Almost Every Simply Typed λ-Term Has a Long β-Reduction Sequence

Ryoma Sin'ya, Kazuyuki Asada, Naoki Kobayashi, and Takeshi Tsukada

The University of Tokyo

Abstract. It is well known that the length of a β-reduction sequence of a simply typed λ-term of order k can be huge; it is as large as k-fold exponential in the size of the λ-term in the *worst* case. We consider the following relevant question about *quantitative* properties, instead of the worst case: *how many* simply typed λ-terms have very long reduction sequences? We provide a partial answer to this question, by showing that asymptotically almost every simply typed λ-term of order k has a reduction sequence as long as $(k-2)$-fold exponential in the term size, under the assumption that the arity of functions and the number of variables that may occur in every subterm are bounded above by a constant. The work has been motivated by quantitative analysis of the complexity of higher-order model checking.

1 Introduction

It is well known that the length of a β-reduction sequence of a simply typed λ-term can be extremely long. Beckmann [1] showed that, for any $k \geq 0$,

$$\max\{\beta(t) \mid t \text{ is a simply typed } \lambda\text{-term of order } k \text{ and size } n\} = \mathbf{exp}_k(\Theta(n))$$

where $\beta(t)$ is the maximum length of the β-reduction sequences of the term t, and $\mathbf{exp}_k(x)$ is defined by: $\mathbf{exp}_0(x) \triangleq x$ and $\mathbf{exp}_{k+1}(x) \triangleq 2^{\mathbf{exp}_k(x)}$. Indeed, the following order-k term [1]:

$$(\mathit{Twice}_k)^n\, \mathit{Twice}_{k-1} \cdots \mathit{Twice}_2 (\lambda x.a\, x\, x)c,$$

where Twice_j is the twice function $\lambda f^{\sigma_{j-1}}.\lambda x^{\sigma_{j-2}}.f(f\,x)$ (with σ_j being the order-j type defined by: $\sigma_0 = \mathsf{o}$ and $\sigma_j = \sigma_{j-1} \to \sigma_{j-1}$), has a β-reduction sequence of length $\mathbf{exp}_k(\Omega(n))$.

Although the worst-case length of the longest β-reduction sequence is well known as above, much is not known about the *average-case* length of the longest β-reduction sequence: *how often* does one encounter a term having a very long β-reduction sequence? In other words, suppose we pick a simply-typed λ-term t of order k and size n *randomly*; then what is the probability that t has a β-reduction sequence longer than a certain bound, like $\mathbf{exp}_k(cn)$ (where c is some constant)? One may expect that, although there exists a term (like the one above) whose reduction sequence is as long as $\mathbf{exp}_k(\Omega(n))$, such a term is rarely encountered.

In the present paper, we provide a partial answer to the above question, by showing that almost every simply typed λ-term of order k has a reduction sequence as long as $(k-2)$-fold exponential in the term size, under a certain assumption. More precisely, we shall show:

$$\lim_{n\to\infty} \frac{\#\big(\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid \beta(t) \geq \mathbf{exp}_{k-2}(n)\}\big)}{\#(\Lambda_n^\alpha(k,\iota,\xi))} = 1$$

where $\Lambda_n^\alpha(k,\iota,\xi)$ is the set of (α-equivalence classes $[-]_\alpha$ of) simply-typed λ-terms such that the term size is n, the order is up to k, the (internal) arity is up to $\iota \geq k$ and the number of variable names is up to ξ (see the next section for the precise definition).

To obtain the above result, we use techniques inspired by the quantitative analysis of *untyped* λ-terms [2,3,4]. For example, David et al. [2] have shown that almost all untyped λ-terms are strongly normalizing, whereas the result is opposite in the corresponding combinatory logic. A more sophisticated analysis is, however, required in our case, for considering only well-typed terms, and also for reasoning about the *length* of a reduction sequence instead of a qualitative property like strong normalization.

This work is a part of our long-term project on the quantitative analysis of the complexity of higher-order model checking [5,6]. The higher-order model checking asks whether the (possibly infinite) tree generated by a ground-type term of the λY-calculus (or, a higher-order recursion scheme) satisfies a given regular property, and it is known that the problem is k-EXPTIME complete for order-k terms [6]. Despite the huge worst-case complexity, practical model checkers [7,8,9] have been built, which run fast for many typical inputs, and have successfully been applied to automated verification of functional programs [10,11,12,13]. The project aims to provide a theoretical justification for it, by studying *how many* inputs actually suffer from the worst-case complexity. Since the problem appears to be hard due to recursion, as an intermediate step towards the goal, we aimed to analyze the variant of the problem considered by Terui [14]: given a term of the simply-typed λ-calculus (without recursion) of type Bool, decide whether it evaluates to true or false (where Booleans are Church-encoded; see [14] for the precise definition). Terui has shown that even for the problem, the complexity is k-EXPTIME complete for order-$(2k+2)$ terms. If, contrary to the result of the present paper, the upper-bound of the lengths of β-reduction sequences *were* small for almost every term, then we could have concluded that the decision problem above is easily solvable for most of the inputs. The result in the present paper does not necessarily provide a negative answer to the question above, because one need not necessarily apply β-reductions to solve Terui's decision problem.

The present work may also shed some light on other problems on typed λ-calculi with exponential or higher worst-case complexity. For example, despite DEXPTIME-completeness of ML typability [15,16], it is often said that the exponential behavior is rarely seen *in practice*. That is, however, based on only empirical studies. Our technique may be used to provide a theoretical justification (or possibly *un*justification).

The rest of this paper is organized as follows. Section 2 states our main result formally. Section 3 analyzes the asymptotic behavior of the number of typed λ-terms of a given size. Section 4 proves the main result. Section 5 discusses related work, and Section 6 concludes the paper.

2 Main Result

In this section we give the precise statement of our main theorem. We denote the cardinality of a set S by $\#(S)$, and the domain and image of a function f by $\mathrm{Dom}(f)$ and $\mathrm{Im}(f)$, respectively.

The set of (*simple*) *types*, ranged over by τ and σ, is given by: $\tau ::= \mathsf{o} \mid \sigma \to \tau$. Let V be a countably infinite set, which is ranged over by x, x_1, x_2, etc. The set of λ-*terms* (or *terms*), ranged over by t, is defined by:

$$t ::= x \mid \lambda \overline{x}^{\tau}.t \mid t\,t \qquad\qquad \overline{x} ::= x \mid *$$

We call elements of $V \cup \{*\}$ *variables*; $V \cup \{*\}$ is ranged over by $\overline{x}, \overline{x}_1, \overline{x}_2$, etc. We call the special variable $*$ an *unused variable*. We sometimes omit type annotations and just write $\lambda\overline{x}.t$ for $\lambda\overline{x}^{\tau}.t$.

Terms of our syntax can be translated to usual λ-terms by regarding elements in $V \cup \{*\}$ as usual variables. We define the notions of free variables, closed terms, and α-equivalence $\sim_\alpha$ through this identification. The α-equivalence class of a term t is written as $[t]_\alpha$. In this paper, we do not consider a term as an α-equivalence class, and we always use $[-]_\alpha$ explicitly. For a term t, we write $\mathbf{FV}(t)$ for the set of all the free variables of t.

For a term t, we define the set $\mathbf{V}(t)$ of variables (except $*$) in t by:

$$\mathbf{V}(x) \triangleq \{x\} \quad \mathbf{V}(\lambda x^{\tau}.t) \triangleq \{x\} \cup \mathbf{V}(t) \quad \mathbf{V}(\lambda *^{\tau}.t) \triangleq \mathbf{V}(t) \quad \mathbf{V}(t_1 t_2) \triangleq \mathbf{V}(t_1) \cup \mathbf{V}(t_2).$$

Note that neither $\mathbf{V}(t)$ nor even $\#(\mathbf{V}(t))$ is preserved by α-equivalence. For example, $t = \lambda x_1.(\lambda x_2.x_2)(\lambda x_3.x_1)$ and $t' = \lambda x_1.(\lambda x_1.x_1)(\lambda *.x_1)$ are α-equivalent, but $\#(\mathbf{V}(t)) = 3$ and $\#(\mathbf{V}(t')) = 1$.

A *type environment* Γ is a finite set of type bindings of the form $x : \tau$ such that if $(x : \tau), (x : \tau') \in \Gamma$ then $\tau = \tau'$; sometimes we regard an environment also as a function. Note that $(* : \tau)$ cannot belong to a type environment; we do not need any type assumption for $*$ since it does not occur in terms. We give the typing rules as follows:

$$\frac{}{x : \tau \vdash x : \tau} \qquad \frac{\Gamma_1 \vdash t_1 : \sigma \to \tau \qquad \Gamma_2 \vdash t_2 : \sigma}{\Gamma_1 \cup \Gamma_2 \vdash t_1 t_2 : \tau}$$

$$\frac{\Gamma' \vdash t : \tau \qquad \Gamma' = \Gamma \text{ or } \Gamma' = \Gamma \cup \{\overline{x} : \sigma\} \qquad \overline{x} \notin \mathrm{Dom}(\Gamma)}{\Gamma \vdash \lambda \overline{x}^{\sigma}.t : \sigma \to \tau}$$

The above typing rules are equivalent to the usual ones for closed terms, and if $\Gamma \vdash t : \tau$ is derivable, then the derivation is unique. Moreover, if $\Gamma \vdash t : \tau$ then $\mathrm{Dom}(\Gamma) = \mathbf{FV}(t)$. Below we consider only well-typed λ-terms. A pair $\langle \Gamma; \tau \rangle$ of Γ

and τ is called a *typing*. We use θ as a metavariable for typings. When $\Gamma \vdash t : \tau$ is derived, we call $\langle\Gamma;\tau\rangle$ a *typing of* a term t, and call t an *inhabitant of* $\langle\Gamma;\tau\rangle$ or a $\langle\Gamma;\tau\rangle$*-term*.

Definition 1 (size, order and internal arity of a term). The *size* of a term t, written $|t|$, is defined by:

$$|\overline{x}| \triangleq 1 \qquad |\lambda\overline{x}^{\tau}.t| \triangleq |t|+1 \qquad |t_1 t_2| \triangleq |t_1|+|t_2|+1.$$

The *order* and *internal arity* of a type τ, written $\mathtt{ord}(\tau)$ and $\mathtt{iar}(\tau)$, are defined respectively by:

$$\mathtt{ord}(\mathtt{o}) \triangleq 0 \qquad \mathtt{iar}(\mathtt{o}) \triangleq 0$$
$$\mathtt{ord}(\tau_1 \to \cdots \to \tau_n \to \mathtt{o}) \triangleq \max\{\mathtt{ord}(\tau_i)+1 \mid 1 \le i \le n\} \qquad (n \ge 1)$$
$$\mathtt{iar}(\tau_1 \to \cdots \to \tau_n \to \mathtt{o}) \triangleq \max(\{n\} \cup \{\mathtt{iar}(\tau_i) \mid 1 \le i \le n\}) \quad (n \ge 1).$$

For a $\langle\Gamma;\tau\rangle$-term t, we define the order and internal arity of $\Gamma \vdash t : \tau$ written $\mathtt{ord}(\Gamma \vdash t : \tau)$ and $\mathtt{iar}(\Gamma \vdash t : \tau)$ by:

$$\mathtt{ord}(\Gamma \vdash t : \tau) \triangleq \max\{\mathtt{ord}(\tau') \mid (\Gamma' \vdash t' : \tau') \text{ occurs in } \Delta\}$$
$$\mathtt{iar}(\Gamma \vdash t : \tau) \triangleq \max\{\mathtt{iar}(\tau') \mid (\Gamma' \vdash t' : \tau') \text{ occurs in } \Delta\}$$

where Δ is the (unique) derivation tree for $\Gamma \vdash t : \tau$.

Note that the notions of size, order, internal arity, and $\beta(t)$ (defined in the introduction) are well-defined with respect to α-equivalence.

Definition 2 (terms with bounds on types and variables). Let $\delta, \iota, \xi \ge 0$ and $n \ge 1$ be integers. We denote by $\mathrm{Types}(\delta,\iota)$ the set of types $\{\tau \mid \mathtt{ord}(\tau) \le \delta, \mathtt{iar}(\tau) \le \iota\}$. For Γ and τ we define:

$$\Lambda_n^{\alpha}(\langle\Gamma;\tau\rangle,\delta,\iota,\xi) \triangleq \{[t]_\alpha \mid \Gamma \vdash t : \tau,\ |t| = n,\ \min_{t' \in [t]_\alpha} \#(\mathbf{V}(t')) \le \xi,$$
$$\mathtt{ord}(\Gamma \vdash t : \tau) \le \delta,\ \mathtt{iar}(\Gamma \vdash t : \tau) \le \iota\}$$
$$\Lambda^{\alpha}(\langle\Gamma;\tau\rangle,\delta,\iota,\xi) \triangleq \bigcup_{n \ge 1} \Lambda_n^{\alpha}(\langle\Gamma;\tau\rangle,\delta,\iota,\xi).$$

Also we define:

$$\Lambda_n^{\alpha}(\delta,\iota,\xi) \triangleq \bigcup_{\tau \in \mathrm{Types}(\delta,\iota)} \Lambda_n^{\alpha}(\langle\emptyset;\tau\rangle,\delta,\iota,\xi) \qquad \Lambda^{\alpha}(\delta,\iota,\xi) \triangleq \bigcup_{n \ge 1} \Lambda_n^{\alpha}(\delta,\iota,\xi).$$

Our main result is the following theorem, which will be proved in Section 4.

Theorem 1. *Let $\delta, \iota, \xi \ge 2$ be integers and let $k = \min\{\delta,\iota\}$. Then,*

$$\lim_{n\to\infty} \frac{\#\big(\{[t]_\alpha \in \Lambda_n^{\alpha}(\delta,\iota,\xi) \mid \beta(t) \ge \mathbf{exp}_{k-2}(n)\}\big)}{\#(\Lambda_n^{\alpha}(\delta,\iota,\xi))} = 1.$$

Remark 1. Note that in the above theorem, the order δ, the internal arity ι and the number ξ of variables are bounded above by a constant, independently of the term size n. It is debatable whether the assumption is reasonable, and a slight change of the assumption may change the result, as is the case for strong normalization of untyped λ-term [2,4]. When λ-terms are viewed as models of functional programs, our rationale behind the assumption is as follows. The assumption that the size of types (hence also the order and the internal arity) is fixed is sometimes assumed in the context of type-based program analysis [17]. The assumption on the number of variables comes from the observation that a large program usually consists of a large number of *small* functions, and that the number of variables is bounded by the size of each function.

Remark 2. Actually, by refining the proof of Theorem 1, we can strengthen the main result as follows: For $\delta, \iota, \xi \geq 2$ and $k = \min\{\delta, \iota\}$, there exists a real number $q > 0$ such that

$$\lim_{n\to\infty} \frac{\#\big(\{[t]_\alpha \in \Lambda_n^\alpha(\delta,\iota,\xi) \mid \beta(t) \geq \mathbf{exp}_{k-1}(n^q)\}\big)}{\#(\Lambda_n^\alpha(\delta,\iota,\xi))} = 1.$$

A proof of this strengthened result is presented in a revised version of this paper under preparation.

