rooted tree. A subtree $\mathcal{T}_*$ is called* faithful *if it contains the root of $\mathcal{T}$ as its own root and if every leaf of $\mathcal{T}_*$ is a leaf of $\mathcal{T}$. We denote the set of all faithful subtrees of $\mathcal{T}$ by $\mathscr{F}(\mathcal{T})$, and the set of all faithful subtrees on n vertices by $\mathscr{F}_n(\mathcal{T})$. Equivalently, a faithful subtree $\mathcal{T}_* \subseteq \mathcal{T}$ can be seen as choosing a number of paths from the root to the leaves of $\mathcal{T}$. We hence denote the set of all faithful subtrees of $\mathcal{T}$ with exactly ℓ leaves by $\mathscr{P}_\ell(\mathcal{T})$.*

The following propositions gives an alternative, combinatorial characterisation of $p[\mathcal{T}]$ which might be of independent interest:

▶ **Proposition 15.** *Write $p[\mathcal{T}](x) = \sum_{j=0}^n a_j x^j$ with $a_j \in \mathbb{Z}$ for $j = 0, 1, ..., n$. Denote by $L(\mathcal{T}_*)$ the set of leaves of $\mathcal{T}_*$. Then,*

$$a_j = \sum_{\mathcal{T}_* \in \mathscr{F}_j(\mathcal{T})} (-1)^{|L(\mathcal{T}_*)|+1} \tag{15}$$

$$= \sum_{\ell \geq 1} (-1)^{\ell+1} \left|\mathscr{F}_j(\mathcal{T}) \cap \mathscr{P}_\ell(\mathcal{T})\right|. \tag{16}$$

In particular, $\deg(p[\mathcal{T}]) = |\mathcal{T}|$ *and* $a_j = 0$ *for* $j \leq \mathrm{dist}(\mathrm{root}, L(\mathcal{T}))$.

▶ **Example 16.** Denote by CBT_n the full complete rooted binary tree on $2^n - 1$ vertices, this being the rooted binary tree having 2^h vertices at every height $h = 0, ..., n-1$. Observe that CBT_1 is the tree consisting of only the root node, and that CBT_{n+1} splits into two copies of CBT_n upon removing the root node. Thus, the associated polynomials satisfy the recurrence relation

$$p[\mathrm{CBT}_{n+1}](x) = x(1 - (1 - p[\mathrm{CBT}_n](x))^2) = 2xp[\mathrm{CBT}_n](x) - xp[\mathrm{CBT}_n](x)^2$$

with $p[\mathrm{CBT}_1](x) = x$. It can be shown that $p[\mathrm{CBT}_n]$ converges pointwise and monotonically decreasing to

$$\varphi(x) := \max\left\{0, 2 - \frac{1}{x}\right\} \tag{17}$$

as $n \to \infty$, for $x \in [0, 1]$.

In light of Proposition 3, this result should not be surprising: The sequence of rooted trees CBT_n satisfies all the conditions, and the function $\varphi(x)$ indeed equals the probability that the root node is contained in an infinite cluster of $\mathsf{Ber}(x)$-site percolation, as can be verified independently, e.g. from [17, p.256].

Continuing with our analysis, the probability of the remaining tree at separation being empty now follows handily from Corollary 12:

$$\lim_{n\to\infty} \mathbf{P}[|\mathrm{CBT}_{n,\mathfrak{S}}| = 0] = \int_0^1 \frac{\varphi(x)}{x}\,\mathrm{d}x = \ln 4 - 1 \approx 0.3863$$

In a similar fashion, we can continue to determine the limiting probability of separation graphs of any size $m \geq 0$: Since there are $C_m = \frac{1}{m+1}\binom{2m}{m}$-many[1] subtrees of the infinite rooted binary tree on m vertices, and each of those has $m+1$ boundary vertices, it can be shown using the dominated convergence theorem that

$$p_m := \lim_{n\to\infty} \mathbf{P}[|\mathrm{CBT}_{n,\mathfrak{S}}| = m] = \binom{2m}{m} \int_0^1 x^{m-1}(1-x)^m \varphi(x)\,\mathrm{d}x. \tag{18}$$

Observe that, in the notation of Theorem 9, the sequence $a_m = \binom{2m}{m}$ has generating function $\frac{1}{\sqrt{1-4x}}$ (cf. [1, OEIS A000984]), thus violating the integrability condition (10) for the proper constant $b = 1$. However, we can check by hand that (18) defines a probability distribution with the following computation, relying on the generating function of $\binom{2m}{m}$:

$$\begin{aligned}\sum_{m=0}^{\infty} p_m &= \int_{1/2}^1 \left(\sum_{m=0}^{\infty} \binom{2m}{m}(x - x^2)^m\right) \cdot \frac{2x-1}{x^2}\,\mathrm{d}x \\ &= \int_{1/2}^1 \frac{1}{\sqrt{1 - 4x + 4x^2}} \cdot \frac{2x-1}{x^2}\,\mathrm{d}x = \int_{1/2}^1 \frac{1}{x^2}\,\mathrm{d}x = 1. \end{aligned} \tag{19}$$

Note also that a random variable X having the probability distribution defined by (18) does not have a finite first moment: Imitating the approach of (19) leads to

$$\mathbf{E}[X] = \sum_{m=0}^{\infty} m\binom{2m}{m} \int_0^1 x^{m-1}(1-x)^m \varphi(x)\,\mathrm{d}x = \int_{1/2}^1 \frac{2(1-x)}{x(2x-1)^2}\,\mathrm{d}x$$

where the integral on the right-hand side diverges.

[1] These are, of course, the Catalan numbers, [1, OEIS A000108].

A more detailed investigation of how $p[\mathcal{T}]$ depends on the tree $\mathcal{T}$ reveals that two trees that only differ in a fringe tree which is rooted far away from the root node will have approximately the same polynomial function $p[\,\cdot\,]$ over $[0,1]$, provided a technical condition holds, see [9, Theorem 25]. Using this, it is possible to verify that the same limiting distribution as in (18) also holds if we consider the sequence of complete binary trees on n vertices (of which the full complete binary trees are merely a subsequence):

Denote by $\mathcal{T}_n$ the complete binary tree on n vertices, this being the binary tree having 2^k vertices at height k for $0 \leq k < \lfloor \lg n \rfloor =: m$, with the remaining $n - 2^m + 1$ vertices at height m in their left-most positions.

▶ **Proposition 17.** *With $\varphi(x)$ defined by equation (17) above, we have $p[\mathcal{T}_n](x) \to \varphi(x)$ uniformly over $[0,1]$, as $n \to \infty$.*

Consequently, the limiting distribution of $|\mathcal{T}_{n,\mathfrak{S}}|$ coincides with that of $|\mathrm{CBT}_{n,\mathfrak{S}}|$ and is given by equation (18).

Observe that by this proposition and by (19), the random variables $|\mathcal{T}_{n,\mathfrak{S}}|$ converge in distribution and are therefore tight. By Remark 10 this means that the limit law obtained by Janson in [20, Theorem 1.1] for $\mathfrak{C}(\mathcal{T}_n)$ holds also for $\mathfrak{S}(\mathcal{T}_n)$. More explicitly, if we denote by $\{x\} := x - \lfloor x \rfloor$ the fractional part of $x \in \mathbb{R}$ then we obtain the following

▶ **Corollary 18.** *Let $n \to \infty$ such that $\{\lg n - \lg\lg n\} \to \gamma \in [0,1]$. Write $f(\gamma) := 2^\gamma - 1 - \gamma$. Let W_γ be a random variable with an infinitely divisible distribution with characteristic function*

$$\mathbf{E}\left[e^{itW_\gamma}\right] = \exp\left(if(\gamma)t + \int_0^\infty \left(e^{itx} - 1 - itx\mathbb{1}_{\{x<1\}}\right) \mathrm{d}\nu_\gamma(x)\right),$$

with the Lévy measure ν_γ being supported on $[0,\infty)$ and having density $\mathrm{d}\nu_\gamma = 2^{\{\lg x + \gamma\}} x^{-2}\,\mathrm{d}x$. Then

$$\frac{\lg^2 n}{n}\left(\mathfrak{S}(\mathcal{T}_n) - \frac{n}{\lg n} - \frac{n\lg\lg n}{\lg^2 n}\right) \xrightarrow{\mathscr{D}} -W_\gamma$$

6 Further questions

We use this final section to present several, deliberately broad questions or remarks that could lead to interesting future research.

1. Determine the asymptotic distributions of $\mathfrak{S}(G^{(n)})$ and $\left|G^{(n)}_{\mathfrak{S}}\right|$ for other families of deterministic and random trees. The author hopes to answer this for conditioned Galton-Watson trees in a follow-up paper.
2. What happens if the roles of S and T are exchanged? For which graphs and which choices of S, T are the random variables $\mathfrak{S}(G;S,T)$ and $\mathfrak{S}(G;T,S)$ equal in distribution?
3. How to evaluate the asymptotic distribution of $\mathfrak{S}$ directly, without relying on previous knowledge of $\mathfrak{C}$ as in Corollary 18?
4. It is easy to see that the edge-cutting process on a graph G is exactly the vertex-cutting process on the line graph of G. This therefore raises the question: How is the separation time on G related to the separation time on the line graph of G?
5. For which sequences of graphs $G^{(n)}$ exhausting a locally finite infinite G (with fixed sources and targets) are the random variables $\left|G^{(n)}_{\mathfrak{S}}\right|$ *not* tight? In this case, what can be said about the structure of the remaining graph?

6. By definition, the separation number is the number of cuts required to separate two subsets $S, T \subseteq V$ from each other. Starting in a graph with a high connectivity (say, by having $k \geq 2$ vertex-disjoint paths from S to T in G) we can ask for the number of cuts required to reduce the connectivity to some $j \leq k$. The case $j = 0$ specializes to the separation number as we defined it.

References

1 OEIS Foundation Inc. (2021). The on-line encyclopedia of integer sequences. URL: `http://oeis.org`.

2 L. Addario-Berry, N. Broutin, and C. Holmgren. Cutting down trees with a Markov chainsaw. *Annals of Applied Probability*, 24(6):2297–2339, 2014.

3 N. Alon, I. Benjamini, and A. Stacey. Percolation on finite graphs and isoperimetric inequalities. *The Annals of Probability*, 32(3):1727–1745, 2004.

4 J. Bertoin. Fires on trees. In *Annales de l'IHP Probabilités et statistiques*, volume 48, 909–921, 2012.

5 G. Berzunza. The cut-tree of large trees with small heights. *Random Structures & Algorithms*, 51(3):404–427, 2017.

6 G. Berzunza, X. S. Cai, and C. Holmgren. The k-cut model in deterministic and random trees. *arXiv*, 2019. `arXiv:1907.02770`.

7 P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, 1st edition, 1968.

8 N. Broutin and M. Wang. Cutting down **p**-trees and inhomogeneous continuum random trees. *Bernoulli*, 23(4A):2380–2433, 2017.

9 F. Burghart. A modification of the random cutting model, 2022. `arXiv:2111.02968`.

10 X. S. Cai and C. Holmgren. Cutting resilient networks-complete binary trees. *The Electronic Journal of Combinatorics*, 26(4), 2019.

11 X. S. Cai, C. Holmgren, L. Devroye, and F. Skerman. k-cut on paths and some trees. *Electronic Journal of Probability*, 24, 2019.

12 V. Campos, V. Chvatal, L. Devroye, and P. Taslakian. Transversals in trees. *Journal of graph theory*, 73(1):32–43, 2013.

13 P. Chassaing and R. Marchand. Merging costs for the additive Marcus-Lushnikov process, and Union-Find algorithms, 2004. `arXiv:math/0406094`.

14 L. Devroye. A note on the probability of cutting a Galton-Watson tree. *Electronic Journal of Probability*, 16:2001–2019, 2011.

15 J. A. Fill, N. Kapur, and A. Panholzer. Destruction of very simple trees. *Algorithmica*, 46(3):345–366, 2006.

16 P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

17 G. R. Grimmett. *Percolation*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, 2nd edition, 1999.

18 C. Holmgren. Random records and cuttings in binary search trees. *Combinatorics, Probability and Computing*, 19(3):391–424, 2010.

19 C. Holmgren. A weakly 1-stable distribution for the number of random records and cuttings in split trees. *Advances in Applied Probability*, 43(1):151–177, 2011.

20 S. Janson. Random records and cuttings in complete binary trees. In *Mathematics and Computer Science III,* 241–253. Springer, 2004.

21 S. Janson. Random cutting and records in deterministic and random trees. *Random Structures & Algorithms*, 29(2):139–179, 2006.

22 D. E. Knuth and A. Schönhage. The expected linearity of a simple equivalence algorithm. *Theoretical Computer Science*, 6(3):281–315, 1978.

23 M. Kuba and A. Panholzer. Analysis of the total costs for variants of the Union-Find algorithm. In *Discrete Mathematics and Theoretical Computer Science*, pages 283–294. Discrete Mathematics and Theoretical Computer Science, 2007.

24 M. Kuba and A. Panholzer. Isolating a leaf in rooted trees via random cuttings. *Annals of Combinatorics*, 12(1):81–99, 2008.

25 M. Kuba and A. Panholzer. Isolating nodes in recursive trees. *Aequationes mathematicae*, 76(3):258–280, 2008.

26 M. Kuba and A. Panholzer. Multiple isolation of nodes in recursive trees. *Online Journal of Analytic Combinatorics*, 9(7), 2014.

27 R. Lyons and Y. Peres. *Probability on trees and networks.* Cambridge University Press, 2017.

28 A. Meir and J. W. Moon. Cutting down random trees. *Journal of the Australian Mathematical Society*, 11(3):313–324, 1970.

29 A. Panholzer. Non-crossing trees revisited: cutting down and spanning subtrees. In *Discrete Mathematics and Theoretical Computer Science,* 265–276, 2003.

30 A. Panholzer. Cutting down very simple trees. *Quaestiones Mathematicae*, 29(2):211–227, 2006.

31 A. C.-C. Yao. On the average behavior of set merging algorithms. In *Proceedings of the eighth annual ACM symposium on Theory of computing,* 192–195, 1976.

Enumeration of d-Combining Tree-Child Networks

Yu-Sheng Chang
Department of Mathematical Sciences, National Chengchi University, Taipei, 116, Taiwan

Michael Fuchs
Department of Mathematical Sciences, National Chengchi University, Taipei, 116, Taiwan

Hexuan Liu
Department of Mathematical Sciences, National Chengchi University, Taipei, 116, Taiwan

Michael Wallner
Institute of Discrete Mathematics and Geometry, TU Wien, Vienna, 1040, Austria

Guan-Ru Yu
Department of Mathematics, National Kaohsiung Normal University, Kaohsiung, 824, Taiwan

Abstract

Tree-child networks are one of the most prominent network classes for modeling evolutionary processes which contain reticulation events. Several recent studies have addressed counting questions for *bicombining tree-child networks* which are tree-child networks with every reticulation node having exactly two parents. In this paper, we extend these studies to *d-combining tree-child networks* where every reticulation node has now $d \geq 2$ parents. Moreover, we also give results and conjectures on the distributional behavior of the number of reticulation nodes of a network which is drawn uniformly at random from the set of all tree-child networks with the same number of leaves.

2012 ACM Subject Classification Mathematics of computing → Discrete mathematics

Keywords and phrases Phylogenetic network, tree-child network, d-combining tree-child network, exact enumeration, asymptotic enumeration, reticulation node, limit law, stretched exponential

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.5

Funding *Michael Fuchs*: Partially supported by MOST-109-2115-M-004-003-MY2.
Michael Wallner: Supported by the Austrian Science Fund (FWF): P 34142.
Guan-Ru Yu: Partially supported by MOST-110-2115-M-017-003-MY3.

1 Introduction and Results

The evolutionary process of, e.g., chromosomes, species, and populations is not always tree-like due to the occurrence of reticulation events caused by meiotic recombination on the chromosome level, specification and horizontal gene transfer on the species level, and sexual recombination on the population level. Because of this, *phylogenetic networks* have been introduced as appropriate models for reticulate evolution. Studying the properties of these networks is at the moment one of the most active areas of research in phylogenetics; see [10, 13].

While algorithmic and combinatorial aspects of phylogenetic networks have been investigated now for a few decades, enumerating and counting phylogenetic networks as well as understanding their "typical shape" are relatively recent areas of research; see [13, page 253] where such questions are only discussed in one short paragraph. However, the last couple of years have seen a lot of progress on these questions, in particular for the class of *tree-child networks*, which is one of the most prominent subclasses amongst the many subclasses of phylogenetic networks; see [1, 5–9, 11, 12].

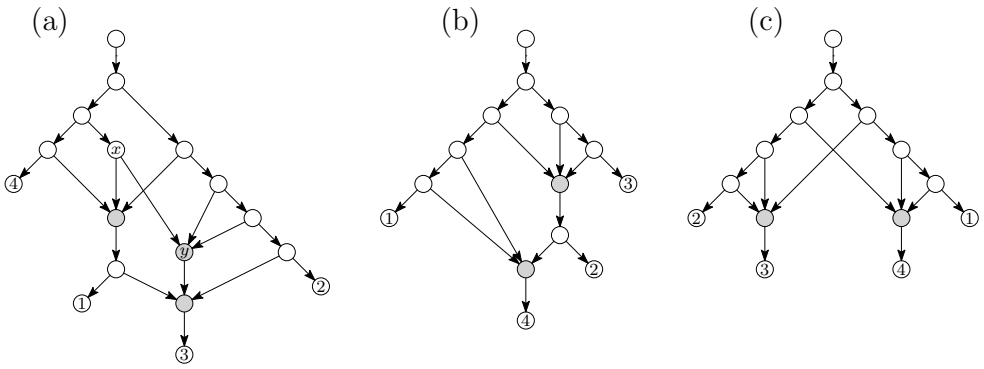


Figure 1 (a) A 3-combining phylogenetic network which is *not* a tree-child network (because both children of the tree node x are reticulation nodes and the only child of the reticulation node y is also a reticulation node); (b) a 3-combining tree-child network; (c) a 3-combining one-component tree-child network.

Most of the studies on tree-child networks have focused on *bicombining tree-child networks* which are tree-child networks where every reticulation event involves exactly two individuals. The purpose of this extended abstract is to discuss extensions of previous results to *multicombining tree-child* networks. More precisely, we will focus on *d-combining tree-child networks* which are tree-child networks whose reticulation events involve exactly $d \geq 2$ individuals. We will highlight similarities and differences between the two cases $d = 2$ and $d > 2$.

Before explaining our results, we will first give precise definitions and fix some notation. We start with the definition of phylogenetic networks.

▶ **Definition 1** (Phylogenetic network). *A (rooted) phylogenetic network with n leaves is a rooted, simple, directed acyclic graph (DAG) with no nodes of in- and outdegree 1 and exactly n nodes of indegree 1 and outdegree 0 (i.e., leaves) which are bijectively labeled with labels from the set $\{1, \ldots, n\}$.*

In this work, we will only consider phylogenetic networks whose nodes all have either indegree 1 *or* outdegree 1 and whose internal nodes (i.e., nodes that are neither leaves nor the root) with indegree 1 all have outdegree 2 (bifurcating case); the latter nodes will be called *tree nodes*. Finally, we will assume that all internal nodes with outdegree 1 have indegree $d \geq 2$ and these nodes will be called *reticulation nodes*; see Figure 1 for examples with $d = 3$. Note that $d = 2$ is the above mentioned bicombining case.

We next recall the definition of tree-child networks.

▶ **Definition 2** (Tree-child network). *A phylogenetic network is called a tree-child network if every non-leaf node has at least one child which is not a reticulation node.*

In other words, a phylogenetic network is a tree-child network if (a) the root is not followed by a reticulation node; (b) a reticulation node is not followed by another reticulation node; and (c) a tree node has at least one child which is not a reticulation node; see Figure 1, (b) for an example. A simple and important subclass of tree-child networks is the class of one-component tree-child networks; see the definition below and Figure 1, (c) for an example.

▶ **Definition 3** (One-component tree-child network). *A tree-child network is called a one-component tree-child network if every reticulation node is directly followed by a leaf.*

One-component networks are more "tree-like" than general tree-child networks. Moreover, they constitute an important building block in the construction of general tree-child networks; see [1] for the bicombining case.

In the following, we denote by $\mathrm{OTC}_{n,k}^{(d)}$ and $\mathrm{TC}_{n,k}^{(d)}$ the number of one-component and general d-combining tree-child networks with n leaves and k reticulation nodes, respectively. Note that the tree-child property implies that $k \le n-1$. Thus, the total numbers of one-component and general d-combining tree-child networks, denoted by $\mathrm{OTC}_n^{(d)}$ and $\mathrm{TC}_n^{(d)}$, satisfy

$$\mathrm{OTC}_n^{(d)} := \sum_{k=0}^{n-1} \mathrm{OTC}_{n,k}^{(d)} \qquad \text{and} \qquad \mathrm{TC}_n^{(d)} := \sum_{k=0}^{n-1} \mathrm{TC}_{n,k}^{(d)}.$$

Now, we are ready to present our results.

First, for one-component tree-child networks, we have the following formula which extends the one for $d = 2$ from Theorem 13 in [1].

▶ **Theorem 4.** *The numbers of one-component d-combining tree-child networks with n leaves and k reticulation nodes for $0 \le k \le n-1$ are given by*

$$\mathrm{OTC}_{n,k}^{(d)} = \binom{n}{k} \frac{(2n+(d-2)k-2)!}{(d!)^k\, 2^{n-k-1}\, (n-k-1)!}$$

and equal to 0 otherwise.

Using this formula, we obtain the following limit law result, where (here and throughout the work) we use $\xrightarrow{w}$ to denote weak convergence.

▶ **Corollary 5.** *Let $R_n^{(d)}$ be the number of reticulation nodes of a one-component d-combining tree-child network picked uniformly at random from the set of all one-component d-combining tree-child networks with n leaves. Then, we have the following limit behavior of $R_n^{(d)}$.*

(i) *For $d = 2$ (bicombining case), we have*

$$\frac{R_n^{(2)} - n + \sqrt{n}}{\sqrt[4]{n/4}} \xrightarrow{w} N(0,1).$$

(ii) *For $d = 3$, we have*

$$n - 1 - R_n^{(3)} \xrightarrow{w} \mathrm{Bessel}(1,2),$$

where Bessel(v,a) denotes the Bessel distribution, i.e.,

$$\mathbb{P}(\mathrm{Bessel}(1,2) = k) = \frac{1}{I_1(2)k!(k+1)!}, \qquad (k \ge 0).$$

Here, $I_v(a) = \left(\frac{a}{2}\right)^v \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+v+1)} \frac{a^{2k}}{4^k}$ is the modified Bessel function of the first kind.

(iii) *For $d \ge 4$, the limit law of $n - 1 - R_n^{(d)}$ is degenerate, i.e.,*

$$n - 1 - R_n^{(d)} \xrightarrow{w} 0.$$

▶ Remark 6. If t denotes the number of tree nodes and N the total number of nodes, then by the handshaking lemma, we have

$$t = n + (d-1)k - 1 \qquad \text{and} \qquad N = 2n + dk. \tag{1}$$

Therefore, we have similar limit distribution results for these numbers as well.

Note that the above result for $d = 2$ is already contained in the proof of Theorem 3 in [9] where even a local limit theorem was proved; see also [8]. Using the above corollary, we now obtain the first order asymptotics of the total number of one-component tree-child networks.

▶ **Corollary 7.**

(i) *For $d = 2$ (bicombining case), we have*

$$\mathrm{OTC}_n^{(2)} \sim \frac{1}{4\pi\sqrt{e}}(n!)^2 2^n e^{2\sqrt{n}} n^{-9/4}.$$

(ii) *For $d = 3$, we have*

$$\mathrm{OTC}_n^{(3)} \sim I_1(2) \cdot \mathrm{OTC}_{n,n-1}^{(3)} \sim \frac{I_1(2)\sqrt{3}}{9\pi}(n!)^3 \left(\frac{9}{2}\right)^n n^{-3},$$

where $I_v(\alpha)$ is as in Corollary 5, (ii).

(iii) *For $d \geq 4$, we have*

$$\mathrm{OTC}_n^{(d)} \sim \mathrm{OTC}_{n,n-1}^{(d)} \sim \frac{d!}{d^{d-1/2}(2\pi)^{(d-1)/2}}(n!)^d \left(\frac{d^d}{d!}\right)^n n^{3(1-d)/2}.$$

Again, the result for the case $d = 2$ is already contained in [9]; this is also the only case of the three above in which we find a stretched exponential in the asymptotics (see [3]).

We next turn to general tree-child networks. Here, in contrast to one-component tree-child networks, we only understand the behavior of $\mathrm{TC}_{n,k}^{(d)}$ for fixed k and for $k = n - 1$ (i.e., for maximally reticulated networks).

First, for fixed k, the first order asymptotics in the bicombining case ($d = 2$) was derived in [5, 6] and with a different method in [7]; see also [12] where the result was re-derived with yet another method which is however based on a (yet) unproven conjecture. The approach from [7] can also be used in the d-combining case leading to the following result.

▶ **Theorem 8.** *For the number of d-combining tree-child networks with n leaves and k reticulation nodes, we have for fixed k, as $n \to \infty$,*

$$\mathrm{TC}_{n,k}^{(d)} \sim \frac{2^{(4-d)k-1}}{d!^k\, k!\sqrt{\pi}} n! 2^n n^{(4-d)k-3/2}.$$

▶ Remark 9. Our approach can also be used to compute tables of $\mathrm{TC}_{n,k}^{(d)}$ for small values of n, k, and d; see the Appendix. Moreover, the approach is also capable of giving exact formulas for $\mathrm{TC}_{n,k}^{(d)}$ and small values of k; see [1, 6] for such formulas in the bicombining case.

Note that the asymptotic order in the above theorem is much smaller than the one obtained for one-component tree-child networks. Thus, the majority of tree-child networks do not have a bounded number of reticulation nodes. In fact, the number of reticulation nodes of a "typical" tree-child network is close to the maximum $n - 1$. More precisely, the following result holds.

▶ **Theorem 10.** *For the number of d-combining tree-child networks with n leaves, we have*

$$\mathrm{TC}_n^{(d)} = \Theta\left(\mathrm{TC}_{n,n-1}^{(d)}\right) = \Theta\left((n!)^d\, \gamma(d)^n\, e^{3a_1\beta(d)n^{1/3}} n^{\alpha(d)}\right),$$

where $a_1 = -2.33810741\ldots$ is the largest root of the Airy function of the first kind and

$$\alpha(d) = -\frac{d(3d-1)}{2(d+1)}, \qquad \beta(d) = \left(\frac{d-1}{d+1}\right)^{2/3}, \qquad \gamma(d) = 4\frac{(d+1)^{d-1}}{(d-1)!}.$$

For the bicombining case, this result was proved in [9] by encoding tree-child networks with n leaves and $n-1$ reticulation nodes by certain words and (asymptotically) counting these words with the method from [3]. In the more general d-combining case, we will use a similar strategy, however, details will be more demanding due to the dependence on the parameter d.

As for the stochastic behavior of the number of reticulation nodes for general tree-child networks, we have a conjecture for the limit laws which we are going to present in Section 4. Note that even for $d = 2$, no limit law for any shape parameter of random tree-child networks has been established yet.

Our conjecture will clarify the behavior of $\mathrm{TC}_{n,k}^{(d)}$ for k close to n. Thus, the behavior of $\mathrm{TC}_{n,k}^{(d)}$ for small and large k is clear. For the remaining range, there is an interesting recent conjecture for the bicombining case in [12], which has been proved for the special case of one-component tree-child networks in [8]. Whether this conjecture can be extended to d-combining tree-child networks is not clear yet; see the comments in Section 4.

We conclude the introduction with an outline of this extended abstract. In the next section, we will consider one-component networks and prove Theorem 4 and Corollaries 5 and 7. In Section 3, we will discuss our results for general networks. Finally, Section 4 will contain the above mentioned conjecture for the limit laws of the number of reticulation nodes of a random d-combining tree-child network and some concluding remarks.

2 One-Component Networks

In this section, we will prove our results for one-component tree-child networks. We start with Theorem 4.

Proof of Theorem 4. Suppose N is a one-component d-combining tree-child network with $n-1$ leaves and $k-1$ reticulation nodes.

Then, we can construct one-component d-combining tree-child networks with n leaves and k reticulation nodes from N by the following three steps: (i) put d new nodes into the *candidate edges* where we call an edge of N a candidate edge if it is not incident to any reticulation node; (ii) create a new reticulation node which is adjacent to the d new nodes; and (iii) add a new leaf as a child of this reticulation node; moreover, label it with a label from $\{1,\dots,n\}$ and increase all (old) labels in N which are at least as large as the new label by $+1$ (if there are any).

Now, note that in step (i), we have

$$\underbrace{n-1+(d-1)(k-1)-1}_{\substack{\#\text{ edges leading to}\\ \text{a tree node; see (1)}}}+\underbrace{n-1}_{\substack{\#\text{ edges leading}\\ \text{to a leaf}}}-\underbrace{(k-1)}_{\substack{\#\text{ edges below}\\ \text{ret. nodes}}}=2n+(d-2)(k-1)-3$$

candidate edges and thus there are

$$\binom{2n+(d-2)k-2}{d}$$

choices of the d nodes. Moreover, in step (iii), there are n choices of the label. Finally, note that the above construction gives each network exactly k times.

Overall, the above arguments give

$$\mathrm{OTC}^{(d)}_{n,k} = \frac{n}{k}\binom{2n+(d-2)k-2}{d}\mathrm{OTC}^{(d)}_{n-1,k-1},$$

and by iteration,

$$\mathrm{OTC}^{(d)}_{n,k} = \binom{n}{k}\frac{(2n+(d-2)k-2)!}{d!^k(2n-k-2)!}\mathrm{OTC}^{(d)}_{n-k,0}.$$

The result follows now by the fact that

$$\mathrm{OTC}^{(d)}_{n-k,0} = (2(n-k)-3)!! = \frac{(2n-2k-2)!}{2^{n-k-1}(n-k-1)!},$$

since $\mathrm{OTC}^{(d)}_{n-k,0}$ is the number of phylogenetic trees with $n-k$ leaves; see, e.g., [13, Section 2.1]. ◀

Now, we can prove the two corollaries from above.

Proof of Corollaries 5 and 7. Since the results for $d = 2$ are already contained in [9] (see also [8]), we can focus on the cases $d \geq 3$.

We start with the case $d = 3$. Note that

$$\mathrm{OTC}^{(3)}_{n,k} = \binom{n}{k}\frac{(2n+k-2)!}{3^k 2^{n-1}(n-k-1)!}, \qquad (0 \leq k \leq n-1)$$

and this sequence is increasing in k. (This is in contrast to $d = 2$ where this sequence has a maximum at $k = n - \sqrt{n+1}$; see [9].) By replacing k by $n-1-k$ and using Stirling's formula, we obtain that

$$\mathrm{OTC}^{(3)}_{n,n-1-k} = \frac{1}{k!(k+1)!}\cdot\frac{n(3n-3)!}{6^{n-1}}\left(1+\mathcal{O}\left(\frac{1+k^2}{n}\right)\right) \tag{2}$$

uniformly for k with $k = o(\sqrt{n})$. Thus, by a standard application of the Laplace method:

$$\mathrm{OTC}^{(3)}_{n} \sim \left(\sum_{k\geq 0}\frac{1}{k!(k+1)!}\right)\cdot\frac{n(3n-3)!}{6^{n-1}} = I_1(2)\cdot\frac{n(3n-3)!}{6^{n-1}}$$

which is the first claim from Corollary 7, (ii); the second follows from this by another application of Stirling's formula. Moreover, since

$$\mathbb{P}(R^{(3)}_n = n-1-k) = \frac{\mathrm{OTC}^{(3)}_{n,n-1-k}}{\mathrm{OTC}^{(3)}_{n}},$$

the result from Corollary 5, (ii) follows from the above two expansions too.

Next, we consider the case $d \geq 4$. The details of the proof are the same as above, with the main difference that the expansion (2) now becomes

$$\mathrm{OTC}^{(d)}_{n,n-1-k} = \left(\frac{d^2 d!}{2d^d}\right)^k\frac{1}{k!(k+1)!}\cdot n^{(3-d)k}\cdot\frac{n(dn-d)!}{d!^{n-1}}\left(1+\mathcal{O}\left(\frac{1+k^2}{n}\right)\right)$$

uniformly for k with $k = o(\sqrt{n})$. This expansion, for $d \geq 4$, contains the (non-trivial decreasing) factor $n^{(3-d)k}$ which is responsible for $\mathrm{OTC}^{(d)}_n$ being now asymptotically dominated by $\mathrm{OTC}^{(d)}_{n,n-1}$ (proving Corollary 7, (iii)) and the limiting distribution of $n-1-R^{(d)}_n$ being degenerate (proving Corollary 5, (iii)). ◀

3 General Networks

In this section, we will discuss the asymptotic enumeration of general d-combining tree-child networks with a fixed (Theorem 8) and a maximal number of reticulation nodes (Theorem 10). Note that the latter dominates asymptotically all networks of a given size.

We start with Theorem 8 on d-combining networks with a fixed number of reticulation nodes. It can be proved by generalizing the approach from [7], which was based on the classification of tree-child networks via *component graphs* from [1]. Component graphs can also be defined for d-combining tree-child networks and then be used to prove Theorem 8; details will be given in the journal version of this work. Moreover, component graphs can also be used as in [1] to (a) compute $\mathrm{TC}_{n,k}^{(d)}$ for small values of n, k, and d (see the Appendix); and (b) give explicit formulas for small values of k.

The remainder of this section is devoted to the proof of Theorem 10, which extends the approaches from [3, 9]. We start with some lemmas which are generalizations of the corresponding results from [9] (and are proved with similar arguments).

The first lemma shows that the asymptotic growth of $\mathrm{TC}_n^{(d)}$ is, up to a constant, determined by the asymptotics of $\mathrm{TC}_{n,n-1}^{(d)}$.

▶ **Lemma 11.** *For $n \to \infty$, we have*

$$\mathrm{TC}_n^{(d)} = \Theta\left(\mathrm{TC}_{n,n-1}^{(d)}\right).$$

Proof. Let N be a d-combining tree-child network with n leaves and k reticulation nodes. A *free tree node* is a tree node whose children are both not reticulation nodes; the edges to these children are called *free edges*. Using a simple counting argument, it is easy to see that N has $2(n-k-1)$ free edges; see [9, Lemma 1] for the case $d = 2$.

Next, we can construct d-combining tree-child networks with n leaves and $k+1$ reticulation nodes by (i) inserting d tree nodes into the root edge of N and a reticulation node into a free edge and (ii) connecting the d new tree nodes to the new reticulation node. Note that each network built in this way is different. Thus,

$$2(n-k-1)\mathrm{TC}_{n,k}^{(d)} \le \mathrm{TC}_{n,k+1}^{(d)}.$$

Iterating this construction yields

$$\mathrm{TC}_{n,k}^{(d)} \le \frac{1}{2^{n-k-1}(n-k-1)!}\mathrm{TC}_{n,n-1}^{(d)} \tag{3}$$

and thus,

$$\mathrm{TC}_{n,n-1}^{(d)} \le \mathrm{TC}_n^{(d)} \le \left(\sum_{j\ge 0}\frac{1}{2^j j!}\right)\cdot \mathrm{TC}_{n,n-1}^{(d)} = \sqrt{e}\cdot \mathrm{TC}_{n,n-1}^{(d)},$$

which proves the claim. ◀

Next, we define a generalization of the class of words from [9] which is used to encode d-combining tree-child networks with a maximal number of reticulation nodes.

▶ **Definition 12.** *Let $\mathcal{C}_n^{(d)}$ denote the class of words built from n letters $\{\omega_1, \ldots, \omega_n\}$ in which each letter occurs exactly $d+1$ times such that in every prefix the letter ω_i has either not yet occurred more than $d-2$ times, or, if it has, then the number of occurrences of ω_i is at least as large as the number of occurrences of ω_j for all $j > i$.*

In [9], a bijection between bicombining tree-child networks with n leaves and k reticulationd nodes whose labels are removed and $\mathcal{C}_{n-1}^{(2)}$ was proved. In fact, this bijection can be extended to d-combining networks. Then, because the networks are all different when labeling the (now empty) leaves by a random permutation, we obtain the following lemma.

▶ **Lemma 13.** *Let $c_n^{(d)}$ be the cardinality of $\mathcal{C}_n^{(d)}$. Then*

$$\mathrm{TC}_{n,n-1}^{(d)} = n! c_{n-1}^{(d)}.$$

Now the recursive nature of this encoding allows us to derive the following counting result.

▶ **Lemma 14.** *We have*

$$c_n^{(d)} = \sum_{m\geq 1} b_{n,m}^{(d)}, \tag{4}$$

where $b_{n,m}^{(d)}$ satisfies the recurrence

$$b_{n,m}^{(d)} = \frac{dn+m-2}{dn+m-d-1} b_{n,m-1}^{(d)} + \binom{dn+m-2}{d-1} b_{n-1,m}^{(d)}, \qquad (n\geq 2, 0\leq m\leq n) \tag{5}$$

with $b_{1,1}^{(d)} = 1$ and $b_{n,m}^{(d)} = 0$ for (i) $n \geq 2$ and $m = -1$; (ii) $n = 1$ and $m = 0$; and (iii) $n < m$.

Proof. First, note that any word in $\mathcal{C}_n^{(d)}$ has a suffix $\omega_n\omega_m\omega_{m+1}\cdots\omega_{n-1}\omega_n$ with $1 \leq m \leq n$. Denote by $b_{n,m}^{(d)}$ the number of these words. Removing the d occurrences of ω_n from these words gives a word of $\mathcal{C}_{n-1}^{(d)}$ with suffix $\omega_m\omega_{m+1}\cdots\omega_{n-1}$, i.e., it has a suffix $\omega_{n-1}\omega_j\omega_{j+1}\ldots\omega_{n-1}$ for $j = 1,\ldots,m$. Reversing this procedure gives

$$b_{n,m}^{(d)} = \binom{dn+m-2}{d-1} \sum_{j=1}^{m} b_{n-1,j}^{(d)}, \tag{6}$$

where the binomial coefficient counts the number of ways of adding back the $d-1$ occurrences of ω_n after two ω_n's have been added, one before the last ω_m and one at the end of the word. By Definition 12 these first $d-1$ occurrences of ω_n may be anywhere. Differencing yields

$$\frac{b_{n,m}^{(d)}}{\binom{dn+m-2}{d-1}} - \frac{b_{n,m-1}^{(d)}}{\binom{dn+m-3}{d-1}} = b_{n-1,m}^{(d)}.$$

This gives the claimed recurrence and the initial conditions are easily checked. ◀

The advantage of the recurrence for $b_{n,m}^{(d)}$ is that we are actually only interested in the asymptotics of $b_{n,n}^{(d)}$ as by (4) and (6) we have

$$b_{n,n}^{(d)} = \binom{(d+1)n-2}{d-1} c_{n-1}^{(d)}.$$

Now we are ready to use the method from [3]. Due to the similarities, we will only discuss the main differences. We start with the following transformation of $(b_{n,m})_{0\leq m\leq n}$ to $(e_{i,j}^{(d)})_{\substack{0\leq i\leq j \\ i-j \text{ even}}}$, which changes the indices and captures the exponential and superexponential terms coming from the binomial coefficient in (5).

▶ **Lemma 15.** *We have*

$$b_{n,m}^{(d)} = \lambda(d)^n (n!)^{d-1} e_{n+m,n-m}^{(d)} \qquad \text{with} \qquad \lambda(d) = \frac{(d+1)^{d-1}}{(d-1)!},$$

where $e_{n,m}^{(d)}$ *satisfies the following recurrence*

$$e_{n,m}^{(d)} = \mu_{n,m}^{(d)} e_{n-1,m+1}^{(d)} + \nu_{n,m}^{(d)} e_{n-1,m-1}^{(d)} \tag{7}$$

with

$$\mu_{n,m}^{(d)} = 1 + \frac{2(d-1)}{(d+1)n + (d-1)m - 2(d+1)} \quad \text{and} \quad \nu_{n,m}^{(d)} = \prod_{i=2}^{d} \left(1 - \frac{2(m+i)}{(d+1)(n+m)}\right)$$

for $n \geq 3$ *and* $m \geq 0$*, where* $e_{n,-1}^{(d)} = e_{2,n}^{(d)} = 0$ *except for* $e_{2,0}^{(d)} = 1/\lambda(d)$*.*

Note that we are interested in $e_{2n,0}^{(d)} = \frac{b_{n,n}^{(d)}}{\lambda(d)^n (n!)^{d-1}}$ because by the previous lemmas we have

$$\mathrm{TC}_n^{(d)} = \Theta\left(\mathrm{TC}_{n,n-1}^{(d)}\right) = \Theta\left(n! c_{n-1}^{(d)}\right) = \Theta\left(n!\, n^{1-d} b_{n,n}^{(d)}\right) = \Theta\left((n!)^d \lambda(d)^n n^{1-d} e_{2n,0}^{(d)}\right). \tag{8}$$

Moreover, observe that for the Theta-result the initial value of $e_{2,0}^{(d)}$ is irrelevant, as it creates only a constant factor. So we may set it to $e_{2,0}^{(d)} = 1$, or any convenient constant. Note that this recurrence is very similar to that of relaxed trees [3, Equation (2)], yet with more complicated factors. Observe also that this is exactly recurrence [9, Equation (10)] for $d = 2$.

Motivated by experiments for large n, we use the following ansatz

$$e_{n,m}^{(d)} \approx h(n) f\left(\frac{m+1}{n^{1/3}}\right),$$

where h and f are some "regular" functions. Next, we substitute $s(n) = h(n)/h(n-1)$ and $m = \kappa n^{1/3} - 1$ into (7). Then, for $n \to \infty$ we get the expansion

$$f(\kappa)s(n) = 2f(\kappa) + \left(f''(\kappa) - \frac{2(d-1)}{d+1}\kappa f(\kappa)\right) n^{-2/3} + \mathcal{O}\left(n^{-1}\right).$$

Hence, we may assume that $s(n) = 2 + c_1 n^{-2/3} + c_2 n^{-1} + \dots$ and this implies that $f(\kappa)$ satisfies the differential equation

$$f''(\kappa) = \left(c_1 + \frac{2(d-1)}{d+1}\kappa\right) f(\kappa)$$

that is solved by the Airy function Ai of the first kind. Additionally, the boundary conditions allow to compute c_1 and we get that

$$f(\kappa) = C\mathrm{Ai}\left(a_1 + B^{1/3}\kappa\right) \qquad \text{where} \qquad B := \frac{2(d-1)}{d+1},$$

$a_1 \approx 2.338$ is the largest root of the Airy function Ai, and C is an arbitrary constant. From this we get that $c_1 = a_1 B^{1/3}$. These heuristic arguments guide us to the following results. The proofs are analogous to [3, 4, 9]; for the details we refer to the accompanying Maple worksheet [2].

▶ **Proposition 16.** *For all $n, m \geq 0$ let*

$$\tilde{X}_{n,m} := \left(1 - \frac{2d-1}{3(d+1)}\frac{m^2}{n} - \frac{3d^2+12-11}{6(d+1)}\frac{m}{n}\right) \mathrm{Ai}\left(a_1 + \frac{B^{1/3}(m+1)}{n^{1/3}}\right) \quad \text{and}$$
$$\tilde{s}_n := 2 + \frac{a_1 B^{2/3}}{n^{2/3}} - \frac{3d^2-5d+4}{3(d+1)n} - \frac{1}{n^{7/6}}.$$

Then, for any $\varepsilon > 0$, there exists an $\tilde{n}_0$ such that

$$\tilde{X}_{n,m}\tilde{s}_n \leq \mu^{(d)}_{n,m}\tilde{X}_{n-1,m+1} + \nu^{(d)}_{n,m}\tilde{X}_{n-1,m-1}$$

for all $n \geq \tilde{n}_0$ and for all $0 \leq m < n^{2/3-\varepsilon}$, where $\mu^{(d)}_{n,m}$ and $\nu^{(d)}_{n,m}$ are as in Lemma 15.

▶ **Proposition 17.** *Choose $\eta > \frac{(2d-1)^2}{18(d+1)^2}$ fixed and for all $n, m \geq 0$ let*

$$\hat{X}_{n,m} := \left(1 - \frac{2d-1}{3(d+1)}\frac{m^2}{n} - \frac{3d^2+12-11}{6(d+1)}\frac{m}{n} + \eta\frac{m^4}{n^2}\right) \mathrm{Ai}\left(a_1 + \frac{B^{1/3}(m+1)}{n^{1/3}}\right) \quad \text{and}$$
$$\hat{s}_n := 2 + \frac{a_1 B^{2/3}}{n^{2/3}} - \frac{3d^2-5d+4}{3(d+1)n} + \frac{1}{n^{7/6}}.$$

Then, for any $\varepsilon > 0$, there exists a constant $\hat{n}_0$ such that

$$\hat{X}_{n,m}\hat{s}_n \geq \mu^{(d)}_{n,m}\hat{X}_{n-1,m+1} + \nu^{(d)}_{n,m}\hat{X}_{n-1,m-1}$$

for all $n \geq \hat{n}_0$ and all $0 \leq m < n^{1-\varepsilon}$.

Proof of Theorem 10. Let us start with the lower bound. We first define a sequence $X_{n,m} := \max\{\tilde{X}_{n,m}, 0\}$ which satisfies the inequality of Proposition 16 for *all* $m \leq n$. Then, we define an explicit sequence $\tilde{h}_n := \tilde{s}_n\tilde{h}_{n-1}$ for $n > 0$ and $\tilde{h}_0 = \tilde{s}_0$. From this, we show by induction that $e^{(d)}_{n,m} \geq C_0\tilde{h}_n X_{n,m}$ for some constant $C_0 > 0$ and all $n \geq \tilde{n}_0$ and all $0 \leq m \leq n$. Hence,

$$\begin{aligned} e^{(d)}_{2n,0} &\geq C_0\tilde{h}_{2n}X_{2n,0} \\ &\geq C_0\prod_{i=1}^{2n}\left(2 + \frac{a_1B^{2/3}}{i^{2/3}} - \frac{3d^2-5d+4}{3(d+1)i} - \frac{1}{i^{7/6}}\right)\mathrm{Ai}\left(a_1 + \frac{B^{1/3}}{n^{1/3}}\right) \\ &\geq C_1(n!)^{d-1}4^n e^{3a_1B^{2/3}n^{1/3}} n^{\frac{d^2+d-2}{2(d+1)}}. \end{aligned}$$

Finally, combining this with (8) we get the lower bound.

The upper bound is similar, yet more technical. The starting point is Proposition 17 and a function $X_{n,m}$ that is valid for *all* $0 \leq m \leq n$. For this purpose we define a sequence $\hat{e}^{(d)}_{n,m}$ such that $\hat{e}^{(d)}_{n,m} := e^{(d)}_{n,m}$ for $0 \leq m \leq n^{1-\varepsilon}$ and $\hat{e}^{(d)}_{n,m} := 0$ otherwise; compare with [3,4]. Then, using tools from lattice path theory and computer algebra, we show that $e^{(d)}_{2n,0} = \mathcal{O}(\hat{e}^{(d)}_{2n,0})$ and that

$$\hat{e}^{(d)}_{2n,0} \leq \hat{C}_1(n!)^{d-1}4^n e^{3a_1B^{2/3}n^{1/3}} n^{\frac{d^2+d-2}{2(d+1)}}.$$

For more details, see the journal version of this work. ◀

4 Conjectures and Remarks

The main purpose of this paper was to extend recent results on bicombining tree-child networks to d-combining tree-child networks. We did this for both one-component tree-child networks as well as general tree-child networks. For one-component d-combining tree-child networks, we proved an exact counting formula for their number with n leaves and k reticulation nodes. As a consequence of this formula, we obtained limit laws for the number of reticulation nodes of a random network and (asymptotic) counting results for their total number. For general d-combining tree-child networks, our knowledge of their counts is less complete. We derived (asymptotic) results for a fixed number of reticulation nodes and the maximal number of reticulation nodes. The latter implied also an (asymptotic) counting result for their total number.

How about limit laws for the number of reticulation nodes of general d-combining tree-child networks? We think that the upper bound (3) is sharp for $d = 2$ and far away from being sharp for $d \geq 3$. More precisely, we believe that the following conjecture holds.

▶ **Conjecture 18.** *Let $T_n^{(d)}$ be the number of reticulation nodes of a d-combining tree-child network picked uniformly at random from the set of all d-combining tree-child networks with n leaves. Then, we have the following limit behavior of $T_n^{(d)}$.*

(i) *For $d = 2$ (bicombining case), we have the weak convergence result:*

$$n - 1 - T_n^{(2)} \xrightarrow{w} \mathrm{Poisson}(1/2),$$

where $\mathrm{Poisson}(\alpha)$ *denotes the Poisson distribution.*

(ii) *For $d \geq 3$, the limit distribution of $n - 1 - T_n^{(d)}$ is degenerate.*

Moreover, the proof of this conjecture should also give the following result.

▶ **Corollary 19.**

(i) *For $d = 2$ (bicombining case), we have $\mathrm{TC}_n^{(2)} \sim \sqrt{e} \cdot \mathrm{TC}_{n,n-1}^{(2)}$.*

(ii) *For $d \geq 3$, we have $\mathrm{TC}_n^{(d)} \sim \mathrm{TC}_{n,n-1}^{(d)}$.*

▶ Remark 20. Note that even with the above result, it is still not possible to give the first-order asymptotics of $\mathrm{TC}_n^{(d)}$ since the approach of [3] is only capable of giving a Theta-result.

In fact, we recently found a method which should allow us to prove these results; details are currently checked. The proofs (if correct) will be presented in the journal version of this paper.

The above limit distribution result would clarify the behavior of the number of general d-combining tree-child networks for a large number, i.e., a number close to n, of reticulation nodes. So, how about the remaining range? (Recall that the number of networks for a small number of reticulation nodes is covered by Theorem 8.)

In this regard, there is a recent interesting conjecture for the bicombining case; see [12]. To give details, denote by $\mathcal{C}_{n,k}^{(2)}$ a class of words which is similar defined as in Definition 12 but with the difference that only k letters appear 3 times while the remaining $n - k$ letters appear 2 times. Let $c_{n,k}^{(2)}$ be their number. Then, it was conjectured in [12], together with some striking consequences, that

$$\mathrm{TC}_{n,k}^{(2)} = \frac{n!}{(n-k)!} c_{n,k}^{(2)}.$$

Can this be extended to d-combining networks? (The obvious generalization by replacing 2 by d does not work.) We do not know the answer to this question yet. However, we have recently managed to define a modification of $\mathcal{C}_{n,k}^{(d)}$ which can be used to encode d-combining tree-child networks. This encoding seems to be useful for the proof of Conjecture 18 and might also shed further light on [12]. Details will be again discussed in the journal version.

Finally, how about extension of our results to multicombining tree-child networks, i.e., tree-child networks where different reticulation nodes may have different number of parents? We think that many of the results of this extended abstract can be generalized to this case, however, notation becomes more cumbersome. We might also include a discussion on this in the journal version of the current paper.

References

1 Gabriel Cardona and Louxin Zhang. Counting and enumerating tree-child networks and their subclasses. *J. Comput. System Sci.*, 114:84–104, 2020.

2 Yu-Sheng Chang, Michael Fuchs, Hexuan Liu, Michael Wallner, and Guan-Ru Yu. Maple worksheet, 2021. `arXiv:2203.07619`.

3 Andrew Elvey Price, Wenjie Fang, and Michael Wallner. Asymptotics of minimal deterministic finite automata recognizing a finite binary language. In *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 159:Art. No. 11, 13 pages of *LIPIcs. Leibniz Int. Proc. Inform.* Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2020.

4 Andrew Elvey Price, Wenjie Fang, and Michael Wallner. Compacted binary trees admit a stretched exponential. *J. Combin. Theory Ser. A*, 177:Paper No. 105306, 40 pages, 2021.

5 Michael Fuchs, Bernhard Gittenberger, and Marefatollah Mansouri. Counting phylogenetic networks with few reticulation vertices: tree-child and normal networks. *Australas. J. Combin.*, 73:385–423, 2019.

6 Michael Fuchs, Bernhard Gittenberger, and Marefatollah Mansouri. Counting phylogenetic networks with few reticulation vertices: exact enumeration and corrections. *Australas. J. Combin.*, 81:257–282, 2021.

7 Michael Fuchs, En-Yu Huang, and Guan-Ru Yu. Counting phylogenetic networks with few reticulation vertices: a second approach. `arXiv:2104.07842`.

8 Michael Fuchs, Hexuan Liu, and Guan-Ru Yu. A short note on the exact counting of tree-child networks. arXiv:2110.03842. `arXiv:2110.03842`.

9 Michael Fuchs, Guan-Ru Yu, and Louxin Zhang. On the asymptotic growth of the number of tree-child networks. *European J. Combin.*, 93:Paper No. 103278, 2021.

10 Daniel H. Huson, Regula Rupp, and Celine Scornavacca. *Phylogenetic Networks: Concepts, Algorithms and Applications*. Cambridge University Press, 1st edition, 2010.

11 Colin McDiarmid, Charles Semple, and Dominic Welsh. Counting phylogenetic networks. *Ann. Comb.*, 19(1):205–224, 2015.

12 Miquel Pons and Josep Batle. Combinatorial characterization of a certain class of words and a conjectured connection with general subclasses of phylogenetic tree-child networks. *Sci. Rep.*, 11:Article number: 21875, 2021.

13 Mike Steel. *Phylogeny – discrete and random processes in evolution*, volume 89 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2016.

A Appendix

Table 1 $\mathrm{TC}^{(2)}_{n,k}$ for $2 \le n \le 8$ and $0 \le k < n$; see also [1].

$n\backslash k$	0	1	2	3	4	5	6	7
2	1	2						
3	3	21	42					
4	15	228	1272	2544				
5	105	2805	30300	154500	309000			
6	945	39330	696600	6494400	31534200	63068400		
7	10395	623385	16418430	241204950	2068516800	9737380800	19474761600	
8	135135	11055240	405755280	8609378400	113376463200	920900131200	4242782275200	8485564550400

Table 2 $\mathrm{TC}^{(3)}_{n,k}$ for $2 \le n \le 7$ and $0 \le k < n$.

$n\backslash k$	0	1	2	3	4	5	6
2	1	2					
3	3	33	150				
4	15	492	7908	55320			
5	105	7725	291420	6179940	57939000		
6	945	132030	9603270	430105320	11292075000	132120450000	
7	10395	2471805	307525050	24586633890	1284266876760	40079165452200	560319972030000

Table 3 $\mathrm{TC}^{(4)}_{n,k}$ for $2 \le n \le 6$ and $0 \le k < n$.

$n\backslash k$	0	1	2	3	4	5
2	1	2				
3	3	48	546			
4	15	942	45132	1243704		
5	105	18375	2394360	227116260	11351644920	
6	945	375705	107314200	23919407460	3724353682560	291451508298720

Table 4 $\mathrm{TC}^{(5)}_{n,k}$ for $2 \le n \le 5$ and $0 \le k < n$.

$n\backslash k$	0	1	2	3	4
2	1	2			
3	3	66	2016		
4	15	1650	242496	28710864	
5	105	39135	17566470	7876446840	2307919133520

Table 5 $\mathrm{TC}^{(6)}_{n,k}$ for $2 \le n \le 5$ and $0 \le k < n$.

$n\backslash k$	0	1	2	3	4
2	1	2			
3	3	87	7524		
4	15	2700	1246740	676431360	
5	105	76515	118491090	262058953860	483098464854720

Random Partitions Under the Plancherel-Hurwitz Measure, High Genus Hurwitz Numbers and Maps

Guillaume Chapuy
Université Paris Cité, IRIF, CNRS, F-75013 Paris, France

Baptiste Louf
Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

Harriet Walsh
Univ Lyon, ENS de Lyon, CNRS, Laboratoire de Physique, F-69342 Lyon, France
Université Paris Cité, IRIF, CNRS, F-75013 Paris, France

Abstract

We study the asymptotic behaviour of random integer partitions under a new probability law that we introduce, the Plancherel–Hurwitz measure. This distribution, which has a natural definition in terms of Young tableaux, is a deformation of the classical Plancherel measure. It appears naturally in the enumeration of Hurwitz maps, or equivalently transposition factorisations in symmetric groups.

We study a regime in which the number of factors in the underlying factorisations grows linearly with the order of the group, and the corresponding maps are of high genus. We prove that the limiting behaviour exhibits a new, twofold, phenomenon: the first part becomes very large, while the rest of the partition has the standard Vershik–Kerov–Logan–Shepp limit shape. As a consequence, we obtain asymptotic estimates for unconnected Hurwitz numbers with linear Euler characteristic, which we use to study random Hurwitz maps in this regime. This result can also be interpreted as the return probability of the transposition random walk on the symmetric group after linearly many steps.

2012 ACM Subject Classification Mathematics of computing → Enumeration; Mathematics of computing → Distribution functions; Mathematics of computing → Random graphs

Keywords and phrases Random partitions, limit shapes, transposition factorisations, map enumeration, Hurwitz numbers, RSK algorithm, giant components

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.6

Funding This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. ERC-2016-STG 716083 "CombiTop") and from ANR COMBINÉ (ANR-19-CE48-0011).
Baptiste Louf: supported by the Knut and Alice Wallenberg foundation.

Acknowledgements We thank Philippe Biane, Jérémie Bouttier and Andrea Sportiello for insightful conversations.

1 Random partitions, Plancherel and Plancherel–Hurwitz measures

Let $n \geq 1$ be an integer and let $\mathfrak{S}_n$ denote the group of permutations on $[n] := \{1, 2, \ldots, n\}$. The famous *RSK algorithm* (Robinson, Schensted, Knuth) associates each permutation $\sigma \in \mathfrak{S}_n$ bijectively to a pair (P, Q) of standard Young tableaux (SYT) of the same shape. It is impossible to overstate the importance of this construction in enumerative and algebraic combinatorics. At the enumerative level, the RSK algorithm gives a bijective proof of the following identity:

$$\sum_{\lambda \vdash n} (f_\lambda)^2 = n!, \tag{1}$$

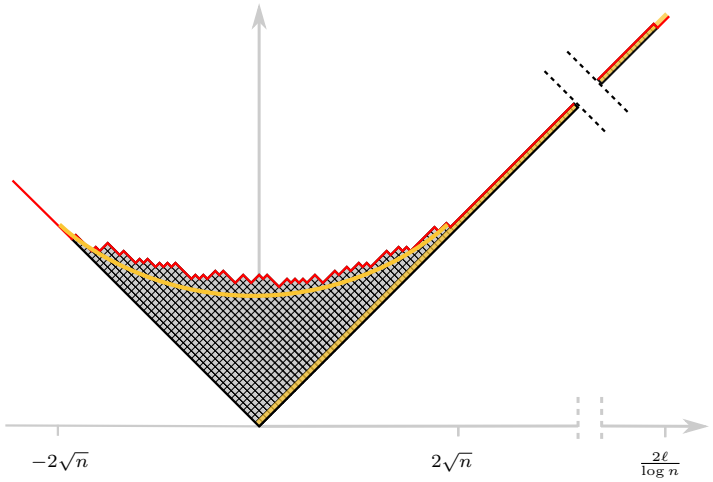


Figure 1 A random partition of $n = 2500$ under the Plancherel–Hurwitz measure $\mathbb{P}_{n,\ell}$ in the high genus regime $\ell = 2\lfloor 1.5n \rfloor$ (sampled via a Metropolis–Hastings algorithm). The twofold asymptotic behaviour is shown in yellow: the first part λ_1 is asymptotic to $\frac{2\ell}{\log(n)}$ and escapes the picture, while the rest of the partition scales in $\sqrt{n}$ with a VKLS limit shape. See Theorem 2. The profile of the partition is in red, while the VKLS limit shape scaled up to $\sqrt{n} \cdot \Omega(x/\sqrt{n})$ is the yellow curve.

where the sum is taken over integer partitions of n, and where f_λ is the number of SYT of shape λ (see Figure 2 or Section 3). If the permutation σ is chosen uniformly at random, the shape λ of the associated tableaux is a random partition of n distributed according to the probability measure

$$\lambda \mapsto \frac{1}{n!}(f_\lambda)^2, \tag{2}$$

which is the *Plancherel measure* of the symmetric group $\mathfrak{S}_n$.

The study of random partitions under the Plancherel measure is an immense subject in itself with many ramifications. One of the classical and most famous results is the fact, due independently to Logan and Shepp [15] and Vershik and Kerov [19], that when n goes to infinity, the diagram of a Plancherel distributed partition converges in some precise sense to a deterministic limit shape (Theorem 7 below) that we call the *VKLS limit shape* following these authors' initials. Other deep results deal with the behaviour of the largest part λ_1, which coincides with the *longest increasing subsequence* of the random permutation σ, which scales as $2\sqrt{n}$ with fluctuations of order $n^{1/6}$ driven by a Tracy–Widom distribution [14]. The book [17] is a delightful introduction to the subject.

In this paper we will be interested in a generalisation of this measure, motivated by the study of transposition factorisation, or Hurwitz maps, discussed in the next section. For an even integer $\ell \geq 0$, we let $H_{n,\ell}$ be the number of factorisations of the identity of $\mathfrak{S}_n$ into ℓ transpositions:

$$H_{n,\ell} = \#\{(\tau_1, \tau_2, \dots, \tau_\ell) \in (\mathfrak{S}_n)^\ell, \tau_1\tau_2\cdots\tau_\ell = \mathrm{id}, \text{ each } \tau_i \text{ is a transposition}\}. \tag{3}$$

The *Frobenius formula* from representation theory of finite groups (see e.g. [12]) together with the combinatorial representation theory of $\mathfrak{S}_n$, gives an explicit expression for the number $H_{n,\ell}$ as a sum over partitions, vastly generalising (1) (which corresponds to $\ell = 0$). Indeed,

$$H_{n,\ell} = \frac{1}{n!}\sum_{\lambda \vdash n} f_\lambda^2 (C_\lambda)^\ell, \tag{4}$$

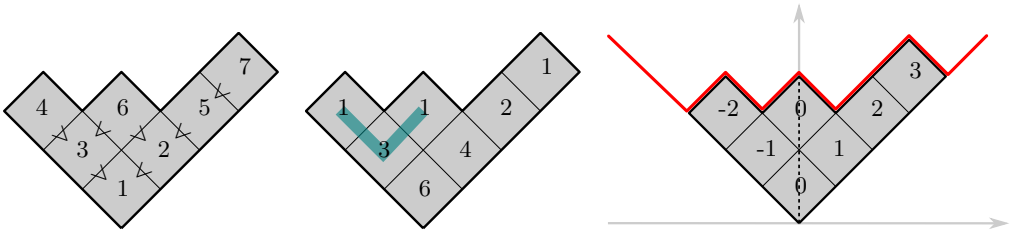


Figure 2 The Young diagram of the partition $(4,2,1) \vdash 7$ (in this paper we use the "Russian" representation where boxes are tilted by 45°). Left, its boxes are filled to produce a SYT of shape $(4,2,1)$; center, they are filled with their hook lengths, showing there are $f_{(4,2,1)} = 7!/(6\cdot 4\cdot 3\cdot 2) = 35$ such tableaux; right, each box is filled with its content, which is the abscissa of its middle point in this representation. The content-sum is $C_{(4,2,1)} = -2-1+0+0+1+2+3 = 3$. The profile is the piecewise linear function represented here in red (coordinate axes are in grey).

where C_λ is a combinatorial quantity, namely the sum of contents of all boxes of the partition λ (see Figure 2). The RHS of this formula naturally gives rise to a certain measure on partitions, which is our main object of study:

▶ **Definition 1** (Main object). *For $n \in \mathbb{Z}_{>0}$, $\ell \in 2\mathbb{Z}_{\geq 0}$, the Plancherel–Hurwitz measure is the probability measure on partitions of n defined by*

$$\mathbb{P}_{n,\ell}(\lambda) := \frac{1}{n! H_{n,\ell}} f_\lambda^2 (C_\lambda)^\ell. \tag{5}$$

The measure $\mathbb{P}_{n,\ell}$ is invariant by conjugation of a partition (vertical reflection of the diagrams), which sends the content-sum C_λ to its opposite. For $\ell > 0$ we will choose to work on the "positive half" of the measure, namely we let

$$\mathbb{P}^+_{n,\ell}(\lambda) := \mathbb{P}_{n,\ell}(\lambda | C_\lambda > 0) = 2 \cdot \mathbf{1}_{C_\lambda > 0} \cdot \mathbb{P}_{n,\ell}(\lambda). \tag{6}$$

A partition distributed under $\mathbb{P}_{n,\ell}$ for $\ell > 0$ can be thought of as a partition distributed under $\mathbb{P}^+_{n,\ell}$ which is reflected about a vertical axis with probability $\frac{1}{2}$.

When $\ell = 0$ the measure $\mathbb{P}_{n,\ell}$ is nothing but the Plancherel measure. Our main result deals instead with the case where ℓ grows linearly with n, corresponding to a high genus for the underlying map (see next section). The *rescaled profile* ψ_λ of a partition λ of n is the real function (piecewise linear with slope ± 1) whose graph follows the contour of the diagram of λ in the coordinates of its tilted diagram representation, rescaled so that each box has area $1/n$ (see Figures 1 and 2).

▶ **Theorem 2** (Main result, see Figure 1). *Fix $\theta > 0$ and let $\lambda \vdash n$ be a random partition under the Plancherel–Hurwitz measure $\mathbb{P}^+_{n,\ell}$ in the "high-genus" regime given by $\ell = \ell(n) := 2\lfloor \theta n \rfloor$. Then, as $n \to \infty$:*

(i) *the first part λ_1 is equivalent to $\frac{2\ell}{\log n}$ (in probability)*

(ii) *the rest of the partition $\tilde{\lambda} = (\lambda_2, \lambda_3, \dots)$ has a VKLS limit shape. Namely, w.h.p.,*

$$\sup_x |\psi_{\tilde{\lambda}}(x) - \Omega(x)| \to 0, \quad \text{with } \Omega(x) = \begin{cases} \frac{2}{\pi}\big(\arcsin \frac{x}{2} + \sqrt{4-x^2}\big), & |x| \leq 2 \\ |x|, & |x| > 2, \end{cases} \tag{7}$$

where $\psi_{\tilde{\lambda}}(x)$ is the rescaled profile of $\tilde{\lambda}$.

We could include λ_1 in the partition $\tilde{\lambda}$ of (7), since the supremum norm in this statement is insensitive to a small number of large parts. However, as Figure 1 indicates, λ_1 is the only part not scaling as $\sqrt{n}$ so we find this formulation more natural. Indeed, we can show that:

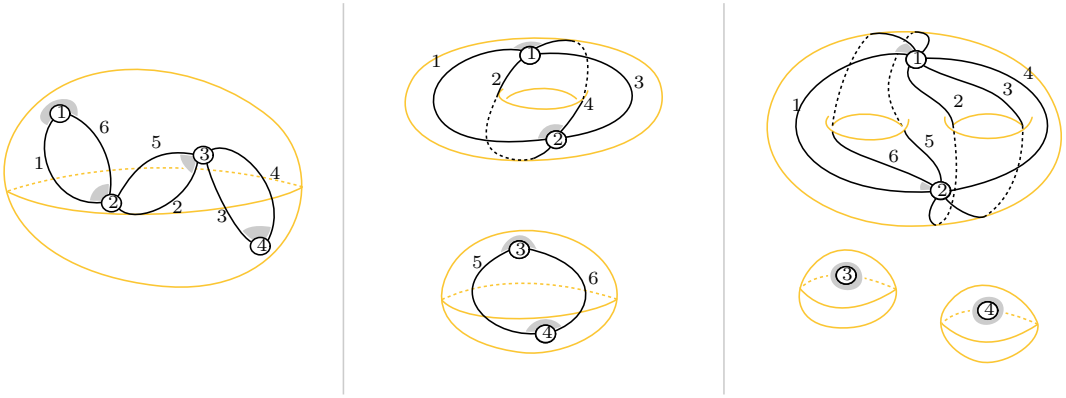


Figure 3 Three pure Hurwitz maps, each with 4 vertices, 6 edges and Euler characteristic $\chi = 0$. Left, the map corresponding to $(12)(23)(34)(34)(23)(12) = \mathrm{id}$ is connected and has genus 0; center, the map corresponding to $(12)^4(34)^2 = \mathrm{id}$ has two connected components, of genus 1 and 0; right, the map corresponding to $(12)^6 = \mathrm{id}$ has three components, of genus 2, 0 and 0.

▶ **Proposition 3.** *Under the hypotheses of Theorem 2 we have*
(iii) *the second part satisfies* $\lambda_2 \leq (e + o(1))\sqrt{n}$ *w.h.p.*

As the figure strongly suggests, the constant e can be replaced by 2 in the previous proposition. We did not have a full proof at the time of writing this extended abstract, and we refer readers to the long version of this paper, to appear shortly.

2 High genus maps, Hurwitz numbers, and random walks

Our original motivation to study the Plancherel–Hurwitz measure comes from the field of enumerative geometry and map enumeration. A *map* is a multigraph embedded on a compact oriented surface with simply connected faces, considered up to homeomorphisms. Equivalently it can be seen as a discrete surface, discretized by a finite number of polygons. Since the pioneering worsk of Tutte on planar maps (e.g. [18]) the enumeration of maps has proven to be particularly interesting, borrowing tools from physics, algebra and geometry and revealing their connections within combinatorics. These tools include matrix integral generating functions discovered by treating maps as Feynman diagrams [5], the topological recursion [11], and recurrence formulas based on integrable hierarchies [13]. Such exact methods have led to the asymptotic enumeration of many types of maps which notably exhibit a universal exponent of $-\frac{5}{2}$, and can extend to surfaces with positive genus (e.g. [7]).

These methods do not, however, extend to maps whose genus grows with the number of polygons. This "high genus" regime is one of the most recent and exciting frontiers in the field, due the inefficiency of existing generating-function or bijective methods, requiring the development of new tools. The first result in this direction was recently obtained by Budzinski and the second author [6], who showed the following estimate for the number of (connected) triangulations of size n on a surface of genus $g \sim \theta n$, by a combination of algebraic, combinatorial, and probabilistic methods:

$$T_{n,g} = n^{2g} \exp[c(\theta)n + o(n)], \quad g \sim \theta n, \tag{8}$$

where $c(\theta) > 0$ is a known continuous function. In this paper we will be interested in a different model of maps:

▶ **Definition 4** (Hurwitz map, see Figure 3). *A Hurwitz map with n vertices and ℓ edges is a map on a (non-necessarily connected) compact oriented surface, with vertices labelled from* 1 *to n and edges labelled from* 1 *to ℓ, such that the labels of edges around each vertex increase (cyclically) counterclockwise. In such a map each vertex is incident to precisely one corner which is an edge-label descent. If moreover each face of the map contains precisely one such corner, the Hurwitz map is called* pure.

It is classical, and easy to see, that Hurwitz maps of parameters n and ℓ are in bijection with tuples of transpositions $(\tau_1, \dots, \tau_\ell)$ in $\mathfrak{S}_n$, while *pure* Hurwitz maps are in bijection with tuples whose product is equal to the identity. The bijection only consists in identifying tranpositions with edges of the map, and their index with the edge-label, see Figure 3 (this construction is a special case and an adaptation of the classical construction of "constellations", see [4, 10]). The reader might find the definition of pure Hurwitz maps rather unnatural, however this model has a legitimate history in the field. In particular they are known [10], in the planar and fixed-genus cases, to belong to the same *universality class* as e.g. triangulations, quadrangulations, etc. (properties such as counting exponents or the existence of bijections are known, convergence to Brownian surfaces is conjectured). We chose this model because among the natural models of maps, it is the one for which the connection to the Plancherel measure is the most combinatorial, and it is therefore a natural candidate to test the idea of using random partition techniques to study high genus random maps.

It is important to insist that our maps are not necessarily connected, which is an important difference with most of the literature. A pure Hurwitz map of parameters n and ℓ has necessarily n faces, and its Euler characteristic χ, its number of components κ, and its generalised genus G (sum of the genera, or number of handles, of each connected component) are related by Euler's formula:

$$\chi = \#\text{vertices} - \#\text{edges} + \#\text{faces} = 2n - \ell = 2\kappa - 2G. \tag{9}$$

This is why we call the regime $\ell \gg 2n$ the "high genus" regime.

By the above correspondence, the number $H_{n,\ell}$ introduced in (3) is the number of pure Hurwitz maps with n vertices and ℓ edges. This number is called an *unconnected Hurwitz number* in the enumerative geometry literature. As a consequence of our analysis of the Plancherel–Hurwitz measure, we obtain the following estimate:

▶ **Theorem 5** (Asymptotics of high genus unconnected Hurwitz numbers). *As in* (3)*, let $H_{n,\ell}$ be the unconnected Hurwitz number counting not necessarily connected pure Hurwitz maps with n vertices and $\ell = \ell(n) = 2\lfloor \theta n \rfloor$ edges, for $\theta > 0$. Then, as $n \to \infty$,*

$$H_{n,\ell} = \left(\frac{n}{\log n}\right)^{2\ell} \exp\left[2(\log(2^{3/2}\theta) - 1)\ell + o(n)\right]. \tag{10}$$

It is tempting to see this theorem as as strong (for our model) as the Budzinski–Louf estimate (8), but unfortunately this is not quite the case. The major difference is that our maps are not necessarily connected. Moreover, we can show that (the proof is omitted in this extended abstract but follows easily from our results)

▶ **Proposition 6.** *For $\theta > 1$, as $n \to \infty$ a uniformly random unconnected Hurwitz map with n vertices and $\ell = 2\lfloor \theta n \rfloor$ edges contains a connected component with at least $\gamma(\theta)\ell$ edges, for some function $\gamma(\theta) > 0$, and $m = o(\ell)$ vertices, w.h.p.*

The fact that the "giant" edge-component in the previous proposition has a sublinear number of vertices seems to rule out the possibility of deducing asymptotics for the *connected* linear-genus regime from our results, at least not without new ideas.

At this point it is worth commenting that in map enumeration, the regime in which the genus is unconstrained, or superlinear, is often much easier to deal with than the linear case. In fact, the Plancherel–Hurwitz measure already appears (with no name) in the "super high genus" regime $\ell > \frac{1}{2} n \log n$, in work of Diaconis and Shahshahani on the transposition random walk on $\mathfrak{S}_n$. They famously showed [8] that when $\ell \geq \frac{1+\epsilon}{2} n \log n$, the walk is strongly mixed after ℓ steps, and the proof essentially consists in showing that the Plancherel–Hurwitz measure is dominated by the trivial partition (n) in this regime. In this context, our result (10) can also be interpreted as an estimate on the return probability of the random walk after linearly many steps – much before the cut-off time, at a time when the Plancherel–Hurwitz measure still has a more subtle behaviour than the trivial partition.

Finally, we note that related measures on partitions were studied by Biane in the context of the factorisation of characters of $\mathfrak{S}_n$ [1], related to the intermediate regime $\ell = 2\lfloor \theta \sqrt{n} \rfloor$ which we do not study in this work. The limit-shape phenomena observed in this reference are different from ours. We leave the study of a possible connection, and more generally the complete study of intermediate (sublinear) values of ℓ to further works.

3 Elements of the classical Plancherel case

Formally, a partition $\lambda = (\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{\ell(\lambda)})$ of n (we write "$\lambda \vdash n$") is a weakly decreasing sequence of $\ell(\lambda)$ positive integers (called parts) which sum to n. We represent it by its Young diagram in Russian convention (Figure 2). A *Standard Young Tableau (SYT)* of shape $\lambda \vdash n$ is a filling of the boxes of its diagram with all the numbers from 1 to n which is increasing along rows and columns. The number f_λ of such tableaux can be calculated by the "hook-length" formula

$$f_\lambda = \frac{n!}{\prod_{\square \in \lambda} h_\lambda(\square)} \tag{11}$$

where the hook length $h_\lambda(\square)$ is the number of boxes in a hook going down and right from the top edge of the diagram to $\square$ and up and right to the top edge (Figure 2 again).

The number f_λ is famously equal to the dimension of the irreducible representation V^λ of the symmetric group indexed by λ, a representation theoretic connection that we have no space to develop here. We will only point out, for interested readers, that the sum of all transpositions in $\mathfrak{S}_n$ acts on this module V^λ as a scalar C_λ, which is explicitly given by the sum of contents as defined in the introduction (Figure 2 again). These two facts, together with classical representation theory, are the main reasons behind the Frobenius formula (4) expressing the count of transposition factorisations in $\mathfrak{S}_n$ with tableaux-theoretic quantities.

For $\ell = 0$ the Plancherel–Hurwitz measure becomes the Plancherel measure $\mathbb{P}_{n,0}(\lambda) = \frac{1}{n!} f_\lambda^2$ which, as said in the introduction, is very well understood.

▶ **Theorem 7** (Vershik–Kerov–Logan–Shepp (VKLS) [15, 19]). *Let $\lambda \vdash n$ be a random partition under the Plancherel measure $\mathbb{P}_{n,0}$. Then, as $n \to \infty$ we have, w.h.p.,*

$$\sup_x |\psi_\lambda(x) - \Omega(x)| = 0 \qquad \text{and} \qquad \lambda_1 \leq 2\sqrt{n} + o(\sqrt{n}),\ \ell(\lambda) \leq 2\sqrt{n} + o(\sqrt{n}) \tag{12}$$

where $\psi_\lambda(x)$ is the rescaled profile of λ and $\Omega(x)$ is the curve defined at (7).

Several proofs exists of this limit shape result. Perhaps the simplest and most conceptual ones use the formulation of the Plancherel measure in the language of fermions and the infinite wedge space, which provides a direct connection with determinantal processes [3, 14]. Such approaches and their generalisations have grown into a vast field of research after the introduction of the theory of Schur processes (see e.g. [2] for an entry point).

In the case $\ell > 0$ that we study here, it is still possible (and natural) to formulate the Plancherel-Hurwitz measure in terms of the infinite wedge, see [16]. This leads to a deep connection with integrable hierachies (the KP and 2-Toda hierarchy in particular), and even to a simple looking recurrence formula to compute the number $H_{n,\ell}$ (more precisely, their connected counterpart, see [9]). However, we do not know how to use either of these tools to approach our problems (for readers familiar with the subject: the connection to determinantal processes in presence of a sandwiched content-sum scalar operator is unclear).

Other, maybe more elementary, proofs of the VKLS theorem are based on a direct scaling of the hook-length formula (11) and variational calculus. We recommend the first chapters of the book [17] as a useful reference for such approaches. A key outcome of such an approach is the following estimate for the Plancherel measure of a partition λ in terms of its rescaled profile ψ_λ, see e.g. [17, Section 1.14]:

$$\mathbb{P}_{n,0}(\lambda) = \frac{1}{n!} f_\lambda^2 = \exp\left[- n\left(1 + 2I_{\text{hook}}(\psi_\lambda) + O\left(\frac{\log n}{\sqrt{n}}\right)\right)\right] \tag{13}$$

where $I_{\text{hook}}(\cdot)$ is an "energy" functional defined by an explicit integral formula. The VKLS limiting curve $\Omega(x)$ introduced in Theorem 2 is the unique continuous function satisfying $\int(\Omega(x) - |x|)dx = 1$ which minimises $I_{\text{hook}}(\Omega)$ (see e.g. [17, Section 1.17]). This implies the limit shape part of the VKLS theorem, since any partition whose profile is "far" from $\Omega(x)$ will appear with an exponentially small probability.

The upper bound on the first part λ_1 in the VKLS theorem does not directly follow from this limit shape analysis. Classical proofs usually depend either on the RSK algorithm or on the random growth process. The fact that neither of these tools exist in the context of factorisations ($\ell > 0$) will make our proofs become harder, see comments in the next section.

4 Proofs of our results

We will now sketch the proofs of Theorems 5 and 2, and Proposition 3. Throughout this section, we consider λ to be a random partition of n distributed by the Plancherel–Hurwitz measure $\mathbb{P}^+_{n,\ell}$ with $\ell = \ell(n) = 2\lfloor \theta n \rfloor$. Heuristically, a random partition under $\mathbb{P}^+_{n,\ell}$ is driven by two different "forces":

1. on the one hand, the "Plancherel entropy": the estimate (13) shows that there is an exponential cost for the partition, in terms of the Plancherel factor f_λ^2, to deviate from the VKLS shape.
2. on the other hand, the "content-sum entropy": the factor $(C_\lambda)^\ell$ can itself become exponentially high, so the partition may prefer to deviate from VKLS if this leads to a much higher content-sum.

Our main theorem shows that the best way for the partition to adapt to this situation, is to "throw" all its contribution to a large content-sum in the first part λ_1, and that after this the rest of the partition maximises the entropy classically. We establish this fact by successive refinements, in several steps.

We now go through the proofs. We will use the notation $Z_n(\Lambda) = \frac{1}{n!}\sum_{\lambda\in\Lambda} f_\lambda^2 (C_\lambda)^\ell$ for any set Λ of partitions of n, such that the partition function is $H_{n,\ell} = Z_n(\{\lambda \vdash n\})$. We also fix $\varepsilon = \frac{1}{100}$ and split any partition $\lambda \vdash n$ into $\lambda = \lambda^+ \sqcup \lambda^-$ where λ^+ denotes the parts that are greater than $n^{1-\varepsilon}$ and λ^- the parts that are less than $n^{1-\varepsilon}$, see Figure 4. We will use the following immediate and convenient bounds.

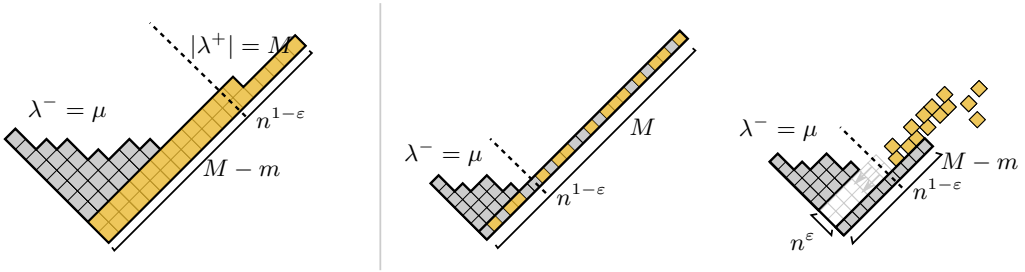


Figure 4 Left, a partition $\lambda \vdash n$ in $\Lambda(\mu, M, m)$, with λ^+ and λ^- indicated. Right, a SYT of shape $\lambda^0 \in \Lambda(\mu, M, 0)$ (the filling of the boxes is not shown) is transformed to a SYT of some shape $\lambda \in \Lambda(\mu, M, m)$ or to something else by the surjective operation used to prove Claim 13.

▶ **Lemma 8** (Useful bounds). *Let $\lambda \vdash n$ with $\lambda^+ = (\lambda_1, \dots, \lambda_p)$, then*

(i) $\frac{1}{n!} f_\lambda^2 \le \frac{n!}{\prod_{i=1}^p (\lambda_i!)^2 (n-|\lambda^+|)!} \le \frac{n^{|\lambda^+|}}{\prod_{i=1}^p (\lambda_i!)^2}$,

(ii) $C_\lambda \le \frac{\lambda_1 n}{2}$,

(iii) $C_\lambda = \sum_{i=1}^p \left(\frac{\lambda_i(\lambda_i - 1)}{2} - (i-1)\lambda_i \right) - p|\lambda^-| + C_{\lambda^-}$.

We now proceed with the succession of lemmas that constitutes the core of our proof.

▶ **Lemma 9** (Bounding the partition function below). *We have*

$$H_{n,\ell} \ge \exp\left[2\ell(\log \ell - \log\log n) - \ell(2 - \log 2) + o(n)\right]. \tag{14}$$

Proof. Let $L := \frac{2\ell}{\log n}$ and $\lambda^* = L \sqcup \mu$ with $\ell(\mu), \mu_1 \le 2\sqrt{n}$ (μ can be taken as the partition of $n - L$ maximising f_μ). Using Lemma 8(iii) we have

$$C_{\lambda^*} = \frac{L(L-1)}{2} - |\mu| + C_\mu$$

from which it is not difficult to show that

$$Z_n(\{\lambda^*\}) \ge \exp\left[2\ell(\log \ell - \log\log n) - \ell(2 - \log 2) + o(n)\right] \tag{15}$$

and this finishes the proof since $H_{n,\ell} = Z_n(\{\lambda \vdash n\}) \ge Z_n(\{\lambda^*\})$. ◀

The following lemma controls the contribution of "big parts" λ^+ in a Plancherel–Hurwitz random partition. The "truncation" threshold $n^{0.99}$ is somewhat arbitrary at this stage and will be improved to $O(\sqrt{n})$ at the very end of our analysis.

Throughout the following, let λ be a random partition under the Plancherel–Hurwitz measure $\mathbb{P}^+_{n,\ell}$ at high genus, with $\ell = 2\lfloor \theta n \rfloor$.

▶ **Lemma 10** (Controlling big parts). *W.h.p., we have $|\lambda^+| \in [\frac{0.4\ell}{\log n}, \frac{6\ell}{\log n}]$ where $L = \frac{2\ell}{\log n}$.*

Proof. Let $R_\lambda = \frac{|\lambda^+| \log n}{\ell}$. For all $\lambda \vdash n$, by Lemma 8(i),

$$\frac{1}{n!} f_\lambda^2 \le \exp\left[-(1 - 2\varepsilon) R_\lambda \ell + \frac{2 R_\lambda \ell}{\log n} \right]. \tag{16}$$

On the other hand, by Lemma 8(ii)-(iii), if $C_\lambda \ge 0$, then

$$C_\lambda^\ell \le \exp\left[2\ell(\log \ell - \log\log n) + \ell\left(\log\left(R_\lambda^2 + \frac{n^{2-\varepsilon} \log^2 n}{\ell^2} \right) - \log 2 \right) \right]. \tag{17}$$

Combining (16) and (17), and using (14), we obtain

$$\frac{Z_n(\{\lambda\})}{H_{n,\ell}} \leq \exp\left[\ell\left(2-(1-2\varepsilon)R_\lambda + \frac{2R_\lambda}{\log n} + \log\left(R_\lambda^2 + \frac{n^{2-\varepsilon}\log^2 n}{\ell^2}\right)\right)\right] \tag{18}$$

hence for n large enough and $\lambda \vdash n$ with $R_\lambda \notin [0.4, 6]$, $\mathbb{P}^+_{n,\ell}(\lambda) \leq \exp(-\ell/100)$, which entails the result since there are $e^{O(\sqrt{n})}$ partitions of n. ◀

▶ **Lemma 11** (Uniqueness of the big part). *W.h.p.,* $\lambda^+ = (\lambda_1)$.

The proof of Lemma 11 requires to compare the contribution of partitions having a single big part, to those having more than one (indeed, because we do not have exact formulas nor precise estimates on our partition functions, we can only rely on "comparison" of probabilities at this stage). We will perform this comparison among partitions having the same set of "small parts" (called μ below).

For non-negative integers M, m and partitions $\mu \vdash n - M$, we let $\Lambda(\mu, M) = \{\lambda | |\lambda^+| = M, \lambda^- = \mu\}$ and $\Lambda(\mu, M, m) = \{\lambda \in \Lambda(\mu, M) | \lambda_1 = M - m\}$. We also use the notation $\lambda^0 = M \sqcup \mu$ so that $\Lambda(\mu, M, 0) = \{\lambda^0\}$. We will need the following two claims, whose proof is postponed to after that of the lemma.

▷ Claim 12. For all $\lambda \in \Lambda(\mu, M, m)$, we have $C_\lambda \leq C_{\lambda^0} - (m-1)\frac{M}{2}$.

▷ Claim 13. If $m > 0$ then, $\sum_{\lambda \in \Lambda(\mu,M,m)} f_\lambda \leq f_{\lambda^0} \exp[m(2\varepsilon \log n + 1)]$.

Proof of Lemma 11. By Lemma 10, we know that, w.h.p., $|\lambda^+| \in [0.4\frac{\ell}{\log n}, 6\frac{\ell}{\log n}]$. We can thus assume this event for the rest of this proof.

We now condition on $|\lambda^+| = M$ and $\lambda^- = \mu$, with given $M \in [0.4\frac{\ell}{\log n}, 6\frac{\ell}{\log n}]$ and $\mu \vdash n - M$. Combining Claims 12 and 13 for $m > 0$, one obtains

$$\frac{Z_n(\Lambda(\mu, M, m))}{Z_n(\{\lambda^0\})} \leq \exp\left[\ell \log\left(1 - \frac{(m-1)M}{2C_{\lambda^0}}\right) + 2m(2\varepsilon \log n + 1)\right]. \tag{19}$$

But we know that $C_{\lambda^0} \leq (1 + o(1))\frac{M^2}{2}$ and $M \leq 6\frac{\ell}{\log n}$. Hence

$$\frac{Z_n(\Lambda(\mu, M, m))}{Z_n(\{\lambda^0\})} \leq \exp\left[-\frac{m \log n}{100}\right] \tag{20}$$

Summing this over all $m > 0$ (recall that $m \geq n^{1-\varepsilon}$ if the set is non-empty), we have

$$\sum_{m>0} Z(\Lambda(\mu, M, m)) = o(Z(\{\lambda^0\})) \tag{21}$$

which is enough to conclude that $\lambda^+ = (\lambda_1)$ w.h.p. ◀

Proof of the claims. The first claim is direct. For the second one, we need to define a proper "redistribution" operation that enables us to compare the contribution of partitions with one big part to others. To do this, we will describe an operation taking as input a SYT of shape λ^0 plus some information, and outputting a SYT of some $\lambda \in \Lambda(\mu, M, m)$, or something else, such that this operation is surjective on $\Lambda(\mu, M, m)$.

Input: A SYT T of shape λ^0.

1. Create n^ε empty rows between the first row of T and the rest,
2. choose m numbers in the first row of T *($\binom{M}{m}$ choices)*,
3. for each of these numbers, choose one of the newly created rows, and move it there *(n^ε choices each time)*,
4. sort each row and delete the empty rows, output the result.

It is easily checked that this procedure can output any SYT of λ for any $\lambda \in \Lambda(\mu, M, m)$ (indeed, for such a λ, λ^+ must have at most $\frac{n}{n^{1-\varepsilon}} = n^\varepsilon$ rows). Hence we have

$$\sum_{\lambda \in \Lambda(\mu,M,m)} f_\lambda \leq \binom{M}{m} n^{\varepsilon m} f_{\lambda^0} \leq f_{\lambda^0} \exp(m(2\varepsilon \log n + 1) \tag{22}$$

where in the last inequality we used the bound $m! \geq (m/e)^m$, along with the facts that $\log M \leq \log n$ and $\log m \geq (1-\varepsilon)\log n$. ◁

We can now collect the fruits of the previous lemmas to obtain our main theorems.

Proof of Theorem 5. The previous lemmas imply that for a Plancherel–Hurwitz distributed partition λ, we have w.h.p. $\lambda_+ = (\lambda_1)$ with $\lambda_1 = O(\frac{n}{\log(n)})$, $C_\lambda = (1+o(1))\frac{\lambda_1^2}{2}$. On the other hand, we have $\frac{1}{n!} f_\lambda^2 \leq \frac{n^{\lambda_1}}{(\lambda_1!)^2}$, hence

$$Z_n(\{\lambda\}) \leq \exp\left[2\ell \log(\lambda_1) - \ell \log 2 - \lambda_1 \log n + o(n)\right]. \tag{23}$$

Now we substitute $\lambda_1 = \frac{R_\lambda \ell}{\log n}$ in the inequality above, and we obtain

$$Z_n(\{\lambda\}) \leq \left(\frac{n}{\log n}\right)^{2\ell} \exp\left[2(\log\theta - 2)\ell\right] \exp\left[\ell(2(\log R_\lambda - \log 2) + 2 - R_\lambda) + o(n)\right]. \tag{24}$$

Now, since for $x > 0$ we always have $2(\log x - \log 2) + 2 - x \leq 0$, we get

$$Z(\{\lambda\}) \leq \left(\frac{\ell}{\log n}\right)^{2\ell} \exp\left[(-2 + \log 2)\ell + o(n)\right]. \tag{25}$$

This, together with the lower bound of Lemma 9, proves Theorem 5 since there are $e^{O(\sqrt{n})}$ partitions of n. ◀

Proof of Theorem 2, part (i). The last argument of the previous proof also implies that

$$\mathbb{P}_{n,\ell}(\lambda) \leq \exp\left[\ell(2(\log R_\lambda - \log 2) + 2 - R_\lambda) + o(n)\right]. \tag{26}$$

Now, the function on positive reals $x \longmapsto 2(\log x - \log 2) + 2 - x$ has a unique maximum at $x = 2$. Any non-negligible deviation of R_λ from this maximum thus entails an exponentially decreasing probability, which is enough to conclude that $R_\lambda = 2 + o(1)$ w.h.p. ◀

Proof of Theorem 2, part (ii). The previous discussions imply that, w.h.p., $C_\lambda = (1 + o(1))2\left(\frac{\ell}{\log n}\right)^2$ and $f_\lambda = \binom{n}{\lambda_1} f_{\tilde{\lambda}} e^{o(n)}$, which, by Theorem 5 and the Plancherel entropy estimate (13), lead to

$$\mathbb{P}_{n,\ell}(\lambda) \leq \exp\left(2n(I_{\text{hook}}(\psi_\lambda) - I_{\text{hook}}(\Omega)) + o(n)\right). \tag{27}$$

This implies, as in the classical Plancherel case (see [17, Section 1.17]), the almost sure convergence in supremum norm to the VKLS limit shape. ◀

It only remains to prove Proposition 3, i.e. to upper bound the size of λ_2. As we said already, the VKLS limit shape result in supremum norm does not imply such a bound, and even in the Plancherel case extra arguments are needed. We find convenient here to refer again to Romik's book where two bounds are given for the largest part in the Plancherel regime:

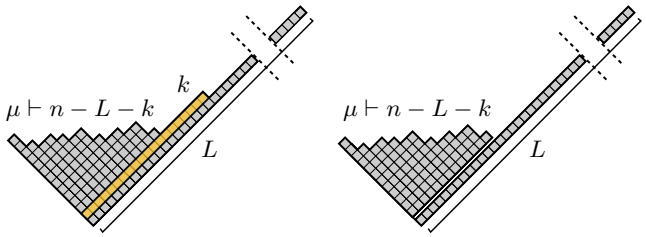


Figure 5 Partitions $L \sqcup k \sqcup \mu \vdash n$ and $L \sqcup \mu \vdash n - k$.

