# Statistical properties of lambda terms

Maciej Bendkowski[*]

Theoretical Computer Science Department
Faculty of Mathematics and Computer Science
Jagiellonian University
ul. Prof. Łojasiewicza 6, 30-348 Kraków, Poland.

maciej.bendkowski@tcs.uj.edu.pl

Olivier Bodini[†]      Sergey Dovgal[†]

Universitè Paris 13
Sorbonne Paris Citè
LIPN CNRS UMR 7030
99 avenue Jean-Baptiste Clèment, 93430 Villetaneuse, France.

{olivier.bodini,dovgal}@lipn.univ-paris13.fr

**Abstract**

We present a quantitative, statistical analysis of random lambda terms in the De Bruijn notation. Following an analytic approach using multivariate generating functions, we investigate the distribution of various combinatorial parameters of random open and closed lambda terms, including the number of redexes, head abstractions, free variables or the De Bruijn index value profile. Moreover, we conduct an average-case complexity analysis of finding the leftmost-outermost redex in random lambda terms showing that it is on average constant. The main technical ingredient of our analysis is a novel method of dealing with combinatorial parameters inside certain infinite, algebraic systems of multivariate generating functions. Finally, we briefly discuss the random generation of lambda terms following a given skewed parameter distribution and provide empirical results regarding a series of more involved combinatorial parameters such as the number of open subterms and binding abstractions in closed lambda terms.

**Mathematics Subject Classifications:** 03B40, 03B70, 05A16, 05A99, 62E20, 68R05, 68W40.

# Contents

# 1  Introduction

Lambda calculus (often abbreviated to $\lambda$-calculus) is a functional calculus proposed by Alonzo Church in the 1930s as an alternative foundation of mathematics. Although the initial plan failed, due to logical inconsistencies discovered in Church's naive system, it was quickly realised that $\lambda$-calculus itself is able to elegantly capture the, by then still informal, notion of computability, see [CH09]. Nowadays, $\lambda$-calculus is considered not only as an important theoretical model of computation, but is also used in practical applications ranging from functional programming languages [PJ87], including the evaluation and testing of functional programming language compilers [CH00, Pal12], to automated theorem provers [BC10].

Despite the extensive use of $\lambda$-terms (i.e. formal expressions of $\lambda$-calculus) as computations in functional programming languages or as components of proof artifacts in various automated theorem provers, quantitative investigations into the combinatorial or statistical properties of $\lambda$-terms were initialised only quite recently. Motivated by the uniformly random (conditioned on size) generation of $\lambda$-terms, Wang [Wan05] explored a combinatorial model of $\lambda$-calculus where $\alpha$-convertible $\lambda$-terms (i.e. terms identical up to bound variable names) are considered equivalent. The central problem of providing asymptotic estimates on the number of $\lambda$-terms in this model remained, however, open. Some time later, David et al. [DGK$^+$13] investigated a similar model of $\lambda$-calculus where variables do not contribute to the term size and showed that asymptotically almost all $\lambda$-terms are strongly normalising. In other words, the fraction of $\lambda$-terms for which all evaluation strategies terminate approaches one as the term size tends to infinity. Likewise, in this model the central problem of giving accurate estimates on the number of $\lambda$-terms of size $n$ remained open. Enumeration problems for restricted classes of closed affine and linear $\lambda$-terms, where binders capture at most and exactly one variable, respectively, were investigated by Bodini, Gardy, Jacquot and Gittenberger [BGGJ13, BGJ13, BG14]. The class of $\lambda$-terms with restricted unary height was considered by Bodini, Gardy and Gittenberger in [BGG11].

The canonical models of Wang and David et al. pose considerable difficulties due to the global, intractable structure of binders (abstractions) and their associated variables, all considered modulo $\alpha$-equivalence. Explicit variable names, though elegant for manual manipulation, introduce also problems with substitution of terms for variables,

especially when computations in $\lambda$-calculus are meant to be automatised. For the latter purpose, De Bruijn proposed an alternative notation of $\lambda$-terms, involving non-negative indices instead of explicit variable names [dB72]. This notation was later adopted by Lescanne [Les13, GL13] who proposed a new combinatorial representation for the enumeration of $\lambda$-terms. Within this new representation, $\lambda$-terms represent entire $\alpha$-equivalence classes in the former models. Consequently, it became possible to enumerate also open terms (i.e. containing free variables) not just closed ones. Let us also remark that independently, a different combinatorial model was proposed by Tromp who considered a binary encoding of $\lambda$-calculus meant for the construction of a compact and efficient self-interpreter with applications to Kolmogorov complexity [Tro06]. Enumeration problems related to the binary $\lambda$-calculus, as well as the effective random generation of $\lambda$-terms, were later studied by Grygiel and Lescanne [GL15].

Investigations into quantitative properties of $\lambda$-terms in the De Bruijn notation were continued by Bendkowski et al. [BGLZ16, BGLZ17] who showed that, in contrast to the canonical representation of David et al., asymptotically almost all $\lambda$-terms are not strongly normalising. In other words, the proportion of terms for which all evaluation strategies terminate approaches zero as the term size tends to infinity. Various size models based on the De Bruijn notation, such as Tromp's binary encoding or the so-called natural size notion introduced by Bendkowski et al. were later generalised in a common framework by Gittenberger and Gołębiewski who provided tight lower and upper asymptotic bounds on the number of closed $\lambda$-terms [GG16]. Recently, the gap between both the lower and bounds was closed by Bodini, Gittenberger and Gołębiewski [BGG18]. Subsequently, efficient sampling methods for closed terms were developed and the enumeration of closed $\lambda$-terms was finally completed.

In the current paper we propose to deepen the quantitative analysis of $\lambda$-calculus in the De Bruijn notation, offering a detailed statistical analysis of random $\lambda$-terms. We investigate the distribution of several combinatorial parameters related to plain (i.e. unrestricted) and closed $\lambda$-terms. Table 1 provides a brief overview of the obtained results.

In the current paper, we provide limit laws and asymptotic estimates using techniques from analytic combinatorics. While plain $\lambda$-terms in De Bruijn size notion can be analysed using classical methods, the respective analysis of closed $\lambda$-terms requires solving infinite systems of algebraic equations. Let us mention that an earlier paper by Drmota, Gittenberger and Morgenbesser [DGM12] deals with infinite algebraic systems which are strongly connected and whose Jacobian can be represented as a sum of a scaled identity operator and an operator whose power is compact. Here, we develop a general tool meant for analysis of infinite algebraic systems that resemble in structure systems for closed $\lambda$-terms, however do not fit into the framework of Drmota, Gittenberger and Morgenbesser. In this context, our result can be considered as a continuation of [DGM12].

The paper is structured as follows. In Section 2 we provide a concise presentation of preliminary notions and techniques. In particular, we discuss the De Bruijn representation of $\lambda$-terms and explain the rules of lambda calculus. We then continue with a fairly standard analysis of basic statistics corresponding to plain $\lambda$-terms (Section 3). In the subsequent section we use our advanced marking techniques and study various

Table 1: Comparison of obtained statistics for random $\lambda$-terms. Listed constants are approximated up to the third decimal point. See Section 2 for details on trivial.

| Parameter | Mean, $\sim$ | | Distribution | |
|---|---|---|---|---|
| | plain | closed | plain | closed |
| Variables | $0.307n$ | | Normal | |
| Abstractions | $0.258n$ | | Normal | |
| Successors | $0.129n$ | | Normal | |
| Redexes | $0.091n$ | | Normal | |
| Index value | $0.420$ | | Geometric | |
| Redex search time | 6.222 | 6.054 | Discrete | Discrete |
| Head abstractions | 0.420 | 1.447 | Geometric | Discrete |
| $m$-openness | 2.019 | 0 | Discrete | trivial |
| Free variables | 5.722 | 0 | Discrete | trivial |
| Unary height profile | $0.122\sqrt{n}$ | | Rayleigh | |
| Natural height profile | $0.412\sqrt{n}$ | | Rayleigh | |

parameters in closed $\lambda$-terms (Section 4). Next, we develop our main technical tool for investigating combinatorial parameters in closed $\lambda$-terms (Section 5) used in the previous section. Afterwards, we provide an empirical evaluation of several statistical properties corresponding to plain, closed, and so-called $h$-shallow $\lambda$-terms, i.e. terms with De Bruijn indices whose value does not exceed $h$ (Section 6). We give empirical histograms and relate the discovered distributions with considered term types, exhibiting some intriguing correlations. Finally, we conclude the paper with remarks and open questions (Section 7). For the sake of completeness, we describe standard analytic tools in Appendix A.

## 2   Preliminaries on lambda calculus

$\lambda$-calculus is a theoretical formalism famously equivalent in expressiveness to Turing machines, see [Bar84]. In this calculus, computations are represented as $\lambda$-*terms* defined by the formal grammar $T ::= x \mid (\lambda x.T) \mid (T\,T)$ in which $x$ belongs to the countable, infinite alphabet of *variables*; $(\lambda x.T)$ is an *abstraction* of variable $x$ in $T$; and $(T\,T)$ denotes an *application* of two $\lambda$-terms. Given an abstraction $(\lambda x.T)$, occurrences of $x$ in $T$ are said to be *bound*. Unbound variable occurrences are said to occur *freely*.

Lambda terms, intended to represent anonymous functions, are executed by means of the iterated process of $\beta$-*reduction*. First, an arbitrary $\beta$-*redex* subterm in form of $(\lambda x.N)M$ is selected (if no such subterm exists, computations are terminated). Next, the selected $\beta$-redex is replaced with $N[x := M]$, i.e. $N$ in which each occurrence of $x$ is substituted, in a *capture-avoiding* manner, by $M$. While substituting $M$ for $x$ in $N$ we have to avoid the unintended situation in which free variable occurrences in $M$ get bound, in other words *captured*, by some abstractions occurring in $N$. For instance, let $N = (\lambda y.x)$ and $M = y$. The term $(\lambda x.N)M$ should not be reduced to $\lambda y.y$ as, by doing so, the free variable occurrence $y$ gets bound due to a coincidental *clash* with the inner

abstraction variable name. Certainly, the arbitrary choice of the formal variable name $y$ should not influence the intended semantics of the represented computation. Following this motivation, $\lambda$-terms differing only in bound variable names are considered equivalent (in other words $\alpha$-*convertible*). In order to avoid potential name clashes, we can therefore *rename* bound variable occurrences before proceeding with $\beta$-reduction. Since there is an infinite supply of available variable names, it is always possible to avoid variable captures. Consequently, we can equivalently $\alpha$-convert $(\lambda x.\lambda y.x)y$ into, say, $(\lambda x.\lambda w.x)y$ and proceed with $(\lambda x.\lambda w.x)y \to_\beta (\lambda w.x)[x := y] = \lambda w.y$.

Though intuitive, explicit variable names pose considerable conceptual and implementation problems. For instance, consider the terms $\lambda x.x$ and $\lambda y.y$. Although syntactically different, both semantically represent the same anonymous identity function as $(\lambda x.x)T \to_\beta T$ and $(\lambda y.y)T \to_\beta T$ for arbitrary $T$. In order to facilitate automatic computations, De Bruijn proposed a different notation for $\lambda$-terms eliminating in effect the troublesome variable names [dB72]. In his notation, variable occurrences are replaced with *indices* represented as non-negative integers. The intention is to view $\lambda$-terms as natural tree-like structures and encode variable occurrences as indices denoting their relative distance to respective variable binders – each index $\underline{\mathtt{n}}$ denotes a variable occurrence $x$ whose relative distance to its binder (in terms of passed lambda symbols) is equal to $n+1$. For instance, $\underline{\mathtt{0}}$ corresponds to a variable occurrence bound to the nearest abstraction on its unique path to the root in the associated tree-like representation of the considered $\lambda$-term. Consequently, $\alpha$-convertible $\lambda$-terms have the same De Bruijn representation. In effect, each $\lambda$-term in the De Bruijn notation represents an entire $\alpha$-equivalence class of $\lambda$-terms in the classic variable notation. For instance, both $\lambda x.x$ and $\lambda y.y$, being $\alpha$-convertible, are represented as $\lambda\underline{\mathtt{0}}$ in the De Bruijn notation. The use of De Bruijn indices significantly simplifies the automatic substitution operation. Due to the lack of explicit variable names, variable captures and name clashes do not pose implementation issues.

*Remark* 2.1. There exists a disagreement in the literature whether to start De Bruijn indices with $\underline{\mathtt{0}}$ or $\underline{\mathtt{1}}$. Although De Bruijn himself assumed the latter [dB72], some authors follow his footsteps, see e.g [GL13, GL15] whereas others do not, including $\underline{\mathtt{0}}$ in the set of admissible indices, see e.g. [GG16, BGLZ17, BT18]. Certainly, neither convention is better than the other. In the current paper, we follow the convention of starting De Bruijn indices with $\underline{\mathtt{0}}$ so the keep consistent with the most recent literature.

**Definition 2.2.** Let $\{\underline{\mathtt{0}}, \underline{\mathtt{1}}, \ldots\}$ be an infinite, denumerable set of available indices. Then, the set $\mathcal{L}_\infty$ of $\lambda$-*terms in the De Bruijn notation* is defined inductively as follows:

  a) Each index $\underline{\mathtt{n}}$ is a $\lambda$-term;

  b) If $N$ and $M$ are $\lambda$-terms, then $(NM)$ is a $\lambda$-term;

  c) If $N$ is a $\lambda$-term, then $(\lambda N)$ is a $\lambda$-term.

Following usual notational conventions, we omit outermost parentheses and drop parentheses from left-associated $\lambda$-terms. For instance, $\lambda x.\lambda y.\lambda z.((xy)z)$ in the classical variable notation is written as $\lambda\lambda\lambda\underline{\mathtt{210}}$.

An index occurrence $\underline{n}$ is said to be *bound* in the term $N$ if there exist at least $n + 1$ lambda symbols on the unique path from $\underline{n}$ to the root of the associated tree-like representation of $N$, see e.g. Figure 1. Otherwise, $\underline{n}$ is said to be occurring *freely* in $N$ and hence corresponds to a free variable in the classical $\lambda$-calculus notation. For convenience, we refer to De Bruijn indices both as indices and variables. If each index occurrence in $N$ is bound, then $N$ is said to be *closed*. Otherwise, it is said to be *open*. And so, $\lambda\lambda\lambda\underline{210}$ is closed whereas $\lambda\lambda\underline{21}$ is not as here $\underline{2}$ is not bound. If prepending $N$ with $m$ lambdas turns it into a closed $\lambda$-term, then $N$ is said to be *m-open*. Certainly, if $N$ is $m$-open, then it is also $(m+1)$-open. Moreover, 0-open $\lambda$-terms correspond exactly to closed $\lambda$-terms. Hence, though $\lambda\lambda\underline{21}$ is not closed, it is 1-open as $\lambda\lambda\lambda\underline{21}$ is a closed $\lambda$-term. Finally, we write that a $\lambda$-term is *plain* if we mean to denote that it is either open or closed, without specifying which case holds.
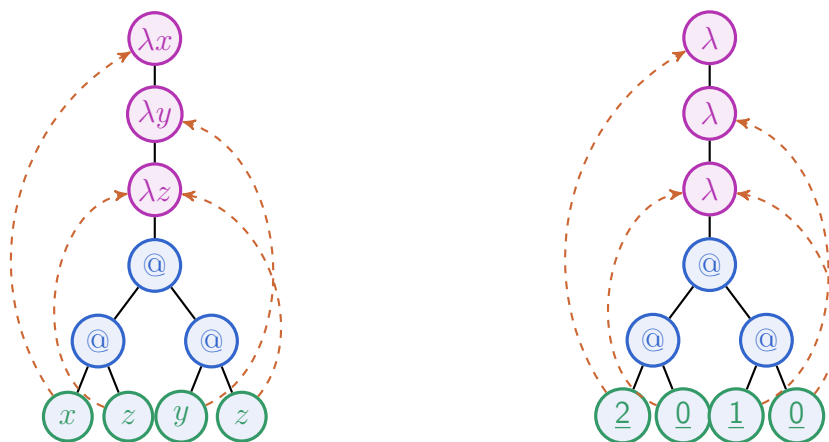


Figure 1: Two tree-like representations associated with the same example $\lambda$-term — $\lambda x.\lambda y.\lambda z.xz(yz)$ and its De Bruijn notation variant $\lambda\lambda\lambda\underline{20}(\underline{10})$. Back pointers to abstractions are included for illustrative purposes only.

## 2.1   Enumeration

In the current paper we follow [BGLZ16, BGLZ17, GG16, BGG18] and investigate the statistical properties of random $\lambda$-terms in the De Bruijn representation. We assume a unary base encoding of indices, i.e. an encoding in which $\underline{n}$ is identified with an $n$-fold application of the *successor* operator $\mathsf{S}$ to zero. Formally, the set $\mathcal{L}_\infty$ of $\lambda$-terms is described by the following formal grammar:

$$\begin{aligned}
\mathcal{L}_\infty &::= \underline{n} \mid (\lambda\mathcal{L}_\infty) \mid (\mathcal{L}_\infty\mathcal{L}_\infty) \\
\underline{n} &::= 0 \mid \mathsf{S}\,\underline{n}.
\end{aligned}$$

(2.1)

In order to enumerate $\lambda$-terms, we have to assign a formal notion of *size* to each term in such a manner that for each available size $n$ the number of terms of size $n$ is finite. Though various size measures are considered in the literature, most notably the general

size model framework of Gittenberger and Gołębiewski [GG16], we assume the simple *natural size notion* [BGLZ16] in which the size of $T$ is equal to the total number of abstractions, applications, successors and zeros occurring in $T$. Formally, we define the size of $T$ inductively as follows:

$$\begin{array}{rclcrcl} |0| & = & 1 & \qquad & |M\,N| & = & |M| + |N| + 1 \\ |\mathsf{S}\,\underline{n}| & = & |\underline{n}| + 1 & & |\lambda M| & = & |M| + 1. \end{array} \qquad (2.2)$$

**Example 2.3.** Note that, in general, $\underline{n}$ is of size $n+1$ as it consists of $n$ successors applied to zero. Consequently, the term $\lambda\lambda\lambda\underline{210}$ is of size 11 as it consists of three $\lambda$ symbols, two applications between $\underline{2},\underline{1}$ and $\underline{0}$, by convention omitted in writing, and indices $\underline{0},\underline{1},\underline{2}$ of total size six.

*Remark* 2.4. It is worth noticing that, with some minor technical overhead, the analysis presented in the current paper extends onto the more general size model framework of Gittenberger and Gołębiewski [GG16] including the assumed natural size notion as a special case. We prefer to avoid technicalities related to the general size notion and so, for the reader's convenience, favour a lucid presentation of the key arguments.

Let $\mathcal{L}_m$ denote the set of $m$-open $\lambda$-terms, see Theorem 2.2 (plain terms can be viewed as "infinitely" open, hence the $\infty$ symbol in the subscript of $\mathcal{L}_\infty$). Like plain $\lambda$-terms (2.1), $\mathcal{L}_m$ can be described in terms of a formal, though now infinite, grammar as follows:

$$\begin{array}{l} \mathcal{L}_m ::= (\lambda\mathcal{L}_{m+1}) \mid (\mathcal{L}_m\mathcal{L}_m) \mid \underline{0},\underline{1},\dots,\underline{\mathsf{m}\text{ - }\mathsf{1}} \\ \mathcal{L}_{m+1} ::= (\lambda\mathcal{L}_{m+2}) \mid (\mathcal{L}_{m+1}\mathcal{L}_{m+1}) \mid \underline{0},\underline{1},\dots,\underline{\mathsf{m}} \\ \qquad \dots \qquad \dots \end{array} \qquad (2.3)$$

An $m$-open $\lambda$-term $T$ can take one of the three forms. Either $T$ is in the form of abstraction followed by an $(m+1)$-open $\lambda$-term; or it is an application of two $m$-open $\lambda$-terms; or, finally, $T$ is one of the indices $\underline{0},\underline{1},\dots,\underline{\mathsf{m}\text{-}\mathsf{1}}$.

Due to the infinite combinatorial specification (2.3) for $\mathcal{L}_m$ standard analytic combinatorics techniques are not readily applicable. Consequently, enumerating closed $\lambda$-terms poses a considerable challenge. In [GG16] a partial solution bounding the asymptotic growth rate of the number of $m$-open $\lambda$-terms of size $n$ was proposed. Although both the lower and upper bounds were of the form $C\rho^n n^{-3/2}$, a typical trait of various tree-like structures, the asymptotic growth rate of $m$-open terms remained open. Remarkably, some time later in their joint paper [BGG18] Bodini, Gittenberger and Gołębiewski closed the remaining gap and confirmed the conjectured $C\rho^n n^{-3/2}$ form of the asymptotic growth of $m$-open $\lambda$-terms. Furthermore, two combinatorial problems related to random closed $\lambda$-terms were studied. Specifically, the number of terms with an *a priori* fixed number of abstractions and the number of terms in $\beta$-normal form, i.e. without $\beta$-redexes. In this context, our contribution is a natural continuation of their work. In addition, we offer a different proof of the asymptotic growth rate of $m$-open $\lambda$-terms.

# 3 Basic statistics of plain lambda terms

In this section we investigate several basic combinatorial parameters related to random plain $\lambda$-terms. Let us start with invoking the combinatorial specification (2.1) describing the set $\mathcal{L}_\infty$ of plain $\lambda$-terms. Recall that $\mathcal{L}_\infty$ is specified as

$$
\begin{aligned}
\mathcal{L}_\infty &::= \underline{\mathsf{n}} \mid \lambda\mathcal{L}_\infty \mid (\mathcal{L}_\infty\mathcal{L}_\infty) \\
\underline{\mathsf{n}} &::= 0 \mid \mathsf{S}\ \underline{\mathsf{n}}.
\end{aligned} \tag{3.1}
$$

Equivalently, the set $\mathcal{L}_\infty$ of $\lambda$-terms can be specified using the following pictorial tree grammar (note the explicit @ symbol for term application):
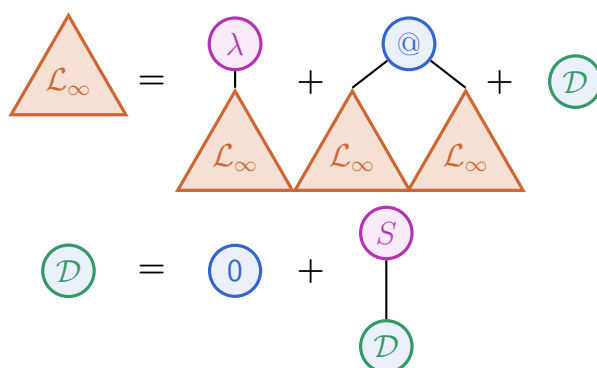


Figure 2: Combinatorial specification for plain $\lambda$-terms.

Following symbolic methods [FS09, Part A: Symbolic Methods] we note that the generating function $D(z)$ corresponding to De Bruijn indices takes the form $D(z) = \dfrac{z}{1-z}$ and so the generating function $L_\infty(z)$ associated with plain $\lambda$-terms satisfies the following functional equation:

$$
L_\infty(z) = zL_\infty(z) + zL_\infty(z)^2 + \frac{z}{1-z}. \tag{3.2}
$$

Solving (3.2) for $L_\infty(z)$ we obtain two formal solutions. Since we know *a priori* that the resulting generating function has non-negative coefficients $[z^n]L_\infty(z)$ we conclude that

$$
L_\infty(z) = \frac{1}{2z}\left(1 - z - \sqrt{(1-z)^2 - \frac{4z^2}{1-z}}\right). \tag{3.3}
$$

In this form, we can easily verify that the radicand expression $(1-z)^2 - \dfrac{4z^2}{1-z}$ carries the single dominant square-root type singularity $\rho$ of $L_\infty(z)$. At this point, a straightforward application of the transfer theorem (see Proposition A.1) gives us access to the asymptotic growth rate of the counting sequence corresponding to plain $\lambda$-terms.

**Proposition 3.1** (see [BGLZ16]). Let $L_\infty(z)$ be the generating function associated with plain $\lambda$-terms (3.3). Then, the number $[z^n]L_\infty(z)$ of plain terms of size $n$ admits the following asymptotic approximation as $n \to \infty$:

$$[z^n]L_\infty(z) \sim C\rho^{-n}n^{-3/2} \tag{3.4}$$

where

$$\rho \approx 0.29559774 \quad \text{and} \quad C \approx 0.606767. \tag{3.5}$$

## 3.1 Joint distribution of variables, abstractions, successors and redexes

Basic marking techniques allow us to investigate limiting distributions of various sub-patterns in plain $\lambda$-terms. In what follows we study the fundamental patterns of variables, abstractions, successors and $\beta$-redexes. Recall that a $\beta$-redex is a $\lambda$-term in form of $(\lambda N)M$ where $N$ and $M$ arbitrary $\lambda$-terms. In other words, redex is a sub-pattern which can be depicted as  .

**Proposition 3.2.** The multivariate generating function $L_\infty(z, \boldsymbol{u})$ for plain $\lambda$-terms with the auxilliary vector $\boldsymbol{u} = (u_{(\mathsf{var})}, u_{(\mathsf{red})}, u_{(\mathsf{suc})}, u_{(\mathsf{abs})})$ marking, respectively, variables, redexes, successors and abstractions in a lambda term, satisfies a system of algebraic functional equations

$$L_\infty(z, \boldsymbol{u}) = u_{(\mathsf{abs})}zL_\infty(z, \boldsymbol{u}) + A(z, \boldsymbol{u})$$
$$A(z, \boldsymbol{u}) = \frac{u_{(\mathsf{var})}z}{1 - u_{(\mathsf{suc})}z} + u_{(\mathsf{red})}u_{(\mathsf{abs})}z^2 L_\infty(z, \boldsymbol{u})^2 + zA(z, \boldsymbol{u})L_\infty(z, \boldsymbol{u}). \tag{3.6}$$
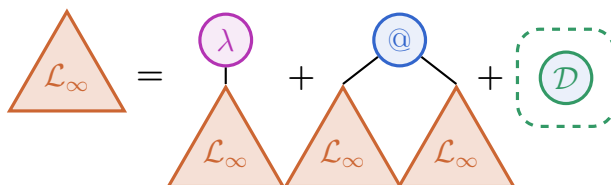


Figure 3: Marking variables in plain terms.

*Proof.* Let us consider a bivariate generating function $L_\infty(z, u)$ in which $[z^n u^k]L_\infty(z, u)$, i.e. the coefficient standing by $z^n u^k$, denotes the number of plain $\lambda$-terms of size $n$ with $k$ variables (equivalently $k$ occurrences of $\mathbf{0}$). Marking all occurrences of $\mathbf{0}$ in the defining equation (3.2) of $L_\infty(z)$, see Figure 3, we obtain the following combinatorial specification for $L_\infty(z, u)$:

$$L_\infty(z, u) = zL_\infty(z, u) + zL_\infty(z, u)^2 + \frac{uz}{1 - z}. \tag{3.7}$$

Consider next the auxiliary class $\mathcal{N}$ consisting of De Bruijn indices and terms in application form. Note that if $N$ is in application form, then either $N$ is a $\beta$-redex in form of $N = (\lambda M)P$, or $N = MP$ where $M$ belongs itself to class $\mathcal{N}$. Furthermore, with $\mathcal{N}$ at

hand, notice that each plain $\lambda$-term $N$ takes either the form $N = \lambda M$ for some $\lambda$-term $M$, or is an element of $\mathcal{N}$. We can therefore write down the following joint combinatorial specification (3.8) for $L_\infty(z, v)$ and $N(z, v)$ corresponding to $\mathcal{L}_\infty$ and $\mathcal{N}$, respectively, using the variable $v$ to mark the redex occurrences in $\mathcal{N}$, see Figure 4:

$$
\begin{aligned}
L_\infty(z, u) &= z L_\infty(z, u) + N(z, u) \\
N(z, u) &= \frac{z}{1 - z} + u z^2 L_\infty(z, u)^2 + z N(z, u) L_\infty(z, u).
\end{aligned}
\tag{3.8}
$$
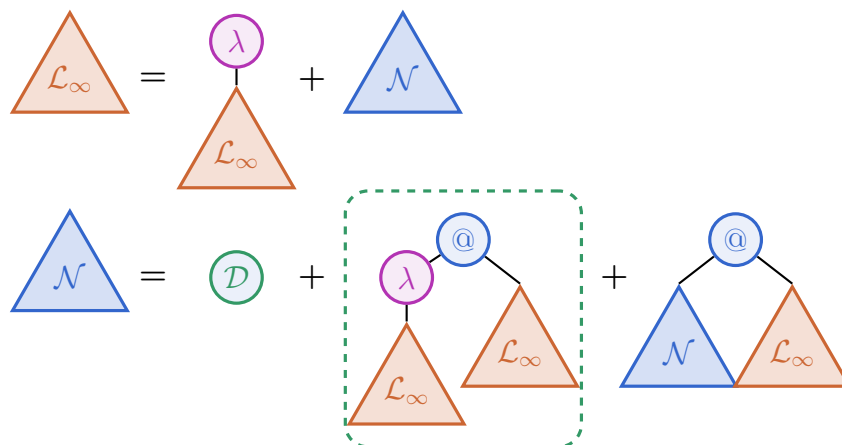


Figure 4: Marking redexes in plain terms.

Finally, in order to simultaneously mark four parameters in plain $\lambda$-terms, we augment the above specification with marking variables for abstractions, variables, and successors. For that purpose we use the following auxiliary variable vector $\boldsymbol{u} = (u_{(\mathsf{var})}, u_{(\mathsf{red})}, u_{(\mathsf{suc})}, u_{(\mathsf{abs})})$ corresponding to respective components of $\boldsymbol{X}_n$, see Figure 5 (cf. Figure 4).