3 Analysis of $\Lambda_n^\alpha(\delta, \iota, \xi)$

To prove our main theorem, we first analyze some formal language theoretic structure and properties of $\Lambda^\alpha(\delta,\iota,\xi)$: in Section 3.1, we construct a regular tree grammar such that there is a size preserving bijection between its tree language and $\Lambda^\alpha(\delta,\iota,\xi)$; in Section 3.2, we show that the grammar has two important properties: irreducibility and aperiodicity. Thanks to those properties, we can obtain a simple asymptotic formula for $\#(\Lambda_n^\alpha(\delta,\iota,\xi))$ using analytic combinatorics [18]. The irreducibility and aperiodicity properties will also be used in Section 4 for adjusting the size and typing of a λ-term.

3.1 $\Lambda^\alpha(\delta, \iota, \xi)$ as a Regular Tree Language

We first recall some basic definitions for regular tree grammars. A *ranked alphabet* Σ is a mapping from a finite set of symbols to the set of natural numbers. For a symbol $a \in \mathrm{Dom}(\Sigma)$, we call $\Sigma(a)$ the *rank* of a. A Σ*-tree* is a tree composed from symbols in Σ according to their ranks: (i) a is a Σ-tree if $\Sigma(a) = 0$, (ii) $a(T_1, \cdots, T_{\Sigma(a)})$ is a Σ-tree if $\Sigma(a) \geq 1$ and T_i is a Σ-tree for each $i \in \{1, \ldots, \Sigma(a)\}$. We use the meta-variable T for trees. The *size* of T, written as $|T|$, is the number of nodes and leaves of T. We denote the set of all Σ-trees by $\mathcal{T}_\Sigma$.

A *regular tree grammar* is a triple $\mathcal{G} = (\Sigma, \mathcal{N}, \mathcal{R})$ where (i) Σ is a ranked alphabet; (ii) $\mathcal{N}$ is a finite set of *non-terminals*; (iii) $\mathcal{R}$ is a finite set of *rewriting rules* of the form $N \longrightarrow a(N_1, \cdots, N_{\Sigma(a)})$ where $a \in \mathrm{Dom}(\Sigma)$, $N \in \mathcal{N}$ and

$N_i \in \mathcal{N}$ for every $i \in \{1, \dots, \Sigma(a)\}$. A $(\Sigma \cup \mathcal{N})$-*tree* T is a tree composed from symbols in $\Sigma \cup \mathcal{N}$ according to their ranks where the rank of every symbol in $\mathcal{N}$ is zero (thus non-terminals appear only in leaves of T). For a tree grammar $\mathcal{G} = (\Sigma, \mathcal{N}, \mathcal{R})$ and a non-terminal $N \in \mathcal{N}$, *the language* $\mathcal{L}(\mathcal{G}, N)$ *of* N is defined by $\mathcal{L}(\mathcal{G}, N) \triangleq \{T \in \mathcal{T}_\Sigma \mid N \longrightarrow^*_{\mathcal{G}} T\}$ where $\longrightarrow^*_{\mathcal{G}}$ denotes the reflexive and transitive closure of the rewriting relation $\longrightarrow_{\mathcal{G}}$. We also define $\mathcal{L}_n(\mathcal{G}, N) \triangleq \{T \in \mathcal{T}_\Sigma \mid N \longrightarrow^* T, |T| = n\}$. We often omit $\mathcal{G}$ and write $N \longrightarrow^* N'$, $\mathcal{L}(N)$, and $\mathcal{L}_n(N)$ for $N \longrightarrow^*_{\mathcal{G}} N'$, $\mathcal{L}(\mathcal{G}, N)$, and $\mathcal{L}_n(\mathcal{G}, N)$ respectively, if $\mathcal{G}$ is clear from the context. We say that N' is *reachable* from N if there exists a $(\Sigma \cup \mathcal{N})$-tree T such that $N \longrightarrow^* T$ and T contains N' as a leaf. A grammar $\mathcal{G}$ is *unambiguous* if, for every pair of a non-terminal N and a tree T, there exists at most one leftmost reduction sequence from N to T.

Definition 3 (grammar of $\Lambda^\alpha(\delta, \iota, \xi)$). Let $\delta, \iota, \xi \geq 0$ be integers and $X_\xi = \{x_1, \cdots, x_\xi\}$ be a subset of V. The regular tree grammar $\mathcal{G}(\delta, \iota, \xi)$ is defined as $(\Sigma(\delta, \iota, \xi), \mathcal{N}(\delta, \iota, \xi), \mathcal{R}(\delta, \iota, \xi))$ where:

$$\begin{aligned}
\Sigma(\delta, \iota, \xi) &\triangleq \{x \mapsto 0 \mid x \in X_\xi\} \cup \{@ \mapsto 2\} \\
&\quad \cup \{\lambda \overline{x}^\tau \mapsto 1 \mid \overline{x} \in \{*\} \cup X_\xi,\ \tau \in \mathrm{Types}(\delta - 1, \iota)\} \\
\mathcal{N}(\delta, \iota, \xi) &\triangleq \{N_{\langle \Gamma; \tau \rangle} \mid \tau \in \mathrm{Types}(\delta, \iota), \mathrm{Dom}(\Gamma) \subseteq X_\xi,\ \mathrm{Im}(\Gamma) \subseteq \mathrm{Types}(\delta - 1, \iota), \\
&\qquad \Gamma \vdash t : \tau \text{ for some } t\ \} \\
\mathcal{R}(\delta, \iota, \xi) &\triangleq \{N_{\langle \{x_i : \tau\}; \tau \rangle} \longrightarrow x_i\} \cup \{N_{\langle \Gamma; \sigma \to \tau \rangle} \longrightarrow \lambda *^\sigma(N_{\langle \Gamma; \tau \rangle})\} \\
&\quad \cup \{N_{\langle \Gamma; \sigma \to \tau \rangle} \longrightarrow \lambda x_i^\sigma(N_{\langle \Gamma \cup \{x_i : \sigma\}; \tau \rangle}) \mid i = \min\{j \geq 1 \mid x_j \notin \mathrm{Dom}(\Gamma)\}, \\
&\qquad \#(\Gamma) < \xi\} \cup \{N_{\langle \Gamma; \tau \rangle} \longrightarrow @(N_{\langle \Gamma_1; \sigma \to \tau \rangle}, N_{\langle \Gamma_2; \sigma \rangle}) \mid \Gamma = \Gamma_1 \cup \Gamma_2\}
\end{aligned}$$

Here, the special symbol $@ \in \mathrm{Dom}(\Sigma(\delta, \iota, \xi))$ corresponds to application. For a technical convenience, the above definition excludes from $\mathcal{N}(\delta, \iota, \xi)$ typings which have no inhabitant. Note that $\Sigma(\delta, \iota, \xi)$, $\mathcal{N}(\delta, \iota, \xi)$ and $\mathcal{R}(\delta, \iota, \xi)$ are finite. To see the finiteness of $\mathcal{N}(\delta, \iota, \xi)$, notice that X_ξ and $\mathrm{Types}(\delta - 1, \iota)$ are finite, hence so is $\{\Gamma \mid \mathrm{Dom}(\Gamma) \subseteq X_\xi,\ \mathrm{Im}(\Gamma) \subseteq \mathrm{Types}(\delta - 1, \iota)\}$. The finiteness of $\mathcal{R}(\delta, \iota, \xi)$ follows immediately from that of $\mathcal{N}(\delta, \iota, \xi)$.

Example 1. Let us consider the case where $\delta = \iota = \xi = 1$. The grammar $\mathcal{G}(1, 1, 1)$ consists of the following.

$$\Sigma(1,1,1) = \{x_1, @, \lambda x_1^{\mathsf{o}}, \lambda *^{\mathsf{o}}\} \qquad \mathcal{N}(1,1,1) = \{N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle}, N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \rangle} N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \to \mathsf{o} \rangle}\}$$

$$\mathcal{R}(1,1,1) = \begin{cases} N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle} \longrightarrow \lambda x_1^{\mathsf{o}}(N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \rangle}) \\ N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \rangle} \longrightarrow x_1 | @(N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \to \mathsf{o} \rangle}, N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \rangle}) | @(N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle}, N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \rangle}) \\ N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \to \mathsf{o} \rangle} \longrightarrow \lambda *^{\mathsf{o}}(N_{\langle \{x_1 : \mathsf{o}\}; \mathsf{o} \rangle}). \end{cases}$$

There is the obvious embedding $e^{(\delta, \iota, \xi)}$ (e for short) from trees in $\mathcal{T}_{\Sigma(\delta, \iota, \xi)}$ into λ-terms. For $N_{\langle \Gamma; \tau \rangle} \in \mathcal{N}(\delta, \iota, \xi)$ we define

$$\pi^{(\delta, \iota, \xi)}_{\langle \Gamma; \tau \rangle} \triangleq [-]_\alpha \circ e : \mathcal{L}\left(N_{\langle \Gamma; \tau \rangle}\right) \to \Lambda^\alpha(\langle \Gamma; \tau \rangle, \delta, \iota, \xi).$$

We sometimes omit the superscript and/or the subscript.

Proposition 1. *For $\delta, \iota, \xi \geq 0$, $\pi_{\langle \Gamma;\tau \rangle}$ is a size-preserving bijection, and $\mathcal{G}(\delta, \iota, \xi)$ is unambiguous.*

The former part of Proposition 1 says that $\mathcal{G}(\delta, \iota, \xi)$ gives a complete representation system of the α-equivalence classes. For $[t]_\alpha \in \Lambda^\alpha(\langle \Gamma; \tau \rangle, \delta, \iota, \xi)$, we define $\nu^{(\delta,\iota,\xi)}_{\langle \Gamma;\tau \rangle}([t]_\alpha)$ (or $\nu([t]_\alpha)$ for short) as $e^{(\delta,\iota,\xi)} \circ \left(\pi^{(\delta,\iota,\xi)}_{\langle \Gamma;\tau \rangle}\right)^{-1}([t]_\alpha)$. The function ν *normalizes* variable names. For example, $t = \lambda x.x(\lambda y.\lambda z.z)$ is normalized to $\nu([t]_\alpha) = \lambda x_1.x_1(\lambda *.\lambda x_1.x_1)$.

Due to technical reasons, we restrict the grammar $\mathcal{G}(\delta, \iota, \xi)$ to $\mathcal{G}^\emptyset(\delta, \iota, \xi)$, which contains only non-terminals reachable from $N_{\langle \emptyset;\sigma \rangle}$ for some σ (see Appendix B for details).

$$\begin{aligned}
\mathcal{N}^\emptyset(\delta, \iota, \xi) &\triangleq \{N_\theta \in \mathcal{N}(\delta, \iota, \xi) \mid N_\theta \text{ is reachable from some } N_{\langle \emptyset;\sigma \rangle} \in \mathcal{N}(\delta, \iota, \xi)\} \\
\mathcal{R}^\emptyset(\delta, \iota, \xi) &\triangleq \{N_\theta \longrightarrow T \in \mathcal{R}(\delta, \iota, \xi) \mid N_\theta \in \mathcal{N}^\emptyset(\delta, \iota, \xi)\} \\
\mathcal{G}^\emptyset(\delta, \iota, \xi) &\triangleq (\Sigma(\delta, \iota, \xi), \mathcal{N}^\emptyset(\delta, \iota, \xi), \mathcal{R}^\emptyset(\delta, \iota, \xi)).
\end{aligned}$$

For $N_\theta \in \mathcal{N}^\emptyset(\delta, \iota, \xi)$, clearly $\mathcal{L}\left(\mathcal{G}^\emptyset(\delta, \iota, \xi), N_\theta\right) = \mathcal{L}\left(\mathcal{G}(\delta, \iota, \xi), N_\theta\right)$. Through the bijection π, we can show that, for any $N_{\langle \Gamma;\tau \rangle} \in \mathcal{N}(\delta, \iota, \xi)$, $N_{\langle \Gamma;\tau \rangle}$ also belongs to $\mathcal{N}^\emptyset(\delta, \iota, \xi)$ if and only if there exists a term in $\Lambda^\alpha(\delta, \iota, \xi)$ whose derivation contains a type judgment of the form $\Gamma \vdash t : \tau$.

3.2 Irreducibility and Aperiodicity

We discuss two important properties of the grammar $\mathcal{G}^\emptyset(\delta, \iota, \xi)$ where $\delta, \iota, \xi \geq 2$: irreducibility and aperiodicity [18].[1]

Definition 4 (irreducibility and aperiodicity). Let $\mathcal{G} = (\Sigma, \mathcal{N}, \mathcal{R})$ be a regular tree grammar. We say that $\mathcal{G}$ is:

- *non-linear* if $\mathcal{R}$ contains at least one rule of the form $N \longrightarrow a(N_1, \cdots, N_{\Sigma(a)})$ with $\Sigma(a) \geq 2$,
- *strongly connected* if for any pair of non-terminals $N_1, N_2 \in \mathcal{N}$, N_1 is reachable from N_2,
- *irreducible* if $\mathcal{G}$ is both non-linear and strongly connected,
- *aperiodic* if for any non-terminal $N \in \mathcal{N}$ there exists an integer $m > 0$ such that $\#(\mathcal{L}_n(N)) > 0$ for any $n > m$.

Proposition 2. *$\mathcal{G}^\emptyset(\delta, \iota, \xi)$ is irreducible and aperiodic for any $\delta, \iota, \xi \geq 2$.*

The following theorem is a minor modification of Theorem VII.5 in [18], which states the asymptotic behavior of an irreducible and aperiodic *context-free specification* (see Appendix C for details). Below, $\sim$ means the *asymptotic equality*, i.e., $f(n) \sim g(n) \iff \lim_{n\to\infty} f(n)/g(n) = 1$.

[1] In [18], irreducibility and aperiodicity are defined for *context-free specifications*. Our definition is a straightforward adaptation of the definition for regular tree grammars.

Theorem 2 ([18]). *Let $\mathcal{G} = (\Sigma, \mathcal{N}, \mathcal{R})$ be an unambiguous, irreducible and aperiodic regular tree grammar. Then there exists a constant $\gamma(\mathcal{G}) > 1$ such that, for any non-terminal $N \in \mathcal{N}$, there exists a constant $C_N(\mathcal{G}) > 0$ such that*

$$\#(\mathcal{L}_n(N)) \sim C_N(\mathcal{G})\gamma(\mathcal{G})^n n^{-3/2}.$$

As a corollary of Proposition 2 and Theorem 2 above, we obtain:

$$\#(\Lambda_n^\alpha(\delta, \iota, \xi)) \sim C\gamma^n n^{-3/2} \tag{1}$$

where $C > 0$ and $\gamma > 1$ are some real constants determined by δ, ι, $\xi \geq 2$. For proving our main theorem, we use a variation of the formula (1) above, stated as Lemma 1 later.