- an elementary bound, based on a first moment calculation and the RSK algorithm, which is enough to establish a bound of the form $(e + o(1))\sqrt{n}$ ([17, Lemma 1.4]).
- a more sophisticated bound based on the Cauchy–Schwarz inequality and on the existence of the corner-growth process for the Plancherel measure, which leads to the sharp bound $(2 + o(1))\sqrt{n}$ ([17, Section 1.19]).

In our context, we unfortunately do not have the analogue of the RSK algorithm (see next section), let alone of the corner growth process. The proof below mimics the first moment argument of the classical proof at the level of tableaux, and together with previous estimates on the partition functions enable us to reach the bound $(e+o(1))\sqrt{n}$. A more subtle approach which tries to mimic the corner growth process as in the second proof should enable us to attain soon the (conjectured) bound $(2 + o(1))\sqrt{n}$ in which case it will appear in the journal version of this paper.

Proof of Proposition 3. Under the Plancherel-Hurwitz measure, if we condition on the first part being $\lambda_1 = \frac{2\ell}{\log n} = L$, the distribution of the second part is

$$\mathbb{P}(\lambda_2 = k | \lambda_1 = L) = \frac{1}{n! Z_n[L]} \sum_{\mu \vdash n-L-k, \mu_1 \leq k} f^2_{L \sqcup k \sqcup \mu} C^{\ell}_{L \sqcup k \sqcup \mu}. \tag{28}$$

where $Z_n[L] \equiv Z_n(\{\lambda \vdash n | \lambda_1 = L\})$. Comparing SYT of shape $L \sqcup k \sqcup \mu \vdash n$ with ones of shape $L \sqcup \mu \vdash n-k$, obtained by removing the second part, and the contents of the partitions, we have

$$f_{L \sqcup k \sqcup \mu} \leq \binom{n}{k} f_{L \sqcup \mu}, \qquad C_{L \sqcup k \sqcup \mu} = C_{L \sqcup \mu} - |\mu| + \frac{k^2}{2} = C_{L \sqcup \mu}(1 + o(1)) \tag{29}$$

and from there we obtain

$$\mathbb{P}(\lambda_2 = k | \lambda_1 = L) \leq \binom{n}{k}^2 \frac{(n-k)!}{n!} \frac{Z_{n-k}[L]}{Z_n[L]} (1 + o(1)). \tag{30}$$

Now, following an application of the identity $n f_\mu = \sum_{\nu : \mu \nearrow \nu} f_\nu$ for $\mu \vdash n$, where "$\mu \nearrow \nu$" means that ν is obtained from μ by adding one box, and using elementary bounds on the variation of the content-sum when a single box is added, it is possible to show that

$$Z_n[L] = Z_{n-1}[L] e^{o(1)}. \tag{31}$$

It follows that

$$\mathbb{P}(\lambda_2 = k | \lambda_1 = L) \leq \frac{n!}{k!^2 (n-k)!} e^{o(k)} \leq \frac{n^k}{(k/e)^{2k}} e^{o(k)}, \tag{32}$$

and, to conclude the proof,

$$\forall \varepsilon > 0, \qquad \lim_{n \to \infty} \mathbb{P}(\lambda_2 = (1 + \varepsilon) e \sqrt{n} | \lambda_1 = L) = 0 \tag{33}$$

and we have $\lambda_2 \leq e(1 + o(1))\sqrt{n}$ w.h.p. as $n \to \infty$. ◀

5 Open questions and perspectives

Maybe the main open question that follows our work is the following: *does there exist an analogue of the RSK algorithm proving combinatorially the identity* (4)*?* If this is the case, then our results about λ_1 and λ_2 probably translate into distributional limit theorems for certain parameters of random factorisations (or random pure Hurwitz maps). To start with, can one identify the "meaning" of the statistic λ_1 on the Hurwitz side?

Another question is, of course, to know if one can use the Plancherel–Hurwitz approach to say anything about *connected* Hurwitz maps of high genus. This would be very interesting. It may also be interesting to combine this approach with the technology of integrable hierarchies, which have been so fruitful but have so far not directly led to precise asymptotic estimates nor limit theorems for connected random maps or Hurwitz numbers of high genus.

References

1 P. Biane. Approximate factorization and concentration for characters of symmetric groups. *International Mathematics Research Notices*, 2001(4):179–192, January 2001.

2 A. Borodin and V. Gorin. Lectures on integrable probability. In *Probability and Statistical Physics in St. Petersburg*, volume 91 of *Proceedings of Symposia in Pure Mathematics*, pages 155–214. AMS, 2016. `arXiv:1212.3351`.

3 A. Borodin, A. Okounkov, and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.*, 13(3):481–515, July 2000.

4 M. Bousquet-Mélou and G. Schaeffer. Enumeration of planar constellations. *Advances in Applied Mathematics*, 24(4):337–368, 2000.

5 E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber. Planar diagrams. *Communications in Mathematical Physics*, 59(1):35–51, 1978.

6 T. Budzinski and B. Louf. Local limits of uniform triangulations in high genus. *Inventiones mathematicae*, 223, January 2021.

7 G. Chapuy, M. Marcus, and G. Schaeffer. A bijection for rooted maps on orientable surfaces. *SIAM J. Discrete Math.*, 23(3):1587–1611, 2009.

8 P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57:159–179, 1981.

9 B. Dubrovin, D. Yang, and D. Zagier. Classical hurwitz numbers and related combinatorics. *Moscow Mathematical Journal*, 17:601–633, October 2017.

10 E. Duchi, D. Poulalhon, and G. Schaeffer. Bijections for simple and double Hurwitz numbers, 2014. `arXiv:1410.6521`.

11 B. Eynard and N. Orantin. Invariants of algebraic curves and topological expansion. *Communications in Number Theory and Physics*, 1, March 2007.

12 W. Fulton. *Young Tableaux: With Applications to Representation Theory and Geometry*. London Mathematical Society Student Texts. Cambridge University Press, 1996.

13 I.P. Goulden and D.M. Jackson. The KP hierarchy, branched covers, and triangulations. *Advances in Mathematics*, 219(3):932–951, 2008.

14 K. Johansson. The longest increasing subsequence in a random permutation and a unitary random matrix model. *Mathematical Research Letters*, 5:63–82, January 1998.

15 B.F. Logan and L.A. Shepp. A variational problem for random Young tableaux. *Advances in Mathematics*, 26(2):206–222, 1977.

16 A. Okounkov. Toda equations for Hurwitz numbers. *Math. Res. Lett.*, 7(4):447–453, 2000.

17 D. Romik. *The Surprising Mathematics of Longest Increasing Subsequences*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2015.

18 W. T. Tutte. A census of planar maps. *Canadian Journal of Mathematics*, 15:249–271, 1963.

19 A.M. Vershik and S.V. Kerov. Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux. *Doklady Akademii Nauk*, 233(6):1024–1027, 1977.

Universal Properties of Catalytic Variable Equations

Michael Drmota
TU Wien, Institute of Discrete Mathematics and Geometry, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

Eva-Maria Hainzl
TU Wien, Institute of Discrete Mathematics and Geometry, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

Abstract

Catalytic equations appear in several combinatorial applications, most notably in the enumeration of lattice paths and in the enumeration of planar maps. The main purpose of this paper is to show that under certain positivity assumptions the dominant singularity of the solution function has a universal behavior. We have to distinguish between linear catalytic equations, where a dominating square-root singularity appears, and non-linear catalytic equations, where we – usually – have a singularity of type 3/2.

2012 ACM Subject Classification Mathematics of computing → Generating functions

Keywords and phrases catalytic equation, singular expansion, univeral asymptotics

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.7

Funding *Michael Drmota*: Research supported by the Austrian Science Foundation FWF, projects F50-02, F55-02, and P35016.
Eva-Maria Hainzl: Research supported by the Austrian Science Foundation FWF, project F55-02.

Acknowledgements We thank the referees for their careful reading and their valuable comments for improving the presentation.

1 Introduction

Catalytic equations have their origin mostly in map enumeration [12] and in lattice path enumeration [3]. Such equations were first solved with the help of the *kernel method* [3, 11, 1] (in the linear case) and with the help of the *quadratic method* [12, 5] (in the quadratic case). Both approaches were unified and extended by Bousquet-Mélou and Jehanne [4]. They considered *general catalytic equations* of the form

$$P(z, u, M(z, u), M_1(z), \dots, M_k(z)) = 0, \tag{1}$$

where $P(z, u, x_0, x_1, \dots x_k)$ is a polynomial and all power series $M(z, u), M_1(z), \dots, M_k(z)$ are characterized by this equation.

The variable "u" is called *catalytic* since it is usually an auxiliary variable that counts an additional (usually combinatorial) parameter which simplifies the recursive decomposition of the structure of interest. In general, one is just interested in the function $M(z, 0)$, $M(z, 1)$ or in $M_1(z)$.

One of the most prominent examples is the counting problem of rooted planar maps that goes back to Tutte [12]. Let $M_k(z)$ denote the generating function of those maps, where the root face has valency $k \geq 0$. Then we have $M_0(z) = 1$ and

$$M_k(z) = z \sum_{j=0}^{k-2} M_j(z) M_{k-j-2}(z) + z \sum_{j=k-1}^{\infty} M_j(z) \qquad (k \geq 1) \tag{2}$$

where the right sum arises if the root edge is not a bridge and the left sum if the deletion of the root edge results in decomposing the map into two components. One is interested in the generating function $M(z) = \sum_{k\geq 0} M_k(z)$ of all planar maps. By introducing the variable u and setting $M(z,u) = \sum_{k\geq 0} M_k(z)u^k$, the infinite system (2) rewrites to the *catalytic equation*

$$M(z,u) = 1 + zu^2 M(z,u)^2 + uz\frac{uM(z,u) - M(z,1)}{u-1}. \tag{3}$$

By using the above mentioned quadratic method the equation can be explicitly solved:

$$M(z) = M(z,1) = \frac{18z - 1 + (1-12z)^{3/2}}{54z^2}.$$

This leads to an explicit formula $M_n = [z^n]\, M(z,1) = \frac{2(2n)!}{(n+2)!n!}3^n$ and to an asymptotic one: $M_n \sim (2/\sqrt{\pi})12^n n^{-5/2}$. Note that the asymptotic behavior is reflected by the dominant singular behavior of $M(z,1)$ at $z_0 = 1/12$. The type of the singularity is $3/2$ which translates to the critical exponent $-5/2 = -1 - 3/2$ by the well known Transfer Lemma [10].

In [4] several applications mostly from map enumeration (different classes of planar maps, constellations, hard particles in planar maps etc.) are given. Bousquet-Mélou and Jehanne [4] considered in particular equations of the form

$$M(z,u) = F_0(z,u) + zQ\left(z, u, M(z,u), \Delta^{(1)}(M(z,u)), \ldots, \Delta^{(k)}(M(z,u))\right), \tag{4}$$

where $F_0(z,u)$ and $Q(z,u,\alpha_0,\alpha_1,\ldots,\alpha_k)$ are polynomials and where we have used the abbreviations

$$\Delta^{(j)}(M(z,u)) = \frac{M(z,u) - M(z,0) - u\,M_u(z,0) - \cdots u^{j-1}M_{u^{j-1}}(z,0)}{u^j} \qquad (j \geq 1).$$

It is convenient to consider just the catalytic variable u at 0. In the above case of planar maps we substitute u by $u+1$ to reduce it to this case.

One main result of [4] it that equations of the form (1) can be solved with the help of proper systems of polynomial equations. Hence, the solutions are always algebraic functions and consequently for every singularity we have a Puiseux expansion. However, this approach does not specify the kind of the Puiseux expansion. There is in principle no restriction on the rational exponents that might occur.

However, if we consider the special case $k = 1$ (in (4))

$$M(z,u) = F_0(z,u) + zQ\left(z, u, M(z,u), \frac{M(z,u) - M(z,0)}{u}\right), \tag{5}$$

where $F_0(z,u)$ and $Q(z,u,\alpha_0,\alpha_1)$ are polynomials with *non-negative coefficients*, then Drmota, Noy, and Yu [7] showed that there is a dichotomy (under natural conditions on Q). If Q is linear in α_0 and α_1 then the dominant singularity is of type $\frac{1}{2}$, that is, a square-root singularity which leads to an asymptotic behavior for the coefficients of the form $\sim c\rho^n n^{-3/2}$. However, in the non-linear case the dominant singularity is of type $\frac{3}{2}$ (as in the above mentioned example of planar maps) which means that the coefficients are asymptotically of the form $\sim c\rho^n n^{-5/2}$.

In what follows we will focus on the case $k = 2$, where F_0 and Q are polynomials with non-negative coefficients and we will show that the results for the case $k = 1$ can be extended. However, there are several (major) differences. Whereas in the case $k = 1$ the catalytic

equation can be solved with the help of a so-called positive system of polynomial equations (see [2, 7]) which determines directly a dominant square-root singularity for the involved solution function this property is widely lost for the cases $k \geq 2$. Thus, it is necessary to develop new methods and concepts in order to deduce the universal singular behavior. Clearly we expect similar properties for all equations of the form (4) as well as for systems of positive catalytic equations but the cases $k > 2$ are even more involved.

2 Main results

The solution method by Bousquet-Mélou and Jehanne [4] for an equation of the form (1) works as follows. One considers the algebraic system of $3k$ equations

$$\begin{aligned} P(z, u_i(z), f_i(z), M_1(z), \dots, M_k(z)) &= 0, \quad 1 \leq i \leq k, \\ P_{x_0}(z, u_i(z), f_i(z), M_1(z), \dots, M_k(z)) &= 0, \quad 1 \leq i \leq k, \\ P_u(z, u_i(z), f_i(z), M_1(z), \dots, M_k(z)) &= 0, \quad 1 \leq i \leq k, \end{aligned} \tag{6}$$

for the $3k$ unknown functions $u_1(z), \dots, u_k(z)$, $f_1(z), \dots, f_k(z)$, $M_1(z), \dots, M_k(z)$. In general it is not clear that such a system is solvable. However, if the catalytic equation is of the form (4) then this is granted and leads to the unknown functions $M_1(z), \dots, M_k(z)$ (see [4]).

In our context we reformulate the catalytic equation slightly to

$$\begin{aligned} &u^2\Delta(z,u) + uM_1(z) + M_0(z) \qquad (7)\\ &\quad = zQ\left(z, u, u^2\Delta(z,u) + uM_1(z) + M_0(z), u\Delta(z,u) + M_1(z), \Delta(z,u)\right) \\ &\quad =: R(z, u, \Delta(z,u), M_1(z), M_0(z)) \qquad (8) \end{aligned}$$

where $\Delta(z,u) = \Delta^{(2)}M(z,u)$, $M_1(z) = M_u(z,0)$, $M_0(z) = M(z,0)$ and $Q(z,u,\alpha_0,\alpha_1,\alpha_2)$ and consequently $R(z,u,y_0,y_1,y_2)$ are polynomials with non-negative coefficients. W.l.o.g. the polynomial part $F_0(z,u)$ can be omitted by substituting $M(z,u) = \widetilde{M}(z,u) + F(0,u)$. In particular we have

$$\begin{aligned} P(z,u,x_0,x_1,x_2) &= zQ\left(z,u,u^2x_0 + ux_1 + x_2, ux_0 + x_1, x_0\right) - u^2x_0 - ux_1 - x_2 \\ &= R(z,u,x_0,x_1,x_2) - u^2x_0 - ux_1 - x_2. \end{aligned}$$

The system (6) now rewrites to

$$\begin{aligned} u_i(z)^2\Delta_i(z) + u_i(z)M_1(z) + M_0(z) &= R\left(z, u_i(z), \Delta_i(z), M_1(z), M_0(z)\right), \quad i = 1,2, \\ u_i(z)^2 &= R_{y_0}\left(z, u_i(z), \Delta_i(z), M_1(z), M_0(z)\right), \quad i = 1,2, \\ 2u_i(z)\Delta_i(z) + M_1(z) &= R_u\left(z, u_i(z), \Delta_i(z), M_1(z), M_0(z)\right), \quad i = 1,2, \end{aligned} \tag{9}$$

for the six indeterminate functions $M_1(z), M_0(z)$, $u_{1,2}(z)$ and $\Delta_{1,2}(z)$ (which correspond to the functions $\Delta_{1,2}(z) = \Delta(z, u_{1,2}(z))$). In order to distinguish between $i = 1$ and $i = 2$ we assume that $u_1(z) > 0$ and $u_2(z) < 0$ for $z > 0$ that are sufficiently small.

We now state our main results that generalize the results of [7] to the case $k = 2$. We say that an algebraic function has a square-root singularity at z_0 if the dominating term in the Puiseux expansion at z_0 is of the form $(z - z_0)^{1/2}$. Similiarly we say that a singulariy at z_0 has type $3/2$ if the dominating term is of the form $(z - z_0)^{3/2}$. Recall that all solutions of (7) are algebraic.

▶ **Theorem 1.** *Suppose that the polynomial Q in the catalytic equation (7) is linear in $(\alpha_0, \alpha_1, \alpha_2)$ and has non-negative coefficients. Suppose further that Q_{α_0} is not a polynomial in u^2 and that u does not divide Q_{α_1}. Then the functions $M(z,0)$ and $M_u(z,0)$ have a common radius of convergence z_0 and a square-root singularity at z_0.*

For the second theorem we will need an extra condition of the term

$$T := R_{uuu}+(3R_{uuy_0}-6)\frac{2u-R_{uy_0}}{R_{y_0y_0}}+3R_{uy_0y_0}\left(\frac{2u-R_{uy_0}}{R_{y_0y_0}}\right)^2+R_{y_0y_0y_0}\left(\frac{2u-R_{uy_0}}{R_{y_0y_0}}\right)^3. \quad (10)$$

▶ **Theorem 2.** *Suppose that the polynomial Q in the catalytic equation (7) is non-linear in $(\alpha_0, \alpha_1, \alpha_2)$ and has non-negative coefficients. Suppose further that $Q_{\alpha_0 u} \neq 0$ and that u does not divide Q_{α_1}. Then the functions $M(z,0)$ and $M_u(z,0)$ have a common radius of convergence z_0, and if $T \neq 0$ at $(z,u) = (z_0, u_1(z_0))$ then both $M(z,0)$ and $M_u(z,0)$ have a singularity of type $3/2$.*

We first comment on the conditions on the polynomial Q. They are just put to simplify the presentation. They exclude degenerate cases that reduce to finite systems or systems or to cases where at least one solution to the curve equation is constant 0 (which can further be reduced to cases with universal laws).

Secondly the condition $T \neq 0$ at $(z,u) = (z_0, u_1(z_0))$ in Theorem 2 seems to be artificial. Actually one always has $T \geq 0$ and we have not found an example yet, where $T = 0$, but it is unclear how the zero case could be excluded. A corresponding condition for the case $k = 1$ always holds, since in this case all (corresponding) summands are positive. Nevertheless, the case $T = 0$ can be also discussed and we would get a dominating singularity of the form $(z - z_0)^{4/3}$.

Finally, as mentioned above for the planar map counting problem, the type of the dominating singularity is reflected in the asymptotic behavior of the coefficients. In order to keep the presentation simple we do not go into these details. We just mention that in the linear case the square-root singularity corresponds to asymptotics of the form $c\, z_0^{-n} n^{-3/2}$ whereas in the non-linear case the singularity of type $3/2$ corresponds to asymptotics of the form $c\, z_0^{-n} n^{-5/2}$. However, in general these kinds of asymptotics hold only in residue classes (compare with the results of [7]).

▶ **Example 3.** Let us consider one-dimensional non-negative lattice paths, where we allow steps of the form ± 1 and ± 2. The generating functions $E_k(z)$ of walks that start at 0 and end at level k satisfy the system of equations

$$\begin{aligned} E_0(z) &= 1 + z(E_1(z) + E_2(z)),\\ E_1(z) &= z(E_0(z) + E_1(z) + E_2(z)),\\ E_k(z) &= z(E_{k-2}(z) + E_{k-1}(z) + E_{k+1}(z) + E_{k+2}(z)) \qquad (k \geq 2). \end{aligned}$$

Hence, the generating function $E(z,u) = \sum_{k\geq 0} E_k(z)u^k$ satisfies

$$E(z,u) = 1 + z(u+u^2)E(z,u) + z\frac{E(z,u)-E(z,0)}{u} + z\frac{E(z,u)-E(z,0)-uE_v(u,0)}{u^2}. \quad (11)$$

This is precisely a linear equation of the form (4) with $k = 2$. Theorem 1 applies directly and implies that the generating function $E_0(z) = E(z,0)$ of excursions has a square-root singularity, compare also with [3] or with the discussion in Section 6.

▶ **Example 4.** 3-Constellations are Eulerian maps, where the faces are bi-colored, black faces have valency 3 whereas white faces have a valency that is a multiple of 3 (more generally one considers m-constellations, see [4]). The corresponding (catalytic) equation for 3-constellations is given by

$$\begin{aligned} C(z,u) = 1 &+ zuC(z,u)^3 + zu(2C(z,u) + C(z,1))\frac{C(z,u)-C(z,1)}{u-1}\\ &+ zu\frac{C(z,u)-C(z,1)-(u-1)C_u(z,1)}{(u-1)^2} \end{aligned}$$

This catalytic equation is almost of the form, where we can apply Theorem 2 due to the additional appearance of $C(z,1)$. However, the polynomial P in (8) has still non-negative coefficients. Thus a slight extension of Theorem 2 applies, where we require the determinant of A in the calculations of Section 5 to be positive at z_0. Furthermore, (10) is satisfied. Consequently, the function $C(z,1)$ has a dominant singularity of type $3/2$, see also the discussion in Section 6

Further examples can be found in [4, 7]. It should be also mentioned that Theorems 1 and 2 can be extended to prove central limit theorems for several parameters that are encoded by an additional variable (see [8, 9, 7]).

3 The Curve Equation

Our first observation is that $M(z,u)$ which is defined by equation (7) is analytic by considering the equation as a fixed point problem in the sequence space of the coefficients. The factor z on the right hand side can be chosen small enough such that the map is a contraction and yields a unique solution that is analytic by uniform convergence. Furthermore, by rewriting (7) into an infinite system (by considering the expansion with respect to u and by iterating this system) it follows that the solution function has non-negative coefficients.

Given that we know there is a unique solution $M(z,u)$ fulfilling the equation, we may regard equation (7) (or (8)) as an equation in z and u and differentiate the equation with respect to u and group the terms with a factor $\partial_u\Delta(z,u)$ into the equation

$$u^2 = R_{y_0}\left(z,u,\Delta(z,u),M_1(z),M_0(z)\right) =: C(z,u)$$

which was proven to have two unique solutions $u_1(z)$ and $u_2(z)$, with $u_1(0) = u_2(0) = 0$. We will refer to $u^2 = C(z,u)$ as the *curve equation*. Note that $C(z,u)$ has non-negative coefficients.

Next we consider $u_1(z)$ and $u_2(z)$. In general, the two series only have a Puiseux expansion at 0 but applying the Weierstrass preparation theorem to equation (7), we can see that locally both $u_i(z)$ are zeros of

$$u_i^2 - C(z,u_i) = K(z,u_i)\left(u_i^2 + a_1(z)u_i + a_2(z)\right) = 0$$

where $K(z,u)$, $a_1(z), a_2(z)$ are analytic functions at 0 with $a_1(0) = a_2(0) = 0$ and $K(0,0) \neq 0$. Note that all these functions are uniquely given. Hence, we can express

$$u_{1,2}(z) = -\frac{a_1(z)}{2} \pm \sqrt{\frac{a_1(z)^2}{4} - a_2(z)} =: g(z) \pm \sqrt{h(z)}. \tag{12}$$

Now the idea is to split u and all power series in u into two parts: one with a factor $\sqrt{h}$ and the other without. That is, for

$$u_{1,2}^2 = (g \pm \sqrt{h})^2 = g^2 + h \pm \sqrt{h}\,2g.$$

we define $(u^2)^+ = g^2 + h$ and $(u^2)^- = 2g$ and further we split

$$\Delta(z,u_{1,2}) = \Delta^+(z,g,h) \pm \sqrt{h}\,\Delta^-(z,g,h).$$

By doing the same with $R(z,u_{1,2},\Delta_{1,2},M_1,M_0)$ and the curve equation

$$g^2 + h \pm \sqrt{h}2g = C^+(z,g,h) \pm \sqrt{h}\,C^-(z,g,h)$$

and considering the unique solutions to the system

$$h = C^+(z,g,h) - g^2, \quad g = \frac{1}{2}C^-(z,g,h) \tag{13}$$

that consequently have to be exactly g and h as defined in (12) we may derive the following result.

▶ **Lemma 5.** *Suppose that $C(z,u)$ is a power series with non-negative coefficients such that z divides $C(z,u)$. Furthermore let $u_{1,2}(z) = g(z) \pm \sqrt{h(z)}$ be the two solutions with $u(0) = 0$ of the equation $u^2 = C(z,u)$. Then $g(z)$ and $h(z)$ are power series with $g(0) = h(0) = 0$ and non-negative coefficients.*

The non-negativity of the coefficients does not follow immediately. In fact, we have to verify that $h'(0), g'(0) > 0$ and subsequently that all higher derivatives $h^{(n)}(0), g^{(n)}(0) > 0$ individually. We will leave out the detailed proof, since it is long and technical and most of the rest of our results do not rely on this fact. The important part is that $u_1(z)$ is positive and monotone increasing for $z > 0$ and that $|u_2(z)| < u_1(z)$ if the curve equation is not a power series in u^2.

4 Proof of Theorem 1 (The Linear Case)

If Q is linear in α_0, α_1 and α_2, we can rewrite (7) to

$$M(z,u) = R_0(z,u) + zR_1(z,u)M(z,u) + zR_2(z,u)\Delta M(z,u) + zR_3(z,u)\Delta^{(2)}M(z,u)$$

where $R_0(z,u), R_1(z,u), R_2(z,u), R_3(z,u)$ are polynomials with non-negative coefficients. Equivalently we have

$$\begin{aligned} &M(z,u)\left(1 - \left(R_1(z,u) + \frac{1}{u}R_2(z,u) + \frac{1}{u^2}R_3(z,u)\right)\right) \\ &\quad = R_0(z,u) - zR_2(z,u)\frac{M(z,0)}{u} - zR_3(z,u)\left(\frac{M_u(z,0)}{u} + \frac{M(z,0)}{u^2}\right). \end{aligned} \tag{14}$$

In this case, the curve equation is a polynomial equation in u and z

$$u^2 - C(z,u) = u^2 - \left(u^2R_1(z,u) + uR_2(z,u) + R_3(z,u)\right)$$

and can be independently solved (and is actually the basic equation of the original kernel method). Subsequently, by using (14) and the two solutions $u_{1,2}(z)$ of the curve equation we get the following linear system of equations

$$M(z,0) + \left(u_1(z) - z\frac{R_2(z,u_1(z))}{1 - zR_1(z,u_1(z))}\right)M_u(z,0) = \frac{R_0(z,u_1(z))}{1 - zR_1(z,u_1(z))}, \tag{15}$$

$$M(z,0) + \left(u_2(z) - z\frac{R_2(z,u_2(z))}{1 - zR_1(z,u_2(z))}\right)M_u(z,0) = \frac{R_0(z,u_2(z))}{1 - zR_1(z,u_2(z))}, \tag{16}$$

to calculate $M(z,0)$ and $M_u(z,0)$. (Of course if these functions are given we can use them to obtain the full solution function $M(z,u)$.)

We start by determining the singular expansions of $u_1(z)$ and $u_2(z)$, where we will use the following lemma to show that $g(z), h(z)$ have a common square root singularity at their radius of convergence z_0 and $u_2(z)$ is regular at z_0.

▶ **Lemma 6.** *Let $C(z,u) = \sum_{k,j\geq 0} c_{k,j} z^k u^j$ be an analytic function with non-negative coefficients and k_1, k_2 and $j_1 < j_2$ such that $c_{k_1,j_1}, c_{k_2,j_2} \neq 0$. Then, for $z, u > 0$ inside the region of convergence, it holds that*

$$C(z,u)C_{uu}(z,u) - C_u(z,u)^2 + \frac{C_u(z,u)C(z,u)}{u} > 0$$

Proof. By assumption we clearly have $|C(z,ue^{i\theta})| \leq C(z,u)e^{-c\theta^2}$ for $z,u>0$, θ sufficiently close to 0 and some constant $c = c(z,u) > 0$ if C is a power series with positive coefficients. Further, by using an exp-log scheme and the Taylor expansion of the logarithm at $\theta = 0$, it holds that

$$C(z,ue^{i\theta}) = \exp\left(\log(C(z,u)) + iu\frac{C_u(z,u)}{C(z,u)}\theta + \frac{u^2}{2\,C(z,u)^2}b(z,u)\theta^2 + o(\theta^2)\right)$$

where $b(z,u) = C_u(z,u)^2 - \frac{C_u(z,u)C(z,u)}{u} - C_{uu}(z,u)C(z,u)$. Thus,

$$|C(z,ue^{i\theta})| = C(z,u)e^{\frac{u^2}{2\,C(z,u)^2}b(z,u)\theta^2 + o(\theta^2)} \leq C(z,u)e^{-c\theta^2}$$

and consequently, that factor $b(z,u)$ has to be negative. ◀

▶ **Lemma 7.** *Suppose that $C(z,u)$ is not a power series in u^2, but a power series with non-negative coefficients with degree ≥ 3 in u, and where z divides $C(z,u)$. Further denote by $u_{1,2}(z) = g(z) \pm \sqrt{h(z)}$ the two solutions with $u(0) = 0$ of the equation $u^2 = C(z,u)$. If $(z_0, u_1(z_0))$ is inside the region of convergence of $C(z,u)$, where z_0 is the smallest $z > 0$ such that $2u_1(z_0) = C_u(z_0, u_1(z_0))$, then the critical exponent of $u_1(z)$ at z_0 is $1/2$, while $u_2(z)$ is regular at z_0. That is, $g(z), h(z)$ have a common square-root singularity at their radius of convergence z_0, and their square-root singularities cancel in the representation $u_2(z) = g(z) - \sqrt{h(z)}$.*

Proof. We certainly have

$$u_1(z_0)^2 = C(z_0, u_1(z_0)), \quad 2u_1(z_0) = C_u(z, u_1(z_0)), \quad C_z(z, u_1(z_0)) > 0.$$

By Lemma 6 it also follows that $2 < C_{uu}(z, u_1(z_0))$. Then by standard arguments using the Weierstrass preparation theorem (compare with [6, Remark 2.20]), we can derive that $u_1(z)$ is locally equal to $u_1(z) = g_1(z) + h_1(z)\sqrt{z - z_0}$, where $g_1(z), h_1(z)$ are analytic functions around z_0 and $h_1(z_0) \neq 0$. Now let us assume that $2u_2(z_0) = C_u(z_0, u_2(z_0))$ as well. This would imply that

$$2(g_0 + \sqrt{h_0}) = C_u(z_0, g_0 + \sqrt{h_0}) \quad \text{and} \quad 2(g_0 - \sqrt{h_0}) = C_u(z_0, g_0 - \sqrt{h_0}),$$

where $g_0 = g(z_0)$ and $h_0 = h(z_0)$. Since

$$\frac{C_u(z, g + \sqrt{h}) + C_u(z, g - \sqrt{h})}{2} = C_g^+(z,g,h)$$

it would follow that $2g_0 = C_g^+(z_0, g_0, h_0)$ and therefore $C_g^+(z_0,g_0,h_0) - C^-(z_0,g_0,h_0) = 0$. At this point we mention that

$$\binom{k}{2\ell}(k - 2\ell) - \binom{k}{2\ell+1} = \binom{k}{2\ell+1}2\ell \geq 0$$

which ensures that

$$\begin{aligned} C_g^+(z,g,h) - C^-(z,g,h) = \\ &= \sum_{k,\ell} C_k(z)\binom{k}{2\ell}(k-2\ell)g^{k-2\ell-1}h^\ell - \sum_{k,\ell} C_k(z)\binom{k}{2\ell+1}g^{k-2\ell-1}h^\ell \\ &= \sum_{k,\ell} C_k(z)\binom{k}{2\ell+1}2\ell g^{k-2\ell-1}h^\ell \end{aligned} \tag{17}$$

Since C has degree ≥ 3 in u and $g \neq 0$ if $C(z,u)$ is not a power series in u^2 this is a contradiction to being 0. Hence, $2u_2(z_0) \neq C_u(z_0, u_2(z_0))$ and consequently $u_2(z)$ is regular at z_0. This implies further that $g(z) = (u_1(z) + u_2(z))/2$ and $h(z) = (u_1(z) - u_2(z))^2/4$ share a square root singularity at z_0. ◀

In a final step we can also detect the singular behavior of $M(z,0)$ and $M_u(z,0)$.

▶ **Lemma 8.** *Suppose that the assumptions of Theorem 1 are satisfied and let $M(z,0)$ and $M_u(z,0)$ be the solutions of the linear system (15)–(16). Then $M(z,0)$ and $M_u(z,0)$ have square-root singularities at z_0.*

Proof. We recall that $M(z,0)$ and $M_u(z,0)$ are given by (15)–(16) and that $u_1(z)$ and $u_2(z)$ are the solutions to the curve equation, where $u_1(z)$ has a square-root singularity at z_0, whereas $u_2(z)$ is regular at z_0. We recall that (15)–(16) can be rewritten as

$$M(z,0) + \left(u_{1,2}(z) - \frac{zR_2(z,u_{1,2}(z))}{1 - zR_1(z,u_{1,2}(z))}\right) M_u(z,0) = \frac{R_0(z,u_{1,2}(z))}{1 - zR_1(z,u_{1,2}(z))}. \tag{18}$$

At this point we rewrite $u_{1,2}(z)$ as $u_{1,2}(z) = g(z) \pm \sqrt{h(z)}$ and split up between the $+$-part and the $-$-part. In particular we have

$$(1)^- = 0 \quad \text{and} \quad \left(u - \frac{zR_2(z,u)}{1 - zR_1(z,u)}\right)^- = 1 - \left(\frac{zR_2(z,u)}{1 - zR_1(z,u)}\right)^-,$$

which leads to

$$\left(1 - \left(\frac{zR_2(z,u)}{1 - zR_1(z,u)}\right)^-\right) M_u(z,0) = \left(\frac{R_0(z,u)}{1 - zR_1(z,u)}\right)^-$$

Now notice that by our conditions that Q_{α_0} is not a polynomial in u^2 the negative part on the right hand side is non-zero. Finally, we obtain

$$M_u(z,0) = \frac{\left(\frac{R_0(z,u)}{1-zR_1(z,u)}\right)^-(z,g(z),h(z))}{1 - \left(\frac{zR_2(z,u)}{1-zR_1(z,u)}\right)^-(z,g(z),h(z))},$$

where the right hand side depends on z, g and h and has non-negative coefficients. Therefore, it immediately follows that $M_u(z,0)$ has a square-root singularity at z_0.

Now we can use the equation (18) to deduce that $M(z,0)$ has *at most* a square-root-singularity at z_0 – it might be that the singularity cancels. However, by considering the original catalytic equation (7) for constant $u = 0$ we have

$$M(z,0) = R_0(z,0) + zR_1(z,0)M(z,0) + zR_2(z,0)M_u(z,0) + zR_3(z,0)M_{uu}(z,0)$$

Thus, it follows that $M(z,0)$ has *at least* a square-root-singularity at z_0, as u is not a factor of $R_2(z,0)$. Consequently, $M(z,0)$ has square-root-singularity at z_0. This completes the proof of the lemma. ◀

5 Proof of Theorem 2 (The Non-Linear Case)

In this section we use the following notation. If an expression like $R(z,u,\Delta(z,u),M_1(z),M_0(z))$ is evaluated along $u_1(z)$ (and $\Delta_1(z) = \Delta(z,u_1(z))$), we just write R. If the expression is evaluated along $u_2(z)$, we will write $\overline{R}$. We also assume that (10) holds and that $C(z,u)$ is not a power series in u^2.

The proof of Theorem 2 itself will, again, mostly be concerned about the singularity of $u_1(z)$. We will first show this at z_0, where the determinant of the Jacobian of the system (9) equals 0. By considering just 5 equations we can compute functions $\Delta_1(z,u_1), M_1(z,u_1), M_0(z,u_1), u_2(z,u_1), \Delta_2(z,u_1)$ that are analytic at $z_0, u_1(z_0)$. By substituting these functions into the 6th equation (the curve equation) we finally get a single equation for the unknown function $u_1 = u_1(z)$. Next we prove that $u_1(z)$ has a square root singularity at z_0 (provided that (10) holds) and that the functions $M_1(z), M_0(z), \Delta_2(z), u_2(z)$ have at most a $3/2$ singularity. Finally we will confirm the $3/2$ singularity analogously to the linear case.

At several points, the derivative of equation (7) with respect to z plays a crucial role. Along $u_{1,2}(z)$ the terms with factor $\partial_z\Delta(z,u)$ cancel again, and we are left with the system

$$\begin{pmatrix} 1-R_{y_2} & u-R_{y_1} \\ \overline{1-R_{y_2}} & \overline{u-R_{y_1}} \end{pmatrix} \begin{pmatrix} M_0'(z) \\ M_1'(z) \end{pmatrix} = \begin{pmatrix} R_z \\ \overline{R_z} \end{pmatrix} \tag{19}$$

We will denote the matrix on the left hand side by A and use it in particular to prove that the critical exponent of $M_1(z)$ and $M_0(z)$ is $3/2$.

The matrix A appears also right in the first step of the proof. We consider the Jacobian matrix of system (9)

$$\begin{pmatrix} A & 0 & 0 \\ C_1 & B_1 & 0 \\ C_2 & 0 & B_2 \end{pmatrix}$$

where A, C_1, C_2, B_1, B_2 are 2×2 matrices such that its determinant decomposes into three factors. These are the determinants of the submatrices

$$A, \quad B_1 = \begin{pmatrix} 2u-R_{uy_0} & -R_{y_0y_0} \\ 2\Delta - R_{uu} & 2u-R_{uy_0} \end{pmatrix}, \quad B_2 = \begin{pmatrix} \overline{2u-R_{uy_0}} & \overline{-R_{y_0y_0}} \\ \overline{2\Delta-R_{uu}} & \overline{2u-R_{uy_0}} \end{pmatrix}$$

and we denote them by

$$D_0 = \det A, \quad D_1 = \det B_1, \quad D_2 = \det B_2.$$

In a first step, we show that D_0 is never 0 and that the submatrix B_2 which corresponds to the equations for $u_2(z)$ and $\Delta_2(z)$ is invertible, if $D_1 = 0$. Further note that it is fairly obvious from the curve equation that the smallest positive z_0 where $u_1(z)$ is singular is bounded by the convergence radius of $M_0(z)$ and $M_1(z)$ and that $u_2(z)$ will be regular for all $0 < z < z_0$.

▶ **Lemma 9.** *Let z_0 be the smallest positive z, where $u_1(z)$ is singular. Then, the determinants D_0, D_1, D_2 evaluated at z_0 satisfy $D_0 \neq 0$, $D_1 = 0$, and $D_2 \neq 0$.*

Proof. First, we consider equation (19) and note that $R_{y_2} = Q_{\alpha_0}$ and $R_{y_1} = uQ_{\alpha_0} + Q_{\alpha_1}$. Since u_1, u_2 fulfill the curve equation, $Q_{u\alpha_0} \neq 0$ and $u_1 > u_2$, we know that

$$0 < 1 - R_{y_2} < 1 - \overline{R_{y_2}}.$$

Assuming that $D_0 = 0$, it would also have to hold that $0 < u - R_{y_1} < \overline{u - R_{y_1}}$. But $M_1(z)$ and $M_2(z)$ have non-negative coefficients. Hence, (19) would imply that $R_z < \overline{R_z}$ which is certainly wrong. Next in order to prove that D_1 and D_2 cannot be both 0 we consider the curve equation first. Its partial derivative with respect to u equals

$$2u = R_{uy_0} + R_{y_0y_0} \cdot \partial_u \Delta$$

If we differentiate equation (7) twice with respect to u, we can see that the terms with factor $\partial_u^2 \Delta(z,u)$ add up again to the curve equation.

$$\left(u^2 - R_{y_0}\right) \partial_u^2 \Delta + 2\left(2u - R_{uy_0}\right) \partial_u \Delta + 2\Delta = R_{uu} + R_{y_0y_0} \left(\partial_u \Delta\right)^2 \tag{20}$$

Hence, we may compute $\partial_u \Delta$ along $u_1(z)$ as

$$\partial_u \Delta = \frac{2u - R_{uy_0} \pm \sqrt{D_1}}{R_{y_0y_0}} \tag{21}$$

and along $u_2(z)$ analogously. That is, along $u_1(z)$ and $u_2(z)$ it holds that

$$C_u(z, u_i(z)) = 2u_i(z) \pm \sqrt{D_i}.$$

But if $D_1 = D_2 = 0$, then we could repeat the calculations at the end of the proof of Lemma 7 and show that $C_u^+ - C^- = 0$ which is impossible. ◀

Equation (21) concerning the partial derivative of $\Delta(z,u)$ further tells us that $2u > R_{uy_0}$ since $\Delta(z,u)$ has non-negative coefficients and $u_1(z) > 0$ for $z > 0$. This means that the following submatrix of the Jacobian of system (9) is invertible and its inverse equals

$$\begin{pmatrix} A & 0 & 0 \\ C_1 & 2u - R_{uy_0} & 0 \\ C_2 & 0 & B_2 \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 & 0 \\ D_1 & (2u - R_{uy_0})^{-1} & 0 \\ -B_2^{-1} C_2 A^{-1} & 0 & B_2^{-1} \end{pmatrix}$$

where C_1, C_2 are generally non-zero matrices that contain the partial derivatives with respect to M_0 and M_1 of the third, fifth and sixth equation of system (9) respectively and D_1 can be computed appropriately. The implicit function theorem yields analytic functions $M_0(z,u_1), M_1(z,u_1), u_2(z,u_1), \Delta_1(z,u_1)$, and $\Delta_2(z,u_2)$ which have partial derivatives with respect to u that are equal to

$$\partial_{u_1} \begin{pmatrix} M_0(z,u_1) \\ M_1(z,u_1) \\ \Delta_1(z,u_1) \\ u_2(z,u_1) \\ \Delta_2(z,u_1) \end{pmatrix} = - \begin{pmatrix} A^{-1} & 0 & 0 \\ D_1 & (2u - R_{uy_0})^{-1} & 0 \\ -B_2^{-1} C_2 A^{-1} & 0 & B_2^{-1} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 2\Delta - R_{uu} \\ 0 \\ 0 \end{pmatrix} \tag{22}$$

Hence, depending on the critical exponent of $u_1(z)$ in its singular expansion at z_0, we can already state that $M_0(z), M_1(z)$ have at most 2α-singularities. The following lemma proves $\alpha = 1/2$ under the vital condition (10).

▶ **Lemma 10.** *Let $u_{1,2}(z) = g(z) \pm \sqrt{h}$ be the two solutions to the curve equation, z_0 be a the smallest positive z where $u_1(z)$ is singular. If (10) holds at $(z,u) = (z_0, u_1(z_0))$ then $u_1(z)$ has a square root singularity at z_0.*

Proof. First we prove the case where a square root singularity appears. By the computations above we have analytic functions $M_0(z,u_1)$, $M_1(z,u_1)$, and $\Delta_1(z,u_1)$ that we can plug into our equation $u^2 = R_{y_0}(z,u,\Delta,M_1,M_0)$. Now we want to show that for $u_0 = \lim_{z\nearrow z_0} u(z)$, it holds that

$$\begin{aligned}
0 &= u_0^2 - R_{y_0}(z_0,u_0,\Delta_1(z_0,u_0),M_1(z_0,u_0),M_0(z_0,u_0))\big|_{z=z_0,u_1=u_0},\\
0 &= 2u_0 - \partial_{u_1} R_{y_0}(z,u_1,\Delta_1(z,u_1),M_1(z,u_1),M_0(z,u_1))\big|_{z=z_0,u_1=u_0},\\
0 &\neq 2 - \partial_{u_1}^2 R_{y_0}(z,u_1,\Delta_1(z,u_1),M_1(z,u_1),M_0(z,u_1))\big|_{z=z_0,u_1=u_0},\\
0 &\neq -\partial_z R_{y_0}(z,u_1,\Delta_1(z,u_1),M_1(z,u_1),M_0(z,u_1))\big|_{z=z_0,u_1=u_0}.
\end{aligned}$$

The square root singularity of $u_1(z)$ then follows by standard arguments. So, any of the following computations are evaluated at (z_0,u_0).

We already computed that the first partial derivative with respect to u of all plugged in functions is 0 except for $\partial_{u_1}\Delta_1(z,u_1)$. Hence, if we consider the second equation which we want to prove and use the fact that $D_1 = 0$ then

$$\partial_{u_1}\Delta_1 = -\frac{2\Delta_1 - R_{uu}}{2u_1 - R_{uy_0}} = \frac{2u_1 - R_{uy_0}}{R_{y_0y_0}} = \partial_u \Delta$$

and further,

$$\begin{aligned}
&2u_1 - R_{uy_0} - R_{y_0y_0}\partial_{u_1}\Delta_1(z,u_1) - R_{y_0y_2}\partial_{u_1}M_0(z,u_1) - R_{y_0y_1}\partial_{u_1}M_1(z,u_1)\\
&\quad = 2u_1 - R_{uy_0} + R_{y_0y_0}\frac{2\Delta_1 - R_{uu}}{2u_1 - R_{uy_0}} = 0.
\end{aligned}$$

For the third equation, we compute the second partial derivatives of $M_1(z,u_1), M_0(z,u_1)$ and $\Delta_1(z,u_1)$ analogously to our computations above. If we only consider the submatrix for the relevant derivatives, we obtain that

$$\begin{aligned}
\partial_{u_1}^2\begin{pmatrix} M_0\\ M_1\\ \Delta_1\end{pmatrix} &= \begin{pmatrix} A^{-1} & 0\\ D_1 & (2u_1 - R_{uy_0})^{-1}\end{pmatrix}\\
&\quad\cdot\begin{pmatrix} R_{uu} + (2R_{uy_0} - 4u_1)(\partial_{u_1}\Delta_1) + R_{y_0y_0}(\partial_{u_1}\Delta_1)^2 - 2\Delta_1\\ 0\\ R_{uuu} + 2R_{uuy_0}\partial_{u_1}\Delta_1 + R_{uy_0y_0}(\partial_{u_1}\Delta_1)^2 - 4\partial_{u_1}\Delta_1\end{pmatrix}
\end{aligned}$$

Now by dividing the first entry in the right vector above by $R_{y_0y_0}$, it is easy to see that at (z_0,u_0) it is equal to 0. That is,

$$\frac{R_{uu} - 2\Delta_1}{2u - R_{uy_0}}\frac{2u - R_{uy_0}}{R_{y_0y_0}} - 2\frac{2u - R_{uy_0}}{R_{y_0y_0}}(\partial_{u_1}\Delta_1) + (\partial_{u_1}\Delta_1)^2 = 0$$

Consequently $(\partial_{u_1}^2 M_0)(z_0,u_0) = (\partial_{u_1}^2 M_1)(z_0,u_0) = 0$ and

$$(\partial_{u_1}^2\Delta_1)(z_0,u_0) = \left(\frac{R_{uuu} + 2R_{uuy_0}\partial_{u_1}\Delta_1 + R_{uy_0y_0}(\partial_{u_1}\Delta_1)^2 - 4\partial_{u_1}\Delta_1}{2u_1 - R_{uy_0}}\right)$$

The expression from the third equation is thus equal to

$$2 - R_{uuy_0} - 2R_{uy_0y_0}(\partial_{u_1}\Delta_1) - R_{y_0y_0y_0}(\partial_{u_1}\Delta_1)^2 - R_{y_0y_0}(\partial_{u_1}^2\Delta_1).$$

If we plug in the expression for $\partial_{u_1}^2\Delta_1$ and multiply the equation by $\partial_{u_1}\Delta_1$, we obtain

$$6(\partial_{u_1}\Delta_1) - R_{uuu} - 3R_{uuy_0}\partial_{u_1}\Delta_1 - 3R_{uy_0y_0}(\partial_{u_1}\Delta_1)^2 - R_{y_0y_0y_0}(\partial_{u_1}\Delta_1)^3$$

which is non-zero by assumption.

What is left to show is that the derivative with respect to z is non-zero. In a first step we compute

$$\partial_z \begin{pmatrix} M_0(z,u_1) \\ M_1(z,u_1) \\ \Delta_1(z,u_1) \end{pmatrix} = \begin{pmatrix} A^{-1} & 0 \\ D_1 & (2u - R_{uy_0})^{-1} \end{pmatrix} \begin{pmatrix} R_z \\ \overline{R_z} \\ R_{zy_0} \end{pmatrix} \tag{23}$$

Note that $A^{-1}(R_z, \overline{R_z})^T = (M_0'(z), M_1'(z))^T > 0$ by (19) and the fact that $M_0(z), M_1(z)$ have non-negative coefficients. Further we have to compute $D_1 = (d_{11}, d_{12})$ which equal

$$\begin{aligned} d_{11} &= \frac{1}{(2u - R_{uy_0})\det A}\Big(R_{uy_2}\overline{(u - R_{y_1})} + (1 - R_{uy_1})\overline{(1 - R_{y_2})}\Big) \\ d_{12} &= \frac{-1}{(2u - R_{uy_0})\det A}\Big(R_{uy_2}(u - R_{y_1}) + (1 - R_{uy_1})(1 - R_{y_2})\Big). \end{aligned}$$

So the right hand side of the fourth inequality that we want to prove is

$$\begin{aligned} &R_{zy_0} + R_{y_0y_0}\partial_z\Delta_1(z,u_1) + R_{y_0y_2}\partial_z M_0(z,u_1) + R_{y_0y_1}\partial_z M_1(z,u_1) \\ &\quad = R_{zy_0} + R_{y_0y_2}M_0'(z) + R_{y_0y_1}M_1'(z) \\ &\qquad + \frac{R_{y_0y_0}}{(2u - R_{uy_0})\det A}\left(R_{uy_2}\overline{(u - R_{y_1})} + (1 - R_{uy_1})\overline{(1 - R_{y_2})}\right) R_z \\ &\qquad - \frac{R_{y_0y_0}}{(2u - R_{uy_0})\det A}\Big(R_{uy_2}(u - R_{y_1}) + (1 - R_{uy_1})(1 - R_{y_2})\Big)\overline{R_z} \\ &\quad = R_{zy_0} + R_{y_0y_2}M_0'(z) + R_{y_0y_1}M_1'(z) \\ &\qquad + \frac{R_{y_0y_0}}{(2u - R_{uy_0})}\left(R_{uy_2}M_0'(z) - (1 - R_{uy_1})M_1'(z)\right). \end{aligned} \tag{24}$$

Now let us do a similar trick as in the computation of $\partial_u\Delta(z,u)$. We consider

$$\begin{aligned} &\partial_z\partial_u\left(u^2\Delta(z,u) + uM_1(z) + M_0(z) - R(z,u,\Delta(z,u),M_1(z),M_0(z))\right) \\ &\quad = (2u - R_{uy_0} - R_{y_0y_0}\partial_u\Delta(z,u))\,\partial_z\Delta(z,u) + \left(u^2 - R_{y_0}\right)\partial_u\partial_z\Delta(z,u) \\ &\qquad + M_1'(z) - R_{zu} - R_{zy_0}\partial_u\Delta(z,u) - R_{uy_1}M_1'(z) - R_{uy_2}M_0'(z) \\ &\qquad - R_{y_0y_1}\partial_u\Delta(z,u)M_1'(z) - R_{y_0y_2}\partial_u\Delta(z,u)M_0'(z) \\ &\quad = (1 - R_{uy_1} - R_{y_0y_1}\partial_u\Delta(z,u))M_1'(z) - (R_{uy_2} + R_{y_0y_2}\partial_u\Delta(z,u))M_0'(z) \\ &\qquad - R_{zu} - R_{zy_0}\partial_u\Delta(z,u) = 0 \end{aligned} \tag{25}$$

The terms with factor $\partial_u\partial_z\Delta(z,u)$ add up to the curve equation and cancel, the ones with factor $\partial_z\Delta(z,u)$ add up to 0 since by (20)

$$\left(2u - R_{uy_0} - \frac{1}{2}R_{y_0y_0}\partial_u\Delta\right) = \frac{R_{uu} - 2\Delta}{2\,\partial_u\Delta} = \frac{1}{2}R_{y_0y_0}\partial_u\Delta$$

Now if we multiply (24) with $\partial_u\Delta(z,u) = \partial_{u_1}\Delta_1(z,u_1)$ we can see by our computation of (25) that it is equal to $R_{zu} > 0$ which was left to show. ◀

We can determine the singularities of $M_0(z)$ and $M_1(z)$ analogously to the linear case. Obviously if $u_1(z)$ has a square root singularity both of them can have at most 3/2-singularities by our computation of $\partial_{u_1} M_0(z,u_1)$ and $\partial_{u_1} M_1(z,u_1)$.
The negative part of the equations of (19) gives a positive equation for $M_1'(z)$

$$M_1'(z) = \partial_z \left(R(z,u,\Delta(z,u),M_1(z),M_0)\right)^-$$

where the right hand side does indeed confirm this 3/2 singularity. Similarly by the positive equation $M_0(z) = zQ(z,0,M_0(z),M_1(z),M_2(z))$ the 3/2 singularity of $M_0(z)$ is proved as well.

Finally, we comment on the case where the condition (10) is not satisfied. In this case, the third equation that we stated in the beginning is satisfied and one may compute analogously to above that $\partial_{u_1}^3 R_{y_0}(z,u_1,\Delta_1(z,u_1),M_1(z,u_1),M_0(z,u_1)) \neq 0$. By the Weierstrass preparation theorem $u_1(z)$ therefore satisfies a cubic equation

$$(u_1(z)-u_0)^3 + a_2(z)(u_1(z)-u_0)^2 + a_1(z)(u_1(z)-u_0) + a_0(z) = 0$$

where $a_i(z), i = 0,1,2$ are analytic functions at z_0 with $a_i(z_0) = 0$ and, since the fourth equation that we stated in the beginning is satisfied, $a_0(z) = (z-z_0)b(z)$ with $b(z_0) \neq 0$. By considering the critical exponents of each of the summands it follows that $u_1(z)$ has a 1/3-singularity, which implies then that the critical exponent of $M_0(z)$ and $M_1(z)$ is 4/3.

6 Examples

In this section, we will illustrate our generic computations in the proof of Theorems 1 and 2 on the examples given in Section 2.

▶ **Example 11** (Example 3 continued)**.** For one-dimensional non-negative lattice paths where we allow steps of the form ± 1 and ± 2 we obtained the functional equation

$$E(z,u) = 1 + z(u+u^2)E(z,u) + z\frac{E(z,u)-E(z,0)}{u} + z\frac{E(z,u)-E(z,0)-uE_v(u,0)}{u^2}. \quad (26)$$

We know that the curve equation

$$u^2 = z(1+u)u^3 + zu + z.$$

has two solutions $u_1(z)$, $u_2(z)$ with $u_1(0) = u_2(0) = 0$ and $u_1(z)$ is singular at $z_0 > 0$. The common zeros (z_0,u_0) of this equation and its partial derivative with respect to u are

$$\left\{(0,0), \left(\frac{1}{4},1\right), \left(-\frac{4}{9}, -\frac{1-\sqrt{15}i}{4}\right), \left(-\frac{4}{9}, -\frac{1+\sqrt{15}i}{4}\right)\right\}$$

Hence, it follows that $z_0 = \frac{1}{4}$ and $u_1(z_0) = 1$. Furthermore, the local expansion of $u_1(z)$ at $z = z_0$ is given by

$$u_1(z) = 1 - \sqrt{8}\sqrt{1-4z} + \cdots.$$

Next we consider the system of equations for $g(z)$ and $h(z)$:

$$\begin{aligned} g^2 + h &= z(g^4 + h^2 + g^3 + 3(2g^2+g)h + g + 1), \\ 2g &= z((4g+1)h + (4g^3+3g^2+1)). \end{aligned}$$

At $z_0 = \frac{1}{4}$ there are only finitely many solutions, however, the only ones with $g_0 + \sqrt{h_0} = 1$ are $g_0 = (\sqrt{5}-1)/4$ and $h_0 = (15-5\sqrt{5})/8$. Consequently we have

$$u_2(z_0) = g_0 - \sqrt{h_0} = \frac{\sqrt{5}-3}{2}.$$

Finally we use the linear system (15)–(16) to obtain the local expansion for $M(z,0)$ and $M_u(z,0)$:

$$M(z,0) = (6-2\sqrt{5}) - \frac{101\sqrt{2}-45\sqrt{10}}{19}\sqrt{1-4z} + \cdots,$$
$$M_u(z,0) = (4\sqrt{5}-8) - \frac{28\sqrt{10}-62\sqrt{2}}{19}\sqrt{1-4z} + \cdots.$$

Since there are no periodicities this implies that

$$[z^n]E(z,0) \sim \frac{101\sqrt{2}-45\sqrt{10}}{38\sqrt{\pi}}\, n^{-3/2}\, 4^n$$

▶ **Example 12** (Example 4 continued)**.** The functional equation for 3-Constellations can be transformed to the equation

$$C(z,v) = 1 + z(v+1)C(z,v)^3 + z(v+1)(2C(z,v)+C(z,0))\frac{C(z,v)-C(z,v)}{v} + z(v+1)\frac{C(z,v)-C(z,0)-vC_v(z,0)}{v^2}$$

by substituting $v = u-1$. The equations for the unknowns $u_i(z), d_i = \Delta^{(2)}C(z,u_i(z)), i = 1,2$ and $m_1(z) = C_v(z,u_i(z)), m_0 = C(z,u_i(z))$ are given by

$$\begin{aligned}
u_i^2 d_i + u_i m_1 + m_0 =& z(u_i+1)((u_i^2 d_i + u_i m_1 + m_0)^3) \\
&+ z(u_i+1)((2u_i^2 d_i + 2u_i m_1 + 3m_0 + 3)(u_i d_i + m_1) + d_i) \\
u_i^2 =& z(u_i+1)(3(u_i^2 d_i + u_i m_1 + m_0 + 1)^2 u_i^2 + 2(u_i d_i + m_1)u_i^2) \\
&+ z(u_i+1)((2u_i^2 d_i + 2u_i m_1 + 3m_0 + 3)u_i + 1), \\
2u_i d_i + m_1 =& z(u_i+1)(3(u_i^2 d_i + u_i m_1 + m_0 + 1)^2(2u_i d_i + m_1) + 2(2u_i d_i + m_1)(u_i d_i + m_1) \\
&+ z(u_i+1)(2u_i^2 d_i + 2u_i m_1 + 3m_0 + 3)d_i) + ((u_i^2 d_i + u_i m_1 + m_0 + 1)^3) \\
&+ z(u_i+1)((2u_i^2 d_i + 2u_i m_1 + 3m_0 + 3)(d_i u_i + m_1) + d_i).
\end{aligned}$$

Numerical computations show, that the smallest positive z_0 where the Jacobian of this system is invertible equals $z_0 \approx 0.0494$. Indeed, the exact value for the singularity is $4/81 = 0,04938...$. The other variables take the approximate values

$$u_1 \approx 0.6867, \quad u_2 \approx -0.1562, \quad d_1 \approx 0.1070, \quad d_2 \approx 0.0433, \quad m_1 \approx 0.1134, \quad m_0 \approx 0.0833.$$

Note that all computations can be worked out although the scheme of Theorem 2 is not strictly satisfied. We need to check the values of the determinants which equal

$$\det A \approx -0.2588, \quad \det B_1 \approx 0, \quad \det B_2 \approx 0.1828.$$

The neccessary condition of Theorem 2 is also satisfied, since the value of the expression equals $T \approx 2.7209$ (after cancellation of a positive factor to simplify computations). Finally we get the asymptotics

$$[z_n]C(z,0) \sim c\, n^{-5/2} \left(\frac{81}{4}\right)^n$$

for $c \approx 0.0731$.

References

1 Andrei Asinowski, Axel Bacher, Cyril Banderier, and Bernhard Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns, the vectorial kernel method, and generating functions for pushdown automata. *Algorithmica*, 82(3):386–428, 2020.

2 Cyril Banderier and Michael Drmota. Formulae and asymptotics for coefficients of algebraic functions. *Comb. Probab. Comput.*, 24(1):1–53, 2015. doi:10.1017/S0963548314000728.

3 Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *Theor. Comput. Sci.*, 281(1-2):37–80, 2002.

4 Mireille Bousquet-Mélou and Arnaud Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *J. Comb. Theory, Ser. B*, 96(5):623–672, 2006. doi:10.1016/j.jctb.2005.12.003.

5 W. G. Brown and W. T. Tutte. On the enumeration of rooted non-separable planar maps. *Can. J. Math.*, 16:572–577, 1964. doi:10.4153/CJM-1964-058-7.

6 Michael Drmota. *Random trees. An interplay between combinatorics and probability*. Wien: Springer, 2009. doi:10.1007/978-3-211-75357-6.

7 Michael Drmota, Marc Noy, and Guan-Ru Yu. Universal singular exponents in catalytic variable equations. *J. Comb. Theory, Ser. A*, 185:33, 2022. Id/No 105522. doi:10.1016/j.jcta.2021.105522.

8 Michael Drmota and Konstantinos Panagiotou. A central limit theorem for the number of degree-k vertices in random maps. *Algorithmica*, 66(4):741–761, 2013.

9 Michael Drmota and Guan-Ru Yu. The Number of Double Triangles in Random Planar Maps. In James Allen Fill and Mark Daniel Ward, editors, *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018)*, volume 110 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 19:1–19:18, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.AofA.2018.19.

10 Philippe Flajolet and Andrew Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990. doi:10.1137/0403019.

11 Helmut Prodinger. The kernel method: a collection of examples. *Sémin. Lothar. Comb.*, 50:b50f, 19, 2003.

12 W. T. Tutte. A census of planar maps. *Can. J. Math.*, 15:249–271, 1963.

Partial Match Queries in Quad-K-d Trees

Amalia Duch
Universitat Politècnica de Catalunya, Barcelona, Spain

Conrado Martínez
Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract

Quad-K-d trees [Bereckzy et al., 2014] are a generalization of several well-known hierarchical K-dimensional data structures. They were introduced to provide a unified framework for the analysis of associative queries and to investigate the trade-offs between the cost of different operations and the memory needs (each node of a quad-K-d tree has arity 2^m for some m, $1 \leq m \leq K$). Indeed, we consider here partial match – one of the fundamental associative queries – for several families of quad-K-d trees including, among others, relaxed K-d trees and quadtrees. In particular, we prove that the expected cost of a random partial match $\hat{P}_n$ that has s out of K specified coordinates in a random quad-K-d tree of size n is $\hat{P}_n \sim \beta \cdot n^\alpha$ where α and β are constants given in terms of K and s as well as additional parameters that characterize the specific family of quad-K-d trees under consideration. Additionally, we derive a precise asymptotic estimate for the main order term of $P_{n,\mathbf{q}}$ – the expected cost of a fixed partial match in a random quad-K-d tree of size n. The techniques and procedures used to derive the mentioned costs extend those already successfully applied to derive analogous results in quadtrees and relaxed K-d trees; our results show that the previous results are just particular cases, and states the validity of the conjecture made in [Duch et al., 2016] to a wider variety of multidimensional data structures.

2012 ACM Subject Classification Theory of computation → Data structures design and analysis; Theory of computation → Design and analysis of algorithms

Keywords and phrases Quadtree, Partial match queries, Associative queries, Multidimensional search, Analysis of algorithms

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.8

Funding This work has been supported by funds from the MOTION Project (Project PID2020-112581GB-C21) of the Spanish Ministery of Science and Innovation MCIN/AEI/10.13039/501100011033.

1 Introduction

Considered as fundamental associative queries, partial match queries have been widely studied in the literature, see for instance [18, 23]. More specifically, given a collection (or file) $\mathcal{F}$ of n records, in which each *record* in $\mathcal{F}$ is an ordered K-tuple of values (the record's attributes or coordinates), a *query* in $\mathcal{F}$ retrieves the records satisfying certain given conditions. If the imposed conditions deal with more than one attribute, then the query is considered *associative*.

In particular, *partial match* (PM hereinafter) queries consist of retrieving from $\mathcal{F}$ all the records with attributes matching some specified attributes of the given query record. Indeed, the analysis of PM queries (either random or fixed) has been carried out in a wide variety of hierarchical multidimensional data structures, see [6, 7, 8, 11, 14, 15, 21] and references therein.

From the point of view of their analysis, it would be of great interest to unify all these results in a comprehensive way. The general framework of quad-K-d trees (introduced in [4]) is an attempt in that direction. A quad-K-d tree is a multidimensional tree in which each node discriminates with respect to some number m, $1 \leq m \leq K$, of coordinates (and thus it

has 2^m subtrees). The number m and the subset of discriminating coordinates (the *type*) is potentially distinct for each node. When all the K coordinates are used to discriminate at each node, the tree is a quadtree [2], in contrast, when exactly one of the K coordinates is used to discriminate at each node, the tree is a K-d tree [1].

In this work we follow the approach of Chern and Hwang [8] to prove (extending the preliminary results in [11]) that the expected cost $\hat{P}_n$ of a random PM query in a random quad-K-d tree of size n is $\hat{P}_n = \beta \cdot n^\alpha +$ lower order terms (l.o.t.), with α and β given in Theorem 3.

We also give (following the steps in [14] and [12]) a precise asymptotic estimate[1] for the main order term of of the expected cost $P_{n,\mathbf{q}} = \nu \cdot f(\mathbf{q}) \cdot n^\alpha +$ l.o.t. of a fixed partial match with query $\mathbf{q}$ in a random quad-K-d tree of size n, where α is the same as for random PM queries. This is formalized in Theorem 6 where we give the explicit form of the constant ν and of $f(\mathbf{q})$.

Our results apply to any family of quad-K-d trees whose nodes have *types* that are independent from one another. This includes, indeed, random relaxed K-d trees and quadtrees.

The paper is organized as follows. In Section 2 we give some preliminaries of quad-K-d trees (Subsection 2.1) and partial matches (Subsection 2.2). We then derive, in Section 3, the expected cost of a random and fixed PM queries in a random quad-K-d tree (Subsections 3.1 and 3.2, respectively). We finish in Section 4 with conclusions and a brief discussion around further work in this research topic.

2 Preliminaries

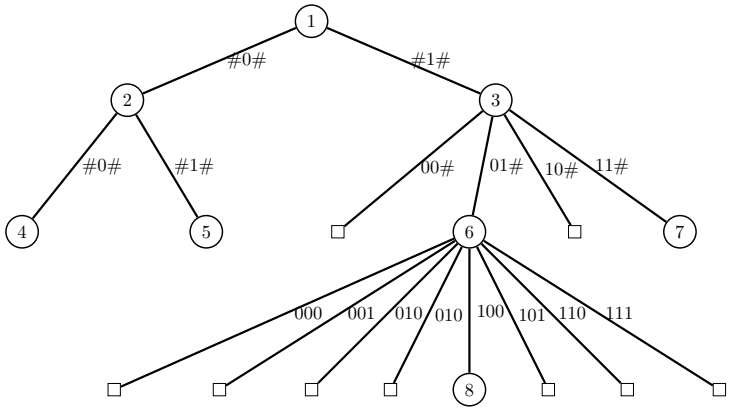
2.1 Quad-K-d Trees

Quad-K-d trees generalize K-dimensional trees [1] and quadtrees [2]. In K-d trees each node $\mathbf{x}$ has a *discriminating* coordinate i, $0 \leq i < K$: all keys $\mathbf{y}$ in the left subtree have their i-th coordinate y_i smaller than x_i. Likewise all keys $\mathbf{z}$ in the the right subtree have their i-th coordinate z_i larger than x_i. Each node thus is associated with a region (*bounding box*) of the domain from which the keys are drawn, and divides that region into two, according to the i-th coordinate being smaller or larger. The choice of which coordinate discriminantes at each node leads to several variants of K-d trees. For quadtrees, each node $\mathbf{x}$ induces a partition of its associated region into 2^K quadrants, and it will have thus 2^K subtrees, one for each quadrant. If we label subtrees with a bitstring $\mathbf{w}$ of length K, the i-th bit of $\mathbf{w}$ indicates if all keys $\mathbf{y}$ in the subtree have $y_i < x_i$ (when $w_i = 0$) or $y_i > x_i$ (when $w_i = 1$).

Quad-K-d trees, being a generalization of both K-dimensional trees and quadtrees, store for each node a subset of m discriminating coordinates, with m ranging from 1 to K, potentially different for each node. We say that such a node is of *t*ype m. Nodes in a K-d tree are all of type 1, nodes in a quadtree are all of type K, and a general quad-K-d tree may contain a mixture of nodes of all possible different types.

The subset of discriminating coordinates of a node is represented by its characteristic function or *coordinate split vector* $\boldsymbol{\delta}$: if $\delta_i = 0$ then the corresponding node does not discriminate with respect to the i-th coordinate, otherwise $\delta_i = 1$ and the node will discriminate

[1] Our result applies under the reasonable but technically hard to establish assumption that the limit of $P_{n,\mathbf{q}}/n^\alpha$ when $n \to \infty$ exists.



Node	Coordinates	Split vector
1	(0.35,0.5,0.17)	010
2	(0.31, 0.45, 0.47)	010
3	(0.29, 0.6, 0.9)	110
4	(0.3, 0.3, 0.3)	001
5	(0.4, 0.47, 0.3)	110
6	(0.2, 0.7, 0.5)	111
7	(0.38, 0.8, 0.89)	011
8	(0.25, 0.68, 0.4)	100

Figure 1 An example of a 3-dimensional quad-K-d tree, omitting some empty subtrees as well as the partition induced by the tree in $[0,1]^3$.

with respect to that coordinate. Thus a node of type m has 2^m subtrees and partitions its associated region of the domain of the keys into 2^m subregions, each subregion associated to a subtree.

Figure 1 shows an example of a 3-dimensional quad-K-d tree. Inside each node appears its label and in the table therein is the 3-dimensional key associated to the node together with its split vector. Next to every edge appears the label (the string $\mathbf{w}$) of the subtree where it points to. When $w_i = 0$ it means that i is a discriminating coordinate ($\delta_i = 1$) and all keys $\mathbf{y}$ in the subtree have $y_i < x_i$. If $w_i = 1$ then i is a discriminating coordinate ($\delta_i = 1$) but now all keys $\mathbf{y}$ in the subtree have $y_i > x_i$. When i is not a discriminating coordinate ($\delta_i = 0$) we indicate so in the subtree/subregion label $\mathbf{w}$ by setting $w_i = \#$.

More formally, a K-dimensional record (or key) is a K-tuple of values $\mathbf{x} = (x_0, \ldots, x_{K-1})$ where each x_i is drawn from a totally ordered domain D_i. The domain $D = D_0 \times \cdots \times D_{K-1}$ is known as the search space and without loss of generality it is usually assumed to be $D = [0,1]^K$. Then, the formal definition of quad-K-d trees is as follows.

▶ **Definition 1** (Bereckzy et al. [3]). *A quad-K-d search tree T of size $n \geq 0$ stores a set of n K-dimensional records, each holding a key $\mathbf{x} = (x_0, \ldots, x_{K-1})$ and a* coordinate split *bitvector $\boldsymbol{\delta} = (\delta_0, \ldots, \delta_{K-1}) \in \{0,1\}^K$. The quad-$K$-d tree T is such that*

- *either it is empty if $n = 0$, or*
- *its root r stores a record with key $\mathbf{x}$, a coordinate split vector $\boldsymbol{\delta}$ that contains exactly m ones (i.e., the root node is of order m), with $1 \leq m \leq K$, and pointers to its 2^m subtrees that store the $n-1$ remaining records as follows: each subtree, let us call it $T_{\mathbf{w}}$, is itself a quad-K-d tree and its associated string $\mathbf{w} = w_0 w_1 \ldots w_{K-1} \in \{0,1,\#\}^K$, is such that for all j, $0 \leq j < K$, $\delta_j = 0 \implies w_j = \#$, and for any key $\mathbf{y} \in T_{\mathbf{w}}$,*
  - *if $\delta_j = 1$ and $w_j = 0$, then $y_j < x_j$*
  - *if $\delta_j = 1$ and $w_j = 1$, then $y_j > x_j$.*

It is worth noting that with this definition, as is the case of binary search trees and other search trees, we do not consider cases in which two or more keys have some identical coordinates; however the definition above can be adapted (and hence the algorithms supported by the data structure) to cover these cases. However, because of our assumptions in the analysis, we can safely disregard this situation since the probability that two records share a coordinate value is 0.

As we have already mentioned, both K-d trees and quadtrees are special cases of quad-K-d trees. In fact, for any quad-K-d tree T of size n, if the split vector $\boldsymbol{\delta}$ associated with every node of T contains all the K coordinates ($\delta_j = 1$ for all j) then T is a quadtree, and if it contains exactly one for every node then T is a relaxed K-d tree [10]. We use here the term *relaxed* K-d trees to stress that the discriminating coordinate of each node is arbitrary, in contrast with standard K-d trees [1], squarish K-d trees [9], and other families of K-d trees where the discriminating coordinate of each node is determined by some fixed rule. Thus relaxed K-d trees are the most general family of K-d trees, of which all other are particular instances. We can have a similar situation with quad-K-d trees: we can define families of quad-K-d trees in which some fixed rule prescribes the types of the nodes and which are the discriminating coordinates of each node. We therefore will use the term *relaxed quad-K-d trees* to emphasise that we consider the most general family of quad-K-d trees, as implied by Definition 1. All over this work, unless otherwise stated, we will refer to relaxed quad-K-d trees as simply quad-K-d trees.

A node holding key $\mathbf{x}$ and split vector $\boldsymbol{\delta}$ in a quad-K-d tree is *of type m* if and only if $\boldsymbol{\delta}$ is of order m, $1 \leq m \leq K$. Every node of type m has 2^m children. A *m-regular* (or m-ary) quad-K-d tree is a quad-K-d tree that has all its internal nodes of type m. Indeed, quadtrees of dimension K are K-regular quad-K-d trees and all variants of K-d trees are 1-regular quad-K-d trees.

The probabilistic analysis of PM queries in quad-K-d trees works under the assumption that the trees under consideration are *random*.

▶ **Definition 2.** *A random relaxed quad-K-d tree of size n is a quad-K-d tree built by n random insertions*[2] *into an initially empty tree, and it additionally satisfies the following two conditions:*

1. *The types of its n nodes are given by n i.i.d. random variables in the set $\{1, \ldots, K\}$. We denote as τ_m the probability that an arbitrary node is of type m.*
2. *For a node of type m, any subset of m coordinates out of K is equally likely to be the set of discriminating coordinates (that is, those for which $\delta_j = 1$). Thus the probability that the discriminating coordinates of a node of type m are $0 \leq i_0 < i_1 < \cdots < i_{m-1} < K$ is $1/\binom{K}{m}$, for any subset $\{i_0, \ldots, i_{m-1}\} \subseteq \{0, \ldots, K-1\}$.*

The previous definition of random quad-K-d trees is equivalent to the conventional random models found in the literature for the particular cases of random quadtrees and random relaxed K-d trees. Several other instances of random quad-K-d trees have been proposed in previous works [4, 11], we refer the reader to these works for a detailed analysis of the space needs and the expected cost of exact searches (IPL) in the different families, relating them to parameters such as the average arity $\overline{d} = \sum_{1 \leq m \leq K} 2^m \tau_m$ and the average order $\overline{m} = \sum_{1 \leq m \leq K} m \tau_m$.

2.2 Partial Match

A PM query is a pair $\langle \mathbf{q}, \mathbf{u} \rangle$, where $\mathbf{q} = (q_0, \ldots, q_{K-1})$ is a K-dimensional key and $\mathbf{u} = (u_0, \ldots, u_{K-1})$ is the *pattern* of the query; each $u_i = S$ (the i-th attributed of the query is *specified*) or $u_i = *$ (the i-th attributed is *unspecified*). Alternatively, we can define a PM

[2] There are several ways to characterize random insertions. The one we will consider here is that every coordinate of the data point to be inserted is independently drawn from some continuous distribution in $[0, 1]$.

query as K-tuple $\mathbf{q} = (q_0, \ldots, q_{K-1})$ where each q_i is a value in the i-th domain D_i or $q_i = *$ to indicate that it is unspecified; we will use most often this second form to represent PM queries.

The PM query $\langle \mathbf{q}, \mathbf{u} \rangle$ is *random* if $\mathbf{q} = (q_0, \ldots, q_{K-1})$ is independently drawn from the same continuous distribution as the data points, otherwise we say that the PM is *fixed*.

The goal of the PM search is to report all data points $\mathbf{x} = (x_0, \ldots, x_{K-1})$ in the tree such that $x_i = q_i$ whenever $q_i \neq *$. The number of specified coordinates will be denoted by s; the interesting cases are when $0 < s < K$. Therefore, to perform a PM search with query $\mathbf{q}$, a quad-K-d tree with root of type m is recursively explored as follows. First, we check whether the root with associated $\mathbf{x}$ matches $\mathbf{q}$ or not, to report it in the former case. Since the node discriminates by m coordinates, let us consider that i of them correspond to specified coordinates in the query ($0 \leq i \leq \min\{s, m\}$). Then, we make recursive calls in all the 2^{m-i} subtrees $T_{\mathbf{w}}$ of the root such that $w_j = \#$ or $q_j = *$, or $w_j = 0$ and $q_j \leq x_j$, or $w_j = 1$ and $q_j > x_j$.

For example, imagine that $K = 5$ and the query is $\mathbf{q} = (0.1, 0.82, *, 0.76, *)$. Moreover, suppose that the root node is of type $m = 3$ with coordinate split vector $\boldsymbol{\delta} = (1, 0, 1, 1, 0)$ and that it contains the key $\mathbf{x} = (0.54, 0.46, 0.39, 0.03, 0.62)$. The 8 subtrees of the root node are labelled as $(0, \#, 0, 0, \#), (0, \#, 0, 1, \#), \ldots, (1, \#, 1, 1, \#)$. It turns out that only $i = 2$ of the discriminating coordinates are specified in the query, the third one is not. Since $0.1 < 0.54$ and $0.76 > 0.03$ we will have to explore the two subtrees $(0, \#, 0, 1, \#)$ and $(0, \#, 1, 1, \#)$, since, even though the third coordinate is a discriminating coordinate of the node, it is not specified in the query: that is, we can't discard the keys of any of these two subtrees.

3 Analysis

The cost of the PM search is measured –as usual in the literature– as the number of nodes visited by the algorithm in the corresponding tree. According to this way of measuring the cost, it is relevant towards our analysis to observe that, except for eventual matches, only matter the relative order of the coordinates of the stored keys.

Indeed, let us call the *rank vector* of a query $\mathbf{q}$ the vector $\mathbf{r}(\mathbf{q}) = (r_0, \ldots, r_{K-1})$ such that $r_i = *$, if $q_i = *$, and r_i is the number of records $\mathbf{x}$ in the collection F such that $x_i \leq q_i$ ($0 \leq r_i \leq n$), if $q_i \neq *$. Then for any two given queries $\mathbf{q}$ and $\mathbf{q}'$ with equal rank vectors $\mathbf{r}(\mathbf{q}) = \mathbf{r}(\mathbf{q}')$ the PM procedure described above will visit exactly the same set of nodes of the tree. In our analysis, we will use rank vectors instead of the queries themselves (as done in [11] and [12]) and consider the cost $\mathcal{P}_{n,\mathbf{r}}$ of a PM query with given rank vector $\mathbf{r}$ in a random quad-K-d tree of size n. We will use $\mathcal{P}_{n,\mathbf{q}}$ for the cost of a PM query with a query $\mathbf{q}$ in a random quad-K-d tree of size n, where each $q_i \in (0, 1)$ or $q_i = *$. From our discussion above $\mathcal{P}_{n,\mathbf{q}} = \mathcal{P}_{n,\mathbf{r}(\mathbf{q})}$. Moreover, because of the symmetries in the model of random quad-K-d trees we can safely assume that the queries are of the form $\mathbf{q} = (q_0, \ldots, q_{s-1}, *, \ldots, *)$ with $q_i \in (0, 1)$ for all i, $0 \leq i < s$, we will abuse of the notation and simply write $\mathbf{q} = (q_0, \ldots, q_{s-1})$ and $\mathbf{r} = (r_0, \ldots, r_{s-1})$, omitting the non-specified coordinates.

3.1 Analysis of Random Partial Match

Let $\hat{P}_n$ denote the expected cost of a random PM query in a random relaxed quad-K-d tree of size n, this corresponds to a query $\langle \mathbf{q}, \mathbf{u} \rangle$ where $\mathbf{q} = (q_0, \ldots, q_{s-1})$ is a random K-dimensional point drawn independently from the same distribution of the data points in the quad-K-d tree.

Here we are going to proceed –as is usual in the literature– obtaining the expected cost P'_n of an "idealised" PM search procedure in which, every time that we recursively invoke it on a subtree, a new random query is generated (but inside the region of the space associated to the particular subtree). The random variable for the cost of this idealised PM search is obviously different from the one that gives the cost of a PM search with a random query; but their respective expected costs coincide: $P'_n = \hat{P}_n$. So, in an abuse of notation, in what follows we will use $\hat{P}_n$ instead of P'_n.

In [11], Roura's continuous master theorem [22] was used to show that the expected cost $\hat{P}_n$ of a random partial match in a random relaxed quad-K-d tree of size n is

$$\hat{P}_n = \Theta(n^\alpha),$$

where α is the unique real solution in the interval $(0,1)$ of the equation

$$\sum_{m=1}^{K} \tau_m \sum_{0\le i\le s} \frac{\binom{m}{i}\binom{K-m}{s-i}}{\binom{K}{s}} \frac{2^m}{(\alpha+1)^{m-i}(\alpha+2)^i} = 1. \tag{1}$$

Here we carry out a more delicate and precise analysis, following the same steps as Chern and Hwang [8] in their analysis of PM in random quadtrees to obtain also the coefficient of the leading term of the cost. We can thus establish the following theorem.

▶ **Theorem 3.** *For any variant of random relaxed quad-K-d trees of size $n \ge 2$, the expected cost $\hat{P}_n$ of a random PM query with s $(0 < s < K)$ specified coordinates, satisfies*

$$\hat{P}_n = \beta_{s,K} n^\alpha + o(n^\alpha),$$

where α is the unique solution in $(0,1)$ of

$$\sum_{m} \tau_m \sum_{i\ge 0} \frac{\binom{K-m}{s-i}\binom{m}{i}}{\binom{K}{s}} \frac{2^m}{(\alpha+1)^{m-i}(\alpha+2)^i} = 1,$$

and

$$\beta_{s,K} = \frac{2^s}{(\alpha_1-1)(1-\alpha_2)\dots(1-\alpha_k)} \frac{1}{\Gamma(\alpha)^{K-s}\Gamma(\alpha+1)^s} \prod_{2\le j\le K} \frac{\Gamma(\alpha-\alpha_j)}{\Gamma(1-\alpha_j)},$$

where $\alpha_1, \alpha_2, \dots \alpha_K$ with $\Re(\alpha_1) \ge \Re(\alpha_2) \ge \dots \ge \Re(\alpha_K)$ are the roots of:

$$\Phi(z) = (z+1)^s z^{K-s} - \sum_{m=1}^{K} 2^m \tau_m \sum_i \frac{\binom{m}{i}\binom{K-m}{s-i}}{\binom{K}{s}} (z+1)^{s-i} z^{K-m-(s-i)},$$

and $\alpha = \alpha_1 - 1$.

Proof. To prove the theorem, we start conditioning on the type m of the root of the random quad-K-d tree. Let $\hat{P}_n^{(m)}$ denote the expected cost of a random PM query, conditional to the root of the quad-K-d tree being of type m; similarly, let $\hat{P}_n^{(i,m)}$ denote the same expected cost now conditional on the root being of type m and that exactly i, $0 \le i \le \min(m,s)$ of the discriminating coordinates of the root are specified in the PM query. Then, since the probability that the root of a random quad-K-d tree is of type m is τ_m we have, for $n > 0$,

$$\begin{aligned}
\hat{P}_n &= \sum_{m=1}^{K} \tau_m \hat{P}_n^{(m)}, \\
\hat{P}_n^{(m)} &= \sum_{0\le i\le m} \mu_{i,m} \hat{P}_n^{(i,m)}, \\
\hat{P}_n^{(i,m)} &= 1 + \sum_{0\le k<n} \pi_{n,k}^{(i,m)} \hat{P}_k,
\end{aligned}$$

where $\mu_{i,m}$ is the probability that exactly i of the m discriminating coordinates are specified (these quantities also depend on s and K), and $\pi_{n,k}^{(i,m)}$ is the average number of recursive calls on root's subtrees of size k when the tree is of size n, its root is of type m and exactly i out of its m discriminating coordinates are specified. It is not difficult to prove that $\mu_{i,m} = \frac{\binom{m}{i}\binom{K-m}{s-i}}{\binom{K}{s}}$ and therefore

$$\hat{P}_n = 1 + \sum_{0\le k<n} \pi_{n,k}\hat{P}_k, \qquad n > 0 \tag{2}$$

where $\pi_{n,k} = \sum_{m=1}^{K} \tau_m \sum_{0\le i\le m} \mu_{i,m}\pi_{n,k}^{(i,m)}$, and $\hat{P}_0 = 0$.

Our next step is to apply the (symmetric) Binomial Transform (see for instance [5, 24] or [20, 5.2.2, p. 136]) to the sequence $\{\hat{P}_n\}_{n\ge0}$, that is,

$$\hat{P}_n^* := \sum_{k=1}^{n} \binom{n}{k} (-1)^k \hat{P}_k, \tag{3}$$

with $\hat{P}_0 = 0$. We can easily prove then the following proposition.

▶ **Proposition 4.** *For $n \ge 2$, the sequence $\hat{P}_n^*$ satisfies the first order recurrence*

$$\hat{P}_n^* - \hat{P}_{n-1}^* = -\mu(n)\hat{P}_{n-1}^*,$$

where $\hat{P}_1^* = -1$ and $\mu(n) := \mu(n,K,s,\tau) = \sum_{m=1}^{K} \tau_m 2^m \sum_{0\le i\le m} \frac{\binom{m}{i}\binom{K-m}{s-i}}{\binom{K}{s}} n^{-m+i}(n+1)^{-i}$.

Proof. As $n \ge 2$, we can replace $\hat{P}_n^*$ and $\hat{P}_{n-1}^*$ by their values as defined in (3); then a few mathematical manipulations yield

$$\begin{aligned}\hat{P}_n^* - \hat{P}_{n-1}^* &= \sum_{k=1}^{n} \binom{n-1}{k-1}(-1)^k \sum_{0\le j<k} \pi_{k,j}\hat{P}_j \\ &= -\sum_{j=0}^{n-1} \hat{P}_j \sum_{k=j}^{n-1} \binom{n-1}{k}(-1)^k \pi_{k+1,j}.\end{aligned}$$

Now, the inner sum of previous equation can be expanded as follows

$$\begin{aligned}\sum_{k=j}^{n-1} \binom{n-1}{k}(-1)^k \pi_{k+1,j} &= \sum_{m-1}^{K} \tau_m 2^m \sum_{0\le i\le m} \frac{\binom{m}{i}\binom{K-m}{s-i}}{\binom{K}{s}} \sum_{k=j}^{n-1} \binom{n-1}{k}(-1)^k \pi_{k+1,j}^{(i,m)} \\ &= \sum_{m=1}^{K} \tau_m 2^m \sum_{0\le i\le m} \frac{\binom{m}{i}\binom{K-m}{s-i}}{\binom{K}{s}} \binom{n-1}{j}(-1)^j n^{-m+i}(n+1)^{-i},\end{aligned}$$

because

$$\sum_{k=j}^{n-1} \binom{n-1}{k}(-1)^k \pi_{k+1,j}^{(i,m)} = \binom{n-1}{j}(-1)^j (n+1)^{-i} n^{-m+i}. \tag{4}$$

To prove this we follow the analogous derivation in [8]; we give it in detail in Appendix A.

Finally, using the definition of $\mu(n)$ given in the statement of the proposition, it follows that for $n \ge 2$,

$$\hat{P}_n^* - \hat{P}_{n-1}^* = -\sum_{j=0}^{n-1} \hat{P}_j \binom{n-1}{j}(-1)^j \mu(n) = -\mu(n)\hat{P}_{n-1}^*. \qquad ◀$$

▶ **Proposition 5.** *For $n \geq 1$, $\hat{P}_n$ satisfies*

$$\hat{P}_n = \sum_{j=1}^{n} \binom{n}{j} (-1)^{j+1} \frac{2^s}{j!^{K-s}(j+1)!^s} (2-\alpha_1)^{\overline{j-1}} (2-\alpha_2)^{\overline{j-1}} \cdots (2-\alpha_K)^{\overline{j-1}},$$

where $x^{\overline{r}} = x \cdot (x+1) \cdot (x+r-1)$ denotes the r-th raising factorial of x [19].

Proof. From Proposition 4, it follows that, for $n \geq 2$,

$$\hat{P}^*_n = (1-\mu(n))\hat{P}^*_{n-1},$$

and iterating,

$$\hat{P}^*_n = (-1)^n \prod_{j=2}^{n} (1-\mu(j))\hat{P}_1 = -\prod_{j=2}^{n} \frac{\Phi(j)}{j^{K-s}(j+1)^s},$$

where $\Phi(z)$ is given in the statement of the theorem and since $\Phi(n)(n+1)^{-s}n^{-(K-s)} = 1-\mu(n)$. Let $\alpha_1, \ldots, \alpha_K$ be the K roots of $\Phi(z)$, then we can write

$$\Phi(z) = (z-\alpha_1)\cdots(z-\alpha_K)$$

and

$$\begin{aligned}\hat{P}^*_n &= -\prod_{j=2}^{n} \frac{(j-\alpha_1)\cdots(j-\alpha_K)}{j^{K-s}(j+1)^s} = -\left(\prod_{j=2}^{n} \frac{j-\alpha_1}{j^{K-s}(j+1)^s}\right)\cdots\left(\prod_{j=2}^{n} \frac{j-\alpha_K}{j^{K-s}(j+1)^s}\right) \\ &= -\frac{2^s}{n!^{K-s}(n+1)!^s}(2-\alpha_1)^{\overline{n-1}}(2-\alpha_2)^{\overline{n-1}}\cdots(2-\alpha_K)^{\overline{n-1}},\end{aligned}$$

Since the binomial transform is an involution, we have

$$\hat{P}_n = \sum_{j=1}^{n} \binom{n}{j} (-1)^j \hat{P}^*_j,$$

and the statement of the proposition follows. ◀

In order to complete the proof of the theorem we obtain the asymptotic behaviour of $\hat{P}_n$ from Proposition 5 using Nørlund-Rice's integrals (see for instance [16, 17]), from where we get the asymptotic estimate given in the theorem. In particular, $\hat{P}_n \sim \beta \cdot n^\alpha$, with the values of α and β given in the statement of the theorem. ◀

Notice that setting $\tau_K = 1$ and $\tau_m = 0$ for $m < K$ in Equation (1) we obtain the indicial equation $\Phi(z) = z^{K-s}(z+1)^s - 2^K = 0$ of quadtrees [8]. Likewise, with $\tau_1 = 1$ and $\tau_m = 0$ for $m > 1$ we obtain

$$\Phi(z) = z^{K-s-1}(z+1)^{s-1}\left(z(z+1) - 2\left[\frac{K-s}{K}(z+1) + \frac{s}{K}z\right]\right),$$

whose roots are $z = 0$, $z = 1$ and the roots of $z(z+1) - 2\frac{K-s}{K}(z+1) + \frac{s}{K}z = 0$, the indicial equation giving the exponent α for relaxed K-d trees (see [10, 21]).

