International Journal of Foundations of Computer Science
© World Scientific Publishing Company

Simplifications of Uniform Expressions Specified by Systems

Florent Koechlin
LIGM, Univ Gustave Eiffel, 5 boulevard Descartes,
Champs-sur-Marne, 777454, France
florent.koechlin@u-pem.fr

Cyril Nicaud
LIGM, Univ Gustave Eiffel, 5 boulevard Descartes,
Champs-sur-Marne, 777454, France
cyril.nicaud@u-pem.fr

Pablo Rotondo
LITIS, Univ Rouen Normandie, 685 avenue de l'université,
Saint-Étienne-du-Rouvray, 76800, France
pablo.rotondo@univ-rouen.fr

Received (Day Month Year)
Accepted (Day Month Year)
Communicated by (xxxxxxxxxx)

In this article, we study the impact of applying simple reduction rules to random syntactic formulas encoded as trees. We assume that there is an operator that has an absorbing pattern and prove that if we use this property to simplify a uniform random expression with n nodes, then the expected size of the result is bounded by a constant. The same holds for higher moments, establishing the lack of expressivity of uniform random expressions. Our framework is quite general as we consider expressions defined by systems of combinatorial equations.

Keywords: Random expressions; simplification of expressions; analytic combinatorics.

1. Introduction

This article is the full version of the extended abstract [13]. It is the sequel of the work started in [12], where we investigate the lack of expressivity of uniform random expressions. In our settings, we use the natural encoding of expressions as trees, which is a convenient way to manipulate them both in theory and in practice. In particular, it allows us to treat many different kinds of expressions at a general level (see Fig. 1 below): regular expressions, arithmetic expressions, boolean formulas, LTL formulas, . . .

Under this encoding, some classical questions are solved using a simple traversal of the tree, e.g. testing whether the language of a regular expression contains the empty word, or formally differentiating a function. Sometimes however, the tree is

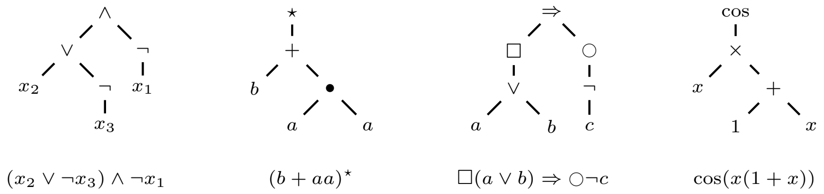

$(x_2 \vee \neg x_3) \wedge \neg x_1$
$(b + aa)^\star$
$\Box(a \vee b) \Rightarrow \bigcirc\neg c$
$\cos(x(1+x))$

Fig. 1. Four expression trees and their associated formulas. From left to right: a logical formula, a regular expression, an LTL formula and a function.

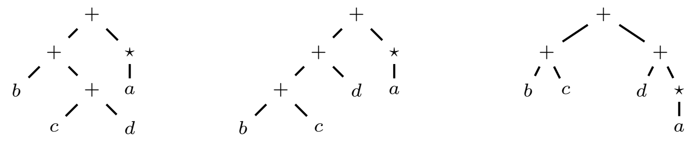
not the best choice, and it is first transformed into an equivalent adequate structure; in the context of formal languages, a regular expression (encoded using a tree) is typically transformed into an automaton, using one of the many known algorithms such as Thompson's construction or Glushkov automaton.

In our settings, we assume that one wants to estimate the efficiency of an algorithm, or a tool, whose inputs are expressions. The classical theoretical framework consists in analyzing the worst case complexity, but there are often some discrepancy between this measure of efficiency and what is observed in practice. A practical approach consists in using benchmarks to test the tool on real data. But in many contexts, having access to good benchmarks is quite difficult. Considering the average complexity of the algorithm is a classical alternative: it is sometimes amenable to a mathematical analysis, and it can be studied experimentally, provided there is a random generator at hand. Going that way, we have to choose a probability distribution on size-n inputs, which can be difficult: we want to study a "realistic" probability distribution that is also mathematically tractable. When no specific random model is available, it is classical to consider the uniform distribution, where all size-n inputs are equally likely. In many frameworks, such as sorting algorithms, studying the uniform distribution yields useful insights on the algorithms.

Following this idea, several works have been undertaken on uniform random expressions, in various contexts. Some are done at a general level: the expected height of a uniform random expression [15] always grows in $\Theta(\sqrt{n})$, if we identify common subexpressions then the expected size of the resulting acyclic graph [8] is in $\Theta(\frac{n}{\sqrt{\log n}})$, ... There are also more specific results on the expected size of the automaton built from a uniform random regular expression, using various algorithms [4,17]. In another setting, the expected cost of the computation of the derivative of a random function was proved to be in $\Theta(n^{3/2})$, both in time and space [9]. There are also many results on random boolean formulas, but the framework is a bit different, see Gardy's survey [10] for a more detailed account on this topic.

In [12], we questioned the model of uniform random expressions. Let us illustrate the main result of [12] on the example of regular expressions over the alphabet $\{a, b\}$. The set $\mathcal{L}_\mathcal{R}$ of regular expressions is inductively defined by

$$\mathcal{L}_\mathcal{R} = a + b + \varepsilon + \overset{\star}{\underset{\mathcal{L}_\mathcal{R}}{|}} + \overset{\bullet}{\underset{\mathcal{L}_\mathcal{R}\;\mathcal{L}_\mathcal{R}}{/\backslash}} + \overset{+}{\underset{\mathcal{L}_\mathcal{R}\;\mathcal{L}_\mathcal{R}}{/\backslash}}. \tag{$\star$}$$



$(b + (c + d)) + a^\star$ $((b + c) + d) + a^\star$ $(b + c) + (d + a^\star)$

Fig. 2. Regular expressions denoting the same language by associativity of the union.

The formula above is an equation on trees, where the size of a tree is its number of nodes. In particular a, b and ε represent trees of size 1, reduced to a leaf, labeled accordingly. As one can see from the specification ($\star$), leaves have labels in $\{a, b, \varepsilon\}$, unary nodes are labeled by $\star$ and binary nodes by either the concatenation $\bullet$ or the union $+$. Observe that the regular expression $\mathcal{P}$ corresponding to $(a + b)^\star$ denotes the regular language $\{a, b\}^\star$ of all possible words. This language is absorbing for the union operation on regular languages. So if we start with a regular expression $\mathcal{R}$ (a tree), identify every occurrence of the pattern $\mathcal{P}$ (a subtree), then rewrite the tree (bottom-up) by using inductively the simplifications $\begin{smallmatrix} & + & \\ / & & \backslash \\ \mathcal{X} & & \mathcal{P} \end{smallmatrix} \to \mathcal{P}$ and $\begin{smallmatrix} & + & \\ / & & \backslash \\ \mathcal{P} & & \mathcal{X} \end{smallmatrix} \to \mathcal{P}$, this results in a *simplified tree* $\sigma(\mathcal{R})$ that denotes the same regular language. Of course, other simplifications could be considered, but we just focus on this particular one. The main theorem of [12] implies that if one takes uniformly at random a regular expression of size n and applies this simplification algorithm, then the expected size of the resulting equivalent expression tends to a constant! It means that the uniform distribution on regular expressions produces a degenerated distribution on regular languages. More generally, we proved that: *For every class of expressions that admits a specification similar to Eq.* ($\star$) *and such that there is an absorbing pattern for some of the operations, the expected size of the simplification of a uniform random expression of size n tends to a constant as n tends to infinity.*[a] This negative result is quite general, as most examples of expressions have an absorbing pattern: for instance $x \wedge \neg x$ is always `false`, and therefore absorbing for $\wedge$.

The statement of the main theorem of [12] is general, as it can be used to discard the uniform distribution for expressions defined inductively as in Eq. ($\star$). However it is limited to that kind of simple specifications. And if we take a closer look at regular expressions in $\mathcal{L}_\mathcal{R}$, we observe that nothing prevents, for instance, useless sequences of nested stars as in $(((a + bb)^\star)^\star)^\star$. It is natural to wonder whether the result of [12] still holds when we forbid two consecutive stars in the specification. We could also use the associativity of the union to prevent different representations of the same language, as depicted in Fig. 2, or many other properties, to try to reduce the redundancy at the combinatorial level.

[a] The starting point of our work is a very specific analysis of and/or formulas established in Nguyên Thê PhD's dissertation [16, Ch 4.4].

This is the question we investigate in this article: does the degeneracy phenomenon of [12] still hold for more advanced combinatorial specifications? More precisely, we now consider specifications made using a system of (inductive) combinatorial equations, instead of only one as in Eq. ($\star$). For instance, we can forbid consecutive stars using the combinatorial system:

$$\begin{cases} \mathcal{L}_{\mathcal{R}} = \overset{\star}{\underset{\mathcal{S}}{|}} + \mathcal{S}, \\ \mathcal{S} = a + b + \varepsilon + \underset{\mathcal{L}_{\mathcal{R}}\ \mathcal{L}_{\mathcal{R}}}{\overset{+}{\wedge}} + \underset{\mathcal{L}_{\mathcal{R}}\ \mathcal{L}_{\mathcal{R}}}{\overset{\bullet}{\wedge}}. \end{cases} \qquad (\star\star)$$

The associativity of the union (Fig. 2) can be taken into account by preventing the right child of any $+$-node from being also labeled by $+$. Clearly, systems cannot be used for forbidding intricated patterns, but they still greatly enrich the families of expressions we can deal with. Moreover that kind of systems, which has strong similarities with context-free grammars, is amenable to analytic techniques as we will see in the sequel; this was for instance used by Lee and Shallit to estimate the number of regular languages in [14].

Our contributions can be described as follows. We consider expressions defined by systems of combinatorial equations and establish a universal degeneracy result: if there is an absorbing pattern, then the expected reduced size of a uniform random expression of size n is upper bounded by a constant as n tends to infinity. The result holds for natural yet technical conditions on the system. Hence, even if we use the system to remove redundancy from the specification (e.g., by forbidding consecutive stars), uniform random expressions still lack expressivity. Technically, we once again rely on the framework of analytic combinatorics for our proofs. However, the generalization to systems induces two main difficulties: First, we are not dealing with the well-known *varieties of simple trees* anymore [7, VII.3], so we have to rely on much more advanced techniques of analytic combinatorics; this is detailed in Section 5. Second, some work is required on the specification itself, to identify suitable hypotheses for our theorem; for instance, it is easy from the specification to prevent the absorbing pattern from appearing as a subtree at all, in which case our statement does not hold anymore, since there is no simplification.

An extended abstract of this work appeared in the proceedings of DLT'20 [13]. This version includes all the technical proofs of our results, and naturally focuses on a detailed study of systems of combinatorial equations using complex analysis.

2. Basic Definitions

For a given positive integer n, $[n] = \{1, \dots, n\}$ denotes the set of the first n positive integers. If E is a finite set, $|E|$ denotes its cardinality.

A *combinatorial class* is a set $\mathcal{C}$ equipped with a *size function* $|\cdot|$ from $\mathcal{C}$ to $\mathbb{N}$ (the size of $C \in \mathcal{C}$ is $|C|$) such that for any $n \in \mathbb{N}$, the set $\mathcal{C}_n$ of size-n elements of $\mathcal{C}$ is finite. Let $C_n = |\mathcal{C}_n|$, the *generating series* $C(z)$ of $\mathcal{C}$ is the formal power series defined by $C(z) = \sum_{C \in \mathcal{C}} z^{|C|} = \sum_{n \geq 0} C_n z^n$. Generating series are tools of choice to study combinatorial objects. When their radius of convergence is not zero, they can be viewed as analytic function from $\mathbb{C}$ to $\mathbb{C}$, and very useful theorems have been developed in the field of analytic combinatorics [7] to, for instance, easily obtain an asymptotic equivalent to C_n. We rely on that kind of techniques in Section 5 to prove our main theorem.