4 Proof of the Main Theorem

We give a proof of Theorem 1 in this section. In the rest of the paper, we denote by $\log^{(2)}(n)$ the 2-fold logarithm: $\log^{(2)}(n) \triangleq \log\log n$. All logarithms are base 2. The outline of the proof is as follows. We prepare a family $(t_n)_{n\in\mathbb{N}}$ of λ-terms such that t_n is of size $\Omega(\log^{(2)}(n))$ and has a β-reduction sequence of length $\mathbf{exp}_k(\Omega(|t_n|))$, i.e., $\mathbf{exp}_{k-2}(\Omega(n))$. Then we show that almost every λ-term of size n contains t_n as a subterm. The latter is shown by adapting (a parameterized version of) the *infinite monkey theorem*[2] for words to simply-typed λ-terms.

To clarify the idea, let us first recall the infinite monkey theorem for words. Let A be an alphabet, i.e., a finite non-empty set of symbols. For a word $w = a_1 \cdots a_n$, we write $|w| = n$ for the length of w. As usual, we denote by A^n the set of all words of length n over A, and by A^* the set of all finite words over A: $A^* = \bigcup_{n\geq 0} A^n$. For two words $w, w' \in A^*$, we say w' is a *subword of* w and write $w' \sqsubseteq w$ if $w = w_1 w' w_2$ for some words $w_1, w_2 \in A^*$. The infinite monkey theorem states that, for any word $w \in A^*$, the probability that a randomly chosen word of size n contains w as a subword tends to one if n tends to infinity.

To prove our main theorem, we need to extend the above infinite monkey theorem to the following parameterized version[3], and then further extend it for simply-typed λ-terms instead of words. We give a proof of the following proposition, because it will clarify the overall structure of the proof of the main theorem.

Proposition 3 (parameterized infinite monkey theorem). *Let A be an alphabet and $(w_n)_n$ be a family of words over A such that $|w_n| = \lceil\log^{(2)}(n)\rceil$. Then, we have:*

$$\lim_{n\to\infty} \frac{\#(\{w \in A^n \mid w_n \sqsubseteq w\})}{\#(A^n)} = 1.$$

[2] It is also called as "Borges's theorem" (*cf.* [18, p.61, Note I.35]).

[3] Although it is a simple extension, we are not aware of literature that explicitly states this parameterized version.

Proof. Let $p(n)$ be $1 - \#(\{w \in A^n \mid w_n \sqsubseteq w\}) / \#(A^n)$, i.e., the probability that a word of size n does *not* contain w_n. We write $s(n)$ for $\lceil \log^{(2)}(n) \rceil$ and $c(n)$ for $\lfloor n/s(n) \rfloor$. Given a word $w = a_1 \cdots a_n \in A^n$, let us partition it to subwords of length $s(n)$ as follows.

$$w = \underbrace{a_1 \cdots a_{s(n)}}_{\text{1-st subword}} \cdots \underbrace{a_{(c(n)-1)s(n)+1} \cdots a_{c(n)s(n)}}_{c(n)\text{-th subword}} a_{c(n)s(n)+1} \cdots a_n$$

Then,

$$\begin{aligned} p(n) &\leq \text{“the probability that none of the } i\text{-th subword is } w_n\text{”} \\ &= \left(\frac{\#\left(A^{s(n)} \setminus \{w_n\}\right)}{\#\left(A^{s(n)}\right)} \right)^{c(n)} = \left(\frac{\#\left(A^{s(n)}\right) - 1}{\#\left(A^{s(n)}\right)} \right)^{c(n)} = \left(1 - \frac{1}{\#(A)^{s(n)}}\right)^{c(n)}. \end{aligned}$$

Since $\left(1 - \frac{1}{\#(A)^{s(n)}}\right)^{c(n)} = \left(1 - \frac{1}{\#(A)^{\lceil \log^{(2)}(n) \rceil}}\right)^{\lfloor n / \lceil \log^{(2)}(n) \rceil \rfloor}$ tends to zero (see Appendix E) if n tends to infinity, we have the required result. $\square$

To prove an analogous result for simply-typed λ-terms, we consider below subcontexts of a given term instead of subwords of a given word. To consider “contexts up to α-equivalence”, in Section 4.1 we introduce the set $\mathcal{U}_n^\nu(\delta, \iota, \xi)$ of “normalized” contexts (of size n and with the restriction by δ, ι and ξ), where $\mathcal{U}_{s(n)}^\nu(\delta, \iota, \xi)$ corresponds to $A^{s(n)}$ above, and give an upper bound of $\#(\mathcal{U}_n^\nu(\delta, \iota, \xi))$. A key property used in the above proof was that any word of length n can be partitioned to sufficiently many subwords of length $\log^{(2)}(n)$. Section 4.2 below shows an analogous result that any term of size n can be decomposed into sufficiently many subcontexts of a given size. Section 4.3 constructs a family of contexts Expl_n^k (called “explosive contexts”) that have very long reduction sequences; $(\mathit{Expl}_n^k)_n$ corresponds to $(w_n)_n$ above. Finally, Section 4.4 proves the main theorem using an argument similar to (but more involved than) the one used in the proof above.

4.1 Normalized Contexts

We first introduce some basic definitions of contexts, and then we define the notion of a normalized context, which is a context normalized by the function ν given in Section 3.1.

The set of *contexts*, ranged over by C, is defined by

$$C ::= [\,] \mid x \mid \lambda \overline{x}^\tau.C \mid CC$$

The *size* of C, written $|C|$, is defined by:

$$|[\,]| \triangleq 0 \qquad |x| \triangleq 1 \qquad |\lambda \overline{x}^\tau.C| \triangleq |C| + 1 \qquad |C_1 C_2| \triangleq |C_1| + |C_2| + 1.$$

We call a context C an *n-context* (and define $\mathtt{hn}(C) \triangleq n$) if C contains n occurrences of $[\,]$. We use the metavariable S for 1-contexts. A *0/1-context* is a term

t or a 1-context S and we use the metavariable u to denote 0/1-contexts. The holes in C occur as leaves and we write $[\,]_i$ for the i-th hole, which is counted in the left-to-right order.

For C, C_1, …, $C_{\mathtt{hn}(C)}$, we write $C[C_1]\dots[C_{\mathtt{hn}(C)}]$ for the context obtained by replacing $[\,]_i$ in C with C_i for each $i \leq \mathtt{hn}(C)$. For C and C', we write $C[C']_i$ for the context obtained by replacing the i-th hole $[\,]_i$ in C with C'. As usual, these substitutions may capture variables; e.g., $(\lambda x.[\,])[x]$ is $\lambda x.x$. We say that C is a *subcontext of* C' and write $C \preceq C'$ if there exist C'', $1 \leq i \leq \mathtt{hn}(C'')$ and $C_1, \cdots, C_{\mathtt{hn}(C)}$ such that $C' = C''[C[C_1]\cdots[C_{\mathtt{hn}(C)}]]_i$.

The set of *context typings*, ranged over by κ, is defined by: $\kappa ::= \theta_1 \cdots \theta_k \Rightarrow \theta$ where $k \in \mathbb{N}$ and θ_i is a typing of the form $\langle \Gamma_i; \tau_i \rangle$ for each $1 \leq i \leq k$ (recall that we use θ as a metavariable for typings). A $\langle \Gamma_1; \tau_1 \rangle \cdots \langle \Gamma_k; \tau_k \rangle \Rightarrow \langle \Gamma; \tau \rangle$-context is a k-context C such that $\Gamma \vdash C : \tau$ is derivable from $\Gamma_i \vdash [\,]_i : \tau_i$. We identify a context typing $\Rightarrow\theta$ with the typing θ, and call a θ-context also a *θ-term*.

From now, we begin to define normalized contexts. First we consider contexts in terms of the grammar $\mathcal{G}^{\emptyset}(\delta, \iota, \xi)$ given in Section 3.1. Let $\delta, \iota, \xi \geq 0$. For $\kappa = \theta_1 \cdots \theta_n \Rightarrow \theta$ such that $N_{\theta_1}, \dots, N_{\theta_n}, N_\theta \in \mathcal{N}(\delta, \iota, \xi)$, a ($\kappa$-)*context-tree* is a tree $\widehat{T}$ in $\mathcal{T}_{\Sigma(\delta,\iota,\xi) \cup \mathcal{N}(\delta,\iota,\xi)}$ such that there exists a reduction $N_\theta \longrightarrow^* \widehat{T}$ and the occurrences of non-terminals in $\widehat{T}$ (in the left-to-right order) are exactly $N_{\theta_1}, \dots, N_{\theta_n}$. We use $\widehat{T}$ as a metavariable for context-trees. We write $\mathcal{L}(\kappa, \delta, \iota, \xi)$ for the set of all κ-context-trees. For $\theta_1 \cdots \theta_n \Rightarrow \theta$-context-tree $\widehat{T}$ and $\theta^i_1 \cdots \theta^i_{k_i} \Rightarrow \theta_i$-context-trees $\widehat{T}_i$ $(i = 1, \dots, n)$, we define the *substitution* $\widehat{T}[\widehat{T}_1] \cdots [\widehat{T}_n]$ as the $\theta^1_1 \cdots \theta^1_{k_1} \cdots \theta^n_1 \cdots \theta^n_{k_n} \Rightarrow \theta$-context-tree obtained by replacing N_{θ_i} in $\widehat{T}$ with $\widehat{T}_i$.

The set $\mathcal{C}^\nu(\kappa, \delta, \iota, \xi)$ of *normalized κ-contexts* is defined by:

$$\mathcal{C}^\nu(\kappa, \delta, \iota, \xi) \triangleq e^{(\delta,\iota,\xi)}_\kappa(\mathcal{L}(\kappa, \delta, \iota, \xi))$$

where $e^{(\delta,\iota,\xi)}_\kappa$ is the obvious embedding from κ-context-trees to κ-contexts that preserves the substitution (i.e., $e^{(\delta,\iota,\xi)}_\kappa(T[T']) = e^{(\delta,\iota,\xi)}_\kappa(T)[e^{(\delta,\iota,\xi)}_\kappa(T')]$). Further, the sets $\mathcal{U}^\nu(\delta, \iota, \xi)$ and $\mathcal{U}^\nu_n(\delta, \iota, \xi)$ of *normalized 0/1-contexts* are defined by:

$$\mathcal{U}^\nu(\delta, \iota, \xi) \triangleq \Big(\bigcup_{N_\theta \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)} \mathcal{C}^\nu(\theta, \delta, \iota, \xi) \Big) \bigcup \Big(\bigcup_{N_\theta, N_{\theta'} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)} \mathcal{C}^\nu(\theta \Rightarrow \theta', \delta, \iota, \xi) \Big)$$

$$\mathcal{U}^\nu_n(\delta, \iota, \xi) \triangleq \{ u \in \mathcal{U}^\nu(\delta, \iota, \xi) \mid |u| = n \}.$$

In our proof of the main theorem, the set $\mathcal{U}^\nu_{s(n)}(\delta, \iota, \xi)$ plays a role corresponding to $A^{s(n)}$ in the word case explained above. Note that in the word case we calculated the limit of some upper bound of $p(n)$; similarly, in our proof, we only need an upper bound of $\#(\mathcal{U}^\nu_n(\delta, \iota, \xi))$, which is given as follows.

Lemma 1 (upper bound of $\#(\mathcal{U}^\nu_n(\delta, \iota, \xi))$). *For any $\delta, \iota, \xi \geq 2$, there exists some constant $\overline{\gamma}(\delta, \iota, \xi) > 1$ such that $\#(\mathcal{U}^\nu_n(\delta, \iota, \xi)) = \mathrm{O}(\overline{\gamma}(\delta, \iota, \xi)^n)$.*

Proof Sketch. Given an unambiguous, irreducible and aperiodic regular tree grammar, adding a new terminal of the form a_N and a new rule of the form

$N \longrightarrow a_N$ for each non-terminal N does not change the unambiguity, irreducibility and aperiodicity. Let $\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi)$ be the grammar obtained by applying this transformation to $\mathcal{G}^{\emptyset}(\delta,\iota,\xi)$. We can regard a tree of $\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi)$ as a normalized context, with a_{N_θ} considered a hole with typing θ. Then, clearly we have

$$\#(\mathcal{U}^{\nu}_{n}(\delta,\iota,\xi)) \leq \#\Big(\cup_{N\in\mathcal{N}^{\emptyset}(\delta,\iota,\xi)}\mathcal{L}_n\left(\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi),N\right)\Big).$$

Thus the required result follows from Theorem 2. □

4.2 Decomposition

As explained in the beginning of this section, to prove the parameterized infinite monkey theorem for terms, we need to decompose a λ-term into sufficiently many subcontexts of the term. Thus, in this subsection, we will define a decomposition function $\widehat{\Phi}_m$ (where m is a parameter) that decomposes a term t into (i) a (sufficiently long) sequence P of 0/1-subcontexts of t such that every component u of P satisfies $|u| \geq m$, and (ii) a "second-order" context E (defined later), which is a remainder of extracting P from t. Figure 1 illustrates how a term is decomposed by $\widehat{\Phi}_3$. Here, the symbols $[\![\,]\!]$ in the second-order context on the right-hand side represents the original position of each subcontext $(\lambda y.[\,])x$, $\lambda z.\lambda *.z$, and $(\lambda *.y)\lambda z.z$.

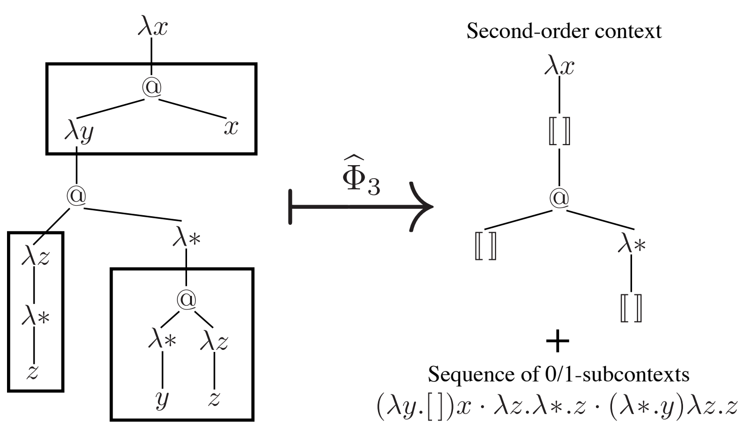

Fig. 1: Example of a decomposition

In order to define $\widehat{\Phi}_m$, let us give a precise definition of second-order contexts. The set of *second-order contexts*, ranged over by E, is defined by:

$$E ::= [\![\,]\!]^{\theta_1\cdots\theta_k\Rightarrow\theta}_n[E_1]\cdots[E_k]\ (n\in\mathbb{N}) \mid x \mid \lambda\overline{x}^{\tau}.E \mid E_1E_2.$$

Intuitively, the second-order context is an expression having holes of the form $[\![\,]\!]^{\kappa}_n$ (called *second-order holes*). In the second-order context $[\![\,]\!]^{\theta_1\cdots\theta_k\Rightarrow\theta}_n[E_1]\cdots[E_k]$,

$[\![\,]\!]_n^{\theta_1\cdots\theta_k\Rightarrow\theta}$ should be filled with a $\theta_1\cdots\theta_k\Rightarrow\theta$-context of size n, yielding a term whose typing is θ. We use the metavariable P for sequences of contexts. For a sequence of contexts $P = C_1 \cdot C_2 \cdots C_\ell$ and $i \leq \ell$, we write $\#(P)$ for the length ℓ, and $P.i$ for the i-th component C_i.