Let $\boldsymbol{u}$ denote the vector of considered marking variables. Such a specification, when converted into a system of functional equations involving the generating functions $L_\infty(z, \boldsymbol{u})$ and $A(z, \boldsymbol{u})$ associated with $\mathcal{L}_\infty$ and $\mathcal{A}$, respectively, yields (3.6). □

**Proposition 3.3.** Let $\boldsymbol{X}_n = (X_{n(\mathsf{var})}, X_{n(\mathsf{red})}, X_{n(\mathsf{suc})}, X_{n(\mathsf{abs})})$ be a random vector denoting

a) the number $X_{n(\mathsf{var})}$ of variables;

b) the number $X_{n(\mathsf{red})}$ of $\beta$-redexes;

c) the number $X_{n(\mathsf{suc})}$ of successors, and

d) the number $X_{n(\mathsf{abs})}$ of abstractions
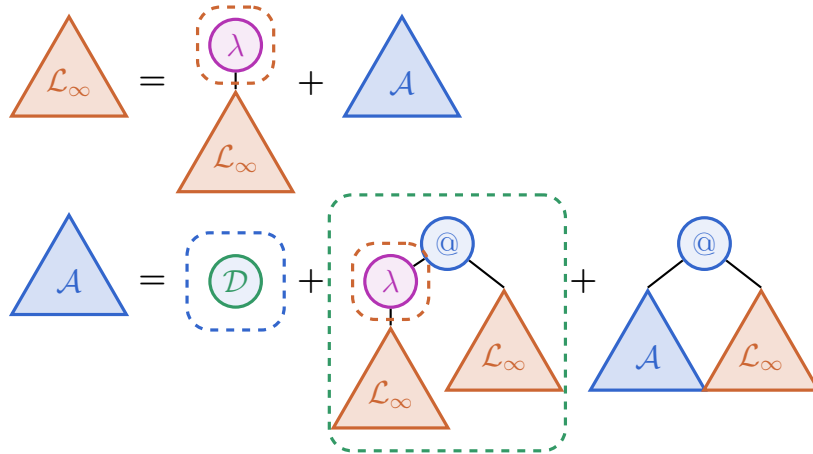
Figure 5: Marking abstractions, variables, successors and redexes in plain $\lambda$-terms.

in a random plane $\lambda$-terms of size $n$. Then, after standardisation, the random vector $\boldsymbol{X}_n$ converges in law to a multivariate Gaussian distribution satisfying (up to numerical approximation)

$$X_n \xrightarrow{d} \mathcal{N}\left( n \begin{pmatrix} 0.307 \\ 0.091 \\ 0.129 \\ 0.258 \end{pmatrix}, n \begin{pmatrix} 0.052 & -0.013 & -0.034 & -0.069 \\ -0.013 & 0.052 & -0.022 & 0.047 \\ -0.034 & -0.022 & 0.145 & -0.076 \\ -0.069 & 0.047 & -0.076 & 0.214 \end{pmatrix} \right). \tag{3.9}$$

*Proof.* We start our investigations with the variable distribution in plain $\lambda$-terms. Consider again the bivariate generating function $L_\infty(z, u)$ with $z$ marking the size of the plain $\lambda$-term, and $u$ marking the number of its variables. Then,

$$L_\infty(z, u) = z L_\infty(z, u) + z L_\infty(z, u)^2 + \frac{uz}{1-z} \tag{3.10}$$

and hence

$$L_\infty(z, u) = \frac{1}{2z}\left( 1 - z - \sqrt{(1-z)^2 - \frac{4uz^2}{1-z}} \right) \tag{3.11}$$

as $L_\infty(z, 1) = L_\infty(z)$.

The dominant singularity $\rho(u)$ of $L_\infty(z, u)$ is the real positive root of the radicand expression $F(z, u) = (1-z)^2 - \dfrac{4uz^2}{1-z}$. Moreover, the singularity has a non-zero derivative at $u = 1$. According to Theorem A.12 (moving singularity framework) this leads to a Gaussian limit law.

The mean and the variance of the resulting normal distribution can be computed by Proposition A.11 using the values $\rho'(1)$ and $\rho''(1)$ from the partial derivatives of $F(z, u)$.

Since $F(\rho(u), u) = 0$, after taking the derivative with respect to $u$ we obtain

$$\rho'(1) = -\frac{\partial_u F(\rho, 1)}{\partial_z F(\rho, 1)} \quad \text{and} \quad \rho''(1) = -\frac{\partial_u^2 F(\rho, u) + 2\rho'(1)\partial_z\partial_u F(\rho, 1) + \rho'(1)^2\partial_z^2 F(\rho, 1)}{\partial_z F(\rho, 1)}.$$
(3.12)

A similar scheme can be adopted for the case of several parameters. We base our proof on the multivariate central limit theorem (see Proposition A.11). Reformulating (3.6) we find that

$$(1 - u_{(\mathsf{abs})}z)L_\infty(z, \boldsymbol{u}) = \left(zu_{(\mathsf{abs})}(u_{(\mathsf{red})} - 1) + 1\right)zL_\infty(z, \boldsymbol{u})^2 + \frac{u_{(\mathsf{var})}z}{1 - u_{(\mathsf{suc})}z}.$$
(3.13)

In this form, the dominant singularity $\rho(\boldsymbol{u})$ of $L_\infty(z, \boldsymbol{u})$ can be computed as the critical value $z = \rho(\boldsymbol{u})$ such that the discriminant of the equation is equal to zero. By applying the multivariate central limit theorem, we obtain the claimed means and covariance matrix, see (3.9).

$\square$

*Remark* 3.4. Let us note that, in the case of limiting distribution for the number of variables in a plain $\lambda$-term, in a general *abcd* size notion, $\rho(u)$ is a root of the polynomial $(1 - z^b)F(z, u)$ whose degree depends on the specific size model parameters $a, b, c, d$ denoting the weights of zero, successor, abstraction and application, respectively, cf. [GG16]. Specifically,

$$(1 - z^b)F(z, u) = (1 - z^b)(1 - z^c)^2 - 4uz^{a+d}.$$
(3.14)

Consequently, for most admissible size notion parameters we cannot explicitly obtain analytic expression of $\rho(u)$. Instead, in order to check the premises of the multivariate central limit theorem we have to work with the implicit equation (3.14).

The main technical obstacle lies in the verification of the requested variability condition $B''(1) + B'(1) - B'(1)^2 \neq 0$, see (A.23). The remaining argumentation is virtually identical to the one presented for the specific case of $a = b = c = d = 1$.

*Remark* 3.5. Arguably, the most interesting part of the covariance matrix (3.9) is the sign of the correlations and the absolute values of associated variances. Note that in the natural size notion, the number of abstractions has greater variance than other constructors. Interestingly, the number of $\beta$-redexes is positively correlated with the number of abstractions. Not surprisingly, all other parameters are negatively correlated.

## 3.2 Head abstractions in plain lambda terms

In this section we turn to head abstractions in plain $\lambda$-terms showing that the corresponding random variable converges to a discrete geometric distribution.

**Proposition 3.6.** Let $X_n$ be a random variable denoting the number of head abstractions in a random plain $\lambda$-term of size $n$. Then, $X_n$ converges in law to a geometric distribution $\mathrm{GEOM}(\rho)$ with parameter $\rho$. Specifically,

$$\mathbb{P}(X_n = h) \xrightarrow[n\to\infty]{} \mathbb{P}(\mathrm{GEOM}(\rho) = h) = (1 - \rho)\rho^h.$$
(3.15)

*Proof.* Note that each $\lambda$-term starts with a (possibly empty) sequence of consecutive head abstractions followed either by a De Bruijn index or an application of two terms (recall that abstractions therein are no longer considered to be head abstractions). Consequently, the set $\mathcal{L}_\infty$ of plain $\lambda$-terms can be specified using the auxiliary class $\mathcal{H}$ of head abstractions as depicted in Figure 6.
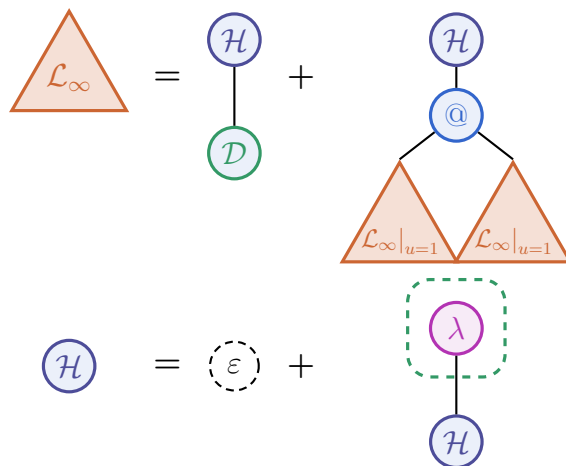


Figure 6: Marking head abstractions in plain terms.

The bivariate generating function $L_\infty(z, u)$ corresponding to plain $\lambda$-terms with marked head abstractions satisfies therefore

$$L_\infty(z, u) = \frac{1}{1 - zu} \left( \frac{z}{1 - z} + z L_\infty(z, 1)^2 \right). \tag{3.16}$$

Note that the dominant singularity $\rho(u)$ of $L_\infty(z, u)$ does not depend on $u$. In fact, $\rho(u) = \rho$ (i.e. the dominant singularity of $L_\infty(z)$, see (3.5)) as, by construction, $L_\infty(z, 1) = L_\infty(z)$.

We can therefore apply Proposition A.8 and, after routine calculations, find that the limit probability generating function $p(u)$ associated with $L_\infty(z, u)$ satisfies

$$p(u) = \frac{1 - \rho}{1 - \rho u} \tag{3.17}$$

which indeed corresponds to the asserted limit geometric distribution (3.15) of $X_n$. $\qquad\square$

*Remark* 3.7. The mean number $\mu_n$ of head abstractions in a random $\lambda$-term of size $n$ satisfies

$$\mu_n \xrightarrow[n \to \infty]{} \frac{\rho}{1 - \rho}. \tag{3.18}$$

The limit mean (3.18) is close to 0.4196. Consequently, sufficiently large plain terms have, on average, less than one head abstraction. Such a result stands in sharp contrast to the canonical representation of David et al. [DGK+13] where the number of head abstractions in a random (closed) $\lambda$-term of size $n$ is at least of order $\sqrt{n/\log n}$; in particular, it is a moderately growing function of $n$.

### 3.3 De Bruijn index values in plain lambda terms

In the current subsection we focus on the distribution of De Bruijn index values in random $\lambda$-terms.

**Proposition 3.8.** Let $X_n$ be a random variable denoting the De Bruijn index value $m$ of an index $\underline{\mathsf{m}}$ taken uniformly at random from a random plain $\lambda$-term of size $n$. Then, $X_n$ converges in law to a geometric distribution $\mathrm{GEOM}(\rho)$ with parameter $\rho$. Specifically,

$$\mathbb{P}(X_n = m) \xrightarrow[n \to \infty]{} \mathbb{P}(\mathrm{GEOM}(\rho) = m) = (1 - \rho)\rho^m. \tag{3.19}$$
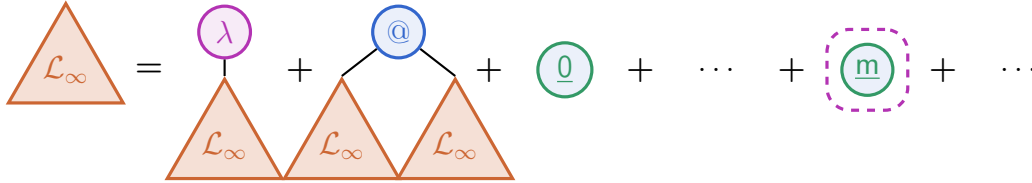


Figure 7: Marking the index $\underline{\mathsf{m}}$ in plain $\lambda$-terms.

*Proof.* Let $V_n^{(m)}$ be a random variable denoting the number of De Bruijn indices $\underline{\mathsf{m}}$ in a plain $\lambda$-term of size $n$. Marking the index $\underline{\mathsf{m}}$ in the specification for plain terms, see Figure 7, we note that the bivariate generating function $L_\infty^{(m)}(z, u)$ associated with $V_n^{(m)}$ satisfies

$$L_\infty^{(m)}(z, u) = z L_\infty^{(m)}(z, u) + z L_\infty^{(m)}(z, u)^2 + \frac{z}{1 - z} + (u - 1)z^{m+1}. \tag{3.20}$$

Denote $\frac{\partial}{\partial u} L_\infty^{(m)}(z, u)|_{u=1}$ as $D_\infty^{(m)}(z)$. Then, taking the partial derivative $\partial u$ at $u = 1$ of both sides of (3.20) we arrive at

$$D_\infty^{(m)}(z) = z D_\infty^{(m)}(z) + 2z D_\infty^{(m)}(z) L_\infty(z) + z^{m+1} \tag{3.21}$$

as $L_\infty^{(m)}(z, 1) = L_\infty(z)$ for each $m \geqslant 0$, cf. (3.20) and (3.3). Note that $[z^n] D^{(m)}(z)$ corresponds to the weighted sum over all plain $\lambda$-terms of size $n$ where each term comes with weight equal to the total number of occurrences of index $\underline{\mathsf{m}}$ in it.

Let $S_\infty(z, w) = \sum_{m \geqslant 0} D_\infty^{(m)}(z) w^m$. Taking the weighted sum over all $m \geqslant 0$ of both sides of (3.21) such that weight corresponding to $m$ is $w^m$ we obtain

$$S_\infty(z, w) = z S_\infty(z, w) + 2z S_\infty(z, w) L_\infty(z) + \frac{z}{1 - zw}. \tag{3.22}$$

Consequently, $[z^n w^m] S_\infty(z, w)$ stands for $[z^n] D^{(m)}(z)$ whereas $[z^n] S_\infty(z, 1)$ denotes the weighted sum of all $\lambda$-terms of size $n$ where each term has weight equal to its total number of variables. In other words, $w$ in $S_\infty(z, w)$ marks the probability mass function corresponding to $X_n$. Solving (3.22) we find that

$$S_\infty(z, w) = \frac{1}{1 - zw} \left( \frac{z}{1 - z - 2z L_\infty(z)} \right) = \frac{z}{1 - zw} \left( \sqrt{(1 - z)^2 - \frac{4z^2}{1 - z}} \right)^{-1} \tag{3.23}$$

where the latter equality follows from the formula (3.3) for $L_\infty(z)$.

Consequently, $S_\infty(z, w)$ admits a Puiseux series expansion required for Proposition A.8. A final computation verifies the asserted geometric limit distribution $\mathrm{GEOM}(p)$ corresponding to the variable $X_n$. □

*Remark* 3.9. The mean value $\mu_n$ of a random index with a random plain terms satisfies

$$\mu_n \xrightarrow[n\to\infty]{} \frac{\rho}{1-\rho}. \tag{3.24}$$

The limit value (3.24), coinciding in the natural size notion with the corresponding mean for head abstractions (3.18), is close to 0.4196. This result stands, again, is sharp contrast to the canonical model of David et al. [DGK+13] in which variables (in closed terms) tend to be arbitrarily far from their binding abstractions.

Let us point out that such a disparity is a consequence of the different combinatorial models for $\lambda$-terms. In the canonical representation, the distance from a variable to its binding abstraction does not contribute to the weight of the corresponding variable (all variables have weight zero). On the other hand, in the De Bruijn representation weights of bound indices are proportional to the distances to their binding abstractions. Consequently, De Bruijn indices in large random $\lambda$-terms tend to be, on average, shallow. This central difference of both combinatorial models leads to remarkably contrasting asymptotic properties, including normalisation of large random $\lambda$-terms, cf. [DGK+13, BGLZ16, Ben17].

## 3.4 Leftmost-outermost redex search

In order to evaluate an expression represented by a $\lambda$-term $M$ we iteratively choose a $\beta$-redex (i.e. a subterm in form of $(\lambda P)Q$) in $M$ and contract it using $\beta$-reduction, see e.g. Figure 8 (cf. Section 2).
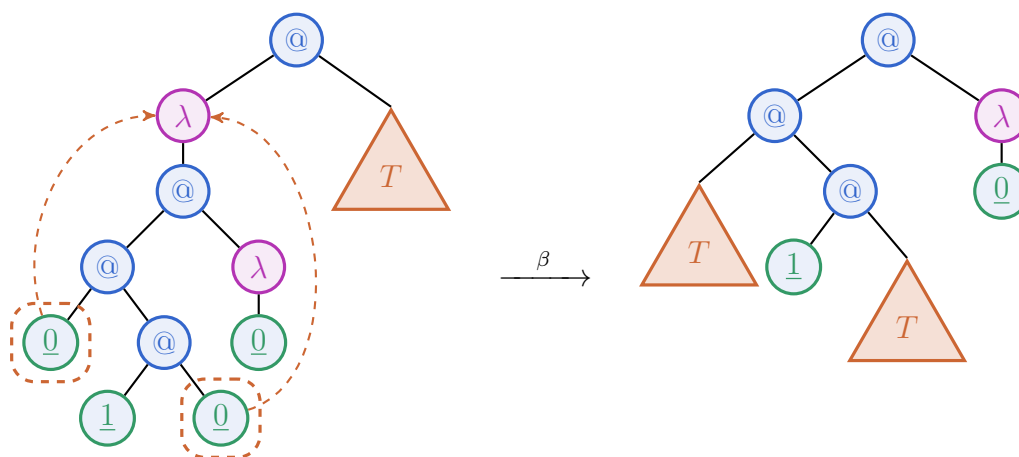


Figure 8: An example of $\beta$-reduction $(\lambda \underline{0}(\underline{1}\underline{0})(\lambda \underline{0}))T \to_\beta T(\underline{1}T)(\lambda \underline{0})$.

The order of evaluation (i.e. the order in which redexes are contracted) has a crucial impact on the computational effect of repeated $\beta$-reduction. Consider $\lambda$-terms $\Omega :=$ $(\lambda\underline{00})(\lambda\underline{00})$ and $P := (\lambda\lambda\underline{0})\Omega$. Note that $\Omega$ can be evaluated *ad infinitum* as $\Omega \to_\beta \Omega$. If we choose the $\Omega$ redex over the main $(\lambda\lambda\underline{0})\Omega$ redex in $P$, then $P \to_\beta P$. Otherwise, if we choose the *leftmost-outermost* redex $(\lambda\lambda\underline{0})\Omega$ instead of $\Omega$ we note that $P \to_\beta \lambda\underline{0}$ since the topmost abstraction in $P$ has no associated indices. We cannot continue $\beta$-reducing $\lambda\underline{0}$ as it contains no more redexes (such terms are in so-called *($\beta$-)normal form*) and so we terminate the evaluation process.

Terms from which it is possible to reach a normal form using repeated $\beta$-reduction are said to be *normalisable*. Since it is possible to emulate computations of arbitrary Turing machines in $\lambda$-calculus by means of normalisable terms, it is undecidable to determine whether a given $\lambda$-term is normalisable or not, see [Bar84]. Remarkably, the following classical result provides a normalising evaluation strategy guaranteed to find normal forms of normalisable $\lambda$-terms.

**Theorem 3.10** (Standardisation theorem, see e.g [Bar84])**.** *Let $N$ be a normalising $\lambda$-term. Then, the iterated process of applying $\beta$-reduction to the leftmost-outermost redex in $N$ leads to the (unique) normal form of $N$.*

When searching for the leftmost-outermost redex in a given $\lambda$-term, we traverse the associated $\lambda$-tree in a depth-first manner favouring left branches of application nodes. Both cases of handling abstractions and indices are trivial – when visiting an abstraction node, we immediately recurse to its subterm looking for the leftmost-outermost redex; indices cannot contain $\beta$-redexes hence after arriving at an index, we terminate the traversal.

The most interesting part of the traversal algorithm lies in visiting application nodes. Suppose that we are currently visiting an application node. If its left branch starts with an abstraction node, we terminate the traversal as we have just found the leftmost-outermost redex. Otherwise, we have two possibilities based on whether the left branch is in $\beta$-normal form or not. If it is, we move into the left branch and, as we cannot find a $\beta$-redex, return from the recursion moving to the corresponding right branch. Otherwise if the left branch is not in $\beta$-normal form, we handle it recursively, however since it contains a $\beta$-redex, we eventually terminate the search before ever returning to visit the right branch.

Let us note that such a traversal algorithm, henceforth abbreviated to LO, introduces some *a priori* non-trivial computational overhead to the execution cost of $\beta$-reduction in $\lambda$-calculus. If carried out on a $\beta$-redex, LO takes constant time to run. In contrast, when carried out on a $\lambda$-term in $\beta$-normal form, LO traverses nearly the whole $\lambda$-term. Such a varying traversal cost poses the natural question of the average-case performance of LO. In what follows, we show that the execution cost of LO, when viewed as a random variable ranging over random $\lambda$-terms, tends to a discrete limit law with constant expectation. Consequently, finding the leftmost-outermost redex introduces, on average, only a constant overhead to the cost of carrying out a single $\beta$-reduction.

**Proposition 3.11.** Let $X_n$ be a random variable denoting the number of nodes visited by the LO traversal algorithm while searching for the leftmost-outermost $\beta$-redex in a random

plain $\lambda$-term of size $n$. Then, $X_n$ converges in law to a discrete limiting distribution with computable probability generating function and constant expectation. The corresponding means $\mu_n$ satisfy (up to numerical approximation)

$$\mu_n \xrightarrow[n \to \infty]{} 6.222262521\ldots \tag{3.25}$$

*Proof.* We start with providing a combinatorial specification for plain $\lambda$-terms marking all nodes that are visited by the leftmost-outermost redex traversal algorithm LO. For that purpose we introduce the following three auxiliary classes:

- $\mathcal{N}$ for the class of $\beta$-normal forms;

- $\mathcal{M}$ for the class of so-called *neutral terms*, and

- $\mathcal{A}$ for the class of De Bruijn indices and plain $\lambda$-terms starting with an application.

In order to give their combinatorial specification we follow the presentation of [BGLZ16] and give a joint description for the class $\mathcal{N}$ of $\beta$-normal forms and the associated class $\mathcal{M}$ of neutral terms. A $\beta$-normal form is either a plain $\lambda$-term starting with an abstraction followed by another $\beta$-normal form, or a neutral term. A neutral term, in turn, is either a De Bruijn index, or an application of a neutral term to a $\beta$-normal form, see Figure 9.
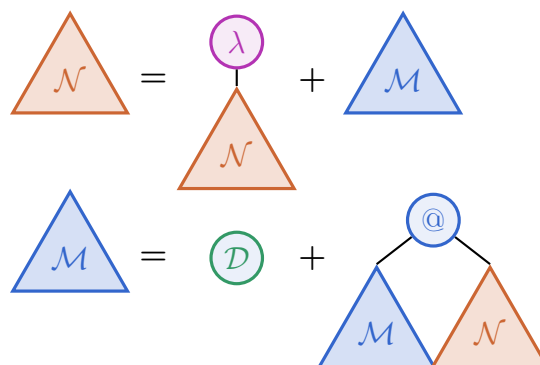


Figure 9: Joint specification for $\beta$-normal forms $\mathcal{N}$ and neutral terms $\mathcal{M}$.

Such a specification provides the following system of functional equations defining the generating functions $N(z)$ and $M(z)$ corresponding to the class of $\beta$-normal forms and neutral terms, respectively:

$$\begin{aligned}
N(z) &= zN(z) + M(z) \\
M(z) &= zM(z)N(z) + \frac{z}{1-z}.
\end{aligned} \tag{3.26}$$

Solving (3.26) we note that $M(z)$ and $N(z)$ satisfy

$$N(z) = \frac{M(z)}{1-z} \quad \text{and} \quad M(z) = \frac{1 - z - \sqrt{(1+z)(1-3z)}}{2z}. \tag{3.27}$$
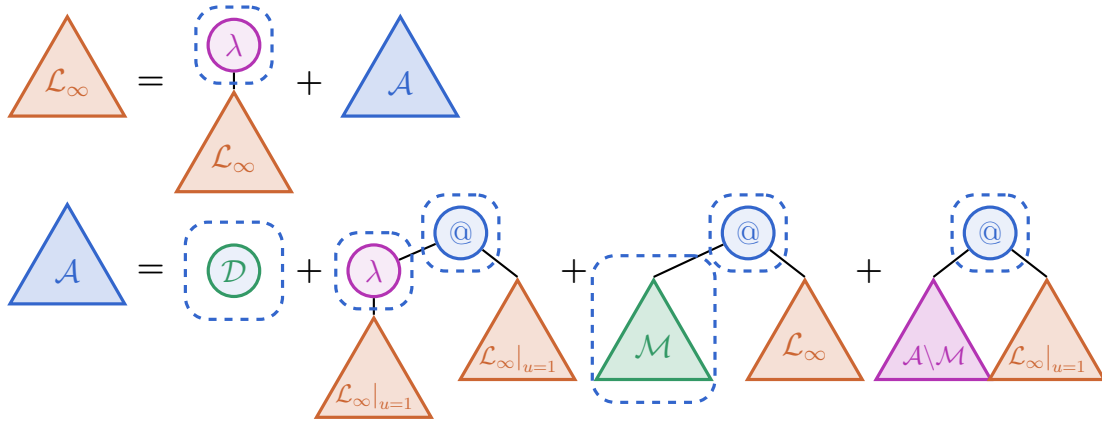
Figure 10: Specification for plain $\lambda$-terms with marked nodes following the execution of the redex finding traversal algorithm LO.

With both $N(z)$ and $M(z)$ at hand, we can now proceed and give the announced specification for plain $\lambda$-terms with marked nodes visited during the execution of LO, see Figure 10.

Let $T$ be a plain $\lambda$-term in the class $\mathcal{A}$. Certainly, if $T$ is a De Bruijn index, we mark only its topmost atom (i.e. the topmost successor or $0$ if $T$ is equal to $\underline{0}$). Otherwise, $T$ starts with an application. If $T$ is a $\beta$-redex, we mark two atoms – the topmost application and the abstraction starting the left branch of $T$. Remaining nodes are left unmarked. If $T$ is not a redex however its left branch is a neutral term, we mark the entire left branch as well as the topmost application. Finally, if the left branch of $T$ is not a neutral term, we take the difference class $\mathcal{A} \setminus \mathcal{M}$ of $\mathcal{A}$ and (marked) neutral terms $\mathcal{M}$ for the left branch. The right branch remains unmarked.

Such a specification determines the following system of functional equations defining the generating functions $L_\infty(z, u)$ and $A(z, u)$ corresponding to classes $\mathcal{L}_\infty$ and $\mathcal{A}$, respectively:

$$
\begin{aligned}
L_\infty(z, u) &= zuL_\infty(z, u) + A(z, u) \\
A(z, u) &= \frac{zu}{1 - z} + z^2 u^2 L_\infty(z, 1)^2 + zuM(zu)L_\infty(z, u) \\
&\quad + zu\left(A(z, u) - M(zu)\right) L_\infty(z, 1).
\end{aligned}
\tag{3.28}
$$

Knowing *a priori* that $L_\infty(z, 1)$ corresponds to the generating function for plain $\lambda$-terms (3.3) we solve system (3.28) and find that $L_\infty(z, u)$ satisfies

$$
L_\infty(z, u) = \frac{zuM(zu)L_\infty(z, 1) - (zuL_\infty(z, 1))^2 - \frac{zu}{1-z}}{zuM(zu) - (1 - zuL_\infty(z, 1))(1 - zu)}.
\tag{3.29}
$$

What remains to finish the proof is to check that $L_\infty(z, u)$ meets the premises of Proposition A.8. Specifically, it admits a single, fixed dominant singularity $\rho(u) = \rho$ and a

corresponding Puiseux expansion in form of

$$L_\infty(z, u) = \alpha(u) - \beta(u)\sqrt{1 - \frac{z}{\rho}} + O\left(\left|1 - \frac{z}{\rho}\right|\right). \tag{3.30}$$

Denote the denominator expression of (3.29) as $F(z, u)$. Note that $F(\rho, 1) > 0$ and hence in a fixed neighbourhood of $u = 1$ the denominator $F(z, u)$ is non-zero. Consequently, $L_\infty(z, u)$ shares its (fixed) dominant singularity with $L_\infty(z, 1)$. The required form of the Puiseux expansion of $L_\infty(z, u)$ follows as a consequence of the Puiseux expansions of both $L_\infty(z, 1)$ and its power $L_\infty(z, 1)^2$. A direct computation gives access to the corresponding probability generating function (omitted for brevity) and also the specific limit mean (3.25) of $X_n$. $\qquad\square$

*Remark* 3.12. The generating function $M(z)$ associated with neutral terms, see (3.26), also corresponds to the well-known class of Motzkin numbers enumerating, *inter alia*, plane unary-binary trees, see e.g. [FS09, Note I.39, p.68]. We refer the curious reader to [BGLZ16] for a size-preserving correspondence between neutral terms of size $n$ and Motzkin trees with $n$ nodes.

## 3.5 Height profile in plain lambda terms

The goal of this section is to obtain some insight into the mean height profile of plain $\lambda$-terms. We distinguish essentially two different notions of height. The first notion which we call *unary height*, takes into account only the number of abstractions above the considered node. The second notion concerns the *natural height* of a tree, i.e. the number of predecessors of a node which can be either abstractions or applications. In both situations, the semi-large powers theorem (see Proposition A.2) can be applied. Consequently, the mean profile is always related to the Rayleigh distribution.

We are interested in the mean profile of different types of nodes. In what follows we consider three types of mean profiles:

- the mean (unary or natural) profile of leaves;

- the mean (unary or natural) profile of abstractions, and

- the mean (unary or natural) profile of applications.

**Proposition 3.13.** Let $H_n$ be a random variable denoting the unary (respectively natural) height of a randomly chosen variable (application or abstraction) in a random plain $\lambda$-term. Then, with $x$ in any compact subinterval of $(0, +\infty)$, $H_n$ admits a limiting Rayleigh distribution

$$\mathbb{P}(H_n = k) \sim \frac{C}{\sqrt{n}} \cdot \frac{x}{2} e^{-x^2/4} \quad \text{where} \quad x = \frac{k}{\sqrt{n}} \cdot C \tag{3.31}$$

with $C \approx 4.30187$ for unary height, and $C \approx 1.27162$ for the natural height. The average value of mean height is $\sqrt{\pi n}/C$ whereas the peak value is attained at $k^* = \sqrt{2n}/C$.