3.2 Analysis of Fixed Partial Match

In this subsection we consider the expected cost of a PM search with rank vector $\mathbf{r}$, or equivalently with a query $\mathbf{q}$, if every coordinate of each data point is independent and uniformly distributed in $(0,1)$, see [11] for a discussion about the two "models" and how results for one translate into results for the other. We shall start stating the main theorem of this section, also one of the major contributions of this extended abstract, and devote the rest of the subsection to schematise its proof.

▶ **Theorem 6.** *Let $\mathbf{r} = (r_0, r_1, \ldots, r_{K-1})$ be a rank vector, where $r_i \in [0..n] \cup \{*\}$, $0 \le i < K$, and such that exactly s of the ranks are specified, that is, we have $r_i \neq *$, for s ranks, $0 < s < K$. Let $z_i = \lim_{n\to\infty} r_i/n$ if $r_i \neq *$ and suppose $z_i \in (0,1)$ for all i such that $r_i \neq *$, $0 \le i < K$.*

Then, for any variant of random quad-K-d trees the expected cost $P_{n,\mathbf{r}}$ of a fixed PM query with rank vector $\mathbf{r}$ in a random quad-K-d tree of size n, is

$$P_{n,\mathbf{r}} = \nu_{s,K} \left(\prod_{r_i \neq *} z_i(1-z_i) \right)^{\alpha/2} n^{\alpha} + l.o.t.,$$

where

$$\nu_{s,K} = \beta_{s,K} \frac{\Gamma^s(\alpha+2)}{\Gamma^{2s}(\alpha/2+1)},$$

and α and $\beta_{s,K}$ are the same as in Theorem 3, provided that

$$\lim_{n\to\infty} \frac{P_{n,\mathbf{r}}}{n^{\alpha}} = f(z_0, \ldots, z_{s-1})$$

exists.

Proof. Let $\mathcal{P}_{n,\mathbf{r}}$ be the cost (number of visited nodes) of a partial match search in a random quad-K-d tree of size n where the query has fixed rank vector $\mathbf{r}$. Our goal is to find the main order asymptotics of $P_{n,\mathbf{r}} = \mathbb{E}\{\mathcal{P}_{n,\mathbf{r}}\}$; more specifically, our goal is to find an explicit function $f(z_0, \ldots, z_{s-1})$ and to show that for $\mathbf{r} = (r_0, \ldots, r_{s-1}, *, \ldots, *)$, if $\lim_{n\to\infty} r_i/n = z_i$, with $z_i \in (0,1)$ for all i, $0 \le i < s$ and

$$\lim_{n\to\infty} \frac{P_{n,\mathbf{r}}}{n^{\alpha}} = f(z_0, \ldots, z_{s-1}),$$

exists, with α the exponent of n in the expected cost of a random partial match search (see previous subsection), then the function f is as stated in the theorem.

The first step in our analysis is to setup a recurrence for $P_{n,\mathbf{r}}$. As in our analysis of random PMs, we condition first on the type m of the root node, so we can write

$$P_{n,\mathbf{r}} = \sum_{m=1}^{K} \tau_m P_{n,\mathbf{r}}^{(m)},$$

where $P_{n,\mathbf{r}}^{(m)} = \mathbb{E}\{\mathcal{P}_{n,\mathbf{r}} \,|\, \text{root of type } m\}$. Now, we need to condition on which coordinates are specified by the query and used to discriminate at the root node. Then, we have

$$P_{n,\mathbf{r}}^{(m)} = \sum_{i=0}^{m} \frac{\binom{K-s}{m-i}}{\binom{K}{m}} \left(1 + \sum_{\substack{\mathbf{w}\in(0+1+\#)^K \\ |\mathbf{w}|_{0,1}=i}} Q_{n,\mathbf{r},\mathbf{w}} \right)$$

where $Q_{n,\mathbf{r},\mathbf{w}}$ is the contribution to the total fixed PM cost of the recursive call on subtree $T_{\mathbf{w}}$ when the root of the quad-K-d tree of size n is of type $m = |\mathbf{w}|$ and it discriminates with respect to i of the coordinates that are specified in $\mathbf{r}$ (namely, the root discriminates with respect to those coordinates for which the bitstring $\mathbf{w}$ indexing $T_{\mathbf{w}}$ is 0 or 1; thus $|\mathbf{w}|_{0,1} = i$ and the remaining bits are '#'). The factor $\binom{K-s}{m-i}$ in the formula above takes care of the $m-i$ unspecified coordinates, as it does not matter which $m-i$ coordinates of the root are discriminating coordinates if the query has them unspecified. Moreover, the first summation actually runs from $i = 0$ to $i = \min(s, m)$ since $\binom{K-s}{m-i} = 0$ whenever $i > m$.

The next step is to express $Q_{n,\mathbf{r},\mathbf{w}}$ in terms of $P_{n',\mathbf{r}'}$, where n' is the size of $T_{\mathbf{w}}$ and it will depend on n, $\mathbf{w}$ and the rank vector $\mathbf{j}$ of the root of the quad-K-d tree. Likewise, the rank vector $\mathbf{r}'$ of the recursive call to PM inside $T_{\mathbf{w}}$ will depend on n, $\mathbf{w}$, $\mathbf{r}$ and $\mathbf{j}$. We might thus write

$$\frac{1}{(n+1)^K} \sum_{\mathbf{j}} \sum_{n'} \sum_{\mathbf{r}'} \pi(n, n', \mathbf{r}, \mathbf{r}') P_{n',\mathbf{r}'},$$

where $\pi(n, n', \mathbf{r}, \mathbf{r}')$ is the probability that the rank vector of the PM search in $T_{\mathbf{w}}$ is $\mathbf{r}'$ and the size of $T_{\mathbf{w}}$ is n'.

The full recurrence is quite involved, but the analysis follows closely that of quadtrees [12] with regards to the discriminating coordinates ($w_k \in \{0, 1\}$) and that of relaxed K-d trees [14] when the coordinates are not discriminating ($w_k = \#$). This is because the PM behaves as if the query were unspecified for that particular coordinate (as if $r_k = *$, despite r_k is actually a value in $[0, n]$). When we pass to the limit both sides of the recurrence, and under the assumption that

$$f(z_0, \ldots, z_{s-1}) = \lim_{n\to\infty} \frac{P_{n,\mathbf{r}}}{n^\alpha}$$

exists, with $z_i = \lim_{n\to\infty} r_i/n$ and $0 < z_i < 1$, the probabilities $\pi(n, n', \mathbf{r}, \mathbf{r}')$ become highly concentrated around the expected values of $\mathbf{r}'$ and thus it becomes considerably simplified. Summations transform into integrals and the recurrence leads to an integral equation for $f(z_0, \ldots, z_{s-1})$.

As we have mentioned before, when the root node of the quad-K-d tree is of type m the behaviour is like that of a quadtree of dimension m; but the choice of discriminating coordinates (equivalently, the coordinate split vector $\boldsymbol{\delta}$) of the node will determine which coordinates will be taken into consideration, and instead of s specified coordinates we shall only have i specified coordinates, with $0 \le i \le \min(s, m)$. Moreover, the identity of the chosen discriminating coordinates matters, as now we are dealing with a fixed partial match – in contrast to random PM where only the distinction between specified or non-specified coordinates is relevant and so they can be handled equally. This is an important aspect that introduces a new degree of difficulty in the analysis of fixed PM in quad-K-d trees. The other difficulty lies in deducing how the rank vector $\mathbf{r}$ changes when the PM recursively continues in a subtree $T_{\mathbf{w}}$, that is, what is the probability that the rank vector is $\mathbf{r}'$ in the recursive call in $T_{\mathbf{w}}$ given that the rank vector was $\mathbf{r}$. In a quadtree all coordinates are discriminating and affect $\mathbf{r}$ in the same way, only depending on the bitvector $\mathbf{w}$ and the key $\mathbf{x}$ at the root. Nevertheless, because of the symmetry of the problem, it is safe to assume that the specified coordinates of the query are the first s coordinates, making the analysis a bit easier.

The situation that we face here is more similar to the case of K-d trees, but in that case there is only a single discriminating coordinate and all the others can be thought of as unspecified at that level, therefore we can condition on the event that i, $0 \le i < K$, is the only specified coordinate. In quad-K-d trees we have to deal with m discriminating

coordinates, and hence we must condition on the event that coordinate split vector is $\boldsymbol{\delta}$, for all $\binom{K}{m}$ possible $\boldsymbol{\delta}$. As the coordinate split vector is implicitly encoded in the labels $\mathbf{w}$ of the subtrees, we can avoid an explicit sum over all posible $\boldsymbol{\delta}$, and just sum over all possible $\mathbf{w}$.

The behaviour of the PM with respect to the $K-s$ unspecified coordinates (the last ones in the query, by assumption) is always the same, whether the root node discriminates or not with respect to any of those coordinates. Thus we can pull out a common factor contributed by the $m-i$ discriminating coordinates which are not specified in the PM query and condition on $\mathbf{w} \in \{0,1,\#\}^s$ with exactly i symbols that are 0 or 1, so we also sum over all possible values of i, $0 \le i \le \min(s,m)$. This is simpler than conditioning over all possible $\mathbf{w} \in \{0,1,\#\}^K$ of which m symbols are 0 or 1, and then condition on i of the first s coordinates being 0 or 1.

Let us begin first with the (identical) contribution of all 2^{m-i} subtrees with a particular choice $\mathbf{k} = (k_0, \ldots, k_{m-i})$ of non-specified discriminating coordinates:

$$\int_0^1 \cdots \int_0^1 f(z_0,\ldots,z_{s-1}) \cdot (u_{k_0}^\alpha + (1-u_{k_0})^\alpha) \cdots (u_{k_{m-i}}^\alpha + (1-u_{k_{m-i}})^\alpha)\, du_{k_0} \cdots du_{k_{m-i}} = \left(\frac{2}{\alpha+1}\right)^{m-i} f(z_0,\ldots,z_{s-1}).$$

We have therefore the following integral equation

$$f(z_0,\ldots,z_{s-1}) = \sum_{m=1}^{K} \tau_m \sum_{i\ge 0} \frac{\binom{K-s}{m-i}}{\binom{K}{m}} \left(\frac{2}{\alpha+1}\right)^{m-i} \times \sum_{\substack{\mathbf{w}\in(0+1+\#)^s \\ |\mathbf{w}|_{0,1}=i}} \Bigg\{ \int_{I_{w_0}(z_0)} \cdots \int_{I_{w_{s-1}}(z_{s-1})} f\Big(\rho_{w_0}(z_0,u_0),\ldots,\rho_{w_{s-1}}(z_{s-1},u_{s-1})\Big) \cdot \Big(\theta_{w_0}(u_0)\cdots\theta_{w_{s-1}}(u_{s-1})\Big)^\alpha du_{s-1}\cdots du_0 \Bigg\}, \quad (5)$$

where $\rho_1(z_i,u_i) = \frac{z_i}{u_i}$, $\rho_0(z_i,u_i) = \frac{1-z_i}{1-u_i}$, $\rho_\#(z_i,u_i) = z_i$, $\theta_1(u) = u$, $\theta_0(u) = 1-u$, $\theta_\#(u) = 1$, $I_1(z) = [z,1]$, $I_0(z) = [0,z]$ and $I_\#(z) = [0,1]$.

Recall that only $i \le m$ coordinates will be specified and discriminating at the root of the quad-K-d tree; the remaining coordinates are either unspecified or non-discriminating and the PM does the same from the point of view of those coordinates; we have thus to sum over all possible choices involving the s specified coordinates: the root does not discriminate for that coordinate ($w_j = \#$), the root discriminates for that coordinate and the rank specified in the PM is less (or equal to) than the rank of the key of the root node ($w_j = 0$) or the root discriminates for that coordinate and the rank specified in the PM is greater than the rank of the key of the root node ($w_j = 1$).

In order to find a solution for the integral equation above, we will use that $f(z_0,\ldots,z_{s-1})$ – if it exists – satisfies the following conditions:

1. $f(z_0,\ldots,z_{s-1})$ is symmetric on all variables, that is, for any i and j,

$$f(z_0,\ldots,z_i,\ldots,z_j,\ldots,z_{s-1}) = f(z_0,\ldots,z_j,\ldots,z_i,\ldots,z_{s-1});$$

and,

2. averaging $P_{n,\mathbf{r}}$ over all possible $\mathbf{r}$ should give us $\hat{P}_n$, hence

$$\int_0^1 \int_0^1 \cdots \int_0^1 f(z_0,\ldots,z_{s-1})\, dz_0 \cdots dz_{s-1} = \beta_{s,K},$$

where $\beta_{s,K}$ is the constant factor of the main order term of $\hat{P}_n$ (see Theorem 3).

We will begin assuming that the integral equation (5) admits a solution on separate variables, that is, $f(z_0, z_1, \ldots, z_{s-1}) = \phi_0(z_0) \cdot \phi_1(z_1) \cdots \phi_{s-1}(z_{s-1})$. And because of the symmetries that f satisifes, namely Condition #1 above, we must have $\phi = \phi_0 = \cdots = \phi_{s-1}$.

Then

$$\begin{aligned}
\phi(z_0)\ldots\phi(z_{s-1}) &= \sum_{m=1}^{K} \tau_m \sum_{i=0}^{m} \frac{\binom{K-s}{m-i}}{\binom{K}{m}} \left(\frac{2}{\alpha+1}\right)^{m-i} \times \sum_{\substack{\mathbf{w}\in(0+1+\#)^s \\ |\mathbf{w}|_{0,1}=i}} \Bigg\{ \\
&\int_{I_{w_0}(z_0)} \cdots \int_{I_{w_{s-1}}(z_{s-1})} \phi(\rho_{w_0}(z_0,u_0))\ldots\phi(\rho_{w_{s-1}}(z_{s-1},u_{s-1}))\times \\
&\qquad \left(\theta_{w_0}(u_0)\cdots\theta_{w_{s-1}}(u_{s-1})\right)^{\alpha} du_{s-1}\cdots du_0 \Bigg\} \\
&= \sum_{m=1}^{K} \tau_m \sum_{i=0}^{m} \frac{\binom{K-s}{m-i}}{\binom{K}{m}} \left(\frac{2}{\alpha+1}\right)^{m-i} \sum_{\substack{\mathbf{w}\in(0+1+\#)^s \\ |\mathbf{w}|_{0,1}=i}} \Bigg\{ \\
&\qquad\qquad \prod_{k=0}^{s-1} \int_{I_{w_k}[z_k]} \phi(\rho_{w_k}(z_k,y))(\theta_{w_k}(z_k,y))^{\alpha} dy \Bigg\}.
\end{aligned}$$

Since we are assuming that $f(z_0, \ldots, z_{s-1}) = \phi(z_0)\cdots\phi(z_{s-1})$ the summation on $\mathbf{w}$ on the right hand side can be written

$$\begin{aligned}
&\sum_{\substack{\mathbf{w}\in(0+1+\#)^s \\ |\mathbf{w}|_{0,1}=i}} \Bigg\{ \int_{I_{w_0}(z_0)} \cdots \int_{I_{w_{s-1}}(z_{s-1})} f\Big(\rho_{w_0}(z_0,u_0),\ldots,\rho_{w_{s-1}}(z_{s-1},u_{s-1})\Big) \\
&\qquad \cdot \left(\theta_{w_0}(u_0)\cdots\theta_{w_{s-1}}(u_{s-1})\right)^{\alpha} du_{s-1}\cdots du_0 \Bigg\} \\
&= \sum_{\substack{\mathbf{w}\in(0+1+\#)^s \\ |\mathbf{w}|_{0,1}=i}} \Bigg\{ \prod_{k=0}^{s} \int_{I_{w_k}(z_k)} \phi(\rho_{w_k}(z_k,u_k)) \left(\theta_{w_k}(u_k)\right)^{\alpha} du_k \Bigg\} \\
&= \sum_{\mathbf{k}=(k_0,\ldots,k_{i-1})\subset\{0,\ldots,s-1\}} \Bigg\{ \prod_{\ell=0}^{i-1} \Bigg(\int_0^{z_{k_\ell}} \phi(z_{k_\ell}/u) u^{\alpha} du \\
&\qquad\qquad + \int_{z_{k_\ell}}^{1} \phi((1-z_{k_\ell})/(1-u))\,(1-u)^{\alpha} du \Bigg) \cdot \prod_{0\le j<s : j\notin\mathbf{k}} \int_0^1 \phi(z_j) du \Bigg\}.
\end{aligned}$$

Making the change of variables in the integrals ($y := z_i/u_i$ or $y := (1-z_i)/(1-u_i)$, as needed), we arrive at

$$\begin{aligned}
\phi(z_0)\cdot\phi(z_1)\cdots\phi(z_{s-1}) &= \sum_{m=1}^{K} \tau_m \sum_{i=0}^{m} \frac{\binom{K-s}{m-i}}{\binom{K}{m}} \left(\frac{2}{\alpha+1}\right)^{m-i} \sum_{\mathbf{k}=(k_0,\ldots,k_{i-1})\subset\{0,\ldots,s-1\}} \Bigg\{ \\
&\prod_{\ell=0}^{i-1} \left(z_{k_\ell}^{\alpha} \int_0^{z_{k_\ell}} \phi(u) \frac{du}{u^{\alpha+2}} + (1-z_{k_\ell})^{\alpha} \int_{z_{k_\ell}}^{1} \phi(u) \frac{du}{(1-u)^{\alpha+2}} \right) \cdot \prod_{0\le j<s : j\notin\mathbf{k}} \phi(z_j) \Bigg\}. \quad (6)
\end{aligned}$$

Let $\phi(z) = \nu(z(1-z))^{\varphi-1}$, with $\varphi = \alpha/2+1$, for an arbitrary constant ν. If we define

$$L[\phi(z)] = z^\alpha \int_z^1 \phi(y)\frac{dy}{y^{\alpha+2}}, \qquad \text{and}$$
$$R[\phi(z)] = (1-z)^\alpha \int_0^z \phi(y)\frac{dy}{(1-y)^{\alpha+2}},$$

then it is easy to show (see [14] for more details) that

$$\phi(z) = \varphi \cdot (L[\phi(z)] + R[\phi(z)])$$

Let $\mathbf{k} = \{k_0, k_1, \ldots, k_{i-1}\} \subset \{0,\ldots,s-1\}$ and $\hat{\mathbf{k}} = \{\hat{k}_0, \hat{k}_1, \ldots, \hat{k}_{s-1-i}\} = \{0,\ldots,s-1\}\backslash\mathbf{k}$, then, for any $\mathbf{k}$ we have

$$\begin{aligned}
&\phi(z_{\hat{k}_0})\cdots\phi(z_{\hat{k}_{s-1-i}}) \cdot \prod_{\ell=0}^{i-1}(L[\phi(z_{k_\ell})] + R[\phi(z_{k_\ell})]) \\
&\quad = \phi(z_{\hat{k}_0})\cdots\phi(z_{\hat{k}_{s-1-i}}) \cdot \left(\frac{1}{\varphi}\right)^i \cdot \phi(z_{k_0})\cdots\phi(z_{k_{i-1}}) \\
&\quad = \phi(z_0)\cdots\phi(z_{s-1}) \cdot \left(\frac{1}{\varphi}\right)^i.
\end{aligned}$$

Therefore the summation in (6) reads

$$\begin{aligned}
&\sum_{\mathbf{k}:\mathbf{k}\subset\{0,\ldots,s-1\}} \phi(z_{\hat{k}_0})\cdots\phi(z_{\hat{k}_{s-1-i}}) \cdot \prod_{\ell=0}^{i-1}(L[\phi(z_{k_\ell})] + R[\phi(z_{k_\ell})]) \\
&\qquad = \binom{s}{i}\left(\frac{1}{\varphi}\right)^i \phi(z_0)\cdots\phi(z_{s-1}).
\end{aligned}$$

Thus we can express the right hand side of (5) as

$$\begin{aligned}
&\sum_{m=1}^K \tau_m \sum_{i=0}^m \frac{\binom{K-s}{m-i}}{\binom{K}{m}}\left(\frac{2}{\alpha+1}\right)^{m-i}\Bigg\{ \\
&\qquad \sum_{\substack{\mathbf{k}:\mathbf{k}\subset\{0,\ldots,s-1\} \\ \hat{\mathbf{k}}=\{0,\ldots,s-1\}\backslash\mathbf{k}}} \phi(z_{\hat{k}_0})\cdots\phi(z_{\hat{k}_{s-1-i}}) \prod_{\ell=0}^{i-1}(L[\phi(z_{k_\ell})] + R[\phi(z_{k_\ell})])\Bigg\} \\
&= \sum_{m=1}^K \tau_m \sum_{i=0}^m \frac{\binom{K-s}{m-i}}{\binom{K}{m}}\left(\frac{2}{\alpha+1}\right)^{m-i}\binom{s}{i}\left(\frac{1}{\varphi}\right)^i \phi(z_0)\cdots\phi(z_{s-1}) \\
&= \phi(z_0)\cdots\phi(z_{s-1}) \cdot \left\{\sum_{m=1}^K \tau_m \sum_{i=0}^m \frac{\binom{K-s}{m-i}\binom{s}{i}}{\binom{K}{m}} \frac{2^m}{(\alpha+1)^{m-i}(\alpha+2)^i}\right\} \\
&\qquad = \phi(z_0)\cdots\phi(z_{s-1}) = f(z_0,\ldots,z_{s-1}),
\end{aligned}$$

where we have used the fact that the summation enclosed in curly braces of the second-to-last line is, by definition, equal to 1 (see (1)).

To conclude we need to find the value of μ. Using Condition #2, the integral of f in the hypercube $[0,1]^s$ must be equal to $\beta_{s,K}$. That is, we must have

$$\left(\int_0^1 \nu\,(z(1-z))^{\alpha/2}\,dz\right)^s = \beta_{s,K},$$

and thus the value of $\nu_{s,K}$ given in Theorem 6 must be

$$\beta_{s,K}\left(\frac{1}{\int_0^1 (z(1-z))^{\alpha/2}\,dz}\right)^s = \beta_{s,K}\frac{\Gamma^s(\alpha+2)}{\Gamma^{2s}(\alpha/2+1)}.$$ ◀

4 Conclusions and Further Work

We have derived in this paper the expected performance of partial match queries in random relaxed quad-K-d trees, for random as well as for fixed queries. Our results show that the behaviour of random relaxed quad-K-d trees is identical to that of quadtrees and relaxed K-d trees (which are two important particular cases): we always have $\hat{P}_n \sim \beta n^\alpha$ for random partial match and $P_{n,\mathbf{q}} \sim \nu \cdot f(\mathbf{q}) \cdot n^\alpha$ for fixed partial matches, with the exponent α and the constant factors β and ν depending only of the dimension K, the number of specified coordinates s in the partial match and the specific family of quad-K-d trees – which is "encoded" in the probabilities $\{\tau_m\}_{1\leq m\leq K}$.

Our unifying analysis confirms, for this very general family of multidimensional trees –the random relaxed quad-K-d trees–, the conjecture made in [14] on the expected performance of fixed partial matches.

We also see these results as a fundamental first step to establish trade-offs between the amount of space needed to store the data structure and the expected performance of exact searches and of partial matches – the last, in turn, is fundamental to derive the performance of other associative queries, such as orthogonal range queries and nearest neighbour queries (see for instance [9, 13]).

Indeed, although it is not possible to find closed forms for α, β and ν (as they involve the roots of a polynomial of degree K), we can still investigate numerically how these parameters behave on a particular family of relaxed quad-K-d trees, characterised by the probabilities τ. In fact, with the general definition of relaxed quad-K-d trees we have introduce "control knobs" like the probability p in binomial split, pseudo-binomial split and geometric split quad-K-d trees (see [4, 11]), which allow us to smoothly transition from random relaxed K-d trees when $p = 0$ (best memory usage, largest IPL and partial match costs) to random quadtrees when $p = 1$ (worst memory usage, smallest IPL and partial match costs). Another example would be to consider m-regular quad-K-d trees (where all nodes are of type m) and study how different performance parameters evolve with m as we go from $m = 1$ to $m = K$.

In practical settings, such a study is relevant as it helps to select the family of quad-K-d trees (that is, fix the probabilities τ_m) which better fit the requirements of a given application.

Another possible line of research is to extend these results to families of quad-K-d trees which are not *relaxed*, like the standard K-d trees [1] or squarish K-d trees [9] or their generalisations to m-ary trees: these are examples of important families of quad-K-d trees where the types of the nodes are not independently chosen and/or the discriminating coordinates within a given node are not randomly assigned.

References

1 J. L. Bentley. Multidimensional Binary Search Trees Used for Associative Searching. *Communications of the ACM*, 18(9):509–517, 1975.

2 J.L. Bentley and R. A. Finkel. Quad trees: A data structure for retrieval on composite keys. *Acta Informatica*, 4:1–9, 1974.

3 Nikolett Bereczky, Amalia Duch, Krisztián Németh, and Salvador Roura. Quad-k-d trees. In *LATIN 2014: Theoretical Informatics - 11th Latin American Symposium, Montevideo, Uruguay, March 31 - April 4, 2014. Proceedings*, pages 743–754, 2014.

4 Nikolett Bereczky, Amalia Duch, Krisztián Németh, and Salvador Roura. Quad-kd trees: A general framework for kd trees and quad trees. *Theor. Comput. Sci.*, 616:126–140, 2016.

5 Khristo N. Boyadzhiev. *Notes on the Binomial Transform.* World Scientific, 2018.

6 N. Broutin, R. Neininger, and H. Sulzbach. A limit process for partial match queries in random quadtrees and 2-d trees. *Annals of Applied Probability*, 23(6):2560–2603, 2013.

7 H.-H. Chern and H.-K. Hwang. Partial match queries in random k-d trees. *SIAM Journal on Computing*, 35(6):1440–1466, 2006.

8 H.-H. Chern and H.-K. Hwang. Partial match queries in random quadtrees. *SIAM Journal on Computing*, 32(4):904–915, 2012.

9 L. Devroye, J. Jabbour, and C. Zamora-Cura. Squarish k-d trees. *SIAM Journal on Computing*, 30:1678–1700, 2000.

10 A. Duch, V. Estivill-Castro, and C. Martínez. Randomized K-dimensional binary search trees. In K. Y. Chwa and O. H. Ibarra, editors, *9th Annual International Symposium on Algorithms and Computation (ISAAC 98)*, volume 1533 of *Lecture Notes in Computer Science*, pages 199–208, 1998.

11 A. Duch, G. Lau, and C. Martínez. Random partial match in quad-k-d trees. In E. Kranakis, G. Navarro, and E. Chávez, editors, *LATIN 2016: Theoretical Informatics (Proceedings of the 12$^{\text{th}}$ Latin American Theoretical Informatics Conference (LATIN))*, volume 9644 of *Lecture Notes in Computer Science*, pages 376–389. Springer-Verlag, 2016.

12 A. Duch, G. Lau, and C. Martínez. Fixed partial match queries in quadtrees. In J.A. Fill amd M.D. Ward, editor, *Proceedings of the 29$^{\text{th}}$ International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA)*, volume 110 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 20:1–20:18, 2018.

13 A. Duch and C. Martínez. On the average performance of orthogonal range search in multidimensional data structures. *Journal of Algorithms*, 44(1):226–245, 2002. doi:{10.1016/S0196-6774(02)00213-4}.

14 Amalia Duch, Gustavo Lau, and Conrado Martínez. On the cost of fixed partial match queries in k-d trees. *Algorithmica*, 75(4):684–723, 2016.

15 Ph. Flajolet and C. Puech. Partial match retrieval of multidimensional data. *Journal of the ACM*, 33(2):371–407, 1986.

16 Philippe Flajolet and Robert Sedgewick. Mellin transforms and asymptotics: Finite differences and rice's integrals. *Theoretical Computer Science*, 144(1&2):101–124, 1995. doi:10.1016/0304-3975(94)00281-M.

17 Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009. URL: http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=9780521898065.

18 V. Gaede and O. Günther. Multidimensional access methods. *ACM Computing Surveys*, 30(2):170–231, 1998.

19 Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science, 2nd Ed.* Addison-Wesley, 1994.

20 D. E. Knuth. *The Art of Computer Programming: Sorting and Searching*, volume 3. Addison–Wesley, 2nd edition, 1998.

21 C. Martínez, A. Panholzer, and H. Prodinger. Partial match queries in relaxed multidimensional search trees. *Algorithmica*, 29(1–2):181–204, 2001.

22 Roura, S. Improved master theorems for divide-and-conquer recurrences. *J. ACM*, 48(2):170–205, 2001.

23 H. Samet. *The Design and Analysis of Spatial Data Structures.* Addison-Wesley, 1990.

24 Neil J.A. Sloane and Simon Plouffe. *The Encyclopedia of Integer Sequences.* Academic Press, 1995.

A Proof of Equation (4)

Our goal here is to prove (4):

$$\sum_{k=j}^{n-1}\binom{n-1}{k}(-1)^k\pi_{k+1,j}^{(i,m)} = \binom{n-1}{j}(-1)^j(n+1)^{-i}n^{-m+i},$$

an intermediate result in the proof of Propositon 4. Recall that $\pi_{k+1,j}^{(i,m)}$ denotes the average number of top-level recursive calls on subtrees of size j when the quad-K-d tree is of size $k+1$, its root is of type m and exactly i discriminating coordinates out of m are specified. But this number is exactly the same as the average number of top-level recursive calls on subtrees of size j on a m-dimensional quadtree of size $k+1$ for a random partial match in which i of the m dimensions are specified; and this probability is given by [8, Eq. (2.1)]

$$\pi_{k+1,j}^{(i,m)} = \binom{k}{j}\int_{[0,1]^m}\mathbf{dx}\; x_0\cdots x_{i-1}(x_0\cdots x_{m-1})^j(1-x_0\cdots x_{m-1})^{k-j},$$

where $\mathbf{dx} := dx_0\cdots dx_{m-1}$.

We can then write the LHS of (4), for a fixed j,

$$\begin{aligned}
&\sum_{k=j}^{n-1}\binom{n-1}{k}(-1)^k\pi_{k+1,j}^{(i,m)}\\
&\quad=\sum_{k=j}^{n-1}\binom{n-1}{k}(-1)^k\binom{k}{j}\int_{[0,1]^m}\mathbf{dx}\; x_0\cdots x_{i-1}(x_0\cdots x_{m-1})^j(1-x_0\cdots x_{m-1})^{k-j}\\
&\quad=\binom{n-1}{j}\sum_{k=j}^{n-1}\binom{n-1-j}{k-j}(-1)^k\times\Bigg\{\\
&\qquad\int_{[0,1]^m}\mathbf{dx}\; x_0\cdots x_{i-1}(x_0\cdots x_{m-1})^j(1-x_0\cdots x_{m-1})^{k-j}\Bigg\}\\
&\quad=\binom{n-1}{j}(-1)^j\Bigg\{\sum_{\ell=0}^{n-1-j}\binom{n-1-j}{\ell}(-1)^\ell\\
&\qquad\times\int_{[0,1]^m}\mathbf{dx}\; x_0\cdots x_{i-1}(x_0\cdots x_{m-1})^j(1-x_0\cdots x_{m-1})^{\ell}\Bigg\}\\
&\quad=\binom{n-1}{j}(-1)^j\int_{[0,1]^m}\mathbf{dx}\; x_0\cdots x_{i-1}(x_0\cdots x_{m-1})^j\Bigg\{\sum_{\ell=0}^{n-1-j}\binom{n-1-j}{\ell}\times\\
&\qquad(-1)^\ell(1-x_0\cdots x_{m-1})^{\ell}\Bigg\}\\
&\quad=\binom{n-1}{j}(-1)^j\int_{[0,1]^m}\mathbf{dx}\; x_0\cdots x_{i-1}(x_0\cdots x_{m-1})^j(x_0\cdots x_{m-1})^{n-1-j}\\
&\quad=\binom{n-1}{j}(-1)^j\int_{[0,1]^m}\mathbf{dx}\; x_0\cdots x_{i-1}(x_0\cdots x_{m-1})^{n-1}\\
&\quad=\binom{n-1}{j}(-1)^j\left(\int_0^1 x^n dx\right)^i\cdot\left(\int_0^1 x^{n-1}dx\right)^{m-i}=\binom{n-1}{j}(-1)^j(n+1)^{-i}n^{-m+i}.
\end{aligned}$$

Improved Error Bounds for the Number of Irreducible Polynomials and Self-Reciprocal Irreducible Monic Polynomials with Prescribed Coefficients over a Finite Field

Zhicheng Gao
School of Mathematics and Statistics, Carleton University, Canada

Abstract

A polynomial is called self-reciprocal (or palindromic) if the sequence of its coefficients is palindromic. In this paper we obtain improved error bounds for the number of irreducible polynomials and self-reciprocal irreducible monic polynomials with prescribed coefficients over a finite field. The improved bounds imply that self-reciprocal irreducible monic polynomials with degree $2d$ and prescribed ℓ leading coefficients always exist provided that ℓ is slightly less than $d/2$.

2012 ACM Subject Classification Mathematics of computing → Generating functions; Mathematics of computing → Enumeration

Keywords and phrases finite fields, irreducible polynomials, prescribed coefficients, generating functions, Weil bounds, self-reciprocal

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.9

Funding Research supported by NSERC (RGPIN 04010-2015) and Carleton University Development Grant (189035).

Acknowledgements I would like to thank the referees for their helpful comments which improve the presentation of the paper.

1 Introduction

The existence of irreducible polynomials over finite fields with restricted coefficients play important roles in coding theory and information theory (see, e.g., [13, 10]). The main objective of this paper is to improve some well-known error bounds on the number of irreducible polynomials and self-reciprocal irreducible monic polynomials with prescribed coefficients. Asymptotic formulas with good error bounds played essential roles in proving the existence of irreducible polynomials and self-reciprocal irreducible monic polynomials with prescribed coefficients. For example, the famous Hansen-Mullen conjecture on the existence of irreducible polynomials with one prescribed coefficient was proved asymptotically by Wan [14] using the Weil bound on character sums. Practical error bounds allow Ham and Mullen [7] to confirm the conjecture for small degrees and finite fields. Panario and Tzanakis [11] used Wan's approach to study the extended Hansen-Mullen conjecture by considering several prescribed coefficients. Garefalakis and Kapetanakis [5] used Wan's approach to prove the existence of self-reciprocal irreducible monic polynomials with one prescribed coefficient. Ha and Pollack [6, 12] obtained bounds for several prescribed coefficients using a different approach based on the circle method. Our approach uses generating functions whose coefficients are from the group algebra defined in terms of the prescribed coefficients.

Throughout the paper, we shall use the following notations.

- $\mathbb{F}_q$ denotes the finite field with q elements and $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$.
- $\mathcal{M}_q$ denotes the set of monic polynomials over $\mathbb{F}_q$ and $\mathcal{M}_q(d) = \{f : f \in \mathcal{M}_q, \deg(f) = d\}$.

- For a polynomial f, $\deg(f)$ denotes the degree of f, and $f^*(x) = x^{\deg(f)} f(1/x)$ is the *reciprocal* of f. If $f^* = f$, then f is called *self-reciprocal* or *palindromic*.
- $\mathcal{I}_q \subseteq \mathcal{M}_q$ denotes the set of irreducible monic polynomials and $\mathcal{I}_q(d) = \mathcal{I}_q \cap \mathcal{M}_q(d)$.
- $\mathcal{S}_q \subseteq \mathcal{I}_q$ denotes the set of self-reciprocal irreducible monic polynomials. Since every odd-degree self-reciprocal polynomial contains the factor $x+1$, we only need to consider $\mathcal{S}_q(d) = \{f : f \in \mathcal{S}_q, \deg(f) = 2d\}$.
- For a generating function F, $[\,]F$ extracts the relevant coefficient from F.

Given non-negative integers ℓ, t, we say that two polynomials $f, g \in \mathcal{M}_q$ are *equivalent* with respect to ℓ, t if

$$\begin{aligned}\left[x^{\deg(f)-j}\right] f(x) &= \left[x^{\deg(g)-j}\right] g(x), 1 \le j \le \ell,\\ \left[x^{j}\right] f(x) &= \left[x^{j}\right] g(x), 0 \le j \le t-1.\end{aligned}$$

Let $\langle f \rangle$ denote the equivalence class represented by f. It is known [15, 8, 9] that the set $\mathcal{E}^{\ell,t}$ of all equivalence classes forms an abelian group under the multiplication $\langle f\rangle\langle g\rangle = \langle fg\rangle$. (When $t > 0$, it is assumed that the constant term is nonzero.)

Given $\varepsilon \in \mathcal{E}^{\ell,t}$ and $\delta \in \mathcal{E}^{\ell,0}$, define

$$I_q(d;\varepsilon) = \left|\{f \in \mathcal{I}_q(d) : \langle f\rangle = \varepsilon\}\right|, \quad S_q(d;\delta) = \left|\{f \in \mathcal{S}_q(d) : \langle f\rangle = \delta\}\right|.$$

The rest of the paper is organized as follows. In Section 2 we state our main results about the bounds for $I_q(d;\varepsilon)$ and $S_q(d;\varepsilon)$. In Section 3 we sketch the proofs of our main results. Section 4 gives some examples to demonstrate the improvement of our bounds over those in [2, 9]. Section 5 concludes the paper.

2 Main results

In the rest of the paper, we shall use the Iverson bracket $[\![P]\!]$ which has value 1 if the predicate P is true and value 0 otherwise. It is easy to see [15, 9] that

$$\left|\mathcal{E}^{\ell,t}\right| = (q - [\![t > 0]\!])q^{\ell+t-1}.$$

For typographical convenience, we shall use $\mathcal{E}^{\ell}$ to denote $\mathcal{E}^{\ell,0}$. For any given q, the following observation [9, Lemma 1.1] will be useful:

$$\mathcal{E}^{\ell,t} \cong \mathcal{E}^{\ell} \times \mathcal{E}^{t-1} \times \mathbb{F}_q^*, \quad t \ge 1. \tag{1}$$

Thus we may focus on the group $\mathcal{E}^{\ell}$. When $t > 0$ and $\varepsilon \in \mathcal{E}^{\ell,t}$, we also write $\varepsilon = (\varepsilon_1, \gamma^m, \varepsilon_2)$ with $\varepsilon_1 \in \mathcal{E}^{\ell}$, $\varepsilon_2 \in \mathcal{E}^{t-1}$, and $1 \le m \le q-1$.

Since $\mathcal{E}^{\ell}$ is abelian, it is isomorphic to a direct product of cyclic groups. Let $\xi_{\ell,1}, \ldots, \xi_{\ell,u_\ell}$ be a fixed minimal set of generators of $\mathcal{E}^{\ell}$, and denote their orders by $r_{\ell,1}, \ldots, r_{\ell,u_\ell}$, respectively. In the rest of the paper, γ denotes a fixed generator of the multiplicative group $\mathbb{F}_q^*$. By (1), each $\varepsilon \in \mathcal{E}^{\ell,t}$ can be written uniquely as

$$\varepsilon = \gamma^{e_0(\varepsilon)} \prod_{h=1}^{u_\ell} \xi_{\ell,h}^{e_{\ell,h}(\varepsilon)} \prod_{i=1}^{u_{t-1}} \xi_{t-1,i}^{e_{t-1,i}(\varepsilon)}, 1 \le e_0(\varepsilon) \le q-1, 1 \le e_{\ell,h}(\varepsilon) \le r_{\ell,u_\ell}, 1 \le e_{t-1,i}(\varepsilon) \le r_{\ell,u_{t-1}}.$$

Thus each $\varepsilon \in \mathcal{E}^{\ell,t}$ can be represented uniquely by either a monic polynomial of degree $\ell + t$ or the *exponent vector* $\vec{e}(\varepsilon) = (e_0(\varepsilon), e_{\ell,1}(\varepsilon), \ldots, e_{\ell.u_\ell}, e_{t-1,1}, \ldots, e_{t-1,u_{t-1}})$. When $t = 0$, it is understood that e_0 is ignored.

Let $\omega_r = \exp(2\pi i/r)$ and $\varepsilon, \varepsilon' \in \mathcal{E}^{\ell,t}$. Define

$$\{\varepsilon^{1/k}\} = \{\delta \in \mathcal{E}^{\ell,t} : \delta^k = \varepsilon\}, \tag{2}$$

$$\mathcal{E}^{\ell,t}(d) = \{\langle f\rangle : f \in \mathcal{M}_q(d)\},$$

$$a(\varepsilon,\varepsilon') = \omega_{q-1}^{e_0(\varepsilon)e_0(\varepsilon')} \prod_{h=1}^{u_\ell} \omega_{r_{\ell,h}}^{e_{\ell,h}(\varepsilon)e_{\ell,h}(\varepsilon')} \prod_{i=1}^{u_{t-1}} \omega_{r_{t-1,i}}^{e_{t-1,i}(\varepsilon)e_{t-1,i}(\varepsilon')}, \tag{3}$$

$$c(d;\varepsilon) = \sum_{\varepsilon' \in \mathcal{E}^{\ell,t}(d)} a(\varepsilon,\varepsilon'), \tag{4}$$

$$P(z;\varepsilon) = 1 + \sum_{d=1}^{\ell+t-1} c(d;\varepsilon) z^d, \tag{5}$$

$$D = \sum_{\varepsilon \neq \langle 1\rangle} \deg(P(z;\varepsilon)). \tag{6}$$

Define

$$D' = \sum_{\delta \in \mathcal{E}^\ell \setminus \{\langle 1\rangle\}} \deg(P(z;\delta,1,\delta)).$$

Since $\deg(P(z;\varepsilon)) \le \ell + t - 1$ and $\deg(P(z;\delta,1,\delta)) \le 2\ell$, we have

$$D \le (\ell + t - 1)(\left|\mathcal{E}^{\ell,t}\right| - 1), \tag{7}$$

$$D' \le 2\ell \left(q^\ell - 1\right). \tag{8}$$

With the above notations, we now state our main results.

▶ **Theorem 1.** *Let $\mathcal{E}$ denote the group $\mathcal{E}^{\ell,t}$ and $\varepsilon \in \mathcal{E}$.*

(a) *We have the following upper bounds:*

$$I_q(d;\varepsilon) \le \frac{1}{|\mathcal{E}|}\frac{q^d - [\![t>0]\!]}{d} + \frac{D}{|\mathcal{E}|}\frac{q^{d/2}}{d} \tag{9}$$

$$\le \frac{1}{|\mathcal{E}|}\frac{q^d}{d} + \frac{(|\mathcal{E}|-1)(\ell+t-1)}{|\mathcal{E}|}\frac{q^{d/2}}{d}. \tag{10}$$

(b) *Assume $\ell + t \le \lceil d/2\rceil - 1$ and let $e_1(q,d) = \min\left\{3.4q^{-d/6}, 0.8\right\}$. We have the following lower bounds:*

$$I_q(d;\varepsilon) \ge \frac{1}{|\mathcal{E}|}\frac{q^d - [\![t>0]\!]}{d} - \left(\frac{D + \left|\{\varepsilon^{1/2}\}\right| [\![2 \mid d]\!]}{|\mathcal{E}|} + e_1(q,d)\right)\frac{q^{d/2}}{d} \tag{11}$$

$$\ge \frac{1}{|\mathcal{E}|}\frac{q^d}{d} - (\ell + t + 1)\frac{q^{d/2}}{d}. \tag{12}$$

▶ Remark.

- The upper bound (10) is given in [9, Theorem 2.4], which follows immediately from (9) and (7). The lower bound (12) is given in [2, Theorem 2.1], which follows immediately from (11), (7) and $|\{\varepsilon^{1/2}\}| \le |\mathcal{E}|$.
- We also note that the upper bound given in [2, Theorem 2.1] is slightly weaker than (10), and the lower bound in [9, Theorem 2.4] is slightly weaker than (12).
- Recall $|\mathcal{E}| = (q - [\![t>0]\!])q^{\ell+t-1}$. When $2 \nmid q$, we also have $\left|\{\varepsilon^{1/2}\}\right| \le 1 + [\![t>0]\!]$.

By (7) and (8), our next theorem improves the error bound in [4, Theorem 3] by a factor of q^ℓ. This improvement enables us to essentially extend the range of ℓ from $d/4$ to $d/2$.

▶ **Theorem 2.** *Let $\varepsilon \in \mathcal{E}^\ell$ and $e_2(q,d) = \min\{7q^{-d/6}, 2\}$. Assume $\ell \le \lceil d/2 \rceil - 1$.*

(a) *We have the following upper bounds:*

$$S_q(d;\varepsilon) \le \frac{1}{2d}q^{d-\ell} + \left(\frac{D'+2D+3[\![2 \mid d]\!]\left|\{\varepsilon^{1/2}\}\right|}{2q^\ell} + e_2(q,d)\right)\frac{q^{d/2}}{d} \tag{13}$$

$$\le \frac{1}{2d}q^{d-\ell} + (2\ell+2.5)\frac{q^{d/2}}{d}. \tag{14}$$

(b) *We have the following lower bounds:*

$$S_q(d;\varepsilon) \ge \frac{1}{2d}q^{d-\ell} - \left(\frac{D'+2D}{2q^\ell} + [\![2 \mid d]\!]\frac{\left|\{\langle 1\rangle^{1/2}\}\right|}{q^\ell} + e_2(q,d)\right)\frac{q^{d/2}}{d} \tag{15}$$

$$\ge \frac{1}{2d}q^{d-\ell} - (2\ell+2)\frac{q^{d/2}}{d}. \tag{16}$$

Consequently $S_q(d;\varepsilon) > 0$ whenever

$$\ell \le \min\left\{\left\lceil\frac{d}{2}\right\rceil - 1, \frac{d}{2} - \log_q(2d+2)\right\}.$$

3 Outline of proofs

Fix ℓ, t and consider the group $\mathcal{E} := \mathcal{E}^{\ell,t}$. In the following, when $t > 0$ and $f(0) = 0$, it is understood that $\langle f\rangle = 0$. Define

$$F(z) = \sum_{f\in\mathcal{M}_q}\langle f\rangle z^{\deg(f)} = \langle 1\rangle + \sum_{d\ge 1}\sum_{f\in\mathcal{M}_q(d)}\langle f\rangle z^d,$$
$$F_q(d;\varepsilon) = d\left[z^d\varepsilon\right]\ln F(z),\ \varepsilon\in\mathcal{E},$$

where

$$\ln F(z) = \sum_{k\ge 1}\frac{(-1)^{k-1}}{k}\left(\sum_{d\ge 1}\sum_{f\in\mathcal{M}_q(d)}\langle f\rangle z^d\right)^k.$$

We note that $F(z)$ is a generating function with coefficients from the group algebra $\mathbb{C}[\mathcal{E}]$. Using the fact that every polynomial is uniquely factored into irreducible polynomials and $\langle f\rangle\langle g\rangle = \langle fg\rangle$, one can obtain [15, Proposition 2] the following equations:

$$F_q(d;\varepsilon) = \sum_{k\mid d}\frac{d}{k}\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}} I_q(d/k;\varepsilon_1), \tag{17}$$

$$I_q(d;\varepsilon) = \frac{1}{d}\sum_{k\mid d}\mu(k)\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}} F_q(d/k;\varepsilon_1). \tag{18}$$

We need a few more notations before stating a formula for $F_q(d;\varepsilon)$ derived in [15]. If $\{\varepsilon^{1/k}\} \ne \emptyset$, we let $\varepsilon^{1/k}$ denote any particular element in $\{\varepsilon^{1/k}\}$. The following simple observations are immediate from (2) and (3).

$$\left\{\varepsilon^{1/k}\right\} = \varepsilon^{1/k}\left\{\langle 1\rangle^{1/k}\right\},$$
$$\left|\left\{\varepsilon^{1/k}\right\}\right| = \left|\left\{\langle 1\rangle^{1/k}\right\}\right| \left[\!\left[\left\{\varepsilon^{1/k}\right\} \neq \emptyset\right]\!\right],$$
$$a(\varepsilon^{-1}, \delta) = a(\varepsilon, \delta^{-1}), \tag{19}$$
$$a(\delta, \varepsilon_1\varepsilon_2) = a(\delta, \varepsilon_1)a(\delta, \varepsilon_2). \tag{20}$$

Set

$$\rho_d(g) := \sum_{\rho} \rho^{-d},$$

where the sum is over all the nonzero roots (with multiplicity) of the polynomial $g \in \mathbb{C}[z]$. Theorem 3 in [15] gives the following formula (written in slightly different notation).

▶ **Proposition 3.** *Let $\mathcal{E}$ denote the group $\mathcal{E}^{\ell,t}$ and let $\varepsilon \in \mathcal{E}$. We have*

$$F_q(d;\varepsilon) = \frac{q^d - [\![t>0]\!]}{|\mathcal{E}|} - \frac{1}{|\mathcal{E}|}\sum_{\delta\in\mathcal{E}\setminus\{\langle 1\rangle\}} a\left(\delta, \varepsilon^{-1}\right)\rho_d(P(z;\delta)). \tag{21}$$

The following lemma simplifies sums involving $F_q(d;\varepsilon)$ over some subgroups of $\mathcal{E}^{\ell,t}$, which play a crucial role in the proofs of Theorems 1 and 2.

▶ **Lemma 4.** *Let $\mathcal{E}$ denote the group $\mathcal{E}^{\ell,t}$.*

(a) *For each $\varepsilon \in \mathcal{E}$, we have*

$$\begin{aligned}
&\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}} F_q\left(d/k;\varepsilon_1\right)\\
&= \frac{\left|\{\varepsilon^{1/k}\}\right|}{|\mathcal{E}|}\left(q^{d/k} - [\![t>0]\!]\right)\\
&\quad - \frac{\left|\{\varepsilon^{1/k}\}\right|}{|\mathcal{E}|}\sum_{\delta\in\mathcal{E}\setminus\{\langle 1\rangle\}} \left[\!\left[\{\delta^{1/k}\} \neq \emptyset\right]\!\right] a\left(\delta, \varepsilon^{-1/k}\right)\rho_{d/k}(P(z;\delta)) \qquad (22)\\
&\leq \frac{\left|\{\varepsilon^{1/k}\}\right|}{|\mathcal{E}|}\left(q^{d/k} - [\![t>0]\!]\right) + (\ell+t-1)q^{d/2k}. \qquad (23)
\end{aligned}$$

(b) *Fix a generator γ of $\mathbb{F}_q^*$ and recall that $\mathcal{E}^\ell$ denotes the group $\mathcal{E}^{\ell,0}$. For each $\varepsilon \in \mathcal{E}^\ell$, we have*

$$\begin{aligned}
&\sum_{m=1}^{q-1}\sum_{\delta\in\mathcal{E}^\ell} F_q\left(d;\varepsilon\delta^{-1},\gamma^m,\delta\right)\\
&= \frac{q^d-1}{q^\ell} - \frac{1}{q^\ell}\sum_{\delta\neq\langle 1\rangle} a\left(\delta,\varepsilon^{-1/k}\right)\rho_d\left(P(z;\delta,1,\delta)\right), \qquad (24)\\
&\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}}\sum_{m=1}^{q-1}\sum_{\delta\in\mathcal{E}^\ell} F_q\left(d/k;\varepsilon_1\delta^{-1},\gamma^m,\delta\right)\\
&= \frac{\left|\{\varepsilon^{1/k}\}\right|}{q^\ell}\left(q^{d/k} - 1 - \sum_{\delta\neq\langle 1\rangle,\{\delta^{1/k}\}\neq\emptyset} a\left(\delta,\varepsilon^{-1/k}\right)\rho_{d/k}\left(P(z;\delta,1,\delta)\right)\right). \qquad (25)
\end{aligned}$$

Proof. (a) The well-known identity

$$\sum_{s=0}^{r-1} \omega_r^{sj} = r[\![r \mid j]\!]$$

immediately leads to

$$\sum_{\delta\in\mathcal{E}} a(\delta,\varepsilon) = |\mathcal{E}|[\![\varepsilon = \langle 1\rangle]\!], \tag{26}$$

$$\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}} a(\varepsilon_1,\delta) = \left|\left\{\varepsilon^{1/k}\right\}\right| a(\varepsilon^{1/k},\delta)[\![\{\delta^{1/k}\}\neq\emptyset]\!]. \tag{27}$$

where $a(\varepsilon^{1/k},\delta)$ is interpreted as 0 if $\{\varepsilon^{1/k}\}=\emptyset$. It follows from (21) and (27) that

$$\begin{aligned}
&\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}} F_q\left(d/k;\varepsilon_1\right)\\
&= \frac{\left|\{\varepsilon^{1/k}\}\right|}{|\mathcal{E}|}\left(q^{d/k} - [\![t>0]\!]\right)\\
&\quad - \frac{1}{|\mathcal{E}|}\sum_{\delta\in\mathcal{E}\setminus\{\langle 1\rangle\}}\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}} a(\varepsilon_1,\delta^{-1})\rho_{d/k}(P(z;\delta))\\
&= \frac{\left|\{\varepsilon^{1/k}\}\right|}{|\mathcal{E}|}\left(q^{d/k} - [\![t>0]\!]\right)\\
&\quad - \frac{\left|\{\varepsilon^{1/k}\}\right|}{|\mathcal{E}|}\sum_{\delta\in\mathcal{E}\setminus\{\langle 1\rangle\}} \left[\!\left[\{\delta^{1/k}\}\neq\emptyset\right]\!\right] a\left(\delta,\varepsilon^{-1/k}\right)\rho_{d/k}(P(z;\delta)),
\end{aligned}$$

which is (22). Now (23) follows by noting

$$\sum_{\delta\in\mathcal{E}\setminus\{\langle 1\rangle\}} \left[\!\left[\{\delta^{1/k}\}\neq\emptyset\right]\!\right] = \frac{|\mathcal{E}|}{\left|\{\langle 1\rangle^{1/k}\}\right|} - 1. \tag{28}$$

To prove part (b), we use (19), (20), (21), and (26) to obtain

$$\begin{aligned}
&\sum_{m=1}^{q-1}\sum_{\delta\in\mathcal{E}^\ell} F_q\left(d;\varepsilon\delta^{-1},\gamma^m,\delta\right)\\
&= \frac{(q-1)q^\ell}{|\mathcal{E}^{\ell,\ell+1}|}\left(q^d-1\right)\\
&\quad - \frac{1}{|\mathcal{E}^{\ell,\ell+1}|}\sum_{(\varepsilon_1,\gamma^n,\varepsilon_2)\neq\langle 1\rangle} a\left(\varepsilon,\varepsilon_1^{-1}\right)\rho_d\left(P(z;\varepsilon_1,\gamma^n,\varepsilon_2)\right)\sum_{m=1}^{q-1}\omega_{q-1}^{-nm}\sum_{\delta\in\mathcal{E}^\ell} a\left(\delta,\varepsilon_1\varepsilon_2^{-1}\right)\\
&= \frac{1}{q^\ell}\left(q^d-1\right) - \frac{1}{q^\ell}\sum_{\delta\neq\langle 1\rangle} a(\varepsilon^{-1},\delta)\rho_d\left(P(z;\delta,1,\delta)\right).
\end{aligned}$$

Now (25) follows from (27). ◀

Proof of Theorem 1. Hsu [9, Theoem 1.3] showed that each (complex) root ρ of $P(z;\delta)$ satisfies

$$|\rho| \geq q^{-1/2}.$$

It follows from (21) that

$$\left|F_q(d;\varepsilon)-\frac{1}{|\mathcal{E}|}\left(q^d-[\![t>0]\!]\right)\right|\le\frac{D}{|\mathcal{E}|}q^{d/2}. \tag{29}$$

By (17), we have

$$I_q(d;\varepsilon)\le\frac{1}{d}F_q(d;\varepsilon).$$

It follows from (29) that

$$I_q(d;\varepsilon)\le\frac{1}{d|\mathcal{E}|}\left(q^d-[\![t>0]\!]\right)+\frac{D}{d|\mathcal{E}|}q^{d/2},$$

which establishes the desired upper bound.

To prove the lower bound, we define

$$\begin{aligned}L_q(d,\ell,t)=&\sum_{k\ge3}[\![k\mid d,\mu(k)=-1]\!]\frac{\left|\{\langle1\rangle^{1/k}\}\right|}{|\mathcal{E}|}q^{d/k-d/2}\\&+(\ell+t-1)\sum_{k\ge2}[\![k\mid d,\mu(k)=-1]\!]\left(1-\frac{\left|\{\langle1\rangle^{1/k}\}\right|}{|\mathcal{E}|}\right)q^{d/2k-d/2}.\end{aligned} \tag{30}$$

Using (18), (23), and (29), we obtain

$$\begin{aligned}I_q(d;\varepsilon)&\ge\frac{1}{d}F_q(d;\varepsilon)-\frac{1}{d}\sum_{k\mid d}[\![\mu(k)=-1]\!]\sum_{\delta\in\{\varepsilon^{1/k}\}}F_q(d/k;\delta)\\&\ge\frac{q^d-[\![t>0]\!]}{d|\mathcal{E}|}-\left(\frac{D+\left|\{\varepsilon^{1/2}\}\right|[\![2\mid d]\!]}{|\mathcal{E}|}+L_q(d,\ell,t)\right)\frac{q^{d/2}}{d}.\end{aligned}$$

We now estimate $L_q(d,\ell,t)$ by truncating the sums in (30) and bounding the remainders by geometric sums. For our purpose, we use

$$\begin{aligned}L_q(d;\ell,t)\le&\sum_{k=3}^{29}[\![k\mid d,\mu(k)=-1]\!]\frac{\left|\{\langle1\rangle^{1/k}\}\right|}{|\mathcal{E}|}q^{d/k-d/2}\\&+(\ell+t-1)\left(1-\frac{1}{|\mathcal{E}|}\right)\sum_{k=2}^{29}[\![k\mid d,\mu(k)=-1]\!]q^{d/2k-d/2}\\&+[\![d\ge30]\!]q^{-d/2}\sum_{1\le j\le d/30}\left(q^j+(\ell+t-1)\left(1-\frac{1}{|\mathcal{E}|}\right)q^{j/2}\right).\end{aligned} \tag{31}$$

Using (31), $\ell+t\le\lceil d/2\rceil-1$, and

$$1-\frac{1}{|\mathcal{E}|}<1,\ \frac{1}{|\mathcal{E}|}\le\frac{\left|\{\langle1\rangle^{1/k}\}\right|}{|\mathcal{E}|}\le1,\ \sum_{1\le j\le d/30}q^j\le\frac{q}{q-1}\left(q^{d/30}-1\right),$$

we obtain (with the help of Maple) $L_q(d,\ell,t)\le\min\{2.8q^{-d/6},0.6\}$ when $q\ge3$.

The case $q=2$ can be treated similarly by observing $|\mathcal{E}|=2^{\ell+t-1}$ and $|\{\langle1\rangle^{1/k}\}|=1$ when k is not a powe of 2. ◀

The following bijection ϕ_d from $\mathcal{E}^\ell$ to itself is used to express $S_q(d;\varepsilon)$ in terms of $I_q(d;\varepsilon)$ [1, 3, 4].

Set $g_0 = 1$. For each positive integer d, let $\phi_d : (g_1, \ldots, g_\ell) \mapsto (f_1, \ldots, f_\ell)$ be defined by

$$f_k = \sum_{j \le k/2} \binom{d+2j-k}{j} g_{k-2j}, \quad 1 \le k \le \ell.$$

The following is [4, Theorem 1], rewritten in more compact notation.

▶ **Proposition 5.** *Suppose $d > 1$ and $\varepsilon \in \mathcal{E}^\ell$. Then*

$$S_q(d;\varepsilon) = \frac{1}{2} \sum_{\varepsilon_1 \in \{\varepsilon^{1/2}\}} S_q(d/2;\varepsilon_1) + I_q(d;\phi_d^{-1}(\varepsilon)) - \frac{1}{2} \sum_{n=0}^{q-2} \sum_{\delta \in \mathcal{E}^\ell} I_q\left(d;\varepsilon\delta^{-1},\gamma^n,\delta\right).$$

Proof of Theorem 2. We first use Theorem 3 and (28) to simplify the following sum:

$$\begin{aligned}
&\sum_{m=1}^{q-1} \sum_{\delta \in \mathcal{E}} I_q(d;\varepsilon\delta^{-1},\gamma^m,\delta) \\
&= \sum_{k \mid d} \frac{\mu(k)}{d} \sum_{m,\delta} \sum_{m_1,\delta_1,\varepsilon_1} F_q(d/k;\varepsilon_1\delta_1^{-1},\gamma^{m_1},\delta_1) [\![\gamma^{km_1} = \gamma^m, \delta_1^k = \delta, \varepsilon_1^k = \varepsilon]\!] \\
&= \sum_{k \mid d} \frac{\mu(k)}{d} \sum_{\varepsilon_1 \in \{\varepsilon^{1/k}\}} \sum_{m_1=1}^{q-1} \sum_{\delta_1 \in \mathcal{E}} F_q(d/k;\varepsilon_1\delta_1^{-1},\gamma^{m_1},\delta_1). \qquad (32)
\end{aligned}$$

Applying Theorem 5, (32), and noting

$$S_q(d/2;\delta) \le I_q(d/2;\delta) \le \frac{2}{d} F_q(d/2;\delta),$$

we obtain

$$\begin{aligned}
S_q(d;\varepsilon) \le{}& \frac{1}{d} \sum_{\varepsilon_1 \in \{\varepsilon^{1/2}\}} F_q(d/2;\varepsilon_1) + I_q(d;\phi_d^{-1}(\varepsilon)) \\
&- \frac{1}{2d} \sum_{m=1}^{q-1} \sum_{\delta \in \mathcal{E}} F_q(d;\varepsilon\delta^{-1},\gamma^m,\delta) \\
&+ \sum_{k \ge 2} \frac{[\![k \mid d, \mu(k) = -1]\!]}{2d} \sum_{\varepsilon_1 \in \{\varepsilon^{1/k}\}} \sum_{m=1}^{q-1} \sum_{\delta \in \mathcal{E}} F_q(d/k;\varepsilon_1\delta^{-1},\gamma^m,\delta).
\end{aligned}$$

It follows from (9), (25), and (27) that

$$\begin{aligned}
S_q(d;\varepsilon) &\le [\![2 \mid d]\!]\frac{\left|\{\varepsilon^{1/2}\}\right|}{q^\ell}\frac{q^{d/2}}{d} + [\![2 \mid d]\!](\ell-1)\left(1-\frac{1}{q^\ell}\right)\frac{q^{d/4}}{d} \\
&\quad + \frac{1}{dq^\ell}\left(q^d + Dq^{d/2}\right) - \frac{1}{2dq^\ell}\left(q^d-1\right) + \frac{D'}{2dq^\ell}q^{d/2} \\
&\quad + [\![2 \mid d]\!]\frac{\left|\{\varepsilon^{1/2}\}\right|}{2q^\ell}\frac{q^{d/2}}{d} + [\![2 \mid d]\!]\ell\left(1-\frac{1}{q^\ell}\right)\frac{q^{d/4}}{d} \\
&\quad + \frac{1}{2dq^\ell}\sum_{k\ge 3}[\![k \mid d, \mu(k)=-1]\!]\left|\{\varepsilon^{1/k}\}\right| q^{d/k} \\
&\quad + \frac{\ell}{d}\sum_{k\ge 3}[\![k \mid d, \mu(k)=-1]\!]\left(1-\frac{1}{q^\ell}\right)q^{d/2k} \\
&\le \frac{1}{d}q^{d-\ell} + \left(\frac{D'+2D+3[\![2 \mid d]\!]\left|\{\varepsilon^{1/2}\}\right|}{2q^\ell} + U_q(d;\ell)\right)\frac{q^{d/2}}{d},
\end{aligned}$$

where

$$\begin{aligned}
U_q(d;\ell) &= \frac{1}{2}q^{-\ell-d/2} + [\![2 \mid d]\!](2\ell-1)\left(1-\frac{1}{q^\ell}\right)q^{-d/4} \\
&\quad + \frac{1}{2q^\ell}\sum_{k\ge 3}[\![k \mid d, \mu(k)=-1]\!]\left|\{\varepsilon^{1/k}\}\right| q^{d/k-d/2} \\
&\quad + \ell\left(1-\frac{1}{q^\ell}\right)\sum_{k\ge 3}[\![k \mid d, \mu(k)=-1]\!]q^{d/2k-d/2}.
\end{aligned}$$

Simple calculations as in the proof of Theorem 1 give $U_q(d;\ell) \le \min\{6.6q^{-d/6}, 1.5\}$. This establishes (13). The bound (14) follows immediately from (13) and (7).

For the lower bound, we use Proposition 5 and (32) to obtain

$$\begin{aligned}
S_q(d;\varepsilon) &\ge I_q(d;\phi_d^{-1}(\varepsilon)) \\
&\quad - \sum_{k\mid d}\frac{[\![k \mid d]\!]}{2d}\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}}\sum_{m=1}^{q-1}\sum_{\delta\in\mathcal{E}^\ell}F_q(d/k;\varepsilon_1\delta^{-1},\gamma^m,\delta) \\
&\ge \frac{1}{d}q^{d-\ell} - \left(\frac{D+\left|\{\langle 1\rangle^{1/2}\}\right|[\![2 \mid d]\!]}{q^\ell} + L_q(d;\ell,0)\right)\frac{q^{d/2}}{d} \\
&\quad - \frac{1}{2d}\sum_{m=1}^{q-1}\sum_{\delta\in\mathcal{E}}F_q(d;\varepsilon\delta^{-1},\gamma^m,\delta) \\
&\quad - \frac{1}{2d}\sum_{k\ge 3}[\![k \mid d, \mu(k)=1]\!]\sum_{\varepsilon_1\in\{\varepsilon^{1/k}\}}\sum_{m=1}^{q-1}\sum_{\delta\in\mathcal{E}}F_q(d/k;\varepsilon_1\delta^{-1},\gamma^m,\delta).
\end{aligned}$$

Applying (24) and (28) again, we obtain

$$\begin{aligned} S_q(d;\varepsilon) \geq \frac{1}{d}q^{d-\ell} - \left(\frac{D + \left|\{\langle 1\rangle^{1/2}\}\right| [\![2 \mid d]\!]}{q^\ell} + L_q(d;\ell,0)\right)\frac{q^{d/2}}{d} \\ - \frac{1}{2dq^\ell}\left(q^d - 1 + D'q^{d/2}\right) \\ - \frac{1}{2d}\sum_{k\geq 6}[\![k \mid d, \mu(k) = 1]\!]\frac{\left|\{\langle 1\rangle^{1/k}\}\right|}{q^\ell}q^{d/k} \\ - \frac{\ell}{d}\left(1 - \frac{1}{q^\ell}\right)\sum_{k\geq 6}[\![k \mid d, \mu(k) = 1]\!]q^{d/2k}. \end{aligned}$$

Define

$$\begin{aligned} L'_q(d;\ell) = L_q(d;\ell,0) + \frac{1}{2}\sum_{k\geq 6}[\![k \mid d, \mu(k) = 1]\!]\frac{\left|\{\langle 1\rangle^{1/k}\}\right|}{q^\ell}q^{d/k-d/2} \\ + \ell\left(1 - \frac{1}{q^\ell}\right)\sum_{k\geq 6}[\![k \mid d, \mu(k) = 1]\!]q^{d/2k-d/2}. \end{aligned}$$

We then have

$$S_q(d;\varepsilon) \geq \frac{1}{2d}q^{d-\ell} - \left(\frac{D'}{2q^\ell} + \frac{D + \left|\{\langle 1\rangle^{1/2}\}\right| [\![2 \mid d]\!]}{q^\ell} + L'_q(d;\ell)\right)\frac{q^{d/2}}{d}. \tag{33}$$

Similar calculations give

$$L'_q(d;\ell) \leq \min\{7q^{-d/6}, 2\}.$$

Now (15) follows from (33). The bound (16) follows immediately from (15), (7), (8), and

$$\left|\{\langle 1\rangle^{1/2}\}\right| \leq q^\ell.$$

Hence $S_q(d;\varepsilon) > 0$ when

$$q^{d/2} > 2(2\ell + 2)q^\ell.$$

Using $2\ell \leq d - 1$ and taking $\log_q$ on both sides, we complete the proof. ◀

4 Examples

In this section, we use some examples to demonstrate that $\frac{D}{|\mathcal{E}^\ell|-1}$ is smaller than $\ell - 1$. For $P(z;\varepsilon)$ defined in (5), let

$$d_j = |\{\varepsilon \in \mathcal{E}^\ell \setminus \{\langle 1\rangle\} : \deg(P(z;\varepsilon)) = j\}|, \quad \vec{d} = (d_1, d_2, \ldots d_{\ell-1}).$$

We note

$$D = \sum_{j=1}^{\ell-1} j d_j.$$

▶ **Example 6.** Consider $q = 2$ and $\ell = 4$. From [15, Example 4], we have

$$\vec{d} = (2, 4, 8),$$
$$D = 2 + 2 \times 4 + 3 \times 8 = 34,$$
$$\frac{D}{|\mathcal{E}^4| - 1} = \frac{34}{2^4 - 1} < 2.3.$$

▶ **Example 7.** Consider $q = 2$ and $\ell = 5$. From [15, Example 6], we have

$$\vec{d} = (2, 4, 8, 16),$$
$$D = 2 + 2 \times 4 + 3 \times 8 + 4 \times 16 = 98,$$
$$\frac{D}{|\mathcal{E}^5| - 1} = \frac{98}{2^5 - 1} < 3.2.$$

▶ **Example 8.** Consider $q = 3$ and $\ell = 3$. From [15, Example 5], we have

$$\vec{d} = (6, 18),$$
$$D = 6 + 2 \times 18 = 42,$$
$$\frac{D}{|\mathcal{E}^3| - 1} = \frac{42}{3^3 - 1} < 1.62.$$

We use the following result from [15, Lemma 1] to produce a few more examples.

▶ **Lemma 9.** *Let $q = p$ be a prime number. The generators of $\mathcal{E}^\ell$ are*

$$\left\{\langle x^j + 1\rangle : \gcd(p, j) = 1, 1 \le j \le \ell\right\},$$

and the order of $\langle x^j + 1\rangle$ is equal to p^{s_j}, where s_j is the smallest positive integer such that $jp^{s_j} > \ell$.

The degree sequence $\vec{d}$ in the following examples are calculated using (3)–(6) with the help of the computer algebra system *Maple*.

▶ **Example 10.** Consider $q = 2$ and $\ell = 6$. By Lemma 9, the group $\mathcal{E}^6$ is generated by $\langle x + 1\rangle$, $\langle x^3 + 1\rangle$, $\langle x^5 + 1\rangle$, of orders 8,4,2, respectively. Hence

$$\vec{d} = (2, 4, 8, 16, 32),$$
$$D = 2 + 2 \times 4 + 3 \times 8 + 4 \times 16 + 5 \times 32 = 258,$$
$$\frac{D}{|\mathcal{E}^6| - 1} = \frac{258}{2^6 - 1} < 4.1.$$

▶ **Example 11.** Consider $q = 2$ and $\ell = 7$. By Lemma 9, the group $\mathcal{E}^7$ is generated by $\langle x + 1\rangle$, $\langle x^3 + 1\rangle$, $\langle x^5 + 1\rangle$, $\langle x^7 + 1\rangle$, of orders 8,4,2,2, respectively. Hence

$$\vec{d} = (2, 4, 8, 16, 32, 64),$$
$$D = 2 + 2 \times 4 + 3 \times 8 + 4 \times 16 + 5 \times 32 + 6 \times 64 = 642,$$
$$\frac{D}{|\mathcal{E}^7| - 1} = \frac{642}{2^7 - 1} < 5.1.$$

▶ **Example 12.** Consider $q = 3$ and $\ell = 4$. By Lemma 9, the group $\mathcal{E}^4$ is generated by $\langle x + 1\rangle$, $\langle x^2 + 1\rangle$, $\langle x^4 + 1\rangle$, of orders 9,3,3, respectively. Hence

$$\vec{d} = (6, 18, 54),$$
$$D = 6 + 2 \times 18 + 3 \times 54 = 204,$$
$$\frac{D}{|\mathcal{E}^4| - 1} = \frac{204}{3^4 - 1} < 2.6.$$

▶ **Example 13.** Consider $q = 3$ and $\ell = 5$. By Lemma 9, the group $\mathcal{E}^5$ is generated by $\langle x+1\rangle$, $\langle x^2+1\rangle$, $\langle x^4+1\rangle$, $\langle x^5+1\rangle$, of orders 9,3,3,3, respectively. Hence

$$\begin{aligned}\vec{d} &= (6, 18, 54, 162),\\ D &= 6 + 2\times 18 + 3\times 54 + 4\times 162 = 852,\\ \frac{D}{|\mathcal{E}^5| - 1} &= \frac{852}{3^5 - 1} < 3.6.\end{aligned}$$

▶ **Example 14.** Consider $q = 3$ and $\ell = 6$. By Lemma 9, the group $\mathcal{E}^6$ is generated by $\langle x+1\rangle$, $\langle x^2+1\rangle$, $\langle x^4+1\rangle$, $\langle x^5+1\rangle$, of orders 9,9,3,3, respectively. Hence

$$\begin{aligned}\vec{d} &= (6, 18, 54, 162, 486),\\ D &= 6 + 2\times 18 + 3\times 54 + 4\times 162 + 5\times 486 = 3282,\\ \frac{D}{|\mathcal{E}^6| - 1} &= \frac{3282}{3^6 - 1} < 4.51.\end{aligned}$$

▶ **Observation.** *The above examples suggest that the degree sequence $\vec{d}$ for the group $\mathcal{E}^\ell$ satisfies*

$$d_j = (q-1)q^j.$$

5 Conclusion

We derived new error bounds for the number of irreducible monic polynomials with prescribed leading and ending coefficients. These bounds improve the bounds in [2, 9]. The new bounds are then used to obtain bounds for the number $S_q(d;\varepsilon)$ of self-reciprocal irreducible monic polynomials of degree $2d$ with ℓ prescribed leading coefficients. The new lower bound for $S_q(d;\varepsilon)$ significantly improves the one obtained in [4] and it implies $S_q(d;\varepsilon) > 0$ when

$$\ell \le \min\left\{\left\lceil\frac{d}{2}\right\rceil - 1, \frac{d}{2} - \log_q(2d+2)\right\}.$$

Some examples are given to demonstrate the improvement of our bounds in Theorem 1 over those in [2, 9]. Our examples show a pattern about the degree sequence $\vec{d}$, which can be used to calculate D exactly.

References

1 L. Carlitz. Some theorems on irreducible reciprocal polynomials over a finite field. *J. Reine Angew. Math.*, 227:212–220, 1967.

2 S. D. Cohen. On irreducible polynomials of certain types in finite fields. *Proc. Camb. Phil. Soc.*, 66:335–344, 1969.

3 S. D. Cohen. Explicit theorems on generator polynomials. *Finite Fields Appl.*, 11:337–357, 2005.

4 Z.C. Gao. Counting self-reciprocal irreducible monic polynomials with prescribed coefficients over a finite field. `arXiv:2109.09006`.

5 T. Garefalakis and G. Kapetanakis. On the hansen-mullen conjecture for self-reciprocal irreducible polynomials. *Finite Fields Appl.*, 69:832–841, 2012.

6 J Ha. Irreducible polynomials with several prescribed coefficients. *Finite Fields Appl.*, 40:10–25, 2016.

7 K.H. Ham and G.L. Mullen. Distribution of irreducible polynomials of small degrees over finite fields. *Math. Comp.*, 67(221):337–341, 1998.

8 D. R. Hayes. The distribution of irreducibles in GF$[q, x]$. *Trans. Amer. Math. Soc.*, 117:101–127, 1965.

9 C.N. Hsu. The distribution of irreducible polynomials in $\mathbb{F}_q[t]$. *J. Number Theory*, 61(1):85–96, 1996.

10 D. Panario. Open problems for polynomials over finite fields and applications. In *Chapter 5 of Proceedings on the Open Problems in Mathematics and Computer Science Conference*, pages 111–126. Springer, 2015.

11 D. Panario and G. Tzanakis. A generalization of the hansen-mullen conjecture on irreducible polynomials over finite fields. *Finite Fields Appl.*, 18:303–315, 2012.

12 P. Pollack. Irreducible polynomials with several prescribed coefficients. *Finite Fields Appl.*, 22:70–78, 2013.

13 D. Bossen S. Hong. On some properties of self-reciprocal polynomials. *IEEE Trans. Inform. Theory*, 21(4):462–464, 1975.

14 D. Wan. Generators and irreducible polynomials over finite fields. *Math. Comp.*, 219:1195–1212, 1997.

15 S. Kuttner Z. Gao and Q. Wang. Counting irreducible polynomials with prescribed coefficients over a finite field. *Finite Fields Appl.*, 80(102023), 2022. doi:10.1016/j.ffa.2022.102023.

Uncovering a Random Tree

Benjamin Hackl
Uppsala University, Sweden
University of Klagenfurt, Austria

Alois Panholzer
TU Wien, Austria

Stephan Wagner
Uppsala University, Sweden

Abstract

We consider the process of uncovering the vertices of a random labeled tree according to their labels. First, a labeled tree with n vertices is generated uniformly at random. Thereafter, the vertices are uncovered one by one, in order of their labels. With each new vertex, all edges to previously uncovered vertices are uncovered as well. In this way, one obtains a growing sequence of forests. Three particular aspects of this process are studied in this extended abstract: first the number of edges, which we prove to converge to a stochastic process akin to a Brownian bridge after appropriate rescaling. Second, the connected component of a fixed vertex, for which different phases are identified and limiting distributions determined in each phase. Lastly, the largest connected component, for which we also observe a phase transition.

2012 ACM Subject Classification Mathematics of computing → Random graphs; Mathematics of computing → Generating functions

Keywords and phrases Labeled tree, uncover process, functional central limit theorem, limiting distribution, phase transition

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.10

Funding *Stephan Wagner*: supported by the Knut and Alice Wallenberg Foundation.

1 Introduction

We consider the process of uncovering the vertices of a random tree: starting either from one of the n^{n-2} unrooted or one of the n^{n-1} rooted unordered labeled trees of size n (i.e., with n vertices) chosen uniformly at random, we uncover the vertices one by one in order of their labels. This yields a growing sequence of (rooted) forests induced by the uncovered vertices, and we are interested in the evolution of these forests from the first vertex to the point that all vertices are uncovered.

This model is motivated by stochastic models known as coalescent models for particle coalescence, most notably the additive and the multiplicative coalescent [2] and the Kingman coalescent [7]. To make the distinction between these classical coalescent models and our model more explicit, let us briefly revisit the additive coalescent model (see [1]) as a prototypical example. This model describes a Markov process on a state space consisting of tuples $(x_1, x_2, \dots)$ with $x_1 \geq x_2 \geq \cdots \geq 0$ and $\sum_{i\geq 0} x_i = 1$ that model the fragmentation of a unit mass into clusters of mass x_i. Pairs of clusters with masses x_i and x_j then merge into a new cluster of mass $x_i + x_j$ at rate $x_i + x_j$. In the corresponding discrete time version of the process, exactly two clusters are merged in every time step. There are various combinatorial settings in which this discrete additive coalescent model appears, for example in the evolution of parking blocks in parking schemes related to "hashing with linear probing" [4], as the dual fragmentation process in the context of the random cutting of trees [6], and in a certain scheme for merging forests by uncovering one edge in every time step [10].

While the latter incarnation of the additive coalescent in which edges are uncovered successively is very much related in spirit to our vertex uncover model, the underlying processes are fundamentally different: these classical coalescent models rely on the fact that exactly two clusters are merged in every time step, which is not the case in our model. When uncovering a new vertex, a more or less arbitrary number of edges (including none at all) can be uncovered. There are coalescent models like the Λ-coalescent, a generalization of the Kingman coalescent [11], which allow for more than two clusters being merged – however, at present we are not aware of any known coalescent model that is able to capture the behavior of the vertex uncover process.

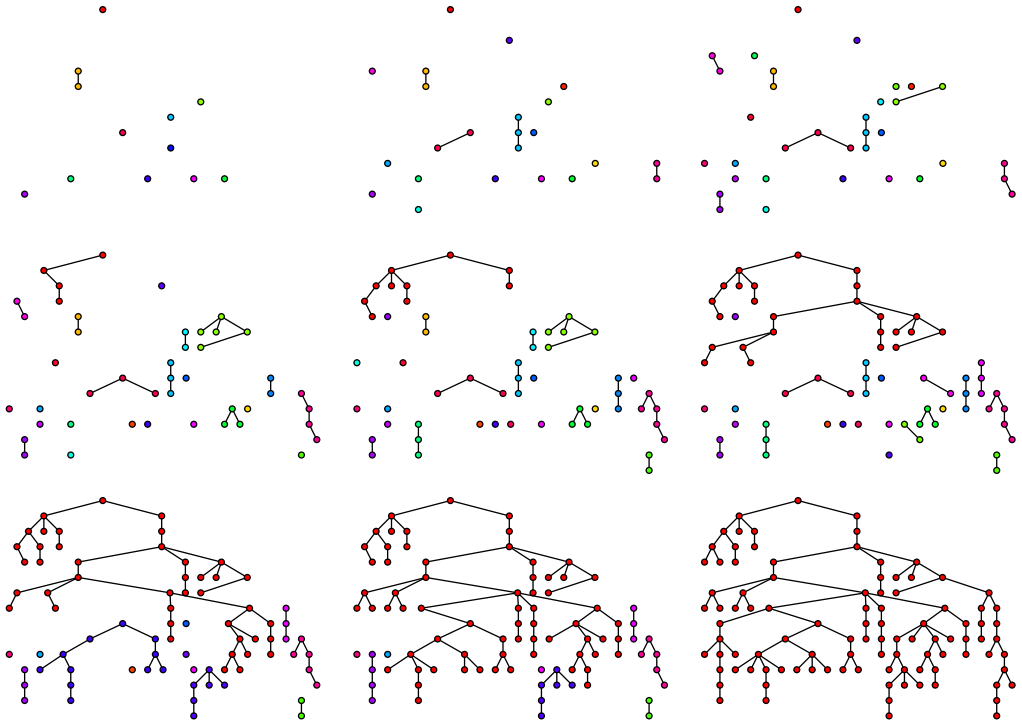


Figure 1 A few snapshots of the *uncover process* applied to a random labeled tree of size 100. From left to right and top to bottom, there are 12, 23, 34, ..., 89, and 100 uncovered vertices in the figures, respectively. Vertex labels are omitted for the sake of readability, and vertices are colored per connected component.

Overview. Different aspects of the uncover process on labeled trees are investigated in this extended abstract. In Section 2, we study the stochastic process given by the number of uncovered edges. The corresponding main result, a full characterization of the process and its limiting behavior, is given in Theorem 5. In this extended abstract, we sketch the proofs of the results in Section 2 – full details can be found in Appendix A.

Sections 3 and 4 are both concerned with cluster sizes, i.e., with the sizes of the connected components that are created throughout the process. In particular, in Section 3 we shift our attention to rooted labeled trees, to study the behavior of the component containing a fixed vertex. The expected size of the root cluster is analyzed in Theorem 9. Furthermore, we show that the number of rooted trees whose root cluster has a given size is given by a rather

simple enumeration formula – which, in turn, manifests in Theorem 11, a characterization of the different limiting distributions for the root cluster size depending on the number of uncovered vertices.

Finally, in Section 4 we use the results on the root cluster to draw conclusions regarding the size of the largest cluster in the tree.

Notation. Throughout this work we use the notation $[n] = \{1, \dots, n\}$ and $[k, \ell] = \{k, k+1, \dots, \ell\}$ for discrete intervals, and $x^{\underline{j}} = x(x-1)\cdots(x-j+1)$ for the falling factorials. The floor and ceiling function are denoted by $\lfloor x \rfloor$ and $\lceil x \rceil$, respectively. Furthermore, we use $\mathcal{T}$ and $\mathcal{T}^\bullet$ for the combinatorial classes of labeled trees and rooted labeled trees, respectively, and $\mathcal{T}_n$ and $\mathcal{T}_n^\bullet$ for the classes of labeled and rooted labeled trees of size n, i.e., with n vertices. Finally, we use $X_n \xrightarrow{d} X$ and $X_n \xrightarrow{p} X$ to denote convergence in distribution resp. probability of a sequence of random variables (r.v.) $(X_n)_{n\geq 0}$ to the r.v. X.

2 The number of uncovered edges

In this section our main interest is the behavior of the number of uncovered edges in the uncover process. We begin by introducing a formal parameter for this quantity.

▶ **Definition 1.** *Let T be a labeled tree with vertex set $V(T) = [n]$. For $1 \leq j \leq n$, we let $k_j(T) := \|T[1, 2, \dots, j]\|$ denote the number of edges in the subgraph of T induced by the vertices in $[j]$. We refer to the sequence $(k_j(T))_{1\leq j\leq n}$ as the* (edge) uncover sequence.

We start with a few simple observations. First, for any labeled tree of order n we have $k_1(T) = 0$, as well as $k_n(T) = n - 1$. Second, as any induced subgraph of a tree is a forest, and as forests have the elementary property that the number of edges together with the number of connected components gives the order of the forest, we find that $j - k_j(T)$ is the number of connected components after uncovering the first j vertices of T. Figure 2 illustrates the progression of the number of edges and the number of connected components for $1 \leq j \leq 1000$ in a randomly chosen labeled tree on 1000 vertices.

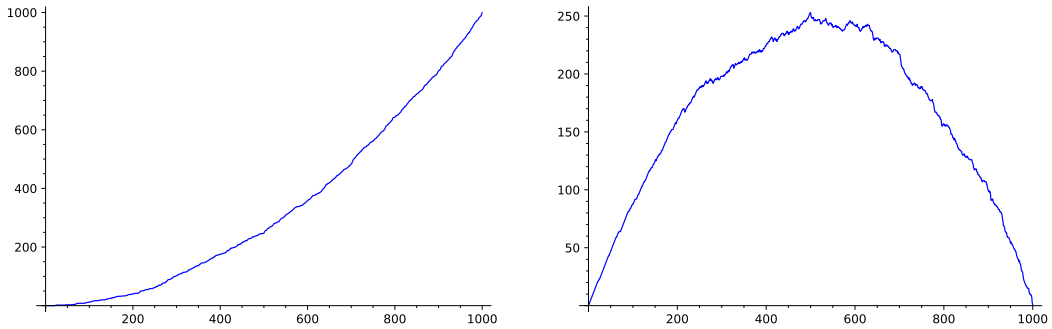


Figure 2 Progression of the number of edges (left) and the number of connected components (right) when sequentially uncovering a random labeled tree on 1000 vertices.

Moreover, the fact that the first j vertices of the tree induce a forest also yields the sharp bound $0 \leq k_j(T) \leq j - 1$ for all $1 \leq j \leq n - 1$. The lower bound is attained by the star with central vertex n, and the upper bound is attained by the (linearly ordered) path. We can also observe that as soon as $k_{n-1}(T) > 0$, the set of edges added in the last uncover step is not determined uniquely. Thus, the star with central vertex n is the only tree which is fully determined by its uncover sequence.

The following theorem provides explicit enumeration formulas for the number of trees with a partially and fully specified uncover sequence, respectively.

▶ **Theorem 2.** *Let r be a fixed positive integer with $1 \le r < n-1$, let $j_1, j_2, \ldots, j_r$ be positive integers with $1 < j_1 < j_2 < \cdots < j_r < n$, and let $a_1, a_2, \ldots, a_r$ be a nondecreasing sequence of nonnegative integers satisfying $a_i \le j_i - 1$ for all $1 \le i \le r$. Additionally, let $j_0 = 1$ and $a_0 = 0$. Then, the number of rooted labeled trees T of order n that satisfy $k_{j_i}(T) = a_i$ for all $1 \le i \le r$ is given by*

$$(n-j_r)^{j_r-a_r-1} n^{n-j_r-1} \times \prod_{i=1}^{r} \left(\sum_{h=0}^{a_i-a_{i-1}} \binom{j_{i-1}-a_{i-1}-1}{h} \binom{j_i-j_{i-1}}{a_i-a_{i-1}-h} (j_i-j_{i-1})^h j_i^{a_i-a_{i-1}-h} \right). \quad (1)$$

We first derive a helpful auxiliary result, namely an explicit formula for the corresponding (multivariate) generating function. The enumeration formula will then follow by extracting the appropriate coefficients.

▶ **Lemma 3.** *In the setting of Theorem 2, the multivariate generating function for the increments in the edge uncover process is given by*

$$E_n(z_1, z_2, \ldots, z_r) = n^{n-j_r-1} \prod_{i=1}^{r} \Big(n - j_r + j_i z_i + \sum_{h=i+1}^{r} (j_h - j_{h-1}) z_h \Big)^{j_i - j_{i-1}}. \quad (2)$$

In other words, the coefficient of the monomial $z_1^{a_1} z_2^{a_2-a_1} \ldots z_r^{a_r-a_{r-1}}$ in the expansion of $E_n(z_1, \ldots, z_r)$ is the number of labeled trees T of order n with $k_{j_i}(T) = a_i$ for all $1 \le i \le r$.

▶ Remark 4. By specifying the integers $j_1, j_2, \ldots, j_r$, the uncover process is effectively partitioned into intervals. This is also reflected by the quantities occurring in the product in (2): the difference $j_i - j_{i-1}$ corresponds to the number of vertices uncovered in the i-th interval, j_i represents the number of vertices uncovered in total up to the i-th interval, and $n - j_r$ corresponds to the number of vertices uncovered in the last interval.

Proof of Lemma 3. We begin by observing that when the process uncovers the vertex with label j, edges to all adjacent vertices whose label is less than j are uncovered as well. To determine the generating function of the edge increments, we assign the weight $x_i y_j$ to the edge connecting vertex i and vertex j with $i < j$, and then consider the generating function for the tree weight $w(T)$ (which is defined as the product of the edge weights); $E_n(z_1, \ldots, z_r) = \sum_{|T|=n} w(T)$.

Following Martin and Reiner [9, Theorem 4] or Remmel and Williamson [12, Equation (8)], the generating function of the tree weights $w(T)$ has the explicit formula

$$\sum_{|T|=n} w(T) = x_1 y_n \prod_{j=2}^{n-1} \Big(\sum_{i=1}^{n} x_{\min(i,j)} y_{\max(i,j)} \Big). \quad (3)$$

As initially observed, edges that are counted by $k_{j_i}(T)$ are precisely those that induce a factor y_ℓ for some $\ell \le j_i$. Thus we make the following substitutions: $x_\ell = 1$ for all ℓ, $y_\ell = z_i$ if and only if $j_{i-1} < \ell \le j_i$ (where[1] $j_0 = 1$), and $y_\ell = 1$ if $\ell > j_r$. To deal with the sum over $y_{\max(i,j)}$, observe that we can rewrite it as

$$\sum_{i=1}^{n} y_{\max(i,j)} = n - j_r + \sum_{i=1}^{j_1} y_{\max(i,j)} + \cdots + \sum_{i=j_{r-1}+1}^{j_r} y_{\max(i,j)}.$$

[1] Observe that y_1 does not occur, since at least one of the ends of every edge has a label greater than 1.

In this form, the different values assumed by the sum when j moves through the ranges $1 < j \leq j_1$, $j_1 < j \leq j_2$, etc. can be determined directly. For some j with $j_{i-1} < j \leq j_i$, the contribution to the product in (3) is

$$n - j_r + j_i z_i + \sum_{h=i+1}^{r} (j_h - j_{h-1}) z_h,$$

and for $j_r < j \leq n-1$ all y-variables are replaced by 1, so that the contribution to the product is a factor n. Putting both of these observations together shows that the right-hand side of (3) can be rewritten as the right-hand side of (2) and thus proves the lemma. ◀

With an explicit formula for the generating function of edge increments in the uncover process at our disposal, an explicit formula for the number of trees with given (partial) uncover sequence follows as a simple consequence.