We define $|[\![\,]\!]_n^\kappa| \triangleq n$. We write $\mathtt{shn}(E)$ for the number of the second-order holes in E. For $i \leq \mathtt{shn}(E)$, we write $E.i$ for the i-th second-order hole (counted in the depth-first left-to-right pre-order). For a context C and a second-order hole $[\![\,]\!]_n^\kappa$, we write $C : [\![\,]\!]_n^\kappa$ if C is a κ-context of size n. For E and $P = C_1 \cdot C_2 \cdot \cdots C_{\mathtt{shn}(E)}$, we write $P : E$ if $C_i : E.i$ for each $i \leq \mathtt{shn}(E)$. We distinguish between second-order contexts with different annotations; for example, $[\![\,]\!]_0^{\langle\{x:\mathtt{o}\};\mathtt{o}\rangle\Rightarrow\langle\{x:\mathtt{o}\};\mathtt{o}\rangle}[x]$, $[\![\,]\!]_2^{\langle\{x:\mathtt{o}\};\mathtt{o}\rangle\Rightarrow\langle\{x:\mathtt{o}\};\mathtt{o}\rangle}[x]$ and $[\![\,]\!]_2^{\langle\{x:\mathtt{o}\to\mathtt{o}\};\mathtt{o}\to\mathtt{o}\rangle\Rightarrow\langle\{x:\mathtt{o}\to\mathtt{o}\};\mathtt{o}\to\mathtt{o}\rangle}[x]$ are different from each other. Note that every term can be regarded as a second-order context E such that $\mathtt{shn}(E) = 0$.

The bracket $[-]$ in a second-order context is just a syntactical representation rather than the substitution operation of contexts. Given E and C such that $\mathtt{shn}(E) \geq 1$ and $C : E.1$, we write $E[\![C]\!]$ for the second-order context obtained by replacing the leftmost second-order hole of E (i.e., $E.1$) with C (and by interpreting the syntactical bracket $[-]$ as the substitution operation). For example, we have: $((\lambda x.[\![\,]\!][x][x])[\![\,]\!])\,[\![\lambda y.y[\,][\,]]\!] = (\lambda x.(\lambda y.y[\,][\,])[x][x])\,[\![\,]\!] = (\lambda x.\lambda y.yxx)[\![\,]\!]$. Formally, it is defined by induction on E as follows:

$$\left([\![\,]\!]_n^{\langle\Gamma_1;\tau_1\rangle\cdots\langle\Gamma_k;\tau_k\rangle\Rightarrow\langle\Gamma;\tau\rangle}[E_1]\cdots[E_k]\right)[\![C]\!] \triangleq C[E_1]\cdots[E_k]$$

$$(\lambda\overline{x}^\tau.E)[\![C]\!] \triangleq \lambda\overline{x}^\tau.(E[\![C]\!]) \qquad (E_1E_2)[\![C]\!] \triangleq \begin{cases}(E_1[\![C]\!])E_2 & (\mathtt{shn}(E_1) \geq 1)\\ E_1(E_2[\![C]\!]) & (\mathtt{shn}(E_1) = 0)\end{cases}$$

where $C[E_1]\cdots[E_n]$ is the second-order context obtained by replacing $[\,]_i$ in C with E_i for each $i \leq n$. We extend the operation $E[\![-]\!]$ to a sequence of context. Given E and a sequence of contexts $P = C_1 \cdot C_2 \cdot \cdots C_\ell$ such that $\ell \leq \mathtt{shn}(E)$ and $C_i : E.i$ for each $i \leq \ell$, we define $E[\![P]\!]$ by induction on P:

$$E[\![\epsilon]\!] \triangleq E \qquad\qquad E[\![C \cdot P]\!] \triangleq (E[\![C]\!])[\![P]\!]$$

Note that $\mathtt{shn}(E[\![C]\!]) = \mathtt{shn}(E) - 1$, so if $P : E$ then $E[\![P]\!]$ is a term. Below we actually consider only second-order contexts whose second-order holes are of the form $[\![\,]\!]_n^\theta$ or $[\![\,]\!]_n^{\theta'\Rightarrow\theta}$.

We are now ready to define the the decomposition function $\widehat{\Phi}_m$. We first prepare an auxiliary function $\Phi_m(t) = (E, u, P)$ such that (i) u is an auxiliary 0/1-subcontext, (ii) $E[\![u \cdot P]\!] = t$, and (iii) the size of each context in P is between m and $2m-1$. It is defined by induction on the size of t as follows:
If $|t| < m$, then $\Phi_m(t) \triangleq ([\![\,]\!], t, \epsilon)$.
If $|t| \geq m$, then:

$$\Phi_m(\lambda\overline{x}^\tau.t_1) \triangleq \begin{cases}(E_1, \lambda\overline{x}^\tau.u_1, P_1) & \text{if } |\lambda\overline{x}^\tau.u_1| < m\\ ([\![\,]\!][E_1], [\,], (\lambda\overline{x}^\tau.u_1)\cdot P_1) & \text{if } |\lambda\overline{x}^\tau.u_1| = m\end{cases}$$
$$\text{where } (E_1, u_1, P_1) = \Phi_m(t_1).$$

$$\Phi_m(t_1 t_2) \triangleq \begin{cases} ([\![\,]\!][(E_1[\![u_1]\!])(E_2[\![u_2]\!])], [\,], P_1 \cdot P_2) & \text{if } |t_i| \geq m \ (i = 1, 2) \\ (E_1, u_1 t_2, P_1) & \text{if } |t_1| \geq m,\ |t_2| < m,\ |u_1 t_2| < m \\ ([\![\,]\!][E_1], [\,], (u_1 t_2) \cdot P_1) & \text{if } |t_1| \geq m,\ |t_2| < m,\ |u_1 t_2| \geq m \\ (E_2, t_1 u_2, P_2) & \text{if } |t_1| < m,\ |t_1 u_2| < m \\ ([\![\,]\!][E_2], [\,], (t_1 u_2) \cdot P_2) & \text{if } |t_1| < m,\ |t_1 u_2| \geq m \end{cases}$$
$$\text{where } (E_i, u_i, P_i) = \Phi_m(t_i) \quad (i = 1, 2).$$

In the rest of this subsection, we show key properties of $\widehat{\Phi}_m$. We say that a 0/1-context u is *good for m* if u is either (i) a λ-abstraction where $|u| = m$; or (ii) an application $u_1 u_2$ where $|u_j| < m$ for each $j = 1, 2$. By the definition of $\widehat{\Phi}_m(t) = (E, P)$, every component u of P is good for m.

For $m \geq 2$, E and $1 \leq i \leq \mathtt{shn}(E)$, we define $\widehat{U}^m_{E.i}(\delta, \iota, \xi)$, $\Lambda^m_E(\delta, \iota, \xi)$, and $\mathcal{B}^m_n(\delta, \iota, \xi)$ by:

$$\begin{aligned} \widehat{U}^m_{E.i}(\delta, \iota, \xi) &\triangleq \{u \in \mathcal{U}^\nu(\delta, \iota, \xi) \mid u : E.i \text{ and } u \text{ is good for } m\} \\ \Lambda^m_E(\delta, \iota, \xi) &\triangleq \{[E[\![u_1 \cdots u_{\mathtt{shn}(E)}]\!]]_\alpha \mid u_i \in \widehat{U}^m_{E.i}(\delta, \iota, \xi) \text{ for } 1 \leq i \leq \mathtt{shn}(E)\} \\ \mathcal{B}^m_n(\delta, \iota, \xi) &\triangleq \{E \mid (E, P) = \widehat{\Phi}_m(\nu([t]_\alpha)) \text{ for some } [t]_\alpha \in \Lambda^\alpha_n(\delta, \iota, \xi)\}. \end{aligned}$$

Intuitively, $\widehat{U}^m_{E.i}(\delta, \iota, \xi)$ is the set of good contexts that can fill $E.i$, $\Lambda^m_E(\delta, \iota, \xi)$ is the set of terms obtained by filling the second-order holes of E with good contexts, and $\mathcal{B}^m_n(\delta, \iota, \xi)$ is the set of second-order contexts that can be obtained by decomposing a term of size n. The following lemma states the key properties of $\widehat{\Phi}_m$.

Lemma 2 (decomposition). *Let $\delta, \iota, \xi \geq 0$ and $2 \leq m \leq n$.*

1. *$\Lambda^\alpha_n(\delta, \iota, \xi)$ is the disjoint union of $\Lambda^m_E(\delta, \iota, \xi)$'s, i.e.,*

$$\Lambda^\alpha_n(\delta, \iota, \xi) = \biguplus_{E \in \mathcal{B}^m_n(\delta, \iota, \xi)} \Lambda^m_E(\delta, \iota, \xi).$$

   Moreover, $\widehat{\Phi}_m(E[\![P]\!]) = (E, P)$ holds for any $P \in \prod_{1 \leq i \leq \mathtt{shn}(E)} \widehat{U}^m_{E.i}(\delta, \iota, \xi)$.
2. *$m \leq |E.i| < 2m$ $(1 \leq i \leq \mathtt{shn}(E))$ for every $E \in \mathcal{B}^m_n(\delta, \iota, \xi)$.*
3. *$\mathtt{shn}(E) \geq n/4m$ for every $E \in \mathcal{B}^m_n(\delta, \iota, \xi)$.*

The second and third properties say that $\widehat{\Phi}_m$ decomposes each term into sufficiently many contexts of appropriate size.

4.3 Explosive Context

Here, we show that each $\widehat{U}^m_{E.i}(\delta, \iota, \xi)$ contains at least one context that has a very long reduction sequence. To this end, we first prepare a special context Expl^m_k that has a long reduction sequence, and shows that at least one element of $\widehat{U}^m_{E.i}(\delta, \iota, \xi)$ contains Expl^m_k as a subcontext.

We define a "duplicating term" $Dup \triangleq \lambda x^{\circ}.(\lambda x^{\circ}.\lambda *^{\circ}.x)xx$, and $Id \triangleq \lambda x^{\circ}.x$. For two terms t, t' and integer $n \geq 1$, we define the "n-fold application" operation $\uparrow^n$ as $t \uparrow^0 t' \triangleq t'$ and $t \uparrow^n t' \triangleq t(t \uparrow^{n-1} t')$. For an integer $k \geq 2$, we define an order-k term

$$\overline{2}_k \triangleq \lambda f^{\tau(k) \to \tau(k)}.\lambda x^{\tau(k)}.f(fx)$$

where $\tau(i)$ is defined by $\tau(2) \triangleq \mathtt{o}$ and $\tau(i+1) \triangleq \tau(i) \to \tau(i)$.

Definition 5 (explosive context). Let $m \geq 1$ and $k \geq 2$ be integers and let

$$t \triangleq \nu \left(\lambda x^{\circ}.\left((\overline{2}_k \uparrow^m \overline{2}_{k-1})\overline{2}_{k-2} \cdots \overline{2}_2 \, Dup(Id \, x^{\dagger})\right)\right)$$

where $x^{\dagger}$ is just variable x but we put $\dagger$ to refer to the occurrence. We define the *explosive context* $Expl^k_m$ (of m-fold and order k) as the 1-context obtained by replacing the "normalized" variable ${x_1}^{\dagger}$ in t with $[\,]$.

We state key properties of $Expl^k_m$ below. The proof of Item 5 is the same as that in [1]. The other items follow by straightforward calculation.

Lemma 3 (explosive).

1. $\emptyset \vdash Expl^k_m[x_1] : \mathtt{o} \to \mathtt{o}$ *is derivable.*
2. $|Expl^k_m| = 8m + 8k - 3$.
3. $\mathtt{ord}(Expl^k_m[x_1]) = k$, $\mathtt{iar}(Expl^k_m[x_1]) = k$ *and* $\#(\mathbf{V}(Expl^k_m)) = 2$.
4. $Expl^k_m \in \mathcal{U}^{\nu}(\delta, \iota, \xi)$ *if* $\delta,\ \iota \geq k$ *and* $\xi \geq 2$.
5. *If a term t satisfies* $Expl^k_m \preceq t$, *then* $\beta(t) \geq \mathbf{exp}_k(m)$ *holds.*

We show that at least one element of $\widehat{U}^m_{E \bullet i}(\delta, \iota, \xi)$ contains $Expl^k_{m'}$ as a subcontext.

Lemma 4. *Let $\delta, \iota, \xi \geq 2$ be integers and $k = \min\{\delta, \iota\}$. There exist integers $b, c \geq 2$ such that, for any $n \geq 1$, $m' \geq b$, $E \in \mathcal{B}^{cm'}_n(\delta, \iota, \xi)$ and $i \in \{1, \dots, \mathtt{shn}(E)\}$, $\widehat{U}^{cm'}_{E \bullet i}(\delta, \iota, \xi)$ contains u' such that $Expl^k_{m'} \preceq u'$.*

Proof Sketch. We pick $u'' \in \widehat{U}^{cm'}_{E \bullet i}(\delta, \iota, \xi)$ and construct u' by replacing some subcontext u_0 of u'' with a 0/1-context of the form $S^{\circ}[Expl^k_{m'}[u^{\circ}]]$. Here S° and u° adjust the context typing and size of $Expl^k_{m'}$ and these can be obtained by using Proposition 2. The subcontext u_0 is chosen so that the goodness of u'' is preserved by this replacement. □

4.4 Proof Sketch of Theorem 1

We are now ready to prove the main theorem; see Appendix E for details. For readability, we omit the parameters (δ, ι, ξ), and write $\Lambda^{\alpha}_n, \mathcal{U}^{\nu}_n, \Lambda^m_E, \widehat{U}^m_{E \bullet i}$ and $\mathcal{B}^m_n$ for $\Lambda^{\alpha}_n(\delta, \iota, \xi), \mathcal{U}^{\nu}_n(\delta, \iota, \xi), \Lambda^m_E(\delta, \iota, \xi), \widehat{U}^m_{E \bullet i}(\delta, \iota, \xi)$ and $\mathcal{B}^m_n(\delta, \iota, \xi)$ respectively.

Let $\overline{p}(n)$ be the probability that a randomly chosen normalized term t in Λ^{α}_n does *not* contain $Expl^k_{\lceil \log^{(2)}(n) \rceil}$ as a subcontext. By Item 5 of Lemma 3, it suffices to show $\lim_{n\to\infty} \overline{p}(n) = 0$. Let b and c be the constants in Lemma 4 and let $n \geq 2^{2^b}$, $m' = \lceil \log^{(2)}(n) \rceil$ and $m = cm'$. Then $m' \geq \log^{(2)}(n) \geq b$.

