

# Towards the Average-Case Analysis of Substitution Resolution in $\lambda$ -Calculus

Maciej Bendkowski

Jagiellonian University, Faculty of Mathematics and Computer Science,  
Theoretical Computer Science Department, ul. Prof. Łojasiewicza 6, 30–348 Kraków, Poland  
maciej.bendkowski@tcs.uj.edu.pl

---

## Abstract

Substitution resolution supports the computational character of  $\beta$ -reduction, complementing its execution with a capture-avoiding exchange of terms for bound variables. Alas, the meta-level definition of substitution, masking a non-trivial computation, turns  $\beta$ -reduction into an atomic rewriting rule, despite its varying operational complexity. In the current paper we propose a somewhat indirect average-case analysis of substitution resolution in the classic  $\lambda$ -calculus, based on the quantitative analysis of substitution in  $\lambda v$ , an extension of  $\lambda$ -calculus internalising the  $v$ -calculus of explicit substitutions. Within this framework, we show that for any fixed  $n \geq 0$ , the probability that a uniformly random, conditioned on size,  $\lambda v$ -term  $v$ -normalises in  $n$  normal-order (i.e. leftmost-outermost) reduction steps tends to a computable limit as the term size tends to infinity. For that purpose, we establish an effective hierarchy  $(\mathcal{G}_n)_n$  of regular tree grammars partitioning  $v$ -normalisable terms into classes of terms normalising in  $n$  normal-order rewriting steps. The main technical ingredient in our construction is an inductive approach to the construction of  $\mathcal{G}_{n+1}$  out of  $\mathcal{G}_n$  based, in turn, on the algorithmic construction of finite intersection partitions, inspired by Robinson’s unification algorithm. Finally, we briefly discuss applications of our approach to other term rewriting systems, focusing on two closely related formalisms, i.e. the full  $\lambda v$ -calculus and combinatory logic.

**2012 ACM Subject Classification** Theory of computation  $\rightarrow$  Lambda calculus; Mathematics of computing  $\rightarrow$  Generating functions

**Keywords and phrases** lambda calculus, explicit substitutions, complexity, combinatorics

**Digital Object Identifier** 10.4230/LIPIcs.FSCD.2019.7

**Funding** *Maciej Bendkowski*: Maciej Bendkowski was partially supported within the Polish National Science Center grant 2016/21/N/ST6/01032.

## 1 Introduction

Traditional, machine-based computational models, such as Turing machines or RAMs, admit a natural notion of an atomic computation step, closely reflecting the actual operational cost of executing the represented computations. Unfortunately, this is not quite the case for computational models based on term rewriting systems with substitution, such as the classic  $\lambda$ -calculus. Given the (traditionally) epiteoretic nature of substitution, the single rewriting rule of  $\beta$ -reduction  $(\lambda x.a)b \rightarrow_{\beta} a[x := b]$  masks a non-trivial computation of resolving (i.e. executing) the pending substitution of  $b$  for occurrences of  $x$  in  $a$ . Moreover, unlike machine-based models,  $\lambda$ -calculus (as other term rewriting systems) does not impose a strict, deterministic evaluation mechanism. Consequently, various strategies for resolving substitutions can be used, even more obfuscating the operational semantics of  $\beta$ -reduction and hence also its operational cost. Those subtle nuances hidden behind the implementation details of substitution resolution are in fact one of the core issues in establishing reasonable cost models for the classic  $\lambda$ -calculus, relating it with other, machine-based computational models, see [14].



© Maciej Bendkowski;

licensed under Creative Commons License CC-BY

4th International Conference on Formal Structures for Computation and Deduction (FSCD 2019).

Editor: Herman Geuvers; Article No. 7; pp. 7:1–7:21

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In order to resolve this apparent inadequacy, Abadi et al. proposed to refine substitution in the classic  $\lambda$ -calculus and decompose it into a series of atomic rewriting steps, internalising in effect the calculus of executing substitutions [1]. Substitutions become first-class citizens and so can be manipulated together with regular terms. Consequently, the general framework of explicit substitutions provides a machine-independent setup for the operational semantics of substitution, based on a finite set of unit rewriting primitives. Remarkably, with the help of linear substitution calculus (a resource aware calculus of explicit substitutions) Accattoli and Dal Lago showed recently that the leftmost-outermost  $\beta$ -reduction strategy is a reasonable invariant cost model for  $\lambda$ -calculus, and hence it is able to simulate RAMs (or equivalent, machine-based models) within a polynomial time overhead [2].

Various subtleties of substitution resolution, reflected in the variety of available calculi of explicit substitutions, induce different operational semantics for executing substitutions in  $\lambda$ -calculus. This abundance of approaches is perhaps the main barrier in establishing a systematic, quantitative analysis of the operational complexity of substitution resolution and, among other things, a term rewriting analogue of classic average-case complexity analysis. In the current paper we propose a step towards filling this gap by offering a quantitative approach to substitution resolution in Lescanne's  $\lambda v$ -calculus of explicit substitutions [15]. In particular, we focus on the following, average-case analysis type of question. Having fixed arbitrary non-negative  $n$ , what is the probability that a (uniformly) random  $\lambda v$ -term of given size is  $v$ -normalisable (i.e. can be reduced to a normal form without explicit substitutions) in exactly  $n$  leftmost-outermost reduction steps? Furthermore, how does this probability distribution change when the term size tends to infinity?

We address the above questions using a two-step approach. First, we exhibit an effective (i.e. computable) hierarchy  $(\mathcal{G}_n)_n$  of unambiguous regular tree grammars with the property that  $\mathcal{G}_n$  describes the language of terms  $v$ -normalising in precisely  $n$  leftmost-outermost  $v$ -rewriting steps. Next, borrowing techniques from analytic combinatorics, we analyse the limit proportion of terms  $v$ -normalising in  $n$  normal-order steps. To that end, we construct appropriate generating functions and provide asymptotic estimates for the number of  $\lambda v$ -terms  $v$ -normalising in  $n$  normal-order reduction steps. As a result, we base our approach on a direct quantitative analysis of the  $v$  term rewriting system, measuring the operational cost of evaluating substitution in terms of the number of leftmost-outermost rewriting steps required to reach a ( $v$ -)normal form.

The paper is structured as follows. In Section 2 we outline  $\lambda v$ -calculus and the framework of regular tree grammars, establishing the necessary terminology for the remainder of the paper. Next, in Section 3, we prepare the background for the construction of  $(\mathcal{G}_n)_n$ . In particular, we sketch its general, intuitive scheme. In Section 4 we introduce the main tool of finite intersection partitions and show that it is indeed constructible in the context of generated reduction grammars. Afterwards, in Section 5, we show how finite intersection partitions can be used in the construction of new productions in  $\mathcal{G}_{n+1}$  based on productions in the grammar  $\mathcal{G}_n$ . Having constructed  $(\mathcal{G}_n)_n$  we then proceed to the main quantitative analysis of  $v$ -calculus using methods of analytic combinatorics, see Section 6. Finally, in Section 7 we discuss broader applications of our technique to other term rewriting systems, based on the examples of  $\lambda v$ -calculus and combinatory logic, and conclude the paper in the final Section 8.

## 2 Preliminaries

### 2.1 Lambda epsilon calculus

$\lambda v$  (lambda epsilon) is a simple, first-order term rewriting system extending the classic  $\lambda$ -calculus based on de Bruijn indices [11] with the calculus of resolving pending substitutions [15, 16]. Its formal terms, so-called  $\lambda v$ -terms, are comprised of de Bruijn indices  $\underline{n}$ ,

application, abstraction, together with an additional, explicit *closure* operator  $[\cdot]$  standing for unresolved substitutions. De Bruijn indices are represented in unary base expansion. In other words,  $\underline{n}$  is encoded as an  $n$ -fold application of the successor operator  $S$  to zero  $\underline{0}$ . Substitutions, in turn, consist of three primitives, i.e. a constant *shift*  $\uparrow$ , a unary *lift* operator  $\uparrow$ , mapping substitutions onto substitutions, and a unary *slash* operator  $/$ , mapping terms onto substitutions. Terms containing closures are called *impure* whereas terms without them are said to be *pure*. Figure 1 summarises the formal specification of  $\lambda v$ -terms and the corresponding rewriting system  $\lambda v$ .

$$\begin{array}{ll}
 t ::= \underline{n} \mid \lambda t \mid tt \mid t[s] & (\lambda a)b \rightarrow a[b/] \quad (\text{Beta}) \\
 s ::= t/ \mid \uparrow(s) \mid \uparrow & (ab)[s] \rightarrow a[s](b[s]) \quad (\text{App}) \\
 \underline{n} ::= \underline{0} \mid S \underline{n}. & (\lambda a)[s] \rightarrow \lambda(a[\uparrow(s)]) \quad (\text{Lambda}) \\
 & \underline{0}[a/] \rightarrow a \quad (\text{FVar}) \\
 & (S \underline{n})[a/] \rightarrow \underline{n} \quad (\text{RVar}) \\
 & \underline{0}[\uparrow(s)] \rightarrow \underline{0} \quad (\text{FVarLift}) \\
 & (S \underline{n})[\uparrow(s)] \rightarrow \underline{n}[s][\uparrow] \quad (\text{RVarLift}) \\
 & \underline{n}[\uparrow] \rightarrow S \underline{n}. \quad (\text{VarShift})
 \end{array}$$

(a) Terms of  $\lambda v$ -calculus.

(b) Rewriting rules.

■ **Figure 1** The  $\lambda v$ -calculus rewriting system.

► **Example 2.1.** Note that the well-known combinator  $K = \lambda xy.x$  is represented in the de Bruijn notation as  $\lambda \lambda \underline{1}$ . The reverse application term  $\lambda xy.yx$ , on the other hand, is represented as  $\lambda \lambda \underline{0} \underline{1}$ . Consequently, in a single  $\beta$ -reduction step, it holds  $(\lambda \lambda \underline{0} \underline{1}) K \rightarrow_{\beta} \lambda(\underline{0}K)$ . In  $\lambda v$ -calculus, however, this single  $\beta$ -reduction is decomposed into a series of small rewriting steps governing both the  $\beta$ -reduction as well as the subsequent substitution resolution. For instance, we have

$$\begin{aligned}
 (\lambda \lambda \underline{0} \underline{1}) K &\rightarrow (\lambda \underline{0} \underline{1}) [K/] \rightarrow (\lambda (\underline{0} \underline{1}) [\uparrow(K/)]) \rightarrow \lambda (\underline{0}[\uparrow(K/)] (\underline{1}[\uparrow(K/)])) \\
 &\rightarrow \lambda (\underline{0} (\underline{1}[\uparrow(K/)])) \rightarrow \lambda (\underline{0} (\underline{0}[K/][\uparrow])) \rightarrow \lambda (\underline{0} (K[\uparrow])).
 \end{aligned} \tag{1}$$

Furthermore,

$$\begin{aligned}
 K[\uparrow] &= (\lambda \lambda \underline{1}) [\uparrow] \rightarrow \lambda ((\lambda \underline{1}) [\uparrow(\uparrow)]) \rightarrow \lambda \lambda (\underline{1}[\uparrow(\uparrow)]) \\
 &\rightarrow \lambda \lambda (\underline{0}[\uparrow(\uparrow)][\uparrow]) \rightarrow \lambda \lambda (\underline{0}[\uparrow]) \\
 &\rightarrow \lambda \lambda \underline{1} = K
 \end{aligned} \tag{2}$$

hence indeed  $(\lambda \lambda \underline{0} \underline{1}) K$  rewrites to  $\lambda(\underline{0}K)$ .

Let us notice that in the presence of the erasing (RVar) and duplicating (App) rewriting rules, not all reduction sequences starting with the same term have to be of equal length. Like in the classic  $\lambda$ -calculus, depending on the considered term, some rewriting strategies might be more efficient than others.