Proof of Theorem 2. It remains to extract the coefficient of $z_1^{a_1} z_2^{a_2-a_1} \cdots z_r^{a_r-a_{r-1}}$, which is done step by step, starting with z_1. ◀

2.1 A closer look at the stochastic process

The exceptionally nice formula for the generating function of edge increments can be used to study the stochastic process that describes the number of uncovered edges in more detail. Let the sequence of random variables $(K_j^{(n)})_{1 \leq j \leq n}$ be the discrete stochastic process modeling the number of uncovered edges after uncovering the first j vertices in a random labeled tree of size n, chosen uniformly at random. The expected number of uncovered edges can be determined by a simple argument: with j uncovered vertices, $\binom{j}{2}$ of the $\binom{n}{2}$ possible positions for the edges have been uncovered. As every position is, due to symmetry and the uniform choice of the labeled tree, equally likely to hold one of the $n-1$ edges, we find

$$\mathbb{E}K_j^{(n)} = (n-1)\frac{\binom{j}{2}}{\binom{n}{2}} = \frac{j(j-1)}{n}. \tag{4}$$

To motivate our investigations further, consider the illustrations in Figure 3. The rescaled deviation from the mean is reminiscent of a stochastic process known as Brownian bridge.

In order to define this process formally, recall first that the Wiener process $(W(t))_{t \in [0,1]}$ is the unique stochastic process that satisfies

- $W(0) = 0$,
- W has independent, stationary increments,
- $W(t) \sim \mathcal{N}(0, t)$ for all $t > 0$, and
- $t \mapsto W(t)$ is almost surely continuous,

see [8, Definition 21.8]. A Brownian bridge can then be defined by setting

$$B(t) = (1-t)W(t/(1-t)), \tag{5}$$

see e.g. [13, Exercise 3.10]. The term "bridge" results from the fact that we have $B(0) = B(1) = 0$.

While the (normalized) deviation from the mean looks like it might converge to a Brownian bridge, we will prove that this is only *almost* the case. The following theorem characterizes the stochastic process. For technical purposes, we set $K_0^{(n)} = 0$ and introduce the linearly interpolated process $(\tilde{K}_t^{(n)})_{t \in [0,1]}$, where

$$\tilde{K}_t^{(n)} := (1 + \lfloor tn \rfloor - tn) K_{\lfloor tn \rfloor}^{(n)} + (tn - \lfloor tn \rfloor) K_{\lceil tn \rceil}^{(n)}, \tag{6}$$

which by construction has continuous paths.

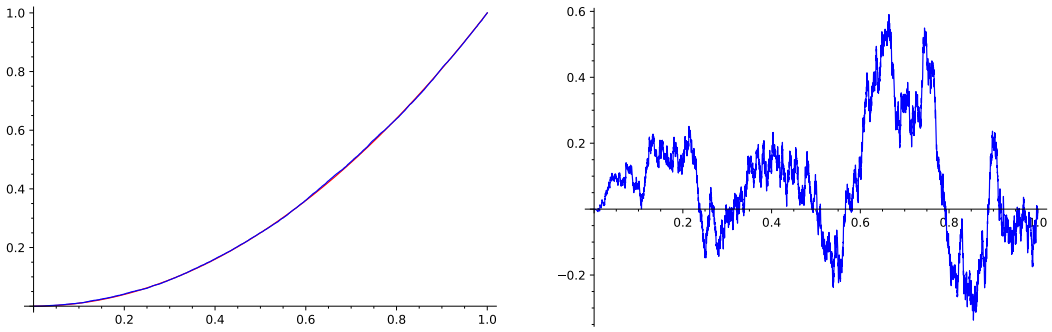


Figure 3 A path of the rescaled stochastic process $(K^{(n)}_{\lfloor tn \rfloor}/(n-1))_{t\in[0,1]}$ (left-hand side) and the corresponding (rescaled) deviation $\left(\frac{\tilde{K}^{(n)}_t - t^2 n}{\sqrt{n}}\right)_{t\in[0,1]}$ for a random labeled tree of size $n = 10000$.

▶ **Theorem 5.** *Let $(Z^{(n)}(t))_{t\in[0,1]}$ be the continuous stochastic process resulting from centering and rescaling the linearly interpolated process $(\tilde{K}^{(n)}_t)_{t\in[0,1]}$ in the form of*

$$Z^{(n)}(t) := \frac{\tilde{K}^{(n)}_t - t^2 n}{\sqrt{n}},$$

and let $(W(t))_{t\in[0,1]}$ be the standard Wiener process. Then, for $n \to \infty$, the rescaled process converges weakly with respect to the sup*-norm on $C([0,1])$ to a limiting process $Z^\infty(t)$ that is given by*

$$Z^\infty(t) = (1-t)W\big(t^2/(1-t)\big). \tag{7}$$

Furthermore, for s, $t \in [0,1]$ with $s < t$, the limiting process satisfies

$$\mathbb{E}Z^\infty(t) = 0, \quad \mathbb{V}Z^\infty(t) = t^2(1-t), \quad \text{and} \quad \operatorname{Cov}(Z^\infty(s), Z^\infty(t)) = s^2(1-t). \tag{8}$$

▶ Remark 6. While the limiting process $(Z^\infty(t))_{t\in[0,1]}$ is not a Brownian bridge (the corresponding variances and covariances as given in (8) do not match), it is closely related. Comparing the characterization of $Z^\infty(t)$ in (7) to (5), we see that the processes only differ by the square in the numerator of the argument of the Wiener process.

Two main ingredients are required to prove this result: a uniform tightness bound on the one hand, and information on the finite-dimensional joint distributions of $(\tilde{K}^{(n)}_t)_{t\in[0,1]}$ on the other hand.

To prove tightness, let us begin by revisiting (2). Given Cayley's well-known tree enumeration formula, the corresponding probability generating function for the complete uncover sequence, i.e., when we choose our integer vector as $\mathbf{j} = (2, 3, \ldots, n-1)$, is

$$P_n(z_2, \ldots, z_{n-1}) = \prod_{i=2}^{n-1} \Big(\frac{1}{n} + \frac{i}{n} z_i + \sum_{h=i+1}^{n-1} \frac{1}{n} z_h \Big). \tag{9}$$

This suggests modeling the process with $n-2$ independent random variables, each representing an edge increment[2]. The factorization suggests that the j-th increment (which corresponds to the factor with $i = j+1$) happens with probability $(j+1)/n$ when the vertex with label $j+1$

[2] We explicitly model edge increments here instead of edges, because with this approach we do not need to care about *which* edge is being uncovered. Our model explicitly only captures the behavior of the number of uncovered edges.

is uncovered, or with probability $1/n$ every time any of the subsequent vertices are uncovered. This probabilistic point of view can be used to construct a recursive characterization for the number of uncovered edges, namely[3]

$$K_{j+1}^{(n)} = K_j^{(n)} + \mathrm{Ber}\Big(\frac{j+1}{n}\Big) + \mathrm{Bin}\Big(j-1-K_j^{(n)}, \frac{1}{n-j}\Big). \tag{10}$$

The Bernoulli variable models the probability that the j-th edge increment is added when uncovering the vertex with label $j+1$, and the binomial variable models all of the remaining, not yet uncovered edge increments.

Now let us consider a centered and rescaled version of the process $(K_j^{(n)})_{1\le j\le n}$ by defining

$$Y_j^{(n)} := \frac{K_j^{(n)} - \frac{j(j-1)}{n}}{n-j}. \tag{11}$$

With the help of the recursive description in (10), one can show that $(Y_j^{(n)})_{1\le j\le n-1}$ is a martingale, see Appendix A.1.

We are now ready to state and prove the first required ingredient.

▶ **Lemma 7.** *For any real $C>1$ and any positive integer n, the random variable $Z^{(n)}(t)$ satisfies the tightness bound*

$$\mathbb{P}\big(\sup_{t\in[0,1]} |Z^{(n)}(t)| \ge C\big) \le 4(C-1)^{-2}, \tag{12}$$

so that for $C\to\infty$, the probability for the process to exceed C in absolute value converges to 0 uniformly in terms of n.

Sketch of proof. One first shows that $\sup_{t\in[0,1]}|Z^{(n)}(t)|$ can be bounded in terms of the deviation of the discrete process $(K_j^{(n)})_{0\le j\le n}$ from its mean. The bound then follows after expressing the discrete process in terms of the martingale $Y_j^{(n)}$ constructed above, partitioning the discrete interval $[0,n]$ appropriately and applying both Doob's L^p-inequality and the union bound. See Appendix A.2 for details. ◀

▶ **Lemma 8.** *Let r be a fixed positive integer, and let $\mathbf{t}=(t_1,\dots,t_r)\in(0,1)^r$. Then for $n\to\infty$, the random vector*

$$\mathbf{K}_{\lfloor \mathbf{t}n\rfloor}^{(n)} := (K_{\lfloor t_1 n\rfloor}^{(n)}, K_{\lfloor t_2 n\rfloor}^{(n)}, \dots, K_{\lfloor t_r n\rfloor}^{(n)})$$

converges, after centering and rescaling, for $n\to\infty$ in distribution to a multivariate normal distribution,

$$\frac{\mathbf{K}_{\lfloor \mathbf{t}n\rfloor}^{(n)} - \mathbb{E}\mathbf{K}_{\lfloor \mathbf{t}n\rfloor}^{(n)}}{\sqrt{n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(\mathbf{0},\Sigma),$$

where the expectation vector $\mathbb{E}\mathbf{K}_{\lfloor \mathbf{t}n\rfloor}^{(n)}$ satisfies

$$\mathbb{E}\mathbf{K}_{\lfloor \mathbf{t}n\rfloor}^{(n)} = n(t_1^2, t_2^2, \dots, t_r^2) + O(1), \tag{13}$$

and the entries of the variance-covariance matrix $\Sigma=(\sigma_{i,j})_{1\le i,j\le r}$ are

$$\sigma_{i,j} = \begin{cases} t_i^2(1-t_j) & \text{if } i\le j,\\ t_j^2(1-t_i) & \text{if } i>j. \end{cases} \tag{14}$$

[3] We slightly abuse notation: formally, we would need to introduce auxiliary variables that are distributed according to the specified binomial and Bernoulli distributions.

Sketch of proof. The factorization of the probability generating function (9) implies that the distribution of the corresponding random vector $\Delta^{(n)}_{\lfloor \mathbf{t}n \rfloor} = (K^{(n)}_{\lfloor t_1 n \rfloor}, K^{(n)}_{\lfloor t_2 n \rfloor} - K^{(n)}_{\lfloor t_1 n \rfloor}, \dots, K^{(n)}_{\lfloor t_r n \rfloor} - K^{(n)}_{\lfloor t_{r-1} n \rfloor})$ is a marginal distribution of the sum of r independent, multinomially distributed random vectors. By the multivariate central limit theorem, $\Delta^{(n)}_{\lfloor \mathbf{t}n \rfloor}$ converges to a multivariate normal distribution – and as a consequence, so does $\mathbf{K}^{(n)}_{\lfloor \mathbf{t}n \rfloor}$.

The variance-covariance matrix of the centered and rescaled vector can be obtained, for example, by using the martingale constructed above. See again Appendix A.2 for details. ◀

Proof of Theorem 5. The proof relies on the well-known result asserting that given tightness of the sequence of corresponding probability measures as well as convergence of the finite-dimensional probability distributions, a sequence of stochastic processes converges to a limiting process (see [3, Theorem 7.1, Theorem 7.5]).

Tightness is implied by the uniform bound (12) derived in Lemma 7. The (limiting) behavior of the finite-dimensional distributions for the original process $(K^{(n)}_{\lfloor tn \rfloor})_{t\in[0,1]}$ is characterized by Lemma 8. This characterization carries over to the linearly interpolated process by an application of Slutsky's theorem [8, Theorem 13.18] after observing that

$$\mathbb{P}\left(\left|Z^{(n)}(t) - \frac{K^{(n)}_{\lfloor tn \rfloor} - t^2 n}{\sqrt{n}}\right| > \varepsilon\right) = \mathbb{P}\left(\frac{|\tilde{K}^{(n)}_t - K^{(n)}_{\lfloor tn \rfloor}|}{\sqrt{n}} > \varepsilon\right) \le \frac{\mathbb{E}((\tilde{K}^{(n)}_t - K^{(n)}_{\lfloor tn \rfloor})^2)}{n\varepsilon^2}$$
$$\le \frac{\mathbb{E}((K^{(n)}_{\lfloor tn \rfloor+1} - K^{(n)}_{\lfloor tn \rfloor})^2)}{n\varepsilon^2} \xrightarrow{n\to\infty} 0,$$

as a mechanical computation shows that $\mathbb{E}((K^{(n)}_{\lfloor tn \rfloor+1} - K^{(n)}_{\lfloor tn \rfloor})^2) = O(1)$.

Note that as the finite-dimensional distributions converge to Gaussian distributions, the limiting process $(Z^\infty(t))_{t\in[0,1]}$ is Gaussian itself – which means that it is fully characterized by its first and second order moments. As a consequence of Lemma 8, we find for all $s, t \in [0,1]$ with $s < t$ that

$$\mathbb{E}Z^\infty(t) = 0, \qquad \mathbb{V}Z^\infty(t) = t^2(1-t), \qquad \mathrm{Cov}(Z^\infty(s), Z^\infty(t)) = s^2(1-t).$$

It can be checked that if $(W(t))_{t\in[0,1]}$ is a standard Wiener process, the Gaussian process $((1-t)W(t^2/(1-t)))_{t\in[0,1]}$ has the same first and second order moments and therefore also the same distribution as Z^∞. ◀

3 Size of the root cluster

We now shift our attention from the number of uncovered edges to the sizes of the connected components (or *clusters*) appearing in the graph throughout the uncover process. It will prove convenient to change our tree model to *rooted* labeled trees, as the nature of rooted trees allows us to focus our investigation on one particular cluster – the one containing the root vertex. In case the root vertex has not yet been uncovered, we will consider the size of the root cluster to be 0. Formally, we let the random variable $R^{(k)}_n$ be the size of the root cluster of a (uniformly) random rooted labeled tree of size n with k uncovered vertices.

Using the symbolic method for labeled structures (cf. [5, Chapter II]), we can set up a formal specification for the corresponding combinatorial classes and subsequently extract functional equations for the associated generating functions. Let $\mathcal{T}^\bullet$ be the class of rooted labeled trees, and let $\mathcal{G}$ be a refinement of $\mathcal{T}^\bullet$ where the vertices can either be covered or uncovered, and where uncovered vertices are marked with a marker U. Finally, let $\mathcal{F}$ be

a further refinement of $\mathcal{G}$ where all uncovered vertices in the root cluster are additionally marked with marker V. A straightforward "top-down approach", i.e., a decomposition of the members of the tree family w.r.t. the root vertex, yields the formal specification

$$\mathcal{F} = \mathcal{Z} * \textsc{Set}(\mathcal{G}) + \mathcal{Z} \times \{U, V\} * \textsc{Set}(\mathcal{F}), \qquad \mathcal{G} = \mathcal{Z} * \textsc{Set}(\mathcal{G}) + \mathcal{Z} \times \{U\} * \textsc{Set}(\mathcal{G}).$$

Introducing the corresponding exponential generating functions $F := F(z,u,v)$ and $G := G(z,u)$, we obtain the characterizing equations

$$F = ze^G + zuve^F, \qquad G = z(1+u)e^G. \tag{15}$$

Of course, $G(z,u) = T^\bullet(z(1+u))$, where $T^\bullet$ is the exponential generating function associated with $\mathcal{T}^\bullet$, the Cayley tree function. Starting with (15), the following results on $R_n^{(k)}$ can be deduced.

▶ **Theorem 9.** *The expectation $\mathbb{E}(R_n^{(k)})$ is, for $0 \le k \le n$ and $n \ge 1$, given by*

$$\mathbb{E}(R_n^{(k)}) = \sum_{j=1}^{k} \frac{j\, k^{\underline{j}}}{n^j}. \tag{16}$$

Depending on the growth of $k = k(n)$, $\mathbb{E}(R_n^{(k)})$ has the following asymptotic behavior:

$$\mathbb{E}(R_n^{(k)}) \sim \begin{cases} \frac{k}{n}, & \text{for } k = o(n), \quad \textit{(k small)}, \\ \frac{\alpha}{(1-\alpha)^2}, & \text{for } k \sim \alpha n, \text{ with } 0 < \alpha < 1, \quad \textit{(k in central region)}, \\ \frac{n^2}{d^2}, & \text{for } k = n - d, \text{ with } d = \omega(\sqrt{n}) \text{ and } d = o(n), \\ & \quad \textit{(k subcritically large)}, \\ \kappa n, \quad \text{with} & \kappa = 1 - ce^{\frac{c^2}{2}} \int_c^\infty e^{-\frac{t^2}{2}}\, dt, \\ & \text{for } k = n - d, \text{ with } d \sim c\sqrt{n} \text{ and } c > 0, \quad \textit{(k critically large)}, \\ n - \sqrt{\frac{\pi}{2}} d\sqrt{n}, & \text{for } k = n - d, \text{ with } d = o(\sqrt{n}), \quad \textit{(k supercritically large)}. \end{cases}$$

Proof. After introducing $E := E(z,u) = \frac{\partial}{\partial v} F(z,u,v)\big|_{v=1} = \sum_{n,k} \frac{n^{n-1}}{n!} \binom{n}{k} \mathbb{E}(R_n^{(k)})$, considering the partial derivative of (15) with respect to v yields

$$E = \frac{1}{1 - \frac{u}{1+u} G} - 1.$$

Extracting coefficients and using $\mathbb{E}(R_n^{(k)}) = \frac{[z^n u^k] E}{[z^n u^k] G}$, we obtain (16). In order to analyze the asymptotic behavior of $\mathbb{E}(R_n^{(k)})$, the integral representation

$$\mathbb{E}(R_n^{(k)}) = \int_0^\infty (x-1)\, e^{-x} \Big(1 + \frac{x}{n}\Big)^k dx,$$

which can be verified in a straightforward way, turns out to be advantageous. Expanding the integrand and distinguishing several cases yields the asymptotic results given in the theorem. ◀

We can even obtain the exact distribution of $R_n^{(k)}$. There are two different approaches we want to briefly sketch: for one, an explicit formula for the generating function $F = F(z,u,v)$ can be found either from manipulating the recursive description (15), or directly by decomposing $\mathcal{F}$ as a tree forming the uncovered root cluster with a forest with covered roots attached. Either way, this yields

$$F = T^\bullet\big(vXe^{-X}\big) + \frac{G}{1+u}, \quad \text{with} \quad X = \frac{uG}{1+u}.$$

Extracting coefficients via an application of the Lagrange inversion formula (see, e.g., [5, Theorem A.2]) then yields an explicit formula for $F_{n,k,m} = n![z^n u^k v^m]F(z,u,v)$, i.e., the number of labeled rooted trees with n vertices of which k are uncovered and m belong to the root cluster (for $0 \le m \le k \le n$ and $n \ge 1$):

$$F_{n,k,m} = \begin{cases} \binom{n-1}{k} n^{n-1}, & m = 0, \\ \binom{n}{m}\binom{n-m-1}{k-m} n^{n-k-1} m^m (n-m)^{k-m}, & m \ge 1. \end{cases}$$

From this formula, the probabilities $\mathbb{P}(R_n^{(k)} = r)$ given below can be obtained directly.

Alternatively, there is also a more combinatorial approach to determine these probabilities: there is an elementary formula enumerating trees where a specified set of vertices forms a cluster.

▷ Claim 10. The number of trees on $[n]$ that do not have any edges between the vertex sets $[r]$ and $[r+1,k]$, where additionally the induced subgraph on $[r]$ is a tree itself, is given by

$$r^{r-1} n^{n-k-1} (n-k)(n-r)^{k-r-1}. \tag{17}$$

Proof of Claim 10. By Cayley's tree enumeration formula, the number of labeled trees on $[r]$ is r^{r-2}. Hence, the statement of the claim is equivalent to

$$r n^{n-k-1} (n-k)(n-r)^{k-r-1} \tag{18}$$

enumerating all rooted forests on $[n]$ whose roots are the vertices in $[r]$ that do not have any edges between $[r]$ and $[r+1,k]$. This can be proved bijectively with a construction similar to the Prüfer code, or by an application of Kirchhoff's Matrix-Tree Theorem [15, Theorem 5.6.8]. Details can be found in the full version of this extended abstract. ◁

As a consequence, the probability $\mathbb{P}(R_n^{(k)} = r)$ can be obtained by multiplying the number of these trees with $r\binom{k}{r}$ (to account for which vertices in $[k]$ form the root cluster and for the choice of the root), and then normalizing by n^{n-1}, the number of labeled rooted trees on n vertices.

▶ **Theorem 11.** *The exact distribution of $R_n^{(k)}$ is characterized by the following probability mass function (p.m.f.), which is given as follows for $0 \le m \le k \le n$ and $n \ge 1$ (and 0 otherwise):*

$$\mathbb{P}(R_n^{(k)} = m) = \begin{cases} 1 - \frac{k}{n}, & \text{for } m = 0, \\ \frac{m^m (n-k)(n-m)^{k-m-1}}{n^k}\binom{k}{m}, & \text{for } 1 \le m \le k < n, \\ 1, & \text{for } m = k = n. \end{cases}$$

Depending on the growth of $k = k(n)$, we obtain the following limiting behavior:

- *k small, i.e., $k = o(n)$:*

  $$R_n^{(k)} \xrightarrow{d} 0.$$

- *k in central region, i.e., $k \sim \alpha n$ with $0 < \alpha < 1$:*

  $$R_n^{(k)} \xrightarrow{d} R_\alpha, \quad \text{where the discrete r.v. } R_\alpha \text{ is characterized by its p.m.f.}$$

  $$\mathbb{P}(R_\alpha = m) =: p_m = \begin{cases} 1 - \alpha, & m = 0, \\ \frac{m^m}{m!}(1-\alpha)\alpha^m e^{-\alpha m}, & m \ge 1, \end{cases}$$

  or alternatively by the probability generating function $p(v) = \sum_{m \ge 0} p_m v^m = \frac{1-\alpha}{1 - T^\bullet(v\alpha e^{-\alpha})}$.

- *k subcritically large, i.e., $k = n - d$ with $d = \omega(\sqrt{n})$ and $d = o(n)$:*

$$\left(\frac{d}{n}\right)^2 \cdot R_n^{(k)} \xrightarrow{d} \textsc{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right),$$

where $\textsc{Gamma}(\frac{1}{2}, \frac{1}{2})$ is a Gamma-distribution characterized by its density $f(x) = \frac{1}{\sqrt{2\pi x}}\, e^{-\frac{x}{2}}$, for $x > 0$.

- *k critically large, i.e., $k = n - d$ with $d \sim c\sqrt{n}$ and $c > 0$:*

$$\frac{1}{n} \cdot R_n^{(k)} \xrightarrow{d} R(c),$$

where the continuous r.v. $R(c)$ is characterized by its density $f_c(x) = \frac{1}{\sqrt{2\pi}} \frac{c}{\sqrt{x}(1-x)^{\frac{3}{2}}}\, e^{-\frac{c^2 x}{2(1-x)}}$, for $0 < x < 1$.

- *k supercritically large, i.e., $k = n - d$ with $d = \omega(1)$ and $d = o(\sqrt{n})$:*

$$\frac{1}{d^2} \cdot \left(n - R_n^{(k)}\right) \xrightarrow{d} D,$$

where the continuous r.v. D is characterized by its density $f(x) = \frac{1}{\sqrt{2\pi}\, x^{\frac{3}{2}}}\, e^{-\frac{1}{2x}}$, $x > 0$.

- *k supercritically large with fixed difference, i.e., $k = n - d$ with d fixed:*

$$n - d - R_n^{(k)} \xrightarrow{d} D(d),$$

where the discrete r.v. $D(d)$ is characterized by the p.m.f.

$$\mathbb{P}(D(d) = j) =: p_j = e^{-d} \cdot \frac{d(d+j)^{j-1}}{j!} \cdot e^{-j}, \quad j \geq 0,$$

or alternatively via the probability generating function $p(v) = \sum_{j \geq 0} p_j v^j = e^{d(T^\bullet(\frac{v}{e}) - 1)}$.

Proof. The probability mass function of $R_n^{(k)}$ follows from the considerations made before the statement of the theorem. Due to its explicit nature, the limiting distribution results stated in Theorem 11 can be obtained in a rather straightforward way by applying Stirling's formula for the factorials after distinguishing several cases. ◀

▶ Remark 12. Of course, for labeled trees, the distribution of $R_n^{(k)}$ matches with the distribution of the cluster size of a random vertex. Furthermore, by conditioning, one can easily transfer the results of Theorem 11 to results for the size $S_n^{(k)}$ of the cluster of the k-th uncovered vertex: $\mathbb{P}(S_n^{(k)} = m) = \mathbb{P}(R_n^{(k)} = m | R_n^{(k)} > 0) = \frac{n}{k} \cdot \mathbb{P}(R_n^{(k)} = m)$, for $m \geq 1$.

4 Size of the largest uncovered component

With knowledge about the behavior of the root cluster at our disposal, we return to non-rooted labeled trees and study the size of the largest cluster. To this aim, we introduce the random variable $X_{n,r}^{(k)}$ which models the number of components of size r after uncovering the vertices 1 to k in a uniformly random labeled tree of size n.

Formally, $X_{n,r}^{(k)} \colon \mathcal{T}_n \to \mathbb{Z}_{\geq 0}$. Note that we have, for all labeled trees $T \in \mathcal{T}_n$,

$$\sum_{r=1}^{n} r \cdot X_{n,r}^{(k)}(T) = k. \tag{19}$$

▶ **Theorem 13.** *Let $n, k, r \in \mathbb{Z}_{\geq 0}$ with $0 \leq r \leq k \leq n$. The expected number of connected components of size r after uncovering k vertices of a labeled tree of size n chosen uniformly at random is*

$$\mathbb{E}X_{n,r}^{(k)} = \binom{k}{r}\left(\frac{r}{n}\right)^{r-1}\left(1-\frac{k}{n}\right)\left(1-\frac{r}{n}\right)^{k-r-1}. \tag{20}$$

Sketch of proof. Observe that $X_{n,r}^{(k)}$ can be written as a sum of Bernoulli random variables

$$X_{n,r}^{(k)} = \sum_{\substack{S\subseteq[k]\\|S|=r}} X_{n,S}^{(k)},$$

with $X_{n,S}^{(k)}$ being 0 or 1 depending on whether or not the vertices in S form a cluster after k uncover steps. By symmetry and linearity of the expected value, we have

$$\mathbb{E}X_{n,r}^{(k)} = \sum_{\substack{S\subseteq[k]\\|S|=r}} \mathbb{E}X_{n,S}^{(k)} = \binom{k}{r}\mathbb{E}X_{n,[r]}^{(k)}.$$

A formula for the expected value on the right-hand side follows from Claim 10, and thus proves the theorem. ◀

In the spirit of the observation in (19), the formula in Claim 10 provides a combinatorial proof for the following summation identity.

▶ **Corollary 14.** *Let $n, k \in \mathbb{Z}_{\geq 0}$ with $0 \leq k \leq n$. Then, the identity*

$$\sum_{r=1}^{k}\binom{k}{r} r^r n^{n-k-1}(n-r)^{k-r-1}(n-k) = kn^{n-2} \tag{21}$$

holds.

Proof. The right-hand side enumerates the vertices in $[k]$ in all labeled trees on n vertices. The left-hand side does the same, with the summands enumerating the vertices in connected components of size r. ◀

▶ Remark 15. Observe that the identity in (21) can be rewritten as

$$\sum_{r=1}^{k}\binom{k}{r} r^r (n-r)^{k-r-1} = \frac{k}{n-k}n^{k-1},$$

which is a specialized form of Abel's Binomial Theorem – a classical, and well-known result; see, e.g., [14].

For a tree $T \in \mathcal{T}$, let $c_{\max}^{(k)}(T)$ denote the largest connected component of T after uncovering the first k vertices.

▶ **Theorem 16.** *Let $n \in \mathbb{Z}_{\geq 0}$, and let $T_n \in \mathcal{T}_n$ be a tree chosen uniformly at random. Then the behavior of the random variable $c_{\max}^{(k)}(T_n)$ as $n \to \infty$ can be described as follows:*

- *for $k = n - d$ with $d = \omega(\sqrt{n})$ (*subcritical case*), we have $c_{\max}^{(k)}(T_n)/n \xrightarrow{p} 0$.*
- *for $k = n - d$ with $d \sim c\sqrt{n}$ for a constant c (*critical case*), the rescaled random variable $c_{\max}^{(k)}(T_n)/n$ converges weakly to a (non-degenerate) continuous limiting distribution,*
- *for $k = n - d$ with $d = o(\sqrt{n})$ (*supercritical case*), we have $c_{\max}^{(k)}(T_n)/n \xrightarrow{p} 1$. With high probability, there is one "giant" component whose size is asymptotically equal to n.*

Sketch of proof. For the subcritical case, we use the expected root cluster size from Theorem 9. Since a cluster of size r contains the root with probability $\frac{r}{n}$, we have

$$\frac{n^2}{d^2} \sim \mathbb{E}R_n^{(n-d)} = \sum_{r=0}^{n-d} \mathbb{E}X_{n,r}^{(n-d)} \cdot r \cdot \frac{r}{n} \geq \sum_{r=m}^{n-d} \mathbb{E}X_{n,r}^{(n-d)} \frac{r^2}{n} \geq \frac{m^2}{n} \sum_{r=m}^{n-d} \mathbb{E}X_{n,r}^{(n-d)}$$
$$\geq \frac{m^2}{n} \mathbb{P}(c_{\max}^{(n-d)}(T_n) \geq m).$$

This implies that

$$\mathbb{P}(c_{\max}^{(n-d)}(T_n) \geq m) = O\Big(\frac{n^3}{d^2 m^2}\Big),$$

so if $m = \epsilon n$ for any fixed $\epsilon > 0$, we have $\mathbb{P}(c_{\max}^{(n-d)}(T_n) \geq m) \to 0$.

In the critical case, we first consider the situation that there is a cluster that contains more than half of the vertices. Clearly, such a cluster must be the largest cluster and the only cluster of its size. Therefore, if $r = \rho n$ with $\rho > \frac{1}{2}$, we have

$$\mathbb{P}(c_{\max}^{(k)}(T_n) = r) = \mathbb{E}X_{n,r}^{(k)} \sim \frac{1}{n} \frac{e^{-\frac{c^2}{2}\frac{\rho}{1-\rho}}}{\sqrt{2\pi}} \frac{c}{(\rho(1-\rho))^{3/2}},$$

using Theorem 13 and Stirling's approximation. Thus, for $\rho > \frac{1}{2}$,

$$\mathbb{P}(c_{\max}^{(k)}(T_n) \geq \rho n) \to \int_{\rho}^{1} \frac{e^{-\frac{c^2}{2}\frac{t}{1-t}}}{\sqrt{2\pi}} \frac{c}{(t(1-t))^{3/2}}\, dt.$$

For $\rho \leq \frac{1}{2}$, we can modify this argument with a generalized version of Theorem 13 for several clusters and the inclusion-exclusion principle to prove convergence of $c_{\max}^{(k)}(T_n)/n$ to a continuous random variable with support $[0, 1]$. Details are left to the full version.

Finally, in the supercritical case, we recall the corresponding case for the size of the root cluster from Theorem 9. Using Markov's inequality yields, for any $\epsilon > 0$,

$$\mathbb{P}(n - R_n^{(k)} \geq \epsilon n) \leq \frac{n - \mathbb{E}(R_n^{(k)})}{\epsilon n} \sim \frac{d\sqrt{n}}{\epsilon n} \xrightarrow{n\to\infty} 0.$$

Thus, the root cluster is the largest cluster of size $\sim n$ with high probability. Translating this from rooted to unrooted trees proves the theorem. ◀

References

1. David Aldous and Jim Pitman. The standard additive coalescent. *Ann. Probab.*, 26(4):1703–1726, 1998. doi:10.1214/aop/1022855879.
2. Jean Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. doi:10.1017/CBO9780511617768.
3. Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, Inc., New York, second edition, 1999. doi:10.1002/9780470316962.
4. P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms*, 21(1):76–119, 2002. doi:10.1002/rsa.10039.
5. Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009. doi:10.1017/CBO9780511801655.
6. Svante Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms*, 29(2):139–179, 2006. doi:10.1002/rsa.20086.

7 J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982. doi:10.1016/0304-4149(82)90011-4.

8 Achim Klenke. *Probability theory—a comprehensive course*. Universitext. Springer, Cham, 2020. Third edition. doi:10.1007/978-3-030-56402-5.

9 Jeremy L. Martin and Victor Reiner. Factorization of some weighted spanning tree enumerators. *J. Combin. Theory Ser. A*, 104(2):287–300, 2003. doi:10.1016/j.jcta.2003.08.003.

10 Jim Pitman. Coalescent random forests. *J. Combin. Theory Ser. A*, 85(2):165–193, 1999. doi:10.1006/jcta.1998.2919.

11 Jim Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999. doi:10.1214/aop/1022677552.

12 Jeffery B. Remmel and S. Gill Williamson. Spanning trees and function classes. *Electron. J. Combin.*, 9(1):Research Paper 34, 24, 2002. URL: http://www.combinatorics.org/Volume_9/Abstracts/v9i1r34.html.

13 Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition, 1999. doi:10.1007/978-3-662-06400-9.

14 John Riordan. *Combinatorial identities*. John Wiley & Sons, Inc., New York, 1968.

15 Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. doi:10.1017/CBO9780511609589.

A Additional details on the proof of Theorem 5

A.1 Computations related to the martingale

With the help of the recursive description in (10), we can show that $(Y_j^{(n)})_{1\le j\le n-1}$ is a martingale by computing

$$\begin{aligned}\mathbb{E}(Y_{j+1}^{(n)}|Y_j^{(n)}) &= \frac{\mathbb{E}(K_{j+1}^{(n)}|K_j^{(n)})}{n-j-1} - \frac{j(j+1)}{n(n-j-1)} = \frac{K_j^{(n)} + \frac{j+1}{n} + \frac{j-1-K_j^{(n)}}{n-j}}{n-j-1} - \frac{j(j+1)}{n(n-j-1)}\\ &= \frac{K_j^{(n)}}{n-j} - \frac{(j-1)j}{n(n-j)} = Y_j^{(n)}.\end{aligned}$$

We can also give an explicit expression for the variance of $Y_k^{(n)}$: recall that by (11), we have $\mathbb{V}Y_k^{(n)} = (n-k)^{-2}\mathbb{V}K_k^{(n)}$. Then, with (10) and the laws of total variance and total expectation we find the recurrence

$$\mathbb{V}K_{k+1}^{(n)} = \Big(1 - \frac{1}{n-k}\Big)^2 \mathbb{V}K_k^{(n)} + \frac{(n-k-1)(2n-k-1)k}{(n-k)n^2},$$

for $1 \le k < n-1$ and $\mathbb{V}K_1^{(n)} = 0$. This allows us to conclude that

$$\mathbb{V}K_k^{(n)} = \sum_{j=1}^{k-1} \Big(\frac{n-k}{n-j-1}\Big)^2 \frac{(n-j-1)(2n-j-1)j}{(n-j)n^2} = \frac{k(k-1)(n-k)}{n^2}, \tag{22}$$

where the sum can be evaluated with the help of partial fractions and telescoping.

A.2 Proofs of auxiliary results

Proof of Lemma 7. In order to obtain the tightness condition, we show first that it can be reduced to an inequality for the martingale from the previous section. To this end, let us write $tn = j + \eta$, with $j \in \mathbb{Z}$ and $\eta \in [0, 1)$. A simple calculation shows that

$$\begin{aligned} Z^{(n)}(t) &= \frac{\tilde{K}_t^{(n)} - t^2 n}{\sqrt{n}} \\ &= \frac{(1-\eta)K_j^{(n)} + \eta K_{j+1}^{(n)} - (j+\eta)^2/n}{\sqrt{n}} \\ &= \frac{(1-\eta)(K_j^{(n)} - j(j-1)/n) + \eta(K_{j+1}^{(n)} - j(j+1)/n) - (j+\eta^2)/n}{\sqrt{n}} \\ &= (1-\eta)\frac{K_j^{(n)} - j(j-1)/n}{\sqrt{n}} + \eta\frac{K_{j+1}^{(n)} - j(j+1)/n}{\sqrt{n}} - \frac{j+\eta^2}{n^{3/2}}. \end{aligned}$$

The final fraction is bounded by 1, since $j + \eta^2 \le j + \eta = tn \le n$. It follows that

$$\sup_{t\in[0,1]} |Z^{(n)}(t)| \le \sup_{0\le j\le n} \Big|\frac{K_j^{(n)} - j(j-1)/n}{\sqrt{n}}\Big| + 1,$$

so

$$\begin{aligned} \mathbb{P}\Big(\sup_{t\in[0,1]} |Z^{(n)}(t)| \ge C\Big) &\le \mathbb{P}\Big(\sup_{0\le j\le n} \Big|\frac{K_j^{(n)} - j(j-1)/n}{\sqrt{n}}\Big| \ge C-1\Big) \\ &= \mathbb{P}\Big(\sup_{1\le j\le n-1} \Big|\frac{Y_j^{(n)}(n-j)}{\sqrt{n}}\Big| \ge C-1\Big). \end{aligned} \tag{23}$$

Note here that we need not consider $j = 0$ and $j = n$ in the supremum, since $K_j^{(n)} - j(j-1)/n = 0$ in either case. Since $(Y_j^{(n)})_{1\le j\le n-1}$ is a martingale, we can use Doob's L^p-inequality [8, Theorem 11.2]. For any real $C > 0$ and any fixed integer k with $1 \le k \le n-1$, we have

$$\mathbb{P}\Big(\sup_{1\le j\le k} |Y_j^{(n)}| \ge C\Big) \le \frac{\mathbb{V}Y_k^{(n)}}{C^2} = \frac{k(k-1)}{C^2(n-k)n^2}.$$

With this, we have all required prerequisites to prove tightness of $Z^{(n)}(t)$. We partition the interval over which the supremum is taken in (23), apply the martingale inequality, and then obtain the desired result after summing over all these upper bounds. For every integer $i > 0$, let $I_i^{(n)} := [2^{-i}n, 2^{-i+1}n] \cap \mathbb{Z}$. We find

$$\begin{aligned} \mathbb{P}\Big(\sup_{n-j\in I_i^{(n)}} \Big|\frac{Y_j^{(n)}(n-j)}{\sqrt{n}}\Big| \ge C-1\Big) &\le \mathbb{P}\Big(\sup_{n-j\in I_i^{(n)}} |Y_j^{(n)} 2^{-i+1}\sqrt{n}| \ge C-1\Big) \\ &= \mathbb{P}\Big(\sup_{n-j\in I_i^{(n)}} |Y_j^{(n)}| \ge \frac{2^{i-1}(C-1)}{\sqrt{n}}\Big) \\ &\le \frac{n}{2^{2i-2}(C-1)^2}\mathbb{V}(Y_{n-\lceil 2^{-i}n\rceil}^{(n)}) \\ &\le \frac{n}{2^{2i-2}(C-1)^2}\cdot\frac{2^i}{n} = \frac{4}{2^i(C-1)^2}, \end{aligned}$$

where in the last inequality we bounded the variance as follows:

$$\mathbb{V}(Y^{(n)}_{n-\lceil 2^{-i}n\rceil}) = \frac{\mathbb{V}(K^{(n)}_{n-\lceil 2^{-i}n\rceil})}{\lceil 2^{-i}n\rceil^2} = \frac{(n-\lceil 2^{-i}n\rceil)(n-\lceil 2^{-i}n\rceil-1)\lceil 2^{-i}n\rceil}{n^2\lceil 2^{-i}n\rceil^2} \leq \frac{2^i}{n}.$$

Finally, the union bound together with the observation that $\sum_{i\geq 1}\frac{4}{2^i(C-1)^2} = 4(C-1)^{-2}$ yields the upper bound in (12) and therefore completes the proof. ◀

Proof of Lemma 8. As a consequence of Lemma 3 and Cayley's well-known enumeration formula for labeled trees of size n, we find that the probability generating function of the number of edge increments after $1 < j_1 < j_2 < \cdots < j_r < n$ steps, respectively, is given by

$$\begin{aligned} P_n(z_1, z_2, \ldots, z_r) &= \frac{E_n(z_1, z_2, \ldots, z_r)}{n^{n-2}} \\ &= \prod_{i=1}^{r}\Big(1 - \frac{j_r}{n} + \frac{j_i}{n}z_i + \sum_{h=i+1}^{r}\frac{j_h - j_{h-1}}{n}z_h\Big)^{j_i - j_{i-1}}, \end{aligned} \tag{24}$$

where $j_0 = 1$ for the sake of convenience. Now observe that $\Delta^{(n)}_{\mathbf{j}}$ can be seen as a marginal distribution of the sum of r independent, multinomially distributed random vectors: write $t_i = j_i/n$ and consider $M_j \sim \mathrm{Multi}(j_i - j_{i-1}, \mathbf{p}_i)$ where

$$\mathbf{p}_i = (p_{i,0}, p_{i,1}, \ldots, p_{i,r}) \in [0,1]^r \quad \text{such that} \quad p_{i,h} = \begin{cases} 1 - t_r & \text{if } h = 0, \\ 0 & \text{if } 0 < h < i, \\ t_i & \text{if } h = i, \\ t_h - t_{h-1} & \text{otherwise.} \end{cases} \tag{25}$$

By construction, the probability generating function of M_i is then given by

$$\Big((1-t_r)z_0 + t_i z_i + \sum_{h=i+1}^{r}(t_h - t_{h-1})\Big)^{j_i - j_{i-1}},$$

so that the probability generating function of the sum $M_1 + \cdots + M_r$ is a product that is very similar (and actually equal if we set $z_0 = 1$, which corresponds to marginalizing out the first component) to (24). In order to make the following arguments formally easier to read, and as the first component is not relevant for us at all, we slightly abuse notation and let M_i for $1 \leq i \leq r$ denote the corresponding marginalized multinomial distributions instead.

For the sake of convenience, we make a slight adjustment: instead of fixing the integer vector $\mathbf{j} = (j_1, \ldots, j_r)$, we fix $\mathbf{t} = (t_1, \ldots, t_r)$ with $0 < t_1 < \cdots < t_r < 1$ and define $\mathbf{j} = \lfloor \mathbf{t}n \rfloor$. Here, n is considered to be sufficiently large so that the conditions for the corresponding integer vector, $1 < \lfloor t_1 n\rfloor < \cdots < \lfloor t_r n\rfloor < n$, are still satisfied.

By the multivariate central limit theorem, it is well-known that a multinomially distributed random vector $M \sim \mathrm{Multi}(n, \mathbf{p})$ converges, for $n \to \infty$ and after appropriate scaling, in distribution to a multivariate normal distribution,

$$\frac{M - n\mathbf{p}}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathrm{diag}(\mathbf{p}) - \mathbf{p}^\top\mathbf{p}). \tag{26}$$

As a consequence, we find that

$$\begin{aligned}\frac{\Delta^{(n)}_{\lfloor n\mathbf{t}\rfloor}-\mathbb{E}\Delta^{(n)}_{\lfloor n\mathbf{t}\rfloor}}{\sqrt{n}} &= \frac{(M_1+\cdots+M_r)-\mathbb{E}(M_1+\cdots+M_r)}{\sqrt{n}}\\ &= (\sqrt{t_1}+O(n^{-1}))\frac{M_1-\mathbb{E}M_1}{\sqrt{\lfloor t_1 n\rfloor}}+\cdots\\ &\quad+(\sqrt{t_r-t_{r-1}}+O(n^{-1}))\frac{M_r-\mathbb{E}M_r}{\sqrt{\lfloor t_r n\rfloor-\lfloor t_{r-1}n\rfloor}}\\ &\xrightarrow[n\to\infty]{d}\sqrt{t_1}\mathcal{N}(\mathbf{0},\Sigma_1)+\cdots+\sqrt{t_r-t_{r-1}}\mathcal{N}(\mathbf{0},\Sigma_r)\\ &= \mathcal{N}(\mathbf{0},t_1\Sigma_1+\cdots+(t_r-t_{r-1})\Sigma_r),\end{aligned}$$

where the variance-covariance matrices are given by

$$\Sigma_j=\operatorname{diag}(\mathbf{p}_j)-\mathbf{p}_j^{\top}\mathbf{p}_j.$$

By a straightforward (linear) transformation consisting of taking partial sums, the random vector of increments $\Delta^{(n)}_{\lfloor \mathbf{t}n\rfloor}$ can be transformed into $\mathbf{K}^{(n)}_{\lfloor \mathbf{t}n\rfloor}$. This proves that $\mathbf{K}^{(n)}_{\lfloor \mathbf{t}n\rfloor}$ converges, after centering and rescaling, to a multivariate normal distribution.

The entries of the corresponding variance-covariance matrix can either be determined mechanically from the entries of $t_1\Sigma_1+\cdots+(t_r-t_{r-1})\Sigma_r$ by taking the partial summation into account, or alternatively, our observations concerning the martingale from Section A.1 can be used. In particular, using (11), we find, for fixed $s,t\in[0,1]$ with $s<t$, that

$$\begin{aligned}\operatorname{Cov}\Big(\frac{K^{(n)}_{\lfloor sn\rfloor}-\mathbb{E}K^{(n)}_{\lfloor sn\rfloor}}{\sqrt{n}},\frac{K^{(n)}_{\lfloor tn\rfloor}-\mathbb{E}K^{(n)}_{\lfloor tn\rfloor}}{\sqrt{n}}\Big) &= \frac{(n-\lfloor tn\rfloor)(n-\lfloor sn\rfloor)}{n}\mathbb{E}(Y^{(n)}_{\lfloor sn\rfloor}Y^{(n)}_{\lfloor tn\rfloor})\\ &= (n(1-t)(1-s)+O(1))\mathbb{E}({Y^{(n)}_{\lfloor sn\rfloor}}^2)\\ &= s^2(1-t)+O(n^{-1}),\end{aligned}$$

where we made use of the martingale property, and that the second moment of $Y^{(n)}_j$ is equal to the variance $n^{-2}j(j-1)/(n-j)$. Ultimately, this verifies (14) and thus completes the proof. ◀

Depth-First Search Performance in a Random Digraph with Geometric Degree Distribution

Philippe Jacquet
Inria Saclay Ile de France, France

Svante Janson
Department of Mathematics, Uppsala University, Uppsala, Sweden

Abstract

We present an analysis of the depth-first search algorithm in a random digraph model with geometric outdegree distribution. We give also some extensions to general outdegree distributions. This problem posed by Donald Knuth in his next to appear volume of *The Art of Computer Programming* gives interesting insight in one of the most elegant and efficient algorithm for graph analysis due to Tarjan.

2012 ACM Subject Classification Theory of computation → Graph algorithms analysis

Keywords and phrases Combinatorics, Depth-First Search, Random Digraphs

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.11

Related Version *Full Version*: `https://hal.archives-ouvertes.fr/hal-03537124` [4]

Funding *Svante Janson*: supported by the Knut and Alice Wallenberg Foundation

Acknowledgements We thank Donald Knuth for posing us questions and conjectures that led to the present paper.

1 Introduction

The motivation of this paper is a new section in Donald Knuth's *The Art of Computer Programming* [5], which is dedicated to Depth-First Search (DFS) in a digraph. Briefly, the DFS starts with an arbitrary vertex, and explores the arcs from that vertex one by one. When an arc is found leading to a vertex that has not been seen before, the DFS explores the arcs from it in the same way, in a recursive fashion, before returning to the next arc from its parent. This eventually yields a tree containing all descendants of the the first vertex (which is the root of the tree). If there still are some unseen vertices, the DFS starts again with one of them and finds a new tree, and so on until all vertices are found. We refer to [5] for details as well as for historical notes. (See also S1–S2 in Section 4.) Note that the digraphs in [5] and here are multi-digraphs, where loops and multiple arcs are allowed. (Although in our random model they are few and usually not important.) The DFS algorithm generates a spanning forest (the *depth-first forest*) in the digraph, with all arcs in the forest directed away from the roots. Our main purpose is to study the distribution of the depth of vertices in the depth-first forest, starting with a random digraph G.

Furthermore, the DFS algorithm in [5] classifies the arcs in the digraph into the following five types, see Figure 1 for examples:

- *loops*;
- *tree arcs*, the arcs in the resulting depth-first forest;
- *back arcs*, the arcs which point to an ancestor of the current vertex in the current tree;
- *forward arcs*, the arcs which point to an already discovered descendant of the current vertex in the current tree;
- *cross arcs*, all other arcs (these point to an already discovered vertex which is neither a descendant nor an ancestor of the current vertex, and might be in another tree).

We will study the numbers of arcs of different types. (See further the exercises in [5].)

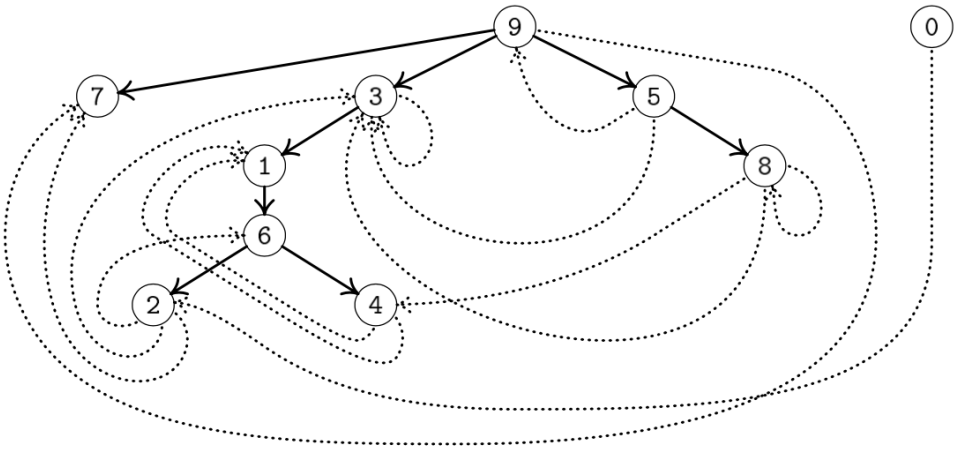


Figure 1 Example of a depth-first forest (jungle) from [5], by courtesy of Donald Knuth. Tree arcs are solid (e.g. ⑨→ ③). For example, ③⇢ ③ is a loop, ②⇢ ③ is a back arc, ⑨⇢ ⑦ is a forward arc, ⑧⇢ ④ and ⓪⇢ ② are cross arcs.

The random digraph model that we consider has n vertices and a given outdegree distribution $\mathbf{P}$. The outdegrees (number of outgoing arcs) of the n vertices are independent random numbers with this distribution. The endpoint of each arc is uniformly selected at random among the n vertices, independently of all other arcs. (Therefore, an arc can loop back to the starting vertex, and multiple arcs can occur.) We consider asymptotics as $n \to \infty$ for a fixed outdegree distribution.

We will focus on the case of a geometric outdegree distribution; the lack-of-memory property in this case leads to interesting features and a simpler analysis. We describe in Section 4 how the results can be extended to general outdegree distributions using a variation of the method. The paper will study the following outdegree distributions in the following order:

- a geometric distribution;
- a shifted geometric distribution (starting from integer 1 instead of zero);
- a general distribution.

▶ Remark 1. Related results are proved by [3] for DFS in an undirected Erdős–Rényi graph $G(n, \lambda/n)$; see also [2]. The main result of [3] shows convergence of the depth profile in the depth-first forest to a certain deterministic limit. Although this is for a different random graph model, DFS on $G(n, \lambda/n)$ is the same as DFS on the Erdős–Rényi digraph $D(n, \lambda/n)$, which is essentially the same as the digraph studied in the present paper with outdegree distribution $\mathrm{Po}(\lambda)$. Hence the result in [3] is essentially the special case $\mathbf{P} = \mathrm{Po}(\lambda)$ of our result for the depths. The proofs are quite different. DFS in the random digraph $D(n, p)$ has also been considered previously, for example in the proof of [6, Theorem 3].

This is an extended abstract of [4], which contains further results and complete proofs not included here; in particular we treat there general outdegree distributions in detail.

1.1 Some notation

We denote the given outdegree distribution by $\mathbf{P}$. We let η, often with subscripts, denote random variables with this distribution. In particular, we denote the outdegree of vertex v by $\eta(v)$. Recall that our standing assumption is that these outdegrees are i.i.d. (independent and identically distributed) with $\eta_v \sim \mathbf{P}$. We let v_t denote the t-th vertex found by the DFS,

and simplify notation by letting $\eta_t := \eta(v_t)$ be its outdegree. It follows from the construction of the DFS that also the random variables η_t, $t = 1, \dots, n$ are i.i.d. with distribution $\mathbf{P}$; this fundamental property will be used repeatedly without further mention.

The mean outdegree, *i.e.*, the expectation $\mathbb{E}\,\eta$ of $\mathbf{P}$, is denoted by λ. We assume throughout, for technical reasons, that the second moment $\mathbb{E}\,\eta^2 < \infty$.

As usual, w.h.p. means *with high probability*, *i.e.*, with probability $1 - o(1)$. We use $\xrightarrow{\mathrm{p}}$ for convergence in probability, and $\xrightarrow{\mathrm{d}}$ for convergence in distribution of random variables.

Moreover, let (a_n) be a sequence of positive numbers, and X_n a sequence of random variables. We write $X_n = o_{\mathrm{p}}(a_n)$ if, as $n \to \infty$, $X_n/a_n \xrightarrow{\mathrm{p}} 0$, *i.e.*, if for every $\varepsilon > 0$, we have $\mathbb{P}(|X_n| > \varepsilon a_n) \to 0$. Furthermore, $X_n = o_{L^2}(a_n)$ means $\mathbb{E}\bigl[|X_n/a_n|^2\bigr] \to 0$, and $X_n = O_{L^2}(a_n)$ means $\mathbb{E}\bigl[|X_n/a_n|^2\bigr] = O(1)$. Note that $X_n = o_{L^2}(a_n)$ implies $X_n = o_{\mathrm{p}}(a_n)$, and that $X_n = O_{L^2}(a_n)$ implies $X_n = o_{L^2}(\omega_n a_n)$ and thus $X_n = o_{\mathrm{p}}(\omega_n a_n)$, for any sequence $\omega(n) \to \infty$. Note also that $X_n = o_{L^2}(a_n)$ implies $\mathbb{E}\,X_n = o(a_n)$, and similarly for $O_{L^2}(a_n)$; thus error terms of this type implies immediately estimates for expectations and second moments. In particular, for the most common case below, $X_n = O_{L^2}(n^{1/2})$ is equivalent to $\mathbb{E}\,X_n = O(a_n)$ and $\operatorname{Var} X_n = O(n)$.

We define $\rho_0(x)$, for $x \geqslant 0$, as the largest solution in $[0, 1)$ to

$$1 - \rho_0 = e^{-x\rho_0}. \tag{1}$$

As is well known, $\rho_0(x)$ is the survival probability of a Galton–Watson process with $\mathrm{Po}(x)$ offspring distribution. We have $\rho_0(x) = 0$ for $x \leqslant 1$ and $0 < \rho_0(x) < 1$ for $x > 1$.

All logarithms are natural.

2 Depth analysis with geometric outdegree distribution

In this section we assume that the outdegree distribution is geometric $\mathrm{Ge}(1-p)$ for some fixed $0 < p < 1$, and thus has mean

$$\lambda := \mathbb{E}\,\eta = \frac{p}{1-p}. \tag{2}$$

When doing the DFS on a random digraph of the type studied in this paper, we generally reveal the outdegree of a vertex as soon as we find it. (See S1–S2 in Section 4.) However, for a geometric outdegree distribution, because of its lack-of-memory property, we do not have to immediately reveal the outdegree when we find a new vertex v. Instead, we only check whether there is at least one outgoing arc (probability p), and if so, we find its endpoint and explore this endpoint if it has not already been visited; eventually, we return to v, and then we check whether there is another outgoing arc (again probability p, by the lack-of-memory property), and so on. This will yield the important Markov property in the construction in the next subsection.

In the following, by a *future arc* from some vertex, we mean an arc that at the current time has not yet been seen by the DFS.

2.1 Depth Markov chain

Our aim is to track the evolution of the search depth as a function of the number t of discovered vertices. Let v_t be the t-th vertex discovered by the DFS ($t = 1, \dots, n$), and let $d(t)$ be the depth of v_t in the resulting depth-first forest, *i.e.*, the number of tree edges that connect the root of the current tree to v_t. The first found vertex v_1 is a root, and thus $d(1) = 0$.

The quantity $d(t)$ follows a Markov chain with transitions ($1 \leqslant t < n$):

(i) $d(t+1) = d(t) + 1$.
This happens if, for some $k \geqslant 1$, v_t has at least k outgoing arcs, the first $k-1$ arcs lead to vertices already visited, and the kth arc leads to a new vertex (which then becomes v_{t+1}). The probability of this is

$$\sum_{k=1}^{\infty} p^k \Big(\frac{t}{n}\Big)^{k-1} \Big(1 - \frac{t}{n}\Big) = \frac{(1-t/n)p}{1-pt/n}. \tag{3}$$

(ii) $d(t+1) = d(t)$, assuming $d(t) > 0$.
This holds if all arcs from v_t lead to already visited vertices, *i.e.*, (i) does not happen, and furthermore, the parent of v_t has at least one future arc leading to an unvisited vertex. These two events are independent. Moreover, by the lack-of-memory property, the number of future arcs from the parent of v_t has the same distribution $\mathrm{Ge}(1-p)$. Hence, the probability that one of these future arcs leads to an unvisited vertex equals the probability in (3). The probability of (ii) is thus

$$\Big(1 - \frac{(1-t/n)p}{1-pt/n}\Big)\frac{(1-t/n)p}{1-pt/n}. \tag{4}$$

(iii) $d(t+1) = d(t) - \ell$, assuming $d(t) > \ell \geqslant 1$.
This happens if all arcs from v_t lead to already visited vertices, and so do all future arcs from the ℓ nearest ancestors of v_t, while the $(\ell+1)$th ancestor has at least one future arc leading to an unvisited vertex. The argument in (ii) generalizes and shows that this has probability

$$\Big(1 - \frac{(1-t/n)p}{1-pt/n}\Big)^{\ell+1}\frac{(1-t/n)p}{1-pt/n}. \tag{5}$$

(iv) $d(t+1) = d(t) - \ell$, assuming $d(t) = \ell \geqslant 0$.
By the same argument as in (ii) and (iii), except that the $(\ell+1)$th ancestor does not exist and we ignore it, we obtain the probability

$$\Big(1 - \frac{(1-t/n)p}{1-pt/n}\Big)^{\ell+1}. \tag{6}$$

Note that (iv) is the case when $d(t+1) = 0$ and thus v_{t+1} is the root of a new tree in the depth-first forest.

We can summarize (i)–(iv) in the formula

$$d(t+1) = \big(d(t) + 1 - \xi_t\big)^+, \tag{7}$$

where $x^+ := \max\{x, 0\}$, and ξ_t is a random variable, independent of the history, with the distribution

$$\mathbb{P}(\xi_t = k) = (1-\pi_t)^k\pi_t, \quad k \geqslant 0, \qquad \text{with} \qquad \pi_t := \frac{(1-t/n)p}{1-pt/n} = 1 - \frac{1-p}{1-pt/n}. \tag{8}$$

In other words, ξ_t has the geometric distribution $\mathrm{Ge}(\pi_t)$. Define

$$\widetilde{d}(t) := \sum_{i=1}^{t-1}(1-\xi_i), \tag{9}$$

and note that (9) is a sum of independent random variables. Then (7) and induction yield

$$d(t) = \widetilde{d}(t) - \min_{1\leqslant j\leqslant t} \widetilde{d}(j), \qquad 1 \leqslant t \leqslant n. \tag{10}$$

▶ Remark 2. Similar formulas have been used for other, related, problems with random graphs and trees, where trees have been coded as walks, see for example [1, Section 1.3]. Note that in our case, unlike e.g. [1], $\widetilde{d}(t)$ may have negative jumps of arbitrary size.

▶ Remark 3. We can also express these relations using generating functions. Let $p(t,z)$ be the probability generating function $\mathbb{E}\, z^{\xi_t}$ of ξ_t, *i.e.*,

$$p(t,z) := \frac{(1-t/n)p}{1-pt/n}\sum_{\ell\geqslant 0}\left(1-\frac{(1-t/n)p}{1-pt/n}\right)^{\ell} z^{\ell} = \frac{(1-t/n)p}{1-pt/n-(1-p)z}, \tag{11}$$

and let $f(t,z) := E[z^{d(t)}]$. We then have the identity, equivalent to (7),

$$f(t+1,z) = \mathcal{N}\big[R(t,z)f(t,z)\big] \tag{12}$$

where $R(t,z) := p(t,1/z)z$ and $\mathcal{N}$ is the operator on power series in $z^{\pm 1}$:

$$\mathcal{N}g(z) = \Pi^{+}g(z) + (\Pi^{-}g)(1) \tag{13}$$

where Π^{+} is the operator which removes the strictly negative powers of z and Π^{-} is the operator which removes the nonnegative powers of z. Thus we have, since $f(1,z)=1$,

$$f(t+1,z) = \mathcal{N}R(t,z)\mathcal{N}R(t-1,z)\mathcal{N}\cdots\mathcal{N}R(1,z). \tag{14}$$

At least some of the results below can be derived using this formalism, but we will not employ it in the present paper.

2.2 Main result for depth analysis

Note first that (9) implies that the expectation of $\widetilde{d}(t)$ is

$$\mathbb{E}\big[\widetilde{d}(t)\big] = \sum_{i=1}^{t-1}(1-\mathbb{E}\,\xi_i) = \sum_{i=1}^{t-1}\Big(1-\frac{1-\pi_i}{\pi_i}\Big) = \sum_{i=1}^{t-1}\Big(1-\frac{1-p}{p(1-i/n)}\Big). \tag{15}$$

Let $\theta := t/n$. We fix a $\theta^* < 1$ and obtain that, uniformly for $\theta \leqslant \theta^*$, recalling (2),

$$\mathbb{E}\big[\widetilde{d}(t)\big] = \int_0^t \Big(1-\frac{1}{\lambda(1-x/n)}\Big)\,\mathrm{d}x + O(1) = n\widetilde{\ell}(\theta) + O(1), \tag{16}$$

where

$$\widetilde{\ell}(\theta) := \int_0^{\theta}\Big(1-\frac{1}{\lambda(1-x)}\Big)\,\mathrm{d}x = \theta + \frac{1}{\lambda}\log(1-\theta). \tag{17}$$

Note that the derivative $\widetilde{\ell}'(\theta) = 1-\lambda^{-1}/(1-\theta)$ is (strictly) decreasing on $(0,1)$, *i.e.*, $\widetilde{\ell}$ is concave. Moreover, if $\lambda > 1$ (*i.e.*, $p > \frac{1}{2}$) which we call the *supercritical* case, then $\widetilde{\ell}'(0) > 0$, and (17) shows that $\widetilde{\ell}(\theta)$ is positive and increasing for $\theta < \theta_0 := 1-\lambda^{-1} = (2p-1)/p$. After the maximum at θ_0, $\widetilde{\ell}(\theta)$ decreases and tends to $-\infty$ as $\theta \nearrow 1$. Hence, there exists a $0 < \theta_1 < 1$ such that $\widetilde{\ell}(\theta_1) = 0$; we then have $\widetilde{\ell}(\theta) > 0$ for $0 < \theta < \theta_1$ and $\widetilde{\ell}(\theta) < 0$ for $\theta > \theta_1$. We will see that in this case the depth-first forest w.h.p. contains a giant tree, of order and height both linear in n, while all other trees are small.

On the other hand, if $\lambda \leqslant 1$ (*i.e.*, $p \leqslant \frac{1}{2}$) (the *subcritical* and *critical* cases), then $\widetilde{\ell}'(0) \leqslant 0$ and $\widetilde{\ell}(\theta)$ is negative and decreasing for all $\theta \in (0,1)$. In this case, we define $\theta_0 := \theta_1 := 0$ and note that the properties just stated for $\widetilde{\ell}$ still hold (rather trivially). We will see that in this case w.h.p. all trees in the depth-first forest are small.

Note that in all cases, θ_1 is the largest solution in $[0,1)$ to

$$\log(1-\theta_1) = -\lambda\theta_1. \tag{18}$$

▶ Remark 4. The equation (18) may also be written $1-\theta_1 = \exp(-\lambda\theta_1)$, which shows that $\theta_1 = \rho_0(\lambda)$, the survival probability of a Galton–Watson process with $\mathrm{Po}(\lambda)$ offspring distribution defined in (1).

We define $\widetilde{\ell}^+(\theta) := [\widetilde{\ell}(\theta)]^+$. Thus, by (17) and the comments above,

$$\widetilde{\ell}^+(\theta) = \begin{cases} \theta + \lambda^{-1}\log(1-\theta), & 0 \leqslant \theta \leqslant \theta_1, \\ 0, & \theta_1 \leqslant \theta \leqslant 1. \end{cases} \tag{19}$$

We can now state one of our main results. Proofs are given in the next subsection.

▶ **Theorem 5.** *We have*

$$\max_{1\leqslant t\leqslant n} \bigl|d(t) - n\widetilde{\ell}^+(t/n)\bigr| = O_{L^2}(n^{1/2}). \tag{20}$$

▶ **Corollary 6.** *The height Υ of the depth-first forest is*

$$\Upsilon := \max_{1\leqslant t\leqslant n} d(t) = \upsilon n + O_{L^2}(n^{1/2}), \tag{21}$$

where

$$\upsilon = \upsilon(p) := \widetilde{\ell}^+(\theta_0) = \begin{cases} 0, & 0 < \lambda \leqslant 1, \\ 1 - \lambda^{-1} - \lambda^{-1}\log\lambda, & \lambda > 1 \end{cases} \tag{22}$$

Moreover, we can show that the height Υ is asymptotically normally distributed. Details are given in the full paper [4].

▶ **Corollary 7.** *The average depth $\overline{d}$ in the depth-first forest is*

$$\overline{d} := \frac{1}{n}\sum_{t=1}^{n} d(t) = \alpha n + O_{L^2}(n^{1/2}), \tag{23}$$

where

$$\alpha = \alpha(p) := \frac{1}{2}\theta_1^2 - \frac{1}{\lambda}\Bigl((1-\theta_1)\log(1-\theta_1) + \theta_1\Bigr) = \frac{\lambda-1}{\lambda}\theta_1 - \frac{1}{2}\theta_1^2. \tag{24}$$

We have $\alpha = 0$ if and only if $\lambda \geqslant 1$, i.e.*, $p \leqslant 1/2$.*

▶ Remark 8. When $p > \frac{1}{2}$, the height is thus linear in n, unlike many other types of random trees. This might imply a rather slow performance of algorithms that operate on the depth-first forest.

2.3 Proofs

Proof of Theorem 5. Since (9) is a sum of independent random variables, $\widetilde{d}(t) - \mathbb{E}\,\widetilde{d}(t)$ $(t = 1, \dots, n)$ is a martingale, and Doob's inequality yields, for all $T \leqslant n$,

$$\mathbb{E}\bigl[\max_{t\leqslant T} |\widetilde{d}(t) - \mathbb{E}\,\widetilde{d}(t)|^2\bigr] \leqslant 4\,\mathbb{E}\bigl[|\widetilde{d}(T) - \mathbb{E}\,\widetilde{d}(T)|^2\bigr] = 4\sum_{i=1}^{T-1}\mathrm{Var}(\xi_i). \tag{25}$$

Fix $\theta^* < 1$, and assume, as we may, that $\theta^* > \theta_1$. Let $T^* := \lfloor n\theta^* \rfloor$, and consider first $t \leqslant T^*$. For $i < T^*$, we have $\mathrm{Var}\,\xi_i = O(1)$, and thus, for $T = T^*$, the sum in (25) is $O(T^*) = O(n)$. Consequently, (25) yields

$$\max_{t\leqslant T^*} \bigl|\widetilde{d}(t) - \mathbb{E}\,\widetilde{d}(t)\bigr| = O_{L^2}(n^{1/2}). \tag{26}$$

Hence, by (16),

$$M^* := \max_{t\leqslant T^*}\bigl|\widetilde{d}(t) - n\widetilde{\ell}(t/n)\bigr| = O_{L^2}(n^{1/2}). \tag{27}$$

For $t \leqslant T^*$, the definition of M^* in (27) implies

$$\Bigl|\min_{1\leqslant j\leqslant t}\widetilde{d}(j) - n\min_{1\leqslant j\leqslant t}\widetilde{\ell}(j/n)\Bigr| \leqslant M^*. \tag{28}$$

Moreover, for $t/n \leqslant \theta_1$, we have $\min_{1\leqslant j\leqslant t}\widetilde{\ell}(j/n) = O(1/n)$, while for $t/n \geqslant \theta_1$, we have $\min_{1\leqslant j\leqslant t}\widetilde{\ell}(j/n) = \widetilde{\ell}(t/n)$. Hence, for all $t \leqslant T^*$,

$$\min_{1\leqslant j\leqslant t}\widetilde{\ell}(j/n) = \widetilde{\ell}(t/n) - \widetilde{\ell}^+(t/n) + O(1/n), \tag{29}$$

and thus, by (28),

$$\Bigl|\min_{1\leqslant j\leqslant t}\widetilde{d}(j) - n\widetilde{\ell}(t/n) + n\widetilde{\ell}^+(t/n)\Bigr| \leqslant M^* + O(1). \tag{30}$$

Finally, by (10), (27) and (30),

$$\bigl|d(t) - n\widetilde{\ell}^+(t/n)\bigr| \leqslant 2M^* + O(1). \tag{31}$$

This holds uniformly for $t \leqslant T^*$, and thus, by (27),

$$\max_{1\leqslant t\leqslant T^*}\bigl|d(t) - n\widetilde{\ell}^+(t/n)\bigr| = O_{L^2}(n^{1/2}). \tag{32}$$

It remains to consider $T^* < t \leqslant n$. Then the argument above does not quite work, because $\pi_t \searrow 0$ and thus $\operatorname{Var}\xi_t \nearrow \infty$ as $t \nearrow n$. We therefore modify ξ_t. We define $\widehat{\pi}_t := \max\{\pi_t, \pi_{T^*}\}$; thus $\widehat{\pi}_t = \pi_t$ for $t \leqslant T^*$ and $\widehat{\pi}_t > \pi_t$ for $t > T^*$. We may thus define independent random variables $\widehat{\xi}_t$ such that $\widehat{\xi}_t \sim \mathrm{Ge}(\widehat{\pi}_t)$ and $\widehat{\xi}_t \leqslant \xi_t$ for all $t < n$. (Thus, $\widehat{\xi}_t = \xi_t$ for $t \leqslant T^*$.) The argument above works for the modified variables for all $t \leqslant n$. Since the modification can only increase $d(t)$, it follows that

$$\max_{T^*<t\leqslant n} d(t) = O_{L^2}(n^{1/2}), \tag{33}$$

which completes the proof since $\widetilde{\ell}(t/n) = 0$ for $t > T^*$. We omit the details. ◀

Proof of Corollary 6. Immediate from Theorem 5 and (17), since we have $\max_t \widetilde{\ell}^+(t/n) = \max_\theta \widetilde{\ell}^+(\theta) + O(1/n)$ and $\max_\theta \widetilde{\ell}^+(\theta) = \widetilde{\ell}^+(\theta_0) = \widetilde{\ell}(\theta_0)$. ◀

Proof of Corollary 7. By Theorem 5,

$$\frac{1}{n}\sum_{t=1}^{n} d(t) = \sum_{t=1}^{n}\widetilde{\ell}^+(t/n) + O_{L^2}\bigl(n^{1/2}\bigr) = n\alpha + O_{L^2}\bigl(n^{1/2}\bigr), \tag{34}$$

where

$$\begin{aligned}\alpha &:= \int_0^1 \widetilde{\ell}^+(x)\,\mathrm{d}x = \int_0^{\theta_1}\widetilde{\ell}(x)\,\mathrm{d}x = \int_0^{\theta_1}\Bigl(x + \lambda^{-1}\log(1-x)\Bigr)\,\mathrm{d}x \\ &= \frac{1}{2}\theta_1^2 - \lambda^{-1}\Bigl((1-\theta_1)\log(1-\theta_1) + \theta_1\Bigr),\end{aligned} \tag{35}$$

which yields (24), using (18). ◀

2.4 The trees in the forest

▶ **Theorem 9.** *Let N be the number of trees in the depth-first forest. Then*

$$N = \psi n + O_{L^2}(n^{1/2}), \tag{36}$$

where

$$\psi = \psi(p) := 1 - \theta_1 - \frac{\lambda}{2}(1-\theta_1)^2. \tag{37}$$

Proof. Let $J_t := \mathbf{1}\{d(t) = 0\}$, the indicator that vertex t is a root and thus starts a new tree. Thus $N = \sum_1^n J_t$.

If $\theta_1 > 0$ (*i.e.*, $\lambda > 1$), then Theorem 5 shows that w.h.p. $d(t) > 0$ in the interval $(1, n\theta_1)$, except possibly close to the endpoints. Thus the DFS will find one giant tree of order $\approx \theta_1 n$, possibly preceded by a few small trees, and, as we will see later, followed by many small trees. To obtain a precise estimate, we note that there exists a constant $c > 0$ such that $\widetilde{\ell}(\theta) \geqslant \min\{c\theta, c(\theta_1 - \theta)\}$ for $\theta \in [0, \theta_1]$. Hence, if $t \leqslant n\theta_1$ and $d(t) = 0$, then $\widetilde{d}(t) \leqslant d(t) = 0$ by (10) and, recalling (27),

$$M^* \geqslant n\widetilde{\ell}(t/n) \geqslant c\min\{t, n\theta_1 - t\}. \tag{38}$$

Consequently, $d(t) = 0$ with $t \leqslant n\theta_1$ implies $t \in [1, c^{-1}M^*] \cup [n\theta_1 - c^{-1}M^*, n\theta_1]$. The number of such t is thus $O(M^* + 1) = O_{L^2}(n^{1/2})$, using (27).

Let $T_1 := \lceil n\theta_1 \rceil$. We have just shown that (the case $\theta_1 = 0$ is trivial)

$$\sum_{t=1}^{T_1 - 1} J_t = O_{L^2}(n^{1/2}). \tag{39}$$

It remains to consider $t \geqslant T_1$. Let

$$\mu_t := \mathbb{E}\,\xi_t = \frac{1-\pi_t}{\pi_t} = \frac{1-p}{p(1-t/n)} = \frac{1}{\lambda(1-t/n)}. \tag{40}$$

For any integer $k \geqslant 0$, the conditional distribution of $\xi_t - k$ given $\xi_t \geqslant k$ equals the distribution of ξ_t. Hence,

$$\mathbb{E}\big[(\xi_t - k)^+\big] = \mathbb{E}\big[\xi_t - k \mid \xi_t \geqslant k\big]\,\mathbb{P}(\xi_t \geqslant k) = \mu_t\,\mathbb{P}(\xi_t - k \geqslant 0). \tag{41}$$

We use again the stochastic recursion (7). Let $\mathcal{F}_t$ be the σ-field generated by $\xi_1, \dots, \xi_{t-1}$. Then $d(t)$ is $\mathcal{F}_t$-measurable, while ξ_t is independent of $\mathcal{F}_t$. Hence, (7) and (41) yield

$$\begin{aligned}
\mathbb{E}\big[d(t+1) \mid \mathcal{F}_t\big] &= \mathbb{E}\big[d(t) + 1 - \xi_t \mid \mathcal{F}_t\big] + \mathbb{E}\big[(\xi_t - 1 - d(t))^+ \mid \mathcal{F}_t\big] \\
&= d(t) + 1 - \mu_t + \mu_t\,\mathbb{P}\big[\xi_t - 1 - d(t) \geqslant 0 \mid \mathcal{F}_t\big] \\
&= d(t) + 1 - \mu_t + \mu_t\,\mathbb{P}\big[d(t+1) = 0 \mid \mathcal{F}_t\big] \\
&= d(t) + 1 - \mu_t + \mu_t\,\mathbb{E}\big[J_{t+1} \mid \mathcal{F}_t\big].
\end{aligned} \tag{42}$$

We write $\Delta d(t) := d(t+1) - d(t)$ and $\overline{J}_t := 1 - J_t$. Then (42) yields

$$\mathbb{E}\big[\Delta d(t) - 1 + \mu_t \overline{J}_{t+1} \mid \mathcal{F}_t\big] = 0. \tag{43}$$

Define

$$\mathcal{M}_t := \sum_{i=1}^{t-1} \mu_i^{-1}\big(\Delta d(i) - 1 + \mu_i \overline{J}_{i+1}\big) = \sum_{i=1}^{t-1}\Big(\mu_i^{-1}\Delta d(i) - \mu_i^{-1} + \overline{J}_{i+1}\Big). \tag{44}$$

Then $\mathcal{M}_t$ is $\mathcal{F}_t$-measurable, and (43) shows that $\mathcal{M}_t$ is a martingale. We have, with $\Delta\mathcal{M}_t := \mathcal{M}_{t+1} - \mathcal{M}_t$, using (7),

$$|\Delta\mathcal{M}_t| \leqslant \mu_t^{-1}\bigl|d(t+1) - d(t) - 1\bigr| + \overline{J}_{t+1} \leqslant \mu_t^{-1}\xi_t + 1, \tag{45}$$

and thus, since $\pi_t \leqslant p < 1$ for all t by (8),

$$\mathbb{E}|\Delta\mathcal{M}_t|^2 \leqslant 2\mu_t^{-2}\,\mathbb{E}\,\xi_t^2 + 2 = 2\Bigl(\frac{\pi_t}{1-\pi_t}\Bigr)^2\frac{1-\pi_t+(1-\pi_t)^2}{\pi_t^2} + 2 = O(1). \tag{46}$$

Hence, uniformly for all $T \leqslant n$,

$$\mathbb{E}\,\mathcal{M}_T^2 = \sum_{t=1}^{T-1}\mathbb{E}\,|\Delta\mathcal{M}_t|^2 = O(T) = O(n). \tag{47}$$

The definition (44) yields

$$\mathcal{M}_n - \mathcal{M}_{T_1} = \sum_{t=T_1}^{n-1}\mu_t^{-1}\Delta d(t) - \sum_{t=T_1}^{n-1}\mu_t^{-1} + \sum_{t=T_1}^{n-1}\overline{J}_{t+1}. \tag{48}$$

By a summation by parts, and interpreting $\mu_n^{-1} := 0$,

$$\sum_{t=T_1}^{n-1}\mu_t^{-1}\Delta d(t) = \sum_{t=T_1+1}^{n}\bigl(\mu_{t-1}^{-1} - \mu_t^{-1}\bigr)d(t) - \mu_{T_1}^{-1}d(T_1). \tag{49}$$

As t increases, μ_t increases by (40), and thus $\mu_{t-1}^{-1} - \mu_t^{-1} > 0$. Hence, (49) implies

$$\begin{aligned}\Bigl|\sum_{t=T_1}^{n-1}\mu_t^{-1}\Delta d(t)\Bigr| &\leqslant \sum_{t=T_1+1}^{n}\bigl(\mu_{t-1}^{-1} - \mu_t^{-1}\bigr)\sup_{i>T_1}|d(t)| + \mu_{T_1}^{-1}|d(T_1)| \leqslant 2\mu_{T_1}^{-1}\sup_{i\geqslant T_1}|d(t)| \\ &= O_{L^2}(n^{1/2})\end{aligned} \tag{50}$$

by (20), since $\widetilde{\ell}^{+}(t/n) = 0$ for $t \geqslant T_1 \geqslant n\theta_1$. Furthermore, (47) shows that $\mathcal{M}_n, \mathcal{M}_{T_1} = O_{L^2}(n^{1/2})$. Hence, (48) yields

$$\sum_{t=T_1+1}^{n} J_t = n - T_1 - \sum_{t=T_1+1}^{n}\overline{J}_t = n - T_1 - \sum_{t=T_1}^{n-1}\mu_t^{-1} + O_{L^2}(n^{1/2}) = n\psi + O_{L^2}(n^{1/2}), \tag{51}$$

where

$$\psi := 1 - \theta_1 - \int_{\theta_1}^{1}\lambda(1-x)\,\mathrm{d}x = 1 - \theta_1 - \frac{\lambda}{2}(1-\theta_1)^2. \tag{52}$$

The result follows by (51) and (39). ◀

The argument in the proof of Theorem 9 shows also the following; we omit the details.

▶ **Theorem 10.** *If $\lambda > 1$ ($p > \frac{1}{2}$), then the largest tree $\mathbf{T}_1$ in the depth-first forest has order $|\mathbf{T}_1| = \theta_1 n + O_{L^2}(n^{1/2})$.*

By different methods (for $\lambda = 1$ adapted from [1]), we can prove the following complements, which show the same behaviour as the largest component in a random Erdős–Rényi graph.

▶ **Theorem 11.**

(i) *If $\lambda = 1$ ($p = \frac{1}{2}$), then $|\mathbf{T}_1|/n^{2/3}$ converges in distribution to a positive random variable.*

(ii) *If $\lambda < 1$ ($p < \frac{1}{2}$), then, for some constants $c, C > 0$, w.h.p. $c\log n \leqslant |\mathbf{T}_1| \leqslant C\log n$.*

The limit distribution in (i) is of the type found in [1] for components in random graphs.

2.5 Types of arcs

Recall from the introduction the classification of the arcs in the digraph G. Since we assume that the outdegrees are $\mathrm{Ge}(1-p)$ and independent, the total number of arcs, M say, has a negative binomial distribution with mean λn, and, by a weak version of the law of large numbers,

$$M = \lambda n + O_{L^2}(n^{1/2}). \tag{53}$$

In the following theorem, we give the asymptotics of the number of arcs of each type.

▶ **Theorem 12.** *Let L, T, B, F and C be the numbers of loops, tree arcs, back arcs, forward arcs, and cross arcs in the random digraph. Then*

$$L = O_{L^2}(1), \tag{54}$$
$$T = \tau n + O_{L^2}(n^{1/2}), \tag{55}$$
$$B = \beta n + O_{L^2}(n^{1/2}), \tag{56}$$
$$F = \varphi n + O_{L^2}(n^{1/2}), \tag{57}$$
$$C = \chi n + O_{L^2}(n^{1/2}), \tag{58}$$

where

$$\tau := \chi := 1 - \psi = \theta_1 + \frac{\lambda}{2}(1-\theta_1)^2, \tag{59}$$
$$\beta := \varphi := \lambda\alpha = (\lambda-1)\theta_1 - \frac{\lambda}{2}\theta_1^2. \tag{60}$$

The equalitites $\tau = \chi$ and $\beta = \varphi$ mean asymptotic equality of the corresponding expectations of numbers of arcs. In fact, there are exact equalities.

▶ **Theorem 13.** *For any n, $\mathbb{E}\,T = \mathbb{E}\,C$ and*

$$\mathbb{E}\,B = \mathbb{E}\,F = \lambda\,\mathbb{E}\,\overline{d} = \beta n + O(n^{1/2}). \tag{61}$$

▶ Remark 14. Knuth [5] conjectures, based on exact formulas for small n, that, much more strongly, B and F have the same distribution for every n. (Note that T and C do not have the same distribution; we have $T \leqslant n-1$, while C may take arbitrarily large values.)

Partial proof of Theorems 12 and 13.
L: A simple argument with generating functions shows that the number of loops at a given vertex v is $\mathrm{Ge}(1-p/(n-np+p))$; these numbers are independent, and thus $L \sim \mathrm{NegBin}\big(n, 1-p/(n-np+p)\big)$ with $\mathbb{E}\,L = p/(1-p) = \lambda = O(1)$ and $\mathrm{Var}(L) = p(1-p+p/n)/(1-p)^2 = O(1)$ [5]. Moreover, it is easily seen that asymptotically, L has a Poisson distribution, $L \xrightarrow{\mathrm{d}} \mathrm{Po}(\lambda)$.

T: This follows immediately from Theorem 9, since $T = n - N$.

B, F: Let v, w be two distinct vertices. If the DFS finds w as a descendant of v, then there will later be $\mathrm{Ge}(1-p)$ arcs from w, and each has probability $1/n$ of being a back arc to v. Similarly, there will be $\mathrm{Ge}(1-p)$ future arcs from v, and each has probability $1/n$ of being a forward arc to w. Hence, if I_{vw} is the indicator that w is a descendant of v, and B_{vw} [F_{vw}] is the number of back [forward] arcs vw, then

$$\mathbb{E}\,B_{wv} = \mathbb{E}\,F_{vw} = \frac{\lambda}{n}\,\mathbb{E}\,I_{vw}. \tag{62}$$

Summing over all pairs of distinct v and w, we obtain

$$\mathbb{E}\, B = \mathbb{E}\, F = \frac{\lambda}{n}\,\mathbb{E}\sum_w \sum_{v\neq w} I_{vw} = \frac{\lambda}{n}\,\mathbb{E}\sum_w d(w) = \lambda\,\mathbb{E}\,\overline{d}, \tag{63}$$

and (61) follows by Corollary 7.

The proofs of (56)–(58) and the equality $\mathbb{E}\, T = \mathbb{E}\, C$ are given in the full paper [4]. ◀

3 Depth, trees and arc analysis in the shifted geometric outdegree distribution

In this section, the outdegree distribution is $\mathrm{Ge}_1(1-p) = 1 + \mathrm{Ge}(1-p)$. Thus we now have the mean

$$\lambda = \frac{1}{1-p}.$$

As in Section 2, the depth $d(t)$ is a Markov chain given by (7), but the distribution of ξ_t is now different. The probability (3) is replaced by $(1-t/n)/(1-pt/n)$, but the number of future arcs from an ancestor is still $\mathrm{Ge}(1-p)$, and, with $\theta := t/n$,

$$\mathbb{P}\big(\xi_t = k\big) = \begin{cases} \overline{\pi}_t := \frac{1-\theta}{1-p\theta}, & k = 0, \\ (1-\overline{\pi}_t)(1-\pi_t)^{k-1}\pi_t, & k \geqslant 1, \end{cases} \tag{64}$$

where $\pi_t = p\overline{\pi}_t$ is as in (8). The probability generating function of ξ_t is, instead of (11),

$$p(t,z) = \overline{\pi}_t + (1-\overline{\pi}_t)\frac{\pi_t z}{1-(1-\pi_t)z} = (1-\theta)\frac{1-(1-p)z}{1-p\theta-(1-p)z}. \tag{65}$$

The rest of the analysis does not change, and the results in Theorems 5–12 still hold, but we get different values for many of the constants.

We now have $\mathbb{E}\,\xi_t = \frac{(1-p)\theta}{p(1-\theta)}$ and instead of (16) we have $\mathbb{E}\,\widetilde{d}(t) = n\widetilde{\ell}(\theta) + O(1)$ where now $\widetilde{\ell}(\theta)$ takes the new value

$$\widetilde{\ell}(\theta) := \frac{1}{p}\theta + \frac{1-p}{p}\log(1-\theta). \tag{66}$$

Note that $\widetilde{\ell}(\theta_1) = 0$ still gives (18), now with $\lambda = 1/(1-p)$, and that $\lambda > 1$ for every p. Differentiating (66) shows that the maximum point $\theta_0 = p > 0$.

Figure 2 shows $\widetilde{\ell}(\theta)$ for both geometric distributions.

Straightforward calculations yield

$$\upsilon := \widetilde{\ell}(p) = 1 + \frac{1-p}{p}\log(1-p), \tag{67}$$

$$\alpha := \frac{1}{p}\left(\frac{\theta_1^2}{2} - \frac{1}{\lambda}((1-\theta_1)\log(1-\theta_1) - \frac{1}{\lambda}\theta_1\right) = \theta_1 - \frac{\theta_1^2}{2p}, \tag{68}$$

$$\psi := 1 - \theta_1 - \frac{\lambda}{2}(1-\theta_1)^2. \tag{69}$$

Note that the value for ψ is the same as in (37). This is no coincidence; we show by the method in Section 4 that this holds for all offspring distributions with the same mean λ.

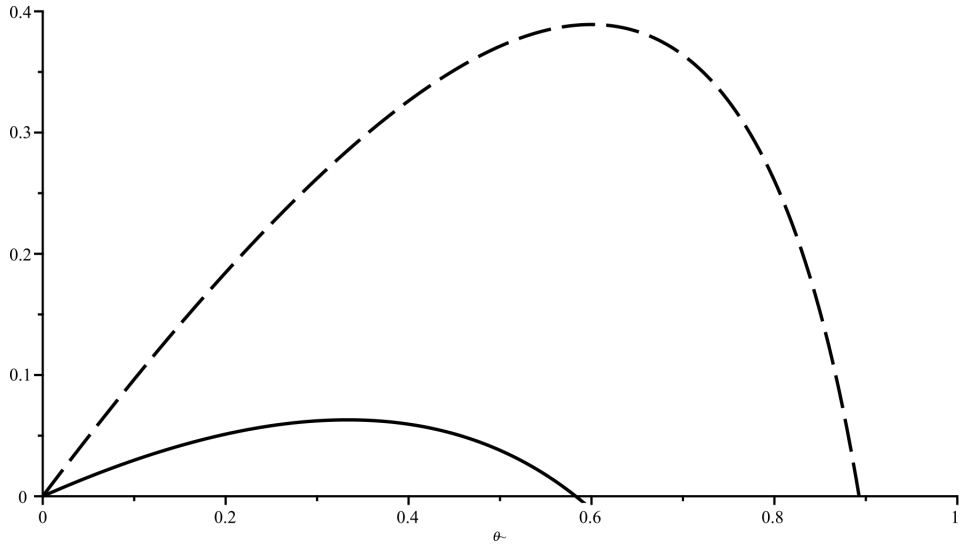


Figure 2 $\widetilde{\ell}(\theta)$, the asymptotic search depth, for geometric distribution (solid) and shifted geometric distribution (dashed) with $p = 0.6$.

Now the expected numbers of back and forward arcs differ since $\mathbb{E}\, B = \lambda\, \mathbb{E}\, \overline{d} \sim \lambda\alpha n$ and $\mathbb{E}\, F = (\lambda - 1)\, \mathbb{E}\, \overline{d} \sim (\lambda - 1)\alpha n$ because the average number of future arcs at a vertex after a descendant have been created is $\lambda - 1$. Thus the equality $\beta = \varphi$ and the equality of the expected number of back and forward arcs in Theorems 12 and 13 was an artefact of the geometric degree distribution.

The estimates (54)–(58) in Theorem 12 hold, with the constants now given by

$$\tau := 1 - \psi = \theta_1 + \frac{\lambda}{2}(1 - \theta_1)^2, \tag{70}$$

$$\beta := \lambda\alpha = \lambda\theta_1 - \frac{\lambda}{2p}\theta_1^2 = \lambda\theta_1 - \frac{\lambda^2}{2(\lambda - 1)}\theta_1^2, \tag{71}$$

$$\varphi := (\lambda - 1)\alpha = (\lambda - 1)\theta_1 - \frac{\lambda}{2}\theta_1^2 = \frac{\lambda}{2} - \tau, \tag{72}$$

$$\chi := \frac{\lambda}{2} - \beta = \frac{\lambda}{2}(1 - \theta_1)^2 + \frac{\lambda}{2(\lambda - 1)}\theta_1^2. \tag{73}$$

Note that τ and φ are as in Theorem 12, while β and χ are different. In particular, $\beta \neq \varphi$ as noted above. Similarly, $\chi \neq \tau$, and thus the equality of the expected numbers of tree arcs and cross arcs in Theorem 13 also was an effect of the geometric distribution.

4 Stack index analysis and forest size for a general outdegree distribution

In this section, we consider a general outdegree distribution $\mathbf{P}$, with mean λ and finite variance. When the outdegree distribution is general, the depth does not longer follow an easy Markov chain, since we should keep track of the number of children seen so far at each level of the branch of the tree toward the current vertex.

Instead we get back a Markov chain if we replace the depth by the stack index $I(t)$. The DFS can be regarded as keeping a stack of unexplored arcs, for which we have seen the start vertex but not the end. The stack evolves as follows:

S1. If the stack is empty, pick a new vertex v that has not been seen before (if there is no such vertex, we have finished). Otherwise, pop the last arc from the stack, and reveal its endpoint v (which is uniformly random over all vertices). If v already is seen, repeat.

S2. (v is now a new vertex) Reveal the outdegree m of v and add to the stack m new arcs from v, with unspecified endpoints. GOTO S1.

Let again v_t be the tth vertex seen by the DFS, and let $I(t)$ be the size of the stack when $v(t)$ is found (but before we add the arcs from v_t). Also let η_t be the outdegree of v_t. Then $I(1) = 0$ and, in analogy with (7),

$$I(t+1) = \bigl(I(t) + \eta_t - 1 - \xi_t\bigr)^+, \qquad 1 \leqslant t < n, \tag{74}$$

where ξ_t is the number of arcs leading to already seen vertices before we find a new one; we have $\mathbb{P}(\xi = k) = (1 - \frac{t}{n})(\frac{t}{n})^k$ and thus $\xi_t \sim \mathrm{Ge}(1 - t/n)$.