By Lemma 2, Λ_n^α can be represented as the disjoint union $\biguplus_{E\in\mathcal{B}_n^m} \Lambda_E^m$. Let $\overline{\Lambda}_E^m$ be the subset of Λ_E^m that does not contain $\mathit{Expl}_{m'}^k$ as a subcontext. By Lemma 4, each of $\widehat{U}_{E.i}^m$ contains at least one element that has $\mathit{Expl}_{m'}^k$ as a subcontext. Furthermore, since $m \leq |E.i| < 2m$, we have $\#\left(\widehat{U}_{E.i}^m\right) \leq \#\left(\mathcal{U}_{2m+d}^\nu\right)$ for some constant d (see Appendix H). Thus, we have

$$\frac{\#\left(\overline{\Lambda}_E^m\right)}{\#(\Lambda_E^m)} \leq \prod_{1\leq i\leq \mathtt{shn}(E)} \left(1 - \frac{1}{\#\left(\widehat{U}_{E.i}^m\right)}\right) \leq \left(1 - \frac{1}{\#\left(\mathcal{U}_{2m+d}^\nu\right)}\right)^{\mathtt{shn}(E)}$$
$$\leq \left(1 - \frac{1}{\#\left(\mathcal{U}_{2m+d}^\nu\right)}\right)^{\frac{n}{4m}} \quad (\because \text{Item 3 of Lemma 2}).$$

Let $q(n)$ be the rightmost expression. Then we have

$$\begin{aligned}\overline{p}(n) =& \frac{\sum_{E\in\mathcal{B}_n^m} \#\left(\overline{\Lambda}_E^m\right)}{\sum_{E\in\mathcal{B}_n^m} \#(\Lambda_E^m)} \leq \frac{\sum_{E\in\mathcal{B}_n^m} (q(n)\#(\Lambda_E^m))}{\sum_{E\in\mathcal{B}_n^m} \#(\Lambda_E^m)}\\ =& \frac{q(n)\sum_{E\in\mathcal{B}_n^m} \#(\Lambda_E^m)}{\sum_{E\in\mathcal{B}_n^m} \#(\Lambda_E^m)} = q(n) \leq \left(1 - \frac{1}{c'\overline{\gamma}(\delta,\iota,\xi)^{2m}}\right)^{\frac{n}{4m}} \quad (\because \text{Lemma 1})\end{aligned}$$

for sufficiently large n. Finally, we can conclude that

$$\overline{p}(n) \leq \left(1 - \frac{1}{c'\overline{\gamma}(\delta,\iota,\xi)^{2c\lceil\log^{(2)}(n)\rceil}}\right)^{\frac{n}{4c\lceil\log^{(2)}(n)\rceil}} \longrightarrow 0 \quad (\text{as } n \longrightarrow \infty)$$

(see Appendix I for the last convergence) as required. □

5 Related Work

As mentioned in Section 1, there are several pieces of work on probabilistic properties of untyped λ-terms [2,3,4]. David et al. [2] have shown that almost all untyped λ-terms are strongly normalizing, whereas the result is opposite for terms expressed in SK combinators. Their former result implies that untyped λ-terms do *not* satisfy the infinite monkey theorem, i.e., for any term t, the probability that a randomly chosen term of size n contains t as a subterm tends to *zero*. Bendkowski et al. [4] proved that almost all terms in de Brujin representation are not strongly normalizing, by regarding the size of an index i is $i + 1$, instead of the constant 1. The discrepancies among those results suggest that this kind of probabilistic property is quite fragile and depends on the definition of the syntax and the size of terms. Thus, the setting of our paper, especially the assumption on the boundedness of internal arities and the number of variables is a matter of debate, and it would be interesting to study how the result changes for different assumptions.

We are not aware of similar studies on *typed* λ-terms. In fact, in their paper about combinatorial aspects of λ-terms, Grygiel and Lescanne [3] pointed out that the combinatorial study of typed λ-terms is difficult, due to the lack of (simple) recursive definition of typed terms. In the present paper, we have avoided the difficulty by making the assumption on the boundedness of internal arities and the number of variables (which is, as mentioned above, subject to a debate though).

In a larger context, our work may be viewed as an instance of the studies of average-case complexity ([19], chapter 10), which discusses "typical-case feasibility". We are not aware of much work on the average-case complexity of problems with hyper-exponential complexity.

6 Conclusion

We have shown that almost every simply-typed λ-term of order k has a β-reduction sequence as long as $(k-2)$-fold exponential in the term size, under a certain assumption. To our knowledge, this is the first result of this kind for typed λ-terms. A lot of questions are left for future work, such as (i) whether our assumption (on the boundness of arities and the number of variables) is reasonable, and how the result changes for different assumptions, (ii) whether our result is optimal (e.g., whether almost every term has a k-fold exponentially long reduction sequence), and (iii) whether similar results hold for Terui's decision problems [14] and/or the higher-order model checking problem [6].

Acknowledgment We would like to thank anonymous referees for useful comments. This work was supported by JSPS KAKENHI Grant Number JP15H05706.

References

1. Beckmann, A.: Exact bounds for lengths of reductions in typed lambda-calculus. Journal of Symbolic Logic **66**(3) (2001) 1277–1285
2. David, R., Grygiel, K., Kozic, J., Raffalli, C., Theyssier, G., Zaionc, M.: Asymptotically almost all λ-terms are strongly normalizing. Logical Methods in Computer Science **9**(2) (2013)
3. Grygiel, K., Lescanne, P.: Counting and generating lambda terms. Journal of Functional Programming **23**(05) (2013) 594–628
4. Bendkowski, M., Grygiel, K., Lescanne, P., Zaionc, M.: A Natural Counting of Lambda Terms. In: SOFSEM 2016: Theory and Practice of Computer Science: 42nd International Conference on Current Trends in Theory and Practice of Computer Science, Harrachov, Czech Republic, January 23-28, 2016, Proceedings. Springer Berlin Heidelberg, Berlin, Heidelberg (2016) 183–194
5. Knapik, T., Niwinski, D., Urzyczyn, P.: Higher-order pushdown trees are easy. In: FoSSaCS 2002. Volume 2303., Springer (2002) 205–222

6. Ong, C.H.L.: On model-checking trees generated by higher-order recursion schemes. In: LICS 2006, IEEE Computer Society Press (2006) 81–90
7. Kobayashi, N.: Model-checking higher-order functions. In: Proceedings of PPDP 2009, ACM Press (2009) 25–36
8. Broadbent, C.H., Kobayashi, N.: Saturation-based model checking of higher-order recursion schemes. In: Proceedings of CSL 2013. Volume 23 of LIPIcs. (2013) 129–148
9. Ramsay, S., Neatherway, R., Ong, C.H.L.: An abstraction refinement approach to higher-order model checking. In: Proceedings of POPL 2014. (2014)
10. Kobayashi, N.: Types and higher-order recursion schemes for verification of higher-order programs, ACM Press (2009) 416–428
11. Kobayashi, N., Sato, R., Unno, H.: Predicate abstraction and CEGAR for higher-order model checking, ACM Press (2011) 222–233
12. Ong, C.H.L., Ramsay, S.: Verifying higher-order programs with pattern-matching algebraic data types, ACM Press (2011) 587–598
13. Sato, R., Unno, H., Kobayashi, N.: Towards a scalable software model checker for higher-order programs. In: Proceedings of PEPM 2013, ACM Press (2013) 53–62
14. Terui, K.: Semantic evaluation, intersection types and complexity of simply typed lambda calculus. In: 23rd International Conference on Rewriting Techniques and Applications (RTA'12). Volume 15 of LIPIcs., Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2012) 323–338
15. Mairson, H.G.: Deciding ML typability is complete for deterministic exponential time. In: POPL, ACM Press (1990) 382–401
16. Kfoury, A.J., Tiuryn, J., Urzyczyn, P.: ML typability is DEXTIME-complete. In: Proceedings of CAAP '90. Volume 431 of Lecture Notes in Computer Science., Springer (1990) 206–220
17. Heintze, N., McAllester, D.A.: Linear-time subtransitive control flow analysis. (1997) 261–272
18. Flajolet, P., Sedgewick, R.: Analytic Combinatorics. 1 edn. Cambridge University Press, New York, NY, USA (2009)
19. Goldreich, O.: Computational Complexity: A Conceptual Perspective. Cambridge University Press (2008)
20. Chomsky, N., Schützenberger, M.: The algebraic theory of context-free languages*. In Braffort, P., Hirschberg, D., eds.: Computer Programming and Formal Systems. Volume 35 of Studies in Logic and the Foundations of Mathematics. Elsevier (1963) 118–161
21. Flajolet, P., Sedgewick, R.: Analytic combinatorics : functional equations, rational and algebraic functions. Research Report RR-4103, INRIA (2001)

A Proof of Proposition 1

Proof. It is trivial that the image of $\pi_{\langle \Gamma;\tau\rangle}$ is contained in $\Lambda^\alpha(\langle \Gamma;\tau\rangle,\delta,\iota,\xi)$ and π preserves size. We omit the proof of the unambiguity, as it is similar to the proof of the injectivity below.

The injectivity, i.e., $e(T)\sim_\alpha e(T')$ implies $T=T'$ for $T,T'\in\mathcal{L}\left(N_{\langle \Gamma;\tau\rangle}\right)$, is shown by induction on the length of the leftmost reduction sequence of T and by case analysis of the rewriting rule used first for the reduction sequence. We use simultaneous induction on all non-terminals $N_{\langle \Gamma;\tau\rangle}$.

Cases of variables and λ-abstraction: These cases are clear because each rewriting rule of this kind is determined by the non-terminal in the left hand side and the terminal in the right hand side.

Case of $N_{\langle \Gamma;\tau\rangle}\longrightarrow @(N_{\langle \Gamma_1;\sigma\to\tau\rangle},N_{\langle \Gamma_2;\sigma\rangle})$ where $\Gamma=\Gamma_1\cup\Gamma_2$: Let the first rewriting rule for T' be $N_{\langle \Gamma;\tau\rangle}\longrightarrow @(N_{\langle \Gamma_1';\sigma'\to\tau\rangle},N_{\langle \Gamma_2';\sigma'\rangle})$ where $\Gamma=\Gamma_1'\cup\Gamma_2'$. Let $T=@(T_1,T_2)$ and $T'=@(T_1',T_2')$. Since

$$e(T_1)e(T_2)=e(T)\sim_\alpha e(T')=e(T_1')e(T_2'),$$

we have $e(T_i)\sim_\alpha e(T_i')$ and hence $\mathbf{FV}(e(T_i))=\mathbf{FV}(e(T_i'))$ for each $i=1,2$. Now we have $\mathbf{FV}(e(T_i))=\mathrm{Dom}(\Gamma_i)$ and $\mathbf{FV}(e(T_i'))=\mathrm{Dom}(\Gamma_i')$ as our typing rules do not allow weakening; hence $\mathrm{Dom}(\Gamma_i)=\mathrm{Dom}(\Gamma_i')$. Since $\Gamma_i,\Gamma_i'\subseteq\Gamma$, we obtain $\Gamma_i=\Gamma_i'$, for each $i=1,2$. Then, we have also $\sigma=\sigma'$. Hence by induction hypothesis, $T_i=T_i'$ for $i=1,2$, and therefore $T=T'$.

Next we show the surjectivity, i.e., for any $[t]_\alpha\in\Lambda^\alpha(\langle \Gamma;\tau\rangle,\delta,\iota,\xi)$ there exists $T\in\mathcal{L}\left(N_{\langle \Gamma;\tau\rangle}\right)$ such that $e(T)\sim_\alpha t$. For δ,ι,ξ with $X_\xi=\{x_1,\cdots,x_\xi\}$ and $N_{\langle \Gamma;\tau\rangle}\in\mathcal{N}(\delta,\iota,\xi)$, we define a "renaming" function $\rho^{(\delta,\iota,\xi)}_{\langle \Gamma;\tau\rangle}$ ($\rho_{\langle \Gamma;\tau\rangle}$, ρ for short) from $\{t\mid [t]_\alpha\in\Lambda^\alpha(\langle \Gamma;\tau\rangle,\delta,\iota,\xi)\}$ to $\mathcal{L}\left(N_{\langle \Gamma;\tau\rangle}\right)$ by induction on t so that we obtain $\pi(\rho(t))=[t]_\alpha$. The induction is simultaneous on all $\langle \Gamma;\tau\rangle$.

$$\rho_{\langle\{x:\tau\};\tau\rangle}(x)\triangleq x$$

$$\rho_{\langle \Gamma;\sigma\to\tau\rangle}(\lambda *^\sigma.t)\triangleq\lambda *^\sigma(\rho_{\langle \Gamma;\tau\rangle}(t))$$
where we have $\Gamma\vdash t:\tau$

$$\rho_{\langle \Gamma;\sigma\to\tau\rangle}(\lambda x^\sigma.t)\triangleq\lambda *^\sigma(\rho_{\langle \Gamma;\tau\rangle}(t))\qquad(\text{if } x\notin\mathbf{FV}(t))$$
where we have $\Gamma\vdash t:\tau$

$$\rho_{\langle \Gamma;\sigma\to\tau\rangle}(\lambda x^\sigma.t)\triangleq\lambda x_i^\sigma(\rho_{\langle \Gamma\cup\{x_i:\sigma\};\tau\rangle}(t[x_i/x]))\qquad(\text{if } x\in\mathbf{FV}(t))$$
where we have $\Gamma\cup\{x:\sigma\}\vdash t:\tau$ and $i\triangleq\min\{j\mid x_j\notin\mathrm{Dom}(\Gamma)\}$

$$\rho_{\langle \Gamma_1\cup\Gamma_2;\tau\rangle}(t_1t_2)\triangleq @(\rho_{\langle \Gamma_1;\sigma\to\tau\rangle}(t_1),t_2\rho_{\langle \Gamma_2;\sigma\rangle}(t_2))$$
where we have $\Gamma_1\vdash t_1:\sigma\to\tau$ and $\Gamma_2\vdash t_2:\sigma$

Then, $\pi(\rho(t))=[t]_\alpha$ follows by straightforward induction on the size of t. □

B Proof of Proposition 2

Proposition 2 follows from the following three lemmas and the property in the definition of $\mathcal{N}^{\emptyset}(\delta,\iota,\xi)$; the non-linearity is clear. We note that, if a regular tree grammar $\mathcal{G}$ is irreducible, $\mathcal{G}$ is aperiodic if and only if there exist a non-terminal $N \in \mathcal{N}$ and an integer $m > 0$ such that $\#(\mathcal{L}_n(N)) > 0$ for any $n > m$.[4]

Lemma 5. *Let $\delta,\iota \geq 2$ and $\xi \geq 1$ be integers. Then for any non-terminal $N_{\langle\emptyset;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)$, $N_{\langle\emptyset;\tau\rangle}$ is reachable from $N_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle}$.*