More specifically, the average number of

- variables at unary height $k$ is asymptotically equal to $2.839\,ke^{-C^2k^2/4n}$;

- variables at natural height $k$ is asymptotically equal to $0.248\,ke^{-C^2k^2/4n}$;

- abstractions at unary height $k$ is asymptotically equal to $2.383\,ke^{-C^2k^2/4n}$;

- abstractions at natural height $k$ is asymptotically equal to $0.208\,ke^{-C^2k^2/4n}$;

- applications at unary height $k$ is asymptotically equal to $2.839\,ke^{-C^2k^2/4n}$;

- applications at natural height $k$ is asymptotically equal to $0.248\,ke^{-C^2k^2/4n}$.

*Proof.* We start with the unary height profile of variables. Consider generating functions $C_k(z,u)$ corresponding to plain $\lambda$-terms with $u$ marking the variables at the unary height $k$. These functions satisfy the following system of equations:

$$\begin{cases} C_0(z,u) = \dfrac{uz}{1-z} + zL_\infty(z) + zC_0(z,u)^2, & \text{if } k = 0; \\ C_k(z,u) = \dfrac{z}{1-z} + zC_{k-1}(z,u) + zC_k(z,u)^2, & \text{if } k > 0. \end{cases} \tag{3.32}$$

Taking partial derivatives of each equation in (3.32) with respect to $u$, we obtain a linear system for derivatives of generating functions. Setting $u = 1$ we can solve this linear system and obtain

$$\left.\frac{\partial}{\partial u}C_k(z,u)\right|_{u=1} = \frac{1}{1-z}\left(\frac{z}{1-2zL_\infty(z)}\right)^{k+1}. \tag{3.33}$$

Furthermore, a direct computation provides the following Puiseux series expansions as $z \to \rho$:

$$\frac{z}{1-2zL_\infty(z)} \sim 1 - \beta\sqrt{1-\frac{z}{\rho}} \quad \text{and} \quad L_\infty(z) \sim \frac{1-\rho}{2\rho} - b_\infty\sqrt{1-\frac{z}{\rho}} \tag{3.34}$$

where $\beta = 2b_\infty \approx 4.301868701457$. Consequently, the numbers $M_{n,k}$ of variables at unary level $k$ in a random plain lambda term of size $n$ satisfy

$$M_{n,k} = \frac{[z^n]\dfrac{1}{1-z}\left(\dfrac{z}{1-2zL_\infty(z)}\right)^{k+1}}{[z^n]L_\infty(z)}. \tag{3.35}$$

An application of the semi-large powers theorem (see Proposition A.2) and the transfer theorem (see Proposition A.1) to the numerator and denominator, respectively, result in the following asymptotic estimate:

$$M_{n,k} \sim \frac{2k}{1-\rho}e^{-\beta^2k^2/4n}. \tag{3.36}$$

After normalising by the total sum $\sum_{k=0}^{n} M_{n,k}$ we obtain the declared limiting distribution.

Next, we turn to the case of natural height profile of variables. Consider generating functions $C_k(z, u)$ where now $u$ marks the variables at the natural height $k$, instead of the unary height. As in the previous case, we obtain a system of equations

$$\begin{cases} C_0(z, u) = \dfrac{uz}{1 - z} + zL_\infty(z) + zL_\infty(z)^2, & \text{if } k = 0; \\ C_k(z, u) = \dfrac{z}{1 - z} + zC_{k-1}(z, u) + zC_{k-1}(z, u)^2, & \text{if } k > 0. \end{cases} \tag{3.37}$$

Again, taking partial derivatives $\partial u$ at $u = 1$ we can solve the resulting system and find that

$$\frac{\partial}{\partial u} C_k(z, u)\bigg|_{u=1} = \frac{z}{1 - z} \left( z + 2zL_\infty(z) \right)^k. \tag{3.38}$$

In this case, a direct computation verifies that the function $z + 2zL_\infty(z)$ admits a Puiseux series expansion in form of $1 - \gamma\sqrt{1 - z/\rho} + O\left(|1 - z/\rho|\right)$ where $\gamma = \beta\rho \approx 1.27162265120953$. Consequently, this estimate yields a Rayleigh distribution with parameter $1.27162265120953$. In particular, the average number $M_{n,k}$ of variables at natural height $k$ in a random plain $\lambda$-term of size $n$ satisfies

$$M_{n,k} \sim \frac{2\rho^2}{1 - \rho} k e^{-\gamma^2 k^2 / 4n}. \tag{3.39}$$

In order to mark remaining nodes, i.e. abstractions and applications, it is sufficient to change the first equation of the system (3.32). Accordingly, only the constant multiple behind the mean tree width at level $k$ changes. For abstractions, we obtain, respectively,

$$C_0(z, u) = \frac{z}{1 - z} + zuL_\infty(z) + zC_0(z, u)^2 \tag{3.40}$$

for unary height whereas

$$C_0(z, u) = \frac{z}{1 - z} + zuL_\infty(z) + zL_\infty(z)^2 \tag{3.41}$$

for natural height. This change gives the constants $2L_\infty(\rho) = \frac{(1-\rho)}{\rho} \approx 2.383$ for unary height, and $2\rho^2 L_\infty(\rho) = \rho(1 - \rho) \approx 0.208$ for the natural height, respectively.

Similarly, marking applications yields a change in the first equation for the generating function

$$C_0(z, u) = \frac{z}{1 - z} + zL_\infty(z) + zuC_0(z, u)^2 \tag{3.42}$$

for unary height, and

$$C_0(z, u) = \frac{z}{1 - z} + zL_\infty(z) + zuL_\infty(z)^2 \tag{3.43}$$

for natural height. We obtain the constants $2L_\infty^2(\rho) = \frac{(1-\rho)^2}{2\rho^2} \approx 2.839$ for unary height, and $2\rho^2 L_\infty^2(\rho) = \frac{(1-\rho)^2}{2} \approx 0.248$ for natural height, respectively.

The mean value is obtained by using the integral approximation for the ratio of sums $kM_{n,k}/\sum_{k=0}^{n} M_{n,k}$ whereas the peak value is obtained by finding the maximum value of $M_{n,k}$ as a function of $k$. $\qquad\square$

# 4 Advanced marking

In the following section we investigate more parameters related to plain and closed $\lambda$-terms. In particular, we consider:

- Several parameters related to closed $\lambda$-terms, resulting in Gaussian limit laws;

- Further parameters whose limiting distributions are discrete, including the leftmost-outermost redex search time in closed terms, the number of free variables in plain terms, the number of head abstractions in closed terms, and mean degree profile in closed terms;

- Finally, the mean height profile of closed terms for several different notions of height.

## 4.1 $m$-openness and the enumeration of closed terms

Recall that a term is said to be $m$-open (see Section 2) if by prepending it with $m$ head abstractions we obtain a closed $\lambda$-term as a result. Following this natural, hierarchical notion, the set $\mathcal{L}_m$ of $m$-open $\lambda$-terms can be specified as

$$
\begin{aligned}
\mathcal{L}_m &::= \lambda \mathcal{L}_{m+1} \mid (\mathcal{L}_m \mathcal{L}_m) \mid \underline{0}, \underline{1}, \ldots, \underline{\mathsf{m - 1}} \\
\mathcal{L}_{m+1} &::= \lambda \mathcal{L}_{m+2} \mid (\mathcal{L}_{m+1} \mathcal{L}_{m+1}) \mid \underline{0}, \underline{1}, \ldots, \underline{\mathsf{m}} \\
&\cdots \qquad \cdots
\end{aligned}
\tag{4.1}
$$

Let $L_m(z)$ denote the generating function associated with the set of $m$-open $\lambda$-terms, i.e. $L_m(z) = \sum_{n \geqslant 0} a_{n,m} z^n$ where $a_{n,m}$ stands for the number of $m$-open lambda terms of size $n$. Using (4.1) we obtain a corresponding infinite system for the functions $L_m(z)$:

$$
\begin{aligned}
L_0(z) &= z L_1(z) + z L_0(z)^2, \\
L_1(z) &= z L_2(z) + z L_1(z)^2 + z, \\
&\cdots, \\
L_m(z) &= z L_{m+1}(z) + z L_m(z)^2 + z \frac{1 - z^m}{1 - z}, \\
&\cdots
\end{aligned}
\tag{4.2}
$$

In [BGG18, Lemma 8] the authors prove that for each $m \geqslant 0$ the generating functions for $m$-open $\lambda$-terms $L_m(z)$ admit Puiseux expansions in form of

$$
L_m(z) \sim a_m - b_m \sqrt{1 - \frac{z}{\rho}}.
\tag{4.3}
$$

Moreover, by the virtue of their proof, we obtain an suitable approximation procedure for the coefficients $a_m$ and $b_m$ by truncating the system (4.2) and replacing the function $L_m(z)$ with $L_\infty(z)$. Furthermore, the estimated coefficients $\widetilde{a}_m$ and $\widetilde{b}_m$ tend to their respective limits with an error of order $O(\frac{1}{\sqrt{m}})$. Using Theorem 5.9 it is possible to prove that $\widetilde{a}_m$ and

$\widetilde{b}_m$ converge to their respective limits exponentially fast. Consequently, the approximation procedure proposed in [BGG18] convergences exponentially fast, as well.

Immediately, this implies that the probability that a random plain $\lambda$-terms is $m$-open, but not $(m-1)$-open is $\dfrac{b_m - b_{m-1}}{b_\infty}$. Certainly, the limiting distribution associated with $m$-openness is discrete.

Note that the probability distribution function of $m$-openness is proportional to the coefficient at $z^n$ in the bivariate generating function

$$L(z, u) = \sum_{k \geqslant 1} u^k \big(L_k(z) - L_{k-1}(z)\big) \sim \sum_{k \geqslant 1} u^k(a_k - a_{k-1}) - \sum_{k \geqslant 1} u^k(b_k - b_{k-1})\sqrt{1 - \frac{z}{\rho}}. \quad (4.4)$$

The mean value corresponding to $m$-openness of plain terms can be calculated as

$$\frac{[z^n]\dfrac{\partial}{\partial u}L(z, u)\Big|_{u=1}}{[z^n]L_\infty(z)} \sim \frac{\sum_{k \geqslant 1} k(b_k - b_{k-1})}{\sum_{k \geqslant 1}(b_k - b_{k-1})} = \frac{\sum_{k \geqslant 0}(b_\infty - b_k)}{b_\infty} = \sum_{k \geqslant 0}\left(1 - \frac{b_k}{b_\infty}\right). \quad (4.5)$$

In order to compute this expectation we use the approximation procedure discussed above. Using the (aptly truncated) recurrence for the coefficients $a_m$ and $b_m$

$$a_m = \frac{1}{2\rho}\left(1 - \sqrt{1 - 4\rho^2\frac{1 - \rho^m}{1 - \rho} - 4\rho^2 a_{m+1}}\right), \quad b_m = \frac{\rho b_{m+1}}{\sqrt{1 - 4\rho^2\frac{1 - \rho^m}{1 - \rho} - 4\rho^2 a_{m+1}}} \quad (4.6)$$

we obtain the numerical approximation for the mean value corresponding to $m$-openness. Numerical approximation yields an estimate 2.01922912627.

## 4.2 Variables, abstractions, successors and redexes in closed terms

In the following section we investigate the joint distribution of several parameters in closed $\lambda$-terms, utilising the novel Theorem 5.9.

**Proposition 4.1.** Let $\boldsymbol{X}_n = (X_{n(\mathsf{var})}, X_{n(\mathsf{red})}, X_{n(\mathsf{suc})}, X_{n(\mathsf{abs})})$ denote a vector of random variables denoting the number of variables, redexes, successors and abstractions in a random closed $\lambda$-term of size $n$, respectively. Then, after standardisation, the random vector $\boldsymbol{X}_n$ converges in law to a multivariate Gaussian distribution with identical parameters as plain terms.

*Proof.* Let us recall that the system of equations from Proposition 3.3 associated with the four parameters that we consider is, in the general class of plain terms, of the form

$$\begin{aligned} L(z, \boldsymbol{u}) &= u_{(\mathsf{abs})}zL(z, \boldsymbol{u}) + A(z, \boldsymbol{u}), \\ A(z, \boldsymbol{u}) &= \frac{u_{(\mathsf{var})}z}{1 - u_{(\mathsf{suc})}z} + u_{(\mathsf{red})}u_{(\mathsf{abs})}z^2 L(z, \boldsymbol{u})^2 + zA(z, \boldsymbol{u})L(z, \boldsymbol{u}) \end{aligned} \quad (4.7)$$

with $\boldsymbol{u} = (u_{(\mathsf{var})}, u_{(\mathsf{red})}, u_{(\mathsf{suc})}, u_{(\mathsf{abs})})$ corresponding to respective components of $\boldsymbol{X}_n$.

In order to compose a similar, infinite system for $m$-open terms, we index respective generating functions in accordance with the natural combinatorial interpretation of $m$-openness; if an abstraction stands before an occurrence of $L(z, \boldsymbol{u})$, its respective index should be increased by one. This leads us to the following system:

$$L_m(z, \boldsymbol{u}) = u_{(\mathsf{abs})} z L_{m+1}(z, \boldsymbol{u}) + A_m(z, \boldsymbol{u}),$$

$$A_m(z, \boldsymbol{u}) = u_{(\mathsf{var})} \frac{z(1 - (u_{(\mathsf{suc})} z)^m)}{1 - u_{(\mathsf{suc})} z} + u_{(\mathsf{red})} u_{(\mathsf{abs})} z^2 L_m(z, \boldsymbol{u}) L_{m+1}(z, \boldsymbol{u}) \qquad (4.8)$$

$$+ z A_m(z, \boldsymbol{u}) L_m(z, \boldsymbol{u}).$$

Equivalently, we can represent (4.8) as

$$\begin{pmatrix} L_m(z, \boldsymbol{u}) \\ A_m(z, \boldsymbol{u}) \end{pmatrix} = \boldsymbol{\mathcal{K}}_m(L_m(z, \boldsymbol{u}), L_{m+1}(z, \boldsymbol{u}), A_m(z, \boldsymbol{u}), z, \boldsymbol{u}), \quad m = 0, 1, 2, \ldots \qquad (4.9)$$

It is straightforward to check that all the conditions of Theorem 5.9 are satisfied. Consequently, the function $L_0(z, \boldsymbol{u})$ admits a Puiseux expansion in form of

$$L_0(z, \boldsymbol{u}) \sim a_0(\boldsymbol{u}) - b_0(\boldsymbol{u}) \sqrt{1 - \frac{z}{\rho(\boldsymbol{u})}} \qquad (4.10)$$

with the same $\rho(\boldsymbol{u})$ as in Proposition 3.3. Therefore, the limiting distribution, after standardisation, is Gaussian with the mean vector and the covariance matrix completely determined by the behaviour of the singularity $\rho(\boldsymbol{u})$ near the point $\boldsymbol{u} = \boldsymbol{1}$. $\qquad \square$

## 4.3 Free variables in plain terms

**Proposition 4.2.** Let $X_n$ be a random variable denoting the number of free variables in a random plain lambda term of size $n$. Then, $X_n$ converges in law to a computable, discrete limiting distribution.

*Proof.* Consider the infinite system of functional equations $(L_m(z, u))_{m=0}^{\infty}$ where $L_m(z, u)$ corresponds to the generating function for plain $\lambda$-terms in which each De Bruijn index whose value $k$ exceeds its unary height at least by $m$, is marked. For example, $L_0(z, u)$ corresponds to plain $\lambda$-terms with marked free variables. Note that

$$\begin{aligned} L_0(z) &= z L_1(z) + z L_0(z)^2 + uz + uz^2 + \cdots, \\ L_1(z) &= z L_2(z) + z L_1(z)^2 + z + uz^2 + \cdots, \\ &\cdots, \\ L_m(z) &= z L_{m+1}(z) + z L_m(z)^2 + z \frac{1 - z^m}{1 - z} + uz \frac{z^m}{1 - z}, \\ &\cdots \end{aligned} \qquad (4.11)$$

Let us apply Theorem 5.9 to this system taking $L_\infty(z)$ as the limiting equation. Certainly, the limiting equation does not depend on the marking variable $u$. Therefore, the singular point $z^*$ also does not depend on $u$. An application of Proposition A.8 finishes the proof. $\qquad \square$

In order to compute the mean value we set $L_m^\square(z) := \dfrac{\partial}{\partial u} L_m(z, u)\Big|_{u=1}$ and represent the respective derivative as

$$L_m^\square(z) \sim c_m - d_m \sqrt{1 - \frac{z}{\rho}}. \tag{4.12}$$

Based on (4.11) we note that

$$L_m^\square(z) = \frac{z}{1 - 2zL_\infty(z)} \left( L_{m+1}^\square(z) + \frac{z^m}{1 - z} \right). \tag{4.13}$$

Since $\frac{z}{1-2zL_\infty(z)} \sim 1 - 2b\sqrt{1 - z/\rho}$ as $z \to \rho$ we establish the following recurrence relation for the coefficients $c_m$ and $d_m$:

$$c_m = c_{m+1} + \frac{\rho^m}{1 - \rho} \quad \text{and} \quad d_m = d_{m+1} + 2b_\infty c_{m+1} + 2b_\infty \frac{\rho^m}{1 - \rho}. \tag{4.14}$$

Once solved, this implies

$$c_m = \frac{\rho^m}{(1 - \rho)^2} \quad \text{and} \quad d_m = \frac{2b_\infty \rho^m}{(1 - \rho)^3}. \tag{4.15}$$

Consequently, the mean value corresponding the number of free variables in a random plain lambda term is equal to $\frac{d_0}{b_\infty} = \frac{2}{(1-\rho)^3} \approx 5.7222625231204$.

## 4.4 Head abstractions in closed terms

**Proposition 4.3.** Let $X_n$ be a random variable denoting the number of head abstractions in a closed $\lambda$-term of size $n$, chosen uniformly at random. Then, $X_n$ converges in law to a computable, discrete limiting distribution. The corresponding expectation is close to 1.447.

*Proof.* Let $L_m(z, u)$ be the bivariate generating function associated with $m$-open lambda where $u$ marks head abstractions. Then, the system $(L_m(z, u))_{m=0}^\infty$ satisfies

$$
\begin{aligned}
L_0(z, u) &= zuL_1(z, u) + zL_0(z, 1)^2, \\
L_1(z, u) &= zuL_2(z, u) + zL_1(z, 1)^2 + z, \\
&\quad \ldots \\
L_m(z, u) &= zuL_{m+1}(z, u) + zL_m(z, 1)^2 + z\frac{1 - z^m}{1 - z}, \\
&\quad \ldots
\end{aligned}
\tag{4.16}
$$

If a $\lambda$-term starts with a head abstraction, then after its removal, the openness of the respective subterm increases by one. Consequently, we include the expression $uzL_{m+1}(z, u)$ in the equation for $L_m(z, u)$. On the other hand, if the $\lambda$-term does not start with a head abstraction, i.e. starts with an application or is itself a De Bruijn index, we do not mark

remaining abstractions as they are no longer head abstractions. Hence, we also include expressions $zL_m^2(z,1)$ and $z\frac{1-z^m}{1-z}$ in the equation corresponding to $L_m(z,u)$.

Having established the system (4.16) we note that $L_0(z,u)$ can be obtained as a limit of the solutions of truncated systems (see the description of the approximation procedure in Theorem 5.9) and this limit is equal to the sum

$$
\begin{aligned}
L_0(z,u) &= zL_0(z,1)^2 + zu\left(zuL_2(z,u) + zL_1(z,1)^2 + z\right) \\
&= zL_0(z,1)^2 + z^2uL_1(z,1)^2 + z^2u + (zu)^2\left(zuL_3(z,u) + zL_2(z,1)^2 + z + z^2\right) \\
&= \dots
\end{aligned}
$$
(4.17)

and so

$$
\begin{aligned}
L_0(z,u) &= z\sum_{m\geqslant 0}(uz)^m L_m(z,1)^2 + \sum_{m\geqslant 1}u^m(z^{m+1} + \dots + z^{2m}) \\
&= z\sum_{m\geqslant 0}(uz)^m L_m(z,1)^2 + \sum_{m\geqslant 1}(uz)^m z\frac{1-z^m}{1-z} \\
&= z\sum_{m\geqslant 0}(uz)^m L_m(z,1)^2 + \frac{uz^2}{(1-z)(1-uz)} - \frac{uz^3}{(1-z)(1-uz^2)}.
\end{aligned}
$$
(4.18)

Denote the final sum in (4.18) as $S(z,u)$. Since for each $m$ the function $L_m(z,1)$ admits a Puiseux series expansion $L_m(z,1) \sim a_m - b_m\sqrt{1 - z/\rho}$, near $z = \rho$ it holds

$$
S(z,u) \sim c(u) + z\sum_{m\geqslant 0}(uz)^m\left(a_m^2 - 2a_m b_m\sqrt{1 - \frac{z}{\rho}}\right)
$$
(4.19)

where $c(u)$ comes from the last two summands of the previous expression. Since $a_m^2 \leqslant a_\infty^2$ and $a_m b_m \leqslant a_\infty b_\infty$, $S(z,u)$ is convergent near $(z,u) = (\rho, 1)$ and the function $L_0(z,u)$ admits a Puiseux series expansion in form of

$$
L_0(z,u) \sim a_0(u) - b_0(u)\sqrt{1 - \frac{z}{\rho}}.
$$
(4.20)

Consequently, $p(u) = b_0(u)/b_0(1)$ is the limiting probability generating function corresponding to the number of head abstractions in closed $\lambda$-terms. The function $b_0(u)$ satisfies

$$
b_0(u) = 2\rho\sum_{m\geqslant 0}(u\rho)^m a_m b_m.
$$
(4.21)

$\square$

## 4.5 De Bruijn index values in closed lambda terms

**Proposition 4.4.** Let $X_n$ be a random variable denoting the De Bruijn index value $m$ of a random index $\underline{m}$ in a random closed $\lambda$-term of size $n$. Then, $X_n$ converges in law to a geometric distribution $\textsc{Geom}(\rho)$ with parameter $\rho$. Specifically,

$$
\mathbb{P}(X_n = h) \xrightarrow[n\to\infty]{} \mathbb{P}(\textsc{Geom}(\rho) = h) = (1-\rho)\rho^h.
$$
(4.22)

*Proof.* Let $L_{m,k}(z, u)$ denote the generating function for $m$-open $\lambda$-terms with $u$ marking the number of occurrences of De Bruijn index $\underline{k}$. Note that $L_{m,k}(z)$ satisfies a functional equation

$$L_{m,k}(z, u) = zL_{m+1,k}(z, u) + zL_{m,k}(z, u)^2 + z\frac{1 - z^m}{1 - z} + (u - 1)z^{k+1}\mathbf{1}_{[k<m]} \qquad (4.23)$$

where $\mathbf{1}_{[\cdot]}$ stands for the Iverson bracket notation.

Taking the partial derivative of (4.23) with respect to $u$ and assigning $u = 1$, we obtain the generating function corresponding to $\lambda$-terms weighted by the number of occurrences of De Bruijn index $\underline{k}$. Denote $\frac{\partial}{\partial u}L_{m,k}(z, u)|_{u=1}$ as $L_{m,k}^{\square}(z)$. Then, taking into account that $L_{m,k}(z, 1) = L_m(z)$ we arrive at

$$L_{m,k}^{\square}(z) = zL_{m+1,k}^{\square}(z) + 2zL_m(z)L_{m,k}^{\square}(z) + z^{k+1}\mathbf{1}_{[k<m]}. \qquad (4.24)$$

Consider the generating function

$$E_m(z, w) = \sum_{k \geqslant 0} L_{m,k}^{\square}(z)w^k. \qquad (4.25)$$

Note that $[z^n]E_m(z, w)$ denotes the probability generating function associated with the distribution of variables in $m$-open $\lambda$-terms (cf. Proposition 3.8). Consequently, summing (4.24) over $k$ we obtain

$$E_m(z, w) = zE_{m+1}(z, w) + 2zL_m(z)E_m(z, w) + z\frac{1 - (wz)^m}{1 - wz}. \qquad (4.26)$$

These equations generate an infinite system for which Theorem 5.9 with a small modification is applicable. Each of the equations of the infinite system is linear, and the generating functions $L_m(z)$ enter the equations as coefficients. This yields the desired behaviour of the Puiseux expansions, because the non-linearity of the components is used only to provide the Puiseux expansion of the limiting equation, which is given in our case by construction.

The condition of exponential convergence holds because the difference between the limiting system and the $m$th equation of the system is equal to

$$2zE_\infty(z, w)\left(L_\infty(z) - L_m(z)\right) + \frac{z}{1 - wz}(wz)^m \qquad (4.27)$$

and decreases at exponential speed. Hence, the limiting distribution of De Bruijn index value is identical to the respective parameter in plain $\lambda$-terms. $\qquad\square$

## 4.6 Leftmost-outermost redex search time in closed terms

**Proposition 4.5.** Let $X_n$ denote the number of vertices visited by depth-first traversal algorithm searching for the leftmost-outermost $\beta$-redex in a random closed $\lambda$-term of size $n$ (see Section 3.4). Then, the random variable $X_n$ converges in law to a computable, discrete limiting distribution.

*Proof.* Recall that the system (3.28) defining the generating function $L_\infty(z, u)$ corresponding to plain terms with $u$ marking visited nodes is written as

$$L_\infty(z, u) = uzL_\infty(z, u) + A(z, u),$$
$$A(z, u) = u\frac{z}{1 - z} + z^2u^2L_\infty(z, 1)^2 + zuM(zu)L_\infty(z, u) + uz(A(z, u) - M(zu))L_\infty(z, 1)$$
$$(4.28)$$

with $M(z)$ being the generating function associated with so-called neutral terms and also Motzkin numbers, see Theorem 3.12:

$$M(z) = \frac{1 - z - \sqrt{(1 + z)(1 - 3z)}}{2z}.$$
$$(4.29)$$
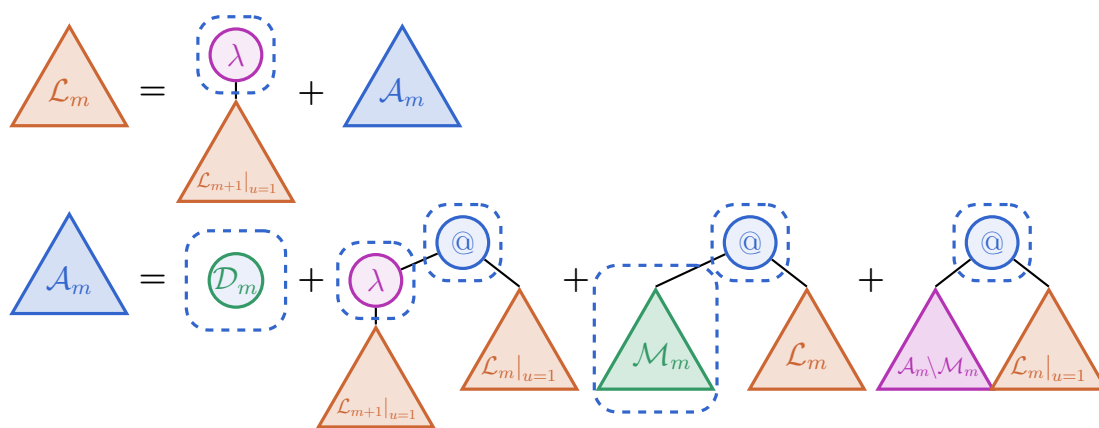


Figure 11: Specification corresponding to redex search time in closed lambda terms.

Note that including indices in (4.28) according to $m$-openness we obtain

$$L_m(z, u) = uzL_{m+1}(z, u) + A_m(z, u),$$
$$A_m(z, u) = uz\frac{1 - z^m}{1 - z} + z^2u^2L_m(z, 1)L_{m+1}(z, 1)$$
$$+ zuM_m(zu)L_m(z, u) + uz(A_m(z, u) - M_m(zu))L_m(z, 1)$$
$$(4.30)$$

where $M_m(z)$ is the generating function for $m$-open neutral lambda terms. The sequence of functions $(M_m(z))_{m=0}^\infty$ can be obtained from the system of equations

$$N_m(z) = zN_{m+1}(z) + M_m(z),$$
$$M_m(z) = zM_m(z)N_m(z) + z\frac{1 - z^m}{1 - z}.$$
$$(4.31)$$

Comparing (4.31) with its limiting counterpart (3.26) we note that the $M_\infty(z) - M_m(z)$ decays at exponential speed as $m \to \infty$ by virtue of Theorem 5.9. As additionally follows from the theorem, the functions $(M_m(z))_{m=0}^\infty$ share the same singularity $\rho = 1/3$.

Next, the system of equations (4.31) can be represented in the form

$$\begin{pmatrix} L_m(z,u) \\ A_m(z,u) \end{pmatrix} = \mathcal{K}_m(L_m(z,u), L_{m+1}(z,u), A_m(z,u), z, u). \tag{4.32}$$

In order to apply Theorem 5.9 to (4.32) we need to replace the condition that the limiting system satisfies the premises of Drmota–Lalley–Woods theorem by an assumption that the limiting system admits Puiseux expansion. It was proven in Proposition 3.11 that $L_\infty(z,u)$ has a fixed singularity $z^*$ which is independent of $u$. Therefore, all the functions $L_m(z,u)$ have a fixed singularity $z^*$ and by applying Proposition A.8, we obtain that the limiting distribution of the redex search time is discrete. $\qquad\square$

## 4.7 Node height profile in closed terms

Like in Section 3.5, in the current section we consider unary and natural height profile of variables, abstractions and applications in closed lambda terms. For this purpose, we provide a variation of the semi-large powers theorem (see Proposition A.2).