In analogy with (9), we define also

$$\widetilde{I}(t) := \sum_{i=1}^{t-1} (\eta_i - 1 - \xi_i) = \sum_{i=1}^{t-1} \zeta_i, \tag{75}$$

where we define $\zeta_t := \eta_t - \xi_t - 1$. Then, as in (10),

$$I(t) = \widetilde{I}(t) - \min_{1 \leqslant j \leqslant t} \widetilde{I}(j). \tag{76}$$

Note that

$$\mathbb{E}\,\zeta_t = \mathbb{E}\,\eta_t - \mathbb{E}\,\xi_t - 1 = \lambda - \frac{t/n}{1 - t/n} - 1 = \lambda - \frac{1}{1 - t/n}. \tag{77}$$

Hence, uniformly in $t/n \leqslant \theta^*$ for any fixed $\theta^* < 1$,

$$\mathbb{E}\,\widetilde{I}(t) = \sum_{i=1}^{t-1} \mathbb{E}\,\zeta_i = (t-1)\lambda - \sum_{i=1}^{t-1} \frac{1}{1 - t/n} = n\widetilde{\iota}(t/n) + O(1), \tag{78}$$

where

$$\widetilde{\iota}(\theta) := \int_0^\theta \Bigl(\lambda - \frac{1}{1-\tau}\Bigr)\,\mathrm{d}\tau = \lambda\theta + \log(1-\theta). \tag{79}$$

Let

$$\widetilde{\iota}^+(\theta) := [\widetilde{\iota}(\theta)]^+ = \begin{cases} \lambda\theta + \log(1-\theta), & 0 \leqslant \theta \leqslant \theta_1, \\ 0, & \theta_1 \leqslant \theta \leqslant 1, \end{cases} \tag{80}$$

where again θ_1 is the largest root in $[0,1)$ of (18), now with $\lambda = \mathbb{E}\,\eta_1$, the mean of $\mathbf{P}$. The proof of Theorem 5 applies with very minor differences, and yields:

▶ **Theorem 15.** *Suppose that the outdegree distribution has finite variance. Then*

$$\max_{1 \leqslant t \leqslant n} \bigl|I(t) - n\widetilde{\iota}^+(t/n)\bigr| = O_{L^2}(n^{1/2}). \tag{81}$$

Moreover, v_{t+1} is a root if and only if $I(t) + \zeta_t = I(t) + \eta_1 - 1 - \xi_t < 0$, cf. (74). The arguments in the proof of Theorem 9 apply with minor differences, and show:

▶ **Theorem 16.** *Theorems 9 and 10 hold for any outdegree distribution with finite variance, with* $\psi := 1 - \theta_1 - \frac{\lambda}{2}(1-\theta_1)^2$.

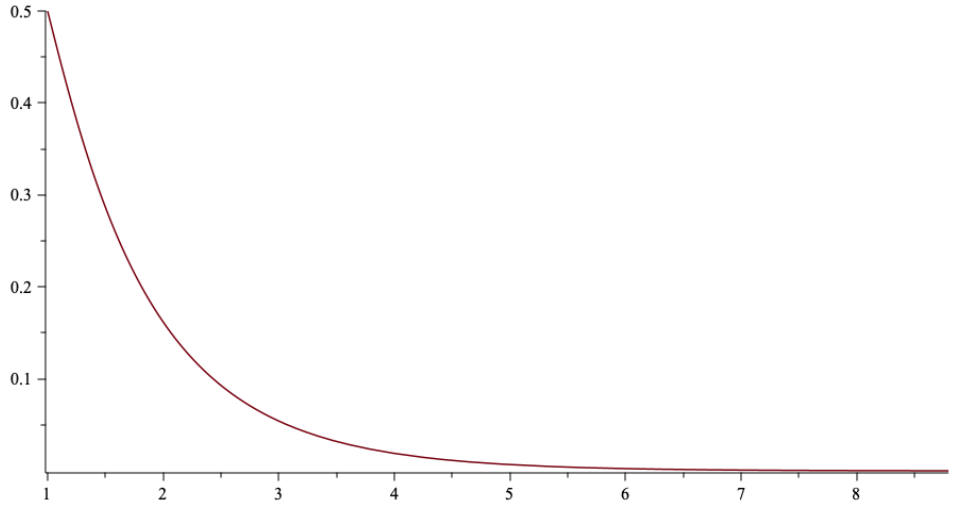


Figure 3 ψ, as function of λ.

Figure 3 shows the parameter ψ as a function of the average degree λ.

Moreover, we have:

▶ **Theorem 17.** *Theorem 11*(i) *holds for any outdegree distribution with mean* $\lambda = 1$ *and positive, finite variance* $\operatorname{Var}\eta$.

Theorem 11(ii) *holds for any outdegree distribution with mean* $0 < \lambda < 1$ *and some finite exponential moment* $\mathbb{E}\, e^{x_0\eta} < \infty$ *for some* $x_0 > 0$.

▶ Remark 18. Theorem 11(i) does not hold for the non-random $\eta = 1$, *i.e.*, when the digraph is a random mapping; in this case the largest trees have sizes of orders $n^{1/2}$, not $n^{2/3}$. (See also [5, Exercise 36 (preliminary number)].)

Moreover, we can use the stack index to find the depth of the nodes, which leads to

▶ **Theorem 19.** *Let* $\rho(\theta)$ *be the survival probability of a Galton–Watson process with offspring distribution obtained from* $\mathbf{P}$ *by thinning, killing each child with probability* θ. *Let*

$$\widetilde{\ell}^{+}(\theta) := \begin{cases} \widetilde{\ell}(\theta) = \int_0^\theta \rho(x)\,\mathrm{d}x, & 0 \leqslant \theta \leqslant \theta_0, \\ \widetilde{\ell}(\breve{\theta}), \ \textit{where } \breve{\theta} \in (0,\theta_0) \textit{ and } \widetilde{\iota}(\breve{\theta}) = \widetilde{\iota}(\theta), & \theta_0 < \theta < \theta_1, \\ 0, & \theta_1 \leqslant \theta \leqslant 1. \end{cases} \tag{82}$$

Then, we have

$$\max_{1\leqslant t\leqslant n} \bigl|d(t) - n\widetilde{\ell}^{+}(t/n)\bigr| = o_{L^2}(n), \tag{83}$$

For the geometric and shifted geometric distributions in Section 2 and 3, we see from (17), (66) and (79) that $\widetilde{\ell}(\theta)$ and $\widetilde{\iota}(\theta)$ are proportional, and thus the expected stack size $\mathbb{E}\,\widetilde{I}(t)$ and depth $\mathbb{E}\,\widetilde{d}(t)$ asymptotically are proportional by a fixed factor independent of t for $t \in (0,\theta_1)$. However, although $\widetilde{\ell}(\theta)$ and $\widetilde{\iota}(\theta)$ always have the same root θ_1, they are in general not proportional; in fact, assuming $\lambda > 1$, this happens only when the outdegree distribution is geometric, or a geometric distribution with $\mathbb{P}(\eta_i = 0)$ modified.

Using the stack index $I(t)$, we can extend many of the results from Section 2 to arbitrary outdegree distributions with finite variance, but not all. Thus some results for geometric distributions remain conjectures in general. In particular, we mention the following. (See the comment after Corollary 6.)

▶ **Conjecture 20.** *We conjecture that as in the geometric case, the height* Υ *is asymptotically normally distributed for any supercritical outdegree distribution with finite variance.*

References

1 David Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997.

2 Sahar Diskin and Michael Krivelevich. On the performance of the Depth First Search algorithm in supercritical random graphs. *arXiv preprint*, 2021. `arXiv:2111.07345`.

3 Nathanaël Enriquez, Gabriel Faraud, and Laurent Ménard. Limiting shape of the depth first search tree in an Erdős-Rényi graph. *Random Structures Algorithms*, 56(2):501–516, 2020.

4 Svante Janson and Philippe Jacquet. Depth-First Search performance in random digraphs. *preprint*, 2022. URL: `https://hal.archives-ouvertes.fr/hal-03537124`.

5 Donald E. Knuth. The Art of Computer Programming, Section 7.4.1.2, 2022. (Preliminary draft, 13 February 2022). URL: `http://cs.stanford.edu/~knuth/fasc12a.ps.gz`.

6 Michael Krivelevich and Benny Sudakov. The phase transition in random graphs: a simple proof. *Random Structures Algorithms*, 43(2):131–138, 2013.

Affirmative Sampling: Theory and Applications

Jérémie Lumbroso
Department of Computer Science, Princeton University, NJ, USA

Conrado Martínez
Department of Computer Science, Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract

Affirmative Sampling is a practical and efficient novel algorithm to obtain random samples of distinct elements from a data stream. Its most salient feature is that the size S of the sample will, on expectation, grow with the (unknown) number n of distinct elements in the data stream. As any distinct element has the same probability to be sampled, and the sample size is greater when the "diversity" (the number of distinct elements) is greater, the samples that Affirmative Sampling delivers are more representative than those produced by any scheme where the sample size is fixed *a priori* – hence its name. Our algorithm is straightforward to implement, and several implementations already exist.

2012 ACM Subject Classification Theory of computation → Data structures design and analysis; Theory of computation → Design and analysis of algorithms; Theory of computation → Sketching and sampling

Keywords and phrases Data streams, Distinct sampling, Random sampling, Cardinality estimation, Analysis of algorithms

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.12

Funding This work has been supported by funds from the MOTION Project (Project PID2020-112581GB-C21) of the Spanish Ministry of Science & Innovation MCIN/AEI/10.13039/501100011033, and by Princeton University, and its Department of Computer Science.

Acknowledgements We want to thank the anonymous reviewers who carefully read the submitted article and made very useful remarks and suggested changes which have allowed us to improve the paper and correct a couple of flawed arguments.

1 Introduction

Drawing random samples from a population is the starting point of any statistical inference, and also in closely related tasks arising in machine learning, information retrieval, data stream analysis, randomized algorithms and many more.

Algorithms to sample elements without replacement from a set of N have a long history [9, 10, 24, 25, 26, 33], but a notable turning point was Vitter's Reservoir Sampling algorithm [32][1] as it was one of the first algorithms not requiring N to be known in advance. After the *reservoir* of size k is filled with the first k elements, it contains at all moments a random sample of size k of the N elements seen so far. This is a useful property that we would also like to have in our sampling algorithms: namely, the ability to return a random sample at any moment.

Despite this, in the context of (large) data streaming, where we deal with streams in which an element can appear many times, Reservoir Sampling may not be the most appropriate tool, as it samples all items with identical probability k/N. If there are a few elements that appear very frequently in the data stream, the reservoir will likely only contain instances of

[1] Refer to Table 1 for a comprehensive comparison of Reservoir Sampling and other existing algorithms, to Affirmative Sampling and its variant both presented in this paper.

these few very frequent elements, and the sample won't be representative enough to make useful statistical inferences (e.g., to infer how many distinct elements there are in the data stream or to make predictions about elements that are not very frequent), or to make these inferences without too much error.

1.1 Novelty of Affirmative Sampling

Our sampling algorithm is novel, as it combines three properties. First, it is an *online algorithm*, that can produce a guaranteed random sample at any time while processing a data stream[2]. Second, it is a *distinct sampling algorithm*, which means it only keeps one copy of a given element, no matter how frequently it may occur in the stream, and therefore avoids being flooded by frequent elements. Third, it is able to *grow the sample size when the number of distinct elements grows*, ensuring we can have a fixed error when inferring population sizes (this is, as far as we know, a completely novel property).

More formally, the data stream $\mathcal{Z} = z_1, \ldots, z_N$ is a (typically very long) sequence of items z_i from some domain or universe $\mathcal{U}$, where there are $n \leq N$ distinct elements (we call n the *cardinality* of $\mathcal{Z}$). Thus underlying $\mathcal{Z}$ we have, on the one hand, the multiset

$$M = M(\mathcal{Z}) = \{x_1^{f_1}, x_2^{f_2}, \ldots, x_n^{f_n}\},$$

where $x_i^{f_i}$ denotes that the i-th distinct element x_i occurs $f_i > 0$ times in $\mathcal{Z}$, and on the other hand, the set X of distinct elements

$$X = X(\mathcal{Z}) = \{x_1, x_2, \ldots, x_n\}.$$

Our goal is to obtain a sample $\mathcal{S} \subset X$ such that any distinct element x_i has the same probability to be sampled as any other. In other words, we want our sampling algorithms to produce *random samples* from the underlying set of distinct elements: any sample $\mathcal{S}$ of size $S = |\mathcal{S}|$ must have probability $1/\binom{n}{S}$ of being returned by the algorithm. Unlike Reservoir Sampling, our algorithm therefore avoids having a sample that could be saturated by frequent elements, by only storing one occurrence of an element in the sample – this is generally called *distinct sampling* (which is the name of a family of algorithms, as well as a specific algorithm [11]).

The problem is challenging because the cardinality n of $\mathcal{Z}$ is unknown, and we would like our algorithm to avoid multiple passes, to spend very little time to process each item in $\mathcal{Z}$ and to use very little memory, e.g., $\Theta(1)$ or $\Theta(\log n)$ (besides the amount of memory for the samples, which will be usually small compared to n, say $\mathcal{O}(\log n)$)[3]. Moreover, the algorithm should not make any statistical assumptions about $\mathcal{Z}$ and it should work online, that is, be able to return a random sample at any given moment of the execution.

Obtaining a random sample of k distinct elements for some fixed value k can be thought as *folklore* (see for instance [6, 8]): using a carefully chosen hash function, it is enough to keep the k distinct elements in $\mathcal{Z}$ with the largest (smallest) hash values seen so far. This method, called bottom-k by several authors has many applications beyond random sampling (see for example [8]). For example, it can be found at the heart of cardinality estimators like in [4, 5]

[2] Nevertheless, the ideas and results in this paper apply in many other contexts.

[3] These requirements, for instance, exclude computing the cardinality n with an initial pass iterating over the stream, and uses some data structure to keep track of elements that have been seen: this would introduce an additional initial pass, which makes the algorithm less time efficient and may not even be an option in online applications; it would require memory proportional to the cardinality, e.g., $\Theta(n)$.

or in the estimation of the similarity of sets [6]. We can assume that the probability that the hash value of some x_i is among the k largest hash values is k/n and thus the corresponding sample will be random. As different occurrences of x_i all will have the same hash value, we will either sample x_i when processing its first occurrence or not all; moreover if some sampled element is evicted from the sample (because some element with a larger hash value is seen and sampled) then that element will not be sampled ever again. The sampling algorithm can thus maintain frequency counters and other useful data about the sample elements because we have them from their very first occurrences; every time we see a new occurrence of an element already in the sample we can update that information accordingly.

▶ **Definition 1.** *We will call* dependable *a sampling algorithm with this property, e.g., that the algorithm can produce exact frequency counts for the elements in its sample.*

Besides this *folklore* algorithm for distinct sampling mentioned above, the first algorithm designed with that purpose in mind is the famed *Distinct Sampling* proposed by Wegman in 1984 and analyzed in depth by Flajolet [11], and popularized by Gibbons[4] [17]. The algorithm was further analyzed by Louchard [28].

The size S of the sample returned by Distinct Sampling is a random variable, and it is used (together with another parameter called the *depth*) to estimate the cardinality n of the data stream. Of course, the samples are random, and they can be used to make statistical inferences about the data stream; Distinct Sampling is also dependable. But the size of the sample is bounded by a fixed value B (the so-called *cache size*), $S \leq B$; because of the way the algorithm works, we have $B/2 \leq \mathbb{E}\{S\} \leq B$ (actually, $B/2 \leq S \leq B$ with high probability).

In this paper we propose a new distinct sampling algorithm called *Affirmative Sampling* (AS, for short). As we have mentioned, the main difference with existing distinct sampling algorithms is that Affirmative Sampling returns random samples of variable size S and $\mathbb{E}\{S\}$ grows with n (despite n is unknown!). This is a very useful property, as larger samples will give more accurate inferences if the size n of the "target" population gets large. As the algorithms previously discussed, AS is dependable (in the meaning introduced by Def. 1) and very easy to understand and implement, with good performance in practice. AS draws upon the ideas behind the cardinality estimator *Recordinality* [20] (briefly reviewed in Subsection 3.3 here) as well as the replacement mechanisms in hiring strategies [2, 20, 19, 21].

1.2 Plan of This Paper

In the first part of the paper (Section 2), we describe AS in detail, discuss its main properties, and then analyze its expected performance (Subsection 2.1). In the second part (Section 3), we discuss and analyze unbiased estimators for the proportion and the absolute number of distinct elements satisfying a certain property (Subsections 3.2 and 3.3, respectively). For example, we might be interested in the proportion or number of elements with relative frequency below 1%, or the number of distinct elements occurring among the last W items seen in the data stream. We also discuss how can we use the algorithm to draw an element close to the median or some other α quantile – assuming some total order to compare elements in the data stream, we want an element that is larger than $\alpha \cdot n$ of the n distinct elements in the data stream (Subsection 3.4). The samples can also be used to accurately estimate, as a by-product, the cardinality n of the data stream (Subsection 3.3).

[4] We will stick to the name *Distinct Sampling* used by Gibbons, instead of *Adaptive Sampling* used by Flajolet and later by other authors.

After that, in Section 4 we discuss variants of the standard AS algorithm; these are simple modifications using slightly different rules to decide when an inspected element becomes part of the sample or not, and in the first case, if some element is evicted from the sample or the sample grows. To investigate these variants we profit from the generality of some of our results (which are not specific to Affirmative Sampling, but to any algorithm drawing random samples of distinct elements) and from the existing literature about the so-called *hiring problem*, a.k.a. *select sets* (see for example [2, 16, 19, 21, 22, 23, 27] and references therein).

We end in Section 5 with some final conclusions and final remarks.

2 The Algorithm

Affirmative Sampling receives a single parameter $k \geq 1$, which we fix in advance, and the data stream $\mathcal{Z}$. If the number n of distinct elements in $\mathcal{Z}$ is equal or smaller than k then the algorithm will read the entire data stream and end with a sample that represents the whole population $X(\mathcal{Z})$. This is rather uninteresting and obviously seldom happening; we can therefore safely assume that $n > k$ (actually, $n \gg k$).

Algorithm 1 Affirmative Sampling: the basic variant.

```
procedure AffirmativeSampling(k, Z)
    fill S with the first k distinct elements (and hash values)
      in the stream Z
    for all remaining z ∈ Z do
        y := hash(z)
        if y < y* ≡ min hash value in S ≡ hash(z*) then
            discard z;
        else if z ∈ S then
            freq[z] := freq[z] + 1; update other statistics of z;
        else if y > k-th largest hash value in S then
            S := S ∪ {z}; freq[z] := 1; ...
        else      ▷ z replaces the element z* with min. hash value
            S := S ∪ {z} \ {z*}; freq[z] := 1; ...
        end if
    end for
    return S
end procedure
```

The initial phase of the algorithm collects in the sample $\mathcal{S}$ the first k distinct elements in the data stream $\mathcal{Z}$, together with their hash values and frequency counts – this entails scanning a subsequence of length $\geq k$ until the $(k+1)$-th distinct element occurs in the data stream. Then we enter the main loop in which we process one by one the remaining items. Let z be the current item. If z is already in $\mathcal{S}$, we just need to increment z's frequency count (and possibly some other associated statistics). Otherwise, we compute a hash value y for z and compare y with the smallest hash value of any element in $\mathcal{S}$. We shall assume here that hash values have enough bits to make the probability of collisions negligible[5] and thus assume

[5] While we cannot rule out collisions on a theoretical basis, we can safely make this assumption on practical grounds.

that every distinct element has a distinct hash value; moreover we will adopt the pragmatic assumption that hash values behave as uniform random variates (see for instance [13]), that is, for any element z and any real value $x \in [0,1]$ we have $\text{hash}(z) \in [0,1]$ and

$$\Pr\{\text{hash}(z) \leq x\} = x.$$

Our first result links the size of the random sample $\mathcal{S}$ returned by AS with the number of records in a random permutation of size n. More specifically with the number of *k-records*; given a permutation σ of size n, we say that i is a (left-to-right) k-record in σ if $\sigma(i) < \sigma(j)$ for at most $k-1$ values of j, $1 \leq j < i$. This is a natural generalization of the well-known notion of records in permutations. Both standard records, $k = 1$, and k-records have been discussed and studied in the literature, see [3] and references therein. Results about the number of k-records in a random permutation were derived in [2] and later largely extended in [21]; although computing the expected number of k-records is a mere exercise, other moments or the probability distribution were not considered in other previous works, as far as we know.

▶ **Lemma 2.** *Let $\mathcal{Y}$ be the permutation of size n induced by the hash values of the corresponding first occurrences of the n distinct elements in a data stream $\mathcal{Z}$. Then the size of the sample $\mathcal{S}$ returned by AS is the number of k-records in $\mathcal{Y}$.*

Proof. Under our assumptions of uniqueness and uniformity of hash values, the sample size increases exactly by one each time that a new distinct element has a hash value that is among the k largest hash values seen so far, that is, when the item is a k-record in the permutation $\mathcal{Y}$. ◀

▶ **Lemma 3.** *Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ be the set of distinct elements in the data stream $\mathcal{Z}$. The sample $\mathcal{S}$ returned by AS is random: any subset of size $|\mathcal{S}|$ has exactly the same probability of being returned. In symbols, for any value of s and any subset $\mathcal{A} \subseteq \mathcal{X}$ of size s*

$$\Pr\{\mathcal{S} = \mathcal{A} \,|\, |\mathcal{S}| = s\} = \frac{1}{\binom{n}{s}}.$$

Equivalently, any element $x \in \mathcal{X}$ has the same probability of being sampled.

Proof. Every time we find a new item z whose hash value is larger than the minimum hash value in $\mathcal{S}$ we either add z to $\mathcal{S}$ or replace the element z^* with minimum hash value with z. Thus, if the sample has size S at any given moment, $\mathcal{S}$ will contain the S distinct elements in $\mathcal{Z}$ with largest hash values. By the way, this also proves that AS is dependable (see Def. 1) – the reasoning is exactly the same as in the case of fixed size samples, returning the k elements with largest hash values seen so far.

Since the hash values are random, any subset $\mathcal{A}$ of s distinct elements has the same probability of being the subset of elements with the s largest hash values. ◀

The size $S := S_{n,k}$ of $\mathcal{S}$ is given by the number of k-records in a random permutation of size n. This random variable is very well understood, including its exact probability distribution (see, for instance [21]). Its expected value and variance are

$$\begin{aligned}
\mathbb{E}\{S\} &= k(H_n - H_k + 1) = k\ln(n/k) + k + \mathcal{O}(1) \\
\mathbb{V}\{S\} &= k(H_n - H_k) - k^2(H_n^{(2)} - H_k^{(2)}) = k\ln(n/k) - k + k^2/n + \mathcal{O}(1),
\end{aligned} \tag{1}$$

where $H_n = \sum_{1 \le j \le n} 1/j = \ln n + \mathcal{O}(1)$ denotes the n-th harmonic number and $H_n^{(2)} = \sum_{1 \le j \ln n} (1/j^2) = \pi^2/6 - 1/n + \mathcal{O}(n^{-2})$ denotes the n-th harmonic number of order 2. The probability generating function of S is [21]

$$\sigma_{n,k}(u) = \sum_{\ell \ge 0} \Pr\{S_{n,k} = \ell\} u^\ell = u^k \frac{\binom{n+k(u-1)}{n}}{\binom{ku}{k}}, \tag{2}$$

from which the expected value and the variance above can be easily derived. Likewise, it's easy to show, albeit cumbersome, that for any $r \ge 0$, $\mathbb{E}\{S^{\underline{r}}\} = \Theta((\log n)^r)$ (we give a proof in Appendix A). From the explicit expression for the probability distribution of S one can also show (see Subsection 3.2)

$$\mathbb{E}\left\{\frac{1}{S}\right\} \sim \frac{1}{\mathbb{E}\{S\}} \sim \frac{1}{k \ln(n/k)}.$$

This last result will be fundamental in our analysis of accuracy of inferences made from the samples drawn using AS (in particular, see Subsection 3.2).

2.1 Complexity

In order to make sure that hash values collide with negligible probability we need that hash values are $\Theta(\log n)$ bits long. A logarithmic number of bits is also necessary for each of the elements in the sample (as they are distinct!), therefore the sample needs total memory $\Theta(|\mathcal{S}| \log n)$; in expectation this is $\Theta(k \log^2 n)$. To support efficient querying and removal of the element with minimum hash value and of the element with k-th largest hash value, one easy and efficient solution is to store the hash values of the sample elements in a min-heap $\mathcal{H}_1$ for the k elements with largest values and another min-heap $\mathcal{H}_2$ for the remaining $|\mathcal{S}| - k$ elements. These two heaps need $\mathcal{O}(|\mathcal{S}| \log n)$ bits. We also need an efficient way to check if a given element belongs or not to the sample; for that we can store the elements of the sample in a hash table (using a hash function which is completely independent and different from the hash function used by the AS algorithm!). That hash table will need $\Theta(|\mathcal{S}| \log n)$ bits. In summary, the memory consumption is a constant factor from the memory needed for the sought result (the size of the sample in bits).

As for the time complexity of the algorithm, the first phase requires time $\mathcal{O}(N_k)$ where $N_k \ge k$ is the number of elements N_k until the first occurrence of the $(k+1)$-th distinct element. At the end of this phase, the min-heap $\mathcal{H}_2$ is empty and the min-heap $\mathcal{H}_1$ contains the first k distinct hash values; it can be constructed in $\Theta(k)$ right at the end of the first phase. Then, during the second phase we process the remaining $N - N_k$ items. Most of them are immediately discarded. Either their hash value is below the minimum hash value in the sample – just check the top of $\mathcal{H}_2$ in time $\Theta(1)$ – or they are already in the sample – this query also needs time $\Theta(1)$ on average using the hash table. Hence these elements can be processed in constant expected time per item. On the other hand, for each of the $S = |\mathcal{S}|$ elements in the table we will have incurred $\mathcal{O}(\log S)$ time to add them; moreover there will be F_n items that were added to the sample and later kicked out from the sample (never to be sampled again), and we also incur time $\mathcal{O}(\log S)$ to add them to the sample and to remove them later. Indeed, when an element has a hash value among the largest k hash values it will be inserted in $\mathcal{H}_1$ (time $\mathcal{O}(\log k)$), but before that the minimum of $\mathcal{H}_1$ must be removed (time $\mathcal{O}(\log k)$) and then added to $\mathcal{H}_2$ (in time $\mathcal{O}(\log(|\mathcal{S}| - k)) = \mathcal{O}(\log |\mathcal{S}|)$). For elements with hash value smaller that the k-th largest hash value but larger than the minimum hash value in the sample (check the top of $\mathcal{H}_1$ and $\mathcal{H}_2$ in constant time), we must remove the minimum in $\mathcal{H}_2$ and add then the new element to $\mathcal{H}_2$ in time $\mathcal{O}(\log |\mathcal{S}|)$.

Let $C(n, N; k)$ denote the complexity of AS with parameter k to process a data stream of length N that contains n distinct elements. Then we can write

$$C(n, N; k) = \Theta(N + k) + \sum_{k<i\leq n} C_i(n, N; k) = \Theta(N) + \sum_{k<i\leq n} C_i(n, N; k)$$

where $C_i(n, N; k)$ is the cost of processing the first occurrence of the i-th distinct element; it is either 0 if the element is discarded (because its hash value is below the minimum hash value in the current sample) or $\mathcal{O}(\log(S_i + 1))$ if it is added to the sample, where S_i denotes the size of the sample just before processing the i-th distinct element. This cost $\mathcal{O}(\log S_i)$ accounts for the addition of the i-th distinct element to the sample (either to $\mathcal{H}_1$ or $\mathcal{H}_2$, which ever corresponds) and the eventual transfer of one element from $\mathcal{H}_1$ to $\mathcal{H}_2$ or the removal of one element from $\mathcal{H}_2$ (and hence from $\mathcal{S}$). We say $C_i = 0$ if the element is discarded because the cost of processing that element is already accounted for in the $\Theta(N)$ term. Likewise the sum starts at $i = k + 1$ because the cost of processing the first k distinct elements is also accounted for within the $\Theta(N + k)$ term. Call Y_i the indicator random for the event "the i-th distinct element is added to the sample". Then

$$\begin{aligned}\mathbb{E}\{C(n, N; k)\} &= \Theta(N) \\ &\quad + \sum_{k<i\leq n} \left\{ \mathbb{E}\{C_i \,|\, Y_i = 1\} \Pr\{Y_i = 1\} + \mathbb{E}\{C_i \,|\, Y_i = 0\} \Pr\{Y_i = 0\} \right\} \\ &= \Theta(N) + \mathcal{O}\left(\sum_{k<i\leq n} \mathbb{E}\{\log S_i\} \Pr\{Y_i = 1\} \right)\end{aligned}$$

The sample size never decreases, hence $S_1 \leq S_2 \leq \cdots S_n \equiv S$; on the other hand $\Pr\{Y_i = 1\} = \mathbb{E}\{Y_i\} = S_i/i$ (it is the probability that the element is among the S_i largest elements of the i elements seen so far) so we can write

$$\begin{aligned}\mathbb{E}\{C(n, N; k)\} &= \Theta(N) + \mathcal{O}\left(\mathbb{E}\{\log S\} \sum_{k<i\leq n} \mathbb{E}\{Y_i\} \right) \\ &= \Theta(N) + \mathcal{O}\left(\mathbb{E}\left\{ \mathbb{E}\{\log S\} \sum_{k<i\leq n} \frac{S_i}{i} \right\} \right) \\ &= \Theta(N) + \mathcal{O}\left(\mathbb{E}\{\log S\} \sum_{k<i\leq n} \mathbb{E}\left\{ \frac{S_i}{i} \right\} \right) \\ &= \Theta(N) + \mathcal{O}\left(\mathbb{E}\{\log S\} \mathbb{E}\{S\} \sum_{k<i\leq n} \frac{1}{i} \right) \\ &= \Theta(N) + \mathcal{O}\left(k \log^2(n/k) \mathbb{E}\{\log S\} \right).\end{aligned}$$

As $\log x$ is concave, Jensen's inequality gives us $\mathbb{E}\{\log S\} \leq \log \mathbb{E}\{S\}$ and putting everything together we find

$$\mathbb{E}\{C(n, N; k)\} = \mathcal{O}\left(N + k \log^2(n/k) \log\log(n/k) \right). \tag{3}$$

3 Making Inferences about the Data Stream

3.1 Frequent Items vs. Frequent Properties

It is important to distinguish between the frequency of an item x and frequent properties P. Consider some property P and the subset $X_P = \{x \in \mathcal{Z} \mid x \text{ satisfies } P\}$ of distinct elements in $\mathcal{Z}$ that satisfy P; we say P is frequent if $n_P = |X_P| = \Theta(n)$. Notice that some properties, like being a very frequent element, can't be a frequent property; for instance, at most $\lfloor 1/c \rfloor$ distinct elements might have relative frequency $\geq c$. As an important consequence, estimating with fixed size samples the proportion of elements in $\mathcal{Z}$ that satisfy P, that is, n_P/n, can be accurately done only if P is frequent.

Distinct Sampling, with its fixed sample size, is thus not a good choice if we want to make inferences about frequent elements (in that case we should sample with probability proportional to the frequency) or the property of interest is not frequent. If not enough elements in the data stream satisfy the property, the samples of variable size drawn by AS can come to the rescue.

3.2 Estimating Proportions

Consider some property $P(x)$ depending only on the occurrences of x in the data stream $\mathcal{Z}$, and let n_P be the number of distinct elements in the data stream satisfying P. We make this restriction as we want to ensure that determining whether x satisfies P or not can be efficiently done and requires little memory.

Let

$$\vartheta_P := \frac{n_P}{n}$$

denote the fraction of elements that satisfy P. If we take a random sample of $S = |\mathcal{S}|$ distinct elements, now with $S > 0$ a random variable, then the probability that there are S_P elements in the sample that satisfy P is given by the hypergeometric distribution

$$\frac{\binom{n_P}{S_P}\binom{n-n_P}{S-S_P}}{\binom{n}{S}}. \tag{4}$$

Let us assume in the computations below and for the remaining of the paper that $n \geq S \geq k \geq 2$, that is, that the sampling algorithm will return at least $k \geq 2$ distinct elements. Otherwise, if the data stream contains less than k distinct element, the sample contains **all** distinct elements in the data stream and their relevant statistics and we can answer queries exactly.

The following theorem is well-known and can be found in many textbook on statistics, see for instance [7].

▶ **Theorem 4.** *Let $\hat{\vartheta}_P := S_P/S$ and assume that $S > 0$. Then*

$$\mathbb{E}\left\{\hat{\vartheta}_P\right\} = \vartheta_P.$$

That is, $\hat{\vartheta}_P$ is an unbiased estimator for ϑ_P, no matter what the probability distribution of $S = |\mathcal{S}|$ is; in particular, it is also true if $S = M$ for some fixed constant $M > 0$.

Proof. We expand $\mathbb{E}\left\{\hat{\vartheta}_P\right\}$ using the formula of total probability conditioning on the events $S=\ell$, and plug-in the hypergeometric distribution for the probability that S_P of the $S=\ell$ elements drawn at random without replacement satisfy the property P. Then

$$\begin{aligned}
\mathbb{E}\left\{\hat{\vartheta}_P\right\} &= \sum_{\ell>0}\mathbb{E}\left\{\frac{S_P}{S}\,\middle|\, S=\ell\right\}\Pr\{S=\ell\} = \sum_{\ell>0}\Pr\{S=\ell\}\sum_{j=0}^{\ell}\frac{j}{\ell}\frac{\binom{n_p}{j}\binom{n-n_p}{\ell-j}}{\binom{n}{\ell}}\\
&= \sum_{\ell>0}\Pr\{S=\ell\}\sum_{j=1}^{\ell}\frac{n_p}{n}\frac{\binom{n_p-1}{j-1}\binom{n-n_p}{\ell-j}}{\binom{n-1}{\ell-1}}\\
&= \frac{n_p}{n}\sum_{\ell>0}\Pr\{S=\ell\}\sum_{j=0}^{\ell-1}\frac{\binom{n_p-1}{j}\binom{n-n_p}{\ell-j-1}}{\binom{n-1}{\ell-1}} = \frac{n_p}{n}\sum_{\ell>0}\Pr\{S=\ell\}\\
&= \frac{n_p}{n} = \vartheta_P,
\end{aligned}$$

where we have used Vandermonde's convolution in the final step to simplify the summation on j, yielding the statement of the theorem. ◀

Quite intuitively, the accuracy of the estimator $\hat{\vartheta}_P$ will depend of the size of the sample. Indeed, for $\mathbb{V}\left\{\hat{\vartheta}_P\right\}$ and now assuming that $n \geq S \geq 2$ we get the following result, which can also be found, for instance, in [31, 7] for the case where the size of the sample S is fixed and not a random variable itself.

▶ **Theorem 5.** *Let $\hat{\vartheta}_P := S_P/S$ and assume that $S \geq 2$. Then*

$$\mathbb{V}\left\{\hat{\vartheta}_P\right\} = \frac{n_p(n-n_p)}{n(n-1)}\cdot\left(\mathbb{E}\left\{\frac{1}{S}\right\}-\frac{1}{n}\right).$$

Proof. We start computing $\mathbb{E}\left\{(\hat{\vartheta}_P)^2\right\}$ following the same steps as in the proof of Theorem 4.

$$\begin{aligned}
\mathbb{E}\left\{\hat{\vartheta}_P^2\right\} &= \sum_{\ell>0}\mathbb{E}\left\{\frac{{S_P}^2}{S}\,\middle|\, S=\ell\right\}\Pr\{S=\ell\} = \sum_{\ell>0}\Pr\{S=\ell\}\sum_{j=0}^{\ell}\frac{j^2}{\ell^2}\frac{\binom{n_p}{j}\binom{n-n_p}{\ell-j}}{\binom{n}{\ell}}\\
&= \sum_{\ell>0}\Pr\{S=\ell\}\sum_{j=1}^{\ell}\frac{j}{\ell}\frac{n_p}{n}\frac{\binom{n_p-1}{j-1}\binom{n-n_p}{\ell-j}}{\binom{n-1}{\ell-1}}\\
&= \frac{n_p}{n}\sum_{\ell>0}\frac{1}{\ell}\Pr\{S=\ell\}\sum_{j=0}^{\ell-1}j\frac{\binom{n_p-1}{j}\binom{n-n_p}{\ell-j-1}}{\binom{n-1}{\ell-1}}\\
&= \frac{n_p}{n}\sum_{\ell>0}\frac{1}{\ell}\Pr\{S=\ell\}\left\{\sum_{j=0}^{\ell-1}(j-1)\frac{\binom{n_p-1}{j}\binom{n-n_p}{\ell-j-1}}{\binom{n-1}{\ell-1}}+\sum_{j=0}^{\ell-1}\frac{\binom{n_p-1}{j}\binom{n-n_p}{\ell-j-1}}{\binom{n-1}{\ell-1}}\right\}\\
&= \frac{n_p}{n}\sum_{\ell>0}\frac{1}{\ell}\Pr\{S=\ell\}\frac{1}{\binom{n-1}{\ell-1}}\left\{(n_P-1)\binom{n-2}{\ell-2}+\binom{n-1}{\ell-1}\right\}\\
&= \frac{n_p}{n}\sum_{\ell>0}\frac{1}{\ell}\Pr\{S=\ell\}\left\{\frac{(n_P-1)(\ell-1)}{n-1}+1\right\}\\
&= \frac{n_P(n_P-1)}{n(n-1)}+\frac{n_P(n-n_P)}{n(n-1)}\mathbb{E}\left\{\frac{1}{S}\right\}
\end{aligned}$$

Hence, assuming that $n \geq S \geq 2$, we have

$$\mathbb{V}\left\{\hat{\vartheta}_P\right\} = \frac{n_p(n-n_p)}{n(n-1)} \cdot \left(\mathbb{E}\left\{\frac{1}{S}\right\} - \frac{1}{n}\right). \qquad \blacktriangleleft$$

If the behavior of the random variable S is smooth enough[6] and $\mathbb{E}\{S\} \to \infty$ when $n \to \infty$ then the accuracy of the estimator $\hat{\vartheta}_P$ will improve, as the variance will decrease and tend to 0 as $n \to \infty$. Indeed, the standard error of $\hat{\vartheta}_P$ satisfies

$$\mathrm{SE}\{\hat{\vartheta}_P\} := \frac{\sqrt{\mathbb{V}\left\{\hat{\vartheta}_P\right\}}}{\vartheta_P} \sim \sqrt{\frac{(1-\vartheta_P)}{\vartheta_P}\,\mathbb{E}\left\{\frac{1}{S}\right\}}. \tag{5}$$

For the samples drawn by AS we have $\mathbb{E}\{1/S\} \sim 1/\mathbb{E}\{S\}$. Indeed, we can start from the bivariate generating function [21] for $S = S_{n,k}$

$$S_k(z,u) = \sum_{\ell\geq 0}\sum_{n\geq 0} \Pr\{S_{n,k} = \ell\}\binom{n}{\ell}u^\ell z^n = \frac{(zu)^k}{(1-z)^{ku+1}}. \tag{6}$$

Integrating $S_k(z,u)/u$ we can recover $\mathbb{E}\{1/S\}$; there is an explicit primitive in terms of the confluent hypergeometric functions, namely Whittaker's $M_{\mu,\nu}(z)$ function [1], and it can be shown that around the singularity $z = 1$ we have

$$\int_0^1 \frac{1}{u} S_k(z,u)\,du \sim_{z\to 1} \frac{z^k}{k(1-z)^{k+1}\ln\left(\frac{1}{1-z}\right)} + \mathcal{O}\left((1-z)^{-k-1}\left(\log\frac{1}{1-z}\right)^{-2}\right),$$

hence, using standard singularity analysis results (see, for instance [14, 15]) we finally get

$$\mathbb{E}\left\{\frac{1}{S}\right\} \sim \frac{1}{\binom{n}{k}}[z^{n-k}]\frac{1}{k(1-z)^{k+1}\ln\left(\frac{1}{1-z}\right)} \sim \frac{1}{k\ln(n/k)}.$$

Then, roughly speaking, the standard error approaches 0 whenever $n/(n_P\,\mathbb{E}\{S\}) \to 0$ and n goes to ∞, which is indeed the case for AS if $n_P = \Omega\left(\frac{n}{\log n}\right)$. In many cases we will be intested in properties such that $n_P = \Theta(n)$, hence the standard error of $\hat{\vartheta}_P$ will decrease and go to 0 as $1/\sqrt{\mathbb{E}\{S\}}$, a rate of decay which we find in many estimators in data stream analysis (e.g., Probabilistic Counting [13], HyperLogLog [12], to name a few) where instead of $\mathbb{E}\{S\}$ we have M, the fixed size of the memory used by those estimators.

3.3 Estimating Cardinalities

We can also estimate the number n_P of elements that satisfy P, not just the proportion n_P/n. For that we can use

$$C_1 := \frac{S_P}{S}\left(k\,(1+1/k)^{S-k+1} - 1\right)$$

or

$$C_2 := \frac{S_P}{S}\frac{S-1}{1-Y_{(S)}},$$

where S_P and S are as in the previous section and $Y_{(S)}$ is the smallest hash value in the sample (the S-th largest hash value in the data stream). Indeed $\mathbb{E}\{C_1\} = \mathbb{E}\{C_2\} = n_p$.

[6] In the sense that $\mathbb{E}\{1/S\} = \mathcal{O}(1/\mathbb{E}\{S\})$; by Jensen's inequality we have only $1/\mathbb{E}\{S\} \leq \mathbb{E}\{1/S\}$.

For the first estimator C_1, recall that the size of the sample $S = |\mathcal{S}|$ returned by AS coincides with the number of k-records in the random permutation of size n induced by the first occurrences of the hash values of the n distinct elements in the data stream. Then the estimator

$$R := k\left(1+\frac{1}{k}\right)^{|\mathcal{S}|-k+1} - 1,$$

called RECORDINALITY, is an unbiased estimator of n (see [20]), that is, $\mathbb{E}\{R\} = n$.

To compute R we would not need to collect the sample $\mathcal{S}$ collected by AS: tracking the k distinct elements in the data stream with largest hash value and how many times we have updated that table would suffice – that is, we would need significantly less memory to achieve the same accuracy to estimate the cardinality. However, if our main goal is to draw a random sample of size growing with n – as drawn by AS– we can get, as a by-product, estimates of the cardinality of the data stream "for free," as well as estimates of the absolute number n_P of elements that satisfy a property P. Indeed,

$$\begin{aligned}
\mathbb{E}\{C_1\} &= \mathbb{E}\left\{\frac{S_P}{S}R\right\} = \sum_{\ell>0}\Pr\{S=\ell\}\,\mathbb{E}\left\{\frac{S_P}{S}R \,\middle|\, S=\ell\right\} \\
&= \sum_{\ell>0}\Pr\{S=\ell\}\sum_{j=0}^{\ell}\frac{j}{\ell}\left(k\left(1+\frac{1}{k}\right)^{\ell-k+1}-1\right)\frac{\binom{n_P}{j}\binom{n-n_P}{\ell-j}}{\binom{n}{\ell}} \\
&= \sum_{\ell>0}\Pr\{S=\ell\}\left(k\left(1+\frac{1}{k}\right)^{\ell-k+1}-1\right)\sum_{j=0}^{\ell}\frac{j}{\ell}\frac{\binom{n_P}{j}\binom{n-n_P}{\ell-j}}{\binom{n}{\ell}} \\
&= \sum_{\ell>0}\Pr\{S=\ell\}\left(k\left(1+\frac{1}{k}\right)^{\ell-k+1}-1\right)\frac{n_P}{n} = \frac{n_P}{n}\,\mathbb{E}\{R\} = n_p.
\end{aligned}$$

Alternatively we can extend the well-know KMV estimator [4, 5, 29] to samples of varying-size. KMV has a parameter $k \geq 2$, fixed in advance, and uses the bottom-k algorithm, that is, collects a table with the k distinct elements with smallest hash values seen so far. Then the largest hash value $Y_{(k)}$ in the table is the k-th smallest hash value in the data stream, that is, the k-th smallest number in a set of n random numbers independently and uniformly drawn in $[0,1]$. Hence

$$Z_k := \frac{k-1}{Y_{(k)}}$$

is an unbiased estimator of n and its standard error is $\mathcal{O}(1/\sqrt{k-2})$.

But once you use AS you have a sample of size $S = |\mathcal{S}|$, with $\mathbb{E}\{S\} = k\ln(n/k)$ and therefore using the smallest hash value in the sample (the S-th largest hash value in the data stream) we can get sharper cardinality estimations. Indeed, if we denote $Y_{(S)}$ the smallest hash value in the sample[7] then $Z := (S-1)/(1-Y_{(S)})$ is an unbiased estimator of

[7] AS is formulated in terms of elements with largest hash values, whereas KMV is formulated in terms of the k-th smallest hash value in the data stream; we have adapted the estimator to work with the k-th largest hash value in the data stream instead. We can therefore use the smallest in the sample $Y_{(S)}$ plugging the value $1 - Y_{(S)}$ in the KMV estimator. Alternatively, AS could be reformulated in terms of the elements with smallest hash values; all its properties would remain unaltered or could easily be adapted, but we have preferred to keep the original formulation in terms of largest hash values, also to honor the inspiration drawn from the RECORDINALITY algorithm.

n, while the standard error goes to 0 as $n \to \infty$ (admittedly, at a very slow rate, i.e., it is $\mathcal{O}(\log^{-1/2} n)$). We do not use here the subscript k in the estimator, as the sample is now of size S, and to distinguish it from the estimator Z_k based in the k-th order statistics of the set of hash values in the sample, for a fixed k.

In what follows we assume that $n \geq k \geq 2$, hence $S \geq k$ (or $\Pr\{S < k\} = 0$). For the expectation we have

$$\begin{aligned}\mathbb{E}\{Z\} &= \sum_{\ell \geq k} \mathbb{E}\{Z \mid S = \ell\} \Pr\{S = \ell\} = \sum_{\ell \geq k} \Pr\{S = \ell\} \mathbb{E}\left\{\frac{(\ell-1)}{1 - Y_{(\ell)}}\right\} \\ &= n \sum_{\ell \geq k} \Pr\{S = \ell\} = n.\end{aligned}$$

Likewise

$$\begin{aligned}\mathbb{E}\left\{Z^2\right\} &= \sum_{\ell \geq k} \mathbb{E}\left\{Z^2 \mid S = \ell\right\} \Pr\{S = \ell\} = \sum_{\ell \geq k} \Pr\{S = \ell\} \mathbb{E}\left\{\frac{(\ell-1)^2}{(1 - Y_{(\ell)})^2}\right\} \\ &= \sum_{\ell \geq k} \Pr\{S = \ell\} \frac{n^{\underline{2}}}{(\ell-2)} = n(n-1)\,\mathbb{E}\left\{\frac{1}{S-2}\right\},\end{aligned}$$

and

$$\mathbb{V}\{Z\} = n(n-1)\,\mathbb{E}\left\{\frac{1}{S-2}\right\} - n^2.$$

It turns out that $\mathbb{E}\{1/(S-2)\} \sim (k \ln(n/k))^{-1}$, hence

$$\mathrm{SE}\{Z\} = \sqrt{\mathbb{E}\{1/(S-2)\}\,(1-1/n) - 1} \sim \frac{1}{\sqrt{k \ln(n/k)}}.$$

Using Z we can also obtain accurate estimates of the number n_P of elements that satisfy P. Following the same steps as in the computation of $\mathbb{E}\{C_1\}$ we obtain

$$\mathbb{E}\{C_2\} = \mathbb{E}\left\{\frac{S_P}{S} Z\right\} = n_P,$$

as we stated at the beginning of this subsection.

3.4 Finding Approximate α-Quantiles

We can use AS to estimate the median and other α-quantiles in the data stream – the rank of an element x being the number of **distinct** elements smaller or equal to x.

Let x^* the α-quantile in the random sample $\mathcal{S}$ returned by AS. That is, $\lceil \alpha \cdot S \rceil$ elements in $\mathcal{S}$ are smaller or equal than x^*. These elements have the property of being $\leq x^*$, and the same proportion of elements in $\mathcal{Z}$ will have that property. That is, the expected value of $\frac{\lceil \alpha \cdot S \rceil}{S} = n_P/n$ is $\alpha + \xi(\alpha, S)$, with $\xi(\alpha, S) \to 0$ as $S \to \infty$. Therefore, the expected rank of x^* in the data stream is $\alpha \cdot n + \xi(\alpha, S) \cdot n \approx \lceil \alpha \cdot n \rceil$.

4 α-Affirmative Sampling: Producing Samples of Size n^α

Instead of using the k-th largest hash value in the sample to decide if a new item is added or not to the sample we can use a different criterion, for instance add it if the hash value is above $(100 \cdot \alpha)\%$ of the hash values. These strategies have also been widely studied (see [16, 22, 19, 23, 27] and the references therein) in the context of the so-called *hiring problem*.

The expected size of the sample when we add elements (not replacing any other, but making the sample grow) if their hash value is above $(100 \cdot \alpha)\%$ of the hash values will be $\Theta(n^\alpha)$. All collected elements with hash value above or equal to the one that defines the "threshold" constitute a random sample of size $\Theta(n^\alpha)$, and there would be no need to use the replacement mechanism, as long as we already get a sample of growing size. However, the algorithm needs to keep not only the elements above the "threshold", but all the elements which were above the "threshold" at some point along the execution of the algorithms. If we apply the replacements mechanism then we guarantee that all collected items, not just those above the threshold, constitute a random sample. This means only a little additional effort, as the expected number of replacements will only be $\Theta(n^\alpha)$ (see [22, 19]). We shall call the resulting algorithm α-*Affirmative Sampling* (α-AS, for short).

The average complexity of processing a data stream of N elements (of which n are distinct) with α-AS will be $\mathcal{O}(N + n^\alpha \log n)$, the analysis follows similar steps to those in Subsection 2.1. The expectation of both $S = |\mathcal{S}|$ and F_n (the number of replacements) is $\Theta(n^\alpha)$. Precise estimates of the constant factor in $\mathbb{E}\{S\}$ and $\mathbb{E}\{F_n\}$ exist, but there are some subtleties that depend on the exact definition of what is the element defining the threshold. For instance, if we add elements whenever their hash value is above the median, it is clear what to do if the size of the sample is odd; but if the size of the sample is even, which is the element that sets the threshold? We refer the reader to [22] or [23] for a more detailed discussion of this issue.

Most results in Section 3 apply, in particular, estimating the proportion of distinct elements that satisfy some property or finding elements close to the median and other quantiles of the data stream. Estimating the absolute number of distinct elements that satisfy P or the cardinality of the data stream is also possible using the generalization of KMV to samples of variable size (that is, the estimators C_2 and Z, respectively, using the smallest hash value in the sample).

To compute the variance and standard error in the estimation of the proportion of distinct elements that satisfy P we would need to know $\mathbb{E}\{1/S\}$; we conjecture that $\mathbb{E}\{1/S\} = \Theta(1/\mathbb{E}\{S\}) = \Theta(n^{-\alpha})$, but because of the lack of information about the distribution of S, we haven't proved it. In the case of the rule "above the median" [22] we have the exact expressions for $\Pr\{S = j\}$. Using the results and techniques there, it is easy to show that

$$\begin{aligned}\mathbb{E}\{1/S\} = \sum_{j\geq 0} j^{-1} \Pr\{S = j\} &\sim \sum_{\ell=1}^{n^{1/2+\epsilon}} \frac{1}{n} e^{-\ell^2/n}(1 + \mathcal{O}(1/\ell)) \\ &\sim \frac{1}{\sqrt{n}} \int_0^\infty e^{-x^2}\, dx = \frac{1}{2}\sqrt{\frac{\pi}{n}},\end{aligned}$$

which is proportional to $1/\mathbb{E}\{S\}$; indeed, $\mathbb{E}\{S\} = \sqrt{\pi n} + \mathcal{O}(1)$ [22].

Likewise, an estimator of the cardinality of the data stream based upon the size S of the sample should be possible if we had a complete characterization of the probability distribution of S; but as far as we know this is only known for the rule "above the median" [22]. In general, we anticipate that $R := \phi \cdot S^{1/\alpha}$ should be an estimator of n, but a detailed knowledge of the probability distribution of S is needed to find the correcting factor ϕ that guarantees that $\mathbb{E}\{R\} \sim n$. In the particular case of the rule "above the median" (the one studied in [22], not the rule in [16] and [27] with $p = 1/2$, there are subtle differences), a suitable unbiased cardinality estimator is $R := S^2/4$. The exact and asymptotic estimation of $\Pr\{S = j\}$ in [22] can be used to show that

Table 1 Sampling Algorithms – a comparative. AS = Affirmative Sampling (this paper). α-AS (this paper, section 4). RS = Reservoir Sampling. BOT = Bottom-k. DS = Distinct Sampling. * = dependable means the exact frequency counts of each element in the sample can be obtained, see Def. 1. ** = conjectured. *** = $S^2/4$ is an unbiased estimator of n, for $\alpha = 1/2$.

Algorithm	AS	α-AS	RS	BOT	DS
Parameter	k	$\alpha \in (0,1)$	k	k	B
Distinct elements	Yes	Yes	No	Yes	Yes
Dependable*	Yes	Yes	No	Yes	Yes
Exp. sample size S	$k\ln(n/k)$	$\Theta(n^\alpha)$	k	k	$\frac{B}{2\ln 2} \approx 0.72\,B$
Std. deviation σ_S	$\sqrt{k\ln(n/k)}$	$\Theta(n^\alpha)$	0	0	$\frac{\sqrt{2}\sqrt{3\ln 2 - 2)B}}{4\ln 2} + \mathcal{O}(1)$
Std. inference err.	$\frac{1}{\sqrt{k\ln(n/k)}}$	$\Theta(n^{-\alpha})^{**}$	No	$1/\sqrt{k}$	$1/\sqrt{B\ln 2}$
Runtime $\mathcal{O}(N+\ldots)$	$k\log^2(n/k)\log\log(n/k)$	$n^\alpha \log n$	$k\log k\log(N/k)$	$k\log k\log(n/k)$	$B\log n$
Memory (in bits)	$\Theta(k\log(n/k)\log n)$	$\Theta(n^\alpha \log n)$	$\Theta(k\log n)$	$\Theta(k\log n)$	$\Theta(B\log n)$
Cardinality estim.	Yes [20]	$S^2/4^{***}$	No	Yes [4, 20, 29]	Yes [11]

$$\mathbb{E}\left\{S^2\right\} = \sum_{j\geq 0} j^2 \Pr\{S=j\} \sim \sum_{\ell=1}^{n^{1/2+\epsilon}} \frac{8\ell^3}{n} e^{-\ell^2/n}(1+\mathcal{O}(1/\ell))$$
$$\sim 8n\int_0^\infty x^3 e^{-x^2}\,dx = 4n,$$

and thus $\mathbb{E}\{R\} = n + o(n)$. The estimator R can also be used to estimate the absolute number n_P of distinct elements satisfying P, namely,

$$C_1 := (S_P \cdot S)/4 \implies \mathbb{E}\{C_1\} \sim n_P.$$

5 Conclusions and Final Remarks

Affirmative Sampling is a simple, elegant and practical algorithm to produce random samples of distinct elements from a "population" of n elements. Contrary to other existing algorithms (which produce random samples of fixed or expected constant size), it is the first distinct sampling algorithm that produces random samples with (expected) size growing with n. In this sense, AS gracefully adapts to the size n of the population and gives us probabilistic guarantees of increasing accuracy of the inferences made, just the opposite of what would happen if the size of the samples is fixed or bounded – this is the case, for instance, of the well known bottom-k and the Distinct Sampling algorithms mentioned in the introduction. Table 1 summarizes the features of the several sampling algorithms introduced in this paper as well as others discussed here.

As we mentioned in our abstract, our algorithm uses primitives that are straightforward to access in most programming languages, and is therefore readily implementable. Several implementations exist, for instance our reference code in Python, see [30].

The idea of replacements can be also used in combination to sampling algorithms such as the famous Reservoir Sampling [32], which returns a random sample of fixed sized from the data stream. As we mentioned in the introduction and in Table 1, Reservoir Sampling samples occurrences z_j in the data stream, not distinct elements. A given element x might appear several times in the sample. Actually, the relative frequency of x within the sample will be, on average, the relative frequency of x in the data stream. To combine replacements with Reservoir Sampling, we just need to change AS so that instead of using the hash value

of each item, we use a random number associated with every element of the data stream. That way, we would be able to collect a random sample of expected size $\sim k \ln(N/k)$ without prior knowledge of the length N of the data stream.

References

1 M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions*. Dover Publ., New York, 1964.

2 M. Archibald and C. Martínez. The hiring problem and permutations. In *Proc. of the 21Int. Col. on Formal Power Series and Algebraic Combinatorics (FPSAC)*, volume AK of *Discrete Mathematics & Theoretical Computer Science (Proceedings)*, pages 63–76, 2009.

3 B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Records*. Wiley series in probability and mathematical statistics. John Wiley & Sons, Inc., New York, 1998.

4 Z. Bar-Yossef, T. S. Jayram, R. Kumar, D. Sivakumar, and L. Trevisan. Counting distinct elements in a data stream. In J. D. P. Rolim and S. P. Vadhan, editors, *Randomization and Approximation Techniques, 6th International Workshop, RANDOM 2002, Cambridge, MA, USA, September 13-15, 2002, Proceedings*, volume 2483 of *Lecture Notes in Computer Science*, pages 1–10. Springer, 2002. doi:10.1007/3-540-45726-7_1.

5 K. S. Beyer, R. Gemulla, P. J. Haas, B. Reinwald, and Y. Sismanis. Distinct-value synopses for multiset operations. *Commun. ACM*, 52(10):87–95, 2009. doi:10.1145/1562764.1562787.

6 A. Z. Broder. On the resemblance and containment of documents. In B. Carpentieri, A. De Santis, U. Vaccaro, and J.A. Storer, editors, *Proc. of the Compression and Complexity of SEQUENCES 1997*, pages 21–29. IEEE Computer Society, 1997. doi:10.1109/SEQUEN.1997.666900.

7 W. G. Cochran. *Sampling Techniques*. John Wiley & Sons, Inc., New York, 3rd edition, 1977.

8 E. Cohen and H. Kaplan. Summarizing data using bottom-k sketches. In Indranil Gupta and Roger Wattenhofer, editors, *Proceedings of the Twenty-Sixth Annual ACM Symposium on Principles of Distributed Computing (PODC 2007)*, pages 225–234. ACM, 2007. doi:10.1145/1281100.1281133.

9 J. Ernvall and O. Nevalainen. An algorithm for unbiased random sampling. *The Computer Journal*, 25(1):45–47, 1982. doi:10.1093/comjnl/25.1.45.

10 C. T. Fan, M. E. Muller, and I. Rezucha. Development of sampling plans by using sequential (item by item) selection techniques and digital computers. *Journal of the American Statistical Association*, 57(298):387–402, 1962. doi:10.2307/2281647.

11 Ph. Flajolet. On adaptive sampling. *Computing*, 43(4):391–400, 1990. doi:10.1007/BF02241657.

12 Ph. Flajolet, É. Fusy, O. Gandouet, and F. Meunier. Hyperloglog: the analysis of a near-optimal cardinality estimation algorithm. In Ph. Jacquet, editor, *Proc. of the 2007 Conference on Analysis of Algorithms (AofA 07)*, volume AH of *Discrete Mathematics & Theoretical Computer Science (Proceedings)*, pages 127–146, 2007.

13 Ph. Flajolet and G. Nigel Martin. Probabilistic counting algorithms for data base applications. *J. Comput. Syst. Sci.*, 31(2):182–209, 1985. doi:10.1016/0022-0000(85)90041-8.

14 Ph. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, 3(1):216–240, 1990. doi:10.1137/0403019.

15 Ph. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. doi:10.1017/CBO9780511801655.

16 J. Gaither and M. D. Ward. Analytic methods for select sets. *Probability in the Engineering and Informational Sciences*, 26:561–568, 2012. doi:10.1017/S0269964812000186.

17 P. B. Gibbons. Distinct sampling for highly-accurate answers to distinct values queries and event reports. In P. M. G. Apers, P. Atzeni, S. Ceri, S. Paraboschi, K. Ramamohanarao, and R. T. Snodgrass, editors, *Proceedings of 27th International Conference on Very Large Data Bases (VLDB 2001)*, pages 541–550. Morgan Kaufmann, 2001.

18 R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison Wesley, 2nd edition, 1994.

19 A. Helmi. *The Hiring Problem and its Algorithmic Applications*. PhD thesis, Dept. Computer Science, Universitat Politècnica de Catalunya, 2013.

20 A. Helmi, J. Lumbroso, C. Martínez, and A. Viola. Counting distinct elements in data streams: the random permutation viewpoint. In N. Broutin and L. Devroye, editors, *Proc. of the 23rd Int. Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA)*, volume AQ of *Discrete Mathematics & Theoretical Computer Science (Proceedings)*, pages 323–338, 2012. doi:10.46298/dmtcs.3002.

21 A. Helmi, C. Martínez, and A. Panholzer. Analysis of the strategy "hiring above the m-th best candidate". *Algorithmica*, 70(2):267–300, 2014. doi:10.1007/s00453-014-9895-3.

22 A. Helmi and A. Panholzer. Analysis of the "hiring above the median" selection strategy for the hiring problem. *Algorithmica*, 66(4):762–803, 2013. doi:10.1007/s00453-012-9727-2.

23 S. Janson. The hiring problem with rank-based strategies. *Electronic Journal of Probability*, 24:1–35, 2019. doi:10.1214/19-EJP382.

24 T.G. Jones. A note on sampling a tape file. *Comm. ACM*, 5(6):343, 1962. doi:10.1145/367766.368159.

25 J. Kawarasaki and M. Bbuya. Random numbers for simple random sampling without replacement. Technical Report 7, Keio University, Dept. Mathematics, 1982.

26 D.E. Knuth. *The Art of Computer Programming: Seminumerical Algorithms*, volume 2. Addison-Wesley, 3 edition, 1997.

27 S. Langowski and M. D. Ward. Moments of select sets. In M. Mishna and J. I. Munro, editors, *Proceedings of the 16th Workshop on Analytic Algorithmics and Combinatorics, ANALCO 2019, San Diego, CA, USA, January 6, 2019*, pages 67–73. SIAM, 2019. doi:10.1137/1.9781611975505.7.

28 G. Louchard. Probabilistic analysis of adaptative sampling. *Random Structures & Algorithms*, 10(1–2):157–168, 1997.

29 J. Lumbroso. An optimal cardinality estimation algorithm based on order statistics and its full analysis. *Discrete Mathematics & Theoretical Computer Science*, 2010.

30 J. Lumbroso and C. Martínez. Affirmative Sampling: Reference Python Implementation, March 2022. doi:10.5281/zenodo.6601690.

31 J. W. Tukey. Some sampling simplified. *Journal of the American Statistical Association*, 45(252):501–519, December 1950. doi:10.2307/2280719.

32 J. S. Vitter. Random sampling with a reservoir. *ACM Trans. Math. Softw.*, 11(1):37–57, 1985. doi:10.1145/3147.3165.

33 J.S. Vitter. Faster methods for random sampling. *Comm. ACM*, 27(7):703–718, 1984. doi:10.1145/358105.893.

A Moments of the sample size S

In this appendix, we will write $S_{n,k}$ for the size of the sample $\mathcal{S}$ generated by AS, when given a data stream that contains n distinct elements and the initial parameter k – we have been not making this dependence on n and k explicit in most of the paper, and have written simply S. As we have discussed $S_{n,k}$ coincides with the number of k-records in a random permutation. Recall the bivariate generating function for $S_{n,k}$ given in (6)

$$S_k(z,u) = \sum_{n\geq k} \binom{n}{k} \sum_{\ell\geq 0} \Pr\left\{S_{n,k}=\ell\right\} z^n u^\ell = \frac{(zu)^k}{(1-z)^{ku+1}}.$$

Differentiating $S_k(z,u)$ r times with respect to u and setting $u=1$, we obtain the generating function

$$S_k^{(r)}(z) = \left.\frac{\partial^r S_k(z,u)}{\partial u^r}\right|_{u=1} = \sum_{n\geq k} \binom{n}{k} \mathbb{E}\left\{S_{n,k}^{\underline{r}}\right\} z^n,$$

where $x^{\underline{r}} := x\cdot(x-1)\cdots(x-r+1)$ denotes the r-th falling factorial of x [18].

The dominant singularity of $S_k^{(r)}(z)$ is located at $z=1$; therefore we have

$$S_k^{(r)}(z) \sim_{z\to 1} \frac{k^r z^k}{(1-z)^{k+1}} \ln\left(\frac{1}{1-z}\right)^r.$$

Applying standard singularity analysis results [14, 15] we obtain an asymptotic estimate for the n-th coefficient, and from there the sought asymptotic estimate for $\mathbb{E}\left\{S_{n,k}^{\underline{r}}\right\}$:

$$[z^n]S_k^{(r)}(z) \sim \frac{k^r n^k}{k!}(\ln n)^r + \text{l.o.t.}$$
$$\mathbb{E}\left\{S_{n,k}^{\underline{r}}\right\} \sim \frac{1}{\binom{n}{k}}[z^n]S_k^{(r)}(z) = k^r(\ln n)^r + \text{l.o.t.}$$

Notice also that $\mathbb{E}\left\{S_{n,k}^{r}\right\} = \Theta(\mathbb{E}\left\{S_{n,k}^{\underline{r}}\right\})$; it is enough to express $S_{n,k}^{r}$ in terms of the falling factorials $S_{n,k}^{\underline{i}}$ $(0\leq i\leq r)$ and the Stirling numbers of the second kind (see, for instance, [18]): for all x and $n\geq 0$

$$x^n = \sum_k \left\{{n \atop k}\right\} x^{\underline{k}}.$$

Mean Field Analysis of an Incentive Algorithm for a Closed Stochastic Network

Bianca Marin Moreno
INRIA Paris, DI-ENS, PSL Research University, Paris, France

Christine Fricker
INRIA Paris, DI-ENS, PSL Research University, Paris, France

Hanene Mohamed
MODAL'X, UMR CNRS 9023, UPL, Université Paris Nanterre, France

Amaury Philippe
Mathematical and Industrial Engineering, Polytechnique Montreal, Montreal, Canada

Martin Trépanier
Mathematical and Industrial Engineering, Polytechnique Montreal, Montreal, Canada

Abstract

The paper deals with a load-balancing algorithm for a closed stochastic network with two zones with different demands. The algorithm is motivated by an incentive algorithm for redistribution of cars in a large-scale car-sharing system. The service area is divided into two zones. When cars stay too long in the low-demand zone, users are encouraged to pick them up and return them in the high-demand zone. The zones are divided in cells called stations. The cars are the network customers. The mean-field limit solution of an ODE gives the large scale distribution of the station state in both clusters for this incentive policy in a discrete Markovian framework. An equilibrium point of this ODE is characterized via the invariant measure of a random walk in the quarter-plane. The proportion of empty and saturated stations measures how the system is balanced. Numerical experiments illustrate the impact of the incentive policy. Our study shows that the incentive policy helps when the high-demand zone observes a lack of cars but a saturation must be prevented especially when the high-demand zone is small.

2012 ACM Subject Classification Mathematics of computing → Queueing theory; Mathematics of computing → Markov processes; Mathematics of computing → Stochastic processes

Keywords and phrases Large scale analysis, mean-field, car-sharing, incentive algorithm, stochastic network, cluster, load balancing, closed Jackson networks, product-form distribution

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.13

Funding *Hanene Mohamed*: This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01)

Acknowledgements The authors would like to thank Communauto for funding and allowing to do this study. They also thank the Natural Science and Engineering Research Council of Canada (NSERC) for funding.

Motivation. Car-sharing, a practice that is gaining ground in urban areas, comes to meet ecological, economic and practical imperatives. For a decade it has been becoming an alternative mode of transportation. The principle is that a given number of vehicles is made available to users at stations or in a public space in a given geographical area to make trips. The user picks up a vehicle if available, makes his trip and then drops it off at his destination.

For the operator, managing such systems is far from simple. The randomness due to the user arrivals as well as to the trips generates an imbalance in the system: Some areas are more or less served by vehicles throughout the day, depending on whether they are residential

areas or in the city center for example. Thus, the users may find themselves without an available vehicle, which alters the efficiency of the system. Rebalancing the network by better distributing the vehicles, in other words, bringing them back where they are needed is a major issue for operators. The usual techniques are either active, such as using trucks to move bikes or drivers for cars, or passive, such as incentive policies that encourage users to move vehicles themselves on their trips. We can cite the example of Velib+ which offered extra time for returning bikes in uphill stations of the Parisian bike-sharing system or the Angel's Rewards bikes program developed in NYC allowing users to earn free day passes and membership extensions.

Gift incentive policy. This paper deals with an incentive policy implemented by Communauto on its free floating car-sharing system in Montreal. In the geographical area, a small zone is identified as a high-demand zone by the operator. Moreover, some cars remain stationary for too long in the rest of the service area with low-demand called the normal zone while users cannot find available cars in the high-demand zone. In order to bring back these stagnant cars from the normal zone to the high-demand zone, Communauto designates them as *gifts* on its application and offers 30 free minutes on the trip if the user returns the *gift* to the high-demand zone. This policy is called here a *gift policy*.

Aim of the paper. The aim is to study the impact of the incentive policy implementing a trip discount to move some cars to a high-demand area. For that, a probabilistic model is proposed for such a system as a large closed stochastic network of interacting particles which are cars and gifts. The service area is divided into cells, called here *stations*, which are nodes of the network, plus extra-nodes containing moving cars and gifts.

Results. We investigate in a Markovian framework the steady state of these stations. Although an invariant measure exists for this irreducible Markov process on a finite state space for a fixed number of stations, it remains untractable. The idea is to deal with the approximation as the number of stations and cars get large together, called the mean-field limit. Indeed, the states of the stations are asymptotically independent and their common distribution is given as a solution of an ODE. See Proposition 1. The equilibrium point of the ODE gives the long-time limit. The special case of a model without incentive policy corresponds to the two-cluster model studied in [6] where the equilibrium point is unique and well determined. Here is a practical application of this framework. For the *gift policy*, Proposition 2 gives a characterization of the equilibrium point as a function of the invariant measure of a random walk in the quarter-plane. It is a first step to address the problem of existence and uniqueness of the equilibrium point.

Performance. Our performance criterion is to minimize the proportion of empty or saturated stations, called for short *problematic*, in order to maximize the efficiency of the system. Since no closed-form solution for the previous invariant measure is derived, we perform a numerical solution of a multidimensional equation for the system with incentive policy. We compare it with the analytical solution of the model without incentives. We study the impact of the policy in the case where everyone follows incentives. This impact is significant when the high-demand zone lacks cars. The risk is to overload it, especially if it is small.

Related works. Mean-field approach comes from physical statistics and is applied for a large class of models of interacting particles. For example biology (neuron networks), economics and social sciences (opinion dynamics). It has applications for stochastic networks as in the

context of load-balancing strategies as Power-of-Choice and others. See references in [10] and [11]. A direct analysis of large networks is difficult because there is either no closed-form expression in general or some but numerically prohibitive. Thus the goal is to find approximations. Mean-field techniques provide an approximation of the evolution of the state distribution of a fixed node as a solution of an ODE. One of the main problems to address is the number of equilibrium points. This issue elucidated, it allows for the performance metrics either explicit expressions or a numerical solution (see [9]), depending on the models. In all these models, the mean-field limit stands for completely connected network while in real systems, the interaction is often weaker. A great difference between the models is the difficulty to address the uniqueness of the equilibrium point. For example, in our study this question is not addressed since numerically solved. Proving existence of multiple equilibrium points can also be hard. In opinion or epidemic models, the set of states of an agent is small (even $\{0, 1\}$) which simplifies the analysis. See [16].

For a direct approach to stochastic network models, let us briefly recall an important class of queueing networks called *closed Jackson networks*. A closed network is a network where customers stay forever in the system with no external arrivals. A Gordon–Newell network, or simply called closed Jackson network, [13] consists of a finite number of nodes, each with one or several servers. The service times of customers at each node are i.i.d. with exponential distribution. The service rate at each node can depend on both the node and its state. Specifically, if there are x_i customers at node i, the service rate is denoted by $\mu_i(x_i)$, where $\mu_i(.)$ is a function from $\mathbb{N} \to \mathbb{R}_+$ with $\mu_i(0) = 0$ and $\mu_i(x) > 0$ for all $x > 0$. At the end of its service, a customer is directed to another node according to a fixed routing matrix. All service times are independent. Note that a Jackson network is the version with external arrivals and a routing matrix including departures to outside (see [14]). Originally, the nodes of such a network are assumed to have infinite capacity. This description can be generalized to a system including nodes with finite capacities. Nevertheless one should describe the way customers must be redirected until reaching to non-saturated nodes, i.e. whose capacity is not yet reached. Consider the blocking-rerouting policy where entering a saturated queue, the customer is rerouted with the routing matrix at infinite speed until he finds an unsaturated node. This blocking-rerouting policy is defined in [1] and used in [8] in a more general version. The main interest of such a class of queueing networks is that the equilibrium distribution is explicit with a product-form expression (see [1]).

It is conventional to model vehicle-sharing systems as closed Jackson networks where the vehicles play the role of the customers. The associated network has two types of nodes: one-server nodes, that describe the stations, and infinite-server nodes, the latter corresponding to the different routes linking the stations. Service times at these nodes are respectively inter-arrival times of users at stations and trip times along the corresponding routes. In a pioneering paper [4], an asymptotic analysis of infinite capacity closed Jackson networks at equilibrium is proposed when the number of nodes and customers tend to $+\infty$ at the same rate. It is applied to vehicle-sharing systems in the infinite capacity case. Note that the asymptotic analysis performed in [12] is done when the number of nodes (stations and routes) is fixed, while the number of customers (vehicles) tends to infinity. Both papers crucially rely on the explicit product-form stationary distribution, which is well-known in the infinite capacity case. The case of vehicle-sharing systems with finite capacity is considered in [8]. The model is identified as a Jackson network with the blocking-rerouting policy previously described. By [1], the invariant distribution has a product-form in this case.

1 The model

1.1 Model description

In this following description and in the whole paper, a car is always a *normal car* and not a gift. We propose a simplified stochastic two-cluster model for car-sharing when including the gift policy. It will be further discussed in Section 4.1. The principle of the model is the following.

- The arrival process of users at any station of cluster i is Poisson with parameter λ_i, where $i \in \{1, 2\}$. As the rate of user arrivals is larger in cluster 1 than cluster 2, we assume that $\lambda_1 > \lambda_2$.
- If the user arrives at a station in cluster 1 where there is an available car, the user picks it up to start a trip. Otherwise he leaves the system.
- Every car parked in cluster 2 becomes a *gift* after a random time with exponential distribution of parameter δ.
- When a user arrives in a station of cluster 2, if there is an available *gift* and an available car in this station, he picks up a gift with probability p, and a car with probability $1 - p$. If there is just one of the resources (gift or car), the user picks it up. Otherwise he leaves the system.
- The car trip is assumed to have an exponential distribution of parameter μ. When a car trip ends, the user chooses cluster i with probability c_i, then he chooses a station at random in this cluster to park the car.
- The *gift* trip is assumed to have an exponential distribution of parameter μ_c. When a *gift* trip ends, the user returns the *gift* car to any station in cluster i with probability q_i. The *gift* parked appears then as a car on the app.
- A station in cluster i has capacity K_i. If the station chosen is full, the user makes another trip until finding a station with an available parking space.

Note that in our model the inter-arrival times of users, trip times and times to become a gift are all independent with exponential distribution. See Figure 1 for an illustration of the model. This modeling was preceded by an analysis of real data. The model will be discussed in 4.1.

1.2 Notations

Let us summarize the notations. For all the following, $i \in \{1, 2\}$ is the cluster type.

- N_i is the number of stations in cluster i.
- $N = \sum_i N_i$ is the total number of stations.
- $\alpha_i = \lim_{N \to +\infty} N_i/N$ is the limiting proportion of stations in cluster i.
- K_i is the capacity of a station in cluster i.
- M is the total number of cars.
- $s = \lim_{N \to +\infty} M/N$ is the limiting mean number of cars per station, called *fleet size parameter*.
- λ_i is the rate of user arrivals at a station in cluster i.
- $1/\mu$ is the mean trip time for a normal car.
- $1/\mu_c$ is the mean trip time for a *gift*.
- δ is the rate at which a car in a station of cluster 2 becomes a *gift*.
- p is the probability that a user takes a gift when cars and gifts are both available.
- q_i is the probability that a user returns the gift in cluster i.
- c_i is the probability that a user returns his normal car in cluster i.

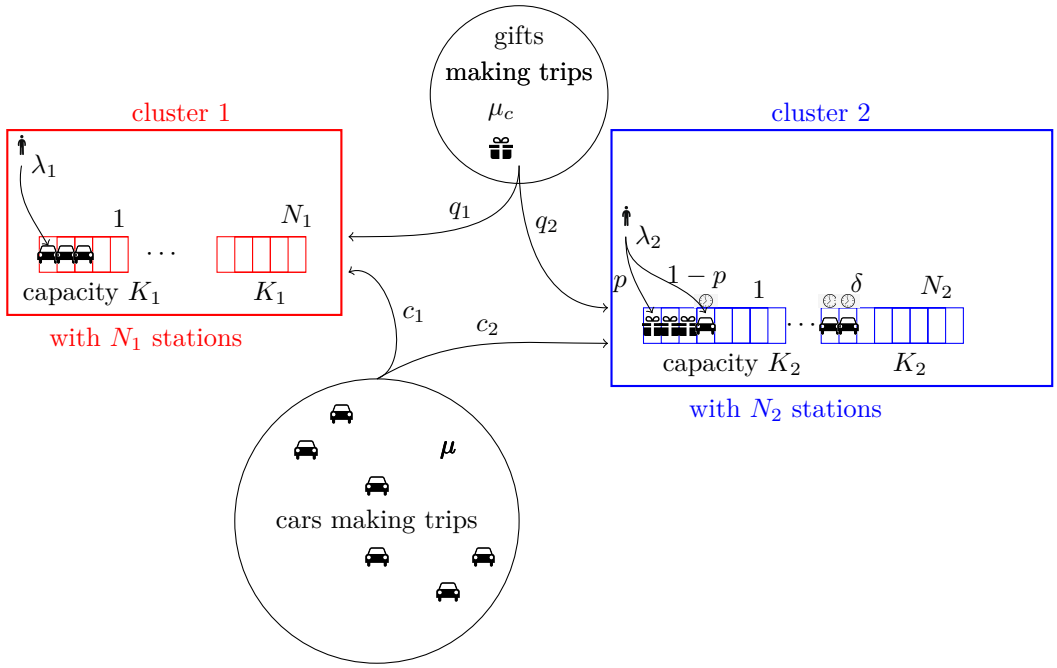


Figure 1 Illustration of the model with gifts.

1.3 Queueing formulation

In this paper, the system can be described as a closed stochastic network. The nodes of the network are a set of $N = N_1 + N_2$ one-server queues of finite capacity (the stations), divided in two clusters, cluster 1 (zone with high demand) with N_1 stations of capacity K_1, cluster 2 (normal zone) with N_2 stations of capacity K_2, plus two infinite-server queues, i.e. the two nodes containing respectively cars and gifts making a trip. The service times at the queues have exponential distribution with parameters respectively λ_1, λ_2, μ and μ_c. According to the queueing vocabulary, there are M customers of two classes: cars and gifts, and a routing matrix given by the previous description.

However, this is not a Jackson network because there are additional transitions since a car in a station of cluster 2 becomes a gift at rate δ and a gift arriving at a station from the infinite-server node becomes a car. It does not fit in this classical framework because of these changes of customer classes. Note that in the case without incentive policy ($\delta \to 0$), the model is a Jackson network since there are no gifts. Such a model (with $\delta \to 0$) is known as a two-cluster bike-sharing system studied in [6] and [8]. Section 2.1.1 is devoted to this case called *model without gifts*.

1.4 The Markov process

The state process is

$$\big(X_{1,n}(t), X_{2,m}(t), C_m(t), Z^N(t),\ 1 \le n \le N_1 \text{ and } 1 \le m \le N_2\big)$$

where

- $X_{1,n}(t)$ is the number of cars at a station n in cluster 1 at time t,
- $X_{2,m}(t)$ is the number of cars at a station m in cluster 2 at time t,
- $C_m(t)$ is the number of *gifts* at a station m (necessarily in cluster 2) at time t and
- $Z^N(t)$ is the number of *gifts* making a trip at time t.

Note that the number of cars making a trip at time t is equal to

$$M - \sum_{n=1}^{N_1} X_{1,n}(t) - \sum_{m=1}^{N_2} \big(X_{2,m}(t) + C_m(t)\big) - Z^N(t).$$

As we deal with a two-cluster model, it is sufficient to study the behavior of one station in each cluster. It amounts to dealing with the empirical measure process

$$(Y^N(t)) = \Big(Y_{1,j}^{N_1}(t), Y_{2,k,l}^{N_2}(t), \frac{Z^N(t)}{N}, j \in \chi_1, (k,l) \in \chi_2\Big)$$

where $Y_{1,j}^{N_1}(t)$ is the proportion of stations with j cars in cluster 1 and $Y_{2,k,l}^{N_2}(t)$ is the proportion of stations with k cars and l *gifts* in cluster 2, defined by

$$Y_{1,j}^{N_1}(t) = \frac{1}{N_1} \sum_{n=1}^{N_1} \mathbf{1}_{\{X_{1,n}(t)=j\}} \text{ and } Y_{2,k,l}^{N_2}(t) = \frac{1}{N_2} \sum_{m=1}^{N_2} \mathbf{1}_{\{\left(X_{2,m}(t),C_m(t)\right)=(k,l)\}}$$

where $\chi_1 = \{j \in \mathbb{N},\ j \le K_1\}$ and $\chi_2 = \{(k,l) \in \mathbb{N}^2,\ k+l \le K_2\}$. Because the inter-arrival times, trip times and times to become a gift have exponential distribution, $(Y^N(t))$ is a Markov process, with finite state space

$$S^N = \Big\{ y = (y_{1,j}, y_{2,k,l}, z)_{\{j\in\chi_1,(k,l)\in\chi_2\}}, y_{1,j} \in \frac{\mathbb{N}}{N_1}, y_{2,k,l} \in \frac{\mathbb{N}}{N_2}, z \in \frac{\mathbb{N}}{N}, \sum_{j\in\chi_1} y_{1,j} = 1,$$
$$\sum_{(k,l)\in\chi_2} y_{2,k,l} = 1, \sum_{j\in\chi_1} j\, y_{1,j} + \sum_{(k,l)\in\chi_2} (k+l)\, y_{2,k,l} + z \le M \Big\}.$$

The inequality in the previous definition of the state space S^N is due to the fact that the number of cars driving has be to added to the left-hand side of the inequality to obtain the total number M of cars in the system. Let us write its transitions from state $y \in S^N$. To simplify the notations, let us denote by

$$E_1 = \sum_{j\in\chi_1} j\, y_{1,j} \quad \text{and} \quad E_2 = \sum_{(k,l)\in\chi_2} (k+l)\, y_{2,k,l} \tag{1}$$

respectively the mean number of cars parked in cluster 1 and the mean number of cars plus gifts parked in cluster 2. Also, let us denote by $(e_{1,j}, e_{2,k,l}, e_0, j \in \chi_1, (k,l) \in \chi_2)$ the canonical basis of $\mathbb{R}^{|\chi_1|+|\chi_2|+1}$, where the cardinality of set A is denoted by $|A|$. The transitions, from state $y = (y_{1,j}, y_{2,k,l}, z) \in S^N$, are due to three events: a user arrival, a gift appearance or a car return. For example, when a user arrives at a station of cluster 2 with k cars and l gifts (for short of type $(2,k,l)$) to take a gift ($l > 0$), the number of gifts decreases by 1. Since there are $y_{2,k,l}N_2$ possible stations, this happens at rate $\lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{l>0\}}\big(p + (1-p)\mathbf{1}_{\{k=0\}}\big)$. Recall that p is the probability for a user arriving to a station in cluster 2 to choose a gift when cars and gifts are available. Thus the corresponding transition is the following.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k,l-1} - e_{2,k,l}) + \frac{e_0}{N} \quad \text{at rate} \quad \lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{l>0\}}\big(p + (1-p)\mathbf{1}_{\{k=0\}}\big).$$

The other transitions are presented in the appendix. These transitions allow us to write the drift of process $(Y^N(t))_t$ which will be useful to the mean-field convergence (Proposition 1).

2 Mean-field limit

Our aim is to investigate the system when M, N_1 and N_2 get large at the same rate, for short, when N gets large. When N tends to $+\infty$, the process $(Y^N(t))$ given by the previous transitions converges in distribution to a deterministic function which is the unique solution of a given ODE. This result is given by the following proposition.

▶ **Proposition 1** (Mean-field convergence). *For $T > 0$, $(Y^N(t))_{t\in[0,T]}$ converges in distribution to the unique solution $(y(t))_{t\in[0,T]}$ of the following ODE with $y(0)$ fixed*

$$
\begin{aligned}
\frac{\mathbf{dy_{1,j}}}{\mathbf{dt}}(\mathbf{t}) &= y_{1,j+1}(t)\lambda_1 \mathbf{1}_{\{j<K_1\}} + y_{1,j-1}(t)\left(\frac{c_1}{\alpha_1}\mu(s-\alpha_1E_1(t)-\alpha_2E_2(t)-z(t)) + \frac{q_1\mu_c}{\alpha_1}z(t)\right)\mathbf{1}_{\{j>0\}} \\
&\quad - y_{1,j}(t)\left(\lambda_1\mathbf{1}_{\{j>0\}} + \frac{c_1}{\alpha_1}\mu(s-\alpha_1E_1(t)-\alpha_2E_2(t)-z(t))\mathbf{1}_{\{j<K_1\}} + \frac{q_1\mu_c}{\alpha_1}z(t)\mathbf{1}_{\{j<K_1\}}\right) \\
\frac{\mathbf{dy_{2,k,l}}}{\mathbf{dt}}(t) &= y_{2,k,l+1}(t)\lambda_2\mathbf{1}_{\{k+l<K_2\}}\left(p+(1-p)\mathbf{1}_{\{k=0\}}\right) + y_{2,k+1,l}(t)\lambda_2\mathbf{1}_{\{k+l<K_2\}}\left(1-p+p\mathbf{1}_{\{l=0\}}\right) \\
&\quad + y_{2,k+1,l-1}(t)\,\delta\,(k+1)\,\mathbf{1}_{\{k<K_2\}} + y_{2,k-1,l}(t)\left(\frac{c_2}{\alpha_2}\,\mu\,(s-\alpha_1E_1(t)-\alpha_2E_2(t)-z(t))\right. \\
&\quad \left. + \frac{q_2\mu_c}{\alpha_2}z(t)\right)\mathbf{1}_{\{k>0\}} - y_{2,k,l}(t)\left(\lambda_2(1-\mathbf{1}_{\{k=0,l=0\}}) + \delta k + \frac{c_2}{\alpha_2}\mu(s-\alpha_1E_1(t)-\alpha_2E_2(t)\right. \\
&\quad \left. - z(t))\mathbf{1}_{\{k+l<K_2\}} + \frac{(1-q)\mu_c}{\alpha_2}z(t)\mathbf{1}_{\{k+l<K_2\}}\right) \\
\frac{\mathbf{dz}}{\mathbf{dt}}(\mathbf{t}) &= -\,q_1\,\mu_c z(t)\sum_{j\in\chi_1} y_{1,j}(t)\,\mathbf{1}_{\{j<K_1\}} + \alpha_2\,\lambda_2 \sum_{(k,l)\in\chi_2} y_{2,k,l}(t)\,\mathbf{1}_{\{l>0\}}\left(p+(1-p)\mathbf{1}_{k=0}\right) \\
&\quad - q_2\mu_c z(t)\sum_{(k,l)\in\chi_2} y_{2,k,l}(t)\mathbf{1}_{\{k+l<K_2\}}.
\end{aligned}
\tag{2}
$$

Recall that, in these equations, s is the limiting number of cars per station and α_i the limiting proportion of stations in cluster i, $i \in \{1,2\}$.

The proof is standard (see [2]). The idea of the proof is that a Markov process can be written as the sum of a martingale term and a drift term in form of an integral on time. When N is large, one can prove that the process is tight. Moreover, the martingale term converges to 0. Then any limiting value satisfies an ODE. The uniqueness of the solution of the ODE gives the convergence of the process.

2.1 The equilibrium point

To investigate the steady-state behavior of the model, we study the equilibrium point $\bar{y}$ of the mean-field ODE written as follows

$$\frac{\mathbf{dy}}{\mathbf{dt}}(\mathbf{t}) = F(y(t))$$

where F comes from Proposition 1. It amounts to finding $\bar{y}$ such that

$$F(\bar{y}) = 0. \tag{3}$$

Note that the vector $\bar{y}$ is of dimension $1 + |\chi_1| + |\chi_2| = 1 + K_1 + K_2(1+K_2)/2$. Finding a closed-form expression of the equilibrium point $\bar{y}$ is out of reach. Let us present two points of view: the first one is based on a nice queueing interpretation which holds for the no-gift case. The second is an analytic approach which should be relevant for the case with gifts but is beyond this work.

2.1.1 The queueing interpretation for the no-gift case

In this case, the existence and uniqueness of the equilibrium point $\bar{y}$ is proved. See [6] for details. In addition, $\bar{y}$ is given by a simple queueing interpretation of the mean-field limit. It gives that the limiting stationary number of cars at a station of cluster i, considered as a $M/M/1/K_i$ queue, has a geometric distribution $\nu_{\rho r_i,K_i}$ on $\{0,\dots,K_i\}$ with parameter ρr_i where for $i=1,2$, $r_i=\Lambda\mu\beta_i/\lambda_i$ with $\beta_i=q_i/\alpha_i$, $\Lambda=1/\max_i(\mu\beta_i/\lambda_i)$ and ρ is the unique solution of the fixed point equation

$$s=\rho\Lambda+\sum_{i=1}^{2}\alpha_i m(\nu_{\rho r_i,K_i}). \tag{4}$$

In the previous equation, we denote by $m(\nu_{\rho,K})$ the mean of the geometric distribution $\nu_{\rho,K}$ on $\{0,\dots,K\}$ with parameter ρ, given by

$$m(\nu_{\rho,K})=\begin{cases}\frac{K}{2} & \text{if } \rho=1\\ \frac{\rho}{1-\rho}-\frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}} & \text{otherwise}\end{cases} \tag{5}$$

because, for $\rho=1$, $\nu_{\rho,K}$ is the uniform distribution on $\{0,\dots,K\}$. It shows that the multidimensional equilibrium point equation (3) amounts to fixed point equation (4) on $\mathbb{R}_+$. This is the purpose of [6, Theorem 1] for the cluster case detailed in [6, Section 2.3].

2.1.2 Characterization of the equilibrium point

Taking into account the gift policy induces a change of classes between normal cars and gifts. This considerably complicates the search for an equilibrium point and changes the nature of the limiting objects involved. The question of existence and uniqueness of a solution of the equilibrium point in equation (3) remains open. For simplicity, let us take the case $p=q_1=1$ in order to highlight the main difficulties of this problem. Remembering that $p=1$ means that, when available, a gift is always chosen over a car in a station of the normal zone, and $q_1=1$ means that all gifts are returned at a station of cluster 1. Heuristically, looking for an equilibrium point $\bar{y}$ means that the right-hand term in the mean-field ODE (2) is null. With obvious notations $\bar{y}=(\bar{y_1},\bar{y_2},\bar{z})$, note first that the number of moving gifts $(Z^N(t)/N)$ is the rescaled number of customers in a $M/M/\infty$ queue introduced by Kelly [15] whose limit is $(z(t))$ which tends to the load parameter (see [17, Section 6.5])

$$\bar{z}=\frac{\alpha_2\lambda_2}{\mu_c}\frac{1-\bar{y}_{2,.,0}}{1-\bar{y}_{1,S}} \tag{6}$$

where $\bar{y}_{1,S}$ is the probability that a station in cluster 1 is saturated and $\bar{y}_{2,.,0}$ the probability that a station in cluster 2 has no gift, i.e. $1-\bar{y}_{1,S}=\sum_{j\in\chi_1}\bar{y}_{1,j}\mathbf{1}_{\{j<K_1\}}$ and $1-\bar{y}_{2,.,0}=\sum_{(k,l)\in\chi_2}\bar{y}_{2,k,l}\mathbf{1}_{\{l>0\}}$. Then a queuing interpretation similar to that for the no-gift case holds. Indeed, at equilibrium, a station of cluster 1 can be considered as a $M/M/1/K_1$ queue, with arrival rate

$$\bar{\gamma_1}=\frac{1}{\alpha_1}\left(c_1\mu(s-\alpha_1\bar{E}_1-\alpha_2\bar{E}_2-\bar{z})+q_1\mu_c\bar{z}\right) \tag{7}$$

where $\bar{E}_i$ are defined by (1) and service rate λ_1. It is well known that its invariant measure is a geometric distribution on $\{0,\dots,K_1\}$ with parameter $\bar{\rho}_1=\bar{\gamma}_1/\lambda_1$, i.e. $\bar{y}_{1,j}=\rho_1^j(1-\rho_1)/(1-\rho_1^{K+1})$ for $0\le j\le K_1$. Note that, plugging equation (6) into (7), $\bar{\rho}_1$ depends on $\bar{y}$, only by $\bar{y}_1$ and $\bar{y}_2$. Moreover $\bar{y}_2=\pi_{\bar{\rho}_2,K_2}$ where

$$\bar{\rho}_2=\frac{1}{\lambda_2\alpha_2}\left(c_2\mu(s-\alpha_1\bar{E}_1-\alpha_2\bar{E}_2-\bar{z})+q_2\mu_c\bar{z}\right)$$

and, for fixed ρ, π_{ρ,K_2} is the invariant measure of the Markov process on χ_2 with matrix jump $Q_{\rho,K}$ given by its non-null non-diagonal terms

$$\begin{cases} Q_{\rho,K}(n, n-e_1) & = \lambda_2 \mathbf{1}_{\{n_1>0\}} \\ Q_{\rho,K}(n, n+e_2) & = \lambda_2 \rho \mathbf{1}_{\{n_2<K_2\}} \\ Q_{\rho,K}(n, n+e_1-e_2) & = \delta \mathbf{1}_{\{n_2>0\}}. \end{cases} \tag{8}$$

In conclusion the equilibrium point $\bar{y}$, solution of a multidimensional fixed point equation, can be expressed as a function of $(\bar{\rho}_1, \bar{\rho}_2)$ solution of a fixed point equation. It is summarized by the following result.