Proof. Let $\tau = \tau_1\to\dots\to\tau_n\to\mathsf{o}$ and $\tau_i = \tau^i_1\to\dots\to\tau^i_{n_i}\to\mathsf{o}$ for $i = 1,\dots,n$. For $i = 1,\dots,n$, we define $T_{\tau_i} \triangleq \lambda *^{\tau^i_1}(\dots\lambda *^{\tau^i_{n_i}}(x_1)\dots)$, and then

$$N_{\langle x_1:\mathsf{o};\tau_i\rangle} \longrightarrow \lambda *^{\tau^i_1}(N_{\langle x_1:\mathsf{o};\tau^i_2\to\dots\to\tau^i_{n_i}\to\mathsf{o}\rangle}) \longrightarrow^* \lambda *^{\tau^i_1}(\dots\lambda *^{\tau^i_{n_i}}(N_{\langle x_1:\mathsf{o};\mathsf{o}\rangle})\dots)$$
$$\longrightarrow \lambda *^{\tau^i_1}(\dots\lambda *^{\tau^i_{n_i}}(x_1)\dots) = T_{\tau_i}$$

and hence

$$\begin{aligned}N_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle} &\longrightarrow \lambda x^{\mathsf{o}}_1(N_{\langle x_1:\mathsf{o};\mathsf{o}\rangle})\\ &\longrightarrow \lambda x^{\mathsf{o}}_1(@(N_{\langle x_1:\mathsf{o};\tau_1\to\mathsf{o}\rangle}, N_{\langle x_1:\mathsf{o};\tau_1\rangle})) \longrightarrow^* \lambda x^{\mathsf{o}}_1(@(N_{\langle x_1:\mathsf{o};\tau_1\to\mathsf{o}\rangle}, T_{\tau_1}))\\ &\longrightarrow^* \lambda x^{\mathsf{o}}_1(@(@(\dots @(N_{\langle\emptyset;\tau\rangle}, T_{\tau_1}),\dots), T_{\tau_n})).\end{aligned}$$

□

Lemma 6. *Let $\delta,\iota,\xi \geq 2$ be integers. Then for any non-terminal $N_{\langle\Gamma;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)$, $N_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle}$ is reachable from $N_{\langle\Gamma;\tau\rangle}$.*

Proof. By the definition of $\mathcal{N}^{\emptyset}(\delta,\iota,\xi)$ (and $\mathcal{N}(\delta,\iota,\xi)$), there exists t such that $\Gamma \vdash t : \tau$. Let $T \triangleq \pi^{-1}([t]_\alpha) \in \mathcal{L}\left(N_{\langle\Gamma;\tau\rangle}\right)$. Now T has at least one variable, say x, and let x has type σ. Since $T \in \mathcal{L}\left(N_{\langle\Gamma;\tau\rangle}\right)$, there exists T' such that T' contains just one occurrence of $N_{\langle\{x:\sigma\};\sigma\rangle}$ and

$$N_{\langle\Gamma;\tau\rangle} \longrightarrow^* T' \longrightarrow T'[x/N_{\langle\{x:\sigma\};\sigma\rangle}] = T.$$

If σ is a function type, say $\sigma = \sigma_1\to\sigma_2$, then since $\xi > 1$, we can apply "η-expansion":

$$\begin{aligned}T' &\longrightarrow T'[\lambda x^{\sigma_1}_i(N_{\langle\{x:\sigma,x_i:\sigma_1\};\sigma_2\rangle})/N_{\langle\{x:\sigma\};\sigma\rangle}] \qquad (i \triangleq \min\{j \mid x_j \notin \mathrm{Dom}(\{x:\sigma\})\})\\ &\longrightarrow T'[\lambda x^{\sigma_1}_i(@(N_{\langle\{x:\sigma\};\sigma\rangle}, N_{\langle\{x_i:\sigma_1\};\sigma_1\rangle}))/N_{\langle\{x:\sigma\};\sigma\rangle}].\end{aligned}$$

Next we focus on $N_{\langle\{x_i:\sigma_1\};\sigma_1\rangle}$ and repeat similar reductions until the types $\sigma,\sigma_1,\dots$ becomes o; note that this ends in finite steps since the orders of $\sigma,\sigma_1,\dots$ are decreasing. Thus we have a tree T_0 that contains $N_{\langle\{y:\mathsf{o}\};\mathsf{o}\rangle}$ such that $N_{\langle\Gamma;\tau\rangle} \longrightarrow^* T_0$. Then we have

$$N_{\langle\Gamma;\tau\rangle} \longrightarrow^* T_0 \longrightarrow T_0[@(N_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle}, N_{\langle\{y:\mathsf{o}\};\mathsf{o}\rangle})/N_{\langle\{y:\mathsf{o}\};\mathsf{o}\rangle}]$$

and thus $N_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle}$ is reachable from $N_{\langle\Gamma;\tau\rangle}$. □

[4] *cf.* Footnote 10 in p. 483 of [18]

Lemma 7. *Let $\delta, \iota \geq 2$ and $\xi \geq 1$ be integers. Then for any integer $n \geq 5$, the non-terminal $N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle}$ of $\mathcal{G}^{\emptyset}(\delta, \iota, \xi)$ satisfies $\mathcal{L}_n\left(N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle}\right) \neq \emptyset$.*

Proof. For simplicity, in this proof we identify trees with terms, which can be justified by the size-preserving functions e in Section 3.1 and ρ in Appendix A. The proof proceeds by induction on n. For $n = 5, 6, 7$, the following terms

$$\lambda x^{\mathsf{o}}.(\lambda *^{\mathsf{o}}.x)x$$
$$\lambda x^{\mathsf{o}}.(\lambda *^{\mathsf{o} \to \mathsf{o}}.x), \lambda *^{\mathsf{o}}.x))$$
$$\lambda x^{\mathsf{o}}.(\lambda *^{\mathsf{o} \to \mathsf{o} \to \mathsf{o}}.x)\lambda *^{\mathsf{o}}.\lambda *^{\mathsf{o}}.x$$

belong to $\mathcal{L}_n\left(N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle}\right)$, respectively. Next, for any $t \in \mathcal{L}_n\left(N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle}\right)$, we have $\lambda x^{\mathsf{o}}.tx \in \mathcal{L}_{n+3}\left(N_{\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle}\right)$. □

In Appendix G, to prove Lemma 4, we will use the following Corollary 1 which is a consequence of Proposition 2.

The following two lemmas are nothing but the irreducibility and the aperiodicity in Proposition 2, respectively, except that we here consider terms/contexts rather than trees. The proofs are given by pulling-back Proposition 2 through π and π^{-1}.

Lemma 8 (irreducibility). *Let $\delta, \iota, \xi \geq 2$ be integers. For $N_{\langle \Gamma; \tau \rangle}, N_{\langle \Gamma'; \tau' \rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$, there exists a $\langle \Gamma'; \tau' \rangle \Rightarrow \langle \Gamma; \tau \rangle$-context $S \in \mathcal{S}^{\nu}(\delta, \iota, \xi)$.*

Lemma 9 (aperiodicity). *Let $\delta, \iota, \xi \geq 2$ be integers. For any integer $n \geq 5$, there exists $\langle \emptyset; \mathsf{o} \to \mathsf{o} \rangle$-term t such that $[t]_\alpha \in \Lambda_n^\alpha(\delta, \iota, \xi)$.*

Corollary 1 (of Lemma 8 and 9). *Let $\delta, \iota, \xi \geq 2$ be integers. There exists a constant $d_0 \geq 1$ such that for any $n \geq d_0$ and for any $N_{\langle \Gamma; \tau \rangle}, N_{\langle \Gamma'; \tau' \rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$, there exist a $\langle \Gamma'; \tau' \rangle \Rightarrow \langle \Gamma; \tau \rangle$-context $S \in \mathcal{S}_n^{\nu}(\delta, \iota, \xi)$ and $\langle \Gamma; \tau \rangle$-term t such that $[t]_\alpha \in \Lambda_n^\alpha(\delta, \iota, \xi)$.*

C Analytic Combinatorics

C.1 Chomsky-Schützenberger Enumeration Theorem for Regular Tree Grammars

We have an obvious translation from a regular tree grammar $\mathcal{G}$ to a context-free grammar G, just by replacing every rewriting rule $N \longrightarrow_{\mathcal{G}} a(N_1, \cdots, N_{\Sigma(a)})$ by $N \longrightarrow_G aN_1 \cdots N_{\Sigma(a)}$, so that the translation preserves unambiguity. Also, this translation preserves size: i.e., the size of trees and the size of words. Therefore, the generating functions for an unambiguous regular tree grammar $\mathcal{G}$ are equal to those for the corresponding context-free grammar G.

Also, there is the well-known translation from a context-free grammar to a system of polynomial equations: it replaces each non-terminal N by the corresponding indeterminate $\overline{\mathbf{F}}_N$, and each terminal a by the variable z, each union $\cup$ by $+$, and each concatenation $\cdot$ by $\times$.

The translation given in Section C.2 is obtained as the composite of the above two translations. The well-known Chomsky-Schützenberger enumeration theorem [20][21, Proposition 8.7] states that the generating function of an unambiguous context-free grammar G is a solution of the corresponding system of polynomial equations.

In the literature of analytic combinatorics [18,21], the notion of a *context-free specification* is discussed, which is just a syntactical representation of a system of polynomial equations. The above translation from context-free grammars to systems of polynomial equations can also be seen as that to context-free specifications (*cf.* Section 6.2 of [21]). Irreducibility and aperiodicity are defined also for context-free specifications in the same way [18, Definition VII. 5], and the translation from regular tree grammars to context-free specifications preserves irreducibility and aperiodicity.

C.2 Polynomial Equations for the Generating Function of $\Lambda^\alpha(\delta, \iota, \xi)$

For a sequence $(a_n)_{n\in\mathbb{N}}$ of numbers, *the (ordinary) generating function of* $(a_n)_n$ is the formal power series: $\sum_{n=0}^{\infty} a_n z^n$. For a set S of α-equivalence classes of terms or of trees, *the generating function of* S is the generating function of $\left(\#\left(\varphi^{-1}(n)\right)\right)_n$ where φ is the size function. A function f that is analytic at 0 is regarded as the generating function of $(f^{(n)}(0)/n!)_n$ by the Taylor expansion. We let $[z^n]\mathbf{F}(z)$ denote the operation of extracting the coefficient of z^n of the generating function $\mathbf{F}(z)$. We often simply write $\mathbf{F}$ instead of $\mathbf{F}(z)$.

For a regular tree grammar $\mathcal{G}$ and a non-terminal N of $\mathcal{G}$, we write $\mathbf{F}_N^{\mathcal{G}}$ (or $\mathbf{F}_N$ for short) for the generating function of $\mathcal{L}(N)$. Also, for δ, ι and ξ, we write $\mathbf{F}_{(\delta,\iota,\xi)}$ for the generating function of $\Lambda^\alpha(\delta, \iota, \xi)$. By Proposition 1, $\mathbf{F}_{(\delta,\iota,\xi)}$ is also the generating function of $\bigcup_{N_{\langle\emptyset;\tau\rangle}\in\mathcal{N}(\delta,\iota,\xi)} \mathcal{L}\left(N_{\langle\emptyset;\tau\rangle}\right) = \bigcup_{N_{\langle\emptyset;\tau\rangle}\in\mathcal{N}^{\emptyset}(\delta,\iota,\xi)} \mathcal{L}\left(N_{\langle\emptyset;\tau\rangle}\right)$.

Now we explain a well-known translation [20] of a regular tree grammar $\mathcal{G}$ to a system of polynomial equations. The definition of the translation itself is simple: Let $\mathcal{G} = (\Sigma, \mathcal{N}, \mathcal{R})$ be a tree grammar. For each non-terminal N, we prepare an indeterminate $\overline{\mathbf{F}}_N$, and let

$$N \longrightarrow a_i(N_1^i, \cdots, N_{\Sigma(a_i)}^i) \qquad (i = 1, \ldots, m_N)$$

be all the rewriting rules of N. Then we define a polynomial P_N as follows:

$$\overline{\mathbf{F}}_N = \sum_{i=1}^{m_N} z\overline{\mathbf{F}}_{N_1^i} \ldots \overline{\mathbf{F}}_{N_{\Sigma(a_i)}^i}$$

where the formal power series $z(= 0 + z + 0z^2 + \ldots)$ is a coefficient of this polynomial. Thus we have obtained a system of polynomial equations $(P_N)_{N\in\mathcal{N}}$.

Chomsky-Schützenberger enumeration theorem [20][21, Proposition 8.7] states that, if $\mathcal{G}$ is unambiguous, $(\mathbf{F}_N)_{N\in\mathcal{N}}$ becomes a solution to this system[5].

[5] We can translate a regular tree grammar also to a context-free grammar or a *context-free specification* [18], which, then, can be (trivially) transformed to the system of polynomial equations. We explain this in Appendix C.

For $\mathcal{G}(1,1,1)$ in Example 1, let us see how this translation works. Applying the translation to $\mathcal{G}(1,1,1)$, we have:

$$\begin{cases} \overline{\mathbf{F}}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle} = z\overline{\mathbf{F}}_{\langle\{x_1:\mathsf{o}\};\mathsf{o}\rangle} \\ \overline{\mathbf{F}}_{\langle\{x_1:\mathsf{o}\};\mathsf{o}\rangle} = z + z\left(\overline{\mathbf{F}}_{\langle\{x_1:\mathsf{o}\};\mathsf{o}\to\mathsf{o}\rangle} \times \overline{\mathbf{F}}_{\langle\{x_1:\mathsf{o}\};\mathsf{o}\rangle}\right) + z\left(\overline{\mathbf{F}}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle} \times \overline{\mathbf{F}}_{\langle\{x_1:\mathsf{o}\};\mathsf{o}\rangle}\right) \\ \overline{\mathbf{F}}_{\langle\{x_1:\mathsf{o}\};\mathsf{o}\to\mathsf{o}\rangle} = z\overline{\mathbf{F}}_{\langle\{x_1:\mathsf{o}\};\mathsf{o}\rangle} \end{cases}$$

For this small example, we can solve this system of equations algebraically as follows. From the above system of equations, one can obtain a one-variable polynomial $2z(\overline{\mathbf{F}}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle})^2 - \overline{\mathbf{F}}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle} + z^2 = 0$ by some suitable substitutions. We have two solutions to this equation:

$$\frac{1+\sqrt{1-8z^3}}{4z} \quad \text{and} \quad \frac{1-\sqrt{1-8z^3}}{4z}.$$

Since $\#(\Lambda_2^\alpha(1,1,1)) = \#(\{[\lambda x^{\mathsf{o}}.x]_\alpha\}) = 1$, $\mathbf{F}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle}$ must be the second solution. Then

$$\mathbf{F}_{(1,1,1)} = \mathbf{F}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle} = \frac{1-\sqrt{1-8z^3}}{4z} = z^2 + 2z^5 + 8z^8 + 40z^{11} + 224z^{14} + \cdots .$$

and

$$\#(\Lambda_n^\alpha(1,1,1)) = [z^n]\mathbf{F}_{(1,1,1)} = [z^n]\mathbf{F}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle} = \mathrm{O}(2^n)$$

where the last equality is obtained from $\mathbf{F}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle} = \frac{1-\sqrt{1-8z^3}}{4z}$ and a well-known fact in analytic combinatorics [18] (i.e., the exponential growth rate 2 (2 in $\mathrm{O}(2^n)$) is given as the multiplicative inverse of the *dominant singularity* (the singularity closest to the origin) of $\mathbf{F}_{\langle\emptyset;\mathsf{o}\to\mathsf{o}\rangle}$).