**Theorem 4.6.** *Let $(f_k(\rho z))_{k \geqslant 0}$ be a sequence of functions analytic in delta-domain $\Delta(R)$ (see Proposition A.1) for some $R > \rho$ admitting Puiseux series expansions in form of*

$$f_k(z) \sim \sigma_k - a_k\sqrt{1 - \frac{z}{\rho}} \tag{4.33}$$

*as $z \to \rho$. Assume there exist $\beta$ and $\widehat{\sigma}$ such that the sequences $(\sigma_k)_{k \geqslant 0}$ and $(a_k)_{k \geqslant 0}$ satisfy*

$$\sum_{j=0}^{k} \frac{a_j}{\sigma_j} \sim \beta k \quad and \quad \lim_{k \to \infty} \prod_{j=0}^{k} \sigma_k \to \widehat{\sigma}. \tag{4.34}$$

*Then, for $x$ in any compact subinterval of $(0, +\infty)$, as $n \to \infty$, it holds*

$$[z^n]\prod_{j=0}^{k} f_j(z) \sim \widehat{\sigma}\frac{\rho^{-n}}{n}S(\beta x) \quad and \quad x = \frac{k}{\sqrt{n}} \tag{4.35}$$

*where $S(x)$ is the Rayleigh function defined in Proposition A.2.*

*Proof.* We recall that in the course of the proof of the semi-large power theorem, see [FS09, Theorem IX.16], the coefficient $[z^n]f(z)^k$ is expressed as the following complex contour integral with the help of Cauchy's integral theorem:

$$[z^n]f(z)^k = \frac{1}{2\pi i}\oint f(z)^k\frac{dz}{z^{n+1}} = \frac{1}{2\pi i}\oint e^{h_{n,k}(z)}\frac{dz}{z}, \quad h_{n,k}(z) = k\log f(z) - n\log z. \tag{4.36}$$

With the change of variables $z = \rho(1-t/n)$ the coefficient $[z^n]f(z)^k$ can be accordingly approximated by the following real integral:

$$[z^n]f(z)^k \sim -\frac{\rho^n}{n}\frac{1}{2\pi i}\int_0^\infty e^{t-ax\sqrt{t}}dt. \tag{4.37}$$

As proven in the referenced literature, this yields the Rayleigh approximation. In the statement of the current theorem, the function $h_{n,k}(z)$, i.e. the logarithm of the sub-integral expression, is replaced by

$$\widetilde{h}_{n,k} = \sum_{j=0}^{k} \log f_j(z) - n \log z. \tag{4.38}$$

Accordingly, with the variable change $z = \rho(1-t/n)$ the coefficient $[z^n] \prod_{j=0}^{k} f_j(z)$ becomes

$$[z^n] \prod_{j=0}^{k} f_j(z) = \prod_{j=0}^{k} \sigma_k \cdot [z^n] \prod_{j=0}^{k} \left(1 - \frac{a_k}{\sigma_k}\sqrt{1 - \frac{z}{\rho}}\right) \sim -\frac{\rho^n \prod_{j=0}^{k} \sigma_k}{n} \frac{1}{2\pi i} \int_0^{\infty} e^{t-\beta x\sqrt{t}} dt \tag{4.39}$$

which has the same form as (4.37), finishing the proof. $\square$

**Proposition 4.7.** Let $H_n$ be a random variable denoting the unary (respectively natural) height of a uniformly random variable in a random closed lambda term. Then, with $x$ in any compact subinterval of $(0, +\infty)$, $H_n$ follows the Rayleigh limiting distribution

$$\mathbb{P}(H_n = k) \sim \frac{C}{\sqrt{n}} \cdot \frac{x}{2} e^{-x^2/4}, \quad \text{where} \quad x = \frac{k}{\sqrt{n}} \cdot C \tag{4.40}$$

with $C \approx 4.30187$ for unary height and $C \approx 1.27162$ for the natural height.

*Proof.* Let $C_{m,k}(z,u)$ denote the bivariate generating function corresponding to $m$-open $\lambda$-terms where variable $u$ marks De Bruijn indices at unary height $k - m$. Certainly, $C_{m,k}(z,1) = L_m(z)$ for each $m$ and $k$. Note that, the functions $(C_{m,k}(z,u))_{m=0}^{\infty}$ satisfy jointly

$$\begin{cases} C_{m,k}(z,u) = z\dfrac{1 - z^m}{1 - z} + zC_{m+1,k}(z,u) + zC_{m,k}(z,u)^2 & \text{if } m < k, \\[2mm] C_{m,k}(z,u) = zu\dfrac{1 - z^m}{1 - z} + zL_{m+1}(z) + zC_{m,k}(z,u)^2 & \text{if } m = k, \\[2mm] C_{m,k}(z,u) = L_m(z) & \text{if } m > k. \end{cases} \tag{4.41}$$

A straightforward induction yields

$$\left.\frac{\partial}{\partial u}C_{0,k}(z,u)\right|_{u=1} = \prod_{j=0}^{k} \frac{z}{1 - 2zL_j(z)} \cdot \frac{1 - z^k}{1 - z}. \tag{4.42}$$

This function is amenable to asymptotic analysis of their coefficients by Theorem 4.6. First show that in the respective Puiseux expansions of the functions

$$\frac{z}{1 - 2zL_j(z)} \sim \sigma_j - c_j\sqrt{1 - z/\rho}$$

the sequence $\sigma_j$ tends to 1 at exponential speed, and the sequence $c_j$ tends to a limit $2b_\infty \approx 4.30187$ again at exponential speed. This holds because in the course of the proof of Theorem 5.9 we have shown that the sequences of coefficients of the Puiseux expansion of $(L_j(z))_{j=0}^\infty$ (respectively, the the sequence of first coefficients $L_j(\rho)$, and the sequence of the second coefficients) tend to their respective limits, i.e. to the coefficients of the Puiseux expansion of $L_\infty(z)$ exponentially fast. Comparing with the Puiseux expansion of $\frac{z}{1-2zL_\infty(z)}$ given in the proof of Proposition 3.13 we obtain the limiting values of the sequences $(\sigma_j)_{j=0}^\infty$ and $(c_j)_{j=0}^\infty$. Since the speed of convergence is exponential, the product $\prod_{j=0}^k \sigma_j$ converges to some $\hat{\sigma}$, and the sum of the ratios $c_j/\sigma_j$ tends to a linear function $\beta k = 2b_\infty k$.

Note that up to a normalising constant, the height profile of other parameters, namely the height profile distribution of abstractions and applications, remains asymptotically the same because from the generating function viewpoint only the multiple in front of the product $\prod_{j=0}^k f_j(z)$ changes (see Proposition 3.13).

In the same manner, there can be obtained Rayleigh distribution for natural height profile of different parameters. For example, in the case of variable height profile, we obtain the system of equations for the family of generating functions $C_{m,k}(z,u)$ for $m$-open lambda terms with variable $u$ marking De Bruijn indices at unary height $k-m$:

$$\begin{cases} C_{m,k}(z,u) = z\dfrac{1-z^m}{1-z} + zC_{m+1,k}(z,u) + zC_{m+1,k}(z,u)^2, & 0 \leqslant m < k; \\[2mm] C_{m,k}(z,u) = uz\dfrac{1-z^m}{1-z} + zL_{m+1}(z) + zL_m(z)^2, & m = k; \\[2mm] C_{m,k}(z,u) = L_m(z), & m > k. \end{cases}$$

This implies

$$\frac{\partial}{\partial u}C_{0,k}(z,u)\bigg|_{u=1} = \prod_{j=1}^k (z + 2zL_j(z)) \cdot z\frac{1-z^k}{1-z}.$$

Using the same argument as in the previous case, and taking into account two first terms of Puiseux expansion of $(z + 2zL_\infty(z))$ (see proof of Proposition 3.13), we obtain again Rayleigh distribution, with the same parameter as for plain lambda terms. □

## 5 Infinite systems of algebraic equations

In this section we present our main technical contribution within analytic combinatorics meant for dealing with certain recursive, infinite systems of generating functions, i.e. Theorem 5.9. Although our results admits broader applications than the one presented in the current paper, for consistency, we focus only on the enumeration and statistical analysis of various combinatorial parameters in closed $\lambda$-terms. Our proof is motivated by the papers [BGG18, GG16] where the authors consider the enumeration problem of closed $\lambda$-terms without additional marking parameters. For that purpose, they construct a series of sequences $(L_{m,N})_{m\geqslant 0}$ that approximate $m$-open $\lambda$-terms with convergence rate of order

$O\left(\frac{1}{\sqrt{N}}\right)$. In what follows we improve this rate to an exponential one. Moreover, we abstract the considered system away from $\lambda$-terms in the De Bruijn notation, allowing for an analysis of more general varieties of combinatorial systems.

Remarkably, in [GG16, Section 5] the authors obtain the asymptotic estimate $\Theta(n^{-3/2}\rho^{-n})$ for the number of closed lambda terms of size $n$ with the difference between upper and lower bounds on the constant multiple within $10^{-7}$. Let us notice that this technique is quite different from the technique used in [BGG18]. Our approach is more similar to the former method, however admits certain improvements. Specifically, we simplify the procedure of improved constant estimation, give a rigorous proof about the exponential convergence rate of the constant multiple, and provide more general and simpler tools based on the properties of the Jacobian of the limiting system (instead of exploiting the particular form of this equation in the case of $\lambda$-terms).

## 5.1 Calculus techniques for formal power series

We start with some basic notation and properties of coefficient-wise inequalities on formal power series and some geometric results on matrices of formal power series. In what follows, we denote the spectral radius of matrix $\boldsymbol{\mathcal{A}}$, i.e. the largest absolute value of its eigenvalues, by $r(\boldsymbol{\mathcal{A}})$.

**Definition 5.1** (Formal power series domination). We say that $f(z)$ *is dominated by* $g(z)$, denoted as $f(z) \preceq g(z)$, if for each $n \geqslant 0$ we have $[z^n]f(z) \leqslant [z^n]g(z)$. For multivariate formal power series

$$f(\boldsymbol{z}) = \sum_{\boldsymbol{n} \geqslant \boldsymbol{0}} a_{\boldsymbol{n}} \boldsymbol{z}^{\boldsymbol{n}} \quad \text{and} \quad g(\boldsymbol{z}) = \sum_{\boldsymbol{n} \geqslant \boldsymbol{0}} b_{\boldsymbol{n}} \boldsymbol{z}^{\boldsymbol{n}} \tag{5.1}$$

where $\boldsymbol{z} = (z_1, \ldots, z_d)$, $\boldsymbol{n} = (n_1, \ldots, n_d)$ and $\boldsymbol{z}^{\boldsymbol{n}} = z_1^{n_1} z_2^{n_2} \ldots z_d^{n_d}$ the domination $f(\boldsymbol{z}) \preceq g(\boldsymbol{z})$ means that for each vector of indices $\boldsymbol{n}$ it holds $a_{\boldsymbol{n}} \leqslant b_{\boldsymbol{n}}$. Certainly, if a combinatorial class $\mathcal{F}$ is included in a combinatorial class $\mathcal{G}$, then the generating functions $f(z)$ and $g(z)$ corresponding to respective classes satisfy $f(z) \preceq g(z)$. The same holds for marked classes and associated multivariate generating functions. Finally, for vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ of identical (however not necessarily finite) dimension, we write $\boldsymbol{A} \preceq \boldsymbol{B}$ to denote a coordinate-wise domination of respective components.

In real analysis, the *squeeze lemma* is a theorem regarding the limit of the sequence which is upper- and lower-bounded by two sequences with the same limit value. The following statement is a variant of this lemma, stated in the context of formal power series admitting coefficient asymptotics suitable for analysis of corresponding limit laws, see Appendix A.2.

**Lemma 5.2** (Squeeze lemma for formal power series). *Let $z \in \mathbb{C}$ and $\boldsymbol{u} = (u_1, \ldots, u_r) \in \mathbb{C}^r$. Assume that $f(z, \boldsymbol{u})$, $(h_m(z, \boldsymbol{u}))_{m \geqslant 0}$, and $(g_m(z, \boldsymbol{u}))_{m \geqslant 0}$ are multivariate formal power series in $(z, \boldsymbol{u})$ with non-negative coefficients, such that for every $n$ and $m$, the functions*

$$[z^n]f(z, \boldsymbol{u}), \ [z^n]h_m(z, \boldsymbol{u}) \ and \ [z^n]g_m(z, \boldsymbol{u}) \tag{5.2}$$

*are polynomials in $\boldsymbol{u}$.*

*Moreover, assume that the following conditions hold:*

- *for each $m \geqslant 0$, we have*

$$h_m(z, \boldsymbol{u}) \preceq f(z, \boldsymbol{u}) \preceq g_m(z, \boldsymbol{u}) \tag{5.3}$$

  *in the sense of multivariate formal power series domination;*

- *there exists a sequence of real positive numbers $C_n$ and functions $\left(\underline{A_m}(\boldsymbol{u})\right)_{m \geqslant 0}$, $\left(\overline{A_m}(\boldsymbol{u})\right)_{m \geqslant 0}$, $B(\boldsymbol{u})$ analytic in a common neighbourhood of $\boldsymbol{u} = \boldsymbol{1}$, such that uniformly for $m \geqslant 0$ and uniformly in a fixed complex vicinity of $\boldsymbol{u} = \boldsymbol{1}$, it holds*

$$\lim_{n \to \infty} \frac{[z^n] h_m(z, \boldsymbol{u})}{C_n \underline{A_m}(\boldsymbol{u}) B(\boldsymbol{u})^n} = 1 \quad and \quad \lim_{n \to \infty} \frac{[z^n] g_m(z, \boldsymbol{u})}{C_n \overline{A_m}(\boldsymbol{u}) B(\boldsymbol{u})^n} = 1; \tag{5.4}$$

- *there exists a function $A(\boldsymbol{u})$ analytic near $\boldsymbol{u} = \boldsymbol{1}$ such that $A(\boldsymbol{1}) \neq 0$ satisfying uniformly in a complex vicinity of $\boldsymbol{u} = \boldsymbol{1}$*

$$\lim_{m \to \infty} \underline{A_m}(\boldsymbol{u}) = \lim_{m \to \infty} \overline{A_m}(\boldsymbol{u}) = A(\boldsymbol{u}). \tag{5.5}$$

*Then, uniformly in a complex vicinity of $\boldsymbol{u} = \boldsymbol{1}$, as $n \to \infty$:*

$$[z^n] f(z, \boldsymbol{u}) \sim C_n A(\boldsymbol{u}) B(\boldsymbol{u})^n. \tag{5.6}$$

*Proof.* We divide the proof into two parts. We start with showing that the statement holds for vectors $\boldsymbol{u}$ whose components are real positive numbers. Next, we extend this property onto all complex components of $\boldsymbol{u}$.

First, take $\boldsymbol{u} \in \mathbb{R}^r$ in the vicinity where the functions $\underline{A_m}(\boldsymbol{u})$, $\overline{A_m}(\boldsymbol{u})$, $B(\boldsymbol{u})$ are analytic. Then, following (5.3) and (5.4), for every positive $\varepsilon$ there exists $N := N(\varepsilon)$, independent of $m$ and $\boldsymbol{u}$, such that

$$\forall n > N \quad \underline{A_m}(\boldsymbol{u})(1 - \varepsilon) \leqslant \frac{[z^n] f(z, \boldsymbol{u})}{C_n B(\boldsymbol{u})^n} \leqslant \overline{A_m}(\boldsymbol{u})(1 + \varepsilon). \tag{5.7}$$

Taking the limit with respect to $m$ we note that condition (5.5) guarantees that for arbitrary small $\varepsilon > 0$ and sufficiently large $N$ (again, independent of $\boldsymbol{u}$) we have

$$\forall n > N \quad A(\boldsymbol{u})(1 - \varepsilon) \leqslant \frac{[z^n] f(z, \boldsymbol{u})}{C_n B(\boldsymbol{u})^n} \leqslant A(\boldsymbol{u})(1 + \varepsilon). \tag{5.8}$$

In other words, for values of $\boldsymbol{u} \in \mathbb{R}^r$ within a fixed vicinity of $\boldsymbol{u} = \boldsymbol{1}$ it holds

$$[z^n] f(z, \boldsymbol{u}) \sim C_n A(\boldsymbol{u}) B(\boldsymbol{u})^n. \tag{5.9}$$

Now, let us consider $\boldsymbol{u} \in \mathbb{C}^r$. Note that since for each $n, m \geqslant 0$ the formal power series $[z^n] f(z, \boldsymbol{u})$, $[z^n] h_m(z, \boldsymbol{u})$ and $[z^n] g_m(z, \boldsymbol{u})$ are polynomials in $\boldsymbol{u}$, they are analytic in

$\mathbb{C}^r$. Moreover, as $\psi_{n,m}(\boldsymbol{u}) := [z^n]f(z, \boldsymbol{u}) - [z^n]h_m(z, \boldsymbol{u})$ is a polynomial with non-negative coefficients, for every $\boldsymbol{u} \in \mathbb{C}^r$ we have $|\psi_{n,m}(\boldsymbol{u})| \leqslant \psi_{n,m}(|\boldsymbol{u}|)$ and, consequently,

$$\left|[z^n]f(z, \boldsymbol{u}) - [z^n]h_m(z, \boldsymbol{u})\right| \leqslant [z^n]f(z, |\boldsymbol{u}|) - [z^n]h_m(z, |\boldsymbol{u}|). \tag{5.10}$$

After dividing both parts by $C_n|A(\boldsymbol{u})||B(\boldsymbol{u})^n|$ we obtain

$$\left|\frac{[z^n]f(z, \boldsymbol{u})}{C_n A(\boldsymbol{u})B(\boldsymbol{u})^n} - \frac{[z^n]h_m(z, \boldsymbol{u})}{C_n A(\boldsymbol{u})B(\boldsymbol{u})^n}\right| \leqslant \frac{[z^n]f(z, |\boldsymbol{u}|)}{C_n|A(\boldsymbol{u})||B(\boldsymbol{u})^n|} - \frac{[z^n]h_m(z, |\boldsymbol{u}|)}{C_n|A(\boldsymbol{u})||B(\boldsymbol{u})^n|}. \tag{5.11}$$

Following condition (5.4) and the estimate (5.9) for $\boldsymbol{u} \in \mathbb{R}^r$, we note that for every $\varepsilon > 0$ there exists $N := N(\varepsilon)$ independent of $m$ and $\boldsymbol{u}$, such that for all $n > N$ we further have

$$\left|\frac{[z^n]f(z, \boldsymbol{u})}{C_n A(\boldsymbol{u})B(\boldsymbol{u})^n} - \frac{A_m(\boldsymbol{u})}{A(\boldsymbol{u})}\right| \leqslant \frac{A(|\boldsymbol{u}|) - A_m(|\boldsymbol{u}|)}{|A(\boldsymbol{u})|} \cdot \frac{B(|\boldsymbol{u}|)^n}{|B(\boldsymbol{u})^n|} + \varepsilon. \tag{5.12}$$

Since $\varepsilon$ does not depend on $m$, we can take the limit with respect to $m$. Given condition (5.5) we note that for sufficiently large $n$ we have

$$\left|\frac{[z^n]f(z, \boldsymbol{u})}{C_n A(\boldsymbol{u})B(\boldsymbol{u})^n} - 1\right| \leqslant \varepsilon. \tag{5.13}$$

Hence, uniformly in a fixed complex vicinity of $\boldsymbol{u} = \boldsymbol{1}$

$$\lim_{n \to \infty} \frac{[z^n]f(z, \boldsymbol{u})}{C_n A(\boldsymbol{u})B(\boldsymbol{u})^n} = 1, \tag{5.14}$$

which finishes the proof. $\qquad\square$

*Remark* 5.3. Using the same technique, higher-order error terms can also be transferred, provided that the function is squeezed between two sequences of formal power series with known Puiseux expansions. In such a situation, higher-order terms correspond to coefficients obtained from the summands of Puiseux expansion

$$f(z, \boldsymbol{u}) \sim \sum_{k \geqslant 0} c_k(\boldsymbol{u}) \left(1 - \frac{z}{\rho(\boldsymbol{u})}\right)^{k/2}. \tag{5.15}$$

The next lemma is a formal power series analogue of Lagrange's mean value theorem.

**Lemma 5.4** (Mean value lemma for formal power series). *Let $f(z)$ and $g(z)$ be two formal power series such that $f(z) \preceq g(z)$. Assume that $\Psi(t) = \sum_{n \geqslant 0} \psi_n t^n$ is a formal power series with non-negative coefficients. Then,*

$$\Psi(g(z)) - \Psi(f(z)) \preceq (g(z) - f(z))\, \Psi'(g(z)). \tag{5.16}$$

*Likewise, the statement holds for vectors of formal power series in a coordinate-wise manner.*

*Proof.* Coefficient-wise subtraction of the left-hand side of (5.16) yields

$$
\begin{aligned}
\Psi(g(z)) - \Psi(f(z)) &= \sum_{n \geqslant 0} \psi_n \left( g(z)^n - f(z)^n \right) \\
&= (g(z) - f(z)) \sum_{n \geqslant 0} \psi_n \left( \sum_{i=0}^{n-1} g(z)^i f(z)^{n-i-1} \right) \\
&\preceq (g(z) - f(z)) \sum_{n \geqslant 0} \psi_n n g(z)^{n-1} \\
&= (g(z) - f(z)) \, \Psi'(g(z)).
\end{aligned}
\tag{5.17}
$$

The first two equalities hold as a consequence of formal power series composition and the identity $a^n - b^n = (a - b) \sum_{i=0}^{n-1} a^i b^{n-i-1}$. The subsequent domination follows from the assumption that $f(z) \preceq g(z)$. □

## 5.2 Forward recursive systems

The following definition of *forward recursive systems* encapsulates the general, abstract features of the infinite systems that we consider in the current paper. Core characteristics of the infinite system corresponding to closed $\lambda$-terms are abstracted and divided into three general conditions which are sufficient to access the asymptotic form of respective coefficients.

**Definition 5.5** (Forward recursive systems)**.** Let $z$ be a formal variable and $\boldsymbol{u} = (u_1, \ldots, u_r)$ be a vector of $r$ formal variables. Consider infinite sequences $\left( \boldsymbol{L}^{\langle m \rangle} \right)_{m \geqslant 0}$ and $\left( \boldsymbol{\mathcal{K}}^{\langle m \rangle} \right)_{m \geqslant 0}$ of $d$-dimensional vectors

$$
\boldsymbol{L}^{\langle m \rangle} = \left( L_1^{\langle m \rangle}, \ldots, L_d^{\langle m \rangle} \right) \quad \text{and} \quad \boldsymbol{\mathcal{K}}^{\langle m \rangle} = \left( K_1^{\langle m \rangle}, \ldots, K_d^{\langle m \rangle} \right)
\tag{5.18}
$$

consisting of formal power series $L_i^{\langle m \rangle}(z, \boldsymbol{u})$ and $K_i^{\langle m \rangle}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u})$ where $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are vectors of $d$ variables and $i = 1, \ldots, d$.

Assume that $\left( \boldsymbol{L}^{\langle m \rangle} \right)_{m \geqslant 0}$ and $\left( \boldsymbol{\mathcal{K}}^{\langle m \rangle} \right)_{m \geqslant 0}$ satisfy

$$
\boldsymbol{L}^{\langle m \rangle} = \boldsymbol{\mathcal{K}}^{\langle m \rangle} \left( \boldsymbol{L}^{\langle m \rangle}, \boldsymbol{L}^{\langle m+1 \rangle}, z, \boldsymbol{u} \right).
\tag{5.19}
$$

Then, we say that the system (5.19) is *forward recursive*.

Furthermore, consider a *limiting system* in form of

$$
\boldsymbol{L}^{\langle \infty \rangle} = \boldsymbol{\mathcal{K}}^{\langle \infty \rangle} \left( \boldsymbol{L}^{\langle \infty \rangle}, \boldsymbol{L}^{\langle \infty \rangle}, z, \boldsymbol{u} \right)
\tag{5.20}
$$

where $\boldsymbol{L}^{\langle \infty \rangle}$ and $\boldsymbol{\mathcal{K}}^{\langle \infty \rangle}$ are $d$-dimensional vectors of formal power series $L_i^{\langle \infty \rangle}(z, \boldsymbol{u})$ and $K_i^{\langle \infty \rangle}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u})$, respectively, and moreover all series $K_i^{\langle \infty \rangle}$ are analytic at $(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u}) = (\boldsymbol{0}, \boldsymbol{0}, 0, \boldsymbol{1})$. In this setting, we say that the system (5.19):

a) is *infinitely nested* if $\boldsymbol{\mathcal{K}}^{\langle m \rangle}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u}) \preceq \boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u})$ for each $m \geqslant 0$;

b) *tends to an irreducible context-free schema* if it is infinitely nested and its corresponding limiting system (5.20) satisfies the premises of the Drmota–Lalley–Woods theorem (see Proposition A.5) i.e. is a polynomial, non-linear system of functional equations which is algebraic positive, proper, irreducible and aperiodic;

c) is *exponentially converging* if there exists a vector $\boldsymbol{A}(z, \boldsymbol{u}) = (A_1(z, \boldsymbol{u}), \ldots, A_d(z, \boldsymbol{u}))$ and a function $B(z, \boldsymbol{u})$ such that:

- for each $m \geqslant 0$ we have

$$\boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{L}^{\langle \infty \rangle}, \boldsymbol{L}^{\langle \infty \rangle}, z, \boldsymbol{u}) - \boldsymbol{\mathcal{K}}^{\langle m \rangle}(\boldsymbol{L}^{\langle \infty \rangle}, \boldsymbol{L}^{\langle \infty \rangle}, z, \boldsymbol{u}) \preceq \boldsymbol{A}(z, \boldsymbol{u}) \cdot B(z, \boldsymbol{u})^m; \quad (5.21)$$

- $A_1(z, \boldsymbol{u}), \ldots, A_d(z, \boldsymbol{u})$ and $B(z, \boldsymbol{u})$ are analytic functions in the disk $|z| < \rho + \varepsilon$ for some $\varepsilon > 0$ and $\boldsymbol{u} = \boldsymbol{1}$ where $\rho$ is the dominant singularity of the limit system (5.20) at point $\boldsymbol{u} = \boldsymbol{1}$. Moreover, at $\boldsymbol{u} = \boldsymbol{1}$ we have $|B(\rho(\boldsymbol{u}), \boldsymbol{u})| < 1$ where $\rho(\boldsymbol{u})$ is the singularity of the limiting system (5.20).

**Example 5.6.** Consider the infinite system corresponding to $m$-open $\lambda$-terms, see (2.3). Recall that the sequence $(L_m(z))_{m \geqslant 0}$ of respective generating functions satisfies

$$
\begin{aligned}
L_0(z) &= zL_1(z) + zL_0(z)^2 \\
L_1(z) &= zL_2(z) + zL_1(z)^2 + z \\
&\cdots \\
L_m(z) &= zL_{m+1}(z) + zL_m(z)^2 + z\frac{1 - z^m}{1 - z} \\
&\cdots
\end{aligned}
\qquad (5.22)
$$

Let us show that (5.22) is an infinitely nested, forward recursive system which tends to an irreducible context-free schema of $L_\infty(z)$ at an exponential convergence rate. Here, each intermediate system $\boldsymbol{L}^{\langle m \rangle}$ consists of a single equation defining $L_m(z)$. Note that there are no additional marking variables $\boldsymbol{u}$. The vectors $\boldsymbol{\mathcal{K}}^{\langle m \rangle}$ are one-dimensional and the corresponding functions $K_m$ are given by

$$K_m(\ell_1, \ell_2, z) := z\ell_2 + z\ell_1^2 + z\frac{1 - z^m}{1 - z}. \qquad (5.23)$$

The limiting system $L_\infty(z)$ satisfies

$$L_\infty(z) = zL_\infty(z) + zL_\infty(z)^2 + \frac{z}{1 - z}. \qquad (5.24)$$

One can easily check that it also satisfies the premises of Proposition A.5; hence, the considered system (5.22) tends to an irreducible context-free schema. Since the trivariate formal power series $K_\infty(\ell_1, \ell_2, z) - K_m(\ell_1, \ell_2, z)$ has non-negative coefficients, the system (5.22) is also infinitely nested. Moreover, the difference between the limiting equation and the $m$th equation computed at $\ell_1 = \ell_2 = L_\infty(z)$ is equal to $\frac{z^{m+1}}{1-z}$ and corresponds to a subset of De Bruijn indices. Certainly, as $m$ tends to infinity, this difference converges to zero exponentially fast.

Given the combinatorial relation between $m$-open $\lambda$-terms and plain terms, we readily obtain the requested condition $L_m(z) \preceq L_\infty(z)$. However, for arbitrary forward recursive systems (4.2) establishing that $\boldsymbol{L}^{\langle m \rangle} \preceq \boldsymbol{L}^{\langle \infty \rangle}$ is no longer so straightforward. In what follows, we prove that for this inequality to hold it is sufficient that the limiting system is well-founded.