▶ **Proposition 2** (Equilibrium point). *An equilibrium point of the ODE is given as*

$$\bar{y} = \left(\nu_{\bar{\rho}_1,K_1}, \pi_{\bar{\rho}_2,K_2}, \frac{\alpha_2\lambda_2}{\mu_c} \frac{1 - \sum_{k=0}^{K_2} \pi_{\bar{\rho}_2,K_2}(k,0)}{\sum_{k=0}^{K_1-1} \nu_{\bar{\rho}_1,K_1}(k)} \right)$$

where $\nu_{\bar{\rho}_1,K_1}$ is the geometric distribution on $\{0,\dots,K_1\}$ with parameter $\bar{\rho}_1$, $\pi_{\bar{\rho}_2,K_2}$ the invariant measure associated to $Q_{\bar{\rho}_2,K_2}$ given by (8) and $(\bar{\rho}_1, \bar{\rho}_2)$ is the solution of the fixed point equation

$$\rho_i = \frac{1}{\lambda_i \alpha_i} \left(c_i\mu(s-E) + (q_i\mu_c - c_i\mu) \frac{\alpha_2\lambda_2}{\mu_c} \frac{1 - \sum_{k=0}^{K_2} \pi_{\rho_2,K_2}(k,0)}{\sum_{k=0}^{K_1-1} \nu_{\rho_1,K_1}(k)} \right), \quad i \in \{1,2\} \tag{9}$$

with $E = \alpha_1 E_1 + \alpha_2 E_2$, E_1 and E_2 being the means associated to ν_{ρ_1,K_1} and π_{ρ_2,K_2}.

Proposition 2 reduces the question of existence and uniqueness of the equilibrium point to a fixed point equation. Indeed, proving the existence and uniqueness of the equilibrium point $\bar{y}$ amounts to finding a unique $(\bar{\rho}_1, \bar{\rho}_2)$ solution of the fixed point equation (9). The further analysis (existence and uniqueness) of this fixed point equation is beyond the scope of the paper. In this direction, a first approach would be to find a closed-form expression for the invariant measure π. An analytical method for generating function F associated to invariant measure π is suggested as an alternative to the probabilistic approach (see [3] for details). Some details are given in the following remark.

▶ Remark 3. Let $\gamma_2 = \lambda_2\rho$. The global balance equation associated to π is

$$\begin{aligned} &\pi_{k,l}(\gamma_2 + \lambda_2(1 - \mathbf{1}_{\{k=l=0\}}) + \delta k) \\ &\quad = \mathbf{1}_{\{k+l<K_2\}}(\pi_{k,l+1}\lambda_2 + \pi_{k+1,l}\lambda_2 \mathbf{1}_{\{l=0\}}) + \pi_{k+1,l-1}\delta(k+1)\mathbf{1}_{\{l>0\}} + \pi_{k-1,l}\gamma_2 \mathbf{1}_{\{k>0\}}. \end{aligned} \tag{10}$$

Using the generating function

$$F(x,y) = \sum_{(k,l)\in\chi_2} \pi_{(k,l)} x^k y^l,$$

the global balance equation (10) yields to a functional equation on F. Although the capacity K_2 is assumed to be finite throughout the whole paper, we present here this functional equation for the case $K_2 = +\infty$ for the sake of simplicity

$$\begin{aligned} F(x,y)\left(\gamma_2(1-x) + \lambda_2\left(1 - \frac{1}{y}\right)\right) \\ = F'_x(x,y)\delta(y-x) + \pi_{0,0}\lambda_2\left(1 - \frac{1}{x}\right) + f(x)\lambda_2\left(\frac{1}{x} - \frac{1}{y}\right) \end{aligned}$$

where $f(x) = \sum_{k=0}^{K_1} \pi_{k,0} x^k$.

Such a functional equation is not similar to the classical case studied in [3] due to the derivative term F'_x and we do not solve it. Without an explicit form, we wonder whether the uniqueness problem could be directly solved. Such track is not explored. No closed-form solution of the invariant measure π is derived. Instead of this, the paper gives in Section 3.2 a numerical solution to equation (2).

3 Performance

In order to evaluate the impact of the incentive algorithm on the system behavior, a usual performance metric is used, i.e. the proportion of stations with no vehicle (car or gift) or no parking space available, called *problematic stations*. It characterizes how far the system is unbalanced.

▶ **Definition 4** (Performance Metric). *Let $\bar{y}$ be the equilibrium point of the mean-field ODE obtained by Proposition 1. The performance metric is the limiting stationary proportion Pb of problematic stations given by*

$$Pb = \alpha_1(\bar{y}_{1,0} + \bar{y}_{1,K_1}) + \alpha_2 \left(\bar{y}_{2,0,0} + \sum_{k=0}^{K_2} \bar{y}_{2,k,K_2-k} \right)$$

where K_i is the station capacity and α_i the limiting proportion of stations for cluster i, $i \in \{1, 2\}$.

The first sum in brackets is the proportion of empty and saturated stations in clusters 1, the first term $\bar{y}_{1,0}$ of stations with no car, the second term $\bar{y}_{1,K_1}$ of saturated stations in the high-demand zone. The second sum in brackets is the proportion of empty and saturated stations in cluster 2, $\bar{y}_{2,0,0}$ of stations with neither cars nor gifts and $\sum_{k=0}^{K_2} \bar{y}_{2,k,K_2-k}$ of saturated stations in the normal zone.

Optimizing the proportion of problematic stations means maximizing the number of transactions and the number of satisfied users. Our aim is to compare the performance with gifts and without gifts. The idea is to vary the fleet size parameter s, which is the limiting ratio of the total number of cars M by the total number of stations N, in order to analyze how much flexibility the gift policy gives to an operator who wants to increase the fleet size without harming the system.

3.1 Analysis of the model without gifts

From Section 2.1.1, the proportion of problematic stations Pb in this case is given by

$$Pb = \sum_{i=1}^{2} \alpha_i \frac{1 - \rho r_i}{1 - (\rho r_i)^{K_i+1}} (1 + (\rho r_i)^{K_i+1})$$

where $\alpha_i = \lim_{N\to\infty} N_i/N$. For $i = 1, 2$, the proportion of problematic stations in cluster i as a function of s is given by the parametric curve

$$\rho \mapsto \left(\rho\Lambda + \sum_{i=1}^{2} \alpha_i m(\nu_{\rho r_i, K_i}), \frac{1 - \rho r_i}{1 - (\rho r_i)^{K_i+1}} (1 + (\rho r_i)^{K_i+1}) \right)$$

where the first term $(1 - \rho r_i)/(1 - (\rho r_i)^{K_i+1})$ is the proportion of empty stations in cluster i and the second term $(\rho r_i)^{K_i+1}(1 - \rho r_i)/(1 - (\rho r_i)^{K_i+1})$ is the proportion of saturated stations in cluster i. As explained in Section 5.2 of [6], the proportion of problematic stations in

cluster i has a minimum $2/(K_i+1)$ for ρr_i equal to 1 i.e. for $\rho = 1/r_i$. Thus, plugging in equation (4), this minimum corresponds to

$$s_i^* = \frac{\Lambda}{r_i} + \sum_{i'=1}^{2} \alpha_{i'} m(\nu_{r_{i'}/r_i, K_{i'}}).$$

where $m(\nu_{\rho,K})$ is defined by equation (5). The following result is the translation, with the notations of the paper, of the result of Section 5.2 of [6]. It gives the fleet size which corresponds to the optimal performance for a given cluster in the system without gift policy.

▶ **Proposition 5** (Optimal performance per cluster without gift policy). *For the model without gifts, the limiting stationary proportion of problematic stations in cluster $i \in \{1,2\}$ is minimal and equal to $2/(K_i+1)$ when*

$$s = s_i^* = \alpha_i \left(\frac{K_i}{2} + \frac{\lambda_i}{\mu q_i} \right) + \alpha_{3-i} \left(\frac{\gamma_{3-i}}{1-\gamma_{3-i}} - \frac{(K_{3-i}+1)\gamma_{3-i}^{K_{3-i}+1}}{1-\gamma_{3-i}^{K_{3-i}+1}} \right)$$

where $\gamma_{3-i} = (q_i \lambda_i \alpha_i)/(q_{3-i}\lambda_{3-i}\alpha_{3-i})$. The last term in brackets must be replaced by $K_{3-i}/2$ for $\gamma_{3-i} = 1$.

Note that, for $s = s_i^*$ which minimizes the proportion of problematic stations in cluster i, the proportion of problematic stations in cluster $i' \neq i$ is not optimal and is exactly $\nu_{r_{i'}/r_i, K_{i'}}(0) + \nu_{r_{i'}/r_i, K_{i'}}(K_{i'})$. Thus minimizing the problematic stations in both clusters simultaneously is not possible.

For the values of Figure 2b and $\alpha_1 = \alpha_2 = 0.5$, Proposition 5 gives $s_1^* = 29.9$ and $s_2^* = 13.1$, and for $\alpha_1 = 0.28$ and $\alpha_2 = 0.72$, $s_1^* = 21.9$ and $s_2^* = 20.4$, which can be checked in Figure 2.

Note the U-shape of the curves plotted in Figure 2b. This shape is typical of these performance curves (cf [5]). Indeed, for small values of the mean number of cars per station, the proportion of empty stations is large and close to 1. Similarly, if the mean number of cars per station is large, the proportion of saturated stations is large and close to 1. Since the performance criterion includes both cases, the U-shape is observed. The contribution of empty and saturated stations to the proportion of problematic stations is illustrated by [6, Figure 2] where the proportions of empty, saturated and problematic stations are plotted.

3.2 Numerical solution

First of all, we numerically obtain the equilibrium point $\bar{y}$ of the mean-field ODE established in Proposition 1, the solution of the fixed point equation (3), as a function of the fleet size parameter s. There are many tools to solve such an equation. We use the Anderson method implemented in Scipy, a Python library.

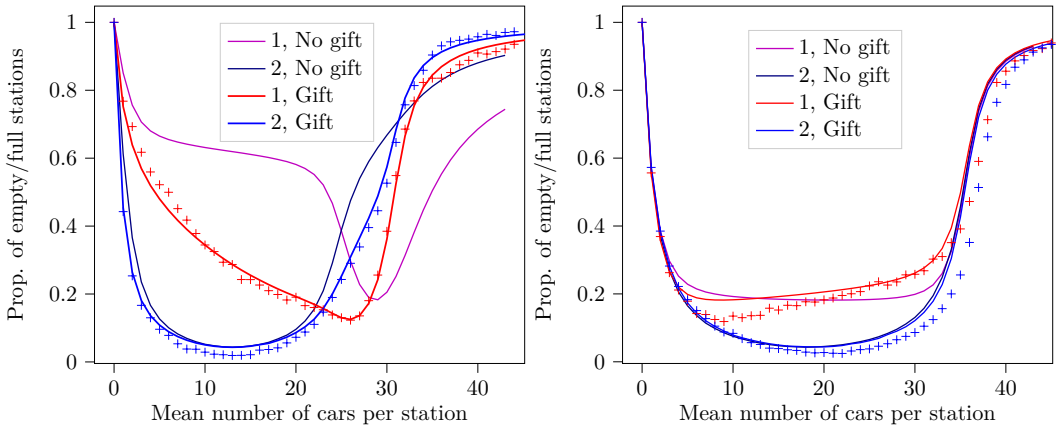
Figure 2a plots the performance Pb numerically obtained as a function of the fleet size parameter s, for the two-cluster model with and without gifts for a naive case: both clusters have the same number of stations, so that $\alpha_1 = \alpha_2 = 0.5$, and everyone follows the gift policy. That means the probability p that a user picks up a gift if gifts and cars are available and the probability q_1 a gift is returned to cluster 1 are equal to 1. All other parameters are given in Figure 2. Figure 2a shows that, for cluster 2, the cases with and without gifts are similar. But, for cluster 1, for this set of parameter values, it seems that an efficient gift policy ($p = q_1 = 1$) would allow an operator to increase the fleet size without harming the system performance and even with improving it. Indeed, for a whole range of values of the

fleet size parameter s, typically $s \leq 20$, the high demand zone suffers from a lack of available cars. About 60% of the stations in the high demand zone are empty for a fleet size parameter s between 10 and 20. The effect of the incentive policy is significant in this case, since the proportion of empty stations in cluster 1 falls under 40% and even reaches 20% for $s = 20$.

Note that the crosses are simulations of the system with $N_1 = 50$ and $N_2 = 50$, the other parameters are given in Figure 2. Compared to the performance curves obtained numerically, it validates that the mean-field limit provides a good approximation for N_1 and N_2 large enough.

Figure 2b plots the performance numerically obtained for the two-cluster model with and without gifts for a more realistic case. The number of stations in the high demand zone is significantly smaller than in the normal one, the ratios are respectively $\alpha_1 = 0.28$ and $\alpha_2 = 0.72$. Figure 2b shows that the performance curves fit for small and large parameter fleet size s for both cases, with and without gifts. In between, there is a plateau where the proportion of problematic stations is close to its minimum. This implies that varying fleet size parameter s around its optimum does not degrade too much the performance which remains close to its optimum. This stability is important for the operator. The minimum proportion of problematic stations should depend on capacities K_1 and K_2, user arrival and trip rates. It is remarkable that the two plateaux correspond to the same values of s. Thus, the stations in cluster 1 do not saturate for s smaller than 30. Despite their small capacity, the high demand in cluster 1 limits the saturation.

In addition, Figure 2b shows that, for a small s, the gift policy slightly improves the performance. It is true until the two curves intersect at $s \simeq 12$. Above this value, on the plateau of cluster 1, the performance is slightly worse with the gift policy. Indeed, gifts seem to saturate cluster 1 and this slightly decreases the system performance. The mean-field approximation is again validated by simulation. See the crosses curve.



(a) $N_1 = N_2 = 50$.

(b) $N_1 = 28$, $N_2 = 72$.

Figure 2 Performance for both clusters (1 for the high-demand zone, and 2 for the normal zone) is numerically computed from equilibrium point equation as a function of the fleet size per station in a system with and without gifts, compared with the simulation curve in crosses. $K_1 = 15$, $K_2 = 45$, $\lambda_1 = 2.6$, $\lambda_2 = 1$, $\mu = \mu_c = 0.65$, $\delta = 1/14$, $c_1 = 0.5$ and $p = q_1 = 1$.

4 Discussion

4.1 Discussion of the model

Discrete Markovian framework. The exponential distributions are assumed to obtain a Markov discrete state process, i.e. the number of gifts and cars in the different stations. It is not true in real systems. This seems to be true for the arrival times of users at a station, but not for the trip times which seem heavy-tailed due to some very long trips. The behavior of the system can be affected by a log-normal trip time distribution compared to an exponential one. As to the threshold, it is deterministic in the real system. Intuitively, the exponential distribution with the same mean for the threshold should not change the behavior of the network. Large stochastic networks with general service time distributions are still largely unexplored. The paper stays in a convenient framework.

Station-based state process. The model does not take into account the detailed moves of the cars, i.e. the fact that a car goes from station i in cluster 1 for example to station j in cluster 2. For the state of stations, these detailed moves do not matter. Mathematically, the detailed routing matrix between stations only affects the stationary behavior of the network by its invariant measure. In other words, this means that the driving cars are indistinguishable and, after their departure, the origin of the trip is no longer important. Only the popularity of stations is significant, expressed as the probability that a car is dropped off at that station.

Space-homogeneity. In order to simplify the presentation, we assume that parameters do not depend on the stations. This mean-field approach can be extended to a completely heterogeneous model. It is out of the scope of the paper.

This modeling was preceded by an analysis of real data which highlighted the low proportion of gifts offered under this incentive policy. Thus we opt for a constant probability p to choose a gift if one is available in the station. Nevertheless, a relevant option is to choose with a probability which takes into account the number of available gifts relative to cars in the station. The study is similar in this case.

Time-homogeneity. In real systems, some parameters, like the arrival rate of users, depend on time. The mathematical model does not take this into account but simulations of the time-inhomogeneous model are performed in Section 4.2.

Reservation. In real car-sharing systems, cars can be booked. It seems that such a study can still be conducted.

4.2 Simulations for time-inhomogeneous arrival process and real trip time distribution

We investigate now the influence of time-inhomogeneity of the parameters, especially the arrival rates of users and also the non exponential trip time distribution. Simulations are performed with arrivals according to a Poisson process with rates λ_1 and λ_2 in the stations of clusters 1 and 2, depending on time, and trip time having the distribution provided by an analysis of real data. To validate the accuracy of this time-inhomogeneous arrivals and a more realistic trip time distribution (see [7]) with Montreal FFCS system dataset, we plot in Figure 3 the average daily rate of arrivals obtained by simulation. It can be compared to that provided by data, plotted in Figure 4.

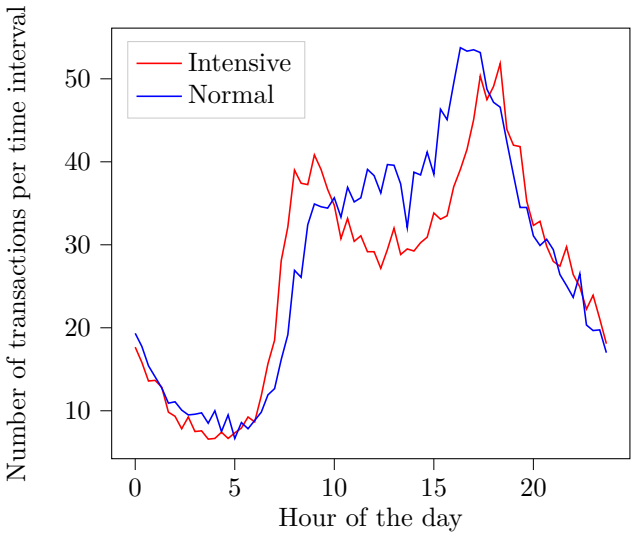


Figure 3 Evolution of a daily arrival rate for the time-inhomogeneous simulation. Time intervals are 20 mn.

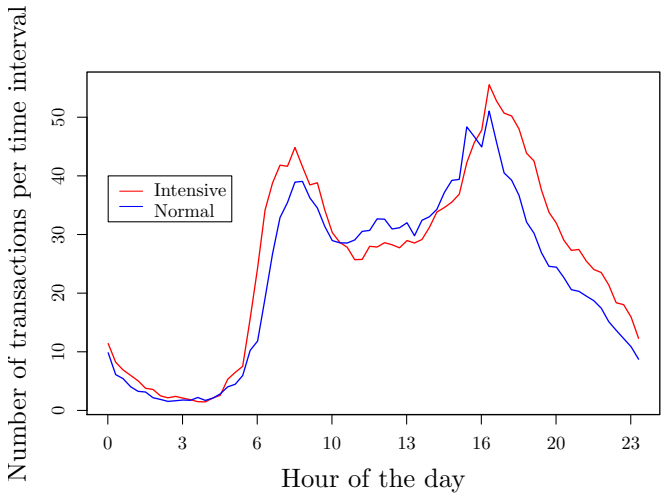


Figure 4 Number of transactions starting in both zones during the day considering intervals of 20 mn.

Figure 5 plots the performance in both cases: time-inhomogeneous arrival rates and so-called real trip time distribution versus time-homogeneous arrival rate and trip time with exponential distribution. Performance is different for both clusters for the whole range of fleet sizes. It can be explained as follows. For large fleet sizes, more arrivals during the day and some longer trips contribute to prevent saturation in stations of both clusters in the time-inhomogeneous case. For small fleet sizes, more arrivals and larger trip times empty the stations, especially in cluster 1. It explains why the time-inhomogeneous case performs better for a mean number of cars per station around 30, but worse when it is smaller than this value.

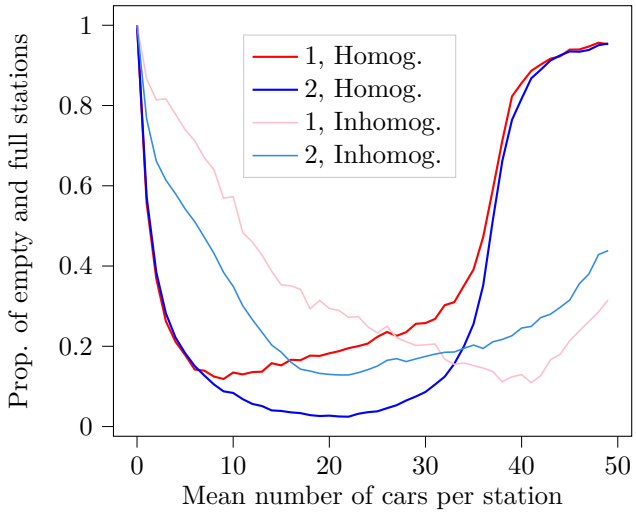


Figure 5 Performance from a time-inhomogeneous and time-homogeneous simulation as a function of fleet size. The thin curve is plotted for time-inhomogeneous arrival rates and real trip time distribution of [7]. It is compared to the thick curve plotted for time-homogeneous arrival rates and trip times with exponential distribution of Figure 2b.

In conclusion, although the homogeneous model helps us to study the influence of parameters such as δ and q, it approximates poorly time-inhomogeneous arrival rates combined with heavy-tailed trip time distribution.

4.3 Future work

The analysis highlights an interesting random walk in the quarter-plane. Its study seems necessary to obtain further analytical results. Another model seems also necessary to analytically obtain the proportion of gifts in the system, to see the price that the operator should pay to implement such a policy.

References

1 A. Economou and D. Fakinos. Product form stationary distributions for queueing networks with blocking and rerouting. *Queueing Systems Theory Appl.*, 30:251–260, 1998.

2 Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

3 G. Fayolle, R. Iasnogorodski, and V. Malyshev. *Random Walks in the Quarter Plane: Algebraic Methods, Boundary Value Problems, Applications to Queueing Systems and Analytic Combinatorics*. Springer, 2003.

4 G. Fayolle and J.-M. Lasgouttes. Asymptotics and scalings for large product-form networks via the central limit theorem. *Markov Process. Related Fields*, 2:317–348, 1996.

5 C. Fricker and N. Gast. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. *Euro journal on transportation and logistics*, 5(3):261–291, 2016.

6 C. Fricker, N. Gast, and H. Mohamed. Mean field analysis for inhomogeneous bike sharing systems. *Discrete Mathematics & Theoretical Computer Science*, DMTCS Proceedings vol. AQ, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12), 2012.

7 C. Fricker, H. Mohamed, T. Popescu, and M. Trépanier. Stochastic modelling of free-floating car-sharing systems. *CIRRELT*, 2021.

8 C. Fricker and D. Tibi. Equivalence of ensembles for large vehicle-sharing models. *The Annals of Applied Probability*, 27(2):883–916, 2017.

9 N. Gast and B. Gaujal. A mean field model of work stealing in large-scale systems. *ACM SIGMETRICS Performance Evaluation Review*, 38(1):13–24, 2010.

10 Nicolas Gast. Expected values estimated via mean-field approximation are 1/n-accurate. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 1(1):1–26, 2017.

11 Nicolas Gast and Benny Van Houdt. A refined mean field approximation. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 1(2):1–28, 2017.

12 D. K. George and C. H. Xia. Asymptotic analysis of closed queueing networks and its implications to achievable service levels. *SIGMETRICS Performance Evaluation Review*, 38:3–5, 2010.

13 W. J. Gordon and G. F. Newell. Closed queueing systems with exponential servers. *Operations Research*, 15(2):254–265, 1967.

14 James R. Jackson. Jobshop-like queueing systems. *Management Science*, 10(1):131–142, 1963.

15 F.P. Kelly. Loss networks. *Annals of Applied Probability*, 1(3):319–378, 1991.

16 Arpan Mukhopadhyay, Ravi R Mazumdar, and Rahul Roy. Binary opinion dynamics with biased agents and agents with different degrees of stubbornness. In *2016 28th International Teletraffic Congress (ITC 28)*, volume 1, pages 261–269. IEEE, 2016.

17 Philippe Robert. *Stochastic Networks and Queues*. Applications of Mathematics: Stochastic Modelling and Applied Probability. Springer, 2003.

A Transitions of the empirical measure process

In this section, we present the detailed transitions of the empirical measure process $(Y^N(t))$ introduced in Section 1.4. The transitions, from state $y = (y_{1,j}, y_{2,k,l}, z) \in S^N$, are given by

- **User arrival.**
  - A user arrival at a station in cluster 2 with k cars and l gifts (for short of type $(2,k,l)$) taking a gift.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k,l-1} - e_{2,k,l}) + \frac{e_0}{N} \quad \text{at rate} \quad \lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{l>0\}} \big(p + (1-p)\mathbf{1}_{\{k=0\}}\big).$$

  - A user arrival at a station of type $(2,k,l)$ taking a normal car.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k-1,l} - e_{2,k,l}) \quad \text{at rate} \quad \lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{k>0\}} \big(1 - p + p\mathbf{1}_{\{l=0\}}\big).$$

  - A user arrival at a station of type $(1,j)$.

$$y \longrightarrow y + \frac{1}{N_1}(e_{1,j-1} - e_{1,j}) \quad \text{at rate} \quad \lambda_1 y_{1,j} N_1 \mathbf{1}_{\{j>0\}}.$$

- **Gift appearance.**
  - A car becoming a gift at a station of type $(2,k,l)$.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k-1,l+1} - e_{2,k,l}) \quad \text{at rate} \quad \delta k N_2 y_{2,k,l}.$$

- **Car return.**
  - A normal car returned at a station of type $(1,j)$.

$$y \longrightarrow y + \frac{1}{N_1}(e_{1,j+1} - e_{1,j}) \quad \text{at rate} \quad c_1 y_{1,j} \mu \Big(M - (E_1 N_1 + E_2 N_2 + zN)\Big) \mathbf{1}_{\{j<K_1\}}.$$

  - A normal car returned at a station of type $(2,k,l)$.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k+1,l} - e_{2,k,l}) \quad \text{at rate} \quad c_2 y_{2,k,l} \mu \Big(M - (E_1 N_1 + E_2 N_2 + zN)\Big) \mathbf{1}_{\{k+l<K_2\}}.$$

- A gift returned at a station of type $(1, j)$.

$$y \longrightarrow y + \frac{1}{N_1}(e_{1,j+1} - e_{1,j}) - \frac{e_0}{N} \quad \text{at rate} \quad q_1 y_{1,j} \mu_c z N \mathbf{1}_{\{j<K_1\}}.$$

- A gift returned at a station of type $(2, k, l)$.

$$y \longrightarrow y + \frac{1}{N_2}(e_{2,k+1,l} - e_{2,k,l}) - \frac{e_0}{N} \quad \text{at rate} \quad q_2 y_{2,k,l} \mu_c z N \mathbf{1}_{\{k+l<K_2\}}.$$

On the Contraction Method with Reduced Independence Assumptions

Ralph Neininger
Institute for Mathematics, Goethe University, 60054 Frankfurt a.M., Germany

Jasmin Straub
Institute for Mathematics, Goethe University, 60054 Frankfurt a.M., Germany

Abstract

Recursive sequences of laws of random variables (and random vectors) are considered where an independence assumption which is usually made within the setting of the contraction method is dropped. This restricts the study to sequences which after normalization lead to asymptotic normality. We provide a general univariate central limit theorem which can directly be applied to problems from the analysis of algorithms and random recursive structures without further knowledge of the contraction method. Also multivariate central limit theorems are shown and bounds on rates of convergence are provided. Examples include some previously shown central limit analogues as well as new applications on Fibonacci matchings.

2012 ACM Subject Classification Theory of computation → Sorting and searching; Theory of computation → Divide and conquer

Keywords and phrases Probabilistic Analysis of Algorithms, random Trees, weak Convergence, Probability Metrics, Contraction Method

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.14

1 Introduction

Sequences $(Y_n)_{n\geq 0}$ of random vectors in $\mathbb{R}^d$, $d \in \mathbb{N}$, are considered which satisfy a distributional recursion

$$Y_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r(n) Y^{(r)}_{I^{(n)}_r} + b_n, \qquad n \geq n_0, \tag{1}$$

where $\stackrel{\mathrm{d}}{=}$ denotes equality in distribution, $n_0 \in \mathbb{N}$, the coefficients $A_1(n), \ldots, A_K(n)$ are random $(d \times d)$-matrices and b_n is a d-dimensional random vector. Such recurrences often arise in the context of divide and conquer methods. Underlying such a recursion is a problem of size n that can be divided into K smaller subproblems of sizes $I^{(n)}_1, \ldots, I^{(n)}_K$, the toll term b_n measuring the "cost" of this division and the merger. Concerning the number of subproblems and the subproblem sizes, we will always make the following assumptions:

- The number K of subproblems is a fixed integer $K \geq 1$. However, extensions to K being random and depending on n are possible.
- The vector $I^{(n)} = (I^{(n)}_1, \ldots, I^{(n)}_K)$ of the subproblem sizes is a random vector in $\{0, \ldots, n\}^K$.

Another integral part of this setting is the assumption that the subproblems are of the same nature as the original problem, or formally:

$$(Y^{(r)}_n)_{n\geq 0} \stackrel{\mathrm{d}}{=} (Y_n)_{n\geq 0} \qquad \text{for } r = 1, \ldots, K. \tag{2}$$

Since this assumption guarantees the self-similarity between the initial structure and the parts into which the structure is decomposed, we will use the term *self-similarity condition* when referring to condition (2). Furthermore, we need some *conditional independence condition*

ensuring that given the subproblem sizes, the subproblems behave independently. To be more precise, within the contraction method, which is sketched below, usually it is assumed that

$$(A_1(n),\dots,A_K(n),b_n,I^{(n)}),(Y_n^{(1)})_{n\geq 0},\dots,(Y_n^{(K)})_{n\geq 0} \quad \text{are independent.} \tag{3}$$

Note, however, that dependencies between the coefficients $A_1(n),\dots,A_K(n),b_n$ and the subproblem sizes $I_1^{(n)},\dots,I_K^{(n)}$ are allowed.

Recurrences of the form (1) come up in various fields, see [11, 7] for many concrete examples ranging from complexity measures of recursive algorithms (e.g., the number of key comparisons used by Quicksort, Mergesort or Quickselect) to parameters of random trees (e.g., the size of tries and m-ary search trees, path lengths in digital search trees, (PATRICIA) tries and m-ary search trees or the number of leaves in quadtrees) to quantities of stochastic geometry (e.g., the number of maxima in right triangles). For all these examples, the contraction method can be used to derive limit laws, i.e., convergence in distribution of the normalized versions of the Y_n towards a limit distribution being either the normal distribution or some other distribution.

In recent years a couple of problems appeared which seemed to fall within the framework above, however, the conditional independence condition (3) was violated. Examples are central limit analogues for the complexity of Quicksort, the composition of cyclic (and other) urns and the number of leaves of random point quadtrees, see [6, 5, 4, 3]. Such applications, and new applications discussed below, can be covered under a weakened independence condition that

$$(A_1(n),\dots,A_K(n),I^{(n)}),(Y_n^{(1)})_{n\geq 0},\dots,(Y_n^{(K)})_{n\geq 0} \quad \text{are independent.} \tag{4}$$

Note that, in contrast to the conditional independence condition (3), we allow dependencies between b_n and $(Y_n^{(1)})_{n\geq 0},\dots,(Y_n^{(K)})_{n\geq 0}$ here. Thus, condition (4) is slightly weaker than condition (3) and will be referred to as *partial conditional independence condition* in the following.

To observe a first implication of the partial conditional independence condition (4) on the setting of the contraction method we sketch the usual approach. We define the normalized sequence $(X_n)_{n\geq 0}$ by

$$X_n := C_n^{-1/2}(Y_n - M_n), \qquad n\geq 0, \tag{5}$$

where M_n is a d-dimensional vector and C_n a positive definite $(d\times d)$-matrix. Essentially, we choose M_n as the mean vector and C_n as the covariance matrix of Y_n if they exist (assuming $\mathrm{Cov}(Y_n)$ being positive definite for all sufficiently large n) or as the leading order terms in expansions of these moments as $n\to\infty$. The normalized quantities satisfy the following modified recursion:

$$X_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n\geq n_0, \tag{6}$$

with

$$A_r^{(n)} := C_n^{-1/2} A_r(n) C^{1/2}_{I_r^{(n)}}, \quad b^{(n)} := C_n^{-1/2}\left(b_n - M_n + \sum_{r=1}^{K} A_r(n) M_{I_r^{(n)}}\right) \tag{7}$$

and self-similarity and independence conditions as above. Then, limits

$$A_r^{(n)} \to A_r, \quad b^{(n)} \to b \tag{8}$$

are identified (in an appropriate sense) and a potential limit X of $(X_n)_{n\geq 0}$ in distribution is identified by satisfying the recursive distributional equation

$$X \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r X_r + b, \tag{9}$$

where $X_1, \dots, X_K$ are independent and identically distributed as X. Under the conditional independence condition (3) it is now justified to require on the right hand side of (9) that moreover $X_1, \dots, X_K$ and $(A_1, \dots, A_K, b)$ are independent. Hence, the distribution of the right hand side of (9) is then fully specified, see [7] for details. However, under the partial conditional independence condition (4) dependencies between b and the $X_1, \dots, X_K$ have to be taken into account, where it is not clear in general how to define the right hand side of (9). For this reason, in the current extended abstract we restrict to the case $b = 0$, and require $X_1, \dots, X_K$ and $(A_1, \dots, A_K)$ to be independent regaining a well-defined right hand side of (9). The assumption $b = 0$ essentially restricts the framework to central limit theorems, although other limit laws, such a stable laws, are still possible as fixed-points, but not covered by our techniques.

Since most applications of the contraction method in the analysis of algorithms are for univariate Y_n, i.e., dimension $d = 1$, and with $A_r^{(n)} = 1$ for all $r = 1, \dots, K$ and all $n \geq 0$ we provide a theorem convenient for applications. Note that the identification of the limits in (8) usually requires control on the expansions of moments related to the M_n and C_n. These expansions are already built into the subsequent theorem so that one does not need to have any knowledge about the underlying contraction method to apply it. For the reader's convenience we also recall the previous conditions (1), (2) and (4) for this frequently occurring univariate case: In dimension $d = 1$ consider the special case of (1) where Y_n satisfies the distributional recursion

$$Y_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} Y^{(r)}_{I_r^{(n)}} + b_n, \qquad n \geq n_0, \tag{10}$$

where $I^{(n)}, (Y_n^{(1)})_{n\geq 0}, \dots, (Y_n^{(K)})_{n\geq 0}$ are independent (this is the partial conditional independence condition), $Y_j^{(r)}$ has the same distribution as Y_j for $j \geq 0$ and $r = 1, \dots, K$, the subproblem sizes $I_r^{(n)}$ are in $\{0, 1, \dots, n\}$ and satisfy $\mathbb{P}(I_r^{(n)} = n) \to 0$ as $n \to \infty$ and all appearing quantities are L_3-integrable. Then, we have the following theorem:

▶ **Theorem 1.** *Let $(Y_n)_{n\geq 0}$ be a sequence of random variables in $\mathbb{R}$ that satisfies recursion (10). Suppose that, for some positive functions f and g and as $n \to \infty$,*

$$\mathbb{E}[Y_n] = f(n) + \mathrm{o}(g^{1/2}(n)), \quad \mathrm{Var}(Y_n) = g(n) + \mathrm{o}(g(n)). \tag{11}$$

Further assume that for all $r = 1, \dots, K$ and as $n \to \infty$, we have the L_3-convergences

$$\frac{g^{1/2}(I_r^{(n)})}{g^{1/2}(n)} \to A_r^*, \qquad \frac{1}{g^{1/2}(n)}\Big(b_n - f(n) + \sum_{r=1}^{K} f(I_r^{(n)})\Big) \to 0 \tag{12}$$

with $(A_1^*)^2 + \cdots + (A_K^*)^2 = 1$ almost surely and $\mathbb{P}(\exists\, r : A_r^* = 1) < 1$. If the technical condition

$$\frac{\mathbb{P}(I_r^{(n)} < \ell)}{g^{3/2}(n)} \to 0 \qquad (n \to \infty) \tag{13}$$

is satisfied for any $\ell \in \mathbb{N}$ and $r = 1, \dots, K$, then we have

$$\frac{Y_n - f(n)}{g^{1/2}(n)} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1) \qquad (n \to \infty).$$

▶ Remark 2. Theorem 1 generalizes the central limit theorem in [7, Corollary 5.2], where the toll function b_n and $(Y_n^{(1)})_{n\geq 0}, \dots, (Y_n^{(K)})_{n\geq 0}$ are additionally assumed to be independent.

In the current extended abstract we also consider the multivariate case of (1). We provide a multivariate central limit theorem under the partial conditional independence condition (4) and study bounds on the rates of convergence, see Theorem 3, in the Zolotarev metric ζ_3, which is the main tool underlying the proofs of the convergence theorems in this extended abstract, see Section 2. In Section 3 a sketch of the proof of Theorem 1 is given. Applications of Theorems 1 and 3 to the analysis of algorithms are given in Section 4.

2 A multivariate CLT including rates of convergence

In this section, we consider a sequence $(Y_n)_{n\geq 0}$ of d-dimensional random vectors satisfying the distributional recursion

$$Y_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r(n) Y^{(r)}_{I_r^{(n)}} + b_n, \qquad n \geq n_0, \tag{14}$$

where $n_0 \in \mathbb{N}$, the coefficients $A_1(n), \dots, A_K(n)$ are random $(d \times d)$-matrices, b_n is a d-dimensional random vector, $I^{(n)} = (I_1^{(n)}, \dots, I_K^{(n)})$ is a random vector in $\{0, \dots, n\}^K$ and all appearing quantities have finite third absolute moments. Furthermore, we assume that the self-similarity condition (2) and the partial conditional independence condition (4) are satisfied.

We assume that there exists some $n_1 \in \mathbb{N}_0$ such that the covariance matrix of Y_n is positive definite for $n \geq n_1$ and define the normalized sequence $(X_n)_{n\geq 0}$ by

$$X_n := C_n^{-1/2}(Y_n - M_n), \qquad n \geq 0, \tag{15}$$

where M_n is chosen as the mean vector of Y_n and C_n as the covariance matrix of Y_n for $n \geq n_1$ (and $C_n = \mathrm{Id}_d$ for $n < n_1$, where Id_d denotes the $d \times d$ identity matrix). The normalized quantities satisfy the following modified recursion:

$$X_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n \geq n_0, \tag{16}$$

with $A_r^{(n)}$ and $b^{(n)}$ given in (7) and self-similarity and independence relations as in (14).

To obtain a convergence result $X_n \to \mathcal{N}(0, \mathrm{Id}_d)$, we assume that the coefficients appearing in (7) converge appropriately, i.e., that there exist L_3-integrable random variables $A_1^*, \dots, A_K^*$ such that, as $n \to \infty$,

$$(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}) \xrightarrow{L_3} (A_1^*, \dots, A_K^*, 0), \tag{17}$$

with $A_1^*(A_1^*)^{\mathrm{T}} + \cdots + A_K^*(A_K^*)^{\mathrm{T}} = \mathrm{Id}_d$ almost surely. Then, from (16), we expect a limit X of X_n to satisfy the distributional fixed-point equation

$$X \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^* X^{(r)}, \tag{18}$$

where $(A_1^*, \dots, A_K^*), X^{(1)}, \dots, X^{(K)}$ are independent and $X^{(r)} \stackrel{\mathrm{d}}{=} X$ for $r = 1, \dots, K$. Under the additional assumption $\sum_{r=1}^{K} \mathbb{E}[\|A_r^*\|_{\mathrm{op}}^3] < 1$, the multivariate standard normal distribution $\mathcal{N}(0, \mathrm{Id}_d)$ is the unique solution of equation (18) in the space $\mathcal{P}_3^d(0, \mathrm{Id}_d)$ of L_3-integrable probability distributions with mean vector 0 and covariance matrix Id_d, see, e.g., [7].

As a tool to derive convergence in distribution of (X_n) in (15) towards $\mathcal{N}(0,\mathrm{Id}_d)$ we use the Zolotarev metric ζ_3, which we only need and only define on the space $\mathcal{P}_3^d(0,\mathrm{Id}_d)$. For $\mathcal{L}(X),\mathcal{L}(Y)\in\mathcal{P}_3^d(0,\mathrm{Id}_d)$ we set

$$\zeta_3(X,Y):=\zeta_3(\mathcal{L}(X),\mathcal{L}(Y)):=\sup_{f\in\mathcal{F}_3}|\mathbb{E}[f(X)-f(Y)]| \tag{19}$$

where

$$\mathcal{F}_3:=\{f\in C^2(\mathbb{R}^d,\mathbb{R}):\|f^{(2)}(x)-f^{(2)}(y)\|\le\|x-y\|\}, \tag{20}$$

with $C^2(\mathbb{R}^d,\mathbb{R})$ denoting the space of twice continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}$ and $f^{(2)}$ denoting the second derivative of $f\in C^2(\mathbb{R}^d,\mathbb{R})$.

In order to obtain such convergence, including a bound on the rate of convergence in ζ_3 we need a quantified version of the convergences in (17). In particular, we assume that

$$\Big\|\sum_{r=1}^K\mathbb{E}[A_r^{(n)}(A_r^{(n)})^{\mathrm{T}}]-\mathrm{Id}_d\Big\|_{\mathrm{op}}+\Big\|\sum_{r=1}^K A_r^{(n)}(A_r^{(n)})^{\mathrm{T}}-\mathrm{Id}_d\Big\|_{3/2}^{3/2}+\big\|b^{(n)}\big\|_3=\mathrm{O}(R(n)) \tag{21}$$

for some monotonically decreasing sequence $R(n)>0$ with $R(n)\to 0$. Furthermore, we assume the technical conditions

$$\big\|\mathbf{1}_{\{I_r^{(n)}<\ell\}}A_r^{(n)}\big\|_2^2+\big\|\mathbf{1}_{\{I_r^{(n)}<\ell\}}A_r^{(n)}\big\|_3^3=\mathrm{O}(R(n)),\quad n\to\infty, \tag{22}$$

for all $\ell\in\mathbb{N}$ and $r=1,\dots,K$ and

$$\big\|\mathbf{1}_{\{I_r^{(n)}=n\}}A_r^{(n)}\big\|_3\to 0,\quad n\to\infty, \tag{23}$$

for all $r=1,\dots,K$. Under these assumptions, with the partial conditional independence condition (4) we have:

▶ **Theorem 3.** *Let $(X_n)_{n\ge0}$ be given as in (15) with $(Y_n)_{n\ge0}$ satisfying the distributional recurrence (14) with the self-similarity condition (2) and the partial conditional independence condition (4). Furthermore, assume that the coefficients $(A_1^{(n)},\dots,A_K^{(n)},b^{(n)})$ defined in (7) converge to $(A_1^*,\dots,A_K^*,0)$ in the L_3 norm as $n\to\infty$ with $\sum A_r^*(A_r^*)^{\mathrm{T}}=\mathrm{Id}_d$ almost surely. If conditions (21), (22) and (23) are satisfied and if*

$$\limsup_{n\to\infty}\sum_{r=1}^K\mathbb{E}\Big[\frac{R(I_r^{(n)})}{R(n)}\big\|A_r^{(n)}\big\|_{\mathrm{op}}^3\Big]<1, \tag{24}$$

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

$$\zeta_3(X_n,\mathcal{N}(0,\mathrm{Id}_d))=\mathrm{O}(R(n)).$$

▶ Remark 4. Condition (21) in Theorem 3 should be compared to a similar result under the stronger conditional independence condition (3) which is stated in Theorem 3 of [8], see the corresponding condition (12) there where only $\|b^{(n)}\|_3^3=\mathrm{O}(R(n))$ is required. In (21) in the present Theorem 3 no additional factor 3 in the exponent appears which is caused by the additional dependencies. However, it is not clear whether Theorem 3 is tight in this respect or if this factor 3 may be regained by some other argument.

▶ Remark 5. In some applications, we are only interested in showing (weak) convergence rather than estimating the rate of convergence. In this case, the formulation of Theorem 3 is more complex than necessary. More specifically, if no rates are needed, we can replace condition (24) by the assumption that

$$\sum_{r=1}^{K} \mathbb{E}[\|A_r^*\|_{\mathrm{op}}^3] < 1.$$

Furthermore, condition (21) can be dropped and instead of condition (22), it is enough to assume that

$$\big\|\mathbf{1}_{\{I_r^{(n)}<\ell\}} A_r^{(n)}\big\|_3 \to 0$$

as $n \to \infty$ for any $\ell \in \mathbb{N}$ and $r = 1, \dots, K$. With these modified conditions, similar arguments as in the proof of Theorem 3 can be used to show that $\zeta_3(X_n, \mathcal{N}(0, \mathrm{Id}_d))$ converges to zero as $n \to \infty$ and hence, X_n converges in distribution to $\mathcal{N}(0, \mathrm{Id}_d)$.

A proof of Theorem 3 is given in the full paper version of this extended abstract.

3 Sketch of the proof of Theorem 1

We assume that all quantities are as in the formulation of Theorem 1.

Sketch of the proof of Theorem 1. Since we have $\mathrm{Var}(Y_n) = g(n) + \mathrm{o}(g(n))$ for some positive function g, we can find some constant $n_1 \in \mathbb{N}_0$ such that $\mathrm{Var}(Y_n)$ is positive for $n \geq n_1$. As before, we define the standardized quantities by

$$X_n := \frac{Y_n - \mu(n)}{\sigma(n)}, \qquad n \geq 0,$$

where $\mu(n) := \mathbb{E}[Y_n]$, $\sigma^2(n) := \mathrm{Var}(Y_n)$ for $n \geq n_1$ and $\sigma(n) = 1$ for $n < n_1$. The statement of the theorem follows directly from the asymptotic expansions in (11) and Slutsky's theorem if we show that the normalized quantities X_n converge in distribution to the standard normal distribution. Thus, it is sufficient to show that the Zolotarev distance $\Delta(n) := \zeta_3(X_n, \mathcal{N}(0,1))$ converges to zero as $n \to \infty$. The sequence $(X_n)_{n\geq 0}$ satisfies the modified recursion

$$X_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n \geq n_0, \tag{25}$$

with $I^{(n)}, (X_n^{(1)})_{n\geq 0}, \dots, (X_n^{(K)})_{n\geq 0}$ independent, $X_j^{(r)}$ identically distributed as X_j for $j \geq 0$ and $r = 1, \dots, K$ and

$$A_r^{(n)} = \frac{\sigma(I_r^{(n)})}{\sigma(n)}, \qquad b^{(n)} = \frac{1}{\sigma(n)}\Big(b_n - \mu(n) + \sum_{r=1}^{K} \mu(I_r^{(n)})\Big).$$

By conditions (11), (12) and (13), we have $A_r^{(n)} \to A_r^*$ and $b^{(n)} \to 0$ in L_3 for $r = 1, \dots, K$. We define

$$Z_n := \sum_{r=1}^{K} A_r^{(n)} \tau_{I_r^{(n)}} N^{(r)},$$

where $I^{(n)}, N^{(1)}, \dots, N^{(K)}$ are independent, the deterministic non-negative sequence $(\tau_n)_{n\geq 0}$ is defined by $\tau_n^2 = \mathrm{Var}(X_n)$ for $n \geq 0$ and $N^{(r)}$ has the standard normal distribution for $r = 1, \dots, K$. Then, Z_n is centered and has variance

$$\mathrm{Var}(Z_n) = \sum_{r=1}^{K} \mathbb{E}\big[(A_r^{(n)})^2 \tau^2_{I_r^{(n)}}\big] = \sum_{r=1}^{K} \Big(\mathbb{E}\big[(A_r^{(n)})^2\big] + \mathbb{E}\big[\mathbf{1}_{\{I_r^{(n)}<n_1\}}(A_r^{(n)})^2\big(\tau^2_{I_r^{(n)}} - 1\big)\big]\Big).$$

We now observe that for any $r = 1, \dots, K$, the latter summand in the above sum converges to zero by condition (13), since this condition and Jensen's inequality imply that $\mathbb{E}\big[\mathbf{1}_{\{I_r^{(n)}<n_1\}}(A_r^{(n)})^2\big] \to 0$. Together with the fact that $A_r^{(n)}$ converges in the L_3 norm (and thus also in L_2) to A_r^* with $(A_1^*)^2 + \dots + (A_K^*)^2 = 1$ almost surely, we obtain that $\mathrm{Var}(Z_n)$ converges to 1. Hence, we can choose a deterministic sequence $(\kappa_n)_{n\geq 0}$ with $\kappa_n \to 0$ such that

$$Z_n^* := (1 + \kappa_n) Z_n$$

has variance 1 for $n \geq n_1$ (where n_1 may need to be enlarged). As a consequence, the distributions of X_n, Z_n^* and $\mathcal{N}(0,1)$ are all in $\mathcal{P}_3^1(0, \mathrm{Id}_1)$ for $n \geq n_1$ and we can apply the triangle inequality to obtain

$$\Delta(n) = \zeta_3(X_n, \mathcal{N}(0,1)) \leq \zeta_3(X_n, Z_n^*) + \zeta_3(Z_n^*, \mathcal{N}(0,1)), \qquad n \geq n_1.$$

With $Q_n := A_1^{(n)} X^{(1)}_{I_1^{(n)}} + \dots + A_K^{(n)} X^{(K)}_{I_K^{(n)}}$ and Lemma 3.4 in [5] we find

$$\begin{aligned}\zeta_3(X_n, Z_n^*) &\leq \zeta_3(Q_n, Z_n) + \|Q_n\|_3^2 \|b^{(n)}\|_3 + \frac{1}{2}\|Q_n\|_3 \|b^{(n)}\|_3^2 + \frac{1}{2}\|b^{(n)}\|_3^3 \\ &\quad + \Big(|\kappa_n| + \frac{1}{2}|\kappa_n|^2 + \frac{1}{2}|\kappa_n|^3\Big)\|Z_n\|_3^3 \\ &= \zeta_3(Q_n, Z_n) + \mathrm{o}(1),\end{aligned}$$

since $b^{(n)}$ converges to zero in the L_3 norm, κ_n converges to zero and $\|Z_n\|_3$ and $\|Q_n\|_3$ are bounded in n, the latter by Lemma 6 (note that we have $\sum \mathbb{E}[(A_r^*)^3] < 1$ by the assumptions $\sum (A_r^*)^2 = 1$ almost surely and $\mathbb{P}(\exists\, r : A_r^* = 1) < 1$ and that the technical condition (23) is satisfied since we assumed $\mathbb{P}(I_r^{(n)} = n) \to 0$ for $r = 1, \dots, K$, thus Lemma 6 is applicable). Conditioning on $I^{(n)}$ implies that, for $n \geq n_1$,

$$\begin{aligned}\zeta_3(Q_n, Z_n) &\leq \left(\sum_{r=1}^{K} \mathbb{P}(I_r^{(n)} = n)\right)\Delta(n) + \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\{n_1 \leq I_r^{(n)} < n\}} (A_r^{(n)})^3 \Delta(I_r^{(n)})\right] \\ &\quad + \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\{I_r^{(n)}<n_1\}} (A_r^{(n)})^3 \sup_{k<n_1} \zeta_3(X_k, \tau_k N^{(r)})\right] \\ &= \mathrm{o}(1)\Delta(n) + \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\{n_1 \leq I_r^{(n)} < n\}} (A_r^{(n)})^3 \Delta(I_r^{(n)})\right] + \mathrm{o}(1),\end{aligned}$$

where we used the assumption $\mathbb{P}(I_r^{(n)} = n) \to 0$ for $r = 1, \dots, K$ and the technical condition (13) in the last step. Furthermore, we have $\zeta_3(Z_n^*, \mathcal{N}(0,1)) \to 0$. This can be seen by showing that $\ell_3(Z_n^*, \mathcal{N}(0,1)) \to 0$ and that $\|Z_n^*\|_3$ is bounded in n, where ℓ_3 denotes the minimal L_3-metric. Collecting all estimates, we obtain that

$$\Delta(n) \leq \mathrm{o}(1)\Delta(n) + \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\{n_1 \leq I_r^{(n)} < n\}} (A_r^{(n)})^3 \Delta(I_r^{(n)})\right] + \mathrm{o}(1). \tag{26}$$

From this, the statement follows by a standard argument (see, e.g., [7, pp. 390–391]). ◀

▶ **Lemma 6.** *Let* $(X_n)_{n\geq 0}$ *be given as in* (15) *with* $(Y_n)_{n\geq 0}$ *satisfying the distributional recurrence* (14)*. Furthermore, assume that the coefficients* $(A_1^{(n)},\dots,A_K^{(n)},b^{(n)})$ *defined in* (7) *converge to* $(A_1^*,\dots,A_K^*,0)$ *in the* L_3 *norm as* $n\to\infty$ *with* $\sum A_r^*(A_r^*)^{\mathrm{T}} = \mathrm{Id}_d$ *almost surely and* $\sum \mathbb{E}[\|A_r^*\|_{\mathrm{op}}^3] < 1$ *and that the technical condition* (23) *is satisfied. Then we have, as* $n\to\infty$, $\|X_n\|_3 = \mathrm{O}(1)$.

The proof of Lemma 6 follows along the same lines as the proof of Lemma 2.3 in [6], generalizing to a more general setting with multivariate quantities and an arbitrary number K of subproblems here. However, in the sketch of the proof of Theorem 1 only the case $d=1$ is needed. Further details can also be found in Straub [12].

4 Applications

As mentioned in the introduction, possible examples of distributional recurrences with dependent toll function, where our results apply can be found in [6, 5, 4, 3]. Since the central limit analogue for the complexity of Quicksort in [6] only contains a convergence result for the Zolotarev metric without a rate of convergence, we take up this example in section 4.1 and use Theorem 3 to rederive this central limit analogue and add a bound on the rate of convergence. Furthermore, in section 4.2, we present another application of Theorems 3 and 1 concerning recent work of Diaconis and Kolesnik [2] on Fibonacci permutations.

4.1 Refined Quicksort asymptotics

We consider the Quicksort algorithm where we set $K_0 = 0$ and denote by K_n, $n\geq 1$, the number of key comparisons needed by Quicksort to sort the list $(U_1,\dots,U_n)$, where $(U_i)_{i\geq 1}$ is a sequence of independent and uniformly on the unit interval distributed random variables. With the normalization

$$C_n := \frac{K_n - \mathbb{E}[K_n]}{n+1}, \qquad n\geq 0,$$

Régnier [9] showed for a suitable version that the sequence $(C_n)_{n\geq 0}$ is a martingale converging almost surely (and in L_p) to some non-degenerate limit C. Rösler [10] showed that C satisfies the distributional fixed-point equation

$$C \overset{\mathrm{d}}{=} UC^{(1)} + (1-U)C^{(2)} + \varphi(U)$$

with U, $C^{(1)}$, $C^{(2)}$ independent, U uniform on the unit interval, $C^{(1)}$ and $C^{(2)}$ having the same distribution as C and $\varphi(u) := 2u\log u + 2(1-u)\log(1-u) + 1$ for $u\in[0,1]$.

The aim of this section is to further quantify the almost sure convergence $C_n\to C$ by analyzing the residual $C_n - C$. Bindjeme and Fill [1, Theorem 1.4] found the L_2-norm of this residual explicitly which implies that

$$\|C_n - C\|_2^2 = 2\,\frac{\log n}{n} + \mathrm{O}\Big(\frac{1}{n}\Big) \tag{27}$$

and in [6, Theorem 1.1] it is shown that

$$\sqrt{\frac{n}{2\log n}}\,(C_n - C) \xrightarrow{\mathrm{d}} \mathcal{N}(0,1)$$

as $n\to\infty$. We now show that the application of Theorem 3 provides a rate of convergence in the Zolotarev metric ζ_3 for the latter convergence without much effort. For this, we need some of the results deduced in [1] and [6]. First of all, we use the notation $Y_n := C_n - C$, $n\geq 0$,

for the residuals. Note that we chose this notation, although differing from the notation used in [1] and [6], to guarantee that the notation is in accordance with the formulation of our theorems. We then observe that the residuals Y_n can be decomposed recursively. Equation (12) in [6] states a sample-pointwise recurrence for the error term Y_n (see also equation (2.6) in [1]), from which we obtain that Y_n satisfies the distributional recursion

$$Y_n \stackrel{\mathrm{d}}{=} \frac{I_n+1}{n+1} Y^{(1)}_{I_n} + \frac{n-I_n}{n+1} Y^{(2)}_{n-1-I_n} + b_n, \qquad n \geq 1, \tag{28}$$

with I_n, $(Y^{(1)}_n)_{n\geq 0}$, $(Y^{(2)}_n)_{n\geq 0}$ independent, I_n uniformly distributed on $\{0,\dots,n-1\}$, $Y^{(r)}_j$ distributed as Y_j for $j \geq 0$ and $r = 1,2$ and some toll function b_n which is not independent of $(Y^{(1)}_n)_{n\geq 0}$ and $(Y^{(2)}_n)_{n\geq 0}$. Since the concrete representation of b_n is not needed in the following, we omit the details here and refer to [1] and [6]. Certainly, recurrence (28) is an instance of recursion (14) with $n_0 = 1$, $K = 2$, $I^{(n)}_1 = I_n$, $I^{(n)}_2 = n-1-I_n$ and $A_r(n) = (I^{(n)}_r + 1)/(n+1)$ for $r = 1,2$.

We denote the normalized residuals by

$$X_n := \frac{Y_n}{\sigma(n)}, \quad n \geq 0,$$

where $\sigma^2(n) := \mathrm{Var}(Y_n) > 0$ for $n \geq 0$. Note that $\sigma^2(n) = 2\log n/n + \mathrm{O}(1/n)$ by (27) and the fact that both components C_n and C of Y_n are centered. The normalized quantities satisfy recursion (16) with the same parameters as above and

$$A^{(n)}_1 = \frac{(I_n+1)\,\sigma(I_n)}{(n+1)\,\sigma(n)}, \quad A^{(n)}_2 = \frac{(n-I_n)\,\sigma(n-1-I_n)}{(n+1)\,\sigma(n)}, \quad b^{(n)} = \frac{b_n}{\sigma(n)}.$$

For these coefficients, we obtain the L_3-convergences (see [6])

$$A^{(n)}_1 \to \sqrt{U_1} =: A^*_1, \quad A^{(n)}_2 \to \sqrt{1-U_1} =: A^*_2, \quad b^{(n)} \to 0.$$

Thus, we are in the situation of Section 2 and now check the conditions of Theorem 3 with $R(n) = \log^{-1/2}(n)$. First of all, Lemma 2.2 in [6] states that, as $n \to \infty$,

$$\|b^{(n)}\|_3 = \mathrm{O}\Big(\frac{1}{\sqrt{\log n}}\Big) = \mathrm{O}(R(n)).$$

The order of $\|(A^{(n)}_1)^2 + (A^{(n)}_2)^2 - 1\|_{3/2}$ can be bound, using [8, Lemma 7], by

$$\begin{aligned}
&\Big\| \sum_{r=1}^{2} (A^{(n)}_r)^2 - 1 \Big\|_{3/2} \\
&\leq \frac{1}{n\log n} \Big\| I_n \log\Big(\frac{I_n}{n}\Big) + (n-1-I_n)\log\Big(\frac{n-1-I_n}{n}\Big) \Big\|_{3/2} + \mathrm{O}\Big(\frac{1}{\log n}\Big) \\
&= \mathrm{O}\Big(\frac{1}{\log n}\Big).
\end{aligned}$$

From this, we also obtain

$$\Big| \sum_{r=1}^{2} \mathbb{E}[(A^{(n)}_r)^2] - 1 \Big| \leq \Big\| \sum_{r=1}^{2} (A^{(n)}_r)^2 - 1 \Big\|_1 \leq \Big\| \sum_{r=1}^{2} (A^{(n)}_r)^2 - 1 \Big\|_{3/2} = \mathrm{O}\Big(\frac{1}{\log n}\Big).$$

Thus, condition (21) is satisfied. Since $A_1^{(n)}$ and $A_2^{(n)}$ are uniformly bounded and I_n is uniform on $\{0,\dots,n-1\}$, the technical conditions (22) and (23) are clearly satisfied. Furthermore, we can use arguments of [8, Lemma 7], the fact that I_n/n converges almost surely to U_1 and the dominated convergence theorem to show that

$$\limsup_{n\to\infty} \sum_{r=1}^{2} \mathbb{E}\left[\frac{R(I_r^{(n)})}{R(n)}\big(A_r^{(n)}\big)^3\right] = 2\,\mathbb{E}[U_1^{3/2}] = \frac{4}{5} < 1,$$

such that all assumptions of Theorem 3 are satisfied and we obtain the following result.

▶ **Theorem 7.** *For the number K_n of key comparisons used by Quicksort to sort the list $(U_1,\dots,U_n)$ with $(U_i)_{i\geq 1}$ independent and uniformly distributed on the unit interval and the almost sure limit C of $C_n = (K_n - \mathbb{E}[K_n])/(n+1)$, we have, as $n\to\infty$,*

$$\zeta_3\Big(\frac{C_n - C}{\sqrt{\mathrm{Var}(C_n - C)}}, \mathcal{N}(0,1)\Big) = \mathrm{O}\Big(\frac{1}{\sqrt{\log n}}\Big).$$

▶ Remark 8. In view of Remark 4 we are not sure whether the bound $\mathrm{O}(\log(n)^{-1/2})$ in Theorem 7 is tight or if $\mathrm{O}(\log(n)^{-3/2})$ may be the correct order.

4.2 Importance sampling for estimating the number of Fibonacci matchings

In this section we refer to the paper [2] by Diaconis and Kolesnik from which we adopt the notation and some of their results. The set $\mathfrak{F}_n$ of *Fibonacci matchings of size n* is defined by

$$\mathfrak{F}_n = \{\pi \in \mathcal{S}_n : |\pi(i) - i| \leq 1 \text{ for } 1 \leq i \leq n\},$$

where $\mathcal{S}_n$ denotes the set of permutations of $\{1,\dots,n\}$ (note that Diaconis and Kolesnik use the notation $\mathcal{F}_{n,1}$ instead of $\mathfrak{F}_n$). The set $\mathfrak{F}_4$ of Fibonacci matchings of size $n=4$ is displayed in Figure 1.

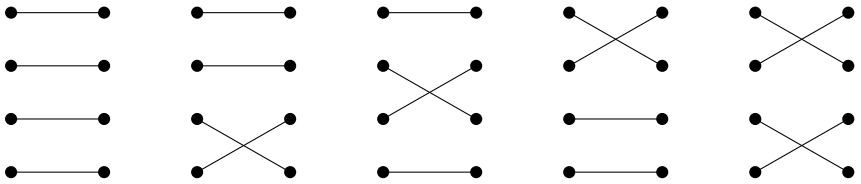


Figure 1 The 5 Fibonacci matchings of size $n = 4$.

Note that the cardinality of the set $\mathfrak{F}_n$ is easily computed by considering whether $\pi(1) = 1$ or $\pi(1) = 2$ and coincides with the $(n+1)$-th Fibonacci number (which explains the name). Although the number of Fibonacci matchings is known, Diaconis and Kolesnik [2] present different importance sampling algorithms for estimating the size of $\mathfrak{F}_n$. These algorithms in each step match the current index with an index chosen uniformly at random among the remaining allowed indices. To be more precise, Diaconis and Kolesnik present three such algorithms differing in the order the indices are matched (see [2] for details):

- The *random algorithm* A_r matches the indices in uniformly random order,
- the *fixed-order algorithm* A_f matches them in fixed order from top to bottom and
- the *greedy algorithm* A_g matches them in a certain greedy order. More precisely, A_g always matches the smallest unmatched index among those indices with the maximal number of remaining choices. This means that algorithm A_g always starts by matching

index 2 uniformly at random with one of 1, 2, 3. If $\pi(2) \in \{1,2\}$ (i.e., either $\pi(2)=1$ and consequently $\pi(1)=2$ or vice versa), then the next index to be matched is index 4 (uniformly among 3, 4, 5), since this is the smallest index with 3 remaining choices. Otherwise, i.e., if $\pi(2)=3$, then the assignments $\pi(1)=1$ and $\pi(3)=2$ are forced and the next index to be matched is index 5 (uniformly among 4, 5, 6).

We summarize some of the results given in [2]: For any Fibonacci matching $\pi \in \mathfrak{F}_n$, we denote by $P_{\mathrm{r}}(\pi)$, $P_{\mathrm{f}}(\pi)$ and $P_{\mathrm{g}}(\pi)$ the probability of π under the algorithm A_{r}, A_{f} and A_{g}, respectively. For Π_n chosen uniformly at random from the set $\mathfrak{F}_n$ of Fibonacci matchings, we have

$$\mathbb{E}[-\log(P_\chi(\Pi_n))] = \mu_\chi n + \mathrm{O}(1), \qquad \mathrm{Var}(-\log(P_\chi(\Pi_n))) = \sigma_\chi^2 n + \mathrm{O}(1), \tag{29}$$

for $\chi = \mathrm{r}, \mathrm{f}, \mathrm{g}$, where $\mu_\chi \in (0.49, 0.51)$ and $\sigma_\chi^2 > 0$ can be computed exactly (we refer to [2] for the concrete values). Furthermore, for $\chi = \mathrm{r}, \mathrm{f}, \mathrm{g}$ and as $n \to \infty$, we have the central limit theorem [2, Theorem 1.1]

$$\frac{-\log(P_\chi(\Pi_n)) - \mu_\chi n}{\sigma_\chi \sqrt{n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1). \tag{30}$$

While this statement can be obtained with [7, Corollary 5.2] for the first two algorithms (see [2, Theorems 3.2 and 3.4]), there is no obvious way of applying [7, Corollary 5.2] in the greedy case due to arising dependencies. Instead, Diaconis and Kolesnik use arguments from renewal theory to show (30) for the greedy algorithm A_{g}. Using the results of the current extended abstract, one can also handle these additional dependencies arising in the greedy case. For this reason we focus on algorithm A_{g} from now on and define the random variable

$$Y_n = -\log(P_{\mathrm{g}}(\Pi_n)),$$

where, as before, Π_n is uniformly distributed on $\mathfrak{F}_n$ and $P_{\mathrm{g}}(\pi)$ denotes the probability of π under the greedy algorithm A_{g} for any $\pi \in \mathfrak{F}_n$. Recall that algorithm A_{g} always starts by matching index 2 uniformly at random with one of 1, 2, 3. Consequently, for a fixed Fibonacci permutation π, the probability that index 2 is matched correctly with $\pi(2)$ by A_{g} equals $1/3$. Depending on the value of $\pi(2)$, the resulting number of indices that are neither matched nor forced is $n-2$ or $n-3$ afterwards. Thus, we obtain that

$$Y_n \overset{\mathrm{d}}{=} Y_{I_1^{(n)}} + \log 3,$$

where $I_1^{(n)}$ takes the values $n-2$ and $n-3$ with probabilities $2|\mathfrak{F}_{n-2}|/|\mathfrak{F}_n|$ and $|\mathfrak{F}_{n-3}|/|\mathfrak{F}_n|$, respectively, and is independent of $(Y_j)_{j\geq 0}$. However, using this recursion, our theorems do not apply, since there is only one subproblem of almost the same size as the original problem (i.e., $A_1^* = 1$).

Instead, to obtain a recursion to which our framework applies, we now divide the permutation at the middle, more precisely at index $k_n = \lfloor n/2 \rfloor$, instead of dividing it at the top. Now, consider whether $\pi(k_n) = k_n - 1$, $\pi(k_n) = k_n$ or $\pi(k_n) = k_n + 1$. In the first case, the resulting subproblem sizes are $k_n - 2$ and $n - k_n$, whereas they are $k_n - 1$ and $n - k_n$ in the second case and $k_n - 1$ and $n - k_n - 1$ in the third case. Hence, we obtain the recursive decomposition

$$Y_n \overset{\mathrm{d}}{=} Y^{(1)}_{I_1^{(n)}} + Y^{(2)}_{I_2^{(n)}} + b_n, \tag{31}$$

where the vector $I^{(n)} = (I_1^{(n)}, I_2^{(n)})$ contains the subproblem sizes and has distribution

$$\mathbb{P}(I_1^{(n)} = i_1, I_2^{(n)} = i_2) = \frac{1}{|\mathfrak{F}_n|} \begin{cases} |\mathfrak{F}_{k_n-2}| \cdot |\mathfrak{F}_{n-k_n}|, & i_1 = k_n - 2, i_2 = n - k_n, \\ |\mathfrak{F}_{k_n-1}| \cdot |\mathfrak{F}_{n-k_n}|, & i_1 = k_n - 1, i_2 = n - k_n, \\ |\mathfrak{F}_{k_n-1}| \cdot |\mathfrak{F}_{n-k_n-1}|, & i_1 = k_n - 1, i_2 = n - k_n - 1, \end{cases}$$

$I^{(n)}$, $(Y_j^{(1)})_{j\geq 0}$ and $(Y_j^{(2)})_{j\geq 0}$ are independent, $Y_j^{(r)}$ has the same distribution as Y_j for $j \geq 0$ and $r = 1, 2$ and b_n is a random variable taking values between 0 and $\log(9/2)$. However, note that b_n is not independent of $(Y_j^{(1)})_{j\geq 0}$ and $(Y_j^{(2)})_{j\geq 0}$, which is the reason why Corollary 5.2 of [7] does not apply. However, we can use Theorem 1 instead: The sequence $(Y_n)_{n\geq 0}$ satisfies recursion (10) as well as condition (11) with $f(n) = \mu_{\mathrm{g}} n$ and $g(n) = \sigma_{\mathrm{g}}^2 n$. Furthermore, conditions (12) and (13) are obviously satisfied with $A_1^* = A_2^* = 1/\sqrt{2}$. Thus, Theorem 1 implies

$$\frac{Y_n - \mu_{\mathrm{g}} n}{\sigma_{\mathrm{g}} \sqrt{n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1),$$

giving another proof of (30) for the greedy case. Note that we can also apply Theorem 3 to derive a bound on the rate of convergence in the limit theorem (30) in the Zolotarev ζ_3 metric. While in the following Theorem 9 we cover algorithm A_{g}, corresponding results can also be derived for the other two algorithms via [8, Theorem 3].

▶ **Theorem 9.** *Let $P(\pi)$ be the probability of π under the random algorithm A_{g} for any Fibonacci permutation π. Further set $Y_n = -\log(P(\Pi_n))$, where Π_n is uniformly distributed on the set $\mathfrak{F}_n$ of Fibonacci matchings of length n. Then, for any $\varepsilon > 0$, as $n \to \infty$, we have*

$$\zeta_3\Big(\frac{Y_n - \mathbb{E}[Y_n]}{\sqrt{\mathrm{Var}(Y_n)}}, \mathcal{N}(0, 1)\Big) = \mathrm{O}(n^{-1/2+\varepsilon}).$$

The proof of Theorem 9 follows easily from Theorem 3 using recurrence (31), $R(n) = n^{-1/2+\varepsilon}$ and noting that

$$\Big\| \sum_{r=1}^{2} (A_r^{(n)})^2 - 1 \Big\|_{3/2}^{3/2} + \big\| b^{(n)} \big\|_3 = \mathrm{O}(n^{-1/2}).$$

References

1. Patrick Bindjeme and James Allen Fill. Exact L^2-distance from the limit for QuickSort key comparisons (extended abstract). In *23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12)*, Discrete Math. Theor. Comput. Sci. Proc., AQ, pages 339–348. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012.
2. Persi Diaconis and Brett Kolesnik. Randomized sequential importance sampling for estimating the number of perfect matchings in bipartite graphs. *Adv. in Appl. Math.*, 131:Article 102247, 2021. doi:10.1016/j.aam.2021.102247.
3. Michael Fuchs, Noela S. Müller, and Henning Sulzbach. Refined asymptotics for the number of leaves of random point quadtrees. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 110 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 23, 16. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.
4. Noela Müller. *Central limit theorems for multicolour urn models*. PhD thesis, Goethe University Frankfurt a.M., 2017. urn:nbn:de:hebis:30:3-453587.
5. Noela Müller and Ralph Neininger. Refined asymptotics for the composition of cyclic urns. *Electron. J. Probab.*, 23:Paper no. 117, 20, 2018. doi:10.1214/18-EJP243.

6 Ralph Neininger. Refined Quicksort asymptotics. *Random Structures Algorithms*, 46(2):346–361, 2015. doi:10.1002/rsa.20497.

7 Ralph Neininger and Ludger Rüschendorf. A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.*, 14(1):378–418, 2004. doi:10.1214/aoap/1075828056.

8 Ralph Neininger and Jasmin Straub. Convergence rates in the probabilistic analysis of algorithms. In *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 159 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 22, 13. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2020.

9 Mireille Régnier. A limiting distribution for quicksort. *RAIRO Inform. Théor. Appl.*, 23(3):335–343, 1989. doi:10.1051/ita/1989230303351.

10 Uwe Rösler. A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.*, 25(1):85–100, 1991. doi:10.1051/ita/1991250100851.

11 Uwe Rösler and Ludger Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29(1-2):3–33, 2001. Average-case analysis of algorithms (Princeton, NJ, 1998). doi:10.1007/BF02679611.

12 Jasmin Straub. *On Rates of Convergence in the Probabilistic Analysis of Algorithms*. PhD thesis, Goethe University Frankfurt a.M., 2021. doi:10.21248/gups.62402.

Polyharmonic Functions in the Quarter Plane

Andreas Nessmann
Technische Universität Wien, Austria
Université de Tours, France

Abstract

In this article, a novel method to compute all discrete polyharmonic functions in the quarter plane for models with small steps, zero drift and a finite group is proposed. A similar method is then introduced for continuous polyharmonic functions, and convergence between the discrete and continuous cases is shown.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains; Mathematics of computing → Markov processes; Mathematics of computing → Generating functions; Mathematics of computing → Combinatorics

Keywords and phrases Polyharmonic functions, Functional equations, Lattice paths, Random walks, Brownian motion, Generating functions, Laplace transforms

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.15

Acknowledgements I would like to thank Kilian Raschel for introducing me to this topic as well as for a lot of valuable input and many fruitful discussions. Also, I would like to thank the anonymous reviewers for their valuable remarks.

1 Introduction and Motivation

Suppose we are given a weighted step set $\mathcal{S} \subseteq \{-1,0,1\}^2$, and we want to count the (weighted) number $q(0,x;n)$ of excursions in the quarter plane of length n from the origin to some point $x=(i,j)$. For the Simple Walk, for instance, we have $\mathcal{S}=\{\uparrow,\rightarrow,\downarrow,\leftarrow\}$, where each step has weight $\frac{1}{4}$. In this case, the number $q(0,x;n)$ can be computed explicitly (see e.g. [4]) via

$$q(0,x;n)=\frac{(i+1)(j+1)n!(n+2)!}{m!(m+i+1)!(m+j+1)!(m+i+j+2)!}, \tag{1}$$

where $m=\frac{n-i-j}{2}$. It is now fairly natural to ask about asymptotics of this expression, or more generally about asymptotics of the number $q(0,x;n)$ for an arbitrary step set $\mathcal{S}$. In particular, we consider (as proposed in [6]) asymptotic expansions of the form

$$q(0,x;n)\sim\gamma^n\sum_{p\geq 1}\frac{v_p(x)}{n^{\alpha_p}}. \tag{2}$$

In case of the Simple Walk, (1) allows us to directly compute

$$v_1=(i+1)(j+1), \tag{3}$$
$$v_2=(i+1)(j+1)(15+4i+2i^2+4j+2j^2), \tag{4}$$
$$\begin{aligned} v_3=&(i+1)(j+1)(317+16i^3+4i^4+168j+100j^2+16j^3\\ &+4j^4+8i(21+4j+2j^2)+4i^2(25+4j+2j^2)). \end{aligned} \tag{5}$$

It should be explicitly noted at this point that expansions of the form (2) are not proven to exist for this type of problem. While for the Simple Walk and a few other examples (e.g. the Diagonal Walk, Tandem Walk, see [5]) this can be shown using an explicit representation

similar as (1), in general it is not so clear (although one-term expansions of this form have been proven for many cases in [9], and more recently, using multivariate analytic techniques, in [8, Thm. 1],[15, 6.1]).
It is now fairly natural to ask about the properties of the v_p; whether they necessarily have a particular structure, if there is a clear relation to our chosen step set, and how to compute them. And indeed, at least the first two questions can be answered fairly easily by utilizing a recursive relation between the $q(0, x; n + 1)$ and $q(0, x; n)$, and showing that each function v_p must be what is called a discrete polyharmonic function of order p.
In the continuous case, a function f is called polyharmonic of degree p if it is a solution of

$$\triangle^p f = 0, \tag{6}$$

where $\triangle$ is a Laplacian operator $\triangle = \frac{1}{2}\left(\sigma_{11}\frac{\partial^2}{\partial x^2} + 2\sigma_{12}\frac{\partial^2}{\partial x \partial y} + \sigma_{22}\frac{\partial^2}{\partial y^2}\right)$. These kinds of functions have already been studied in the late 19th century, notably by E. Almansi, who proved in [1] that in a star-shaped domain containing the origin, any polyharmonic function of degree n can be written as

$$f(x) = \sum_{k=0}^{n} |x|^{2k} h_k(x), \tag{7}$$

where the h_k are harmonic (polyharmonic of degree 1). In particular harmonic and biharmonic functions have by now seen plenty of applications in physics, see e.g. [14].
The discrete setting on the other hand has gained interest comparably recently. A (discrete) function defined on a graph is called polyharmonic if it satisfies (6) as well, but with a discretised version of the Laplacian. For this discretisation, given transition probabilities $p_{x,y}$ from any point x to any point y, one lets

$$\triangle f(x) = \sum_{y} p_{x,y} f(y) - f(x). \tag{8}$$

There have been some results on polyharmonic functions on trees recently [7, 16], and polyharmonic functions on subdomains of $\mathbb{Z}^d$ have become an object of interest linked in particular to the study of discrete random walks. In our case, this subdomain will be the quarter plane and our walk homogeneous, i.e. the transition probabilities $p_s := p_{x,x+s}$, where the steps s are given by the set $\mathcal{S}$ of allowed steps, will be independent of x. The discrete Laplacian thus reads

$$\triangle f(i, j) = \sum_{(u,v)\in\mathcal{S}} p_{u,v} f(i + u, j + v) - f(i, j). \tag{9}$$

It is not at all obvious, however, how polyharmonic functions in general can be found. In [17], a way to construct harmonic functions for zero-drift models with small steps via a boundary value problem is given. This is utilized in [12] to give a complete discription of harmonic functions for symmetric step sets with small negative steps; the methods used therein can be applied to the case considered here with only minor adjustments. Very recently in [6], the authors propose a way to extend this method from harmonic to polyharmonic functions, provided one can compute a so-called "decoupling function", which was first introduced by W. T. Tutte in [19], and is discussed further in [2]. There, this concept is utilized to give remarkably succinct proofs of the algebraicity (or D-algebraicity) of the counting function of some models in the quarter plane.

Generally, instead of working directly with a polyharmonic function $h(u,v)$, one prefers to consider its generating function $H(x,y) := \sum_{i,j} x^{i+1}y^{j+1}h(i,j)$. The main reason to do so is the functional equation

$$K(x,y)H(x,y) = K(x,0)H(x,0) + K(0,y)H(0,y) - K(0,0)H(0,0) - xy\,[\triangle H]\,(x,y), \tag{10}$$

which can be shown by straightforward computation to be satisfied by this generating function (note that we have $\triangle H = 0$ for harmonic H). Here, $K(x,y)$, which will be defined in Section 2.1, is the same kernel that usually appears in the study of random walks, and even the resulting functional equations for counting walks or the stationary distribution look strikingly similar, see e.g. [10, 5].

This article aims to generalize and complete the notions introduced in [6], and thus give a description of all discrete polyharmonic functions for walks with small steps, zero drift and finite group. The main tool to do so will be the explicit computation of decoupling functions. The structure will be roughly as follows:

- In Section 2, an algorithm to construct discrete polyharmonic functions is presented (Thm. 7). In addition, it is shown that all possible discrete polyharmonic functions can be constructed in this manner (Thm. 8).
- In Section 3.1, the same construction is done for the Laplace transforms of continuous polyharmonic functions (Thm. 11), again using a functional equation approach presented for biharmonic functions in [6].
- In Section 4 the relation between the discrete and continuous cases is briefly discussed; it is shown that discrete polyharmonic functions converge towards continuous ones in the sense of generating functions/Laplace transforms (Thm. 14).
- Lastly, Section 5 gives a brief overview of some open questions and ongoing research.

2 The Discrete Case

2.1 Preliminaries

The following only serves as a very brief overview; for a more thorough introduction see e.g. [17, 10]. Consider a homogeneous random walk in $\mathbb{Z}\times\mathbb{Z}$ with a step set $\mathcal{S}$ and transition probabilities $p_{i,j}$. From now on, we will make the following assumptions:

- **(i)** The walk consists of small steps only, i.e. $\mathcal{S} \subseteq \{-1,0,1\}^2$.
- **(ii)** The walk is non-degenerate, that is, the list $p_{1,1}, p_{1,0}, p_{1,-1}, p_{0,-1}, p_{-1,-1}, p_{-1,0}, p_{-1,1}, p_{0,1}$ does not contain three consecutive 0s.
- **(iii)** The walk has zero drift, meaning that $\sum_{(i,j)\in\mathcal{S}} ip_{i,j} = \sum_{(i,j)\in\mathcal{S}} jp_{i,j} = 0$.
- **(iv)** Any polyharmonic function considered is supposed to have Dirichlet boundary conditions, i.e. it is 0 outside of the quarter plane. This is due to the probabilistic interpretation of (2); there can be no paths which start outside but always are inside the quarter plane.

A standard object appearing in a variety of functional equations around random walks (besides those below for example when one wants to compute a stationary distribution, or for counting walks, see e.g. [10]) is the *kernel* of the walk, which is given by

$$K(x,y) = xy\left(\sum_{(i,j)\in\mathcal{S}} p_{i,j}x^{-i}y^{-j} - 1\right). \tag{11}$$

In [10], this kernel is examined quite thoroughly, and we will in the following state a few of their results.

As we consider non-degenerate walks with small steps, our kernel will necessarily be quadratic in both x and y. Letting

$$K(x,y) = a(x)y^2 + b(x)y + c(x) = \tilde{a}(y)x^2 + \tilde{b}(y)x + \tilde{c}(y), \tag{12}$$

we can use the quadratic formula to find solutions of $K(\cdot, y) = 0$, which are given by

$$X_\pm(y) = \frac{-\tilde{b}(y) \pm \sqrt{\tilde{b}(y)^2 - 4\tilde{a}(y)\tilde{c}(y)}}{2\tilde{a}(y)}. \tag{13}$$

One can define $Y_\pm$ in the same fashion by swapping x, y. Letting $D := \tilde{b}(y)^2 - 4\tilde{a}(y)\tilde{c}(y)$, then one can show [17, 2.5], [10, 2.3.2] that $D(y) = 0$ has 3 solutions: the double root $y = 1$, a solution $y_1 \in [-1, 1)$, and a solution $y_4 \in (1, \infty) \cup (-\infty, -1]$. Consequently, one can see that for $y \in [y_1, 1]$, we have $X_+(y) = \overline{X_-(y)}$. This is in particular used in the computation of harmonic functions, as in [17] or [12]. The idea is to define the domain $\mathcal{G}$ as the area bounded by the curve $X_\pm([y_1, 1])$, and notice that the functional equation (10) leads to the boundary value problem

$$K(x,0)H(x,0) - K(\overline{x},0)H(\overline{x},0) = 0 \tag{14}$$

on $\partial\mathcal{G} \setminus \{1\}$, while $K(x,0)H(x,0)$ is analytic in the interior of $\mathcal{G}$ and continuous on $\overline{\mathcal{G}} \setminus \{1\}$ (cf [6, 17]). One can then construct a mapping $\omega : \mathbb{C} \to \bar{\mathbb{C}}$ which is a fundamental solution of the above BVP, in the sense that any other solution can be written as some entire function applied to ω [17]. This ω then has the properties

$$\omega(0) = 0, \quad \omega(X_+(y)) = \omega(X_-(y)) \quad \forall y \in [y_1, 1], \quad \frac{\partial\omega}{\partial x}(x) \neq 0 \quad \forall x \in \mathcal{G}^\circ. \tag{15}$$

In particular, ω is a conformal mapping of the domain $\mathcal{G}$. Furthermore, it has a pole-like singularity of order π/θ at $x = 1$, where θ is the inner angle at which $\partial\mathcal{G}$ intersects the x-axis. Alternatively, θ can be computed via

$$\theta = \arccos\left(-\frac{\sum ijp_{i,j}}{\sqrt{\sum i^2 p_{i,j}}\sqrt{\sum j^2 p_{i,j}}}\right), \tag{16}$$

see e.g. [17, 2.15]. This angle is also closely related to the group of a walk, which will be introduced below in Section 2.4. Additionally, we can see that $\omega \circ X_+$ is a conformal mapping of a region $\mathcal{G}'$ obtained by swapping x, y (by [10, Cor. 5.3.5]), and it has the same behaviour around 1 as ω. Finally, ω turns out to be an invariant in the sense of [2, Def. 4.3].

2.2 Discrete Polyharmonic Functions

We will start with a few elementary properties. Denote in the following by $\mathcal{H}_n$ the space of real-valued discrete n-polyharmonic functions. Clearly, $\mathcal{H}_n$ is a $\mathbb{R}$-vector space. Now, given any $\widehat{H}_n \in \mathcal{H}_n$, we can identify it with the sequence $\left(\widehat{H}_n, \widehat{H}_{n-1}, \dots, \widehat{H}_1\right)$, where $\triangle\widehat{H}_{k+1} = \widehat{H}_k$, and $\triangle\widehat{H}_1 = 0$. It is clear that any such sequence is uniquely defined by $\widehat{H}_n$. Now suppose that we have $\widehat{H}_n, \widehat{H}'_n \in \mathcal{H}_n$, such that, with their sequence representation as above, $\widehat{H}_1 = \widehat{H}'_1$. In this case, we have

$$\triangle^{n-1}\left[\widehat{H}_n - \widehat{H}'_n\right] = \widehat{H}_1 - \widehat{H}'_1 = 0, \tag{17}$$

thus $\widehat{H}_n - \widehat{H}'_n \in \mathcal{H}_{n-1}$. Therefore, provided that for each $\widehat{H}_n \in \mathcal{H}_n$ we can find a corresponding $\widehat{H}_{n+1} \in \mathcal{H}_{n+1}$, which will be shown below in Thms. 7 and 14, by simple induction one can prove the following Lemma:

▶ **Lemma 1.** *Let $\mathcal{H}_n$ be the space of real-valued, discrete n-polyharmonic functions in the quarter plane, subject to $\widehat{H}_n(0,\cdot) = \widehat{H}_n(\cdot,0) \equiv 0$. Then we have an isomorphy of vector spaces*

$$\mathcal{H}_n \cong (\mathcal{H}_1)^n . \tag{18}$$

In particular, if we are given any $\widehat{H}_n \in \mathcal{H}_n$, and we want to find all corresponding $\widehat{H}_{n+1} \in \mathcal{H}_{n+1}$, then this means that instead it suffices to find a single $\widehat{H}_{n+1}$ with this property as well as all harmonic functions, because any other such $\widehat{H}'_{n+1}$ can be written as $\widehat{H}_{n+1} + \widehat{G}_1$, for some $\widehat{G}_1 \in \mathcal{H}_1$. This property enables us to completely classify discrete harmonic functions. We already know (see e.g. [17],[6], using the idea of the BVP outlined above), that for any polynomial $P(x) \in \mathbb{R}[x]$, we can construct (the GF of) a harmonic function via

$$H(x,y) = \frac{P(\omega(x)) - P(\omega(X_+(y)))}{K(x,y)}, \tag{19}$$

where ω is the conformal mapping introduced in Section 2.1. We will now show what is, in a sense, the opposite direction of the above statement. The following theorem (as well as its proof) is an analogue to [12, Thm. 2], where a similar result is shown for the case of symmetric walks with small negative steps.

▶ **Theorem 2.** *For any discrete harmonic function with generating function $H(x,y)$, there is a unique formal power series $P(x)$ such that (19) holds. In particular, we have an isomorphism*

$$\mathcal{H}_1 \cong \mathbb{R}[[x]]. \tag{20}$$

Proof. One constructs explicitly a basis $\left(H_1^k\right)_{k\in\mathbb{N}}$ via

$$H_1^k(x,y) := \frac{P_k(\omega(x)) - P_k(\omega(X_+))}{K(x,y)}, \tag{21}$$

where $P_{2k}(x) = x^k(x-d_0)^k, P_{2k+1} = x^{k+1}(x-d_0)^k$ for $d_0 = \omega(X_+(0))$. See App. A. ◀

If we compare the functional equation (10) for harmonic and polyharmonic functions, then the only difference lies in the additional term of $xyH_n(x,y)$ on the right-hand side not vanishing for the latter. In terms of a BVP, this means that we now want to solve

$$K(X_+,0)\,H_n(X_+,0) - K(X_-,0)\,H_n(X_-,0) = X_+ yH_{n-1}(X_+,y) - X_- yH_{n-1}(X_-,y). \tag{22}$$

In an ideal world, the right-hand side of the latter equation would be 0 as in the harmonic case, and this is indeed what happens for the Simple Walk. In this case, we can proceed as before, and obtain an explicit formula for polyharmonic functions.

2.3 Example: the Simple Walk

The Simple Walk has the step set $\mathcal{S} = \{\uparrow,\rightarrow,\downarrow,\leftarrow\}$, each with probability $\frac{1}{4}$. We have

$$K(x,y) = \frac{xy}{4}\left(x+y+x^{-1}+y^{-1}\right) - xy, \quad \omega(x) = \frac{-2x}{(1-x)^2}, \quad \omega(X_+) = -\omega(y). \tag{23}$$

As it turns out that the right-hand side of (22) keeps vanishing, one can iteratively construct polyharmonic functions via $H_{n+1}(x,y) := \frac{xyH_n(x,y) - X_+ yH_n(X_+,y)}{K(x,y)}$. This allows us to find an explicit expression for all resulting polyharmonic functions. It appears that this property is directly tied to the fact that $\pi/\theta = 2$, where θ is given by (16).

▶ **Theorem 3.**

1. *The functions defined by*

$$H_m^k(x,y) = 2^{m-1}\omega(X_0)^{m-1}\frac{\omega(x)-\omega\left(X_+\right)}{K(x,y)}\left[\sum_{j=0}^{k-1} s_m(j)\omega\left(X_+\right)^j\omega(x)^{k-j-1}\right], \tag{24}$$

where $s_l : \mathbb{N} \to \mathbb{N}$ is defined inductively via $s_1(j) = 1, s_{l+1}(j) = \sum_{i=1}^{j+1} s_l(j)$, are polyharmonic functions with $\triangle H_{m+1}^k = H_m^k$.

2. *Given any formal power series $P(x) = \sum a_n x^n$ and any m, the limit*

$$H_m^P(x,y) := \lim_{n\to\infty}\sum_{j=0}^{n} a_j H_m^j \tag{25}$$

exists, and we again have $\triangle H_{m+1}^P = H_m^P$. In particular, any discrete m-polyharmonic function can be written as H_m^P for some P.

Proof. See App. B. ◀

Using Thm. 3, we can e.g. directly compute $H_1^1 = -\frac{8}{(1-x)^2(1-y)^2}$, $H_2^1 = -\frac{32y}{(x-1)^2(y-1)^4}$, $H_3^1 = -\frac{128y^2}{(x-1)^2(y-1)^6}$.