$\lambda v$  enjoys a series of pleasing properties. Most notably,  $\lambda v$  is confluent, correctly implements  $\beta$ -reduction of the classic  $\lambda$ -calculus, and preserves strong normalisation of closed terms [3]. Moreover, the  $v$  fragment, i.e.  $\lambda v$  without the (Beta) rule, is terminating. In

other words, each  $\lambda v$ -term is  $v$ -normalising as can be shown using, for instance, polynomial interpretations [9]. In the current paper we focus on the normal-order (i.e. leftmost-outermost) evaluation strategy of  $v$ -reduction. For convenience, we assume the following notational conventions. We use lowercase letters  $a, b, c, \dots$  to denote arbitrary terms and  $s$  (with or without subscripts) to denote substitutions. Moreover, we write  $a \downarrow_n$  to denote the fact that  $a$  normalises to its  $v$ -normal form in  $n$  normal-order  $v$ -reduction steps. Sometimes, for further convenience, we also simply state that  $t$  normalises in  $n$  steps, without specifying the assumed evaluation strategy nor the specific rewriting steps and normal form.

## 2.2 Regular tree languages

We base our main construction in the framework of regular tree languages. In what follows, we outline their basic characteristics and that of corresponding regular tree grammars, introducing necessary terminology. We refer the curious reader to [10, Chapter II] for a more detailed exposition.

► **Definition 2.2** (Ranked alphabet). *A ranked alphabet  $\mathcal{F}$  is a finite set of function symbols endowed with a corresponding arity function  $\text{arity}: \mathcal{F} \rightarrow \mathbb{N}$ . We use  $\mathcal{F}_n$  to denote the set of function symbols of arity  $n$ , i.e. function symbols  $f \in \mathcal{F}$  such that  $\text{arity}(f) = n$ . Function symbols of arity zero are called constants. As a notational convention, we use lowercase letters  $f, g, h, \dots$  to denote arbitrary function symbols.*

► **Definition 2.3** (Terms). *Let  $X$  be a finite set of variables. Then, the set  $\mathcal{T}_{\mathcal{F}}(X)$  of terms over  $\mathcal{F}$  is defined inductively as follows:*

- $X, \mathcal{F}_0 \subset \mathcal{T}_{\mathcal{F}}(X)$ ;
- If  $f \in \mathcal{F}_n$  and  $\alpha_1, \dots, \alpha_n \in \mathcal{T}_{\mathcal{F}}(X)$ , then  $f(\alpha_1, \dots, \alpha_n) \in \mathcal{T}_{\mathcal{F}}(X)$ .

*Terms not containing variables, in other words elements of  $\mathcal{T}_{\mathcal{F}}(\emptyset)$ , are called ground terms.*

*As a notational convention, we use lowercase Greek letters  $\alpha, \beta, \gamma, \dots$  to denote arbitrary terms. Whenever it is clear from the context, we use the word term both to refer to the above structures as well as to denote  $\lambda v$ -terms.*

► **Definition 2.4** (Regular tree grammars). *A regular tree grammar  $\mathcal{G} = (S, \mathcal{N}, \mathcal{F}, \mathcal{P})$  is a tuple consisting of:*

- an axiom  $S \in \mathcal{N}$ ;
- a finite set  $\mathcal{N}$  of non-terminal symbols;
- a ranked alphabet  $\mathcal{F}$  of terminal symbols such that  $\mathcal{F} \cap \mathcal{N} = \emptyset$ ; and
- a finite set  $\mathcal{P}$  of productions in form of  $N \rightarrow \alpha$  such that  $N \in \mathcal{N}$  and  $\alpha \in \mathcal{T}_{\mathcal{F}}(\mathcal{N})$ .

*A production  $N \rightarrow \alpha$  is self-referencing if  $N$  occurs in  $\alpha$ . Otherwise, if  $N$  does not occur in  $\alpha$ , we say that  $n \rightarrow \alpha$  is regular. As a notational convention, we use capital letters  $X, Y, Z \dots$  to denote arbitrary non-terminal symbols.*

► **Definition 2.5** (Derivation relation). *The derivation relation  $\rightarrow_{\mathcal{G}}$  associated with the grammar  $\mathcal{G} = (S, \mathcal{N}, \mathcal{F}, \mathcal{P})$  is a relation on pairs of terms in  $\mathcal{T}_{\mathcal{F}}(\mathcal{N})$  satisfying  $\alpha \rightarrow_{\mathcal{G}} \beta$  if and only if there exists a production  $N \rightarrow \gamma$  in  $\mathcal{P}$  such that after substituting  $\gamma$  for some occurrence of  $N$  in  $\alpha$  we obtain  $\beta$ . Following standard notational conventions, we use  $\xrightarrow{*}_{\mathcal{G}}$  to denote the transitive closure of  $\rightarrow_{\mathcal{G}}$ . Moreover, if  $\mathcal{G}$  is clear from the context, we omit it in the subscript of the derivation relations and simply write  $\rightarrow$  and  $\xrightarrow{*}$ .*

*A regular tree grammar  $\mathcal{G}$  with axiom  $S$  is said to be unambiguous if and only if for each ground term  $\alpha \in \mathcal{T}_{\mathcal{F}}(\emptyset)$  there exists at most one derivation sequence in form of  $S \rightarrow \gamma_1 \rightarrow \dots \rightarrow \gamma_n = \alpha$ . Likewise,  $N$  is said to be unambiguous in  $\mathcal{G}$  if and only if for each ground term  $\alpha \in \mathcal{T}_{\mathcal{F}}(\emptyset)$  there exists at most one derivation sequence in form of  $N \rightarrow \gamma_1 \rightarrow \dots \rightarrow \gamma_n = \alpha$ .*

► **Definition 2.6** (Regular tree languages). *The language  $L(\mathcal{G})$  generated by  $\mathcal{G}$  is the set of all ground terms  $\alpha$  such that  $S \xrightarrow{*} \alpha$  where  $S$  is the axiom of  $\mathcal{G}$ . Similarly, the language generated by term  $\alpha \in \mathcal{T}_{\mathcal{F}}(\mathcal{N})$  in  $\mathcal{G}$ , denoted as  $L_{\mathcal{G}}(\alpha)$ , is the set of all ground terms  $\beta$  such that  $\alpha \xrightarrow{*} \beta$ . Finally, a set  $\mathcal{L}$  of ground terms is said to be a regular tree language if there exists a regular tree grammar  $\mathcal{G}$  such that  $L(\mathcal{G}) = \mathcal{L}$ .*

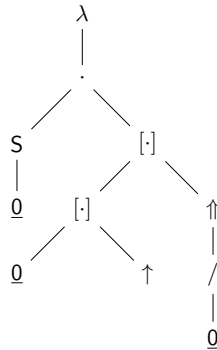
► **Example 2.7.** The set of  $\lambda v$ -terms is an example of a regular tree language. The corresponding regular tree grammar  $\Lambda = (T, \mathcal{N}, \mathcal{F}, \mathcal{P})$  consists of

- a set  $\mathcal{N}$  of three non-terminal symbols  $T, S, N$  intended to stand for  $\lambda v$ -terms, substitutions, and de Bruijn indices, respectively, with  $T$  being the axiom of  $\Lambda$ ;
- a set  $\mathcal{F}$  of terminal symbols, comprised of all the symbols of the  $\lambda v$ -calculus language, i.e. term application and abstraction, closure  $[\cdot]$ , slash  $\cdot/$ , lift  $\uparrow(\cdot)$  and shift  $\uparrow$  operators, and the successor  $S(\cdot)$  with the constant  $\underline{0}$ ; and
- a set  $\mathcal{P}$  of productions

$$\begin{aligned}
 T &\rightarrow N \mid \lambda T \mid TT \mid T[S] \\
 S &\rightarrow T/ \mid \uparrow(S) \mid \uparrow \\
 N &\rightarrow \underline{0} \mid SN.
 \end{aligned}
 \tag{3}$$

Let us notice that (3) consists of five self-referencing productions, three for  $T$  and one for each  $S$  and  $N$ . Moreover,  $L(N) \subset L(T)$  as  $\mathcal{P}$  includes a production  $T \rightarrow N$ .

► **Example 2.8.** Each  $\lambda v$ -term admits a natural tree-like structure. The following example depicts the tree representation of the term  $\lambda \underline{1}[0[\uparrow][\uparrow(\underline{0}/)]]$ . Note that the (conventionally) implicit term application is represented as an explicit binary node  $(\cdot)$ .



■ **Figure 2** Tree representation of  $\lambda \underline{1}[0[\uparrow][\uparrow(\underline{0}/)]]$ .

### 3 Reduction grammars

We conduct our construction of  $(\mathcal{G}_n)_n$  in an inductive, incremental fashion. Starting with  $\mathcal{G}_0$  corresponding to the set of *pure terms* (i.e.  $\lambda v$ -terms without closures) we build the  $(n + 1)$ st grammar  $\mathcal{G}_{n+1}$  based on the structure of the  $n$ th grammar  $\mathcal{G}_n$ . First-order rewriting rules of  $\lambda v$ -calculus guarantee a close structural resemblance of both their left- and right-hand sides, see Figure 1b. Consequently, with  $\mathcal{G}_n$  at hand, we can analyse the right-hand sides of  $v$  rewriting rules and match them with productions of  $\mathcal{G}_n$ . Based on their structure, we then determine the structure of productions of  $\mathcal{G}_{n+1}$  which correspond to respective left-hand sides. Although such a general idea of constructing  $\mathcal{G}_{n+1}$  out of  $\mathcal{G}_n$  is quite straightforward, its implementation requires some careful amount of detail.

For that reason, we make the following initial preparations. Each grammar  $\mathcal{G}_n$  uses the same, global, ranked alphabet  $\mathcal{F}$  corresponding to  $\lambda v$ -calculus, see Example 2.7. The standard non-terminal symbols  $T, S$ , and  $N$ , together with their respective productions (3), are *pre-defined* in each grammar  $\mathcal{G}_n$ . In addition,  $\mathcal{G}_n$  includes  $n + 1$  non-terminal symbols  $G_0, \dots, G_n$  (the final one being the axiom of  $\mathcal{G}_n$ ) with the intended meaning that for all  $0 \leq k \leq n$ , the language  $L_{\mathcal{G}_n}(G_k)$  is equal to the set of terms  $v$ -normalising in  $k$  normal-order steps. In this manner, building  $(\mathcal{G}_n)_n$  amounts to a careful, incremental extension process, starting with the initial grammar  $\mathcal{G}_0$  comprised of the following extra, i.e. not included in (3), productions:

$$G_0 \rightarrow N \mid \lambda G_0 \mid G_0 G_0. \quad (4)$$

In order to formalise the above preparations and the overall presentation of our construction, we introduce the following, abstract notions of  $v$ -reduction grammar and later also their simple variants.

► **Definition 3.1** (*v*-reduction grammars). *Let  $\Lambda = (T, \mathcal{N}, \mathcal{F}, \mathcal{P})$  be the regular tree grammar corresponding to  $\lambda v$ -terms, see Example 2.7. Then, the regular tree grammar*

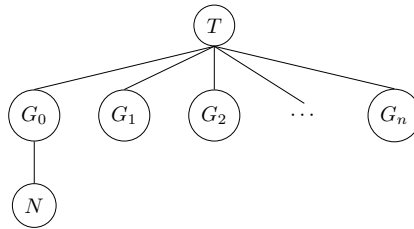
$$\mathcal{G}_n = (G_n, \mathcal{N}_n, \mathcal{F}, \mathcal{P}_n) \quad (5)$$

with

- $\mathcal{N}_n = \mathcal{N} \cup \{G_0, G_1, \dots, G_n\}$ ; and
- $\mathcal{P}_n$  being a set of productions such that  $\mathcal{P} \subset \mathcal{P}_n$

is said to be a *v*-reduction grammar, *v*-RG in short, if all non-terminal symbols  $\mathcal{N}_n$  are unambiguous in  $\mathcal{G}_n$ , for all  $0 \leq k \leq n$  the language  $L(G_k)$  is equal to the set of  $\lambda v$ -terms  $v$ -normalising in  $k$  normal-order  $v$ -steps, and finally  $L(T)$ ,  $L(S)$  and  $L(N)$  are equal to the sets of  $\lambda v$ -terms, substitutions and de Bruijn indices, respectively.