If $C(z) = \sum_{n\geq 0} C_n z^n$ is a formal power series, let $[z^n]C(z)$ denote its n-th coefficient C_n. Let ξ be a *parameter* on the combinatorial class $\mathcal{C}$, that is, a mapping from $\mathcal{C}$ to $\mathbb{N}$. Typically, ξ stands for some statistic on the objects of $\mathcal{C}$: the number of cycles in a permutation, the number of leaves in a tree, ... We define the *bivariate generating series* $C(z,u)$ associated with $\mathcal{C}$ and ξ by: $C(z,u) = \sum_{C\in\mathcal{C}} z^{|C|}u^{\xi(C)} = \sum_{k,n\geq 0} C_{n,k} z^n u^k$, where $C_{n,k}$ is the number of size-n elements C of $\mathcal{C}$ such that $\xi(C) = k$. In particular, $C(z) = C(z,1)$. Bivariate generating series are useful to obtain information on ξ, such as its expectation or higher moments.[b] Indeed, if $\mathbb{E}_n[\xi]$ denotes the expectation of ξ for the uniform distribution on $\mathcal{C}_n$, i.e. where all the elements of size n are equally likely, a direct computation yields:

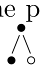
$$\mathbb{E}_n[\xi] = \frac{[z^n]\partial_u C(z,u)\big|_{u=1}}{[z^n]C(z)}, \tag{1}$$

where $\partial_u C(z,u)\big|_{u=1}$ consists in first differentiating $C(z,u)$ with respect to u, and then setting $u = 1$.

In the sequel, the combinatorial objects we study are trees, and we will have methods to compute the generating series directly from their specifications. Then, powerful theorems from analytic combinatorics will be used to estimate the expectation, using Eq. (1). So we delay the automatic construction and the analytic treatment to their respective sections.

3. Combinatorial Systems of Trees

3.1. *Definition of combinatorial expressions and of systems*

In the sequel the only combinatorial objects we consider are *plane trees*. These are trees embedded in the plane, which means that the order of the children matters: the two trees and  are different. Every node is labeled by an element in a set of symbols and the *size* of a tree is its number of nodes.

More formally, let S be a finite set, whose elements are *operator symbols*, and let a be a mapping from S to $\mathbb{N}$. The value $a(s)$ is called the *arity* of the operator s (we do not use the term *degree*, because if the tree is viewed as a graph, the degree of a node is its arity plus one, except for the root). An *expression over S* is a plane tree where each node of arity i is labeled by an element $s \in S$ such that $a(s) = i$, in particular, leaves symbols have arity 0.

Example 1. *In Fig. 1, the first tree is an expression over* $S = \{\wedge, \vee, \neg, x_1, x_2, x_3\}$ *with* $a(\wedge) = a(\vee) = 2$*,* $a(\neg) = 1$ *and* $a(x_1) = a(x_2) = a(x_3) = 0$*.*

[b] Recall that the j-th moment of a random variable ξ is by definition $\mathbb{E}[\xi^j]$ (when it exists). In particular, the first moment of ξ is its expectation.

An *incomplete expression over* S is an expression where (possibly) some leaves are labeled with a new symbol $\square$ of arity 0. Informally, such a tree represents part of an expression, where the $\square$-nodes need to be completed by being substituted by an expression. An incomplete expression with no $\square$-leaf is called a *complete* expression, or just an expression. If T is an incomplete expression over S, its *arity* $a(T)$ is its number of $\square$-leaves. It is consistent with the definition of the arity of a symbol, by viewing a symbol s of arity $a(s)$ as an incomplete expression made of a root labeled by s with $a(s)$ $\square$-children: $\wedge$ is viewed as $\overset{\wedge}{\underset{\square\ \square}{/\backslash}}$. Let $\mathcal{T}_\square(S)$ and $\mathcal{T}(S)$ be the set of incomplete and complete expressions over S. As incomplete expressions can be complete, we have $\mathcal{T}(S) \subseteq \mathcal{T}_\square(S)$.

If T is an incomplete expression over S of arity t, and $T_1, \ldots, T_t$ are expressions over S, we denote by $T[T_1, \ldots, T_t]$ the expression obtained by substituting the i-th $\square$-leaf in depth-first order by T_i, for $i \in [t]$. This notation is generalized to sets of expressions: if $\mathcal{T}_1, \ldots, \mathcal{T}_t$ are sets of expressions then $T[\mathcal{T}_1, \ldots, \mathcal{T}_t] = \{T[T_1, \ldots, T_t] : T_1 \in \mathcal{T}_1, \ldots, T_t \in \mathcal{T}_t\}$.

A *rule* of dimension $m \geq 1$ over S is an incomplete expression $T \in \mathcal{T}_\square(S)$ where each $\square$-node is labelled by an integer of $[m]$. Alternatively, a rule can be seen as a tuple $\mathcal{M} = (T, i_1, \ldots, i_t)$, where T is an incomplete expression of arity t and $i_1, \ldots, i_t$ are the values labelling its $\square$-leaves in depth-first order. The arity $a(\mathcal{M})$ of a rule $\mathcal{M}$ is the arity of its incomplete expression, and $\textsc{ind}(\mathcal{M}) = (i_1, \ldots, i_t)$ is the tuple of integer values obtained by a depth-first traversal of $\mathcal{M}$. A *combinatorial system of trees* $\mathcal{E} = \{E_1, \ldots, E_m\}$ of dimension m over S is a system of m class equations describing complete trees in $\mathcal{T}(S)$: each E_i is a non-empty finite set of rules over S, and the system in variables $\mathcal{L}_1, \ldots, \mathcal{L}_m$ is:

$$\begin{cases} \mathcal{L}_1 = \displaystyle\bigcup_{(T, i_1, \ldots, i_t) \in E_1} T[\mathcal{L}_{i_1}, \ldots, \mathcal{L}_{i_t}] \\ \quad \vdots \\ \mathcal{L}_m = \displaystyle\bigcup_{(T, i_1, \ldots, i_t) \in E_m} T[\mathcal{L}_{i_1}, \ldots, \mathcal{L}_{i_t}]. \end{cases} \tag{2}$$

Example 2. *To specify the system given in Eq.* $(\star\star)$ *using our formalism, we have* $m = 2$*. Its tuples representation is:* $E_1 = \left\{\left(\overset{\star}{\underset{\square}{|}}, 2\right), (\square, 2)\right\}$*, and* $E_2 = \left\{\left(\overset{\bullet}{\underset{\square\ \square}{/\backslash}}, 1, 1\right), \left(\overset{+}{\underset{\square\ \square}{/\backslash}}, 1, 1\right), (a), (b), (\varepsilon)\right\}$*, and its equivalent tree representation is* $E_1 = \left\{\overset{\star}{\underset{\boxed{2}}{|}}, \boxed{2}\right\}$*, and* $E_2 = \left\{\overset{\bullet}{\underset{\boxed{1}\ \boxed{1}}{/\backslash}}, \overset{+}{\underset{\boxed{1}\ \boxed{1}}{/\backslash}}, a, b, \varepsilon\right\}$*, which corresponds to Eq.* $(\star\star)$ *with* $\mathcal{L}_{\mathcal{R}} = \mathcal{L}_1$ *and* $\mathcal{S} = \mathcal{L}_2$*. In practice, we prefer descriptions as in Eq.* $(\star\star)$*, which are easier to read, but the tuple formalism is more convenient for the proofs.*

3.2. *Generating series*

If the system is not ambiguous, that is, if $\mathcal{L}_1, \ldots, \mathcal{L}_m$ is the[c] solution of the system and every tree in every $\mathcal{L}_i$ can be uniquely built from the specification, then the system can be directly translated into a system of equations on the generating series. This is a direct application of the *symbolic method* in analytic combinatorics [7, Part A] and we get the system

$$\begin{cases} L_1(z) = \displaystyle\sum_{(T,i_1,\ldots,i_{a(T)})\in E_1} z^{|T|} L_{i_1}(z)\cdots L_{i_{a(T)}}(z) \\ \quad\vdots \\ L_m(z) = \displaystyle\sum_{(T,i_1,\ldots,i_{a(T)})\in E_m} z^{|T|} L_{i_1}(z)\cdots L_{i_{a(T)}}(z). \end{cases} \tag{3}$$

where $L_i(z)$ is the generating series of $\mathcal{L}_i$. If the system is ambiguous, the $L_i(z)$'s still have a meaning: each expression of $\mathcal{L}_i$ accounts for the number of ways it can be derived from the system. When the system is unambiguous, there is only one way to derive each expression, and $L_i(z)$ is the generating series of $\mathcal{L}_i$.

3.3. *Designing practical combinatorial systems of trees*

Systems of trees such as Eq. (2) are not always well-founded. Sometimes they are, but still contain unnecessary equations. It is not the topic of this article to fully characterize when a system is correct, but we nonetheless need sufficient conditions to ensure that our results hold: in this section, we just present examples to underline some bad properties that might happen. For a more detailed account on combinatorial systems, the reader is referred to [1, 9, 20].

Ambiguity. As mentioned above, the system can be ambiguous, in which case the combinatorial system cannot directly be translated into a system of generating series. This is the case for the system $\left\{\mathcal{L}_1 = a + \overset{\star}{\underset{\mathcal{L}_1}{|}} + \overset{\star}{\underset{\mathcal{L}_2}{|}} ; \mathcal{L}_2 = \overset{\star}{\underset{\mathcal{L}_1}{|}} + a + b + \varepsilon\right\}$ as the expression $\overset{\star}{\underset{a}{|}}$ can be produced in two ways for the component $\mathcal{L}_1$.

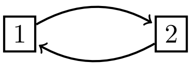
Empty components. Some specifications produce empty $\mathcal{L}_i$'s. For instance, consider the system $\left\{\mathcal{L}_1 = \overset{\bullet}{\underset{\mathcal{L}_1\ \mathcal{L}_2}{/\backslash}};\ \mathcal{L}_2 = a + b + \varepsilon + \mathcal{L}_1\right\}$: its only solution is $\mathcal{L}_1 = \emptyset$ and $\mathcal{L}_2 = \{a, b, \varepsilon\}$.