Proceeding to compute the generating functions $V_{1,2,3}$ of $v_{1,2,3}$ as given in (3)–(5), we obtain

$$V_1 = 64H_1^1, V_2 = \frac{3}{8}H_2^1 - \frac{3}{8}H_1^2 + 60H_1^1, V_3 = -24H_3^1 + 24H_2^2 + 72H_2^1 - 30H_1^3 - 72H_1^2 + 5072H_1^1.$$

It is somewhat striking here that the only p-polyharmonic part contained in v_p is H_p^1, which is in some manner the simplest possible. At this stage there is neither a proof that this is always true nor a counter-example.

2.4 Decoupling

While the computation for the Simple Walk turned out to be fairly simple, this was mainly due to the right-hand side of (22) consistently vanishing. This does not happen in general. For the Tandem Walk, for instance, we arrive at

$$K(X_+,0)H_1^1(X_+,0) - K(X_-,0)H_1^1(X_+,0) = \frac{y^3\sqrt{1-4y}}{(y-1)^5}. \tag{26}$$

The direct approach using the BVP like in the harmonic case does not generally yield an explicit solution as easily as before. Instead, we choose a more combinatorial approach and utilize what is called a decoupling function in [2, Def. 4.7].

▶ **Definition 4.** *Let $M(x,y)$ be an expression in x, y. If we can find $F(x), G(y)$ such that*

$$F(x) + G(y) \equiv M(x,y) \mod K(x,y), \tag{27}$$

*then we say that F is a **decoupling function** of M.*

These decoupling functions are closely related to the concept of invariants as in [2, Def. 4.3]. An example of a decoupling function will for instance be given in Section 2.5. Let in the following H' be polyharmonic, and H be such that $\triangle H = H'$. By substitution into (22), we directly find that for $F(x)$ a decoupling function of $xyH'(x,y)$ we have

$$K(X_+,0)H(X_+,0) - F(X_+) - [K(X_-,0)H(X_-,0) - F(X_-)] = 0. \tag{28}$$

In other words, if one knows how to compute a decoupling function of $xyH'(x,y)$, then one can again let $K(x,0)H(x,0) - F(x) = P(\omega)$; by the same arguments as for the BVP above one will then eventually arrive at a solution for $H(x,y)$. In [6, App. C], a decoupling function is guessed in order to compute a biharmonic function for the Tandem Walk. It turns out, however, that such a decoupling function can be explicitly computed for any model as long as a certain group of the corresponding step set is finite. This group is generated by the mappings

$$\Phi: \begin{cases} x & \mapsto x^{-1}\frac{\tilde{c}(y)}{\tilde{a}(y)}, \\ y & \mapsto y, \end{cases} \qquad \Psi: \begin{cases} x & \mapsto x, \\ y & \mapsto y^{-1}\frac{c(x)}{a(x)}. \end{cases} \tag{29}$$

One can easily see that Φ, Ψ are involutions, and depending on the order of $\Theta := \Phi \circ \Psi$, the group can be either finite or infinite. This group has been of interest in the study of random walks for some time now, see e.g. [5, 18].

▶ **Theorem 5** (see [2, Thm. 4.11]). *Suppose our step set has a finite group of order $2n$, and $M(x,y)$ is such that*

$$\sum_{\gamma\in\mathcal{G}} \operatorname{sgn}(\gamma)\gamma\left(M(x,y)\right) = 0. \tag{30}$$

Then a decoupling function of $M(x,y)$ is given by

$$F(x) = -\frac{1}{n}\sum_{i=1}^{n-1}\theta^i\left[M(x,Y_+) + M(x,Y_-)\right]. \tag{31}$$

In the following, we will show that $xyH_n(x,y)$ will turn out to have an orbit sum of 0 for any polyharmonic H_n. This is in particular independent of whether or not the given model has a vanishing orbit sum as in [2].

▶ **Corollary 6.** *Suppose the group of the step set is finite and has a series representation around $(0,0)$. Then any rational function $M(x,y)$ of the form*

$$M(x,y) = xy\frac{u(x)+v(y)}{K(x,y)} \tag{32}$$

allows for a decoupling function via (31).

Proof. This follows directly from the fact that the denominator $\frac{1}{xy}K(x,y)$ is invariant under $\mathcal{G}$. Alternating orbit summation over the numerator leads to a telescopic sum. ◀

If a model has a finite group, then it can be shown that $\pi/\theta \in \mathbb{Q}$ (cf [10, 7.1]). While the following is only stated for $\pi/\theta \in \mathbb{Z}$, it is possible to extend the statements for an arbitrary finite group. This relaxation of conditions, however, adds a lot more technicalities as the resulting functions are not rational anymore.

▶ **Theorem 7.** *Suppose our step set has finite group and $\pi/\theta \in \mathbb{Z}$. Let $H_1^k(x,y)$ be defined by (21). We can then define inductively*

$$H_n^k(x,y) = \frac{xyH_{n-1}^k(x,y) - F_{n-1}^k(x) - \left[X_+yH_{n-1}^k(X_+,y) - F_{n-1}^k(X_+,y)\right]}{K(x,y)}, \tag{33}$$

where $F_n^k(x)$ is the decoupling function of $xyH_n^k(x,y)$ defined by (31). Then, $H_n^k(x,y)$ is a rational function in $\mathcal{H}_n$ for all n,k, which satisfies $\triangle H_{n+1}^k = H_n^k$ as well as (30). For each n,k we can write

$$H_n^k(x,y) = \frac{p_{n,k}(x,y)}{(1-x)^\alpha(1-y)^\alpha}, \tag{34}$$

where $p_{n,k}$ is a polynomial and $\alpha = k\cdot\pi/\theta + 2(n-1)$.

Proof. See App. C. ◀

The downside of this construction is that we do not know for sure that for any k, the sum $\sum_{n=1}^\infty H_n^k$ converges. This property would be very useful in the proof that we can indeed find all polyharmonic functions. However, utilizing the functional equation (10) and proceeding similarly as in the proof of Thm. 2, one can for each induction step find a harmonic function $H_{1,n}^k$ such that $\widehat{H_n^k} := H_n^k + H_{1,n}^k$ has order at least $\left[\frac{k}{2}\right]$ at 0. The thusly defined $\widehat{H_n^k}$ therefore satisfy the conditions of the following theorem.

▶ **Theorem 8.** *Let $\left(H_n^k\right)_{n,k\in\mathbb{N}}$ be a family of discrete polyharmonic functions, such that*
1. *$H_1^k = H_1^k$ is given by (21),*
2. *$\triangle H_{n+1}^k = H_n^k$,*
3. *For any n and any sequence (a_n), the sum $\sum_{n=1}^\infty a_n H_n^k$ converges.*

Then, given any $H_n \in \mathcal{H}_n$, we can find $a_{i,j}, 1\le i\le n, j\in\mathbb{N}$, such that

$$H_n = \sum_{i=1}^n\sum_{j=1}^\infty a_{i,j}H_i^j. \tag{35}$$

Proof. By induction. For $n=1$, the statement is nothing but Thm. 2. Now assume the theorem holds for n, and let $H_{n+1}\in\mathcal{H}_{n+1}$. By definition, we must then have $H_n := \triangle H_{n+1} \in \mathcal{H}_n$, so we can write $H_n = \sum_{i=1}^n\sum_{j=1}^\infty a_{i,j}H_i^j$. By construction, $\overline{H_{n+1}} := \sum_{i=1}^n\sum_{j=1}^\infty a_{i,j}H_{i+1}^j$ is then a $n+1$-polyharmonic function with $\triangle\overline{H_{n+1}} = H_n$. Thus, $H_{n+1} - \overline{H_{n+1}} \in \mathcal{H}_1$, and an application of Thm. 2 and relabeling of the coefficients immediately yields the statement. ◀

As an immediate consequence, we can state that any discrete polyharmonic function can be expressed as a countable sum of our H_n^k as defined by (33).

2.5 Example: the Tandem Walk

To illustrate the results from Section 2.4, consider the Tandem Walk, which has the step set $\mathcal{S} = \{\rightarrow,\downarrow,\nwarrow\}$, with weights $\frac{1}{3}$ each. We have

$$K(x,y) = \frac{xy}{3}\left(x^{-1}+y+xy^{-1}\right) - xy, \quad \omega(x) = \frac{27x^2}{4(x-1)^3}, \quad \omega\left(X_+\right) = \frac{-27y}{4(y-1)^3}. \tag{36}$$

We directly obtain $H_1^1 = \frac{\omega(x)-\omega(X_+)}{K(x,y)} = \frac{81(xy-1)}{4(x-1)^3(y-1)^3}$, leading to the harmonic function $h(i,j) = (i+1)(j+1)(i+j+2)$. Using (31), we obtain the decoupling function $F_1(x) = -\frac{81x^3}{4(1-x)^5}$. Note that this decoupling function is not the same one as is given in [6, App. C], where instead (after scaling) $F_1' = \frac{-81x^3}{4(1-x)^6}$ is given. This goes to show that the choice of a decoupling function is, due to the invariance property in (15), unique only up to functions of ω; in this particular case we have (up to a multiplicative constant) $F_1'(x) - F_1(x) = \omega(x)^2$.

Using F_1 in (33) directly gives us the biharmonic function $H_2^1 = -\frac{243(xy-1)(x+y+xy(x+y-4))}{(x-1)^5(y-1)^5}$. Once again using (31) gives us the next decoupling function $F_2(x) = \frac{81x^2(x+2)}{4(x-1)^7}$, which we can then use to compute $H_3^1 = \frac{p(x,y)}{(x-1)^7(y-1)^7}$, for $p(x,y)$ a somewhat unwieldy polynomial of degree 9.

3 The Continuous Case

We now consider solutions of (6) with the usual Laplacian. First, it needs to be made clear in which way the latter corresponds to a given step set. So instead of a discrete random walk on $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$, we can also consider a Brownian motion on $\mathbb{R}^+ \times \mathbb{R}^+$. Any such Brownian motion is defined by its covariance matrix $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$, and its infinitesimal generator is the Laplacian

$$\triangle = \frac{1}{2}\left(\sigma_{11}\frac{\partial^2}{\partial x^2} + 2\sigma_{12}\frac{\partial^2}{\partial x \partial y} + \sigma_{22}\frac{\partial^2}{\partial y^2}\right), \tag{37}$$

consequently any polyharmonic function with respect to this Brownian motion satisfies (6). It can be shown that this Brownian motion is the scaling limit of any non-degenerate discrete random walk with small steps and zero drift such that $\mathbb{E}X^2 = \sigma_{11}, \mathbb{E}XY = \sigma_{12}, \mathbb{E}Y^2 = \sigma_{22}$ [13]. While it is possible to compute solutions of (6) explicitly, for instance using polar coordinates as in [6, 11], one can also follow an approach of [17, App. A]. There, instead of a generating function, the authors consider the Laplace transform and obtain the following functional equation:

$$\gamma(x,y)L(h)(x,y) = \frac{1}{2}\left(\sigma_{11}L_1(h)(y) + \sigma_{22}L_2(h)(x)\right) + L\left(\triangle(h)\right)(x,y), \tag{38}$$

where

$$\gamma(x,y) = \frac{1}{2}\left(\sigma_{11}x^2 + 2\sigma_{12}xy + \sigma_{22}y^2\right), \quad L(f)(x,y) = \int_0^\infty \int_0^\infty e^{-ux-vy} f(u,v)\mathrm{d}u\mathrm{d}v, \tag{39}$$

$$L_1(f)(y) = \int_0^\infty \frac{\partial f}{\partial x}(0,v)e^{-vy}\mathrm{d}v, \qquad L_2(f)(x) = \int_0^\infty \frac{\partial f}{\partial y}(u,0)e^{-ux}\mathrm{d}u, \tag{40}$$

see also [6, 2.2]. For any polyharmonic function h_n, this yields

$$\gamma(x,y)L(h_n)(x,y) = \frac{1}{2}\left(\sigma_{11}L_1(h_n)(y) + \sigma_{22}L_2(h_n)(x)\right) + L(h_{n-1})(x,y), \tag{41}$$

where $h_{n-1} = \mathcal{L}(h_n)$ (for $n = 1$ we let $h_0 = 0$). This functional equation, which is very similar to (10), can now be utilized in order to compute continuous polyharmonic functions. For the harmonic and biharmonic case this has already been done in [6, 2.2], where it is also mentioned that a similar method should work to compute higher order polyharmonic functions. Such a method shall be presented in the following.

3.1 Continuous Polyharmonic Functions

The idea used in computing harmonic functions in [17, 6] is very much the same as in the discrete setting. We can rewrite and obtain as in Section 2.1

$$\gamma(x,y) = \frac{1}{2}\left(\sigma_{11}x^2 + 2\sigma_{12}xy + \sigma_{22}y^2\right), \quad x_\pm = c_\pm y, \quad \widehat{\omega}(x) = \frac{1}{x^{\pi/\theta}}, \tag{42}$$

where $c_\pm = ce^{\pm i\theta}$, and eventually arrive at the following result.

▶ **Proposition 9** ([6, Th. 2.4]). *For any polynomial $P(x)$, the function*

$$L(h^P)(x,y) := \frac{P(\omega(x)) - P(\omega(c_+ y))}{\gamma(x,y)} \tag{43}$$

is the Laplace transform of a harmonic function.

The equivalent for the BVP for continuous polyharmonic functions now reads

$$\sigma_{22} L_2(h_n)(c_+ y) - \sigma_{22} L_2(h_n)(c_- y) = L(h_{n-1}(c_+ y, y)) - L(h_{n-1}(c_- y, y)), \tag{44}$$

where as in Section 2.4 the difference to the harmonic case is the right-hand side not necessarily being equal to 0. Our goal will now be to construct a decoupling function $f_{n-1}(x)$, such that

$$f_{n-1}(c_+ y) - f_{n-1}(c_- y) = L(h_{n-1})(c_+ y, y) - L(h_{n-1})(c_- y, y). \tag{45}$$

The computation of such a decoupling function here turns out to be a lot simpler than in the discrete case, as can be seen in the example of the scaling limit of the Tandem Walk.

3.2 Example: the Scaling Limit of the Tandem Walk

For the scaling limit of the Tandem Walk, we have

$$\gamma(x,y) = \frac{1}{3}\left(x^2 - xy + y^2\right), \quad c_\pm = \frac{1 \pm i\sqrt{3}}{2}, \quad \widehat{\omega}(x) = \frac{1}{x^3}. \tag{46}$$

Selecting $L(h_1)(x,y) = \frac{\widehat{\omega}(x) - \widehat{\omega}(c_+ y)}{\gamma(x,y)} = \frac{3(x+y)}{x^3 y^3}$, (45) takes the form $f_1(c_+ y) - f_1(c_- y) = \frac{3i\sqrt{3}}{y^5}$. Since the right-hand side is homogeneous (which, in fact, is generally true, seeing as both $\widehat{\omega}$ and γ are homogeneous), the ansatz $f_1(x) = \frac{\alpha}{x^5}$ is very reasonable. By a quick computation, one obtains that α must be -1. Everything works out, we obtain $f_1(x) = \frac{-3}{x^5}$ and a biharmonic function

$$L(h_2)(x,y) = \frac{L(h_1)(x,y) - f_1(x) - [L(h_1)(c_+ y, y) - f_1(c_+ y)]}{\gamma(x,y)} = \frac{9(x+y)(x^2+y^2)}{x^5 y^5}. \tag{47}$$

3.3 Decoupling

In the following, we let (analogously to the discrete case)

$$L(h_1^k)(x,y) := \frac{\widehat{\omega}(x)^k - \widehat{\omega}(c_+ y)^k}{\gamma(x,y)}. \tag{48}$$

To compute polyharmonic functions, all we need are decoupling functions, i.e. that we can always proceed as in the example above. Given a (homogeneous) polyharmonic function $L(h)$, this can be seen to directly depend on its degree. If $\deg L(h)$ is divisible by π/θ, then our ansatz would not work, for then $c_+^{-\deg L(h)} = c_-^{-\deg L(h)}$. However, this turns out not to matter: were one to continue the above example of the Tandem Walk until the first case where this could be an issue, that is, computing f_3, then one would see that we already have $L(h_3)(c_+ y, y) = L(h_3)(c_- y, y)$, meaning we can select $f_3 = 0$. This is not a coincidence and will always happen in those cases. Since this is essential in order to continue the procedure but our proof relies heavily on convergence of discrete polyharmonic functions, it will be stated here and be proven in Section 4.

▶ **Lemma 10.** *The procedure starting from a $L(h_1^k)$ described in Section 3.2 always works, meaning that there is a constant α such that we can find a decoupling function of the form*

$$f_n(x) = \frac{\alpha}{x^{\deg L(h_n)}}. \tag{49}$$

Utilizing the above Lemma, it is now easy to prove the continuous analogue of Thm. 7.

▶ **Theorem 11.** *Let $L(h_1^k)(x,y)$ be defined by (48). We can then define inductively*

$$L(h_n^k)(x,y) = \frac{L(h_{n-1})(x,y) - f_{n-1}(x) - [L(h_{n-1})(c_+y,y) - f_{n-1}(c_+y)]}{\gamma(x,y)}, \tag{50}$$

where $f_n(x)$ is a decoupling function as in Section 3.2. Then, $L(h_n^k)(x,y)$ is the Laplace transform of an n-harmonic function, such that $\mathcal{L}h_n^k = h_{n-1}^k$. For each n,k we can write

$$L(h_n^k)(x,y) = \frac{q_{n,k}(x,y)}{x^\alpha y^\alpha}, \tag{51}$$

for $\alpha = k\pi/\theta + 2n$ and $q_{n,k}(x,y) \in \mathbb{C}\left[x,y,x^{\pi/\theta},y^{\pi/\theta}\right]$. In particular, if $\pi/\theta \in \mathbb{Z}$, then $q_{n,k}(x,y)$ is a polynomial. Furthermore, $q_{n,k}$ is homogeneous of degree $k\pi/\theta - 2 + 4n$.

Proof. For $n = 1$, the statement can be checked directly. The rest follows inductively using Lemma 10. ◀

4 Convergence

Having continuous and discrete polyharmonic functions associated with a step set, it is natural to assume there would be some kind of connection between them. And indeed, this turns out to be the case, in a similar fashion as has been shown for harmonic functions of symmetric walks in [12, Th. 1]. Comparing the Laplace transform and the generating functions of some functions h and g, we have

$$L(h)(x,y) = \int_0^\infty \int_0^\infty e^{-ux-vy} h(u,v)\mathrm{d}u\mathrm{d}v, \quad G(x,y) = \sum_{i=0}^\infty \sum_{j=0}^\infty x^i y^j g(i,j). \tag{52}$$

It is therefore fairly natural to consider expressions of the form $H\left(e^{-x},e^{-y}\right)$. To transition from the discrete to the continuous setting we will also need some scaling parameter, which eventually leads to us considering limits of the form $\lim_{\mu\to 0}\mu^\alpha H\left(e^{-\mu x},e^{-\mu y}\right)$ for some constant α. That this kind of limit is reasonable is further shown by the following relations between expressions used in the discrete and continuous cases respectively.
The following Lemma 12 can be proven by straightforward computation.

▶ **Lemma 12.** *Let the discrete and continuous kernels K and γ belong to the same step set. We then have*

$$\lim_{\mu\to 0} \frac{K\left(e^{-\mu x},e^{-\mu y}\right)}{\mu^2} = \gamma(x,y), \quad \lim_{\mu\to 0} X_\pm\left(e^{-\mu x}\right) = 1 - c_\pm \mu x + \mathcal{O}(\mu^2). \tag{53}$$

Using Lemma 12, the strategy in order to show a general convergence of polyharmonic functions is quite simple: we use the fact that the recursive definitions (33) and (50) have the same structure, and take the limit of each term separately. All that remains is to consider decoupling functions. However, using once again Lemma 12, this turns out to be rather straightforward, too.

▶ **Lemma 13.** *Suppose we have a discrete and continuous polyharmonic function H and $L(h)$, and a constant α such that*

$$\lim_{\mu\to 0} \mu^\alpha H\left(e^{-\mu x},e^{-\mu y}\right) = L(h)(x,y). \tag{54}$$

Then, for any decoupling function $F(x)$ of $xyH(x,y)$,

$$f(x) := \lim_{\mu\to 0} \mu^{\alpha} F\left(e^{-\mu x}\right) \tag{55}$$

is a decoupling function of $L(h)(x,y)$.

Proof. By taking the corresponding limit of (28). ◀

We can now formulate and prove the following theorem, which shows convergence between the H_n^k and the $L(h_n^k)$ defined in Sections 2.4 and 3.3 respectively. In doing so, we will also prove Lemma 10. Since we will be using Thm. 11 to do so, which in turn utilizes the former, it is worth taking a moment to make sure that in each induction step in the proof of Thm. 14 for some fixed $n+1$, we use the statement of Thm. 11 for n, and prove Lemma 10 for $n+1$. We therefore do not enter any circular reasoning.

▶ **Theorem 14.** *Let $\pi/\theta \in \mathbb{Z}$ and $H_n^k, L(h_n^k)$ be defined by (33), (50) respectively. Then*

$$\lim_{\mu\to 0} \mu^{k\pi/\theta+2n} H\left(e^{-\mu x}, e^{-\mu y}\right) = \alpha_{n,k} L\left(h_n^k\right)(x,y) \tag{56}$$

for some constants $\alpha_{n,k} \neq 0$.

Proof of Thm. 14 and Lemma 10. For $n = 1$ the statement can be checked by direct computation. Now let the statement be true for some n. By Lemma 13, we know that we can define a decoupling function of $L(h_n^k)(x,y)$ via $f_n^k(x) = \frac{1}{\alpha_{n,k}} \lim_{\mu\to 0} \mu^{k\pi/\theta+2n} F_n^k\left(e^{-\mu x}\right)$. From the proof of Thm. 7, we know that $F_n^k(x)$ has the form $P_{n,k}(x)/(1-x)^{2n+k\pi/\theta}$ for some polynomial $P_{n,k}(x)$. Therefore, the ansatz used in Ex. 3.2 must yield a solution, so Lemma 10 is proven. Having now completely proven Thm. 11 for $n+1$, we can simply take piecewise limits of (33), (50), and obtain the statement. ◀

5 Outlook/Open Questions

While Thm. 7 is formulated only for the case $\pi/\theta \in \mathbb{N}$ here, the construction of polyharmonic functions works in essentially the same manner for $\pi/\theta \in \mathbb{Q}$ (albeit the process becomes a bit more technical as we need to work with algebraic instead of rational functions). This will be addressed in a future paper, together with more complete proofs of the theorems above.

- Seeing as for any finite group we have $\pi/\theta \in \mathbb{Q}$ [10, 7.1] (though the reverse does not hold), it follows that we can construct arbitrary polyharmonic functions for any walk with drift 0, small steps and finite group. This directly leads to the question of what to do in the infinite group case. There are some examples where one can show by a direct ansatz that a decoupling function of reasonably nice shape cannot exist, and it can indeed be conjectured that a decoupling function can be computed if and only if the group of the walk is finite. Should this hold, then of course the next question would be how discrete polyharmonic functions could be computed in the infinite group case.
- Another open question regards the positivity of harmonic functions. For drift 0 walks, we know [3] that there is a unique (up to multiples) positive harmonic function. In all known examples this harmonic function is given by H_1^1 as in (21), and for the symmetric case it was already conjectured in [12, Conj. 1] that this might be true in general. To the author's knowledge there has not yet been a general result in that direction.
- While the above gives a description of all discrete polyharmonic function, it is still not clear which ones appear in expansions of the form (2). If the conjecture about the positive harmonic function were to be true, then the harmonic function v_1 would have to be a

multiple of H_1^1. One could then proceed to ask whether this turns out to be true in a more general setting, i.e. if the p-harmonic part of v_p as given by (2) is always a multiple of H_p^1, as is the case in example 2.3.

- In all known (at least to the author) enumeration problems of lattice paths with small steps and zero drift in the quarter plane that we can compute higher-order asymptotics of, an expansion as in (2) exists. This leads to the question of whether there are models for which there is no such expansions, or if there are conditions under which the existence of the latter can be shown in general.

References

1. Emilio Almansi. Sull'integrazione dell'equazione differenziale $\triangle 2n = 0$. *Annali di Matematica Pura ed Applicata*, 2:1–51, 1899.
2. Olivier Bernardi, Mireille Bousquet-Mélou, and Kilian Raschel. Counting quadrant walks via Tutte's invariant method. *Combinatorial Theory (to appear)*, 2021.
3. Aymen Bouaziz, Sami Mustapha, and Mohamed Sifi. Discrete harmonic functions on an orthant in zd. *Electronic Communications in Probability, Institute of Mathematical Statistics (IMS), 10.1214/ECP.v20-4249. hal-01214538*, 20:1–13, 2015.
4. Mireille Bousquet-Mélou. Counting walks in the quarter plane. *Mathematics and Computer Science: Algorithms, trees, combinatorics and probabilities, Trends in Mathematics*, pages 49–67, 2002. doi:10.48550/arXiv.1708.06192.
5. Mireille Bousquet-Mélou and Marni Mishna. Walks with small steps in the quarter plane. *Contemp. Math.*, 520:1–40, 2010.
6. François Chapon, Éric Fusy, and Kilian Raschel. Polyharmonic functions and random processes in cones. In *DMTCS Proceedings of AofA no. 9*, pages 1–19, 2020. doi:10.4230/LIPIcs.AofA.2020.9.
7. Joel Cohen, Flavia Colonna, Kohur GowriSankaran, and David Singman. Polyharmonic functions on trees. *American Journal of Mathematics, 124, 999-1043*, 124:999–1043, 2002. doi:10.48550/arXiv.1802.06239.
8. Julien Courtiel, Stephen Melczer, Marni Mishna, and Kilian Raschel. Weighted lattice walks and universality classes. *Journal of Combinatorial Theory, Series A*, 152, 2017. doi:10.1016/j.jcta.2017.06.008.
9. Denis Denisov and Vitali Wachtel. Random walks in cones. *Annals of Probability*, 43(3):992–1044, 2015. doi:10.1214/13-AOP867.
10. Guy Fayolle, Roudolf Iasnogorodski, and Vadim Malyshev. *Random Walks in the Quarter Plane*. Springer International Publishing AG, 2017.
11. David J. Griffiths. *Introduction to Electrodynamics 4th ed.* Cambridge University Press, 2013.
12. Viet Hung Hoang, Kilian Raschel, and Pierre Tarrago. Constructing discrete harmonic functions in wedges. *arXiv preprint*, 2020. arXiv:2012.08947v1.
13. Gregory F. Lawler and Vlada Limic. *Random Walk: A Modern Introduction*. Cambridge University Press, 2012. doi:10.1017/CBO9780511750854.
14. Sergey A. Lurie and Valery V. Vasiliev. *The Biharmonic Problem in the Theory of Elasticity*. Taylor & Francis Ltd, 1995.
15. Stephen Melczer. *An Invitation to Analytic Combinatorics*. Springer International Publishing AG, 2021.
16. Massimo A. Picardello and Wolfgang Woess. Boundary representations of λ-harmonic and polyharmonic functions on trees. *Potential Analysis*, 51:541–561, 2019.
17. Kilian Raschel. Random walks in the quarter plane, discrete harmonic functions and conformal mappings. *Stochastic Processes and their Applications 124, 3147-3178*, 124:3147–3178, 2014.
18. Michael Singer and Charlotte Hardouin. On differentially algebraic generating series for walks in the quarter plane. *Selecta Mathematica*, 27, 2021. doi:10.1007/s00029-021-00703-9.
19. William T. Tutte. Chromatic sums revisited. *Aequationes Math.*, 50(1-2):95–134, 1995. doi:10.1007/978-3-0348-9096-0_7.

A Proof of Thm. 2

▶ **Theorem 2.** *For any discrete harmonic function with generating function $H(x,y)$, there is a unique formal power series $P(x)$ such that (19) holds. In particular, we have an isomorphism*

$$\mathcal{H}_1 \cong \mathbb{R}[[x]]. \tag{20}$$

Proof (outline). The arguments are mostly the same as in [12, Thm. 2]. From (10), it follows that $K(x,y)H(x,y)$ is already uniquely defined by the (univariate) boundary terms $K(x,0)H(x,0)$ and $K(0,y)H(0,y)$. It suffices to consider two cases:

1. $K(0,0) = 0$:
   In this case, we have $X_+(0) = 0$, we can therefore substitute X_+ into a power series. Doing so in (10) gives

   $$0 = K(X_+,0)H(X_+,0) + K(0,y)H(0,y). \tag{57}$$

   Utilizing this to substitute for $K(0,y)H(0,y)$ in (10), we obtain

   $$K(x,y)H(x,y) = \underbrace{K(x,0)H(x,0)}_{=:P(x)} - \underbrace{K(X_+,0)H(X_+,0)}_{=:P(X_+)}. \tag{58}$$

   Setting

   $$H_1^m(x,y) = \frac{\omega(x)^m - \omega(X_+)^m}{K(x,y)}, \tag{59}$$

   and utilizing that around 0 we have (after scaling and potentially switching x,y) $\omega(x) = \frac{x(1+p(x))}{(1-x)^{\pi/\theta}}$ (see [10, 5.3]; use that our walk is not singular), we can iteratively compute coefficients a_k such that $\sum a_j \omega(x)^k = P(x)$. To see that at the end we indeed obtain a power series, one can apply the Weierstraß preparation theorem.
2. $K(0,0) \neq 0$:
   In this case, the previous approach does not work anymore since substitution of X_+ into an arbitrary power series fails. Instead, let now $\omega(x) = \sum x^n c^n, \omega(X_+) = \sum y^n d^n$. We know that $c_1, d_1 \neq 0, c_0 = 0$ (see [10, 5.3], and notice that $p_{-1,-1} \neq 0$).
   We can now proceed by defining

   $$P_{2m}(z) = z^m (z - d_0)^m, \tag{60}$$
   $$P_{2m+1}(z) = z^{m+1}(z - d_0)^m. \tag{61}$$

   Letting

   $$H_1^m(x,y) := \frac{P_m(\omega(x)) - P_m(\omega(X_+))}{K(x,y)}, \tag{62}$$

   one can check that the monomial with non-zero coefficient with minimal degree in the series representation of $H_1^m(x,y)$ around 0 occurs for $k = l = m$ for m even, and $k = l+1 = m$ otherwise. Note here that $\omega(x), \omega(X_0)$ have non-vanishing derivative at 0 as $0 \in \mathcal{G}^\circ$, see [6, 5.3]. From there, given arbitrary power series $Q(x), R(y)$ with $Q(0) = R(0)$, one can again iteratively build coefficients a_n such that $\sum a_n P_n(\omega(x)) = Q(x)$, $\sum b_n P_n(\omega(X_+)) = R(y)$. We have thus constructed a harmonic functions with boundary terms $Q(x), R(y)$; since these were arbitrary we are done. Note that, since $K(0,0) \neq 0$, $1/K(x,y)$ can be written as a power series around 0. ◀

B Proof of Thm. 3

▶ **Theorem 3.**

1. *The functions defined by*

$$H_m^k(x,y) = 2^{m-1}\omega(X_0)^{m-1}\frac{\omega(x)-\omega(X_+)}{K(x,y)}\left[\sum_{j=0}^{k-1} s_m(j)\omega(X_+)^j\,\omega(x)^{k-j-1}\right], \tag{24}$$

where $s_l : \mathbb{N} \to \mathbb{N}$ is defined inductively via $s_1(j) = 1, s_{l+1}(j) = \sum_{i=1}^{j+1} s_l(j)$, are polyharmonic functions with $\triangle H_{m+1}^k = H_m^k$.

2. *Given any formal power series $P(x) = \sum a_n x^n$ and any m, the limit*

$$H_m^P(x,y) := \lim_{n\to\infty}\sum_{j=0}^{n} a_j H_m^j \tag{25}$$

exists, and we again have $\triangle H_{m+1}^P = H_m^P$. In particular, any discrete m-polyharmonic function can be written as H_m^P for some P.

Proof. For the first part, we use induction over m. For $m = 1$, the statement is nothing but (19). Now, assume the statement holds for $1, \dots, m-1$, and pick an arbitrary k. Due to the invariance property (15), and using that

$$xy\frac{\omega(x)-\omega(X_+)}{K(x,y)} = \frac{-8xy}{(1-x)^2(1-y)^2} = 2\omega(x)\omega(X_+), \tag{63}$$

we see that $X_+ yH_1^k(X_+,y) = X_- yH_1^k(X_-,y)$, since we can write $X_+ yH_1^k(X_+,y)$ as a polynomial in ω, $\omega(X_+)$. Using (10), we can deduce that a m-harmonic function with $\triangle H_m^k = H_{m-1}^k$ is given by

$$H_m^k(x,y) = \frac{xyH_{m-1}^k(x,y) - X_+ yH_{m-1}^k(X_+,y)}{K(x,y)}. \tag{64}$$

After a short computation one obtains

$$\begin{aligned} &xyH_m^k(x,y) - X_+ yH_m^k(X_+,y) \\ &\qquad = 2^m\omega(X_+)^m\left[\omega(x)\sum_{j=0}^{k-1} s_m(j)\omega(X_+)^j\,\omega(x)^{k-1-j} - \omega(X_+)^k\sum_{j=0}^{k-1} s_m(j)\right]. \end{aligned} \tag{65}$$

Using the algebraic identity

$$a\left[\sum_{j=0}^{k-1} c_j a^{k-j-1}b^j\right] - b^n\sum_{j=0}^{k-1} c_j = (a-b)\sum_{j=0}^{k-1}\left(\sum_{i=1}^{j+1} c_i\right)a^{k-j-1}b^j \tag{66}$$

for $a = \omega(x)$ and $b = \omega(X_+)$ then yields the statement.

For the second part, to see the existence of the limit it suffices to notice that the minimal degree of any non-zero coefficient is at least m (note that $\omega(0) = \omega(X_+(0)) = 0$). In the same fashion we can take the limit on both sides of (10) and see that both sides converge to the same power series.

To see that we can in this manner indeed produce all possible polyharmonic functions we proceed as in the proof of Thm. 8. ◀

C Proof of Thm. 7

▶ **Theorem 7.** *Suppose our step set has finite group and $\pi/\theta \in \mathbb{Z}$. Let $H_1^k(x,y)$ be defined by (21). We can then define inductively*

$$H_n^k(x,y) = \frac{xyH_{n-1}^k(x,y) - F_{n-1}^k(x) - \left[X_+ yH_{n-1}^k(X_+,y) - F_{n-1}^k(X_+,y)\right]}{K(x,y)}, \tag{33}$$

where $F_n^k(x)$ is the decoupling function of $xyH_n^k(x,y)$ defined by (31). Then, $H_n^k(x,y)$ is a rational function in $\mathcal{H}_n$ for all n,k, which satisfies $\triangle H_{n+1}^k = H_n^k$ as well as (30). For each n,k we can write

$$H_n^k(x,y) = \frac{p_{n,k}(x,y)}{(1-x)^\alpha(1-y)^\alpha}, \tag{34}$$

where $p_{n,k}$ is a polynomial and $\alpha = k\cdot\pi/\theta + 2(n-1)$.

Proof (outline). Consider first the case $n=1$. $H_1^k(x,y)$ being rational follows immediately from $\pi/\theta \in \mathbb{Z}$, and thus ω being rational (see [17, (3.12)]). As by construction the numerator $N_1^k(x,y)$ of $xyH_1^k(x,y)$ as defined in (33) satisfies $N_1^k(X_\pm) = 0$, it must be a multiple of $K(x,y)$, thus the only poles of H_1^k can be those coming from $\omega(x), \omega(X_+)$. Since $X_+ = 1$ only if $y=1$, the statement follows from (31). The existence of a decoupling function follows immediately from Prop. 6, and since $X_\pm(1) = Y_\pm(1) = 1$, we can conclude that $F_1^k(x)$ has its only pole at 1. We have thus shown the theorem for $n=1$, except for the computation of α which will be done at the end.

Now let $n \geq 2$ and assume the theorem holds for $n-1$. We first want to show that $F_n^k(x)$ (as defined by (31))is a decoupling function. We check the orbit sum criterion (30). Using (33), we can utilize that by induction hypothesis we already know that (30) is satisfied for $xyH_{n-1}^k(x,y)$; dividing by $\frac{1}{xy}K(x,y)$ does not change this, nor does substituting X_+ for x in the numerator. Therefore, it remains to show that $xy\left(F_{n-1}^k(x) - F_{n-1}^k(X_+)\right)/K(x,y)$ admits a decoupling function, but this is an immediate consequence of Cor. 6. In the same manner as for $n=1$, we conclude that $F_n^k(x)$ has its only pole at $x=1$. For the term $X_+yH_{n+1}^k(X_+,y) - F_{n-1}^k(X_+,y)$, note that it is nothing but the $G(y)$ in (27), for which an explicit formula similar to (31) is given in [2, Th. 4.11]. One can easily see that the arguments for $F_n^k(x)$ can be repeated directly for this $G(y)$, and thus this expression too can have its only pole at $y=1$.

$H_n^k(x,y)$ having the form given by (34) (for now with any α) again follows from the fact that its numerator vanishes at $x = X_\pm$, and thus must contain a factor $K(x,y)$. Lastly, to check that $\triangle H_{n+1} = H_n$ it suffices to substitute into the functional equation (10).

It remains to show that the order of the poles at $x,y=1$ is at most $k\cdot\pi/\theta + 2(n-1)$. For $n=1$ this can again be verified directly; afterwards it follows from induction: by a computation one can see that the order of the pole of $F(x)$ compared to the one at $x=1$ of $xyH(x,y)$ increases at most by 2, and by a similar argument for the $G(y)$ in (27) (see [2, Th. 4.11] for an explicit formula) one can show the same for $X_+yH(X_+,y) - F(X_+) = G(y)$. Using (33) finally yields the statement. ◀

Automorphisms of Random Trees

Christoffer Olsson
Department of Mathematics, Uppsala University, Sweden

Stephan Wagner
Department of Mathematics, Uppsala University, Sweden

Abstract

We study the size of the automorphism group of two different types of random trees: Galton–Watson trees and Pólya trees. In both cases, we prove that it asymptotically follows a log-normal distribution. While the proof for Galton–Watson trees mainly relies on probabilistic arguments and a general result on additive tree functionals, generating functions are used in the case of Pólya trees.

2012 ACM Subject Classification Mathematics of computing → Random graphs; Mathematics of computing → Generating functions

Keywords and phrases random tree, Galton–Watson tree, Pólya tree, automorphism group, central limit theorem

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.16

Funding *Christoffer Olsson*: Supported by the Knut and Alice Wallenberg Foundation.
Stephan Wagner: Supported by the Knut and Alice Wallenberg Foundation.

1 Introduction

The automorphism group is a fundamental object associated with a graph as it encodes information about its symmetries. Furthermore, counting mathematical objects up to symmetry is a classical subject in combinatorics which naturally relates to the automorphism group. An example is the case of graphs, where the number of different labelings of a graph G of order n is given by $\frac{n!}{|\operatorname{Aut} G|}$. In this paper we study properties of the automorphism groups associated with random trees, in particular Galton–Watson trees and Pólya trees. We show that the size of the automorphism group follows a log-normal distribution with parameters depending on tree type. The size of the automorphism group has previously been studied in special cases of Galton–Watson trees: binary trees (expected values and limiting distribution: [2]), labeled trees (limiting distribution: [6] and expected value: [16]), binary and ternary trees (expected values: [11] and [12]). It has also been studied for some other types of trees than those considered here: specifically, random recursive trees (expected value: [10]), and d-ary increasing trees (limiting distribution and moments: [13]).

For any rooted tree T, we have a recursive formula for the size of its automorphism group. Let $T_1, T_2, \dots, T_k$ be root branches with multiplicities $m_1, m_2, \dots, m_k$. Then we have

$$|\operatorname{Aut} T| = \prod_{i=1}^{k} m_i! |\operatorname{Aut} T_i|^{m_i},$$

derived from the fact that the automorphism group of a rooted tree is obtained from symmetric groups by iterated direct and wreath products (see [1], Proposition 1.15). In other words, the tree is invariant under the automorphisms of each of the root branches as well as under permutation of isomorphic branches. By taking logarithms, we find that $\log |\operatorname{Aut} T|$ is an *additive functional* of the tree, which is a real-valued function $F(T)$ that satisfies

$$F(T) = f(T) + \sum_{i=1}^{r} F(T_i),$$

where we sum over the r root branches and $f(T)$ is another function called the *toll* of the additive functional. In our case, the toll function is $\sum \log(m_i!)$. We note that we can rewrite the definition as

$$F(T) = f(T) + \sum_{i=1}^{k} m_i F(T_i),$$

where we are now summing over unique root branches $T_1, T_2, \ldots, T_k$ with multiplicities $m_1, m_2, \ldots, m_k$. We will show asymptotic normality of $\log |\operatorname{Aut} \mathcal{T}_n|$, which implies asymptotic log-normality of $|\operatorname{Aut} \mathcal{T}_n|$, where $\mathcal{T}_n$ denotes a random tree on n vertices. Limit theorems for additive functionals have been proven for various classes of random trees under different conditions, see [4, 5, 8, 13–15]. We will specifically make use of a general result due to Ralaivaosaona, Šileikis and the second author [14] that is based on earlier work by Janson [8].

Recall now that a *Galton–Watson tree* is a growth model where we start with one vertex, the root, and the number of children it has is given by a (discrete) random variable ξ, supported on some subset of the non-negative integers that includes at least zero and some number greater than one. The tree grows by letting each of the vertices have children of their own according to the offspring distribution ξ, independently of all other vertices. Different distributions for ξ give rise to different types of Galton–Watson trees. We are especially interested in the case of *critical* Galton–Watson trees, for which $\mathbb{E}\,\xi = 1$, as well as *conditioned* Galton–Watson trees where we condition on the size of the tree, i.e., we pick one of all possible Galton–Watson trees on n vertices at random. A related notion is that of the *size-biased* Galton–Watson tree, which has two different types of vertices. The *normal* vertices have the same offspring distribution ξ as before, while the *special* vertices get offspring according to the size-biased distribution $\hat{\xi}$ defined by $\mathbb{P}(\hat{\xi} = k) = k\,\mathbb{P}(\xi = k)$. We start the growth process with the root being special, and for each special vertex we choose exactly one of its children, uniformly at random, to be special as well. This means that the size-biased Galton–Watson tree has an infinite *spine* of special vertices, with non-biased unconditioned Galton–Watson trees attached to it. Conditioned Galton–Watson trees are closely connected to, and a special case of, *simply generated families* of trees (indeed, we can see them as two sides of the same coin, one being probabilistic and the other being combinatorial, see [3, Section 1.2.7]). Examples of Galton–Watson (and simply generated) trees are plane trees, labeled trees, d-ary trees, etc. The book [3] gives a general introduction to different types of random trees.

Pólya trees are rooted, unordered, unlabeled trees. They are not Galton–Watson trees, even though they have many similar properties, and cannot be interpreted as growth processes so we will need other methods to deal with them. The trees can be characterized by their generating function $P(x) = \sum_{T \in \mathcal{P}} x^{|T|}$, which satisfies

$$P(x) = x \exp\left(\sum_{k=1}^{\infty} \frac{P(x^k)}{k} \right). \tag{1}$$

We use $\mathcal{T}$ to denote Galton–Watson trees, $\mathcal{T}_n$ to denote conditioned Galton–Watson trees on n vertices and $\hat{\mathcal{T}}$ to denote size-biased trees. Similarly, we use T,T_n and $\hat{T}$ to denote specific realizations of the respective trees. Furthermore, we will use $\mathcal{P}$ and $\mathcal{P}_n$ to denote Pólya trees and Pólya trees of size n, respectively.

1.1 Results

In this paper, we prove the following theorem on the automorphism group of Galton–Watson trees.

▶ **Theorem 1.** *Let $\mathcal{T}_n$ be a conditioned Galton–Watson tree of order n with offspring distribution ξ, where $\mathbb{E}\,\xi = 1$, $0 < \operatorname{Var}\xi < \infty$ and $\mathbb{E}\,\xi^5 < \infty$. Then there exist constants μ and $\sigma^2 \geq 0$, depending on $\mathcal{T}$, such that*

$$\frac{\log|\operatorname{Aut}\mathcal{T}_n| - \mu n}{\sqrt{n}} \xrightarrow{d} \mathrm{N}(0,\sigma^2).$$

The condition on $\mathbb{E}\,\xi^5$ is needed for technical purposes and is valid for combinatorially significant examples such as labeled trees, plane trees and d-ary trees. The exponent 5 is probably not best possible, but required to apply the general result on additive functionals that our proof is based on.

The mean constant μ and even more so the variance constant σ^2 do not seem easy to compute numerically in general. In the appendix, we show how to derive the numerical values $\mu = 0.0522901\ldots$ and $\sigma^2 = 0.0394984\ldots$ in the special case of labeled trees, where a connection to Pólya trees can be exploited.

We can also prove the following, similar result, for the class of Pólya trees.

▶ **Theorem 2.** *Let $\mathcal{P}_n$ be a Pólya tree of order n. Then, $\mathbb{E}(\log|\operatorname{Aut}\mathcal{P}_n|) = \mu n + O(1)$ and $\operatorname{Var}(\log|\operatorname{Aut}\mathcal{P}_n|) = \sigma^2 n + O(1)$, with $\mu = 0.1373423\ldots$ and $\sigma^2 = 0.1967696\ldots$. Furthermore, we have*

$$\frac{\log|\operatorname{Aut}\mathcal{P}_n| - \mu n}{\sqrt{n}} \xrightarrow{d} \mathrm{N}(0,\sigma^2).$$

Even though both proofs rely, at their cores, on the same idea of approximating the additive functionals by simpler ones, they are fairly different at a glance. We prove Theorem 1 in Section 2 and Theorem 2 in Section 3.

2 The automorphism group of Galton–Watson trees

To prove asymptotic normality of $\log|\operatorname{Aut}\mathcal{T}_n|$, we will show that it is in fact an almost local additive functional, as defined in [14]. Intuitively, "almost local" means that looking at the first M levels of the tree gives us substantial (albeit not perfect) information about the value of the toll function at the root. We will let $\mathcal{T}^{(M)}$ denote the restriction of a Galton–Watson tree to its first M levels, where the root is at level 0, with similar definitions for the other classes of trees. The theorem we will use is the following.

▶ **Theorem 3** ([14])**.** *Let $\mathcal{T}_n$ be a conditioned Galton–Watson tree of order n with offspring distribution ξ, with $\mathbb{E}\,\xi = 1$ and $0 < \sigma^2 := \operatorname{Var}\xi < \infty$. Assume further that $\mathbb{E}\,\xi^{2\alpha+1} < \infty$ for some integer $\alpha \geq 0$. Consider a functional F of finite rooted ordered trees with the property that*

$$f(T) = O(\deg(T)^\alpha),$$

where f is the toll function associated with the functional.

Furthermore, assume that there exists a sequence $(p_M)_{M\geq 1}$ of positive numbers with $p_M \to 0$ as $M \to \infty$, such that

- *for every integer $M \geq 1$,*

$$\mathbb{E}\left|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)})|\hat{\mathcal{T}}^{(M)}\right)\right| \leq p_M,$$

for all $N \geq M$,

- *there is a sequence of positive integers* $(M_n)_{n\geq 1}$ *such that for large enough* n,

$$\mathbb{E}\,|f(\mathcal{T}_n) - f(\mathcal{T}_n^{(M)})| \leq p_{M_n}.$$

If $a_n = n^{-1/2}(n^{\max\{\alpha,1\}}p_{M_n} + M_n^2)$ *satisfies*

$$\lim_{n\to\infty} a_n = 0, \text{ and } \sum_{n=1}^{\infty} \frac{a_n}{n} < \infty,$$

then

$$\frac{F(\mathcal{T}_n) - \mu n}{\sqrt{n}} \xrightarrow{d} N(0,\gamma^2),$$

where $\mu = \mathbb{E}\, f(\mathcal{T})$ *and* $0 \leq \gamma^2 < \infty$.

The proof shows that the result still holds if we replace $(F(\mathcal{T}_n) - \mu n)/\sqrt{n}$ by $(F(\mathcal{T}_n) - \mathbb{E}\, F(\mathcal{T}_n))/\sqrt{n}$.

2.1 Galton–Watson trees isomorphic up to a certain level

In applying Theorem 3, we are led to consider the probability that two Galton–Watson trees are of height $\geq M$ and isomorphic. We use $\mathcal{C}$ to denote the set of isomorphism classes of Galton–Watson trees as well as $\mathcal{C}^M$ to denote the set of isomorphism classes of trees of height M (i.e., trees that have $M+1$ generations). The definitions extend to conditioned Galton–Watson trees as $\mathcal{C}_n$ and $\mathcal{C}_n^M$, respectively. We start with the following lemma.

▶ **Lemma 4.** *There exists some constant* $0 < c < 1$ *such that*

$$\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \leq c^M,$$

uniformly for all isomorphism classes $C \in \mathcal{C}^M$.

Proof. We say that a level L of a tree T agrees with C if it has the right number of vertices and the offsprings $\xi_1, \xi_2, \ldots, \xi_l$ agree with the offsprings of the same level in C, up to permutation. Let $L_1, L_2, \ldots$ denote the levels of the Galton–Watson tree $\mathcal{T}$. Then the probability is bounded by

$$\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \leq \prod_{i=0}^{M-1} \mathbb{P}(L_i \text{ agrees with } C | L_1, L_2, \ldots, L_{i-1}), \tag{2}$$

where we note that, by truncation, the M-th level will always agree with C, as long as the previous ones do. We can bound each factor in (2) by the probability of the level having the correct number of leaves, conditioned on the previous levels. This random variable follows a binomial distribution with probability $p = \mathbb{P}(\xi = 0)$. It is therefore sufficient to prove a bound $0 < c < 1$ (uniform in both l and k) on the probability that a binomial variable $X_l \sim \mathrm{Bin}(l,p)$ takes a specific value k.

We can in fact bound X_l in terms of p, since if we write X_l as a sum of Bernoulli variables $X_l = Y_1 + Y_2 + \ldots + Y_l$ we have

$$\begin{aligned}\mathbb{P}(Y_1 + Y_2 + \ldots + Y_l = k) &= \sum_r \mathbb{P}(Y_1 + Y_2 + \ldots + Y_{l-1} = r)\,\mathbb{P}(Y_l = k - r) \\ &\leq \sum_r \mathbb{P}(Y_1 + Y_2 + \ldots + Y_{l-1} = r) \max_y \mathbb{P}(Y_l = y) = \max_y \mathbb{P}(Y_l = y) = \max\{p, 1-p\}.\end{aligned}$$

We can thus take $c = \max\{p, 1-p\}$ as a uniform bound for all levels, and now (2) gives the result. ◀

We now see that for two independent trees $\mathcal{T}_1, \mathcal{T}_2$ we have

$$\mathbb{P}(\mathcal{T}_1^{(M)}, \mathcal{T}_2^{(M)} \text{ isomorphic and of height} \geq M) = \sum_{C \in \mathcal{C}^M} \mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)^2 \tag{3}$$

$$\leq \max_{C \in \mathcal{C}^M} \{\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)\} \sum_{C \in \mathcal{C}^M} \mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \tag{4}$$

$$= \max_{C \in \mathcal{C}^M} \{\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)\}. \tag{5}$$

Combining this with Lemma 4, we get the following corollary.

▶ **Corollary 5.** *Let $\mathcal{T}_1, \mathcal{T}_2$ be two independent Galton–Watson trees. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}_1^{(M)}, \mathcal{T}_2^{(M)} \text{ isomorphic and of height} \geq M) \leq c^M.$$

In fact, the argument in (3) also works when one of the trees is the size-biased tree $\hat{\mathcal{T}}$, which lets us bound the probability that a Galton–Watson tree and the size-biased tree are isomorphic up to level M in terms of the maximum probability that the Galton–Watson tree belongs to a specific isomorphism class. This gives another corollary, which we will need later on.

▶ **Corollary 6.** *Let $\mathcal{T}$ be a Galton–Watson tree and $\hat{\mathcal{T}}$ be the size-biased tree, assumed to be independent of $\mathcal{T}$. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}^{(M)}, \hat{\mathcal{T}}^{(M)} \text{ isomorphic and of height} \geq M) \leq c^M.$$

We can obtain similar bounds on the probability that two conditioned Galton–Watson trees are isomorphic up to level M. We start by extending Lemma 4 to the conditioned case.

▶ **Lemma 7.** *Let $\mathcal{T}_n$ be a conditioned Galton–Watson tree of size n. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}_n^{(M)} \text{ belongs to } C) = O\left(n^{\frac{5}{2}} c^M\right),$$

uniformly for all isomorphism classes $C \in \mathcal{C}_n^M$.

The proof uses breadth-first exploration and the cycle lemma, a standard trick in the field, and is deferred to the appendix. Furthermore, using calculations similar to (3), we obtain the following corollary.

▶ **Corollary 8.** *Let $\mathcal{T}_{n_1}, \mathcal{T}_{n_2}$ be two independent conditioned Galton–Watson trees. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}_{n_1}^{(M)}, \mathcal{T}_{n_2}^{(M)} \text{ isomorphic and of height} \geq M) = O\left(n^{\frac{5}{2}} c^M\right).$$

We are now ready to apply the central limit theorem for additive functionals.

2.2 Applying the CLT for almost local additive functionals

By Stirling's approximation, we can bound $f(T) \leq \log \deg(T)! = O(\deg(T)^{1+\epsilon})$ for any $\epsilon > 0$ so that the functional satisfies the degree condition of Theorem 3 with $\alpha = 2$. For the expectations, there are two conditions to check, one for the size-biased Galton–Watson tree and one for the conditioned Galton–Watson tree, and in each case the difference inside the

expectation can only be non-zero if (at least) two branches are isomorphic up to level M but non-isomorphic when we take all levels into account. We can therefore reduce the problem to studying trees that are isomorphic up to the M-th level.

We note that if l root branches are isomorphic up to level M, this contributes at most $\log(l!) \leq \binom{l}{2}$ to the difference inside the expectation. Therefore, the contribution of a random tree can be bounded by the sum of indicators

$$\sum_{T_i, T_j \text{ root branches}} I(T_i^{(M)}, T_j^{(M)} \text{ isomorphic and of height} \geq M),$$

where we sum over distinct branches. We can thus bound the expectation for the conditioned Galton–Watson tree in the following way.

$$\begin{aligned}
&\mathbb{E}\,|f(\mathcal{T}_n) - f(\mathcal{T}_n^{(M)})| \leq \mathbb{E}\left(\sum_{\substack{\mathcal{T}_i, \mathcal{T}_j \\ \text{root branches}}} I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ are iso. with height} \geq M)\right) \\
&= \sum_{k\geq 2} \mathbb{P}(\deg(\mathcal{T}_n) = k) \sum_{n_1, n_2 \geq 1} \mathbb{P}(|\mathcal{T}_i| = n_1 | \deg(\mathcal{T}_n) = k)\, \mathbb{P}(|\mathcal{T}_j| = n_2 | \deg(\mathcal{T}_n) = k) \\
&\qquad \cdot \binom{k}{2} \mathbb{E}\left(I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ are iso. with height} \geq M) \middle| |\mathcal{T}_i| = n_1, |\mathcal{T}_j| = n_2\right) \\
&= O\left(\sum_{k\geq 2} \mathbb{P}(\deg(\mathcal{T}_n) = k) \binom{k}{2} n^{\frac{5}{2}} c^M\right) = O\left(n^{\frac{5}{2}} c^M \sum_{k\geq 2} k\, \mathbb{P}(\xi = k) \binom{k}{2}\right) = O(n^{\frac{5}{2}} c^M)
\end{aligned}$$

where we use the law of total expectation, the fact that $\mathbb{P}(\deg(\mathcal{T}_n) = k) \leq ck\,\mathbb{P}(\xi = k)$ for all k and n, where c is constant [7, (2.7)], and the assumption on moments of the offspring distribution.

The difference $|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}(f(\hat{\mathcal{T}}^{(N)})|\hat{\mathcal{T}}^{(M)})|$ must also be zero unless some branches are isomorphic up to level M, and reasoning similar to above lets us bound its expectation in the following way.

$$\begin{aligned}
&\mathbb{E}\left|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)})|\hat{\mathcal{T}}^{(M)}\right)\right| \\
&= \sum_{k\geq 2} kP(\xi = k) \cdot \left(\mathbb{E}\left(\sum_{\substack{\mathcal{T}_i, \mathcal{T}_j \text{ non-special} \\ \text{root branches}}} I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ are iso. with height} \geq M)\right)\right. \\
&\qquad\qquad \left. + \mathbb{E}\left(\sum_{\substack{\mathcal{T} \text{ non-special root branch} \\ \hat{\mathcal{T}} \text{ special root branch}}} I(\mathcal{T}^{(M)}, \hat{\mathcal{T}}^{(M)} \text{ are iso. with height} \geq M)\right)\right).
\end{aligned}$$

Furthermore, this is equal to

$$\begin{aligned}
&\sum_{k\geq 3} kP(\xi = k) \binom{k-1}{2} \mathbb{P}(\mathcal{T}_1^{(M)}, \mathcal{T}_2^{(M)} \text{ isomorphic and of height} \geq M) \\
&\qquad + \sum_{k\geq 2} kP(\xi = k)(k-1)\,\mathbb{P}(\mathcal{T}^{(M)}, \hat{\mathcal{T}}^{(M)} \text{ isomorphic and of height} \geq M) = O(c^M),
\end{aligned}$$

by Corollaries 5 and 6 (the constant c is the same for both of these corollaries since they both rely on Lemma 4) as well as assumptions on moments of the offspring distribution.

We now set $p_M = Kc_1^M$, for $c < c_1 < 1$ and some suitable constant K, as well as $M_n = A \log n$, for some positive constant A that is large enough to make $n^{5/2} c^{M_n} \leq c_1^{M_n}$ for all n and $A \log c_1 < -3/2$. Then, the expectations mentioned in Theorem 3 are bounded by p_M and p_{M_n}, respectively. Furthermore, the sequence a_n goes to 0 and satisfies $\sum a_n/n < \infty$. Thus, we can apply Theorem 3 to show that $\log |\operatorname{Aut} \mathcal{T}_n|$ is asymptotically normal, which completes the proof of Theorem 1.

3 The automorphism group of Pólya trees

Since Theorem 3 is not available for Pólya trees, we want to prove asymptotic normality by using generating functions and singularity analysis. Thus, we define the generating function of $F(\mathcal{P}_n) = \log |\operatorname{Aut} \mathcal{P}_n|$ to be

$$P(x,t) = \sum_{T \in \mathcal{P}} e^{t \log |\operatorname{Aut} T|} x^{|T|}. \tag{6}$$

Note that $P(x,0) = P(x)$, as defined in the introduction. We will use $\rho = 0.33832\ldots$ to denote the dominant singularity of the generating function and recall that $P(\rho) = 1$ (see [3, Remark 3.9]). We now let $\mathcal{B}(T)$ denote the set of root branches of a particular tree, and $\mathcal{B}_I(T)$ denote the set of unique root branches up to isomorphism. Furthermore, we let $\operatorname{mult}(B)$ be the number of occurrences as root branches of a particular tree B. Observe that for Pólya trees there is exactly one tree in every isomorphism class so it will not be necessary to introduce separate notation for such classes.

By considering only the terms corresponding to the star on n vertices, for each n, we obtain

$$\sum_n (n-1)!^t x^n.$$

This is not analytic for any choice of $t > 0$ and, thus, neither is the original generating function. This is the main obstacle in proving asymptotic normality. To circumvent this problem, we will introduce a cut-off, ignoring the contribution of highly symmetric vertices. This is similar to the proof, in [14], of Theorem 3, but there the cut-off is in terms of the size of the tree instead of symmetric vertices. We can then use the following approximation result to extend the result from the cut-off random variables to the full additive functional.

▶ **Lemma 9.** *Let $(X_n)_{n\geq 1}$ and $(W_{n,N})_{n,N\geq 1}$ be sequences of centered random variables. If we have*

1. *$W_{n,N} \xrightarrow{d}_n W_N$ and $W_N \xrightarrow{d} W$ for some random variables $W, W_1, W_2, \ldots$, and*
2. *$\operatorname{Var}(X_n - W_{n,N}) \underset{N}{\longrightarrow} 0$ uniformly in n,*

then $X_n \xrightarrow{d} W$.

This result follows e.g. from [9], Theorem 4.28. We will apply Lemma 9 to variables X_n defined by

$$\frac{\log |\operatorname{Aut} \mathcal{P}_n| - \mathbb{E}(\log |\operatorname{Aut} \mathcal{P}_n|)}{\sqrt{n}},$$

and $W_{n,N}$ being the, similarly normalized, random variable for the additive functional $F^{\leq N}(T)$, defined by having the toll function:

$$f^{\leq N}(T) = \sum_{B \in \mathcal{B}_I(T)} I(\operatorname{mult}(B) \leq N) \log(\operatorname{mult}(B)!).$$

We note that $F(T) - F^{\leq N}(T) = F^{>N}(T)$ for an additive functional defined by

$$f^{>N}(T) = \sum_{B \in \mathcal{B}_I(T)} I(\text{mult}(B) > N) \log(\text{mult}(B)!),$$

so that we will, in fact, be interested in $\text{Var}(F^{>N}(T_n))$ for the second condition of Lemma 9. By straightforward modifications of (6), we can define generating functions $P^{\leq N}(x,t)$ and $P^{>N}(x,t)$ for the corresponding cut-off functionals.

3.1 Mean and variance

We can now derive moments for the additive functionals $F, F^{\leq N}, F^{>N}$ with the help of generating functions and singularity analysis. The calculations are essentially the same in all cases so, to simplify the exposition, we perform them only for F and indicate in the end how the results differ.

Due to general principles of generating functions, studying the mean and variance corresponds to studying $P_t(x,0)$ and $P_{tt}(x,0)$. According to calculations for general additive functionals from [15], we can write

$$P_t(x,0) = xP_x(x,0)\frac{\sum_T f(T)x^{|T|} + P(x,0)\sum_{k\geq 2} P_t(x^k,0)}{P(x,0)(1+\sum_{k\geq 2} x^k P_x(x^k,0))}, \tag{7}$$

$$P_{tt}(x,0) = \frac{xP_x(x,0)}{P(x,0)(1+\sum_{k\geq 2} x^k P_x(x^k,0))}\Bigg(P(x,0)\Big(P_t(x,0) + \sum_{k\geq 2}\frac{P_t(x^k,0)}{k}\Big)^2 \tag{8}$$

$$+ P(x,0)\sum_{k\geq 2}\frac{P_{tt}(x^k,0)}{k} + \sum_T x^{|T|} f(T)(2F(T) - f(T))\Bigg), \tag{9}$$

for the first and second derivative. To perform singularity analysis, we must first find singular expansions for these expressions. To this end, we study the sums involved in them separately.

The derivatives involving higher powers of x are analytic in a larger region than $P(x,0)$, since $\rho < 1$ so that $\rho^m < \rho$ for $m \geq 2$. Now, note that we can rewrite

$$2F(T) - f(T) = 2\sum_{B\in\mathcal{B}(T)} F(B) + f(T),$$

so that we can study $\sum x^{|T|} f(T) \sum F(B)$ and $\sum x^{|T|} f(T)^2$, as well as $\sum x^{|T|} f(T)$. It turns out that we can factor each of these expressions as $P(x,0)$ times some function that is analytic in a larger radius than ρ. For the sum in the expression for the mean, we have

$$\sum_T x^{|T|} f(T) = \sum_T x^{|T|} \sum_{B\in\mathcal{B}_I(T)} \log(\text{mult}(B)!)$$

$$= \sum_{B\in\mathcal{P}}\sum_{m=1}^{\infty}\log(m!)\sum_{T:\text{mult}(B)=m} x^{|T|} = \sum_{B\in\mathcal{P}}\sum_{m=1}^{\infty}\log(m!)x^{m|B|}(P(x,0) - x^{|B|}P(x,0))$$

$$= P(x)\sum_{B\in\mathcal{P}}\sum_{m=1}^{\infty}\log(m!)x^{m|B|}(1-x^{|B|}) = P(x)\sum_B\sum_{m=2}^{\infty}\log(m)x^{m|B|},$$

where we note that $P(x,0) - x^{|B|}P(x,0)$ equals the generating function for Pólya trees without B as a root branch. For real positive x with $x \leq 1-\epsilon$ for fixed $\epsilon > 0$, we can bound

$$\sum_B\sum_{m=2}^{\infty}\log(m)x^{m|B|} = O(\sum_B x^{2|B|}).$$

The extra power of 2 means that the sum converges for $x < \sqrt{\rho}$, so by the Weierstrass M-test, we have analyticity in a larger region than for the original generating function $P(x)$.

For the sum involving $\sum F(B)$, we have

$$\sum_T x^{|T|} \left(\sum_{B \in \mathcal{B}_I(T)} \log(\mathrm{mult}(B)!) \right) \left(\sum_{B \in \mathcal{B}_I(T)} \mathrm{mult}(B) F(B) \right)$$
$$= \sum_{B \in \mathcal{P}} F(B) \sum_{m=1}^{\infty} m \log(m!) \sum_{T : \mathrm{mult}(B) = m} x^{|T|}$$
$$+ \sum_{\substack{B_1, B_2 \in \mathcal{P}: \\ B_1 \neq B_2}} \sum_{m_1, m_2 \geq 1} \log(m_1!) m_2 F(B_2) \sum_{\substack{T : \mathrm{mult}(B_1) = m_1 \\ \mathrm{mult}(B_2) = m_2}} x^{|T|}.$$

Using the fact that $\sum_B F(B) x^{m|B|} = P_t(x^m, 0)$ and performing calculations similar to above, the first sum can be seen to be

$$P(x,0) \sum_{m=2}^{\infty} \log(m! m^{m-1}) P_t(x^m, 0),$$

where the sum is analytic in a larger region than the original function. To deal with the other sum, we first rewrite

$$\sum_{\substack{T : \mathrm{mult}(B_1) = m_1 \\ \mathrm{mult}(B_2) = m_2}} x^{|T|} = P(x,0) x^{m_1 |B_1|} (1 - x^{|B_1|}) x^{m_2 |B_2|} (1 - x^{|B_2|}).$$

Then, we note that

$$\sum_{\substack{B_1 : \\ B_1 \neq B_2}} F(B_1) \sum_{m_1=1}^{\infty} m_1 x^{m_1 |B_1|} (1 - x^{|B_1|})$$
$$= \sum_{m_1=1}^{\infty} \sum_{\substack{B_1 : \\ B_1 \neq B_2}} F(B_1) x^{m_1 |B_1|} = \sum_{j=1}^{\infty} P_t(x^j, 0) - \sum_{j=1}^{\infty} F(B_2) x^{j|B_2|}.$$

These observations let us rewrite the larger sum as

$$P(x,0) \left(\sum_{j=1}^{\infty} P_t(x^j, 0) \right) \sum_{m=1}^{\infty} \log(m!) x^{m|B|} (1 - x^{|B|})$$
$$- P(x,0) \sum_B F(B) \sum_{m=1}^{\infty} \log(m!) x^{m|B|} (1 - x^{|B|}) \sum_{j=1}^{\infty} x^{j|B|}.$$

The first of these two sums can now be dealt with using calculations identical to those performed earlier, and further simplifications for the second sum let us rewrite the whole expression as

$$P(x,0) \left(\left(P_t(x,0) + \sum_{m=2}^{\infty} P_t(x^m, 0) \right) \sum_B \sum_{m=2}^{\infty} \log(m) x^{m|B|} - \sum_m \log(m!) P_t(x^{m+1}, 0) \right).$$

The sum $\sum x^{|T|} f(T)^2$ can be dealt with using similar techniques and we conclude that we can rewrite (7) as

$$P_t(x,0) = xP_x(x,0)\frac{H(x)+\sum_{k\geq 2} P_t(x^k,0)}{(1+\sum_{k\geq 2} x^k P_x(x^k,0))},$$

$$P_{tt}(x,0) = \frac{xP_x(x,0)}{(1+\sum_{k\geq 2} x^k P_x(x^k,0))}\Bigg(\Big(P_t(x,0)+\sum_{k\geq 2}\frac{P_t(x^k,0)}{k}\Big)^2 + \sum_{k\geq 2}\frac{P_{tt}(x^k,0)}{k} + 2(P_t(x,0)H(x)+K(x))+L(x)\Bigg), \tag{10}$$

for functions $H(x)$, $K(x)$ and $L(x)$ that are analytic in a larger region than $P(x,0)$. This puts us in a situation where we can perform singularity analysis to find the moments. Numerical computations yield $\mu = 0.1373423\ldots$ and $\sigma^2 = 0.1967696\ldots$.

If we instead consider $F^{\leq N}(T)$ or $F^{>N}(T)$, the extra indicator function introduced in the expression will carry trough the calculations and affect the indices in the sums. In the sums with index m above, we will sum up to $m = N$ in the first case and sum from $m = N+1$ to infinity in the second. In particular, for $F^{>N}(T)$, the corresponding analytic functions $H^{>N}(x)$,$K^{>N}(x)$ and $L^{>N}(x)$ will converge to zero within their region of convergence, if we let $N \to \infty$.

3.2 Asymptotic normality for $\log|\operatorname{Aut}\mathcal{P}_n|$

For Pólya trees we have the symbolic decomposition

$$\mathcal{P} = \bullet \times \bigotimes_{T\in\mathcal{P}} (\emptyset \uplus \{T\} \uplus \{T,T\} \uplus \ldots),$$

reflecting the fact that a Pólya tree consists of a tree and a multiset of branches. Taking automorphisms into account, this translates to

$$P(x,t) = x\prod_{T\in\mathcal{P}}\left(\sum_{n=0}^{\infty} x^{n|T|} n!^t |\operatorname{Aut} T|^{nt}\right),$$

by general principles for generating functions. For the case of the cut-off functional $F^{\leq N}(T)$, which is the case we will be interested in, we have

$$P^{\leq N}(x,t) = x\prod_{T\in\mathcal{P}}\left(\sum_{n=0}^{\infty} x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\right).$$

We can manipulate this as follows:

$$P^{\leq N}(x,t) = x\exp\left(\sum_{T\in\mathcal{P}}\log\left(\sum_{n=0}^{\infty} x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\right)\right)$$

$$= x\exp\left(\sum_{T\in\mathcal{P}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\left(\sum_{n=1}^{\infty} x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\right)^k\right)$$

$$= x\exp\Big(\sum_{T\in\mathcal{P}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\sum_{\substack{\lambda_1+\lambda_2\\+\ldots=k}}\binom{k}{\lambda_1,\lambda_2,\ldots}\prod_{n=1}^{\infty}\big(x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\big)^{\lambda_n}\Big). \tag{11}$$

We now write integer partitions as sequences $\lambda = (\lambda_1, \lambda_2, \ldots)$, where λ_i is the number of i's in the partition. The total number of summands is denoted by $|\lambda| = \lambda_1 + \lambda_2 + \ldots$, and we write $\lambda \vdash j$ to denote that λ is a partition of j, i.e. $j = \lambda_1 + 2\lambda_2 + 3\lambda_3 + \ldots$. We can now rearrange the terms in the exponent of (11) to get

$$\begin{aligned} &x \exp\left(\sum_{T \in \mathcal{P}} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{j=1}^{\infty} \sum_{\substack{\lambda_1+\lambda_2+\ldots=k \\ \lambda_1+2\lambda_2+\ldots=j}} \binom{k}{\lambda_1, \lambda_2, \ldots} x^{j|T|} e^{jtF^{\leq N}(T)} \prod_{n=1}^{\infty} n!^{\lambda_n t I(n \leq N)} \right) \\ &= x \exp\left(\sum_{j=1}^{\infty} \sum_{\lambda \vdash j} \frac{(-1)^{|\lambda|-1}}{|\lambda|} \binom{|\lambda|}{\lambda_1, \lambda_2, \ldots} \left(\prod_{n=1}^{N} n!^{\lambda_n t} \right) \sum_{T \in \mathcal{P}} x^{j|T|} e^{jtF^{\leq N}(T)} \right) \\ &= x \exp\left(\sum_{j=1}^{\infty} \sum_{\lambda \vdash j} \frac{(-1)^{|\lambda|-1}}{|\lambda|} \binom{|\lambda|}{\lambda_1, \lambda_2, \ldots} \left(\prod_{n=1}^{N} n!^{\lambda_n t} \right) P^{\leq N}(x^j, jt) \right). \end{aligned}$$

For convenience, we can define

$$c_N(j,t) = j \sum_{\lambda \vdash j} \frac{(-1)^{|\lambda|-1}}{|\lambda|} \binom{|\lambda|}{\lambda_1, \lambda_2, \ldots} \left(\prod_{n=1}^{N} n!^{\lambda_n t} \right),$$

and arrive at the functional equation

$$P^{\leq N}(x,t) = x \exp\left(P^{\leq N}(x,t) + \sum_{j=2}^{\infty} \frac{c_N(j,t)}{j} P^{\leq N}(x^j, jt) \right). \tag{12}$$

Note that $c_N(j,0) = 1$, so that we recover the functional equation (1) from the introduction if we set $t = 0$. We can make completely analogous calculations for $P(x,t)$ to obtain a functional equation for the original functional $\log |\operatorname{Aut} \mathcal{P}_n|$, as well, but recall that $P(x,t)$ is not analytic for $t > 0$.

As a crude upper bound, each of the n vertices contributes at most $\log N!$ to the total value of the additive functional. Therefore, we see that $F^{\leq N}(T) = O(n)$ and, if we restrict to $|t| < \delta$ for some suitable $\delta > 0$,

$$G(x,y,t) := x \exp\left(y + \sum_{j=2}^{\infty} \frac{c(j,t)}{j} T(x^j, jt) \right)$$

is analytic in a region containing $x = \rho$, $y = \tau$. Theorem 2.23 in [3] now gives asymptotic normality for $F^{\leq N}(T)$, i.e. $W_N \sim \mathrm{N}(0, \sigma_N^2)$ for some constant σ_N^2.

Note that

$$\operatorname{Var}(X_n - W_{n,N}) = \frac{\operatorname{Var}(F(\mathcal{P}_n) - F^{\leq N}(\mathcal{P}_n))}{n}.$$

Since $F(T) - F^{\leq N}(T) = F^{>N}(T)$, we want to show that $\operatorname{Var}(F^{>N}(T_n))/n \to 0$ when $N \to \infty$ which leads us to study $P_{tt}^{>N}(x,t)$. The reasoning from the last section shows that coefficients in Taylor expansions of $H^{>N}(x)$, $K^{>N}(x)$ and $L^{>N}(x)$ around $x = \rho$ go to zero as $N \to \infty$. By dominated convergence, the same is true for the expressions

$$\sum \frac{P_t(x^k, 0)}{k} \text{ and } \sum \frac{P_{tt}(x^k, 0)}{k},$$

since all terms of P_t and P_{tt} involve powers of $F^{>N}(T)$ and this goes to zero for any fixed tree as $N \to \infty$. By studying (10) (except with $P_{tt}^{>N}(x,t)$ instead of $P_{tt}(x,t)$) we see that all the coefficients in the singular expansion of $P_{tt}^{>N}(x,t)$ depend on these quantities. Therefore, the expansion must be of the type

$$a_N \left(1 - \frac{x}{\rho}\right)^{-3/2} + b_N \left(1 - \frac{x}{\rho}\right)^{-1} + c_N \left(1 - \frac{x}{\rho}\right)^{-1/2} + O_N(1),$$

where each coefficient, as well as the error, goes to zero with N.

Performing singularity analysis, where we also subtract $\mathbb{E}(F^{>N}(\mathcal{P}_n))^2$ to get the variance, and dividing by n, gives us that

$$\operatorname{Var}(X_n - W_{n,N}) = \gamma_N^2 + O_N\left(\frac{1}{n}\right),$$

for some constant γ_N that goes to 0 as $N \to \infty$. Moreover, the O-term is uniform in N. This implies that the variance of $X_n - W_{n,N}$ goes to zero, uniformly in n so that the approximation lemma applies. Thus, we can conclude asymptotic normality for $\log|\operatorname{Aut}\mathcal{P}_n|$ from the asymptotic normality of $F^{\leq N}(\mathcal{P}_n)$ and finish the proof.

References

1 László Babai. Automorphism groups, isomorphism, reconstruction. In *Handbook of combinatorics, Vol. 1, 2*, pages 1447–1540. Elsevier Sci. B. V., Amsterdam, 1995.

2 Miklós Bóna and Philippe Flajolet. Isomorphism and symmetries in random phylogenetic trees. *J. Appl. Probab.*, 46(4):1005–1019, 2009. doi:10.1239/jap/1261670685.

3 Michael Drmota. *Random trees*. SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability. doi:10.1007/978-3-211-75357-6.

4 Cecilia Holmgren and Svante Janson. Limit laws for functions of fringe trees for binary search trees and random recursive trees. *Electron. J. Probab.*, 20:no. 4, 51, 2015. doi:10.1214/EJP.v20-3627.

5 Cecilia Holmgren, Svante Janson, and Matas Šileikis. Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. *Electron. J. Combin.*, 24(2):Paper No. 2.51, 49, 2017. doi:10.37236/6374.

6 Mikhail Isaev, Angus Southwell, and Maksim Zhukovskii. Distribution of tree parameters by martingale approach, 2019. arXiv:1912.09838.

7 Svante Janson. Random cutting and records in deterministic and random trees. *Random Structures & Algorithms*, 29(2):139–179, 2006. doi:10.1002/rsa.20086.

8 Svante Janson. Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees. *Random Structures Algorithms*, 48(1):57–101, 2016. doi:10.1002/rsa.20568.

9 Olav Kallenberg. *Foundations of modern probability*. Springer, New York, 1997.

10 David Matthews. *Automorphisms of Random Recursive Trees*. PhD thesis, University of Southhampton, 2017. URL: https://eprints.soton.ac.uk/415900.

11 Kathleen A. McKeon. The expected number of symmetries in locally-restricted trees. I. In *Graph theory, combinatorics, and applications. Vol. 2 (Kalamazoo, MI, 1988)*, Wiley-Intersci. Publ., pages 849–860. Wiley, New York, 1991.

12 Kathleen A. McKeon. The expected number of symmetries in locally restricted trees. II. *Discrete Appl. Math.*, 66(3):245–253, 1996. doi:10.1016/0166-218X(94)00164-9.

13 Dimbinaina Ralaivaosaona and Stephan Wagner. A central limit theorem for additive functionals of increasing trees. *Combin. Probab. Comput.*, 28(4):618–637, 2019. doi:10.1017/s0963548318000585.

14 Dimbinaina Ralaivaosaona, Matas Šileikis, and Stephan Wagner. A central limit theorem for almost local additive tree functionals. *Algorithmica*, 82(3):642–679, 2020. doi:10.1007/s00453-019-00622-4.

15 Stephan Wagner. Central limit theorems for additive tree parameters with small toll functions. *Combinatorics, Probability and Computing*, 24(1):329–353, 2015. doi:10.1017/S0963548314000443.

16 Le Yu. *Automorphisms of random trees*. PhD thesis, Drexel University, 2012. URL: https://researchdiscovery.drexel.edu/esploro/outputs/doctoral/Automorphisms-of-random-trees/991014632289304721.

A Omitted proofs and calculations

A.1 Proof of Lemma 7

Order the offsprings $\xi_1, \xi_2, \dots$ of T_n in breadth-first order and consider the sums

$$S_m = \sum_{i=1}^{m} (\xi_i - 1) \quad \text{for } 1 \le m \le n.$$

In each step, $1 \le i \le m$, we are deleting 1 for the current vertex while adding the number of children it has. For a conditioned Galton–Watson tree of size n, we necessarily have

$$\begin{aligned} S_m &> -1 \quad \text{for } 1 \le m < n, \\ S_n &= -1, \end{aligned}$$

since we are adding 1 for all vertices except the root, but deleting 1 for all vertices including the root. Using this, we can formulate the probability we seek to bound in the following way.

$$\mathbb{P}(T_n^{(M)} \text{ belongs to } C) = \frac{\mathbb{P}(\{T' \text{ belongs to } C\} \cap \{S_1, S_2, \dots, S_{n-1} > -1, S_n = -1\})}{\mathbb{P}(S_1, S_2, \dots, S_{n-1} > -1, S_n = -1)},$$

where T' is a Galton–Watson tree with offsprings $\xi_1, \xi_2, \dots, \xi_k$, and k is the number of vertices of each tree in C excluding the last level (since we truncate at level M the number of children on this level is of no interest to us). Since the trees in C are isomorphic they will all have the same number of vertices.

Let l_M be the number of vertices at the last level of each tree in C (again, equal due to isomorphism). Then we have

$$\sum_{i=1}^{n} (\xi_i - 1) = \sum_{i=1}^{k} (\xi_i - 1) + \sum_{i=k+1}^{n} (\xi_i - 1) = l_M - 1 + \sum_{i=k+1}^{n} (\xi_i - 1).$$

By the conditions set on S_m, we draw the conclusion that

$$\begin{aligned} S'_m &:= \sum_{i=k+1}^{k+m} (\xi_i - 1) > -l_M \quad \text{for } 1 \le m < n - k, \\ S'_{n-k} &:= \sum_{i=k+1}^{n} (\xi_i - 1) = -l_M. \end{aligned}$$

By independence, we now have

$$\begin{aligned} &\frac{\mathbb{P}(\{T' \text{ belongs to } C\} \cap \{S_1, S_2, \dots, S_{n-1} > -1, S_n = -1\})}{\mathbb{P}(S_1, S_2, \dots, S_{n-1} > -1, S_n = -1)} \\ &\qquad = \frac{\mathbb{P}(T' \text{ belongs to } C)\, \mathbb{P}(S'_1, S'_2, \dots, S'_{n-k-1} > -l_M, S'_{n-k} = -l_M)}{\mathbb{P}(S_1, S_2, \dots, S_{n-1} > -1, S_n = -1)}, \end{aligned}$$

and using the cycle lemma we find that this equals

$$\frac{\frac{l_M}{n-k}\,\mathbb{P}(S'_{n-k}=-l_M)}{\frac{1}{n}\,\mathbb{P}(S_n=-1)}\,\mathbb{P}(T' \text{ belongs to } C).$$

The probability $\mathbb{P}(S'_{n-k}=-l_M)$ is bounded by 1, and S_n satisfies a local limit theorem. If we also bound $l_M \le n$ as well as $n-k \ge 1$ (k is the number of vertices up to level $M-1$, and by definition there must be at least one vertex at level M) and use Lemma 4 (note that $\mathcal{C}_{n,M}$ is a subset of $\mathcal{C}^M$), we arrive at

$$\mathbb{P}(T_n^{(M)} \text{ belongs to } C) = O\left(n^{\frac{5}{2}} c^M\right),$$

which is what we wanted to prove.

A.2 Mean and variance for labeled trees

In this appendix, we show how the constants μ and σ^2 in Theorem 1 can be computed for labeled trees with fairly good accuracy. To this end, we use the generating function approach from Section 3. Recall that the bivariate generating function

$$P(x,t) = \sum_{T\in\mathcal{P}} e^{t\log|\operatorname{Aut} T|} x^{|T|} = \sum_{T\in\mathcal{P}} |\operatorname{Aut} T|^t x^{|T|}$$

satisfies (letting $N\to\infty$ in (12))

$$P(x,t) = x\exp\left(\sum_{j=1}^{\infty}\frac{c(j,t)}{j}P(x^j,jt)\right)$$

with

$$c(j,t) = j\sum_{\lambda\vdash j}\frac{(-1)^{|\lambda|-1}}{|\lambda|}\binom{|\lambda|}{\lambda_1,\lambda_2,\ldots}\left(\prod_{n=1}^{\infty} n!^{\lambda_n t}\right).$$

We can rewrite this in terms of an analogously defined exponential generating function for rooted labeled trees. Set

$$R(x,t) = \sum_{T\in\mathcal{R}} |\operatorname{Aut} T|^t \frac{x^{|T|}}{|T|!},$$

the sum now being over the set $\mathcal{R}$ of all rooted labeled trees. Since the number of distinct ways to label a Pólya tree T is $|T|!/|\operatorname{Aut} T|$, we have the relation

$$R(x,t) = P(x,t-1),$$

so the functional equation for Pólya trees immediately translates to a functional equation for labeled trees:

$$R(x,t) = x\exp\left(\sum_{j=1}^{\infty}\frac{c(j,t-1)}{j}R(x^j,jt-j+1)\right). \tag{13}$$

When $t=0$, one verifies easily (compare the calculations below for the derivative with respect to t) that $c(j,-1)=0$ for $j>1$ and $c(1,-1)=1$, so the functional equation reduces to $R(x,0)=x\exp(R(x,0))$ as expected.

In order to determine the desired moments, we need to consider the derivatives with respect to t. To this end, note first that

$$\sum_{j\geq 0} y^j \sum_{\lambda\vdash j} \prod_{k\geq 1} \frac{x_k^{\lambda_k}}{\lambda_k! k!^{\lambda_k}} = \prod_{k\geq 1} \sum_{\lambda_k\geq 0} \frac{x_k^{\lambda_k} y^{k\lambda_k}}{\lambda_k! k!^{\lambda_k}} = \prod_{k\geq 1} \exp\Big(\frac{x_k y^k}{k!}\Big) = \exp\Big(\sum_{k\geq 1} \frac{x_k y^k}{k!}\Big).$$

Differentiating with respect to x_m and plugging in $x_1 = x_2 = \cdots = x$ yields

$$\sum_{j\geq 0} y^j \sum_{\lambda\vdash j} x^{|\lambda|-1} \lambda_m \prod_{k\geq 1} \frac{1}{\lambda_k! k!^{\lambda_k}} = \frac{y^m}{m!} \exp\Big(\sum_{k\geq 1} \frac{xy^k}{k!}\Big) = \frac{y^m}{m!} \exp(x(e^y-1)).$$

Consequently,

$$\sum_{\substack{\lambda\vdash j\\ |\lambda|=r}} \lambda_m \prod_{k\geq 1} \frac{1}{\lambda_k! k!^{\lambda_k}} = [x^{r-1}y^j] \frac{y^m}{m!} \exp(x(e^y-1)) = [y^{j-m}] \frac{(e^y-1)^{r-1}}{(r-1)!m!}.$$

By definition, we have

$$\frac{d}{dt}\frac{c(j,t)}{j} = \sum_{\lambda\vdash j} \frac{(-1)^{|\lambda|-1}}{|\lambda|} \binom{|\lambda|}{\lambda_1, \lambda_2, \ldots} \Big(\prod_{n=1}^{\infty} n!^{\lambda_n t}\Big) \sum_{m=1}^{\infty} \lambda_m \log(m!),$$

which therefore becomes

$$\begin{aligned}
\frac{d}{dt}\frac{c(j,t)}{j}\Big|_{t=-1} &= \sum_{r=1}^{\infty}\sum_{m=1}^{\infty} (-1)^{r-1}(r-1)! \sum_{\substack{\lambda\vdash j\\ |\lambda|=r}} \lambda_m \log(m!) \prod_{k\geq 1} \frac{1}{\lambda_k! k!^{\lambda_k}} \\
&= \sum_{r=1}^{\infty}\sum_{m=1}^{\infty} (-1)^{r-1}(r-1)! \log(m!) [y^{j-m}] \frac{(e^y-1)^{r-1}}{(r-1)!m!} \\
&= \sum_{m=1}^{\infty} \frac{\log(m!)}{m!} [y^{j-m}] e^{-y} = \sum_{m=1}^{j} \frac{\log(m!)}{m!} \frac{(-1)^{j-m}}{(j-m)!} \\
&= \frac{1}{j!} \sum_{m=1}^{j} (-1)^{j-m} \binom{j}{m} \log(m!) = \frac{1}{j!} \sum_{m=1}^{j} (-1)^{j-m-1} \binom{j-1}{m-1} \log(m).
\end{aligned}$$

Let us write $d(j)$ for this expression. Differentiating (13) with respect to t and setting $t = 0$, we get

$$\begin{aligned}
R_t(x,0) &= x \exp\left(\sum_{j=1}^{\infty} \frac{c(j,-1)}{j} R(x^j, 1-j)\right) \\
&\quad \times \sum_{j=1}^{\infty} \Big(c(j,-1) R_t(x^j, 1-j) + \frac{d}{dt}\frac{c(j,t)}{j}\Big|_{t=-1} R(x^j, 1-j)\Big) \\
&= R(x,0)\Big(R_t(x,0) + \sum_{j=1}^{\infty} d(j) R(x^j, 1-j)\Big).
\end{aligned}$$

This can be solved for $R_t(x,0)$:

$$R_t(x,0) = \frac{R(x,0)}{1-R(x,0)} \sum_{j=2}^{\infty} d(j) R(x^j, 1-j).$$

Here, we are using the fact that $d(1) = 0$. Now note that $d(j)$ rapidly goes to 0 and that the functions $R(x^j, 1-j)$ are all analytic in a larger region than $R(x, 0)$. Therefore, we can directly apply singularity analysis, based on the well-known singular expansion

$$R(x, 0) = 1 - \sqrt{2(1 - ex)} + \cdots$$

of $R(x, 0)$ at its singularity $\frac{1}{e}$, which yields

$$R_t(x, 0) \sim \frac{1}{\sqrt{2(1 - ex)}} \sum_{j=2}^{\infty} d(j) R(e^{-j}, 1 - j).$$

The infinite series converges rapidly, allowing for a fairly accurate numerical computation. The mean constant μ in this special case is found to be $\mu = 0.0522901\ldots$, and similar calculations for the second derivative yield the variance constant $\sigma^2 = 0.0394984\ldots$.

The Number of Sources and Isolated Vertices in Random Directed Acyclic Graphs

Dimbinaina Ralaivaosaona
University of Stellenbosch, South Africa

Abstract

For a positive integer n and a real number $p \in (0,1)$, a random directed acyclic digraph $\mathbb{D}_{ac}(n,p)$ is obtained from the binomial random digraph model $\mathbb{D}(n,p)$ conditioned to be acyclic, i.e., directed cycles are forbidden. In the binomial random digraph model $\mathbb{D}(n,p)$, every possible directed edge (excluding loops) occurs independently with probability p. Sources and sinks are among the most natural characteristics of directed acyclic graphs. We investigate the distribution of the number of sources in $\mathbb{D}_{ac}(n,p)$ when p is of the form λ/n, where λ is a fixed positive constant. Because of symmetry, the number of sinks will have the same distribution as the number of sources. Our main motivation is to understand how this distribution changes as we pass through the critical point $p = 1/n$. Since we are in the sparse regime, it makes sense to include the number of isolated vertices as well. In a directed graph an isolated vertex can be regarded as a vertex that is both a source and a sink. We prove asymptotic normality for each of these parameters when $p = \lambda/n$. Our method is based on the analysis of a multivariate generating function from a work of Gessel.

2012 ACM Subject Classification Mathematics of computing → Random graphs; Mathematics of computing → Generating functions

Keywords and phrases Directed acyclic graph, generating function, central limit theorem

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.17

Funding This work is based on the research supported in part by the National Research Foundation of South Africa (grant number: 129413).

Acknowledgements I want to thank the anonymous referees for their helpful comments which helped to improve the presentation of this paper.

1 Introduction

For a positive integer n we consider directed graphs (digraphs) on the vertex set $[n] = \{1, 2, \ldots, n\}$ where loops and multi-edges are forbidden. In a digraph, a vertex v is called a source if it has an in-degree zero and a sink if it has an out-degree zero. If we only consider directed acyclic graphs (DAGs), it is well known that a non-empty acyclic digraph has at least one source and one sink. So, we would like to investigate the distributions of the number of these vertices in random DAGs.

The distribution of the number of isolated vertices and its generalisation, the number of vertices of a given degree, in the random (undirected) graphs are well-covered topics in the literature, see for example [1, 2, 6, 18]. It makes sense to extend these results to other graph-like structures. Investigating the number of isolated vertices, sources, and sinks should be the starting point for the case of random DAGs.

The model that we consider in this paper is constructed in the following way: for $p \in (0,1)$, we first consider the binomial random digraph model $\mathbb{D}(n,p)$, where each of the $n(n-1)$ possible directed edges occurs independently with probability p. Then, the random acyclic digraph $\mathbb{D}_{ac}(n,p)$ is simply $\mathbb{D}(n,p)$ conditioned to be acyclic. Due to limited space, in this paper, we restrict ourselves to the sparse case where $p = \lambda/n$, for which $\lambda > 0$ is fixed. It is known that the model $\mathbb{D}(n,p)$ exhibits a phase transition around the critical point $p = 1/n$. This phase transition has also been analysed in the literature, see for example [7, 10, 11].