► **Definition 3.2** (Partial order of sorts). *The partial order of sorts  $(\mathcal{N}_n, \preceq)$  is a partial order (i.e. reflexive, transitive, and anti-symmetric relation) on non-terminal symbols  $\mathcal{N}_n$  satisfying  $X \preceq Y$  if and only if  $L_{\mathcal{G}_n}(X) \subseteq L_{\mathcal{G}_n}(Y)$ . For convenience, we write  $X \wedge Y$  to denote the greatest lower bound of  $\{X, Y\}$ . Figure 3 depicts the partial order  $(\mathcal{N}_n, \preceq)$ .*



■ **Figure 3** Hasse diagram of the partial order  $(\mathcal{N}_n, \preceq)$ .

► **Remark 3.3.** Let us notice that given the interpretation of  $L(G_0), \dots, L(G_n)$ , the partial order of sorts  $(\mathcal{N}_n, \preceq)$  captures all the inclusions among the non-terminal languages within  $\mathcal{G}_n$ . However, in addition, if  $X$  and  $Y$  are not comparable through  $\preceq$ , then  $L(X) \cap L(Y) = \emptyset$  as each term  $v$ -normalises in a unique, determined number of steps.

► **Definition 3.4** (Simple  $v$ -reduction grammars). A  $v$ -RG  $\mathcal{G}_n$  is said to be simple if all its self-referencing productions are either productions of the regular tree grammar corresponding to  $\lambda v$ -terms, see Example 2.7, or are of the form

$$G_k \rightarrow \lambda G_k \mid G_0 G_k \mid G_k G_0 \quad (6)$$

and moreover, for all regular productions in form of  $G_k \rightarrow \alpha$  in  $\mathcal{G}_n$ , it holds  $\alpha \in \mathcal{T}_{\mathcal{F}}(\mathcal{N} \cup \{G_0, \dots, G_{k-1}\})$ , i.e.  $\alpha$  does not reference non-terminals other than  $G_0, \dots, G_{k-1}$ .

► **Remark 3.5.** Let us note that, in general,  $v$ -reduction grammars do not have to be simple. Due to the erasing (RVar) rewriting rule, it is possible to construct, *inter alia*, more involved self-referencing productions. Nonetheless, for technical convenience, we will maintain the simplicity of constructed grammars  $(\mathcal{G}_n)_n$ .

Also, let us remark that the above definition of simple  $v$ -reduction grammars asserts that if  $\mathcal{G}_{n+1}$  is a simple  $v$ -RG, then, by a suitable truncation, it is possible to obtain all of the  $v$ -reduction grammars  $\mathcal{G}_0$  up to  $\mathcal{G}_n$ . Consequently,  $\mathcal{G}_{n+1}$  contains, in a proper sense, all the proceeding grammars  $\mathcal{G}_0, \dots, \mathcal{G}_n$ .

## 4 Finite intersection partitions

The main technical ingredient in our construction of  $(\mathcal{G}_n)_n$  are finite intersection partitions.

► **Definition 4.1** (Finite intersection partition). Assume that  $\alpha, \beta$  are two terms in  $\mathcal{T}_{\mathcal{F}}(X)$ . Then, a finite intersection partition, FIP in short, of  $\alpha$  and  $\beta$  is a finite set  $\Pi(\alpha, \beta) = \{\pi_1, \dots, \pi_n\} \subset \mathcal{T}_{\mathcal{F}}(X)$  such that  $L(\pi_i) \cap L(\pi_j) = \emptyset$  whenever  $i \neq j$ , and moreover  $\bigcup_i L(\pi_i) = L(\alpha) \cap L(\beta)$ .

Let us note that, *a priori*, it is not immediately clear if  $\Pi(\alpha, \beta)$  exists for  $\alpha$  and  $\beta$  in the term algebra  $\mathcal{T}(\mathcal{N}_n)$  associated with a simple  $v$ -RG  $\mathcal{G}_n$  nor whether there exists an algorithmic method of its construction. The following result states that both questions can be settled in the affirmative.

► **Lemma 4.2** (Constructible finite intersection partitions). Let  $\mathcal{G}_n$  be a simple  $v$ -reduction grammar. Assume that  $\alpha, \beta$  are two (not necessarily ground) terms in  $\mathcal{T}(\mathcal{N}_n)$  where  $\mathcal{N}_n$  is the set of non-terminal symbols of  $\mathcal{G}_n$ . Then,  $\alpha$  and  $\beta$  have a computable finite intersection partition  $\Pi(\alpha, \beta)$ .

Figure 4 provides a functional pseudocode description of  $\mathbf{fip}_k$  constructing  $\Pi(\alpha, \beta)$  for arbitrary terms  $\alpha, \beta$  within the scope of a simple  $v$ -RG  $\mathcal{G}_k$ . A corresponding proof of correctness is given in Appendix A.

Our finite intersection partition algorithm resembles a variant of Robinson's unification algorithm [17] applied to many-sorted term algebras with a tree hierarchy of sorts, as investigated by Walther, cf. [19]. It becomes even more apparent once the correspondence between sorts, as stated in the language of many-sorted term algebra, and the tree-like hierarchy of non-terminal symbols in  $v$ -reduction grammars is established, see Figure 3.

## 5 The construction of simple $v$ -reduction grammars

Equipped with constructible, finite intersection partitions, we are now ready to describe the generation procedure for  $(\mathcal{G}_n)_n$ . We begin with establishing a convenient *verbosity* invariant maintained during the construction of  $(\mathcal{G}_n)_n$ .



```

1   fun fipk α β :=
2     match α, β with
3     | f(α1, ..., αn), g(β1, ..., βm) ⇒
4       if f ≠ g ∨ n ≠ m then ∅ (* symbol clash *)
5       else if n = m = 0 then {f}
6       else let Πi := fipk αi βi, ∀ 1 ≤ i ≤ n
7             in {f(π1, ..., πn) | (π1, ..., πn) ∈ Π1 × ... × Πn}
8
9     | X, Y ⇒
10      {X ∧ Y | X ≤ Y ∨ Y ≤ X}
11
12     | f(α1, ..., αn), X ⇒
13      fipk X f(α1, ..., αn) (* flip arguments *)
14
15     | X, f(α1, ..., αn) ⇒
16      let Πγ := fipk γ f(α1, ..., αn), ∀ (X → γ) ∈ Gk
17      in ∪(X → γ) ∈ Gk Πγ
18   end.

```

■ **Figure 4** Pseudocode of the  $\text{fip}_k$  procedure computing  $\Pi(\alpha, \beta)$ .

► **Definition 5.1** (Closure width). *Let  $\alpha$  be a term in  $\mathcal{T}_{\mathcal{F}}(X)$  for some finite set  $X$ . Then,  $\alpha$  has closure width  $w$ , if  $w$  is the largest non-negative integer such that  $\alpha$  is of form  $\chi[\sigma_1] \cdots [\sigma_w]$  for some term  $\chi$  and substitutions  $\sigma_1, \dots, \sigma_w$ . For convenience, we refer to  $\chi$  as the head of  $\alpha$  and to  $\sigma_1, \dots, \sigma_w$  as its tail.*

► **Definition 5.2** (Verbose  $v$ -reduction grammars). *A  $v$ -RG  $\mathcal{G}_n$  is said to be verbose if none of its productions is of form  $X \rightarrow G_k[\sigma_1] \cdots [\sigma_w]$  for some arbitrary non-negative  $w$  and  $k$ .*

Simple, verbose  $v$ -reduction grammars admit a neat structural feature. Specifically, their productions preserve closure width of generated terms.

► **Lemma 5.3.** *Assume that  $\mathcal{G}_n$  is a simple, verbose  $v$ -RG. Then, for each production  $G_n \rightarrow \chi[\sigma_1] \cdots [\sigma_w]$  in  $\mathcal{G}_n$  such that its right-hand side is of closure width  $w$ , either  $\chi = N$  or  $\chi$  is in form  $\chi = f(\alpha_1, \dots, \alpha_m)$  for some non-closure function symbol  $f$  of arity  $m$ .*

**Proof.** See Appendix B. ◀

► **Lemma 5.4.** *Assume that  $\mathcal{G}_n$  is a simple, verbose  $v$ -RG. Then, for each production  $G_n \rightarrow \chi[\sigma_1] \cdots [\sigma_w]$  in  $\mathcal{G}_n$  such that its right-hand side is of closure width  $w$ , and ground term  $\delta \in L(\chi[\sigma_1] \cdots [\sigma_w])$  it holds that  $\delta$  is of closure width  $w$ .*

**Proof.** Direct consequence of Lemma 5.3. ◀

The following  $\varphi$ -matchings are the central tool used in the construction of new reduction grammars. Based on finite intersection partitions,  $\varphi$ -matchings provide a simple *template recognition* mechanism which allows us to match productions in  $\mathcal{G}_n$  with right-hand sides of  $v$  rewriting rules.



► **Definition 5.5** ( $\varphi$ -matchings). Let  $\mathcal{G}_n$  be a simple, verbose  $v$ -RG and  $\varphi = \chi[\tau_1] \cdots [\tau_d] \in \mathcal{T}(\mathcal{N}_n)$  be a template (term) of closure width  $d$ . Assume that  $X \rightarrow \gamma[\sigma_1] \cdots [\sigma_w]$  is a production of  $\mathcal{G}_n$  which right-hand side has closure width  $w \geq d$ . Furthermore, let

$$\Delta_\varphi(\gamma[\sigma_1] \cdots [\sigma_w]) = \Pi(\gamma[\sigma_1] \cdots [\sigma_w], \chi[\tau_1] \cdots [\tau_d] \underbrace{[S] \cdots [S]}_{w-d \text{ times}}) \quad (7)$$

be the set of  $\varphi$ -matchings of  $\gamma[\sigma_1] \cdots [\sigma_w]$ .

Then, the set  $\Delta_\varphi^n$  of  $\varphi$ -matchings of  $\mathcal{G}_n$  is defined as

$$\Delta_\varphi^n = \bigcup \{ \Delta_\varphi(\gamma) \mid G_n \rightarrow \gamma \in \mathcal{G}_n \}. \quad (8)$$

For further convenience, we write  $\varphi_{(i)}$  to denote the template  $\varphi = \chi[\tau_1] \cdots [\tau_d]$  with  $i$  copies of  $[S]$  appended to its original tail, i.e.  $\varphi_{(i)} = \chi[\tau_1] \cdots [\tau_d] \underbrace{[S] \cdots [S]}_{i \text{ times}}$ .

■ **Table 1**  $v$ -rewriting rules with respective templates and production schemes.