Cyclic unit-dependency. The *unit-dependency graph* $\mathcal{G}_\square(\mathcal{E})$ of a system $\mathcal{E}$ is the directed graph of vertex set $[m]$, with an edge $i \to j$ whenever $(\square, j) \in E_i$. Such a rule is called a *unit rule*. It means that $\mathcal{L}_i$ directly depends on $\mathcal{L}_j$. For instance $\mathcal{L}_\mathcal{R}$ directly depends on $\mathcal{S}$ in Eq. ($\star\star$). We can work with systems having unit

[c] In all generalities, there can be several solutions to a system, but the conditions we will add prevent this from happening.

dependencies, provided the unit-dependency graph is acyclic. If it is not, then the equations forming a cycle are useless or badly defined for our purposes. Consider for instance the system and its unit-dependency graph:

$$\begin{cases} \mathcal{L}_1 = \quad \mathcal{L}_2 + \overset{\star}{\underset{\mathcal{L}_1}{|}} \\ \mathcal{L}_2 = a + b + \varepsilon + \mathcal{L}_1 \end{cases}$$



The unit-dependency graph is not acyclic, and there are infinitely many ways to derive a from $\mathcal{L}_2$: $\mathcal{L}_2 \to a$, $\mathcal{L}_2 \to \mathcal{L}_1 \to \mathcal{L}_2 \to a\ldots$

Not strongly connected. The *dependency graph* $\mathcal{G}(\mathcal{E})$ of the system $\mathcal{E}$ is the directed graph of vertex set $[m]$, with an edge $i \to j$ whenever there is a rule $\mathcal{M} \in E_i$ such that $j \in \textsc{ind}(\mathcal{M})$: $\mathcal{L}_i$ depends on $\mathcal{L}_j$ in the specification. Some parts of the system may be unreachable from other parts, which may bring up difficulties. A sufficient condition to prevent this from happening is to ask for the dependency graph to be strongly connected; it is not necessary, but this restriction will also be useful in the proof of our main theorem (non-strongly connected systems are discussed in the conclusion). In Fig. 3 is depicted a system and its associated graph.

$$\begin{cases} \mathcal{L}_1 = \quad \overset{\star}{\underset{\mathcal{L}_2}{|}} + \overset{\star}{\underset{\mathcal{L}_3}{|}} \\ \mathcal{L}_2 = a + b + \varepsilon + \overset{\bullet}{\underset{\mathcal{L}_4\ \mathcal{L}_4}{\wedge}} \\ \mathcal{L}_3 = \quad \overset{+}{\underset{\mathcal{L}_4\ \mathcal{L}_1}{\wedge}} + \overset{+}{\underset{\mathcal{L}_4\ \mathcal{L}_2}{\wedge}} \\ \mathcal{L}_4 = \quad \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \end{cases}$$

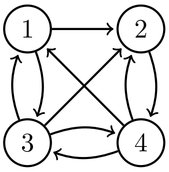


Fig. 3. A system and its associated dependency graph, which is strongly connected.

4. Settings, working hypothesis and simplifications

4.1. *Framework*

In this section, we describe our framework: we specify the kind of systems we are going to work with, and the settings for describing syntactic simplifications.

Let $\mathcal{E}$ be a combinatorial system of trees over S of dimension m of solution $(\mathcal{L}_1, \ldots, \mathcal{L}_m)$. A set of expressions $\mathcal{L}$ over S is *defined by* $\mathcal{E}$ if there exists a non-empty subset I of $[m]$ such that $\mathcal{L} = \cup_{i \in I} \mathcal{L}_i$.

From now on we assume that we are using a system $\mathcal{E}$ of dimension m over S and that S contains an operator $\circledast$ of arity at least 2. We furthermore assume that there is a complete expression $\mathcal{P}$, such that when interpreted, every expression of root $\circledast$ having $\mathcal{P}$ as a child is equivalent to $\mathcal{P}$: the interpretation of $\mathcal{P}$ is absorbing for the operator associated with $\circledast$. The expression $\mathcal{P}$ is the *absorbing pattern* and $\circledast$ is the *absorbing operator.*

Example 3. *Our main example is $\mathcal{L}$ defined by the system of Eq. ($\star\star$) with $\mathcal{L} = \mathcal{L}_\mathcal{R}$,* *the regular expressions with no two consecutive stars. As regular expressions, they are interpreted as regular languages. Since the language* $(a+b)^\star$ *is absorbing for the union, we set the associated expression as the absorbing pattern* $\mathcal{P}$ *and the operator symbol* $+$ *as the absorbing operator.*

The *simplification* of a complete expression T is the complete expression $\sigma(T)$ obtained by applying bottom-up the rewritting rule, where a is the arity of $\circledast$:

$$\begin{array}{c}\circledast\\ /\ \backslash\\ C_1\cdots C_a\end{array} \rightsquigarrow \mathcal{P}\,, \text{ whenever } C_i = \mathcal{P} \text{ for some } i \in \{1,\dots,a\}.$$

More formally, the simplification $\sigma(T, \mathcal{P}, \circledast)$ of T, or just $\sigma(T)$ when the context is clear, is inductively defined by: $\sigma(T) = T$ if T has size 1 and

$$\sigma((\oplus, C_1, \dots, C_d)) = \begin{cases} \mathcal{P} & \text{if } \oplus = \circledast \text{ and } \exists i, \sigma(C_i) = \mathcal{P}, \\ (\oplus, \sigma(C_1), \dots, \sigma(C_d)) & \text{otherwise.} \end{cases}$$

A complete expression T is *fully reducible* when $\sigma(T) = \mathcal{P}$.

We also need some conditions on the system $\mathcal{E}$. Some of them come from the discussion of Section 3.3, others are needed for the techniques from analytic combinatorics used in our proofs. A system $\mathcal{E}$ satisfies the hypothesis (**H**) when:

- ($\mathbf{H}_1$) The graph $\mathcal{G}(\mathcal{E})$ is strongly connected and $\mathcal{G}_\square(\mathcal{E})$ is acyclic.
- ($\mathbf{H}_2$) The system is *aperiodic*: there exists N such that for all $n \geq N$, there is at least one expression of size n in every coordinate of the solution $(\mathcal{L}_1, \dots, \mathcal{L}_m)$ of the system.
- ($\mathbf{H}_3$) For some j, there is a rule $T \in E_j$ of root $\circledast$, having at least two children T' and T'' such that: there is a way to produce a fully reducible expression from T' and $a(T'') \geq 1$.
- ($\mathbf{H}_4$) The system is *not linear*: there is a rule of arity at least 2.
- ($\mathbf{H}_5$) The system is *unambiguous*: each complete expression can be built in at most one way.

Conditions ($\mathbf{H}_1$) and ($\mathbf{H}_5$) were already discussed in Section 3.3. Condition ($\mathbf{H}_4$) prevents the system from generating only lists, i.e. trees whose internal nodes have arity 1, or more generally families that grow linearly as in $\mathcal{L} = \begin{array}{c}+\\ /\backslash\\ \mathcal{L}\ a\end{array} + b$, which are degenerated. Without Condition ($\mathbf{H}_3$) the system could be designed in a way that prevents simplifications, in which case our result does not hold, of course. Finally, Conditions ($\mathbf{H}_1$) and ($\mathbf{H}_2$) are necessary to keep the analysis manageable.

4.2. *Proper systems and system iteration*

A combinatorial system of trees $\mathcal{E}$ is said to be *proper* if it contains no unit rules and when the $\square$-leaves of all its rules have depth one (they are children of a root). The goal of this section is to prepare $\mathcal{E}$ for analytic analysis, by making it proper. Indeed the lack of unit rules is crucial for applying Drmota's theorem, and the depth one □-leaves make it possible to specify more easily the reduction process by focusing only on the roots of the tree rules.

One key tool to transform a system into a proper one is the notion of *system iteration*, which consists in substituting in each rule simultaneously every □-leaf labelled by an integer i by all the rules of E_i. For instance, if we iterate once our recurring system $\{\mathcal{L}_1 = \begin{smallmatrix}\star\\|\\\mathcal{L}_2\end{smallmatrix} + \mathcal{L}_2;\ \mathcal{L}_2 = a + b + \varepsilon + \begin{smallmatrix}+\\/\backslash\\\mathcal{L}_1\ \mathcal{L}_1\end{smallmatrix} + \begin{smallmatrix}\bullet\\/\backslash\\\mathcal{L}_1\ \mathcal{L}_1\end{smallmatrix}\}$, we get[d]

$$\begin{cases}\mathcal{L}_1 = \begin{smallmatrix}\star\\|\\a\end{smallmatrix} + \begin{smallmatrix}\star\\|\\b\end{smallmatrix} + \begin{smallmatrix}\star\\|\\\varepsilon\end{smallmatrix} + \begin{smallmatrix}\star\\|\\+\\/\backslash\\\mathcal{L}_1\ \mathcal{L}_1\end{smallmatrix} + \begin{smallmatrix}\star\\|\\\bullet\\/\backslash\\\mathcal{L}_1\ \mathcal{L}_1\end{smallmatrix} + a + b + \varepsilon + \begin{smallmatrix}+\\/\backslash\\\mathcal{L}_1\ \mathcal{L}_1\end{smallmatrix} + \begin{smallmatrix}\bullet\\/\backslash\\\mathcal{L}_1\ \mathcal{L}_1\end{smallmatrix}\\[2em] \mathcal{L}_2 = a + b + \varepsilon + \begin{smallmatrix}+\\/\backslash\\\mathcal{L}_2\ \mathcal{L}_2\end{smallmatrix} + \begin{smallmatrix}&+&\\/&&\backslash\\\star&&\mathcal{L}_2\\|&&\\\mathcal{L}_2&&\end{smallmatrix} + \begin{smallmatrix}&+&\\/&&\backslash\\\mathcal{L}_2&&\star\\&&|\\&&\mathcal{L}_2\end{smallmatrix} + \begin{smallmatrix}&+&\\/&&\backslash\\\star&&\star\\|&&|\\\mathcal{L}_2&&\mathcal{L}_2\end{smallmatrix} + \begin{smallmatrix}\bullet\\/\backslash\\\mathcal{L}_2\ \mathcal{L}_2\end{smallmatrix} + \begin{smallmatrix}&\bullet&\\/&&\backslash\\\star&&\mathcal{L}_2\\|&&\\\mathcal{L}_2&&\end{smallmatrix} + \begin{smallmatrix}&\bullet&\\/&&\backslash\\\mathcal{L}_2&&\star\\&&|\\&&\mathcal{L}_2\end{smallmatrix} + \begin{smallmatrix}&\bullet&\\/&&\backslash\\\star&&\star\\|&&|\\\mathcal{L}_2&&\mathcal{L}_2\end{smallmatrix}.\end{cases}$$

Formally, if we iterate $\mathcal{E} = \{E_1, \dots, E_m\}$ once, then for all $i \in [m]$ we have

$$\mathcal{L}_i = \bigcup_{(T,i_1,\dots,i_t)\in E_1} T\left[\bigcup_{(T_1,\boldsymbol{j_1})\in E_{i_1}} T_1[\mathcal{L}_{j_{1,1}}, \dots, \mathcal{L}_{j_{1,t_1}}], \dots, \bigcup_{(T_t,\boldsymbol{j_t})\in E_{i_t}} T_t[\mathcal{L}_{j_{t,1}}, \dots, \mathcal{L}_{j_{t,t_t}}]\right]$$

where $\boldsymbol{j_1} = (j_{1,1}, \dots, j_{1,t_1}), \dots, \boldsymbol{j_t} = (j_{t,1}, \dots, j_{t,t_t})$.
Let $\mathcal{E}^2$ denote the system obtained after iterating $\mathcal{E}$ once; it is called the *system of order 2* (from $\mathcal{E}$). More generally $\mathcal{E}^t$ is the system of order t obtained by iterating $t-1$ times the system $\mathcal{E}$. From the definition we directly get:

Lemma 4. *If $\mathcal{L}$ is defined by a system $\mathcal{E}$, it is also defined by all its iterates $\mathcal{E}^t$. Moreover, if $\mathcal{E}$ satisfies* (**H**)*, every $\mathcal{E}^t$ also satisfies* (**H**)*, except that $\mathcal{G}(\mathcal{E}^t)$ may not be strongly connected.*

In order to transform $\mathcal{E}$ into a proper system, we proceed in two steps: we first prove that the unit rules can be removed in Lemma 5, then explain how to bring up the □-leaves in Lemma 6.