**Lemma 5.7.** *Let $\mathcal{S}$ be an infinitely nested, forward recursive system* (5.19)*. Assume that the coefficients of the formal power series $\mathcal{K}^{\langle \infty \rangle}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u})$ corresponding to the limiting system are non-negative and the limiting system* (5.20) *is well-founded (i.e. algebraic proper in the sense of Theorem A.3) and has a non-zero solution $\boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u})$. Finally, assume that $\mathcal{K}^{\langle \infty \rangle}(\boldsymbol{0}, \boldsymbol{0}, 0, \boldsymbol{u}) = \boldsymbol{0}$. Then,*

$$\boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u}) \preceq \boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u}). \tag{5.25}$$

*Proof.* Consider the vectors $\boldsymbol{\mathcal{L}} = (\boldsymbol{L}^{\langle 0 \rangle}, \boldsymbol{L}^{\langle 1 \rangle}, \ldots)$ and $\boldsymbol{\mathcal{L}}^+ = (\boldsymbol{L}^{\langle \infty \rangle}, \boldsymbol{L}^{\langle \infty \rangle}, \ldots)$ consisting of aptly concatenated and flattened systems $\left( \boldsymbol{L}^{\langle m \rangle} \right)_{m \geqslant 0}$ and $\boldsymbol{L}^{\langle \infty \rangle}$, respectively. Intuitively, $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{L}}^+$ are in a sense *vectors of vectors*, but for convenience we call them just *vectors*. Note that both $\boldsymbol{\mathcal{L}}(z, \boldsymbol{u})$ and $\boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u})$ satisfy

$$\begin{aligned} \boldsymbol{\mathcal{L}}(z, \boldsymbol{u}) &= \boldsymbol{\Phi}(\boldsymbol{\mathcal{L}}(z, \boldsymbol{u}), z, \boldsymbol{u}) \\ \boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u}) &= \boldsymbol{\Psi}(\boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u}), z, \boldsymbol{u}) \end{aligned} \tag{5.26}$$

where

$$\begin{aligned} \boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) &= (\mathcal{K}^{\langle 0 \rangle}(\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1, z, \boldsymbol{u}), \mathcal{K}^{\langle 1 \rangle}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, z, \boldsymbol{u}), \ldots) \\ \boldsymbol{\Psi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) &= (\mathcal{K}^{\langle \infty \rangle}(\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1, z, \boldsymbol{u}), \mathcal{K}^{\langle \infty \rangle}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, z, \boldsymbol{u}), \ldots) \end{aligned} \tag{5.27}$$

with $\boldsymbol{\lambda}$ taken as a flattening concatenation of $d$-dimensional vectors of free variables $(\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1, \ldots)$.

Since for each $m$ we have $\mathcal{K}^{\langle m \rangle}(\boldsymbol{\lambda}_m, \boldsymbol{\lambda}_{m+1}, z, \boldsymbol{u}) \preceq \mathcal{K}^{\langle \infty \rangle}(\boldsymbol{\lambda}_m, \boldsymbol{\lambda}_{m+1}, z, \boldsymbol{u})$ it also holds

$$\boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) \preceq \boldsymbol{\Psi}(\boldsymbol{\lambda}, z, \boldsymbol{u}). \tag{5.28}$$

The idea of the current proof is to consider the difference $\boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u}) - \boldsymbol{\mathcal{L}}(z, \boldsymbol{u})$ and show that it is non-negative. According to (5.26) this difference can be represented as

$$\boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u}) - \boldsymbol{\mathcal{L}}(z, \boldsymbol{u}) = \boldsymbol{\Psi}(\boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u}), z, \boldsymbol{u}) - \boldsymbol{\Phi}(\boldsymbol{\mathcal{L}}(z, \boldsymbol{u}), z, \boldsymbol{u}). \tag{5.29}$$

Since $\boldsymbol{\Psi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) \succeq \boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$, the formal power series $\boldsymbol{\Psi}$ can be represented as a sum $\boldsymbol{\Psi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) = \boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) + \boldsymbol{\Theta}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ with $\boldsymbol{\Theta}(\boldsymbol{\lambda}, z, \boldsymbol{u}) \succeq \boldsymbol{0}$. Hence, the difference (5.29) becomes

$$\boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u}) - \boldsymbol{\mathcal{L}}(z, \boldsymbol{u}) = \boldsymbol{\Theta}(\boldsymbol{\mathcal{L}}^+, z, \boldsymbol{u}) + \left( \boldsymbol{\Phi}(\boldsymbol{\mathcal{L}}^+, z, \boldsymbol{u}) - \boldsymbol{\Phi}(\boldsymbol{\mathcal{L}}, z, \boldsymbol{u}) \right). \tag{5.30}$$

At this point, our tactic is to apply an analog of the mean value theorem to the right-hand side difference and obtain an equation of the form

$$\boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u}) - \boldsymbol{\mathcal{L}}(z, \boldsymbol{u}) = \boldsymbol{\Theta}(\boldsymbol{\mathcal{L}}^+, z, \boldsymbol{u}) + \boldsymbol{\mathcal{J}} \cdot \left( \boldsymbol{\mathcal{L}}^+ - \boldsymbol{\mathcal{L}} \right) \tag{5.31}$$

and consequently

$$\mathcal{L}^+(z, \boldsymbol{u}) - \mathcal{L}(z, \boldsymbol{u}) = (\mathbf{I} - \boldsymbol{\mathcal{J}})^{-1} \boldsymbol{\Theta} = \sum_{k \geqslant 0} \boldsymbol{\mathcal{J}}^k \boldsymbol{\Theta} \succeq \mathbf{0} \tag{5.32}$$

where $\boldsymbol{\mathcal{J}}$ is some non-negative operator whereas $\mathbf{I}$ is the corresponding identity. The rest of the proof is dedicated to formalising the above approach, in particular showing that the Neumann series $\sum_{k \geqslant 0} \boldsymbol{\mathcal{J}}^k$ is well-defined.

We start by noticing that due to the well-foundedness of the limiting system (5.20) we have $\mathcal{L}^+(0, \boldsymbol{u}) = \mathbf{0}$. Since $\boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) \preceq \boldsymbol{\Psi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ there also holds $\boldsymbol{\Phi}(\mathbf{0}, 0, \boldsymbol{u}) = \mathbf{0}$. Furthermore, since there exists a unique formal power series solution of the equation $\mathcal{L}(0, \boldsymbol{u}) = \boldsymbol{\Phi}(\mathcal{L}(0, \boldsymbol{u}), 0, \boldsymbol{u})$ and $\mathcal{L}(0, \boldsymbol{u}) = \mathbf{0}$ satisfies this equation, we note that indeed $\mathcal{L}(0, \boldsymbol{u}) = \mathbf{0}$.

Consider two formal infinite-dimensional variables $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^+$ which are both flattened concatenations of $d$-dimensional vectors of free variables. Let us show that the difference $\boldsymbol{\Phi}(\boldsymbol{\lambda}^+, z, \boldsymbol{u}) - \boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ can be represented as

$$\boldsymbol{\Phi}(\boldsymbol{\lambda}^+, z, \boldsymbol{u}) - \boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u}) = \boldsymbol{\mathcal{J}}(z, \boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+)(\boldsymbol{\lambda}^+ - \boldsymbol{\lambda}) \tag{5.33}$$

where $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{J}}(z, \boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+)$ is some non-negative operator (i.e. infinite-dimensional matrix). Moreover, after substituting $\boldsymbol{\lambda} = \mathcal{L}(z, \boldsymbol{u})$ and $\boldsymbol{\lambda}^+ = \mathcal{L}^+(z, \boldsymbol{u})$ into $\boldsymbol{\mathcal{J}}$, there exists a non-negative integer $K > 0$ such that $\boldsymbol{\mathcal{J}}^K$ is element-wise divisible by $z$. Since we have established that both $\mathcal{L}^+(0, \boldsymbol{u}) = \mathbf{0}$ and $\mathcal{L}(0, \boldsymbol{u}) = \mathbf{0}$, the latter condition is equivalent to the nilpotency of the operator $\boldsymbol{\mathcal{J}}$ evaluated at $\boldsymbol{\lambda} = \mathcal{L}(z, \boldsymbol{u})$, $\boldsymbol{\lambda}^+ = \mathcal{L}^+(z, \boldsymbol{u})$ and $z = 0$.

Note that the function $\boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ is a sum of (finite) monomials in formal variables $(\boldsymbol{\lambda}, z, \boldsymbol{u})$; although $\boldsymbol{\lambda}$ is infinitely-dimensional, each of the monomials involves only finitely many factors of $\boldsymbol{\lambda}$. Let us consider the difference of arbitrary monomials in form of

$$x_1^{n_1} \cdots x_k^{n_k} - y_1^{n_1} \cdots y_k^{n_k}. \tag{5.34}$$

Note that we can rewrite (5.34) as

$$\begin{aligned}
x_1^{n_1} \cdots x_k^{n_k} - y_1^{n_1} \cdots y_k^{n_k} &= \left(x_1^{n_1} \cdots x_k^{n_k} - y_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}\right) \\
&\quad + \left(y_1^{n_1} x_2^{n_2} \cdots x_k^{n_k} - y_1^{n_1} y_2^{n_2} x_3^{n_3} \cdots x_k^{n_k}\right) + \cdots \\
&\quad + \left(y_1^{n_1} \cdots y_{k-1}^{n_{k-1}} x_k^{n_k} - y_1^{n_1} \cdots y_k^{n_k}\right) \\
&= (x_1 - y_1) \, x_2^{n_2} \cdots x_k^{n_k} \frac{(x_1^{n_1} - y_1^{n_1})}{x_1 - y_1} \\
&\quad + (x_2 - y_2) \, y_1^{n_1} x_3^{n_3} \cdots x_k^{n_k} \frac{(x_2^{n_2} - y_2^{n_2})}{x_2 - y_2} + \cdots \\
&\quad + (x_k - y_k) \, y_1^{n_1} \cdots y_{k-1}^{n_{k-1}} \frac{(x_k^{n_k} - y_k^{n_k})}{x_k - y_k}.
\end{aligned} \tag{5.35}$$

Note that each factor $\dfrac{(x_i^{n_i} - y_i^{n_i})}{x_i - y_i}$ in the final sum is in fact a polynomial $\sum\limits_{j=0}^{n_i-1} x_i^j y_i^{n_i-j-1}$. Therefore, the difference $x_1^{n_1} \cdots x_k^{n_k} - y_1^{n_1} \cdots y_k^{n_k}$ can be represented as a scalar product

of $\boldsymbol{x} - \boldsymbol{y} := (x_i - y_i)_{i=1}^k$ and a vector of formal power series in $(\boldsymbol{x}, \boldsymbol{y})$. Furthermore, the difference $\boldsymbol{\Phi}(\boldsymbol{\lambda}^+, z, \boldsymbol{u}) - \boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ consists of the sums of such differences of monomials multiplied by appropriate non-negative coefficients. Grouping these differences together, we obtain the desired form (5.33).

Next, as an intermediate step, let us now show that the Jacobian operator $\dfrac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\lambda}}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ is nilpotent at $(z, \boldsymbol{\lambda}) = (0, \boldsymbol{0})$. For convenience, set

$$J_1(\boldsymbol{u}) := \left. \frac{\partial \boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u})}{\partial \boldsymbol{\ell}_1} \right|_{\substack{z=0 \\ \boldsymbol{\ell}_1 = \boldsymbol{0} \\ \boldsymbol{\ell}_2 = \boldsymbol{0}}} \quad \text{and} \quad J_2(\boldsymbol{u}) := \left. \frac{\partial \boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, z, \boldsymbol{u})}{\partial \boldsymbol{\ell}_2} \right|_{\substack{z=0 \\ \boldsymbol{\ell}_1 = \boldsymbol{0} \\ \boldsymbol{\ell}_2 = \boldsymbol{0}}}. \tag{5.36}$$

Then,

$$J_1(\boldsymbol{u}) + J_2(\boldsymbol{u}) = \left. \frac{\partial \boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{\ell}, \boldsymbol{\ell}, z, \boldsymbol{u})}{\partial \boldsymbol{\ell}} \right|_{\substack{z=0 \\ \boldsymbol{\ell} = \boldsymbol{0}}}. \tag{5.37}$$

Since the limiting system is well-founded (see Theorem A.3) the sum $J_1(\boldsymbol{u}) + J_2(\boldsymbol{u})$ is nilpotent. Moreover, since each of the matrices $J_1(\boldsymbol{u})$ and $J_2(\boldsymbol{u})$ is non-negative, there exists $K$ such that all the summands of the expanded binomial $(J_1(\boldsymbol{u}) + J_2(\boldsymbol{u}))^K$ are zero.

On the other hand, note that following the definition of $\boldsymbol{\Psi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ its Jacobian operator $\dfrac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\lambda}}$ evaluated at $(z, \boldsymbol{\lambda}) = (0, \boldsymbol{0})$ admits the following block structure:

$$\left. \frac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\lambda}} \right|_{\substack{z=0 \\ \boldsymbol{\lambda} = \boldsymbol{0}}} = \begin{bmatrix} J_1 & J_2 & & \\ & J_1 & J_2 & \\ & & J_1 & J_2 \\ & & & \ddots \end{bmatrix}. \tag{5.38}$$

If we take the $K$th power of this operator, it will have a block structure in which each block element consists of a sum of certain summands from the binomial expansion of $(J_1 + J_2)^K$. Since the latter is a zero matrix and the summands corresponding to the blocks are non-negative and dominated by $(J_1 + J_2)^K$, all such summands are also zero. This implies that the Jacobian operator $\dfrac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\lambda}}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ evaluated at $(z, \boldsymbol{\lambda}) = (0, \boldsymbol{0})$ is nilpotent with a nilpotence index at most $K$, i.e. the nilpotence index of the Jacobian operator $\dfrac{\partial \boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{\ell}, \boldsymbol{\ell}, z, \boldsymbol{u})}{\partial \boldsymbol{\ell}}$ evaluated at $(z, \boldsymbol{\ell}) = (0, \boldsymbol{0})$.

Now, let us show that the infinitely-dimensional matrix $\boldsymbol{\mathcal{J}}(z, \boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+)$ evaluated at $(z, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+) = (0, \boldsymbol{0}, \boldsymbol{0})$ is equal to the Jacobian operator $\dfrac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\lambda}}$ evaluated at $(z, \boldsymbol{\lambda}) = (0, \boldsymbol{0})$. Recall that the operator $\boldsymbol{\mathcal{J}}$ is determined by the differences of monomials in $\boldsymbol{\Phi}(\boldsymbol{\lambda}^+, z, \boldsymbol{u}) - \boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$. Monomials that have degree zero in $\boldsymbol{\lambda}$ or $\boldsymbol{\lambda}^+$ cancel out because they depend only on the arguments $z$ and $\boldsymbol{u}$. Likewise, monomials with degree two or more in $\boldsymbol{\lambda}$ or $\boldsymbol{\lambda}^+$ vanish after the substitution $\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ = \boldsymbol{0}$. The only type of the terms that do not turn to zero upon substitution $\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ = \boldsymbol{0}$ are terms coming from differences of monomials linear

in $\boldsymbol{\lambda}$ or $\boldsymbol{\lambda}^+$. Note that such terms have the same contribution to the infinitely-dimensional matrix $\boldsymbol{\mathcal{J}}$ as the corresponding terms of the infinitely-dimensional matrix $\dfrac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\lambda}}$. Hence, $\boldsymbol{\mathcal{J}}(z, \boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+)$ evaluated at $(z, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+) = (0, \boldsymbol{0}, \boldsymbol{0})$ is indeed equal to the Jacobian operator $\dfrac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\lambda}}$ evaluated at $(z, \boldsymbol{\lambda}) = (0, \boldsymbol{0})$.

The nilpotence of $\boldsymbol{\mathcal{J}}$ evaluated at $(z, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+) = (0, \boldsymbol{0}, \boldsymbol{0})$ follows from the fact that $\boldsymbol{\Psi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$ is dominating $\boldsymbol{\Phi}(\boldsymbol{\lambda}, z, \boldsymbol{u})$, and therefore, the corresponding Jacobian operator $\dfrac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\lambda}}$ (respectively its $N$th power) is dominating the operator $\dfrac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\lambda}}$ (respectively its $N$th power). For this reason, the latter Jacobian operator $\dfrac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\lambda}}$, evaluated at $(z, \boldsymbol{\lambda}, \boldsymbol{\lambda}^+) = (0, \boldsymbol{0}, \boldsymbol{0})$ is nilpotent with the nilpotence index at most the corresponding nilpotence operator of the former Jacobian operator $\dfrac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\lambda}}$.

And so, we have established that $\boldsymbol{\mathcal{J}}$ evaluated at $\boldsymbol{\lambda} = \boldsymbol{\mathcal{L}}(z, \boldsymbol{u})$, $\boldsymbol{\lambda}^+ = \boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u})$ and $z = 0$ is nilpotent. Equivalently, it meas that after substituting $\boldsymbol{\lambda} = \boldsymbol{\mathcal{L}}(z, \boldsymbol{u})$ and $\boldsymbol{\lambda}^+ = \boldsymbol{\mathcal{L}}^+(z, \boldsymbol{u})$ into $\boldsymbol{\mathcal{J}}$ there exists a non-negative integer $K > 0$ such that $\boldsymbol{\mathcal{J}}^K$ is element-wise divisible by $z$. Consequently, each coefficient in $z$ of the formal sum $\sum_{j \geqslant 0} \boldsymbol{\mathcal{J}}^j$ is finite. Indeed, for each integer $N \geqslant 0$, the coefficient at $z^N$ in this formal series is a sum of coefficients at $z^N$ in the finite sum $\sum_{j=0}^{K \cdot N} \boldsymbol{\mathcal{J}}^j$. Moreover, since $\boldsymbol{\mathcal{J}}$ is non-negative, this sum is also non-negative. Finally, this infinite formal series is equal to $(\boldsymbol{I} - \boldsymbol{\mathcal{J}})^{-1}$ where $\boldsymbol{I}$ is the identity operator of appropriate dimension. $\qquad \square$

*Remark* 5.8. The condition that $\boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{0}, \boldsymbol{0}, 0, \boldsymbol{u}) = \boldsymbol{0}$ can be omitted but we keep it for the simplicity of the proof. For the above proof, it is enough to guarantee that each coefficient in $z$ of the infinite formal sum $\sum_{j \geqslant 0} \boldsymbol{\mathcal{J}}^j$ is finite, which is equivalent to saying that some power of $\boldsymbol{\mathcal{J}}$ is divisible by $z$. More details on well-founded systems can be found in [PSS12].

## 5.3 Coefficient transfer for infinite systems

Finally, we give our main theorem on the transfer of coefficients for infinitely nested forward-recursive systems.

**Theorem 5.9.** *Let $\mathcal{S}$ be an infinitely nested, forward recursive system (5.19) which tends to an irreducible context-free schema at an exponential convergence rate. Then, the respective solutions $L_j^{\langle m \rangle}(z, \boldsymbol{u})$ of $\mathcal{S}$ admit for each $m \geqslant 0$ an asymptotic expansion of their coefficients as $n \to \infty$ in form of*

$$[z^n] L_j^{\langle m \rangle}(z, \boldsymbol{u}) \sim [z^n] \sum_{k \geqslant 0} c_{j,k}^{\langle m \rangle}(\boldsymbol{u}) \left(1 - \frac{z}{\rho(\boldsymbol{u})}\right)^{k/2} \tag{5.39}$$

*where $\rho(\boldsymbol{u})$ is the dominant singularity of the corresponding limiting system (5.20) and the coefficients $c_{j,k}^{\langle m \rangle}(\boldsymbol{u})$ are analytic at $\boldsymbol{u} = \boldsymbol{1}$. Furthermore, $\rho(\boldsymbol{u})$ is analytic near $\boldsymbol{u} = \boldsymbol{1}$.*

The coefficients $c_{j,k}^{\langle m \rangle}(\boldsymbol{u})$ can be approximated by taking first $(M-1)$ equations from (5.19) and replacing the $M$th equation by its following limit variant:

$$\boldsymbol{L}^{\langle M \rangle} = \boldsymbol{\mathcal{K}}^{\langle \infty \rangle}\left( \boldsymbol{L}^{\langle M \rangle}, \boldsymbol{L}^{\langle M \rangle}, z, \boldsymbol{u} \right). \tag{5.40}$$

*Such a truncated system can be solved recursively. Consequently, the coefficients of respective Puiseux expansions (5.39) are estimated with an error which is exponentially small in $M$.*

*Remark* 5.10.  • The condition that the system $\mathcal{S}$ tends to an irreducible context-free schema can be replaced by a weaker condition asserting that the limiting system (5.20) admits a suitable Puiseux expansion.

• Instead of limiting systems with square-root type singularities, it is also possible to consider rational systems or other types of systems. The same set of conditions is sufficient to establish the transfer of behaviours around the dominant singular point.

*Proof.* (Theorem 5.9)  We divide the proof into three conceptual parts.

a) First, we show that each component of the difference vector $\boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u}) - \boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u})$ can be upper bounded by a Puiseux series expansion whose coefficients decay exponentially fast as $m$ tends to infinity;

b) Next, we show that if for $m \geqslant 1$ there exist coordinate-wise upper and lower bounds

$$\underline{\boldsymbol{L}^{\langle m \rangle}}(z, \boldsymbol{u}) \preceq \boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u}) \preceq \overline{\boldsymbol{L}^{\langle m \rangle}}(z, \boldsymbol{u}), \tag{5.41}$$

then the vector of functions $\boldsymbol{L}^{\langle m-1 \rangle}(z, \boldsymbol{u})$ obtained from the infinite system $\mathcal{S}$ admits upper and lower bounds $\overline{\boldsymbol{L}^{\langle m-1 \rangle}}(z, \boldsymbol{u})$ and $\underline{\boldsymbol{L}^{\langle m-1 \rangle}}(z, \boldsymbol{u})$ satisfying

$$\overline{\boldsymbol{L}^{\langle m-1 \rangle}}(z, \boldsymbol{u}) - \underline{\boldsymbol{L}^{\langle m-1 \rangle}}(z, \boldsymbol{u}) \preceq \boldsymbol{\mathcal{R}}(z, \boldsymbol{u})\left( \overline{\boldsymbol{L}^{\langle m \rangle}}(z, \boldsymbol{u}) - \underline{\boldsymbol{L}^{\langle m \rangle}}(z, \boldsymbol{u}) \right) \tag{5.42}$$

for some matrix $\boldsymbol{\mathcal{R}}(z, \boldsymbol{u})$ with spectral radius satisfying $r(\boldsymbol{\mathcal{R}}(z, \boldsymbol{u})) \leqslant 1$ for $z \in [0, \rho(\boldsymbol{u})]$ where $\rho(\boldsymbol{u})$ is the dominant singularity of $\boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u})$;

c) Finally, we combine two previous results and prove that the difference between the Puiseux coefficients of upper and lower bounds of $\boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u})$ can be reduced to zero.

*First part.* According to Theorem 5.7 we have $\boldsymbol{L}^{\langle m \rangle} \preceq \boldsymbol{L}^{\langle \infty \rangle}$. Following the functional definitions of $\boldsymbol{L}^{\langle \infty \rangle}$ and $\boldsymbol{L}^{\langle m \rangle}$ from the infinite system of equations, their difference can be represented as

$$\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m \rangle} = \boldsymbol{\mathcal{K}}^{\langle \infty \rangle}(\boldsymbol{L}^{\langle \infty \rangle}, \boldsymbol{L}^{\langle \infty \rangle}, z, \boldsymbol{u}) - \boldsymbol{\mathcal{K}}^{\langle m \rangle}(\boldsymbol{L}^{\langle m \rangle}, \boldsymbol{L}^{\langle m+1 \rangle}, z, \boldsymbol{u}). \tag{5.43}$$

For convenience, henceforth we omit the arguments $z$ and $\boldsymbol{u}$. Moreover, $\boldsymbol{\mathcal{K}}^{\langle \infty \rangle}$ and $\boldsymbol{\mathcal{K}}^{\langle m \rangle}$ become functions of two vector arguments

$$\boldsymbol{\mathcal{K}}^{\langle \infty \rangle} \colon \mathbb{C}^d \times \mathbb{C}^d \to \mathbb{C}^d \quad \text{and} \quad \boldsymbol{\mathcal{K}}^{\langle m \rangle} \colon \mathbb{C}^d \times \mathbb{C}^d \to \mathbb{C}^d. \tag{5.44}$$

In addition, we use the nabla notation to denote the Jacobian operator

$$\nabla_{\boldsymbol{x}}\mathcal{K}(\boldsymbol{x},\boldsymbol{y}) = \left(\frac{\partial}{\partial x_1}\mathcal{K},\ldots,\frac{\partial}{\partial x_d}\mathcal{K}\right) \quad\text{and}\quad \nabla_{\boldsymbol{y}}\mathcal{K}(\boldsymbol{x},\boldsymbol{y}) = \left(\frac{\partial}{\partial y_1}\mathcal{K},\ldots,\frac{\partial}{\partial y_d}\mathcal{K}\right). \quad (5.45)$$

We start with the following subtraction-addition trick. Each component of the vector difference is evaluated through Theorem 5.4 (mean value lemma for formal power series) and then upper-bounded by the value of the functional $\mathcal{K}^{\langle\infty\rangle}$ at $\boldsymbol{L}^{\langle\infty\rangle}$. Specifically,

$$\begin{aligned}
\boldsymbol{L}^{\langle\infty\rangle} - \boldsymbol{L}^{\langle m\rangle} &= \mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle\infty\rangle},\boldsymbol{L}^{\langle\infty\rangle}) - \mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle m\rangle},\boldsymbol{L}^{\langle\infty\rangle}) + \mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle m\rangle},\boldsymbol{L}^{\langle\infty\rangle}) \\
&\quad - \mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle m\rangle},\boldsymbol{L}^{\langle m+1\rangle}) + \mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle m\rangle},\boldsymbol{L}^{\langle m+1\rangle}) - \mathcal{K}^{\langle m\rangle}(\boldsymbol{L}^{\langle m\rangle},\boldsymbol{L}^{\langle m+1\rangle}) \\
&\preceq \nabla_{\boldsymbol{x}}\mathcal{K}^{\langle\infty\rangle}(\boldsymbol{x},\boldsymbol{y})\Big|_{\substack{\boldsymbol{x}=\boldsymbol{L}^{\langle\infty\rangle}\\\boldsymbol{y}=\boldsymbol{L}^{\langle\infty\rangle}}}(\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m\rangle}) + \nabla_{\boldsymbol{y}}\mathcal{K}^{\langle\infty\rangle}(\boldsymbol{x},\boldsymbol{y})\Big|_{\substack{\boldsymbol{x}=\boldsymbol{L}^{\langle\infty\rangle}\\\boldsymbol{y}=\boldsymbol{L}^{\langle\infty\rangle}}}(\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m+1\rangle}) \\
&\quad + \left(\mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle m\rangle},\boldsymbol{L}^{\langle m+1\rangle}) - \mathcal{K}^{\langle m\rangle}(\boldsymbol{L}^{\langle m\rangle},\boldsymbol{L}^{\langle m+1\rangle})\right).
\end{aligned}$$

Since $\mathcal{K}^{\langle m\rangle} \preceq \mathcal{K}^{\langle\infty\rangle}$, the final difference in the above sum is a vector of formal power series with non-negative coefficients. Consequently, the last summand can be bounded by $\mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle\infty\rangle},\boldsymbol{L}^{\langle\infty\rangle}) - \mathcal{K}^{\langle m\rangle}(\boldsymbol{L}^{\langle\infty\rangle},\boldsymbol{L}^{\langle\infty\rangle})$. By the condition of exponential convergence, this difference can be even further bounded by a vector of exponentially decaying functions $\boldsymbol{A}(z,\boldsymbol{u})(B(z,\boldsymbol{u}))^m$. For brevity, set

$$\partial_1\mathcal{K} := \nabla_{\boldsymbol{x}}\mathcal{K}^{\langle\infty\rangle}(\boldsymbol{x},\boldsymbol{y})\Big|_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{L}^{\langle\infty\rangle},\boldsymbol{L}^{\langle\infty\rangle})} \quad\text{and}\quad \partial_2\mathcal{K} := \nabla_{\boldsymbol{y}}\mathcal{K}^{\langle\infty\rangle}(\boldsymbol{x},\boldsymbol{y})\Big|_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{L}^{\langle\infty\rangle},\boldsymbol{L}^{\langle\infty\rangle})}.$$

Note that $\partial_1\mathcal{K},\partial_2\mathcal{K}\in\mathbb{C}^{d\times d}$. Now, the upper bound on $\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m\rangle}$ can be stated as

$$(\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m\rangle}) \preceq \boldsymbol{A}(z,\boldsymbol{u})B(z,\boldsymbol{u})^m + \partial_1\mathcal{K}\cdot(\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m\rangle}) + \partial_2\mathcal{K}\cdot(\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m+1\rangle}) \quad (5.46)$$

or, equivalently,

$$(\boldsymbol{I}_d-\partial_1\mathcal{K})(\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m\rangle}) \preceq \boldsymbol{A}(z,\boldsymbol{u})B(z,\boldsymbol{u})^m + \partial_2\mathcal{K}\cdot(\boldsymbol{L}^{\langle\infty\rangle}-\boldsymbol{L}^{\langle m+1\rangle}) \quad (5.47)$$

where $\boldsymbol{I}_d\in\mathbb{C}^{d\times d}$ is the $d\times d$ identity matrix.

Let us show that the inverse matrix $(\boldsymbol{I}_d-\partial_1\mathcal{K})^{-1}$ exists and has non-negative coefficients in the sense of formal power series. As discussed in the proof of Theorem 5.7, the matrix $\partial_1\mathcal{K}$ is nilpotent at $z=0$. Equivalently, there exists a non-negative integer $K$ such that $(\partial_1\mathcal{K})^K$ is divisible by $z$. It means that the formal series $\sum_{j\geqslant 0}(\partial_1\mathcal{K})^j$ is well-defined and hence so is the formal inverse $(\boldsymbol{I}_d-\partial_1\mathcal{K})^{-1}$. Moreover, since $\partial_1\mathcal{K}$ has non-negative coefficients, the same holds for the investigated inverse matrix $(\boldsymbol{I}_d-\partial_1\mathcal{K})^{-1}$.