	Rewriting rule	Template $\varphi$	Production scheme $\Delta_\varphi^n \mapsto \gamma$
(App)	$(ab)[s] \rightarrow a[s](b[s])$	$T[S](T[S])$	$\alpha[\tau_1](\beta[\tau_2])[\sigma_1] \cdots [\sigma_w] \mapsto (\alpha\beta)[\tau][\sigma_1] \cdots [\sigma_w]^\dagger$
(Lambda)	$(\lambda a)[s] \rightarrow \lambda(a[\uparrow(s)])$	$\lambda(T[\uparrow(S)])$	$\lambda(\alpha[\uparrow(\sigma)])[\sigma_1] \cdots [\sigma_w] \mapsto (\lambda\alpha)[\sigma][\sigma_1] \cdots [\sigma_w]$
(FVar)	$\mathcal{Q}[a/] \rightarrow a$	$T$	see Remark 5.8
(RVar)	$(S \underline{n})[a/] \rightarrow \underline{n}$	$N$	$\alpha[\sigma_1] \cdots [\sigma_w] \mapsto (S \alpha)[T/][\sigma_1] \cdots [\sigma_w]$
(FVarLift)	$\mathcal{Q}[\uparrow(s)] \rightarrow \mathcal{Q}$	$\mathcal{Q}$	$\mathcal{Q}[\sigma_1] \cdots [\sigma_w] \mapsto \mathcal{Q}[\uparrow(S)][\sigma_1] \cdots [\sigma_w]$
(RVarLift)	$(S \underline{n})[\uparrow(s)] \rightarrow \underline{n}[s][\uparrow]$	$N[S][\uparrow]$	$\alpha[\sigma][\uparrow][\sigma_1] \cdots [\sigma_w] \mapsto (S \alpha)[\uparrow(\sigma)][\sigma_1] \cdots [\sigma_w]$
(VarShift)	$\underline{n}[\uparrow] \rightarrow S \underline{n}$	$S N$	$(S \alpha)[\sigma_1] \cdots [\sigma_w] \mapsto \alpha[\uparrow][\sigma_1] \cdots [\sigma_w]$

†For each  $\tau \in \Pi(\tau_1, \tau_2)$ , see Remark 5.7.

In what follows we use computable intersection partitions in our iterative construction of  $(\mathcal{G}_n)_n$ . Recall that if  $\mathcal{G}_n$  is simple then, *inter alia*, self-referencing productions starting with the non-terminal  $G_n$  take the form

$$G_n \rightarrow \lambda G_n \mid G_0 G_n \mid G_n G_0. \quad (9)$$

If  $t \downarrow_n$  (for  $n \geq 1$ ) but it does not start with a head  $v$ -redex, then it must be of form  $t = \lambda a$  or  $t = ab$ . In the former case, it must hold  $a \downarrow_n$ ; hence the pre-defined production  $G_n \rightarrow \lambda G_n$  in  $\mathcal{G}_n$ . In the latter case, it must hold  $a \downarrow_k$  whereas  $b \downarrow_{n-k}$  for some  $0 \leq k \leq n$ . And so, it means that we have to include productions in form of  $G_n \rightarrow G_k G_{n-k}$  for all  $0 \leq k \leq n$  in  $\mathcal{G}_n$ ; in particular, the already mentioned two self-referencing productions, see (9).

Remaining terms have to start with head redexes. Each of these head  $v$ -redexes is covered by a dedicated set of productions. The following Lemma 5.6 demonstrates how  $\varphi$ -matchings and, in particular, finite intersection partitions can be used for that purpose.

► **Lemma 5.6.** Let  $\varphi = \lambda(T[\uparrow(S)])$  be the template corresponding to the (Lambda) rewriting rule, see Table 1, and  $t = (\lambda a)[s][s_1] \cdots [s_w]$ . Then,  $t \downarrow_{n+1}$  if and only if there exists a unique term  $\pi = (\lambda(\alpha[\uparrow(\sigma)]))[\sigma_1] \cdots [\sigma_w] \in \Delta_\varphi^n$  such that  $t \in L((\lambda\alpha)[\sigma][\sigma_1] \cdots [\sigma_w])$ .

**Proof.** See Appendix B. ◀

Let us remark that almost all of the rewriting rules of  $\lambda v$  exhibit a similar construction scheme; the exceptional (App) and (FVar) rewriting rules are discussed in Remark 5.7 and Remark 5.8, respectively. Given a rewriting rule, we start with the respective template  $\varphi$  (see Table 1) and generate all possible  $\varphi$ -matchings in  $\mathcal{G}_n$ . Intuitively, such an operation extracts unambiguous sublanguages out of each production in  $\mathcal{G}_n$  which match the right-hand side of the considered rewriting rule. Next, we consider each term  $\pi \in \Delta_\varphi^n$  and establish new productions  $G_{n+1} \rightarrow \gamma$  in  $\mathcal{G}_{n+1}$  out of  $\pi$ . Assuming that  $\mathcal{G}_n$  is a simple and verbose  $v$ -RG, the novel productions generated by means of  $\Delta_\varphi^n$  cover all the  $\lambda v$ -terms reducing in  $n + 1$  normal-order steps, starting with the prescribed head rewriting rule. Since the head of each so constructed production either starts with a function symbol or is equal to  $N$ , cf. Table 1, the outcome grammar is necessarily verbose. Moreover, if we complete the production generation for all rewriting rules, by construction, the grammar  $\mathcal{G}_{n+1}$  must be, at the same time, simple. Consequently, the construction of the hierarchy  $(\mathcal{G}_n)_n$  amounts to an inductive application of the above construction scheme.

► **Remark 5.7.** While following the same pattern for the (App) rule, we notice that the corresponding construction requires a slight modification. Specifically, while matching  $\varphi = T[S](T[S])$  with a right-hand side  $\gamma$  of a production  $G_n \rightarrow \gamma$  in  $\mathcal{G}_n$  we cannot conclude that  $\pi \in \Delta_\varphi^n$  takes the form  $\pi = \alpha[\sigma](\beta[\sigma])[\sigma_1] \cdots [\sigma_w]$ . Note that, in fact, we know that  $\pi = \alpha[\tau_1](\beta[\tau_2])[\sigma_1] \cdots [\sigma_w]$  however perhaps  $\tau_1 \neq \tau_2$ . Nonetheless, we can still compute  $\Pi(\tau_1, \tau_2)$  and use  $\tau \in \Pi(\tau_1, \tau_2)$  to generate a finite set of terms in form of  $\pi = \alpha[\tau](\beta[\tau])[\sigma_1] \cdots [\sigma_w]$ . Using those terms, we can continue with our construction and establish a set of new productions in form of  $G_n \rightarrow (\alpha\beta)[\tau][\sigma_1] \cdots [\sigma_w]$  meant to be included in  $\mathcal{G}_{n+1}$ .

► **Remark 5.8.** Let us also remark that the single rewriting rule which has a template  $\varphi$  not retaining closure width is (FVar). In consequence, the utility of  $\Delta_T(\gamma)$  is substantially limited. If  $t = \underline{Q}[a/][s_1] \cdots [s_w] \downarrow_{n+1}$ , then  $t \rightarrow t' = a[s_1] \cdots [s_w]$  which, in turn, satisfies  $t' \downarrow_n$ . Note that if  $\gamma$  is the right-hand side of a unique production  $G_n \rightarrow \gamma$  in  $\mathcal{G}_n$  generating  $t'$ , then we can match  $T$  with *any* non-empty prefix of  $\gamma$ . The length of the chosen prefix influences what initial part  $\alpha$  of  $\gamma$  is going to be placed under the closure in  $G_{n+1} \rightarrow \underline{Q}[\alpha/][\sigma_1] \cdots [\sigma_w]$ .

This motivates the following approach. Let  $G_n \rightarrow \gamma' = \chi[\sigma_1] \cdots [\sigma_w]$  be an arbitrary production in  $\mathcal{G}_n$  of closure width  $w$ . If  $t' \in L(\gamma')$  and  $t \rightarrow t'$  in a single head (FVar)-reduction, then  $t \in L(\underline{Q}[\chi[\sigma_1] \cdots [\sigma_d]/][\sigma_{d+1}] \cdots [\sigma_w])$  for some  $0 \leq d \leq w$ . Therefore, in order to generate all productions in  $\mathcal{G}_{n+1}$  corresponding to  $\lambda v$ -terms  $v$ -normalising in  $n + 1$  steps, starting with a head (FVar)-reduction, we have to include all productions in form of  $G_{n+1} \rightarrow \underline{Q}[\chi[\sigma_1] \cdots [\sigma_d]/][\sigma_{d+1}] \cdots [\sigma_w]$  for each production  $G_n \rightarrow \gamma'$  in  $\mathcal{G}_n$ .

Finally, note that it is, again, possible to optimise the (FVar) construction scheme with respect to the number of generated productions. For each  $G_n \rightarrow \gamma$  in  $\mathcal{G}_n$  the above scheme produces, *inter alia*, a production  $G_{n+1} \rightarrow \underline{Q}[\gamma/]$ . Note that we can easily merge them into a single production  $G_{n+1} \rightarrow \underline{Q}[G_n/]$  instead.

Such a construction leads us to the following conclusion.

► **Theorem 5.9.** *For all  $n \geq 0$  there exists a constructible, simple  $v$ -RG  $\mathcal{G}_n$ .*

► **Example 5.10.** The following example demonstrates the construction of  $\mathcal{G}_1$  out of  $\mathcal{G}_0$ . Note that  $\mathcal{G}_1$  includes the following productions associated with the axiom  $G_1$ :

$$\begin{aligned} G_1 \rightarrow & \lambda G_1 \mid G_0 G_1 \mid G_1 G_0 \\ & \mid \underline{Q}[(G_0 G_0)/] \mid \underline{Q}[\lambda G_0/] \mid \underline{Q}[N/] \\ & \mid (S N)[T/] \mid \underline{Q}[\uparrow(S)] \mid N[\uparrow]. \end{aligned} \tag{10}$$

The first three productions are included by default. The next three productions are derived from the (FVar) rule applied to all the productions of  $G_0 \rightarrow \gamma$  in  $\mathcal{G}_0$ . The final three productions are obtained by (RVar), (FVarLift), and (VarShift), respectively.

## 6 Analytic combinatorics and simple $v$ -reduction grammars

Having established an effective hierarchy  $(\mathcal{G}_k)_k$  of simple  $v$ -reduction grammars, we can now proceed with their quantitative analysis. Given the fact that regular tree grammars represent well-known algebraic tree-like structures, our analysis is in fact a standard application of algebraic singularity analysis of respective generating functions [12, 13]. The following result provides the main tool of the current section.

► **Proposition 6.1** (Algebraic singularity analysis, see [13], Theorem VII.8). *Assume that  $f(z) = (\sqrt{1 - z/\zeta})g(z) + h(z)$  is an algebraic function, analytic at 0, and has a unique dominant singularity  $z = \zeta$ . Moreover, assume that  $g(z)$  and  $h(z)$  are analytic in the disk  $|z| < \zeta + \varepsilon$  for some  $\varepsilon > 0$ . Then, the coefficients  $[z^n]f(z)$  in the series expansion of  $f(z)$  around the origin, satisfy the following asymptotic estimate*

$$[z^n]f(z) \sim \zeta^{-n} \frac{n^{-3/2}g(\zeta)}{\Gamma(-\frac{1}{2})}. \quad (11)$$

In order to analyse the number of  $\lambda v$ -terms normalising in  $k$  steps, we execute the following plan. First, we use the structure (and unambiguity) of  $\mathcal{G}_k$  to convert it by means of symbolic methods into a corresponding generating function  $G_k(z) = \sum g_n^{(k)} z^n$  in which the integer coefficient  $g_n^{(k)}$  standing by  $z^n$  in the series expansion of  $G_k(z)$ , also denoted as  $[z^n]G_k(z)$ , is equal to the number of  $\lambda v$ -terms of size  $n$  normalising in  $k$  steps. Next, we show that so obtained generating functions fit the premises of Proposition 6.1. We start with establishing an appropriate size notion for  $\lambda v$ -terms. For technical convenience, we assume the following *natural size notion*, equivalent to the number of constructors in the associated term algebra  $\mathcal{T}_{\mathcal{F}}(\emptyset)$ , see Figure 5.

$$\begin{array}{ll} |\underline{n}| = n + 1 & |a| = 1 + |a| \\ |\lambda a| = 1 + |a| & |\uparrow(s)| = 1 + |s| \\ |ab| = 1 + |a| + |b| & |\uparrow| = 1. \\ |a[s]| = 1 + |a| + |s| & \end{array}$$

■ **Figure 5** Natural size notion for  $\lambda v$ -terms, cf. Figure 2.

The following results exhibit the closed-form of generating functions corresponding to pure terms as well as the general class of  $\lambda v$ -terms and explicit substitutions.