Lemma 5 (elimination of unit rules) *If the system $\mathcal{E}$ verifies* (**H**)*, then there is an iteration $\mathcal{E}'$ of $\mathcal{E}$, defining the same expression trees, that still verifies* (**H**) *and such that there is no unit rule in any $E' \in \mathcal{E}'$.*

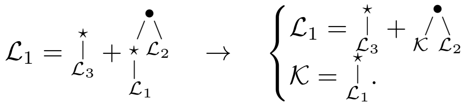
Proof. By Lemma 4, all (**H**)-conditions are still satisfied after iteration except, possibly, the strong connectedness. We want to prove that there is an iteration of $\mathcal{E}$ that also preserves strong connectedness. Observe that the edges of the dependency graph of the system $\mathcal{E}^k$ correspond to paths of length k in the original dependency

[d] Observe that the iterated system is not strongly connected anymore. It also yields two ways of defining the set of expressions using only one equation: it is very specific to this example, no such property holds in general.

graph $G(\mathcal{E})$. Therefore, as $\mathcal{G}_{\Box}(\mathcal{E})$ is acyclic by $(\mathbf{H}_1)$, if k is larger than the number of set equations m, then the iterate $\mathcal{E}^k$ does not contain unit rules anymore.

Let $N > m$ be an integer that is coprime with the lengths of the elementary cycles of $\mathcal{G}(\mathcal{E})$. As explained above, $\mathcal{E}^N$ does not contain any unit rule since $N > m$. We are going to prove that the graph $\mathcal{G}(\mathcal{E}^N)$ associated with the system iterated N times is strongly connected. Consider two indices $i, j \in [m]$. By the strong-connectedness of $\mathcal{G}(\mathcal{E})$, both of them belong to an elementary cycle. If i and j belong to the same elementary cycle $\mathcal{C}$ of length ℓ, they will still be in the same cycle in $\mathcal{G}^N$ iterations because N is coprime with ℓ. Suppose now that i and j belong to different elementary cycles $\mathcal{C}_i$ and $\mathcal{C}_j$. Using the previous statement, we just have to prove that there is a path from i to $\mathcal{C}_j$ in $\mathcal{G}(\mathcal{E}^N)$. The original graph $\mathcal{G}(\mathcal{E})$ is strongly connected, thus there is a simple path connecting i to j with less than m edges. Let t be the length of this simple path. Since $N > m > t$, once we arrive at j we can continue for $N - t$ steps in the elementary cycle $\mathcal{C}_j$. This means that i has an edge to some node of $\mathcal{C}_j$ in the graph $\mathcal{G}(\mathcal{E}^N)$, concluding the proof. □

The following lemma explains how a system $\mathcal{E}$ can be transformed into a new one $\mathcal{E}'$ that has all its □-leaves at depth exactly 1. The idea is to decrease the height of the □-leaves by splitting the rules into smaller ones, like in this example:



For convenience, we choose for this transformation to describe the system by its tree representation: every tuple $(T, i_1, \ldots, i_t) \in E_j$ is represented as a single tree T, where the k-th □-leaf has been labelled by the number i_k, where k corresponds to the order of the leaf in depth-first order. This labelled leaf is denoted by $\boxed{i_k}$ in the tree. It allows us to manipulate the rules derived from subtrees of the elements of E_j more easily, without having to explicitly give the indices. With this notation, we recall that $\textsc{ind}(T) = (i_1, \ldots, i_t)$ denotes the tuple of label-leaves $i_k \in [m]$ obtained when reading T in depth-first order.

Lemma 6 (bringing up the squares) *Every system that satisfies* $(\mathbf{H})$ *is equivalent to a system that satisfies* $(\mathbf{H})$ *and where every* □*-leaf is at depth* 1.

Proof. By Lemma 5, we assume that the system contains no unit rule. We define a new system $\mathcal{E}'$ from our original one $\mathcal{E}$ as follows: let $T \in E_j(\mathcal{E})$, for some $j \in [m]$, be a rule with a □-leaf at depth more than 1. Let $T_1, \ldots, T_s$ be the children of the root of T that have arity at least 1 and that are not □-leaves. Let $\tilde{T}$ be the tree obtained from T by replacing T_i by the new □-leaf $\boxed{m+i}$. The new system is obtained from $\mathcal{E}$ by replacing T by $\tilde{T}$ in E_j, and adding $E_{m+i}(\mathcal{E}') = \{T_i\}$. Note that $\mathcal{E}'$ has $m + s$ components now, and that we do not introduce unit rules.

This process is repeated as long as there are rules T having children of arity at least 1 that are not □-leaves: the process halts since the sum of the depths (summed over all rules) of the □-leaves that are at depth strictly greater than 1 decreases at each iteration. After each repetition, the number of equations increases. However, the first m equations still describe the languages $\mathcal{L}_1, \ldots, \mathcal{L}_m$, by direct induction.

Observe also that $(\mathbf{H})$ is clearly satisfied after each iteration. This concludes the proof. □

As the construction of Lemma 6 does not introduce any unit-rule, the following proposition is a direct consequence of applying Lemma 5 then Lemma 6.

Proposition 7. *If $\mathcal{L}$ is defined by a system $\mathcal{E}$ that satisfies $(\mathbf{H})$, then there exists a proper system $\mathcal{E}'$ that satisfies $(\mathbf{H})$ such that $\mathcal{L}$ is defined by $\mathcal{E}'$.*

5. Main result

Our main result establishes the degeneracy of uniform random expressions when there is an absorbing pattern, in our framework:

Theorem 8. *Let $\mathcal{E}$ be a combinatorial system of trees over S, of absorbing operator $\circledast$ and of absorbing pattern $\mathcal{P}$, that satisfies $(\mathbf{H})$. If $\mathcal{L}$ is defined by $\mathcal{E}$ then there exists a constant $C > 0$ such that, for the uniform distribution on size-n expressions in $\mathcal{L}$, the expected size of the simplification of a random expression is at most C.*

The remainder of this section is devoted to the proof of Theorem 8. Thanks to Proposition 7, we can assume that $\mathcal{E}$ is a proper system. By Condition $(\mathbf{H}_5)$, it is unambiguous so we can directly obtain a system of equations for the associated generating series, as explained in Section 3.2. From now on, for readability and succinctness, we use the vector notation (with bold characters): $\boldsymbol{L}(z)$ denotes the vector $(L_1(z), \ldots, L_m(z))$, and we rewrite the system of Eq. (3) in the more compact form

$$\boldsymbol{L}(z) = z\,\boldsymbol{\phi}(z; \boldsymbol{L}(z)), \tag{4}$$

where $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_m)$ and $\phi_i(z; \boldsymbol{y}) = \sum_{(T, i_1, \ldots, i_{a(T)}) \in E_i} z^{|T|-1} \prod_{j=1}^{a(T)} y_{i_j}$. Note that the factor z in Eq. (4) corresponds to counting the root of each rule (there is always a root since unit rules have been eliminated).

For the proof we rely on Eq. (1) to estimate the asymptotic expected size of the trees after reduction, by introducing the bivariate generating series $\boldsymbol{L}(z, u) = (L_1(z, u), \ldots, L_m(z, u))$ associated with the size of the simplified expression. The study of the denominator is simpler and can be found in Section 5.1. Then, the study of the numerator, which is the central part of the proof, is done in Section 5.2.

5.1. *Analysis of the denominator*

Proposition 9 below describes the number of trees of each type $\mathcal{L}_j$. As announced, this corresponds to the denominator of Eq. (1). Here $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](z; \boldsymbol{y})$ is the Jacobian matrix of the system, which is the $m \times m$ matrix such that $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](z;\boldsymbol{y})_{i,j} = \partial_{y_j}\phi_i(z;\boldsymbol{y})$. As usual, it plays a central role in the analysis of systems.

We will use several times a classical result from Analytic Combinatorics to obtain the asymptotics of the coefficients of generating series: the *Transfer Theorem* [7, Ch VI.3]. This theorem states that, if $L(z)$ of dominant singularity ρ verifies some analytic conditions, and $L(z) \sim_{z\to\rho} \lambda(1-z/\rho)^{-\alpha}$ with $\alpha \notin \{0,-1,-2,\ldots\}$, then $[z^n]L(z) \sim \lambda\rho^{-n}n^{\alpha-1}/\Gamma(\alpha)$, where Γ is Euler's gamma-function, the generalization of the factorial. In our case, we will mainly apply this Theorem with $\alpha = -1/2$.

Proposition 9. *As $\mathcal{E}$ satisfies* $(\mathbf{H})$*, the solution $\boldsymbol{L}(z)$ of the system of equations* (4) *is such that all its coordinates $L_j(z)$ share the same dominant singularity $\rho \in (0,1]$, and we have $\tau_j := L_j(\rho) < \infty$. The singularity ρ and $\boldsymbol{\tau} = (\tau_j)_j$ verify the characteristic system $\{\boldsymbol{\tau} = \rho\,\boldsymbol{\phi}(\rho;\boldsymbol{\tau}), \det(\mathtt{Id}_{m\times m} - \rho\,\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](\rho;\boldsymbol{\tau})) = 0\}$. Moreover, for every j, there exist two functions $g_j(z)$ and $h_j(z)$, analytic at $z=\rho$, such that locally around $z=\rho$, with $z \notin [\rho,+\infty)$,*

$$L_j(z) = g_j(z) - h_j(z)\sqrt{1-z/\rho}\,, \quad \text{with } h_j(\rho) \neq 0\,.$$

Lastly, we have the asymptotics $[z^n]L_j(z) \sim C_j\rho^{-n}/n^{3/2}$ for some positive C_j.

Proof. We just need to check that the hypotheses of Drmota's multidimensional theorem [6, Theorem 2.33] are verified.

Following the notation of Drmota [6], we rewrite the system $[L_1,\ldots,L_m] = \mathbf{f}(z;L_1,\ldots,L_m)$, where $\mathbf{f}(z;y_1,\ldots,y_m) := z\,\boldsymbol{\phi}(z;y_1,\ldots,y_m)$.

$\mathbf{f}(z,\mathbf{y})$ is analytic at $(z,\mathbf{y}) = (0,\mathbf{0})$, since each component is a polynomial, and $\mathbf{f}(0,\mathbf{0}) \equiv \mathbf{0}$. The dependency graph of $\mathbf{f}$ in $\mathbf{y}$ is strongly connected by $(\mathbf{H}_1)$. The Taylor coefficients of $\mathbf{f}$ are non negative, $\mathbf{f}(0,\mathbf{y}) \equiv 0$ as we have z as factor, $\mathbf{f}(z,\mathbf{0}) \not\equiv 0$ since there are leaves in at least one equation, and the system is not linear due to $(\mathbf{H}_4)$. Finally the Taylor coefficients $L_{i,n} = [z^n]L_i(z)$ of the solutions of the system are non-zero for n big enough by $(\mathbf{H}_2)$, so the $L_i(z)$'s are aperiodic.

Then by Drmota's theorem (Theorem 2.33 in [6], refined in [3]), all the L_j's have a same unique singularity ρ at their common convergence radius (by Pringsheim's Theorem [7, Theorem IV.6]), with $0 < \rho \le 1$, and $L_j(\rho) = \tau_j < \infty$. As $\mathbf{f}$ is a polynomial, $(\rho,\boldsymbol{\tau})$ is a characteristic point lying inside the radius of convergence of $\mathbf{f}$, so that $\boldsymbol{\tau} = \rho\,\boldsymbol{\phi}(\rho;\boldsymbol{\tau})$ and $0 = \det(\mathtt{Id}_{m\times m} - \rho\,\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](\rho;\boldsymbol{\tau}))$.