Now, let us focus on the behaviour of $(\boldsymbol{I}_d-\partial_1\mathcal{K})^{-1}$ as a function near $z=\rho(\boldsymbol{u})$. As follows from Proposition A.7 applied to the system of equations $\boldsymbol{L}^{\langle\infty\rangle}=\mathcal{K}^{\langle\infty\rangle}(\boldsymbol{L}^{\langle\infty\rangle},\boldsymbol{L}^{\langle\infty\rangle},z,\boldsymbol{u})$, for each real $0<z<\rho(\boldsymbol{u})$ we have the following inequality:

$$r(\partial_1\mathcal{K}+\partial_2\mathcal{K}) < 1. \quad (5.48)$$

By Perron–Frobenius theorem (see e.g. [Drm09, section 2.2.5] and references therein) the spectral radius of a matrix with positive entries is monotonic in its coefficients. Hence, for all real $0 < z < \rho(\boldsymbol{u})$

$$r(\partial_1 \boldsymbol{\mathcal{K}}) < 1 \tag{5.49}$$

and so, in the same interval

$$(\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} = \sum_{j \geqslant 0} (\partial_1 \boldsymbol{\mathcal{K}})^j. \tag{5.50}$$

Moreover, due to the continuity of the spectral radius, the same identity can be extended to some complex neighbourhood of $\boldsymbol{u} = \boldsymbol{1}$.

Consequently, we can multiply both sides of (5.47) by $(\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1}$ and obtain

$$\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m \rangle} \preceq (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \boldsymbol{A}(z, u) B(z, \boldsymbol{u})^m + (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \partial_2 \boldsymbol{\mathcal{K}} (\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m+1 \rangle}) \tag{5.51}$$

Let us denote $\boldsymbol{\delta}_m := (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \boldsymbol{A}(z, \boldsymbol{u}) B(z, \boldsymbol{u})^m$ and $\boldsymbol{\mathcal{R}} := (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \partial_2 \boldsymbol{\mathcal{K}}$. Note that the inequality $\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m \rangle} \preceq \boldsymbol{\delta}_m + \boldsymbol{\mathcal{R}}(\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m+1 \rangle})$ can be further iterated for increasing values of $m$. In doing so, we find that

$$\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m \rangle} \preceq \boldsymbol{\delta}_m + \boldsymbol{\mathcal{R}} \boldsymbol{\delta}_{m+1} + \boldsymbol{\mathcal{R}}^2 \boldsymbol{\delta}_{m+2} + \cdots \tag{5.52}$$

Hence, the difference $\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m \rangle}$ can be bounded by the tail of a geometric progression which appears as a summation of the formal Neumann series

$$\boldsymbol{L}^{\langle \infty \rangle} - \boldsymbol{L}^{\langle m \rangle} \preceq B(z, u)^m \sum_{k \geqslant 0} (B(z, \boldsymbol{u}) \boldsymbol{\mathcal{R}})^k (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \boldsymbol{A}(z, \boldsymbol{u}). \tag{5.53}$$

Let us now focus on the above formal Neumann series. Note that applying Proposition A.5 (Drmota–Lalley–Woods theorem) to the limiting system (5.20), we obtain that the vector of functions $\boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u})$ admits a coordinate-wise Puiseux expansion in form of

$$\boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u}) \sim \boldsymbol{\ell}_0(\boldsymbol{u}) - \boldsymbol{\ell}_1(\boldsymbol{u}) \sqrt{1 - \frac{z}{\rho(\boldsymbol{u})}} \tag{5.54}$$

where functions $\boldsymbol{\ell}_0(\boldsymbol{u}), \boldsymbol{\ell}_1(\boldsymbol{u}), \rho(\boldsymbol{u})$ are analytic near $\boldsymbol{u} = 1$. Likewise, both matrices $\partial_1 \boldsymbol{\mathcal{K}}$ and $\partial_2 \boldsymbol{\mathcal{K}}$ admit Puiseux expansions of the same kind.

Let us prove that coordinate-wise Puiseux expansions of the matrix $\boldsymbol{\mathcal{R}}$ near the singular point $z = \rho(\boldsymbol{u})$ have the form

$$\boldsymbol{\mathcal{R}} = (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \partial_2 \boldsymbol{\mathcal{K}} \sim \boldsymbol{\mathcal{R}}_0 - \boldsymbol{\mathcal{R}}_1 \sqrt{1 - \frac{z}{\rho(\boldsymbol{u})}}, \quad z \to \rho(\boldsymbol{u}) \tag{5.55}$$

where the spectral radius of $\boldsymbol{\mathcal{R}}_0$ satisfies $r(\boldsymbol{\mathcal{R}}_0) = 1$. According to Perron–Frobenius theorem, since the coefficients of $\boldsymbol{\mathcal{R}}$ are non-negative, the eigenvalue of the matrix $\boldsymbol{\mathcal{R}}$ with the largest absolute value, i.e. the eigenvalue corresponding to the spectral radius of $\boldsymbol{\mathcal{R}}$, is the largest real positive solution $\lambda$ of the characteristic equation

$$\det \left( (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \partial_2 \boldsymbol{\mathcal{K}} - \lambda \boldsymbol{I}_d \right) = 0. \tag{5.56}$$

Since the determinant of a matrix product is equal to the product of respective determinants and $\det\left(\boldsymbol{I}_d - \partial_1\boldsymbol{\mathcal{K}}\right) \neq 0$, as $\boldsymbol{I}_d - \partial_1\boldsymbol{\mathcal{K}}$ is invertible, the above condition is equivalent to

$$\det\left(\partial_2\boldsymbol{\mathcal{K}} + \lambda\partial_1\boldsymbol{\mathcal{K}} - \lambda\boldsymbol{I}_d\right) = 0 \quad \text{and also} \quad \det(\partial_1\boldsymbol{\mathcal{K}} + \lambda^{-1}\partial_2\boldsymbol{\mathcal{K}} - \boldsymbol{I}_d) = 0. \qquad (5.57)$$

Let us show that the largest positive real solution (as a function of $z < \rho(\boldsymbol{u})$) of this equation, does not exceed 1, with equality when $z = \rho(\boldsymbol{u})$. Assume, by contrary, that $\lambda > 1$. The matrix $(\partial_1\boldsymbol{\mathcal{K}} + \lambda^{-1}\partial_2\boldsymbol{\mathcal{K}})$ is a matrix with non-negative coefficients, whose coefficients are strictly smaller than the coefficients of the matrix $(\partial_1\boldsymbol{\mathcal{K}} + \partial_2\boldsymbol{\mathcal{K}})$. By Perron–Frobenius theorem (see e.g. [Drm09, section 2.2.5] and references therein), the spectral radius of a matrix with positive coefficients is monotonic in its coefficients, so for $\lambda > 1$

$$r(\partial_1\boldsymbol{\mathcal{K}} + \lambda^{-1}\partial_2\boldsymbol{\mathcal{K}}) < r(\partial_1\boldsymbol{\mathcal{K}} + \partial_2\boldsymbol{\mathcal{K}}) = 1. \qquad (5.58)$$

Therefore, the characteristic equation cannot have a solution $\lambda > 1$. Moreover, the spectral radius of $\boldsymbol{\mathcal{R}}_0$ is equal to the spectral radius of $\boldsymbol{\mathcal{R}}$ when $z = \rho(\boldsymbol{u})$ because in this case, the two matrices coincide. Therefore, $r(\boldsymbol{\mathcal{R}}_0) = 1$.

Moving back to the upper bound (5.53), according to the exponential convergence condition in Theorem 5.5, in a complex vicinity of $\boldsymbol{u} = \boldsymbol{1}$, the absolute value of the function $B(z, \boldsymbol{u})$ is strictly smaller than 1, hence the inverse matrix $(\boldsymbol{I}_d - B(z, u)\boldsymbol{\mathcal{R}})^{-1}$ in (5.53) exists. Moreover, since

$$A^{-1} = \frac{1}{\det(A)} \cdot \mathbf{adj}(A) \qquad (5.59)$$

where $\mathbf{adj}(A)$ is the adjugate matrix of $A$, each element of the inverse matrix $(\boldsymbol{I}_d - B(z, u)\boldsymbol{\mathcal{R}})^{-1}$ can be represented as a ratio of a sum of products of functions admitting Puiseux series expansions in form of $\boldsymbol{a}(\boldsymbol{u}) - \boldsymbol{b}(\boldsymbol{u})\sqrt{1 - z/\rho(\boldsymbol{u})} + O(|1 - z/\rho(\boldsymbol{u})|)$ and a non-zero determinant of $\boldsymbol{I}_d - B(z, u)\boldsymbol{\mathcal{R}}$. It means that each coordinate in the inverse matrix $(\boldsymbol{I}_d - B(z, u)\boldsymbol{\mathcal{R}})^{-1}$ also admits a Puiseix series expansion of similar form. Thus, the Neumann series in (5.53) converges and we obtain

$$\boldsymbol{L}^{\langle\infty\rangle} - \boldsymbol{L}^{\langle m\rangle} \preceq B(z, \boldsymbol{u})^m \times (\boldsymbol{I}_d - B(z, \boldsymbol{u})\boldsymbol{\mathcal{R}})^{-1}(\boldsymbol{I}_d - \partial_1\boldsymbol{\mathcal{K}})^{-1}\boldsymbol{A}(z, \boldsymbol{u}). \qquad (5.60)$$

From (5.60) we now note that the difference $\boldsymbol{L}^{\langle\infty\rangle} - \boldsymbol{L}^{\langle m\rangle}$ can be bounded by a vector of functions having the same singularity $\rho(\boldsymbol{u})$ as the components of the vector $\boldsymbol{L}^{\langle\infty\rangle}$. The Puiseux coefficients of this upper bound decay exponentially fast as $m \to \infty$. Moreover, these coefficients are analytic functions near $\boldsymbol{u} = \boldsymbol{1}$.

*Second part.* Assume that function $\boldsymbol{L}^{\langle m\rangle}(z, \boldsymbol{u})$ admits some upper and lower bounds and denote by $\Delta_m(z, \boldsymbol{u})$ the difference between these bounds:

$$\underline{\boldsymbol{L}^{\langle m\rangle}}(z, \boldsymbol{u}) \preceq \boldsymbol{L}^{\langle m\rangle}(z, \boldsymbol{u}) \preceq \overline{\boldsymbol{L}^{\langle m\rangle}}(z, \boldsymbol{u}), \quad \Delta^{\langle m\rangle}(z, \boldsymbol{u}) := \overline{\boldsymbol{L}^{\langle m\rangle}}(z, \boldsymbol{u}) - \underline{\boldsymbol{L}^{\langle m\rangle}}(z, \boldsymbol{u}). \quad (5.61)$$

Then, another pair of upper and lower bound can be established for $\boldsymbol{L}^{\langle m-1\rangle}$ from (5.19) (infinite system of equations) with the difference $\Delta^{\langle m-1\rangle}$ satisfying

$$\Delta^{\langle m-1\rangle} = \boldsymbol{\mathcal{K}}^{\langle m-1\rangle}(\overline{\boldsymbol{L}^{\langle m-1\rangle}}, \overline{\boldsymbol{L}^{\langle m\rangle}}) - \boldsymbol{\mathcal{K}}^{\langle m-1\rangle}(\underline{\boldsymbol{L}^{\langle m-1\rangle}}, \underline{\boldsymbol{L}^{\langle m\rangle}}) \preceq \partial_1\boldsymbol{\mathcal{K}} \cdot \Delta^{\langle m-1\rangle} + \partial_2\boldsymbol{\mathcal{K}} \cdot \Delta^{\langle m\rangle}.$$
$$(5.62)$$

That is, repeating the argument that allows to multiply both sides of the inequality by the inverse matrix, we obtain

$$\Delta^{\langle m-1\rangle} \preceq (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \partial_2 \boldsymbol{\mathcal{K}} \cdot \Delta^{\langle m\rangle}. \tag{5.63}$$

As we discovered in the first part, the matrix $\boldsymbol{\mathcal{R}} := (\boldsymbol{I}_d - \partial_1 \boldsymbol{\mathcal{K}})^{-1} \partial_2 \boldsymbol{\mathcal{K}}$ has spectral radius 1 at the singular point $z = \rho(\boldsymbol{u})$.

*Third part.* As a result of the first part, we know that $\boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u}) - \boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u})$ can be bounded in the following manner:

$$\boldsymbol{0} \preceq \boldsymbol{L}^{\langle \infty \rangle}(z, \boldsymbol{u}) - \boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u}) \preceq B(z, \boldsymbol{u})^m (\boldsymbol{I}_d - B(z, u)\boldsymbol{\mathcal{R}})^{-1} \boldsymbol{A}(z, \boldsymbol{u}). \tag{5.64}$$

Let us assign

$$\Delta_0^{\langle m\rangle} := B(z, \boldsymbol{u})^m (\boldsymbol{I}_d - B(z, u)\boldsymbol{\mathcal{R}})^{-1} \boldsymbol{A}(z, \boldsymbol{u}) \tag{5.65}$$

for the difference between upper and lower bounds for the vector of functions $\boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u})$. Next, using the result of the second part, we construct a family of differences $\Delta_k^{\langle m\rangle}$ between upper and lower bounds for $\boldsymbol{L}^{\langle m\rangle}(z, \boldsymbol{u})$, so that for every $k, m \geqslant 0$ it holds

$$\underline{\boldsymbol{L}_k^{\langle m\rangle}}(z, \boldsymbol{u}) \preceq \boldsymbol{L}^{\langle m\rangle}(z, \boldsymbol{u}) \preceq \overline{\boldsymbol{L}_k^{\langle m\rangle}}(z, \boldsymbol{u}) \quad \text{and} \quad \Delta_k^{\langle m\rangle} := \overline{\boldsymbol{L}_k^{\langle m\rangle}}(z, \boldsymbol{u}) - \underline{\boldsymbol{L}_k^{\langle m\rangle}}(z, \boldsymbol{u}). \tag{5.66}$$

The family of upper and lower bounds is defined using the procedure described in the second part. More specifically, for every $m \geqslant 1$ and $k \geqslant 0$ these bounds satisfy the equations

$$
\begin{aligned}
\overline{\boldsymbol{L}_{k+1}^{\langle m-1\rangle}}(z, \boldsymbol{u}) &:= \boldsymbol{\mathcal{K}}^{\langle m-1\rangle}(\overline{\boldsymbol{L}_{k+1}^{\langle m-1\rangle}}(z, \boldsymbol{u}), \overline{\boldsymbol{L}_k^{\langle m\rangle}}(z, \boldsymbol{u}), z, \boldsymbol{u}); \\
\underline{\boldsymbol{L}_{k+1}^{\langle m-1\rangle}}(z, \boldsymbol{u}) &:= \boldsymbol{\mathcal{K}}^{\langle m-1\rangle}(\underline{\boldsymbol{L}_{k+1}^{\langle m-1\rangle}}(z, \boldsymbol{u}), \underline{\boldsymbol{L}_k^{\langle m\rangle}}(z, \boldsymbol{u}), z, \boldsymbol{u}).
\end{aligned}
\tag{5.67}
$$

According to the second part, the differences $\Delta_k^{\langle m\rangle}$ satisfy formal power series inequalities $\Delta_{k+1}^{\langle m-1\rangle} \preceq \boldsymbol{\mathcal{R}} \Delta_k^{\langle m\rangle}$. By iteration, we thus obtain

$$\Delta_m^{\langle 0\rangle} \preceq \boldsymbol{\mathcal{R}}^m \Delta_0^{\langle m\rangle}. \tag{5.68}$$

Since the spectral radius of $\boldsymbol{\mathcal{R}}$ is bounded by 1, and $\Delta_0^{\langle m\rangle}$ is exponentially small in $m$, the values of Puiseux coefficients of $\boldsymbol{L}^{\langle 0\rangle}(z, \boldsymbol{u})$ can be approximated within an exponentially small in $m$ gap, for arbitrarily large value of $m$.

Finally, we note that the functions $\underline{\boldsymbol{L}_m^{\langle 0\rangle}}(z, \boldsymbol{u})$ and $\overline{\boldsymbol{L}_m^{\langle 0\rangle}}(z, \boldsymbol{u})$ have Puiseux expansions of type

$$f_m(z, \boldsymbol{u}) \sim c_m(\boldsymbol{u}) - a_m(\boldsymbol{u})\sqrt{1 - \frac{z}{\rho(\boldsymbol{u})}} \tag{5.69}$$

in a certain delta-domain. According to Proposition A.1 (transfer theorem), their coefficients admit the following asymptotic estimate:

$$f_m(z, \boldsymbol{u}) \sim C_n A_m(\boldsymbol{u}) B(\boldsymbol{u})^n. \tag{5.70}$$

A final application of Theorem 5.2 (squeeze lemma for formal power series) combined with Theorem 5.3 finishes the proof. □

*Remark* 5.11. In Theorem 5.9 we prove a so-called *weak* transfer theorem, i.e. prove that the asymptotics of the coefficients of each $\boldsymbol{L}^{\langle m \rangle}$ can be obtained by taking the asymptotic expansion of the correseponding Puiseux expansion. A stronger version would suggest that the generating functions $\boldsymbol{L}^{\langle m \rangle}(z, \boldsymbol{u})$ can be analytically continued beyond the circle of convergence of corresponding formal power series, in a certain delta-domain. However, for our analysis, the presented weak variant is enough. The techniques presented above, can be further extended to obtain a stronger transfer theorem, by computing the Taylor series expansions at points $z_0$ inside the circle of convergence.

# 6 Empirical results

The analysis of various combinatorial parameters corresponding to plain $\lambda$-terms can be approached using standard proof templates, typical for algebraic structures. Starting from a combinatorial specification associated with the investigated parameter, its analysis follows as an, either direct or indirect, examination of the resulting system of multivariate generating functions and related singularities, see [FS09, Chapter IX].

Alas, such a general approach to the analysis of combinatorial parameters related to plain $\lambda$-terms does not readily transcend to closed $\lambda$-terms. Standard analytic tools, such as Bender and Richmond's multivariate central limit theorem (see Proposition A.11) or the Drmota–Lalley–Woods theorem (see Proposition A.5) require that the investigated system of generating functions is, *inter alia*, finite. Although this is true for plain terms, closed $\lambda$-terms give rise to a more involved, infinite system of generating functions (2.3) based on the hierarchical notion of $m$-openness. Consequently, closed $\lambda$-terms escape the usual course of parameter analysis, successfully carried out in the case of plain terms, see Section 3.

In the current section we present two, somewhat complementary, empirical approaches to the analysis of combinatorial parameters related to closed $\lambda$-terms. We start with an experimental scheme based on the recent development of efficient Boltzmann samplers for closed $\lambda$-terms due to Gittenberger, Bodini and Gołębiewski [BGG18]. We generate large, uniformly random (conditioned on size) closed $\lambda$-terms and collect empirical data for various interesting parameters related to generated terms. The second approach is based on the empirical evaluation of, appropriately truncated, systems of multivariate generating functions. We compute the coefficients of the corresponding formal power series and consequently analyse the distribution of investigated parameters for relatively small term sizes.

The benefits of such empirical approaches are threefold. Firstly, the empirical parameter distribution for large random $\lambda$-terms is closely related to its theoretical, limiting counterpart; hence, develops solid intuitions underlying the successful analysis of a broad class of combinatorial parameters, see Section 4. Secondly, the empirical data for various term sizes provides insight in the convergence rates at which the considered random variables tend to respective limit laws, relating them in effect with practically attainable term sizes. Finally, experimental results for large closed $\lambda$-terms provide strong evidence for conjectures regarding practical, though even more advanced parameters escaping the

restrictions of our analysis.

## 6.1 Empirical evaluation of Boltzmann samplers

Boltzmann samplers are a prominent sampling framework meant for the random generation of large combinatorial structures [DFLS04]. Given an admissible combinatorial specification, it is possible to construct an appropriate sampler whose outcome are uniformly random (conditioned on size) structures built according to the input specification. Although Boltzmann samplers are guaranteed to return uniformly random structures, their eventual size is not deterministic, but instead random. Following a proper calibration during the construction of Boltzmann samplers, the randomness of their output structure size can be controlled so that it is centred around a given (not necessarily finite) expectation. A final rejection phase, dismissing structures of undesired properties, such as for instance inadmissible size, provides the means of controlling the generated structures.

In order to conduct our experiments, we have implemented three kinds of Boltzmann samplers[1] for plain, closed and so-called $h$-shallow $\lambda$-terms, i.e. closed terms in which De Bruijn indices do not exceed the shallowness bound $h$, see e.g. [GG16, Ben17]. The respective sampler for closed $\lambda$-terms follows the ideas of [BGG18]. For each investigated type of terms we sample $k = 50,000$ terms of sizes in the interval $[20,000; 50,000]$ using a dedicated singular Boltzmann sampler (i.e. a Boltzmann sampler with unbounded outcome size expectation, see e.g. [BGR15]) whose controlling parameter is calculated numerically with accuracy of order $10^{-9}$. For closed $h$-shallow terms, we fix $h = 30$. Finally, we record several combinatorial parameters related to so obtained terms and use them in the subsequent evaluation.

For each considered parameter and term type, we plot a histogram relating the collected samples and the respective parameter values. The $x$-axis denotes the parameter value (either raw or aptly normalised) whereas the $y$-axis corresponds to the number of samples attaining the associated value. Averages, variances and standard deviations corresponding to the investigated parameter are rounded up to the fifth decimal point and summarised in respective tables. For brevity, we include only histograms for plain and closed $\lambda$-terms. We comment on the missing $h$-shallow $\lambda$-terms at the end of the current section.

### 6.1.1 Head abstractions

We start with head abstractions, see Section 3.2. Figure 12 depicts the distribution histogram of the obtained data sets. The corresponding numerical approximations of averages, variances and standard deviations are listed in Table 2.

*Remark* 6.1. The distribution corresponding to plain terms resembles a geometric law, as expected by Proposition 3.6. The observed average corresponds closely to the theoretical limit average (3.18). Notably, head abstractions in closed $\lambda$-terms do not follow the same distribution as their plain counterpart.

---

[1]see https://hackage.haskell.org/package/lambda-sampler.
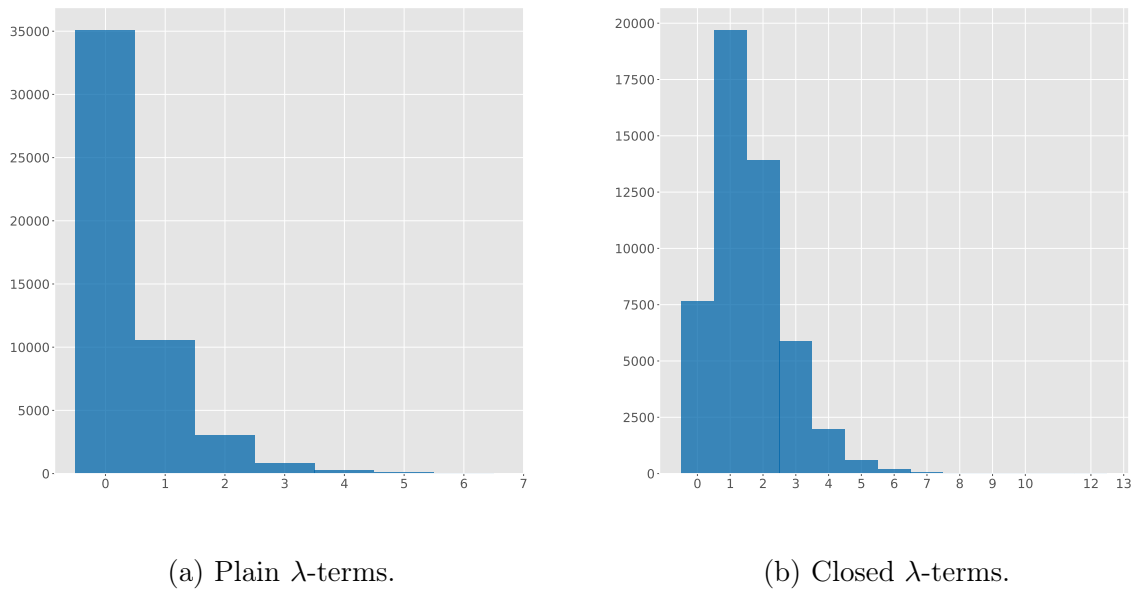
(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 12: Distribution histograms for head abstractions in plain and closed $\lambda$-terms.

Table 2: The average, variance and standard deviation corresponding to head abstractions in plain, closed and $h$-shallow $\lambda$-terms within the obtained data sets.

|  | Average | Variance | Std. dev. |
|---|---|---|---|
| plain | 0.42202 | 0.59632 | 0.77222 |
| closed | 1.55712 | 1.30818 | 1.14376 |
| $h$-shallow | 1.56644 | 1.28803 | 1.13491 |

### 6.1.2   Free variables in plain lambda terms

When viewed as programs of an abstract programming language, variables in $\lambda$-terms become formal arguments bound to abstractions introducing them in respective name scopes (namespaces). In this perspective, free variable occurrences correspond to expressions available in the global namespace such as for instance predefined operators or constants, see e.g. [PJ87].

Figure 13 depicts the distribution histogram of the obtained data set for plain terms (recall that closed $\lambda$-terms contain, by definition, no free variable occurrences). The corresponding numerical approximations of averages, variances and standard deviations are listed in Table 3.

*Remark* 6.2. The empirical distribution of free variables in plain terms resembles closely a geometric law $\mathbb{P}(\text{GEOM} = k) = (1 - p)^k p$ with parameter $p \approx 0.14815$. Note however that the observed distribution is not geometric; its global maximum is not attained at the parameter value corresponding to the lack of free variables, see Figure 13. Remarkably, the observed value $p$ is a quite good estimate for the probability that a sufficiently
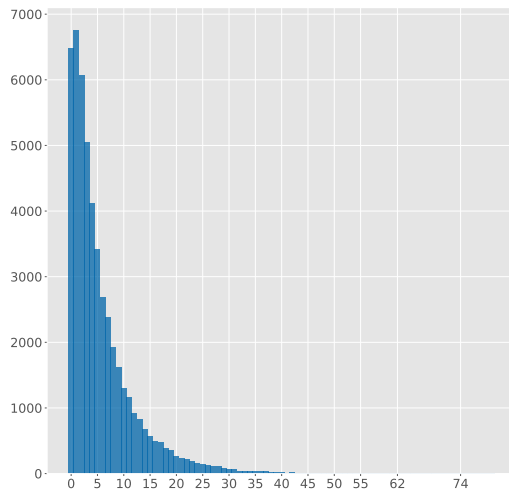
Figure 13: Distribution histogram for free variables in plain $\lambda$-terms.

Table 3: The average, variance and standard deviation corresponding to free variables in plain $\lambda$-terms within the obtained data set.

|  | Average | Variance | Std. dev. |
|---|---|---|---|
| plain | 5.74976 | 43.40258 | 6.58806 |

large random plain $\lambda$-term is closed. Using the numerical estimate for the multiplicative constant $C$ in the asymptotic growth rate $C \cdot n^{-3/2} \rho^{-n}$ corresponding to closed $\lambda$-terms [GG16, Section 6.1] one can find that this probability is in fact close to 0.12840.

### 6.1.3 Leftmost-outermost redex search

The next parameter we evaluate is the cost of finding the leftmost-outermost $\beta$-redex in plain and closed $\lambda$-terms, see Section 3.4. Figure 14 depicts the distribution histogram of the obtained data sets. Corresponding numerical approximations of averages, variances and standard deviations are given in Table 4.

Table 4: The average, variance and standard deviation corresponding the search cost of the leftmost-outermost redex in plain, closed and $h$-shallow $\lambda$-terms within the obtained data sets.

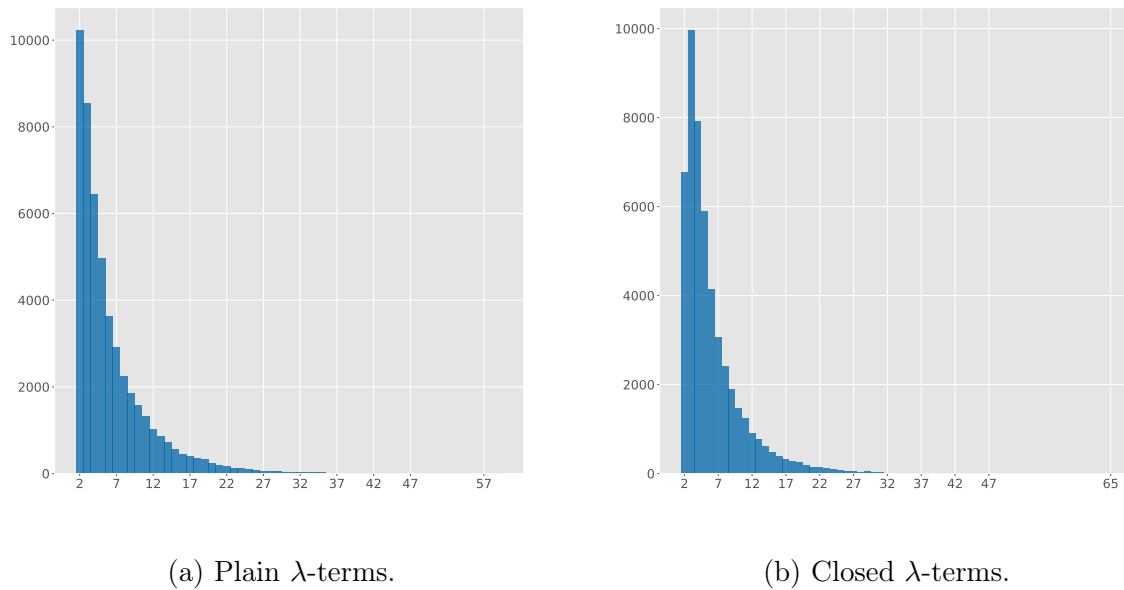|  | Average | Variance | Std. dev. |
|---|---|---|---|
| plain | 6.22118 | 26.72478 | 5.1696 |
| closed | 6.06994 | 21.78197 | 4.66712 |
| $h$-shallow | 6.12196 | 22.40393 | 4.73328 |

(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 14: Distribution histograms for leftmost-outermost redex search time in plain and closed $\lambda$-terms.

*Remark* 6.3. The empirical histograms corresponding to plain and closed terms do not follow the same law, similarly to the case of head abstractions, see Section 6.1.1. However, here the observed variance is in both cases significantly larger. Note that all discovered averages are remarkably close to each other.