C.3 Deduction of Theorem 2

In this subsection we fill the gap between Theorem VII.5 in [18] and Theorem 2. Theorem VII.5 states asymptotic behavior of an irreducible and aperiodic context-free specification, and as its direct consequence via the above translation, we have the following corollary.

Corollary 2 ([18]). *Let $\mathcal{G} = (\Sigma, \mathcal{N}, \mathcal{R})$ be an unambiguous, irreducible and aperiodic regular tree grammar. Then for any non-terminal $N \in \mathcal{N}$, there exist constants $\gamma_N(\mathcal{G}) > 1$ and $C_N(\mathcal{G}) > 0$ such that*

$$\#(\mathcal{L}_n(N)) \sim C_N(\mathcal{G})\gamma_N(\mathcal{G})^n n^{-3/2}.$$

The above corollary is weaker than Theorem 2: in Theorem 2 the constant $\gamma(\mathcal{G})$ depends on only $\mathcal{G}$, while in Corollary 2 the constant $\gamma_N(\mathcal{G})$ depends on $\mathcal{G}$ and N.

This gap can be easily filled as follows. Below we use the notion of substitution of context-trees in Section 4.1. Let $N_1, N_2 \in \mathcal{N}$. Since $\mathcal{G}$ is irreducible, by Lemma, there exists a $N_1 \Rightarrow N_2$-context-tree $\widehat{T}$. Since the function $\widehat{T}[-] : \mathcal{L}(N_1) \to \mathcal{L}(N_2)$

is injective, and $|\widehat{T}[T]| = |\widehat{T}| + |T|$ for any $T \in \mathcal{L}(N_1)$, we have $\#(\mathcal{L}_n(N_1)) \leq \#\left(\mathcal{L}_{|\widehat{T}|+n}(N_2)\right)$. Hence for sufficiently large n, the following holds:

$$C_{N_1}(\mathcal{G})\gamma_{N_1}(\mathcal{G})^n n^{-3/2} \leq C_{N_2}(\mathcal{G})\gamma_{N_2}(\mathcal{G})^{|\widehat{T}|+n}(|\widehat{T}|+n)^{-3/2}.$$

Then $\gamma_{N_1}(\mathcal{G}) \leq \gamma_{N_2}(\mathcal{G})$ must hold. By symmetry, we have $\gamma_{N_2}(\mathcal{G}) \leq \gamma_{N_1}(\mathcal{G})$, and so $\gamma_{N_1}(\mathcal{G}) = \gamma_{N_2}(\mathcal{G})$. Thus all the non-terminals in $\mathcal{N}$ have the same growth rate, which is $\gamma(\mathcal{G})$.

D Proof of Lemma 1 (upper bound of $\#\big(\mathcal{U}^{\nu}_n(\delta,\iota,\xi)\big)$)

To prove Lemma 1, we define a regular tree grammar that represents all subcontexts of $\Lambda^{\alpha}(\delta,\iota,\xi)$.

Definition 6 (Generalized grammar of $\Lambda^{\alpha}(\delta,\iota,\xi)$ and its analysis). Let $\delta,\iota,\xi \geq 0$ be integers. The *generalized grammar*

$$\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi) = (\overline{\Sigma}(\delta,\iota,\xi), \mathcal{N}^{\emptyset}(\delta,\iota,\xi), \overline{\mathcal{R}}^{\emptyset}(\delta,\iota,\xi))$$

of $\Lambda^{\alpha}(\delta,\iota,\xi)$ is defined as an extension of $\mathcal{G}^{\emptyset}(\delta,\iota,\xi)$ by adding new symbols and rules as follows:

– $\overline{\Sigma}(\delta,\iota,\xi) \triangleq \Sigma(\delta,\iota,\xi) \cup \{h_{\langle\Gamma;\tau\rangle} \mapsto 0 \mid N_{\langle\Gamma;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)\}$.
– $\overline{\mathcal{R}}^{\emptyset}(\delta,\iota,\xi) \triangleq \mathcal{R}^{\emptyset}(\delta,\iota,\xi) \cup \{N_{\langle\Gamma;\tau\rangle} \longrightarrow h_{\langle\Gamma;\tau\rangle} \mid N_{\langle\Gamma;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)\}$.

We call a terminal symbol of the form $h_{\langle\Gamma;\tau\rangle}$ a *hole-terminal*. Intuitively, $h_{\langle\Gamma;\tau\rangle}$ represents a hole $[\,]$ with typing annotation $\langle\Gamma;\tau\rangle$, and so there is the obvious embedding $e_h^{(\delta,\iota,\xi)}$ (e_h for short) from $\overline{\Sigma}(\delta,\iota,\xi)$-trees to contexts such that $e_h(h_{\langle\Gamma;\tau\rangle}) \triangleq [\,]$. Note that this embedding e_h drops the type annotations of hole-terminals. We call a tree $T \in \mathcal{L}\left(N_{\langle\Gamma;\tau\rangle}\right)$ where exactly one hole-terminal occurs, say $h_{\langle\Gamma';\tau'\rangle}$, a *$\langle\Gamma';\tau'\rangle \Rightarrow \langle\Gamma;\tau\rangle$-tree*. It is easy to show that, for a $\langle\Gamma';\tau'\rangle \Rightarrow \langle\Gamma;\tau\rangle$-tree T, the context $e_h(\langle\Gamma';\tau'\rangle \Rightarrow \langle\Gamma;\tau\rangle)$ is an inhabitant of $\langle\Gamma';\tau'\rangle \Rightarrow \langle\Gamma;\tau\rangle$. In the previous section, we defined the size of a tree T as the number of nodes and leaves of T. In this section, however, we re-define the size of a $\overline{\Sigma}(\delta,\iota,\xi)$-tree T, written as $|T|$, is the number of nodes and *non-hole-terminal leaves* of T. This definition of the size is necessary for a size-consistency between $\overline{\Sigma}(\delta,\iota,\xi)$-trees and contexts. For $\delta,\iota,\xi \geq 0$ and $n \geq 0$, we define

$$\mathcal{L}_n\left(\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi)\right) \quad \triangleq \bigcup_{N_{\langle\Gamma;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)} \mathcal{L}_n\left(\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi), N_{\langle\Gamma;\tau\rangle}\right).$$

Note that above all non-terminals $N_{\langle\Gamma;\tau\rangle}$ are involved rather than $N_{\langle\emptyset;\tau\rangle}$, since the purpose of this grammar is to (approximately) represent all the *sub*-contexts of closed terms.

We can prove that $\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi)$ has similar properties with of $\mathcal{G}^{\emptyset}(\delta,\iota,\xi)$. In general, adding new symbol of the form a_N for each non-terminal N and adding new rule of the form $N \longrightarrow a_N$ does not change unambiguity, irreducibility and aperiodicity. Item 3 in the following lemma is followed from Theorem VII.5 in [18] (see Appendix C).

Lemma 10. *Let $\delta, \iota, \xi \geq 2$ be integers. Then the following holds.*

1. *$\overline{\mathcal{G}}^{\emptyset}(\delta, \iota, \xi)$ is unambiguous, irreducible and aperiodic.*
2. *There exists a constant $\overline{\gamma}(\delta, \iota, \xi) \geq \gamma(\delta, \iota, \xi)$ such that, for any non-terminal $N_{\langle \Gamma;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$ of $\overline{\mathcal{G}}^{\emptyset}(\delta, \iota, \xi)$*

$$\#\Big(\mathcal{L}_n\left(\overline{\mathcal{G}}^{\emptyset}(\delta, \iota, \xi), N_{\langle \Gamma;\tau\rangle}\right)\Big) \sim \overline{C}^{(\delta,\iota,\xi)}_{N_{\langle \Gamma;\tau\rangle}} \overline{\gamma}(\delta, \iota, \xi)^n n^{-3/2}$$

holds for some constant $\overline{C}^{(\delta,\iota,\xi)}_{N_{\langle \Gamma;\tau\rangle}} > 0$ determined by δ, ι, ξ and $N_{\langle \Gamma;\tau\rangle}$.

Now, to prove Lemma 1, it suffices to show that the following lemma.

Lemma 11. *For $\delta, \iota, \xi \geq 0$, e_h is a surjection from $\mathcal{L}_n\left(\overline{\mathcal{G}}^{\emptyset}(\delta, \iota, \xi)\right)$ to $\mathcal{U}^{\nu}_n(\delta, \iota, \xi)$.*

Lemma 11 immediately follows from the following lemma.

Lemma 12. *Suppose that $T \in \mathcal{L}\left(\mathcal{G}^{\emptyset}(\delta, \iota, \xi), N_{\langle \Gamma_0;\tau_0\rangle}\right)$ and $S'[S[t']] = e(T)$, where $\mathtt{ct}_t(S) = \langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma;\tau\rangle$. Then, $N_{\langle \Gamma;\tau\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$ and there exists a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma;\tau\rangle$-tree T' such that $e_h(T') = S$.*

Proof. The proof proceeds by induction on the length of the leftmost reduction sequence of T and by case analysis on the rewriting rule used first for the reduction sequence and that on S' and S. Recall that to show $N_{\langle \Gamma;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$ is to show that $N_{\langle \Gamma;\tau\rangle}$ is reachable from a non-terminal of the form $N_{\langle \emptyset;\tau'\rangle}$. We note that, if $S' = [\,]$, then $N_{\langle \Gamma;\tau\rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$ since $\langle \Gamma;\tau\rangle = \langle \Gamma_0;\tau_0\rangle$; below, the required condition that $N_{\langle \Gamma;\tau\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$ can be shown by this condition $S' = [\,]$ or by induction hypothesis directly.

We first consider the case that $S = S' = [\,]$. We have $\langle \Gamma_0;\tau_0\rangle = \langle \Gamma;\tau\rangle = \langle \Gamma';\tau'\rangle$ and because of the rule $N_{\langle \Gamma;\tau\rangle} \longrightarrow h_{\langle \Gamma;\tau\rangle}$, $T' \triangleq h_{\langle \Gamma;\tau\rangle}$ is a $\langle \Gamma;\tau\rangle \Rightarrow \langle \Gamma;\tau\rangle$-tree and $e_h(T') = [\,] = S$. From now, we perform the case analysis on the rewriting rule used first for the reduction sequence of T.

Case of $N_{\langle \{x_i:\tau_0\};\tau_0\rangle} \longrightarrow x_i$: Since $e(T) = x_i$, $S' = S = [\,]$.

Case of $N_{\langle \Gamma_0;\sigma_0\to\tau_0\rangle} \longrightarrow \lambda *^{\sigma_0}(N_{\langle \Gamma_0;\tau_0\rangle})$: Let $T = \lambda *^{\sigma_0}(T_1)$; then $e(T) = \lambda *^{\sigma_0}.e(T_1) = S'[S[t']]$. Now S' is either $\lambda *^{\sigma_0}.S'_1$ or $[\,]$. In the former case, by induction hypothesis for T_1 with $S'_1[S[t']] = e(T_1)$, we directly obtain the required properties: $N_{\langle \Gamma;\tau\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in \mathcal{N}^{\emptyset}(\delta, \iota, \xi)$ and there exists a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma;\tau\rangle$-tree T' such that $e_h(T') = S$.

In the case that $S' = [\,]$, by the typing rules of terms, $\langle \Gamma_0;\sigma_0\to\tau_0\rangle = \langle \Gamma;\tau\rangle$, and S is either $\lambda *^{\sigma_0}.S_1$ or $[\,]$. In the former case, by the typing rules of terms, we have $\mathtt{ct}_t(S_1) = \langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0;\tau_0\rangle$. Then, by induction hypothesis for T_1 with $([\,])[S_1[t']] = e(T_1)$, we obtain the following properties: $N_{\langle \Gamma_0;\tau_0\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in$

$\mathcal{N}^{\emptyset}(\delta,\iota,\xi)$ and there exists a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0;\tau_0\rangle$-tree T_1' such that $e_h(T_1') = S_1$. By combining

$$N_{\langle \Gamma_0;\sigma_0\to\tau_0\rangle} \longrightarrow \lambda *^{\sigma_0}(N_{\langle \Gamma_0;\tau_0\rangle}) \quad \text{and} \quad T_1' : \langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0;\tau_0\rangle\text{-tree}$$

$T' \triangleq \lambda *^{\sigma_0}(T_1')$ is a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0;\sigma_0\to\tau_0\rangle$-tree and

$$e_h(T') = \lambda *^{\sigma_0}.e_h(T_1') = \lambda *^{\sigma_0}.S_1 = S.$$

Case of $N_{\langle \Gamma_0;\sigma_0\to\tau_0\rangle} \longrightarrow \lambda x_i^{\sigma_0}(N_{\langle \Gamma_0\cup\{x_i:\sigma_0\};\tau_0\rangle})$ where $\#(\Gamma_0) < \xi$ and $i = \min\{j \mid x_j \notin \mathrm{Dom}(\Gamma_0)\}$: Let $T = \lambda x_i^{\sigma_0}(T_1)$; then $e(T) = \lambda x_i^{\sigma_0}.e(T_1) = S'[S[t']]$. Now S' is either $\lambda x_i^{\sigma_0}.S_1'$ or $[\,]$. In the former case, by induction hypothesis for T_1 with $S_1'[S[t']] = e(T_1)$, we directly obtain the required properties: $N_{\langle \Gamma;\tau\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)$ and there exists a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma;\tau\rangle$-tree T' such that $e_h(T') = S$.