Finally for every j, there exist two functions $g_j(z)$ and $h_j(z)$ that are analytic around $z=\rho$ and that satisfy $L_j(z) = g_j(z) - h_j(z)\sqrt{1-z/\rho}$ locally around $z=\rho$, with $z \notin [\rho,+\infty)$ and $h_j(\rho) \neq 0$. The Transfer Theorem then yields the asymptotics, concluding the proof. □

5.2. *Analysis of the numerator*

In this section, instead of an asymptotic equivalent we just establish an upper bound for the numerator: for each $j \in [m]$, $[z^n]\partial_u L_j(z,u)|_{u=1} \le \alpha\rho^{-n}n^{-3/2}$, for some positive α. This is sufficient to prove our main theorem. We proceed in two steps, corresponding to Propositions 10 and 11 below. The formal proofs can be found in Section 6; we just give an overview of the main ideas for now.

First we split the system $\boldsymbol{\phi}$ into $\boldsymbol{\phi} = \underline{\boldsymbol{\phi}} + \boldsymbol{A} + \boldsymbol{B}$ where: $\underline{\boldsymbol{\phi}}$ corresponds to the rules of $\boldsymbol{\phi}$ whose root is not $\circledast$ and $\boldsymbol{B}$ gathers the rules of root $\circledast$ with a constant fully reducible child; if necessary, we iterate the system to ensure that $\boldsymbol{B}$ is not constant as a function of $\boldsymbol{y}$. This leads to a recursive equation defining $\boldsymbol{L}(z,u)$ in terms of two key classes: the fully reducible expressions and its complement. For example, note that the rules from $\boldsymbol{B}$ always produce fully reducible expressions. Differentiating this equation yields the Jacobian matrices for $\underline{\boldsymbol{\phi}}$ and $\boldsymbol{A}$. We then bound the generating series of the key classes with the whole series $\boldsymbol{L}(z)$, whose behaviour is well-known thanks to Proposition 9. We obtain the following bound:

Proposition 10. *For some $C > 0$, the following coordinate-wise bound holds:*

$$[z^n]\Big\{\partial_u \boldsymbol{L}(z,u)\big|_{u=1}\Big\} \leq C \cdot [z^n]\Big\{\big(\mathtt{Id}_{m\times m} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))\big)^{-1} \cdot \boldsymbol{L}(z)\Big\}.$$

Then we switch to the analysis of the right hand term in the inequality of Proposition 10. It looks complicated, but its dominant singularities are easier to study than those of $\partial_u \boldsymbol{L}|_{u=1}$. We do so by examining the spectrum of the matrix $J(z) = \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))$. Even if the matrix $(\mathtt{Id}_{m\times m} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](z;\boldsymbol{L}(z)))$ is not invertible at $z = \rho$ (see Lemma 19), the fact that $\underline{\boldsymbol{\phi}} + \boldsymbol{A} = \boldsymbol{\phi} - \boldsymbol{B}$, where the Jacobian of $\boldsymbol{B}$ is not $\boldsymbol{0}$, will imply that $\big(\mathtt{Id}_{m\times m} - z \cdot J(z)\big)$ is invertible at $z = \rho$ (see Lemma 20). Thus the behaviour at the dominant singularity comes purely from that of $\boldsymbol{L}(z)$, and not from the inverse of the matrix, as stated in the following proposition.

Proposition 11. *The dominant singularity of $\boldsymbol{F}\colon z \mapsto (\mathtt{Id} - z \cdot J(z))^{-1} \cdot \boldsymbol{L}(z)$ is $\rho = \rho_{\boldsymbol{L}}$. Furthermore, for every index j, around $z = \rho$ there exist analytic functions $\tilde{g}_j, \tilde{h}_j$ such that $F_j(z) = \tilde{g}_j(z) - \tilde{h}_j(z)\sqrt{1 - z/\rho}$ with $\tilde{h}_j(\rho) \neq 0$. Moreover, we have the asymptotics $[z^n]F_j(z) \sim D_j \rho^{-n} n^{-3/2}$, for some positive D_j.*

5.3. *Proof of the main theorem*

We use the results we just obtained on Eq. (1). By Propositions 10 and 11, we have

$$[z^n]\partial_u L_i(z,u)|_{u=1} = O\left(\frac{\rho^{-n}}{n^{3/2}}\right),$$

for every $i \in [m]$. By Proposition 9, we have the same asymptotics (with different multiplicative constants) for $[z^n]L_i(z) \sim C_i \rho^{-n}/n^{3/2}$. This concludes the proof of Theorem 8.

6. Proof of the two technical propositions

6.1. *Proof of Proposition 10*

6.1.1. *Symbolic part: bivariate generating series*

To better understand the process of reduction, we need to identify several relevant subclasses of trees. First, for every $j \in [m]$, we separate each $\mathcal{L}_j$ into a disjoint union of two classes $\mathcal{R}_j := \{T \in \mathcal{L}_j : \sigma(T) = \mathcal{P}\}$ and $\mathcal{G}_j := \mathcal{L}_j \setminus \mathcal{R}_j$, where $\mathcal{R}_j$ corresponds to the trees $T \in \mathcal{L}_j$ that reduce completely to $\mathcal{P}$. An expression of $\mathcal{R}_j$ is *fully reducible*.

To prove our theorem, we use bivariate generating series, for all $i \in [m]$:

$$L_i(z,u) = \sum_{L \in \mathcal{L}} z^{|L|} u^{|\sigma(L)|}\ ;\ R_i(z,u) = \sum_{R \in \mathcal{R}_i} z^{|R|} u^{|\sigma(R)|}\ ;\ G_i(z,u) = \sum_{G \in \mathcal{G}_i} z^{|G|} u^{|\sigma(G)|}\ .$$

Let $\boldsymbol{L}(z,u)$, $\boldsymbol{R}(z,u)$, and $\boldsymbol{G}(z,u)$ be the corresponding vectors of generating series.

To construct these classes and relate to the generating series, for every $j \in [m]$, we classify the tree rules from E_j into the partition $E_j = \underline{E}_j \uplus \mathcal{A}_j \uplus \mathcal{B}_j$, where:

- $\underline{E}_j$ contains all the trees of E_j whose root is not $\circledast$;
- $\mathcal{B}_j$ is the set of all $T \in E_j$ having $\circledast$ as root and such that one of the children of the root is a fully reducible complete tree T', i.e. $\sigma(T') = \mathcal{P}$;
- $\mathcal{A}_j$ is the set of all other elements of $T \in E_j$, their root is necessarily $\circledast$.

We introduce multivariate generating series for the decomposition of the rules, counting in u the size of the reduced *incomplete* tree (not counting the root). These generating series will actually be polynomials in $z, u, y_1, \dots, y_m$ as we have finitely many rules. For the class $\mathcal{A}_i$ we define:

$$A_i(z,u;y_1,\dots,y_m) = \sum_{T \in \mathcal{A}_i} z^{|T|-1} u^{|\sigma(T)|-1} \prod_{j \in \mathtt{ind}(T)} y_j\ ,$$

where we have extended the reduction σ to non-complete tree rules by setting σ to be the identity over the $\square$-nodes, so that, in particular, for every tree $T \in \mathcal{B}_j$, $\sigma(T) = \mathcal{P}$. Similarly we define the polynomials $B_i(z,u;y_1,\dots,y_m)$ (resp. $\underline{\phi}_i(z,u;y_1,\dots,y_m)$) in the same manner by taking the sum over all $T \in \mathcal{A}_i$ (resp. all $T \in \underline{E}_i$).

Let also $\boldsymbol{A}(z,u;y_1,\dots,y_m)$, $\boldsymbol{B}$ and $\underline{\boldsymbol{\phi}}$ denote the corresponding vectors of generating series. Notice that plugging $u = 1$, we obtain the relation $\boldsymbol{\phi} = \underline{\boldsymbol{\phi}} + \boldsymbol{A} + \boldsymbol{B}$.

To get a formula for $\boldsymbol{L}(z,u)$ we apply the technique of marking, see our previous work [12] for an explanation of this technique for tree simplifications.

Lemma 12. *The following equation is satisfied by* $\boldsymbol{L}(z,u)$*:*

$$\boldsymbol{L}(z,u) = u^p\left(\boldsymbol{R}(z) - \boldsymbol{P}(z)\right) + zu\left(\underline{\boldsymbol{\phi}}(z,u;\boldsymbol{L}(z,u)) + \boldsymbol{A}(z,u;\ \boldsymbol{G}(z,u))\right)\ ,$$

where $\boldsymbol{P}(z) = (a_1 z^p, \dots, a_m z^p)$ *where* a_i *is either* 0 *or* 1.

Proof. We mark with u the nodes that are kept for the reduced tree $\sigma(T)$. To know if the simplification applies, we exploit our classification of the rules from E_j. Recall that the system is proper, so that every $\square$-leaf is at depth 1.

- For the trees in $T \in \underline{E}_j$ there is no possible reduction at the root, no matter the children. Hence their roots are still here after reduction and counts for zu, leading to the term $zu\underline{\phi}_j(z, u; \boldsymbol{L}(z, u))$.
- Since the $\square$-leaves are at depth 1, the only way to avoid a reduction at the root for the trees in $\mathcal{A}_j$ is by replacing every $\square$-leaf by a tree in $\boldsymbol{\mathcal{G}}$, thus we add a term $zu\boldsymbol{A}(z, u; \boldsymbol{G}(z, u))$.
- For the trees in $\mathcal{B}_j$, there is no way to avoid reduction at the root.

The trees that are fully reducible come from $\mathcal{B}_j$, and also from $\mathcal{A}_j$ (whenever one child is fully reducible). Together they contribute to a term $R_j(z, u) = u^p R_j(z)$ in the equation. Note that to avoid counting $\mathcal{P}$ twice, we have to be careful and subtract a term $u^p P_j(z)$, where $P_j(z) = z^p$ if $\mathcal{P}$ is constructible from $\underline{E}_j$ or $\mathcal{A}_j$, and $P_j(z) = 0$ otherwise. This yields the announced equation for $\boldsymbol{L}(z, u)$. □

6.1.2. *Coefficient-wise inequalities for the formal power series*

In the proof we produce inequalities for $\partial_u \boldsymbol{L}(z, u)\big|_{u=1}$ seen as a formal power series. We introduce the notation $\preceq$ to deal more conveniently with inequalities that work entry-wise for matrices, and coefficient-wise for formal power series.

Definition 13. *Given two formal power series $F(z) = \sum_{n\in\mathbb{N}} f_n z^n$ and $G(z) = \sum_{n\in\mathbb{N}} g_n z^n$ with real coefficients, we write $F(z) \preceq G(z)$ if and only if for every $n \geq 0$, we have $f_n \leq g_n$.*
We extend the notation to matrices of power series. Consider $\boldsymbol{M}(z) = (M_{i,j}(z))_{i\in[m],j\in[n]}$, $\boldsymbol{N}(z) = (N_{i,j}(z))_{i\in[m],j\in[n]}$ two $m \times n$ matrices of power series, we note $\boldsymbol{M} \preceq \boldsymbol{N}$ if for every pair $i, j \geq 0$, $M_{i,j}(z) \preceq N_{i,j}(z)$.

We summarize the properties of the relation $\preceq$ in the following lemma.