► **Proposition 6.2** (see [5]). *Let  $L_\infty(z)$  denote the generating function corresponding to the set of  $\lambda$ -terms in  $v$ -normal form (i.e. without  $v$ -redexes). Then,*

$$L_\infty(z) = \frac{1 - z - \sqrt{\frac{1 - 3z - z^2 - z^3}{1 - z}}}{2z}. \quad (12)$$

► **Proposition 6.3** (see [7]). *Let  $T(z)$ ,  $S(z)$  and  $N(z)$  denote the generating functions corresponding to  $\lambda v$ -terms, substitutions, and de Bruijn indices, respectively. Then,*

$$T(z) = \frac{1 - \sqrt{1 - 4z}}{2z} - 1, \quad S(z) = \frac{1 - \sqrt{1 - 4z}}{2z} \left( \frac{z}{1 - z} \right) \quad \text{and} \quad N(z) = \frac{z}{1 - z}. \quad (13)$$

With the above basic generating functions, we can now proceed with the construction of generating functions corresponding to simple  $\nu$ -reduction grammars. See Appendix C for respective proofs.

► **Proposition 6.4** (Constructible generating functions). *Let  $\Phi_k$  denote the set of regular productions in  $\mathcal{G}_k$ . Then, for all  $k \geq 1$  there exists a generating function  $G_k(z)$  such that  $[z^n]G_k(z)$  (i.e. the coefficient standing by  $z^n$  in the power series expansion of  $G_k(z)$ ) is equal to the number of terms of size  $n$  which  $\nu$ -normalise in  $k$  normal-order reduction steps, and moreover  $G_k(z)$  admits a closed-form of the following shape:*

$$G_k(z) = \frac{1}{1 - z - 2zL_\infty(z)} \sum_{G_k \rightarrow \gamma \in \Phi_k} G_\gamma(z) \quad (14)$$

where

$$G_\gamma(z) = z^{\zeta(\gamma)} T(z)^{\tau(\gamma)} S(z)^{\sigma(\gamma)} N(z)^{\nu(\gamma)} \prod_{0 \leq i < k} G_i(z)^{\rho_i(\gamma)} \quad (15)$$

and all  $\zeta(\gamma)$ ,  $\tau(\gamma)$ ,  $\sigma(\gamma)$ ,  $\nu(\gamma)$ , and  $\rho_i(\gamma)$  are non-negative integers depending on  $\gamma$ .

► **Theorem 6.5.** *For all  $k \geq 1$ , the coefficients  $[z^n]G_k(z)$  satisfy the asymptotic estimate*

$$[z^n]G_k(z) \sim c_k \cdot 4^n n^{-3/2}. \quad (16)$$

Consider the following *asymptotic density* of  $\lambda\nu$ -terms  $\nu$ -normalisable in  $k$  normal-order reduction steps in the set of all  $\lambda\nu$ -terms:

$$\mu_k = \lim_{n \rightarrow \infty} \frac{[z^n]G_k(z)}{[z^n]T(z)}. \quad (17)$$

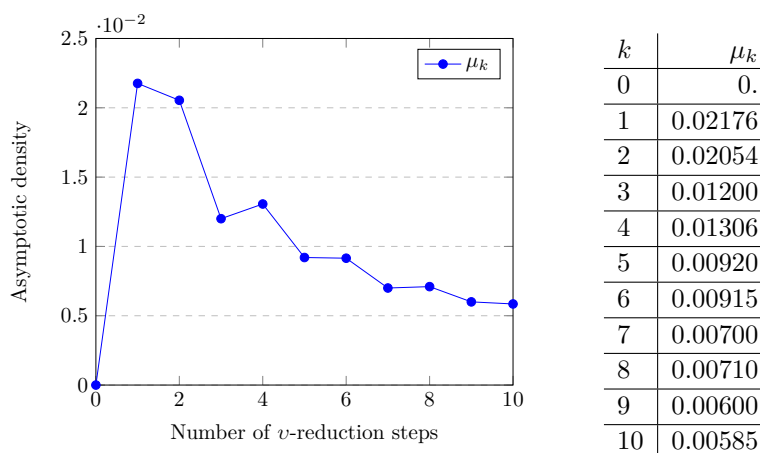
In other words, the limit  $\mu_k$  of the probability that a uniformly random  $\lambda\nu$ -term of size  $n$  normalises in  $k$  steps as  $n$  tends to infinity. Note that for each  $k \geq 1$ , the asymptotic density  $\mu_k$  is positive as both  $[z^n]G_k(z)$  and  $[z^n]T(z)$  admit the same (up to a multiplicative constant) asymptotic estimate. Moreover, it holds  $\mu_k \xrightarrow[k \rightarrow \infty]{} 0$  as the sum  $\sum_k \mu_k$  is increasing and necessarily bounded above by one.

Figure 6 provides the specific asymptotic densities  $\mu_0, \dots, \mu_{10}$  obtained by means of a direct construction and symbolic computations<sup>1</sup> (numerical values are rounded up to the fifth decimal point).

► **Remark 6.6.** Theorem 5.9 and Theorem 6.5 are effective in the sense that both the symbolic representation and the symbolic asymptotic estimate of respective coefficients are computable. Since  $\Gamma(-1/2) = -2\sqrt{\pi}$  is the sole transcendental number occurring in the asymptotic estimates, and cancels out when asymptotic densities are considered, cf. (17), we immediately note that for each  $k \geq 0$ , the asymptotic density of terms  $\nu$ -normalising in  $k$  steps is necessarily an algebraic number.

► **Remark 6.7.** Each  $\lambda\nu$ -term is  $\nu$ -normalising in some (determined) number of normal-order reduction steps. However, it is not clear whether  $\sum_k \mu_k = 1$  as asymptotic density is, in general, not countably additive. Let us remark that if this sum converges to one, then the random variable  $X_n$  denoting the number of normal-order  $\nu$ -reduction steps required to normalise a random  $\lambda\nu$ -term of size  $n$  (i.e. the average-case cost of resolving pending substitutions in a random term of size  $n$ ) converges pointwise to a discrete random variable  $X$  defined as  $\mathbb{P}(X = k) = \mu_k$ . Alas, our current analysis does not determine an answer to this problem.

<sup>1</sup> Corresponding software is available at <https://github.com/maciej-bendkowski/towards-acasrlc>.



■ **Figure 6** Asymptotic densities of terms  $v$ -normalising in  $k$  normal-order reduction steps. In particular, we have  $\mu_0 + \dots + \mu_{10} \approx 0.11162$ .

## 7 Applications to other term rewriting systems

Let us note that the presented construction of reduction grammars does not depend on specific traits immanent to  $\lambda v$ , but rather on certain more general features of its rewriting rules. The key ingredient in our approach is the ability to compute finite intersection partitions for arbitrary terms within the scope of established reduction grammars, which themselves admit certain neat structural properties. Using finite intersection partitions, it becomes possible to generate new productions based on the structural resemblance of both the left-hand and right-hand sides of associated rewriting rules. In what follows we sketch the application of the presented approach to other term rewriting systems, focusing on two examples, i.e.  $\lambda v$ -calculus and combinatory logic. Although the technique is similar, some careful amount of details is required.

### 7.1 $\lambda v$ -calculus

In order to characterise the full  $\lambda v$ -calculus, we need a few adjustments to the already established construction of  $(\mathcal{G}_k)_k$  corresponding to the  $v$  fragment. Clearly, we have to establish a new production construction scheme associated with the (Beta) rewriting rule. Consider the corresponding template  $\varphi = T[T/]$ . Like the respective template for (FVar), cf. Remark 5.8, the current template  $\varphi$  does not retain closure width of generated ground terms. However, at the same time it is also amenable to similar treatment as (FVar).

Let  $t = (\lambda a)b[s_1] \dots [s_w] \downarrow_{n+1}$ . Note that  $t \rightarrow a[b/][s_1] \dots [s_w] \downarrow_n$ . Consequently, one should attempt to match  $\varphi$  with all possible *prefixes* of  $\gamma$  in all productions  $G_n \rightarrow \gamma$  instead of merely their heads, as in the case of  $\Delta_\varphi^n$ . Let  $\gamma = \chi[\sigma_1] \dots [\sigma_d]$  be of closure width  $d$  and  $\delta = \chi[\sigma_1] \dots [\sigma_i]$  be its prefix ( $1 \leq i \leq d$ ). Note that, effectively,  $\Pi(\delta, T[T/])$  checks if  $[\sigma_i]$  takes form  $\beta/$  for some term  $\beta$ . Hence, for each  $\pi \in \Pi(\delta, T[T/])$  we have  $\pi = \alpha[\tau_1] \dots [\tau_{i-1}][\beta/]$  for some terms  $\alpha, \tau_1, \dots, \tau_{i-1}$ , and  $\beta$ . Out of such a partition  $\pi$  we can then construct the production  $G_{n+1} \rightarrow (\lambda \alpha[\tau_1] \dots [\tau_{i-1}])\beta[\tau_{i+1}] \dots [\tau_d]$  in  $\mathcal{G}_{n+1}$ .

However, with the new scheme for (Beta) we are not yet done with our construction. Note that due to the new head redex type, we are, *inter alia*, forced to change the pre-defined set of productions corresponding to  $\lambda v$ -terms without a head redex. Indeed, note that if  $t$  does not start with a head redex, then it must be of form  $(\lambda a)$  or  $(ab)$  where in the latter case

$a$  cannot start with a leading abstraction. This constraint suggests the following approach. We split the set of productions in form of  $G_n \rightarrow \gamma$  into two categories, i.e. productions whose right-hand sides are of form  $\lambda\gamma'$  and the remainder ones. Since both sets are disjoint, we can group them separately and introduce two auxiliary non-terminal symbols  $G_n^{(\lambda)}$  and  $G_n^{(-\lambda)}$  for terms starting with a head abstraction and for those without a head abstraction, respectively, with the additional productions  $G_n \rightarrow G_n^{(\lambda)} \mid G_n^{(-\lambda)}$ . In doing so, it is possible to pre-define the all the productions corresponding to terms without head redexes using productions in form of

$$G_n \rightarrow \lambda G_n \mid G_k^{(-\lambda)} G_{n-k} \quad \text{where} \quad 0 \leq k \leq n. \quad (18)$$

This operation, however, requires a minor adjustment in the formal argument concerning the termination and correctness of constructed finite intersection partitions, see Appendix A and, in particular, Definition A.1). Afterwards, it becomes possible to reuse already established techniques in the construction of new productions in  $\mathcal{G}_{n+1}$  out of  $\mathcal{G}_n$ .

## 7.2 SK-combinators

Using a similar approach, it becomes possible to construct appropriate reduction grammars for  $SK$ -combinators. In particular, our current technique (partially) subsumes, and also simplifies, the results of [6, 4]. With merely two rewriting rules in form of  $Kxy \rightarrow x$  and  $Sxyz \rightarrow xz(yz)$  we can use the developed finite intersection partitions and  $\varphi$ -matchings to construct a hierarchy  $(\mathcal{G}_n)_n$  of normal-order reduction grammars for  $SK$ -combinators. The rewriting rule corresponding to  $K$  is similar to (FVar) whereas the respective rule for  $S$  resembles the (App) rule; as in this case, we have to deal with variable duplication on the right-hand side of the rewriting rule. Instead of closure width, we use a different normal form of terms, and so also productions, based on the sole binary constructor of term application. Consequently, a combinator is of *application width*  $w$  if it takes the form  $X\alpha_1 \dots \alpha_w$  for some primitive combinator  $X \in \{S, K\}$ . Consider the more involved case of productions corresponding to head  $S$ -redexes. Let  $t = Sxyz\alpha_1 \dots \alpha_w$  be a term of application width  $w + 3$  where  $w \geq 0$ . Note that

$$Sxyz\delta_1 \dots \delta_w \rightarrow xz(yz)\delta_1 \dots \delta_w. \quad (19)$$

Let us rewrite the right-hand side of (19) as  $t' = Xx_1 \dots x_k z(yz)\delta_1 \dots \delta_w$  where  $x = Xx_1 \dots x_k$  and  $X$  is a primitive combinator. Assume that  $\gamma$  is the right-hand side of the unique production  $G_n \rightarrow \gamma$  in  $\mathcal{G}_n$  such that  $t' \in L(\gamma)$ . Note that the shape of  $t'$  suggests a construction scheme similar to the already discussed (App), see Remark 5.7, where we first have to match the pattern  $\varphi = T(TT)$  with some arguments of  $\gamma$  and subsequently attempt to extract a finite intersection partition  $\Pi(\alpha, \beta)$  of respective subterms  $\alpha$  and  $\beta$  so that for each  $\pi \in \Pi(\alpha, \beta)$  we have  $z \in L(\pi)$ . With appropriate terms at hand, we can then construct corresponding productions in the next grammar  $\mathcal{G}_{n+1}$ .