### 6.1.4 Open subterms

In the current subsection we are interested in the degree to which variables connect various subterms in random $\lambda$-terms. In other words, in the proportion of open subterms in a random term. Let us note that such a parameter provides some insight in the extend to which subterms, viewed as components of the entire term (i.e. functional program) depend syntactically on each other. In this perspective, terms consisting of many closed subterms correspond to programs with equally many independent subprograms. On the other hand, $\lambda$-terms with few open subterms represent computations in which various parts of the program depend on each other through common variable usage. Since each index $\underline{n}$ consists of $n$ open proper subterms which do not contribute to the intended degree of functional dependence (variables in functional programs are atomic expressions) for our current considerations we assume that $\underline{n}$ is itself atomic, i.e. does not have proper subterms. Consequently, the total number of subterms of a given $\lambda$-term $T$ becomes equal to the number of indices, applications and abstractions occurring in $T$, without accounting for successors.

The resulting distribution histograms for plain and closed $\lambda$-terms are depicted in Figure 15. The $x$-axis denotes the (normalised with respect to the total number of subterms

as defined above) number of open subterms whereas the $y$-axis corresponds to the number of terms attaining the corresponding value. Respective averages, variances and standard deviations are listed in Table 5.
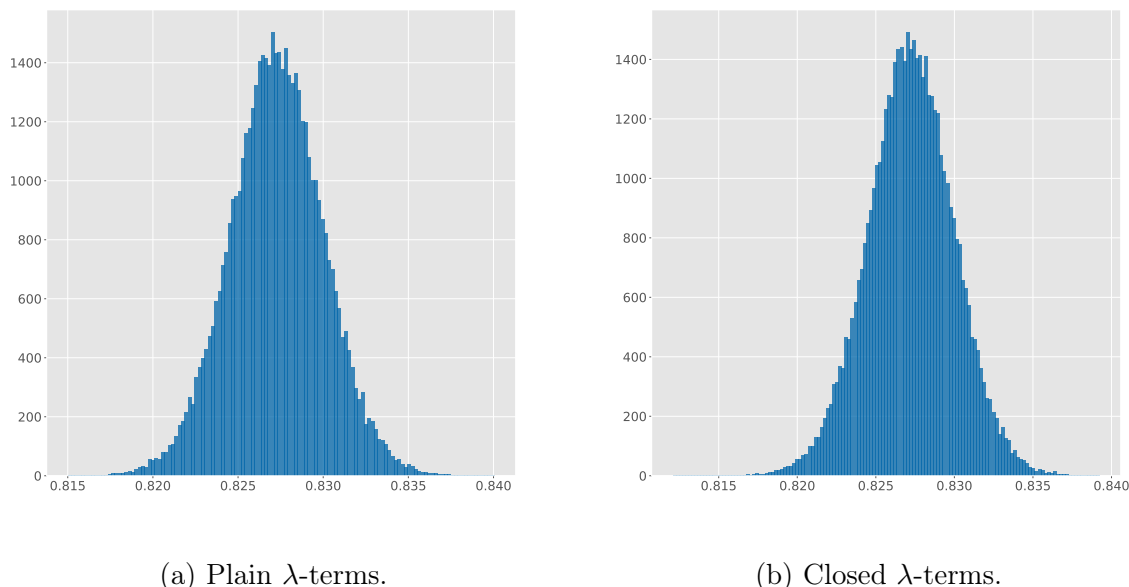


(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 15: Distribution histograms for open subterms in plain and closed $\lambda$-terms.

Table 5: The average, variance and standard deviation corresponding the proportion of open subterms in plain, closed and $h$-shallow $\lambda$-terms within the obtained data sets.

|  | Average | Variance | Std. dev. |
|---|---|---|---|
| plain | 0.82728 | 0.00001 | 0.00276 |
| closed | 0.82723 | 0.00001 | 0.00275 |
| $h$-shallow | 0.82724 | 0.00001 | 0.00276 |

*Remark* 6.4. Both empirical distributions resemble Gaussian laws with respective means and variances, see Table 5. The distributions seem to be concentrated around their means whereas the observed variances and standard deviations are strikingly modest and (almost) identical, though still positive. The high expectations in the order of 0.8272 suggest that the vast majority of subterms in a random $\lambda$-term is open, independently of whether it is itself open or closed.

### 6.1.5 Binding abstractions

The next parameter we consider is the proportion of abstractions binding variables. Figure 16 depicts the empirical distribution histograms for plain and closed $\lambda$-terms, respect-

ively. The $x$-axis denotes the proportion of binding abstractions among all abstractions. Corresponding averages, variances and standard deviations are presented in Table 6.
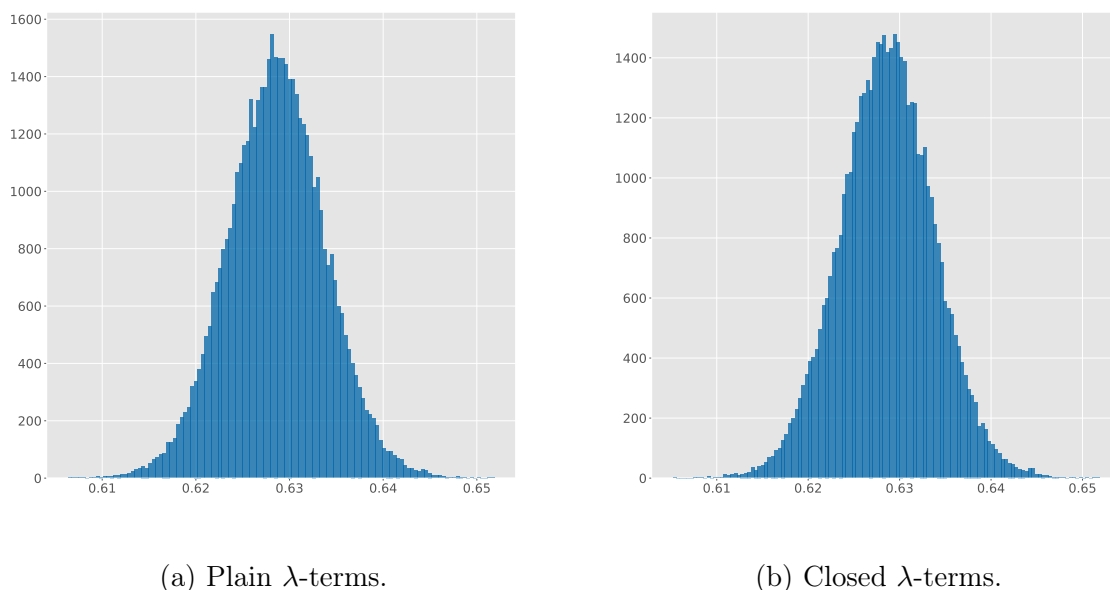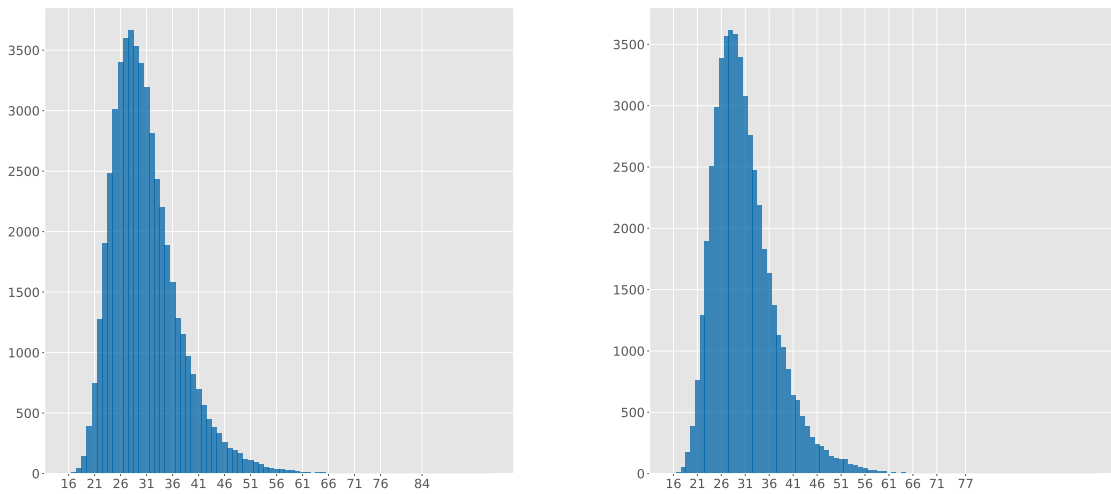


(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 16: Distribution histograms for binding abstractions in plain and closed $\lambda$-terms.

Table 6: The average, variance and standard deviation corresponding to the proportion of binding abstractions in plain, closed and $h$-shallow $\lambda$-terms within the obtained data sets.

|  | Average | Variance | Std. dev. |
| --- | --- | --- | --- |
| plain | 0.62862 | 0.00003 | 0.00521 |
| closed | 0.62863 | 0.00001 | 0.00518 |
| $h$-shallow | 0.62862 | 0.00001 | 0.00512 |

*Remark* 6.5. Like in the case of open subterms, the distribution of binding abstractions in plain and closed $\lambda$-terms, respectively, mirrors a Gaussian law. Corresponding variances and standard deviations, though positive, are minuscule. The listed averages of order 0.6286 suggest that most abstractions are binding, both in plain and closed terms.

### 6.1.6 Maximal number of variables bound to a single abstraction

In the current subsection we turn to extremal statistics related to binding abstractions. Specifically, we investigate the maximal number of variables bound to a single abstraction in random $\lambda$-terms. Figure 17 illustrates the related distribution histogram for plain and closed $\lambda$-terms, respectively. Numerical approximations of averages, variances and standard deviations are given in Table 7.

(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 17: Distribution histograms for the maximal number of variables bound to a single abstraction in plain and closed $\lambda$-terms.

Table 7: The average, variance and standard deviation corresponding the maximal number of variables bound to a single abstraction in plain, closed and $h$-shallow $\lambda$-terms within the obtained data sets.

|  | Average | Variance | Std. dev. |
|---|---|---|---|
| plain | 30.83028 | 43.08992 | 6.56429 |
| closed | 30.8323 | 43.12418 | 6.5669 |
| $h$-shallow | 30.87194 | 43.46242 | 6.5926 |

*Remark* 6.6. Obtained distributions and moments are reminiscent of a double-exponential distribution related to extreme parameters in various combinatorial structures, see [FS09, p. 311]. Notably, $\log_{1/\rho} n$ for $n \in [20,000; 50,000]$ lies in the interval $[8.1259, 8.8777]$ whereas as the same time $\sqrt{n} \geqslant 141.421$. Consequently, the obtained average suggests a similar $C \cdot \log_{1/\rho} n$ type of behaviour occurring, for instance, in the analysis of longest runs in random words, see [FS09, Example V.4]. Nonetheless, unlike the limit distribution of longest runs, the discovered distribution does not suggest a limit concentration due to the significant values of the observed variances.

### 6.1.7 $m$-openness of plain lambda terms

The central notion of $m$-openness plays an important rôle in the analysis of closed $\lambda$-terms. Given the complete analysis of $m$-open terms [BGG18] and efficient techniques allowing to obtain numerical estimates for the relative asymptotic density of $m$-open terms in

the set of all plain $\lambda$-terms [GG16] it is possible to approximate the limit distribution of the $m$-openness parameter in plain terms. However, in order to avoid the laborious computations involved in this approach, we offer a simple Monte Carlo approximation scheme. Figure 18 depicts the distribution histogram of the obtained data set for plain $\lambda$-terms. The corresponding average, variance and standard deviation are listed in Table 8.
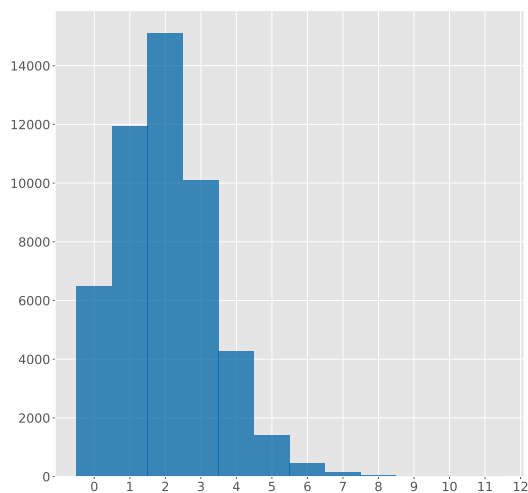


Figure 18: Distribution histogram for $m$-openness of plain $\lambda$-terms.

Table 8: The average, variance and standard deviation corresponding to the $m$-openness of plain $\lambda$-terms within the obtained data set.

|       | Average | Variance | Std. dev. |
|-------|---------|----------|-----------|
| plain | 2.01856 | 1.82538  | 1.35106   |

### 6.1.8  Generalised $m$-openness of plain and closed terms

Though the notion of $m$-openness is defined only for non-negative values of $m$, we propose a natural generalisation to all integers in the following manner. We say that a closed $\lambda$-term $T$ is $m$-closed[2] for $m \geqslant 0$ if it is possible to discard $m$ head abstractions of $T$ and still retain a closed $\lambda$-term. In other words, a closed term is $m$-closed if its top $m$ head abstractions are non-binding.

Note that this new parameter provides a degree of term closeness; the higher the factor $m$ of an $m$-closed $\lambda$-term, the more closed it is. Like in the case of $m$-open $\lambda$-terms, closed $\lambda$-terms are 0-closed. Moreover, if a term is $(m + 1)$-closed, then it is also $m$-closed.

---

[2]Though this notion extends $m$-openness to negative values of $m$, we prefer to avoid the term $m$-open while referring to closed terms and instead use the term $m$-close. Hence, a term is $m$-closed if it is, intuitively, $(-m)$-open.

Figure 19 depicts the distribution histogram of the obtained data sets for plain and closed $\lambda$-terms, respectively. The $x$-axis denotes the generalised $m$-openness factor whereas the $y$-axis corresponds to the number of terms attaining the corresponding cost value. Numerical approximations of averages, variances and standard deviations are listed in Table 9.
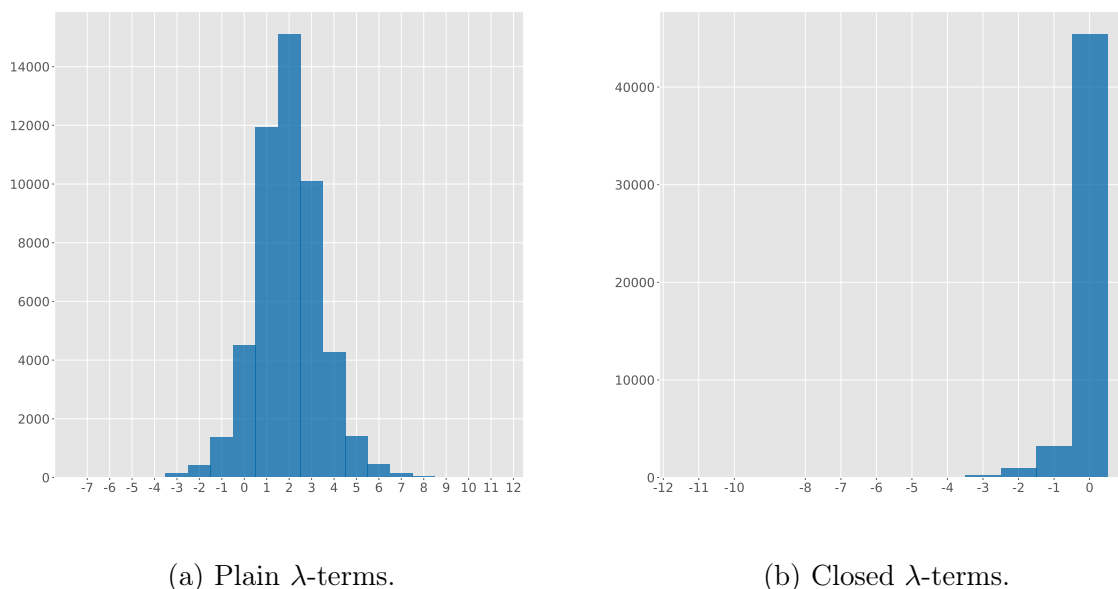


(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 19: Distribution histograms for the generalised $m$-openness in plain and closed $\lambda$-terms.

Table 9: The average, variance and standard deviation corresponding the generalised $m$-openness factor in plain, closed and $h$-shallow $\lambda$-terms within the obtained data sets.

|  | Average | Variance | Std. dev. |
|---|---|---|---|
| plain | 1.96196 | 2.15235 | 1.46709 |
| closed | $-0.43814$ | 0.61569 | 0.78466 |
| $h$-shallow | $-0.44054$ | 0.62834 | 0.79268 |

*Remark* 6.7. The empirical distribution corresponding generalised $m$-openness of closed $\lambda$-terms suggests that the vast majority of closed terms are 0-closed but not 1-closed. Given the related distribution of head abstractions, see Figure 12, such a result suggests that if a closed term has a leading head abstraction, it is more likely that it is in fact binding.

Let us note that all of the presented empirical histograms exhibit a close correspondence between parameters of closed $\lambda$-terms and their corresponding equivalents in $h$-shallow terms. Such a result should not be surprising given the exponential convergence

speed at which $h$-shallow $\lambda$-terms of size $N$ tend to closed $\lambda$-terms of size $N$ as $h \to \infty$, see Section 5, cf. [BGG18]. Consequently, virtually all presented histograms for closed $\lambda$-terms are also correct histograms for $h$-shallow terms.

## 6.2   Empirical speed of convergence from generating functions

The probability generating function $p_n(u)$ of a random variable $X_n$ corresponding to a certain parameter inside a random $\lambda$-term of size $n$, plain or closed, can be obtained by evaluating the bivariate generating function $L(z, u)$ and taking the respective ratio of coefficients, see Appendix A.2:

$$p_n(u) = \frac{[z^n]L(z, u)}{[z^n]L(z, 1)}. \tag{6.1}$$

Remarkably, for statistics whose limit laws are discrete distributions, the convergence is already evident at $n = 20$. We empirically evaluate the coefficients of generating functions whenever the associated functional equations are available.

For closed $\lambda$-terms, the corresponding systems of functional equations are infinite and hence cannot be directly evaluated. Fortunately, truncating the infinite system at the $k$th level and replacing the $k$th function $L_k(z, u)$ by its infinite counterpart $L_\infty(z, u)$ yields an exponentially small difference in the dominant term coefficients, see Section 5. In what follows, we truncate the systems at height $k = 15$.

### 6.2.1   Head abstractions

We return to head abstractions in plain and closed $\lambda$-terms. The distribution of this parameter for large term sizes was discussed in Section 6.1.1. Here, we study this parameter for small term sizes. The equations for corresponding generating functions are presented in Sections 3.2 and 4.4.
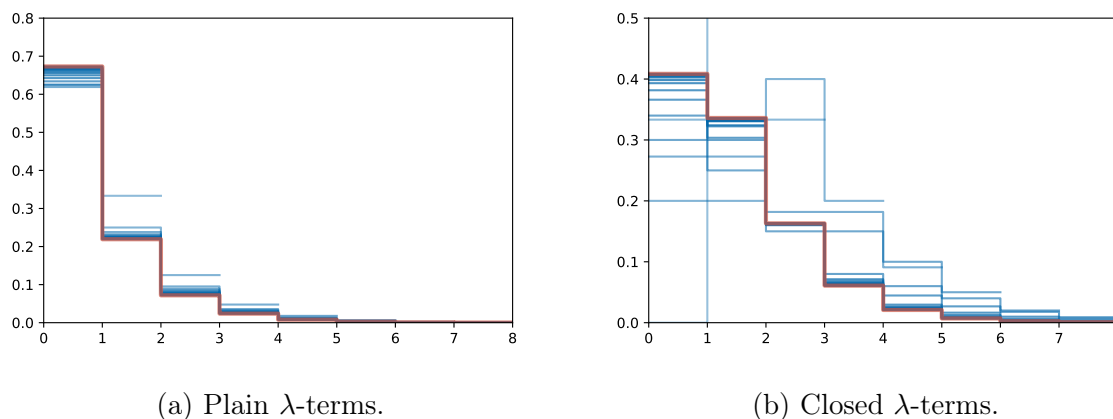


(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 20: Distribution histograms for the number of head abstractions in plain and closed $\lambda$-terms.

The convergence process for $n \in [1, 15]$ is depicted in Figure 20. Head abstractions in plain terms admit a geometric distribution while the number of head abstractions in closed lambda terms converges to a certain computable, discrete distribution.

### 6.2.2 Leftmost-outermost redex search

As already mentioned in Section 6.1, the distribution associated with the cost of finding the leftmost-outermost $\beta$-redex in plain and closed $\lambda$-terms tends to a discrete limiting distribution. The functional systems for bivariate generating functions can be found in Sections 3.4 and 4.6.

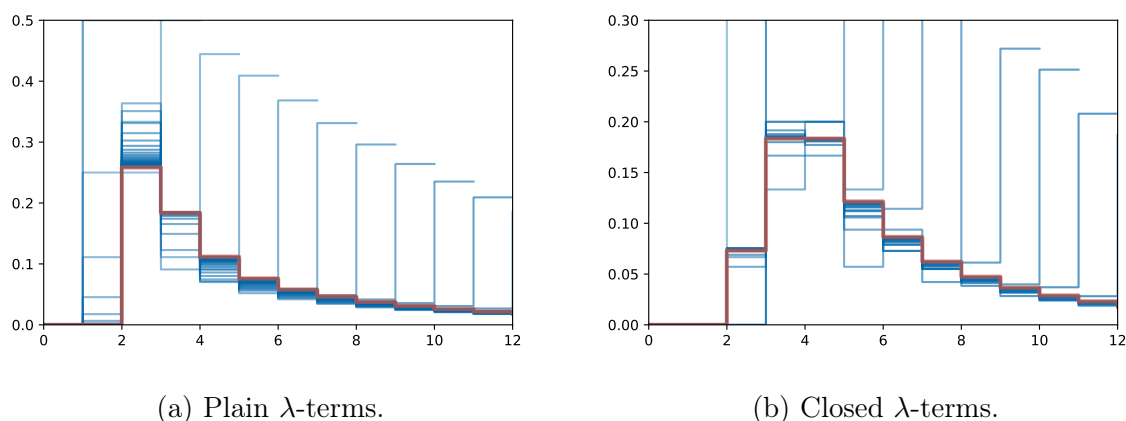

(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 21: Distribution histograms associated with the cost of redex search time in plain and closed $\lambda$-terms.

We depict the convergence process in Figure 21. For both plain and closed $\lambda$-terms, a remarkable detail can be observed. For $\lambda$-terms of small size, the distribution has a peak at the value corresponding to the term size. Note that such a phenomenon is related to $\beta$-normal forms, witnessing the worst-case time complexity of the traversal algorithm. For larger terms, the proportion of $\beta$-normal forms in all lambda terms becomes exponentially negligible.

### 6.2.3 Free variables and $m$-openness in plain terms

Next, we consider two parameters of plain $\lambda$-terms that require advanced marking techniques. The functional equations for the cases of free variables and $m$-openness are given in subsequent Sections 4.1 and 4.3.

The convergence process is depicted in Figure 22. Let us remark that the convergence process corresponding to the number of free variables proceeds more slowly than all the other statistics that we consider in the current paper.
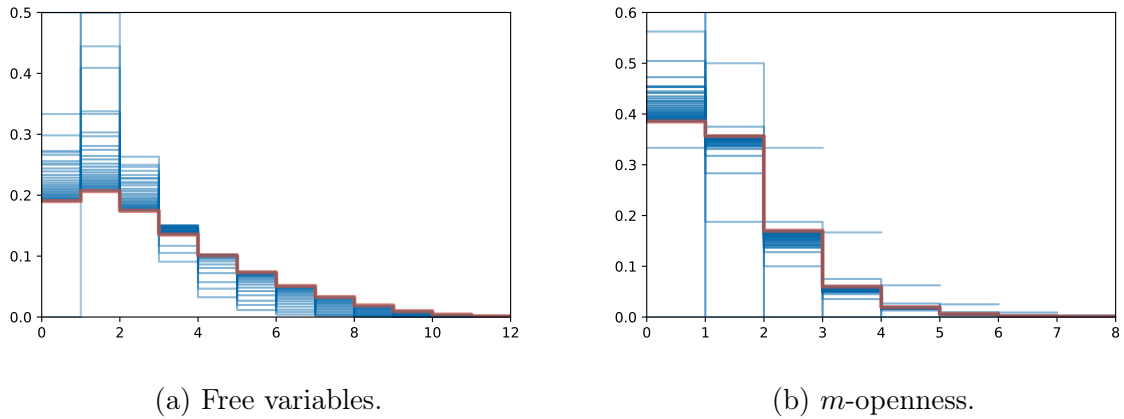
(a) Free variables.

(b) $m$-openness.

Figure 22: Distribution histograms for the number of free variables and $m$-openness in plain $\lambda$-terms.

### 6.2.4 De Bruijn index profile

The final pair of plots concerns the profile of variables or, in other words, the profile of de Bruin indices in random plain and closed $\lambda$-terms. In Sections 3.3 and 4.5 we show that both parameters tend to geometric limiting distributions. Figure 23 depicts the probability distributions for lambda terms of the size in the interval $n \in [1, 15]$.
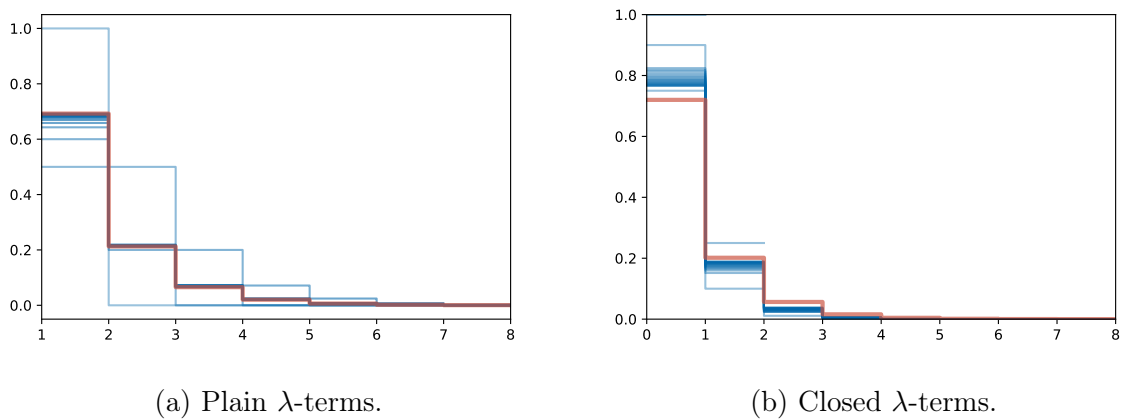


(a) Plain $\lambda$-terms.

(b) Closed $\lambda$-terms.

Figure 23: Distribution histograms for De Bruijn index profile in plain and closed $\lambda$-terms.

## 7    Conclusions

We investigated the statistical properties of $\lambda$-terms in the De Bruijn notation, providing some insight into their internal, quantitative characteristics. In essence, our results suggest that random $\lambda$-terms, both plain and closed, exhibit typical traits of various tree-

like structures. For instance, the distribution of sub-patterns inside random $\lambda$-terms is typically Gaussian whereas their corresponding finding time tends to a discrete limit distribution. Similarly, the height profile of random terms follows the Rayleigh distribution.

Nonetheless, some of the investigated parameters do not have analogues in other tree-like structures such as, for instance, $m$-openness or the number of free variables. In both cases we have established a discrete limiting distribution. Remarkably, we have not discovered substantially different statistical traits of plain and closed lambda terms; however, we found that among the statistics with discrete limiting distributions, the distribution in closed terms is often different from the associated distribution in plain terms.

Given the general algorithmic frameworks meant for the construction of effective exact- and approximate-size combinatorial samplers, such as Boltzmann samplers [DFLS04] and the recursive method [NW78, FZC94], presented parameter specifications provide a novel source of effective sampling methods for $\lambda$-terms with additional control over their specific combinatorial parameters. In this context, most parameter specifications associated with plain terms are finite and hence also readily applicable. Remaining, infinite specifications are a bit more involved. Nonetheless, an appropriate truncation of the specification followed by a final rejection phase allows to discard inadmissible terms. The exponential convergence of intermediate, truncated specifications rationalises such an approach and provides effective samplers for corresponding $\lambda$-terms. Let us also remark that so generated terms do not have to be restricted to their natural parameter distributions. It is possible to gain an additional control over the expected parameter values using a dedicated tuning procedure which distorts the intrinsic parameter distribution and hence allows for a skewed parameter distribution sampler construction [BBD18]. Consequently, the presented analysis provides means for an effective construction of various samplers for (plain or closed) $\lambda$-terms with additional control over their parameter distribution.

Few more aspects of the parameter analysis of $\lambda$-terms remain untouched. For instance, our empirical data suggests that the distribution of binding abstractions, both in plain and closed terms, is Gaussian. The same holds for the number of open subterms. Alas, the theoretical verification of our empirical findings is left open. Moreover, we have not investigated other, well-known parameters. Let us mention, for instance, the height distribution of random closed terms, or the distribution of certain extremal statistics, such as the maximal De Bruijn index, longest lambda run, or the maximal number of variables bound to a single abstraction. We conjecture that the behaviour of these parameters in closed terms does not substantially differ from the behaviour of respective parameters in plain terms. Finally, the question of generalised $m$-openness also has not been settled and the corresponding techniques are still to be developed.

Arguably, from the viewpoint of analytic combinatorics, our novel result complements the existing result of Drmota, Gittenberger and Morgenbesser [DGM12] on infinite systems. In our formulation, the infinite system is not required to be strongly connected. Consequently, we conjecture that the properties of the Jacobian operator of the infinite system are not sufficient to deduce the result, in contrast with the mentioned paper. In other words, we conjecture that the condition of exponential convergence is essential.