Lemma 14 (Toolbox) *The following properties hold:*

- *Let $0 \preceq F(z), G(z), H(z), J(z)$ four series with non-negative coefficients, such that $F(z) \preceq G(z)$ and $H(z) \preceq J(z)$. Then $F(z) + H(z) \preceq G(z) + J(z)$ and $F(z)H(z) \preceq G(z)J(z)$.*
- *let $p(z, y_1, \ldots, y_n)$ a polynomial in $z, y_1, \ldots, y_n$ with non-negative coefficients, and $0 \preceq F_i(z) \preceq G_i(z)$ for $i \in [n]$. Then $p(z, F_1(z), \ldots, F_n(z)) \preceq p(z, G_1(z), \ldots, G_n(z))$.*
- *let $\mathbf{0}_{m\times n} \preceq \boldsymbol{M}(z) \preceq \boldsymbol{N}(z)$, and $\mathbf{0}_{n\times 1} \preceq \boldsymbol{F}(z) \preceq \boldsymbol{G}(z)$, then $\mathbf{0}_{m\times 1} \preceq \boldsymbol{M}(z)\boldsymbol{F}(z) \preceq \boldsymbol{N}(z)\boldsymbol{G}(z)$.*

Lemma 15. *There is a constant $C > 0$ such that*

$$\partial_u \boldsymbol{L}\big|_{u=1} \preceq C \cdot \boldsymbol{L}(z) + z\,\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z; \boldsymbol{L}(z)) \cdot \partial_u \boldsymbol{L}\big|_{u=1}.$$

Proof. Differentiating the equation on $\boldsymbol{L}(z,u)$ given by Lemma 12, we have:

$$\begin{aligned}\partial_u \boldsymbol{L}\Big|_{u=1} = p\,(\boldsymbol{R}(z) - \boldsymbol{P}(z)) &+ z\,\big(\underline{\boldsymbol{\phi}}(z,1;\boldsymbol{L}(z)) + \boldsymbol{A}(z,1;\ \boldsymbol{G}(z))\big)\\ &+ z\,\big(\partial_u\underline{\boldsymbol{\phi}}(z,1;\boldsymbol{L}(z)) + \partial_u\boldsymbol{A}(z,1;\ \boldsymbol{G}(z))\big)\\ &+ z\,\big(\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}}](z,1;\boldsymbol{L}(z))\cdot\partial_u\boldsymbol{L}\Big|_{u=1} + \mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{A}](z,1;\boldsymbol{G}(z))\cdot\partial_u\boldsymbol{G}\Big|_{u=1}\big)\ .\end{aligned}$$

Observe that:

- the combinatorial definition implies that $\boldsymbol{R}(z) - \boldsymbol{P}(z) \preceq \boldsymbol{L}(z)$.
- As the components of $\boldsymbol{A}$ are polynomials with non-negative coefficients and $\boldsymbol{G}(z) \preceq \boldsymbol{L}(z)$, we have $\boldsymbol{A}(z,1;\boldsymbol{G}(z)) \preceq \boldsymbol{A}(z,1;\boldsymbol{L}(z))$ and hence:
$$\begin{aligned}&z\big(\underline{\boldsymbol{\phi}}(z,1;\boldsymbol{L}(z)) + \boldsymbol{A}(z,1;\ \boldsymbol{G}(z))\big)\\ \preceq\, &z\big(\underline{\boldsymbol{\phi}}(z,1;\boldsymbol{L}(z)) + \boldsymbol{A}(z,1;\ \boldsymbol{L}(z)) + \boldsymbol{B}(z,1;\ \boldsymbol{L}(z))\big) = \boldsymbol{L}(z)\,.\end{aligned}$$
Similarly, we have $\partial_u\boldsymbol{A}(z,1;\boldsymbol{G}(z)) \preceq \partial_u\boldsymbol{A}(z,1;\boldsymbol{L}(z))\,.$
- For the same reason, $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{A}](z,1;\boldsymbol{G}(z)) \preceq \mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{A}](z,1;\boldsymbol{L}(z))$
- Finally, $\partial_u\boldsymbol{G}|_{u=1} \preceq \partial_u\boldsymbol{L}|_{u=1}$.

Using these three inequalities in the equation for $\partial_u\boldsymbol{L}|_{u=1}$ above, we obtain

$$\begin{aligned}\partial_u\boldsymbol{L}\Big|_{u=1} \preceq (p+1)\boldsymbol{L}(z) + z\partial_u[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z,1;\boldsymbol{L}(z))&\\ + z\ \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z,1;\boldsymbol{L}(z))\cdot\partial_u\boldsymbol{L}\Big|_{u=1}\,.&\end{aligned}$$

As $\underline{\boldsymbol{\phi}} + \boldsymbol{A}$ consists of polynomial entries in u (also z and $\boldsymbol{y}$), we introduce d their maximum degree in u. Notice that $d+1$ corresponds to the maximum number of nodes coming from a rule reduced by σ. We observe that $z\partial_u[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z,1;\boldsymbol{L}(z)) \preceq d\cdot z[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z,1;\boldsymbol{L}(z))$, and in turn the right-hand side is less than $d\cdot\boldsymbol{L}(z)$, as a consequence of Eq. (4). Thus

$$\partial_u\boldsymbol{L}\Big|_{u=1} \preceq (p+1+d)\cdot\boldsymbol{L}(z) + z\ \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z,1;\boldsymbol{L}(z))\cdot\partial_u\boldsymbol{L}\Big|_{u=1}\,,$$

proving the result for $C = p+1+d$. □

Next, we would like to get rid of the factor $\partial_u\boldsymbol{L}|_{u=1}$ that appears on the right-hand side. We cannot really do this by subtracting, as some coefficients would then become negative, and the toolbox we use is for series with non-negative coefficients. Instead, we rely on the following lemma:

Lemma 16. *Let* $\boldsymbol{0}_{m\times 1} \preceq \boldsymbol{v}(z), \boldsymbol{b}(z)$, $\boldsymbol{0}_{m\times m} \preceq \boldsymbol{M}(z)$, *and* $\boldsymbol{v}(z) \preceq \boldsymbol{b}(z) + z\boldsymbol{M}(z)\boldsymbol{v}(z)$ *then we have the following inequality*

$$\boldsymbol{v}(z) \preceq (\mathtt{Id} - z\boldsymbol{M}(z))^{-1}\cdot\boldsymbol{b}(z)\,,$$

where $(\mathtt{Id} - z\boldsymbol{M}(z))^{-1} = \sum_{k\geq 0} z^k(\boldsymbol{M}(z))^k$.

Remark 17. *Note that the series* $\sum z^k(\boldsymbol{M}(z))^k$ *converges in the ring of formal power series, because the coefficients of degree* n *remain fixed after we have summed the first* $n+1$ *terms of the series.*

Proof. Applying the original inequality k times on the right-hand side we obtain

$$\boldsymbol{v}(z) \preceq \boldsymbol{b}(z) + z\boldsymbol{M}(z)\boldsymbol{b}(z) + \ldots + (z\boldsymbol{M}(z))^k\boldsymbol{b}(z) + (z\boldsymbol{M}(z))^{k+1}\boldsymbol{v}(z)\,,$$

As $[z^k]\left(z\boldsymbol{M}(z))^{k+1}\boldsymbol{v}(z)\right) = \mathbf{0}_{m\times 1}$, this means that

$$\begin{aligned}[z^k]\,(\boldsymbol{v}(z)) &\leq [z^k]\left(\boldsymbol{b}(z) + z\boldsymbol{M}(z)\boldsymbol{b}(z) + \ldots + (z\boldsymbol{M}(z))^k\boldsymbol{b}(z)\right)\\ &= [z^k]\left(\mathtt{Id} - z\boldsymbol{M}(z)\right)^{-1}\boldsymbol{b}(z),\end{aligned}$$

for every k, so by definition $\boldsymbol{v}(z) \preceq \left(\mathtt{Id} - z\boldsymbol{M}(z)\right)^{-1}\cdot\boldsymbol{b}(z)$. □

Lemma 15, together with Lemma 16, prove Proposition 10, namely the bound

$$[z^n]\Big\{\partial_u\boldsymbol{L}(z,u)\big|_{u=1}\Big\} \leq C\cdot[z^n]\Big\{\left(\mathtt{Id}_{m\times m} - z\cdot\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}}+\boldsymbol{A}](z;\boldsymbol{L}(z))\right)^{-1}\cdot\boldsymbol{L}(z)\Big\}\,.$$

6.2. *Proof of Proposition 11*

At this point we will leave the world of formal power series and show that we can interpret the right-hand side of the inequality in Proposition 10 in the world of analytic functions, showing that its coefficients have the same order of asymptotic growth as those of $\boldsymbol{L}(z)$. This will imply our main Theorem.

Proposition 9 characterizes the behaviour of $\boldsymbol{L}(z)$. We now describe the properties of the quasi-inverse $\boldsymbol{J}(z) = \left(\mathtt{Id} - z\cdot\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}}+\boldsymbol{A}](z;\boldsymbol{L}(z))\right)^{-1}$. We show that the dominant singularities of $\boldsymbol{J}(z)$ come from the singularities of $\boldsymbol{L}(z)$ and not from the inversion of the matrix. For this it is necessary to have $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{B}] \neq \mathbf{0}_{m\times m}$, which we may assume thanks to the following lemma.

Lemma 18. *We may suppose that* $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{B}] \neq \mathbf{0}_{m\times m}$*, i.e. at least one coefficient of the Jacobian matrix is not 0.*

Proof. By hypothesis $(\mathbf{H}_3)$, there is a rule T in some E_j whose root is $\circledast$, and such that T has two children T', T'' verifying that T' can generate a fully reducible tree T_R, and $a(T'') \geq 1$. By Lemma 5, iterating the system sufficiently many times leads to an equivalent system $\tilde{\mathcal{E}}$, with a rule of the form $\tilde{T} \in \tilde{E}_j$, where $\tilde{T}$ has among its children two trees T_1, T_2, with $T_1 = T_R$ fully reducible, and $a(T_2) \geq 1$ (by strong–connectedness). So this rule is in $\tilde{\mathcal{B}}_j$ and T_2 contains at least one □-leaf $\boxed{i}$, so that $\partial_{y_i}\tilde{B}_j$ is not 0.

Note that after iteration the □-leaves are not at depth 1 anymore, but this is easily fixed as in the proof of Lemma 6, whose construction preserves the property of $\tilde{\boldsymbol{B}}$. This concludes the proof. □

The radius of convergence of $\boldsymbol{J}(z)$ is related to the spectral radius $\mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}}+\boldsymbol{A}](\rho;\boldsymbol{L}(\rho))\right)$ of $\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}}+\boldsymbol{A}](\rho;\boldsymbol{L}(\rho))$, which we now study.

Lemma 19. *The spectral radius of the Jacobian* $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](\rho;\boldsymbol{L}(\rho))$ *is* $1/\rho$*, i.e.,*

$$\mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](\rho\,;\boldsymbol{L}(\rho))\right) = 1/\rho\,.$$

Proof. We must have $\mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](\rho\,;\boldsymbol{L}(\rho))\right) \geq 1/\rho$ in any case, because $1/\rho$ is an eigenvalue of the matrix, by the statement of Proposition 9 about the characteristic system. Assume for the sake of contradiction that we had the strict inequality. The function $f(t) := t \cdot \mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](t\,;\boldsymbol{L}(t))\right)$ is continuous [22], because the entries of the matrix are continuous (note that $\boldsymbol{L}(t)$ is continuous on $[0,\rho]$). Further, we have $f(0) = 0$, $f(\rho) > 1$. Thus there is $t_\rho \in (0,\rho)$ such that $f(t_\rho) = 1$. Thus the matrix $M = t_\rho \mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](t_\rho\,;\boldsymbol{L}(t_\rho))$ has spectral radius 1. The matrix M is non-negative and its underlying digraph is strongly-connected by our hypotheses on $\boldsymbol{L}$. Hence, by the celebrated Perron-Frobenius Theorem (see [11], Theorem 8.8.1), we conclude that M admits 1 as an eigenvalue.