## 8 Conclusions

Quantitative aspects of term rewriting systems are not well studied. A general complexity analysis was undertaken by Choppy, Kaplan, and Soria who considered a class of confluent, terminating term rewriting systems in which the evaluation cost, measured in the number of rewriting steps required to reach the normal form, is independent of the assumed evaluation strategy [8]. More recently, typical evaluation cost of normal-order reduction in combinatory

logic was studied by Bendkowski, Grygiel and Zaionc [6, 4]. Using quite different, non-analytic methods, Sin'Ya, Asada, Kobayashi and Tsukada considered certain asymptotic aspects of  $\beta$ -reduction in the simply-typed variant of  $\lambda$ -calculus showing that, typically,  $\lambda$ -terms of order  $k$  have  $(k - 1)$ -fold exponentially long  $\beta$ -reduction sequences [18].

Arguably, the main novelty in the presented approach lies in the algorithmic construction of reduction grammars  $(\mathcal{G}_k)_k$  based on finite intersection partitions, assembled using a general technique reminiscent of Robinson's unification algorithm applied to many-sorted term algebras, cf. [17, 19]. Equipped with finite intersection partitions, the construction of  $\mathcal{G}_{k+1}$  out of  $\mathcal{G}_k$  follows a stepwise approach, in which new productions are established on a per rewriting rule basis. Consequently, the general technique of generating reduction grammars does not depend on specific features of  $\lambda\nu$ , but rather on more general traits of certain first-order rewriting systems. Nonetheless, the full scope of our technique is yet to be determined.

Although the presented construction is based on the leftmost-outermost reduction scheme, it does not depend on the specific size notion associated with  $\lambda\nu$ -terms; in principle, more involved size models can be assumed and analysed. The assumed evaluation strategy, size notion, as well as the specific choice of  $\lambda\nu$  are clearly arbitrary and other, equally perfect choices for modelling substitution resolution could have been made. However, due to merely eight rewriting rules forming  $\lambda\nu$ , it is one of the conceptually simplest calculus of explicit substitutions. Together with the normal-order evaluation tactic, it is therefore one of the simplest to investigate in quantitative terms and to demonstrate the finite intersection partitions technique.

Due to the unambiguity of constructed grammars  $(\mathcal{G}_k)_k$  it is possible to automatically establish their corresponding combinatorial specifications and, in consequence, obtain respective generating functions encoding sequences  $\left(g_n^{(k)}\right)_n$  comprised of numbers  $g_n^{(k)}$  associated with  $\lambda\nu$ -terms of size  $n$  which reduce in  $k$  normal-order rewriting steps to their  $\nu$ -normal forms. Singularity analysis provides then the means for systematic, quantitative investigations into the properties of substitution resolution in  $\lambda\nu$ , as well as its machine-independent operational complexity. Finally, with generating functions at hand, it is possible to undertake a more sophisticated statistical analysis of substitution (in particular  $\nu$ -normalisation) using available techniques of analytic combinatorics, effectively analysing the average-case cost of  $\lambda$ -calculus and related term rewriting systems.

It should be noted that such an analysis might provide some novel insight into the combinatorial structure of substitution and, in the long-term perspective, provide a theoretical model for the operational, average-case analysis of substitution in modern functional programming languages. In view of these objectives, we conclude the paper with the following conjecture, postulating the existence of a limit distribution associated with the average-case cost of substitution resolution, and its close relation with the constructible hierarchy of reduction grammars and established series of asymptotic densities.

► **Conjecture.** *Let  $\mu_k$  denote the asymptotic density of  $\lambda\nu$ -terms  $\nu$ -normalising in  $k$  leftmost-outermost  $\nu$ -reduction steps, cf. Remark 6.7. Then, it holds*

$$\sum_{k \geq 0} \mu_k = 1. \quad (20)$$

Consequently, the sequence  $(X_n)_n$  of random variables corresponding to the number of  $\nu$ -reductions required to normalise a uniformly random  $\lambda\nu$ -term of size  $n$  (i.e. the average-case cost of resolving all pending substitutions in a random  $\lambda\nu$ -term of size  $n$ ) converges pointwise to a random variable  $X$  satisfying  $\mathbb{P}(X = k) := \mu_k$ .



## References

- 1 M. Abadi, L. Cardelli, P.-L. Curien, and J.-J. Lévy. Explicit substitutions. *Journal of Functional Programming*, 1(4):375–416, 1991. doi:10.1017/S095679680000186.
- 2 B. Accattoli and U. Dal Lago. (Leftmost-Outermost) Beta Reduction is Invariant, Indeed. *Logical Methods in Computer Science*, 12(1), 2016. doi:10.2168/LMCS-12(1:4)2016.
- 3 Z.-E.-A. Benaissa, D. Briaud, P. Lescanne, and J. Rouyer-Degli.  $\lambda\nu$ , a calculus of explicit substitutions which preserves strong normalisation. *Journal of Functional Programming*, 6(5):699–722, 1996. doi:10.1017/S0956796800001945.
- 4 M. Bendkowski. Normal-order reduction grammars. *Journal of Functional Programming*, 27, 2017. doi:10.1017/S0956796816000332.
- 5 M. Bendkowski, K. Grygiel, P. Lescanne, and M. Zaionc. Combinatorics of  $\lambda$ -terms: a natural approach. *Journal of Logic and Computation*, 27(8):2611–2630, 2017. doi:10.1093/logcom/exx018.
- 6 M. Bendkowski, K. Grygiel, and M. Zaionc. On the likelihood of normalization in combinatory logic. *Journal of Logic and Computation*, 2017. doi:10.1093/logcom/exx005.
- 7 M. Bendkowski and P. Lescanne. Combinatorics of Explicit Substitutions. In *Proceedings of the 20th International Symposium on Principles and Practice of Declarative Programming, PPDP '18*, pages 7:1–7:12. ACM, 2018. doi:10.1145/3236950.3236951.
- 8 C. Choppy, S. Kaplan, and M. Soria. Complexity Analysis of Term-Rewriting Systems. *Theoretical Computer Science*, 67(2&3):261–282, 1989. doi:10.1016/0304-3975(89)90005-4.
- 9 A. Cichon and P. Lescanne. Polynomial interpretations and the complexity of algorithms. In *Automated Deduction—CADE-11*, pages 139–147. Springer Berlin Heidelberg, 1992.
- 10 H. Comon, M. Dauchet, R. Gilleron, Ch. Löding, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi. *Tree Automata Techniques and Applications*, 2007. Release October, 12th 2007. URL: <http://www.grappa.univ-lille3.fr/tata>.
- 11 N. 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.
- 12 Ph. Flajolet and A. M. Odlyzko. Singularity Analysis of Generating Functions. *SIAM Journal on Discrete Mathematics*, 3(2):216–240, 1990.
- 13 Ph. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 1 edition, 2009.
- 14 J. L. Lawall and H. G. Mairson. Optimality and Inefficiency: What Isn't a Cost Model of the Lambda Calculus? In *Proceedings of the 1st ACM SIGPLAN International Conference on Functional Programming, ICFP '96*, pages 92–101. ACM, 1996. doi:10.1145/232627.232639.
- 15 P. Lescanne. From  $\lambda\sigma$  to  $\lambda\nu$ : A Journey Through Calculi of Explicit Substitutions. In *Proceedings of the 21st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 60–69. ACM, 1994.
- 16 P. Lescanne. The lambda calculus as an abstract data type. In M. Haverdaen, O. Owe, and O.-J. Dahl, editors, *Recent Trends in Data Type Specification*, pages 74–80. Springer Berlin Heidelberg, 1996.
- 17 J. A. Robinson. A Machine-Oriented Logic Based on the Resolution Principle. *J. ACM*, 12(1):23–41, January 1965. doi:10.1145/321250.321253.
- 18 R. Sin'Ya, K. Asada, N. Kobayashi, and T. Tsukada. Almost Every Simply Typed  $\lambda$ -Term Has a Long  $\beta$ -Reduction Sequence. In *Proceedings of the 20th International Conference on Foundations of Software Science and Computation Structures*, volume 10203, pages 53–68. Springer-Verlag New York, Inc., 2017. doi:10.1007/978-3-662-54458-7\_4.
- 19 Ch. Walthers. Many-sorted Unification. *J. ACM*, 35(1):1–17, January 1988. doi:10.1145/42267.45071.

## A

 Correctness of the finite intersection partition algorithm

In order to prove the correctness of  $\mathbf{fip}_k$  constructing  $\Pi(\alpha, \beta)$  for arbitrary terms  $\alpha, \beta$  within the scope of a simple  $v$ -RG  $\mathcal{G}_k$ , see Lemma 4.2, we resort to the following technical notion of *term potential* used to embed terms into the well-founded set of natural numbers.

► **Definition A.1** (Term potential). *Let  $\alpha \in \mathcal{T}(\mathcal{N}_n)$  be a term and  $\mathcal{G}_n$  be a simple  $v$ -RG. Let  $\text{Prod}_{\mathcal{G}_n}(X)$  denote the set consisting of right-hand sides of regular productions in form of  $X \rightarrow \beta$  in  $\mathcal{G}_n$ . Then, the potential  $\pi(\alpha)$  of  $\alpha$  in  $\mathcal{G}_n$  is defined inductively as follows:*

- If  $\alpha = f(\alpha_1, \dots, \alpha_m)$ , then  $\pi(\alpha) = 1 + \sum_{i=1}^m \pi(\alpha_i)$ ;
- If  $\alpha = X \in \{T, S, N\}$ , then  $\pi(\alpha) = 1 + \max\{\pi(\gamma) \mid \gamma \in \text{Prod}_{\mathcal{G}_n}(X)\}$ ;
- If  $\alpha = G_k$  for some  $0 \leq k \leq n$ , then  $\pi(\alpha) = 1 + \max\{\pi(\gamma) \mid \gamma \in \bigcup_{i=0}^k \text{Prod}_{\mathcal{G}_n}(G_i)\}$ .

Let us note that  $\pi$  is well-defined as, by assumption,  $\text{Prod}_{\mathcal{G}_n}(X) \neq \emptyset$  for all simple  $v$ -reduction grammars  $\mathcal{G}_n$ ; otherwise  $L(G_k)$  could not span the whole set of  $\lambda v$ -terms  $v$ -normalising in  $k$  normal-order steps. Moreover,  $\pi$  has the following crucial properties:

- For each term  $\alpha$  we have  $\pi(\alpha) \geq 1$ ;
- If  $\alpha$  is a proper subterm of  $\beta$ , then  $\pi(\alpha) < \pi(\beta)$ ;
- If  $X \rightarrow \alpha$  is a regular production, then  $\pi(\alpha) < \pi(X)$ ; and
- For each  $G_i$  and  $G_j$ , it holds  $\pi(G_i) \leq \pi(G_j)$  whenever  $i \leq j$ .

► **Example A.2.** Note that the term potential of  $N$  associated with de Bruijn indices is equal to  $\pi(N) = 2$  as  $\pi(\underline{0}) = 1$ . Since  $T \rightarrow N$  is the single regular production starting with  $T$  on its left-hand side, the potential  $\pi(T)$  is therefore equal to 3. Consequently, we also have  $\pi(S) = 5$  as witnessed by the regular production  $S \rightarrow T/$ . Finally, since  $\pi(N) = 2$  it holds  $\pi(G_0) = 3$  and so, for instance, we also have  $\pi(G_0G_0) = 7$ .