Moreover, we discovered that it is possible to rewrite the infinite system defining closed λ-terms as a strongly connected system, however the Jacobian of the resulting system is not compact. Alas, the framework [DGM12] is not applicable.

Consequently, we finish the paper with an even more general question: what can be stated about the properties of infinite systems which are either not strongly connected or have a non-compact Jacobian operator?

# References

[Bar84]  Henk P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, volume 103. North Holland, revised edition, 1984.

[BBD18]  Maciej Bendkowski, Olivier Bodini, and Sergey Dovgal. *Polynomial tuning of multiparametric combinatorial samplers* Proceedings of the Fifteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO), pages 92–106. Society for Industrial and Applied Mathematics (SIAM), 2018.

[BBY10]  Jason P. Bell, Stanley N. Burris, and Karen A. Yeats. Characteristic points of recursive systems. *The Electronic Journal of Combinatorics*, 17(1):121, 2010.

[BC10]  Yves Bertot and Pierre Castran. *Interactive Theorem Proving and Program Development: Coq'Art The Calculus of Inductive Constructions*. Springer Publishing Company, Incorporated, 1st edition, 2010.

[Ben17]  Maciej Bendkowski. *Quantitative aspects and generation of random lambda and combinatory logic terms*. PhD thesis, Jagiellonian University, Kraków, Poland, May 2017.

[BFSS01]  Cyril Banderier, Philippe Flajolet, Gilles Schaeffer, and Michele Soria. Random maps, coalescing saddles, singularity analysis, and Airy phenomena. *Random Structures & Algorithms*, 19(3-4):194–246, 2001.

[BG14]  Olivier Bodini and Bernhard Gittenberger. *On the asymptotic number of BCK(2)-terms.* Proceedings of the Eleventh Workshop on Analytic Algorithmics and Combinatorics (ANALCO), pages 25–39. Society for Industrial and Applied Mathematics (SIAM), 2014.

[BGG11]  Olivier Bodini, Danièle Gardy, and Bernhard Gittenberger. *Lambda terms of bounded unary height*. Proceedings of the Eighth Workshop on Analytic Algorithmics and Combinatorics (ANALCO), pages 23-32. Society for Industrial and Applied Mathematics (SIAM), 2011.

[BGG18]  Olivier Bodini, Bernhard Gittenberger, and Zbigniew Gołębiewski. Enumerating lambda terms by weighted length of their De Druijn representation. *Discrete Applied Mathematics*, 239:45–61, 2018.

[BGGJ13]  Olivier Bodini, Danièle Gardy, Bernhard Gittenberger, and Alice Jacquot. Enumeration of generalized BCI lambda-terms. *Electronic Journal of Combinatorics*, 20(4), 2013.

[BGJ13]     Olivier Bodini, Danièle Gardy, and Alice Jacquot. Asymptotics and random sampling for BCI and BCK lambda terms. *Theoretical Computer Science*, 502:227 – 238, 2013. Generation of Combinatorial Structures.

[BGLZ16]   Maciej Bendkowski, Katarzyna Grygiel, Pierre Lescanne, and Marek Zaionc. A natural counting of lambda terms. In Rusins Martins Freivalds, Gregor Engels, and Barbara Catania, editors, *SOFSEM 2016: Theory and Practice of Computer Science - 42nd International Conference on Current Trends in Theory and Practice of Computer Science, Harrachov, Czech Republic, January 23-28, 2016, Proceedings*, volume 9587 of *Lecture Notes in Computer Science*, pages 183–194. Springer, 2016.

[BGLZ17]   Maciej Bendkowski, Katarzyna Grygiel, Pierre Lescanne, and Marek Zaionc. Combinatorics of $\lambda$-terms: a natural approach. *Journal of Logic and Computation*, 27(8):2611–2630, 2017.

[BGR15]     Olivier Bodini, Antoine Genitrini, and Nicolas Rolin. Pointed versus singular Boltzmann samplers: a comparative analysis. *Pure Mathematics and Applications*, 25(2):115–131, 2015.

[BR83]      Edward A. Bender and L. Bruce Richmond. Central and local limit theorems applied to asymptotic enumeration II: Multivariate generating functions. *Journal of Combinatorial Theory, Series A*, 34(3):255 – 265, 1983.

[BT18]      Olivier Bodini and Paul Tarau. On uniquely closable and uniquely typable skeletons of lambda terms. In Fabio Fioravanti and John P. Gallagher, editors, *Logic-Based Program Synthesis and Transformation*, pages 252–268, Cham, 2018. Springer International Publishing.

[CH00]      Koen Claessen and John Hughes. Quickcheck: A lightweight tool for random testing of Haskell programs. In *Proceedings of the Fifth ACM SIGPLAN International Conference on Functional Programming*, pages 268–279. ACM, 2000.

[CH09]      Felice Cardone and J. Roger Hindley. Lambda-calculus and combinators in the 20th century. In *Logic from Russell to Church*, volume 5 of *Handbook of the History of Logic*, pages 723–817. Elsevier, 2009.

[dB72]      Nicolaas G. de Bruijn. Lambda calculus notation with nameless dummies, a tool for automatic formula manipulation, with application to the Church-Rosser theorem. *Indagationes Mathematicae (Proceedings)*, 75(5):381–392, 1972.

[DFLS04]    Philippe Duchon, Philippe Flajolet, Guy Louchard, and Gilles Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability and Computing*, 13:2004, 2004.

[DGK$^+$13]  René David, Katarzyna Grygiel, Jakub Kozik, Christophe Raffalli, Guillaume Theyssier, and Marek Zaionc. Asymptotically almost all $\lambda$-terms are strongly normalizing. *Logical Methods in Computer Science*, 9:1–30, 2013.

[DGM12]   Michael Drmota, Bernhard Gittenberger, and Johannes F. Morgenbesser. Infinite Systems of Functional Equations and Gaussian Limiting Distributions. In *23rd International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'12)*, volume DMTCS Proceedings vol. AQ, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12) of *DMTCS Proceedings*, pages 453–478, Montreal, Canada, 2012. Discrete Mathematics and Theoretical Computer Science.

[Drm97]   Michael Drmota. Systems of functional equations. *Random Structures and Algorithms*, 10(1-2):103–124, 1997.

[Drm09]   Michael Drmota. *Random Trees: An Interplay between Combinatorics and Probability*. Springer Science & Business Media, 2009.

[FS09]    Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 1 edition, 2009.

[FZC94]   Philippe Flajolet, Paul Zimmermann, and Bernard Van Cutsem. A calculus for the random generation of labelled combinatorial structures. *Theoretical Computer Science*, 132(1):1–35, 1994.

[GG16]    Bernhard Gittenberger and Zbigniew Gołębiewski. On the Number of Lambda Terms With Prescribed Size of Their De Bruijn Representation. In Nicolas Ollinger and Heribert Vollmer, editors, *33rd Symposium on Theoretical Aspects of Computer Science (STACS 2016)*, volume 47 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 40:1–40:13. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2016.

[GL13]    Katarzyna Grygiel and Pierre Lescanne. Counting and generating lambda terms. *Journal of Functional Programming*, 23(5):594-628, 2013.

[GL15]    Katarzyna Grygiel and Pierre Lescanne. Counting and generating terms in the binary lambda calculus. *Journal of Functional Programming*, 25:e24, 2015.

[Heu07]   Clemens Heuberger. Hwang's quasi-power-theorem in dimension two. *Quaestiones Mathematicae*, 30(4):507–512, 2007.

[HK16]    Clemens Heuberger and Sara Kropf. On the higher dimensional quasi-power theorem and a Berry-Esseen inequality. In *27th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms Kraków, Poland, July 4-8, 2016*, 2016.

[Hwa98]   Hsien-Kuei Hwang. On convergence rates in the central limit theorems for combinatorial structures. *European Journal of Combinatorics*, 19(3):329–343, 1998.

[Joy81]   André Joyal. Une théorie combinatoire des séries formelles. *Advances in mathematics*, 42(1):1–82, 1981.

[Lal93]   Steven P. Lalley. Finite range random walk on free groups and homogeneous trees. *The Annals of Probability*, pages 2087–2130, 1993.

[Les13]     Pierre Lescanne. On counting untyped lambda terms. *Theoretical Computer Science*, 474:80–97, 2013.

[NW78]      Albert Nijenhuis and Herbert S. Wilf. *Combinatorial Algorithms*. Academic Press, 2 edition, 1978.

[Pal12]     Michał H. Pałka. *Random Structured Test Data Generation for Black-Box Testing*. PhD thesis, Chalmers University of Technology, 2012.

[PJ87]      Simon L. Peyton Jones. *The Implementation of Functional Programming Languages*. Prentice-Hall, Inc., 1987.

[PSS12]     Carine Pivoteau, Bruno Salvy, and Michele Soria. Algorithms for combinatorial structures: Well-founded systems and Newton iterations. *Journal of Combinatorial Theory, Series A*, 119(8):1711–1773, 2012.

[Tro06]     John Tromp. Binary lambda calculus and combinatory logic. In *Kolmogorov Complexity and Applications*, volume 06051. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2006.

[Wan05]     Jue Wang. Generating random lambda calculus terms. Technical report, Boston University, 2005.

[Wil06]     Herbert S. Wilf. *Generatingfunctionology*. A. K. Peters, Ltd., 2006.

[Woo97]     Alan R. Woods. Coloring rules for finite trees, and probabilities of monadic second order sentences. *Random structures and algorithms*, 10(4):453–485, 1997.

# A    Standard tools

## A.1    Analytic tools

In the following subsection we present standard, analytic combinatorics tools following the exposition of Flajolet and Sedgewick [FS09]. We also assume conventional notation corresponding to generating functions, their coefficients and asymptotic expansions. We refer the reader to [FS09, Wil06] for a detailed exposition.

For our purposes, combinatorial parameter analysis outlines as follows:

- Let $a_{n,k}$ denote the number of plain (closed) lambda terms of size $n$ for which the investigated parameter takes value $k$. Note that we do not assume that the numbers $a_{n,k}$ are *a priori* known. With such a two-dimensional sequence of numbers we associate a bivariate generating function

$$A(z, u) := \sum_{n,k \geqslant 0} a_{n,k} z^n u^k; \tag{A.1}$$

In order to simultaneously study several different parameters of interest, we introduce *multivariate generating functions* in form of

$$A(z, \boldsymbol{u}) = \sum_{n,\boldsymbol{k} \geqslant 0} a_{n,\boldsymbol{k}} z^n \boldsymbol{u}^{\boldsymbol{k}} \tag{A.2}$$

where $\boldsymbol{u} = (u_1, \ldots, u_d)$ is a $d$-dimensional variable, $\boldsymbol{k} = (k_1, \ldots, k_d)$ is a $d$-dimensional index satisfying $k_i \geqslant 0$, $\boldsymbol{u^k} := u_1^{k_1} u_2^{k_2} \cdots u_d^{k_d}$, and $a_{n,\boldsymbol{k}}$ denotes the number of plain (closed) lambda terms of size $n$ for which the investigated parameter values equal $k_1, k_2, \ldots, k_d$, respectively;

- Considered combinatorial parameters (patterns) inside plain or closed $\lambda$-terms are described in terms of admissible combinatorial specifications (sometimes infinite, as in the case of closed terms);

- So obtained specifications are then converted into systems of equations involving multivariate generating functions where additional variables $\boldsymbol{u} = (u_1, u_2, \ldots, u_d)$ *mark* associated combinatorial patterns;

- In the case of plain lambda terms, the resulting systems of equations are solved, usually approximately, in terms of standard complex-valued functions like $f(z) = \sqrt{1 - z}$. The coefficients of associated generating functions depend on the marking variables $\boldsymbol{u} = (u_1, u_2, \ldots, u_d)$. In the case of closed lambda terms, novel tools developed in Section 5 are applied;

- Finally, an application of Flajolet and Odlyzko's transfer theorem provides access to probability generating functions of the limiting probability distributions. In consequence, properties of investigated combinatorial parameters become readily available.

### A.1.1 Asymptotic expansions

In order to access the asymptotic form of the coefficients of corresponding generating functions, we view them as functions analytic at the origin of the complex plane and examine their singularities, in particular so-called *dominant singularities* of smallest possible absolute value. Typically, at this point, an analytic continuation of the formal power series outside its circle of convergence is required. The following, usual domain in which the function is considered, is called a *delta-domain*.

**Proposition A.1** (Transfer theorem [FS09, Section VI.3]). Suppose that $f(z/\rho)$ is a function analytic in the so-called *delta-domain* $\Delta(R, \phi)$ for some $R > 1$ and $0 < \phi < \frac{\pi}{2}$, where

$$\Delta(R, \phi) = \{\zeta \colon |\zeta| < R, \ \zeta \neq 1, \ \arg(\zeta - 1) > \phi\}. \tag{A.3}$$

Suppose that as $z \to \rho$, for $z/\rho \in \Delta(R, \phi)$, it holds

$$f(z) = h(z) - g(z) \left(1 - \frac{z}{\rho}\right)^{-\alpha} + O\left(\left|1 - \frac{z}{\rho}\right|^{-\beta}\right) \tag{A.4}$$

where $\alpha, \beta \in \mathbb{C} \setminus \mathbb{Z}_{\leqslant 0}$, and $h(z)$ and $g(z)$ are functions analytic in $|z| < R$.

Then, as $n \to \infty$, the coefficients $[z^n]f(z)$ admit an asymptotic approximation in form of

$$[z^n]f(z) \sim g(\rho) \cdot \rho^{-n} \cdot \frac{n^{\alpha-1}}{\Gamma(\alpha)} + O\left(\rho^{-n}n^{\beta-1}\right) \tag{A.5}$$

where $\Gamma \colon \mathbb{C} \setminus \mathbb{Z}_{\leqslant 0} \to \mathbb{C}$ is the Gamma function defined as $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$.

In particular, if for $z/\rho \in \Delta(R, \phi)$, as $z \to \rho$, we have

$$f_1(z) \sim h_1(z) - g_1(z)\sqrt{1 - \frac{z}{\rho}} \quad \text{and} \quad f_2(z) \sim \frac{g_2(z)}{\sqrt{1 - \frac{z}{\rho}}}, \tag{A.6}$$

then we obtain, as $n \to \infty$, the following coefficient expansions:

$$[z^n]f_1(z) \sim \frac{g_1(\rho)\rho^{-n}}{2\sqrt{\pi}n^{3/2}} \quad \text{and} \quad [z^n]f_2(z) \sim \frac{g_2(\rho)\rho^{-n}}{\sqrt{\pi}n^{1/2}}. \tag{A.7}$$

**Proposition A.2** (Semi-large powers theorem, [FS09, Theorem IX.16], [BFSS01])**.** Suppose that $f(z/\rho)$ is a function analytic in delta-domain $\Delta(R, \phi)$, see Proposition A.1, for some $R > 1$, and $f(z)$ admits asymptotic expansion as $z \to \rho$ for $z/\rho$ staying in $\Delta(R, \phi)$:

$$f(z) \sim 1 - a\sqrt{1 - \frac{z}{\rho}}. \tag{A.8}$$

Then, for $x$ in any compact subinterval of $(0, +\infty)$ the coefficient standing by $z^n$ in $f(z)^k$ admits an asymptotic estimate

$$[z^n]f(z)^k \sim \frac{\rho^{-n}}{n}S(ax) \tag{A.9}$$

where $S(x)$ is the Rayleigh function satisfying

$$S(x) = \frac{xe^{-x^2/4}}{2\sqrt{\pi}} \quad \text{and} \quad x = \frac{k}{\sqrt{n}}. \tag{A.10}$$

### A.1.2 Algebraic systems

The following theorem, commonly known as the Drmota–Lalley–Woods theorem, is a fundamental result obtained independently by several authors [Drm97, Woo97, Lal93] in order to establish limit laws in various families of tree structures specified by context-free grammars. In our exposition, we reference Drmota's book [Drm09, Section 2.2.5], and the papers [Drm97, DGM12, PSS12, BBY10].

**Definition A.3.** Consider a polynomial system of equations

$$\boldsymbol{y} = \boldsymbol{\Phi}(z, \boldsymbol{y}, \boldsymbol{u}) \tag{A.11}$$

which is a vector notation for $(y_j = \Phi_j(z, y_1, \ldots, y_m, u_1, \ldots, u_d))$ with $j$ ranging over $1, \ldots, m$. Assume that $\boldsymbol{\Phi}(0, \boldsymbol{0}, \boldsymbol{0}) = \boldsymbol{0}$. Then,

- $\mathbf{\Phi}(z, \boldsymbol{y}, \boldsymbol{u})$ is said to be *non-linear* if at least one of its component polynomials $\Phi_j$ is non-linear in one of the formal variables $y_1, \ldots, y_m$;

- $\mathbf{\Phi}(z, \boldsymbol{y}, \boldsymbol{u})$ is said to be *algebraic positive* if all of its component polynomials $\Phi_j$ have non-negative coefficients;

- $\mathbf{\Phi}(z, \boldsymbol{y}, \boldsymbol{u})$ is said to be *algebraic proper* if it admits a unique formal power series solution to which the iteration

$$\boldsymbol{y}_0(z, \boldsymbol{u}) = \mathbf{0} \quad \text{and} \quad \boldsymbol{y}_{k+1}(z, \boldsymbol{u}) = \mathbf{\Phi}(z, \boldsymbol{y}_k(z, \boldsymbol{u}), \boldsymbol{u}) \tag{A.12}$$

considered in the metric space of formal power series with valuation, converges as $k \to \infty$, and the Jacobian matrix $\dfrac{\partial \mathbf{\Phi}}{\partial \boldsymbol{y}}$ is nilpotent at $(z, \boldsymbol{y}) = (0, \mathbf{0})$.

- $\mathbf{\Phi}(z, \boldsymbol{y}, \boldsymbol{u})$ is said to be *algebraic irreducible* if its dependency graph (i.e. a graph whose vertices are the integers $1, \ldots, m$ and there exists a directed edge $k \to j$ if $y_j$ figures in a monomial of $\Phi_k$) is strongly connected;

- $\mathbf{\Phi}(z, \boldsymbol{y}, \boldsymbol{u})$ is said to be *algebraic aperiodic* if each of its component solutions $y_j(z, \mathbf{1})$ for $j = 1, \ldots, m$ is aperiodic in the sense that the greatest common divisor of the pairwise differences of the set of exponent indices of $z$ within $y_j(z, \mathbf{1})$ is equal to 1.

*Remark* A.4. The notion of algebraic properness of systems, also referred to as *well-foundedness*, is extensively studied in [PSS12, Section 5]. As discussed in [PSS12, Joy81], the system has combinatorial meaning only if the Jacobian is nilpotent, i.e. if the recursive definition is well-defined and allows to inductively construct all the instances of combinatorial species.

Let us note that the condition $\mathbf{\Phi}(0, \mathbf{0}, \mathbf{0}) = \mathbf{0}$ is a technical assumption of the Drmota–Lalley–Woods theorem. Pivoteau, Salvy and Soria consider, *inter alia*, well-founded systems for which $\boldsymbol{y}(0, \mathbf{0}) \neq \mathbf{0}$. One possible characterisation of such systems is the condition that the limit of a suitable iterative approximation procedure yields the solution of the initial functional system.

The assumption that a system is *polynomial* can be replaced by a more general assumption that the functions are *analytic*, see [Drm97]. For a detailed and non-trivial study of the conditions regarding analytic functions and the configuration of the critical points in this more general case, see [BBY10].

**Proposition A.5** (Irreducible positive polynomial systems)**.** Let

$$\boldsymbol{y} = \mathbf{\Phi}(z, \boldsymbol{y}, \boldsymbol{u}) = \left( y_j = \Phi_j(z, y_1, \ldots, y_m, \boldsymbol{u}) \right), \quad j = 1, \ldots, m \tag{A.13}$$

be a non-linear polynomial system of equations which is algebraic positive, proper, and irreducible. Then there exists $\varepsilon > 0$ such that all component solutions $y_j(z, \boldsymbol{u})$ admit representation of the form

$$y_j(z, \boldsymbol{u}) = h_j\left( \sqrt{1 - \frac{z}{\rho(\boldsymbol{u})}}, \boldsymbol{u} \right) = \sum_{k \geqslant 0} c_{k,j}(\boldsymbol{u}) \left( 1 - \frac{z}{\rho(\boldsymbol{u})} \right)^{k/2} \tag{A.14}$$

for $\boldsymbol{u}$ in a neighbourhood of $\boldsymbol{1}$, $|z - \rho(\boldsymbol{u})| < \varepsilon$ and $\arg(z - \rho(\boldsymbol{u})) \neq 0$, where $c_{k,j}(\boldsymbol{u})$ and $\rho(\boldsymbol{u})$ are analytic functions of $\boldsymbol{u}$, and the functions $h_j(t, \boldsymbol{u})$ are analytic at $(t, \boldsymbol{u}) = (0, \boldsymbol{1})$. In addition, if the system is algebraic aperiodic, then all $y_j$ have $\rho(\boldsymbol{u})$ as their unique dominant singularity, and there exist constants $0 < \delta < \pi/2$ and $\eta > 0$ such that $\boldsymbol{y}(z, \boldsymbol{u})$ is analytic in a region of the form

$$\Delta := \{z \colon |z| < \rho(\boldsymbol{1}) + \eta, |\arg(z/\rho(\boldsymbol{u}) - 1)| > \delta\}. \tag{A.15}$$

*Remark* A.6. The above Drmota–Lalley–Woods theorem has been further generalised by Drmota, Gittenberger and Morgenbesser in the case of strongly connected systems with infinitely many equations [DGM12]. In their generalisation, the authors require that the Jacobian of the system (or some of its power) is a compact operator. Alas, as the system corresponding to closed $\lambda$-terms is not strongly connected, it does not fit into their framework. In the current paper we introduce a new condition of *exponential convergence* which is independent of the Jacobian and conjecture that it is crucial for obtaining the respective Puiseux expansions of generating functions.

**Proposition A.7** (Differential condition for the systems of equations [DGM12, see proof of Theorem 1]). Let $\boldsymbol{y} = \boldsymbol{\Phi}(z, \boldsymbol{y})$ be a non-linear system of polynomial equations $y_j = \Phi_j(z, y_1, \ldots, y_m)$ with $j$ ranging over $1, \ldots, m$. Assume that $\boldsymbol{y} = \boldsymbol{\Phi}(z, \boldsymbol{y})$ is algebraic positive, proper and irreducible. Let $\rho$ be the common singularity of its solution vector $y_j$. Then, the spectral radius (largest absolute value of its eigenvalues) of the Jacobian matrix $\dfrac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{y}}$ is a strictly increasing function of $z$ on the interval $[0, \rho]$ and is bounded from above by 1, with the equality holding if and only if $z = \rho$.

## A.2 Limit laws

Consider a bivariate generating function $L(z, u)$ with non-negative coefficients and a sequence of random variables $(X_n)_{n \geq 0}$ such that

$$L(z, u) = \sum_{n,k \geq 0} a_{n,k} z^n u^k \quad \text{and} \quad \mathbb{P}(X_n = k) = \frac{a_{n,k}}{\sum_{j \geq 0} a_{n,j}}. \tag{A.16}$$

We say that $X_n$ *is associated with variable* $u$. In order to understand the limiting behaviour of $X_n$ we investigate the *probability generating function $p_n(u)$* of $X_n$ defined as

$$p_n(u) := \sum_{k \geq 0} \mathbb{P}(X_n = k) u^k = \frac{[z^n] L(z, u)}{[z^n] L(z, 1)}. \tag{A.17}$$

Once accessed, $p_n(u)$ proves extremely useful in establishing the traits of $X_n$ as $n$ tends to infinity. In what follows, we focus on two types of limiting distributions. The first type is related to the case of a so-called *fixed* singularity, which results in discrete limit law; the second type is related to so-called *moving* singularity, and typically results in a Gaussian limit law.

### A.2.1 Discrete limit laws

**Proposition A.8** ( [FS09, Section IX.2]). Suppose that bivariate power series $L(z, u)$ admits in a complex neighbourhood of $u = 1$ a Puiseux series expansion in form of

$$L(z, u) = \alpha(u) - \beta(u)\sqrt{1 - \frac{z}{\rho}} + O\left(\left|1 - \frac{z}{\rho}\right|\right) \tag{A.18}$$

as $z \to \rho$ uniformly in delta-domain $\Delta(R)$ for some $R > \rho$ (see Proposition A.1). Then, the random variable $X_n$ associated with the marking variable $u$ converges in distribution to a discrete limiting distribution with probability generating function

$$p(u) = \lim_{n \to \infty} p_n(u) = \frac{\beta(u)}{\beta(1)}. \tag{A.19}$$

The corresponding mean values satisfy

$$\lim_{n \to \infty} \mathbb{E}X_n = \frac{\beta'(1)}{\beta(1)}. \tag{A.20}$$

### A.2.2 Central limit theorem

*Remark* A.9. In 1983, Bender and Richmond [BR83] proved a multi-dimensional variant of the central limit theorem for probability generating functions taking the quasi-power form $p_n(\boldsymbol{u}) \sim A(\boldsymbol{u})B(\boldsymbol{u})^n$. This line of research was later continued by Hwang [Hwa98] who established precise rates of convergence in the one-dimensional case. The two-dimensional case by was next investigated by Heuberger [Heu07]. More recently, in 2016, the full multi-dimensional version has been resolved by Heuberger and Kropf [HK16] using a multi-dimensional version of the Berry–Esseen inequality. Although we do not touch on the rates of convergence in the current paper, let us mention that they can be obtained using the above results.

In order to formulate the multivariate central limit theorem, it is convenient to introduce the notion of logarithmic derivative which enters the mean value and the covariance matrix of the resulting random variable.

**Definition A.10.** The logarithmic derivative of $A(z, u)$ is given by the expression

$$\frac{\partial}{\partial \log u} A(z, u) := \left.\frac{\partial}{\partial \eta} A(z, e^\eta)\right|_{\eta = \log u} = u\frac{\partial}{\partial u} A(z, u). \tag{A.21}$$

**Proposition A.11** (Multivariate central limit theorem, [BR83, Theorem 1]). Let $(\boldsymbol{X}_n)_{n=1}^\infty$ be a sequence of coordinate-wise non-negative $d$-dimensional discrete random vectors with probability generating functions $p_n(\boldsymbol{u}) := \mathbb{E}(\boldsymbol{u}^{X_n})$, $\boldsymbol{u} = (u_1, u_2, \ldots, u_d)$. Suppose that uniformly in a fixed complex neighbourhood of $\boldsymbol{u} = \boldsymbol{1}$ one has

$$p_n(\boldsymbol{u}) \sim A(\boldsymbol{u}) \cdot B(\boldsymbol{u})^n \tag{A.22}$$

where $A(\boldsymbol{u})$ is uniformly continuous and $B(\boldsymbol{u})$ has a quadratic Taylor series expansion with error term $O\big(\sum |u_k - 1|^3\big)$. Assume that $B(\boldsymbol{u})$ satisfies the following *variability condition*:

$$\det \left[ \frac{\partial^2 \log B(\boldsymbol{u})}{\partial \log u_i\, \partial \log u_j} \right]_{i,j} > 0. \tag{A.23}$$

Then, the sequence of random variables $\boldsymbol{X}_n$, after standardization, converges in law to Gaussian random variable satisfying

$$\frac{\boldsymbol{X}_n - \mathbb{E}\boldsymbol{X}_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}\left(\boldsymbol{0}, \Sigma\right). \tag{A.24}$$

The mean vector and the covariance matrix satisfy

$$\mathbb{E}\boldsymbol{X}_n \sim n \cdot \left. \frac{\partial B}{\partial \boldsymbol{u}} \right|_{\boldsymbol{u}=\boldsymbol{1}} \quad \text{and} \quad \operatorname{Cov} \boldsymbol{X}_n \sim n \cdot \left. \left[ \frac{\partial^2 \log B(\boldsymbol{u})}{\partial \log u_i \partial \log u_j} \right]_{i,j} \right|_{\boldsymbol{u}=1}. \tag{A.25}$$

In the one-dimensional case (A.25) simplifies to

$$\mathbb{E}(X_n) \sim nB'(1) \quad \text{and} \quad \mathbb{V}(X_n) \sim n\left( B''(1) + B'(1) - B'(1)^2 \right). \tag{A.26}$$

*Remark* A.12. Typically, when the singularity is moving (see Proposition A.5) the bivariate generating function takes the form

$$A(z, \boldsymbol{u}) = \alpha(\boldsymbol{u}) - \beta(\boldsymbol{u})\sqrt{1 - \frac{z}{\rho(\boldsymbol{u})}} + O\left( \left| 1 - \frac{z}{\rho(\boldsymbol{u})} \right| \right) \tag{A.27}$$

uniformly as $z \to \rho(\boldsymbol{u})$ for $\boldsymbol{u}$ in a vicinity of $\boldsymbol{1}$.

Consequently, the probability generating function takes form

$$p_n(\boldsymbol{u}) \sim \frac{\beta(\boldsymbol{u})}{\beta(\boldsymbol{1})} \left( \frac{\rho(\boldsymbol{1})}{\rho(\boldsymbol{u})} \right)^n. \tag{A.28}$$

In this form, the probability generating function satisfies the premises of the multivariate quasi-power theorem (see Theorem A.9) and so one can also obtain the speed of convergence. In our situations this speed is typically of order $O\left( \frac{1}{\sqrt{n}} \right)$.

*Remark* A.13. For convenience, we say that a random vector $\boldsymbol{X}_n$ converges in law to multivariate Gaussian distribution with mean $n\boldsymbol{\mu}$ and variance $n\Sigma$ writing

$$\boldsymbol{X}_n \xrightarrow{d} \mathcal{N}(n\boldsymbol{\mu}, n\Sigma) \tag{A.29}$$

to denote that $\dfrac{\boldsymbol{X}_n - n\boldsymbol{\mu}}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \Sigma)$.