Thus we have a characteristic point $(t_\rho, \boldsymbol{L}(t_\rho))$ for our original system

$$\begin{cases} \boldsymbol{L} &= z \cdot \boldsymbol{\phi}(z;\boldsymbol{L}) \\ 0 &= \det\left(\mathtt{Id} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](z;\boldsymbol{L})\right)\,, \end{cases}$$

with $t_\rho < \rho$ and $L_i(t_\rho) < L_i(\rho)$ for all i, in contradiction to Lemma 13 from [3] which characterizes the characteristic points that are different from $(\rho, \boldsymbol{L}(\rho))$. □

Lemma 20. *Provided that* $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{B}] \neq \boldsymbol{0}_{m\times m}$*, we have the strict inequality*

$$\mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](\rho;\boldsymbol{L}(\rho))\right) < 1/\rho\,.$$

Proof. The vector $\boldsymbol{L}(\rho)$ has every component strictly positive, thus the Jacobian $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{B}](\rho;\boldsymbol{L}(\rho))$ has some non-zero entry. Observe that we have the equality $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}] = \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}}] + \mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{A}] + \mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{B}]$. Hence entry-wise we have

$$\Big[\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](\rho\,;\boldsymbol{L}(\rho))\Big]_{i,j} \leq \Big[\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](\rho;\boldsymbol{L}(\rho))\Big]_{i,j}\,,$$

with strict inequality for at least one entry.

By Lemma 19, the spectral radius of the matrix $\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{\phi}](\rho;\boldsymbol{L}(\rho))$ is $1/\rho$. Furthermore, it is non-negative and its underlying directed graph is strongly-connected by our hypothesis on $\boldsymbol{L}$, hence it follows from a classical result in Algebraic Graph Theory ([11], Theorem 8.8.1 (b), page 178) that reducing at least one entry reduces the spectral radius strictly. □

Corollary 21. *At* $z = \rho$ *we have* $\det\Big(\mathtt{Id} - \rho \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](\rho;\boldsymbol{L}(\rho))\Big) \neq 0\,.$

Lemma 20 implies that we may interpret the right-hand side of our inequality in Proposition 10, not just as a formal power series, but also as an analytic function on $|z| < \rho$ which can be continuously extended at $|z| = \rho$.

Corollary 22. *The series*

$$\left(\mathtt{Id} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))\right)^{-1} = \sum_{k\geq 0} z^k \left(\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))\right)^k$$

defines an analytic function on the closed disk $|z| \leq \rho$*, with the only possible exception of* $z = \rho$ *where it can be extended continuously.*

Proof.

Consider w with $|w| \leq \rho$, $w \neq \rho$. We recall from Drmota's article [5, Lemma 1, page 14] that we also have that ρ is the only singularity of $\boldsymbol{L}(z)$ on the disk $|z| \leq \rho$. Thus $\boldsymbol{L}(z)$ is analytic on $z = w$, and defined (and continuous) in some neighborhood of it. We remark that, as the coefficients are non-negative and $|w| \leq \rho$, [e]

$$\mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](w; \boldsymbol{L}(w))\right) \leq \mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](\rho; \boldsymbol{L}(\rho))\right) < 1/\rho\,.$$

Thus by continuity of the function $z \mapsto |z| \cdot \mathtt{sp}\left(\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z; \boldsymbol{L}(z))\right)$ (see [22]), the spectral radius of $z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z; \boldsymbol{L}(z))$ remains strictly less than 1 on some neighbourhood of w. Hence the geometric series converges uniformly on a neighborhood of w, and our quasi-inverse is analytic there.

For the continuity at $z = \rho$ we notice that the geometric series converges uniformly for $|z| \leq \rho$ and that $\boldsymbol{L}(z)$ is continuous on the closed disc. □

To finish the proof we need to show that the inverse is closely related to $\boldsymbol{L}(z)$. We recall the following useful result from Matrix Theory:

Theorem 23 (Cayley-Hamilton) *Let M be a square matrix over a commutative ring. Its characteristic polynomial $p_M(\lambda)$ is defined by $p_M(\lambda) := \det(\lambda \cdot \mathtt{Id} - M)$. Then M cancels its characteristic polynomial: $p_M(M) = \mathbf{0}$.*

A direct corollary of this Theorem is a formula for the inverse of M when we place ourselves over a commutative field. Let $p_M(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \ldots + a_0$, we remark first that $a_0 = (-1)^m \det M$. Then we have

$$M^{-1} = (-1)^{m+1} \frac{M^{m-1} + a_{m-1}M^{m-2} + \ldots + a_1}{\det M}\,. \tag{5}$$

Moreover, the a_i's are polynomials in the entries of M.

Remark 24. *Plugging $M = \mathtt{Id} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z; \boldsymbol{L}(z))$ in Eq. (5), we notice that the coefficients a_i are polynomials in z and $\boldsymbol{L}(z)$, while M^i has entries which are also polynomial in z and $\boldsymbol{L}(z)$. Similarly for the determinant in the denominator of Eq. (5), the entries are also polynomial in z and $\boldsymbol{L}(z)$.*

Now we need the following very useful technical lemma to exploit the expansion of $\boldsymbol{L}(z)$ near $z = \rho$.

Lemma 25. *Let the generating series $A(z)$ and $B(z)$ and have the expansion valid around $z = \rho$*

$$A(z) = g_A(z) - h_A(z)\sqrt{1 - \tfrac{z}{\rho}}\,, \quad B(z) = g_B(z) - h_B(z)\sqrt{1 - \tfrac{z}{\rho}}\,,$$

where $g_A(z), g_B(z), h_A(z), h_B(z)$ are analytic at $z = \rho$. Then the following assertions hold

[e] The spectral radius increases as we increase the absolute value of the entries. This can be seen using Gelfand's formula $\mathtt{sp}(B) = \lim_p \|B^p\|^{1/p}$, see [21, pp.209–212].

(1) The sum $A(z) + B(z)$ and difference $A(z) - B(z)$ also have expansions $A(z)+B(z) = g_{A+B}(z) - h_{A+B}(z)\sqrt{1-z/\rho}$ and $A(z)-B(z) = g_{A-B}(z) - h_{A-B}(z)\sqrt{1-z/\rho}$, around $z = \rho$, with $g_{A+B}(z)$, $h_{A+B}(z)$, $g_{A-B}(z)$, $h_{A-B}(z)$ analytic at $z = \rho$.

(2) The product $A(z) \cdot B(z)$ also has a expansion $g_{A\cdot B}(z) - h_{A\cdot B}(z)\sqrt{1-z/\rho}$, around $z = \rho$, with $g_{A\cdot B}(z)$ and $h_{A\cdot B}(z)$ analytic analytic at $z = \rho$.

(3) Let $H(y)$ be analytic at $y = A(\rho)$ and suppose $H'(A(\rho)) \neq 0$, then the composition has a expansion $H(A(z)) = g_{H\circ A}(z) - h_{H\circ A}(z)\sqrt{1-z/\rho}$ around $z = \rho$, with $g_{H\circ A}$ and $h_{H\circ A}$ analytic at $z = \rho$.

(4) If we suppose that $B(\rho) \neq 0$, the quotient $A(z)/B(z)$ has a singular expansion $g_{A/B}(z) - h_{A/B}(z)\sqrt{1-z/\rho}$ around $z = \rho$, with $g_{A/B}(z)$ and $h_{A/B}(z)$ analytic at $z = \rho$.

Proof. Only the case of the division requires further proof as the sum, difference and product are immediate and (3) is a special case of a lemma from Drmota [6, Lemma 2.26].

As $g_B(\rho) \neq 0$ we may write

$$\frac{A(z)}{B(z)} = \frac{g_A(z) - h_A(z)\sqrt{1-\frac{z}{\rho}}}{g_B(z) - h_B(z)\sqrt{1-\frac{z}{\rho}}} = \frac{g_A(z)/g_B(\rho) - h_A(z)/g_B(\rho)\sqrt{1-\frac{z}{\rho}}}{1 + (g_B(z)/g_B(\rho) - 1) - h_B(z)/g_B(\rho)\sqrt{1-\frac{z}{\rho}}}\,.$$

The function $H(y) = \frac{1}{1+y} = 1 - y + y^2 - y^3 \pm \ldots$ is analytic for $|y| < 1$. Since $\delta\colon z \mapsto (g_B(z)/g_B(\rho) - 1) - h_B(z)/g_B(\rho)\sqrt{1-\frac{z}{\rho}}$ is 0 at $z = \rho$ the result follows because the composition $H(\delta(z))$ satisfies 3. and then we apply 2. with the numerator. □

Lemma 26. *The function $z \mapsto \big(\mathtt{Id} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))\big)^{-1}$ has entries that are locally of the form $\tilde{g}(z) - \tilde{h}(z)\sqrt{1-z/\rho}$ around $z = \rho$, with $\tilde{g}$ and $\tilde{h}$ analytic at $z = \rho$.*

Proof. Set $M = \mathtt{Id} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))$ in Eq. (5).

Observe that the coefficients a_i are polynomials in z and $\boldsymbol{L}(z)$, while M^i has entries which are polynomial too in z and $\boldsymbol{L}(z)$, and similarly for the determinant in the denominator of Eqn. (5). Hence it follows, by Proposition 9 and the closure properties of Lemma 25 that the numerator and the denominator of the entries of M^{-1} are of the form claimed. As the denominator is not 0 at $z = \rho$ by Corollary 21, the closure properties imply the result for the quotients. □

Proof of Proposition 11. We consider

$$\boldsymbol{F}(z) := \Big(\mathtt{Id} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))\Big)^{-1} \cdot \boldsymbol{L}(z)\,,$$

We claim that the radius of convergence of the components of $\boldsymbol{F}(z)$ must be ρ. Indeed, from Proposition 10 we deduce that

$$\boldsymbol{L}(z) \preceq \partial_u \boldsymbol{L}|_{u=1} \preceq C \cdot \boldsymbol{F}(z)\,.$$

where $C > 0$ is the constant from Proposition 10. Hence $[z^n]L_i(z) \le [z^n]C \cdot F_i(z)$ and the radius of convergence of F_i can only be smaller or equal to that of L_i, namely ρ. As we know that $\boldsymbol{F}(z)$ is analytic for $|z| < \rho$, due to Corollary 22, the radius of convergence must be ρ. By Pringsheim's Theorem [7], this means that ρ is a singularity of the entries of $\boldsymbol{F}$. And $z = \rho$ is the only possible singularity on the circle $|z| = \rho$, by Corollary 22.

By Lemma 26, along with the closure properties in Lemma 25, $\boldsymbol{F}(z)$ has entries $F_i(z)$ which are of the form $F_i(z) = \tilde{g}_i(z) - \tilde{h}_i(z)\sqrt{1 - z/\rho}$ around $z = \rho$, with $\tilde{g}_i(z)$ and $\tilde{h}_i(z)$ analytic at $z = \rho$. Moreover, we must have $\tilde{h}_i(\rho) \neq 0$ for all i, otherwise, by the Transfer Theorem $C \cdot [z^n]F_i(z)$ would be asymptotically negligible towards $[z^n]L_i(z)$, a contradiction.

Thus the Transfer Theorem yields that $[z^n]F_i(z) \sim D_i \frac{\rho^{-n}}{n^{3/2}}$, for $D_i > 0$. □

7. Higher moments: expected run-time of polynomial algorithms

Our main result can be extended to all the moments of the random variable corresponding to the size of the reduction.