Productions of a simple  $\mathcal{G}_n$  cannot reference non-terminals other than  $G_0, \dots, G_n$ . Since the potential of  $G_{k+1}$  is defined in terms of the potential of its regular productions, this means that  $\pi(G_{k+1})$  depends, in an implicit manner, on the potentials  $\pi(G_0), \dots, \pi(G_k)$ . Note that this constitutes a traditional inductive definition. In order to compute the potential of a given term  $\alpha$ , we start with computing the potential of associated non-terminals. In particular, we find the values  $\pi(G_0), \dots, \pi(G_n)$  in ascending order. Afterwards, we can recursively decompose  $\alpha$  and calculate its potential based on the potential of non-terminal symbols occurring in  $\alpha$ . Note that the same scheme holds, in particular, for the right-hand sides of self-referencing productions.

► **Definition A.3** (Conservative productions). *A self-referencing production  $X \rightarrow f(\alpha_1, \dots, \alpha_n)$  is said to be conservative if  $\pi(\alpha_i) \leq \pi(X)$  for all  $1 \leq i \leq n$ .*

► **Remark A.4.** Conservative productions play a central role in the algorithmic construction of finite intersection partitions  $\Pi(\alpha, \beta)$ . In particular, let us remark that all self-referencing productions of a simple  $v$ -RG  $\mathcal{G}_n$ , as listed in Figure 7, are at the same time conservative.

With the technical notions of term potential and conservative productions, we are now ready to prove the correctness of  $\mathbf{fip}_k$ , see Figure 4.

**Proof of Lemma 4.2.** Induction over the total potential of  $\alpha$  and  $\beta$ .

Let us start with the base case  $\pi(\alpha) + \pi(\beta) = 2$ . Note that both  $\alpha$  and  $\beta$  have to be constant ground terms (the potential of non-terminals in  $\mathcal{N}_k$  is at least 2). If  $\alpha \neq \beta$ , then certainly  $L(\alpha) \cap L(\beta) = \emptyset$  and so  $\Pi(\alpha, \beta) = \emptyset$ , see Line 4. Otherwise if  $\alpha = \beta$ , then both

$$\begin{array}{ll}
G_k \rightarrow \lambda G_k \mid G_0 G_k \mid G_k G_0 & S \rightarrow \uparrow(S) \\
T \rightarrow \lambda T \mid TT \mid T[S] & N \rightarrow \mathbf{S} N
\end{array}$$

■ **Figure 7** Self-referencing productions of simple  $v$ -reduction grammars.

$L(\alpha) \cap L(\beta) = \Pi(\alpha, \beta) = \{\alpha\} = \{\beta\}$ ; hence,  $\mathbf{fip}_k$  returns a correct intersection partition  $\Pi(\alpha, \beta)$ , see Line 5. And so, assume that  $\pi(\alpha) + \pi(\beta) > 2$ . Depending on the joint structure of both  $\alpha$  and  $\beta$  we have to consider three cases.

**Case 1.** Suppose that  $\alpha = f(\alpha_1, \dots, \alpha_n)$  and  $\beta = g(\beta_1, \dots, \beta_m)$  for some function symbols  $f$  and  $g$  with  $n, m \geq 0$ . Certainly, if either  $f \neq g$  or  $n \neq m$ , then  $L(\alpha) \cap L(\beta) = \emptyset$ . In consequence,  $\emptyset$  is the sole valid FIP of both  $\alpha$  and  $\beta$ , see Line 4.

So, let us assume that both  $f = g$  and  $n = m$ . Moreover, we can also assume that  $n, m \geq 1$  as the trivial case  $n = m = 0$  cannot occur under the working assumption  $\pi(\alpha) + \pi(\beta) > 2$ . Take an arbitrary  $\delta \in L(\alpha) \cap L(\beta)$ . Note that  $\delta$  takes the form of  $\delta = f(\delta_1, \dots, \delta_n)$  for some ground terms  $\delta_1, \dots, \delta_n$ . By induction, for all  $1 \leq i \leq n$ , the recursive call  $\mathbf{fip}_k \alpha_i \beta_i$  yields a finite intersection partition  $\Pi(\alpha_i, \beta_i)$  of  $\alpha_i$  and  $\beta_i$ . Since  $\delta_i \in L(\alpha_i) \cap L(\beta_i)$  there exists a unique  $\pi_i \in \Pi(\alpha_i, \beta_i)$  such that  $\delta_i \in L(\pi_i)$ . Accordingly, for  $\delta = f(\delta_1, \dots, \delta_n)$  there exists a unique term  $\pi = f(\pi_1, \dots, \pi_n)$  in  $\mathbf{fip}_k \alpha \beta$  such that  $\delta \in L(\pi)$ .

Conversely, take an arbitrary ground term  $\delta = f(\delta_1, \dots, \delta_n) \in L(\pi)$  for some  $\pi \in \mathbf{fip}_k \alpha \beta$ . Note that  $\pi$  takes the form  $\pi = f(\pi_1, \dots, \pi_n)$ , see Line 7. Since  $\delta \in L(\pi)$ , we know that  $\delta_i \in L(\pi_i)$ , for each  $1 \leq i \leq n$ . Moreover,  $\pi_i \in \Pi_i$  by the construction of  $\mathbf{fip}_k \alpha \beta$ , see Line 6. Following the inductive hypothesis that  $\Pi_i$  is a FIP of  $\alpha_i$  and  $\beta_i$ , we notice that  $L(\pi_i) \subseteq L(\alpha_i) \cap L(\beta_i)$ . Consequently,  $\delta_i \in L(\alpha_i) \cap L(\beta_i)$  and so  $\delta \in L(\alpha) \cap L(\beta)$ .

**Case 2.** Suppose that  $\alpha = X$  and  $\beta = Y$  are two non-terminal symbols in  $\mathcal{N}_k$ , see Line 9. Let us consider the sort poset  $(\mathcal{N}_k, \preceq)$  associated with  $\mathcal{G}_k$ . Assume that  $X$  and  $Y$  are comparable through  $\preceq$  (w.l.o.g. let  $X \preceq Y$ ). Consequently,  $L(X) \subseteq L(Y)$  and so  $X \wedge Y = X$ . Clearly,  $\{X\}$  is a valid FIP, see Line 10. On the other hand, if  $X$  and  $Y$  are incomparable in the sort poset associated with  $\mathcal{G}_k$ , it means that  $L(X) \cap L(Y) = \emptyset$  and so  $\Pi(\alpha, \beta) = \emptyset$ , see Remark 3.3.

**Case 3.** Suppose w.l.o.g. that  $\alpha = X$  and  $\beta$  takes the form  $\beta = f(\beta_1, \dots, \beta_n)$  with  $n \geq 0$ . Note that  $\mathbf{fip}_k$  flips its arguments if necessary, see Line 13. Take an arbitrary  $\delta \in L(\alpha) \cap L(\beta)$ . Note that from the form of  $\beta$  we know that  $\delta = f(\delta_1, \dots, \delta_n)$  for some ground terms  $\delta_1, \dots, \delta_n$  ( $n \geq 0$ ). Since  $\alpha = X$  is a non-terminal symbol which, by assumption, is unambiguous in  $\mathcal{G}_k$ , there exists a unique production  $X \rightarrow \gamma$  such that  $\delta \in L(\gamma)$ .

If  $X \rightarrow \gamma$  is regular (i.e.  $X$  does not occur in  $\gamma$ ), then  $\pi(\gamma) < \pi(X)$  and so  $\pi(\gamma) + \pi(\beta) < \pi(X) + \pi(\beta)$ . Hence, by induction,  $\mathbf{fip}_k \gamma \beta$  constructs a finite intersection partition  $\Pi(\gamma, \beta)$  with a unique  $\pi \in \Pi(\gamma, \beta) \subset \Pi(X, \beta)$  such that  $\delta \in L(\pi)$ , see Line 17.

Let us therefore assume that  $X \rightarrow \gamma$  is not regular, but instead self-referencing (i.e.  $X$  occurs in  $\gamma$ ). In such a case  $\pi(X) \leq \pi(\gamma)$  and so we cannot directly apply the induction hypothesis to  $\mathbf{fip}_k \gamma \beta$ . Note however, that since  $\delta \in L(\gamma) \cap L(\beta)$  and  $\gamma \neq X$ , the term  $\gamma$  must be of form  $\gamma = f(\gamma_1, \dots, \gamma_n)$  as otherwise  $\delta \notin L(\gamma)$ . Furthermore if  $n = 0$ , then trivially  $\gamma = \beta = f$ , see Line 5. Hence, let us assume that  $n \geq 1$ . It follows that  $\mathbf{fip}_k$  proceeds to construct finite intersection partitions for respective pairs of arguments  $\gamma_i$  and  $\beta_i$ . However, since  $X \rightarrow f(\gamma_1, \dots, \gamma_n)$  is conservative (see Remark A.4), it holds  $\pi(\gamma_i) \leq \pi(X)$ . At the same time,  $\pi(\beta_i) < \pi(\beta)$ ; hence, by induction we can argue

that  $\mathbf{fip}_k \gamma_i \beta_i$  constructs a proper intersection partition  $\Pi(\gamma_i, \beta_i)$  for each  $1 \leq i \leq n$ , see Line 7. There exists therefore a unique term  $\pi = f(\pi_1, \dots, \pi_n) \in \mathbf{fip}_k X \beta$  such that  $\delta \in L(\pi)$ , see Line 7 and Line 17.

Conversely, take an arbitrary ground term  $\delta = f(\delta_1, \dots, \delta_n) \in L(\pi)$  for some  $\pi \in \mathbf{fip}_k X \beta$  ( $n \geq 0$ ). By definition,  $\mathbf{fip}_k$  proceeds to invoke itself on pairs of arguments  $\gamma$  and  $\beta$  where  $\gamma$  is the right-hand side of a production  $X \rightarrow \gamma$  in  $\mathcal{G}_k$ , see Line 16, and returns the set-theoretic union of recursively obtained outcomes. There exists therefore some  $\gamma$  such that  $\pi \in \mathbf{fip}_k \gamma \beta$ . If  $X \rightarrow \gamma$  is regular, then by induction,  $\mathbf{fip}_k \gamma \beta$  constructs a FIP for both  $\gamma$  and  $\beta$ . Consequently, it holds  $\delta \in L(\pi) \subset L(\gamma) \cap L(\beta) \subset L(X) \cap L(\beta)$ . Assume therefore that  $X \rightarrow \gamma$  is not regular, but instead self-referencing. As before, we cannot directly argue about  $\mathbf{fip}_k \gamma \beta$  since the total potential of  $\gamma$  and  $\beta$  exceeds the potential of  $X$  and  $\beta$ . However, since  $\pi \in \mathbf{fip}_k \gamma \beta$  and  $\gamma \neq X$ , we note that  $\gamma$  takes form  $\gamma = f(\gamma_1, \dots, \gamma_n)$ , see Line 4. If  $f$  is a constant symbol, then certainly  $\mathbf{fip}_k \gamma \beta$  outputs a proper FIP. Otherwise,  $\mathbf{fip}_k$  proceeds to invoke itself recursively on respective pairs of arguments  $\gamma_i$  and  $\beta_i$ . Since  $X \rightarrow f(\gamma_1, \dots, \gamma_n)$  is conservative, we know that, by induction,  $\mathbf{fip}_k \gamma_i \beta_i$  constructs finite intersection partitions  $\Pi(\gamma_i, \beta_i)$  for all pairs  $\gamma_i$  and  $\beta_i$ . Certainly,  $\delta_i \in L(\pi_i)$  for some  $\pi_i \in \Pi(\gamma_i, \beta_i)$ ; hence  $\delta \in L(\pi)$  where  $\pi = f(\pi_1, \dots, \pi_n) \in \mathbf{fip}_k \gamma \beta$ . It follows that  $\delta \in L(\pi) \cap L(\beta) \subset L(X) \cap L(\beta)$ , which finishes the proof.  $\blacktriangleleft$

► **Remark A.5.** Note that the termination of  $\mathbf{fip}_k$  is based on the fact that all self-referencing productions of simple  $v$ -reduction grammars are at the same time conservative. Indeed,  $\mathbf{fip}_k$  does not terminate in the presence of non-conservative productions. Consider the non-conservative production  $X \rightarrow f(f(X))$ . Note that

$$\begin{aligned} \mathbf{fip}_k(f(f(X)), f(X)) &\rightarrow \mathbf{fip}_k(f(X), X) \\ &\rightarrow \mathbf{fip}_k(X, f(X)) \\ &\rightarrow \mathbf{fip}_k(f(f(X)), f(X)) \\ &\rightarrow \dots \end{aligned} \tag{21}$$

## B The construction of $(\mathcal{G}_n)_n$

Simple, verbose  $v$ -reduction grammars satisfy the neat structural property of retaining closure width of generated terms. For that reason, we maintain both simplicity and verbosity as an invariant during the construction of  $(\mathcal{G}_n)_n$ .