For $p = \frac{\lambda}{n}$ an asymptotic formula for the probability that $\mathbb{D}(n,p)$ is acyclic is given in [3] following the approach in [14]: if $\lambda > 0$ is a constant, then

$$\mathbb{P}(\mathbb{D}(n,p) \text{ acyclic}) \sim \begin{cases} (1-\lambda)e^{\lambda} & \text{if } \lambda < 1 \\ C_1 n^{-1/3} & \text{if } \lambda = 1 \\ C_2(\lambda) n^{-1/3} e^{-a(\lambda)n + b(\lambda)n^{1/3}} & \text{if } \lambda > 1, \end{cases} \tag{1}$$

where C_1 is a positive constant, and $C_2(\lambda)$, $a(\lambda)$, and $b(\lambda)$ are positive numbers depending only on $\lambda \geq 1$. These terms are explicitly defined in [3, Sec. 6].

There are a few different random digraph models. For instance, the model considered in [14] is obtained from the random binomial undirected graph $G(n, 2p)$ where a direction is given to each existing edge independently with probability $\frac{1}{2}$. Although, it would also be possible to study such a model with our method, the model in the present paper, when conditioned to be a DAG, gives a more natural generalisation to the uniform random DAG model where a digraph is chosen uniformly at random from the set of all DAGs on the vertex set $[n]$. With our notation the latter model is equivalent to $\mathbb{D}_{ac}(n, \frac{1}{2})$.

The number of sources in the random acyclic digraph model $\mathbb{D}_{ac}(n, \frac{1}{2})$ was already studied by Liskovets [9]. It was shown that the number of sources in $\mathbb{D}_{ac}(n, \frac{1}{2})$ has a discrete limiting distribution as $n \to \infty$. More precisely, if denote by $p(n,k)$ the probability that a uniform random DAG on $[n]$ has exactly k sources, then as $n \to \infty$ we have

$$p(n,k) \sim \frac{\varrho^k \phi(2^{-k}\varrho)}{k! 2^{\binom{k}{2}}}, \quad \text{where} \quad \phi(x) = \sum_{n=0}^{\infty} \frac{(-x)^n}{n! 2^{\binom{n}{2}}}, \tag{2}$$

and $\varrho \approx 1.4880785\ldots$ is the smallest positive solution of the equation $\phi(x) = 0$. The function $\phi(x)$ and its zero ϱ already appeared in earlier results on the enumeration of DAGs, see the work of Robinson [15], Liskovets [8], and Stanley [16]. Returning to the structure of the uniform random DAG $\mathbb{D}_{ac}(n, \frac{1}{2})$, McKay [12] showed that the height is asymptotically normally distributed with mean and variance asymptotically equal to Cn and $C'n$ respectively, where $C \approx 0.764334$ and $C' \approx 0.145210$. It is reasonable to expect that similar results would hold for the number of sources and the height of $\mathbb{D}_{ac}(n,p)$ for fixed $p \in (0,1)$. However, it is not clear how these parameters behave when p tends to zero. Let us now state our result for the number of sources when p is of the form λ/n.

▶ **Theorem 1.** *Let $S(D)$ denote the number of sources in an acyclic digraph D. Define*

$$\mu(\lambda) = \begin{cases} e^{-\lambda} & \text{if } \lambda < 1 \\ \dfrac{1}{e\lambda} & \text{if } \lambda \geq 1, \end{cases} \quad \text{and} \quad \sigma^2(\lambda) = \begin{cases} e^{-\lambda}(1-e^{-\lambda}) & \text{if } \lambda < 1 \\ \dfrac{e-1}{e^2\lambda} & \text{if } \lambda \geq 1. \end{cases}$$

Then, for a fixed $\lambda > 0$, the expectation of the number of sources in a random acyclic digraph $\mathbb{D}_{ac}(n, \lambda/n)$ satisfies the asymptotic estimate $\mathbb{E}(S(\mathbb{D}_{ac}(n,\lambda/n))) \sim \mu(\lambda)n$ as $n \to \infty$. Moreover, we have

$$\frac{S(\mathbb{D}_{ac}(n,\lambda/n)) - \mu(\lambda)n}{\sqrt{\sigma^2(\lambda)n}} \xrightarrow{d} \mathcal{N}(0,1).$$

Since we are interested in the structure of $\mathbb{D}_{ac}(n,p)$ in the sparse regime, it makes sense to look at the number of isolated vertices. We obtain the following theorem.

▶ **Theorem 2.** *Let $I(D)$ denote the number of isolated vertices in an acyclic digraph D. Define*

$$\mu^*(\lambda) = \begin{cases} e^{-2\lambda} & \text{if } \lambda < 1, \\ \frac{e^{-(\lambda+1)}}{\lambda} & \text{if } \lambda \geq 1, \end{cases}$$

and

$$\sigma^*(\lambda)^2 = \begin{cases} e^{-2\lambda}(1 + (2\lambda - 1)e^{-2\lambda}) & \text{if } \lambda < 1 \\ \lambda^{-1}e^{-(\lambda+1)}(1 + e^{-(\lambda+1)}) & \text{if } \lambda \geq 1. \end{cases}$$

Then, for a fixed $\lambda > 0$, the expectation of the number of isolated vertices in a random acyclic digraph $\mathbb{D}_{ac}(n, \lambda/n)$ satisfies the asymptotic estimate $\mathbb{E}(I(\mathbb{D}_{ac}(n, \lambda/n))) \sim \mu^*(\lambda)n$ as $n \to \infty$. Moreover, we have

$$\frac{I(\mathbb{D}_{ac}(n, \lambda/n)) - \mu^*(\lambda)n}{\sqrt{\sigma^*(\lambda)^2 n}} \xrightarrow{d} \mathcal{N}(0, 1).$$

It is not difficult to show that main terms in asymptotic expansions of the expectations $\mathbb{E}(S(\mathbb{D}(n, \lambda/n)))$ and $\mathbb{E}(I(\mathbb{D}(n, \lambda/n)))$, for the unconditioned random digraph $\mathbb{D}(n, \lambda/n)$, are $e^{-\lambda}n$ and $e^{-2\lambda}n$, respectively for any fixed $\lambda \geq 0$. So, it appears that conditioning on the event that the random digraph is acyclic does not affect these expectations for $\lambda \in (0, 1]$, at least asymptotically. This is, however, not surprising because we know as we see in (1) that $\mathbb{D}(n, \lambda/n)$ is acyclic with positive probability for $\lambda < 1$, and we expect that the random variables $S(\mathbb{D}(n, \lambda/n))$ and $I(\mathbb{D}(n, \lambda/n))$ are concentrated around their expectations. For $\lambda > 1$ these expectations are higher in the random acyclic digraph model. It would be interesting to know what happens when we allow $\lambda \to \infty$ as $n \to \infty$. Extending the above results to such a case seems to be possible but it requires more work; it would certainly be too long for this extended abstract.

Throughout this paper, we adopt the notations and abbreviations of [14] since most of the asymptotic analysis that we need to prove Theorem 1 and Theorem 2 are based on the analytic method developed in that paper. This paper is organised as follows: in Section 2 we present and prove a result of Gessel [4] about a multivariate generating function that includes the number of sources, sinks, and isolated vertices. Section 3 consists of collections of asymptotic results from [14] and some of their consequences. The proofs of Theorem 1 and Theorem 2 are presented in Section 4 and Section 5 respectively.

2 Generating functions

We use the so-called *graphic generating function* for the enumeration of acyclic digraphs. If we denote by $e(D)$ the number of (directed) edges in a digraph D, then we define

$$A_n(y) = \sum_D y^{e(D)},$$

where the sum is taken over all acyclic digraphs on $[n]$. The corresponding graphic generating function is

$$A(x, y) = \sum_{n=0}^{\infty} \frac{A_n(y)x^n}{n!(1 + y)^{\binom{n}{2}}}.$$

It turns out that this bivariate generating can be written in the following way:

$$A(x,y)=\frac{1}{\phi(x,y)},\quad \text{where}\quad \phi(x,y)=\sum_{n=0}^{\infty}\frac{(-x)^n}{n!(1+y)^{\binom{n}{2}}}, \tag{3}$$

see [15]. Observe, that for $y>0$, the power series in the definition of $\phi(x,y)$ converges for $x\in\mathbb{C}$. Hence, $\phi(x,y)$ can be regarded as an entire function of x for any $y>0$.

To include counts on the number of sources, sinks, and isolated vertices, we have to define two related parameters. For an acyclic digraph D, let $S_o(D)$ and $S_i(D)$, respectively, be the number of sources in D that are not isolated vertices and the number of sinks that are not isolated vertices. Furthermore, let $I(D)$ be the number of isolated vertices of D. Then, consider the generating function

$$A_n(y,u_1,u_2,u_3)=\sum_{D} y^{e(D)}u_1^{S_o(D)}u_2^{S_i(D)}u_3^{I(D)},$$

where the sum is taken over all acyclic digraphs on the vertex set $[n]$. The corresponding graphic generating function is given by

$$A(x,y,u_1,u_2,u_3)=\sum_{n=0}^{\infty}\frac{x^n}{n!(1+y)^{\binom{n}{2}}}A_n(y,u_1,u_2,u_3)x^n.$$

If we denote by $a_{n,m}(k_1,k_1,k_3)$ the number of acyclic digraphs D on the vertex set $[n]$ with $e(D)=m$, $S_o(D)=k_1$, $S_i(D)=k_2$ and $I(D)=k_3$, then the polynomial $A_n(y,u_1,u_2,u_3)$ can be written as follows:

$$A_n(y,u_1,u_2,u_3)=\sum_{m,k_1,k_2,k_3} a_{n,m}(k_1,k_2,k_3)u_1^{k_1}u_2^{k_2}u_3^{k_3}y^m. \tag{4}$$

The next lemma allows us to obtain the joint probability generating function for our parameters in terms of $A_n(y,u_1,u_2,u_3)$.

▶ **Lemma 3.** *We have*

$$\mathbb{E}\left(u_1^{S_o(\mathbb{D}_{ac}(n,p))}u_2^{S_i(\mathbb{D}_{ac}(n,p))}u_3^{I(\mathbb{D}_{ac}(n,p))}\right)=\frac{A_n\left(\frac{p}{1-p},u_1,u_2,u_3\right)}{A_n\left(\frac{p}{1-p}\right)}.$$

The proof of this lemma is straightforward, so we leave it as an exercise to the reader.

We need to express $A(x,y,u_1,u_2,u_3)$ in more a manageable form in order to obtain any useful estimate of $A_n(y,u_1,u_2,u_3)$ as $n\to\infty$. Fortunately for us, this was already done by Gessel in [4]. Since this result is our main ingredient and some part of the proof was omitted in [4], we give a full proof here for completeness. Before we begin let us state a general property of graphic generating functions. Given infinite sequences $(a_n^{(i)})_n$, for $i\in\{1,2,3\}$, if $f_i(x)$ denotes the graphic generating function associated with the sequence $(a_n^{(i)})_n$, that is

$$f_i(x)=\sum_{n=0}^{\infty}\frac{a_n^{(i)}}{n!(1+y)^{\binom{n}{2}}}x^n,$$

and if $c_n=n!(1+y)^{\binom{n}{2}}\times[x^n]\Big(f_1(x)f_2(x)f_3(x)\Big)$, then the three-term convolution formula is

$$c_n=\sum_{j+k+l=n}(1+y)^{jk+jl+kl}\binom{n}{k,l,j}a_j^{(1)}a_k^{(2)}a_l^{(3)}. \tag{5}$$

▶ **Lemma 4** (Theorem 2 in Gessel [4]). *We have*

$$\sum_{n=0}^{\infty} A_n(y,u_1,u_2,u_1+u_2-1)\frac{x^n}{(1+y)^{\binom{n}{2}}n!} = \frac{\phi((1-u_1)x,y)\phi((1-u_2)x,y)}{\phi(x,y)},$$

and

$$\sum_{n=0}^{\infty} A_n(y,u_1,u_2,u_3)\frac{x^n}{n!} = e^{(u_3-u_1-u_2+1)x}\sum_{n=0}^{\infty} A_n(y,u_1,u_2,u_1+u_2-1)\frac{x^n}{n!}.$$

Proof. Consider the set of quadruples (S_1,S_2,D,E) where S_1 and S_2 are disjoint subsets of $[n]$, D is an acyclic digraph on $[n]\setminus(S_1\cup S_2)$ and E a set of directed edges consisting only of edges from S_1 to $V(D)\cup S_2$, or from $V(D)$ to S_2. Such a quadruple can be interpreted as an acyclic digraph where every vertex in S_1 is a source, and every vertex in S_2 is a sink. If each quadruple is weighted by $y^{|E|+e(D)}u_1^{|S_1|}u_2^{|S_2|}$, then the total weight over all possible quadruples is

$$\sum_{j+k+l=n}(1+y)^{jk+jl+kl}\binom{n}{k,l,j}u_1^j u_2^k A_l(y).$$

On the other hand, consider the set of acyclic digraphs D' on $[n]$ with four distinguished subsets of $[n]$: the first is a subset of the *strictly* sources of D' (no isolated vertices), a subset of the strictly sinks, and a pair of disjoint subsets of the isolated vertices. There is a correspondence between such a set with the set of quadruples (S_1,S_2,D,E): the subset of the strictly sources and the first subset of the isolated vertices of D' form the set S_1, while the subset of the strictly sinks together with the second subset of the isolated vertices form the set S_2, and E is the set edges of D' that go from S_1 to S_2, S_1 to $[n]\setminus(S_1\cup S_2)$, or from $[n]\setminus(S_1\cup S_2)$ to S_2. By taking this correspondence into consideration when calculating the total weight formula above, we deduce that

$$A_n(y,1+u_1,1+u_2,1+u_1+u_2) = \sum_{j+k+l=n}(1+y)^{jk+jl+kl}\binom{n}{k,l,j}u_1^j u_2^k A_l(y).$$

This is a three-term convolution formula, as defined in (5). Therefore, we deduce that

$$A(x,y,1+u_1,1+u_2,1+u_1+u_2) = \phi(-u_1x)\phi(-u_2x)A(x,y).$$

The first equation in the lemma follows by using (3) and by shifting the variables u_1 and u_2 by -1.

For the second equation, by removing all isolated vertices, each acyclic digraph D on $[n]$ can be associated with a pair (S,D') where S is a subset of $[n]$ (the set of isolated vertices of D) and D' is an acyclic digraph on $[n]\setminus S$ which has no isolated vertices. Thus, we obtain

$$A_n(y,u_1,u_2,u_3) = \sum_{j=0}^{n}\binom{n}{j}u_3^j A_{n-j}(y,u_1,u_2,0).$$

This is the standard two-term convolution formula for the product of two exponential generating functions. So, we obtain

$$\sum_{n=0}^{\infty} A_n(y,u_1,u_2,u_3)\frac{x^n}{n!} = e^{u_3x}\sum_{n=0}^{\infty} A_n(y,u_1,u_2,0)\frac{x^n}{n!}.$$

In particular, we have

$$\sum_{n=0}^{\infty} A_n(y,u_1,u_2,u_1+u_2-1)\frac{x^n}{n!} = e^{(u_1+u_2-1)x}\sum_{n=0}^{\infty} A_n(y,u_1,u_2,0)\frac{x^n}{n!}.$$

Eliminating the series involving $A_n(y,u_1,u_2,0)$ gives the formula in the lemma. ◀

3 Asymptotic analysis

The results in this section are mainly drawn from [14] or are consequences of the results in [14]. So, let us first summarise the notations and abbreviations that we used in [14]. Throughout, y is a positive number that tends to zero as $n \to \infty$. In our case, in view of Lemma 3, $y = p/(1-p)$ where $p = \lambda/n$. Moreover, it is convenient to use the abbreviation $\alpha = \log(1+y)$ and $\beta = \sqrt{1+y}$. So, $y \sim \alpha$ as $n \to \infty$. In particular, $\alpha \to 0^+$ as $n \to \infty$. In addition, for a complex number x, we define $w = w(x) = W_0(-x\alpha\beta)$, where W_0 is the principal branch of the Lambert W function. It is useful to keep these notations and abbreviations in mind when reading the rest of this paper.

3.1 Estimates of $\phi(x,y)$

We begin by providing asymptotic estimates for $\phi(x,y)$ where $y \to 0^+$ and x is complex. We obtain the next lemma with a minor modification of a similar result in [14], the reader can also consult [3] which contains more details.

▶ **Lemma 5.** *The function $\phi(x,y)$ satisfies the following asymptotic formulas as $y \to 0^+$, both estimates are uniform in x:*

- *If $x = o(\alpha^{-1})$, then*

$$\phi(x,y) \sim e^{\frac{1}{2\alpha}(w^2+2w)}. \tag{6}$$

- *If $x = \mathcal{O}(\alpha^{-1})$ and $w = w(x)$ is bounded away from -1, then*

$$\phi(x,y) \sim 2^{5/6}\pi^{1/2}w^{-1/3}\alpha^{-1/6}\mathrm{Ai}(R)e^{\frac{2}{3}R^{3/2}+\frac{1}{2\alpha}(w^2+2w)}, \tag{7}$$

where $R = 2^{-2/3}(1+w)^2w^{-4/3}\alpha^{-2/3}$, and $\mathrm{Ai}(z)$ is the Airy function.

3.2 Coefficient extraction

We are going to need estimates of $A_m(y)$ for a certain range of values of m close to n. For our purposes, it suffices to consider m to be of the form $n + o(n^{2/3})$. By the Cauchy integral formula, we have

$$\frac{A_m(y)}{m!(1+y)^{\binom{m}{2}}} = \frac{1}{2\pi i}\oint_{|x|=\rho}\frac{1}{\phi(x,y)x^{m+1}}dx. \tag{8}$$

The value of $\rho > 0$ will be chosen depending on the sign of $\lambda - 1$.

▶ **Lemma 6.** *Let λ be a fixed value in the interval $(0,1)$. If $m = n + \mathcal{O}(n^{2/3-\epsilon})$ for a constant $\epsilon > 0$, then as $n \to \infty$,*

$$A_m(y) \sim m!(1+y)^{\binom{m}{2}}\ \frac{1-\alpha m}{\sqrt{2\pi m}}\rho^{-m}e^{-\frac{1}{2\alpha}(\alpha^2m^2-2\alpha m)},$$

where $\rho = \frac{m}{\beta}e^{-\alpha m}$.

Proof. With the abbreviation $w(x) = W_0(-x\alpha\beta)$, let

$$h(t) = -\frac{1}{2\alpha}\left(w^2(\rho e^{it}) + 2w(\rho e^{it})\right).$$

Hence, if $\alpha\rho$ is bounded away from e^{-1}, then by the second part of Lemma 5, the Equation (8) yields

$$\frac{A_m(y)}{m!(1+y)^{\binom{m}{2}}} = (1+o(1))\rho^{-m}\int_{-\pi}^{\pi} P(\alpha, w(\rho e^{it}))e^{h(t)-imt}\,dt,$$

where $P(\alpha, w(\rho e^{it})) = 2^{-11/6}\pi^{-3/2}w^{1/3}\alpha^{1/6}\mathrm{Ai}(R)^{-1}e^{-\frac{2}{3}R^{3/2}}$ and $R = 2^{-2/3}(1+w)^2w^{-4/3}\alpha^{-2/3}$. Since, in our case $\rho = \mathcal{O}(ne^{-\lambda})$, $w(\rho e^{it})$ is bounded away from -1. Hence, $|R| \to \infty$ as $\alpha \to 0^+$ for $t \in [-\pi, \pi]$. Recall the following well known asymptotic formula for the Airy function: for any $\varepsilon > 0$,

$$\mathrm{Ai}(z) \sim \frac{e^{-\frac{2}{3}z^{3/2}}}{2\sqrt{\pi}z^{1/4}} \quad \text{as } |z| \to \infty, \text{ and } |\mathrm{Arg}(z)| \leq \pi - \varepsilon, \tag{9}$$

see [13, (9.7.5)]. We can prove from this that $P(\alpha, w(\rho e^{it}))$ is bounded uniformly for $t \in [-\pi, \pi]$. In fact, skipping the details, the estimate of $P(\alpha, w(\rho e^{it}))$ simplifies to

$$P(\alpha, w(\rho e^{it})) \sim \frac{1}{2\pi}(1 + w(\rho e^{it}))^{1/2},$$

uniformly for $t \in [-\pi, \pi]$.

In order to apply the saddle point method, we need to evaluate the first few derivatives of h. We have

$$h'(t) = -\frac{i}{\alpha}w(\rho e^t), \quad h''(t) = \frac{1}{\alpha}\frac{w(\rho e^t)}{1+w(\rho e^t)}, \text{ and } \quad h^{(3)}(t) = \frac{i}{\alpha}\frac{w(\rho e^t)}{(1+w(\rho e^t))^3}.$$

The saddle point equation is $h'(0) - im = 0$ which is equivalent to $w(\rho) = -\alpha m$ or $\rho = \frac{m}{\beta}e^{-\alpha m}$. Moreover, $h''(0) = \frac{-m}{1-\alpha m}$ which is of order α^{-1} under our assumptions, and $h^{(3)}(t) = \mathcal{O}(\alpha^{-1})$. The standard saddle point method applies in this case: the integral on the right-hand side of is divided into two parts

$$\int_{-\alpha^{-c}}^{\alpha^c} P(\alpha, w(\rho e^{it}))e^{h(t)-imt}\,dt + \int_{|t|\geq\alpha^c} P(\alpha, w(\rho e^{it}))e^{h(t)-imt}dt$$

where c is a fixed number in the interval $(1/3, 1/2)$. When we apply the saddle point method, we call the first integral the local integral and the rest the tail. By Taylor expanding $h(t) - imt$ and $P(\alpha, w(\rho e^{it}))$ for $|t| \leq \alpha^c$, we deduce the following estimate for the local integral:

$$\int_{-\alpha^{-c}}^{\alpha^c} P(\alpha, w(\rho e^{it}))e^{h(t)-imt}\,dt = (1+o(1))P(\alpha, w(\rho))e^{h(0)}\int_{-\alpha^{-c}}^{\alpha^c} e^{\frac{1}{2}h''(0)t^2}dt. \tag{10}$$

The integral on the right-hand side can be approximated by a Gaussian integral, by extending its range to $(-\infty, \infty)$, with an error term smaller than any power of α. Hence, we deduce that

$$\int_{-\alpha^{-c}}^{\alpha^c} P(\alpha, w(\rho e^{it}))e^{h(t)-imt}\,dt = (1+o(1))P(\alpha, w(\rho))e^{h(0)}\sqrt{\frac{2\pi}{-h''(0)}}.$$

On the other hand, it is not difficult to show that $\mathrm{Re}(h(t) - h(0))$ is negative for $\alpha^c \le |t| \le \pi$ (for small enough α). In fact, we can show that there exists a positive constant $C > 0$ (independent of α and t) such that $\mathrm{Re}(h(t) - h(0)) \le -C\alpha^{2c-1}$ for $\alpha^c \le |t| \le \pi$. This is enough to prove that the tail integral is much smaller than the local integral, and therefore, it can be neglected. Thus, we we obtain

$$\frac{A_m(y)}{m!(1+y)^{\binom{m}{2}}} \sim P(\alpha, w(\rho))e^{h(0)}\sqrt{\frac{2\pi}{-h''(0)}}.$$

Expressing everything in terms of m gives us the estimate in the statement of the lemma. ◀

Next, we consider the critical case.

▶ **Lemma 7.** *If $\lambda = 1$ and $m = n + \mathcal{O}(n^{2/3-\epsilon})$ for some constant $\epsilon > 0$, then as $n \to \infty$ we have*

$$A_m(y) \sim m!(1+y)^{\binom{m}{2}} \frac{\alpha^{2/3}Ai(0)}{2\pi\phi(\rho, y)\rho^m} \int_{-\infty}^{\infty} \frac{1}{Ai(-2^{1/3}it)}\, dt,$$

where $\rho = \frac{1}{e\alpha\beta}$.

Proof. This is a direct application of [14, Lemma 9, Eq. (27)]. Since $p = \frac{1}{n}$, $y = \frac{p}{1-p}$ and $\alpha = \log(1+y)$. We can show that if $m = n + \mathcal{O}(n^{2/3-\epsilon})$, then $m = \alpha^{-1} + o(\alpha^{-2/3})$. So, [14, Lemma 9, Eq. (27)] applies with $b = 0$. ◀

Finally, for the super-critical case, we have the following result:

▶ **Lemma 8.** *If $\lambda > 1$ and $m = n + \mathcal{O}(n^{2/3-\epsilon})$ for some constant $\epsilon > 0$ then as $n \to \infty$ we have*

$$A_m(y) = -m!(1+y)^{\binom{m}{2}} \frac{1}{\varrho_1(y)^{m+1}\phi_x(\varrho_1(y), y)} + O\left(\frac{\alpha^{2/3}}{|\phi(\rho, y)|\rho^m}\right).$$

where $\varrho_1(y)$ is the smallest x-solution of the equation $\phi(x, y) = 0$, $\phi_x(\varrho_1(y), y)$ is $\partial_x\phi(x, y)|_{x=\varrho_1(y)}$, and ρ is of the form

$$\rho = \frac{1}{e}y^{-1} - \frac{b}{2^{1/3}e}\, y^{-1/3}.$$

The constant b can be any fixed number in the interval (a_2, a_1), where a_j is the zero of the Airy function $\mathrm{Ai}(z)$ that is j-th closest to 0.

Proof. Once again m is of the form $m = \alpha^{-1} + o(\alpha^{-2/3})$, so the argument in [14, Sec. 3.2.3] remains valid. Recall from [14, Theorem 1] that the j-th zero of $\phi(x, y)$ satisfies the asymptotic formula

$$\varrho_j(y) = \frac{1}{e}y^{-1} - \frac{a_j}{2^{1/3}e}\, y^{-1/3} - \frac{1}{6e} + O(y^{1/3}), \quad \text{as } y \to 0^+, \tag{11}$$

where a_j is the zero of the Airy function $\mathrm{Ai}(z)$ that is j-th closest to 0. So, the choice of ρ guarantees that the circle $|x| = \rho$ contains only one pole of $A(x, y)$. ◀

4 The number of sources

We are now ready to prove Theorem 1. This section consists entirely of the proof of Theorem 1. We shall begin with the estimate of the average.

4.1 Estimate of the average

When considering the number of sources, the corresponding generating function is $A(x, y, u, 1, u)$, where the variable u indicates the number of sources. By Lemma 4,

$$A(x, y, u, 1, u) = \frac{\phi((1-u)x, y)}{\phi(x, y)}.$$

Differentiating this once with respect to u, yields

$$\partial_u A(x, y, u, 1, u)|_{u=1} = \frac{x}{\phi(x, y)} = xA(x, y).$$

Using the formula in Lemma 3, the coefficient of x^n in $\partial_u A(x, y, u, 1, u)|_{u=1}$ is

$$\frac{\partial_u A_n(y, u, 1, u)|_{u=1}}{n!(1+y)^{\binom{n}{2}}} = \mathbb{E}(S(\mathbb{D}_{ac}(n, p)))A_n(y).$$

We know that the coefficient $[x^n](xA(x, y)) = \frac{A_{n-1}(y)}{n!(1+y)^{\binom{n}{2}}}$. Hence, we obtain the following exact formula for $\mathbb{E}(S(\mathbb{D}_{ac}(n, p)))$:

$$\mathbb{E}(S(\mathbb{D}_{ac}(n, p))) = n(1+y)^{n-1}\frac{A_{n-1}(y)}{A_n(y)}. \tag{12}$$

This formula has a simple combinatorial explanation; the term $(1+y)^{n-1}A_{n-1}(y)$ is the generating function for the acyclic digraphs on $[n]$ with one marked source. Lemmas 6–8 can then be used to estimate $A_{n-1}(y)$ and $A_n(y)$ for $y = \frac{p}{1-p} = \frac{\lambda}{n-\lambda}$, and the estimate of $\mathbb{E}(S(\mathbb{D}_{ac}(n, p)))$ in Theorem 1 follows easily. The calculations were done in SageMath [17].

4.2 Asymptotic normality

For the rest of this section we shall slightly abuse notation and simply abbreviate $A_n(y, u, 1, u)$ by $A_n(y, u)$. This should not create confusion as this notation does not appear anywhere else in the paper. To prove the central limit theorem, we need to estimate $A_n(y, u)$ when u is a complex number of the form $1 + \mathcal{O}(n^{-1/2})$. It is convenient to write $v = u - 1$. So, by definition, we have

$$\frac{A_n(y, u)}{n!(1+y)^{\binom{n}{2}}} = \frac{1}{2\pi i}\oint_{|x|=\rho}\frac{\phi(-vx, y)}{\phi(x, y)x^{m+1}}dx.$$

We can then apply the saddle point method to estimate the integral on the right-hand side just as we did in Section 3 but with the extra term $\phi(-vx, y)$ in the integrand. Again the three cases $\lambda < 1$, $\lambda = 1$, and $\lambda > 1$ must be separated. The result is given in the next lemma.

▶ **Lemma 9.** *We have*

$$\frac{A_n(y, u)}{A_n(y)} \sim e^{g(n,\lambda,v)} \quad \text{as} \quad n \to \infty$$

uniformly $v = \mathcal{O}(n^{-1/2})$, *where*

$$g(n, \lambda, v) = \begin{cases} ne^{-\lambda}(v - \frac{1}{2}e^{-\lambda}v^2) & \text{if} \quad \lambda < 1 \\ \frac{n}{e\lambda}(v - \frac{1}{2e}v^2) & \text{if} \quad \lambda \geq 1. \end{cases}$$

Proof (Sketch). We sketch a proof which relies heavily on methods in [14]. However, the approach is very similar to the estimate of the integral in Equation (8). We will be using the same choices of ρ as in Subsection 3.2, depending on the sign of $\lambda - 1$. Hence, in all the cases, if $x = v\rho e^{it}$ where $t \in [-\pi, \pi]$, then $w(-vx) = W_0(v\alpha\beta) = o(1)$ for $v = \mathcal{O}(\sqrt{\alpha})$. Thus, we may use (6) of Lemma 5 to estimate $\phi(-vx, y)$. By Taylor approximation

$$w(-v\rho e^{it})^2 + 2w(-v\rho e^{it}) = 2\alpha v\rho\beta e^{it} - v^2\alpha^2\beta\rho^2 e^{2it} + \mathcal{O}(\alpha^{3/2}).$$

Recalling that $\beta = e^{\alpha/2}$, we can substitute β in the above estimate by $1 + \mathcal{O}(\alpha)$. Therefore, we obtain the following estimate of $\phi(-vx, y)$:

$$\phi(-vx, y) \sim e^{v\rho e^{it} - \frac{1}{2}v^2\alpha\rho^2 e^{2it}} \tag{13}$$

uniformly $v = \mathcal{O}(\sqrt{\alpha})$. Noting that the second term in the exponent is a bounded term, it does not have much effect on the application of the method. However, the first term does affect the local integrals. Skipping the details, the results are summarised as follows:

- For $\lambda < 1$ we choose $\rho = \frac{n}{\beta}e^{-\alpha n}$. When calculating the local integral in the current case, instead of the Gaussian integral (10), we have

  $$P(\alpha, w(\rho))e^{v\rho - \frac{1}{2}v^2\alpha\rho^2} e^{h(0)} \int_{-\alpha^c}^{\alpha^c} e^{v\rho it + \frac{1}{2}h''(0)t^2}\, dt,$$

  where P and h are as defined in the proof of Lemma 6. The range of the integral can be extended to $(-\infty, \infty)$ with an error term smaller than any power of α. This leads to

  $$A_n(y, u) \sim A_n(y)e^{v\rho - \frac{1}{2}v^2\rho^2\frac{\alpha}{\lambda}}.$$

- For $\lambda = 1$, we choose $\rho = \frac{1}{e\alpha\beta}$. The estimate of the integral is based on [14, Lemma 9, Eq. (27)]. There, the length of the range of the local integral is much shorter, $\mathcal{O}(\alpha^c)$, where $c \in (1/2, 2/3)$. For t in that range, the estimate in (13) simplifies further to

  $$\phi(-v\rho e^{it}, y) \sim e^{v\rho - \frac{1}{2}v^2\alpha\rho^2}.$$

  Therefore, we get

  $$A_n(y, u) \sim A_n(y)e^{v\rho - \frac{1}{2}v^2\alpha\rho^2}$$

- For $\lambda > 1$ we choose $\rho = \frac{1}{e}y^{-1} - \frac{b}{2^{1/3}e}\, y^{-1/3}$ exactly as in Lemma 8. The Cauchy integral formula is used to obtain the main term, and [14, Lemma 9] to estimate the error. We get

  $$A_n(y, u) \sim A_n(y)e^{v\varrho_1(y) - \frac{1}{2}v^2\alpha\varrho_1(y)^2}$$

Expressing everything in terms n and v gives the estimate in the lemma. ◀

The central limit theorem in Theorem 1 follows directly from Lemma 9 using Hwang's quasi-power theorem [5].

5 The number of isolated vertices

5.1 Estimate of the average

The generating function in this case is $A(x, y, 1, 1, u)$, i.e., the variable u indicates the number of isolated vertices. Let us abbreviate $A_n(y, 1, 1, u)$ by $A_n^*(y, u)$. Hence, we have

$$\sum_{n=0}^{\infty} A_n^*(y, u)\frac{x^n}{n!} = e^{(u-1)x}\sum_{n=0}^{\infty} A_n(y)\frac{x^n}{n!} \tag{14}$$

Differentiating once with respect to u and substituting $u = 1$, we get $\partial_u A_n^*(y,u)|_{u=1} = nA_{n-1}(y)$. Therefore, we get

$$\mathbb{E}(I(\mathbb{D}_{ac}(n,p))) = n\frac{A_{n-1}(y)}{A_n(y)}.$$

Comparing the latter with the expression of $\mathbb{E}(S(\mathbb{D}_{ac}(n,p)))$ that we obtained in the previous section, we see that the only difference is the term $(1+y)^{n-1}$. The combinatorial interpretation is that the term $A_{n-1}(y)$ is the generating function for the acyclic digraphs on $[n]$ with one marked isolated vertex. The estimate of the mean in Theorem 2 follows from easily the above formula.

5.2 Asymptotic normality

We estimate the quotient $\frac{A_n^*(y,u)}{A_n(y)}$ when u is a complex number of the form $1 + \mathcal{O}(n^{-1/2})$. Again, we write $v = u - 1$. We begin with the follow observation:

▶ **Lemma 10.** *Let $a \neq 0$ be a fixed real constant, then the following estimate holds uniformly for $v = \mathcal{O}(n^{-1/2})$ as $n \to \infty$:*

$$(1 - av\mu^*(\lambda))^{-n/a} \sim \sum_{0\leq j<n^{3/5}} \binom{n}{j} (v\mu^*(\lambda))^j e^{(a+1)\frac{j^2}{2n}}, \tag{15}$$

and

$$e^{nv\mu^*(\lambda)} \sim \sum_{0\leq j<n^{3/5}} \binom{n}{j} (v\mu^*(\lambda))^j e^{\frac{j^2}{2n}}.$$

Proof. First we take the binomial expansion of the term on the left-hand side of (15), we obtain

$$(1 - av\mu^*(\lambda))^{-n/a} = \sum_{j=0}^{\infty} \binom{-\frac{n}{a}}{j} (-a)^j (v\mu^*(\lambda))^j. \tag{16}$$

Then, we simplify the summand of the above series as follows:

$$\begin{aligned}
\binom{-\frac{n}{a}}{j}(-a)^j &= \frac{n(n+a)(n+2a)\cdots(n+a(j-1))}{j!} \\
&= \binom{n}{j} \frac{(1+\frac{a}{n})(1+\frac{2a}{n})\cdots(1+\frac{(j-1)a}{n})}{(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{(j-1)}{n})} \\
&= \binom{n}{j} e^{(a+1)\frac{j^2}{2n}+\mathcal{O}(\frac{j}{n})}.
\end{aligned}$$

Thus, the series on the right-hand side of (16) is asymptotically equal to

$$\sum_{0\leq j<n^{3/5}} \binom{n}{j} (v\mu^*(\lambda))^j e^{(a+1)\frac{j^2}{2n}} + \sum_{j\geq n^{3/5}} \binom{-\frac{n}{a}}{j} (-a)^j (v\mu^*(\lambda))^j.$$

We need to show that the contribution from the second term is significantly smaller. To that end, observe that

$$(1 - av\mu^*(\lambda))^{-n/a} = e^{v\mu^*(\lambda)n+\mathcal{O}(1)},$$

for $v = \mathcal{O}(n^{-1/2})$, which gives us the order magnitude of the main term. On the other hand, the bound $v = \mathcal{O}(n^{-1/2})$ and the Stirling's formula yield

$$\sum_{j \geq n^{3/5}} \binom{-\frac{n}{a}}{j} (-a)^j (v\mu^*(\lambda))^j = \mathcal{O}\left(e^{-(\frac{1}{10}+o(1))n^{3/5}\log(n)}\right)$$

which is asymptotically much smaller than the main term $(1 - av\mu^*(\lambda))^{-n/a}$.

The second estimate is done in a similar manner. Notice that

$$\binom{n}{j}(v\mu^*(\lambda))^j e^{\frac{j^2}{2n}} = (1+o(1))\frac{(nv\mu^*(\lambda))^j}{j!}$$

uniformly for $j < n^{3/5}$. Summing over $j < n^{3/5}$ and using the truncation argument above completes the proof of the lemma. ◀

▶ **Lemma 11.** *We have*

$$\frac{A_n^*(y,u)}{A_n(y)} = (1+o(1))e^{\mu^*(\lambda)vn + \frac{1}{2}a(\lambda)\mu^*(\lambda)^2v^2n},$$

as $n \to \infty$, uniformly for $v = \mathcal{O}(n^{-1/2})$, where

$$a(\lambda) = \begin{cases} 2\lambda - 1 & \text{if } \lambda < 1 \\ \lambda & \text{if } \lambda \geq 1. \end{cases}$$

Proof. If we extract the coefficient of x^n from the left and right sides of Equation (14) and dividing both by $A_n(y)$, we obtain

$$\frac{A_n^*(y,u)}{A_n(y)} = \sum_{j=0}^{n} \binom{n}{j} v^j \frac{A_{n-j}(y)}{A_n(y)}.$$

It is easy to show that $A_{n-j}(y) \leq A_n(y)$ for $y \geq 0$ from that fact that any acyclic digraph on $n-j$ vertices can be extended to an acyclic digraph on n vertices with same number of edges by simply adding j isolated vertices. This implies that the quotient $\frac{A_{n-j}(y)}{A_n(y)}$ is bounded above by 1, and the same argument we used in the proof of the previous lemma yields

$$\frac{A_n^*(y,u)}{A_n(y)} \sim \sum_{0 \leq j < n^{3/5}} \binom{n}{j} v^j \frac{A_{n-j}(y)}{A_n(y)},$$

as $n \to \infty$ uniformly for $v = \mathcal{O}(n^{-1/2})$. By making use of Lemmas 6–8, we obtain

$$\frac{A_{n-j}(y)}{A_n(y)} = (1+o(1))\mu^*(\lambda)^j \times \begin{cases} e^{\lambda j^2/n} & \text{if } \lambda < 1 \\ e^{\frac{1}{2}(\lambda+1)j^2/n} & \text{if } \lambda \geq 1 \end{cases}$$

uniformly for $0 \leq j \leq n^{3/5}$. Putting two latter estimates together and applying Lemma 10 for the appropriate values of a complete the proof the lemma. ◀

The central limit theorem in Theorem 2 follows easily from Lemma 11.

6 Conclusion

We consider a random DAG model $\mathbb{D}_{ac}(n,p)$ on the vertex set $[n] = \{1, 2, \dots, n\}$ which naturally generalises the uniform random DAG model on $[n]$. We established that if p is of the form λ/n, where $\lambda > 0$ is fixed, then the number of sources, sinks and isolated vertices are all asymptotically normal with means proportional to n. For further investigation, it would be interesting to know the typical shape of such a random DAG. We could look at the distribution of the height, which was considered by McKay [12] for the case $p = \frac{1}{2}$. We expect the height to be much smaller than n for $p = \lambda/n$, but it seems that even an estimate of the expectation of this parameter would require a significant amount of work. Alternatively, we could also investigate the number of vertices at a given level (the level of a vertex v in a DAG is the length of the longest directed path from a source to v). McKay also considered this parameter for $p = \frac{1}{2}$ in [12]. The case where p is tending to zero seems to be more challenging. It is not even clear if the graphic generating functions for these parameters can be written in forms that we can analyse. Hence, these problems are left for future work.

References

1 A. D. Barbour. Poisson convergence and random graphs. *Math. Proc. Cambridge Philos. Soc.*, 92(2):349–359, 1982. doi:10.1017/S0305004100059995.

2 A. D. Barbour, MichałKaroński, and Andrzej Ruciński. A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B*, 47(2):125–145, 1989. doi:10.1016/0095-8956(89)90014-2.

3 É. de Panafieu, S. Dovgal, D. Ralaivaosaona, V. Rasendrahasina, and S. Wagner. The birth of the strong components, 2020. arXiv:2009.12127.

4 I. M. Gessel. Counting acyclic digraphs by sources and sinks. *Discrete Mathematics*, 160(1):253–258, 1996.

5 H.-K. Hwang. On convergence rates in the central limit theorems for combinatorial structures. *European Journal of Combinatorics*, 19(3):329–343, 1998. doi:10.1006/eujc.1997.0179.

6 M. Karoński and A. Ruciński. Poisson convergence and semi-induced properties of random graphs. *Math. Proc. Cambridge Philos. Soc.*, 101(2):291–300, 1987. doi:10.1017/S0305004100066664.

7 R. M. Karp. The transitive closure of a random digraph. *Random Structures & Algorithms*, 1(1):73–93, 1990.

8 V. A. Liskovec. On the enumeration of strongly connected directed graphs. *Dokl. Akad. Nauk BSSR*, 17:1077–1080, 1163, 1973.

9 V. A. Liskovets. On the number of maximal vertices of a random acyclic digraph. *Theory of Probability & Its Applications*, 20(2):401–409, 1976.

10 T. Łuczak. The phase transition in the evolution of random digraphs. *Journal of Graph Theory*, 14:217–223, 1990.

11 T. Łuczak and T. Seierstad. The critical behavior of random digraphs. *Random Structures & Algorithms*, 35:271–293, 2009.

12 B. D. Mckay. On the shape of a random acyclic digraph. *Mathematical Proceedings of the Cambridge Philosophical Society*, 106(3):459–465, 1989. doi:10.1017/S0305004100068195.

13 F. W. Olver, D. W. Lozier, Ronald F. Boisvert, and C. W. Clark. *NIST Handbook of Mathematical Functions*. Cambridge University Press, USA, 1st edition, 2010.

14 D. Ralaivaosaona, V. Rasendrahasina, and S. Wagner. On the probability that a random digraph is acyclic. In *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2020)*, volume 159 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 25:1–25:18, 2020.

15 R. W. Robinson. Counting labelled acyclic digraphs. In *New Directions in Graph Theory (Ed. F. Harary)*. New York: Academic Press, 1973.

16 R. P. Stanley. Acyclic orientations of graphs. *Discrete Mathematics*, 5(2):171–178, 1973.

17 The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 8.9)*, 2019. https://www.sagemath.org.

18 P. Zbigniew. On the number of vertices of given degree in a random graph. *J. Graph Theory*, 8(1):167–170, 1984. doi:10.1002/jgt.3190080119.

Parking Functions, Multi-Shuffle, and Asymptotic Phenomena

Mei Yin
Department of Mathematics, University of Denver, Denver, CO, USA

Abstract

Given a positive integer-valued vector $\mathbf{u} = (u_1, \ldots, u_m)$ with $u_1 < \cdots < u_m$, a $\mathbf{u}$-parking function of length m is a sequence $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_m)$ of positive integers whose non-decreasing rearrangement $(\lambda_1, \ldots, \lambda_m)$ satisfies $\lambda_i \leq u_i$ for all $1 \leq i \leq m$. We introduce a combinatorial construction termed a parking function multi-shuffle to generic $\mathbf{u}$-parking functions and obtain an explicit characterization of multiple parking coordinates. As an application, we derive various asymptotic probabilistic properties of a uniform $\mathbf{u}$-parking function of length m when $u_i = cm + ib$. The asymptotic scenario in the generic situation $c > 0$ is in sharp contrast with that of the special situation $c = 0$.

2012 ACM Subject Classification Mathematics of computing → Probability and statistics; Theory of computation → Randomness, geometry and discrete structures

Keywords and phrases Parking function, Multi-shuffle, Asymptotic expansion, Abel's multinomial theorem

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.18

Related Version *Full Version*: `https://arxiv.org/abs/2112.02251`

Funding Research partially supported by the University of Denver's Faculty Research Fund 84688-145601.

1 Introduction

Parking functions were introduced by Konheim and Weiss [10], under the name of "parking disciplines," to study the following problem. Consider a parking lot with n parking spots placed sequentially along a one-way street. In order, a line of $m \leq n$ cars enters the lot. The ith car drives to its preferred spot π_i and parks there if possible, and otherwise takes the next available spot if it exists. The sequence of preferences $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_m)$ is called a *parking function* if all cars successfully park. We denote the set of parking functions by $\mathrm{PF}(m, n)$, where m is the number of cars and n is the number of parking spots.

Nowadays, parking functions are an established area of research in combinatorics, with connections to labeled trees and forests [3], hyperplane arrangements, interval orders, and plane partitions [16, 17], diagonal harmonics and (q, t)-analogs of Catalan numbers [7], abelian sandpiles [4], to mention a few. Properties of random parking functions have also been of interest to statisticians and probabilists [5]. We refer to Knuth [9, Section 6.4] and Yan [20] for a comprehensive survey.

Given a positive integer-valued vector $\mathbf{u} = (u_1, \ldots, u_m)$ with $u_1 < \cdots < u_m$, a $\mathbf{u}$-*parking function* of length m is a sequence $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_m)$ of positive integers whose non-decreasing rearrangement $(\lambda_1, \ldots, \lambda_m)$ satisfies $\lambda_i \leq u_i$ for all $1 \leq i \leq m$. Denote the set of $\mathbf{u}$-parking functions by $\mathrm{PF}(\mathbf{u})$. There is a similar interpretation for $\mathbf{u}$-parking functions in terms of the parking scenario depicted above: One wishes to park m cars in a one-way street with u_m spots, but only m spots, at positions $u_1, \ldots, u_m$, are still empty [12]. We recognize that the parking function $\mathrm{PF}(m, n)$ is a special case of the more general $\mathbf{u}$-parking functions with $u_i = n - m + i$.

In our previous work on $\mathrm{PF}(m,n)$ [8] [21], we introduced an original combinatorial construction which we term a *parking function multi-shuffle*, and it facilitated an investigation of the properties of a parking function chosen uniformly at random from $\mathrm{PF}(m,n)$. This paper will delve into the essence of the multi-shuffle combinatorial construction on parking functions and introduce the concept to generic **u**-parking functions, thus allowing for an explicit characterization of multiple coordinates of **u**-parking functions. As an application, we will derive various asymptotic probabilistic properties of a uniform (a,b)-parking function, which is a class of **u**-parking functions where $u_i = a + (i-1)b$ for some positive integers a and b. Denote the set of (a,b)-parking functions of length m by $\mathrm{PF}(a,b,m)$. It coincides with $\mathrm{PF}(m,n)$ when $a = n-m+1$ and $b = 1$. In view of and generalizing this correspondence, we will study the asymptotics of $\mathrm{PF}(a,b,m)$ when $b \geq 1$ is any integer and $a = cm + b$ for some $c \geq 0$. We find that for large m, various probabilistic quantities display strikingly different asymptotic tendencies in the generic situation $c > 0$ (corresponding to $m \lesssim n$) vs. the special situation $c = 0$ (corresponding to $m = n$), including the boundary behavior of a single coordinate and all moments of multiple coordinates.

Our asymptotic calculation utilizes the multi-dimensional Cauchy product of the *tree function* $F(z)$, which is a variant of the Lambert function, the *Gončarov polynomials* $g_m(x; a_0, a_1, \ldots, a_{m-1})$ for $m = 0, 1, \ldots$, which form a natural basis for working with **u**-parking functions, as well as *Abel's multinomial theorem*. In particular, our new perspective on parking functions leads to asymptotic moment calculations for multiple coordinates of (a,b)-parking functions that complement the work of Kung and Yan [11], where the explicit formulas for the first and second factorial moments and a general form for the higher factorial moments of sums of (a,b)-parking functions were given.

This paper is organized as follows. Section 2 illustrates the notion of **u**-*parking function multi-shuffle* that decomposes a **u**-parking function into smaller components (Definition 3). This construction offers an explicit characterization of multiple coordinates of **u**-parking functions (Theorems 5 and 7). Theorem 7 enumerates **u**-parking functions in connection with Gončarov polynomials, but we also provide an alternative description of **u**-parking functions in Proposition 8. Section 3 uses the multi-shuffle construction introduced in Section 2 to investigate various properties of a parking function chosen uniformly at random from $\mathrm{PF}(cm+b, b, m)$. When the parking preferences $\pi_1, \ldots, \pi_l$ are exactly b spots apart, a simplified characterization of $(cm+b, b)$-parking functions is given in Section 3.1. Building upon Theorem 7 and Proposition 10, we compute asymptotics of all moments of multiple coordinates in Theorem 14 in the generic situation $c > 0$. The asymptotic mixed moments in the generic situation $c > 0$ are contrasted with that of the special situation $c = 0$ in Section 3.2. We then focus on the boundary behavior of a single coordinate in Section 3.3. We find that in the generic situation $c > 0$ on the right end it approximates a Borel distribution with parameter $b/(b+c)$ while on the left end it deviates from the constant value in a rescaled Poisson fashion (Corollaries 19 and 20). This asymptotic tendency differs from that in the special situation $c = 0$, where the boundary behavior of a single coordinate on the left and right ends both approach Borel(1) (Corollaries 19 and 21).

Notations

Let $\mathbb{N}$ be the set of positive integers. For $m, n \in \mathbb{N}$, we write $[m,n]$ for the set of integers $\{m, \ldots, n\}$ and $[n] = [1,n]$. For vectors $\mathbf{a}, \mathbf{b} \in [n]^m$, denote by $\mathbf{a} \leq_C \mathbf{b}$ if $a_i \leq b_i$ for all $i \in [m]$; this is the component-wise partial order on $[n]^m$. In a similar fashion, denote by $\mathbf{a} <_C \mathbf{b}$ if $a_i \leq b_i$ for all $i \in [m]$ and there is at least one $j \in [m]$ such that $a_j < b_j$. For $\mathbf{b} \in [n]^m$, we write $[\mathbf{b}]$ for the set of $\mathbf{a} \in [n]^m$ with $\mathbf{a} \leq_C \mathbf{b}$.

2 u-parking function multi-shuffle

In this section we explore the properties of generic $\mathbf{u}$-parking functions through a $\mathbf{u}$-parking function multi-shuffle construction. We will write our results in terms of parking coordinates $\pi_1, \ldots, \pi_l$ for explicitness, where $1 \leq l \leq m$ is any integer. But due to permutation symmetry, they may be interpreted for any coordinates. Temporarily fix $\pi_{l+1}, \ldots, \pi_m$. Let

$$A_{\pi_{l+1},\ldots,\pi_m} = \{\mathbf{v} = (v_1, \ldots, v_l) : (v_1, \ldots, v_l, \pi_{l+1}, \ldots, \pi_m) \in \mathrm{PF}(\mathbf{u})\},$$

where $\mathbf{v}$ is in non-decreasing order. Following terminology in [1], we will call such $\mathbf{v}$'s $\mathbf{u}$-parking completions for $\pi_{l+1}, \ldots, \pi_m$.

▶ **Proposition 1.** *Take $1 \leq l \leq m$ any integer. Suppose that $\mathbf{v}$ is in non-decreasing order and is a maximal $\mathbf{u}$-parking completion for the fixed $\pi_{l+1}, \ldots, \pi_m$. Then for all $1 \leq i \leq l$, we have $v_i = u_{k_i}$ with $1 \leq k_1 < \cdots < k_l \leq m$, and $v_i \neq \pi_{l+1}, \ldots, \pi_m$.*

To identify the maximal $\mathbf{v}$ in $A_{\pi_{l+1},\ldots,\pi_m}$, we arrange π_i for $l+1 \leq i \leq m$ in non-decreasing order, denoted by $\pi_{(l+1)} \leq \cdots \leq \pi_{(m)}$. Set $n_l = 0$. We find the minimum index n_i in order, starting with n_{l+1}, such that $n_i > n_{i-1}$ and $u_{n_i} \geq \pi_{(i)}$ for each $l+1 \leq i \leq m$. If such u_{n_i}'s cannot be located, then $A_{\pi_{l+1},\ldots,\pi_m}$ is empty. Otherwise excluding these u_{n_i}'s from $\mathbf{u}$ gives the optimal $\mathbf{v}$. From the parking scheme, if $\mathbf{v} \in A_{\pi_{l+1},\ldots,\pi_m}$, then $\mathbf{w} \in A_{\pi_{l+1},\ldots,\pi_m}$ for all $\mathbf{w} \leq_C \mathbf{v}$, where $\leq_C$ is the component-wise partial order. This implies that if $A_{\pi_{l+1},\ldots,\pi_m}$ is non-empty, then there is a unique maximal $\mathbf{u}$-parking completion $\mathbf{v} \in [u_m]^l$ with $v_i = u_{k_i}$ for all $1 \leq i \leq l$, where $1 \leq k_1 < \cdots < k_l \leq m$, and $A_{\pi_{l+1},\ldots,\pi_m} = [\mathbf{v}]$. Therefore given the last $m - l$ parking preferences, it is sufficient to identify the largest feasible first l preferences (if they exist).

▶ **Example 2.** Take $\mathbf{u} = (u_1, u_2, u_3, u_4) = (2, 3, 5, 8)$, $\pi_3 = 6$, and $\pi_4 = 2$. Then $A_{\pi_3,\pi_4} = [\mathbf{v}] = [(u_2, u_3)] = [(3, 5)]$. See illustration below.

$$\begin{array}{cccc} \pi_{(3)} & 2 & \leq & 2 & u_1 \\ v_1 & & & 3 & u_2 \\ v_2 & & & 5 & u_3 \\ \pi_{(4)} & 6 & \leq & 8 & u_4 \end{array}$$

We will now introduce an original combinatorial construction which we term a parking function multi-shuffle to generic $\mathbf{u}$-parking functions.

▶ **Definition 3** ($\mathbf{u}$-parking function multi-shuffle). *Take $1 \leq l \leq m$ any integer. Let $\mathbf{v} = (v_1, \ldots, v_l) \in [u_m]^l$ be in increasing order with $v_i = u_{k_i}$ for all $1 \leq i \leq l$, where $1 \leq k_1 < \cdots < k_l \leq m$. Say that $\pi_{l+1}, \ldots, \pi_m$ is a* $\mathbf{u}$-parking function multi-shuffle *of $l+1$ $\mathbf{u}$-parking functions $\boldsymbol{\alpha}_1 \in PF(u_1, \ldots, u_{k_1-1}), \boldsymbol{\alpha}_2 \in PF(u_{k_1+1} - u_{k_1}, \ldots, u_{k_2-1} - u_{k_1}), \ldots, \boldsymbol{\alpha}_l \in PF(u_{k_{l-1}+1} - u_{k_{l-1}}, \ldots, u_{k_l-1} - u_{k_{l-1}})$, and $\boldsymbol{\alpha}_{l+1} \in PF(u_{k_l+1} - u_{k_l}, \ldots, u_m - u_{k_l})$ if $\pi_{l+1}, \ldots, \pi_m$ is any permutation of the union of the $l+1$ words $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 + (u_{k_1}, \ldots, u_{k_1}), \ldots, \boldsymbol{\alpha}_{l+1} + (u_{k_l}, \ldots, u_{k_l})$. (If $k_{j-1} = k_j - 1$ for some j, we take the corresponding $\boldsymbol{\alpha}_j$ as empty.) We will denote this by $(\pi_{l+1}, \ldots, \pi_m) \in MS(\mathbf{v})$.*

▶ **Example 4.** Take $\mathbf{u} = (u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8) = (3, 4, 5, 6, 7, 8, 9, 10)$ and $\mathbf{v} = (u_4, u_6) = (6, 8)$. Take $\boldsymbol{\alpha}_1 = (2, 1, 2) \in \mathrm{PF}(u_1, u_2, u_3) = \mathrm{PF}(3, 4, 5)$, $\boldsymbol{\alpha}_2 = (1) \in \mathrm{PF}(u_5 - u_4) = \mathrm{PF}(1)$, and $\boldsymbol{\alpha}_3 = (2, 1) \in \mathrm{PF}(u_7 - u_6, u_8 - u_6) = \mathrm{PF}(1, 2)$. Then $(2, \overline{7}, 2, \underline{9}, \underline{10}, 1) \in \mathrm{MS}(6, 8)$ is a multi-shuffle of the three words $(2, 1, 2)$, (7), and $(10, 9)$.

Take $\mathbf{u} = (u_1, u_2, u_3, u_4, u_5, u_6, u_7) = (2, 4, 5, 6, 8, 12, 15)$ and $\mathbf{v} = (u_2, u_5) = (4, 8)$. Take $\boldsymbol{\alpha}_1 = (1) \in \mathrm{PF}(u_1) = \mathrm{PF}(2)$, $\boldsymbol{\alpha}_2 = (1, 2) \in \mathrm{PF}(u_3 - u_2, u_4 - u_2) = \mathrm{PF}(1, 2)$, and $\boldsymbol{\alpha}_3 = (5, 3) \in \mathrm{PF}(u_6 - u_5, u_7 - u_5) = \mathrm{PF}(4, 7)$. Then $(\underline{11}, \underline{13}, \overline{5}, 1, \overline{6}) \in \mathrm{MS}(4, 8)$ is a multi-shuffle of the three words (1), $(5, 6)$, and $(13, 11)$.

The **u**-parking function multi-shuffle allows for an explicit characterization of multiple coordinates of **u**-parking functions. It connects the identification of the maximal element in $A_{\pi_{l+1},\dots,\pi_m}$ to the decomposition of $\pi_{l+1},\dots,\pi_m$ into a multi-shuffle.

▶ **Theorem 5.** *Take $1 \le l \le m$ any integer. Let $\mathbf{v} = (v_1,\dots,v_l) \in [u_m]^l$ be in increasing order with $v_i = u_{k_i}$ for all $1 \le i \le l$, where $1 \le k_1 < \cdots < k_l \le m$. Then $A_{\pi_{l+1},\dots,\pi_m} = [\mathbf{v}]$ if and only if $(\pi_{l+1},\dots,\pi_m) \in MS(\mathbf{v})$.*

▶ **Example 6** (Continued from Example 4)**.** Take $\mathbf{u} = (3,4,5,6,7,8,9,10)$, then $A_{2,7,2,9,10,1} = [(6,8)]$ is equivalent to $(2,7,2,9,10,1) \in \mathrm{MS}(6,8)$. Take $\mathbf{u} = (2,4,5,6,8,12,15)$, then $A_{11,13,5,1,6} = [(4,8)]$ is equivalent to $(11,13,5,1,6) \in \mathrm{MS}(4,8)$. See illustration below.

$\pi_{(3)}$	1	$\le$	3	u_1
$\pi_{(4)}$	2	$\le$	4	u_2
$\pi_{(5)}$	2	$\le$	5	u_3
v_1			6	u_4
$\pi_{(6)}$	7	$\le$	7	u_5
v_2			8	u_6
$\pi_{(7)}$	9	$\le$	9	u_7
$\pi_{(8)}$	10	$\le$	10	u_8

$\pi_{(3)}$	1	$\le$	2	u_1
v_1			4	u_2
$\pi_{(4)}$	5	$\le$	5	u_3
$\pi_{(5)}$	6	$\le$	6	u_4
v_2			8	u_5
$\pi_{(6)}$	11	$\le$	12	u_6
$\pi_{(7)}$	13	$\le$	15	u_7

Of relevance to our investigation, we also utilize the Gončarov polynomials. Let $(a_0, a_1, \dots)$ be a sequence of numbers. The Gončarov polynomials $g_m(x; a_0, a_1, \dots, a_{m-1})$ for $m = 0, 1, \dots$ are the basis of solutions to the Gončarov interpolation problem in numerical analysis. They are defined by the biorthogonality relation:

$$\epsilon(a_i) D^i g_m(x; a_0, a_1, \dots, a_{m-1}) = m! \delta_{im},$$

where $\epsilon(a_i)$ is evaluation at a_i, D is the differentiation operator, and δ_{im} is the Kronecker delta. The Gončarov polynomials satisfy many nice algebraic and analytic properties, making them very useful in analysis and combinatorics. Specifically, we list two properties of Gončarov polynomials below:

1. Determinant formula.

$$g_m(x; a_0, a_1, \dots, a_{m-1}) = m! \begin{vmatrix} 1 & a_0 & \frac{a_0^2}{2!} & \frac{a_0^3}{3!} & \cdots & \frac{a_0^{m-1}}{(m-1)!} & \frac{a_0^m}{m!} \\ 0 & 1 & a_1 & \frac{a_1^2}{2!} & \cdots & \frac{a_1^{m-2}}{(m-2)!} & \frac{a_1^{m-1}}{(m-1)!} \\ 0 & 0 & 1 & a_2 & \cdots & \frac{a_2^{m-3}}{(m-3)!} & \frac{a_2^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & a_{m-1} \\ 1 & x & \frac{x^2}{2!} & \frac{x^3}{3!} & \cdots & \frac{x^{m-1}}{(m-1)!} & \frac{x^m}{m!} \end{vmatrix}.$$

2. Shift invariance.

$$g_m(x + y; a_0 + y, a_1 + y, \dots, a_{m-1} + y) = g_m(x; a_0, a_1, \dots, a_{m-1}).$$

We note that the number of **u**-parking functions of length m is $|\mathrm{PF}(\mathbf{u})| = (-1)^m g_m(0; u_1, \dots, u_m)$. For a full discussion of the connection between Gončarov polynomials and **u**-parking functions, we refer to Kung and Yan [12].

▶ **Theorem 7.** *Take $1 \le l \le m$ any integer. Let $\mathbf{w} = (w_1, \dots, w_l) \in [u_m]^l$ be in non-decreasing order. The number of parking functions $\boldsymbol{\pi} \in PF(\mathbf{u})$ with $\pi_1 = w_1, \dots, \pi_l = w_l$ is*

$$(-1)^{m-l} \sum_{\mathbf{s} \in S_l(\mathbf{w})} \binom{m-l}{\mathbf{s}} \prod_{i=1}^{l+1} g_{s_i}(u_{s_1+\cdots+s_{i-1}+i-1}; u_{s_1+\cdots+s_{i-1}+i}, \dots, u_{s_1+\cdots+s_i+i-1}),$$

where $s_0 = u_0 = 0$, $g_{s_i}(\cdot)$ are the Gončarov polynomials, and

$$S_l(\mathbf{w}) = \left\{ \mathbf{s} = (s_1, \dots, s_{l+1}) \in \mathbb{N}^{l+1} \mid \substack{u_{s_1+\cdots+s_i+i} \ge w_i \ \forall i \in [l] \\ s_1+\cdots+s_{l+1}=m-l} \right\}.$$

Note that this quantity stays constant if all $w_i \le u_i$ and decreases as each w_i increases past u_i as there are fewer resulting summands.

For the special case $l = 0$ and $\mathbf{w} = ()$ (where no parking preferences are specified), we recover the total number of $\mathbf{u}$-parking functions $|\operatorname{PF}(\mathbf{u})| = (-1)^m g_m(0; u_1, \dots, u_m)$. We describe an alternative characterization of this number in the following.

▶ **Proposition 8.** *The number of $\mathbf{u}$-parking functions $|PF(\mathbf{u})|$ satisfies*

$$|PF(\mathbf{u})| = \sum_{\mathbf{s} \in C(m)} \binom{m}{\mathbf{s}} \prod_{i=1}^{u_m - m + 1} (s_i + 1)^{s_i - 1},$$

where $C(m)$ consists of compositions of m: $\mathbf{s} = (s_1, \dots, s_{u_m - m + 1}) \models m$ with $\sum_{i=1}^{u_m-m+1} s_i = m$, subject to $s_1 + \cdots + s_{u_i - i + 1} \ge i$ for all $1 \le i \le m$.

▶ **Example 9.** Take $\mathbf{u} = (2, 5)$. Then $(\pi_1, \pi_2) \in \operatorname{PF}(\mathbf{u})$ satisfies

$$(\pi_1, \pi_2) \in A := \{(1,1), (1,2), (1,3), (1,4), (1,5), (2,1), (2,2), (2,3), (2,4), (2,5), (3,1), (3,2), (4,1), (4,2), (5,1), (5,2)\}.$$

From Proposition 8,

$$\begin{aligned} |A| = &\binom{2}{2,0,0,0} 3^1 1^{-1} 1^{-1} 1^{-1} + \binom{2}{0,2,0,0} 1^{-1} 3^1 1^{-1} 1^{-1} + \binom{2}{1,1,0,0} 2^0 2^0 1^{-1} 1^{-1} \\ &+ \binom{2}{1,0,1,0} 2^0 1^{-1} 2^0 1^{-1} + \binom{2}{1,0,0,1} 2^0 1^{-1} 1^{-1} 2^0 \\ &+ \binom{2}{0,1,1,0} 1^{-1} 2^0 2^0 1^{-1} + \binom{2}{0,1,0,1} 1^{-1} 2^0 1^{-1} 2^0 \\ = &\ 3 + 3 + 2 + 2 + 2 + 2 + 2 = 16. \end{aligned}$$

3 Properties of random (a, b)-parking functions

In general, there are no nice closed-form expressions for Gončarov polynomials related to $\mathbf{u}$-parking functions, but such expressions exist for a specific class of $\mathbf{u}$-parking functions. When the entries of the vector $\mathbf{u}$ form an arithmetic progression: $u_i = a + (i-1)b$ for some positive integers a and b, we get *Abel polynomials*:

$$g_m(x; a, a+b, \dots, a+(m-1)b) = (x-a)(x-a-mb)^{m-1}. \tag{3.1}$$

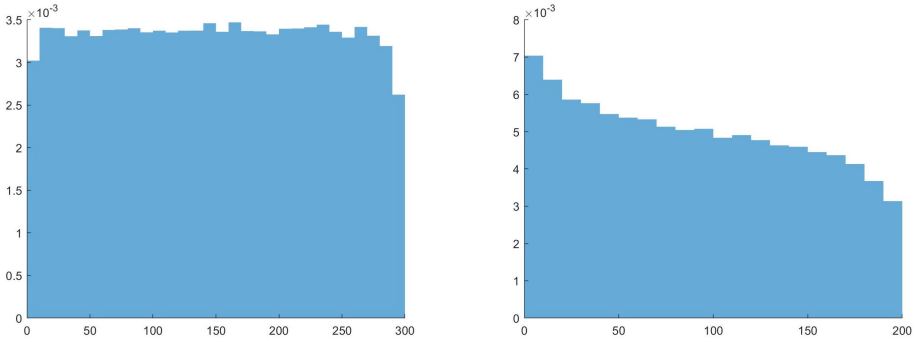


Figure 1 The distribution of π_1 (the first parking coordinate) in 100,000 samples of $(cm+b, b)$-parking functions chosen uniformly at random, where $m = 100$ and $b = 2$. In the left plot $c = 1$ and in the right plot $c = 0$.

We call these **u**-parking functions (a, b)-parking functions, and denote the set of (a, b)-parking functions of length m by $\mathrm{PF}(a, b, m)$. Using (3.1), the number of (a, b)-parking functions of length m is $|\mathrm{PF}(a, b, m)| = (-1)^m g_m(0; a, a+b, \ldots, a+(m-1)b) = a(a+mb)^{m-1}$.

In this section we use the multi-shuffle construction introduced in Section 2 to investigate various properties of a parking function chosen uniformly at random from $\mathrm{PF}(a, b, m)$. As stated in the introduction, taking $b = 1$, $c = n/m - 1$, and $a = cm + b$, an (a, b)-parking function of length m depicts the scenario of parking m cars in n spots sequentially along a one-way street. Therefore among all possible a and b, of particular interest to us is when $a = cm + b$ for some $c \geq 0$. We will write our results in terms of coordinates $\pi_1, \ldots, \pi_l$ of parking functions, where $1 \leq l \leq m$ is any integer. However, the parking coordinates satisfy permutation symmetry, so the statements in this section may be interpreted for any coordinates.

Before proceeding with the calculations, we outline an effective method for generating a random (a, b)-parking function of length m. The algorithm is suggested by Stanley's generalization [16] of Pollak's circle argument for parking functions [6]. To select $\pi \in \mathrm{PF}(a, b, m)$ uniformly at random:

1. Pick an element $\pi \in (\mathbb{Z}/(a+mb)\mathbb{Z})^m$, where the equivalence class representatives are taken in $1, \ldots, a+mb$.
2. For $k \in \{0, \ldots, a+mb-1\}$, record k if $\pi + k(1, \ldots, 1) \in \mathrm{PF}(a, b, m)$ (modulo $a+mb$), where $(1, \ldots, 1)$ is a vector of length m. There should be exactly a such k's.
3. Pick one k from (2) uniformly at random. Then $\pi + k(1, \ldots, 1)$ is an (a, b)-parking function of length m taken uniformly at random.

Figure 1 shows a histogram of the values of π_1 based on 100,000 random samples of $\mathrm{PF}(cm+b, b, m)$ for $m = 100$ and $b = 2$. The left plot is for $c = 1$ and the right plot is for $c = 0$. A closed formula for the distribution of π_1 as well as its asymptotic approximation will be provided in Section 3.3.

3.1 Mixed moments of multiple coordinates

In this subsection we study the asymptotics of the generic mixed moments of (a, b)-parking functions of length m when $a = cm + b$ for some $c > 0$ via a tree function approach. Building upon Theorem 7 and Proposition 8, we first count the number of (a, b)-parking functions of length m where the specified parking preferences of the first l cars are exactly b spots apart. This calculation will be central in deriving an asymptotic formula for the mixed moments.

▶ **Proposition 10.** *Take $1 \le l \le m$ any integer. Let $0 \le k \le m-l$. The number of parking functions $\boldsymbol{\pi} \in PF(a,b,m)$ with $\pi_1 = a+kb, \dots, \pi_l = a+(k+l-1)b$ is*

$$a \sum_{s=0}^{m-l-k} \binom{m-l}{s} (a+(m-l-s)b)^{m-s-l-1} b^s l (s+l)^{s-1}.$$

The following technical lemma will also be needed in the derivation of asymptotic mixed moments.

▶ **Lemma 11.** *Take $l \ge 1$ any integer and m large. Take $b \ge 1$ any integer and $a = cm+b$ for some $c > 0$. Set $u_i = a+(i-1)b$ for $i \ge 1$ and $u_0 = 0$. For $1 \le i \le l$, take $p_i \ge 1$ any integer and $S_i \sim m$ with $S_1 < \cdots < S_l$. Let $\sum_{j=u_s+1}^{u_{s+1}} j^{p_i} = f_i(s)$ for $s \ge 0$. Then*

$$\sum_{\substack{\#\{i:\, s_i \le S_k\} \ge k \\ \forall k \in [l]}} \prod_{i=1}^{l} f_i(s_i) = \frac{(cm+(S_l+1)b)^{\sum_{i=1}^{l} p_i + l}}{\prod_{i=1}^{l}(p_i+1)} \left(1 + \frac{1}{(b+c)m}\left(\frac{\sum_{i=1}^{l} p_i + l}{2}\right) + O\left(\frac{1}{m^2}\right)\right).$$

We are now ready to establish an asymptotic result for the mixed moments of two coordinates. The key idea is to break apart the parking preferences of the first two cars into blocks and utilize properties of the tree function $F(z) = \sum_{s=0}^{\infty} (s+1)^{s-1} \frac{z^s}{s!}$ in the asymptotic investigation.

▶ **Theorem 12.** *Take $p, q \ge 1$ any integer and m large. Take $b \ge 1$ any integer and $a = cm+b$ for some $c > 0$. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from $PF(a,b,m)$, we have*

$$\mathbb{E}(\pi_1^p) = \frac{((b+c)m)^p}{p+1}\left(1 + \frac{1}{c(b+c)m}\left(\frac{c(p+1)}{2} - b^2 p\right) + O\left(\frac{1}{m^2}\right)\right),$$

and

$$\mathbb{E}(\pi_1^p \pi_2^q) = \frac{((b+c)m)^{p+q}}{(p+1)(q+1)}\left(1 + \frac{1}{c(b+c)m}\left(\frac{c(p+q+2)}{2} - b^2(p+q)\right) + O\left(\frac{1}{m^2}\right)\right).$$

▶ **Proposition 13.** *Take m large. Take $b \ge 1$ any integer and $a = cm+b$ for some $c > 0$. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from $PF(a,b,m)$, we have*

$$\mathbb{V}\mathrm{ar}(\pi_1) \sim \frac{((b+c)m)^2}{12} - \frac{b^2(b+c)m}{6c}, \qquad \mathbb{C}\mathrm{ov}(\pi_1, \pi_2) \sim -\frac{b^2(b+c)^2}{4c^2}.$$

Extending the asymptotic expansion approach in the proof of Theorem 12, we have the following more general result.

▶ **Theorem 14.** *Take $l \ge 1$ any integer and m large. Take $b \ge 1$ any integer and $a = cm+b$ for some $c > 0$. For $1 \le i \le l$, take $p_i \ge 1$ any integer. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from $PF(a,b,m)$, we have*

$$\mathbb{E}\left(\prod_{i=1}^{l} \pi_i^{p_i}\right) = \frac{((b+c)m)^{\sum_{i=1}^{l} p_i}}{\prod_{i=1}^{l}(p_i+1)}\left(1 + \frac{1}{c(b+c)m}\left(\frac{c\left(\sum_{i=1}^{l} p_i + l\right)}{2} - b^2 \sum_{i=1}^{l} p_i\right) + O\left(\frac{1}{m^2}\right)\right).$$

3.2 The special situation $c = 0$

In this subsection we study the asymptotics of the generic mixed moments of (b, b)-parking functions of length m via Abel's multinomial theorem. Indeed, the asymptotic moment calculations in Section 3.1 could as well be approached via Abel's multinomial theorem. Unlike the tree function method which fails for the case $c = 0$ due to divergence, Abel's multinomial theorem applies broadly for $c \geq 0$. However calculation-wise it is in general more cumbersome to apply Abel's multinomial theorem as compared with the tree function method, so we only use this alternative approach when $c = 0$ and so $a = b$.

▶ **Theorem 15** (Abel's multinomial theorem, derived from Pitman [14] and Riordan [15]). *Let*

$$A_n(x_1, \ldots, x_m; p_1, \ldots, p_m) = \sum_{\mathbf{s} \models n} \binom{n}{\mathbf{s}} \prod_{j=1}^{m} (x_j + s_j)^{s_j + p_j},$$

where $\mathbf{s} = (s_1, \ldots, s_m)$ *and* $\sum_{i=1}^{m} s_i = n$. *Then*

$$A_n(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m; p_1, \ldots, p_i, \ldots, p_j, \ldots, p_m) = A_n(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m; p_1, \ldots, p_j, \ldots, p_i, \ldots, p_m).$$

$$A_n(x_1, \ldots, x_m; p_1, \ldots, p_m) = \sum_{i=1}^{m} A_{n-1}(x_1, \ldots, x_{i-1}, x_i + 1, x_{i+1}, \ldots, x_m; p_1, \ldots, p_{i-1}, p_i + 1, p_{i+1}, \ldots, p_m).$$

$$A_n(x_1, \ldots, x_m; p_1, \ldots, p_m) = \sum_{s=0}^{n} \binom{n}{s} s!(x_1 + s) A_{n-s}(x_1 + s, x_2, \ldots, x_m; p_1 - 1, p_2, \ldots, p_m).$$

Moreover, the following special instances hold via the basic recurrences listed above:

$$A_n(x_1, \ldots, x_m; -1, \ldots, -1) = (x_1 \cdots x_m)^{-1}(x_1 + \cdots + x_m)(x_1 + \cdots + x_m + n)^{n-1}.$$

$$A_n(x_1, \ldots, x_m; -1, \ldots, -1, 0) = (x_1 \cdots x_m)^{-1} x_m (x_1 + \cdots + x_m + n)^{n}.$$

Take $l \geq 1$ any integer and m large. Take $b \geq 1$ any integer and $a = cm + b$ for some $c \geq 0$. For $1 \leq i \leq l$, take $p_i \geq 1$ any integer. In computing $\mathbb{E}(\prod_{i=1}^{l} \pi_i^{p_i})$, we recognize from Lemma 11 that asymptotically

$$\begin{aligned}
&\sum \left(\prod_{i=1}^{l} \pi_i^{p_i} \right) \{ \boldsymbol{\pi} \in \mathrm{PF}(a, b, m) : \pi_i \text{ specified } \forall i \in [l] \} \\
&= \frac{(cm + b) b^{m + \sum_{i=1}^{l} p_i - 1}}{\prod_{i=1}^{l} (p_i + 1)} \sum_{s_1 = 0}^{m-l} \cdots \sum_{s_l = 0}^{m - l - s_1 - \cdots - s_{l-1}} \binom{m - l}{s_1, \ldots, s_l, m - l - s_1 - \cdots - s_l}. \\
&\cdot (m - l - s_1 - \cdots - s_l + 1 + \frac{cm}{b})^{m - l - 1 - s_1 - \cdots - s_l} \prod_{i=1}^{l} (s_i + 1)^{s_i - 1}. \\
&\cdot (m - s_l + \frac{cm}{b})^{\sum_{i=1}^{l} p_i + l} \left(1 + \frac{1}{(b + c)m - s_l b} \left(\frac{\sum_{i=1}^{l} p_i + l}{2} \right) + O\left(m^{-2}\right) \right). \qquad (3.2)
\end{aligned}$$

Using Abel's multinomials, the leading order terms in (3.2) may be represented as

$$\frac{(cm+b)b^{m+\sum_{i=1}^{l} p_i - 1}}{\prod_{i=1}^{l}(p_i+1)} \left(A_{m-l}\left(1+\frac{cm}{b}, \underbrace{1,\dots,1}_{l \text{ 1's}}; \sum_{i=1}^{l} p_i + l - 1, \underbrace{-1,\dots,-1}_{l \text{ -1's}}\right) \right.$$
$$+ (l-1)\left(\sum_{i=1}^{l} p_i + l\right) A_{m-l}\left(1+\frac{cm}{b}, \underbrace{1,\dots,1}_{l \text{ 1's}}; \sum_{i=1}^{l} p_i + l - 2, 0, \underbrace{-1,\dots,-1}_{l-1 \text{ -1's}}\right)$$
$$\left. + \frac{1}{2b}\left(\sum_{i=1}^{l} p_i + l\right) A_{m-l}\left(1+\frac{cm}{b}, \underbrace{1,\dots,1}_{l \text{ 1's}}; \sum_{i=1}^{l} p_i + l - 2, \underbrace{-1,\dots,-1}_{l \text{ -1's}}\right)\right).$$

This is a general formula that works for any a, b, and l. When $c = 0$ and so $a = b$, taking $l = 1, 2$, we have

$$\mathbb{E}(\pi_1) \sim b\left(\frac{m}{2} - \frac{\sqrt{2\pi}}{4} m^{1/2} + \frac{7}{6}\right) + \frac{1}{2}.$$

$$\mathbb{E}(\pi_1^2) \sim b^2\left(\frac{m^2}{3} - \frac{\sqrt{2\pi}}{4} m^{3/2} + \frac{4}{3}m\right) + \frac{b}{2}m.$$

$$\mathbb{E}(\pi_1\pi_2) \sim b^2\left(\frac{m^2}{4} - \frac{\sqrt{2\pi}}{4} m^{3/2} + \frac{3}{2}m\right) + \frac{b}{2}m.$$

These asymptotic results are in sharp contrast with the case $a = cm + b$ where $c > 0$. As $c \to 0$, the correction terms blow up, contributing to the different asymptotic orders between the generic situation $a \gtrsim b$ (corresponding to $c > 0$) and the special situation $a = b$ (corresponding to $c = 0$). Paralleling Proposition 13, the following asymptotics are immediate.

▶ **Proposition 16.** *Take m large. Take $b \geq 1$ any integer. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from PF(b, b, m), we have*

$$\mathbb{V}\mathrm{ar}(\pi_1) \sim \frac{b^2}{12}m^2 + \frac{b^2(4-3\pi)}{24}m, \qquad \mathbb{C}\mathrm{ov}(\pi_1, \pi_2) \sim \frac{b^2(8-3\pi)}{24}m.$$

For $\boldsymbol{\pi} = (\pi_1, \dots, \pi_m) \in \mathrm{PF}(a, b, m)$, the (a, b)-displacement $\mathrm{disp}^{(a,b)}(\boldsymbol{\pi})$ is defined as

$$\mathrm{disp}^{(a,b)}(\boldsymbol{\pi}) = b\binom{m}{2} + am - (\pi_1 + \cdots + \pi_m).$$

Figure 2 shows a histogram of the displacement based on 100,000 random samples of PF$(cm + b, b, m)$ for $m = 100$ and $b = 2$. The left plot ($c = 1$) approximates a normal distribution and the right plot ($c = 0$) approximates an Airy distribution. The displacement definition is in connection with the displacement enumerator of (a, b)-parking functions. Note that the set of (a, b)-parking functions of length m is in bijection with the set of length-a sequences of rooted b-forests on m vertices, and there are related formulations for the (a, b)-inversion and inversion enumerator of length-a sequences of rooted b-forests. See Yan [19] for more details.

▶ **Theorem 17.** *Take m large. Take $b \geq 1$ any integer and $a = cm + b$ for some $c > 0$. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from PF(a, b, m), we have*

$$\mathbb{E}(\mathit{disp}^{(a,b)}(\boldsymbol{\pi})) \sim \frac{cm^2}{2} + \left(\frac{b^2}{2c} + \frac{b}{2} - \frac{1}{2}\right)m, \qquad \mathbb{V}\mathrm{ar}(\mathit{disp}^{(a,b)}(\boldsymbol{\pi})) \sim \frac{(b+c)^2}{12}m^3.$$

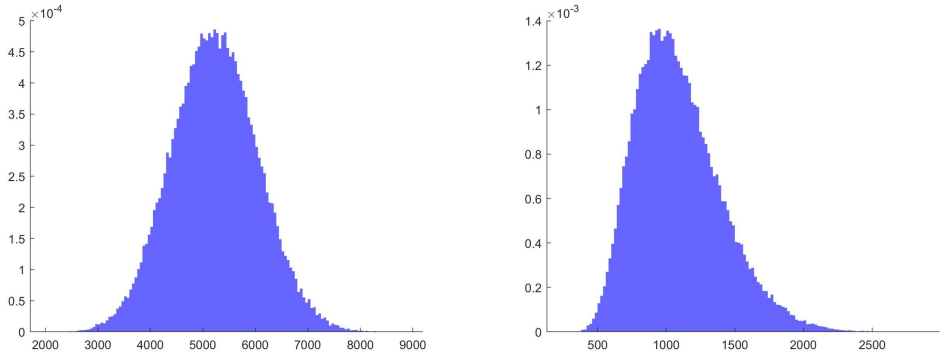


Figure 2 The distribution of displacement in 100,000 samples of $(cm+b, b)$-parking functions chosen uniformly at random, where $m = 100$ and $b = 2$. In the left plot $c = 1$ and in the right plot $c = 0$.

On the other hand, when $c = 0$ and so $a = b$,

$$\mathbb{E}(disp^{(a,b)}(\boldsymbol{\pi})) \sim b\frac{\sqrt{2\pi}}{4}m^{3/2} - \left(\frac{2b}{3} + \frac{1}{2}\right)m, \qquad \mathbb{V}\mathrm{ar}(disp^{(a,b)}(\boldsymbol{\pi})) \sim \frac{b^2(10-3\pi)}{24}m^3.$$

3.3 Boundary behavior of a single coordinate

In this subsection we examine the boundary behavior of a single coordinate of (a, b)-parking functions of length m when $a = cm + b$ for some $c \geq 0$. As in the case of the distribution of multiple coordinates, the asymptotic tendency in the generic situation $c > 0$ and the special situation $c = 0$ are strikingly different. Calculational techniques (tree function, Abel's multinomial theorem) employed in Sections 3.1 and 3.2 will be used in our investigation, with details omitted.

▶ **Proposition 18.** *Let $1 \leq j \leq a+(m-1)b$. The number of parking functions $\boldsymbol{\pi} \in PF(a, b, m)$ with $\pi_1 = j$ is*

$$ab \sum_{s:\, a+sb\geq j} \binom{m-1}{s}(a+sb)^{s-1}((m-s)b)^{m-2-s}.$$

Note that this quantity stays constant for $j \leq a$ and decreases as j increases past a as there are fewer resulting summands.

Recall from Proposition 1 that if $A_{\pi_2,\ldots,\pi_m} = [v]$ is non-empty, then $v = a + sb$ for some $0 \leq s \leq m-1$. Let X be a random variable satisfying the Borel distribution with parameter μ $(0 \leq \mu \leq 1)$, that is, with pmf given by, for $j = 1, 2, \ldots,$

$$\mathbb{P}_\mu(X = j) = \frac{e^{-\mu j}(\mu j)^{j-1}}{j!}.$$

Denote by $\mathbb{Q}_\mu(j) = \mathbb{P}_\mu(X \geq j)$. We refer to Stanley [18] for some nice properties of this discrete distribution.

▶ **Corollary 19.** *Fix j and take m large relative to j. Take $b \geq 1$ any integer and $a = cm + b$ for some $c \geq 0$. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from $PF(a, b, m)$, we have*

$$\mathbb{P}(\pi_1 = a + (m-1)b - j) \sim \frac{1 - \mathbb{Q}_{b/(b+c)}(\lfloor j/b \rfloor + 2)}{(b+c)m},$$

where $\mathbb{Q}_{b/(b+c)}(l) = \mathbb{P}_{b/(b+c)}(X \geq l)$ is the tail distribution function of Borel-$b/(b+c)$.

▶ **Corollary 20.** *Fix j and take m large relative to j. Take $b \geq 1$ any integer and $a = cm + b$ for some $c > 0$. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from $PF(a, b, m)$, we have*

$$\mathbb{P}(\pi_1 = 1) = \cdots = \mathbb{P}(\pi_1 = a) \sim \frac{1}{(b+c)m},$$

and

$$\mathbb{P}(\pi_1 = a + j) \sim e^{\frac{cm}{be} - \frac{b}{b+c}} \left(\frac{b}{b+c}\right)^{m-1} \frac{1}{cm^2} \left(\mathbb{P}(Y \geq \lceil j/b \rceil) - 1\right) + \frac{1}{(b+c)m},$$

where Y is a Poisson$((cm)/(be))$ random variable.

▶ **Corollary 21.** *Fix j and take m large relative to j. Take $b \geq 1$ any integer. For parking function $\boldsymbol{\pi}$ chosen uniformly at random from $PF(b, b, m)$, we have*

$$\mathbb{P}(\pi_1 = 1) = \cdots = \mathbb{P}(\pi_1 = b) \sim \frac{2}{bm},$$

and

$$\mathbb{P}(\pi_1 = b + j) \sim \frac{1 + \mathbb{Q}_1(\lceil j/b \rceil + 1)}{bm},$$

where $\mathbb{Q}_1(l) = \mathbb{P}_1(X \geq l)$ is the tail distribution function of Borel-1.

4 Final remarks

This paper is a part of an ongoing research direction to answer what random parking functions and their generalizations look like. We consider a generalization of parking functions known as **u**-parking functions, which are defined for an arbitrarily fixed positive integer-valued vector **u**. Since **u**-parking functions include ordinary parking functions, the results in this paper can be specialized to recover other recent results about ordinary parking functions. As an application, we study a specific class of **u**-parking functions when $\mathbf{u} = (a, a + b, \ldots, a + (m-1)b)$ is a vector of length m. Such **u**-parking functions are commonly referred to as (a, b)-parking functions and have generated much interest because of their association with particularly nice Gončarov polynomials. We identify a striking contrast between the generic situation $a \gtrsim b$ and the special situation $a = b$ in various asymptotic probabilistic quantities.

Our asymptotic investigations in this work have concentrated mostly on the law of multiple coordinates and displacement of uniformly random (a, b)-parking functions, but there are many other structural properties of parking functions worth studying. For example, descent patterns, cycle counts, and equality processes, among others. The asymptotic distribution of some of these quantities have been explored in [2, 5, 13] for ordinary parking functions, but more generalized models are yet to be examined. Such models include **u**-parking functions studied in this paper, as well as parking functions on mappings, on graphs, and the effect of group action on parking functions. These topics have amazing connections to many areas in math and computer science. Particularly for the field of analytic combinatorics, it would be interesting to see if there is a generic generating function approach to these random combinatorial structures. The multi-dimensional tree function method utilized in this paper is only one small step on a long journey.

References

1 Ayomikun Adeniran, Steve Butler, Galen Dorpalen-Barry, Pamela E. Harris, Cyrus Hettle, Qingzhong Liang, Jeremy L. Martin, and Hayan Nam. Enumerating parking completions using join and split. *Electron. J. Comb.*, 27(2), 2020. URL: https://www.combinatorics.org/ojs/index.php/eljc/article/view/v27i2p44.

2 Peter J. Cameron, Daniel Johannsen, Thomas Prellberg, and Pascal Schweitzer. Counting defective parking functions. *Electron. J. Comb.*, 15(1), 2008. URL: http://www.combinatorics.org/Volume_15/Abstracts/v15i1r92.html.

3 Philippe Chassaing and Jean-François Marckert. Parking functions, empirical processes, and the width of rooted labeled trees. *Electron. J. Comb.*, 8(1), 2001. URL: http://www.combinatorics.org/Volume_8/Abstracts/v8i1r14.html.

4 Robert Cori and Dominique Rossin. On the sandpile group of dual graphs. *Eur. J. Comb.*, 21(4):447–459, 2000. doi:10.1006/eujc.1999.0366.

5 Persi Diaconis and Angela Hicks. Probabilizing parking functions. *Adv. Appl. Math.*, 89:125–155, 2017. doi:10.1016/j.aam.2017.05.004.

6 D. Foata and J. Riordan. Mappings of acyclic and parking functions. *Aequationes Math.*, 10:10–22, 1974.

7 J. Haglund, M. Haiman, N. Loehr, J. B. Remmel, and A. Ulyanov. A combinatorial formula for the characters of the diagonal coinvariants. *Duke Math. J.*, 126:195–232, 2005.

8 R. Kenyon and M. Yin. Parking functions: From combinatorics to probability. *arXiv preprint*, 2021. arXiv:2103.17180.

9 Donald Ervin Knuth. *The Art of Computer Programming: Sorting and Searching*, volume 3. Addison-Wesley, 1998. URL: https://www.worldcat.org/oclc/312994415.

10 A. G. Konheim and B. Weiss. An occupancy discipline and applications. *SIAM J. Appl. Math.*, 14(6):1266–1274, 1966. URL: http://www.jstor.org/stable/2946240.

11 Joseph P. S. Kung and Catherine H. Yan. Expected sums of general parking functions. *Ann. Comb.*, 7:481–493, 2003. doi:10.1007/s00026-003-0198-7.

12 Joseph P. S. Kung and Catherine H. Yan. Gončarov polynomials and parking functions. *J. Comb. Theory, Ser. A*, 102(1):16–37, 2003. doi:10.1016/S0097-3165(03)00009-8.

13 J. E. Paguyo. Cycle structure of random parking functions. *arXiv preprint*, 2022. arXiv:2202.08829.

14 Jim Pitman. Forest volume decompositions and abel-cayley-hurwitz multinomial expansions. *J. Comb. Theory, Ser. A*, 98(1):175–191, 2002. doi:10.1006/jcta.2001.3238.

15 J. Riordan. *Combinatorial Identities*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1968. URL: https://books.google.com/books?id=X6ccBWwECP8C.

16 Richard P. Stanley. Parking functions and noncrossing partitions. *Electron. J. Comb.*, 4(2), 1997. URL: http://www.combinatorics.org/Volume_4/Abstracts/v4i2r20ab.html.

17 Richard P. Stanley. Hyperplane arrangements, parking functions and tree inversions. In *Mathematical essays in honor of Gian-Carlo Rota*, volume 161 of *Progr. Math.*, pages 359–375. Birkhäuser Boston, Boston, MA, 1998.

18 Richard P. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, 1999. URL: http://math.mit.edu/~rstan/ec/.

19 Catherine H. Yan. Generalized parking functions, tree inversions, and multicolored graphs. *Adv. Appl. Math.*, 27(2–3):641–670, 2001. doi:10.1006/aama.2001.0754.

20 Catherine H. Yan. Parking functions. In *Handbook of Enumerative Combinatorics*, Discrete Math. Appl., pages 835–893. CRC Press, Boca Raton, FL, 2015.

21 M. Yin. Parking functions: Interdisciplinary connections. *arXiv preprint*, 2021. arXiv:2107.01767.