Theorem 27. *Let $\mathcal{E}$ be a combinatorial system of trees over S, of absorbing operator $\circledast$ and of absorbing pattern $\mathcal{P}$, that satisfies* (**H**). *If $\mathcal{L}$ is defined by $\mathcal{E}$, for the uniform distribution on size-n expressions in $\mathcal{L}$, every moment of order t of the size of a reduced expression is bounded from above by a constant C_t.*

Proof. The proof follows the same principles as the proof of Theorem 8. It is based on an analogue of Eq (1): considering the bivariate generating series $C(z,u) = \sum_{C \in \mathcal{C}} z^{|C|} u^{\xi(C)}$, we have

$$\mathbb{E}_n\big[\xi^k\big] = \frac{[z^n](u\,\partial_u)^k C(z,u)\big|_{u=1}}{[z^n]C(z)}\,, \tag{6}$$

where $(u\,\partial_u)^k$ means that k times we: differentiate in u and then multiply by u. The study of the numerator proceeds by induction on k and is sketched below. Note that the base case, with $k = 1$, corresponds to the expected value.

Differentiating the equation for $L(z,u)$ once we find

$$\begin{aligned}\partial_u \boldsymbol{L}(z,u) = p u^{p-1}\left(\boldsymbol{R}(z) - \boldsymbol{P}(z)\right) &+ z\left(\underline{\boldsymbol{\phi}}(z,u;\boldsymbol{L}(z,u)) + \boldsymbol{A}(z,u;\ \boldsymbol{G}(z,u))\right)\\ &+ zu\left(\partial_u\underline{\boldsymbol{\phi}}(z,u;\boldsymbol{L}(z,u)) + \partial_u\boldsymbol{A}(z,u;\ \boldsymbol{G}(z,u))\right)\\ &+ zu\mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}}](z,u;\boldsymbol{L}(z,u)) \cdot \partial_u\boldsymbol{L}(z,u)\\ &+ zu\mathtt{Jac}_{\boldsymbol{y}}[\boldsymbol{A}](z,u;\boldsymbol{G}(z,u)) \cdot \partial_u\boldsymbol{G}(z,u)\,.\end{aligned}$$

We may generalize $\preceq$ to two-variable power series and as in Proposition 10 we get

$$\begin{aligned} u\,\partial_u \boldsymbol{L}(z,u) \preceq\ & p u^p \boldsymbol{L}(z) + \boldsymbol{L}(z,u) \\ & + z u^2 \left(\partial_u \underline{\boldsymbol{\phi}}(z,u;\boldsymbol{L}(z,u)) + \partial_u \boldsymbol{A}(z,u;\ \boldsymbol{L}(z,u))\right) \\ & + zu \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z,u;\boldsymbol{L}(z,u)) \cdot u\partial_u \boldsymbol{L}(z,u)\,. \end{aligned}$$

Observe that differentiating in u does not affect the inequalities $\preceq$ for the formal power series as the coefficients are positive. Thus we prove by induction that

$$\begin{aligned} (u\,\partial_u)^k \boldsymbol{L}(z,u) \preceq\ & \boldsymbol{p}_k\big(z,u,\boldsymbol{L}(z),\boldsymbol{L}(z,u),(u\partial_u)\boldsymbol{L}(z,u),\dots,(u\partial_u)^{k-1}\boldsymbol{L}(z,u)\big) \\ & + zu \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z,u;\boldsymbol{L}(z,u)) \cdot (u\partial_u)^k \boldsymbol{L}(z,u)\,, \end{aligned}$$

where the $\boldsymbol{p}_k$'s are polynomials in their entries with non-negative coefficients.

Now taking $u = 1$, which maintains the inequalities, we obtain

$$\begin{aligned} (u\,\partial_u)^k \boldsymbol{L}(z,u)\big|_{u=1} \preceq\ & \boldsymbol{p}_k\big(z,1,\boldsymbol{L}(z),\boldsymbol{L}(z),\dots,(u\partial_u)^{k-1}\boldsymbol{L}(z,u)\big|_{u=1}\big) \\ & + z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z)) \cdot (u\partial_u)^k \boldsymbol{L}(z,u)\big|_{u=1}\,, \end{aligned}$$

thus Lemma 16 implies that

$$\begin{aligned} & (u\,\partial_u)^k \boldsymbol{L}(z,u)\big|_{u=1} \preceq \\ & \quad \big(\mathtt{Id} - z \cdot \mathtt{Jac}_{\boldsymbol{y}}[\underline{\boldsymbol{\phi}} + \boldsymbol{A}](z;\boldsymbol{L}(z))\big)^{-1} \boldsymbol{p}_k\big(z,1,\boldsymbol{L}(z),\ \boldsymbol{L}(z),\dots,(u\partial_u)^{k-1}\boldsymbol{L}(z,u)\big|_{u=1}\big) \end{aligned}$$

The proof follows by induction. Indeed, once we have proved that the derivatives $(u\partial_u)^j \boldsymbol{L}(z,u)\big|_{u=1}$, $j < k$, are all bounded by functions having only ρ as a singularity on the circle $|z| = \rho$, and having the right local behaviour around $z = \rho$, the result then follows for $(u\,\partial_u)^k \boldsymbol{L}(z,u)$ by Proposition 11, which characterizes the function $(\mathtt{Id} - z \cdot J(z))^{-1}$, and the closure properties of Lemma 25. □

Besides its intrinsic mathematical interest, Theorem 27 yields a direct analysis of all polynomial-time algorithms for random expressions, as stated below.

Corollary 28. *Let $\mathcal{A}$ be a polynomial-time algorithm (in the worst case) whose inputs are expressions specified as in the statement of Theorem 27. If one first reduces the expression, which can be done in linear time, before applying $\mathcal{A}$, then the expected running time of applying $\mathcal{A}$ is bounded from above by a constant.*

8. Conclusion and discussion

To summarize our contributions in one sentence, we proved in this article that even if we use systems to specify them, uniform random expressions lack expressivity as they are drastically simplified as soon as there is an absorbing pattern. This confirms and extends our previous result [12], which holds for much more simple specifications only. It questions the relevance of uniform distributions in this context, both for experiments and for theoretical analysis.

Roughly speaking, the intuition behind the surprising power of this simple simplification is that, on the one hand the absorbing pattern appears a linear number of times, while on the other, the shape of uniform trees facilitates the pruning of huge chunks of the expression.

A natural improvement would be to obtain that the expectation tends to a constant, instead of being bounded by a constant, or even a characterization of the limit distribution. In another direction, using infinitely many rules is probably possible, under some analytic conditions, and there are other hypotheses that may be weakened: it is not difficult for instance to ask that the dependency graph has one large strongly connected component (all others having size one)[f], periodicity is also manageable, ... All of these generalizations introduce technical difficulties in the analysis, but we think that in most natural cases, unless we explicitly design the specification to prevent the simplifications from happening sufficiently often, the uniform distribution is degenerated when interpreting the expression: this phenomenon can probably be considered as inherent in this framework.

One generalization that seems to exhibit a different behavior is when the specification itself depends on n. Some preliminary results were obtained recently by the second and third authors [19], for the specific case of regular expressions of size n on an alphabet whose cardinality also depends on n.

In our opinion, instead of generalizing the kind of specification even more, the natural continuation of this work is to investigate non-uniform distributions. The first candidate that comes in mind is what is called BST-like distributions, where the size of the children are distributed as in a binary search tree: that kind of distribution is really used to test algorithms, and it is probably mathematically tractable [18], even if it implies dealing with systems of differential equations.

Acknowledgments. The third author is funded by the Projet RIN Alenor (Regional Project from French Normandy).

References

[1] Alfred V. Aho and Jeffrey D. Ullman. *The Theory of Parsing, Translation, and Compiling.* Prentice-Hall, Inc., USA, 1972.

[2] Cyril Banderier and Michael Drmota. Formulae and asymptotics for coefficients of algebraic functions. *Combinatorics, Probability & Computing*, 24(1):1–53, 2015.

[3] Jason P. Bell, Stanley Burris, and Karen A. Yeats. Characteristic points of recursive systems. *Electr. J. Comb.*, 17(1), 2010.

[4] Sabine Broda, António Machiavelo, Nelma Moreira, and Rogério Reis. Average size of automata constructions from regular expressions. *Bulletin of the EATCS*, 116, 2015.

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

[6] Michael Drmota. *Random Trees: An Interplay Between Combinatorics and Probability.* Springer Publishing Company, Incorporated, 1st edition, 2009.

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

[f] The general case with no constraint on the dependency graph can be really intricate, starting with the asymptotics that may behave differently [2].

[8] Philippe Flajolet, Paolo Sipala, and Jean-Marc Steyaert. Analytic variations on the common subexpression problem. In *Automata, Languages and Programming, 17th International Colloquium, ICALP90, Warwick University, England, UK, July 16-20, 1990, Proceedings*, volume 443 of *Lecture Notes in Computer Science*, pages 220–234. Springer, 1990.

[9] Philippe Flajolet and Jean-Marc Steyaert. A complexity calculus for recursive tree algorithms. *Mathematical Systems Theory*, 19(4):301–331, 1987.

[10] Daniele Gardy. Random boolean expressions. *Discrete Mathematics & Theoretical Computer Science*, DMTCS Proceedings vol. AF, Computational Logic and Applications (CLA '05):1–36, 2005.

[11] Christopher D. Godsil and Gordon F. Royle. *Algebraic Graph Theory.* Graduate texts in mathematics. Springer, 2001.

[12] Florent Koechlin, Cyril Nicaud, and Pablo Rotondo. Uniform random expressions lack expressivity. In Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen, editors, *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 51:1–51:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[13] Florent Koechlin, Cyril Nicaud, and Pablo Rotondo. On the degeneracy of random expressions specified by systems of combinatorial equations. In Natasa Jonoska and Dmytro Savchuk, editors, *Developments in Language Theory - 24th International Conference, DLT 2020, Tampa, FL, USA, May 11-15, 2020, Proceedings*, volume 12086 of *Lecture Notes in Computer Science*, pages 164–177. Springer, 2020.

[14] Jonathan Lee and Jeffrey Shallit. Enumerating regular expressions and their languages. In Michael Domaratzki, Alexander Okhotin, Kai Salomaa, and Sheng Yu, editors, *Implementation and Application of Automata, 9th International Conference, CIAA 2004, Kingston, Canada, July 22-24, 2004*, volume 3317 of *Lecture Notes in Computer Science*, pages 2–22. Springer, 2004.

[15] A Meir and J.W Moon. On an asymptotic method in enumeration. *Journal of Combinatorial Theory, Series A*, 51(1):77 – 89, 1989.

[16] Michel Nguyên-Thê. *Distribution of Valuations on Trees.* Theses, Ecole Polytechnique X, February 2004.

[17] Cyril Nicaud. On the Average Size of Glushkov's Automata. In Adrian-Horia Dediu, Armand-Mihai Ionescu, and Carlos Martín-Vide, editors, *Language and Automata Theory and Applications, Third International Conference, LATA 2009, Tarragona, Spain, April 2-8, 2009. Proceedings*, volume 5457 of *Lecture Notes in Computer Science*, pages 626–637. Springer, 2009.

[18] Cyril Nicaud, Carine Pivoteau, and Benoît Razet. Average analysis of Glushkov automata under a bst-like model. In Kamal Lodaya and Meena Mahajan, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2010, December 15-18, 2010, Chennai, India*, volume 8 of *LIPIcs*, pages 388–399. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2010.

[19] Cyril Nicaud and Pablo Rotondo. Random regular expression over huge alphabets. 2020. Submitted.

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

[21] Kosaku Yosida. *Functional analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

[22] Mishael Zedek. Continuity and location of zeros of linear combinations of polynomials. *Proceedings of the American Mathematical Society*, 16(1):78–84, 1965.