**Proof of Lemma 5.3.** Suppose that neither  $\chi = N$  nor  $\chi = f(\alpha_1, \dots, \alpha_m)$ . Since  $\mathcal{G}_n$  is simple, it follows that either  $\chi = T$  or  $\chi = G_k$  for some  $0 \leq k \leq n$ . However, due to the verbosity of  $\mathcal{G}_n$  we know that  $\chi \neq G_k$  and so it must hold  $\chi = T$ . Consider the following inductive family of terms:

$$\delta_1 = \mathbb{Q}[\uparrow(\uparrow)] \quad \text{whereas} \quad \delta_{n+1} = \mathbb{Q}[\delta_n/]. \tag{22}$$

By construction, we note that  $\delta_n \downarrow_n$ . Let  $s_1, \dots, s_w$  be substitutions satisfying  $s_i \in L(\sigma_i)$ . Note that  $\delta := \delta_{n+1}[s_1] \cdots [s_w] \in L(T[\sigma_1] \cdots [\sigma_w])$ ; hence, simultaneously  $\delta$  reduces in  $n$  steps, as  $\delta \in L(G_n)$ , and in at least  $n+1$  steps, contradiction.  $\blacktriangleleft$

The following result demonstrates a simple scheme of how  $\varphi$ -matchings can be exploited during the construction of new productions generating terms with specific head redexes.

**Proof of Lemma 5.6.** Let  $t = (\lambda a)[s][s_1] \cdots [s_w] \downarrow_{n+1}$  where  $w \geq 0$ . Since  $t$  admits a head  $v$ -redex, we note that  $t \rightarrow t' = (\lambda(a[\uparrow(s)]))[s_1] \cdots [s_w] \downarrow_n$ . By assumption,  $\mathcal{G}_n$  is simple,

hence there exists a unique production  $G_n \rightarrow \gamma$  in  $\mathcal{G}_n$  such that  $t' \in L(\gamma)$ . Consider the set  $\Delta_\varphi^n$ . Since  $G_n \rightarrow \gamma$  is the unique production satisfying  $t' \in L(\gamma)$ , it follows that for each production  $G_n \rightarrow \gamma'$  in  $\mathcal{G}_n$  such that  $\gamma' \neq \gamma$  and all  $\pi \in \Delta_\varphi(\gamma')$  it holds  $t' \notin L(\pi)$ . Let us therefore focus on the set  $\Delta_\varphi(\gamma)$  of  $\varphi$ -matchings limited to  $\gamma$ .

By assumption,  $\mathcal{G}_n$  is not only simple but also verbose. Consequently, we know that  $\gamma$  retains the closure width of generated terms, see Lemma 5.4. It follows that  $\gamma$  has closure width  $w$  and takes the form  $\gamma = \chi[\tau_1] \cdots [\tau_w]$ . Certainly,  $t' \in L(\varphi_{\langle w \rangle})$ . Moreover,  $\Delta_\varphi(\gamma) = \Pi(\gamma, \varphi_{\langle w \rangle})$ . There exists therefore a unique  $\pi \in \Pi(\gamma, \varphi_{\langle w \rangle})$  such that  $t' \in L(\pi)$ . Given the fact that the head of  $\varphi_{\langle w \rangle}$  is equal to  $\varphi = \lambda(T[\uparrow(S)])$  we note that  $\pi$  must be of form  $\pi = \lambda(\alpha[\uparrow(\sigma)])[\sigma_1] \cdots [\sigma_w]$ . However, since  $t' = (\lambda a[\uparrow(s)])[s_1] \cdots [s_w] \in L(\pi)$  it also means that  $a \in L(\alpha)$ ,  $s \in L(\sigma)$ , and  $s_i \in L(\sigma_i)$  for all  $1 \leq i \leq w$ . Consequently,  $t \in L((\lambda\alpha)[\sigma][\sigma_1] \cdots [\sigma_w])$  as required.

Conversely, let  $\pi = \lambda(\alpha[\uparrow(\sigma)])[\sigma_1] \cdots [\sigma_w]$  be the unique term in the  $\varphi$ -matching family  $\Delta_\varphi^n$  such that  $t \in L((\lambda\alpha)[\sigma][\sigma_1] \cdots [\sigma_w])$ . Note that  $a \in L(\alpha)$ ,  $s \in L(\sigma)$ , and  $s_i \in L(\sigma_i)$  for all  $1 \leq i \leq w$ . Since  $t$  has a head  $v$ -redex, after a single reduction step  $t$  reduces to  $t' = \lambda(a[\uparrow(s)])[s_1] \cdots [s_w] \in L(\pi)$ . By construction of  $\Delta_\varphi^n$ , it means that there exists a production  $G_n \rightarrow \gamma$  in  $\mathcal{G}_n$  such that  $L(\pi) \subset L(\gamma)$  and hence  $t' \downarrow_n$ . Certainly, it follows that  $t \downarrow_{n+1}$ .  $\blacktriangleleft$

## C Quantitative analysis of $(\mathcal{G}_n)_n$

Techniques of analytic combinatorics provide systematic means of investigating various discrete structures through a direct analysis of their corresponding generating functions [12, 13]. Although the presented analysis is quite straightforward, it requires a fair amount of background knowledge. We refer the unfamiliar reader to Flajolet and Sedgewick's excellent textbook [13] for a thorough introduction to the subject.

**Proof of Proposition 6.4.** Let  $G_k \rightarrow \gamma$  be a regular production in  $\mathcal{G}_k$ . Since by construction  $\mathcal{G}_k$  is simple, we know that  $\gamma \in \mathcal{T}_{\mathcal{F}}(\mathcal{N} \cup \{G_0, \dots, G_{k-1}\})$ . Following symbolic methods [13, Part A, Symbolic Methods] we can therefore convert each non-terminal  $X \in \mathcal{N} \cup \{G_0, \dots, G_{k-1}\}$  occurring in  $\gamma$  into an appropriate generating function  $X(z)$ . Likewise, we can convert each function symbol occurrence  $f$  into an appropriate monomial  $z$ , see Figure 5. Finally, we group respective monomials together, and note that the generating function  $G_\gamma(z)$  corresponding to  $\gamma$  takes the form (15). Respective exponents denote the number of occurrences of their associated symbols.

Consider the remaining self-referencing productions  $G_k \rightarrow \delta$ . Again, since  $\mathcal{G}_k$  is simple, we know that  $\delta$  takes the form  $\lambda G_k$ ,  $G_0 G_k$  or (symmetrically)  $G_k G_0$ . And so, as each  $X \in \mathcal{N}_n$  is unambiguous in  $\mathcal{G}_k$ , by symbolic methods, it follows that  $G_k(z)$  satisfies the following functional equation:

$$G_k(z) = zG_k(z) + 2zG_0(z)G_k(z) + \sum_{G_k \rightarrow \gamma \in \Phi_k} G_\gamma(z). \quad (23)$$

Note that as no  $G_\gamma(z)$  references the left-hand side  $G_k(z)$ , equation (23) is in fact linear in  $G_k(z)$ . Furthermore, as  $G_0(z) = L_\infty(z)$  we finally obtain the requested form of  $G_k(z)$ , see (14).  $\blacktriangleleft$

Equipped with Proposition 6.4 we are now ready to prove the main quantitative result of the current paper.

**Proof of Theorem 6.5.** We claim that for each  $k \geq 1$  the generating function  $G_k(z)$  can be represented as  $G_k(z) = \sqrt{1 - 4z}P(z) + Q(z)$  where both  $P(z)$  and  $Q(z)$  are functions analytic in the disk  $|z| < \frac{1}{4} + \varepsilon$  for some positive  $\varepsilon$ . The asserted asymptotic estimate follows then as a straightforward application of algebraic singularity analysis, see Proposition 6.1.

We start with showing that each  $G_k(z)$  includes a summand in form of  $\sqrt{1 - 4z}\overline{P}(z) + \overline{Q}(z)$  such that both  $\overline{P}(z)$  and  $\overline{Q}(z)$  are analytic in a large enough disk containing (properly)  $|z| < \frac{1}{4}$ . Afterwards, we argue that no summand has singularities in  $|z| < \frac{1}{4}$ . Standard closure properties of analytic functions with single dominant, square-root type singularities guarantee the required representation of  $G_k(z)$ .

Let  $\varphi = N$  be the template corresponding to the (RVar) rule, see Table 1. Note that since  $\mathcal{G}_0$  includes the production  $G_0 \rightarrow N$ , the set of  $\varphi$ -matchings  $\Delta_\varphi^0$  consists of the single term  $N$ . Hence, due to the respective production construction, it means that  $G_1 \rightarrow (SN)[T/]$  is a production of  $\mathcal{G}_1$ , cf. Example 5.10. Moreover, as a consequence of the construction associated with the (FVar) rule, for each  $G_k \rightarrow \gamma$  in  $\mathcal{G}_k$  there exists a production  $G_{k+1} \rightarrow \underline{0}[\gamma/]$  in the subsequent grammar  $\mathcal{G}_{k+1}$ . And so, each  $\mathcal{G}_k$  includes among its productions one production in form of

$$G_k \rightarrow \underbrace{\underline{0}[\underline{0}[\dots\underline{0}[(SN)[T/]]\dots/]]}_{k-1 \text{ times}} \tag{24}$$

Denote the right-hand side of the above production as  $\gamma$ . Note that the associated generating function  $G_\gamma(z)$ , cf. (23), must therefore take form

$$G_\gamma(z) = z^{\zeta(\gamma)}T(z)N(z) \quad \text{and so} \quad G_\gamma(z) = \sqrt{1 - 4z}\overline{P}(z) + \overline{Q}(z) \tag{25}$$

where both  $P(z)$  and  $Q(z)$  are analytic in  $|z| < \frac{1}{4} + \varepsilon$  for some (determined)  $\varepsilon > 0$ .

In order to show that no production admits a corresponding generating function with singularities in the disk  $|z| < \frac{1}{4}$  we note that the single dominant singularity of  $G_0(z)$ , and so at the same time  $L_\infty(z)$ , is equal to the smallest positive real root  $\rho$  of  $1 - 3z - z^2 - z^3$  which satisfies  $\frac{1}{4} < \rho \approx 0.295598$ , see Proposition 6.2. Due to the form of basic generating functions corresponding to  $T$ ,  $S$  and  $N$ , see Proposition 6.3, we further note that other singularities must lie on the unit circle  $|z| = 1$ . And so, each  $G_k(z)$  admits the asserted form

$$G_k(z) = \sqrt{1 - 4z}P(z) + Q(z) \tag{26}$$

for some functions analytic in a disk  $|z| < \frac{1}{4} + \varepsilon$ . ◀