In the case that $S' = [\,]$, by the typing rules of terms, $\langle \Gamma_0;\sigma_0\to\tau_0\rangle = \langle \Gamma;\tau\rangle$, and S is either $\lambda x_i^{\sigma_0}.S_1$ or $[\,]$. In the former case, by the typing rules of terms, we have $\mathtt{ct}_t(S_1) = \langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0 \cup \{x_i : \sigma_0\};\tau_0\rangle$. Then, by induction hypothesis for T_1 with $([\,])[S_1[t']] = e(T_1)$, we obtain the following properties: $N_{\langle \Gamma_0\cup\{x_i:\sigma_0\};\tau_0\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)$ and there exists a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0 \cup \{x_i : \sigma_0\};\tau_0\rangle$-tree T_1' such that $e_h(T_1') = S_1$. By combining

$$N_{\langle \Gamma_0;\sigma_0\to\tau_0\rangle} \longrightarrow \lambda x_i^{\sigma_0}(N_{\langle \Gamma_0\cup\{x_i:\sigma_0\};\tau_0\rangle}) \quad \text{and} \quad T_1' : \langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0\cup\{x_i\!:\sigma_0\};\tau_0\rangle\text{-tree}$$

$T' \triangleq \lambda x_i^{\sigma_0}(T_1')$ is a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0;\sigma_0\to\tau_0\rangle$-tree and

$$e_h(T') = \lambda x_i^{\sigma_0}.e_h(T_1') = \lambda x_i^{\sigma_0}.S_1 = S.$$

Case of $N_{\langle \Gamma_0;\tau_0\rangle} \longrightarrow @(N_{\langle \Gamma_1;\sigma_0\to\tau_0\rangle} N_{\langle \Gamma_2;\sigma_0\rangle})$ where $\Gamma = \Gamma_1 \cup \Gamma_2$: Let $T = @(T_1, T_2)$; then $e(T) = e(T_1)e(T_2) = S'[S[t']]$. Now S' is either $S_1't_2'$, $t_1'S_2'$ or $[\,]$.

In the first case, we can use the induction hypotheses for T_1 with $S_1'[S[t']] = e(T_1)$. we directly obtain the required properties: $N_{\langle \Gamma;\tau\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)$ and there exists a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma;\tau\rangle$-tree T' such that $e_h(T') = S$. The second case is similar.

In the case that $S' = [\,]$, by the typing rules of terms, $\langle \Gamma_0;\tau_0\rangle = \langle \Gamma;\tau\rangle$, and S is either S_1t_2, t_1S_2 or $[\,]$. In the first case, by the typing rules of terms, we have

$$\Gamma_1' \vdash S_1[t'] : \sigma_0'\to\tau_0 \qquad \Gamma_2' \vdash t_2 : \sigma_0'$$

as well as

$$\mathtt{ct}_t(S_1) = \langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_1';\sigma_0'\to\tau_0\rangle$$

where $\Gamma_1' \cup \Gamma_2' = \Gamma_0$. Since $T_1 \in \mathcal{L}\left(N_{\langle \Gamma_1;\sigma_0\to\tau_0\rangle}\right)$, $\Gamma_1 \vdash e(T_1) : \sigma_0\to\tau_0$. Now $e(T_1) = S_1[t']$, and hence

$$\mathrm{Dom}(\Gamma_1') = \mathbf{FV}(S_1[t']) = \mathbf{FV}(e(T_1)) = \mathrm{Dom}(\Gamma_1).$$

and therefore $\Gamma_1' = \Gamma_1$ since Γ_1', $\Gamma_1 \subseteq \Gamma$. Then, we also have $\sigma_0' = \sigma_0$. Now, we can use induction hypothesis for T_1 with $([\,])[S_1[t']] = e(T_1)$, and obtain the following properties: $N_{\langle \Gamma_1;\sigma_0\to\tau_0\rangle}$, $N_{\langle \Gamma';\tau'\rangle} \in \mathcal{N}^{\emptyset}(\delta,\iota,\xi)$ and there exists a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_1;\sigma_0\to\tau_0\rangle$-tree T_1' such that $e_h(T_1') = S_1$. By combining the following results

$$N_{\langle \Gamma_0;\tau_0\rangle} \longrightarrow @(N_{\langle \Gamma_1;\sigma_0\to\tau_0\rangle} N_{\langle \Gamma_2;\sigma_0\rangle}) \qquad T_1' : \langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_1;\sigma_0\to\tau_0\rangle\text{-tree}$$
$$e(T_2) = t_2 \qquad T_2 \in \mathcal{L}\left(\mathcal{G}^{\emptyset}(\delta,\iota,\xi), N_{\langle \Gamma_2;\sigma_0\rangle}\right)$$

we find that $T' \triangleq @(T_1', T_2)$ is a $\langle \Gamma';\tau'\rangle \Rightarrow \langle \Gamma_0;\tau_0\rangle$-tree. Also we have

$$e_h(T') = \lambda x_i^{\sigma_0}.e_h(T_1') = \lambda x_i^{\sigma_0}.S_1 = S.$$

The case that $S = t_1 S_2$ is similar.

Since there is a surjection from $\mathcal{L}_n\left(\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi)\right)$ to $\mathcal{U}_n^{\nu}(\delta,\iota,\xi)$, we have

$$\#\left(\mathcal{L}_n\left(\overline{\mathcal{G}}^{\emptyset}(\delta,\iota,\xi)\right)\right) \geq \#(\mathcal{U}_n^{\nu}(\delta,\iota,\xi))\,. \tag{2}$$

This ends the proof of Lemma 1.

E Proof of Lemma 2 (decomposition)

In order to prove Lemma 2, we give a complete definition of Φ_m (note that the definition of Φ_m given in Section 4 is a simplified version which ignores the context-typing/size annotations of second-order holes). We first prove Item 2 and 3 of Lemma 2, and then prove Item 1. In the rest of this section, we always assume $n \geq 1$ and $m \geq 2$.

E.1 Complete Definition of Φ_m

The set of *second-order context typings* ζ and typing rules for E are given as follows.

$$\zeta ::= \kappa_1 \cdots \kappa_k \Rightarrow \theta \quad (k \in \mathbb{N})$$

$$\frac{\vdash E_i : \kappa_1^i \cdots \kappa_{k_i}^i \Rightarrow \theta_i \quad (i \in \{1,\ldots,k\}) \qquad \kappa = \theta_1 \cdots \theta_k \Rightarrow \theta}{\vdash [\![\,]\!]_n^{\kappa}[E_1]\cdots[E_k] : \kappa\, \kappa_1^1 \cdots \kappa_{k_1}^1 \cdots \kappa_1^k \cdots \kappa_{k_k}^k \Rightarrow \theta}$$

$$\frac{}{\vdash x : \langle x:\tau;\tau\rangle} \qquad \frac{\vdash E_1 : \kappa_1^1 \cdots \kappa_{k_1}^1 \Rightarrow \langle \Gamma_1;\sigma\to\tau\rangle \qquad \vdash E_2 : \kappa_1^2 \cdots \kappa_{k_1}^2 \Rightarrow \langle \Gamma_2;\sigma\rangle}{\vdash E_1 E_2 : \kappa_1^1 \cdots \kappa_{k_1}^1 \kappa_1^2 \cdots \kappa_{k_1}^2 \Rightarrow \langle \Gamma_1 \cup \Gamma_2;\tau\rangle}$$

$$\frac{\vdash E : \kappa_1 \cdots \kappa_k \Rightarrow \langle \Gamma';\tau\rangle \qquad \Gamma' = \Gamma \text{ or } \Gamma \cup \{\overline{x}:\sigma\} \qquad \overline{x} \notin \mathrm{Dom}(\Gamma)}{\vdash \lambda \overline{x}^{\sigma}.E : \kappa_1 \cdots \kappa_k \Rightarrow \langle \Gamma;\sigma\to\tau\rangle}$$

Below we consider only well-typed second-order contexts. We identify a second-order context typing $\Rightarrow\theta$ with θ.

We define the size of E as follows:

$$|[\![\,]\!]_n^\kappa[E_1]\cdots[E_k]| \triangleq n + |E_1| + \cdots + |E_k| \qquad |x| \triangleq 1$$
$$|\lambda\overline{x}^\tau.E| \triangleq |E| + 1 \qquad |E_1E_2| \triangleq |E_1| + |E_2| + 1$$

Note that $|E[\![C]\!]| = |E[\![P]\!]| = |E|$.

We define $\bar{\kappa}([\![\,]\!]_n^\kappa) \triangleq \kappa$. For a $\langle\Gamma;\tau\rangle$-term t and a 0/1-subcontext occurrence u of t, we write $\bar{\kappa}(u \preceq t : \langle\Gamma;\tau\rangle)$ for the context-typing of u obtained from the derivation of $\Gamma \vdash t : \tau$. For E such that $\mathtt{shn}(E) \geq 1$, we write $E{\upharpoonleft}_n^\kappa$ for the second-order context obtained by replacing $E.1$ with $[\![\,]\!]_n^\kappa$. We omit κ if $\kappa = \bar{\kappa}(E.1)$, and n if $n = |E.1|$. We write $E{\upharpoonleft}(u \preceq t)$ (or $E{\upharpoonleft}u$) for $E{\upharpoonleft}_{|u|}^{\bar{\kappa}(u\preceq t:\langle\Gamma;\tau\rangle)}$, if Γ, τ (and t) are clear from the context.

Now, we define "decomposition" of a term. Below we consider only second-order contexts whose second-order holes are of the form $[\![\,]\!]_n^{\langle\Gamma;\tau\rangle}$ or $[\![\,]\!]_n^{\langle\Gamma';\tau'\rangle\Rightarrow\langle\Gamma;\tau\rangle}$.

For a $\langle\Gamma;\tau\rangle$-term t, we define $\Phi_m(t) = (E, u, P)$ with the following properties

- $u \cdot P : E$ and $E[\![u \cdot P]\!] = t$
- u occurs at the root of t: E is of the form $[\![\,]\!]_n^\kappa[E_1]\cdots[E_k]$ where $k \leq 1$
- $|u| < m$.

We define $\Phi_m(t)$ by induction on the size of t as follows; in this definition we use the above properties (for the operation $-{\upharpoonleft}(- \preceq -)$) and thus the induction is simultaneous.

- If $|t| < m$, then $\Phi_m(t) \triangleq ([\![\,]\!]_{|t|}^{\langle\Gamma;\tau\rangle}, t, \epsilon)$.
- If $|t| \geq m$, then:

$$\Phi_m(\lambda\overline{x}^\tau.t_1) \triangleq \begin{cases} (E_1{\upharpoonleft}(\lambda\overline{x}^\tau.u_1 \preceq \lambda\overline{x}^\tau.t_1), \lambda\overline{x}^\tau.u_1, P_1) \\ \qquad \text{if } \Phi_m(t_1) = (E_1, u_1, P_1) \text{ and } |\lambda\overline{x}^\tau.u_1| < m \\ ([\![\,]\!]_0^{\langle\Gamma;\tau\rangle\Rightarrow\langle\Gamma;\tau\rangle}[E_1{\upharpoonleft}(\lambda\overline{x}^\tau.u_1 \preceq \lambda\overline{x}^\tau.t_1)], [\,], (\lambda\overline{x}^\tau.u_1) \cdot P_1) \\ \qquad \text{if } \Phi_m(t_1) = (E_1, u_1, P_1) \text{ and } |\lambda\overline{x}^\tau.u_1| = m \end{cases}$$

$$\Phi_m(t_1t_2) \triangleq \begin{cases} ([\![\,]\!]_0^{\langle\Gamma;\tau\rangle\Rightarrow\langle\Gamma;\tau\rangle}[(E_1[\![u_1]\!])(E_2[\![u_2]\!])], [\,], P_1 \cdot P_2) \\ \qquad \text{if } \Phi_m(t_i) = (E_i, u_i, P_i) \text{ and } |t_i| \geq m \text{ for each } i \in \{1, 2\} \\ (E_1{\upharpoonleft}(u_1t_2 \preceq t_1t_2), u_1t_2, P_1) \\ \qquad \text{if } \Phi_m(t_1) = (E_1, u_1, P_1),\ |t_1| \geq m,\ |t_2| < m \text{ and } |u_1t_2| < m \\ ([\![\,]\!]_0^{\langle\Gamma;\tau\rangle\Rightarrow\langle\Gamma;\tau\rangle}[E_1{\upharpoonleft}(u_1t_2 \preceq t_1t_2)], [\,], (u_1t_2) \cdot P_1) \\ \qquad \text{if } \Phi_m(t_1) = (E_1, u_1, P_1)\ |t_1| \geq m,\ |t_2| < m \text{ and } |u_1t_2| \geq m \\ (E_2{\upharpoonleft}(t_1u_2 \preceq t_1t_2), t_1u_2, P_2) \\ \qquad \text{if } \Phi_m(t_2) = (E_2, u_2, P_2),\ |t_1| < m \text{ and } |t_1u_2| < m \\ ([\![\,]\!]_0^{\langle\Gamma;\tau\rangle\Rightarrow\langle\Gamma;\tau\rangle}[E_2{\upharpoonleft}(t_1u_2 \preceq t_1t_2)], [\,], (t_1u_2) \cdot P_2) \\ \qquad \text{if } \Phi_m(t_2) = (E_2, u_2, P_2),\ |t_1| < m \text{ and } |t_1u_2| \geq m \end{cases}$$

It is trivial that Φ_m is well-defined, and E, u, and each component of P are well-typed. The decomposition function $\widehat{\Phi}_m$ is then defined by $\widehat{\Phi}_m(t) \triangleq (E[\![u]\!], P)$ where $(E, u, P) = \Phi_m(t)$.

E.2 Proof of Item 2 and 3 of Lemma 2

The following lemma can be shown by straightforward induction of the size of t. Item 4 is Item 2 of Lemma 2.

Lemma 13. *For $\Phi_m(t) = (E, u, P)$,*

1. $u \cdot P : E$ *and* $E[\![u \cdot P]\!] = t$
2. u *occurs at the root of* t: E *is of the form* $[\![\,]\!]^{\kappa}_{n}[E_1] \cdots [E_k]$ *where* $k \leq 1$
3. $|u| < m$
4. $P = u_1 \cdots u_\ell$ *where* $m \leq |u_i| < 2m$ *for each* $1 \leq i \leq \ell$.

Now we prove Item 3 of Lemma 2. Recall that we write $\#(P)$ for the length of a sequence of contexts P. Note that $\mathtt{shn}(E) - 1 = \#(P)$ always holds for any t and $\Phi_m(t) = (E, u, P)$.

Lemma 14. *For any term t and m such that $2 \leq m \leq |t|$, if $\Phi_m(t) = (E, u, P)$, then*

$$\#(P) \geq \frac{|t|}{4m}.$$

Proof. We show below that

$$|u| + 4m\#(P) - 2m \geq |t|$$

by induction on t. Then it follows that

$$4m\#(P) \geq |t| + 2m - |u| > |t| + 2m - m > |t|,$$

which implies $\#(P) > \frac{|t|}{4m}$ as required.

– Case $t = \lambda\overline{x}^{\tau}.t_1$:
  By the assumption $|t| \geq m$, $|t_1| \geq m - 1$. If $|t_1| = m - 1$, then $\Phi_m(t_1) = ([\![\,]\!], t_1, \epsilon)$. Thus, $\Phi_m(t) = ([\![\,]\!][[\![\,]\!]], [\,], \lambda\overline{x}^{\tau}.t_1)$. Therefore, we have

$$|u| + 4m\#(P) - 2m = 0 + 4m - 2m = 2m \geq m = |t|$$

  as required.
  If $|t_1| \geq m$ and $\Phi_m(t_1) = (E_1, u_1, P_1)$, then we have

$$|u_1| + 4m\#(P_1) - 2m \geq |t_1|$$

  by the induction hypothesis. If $|\lambda\overline{x}^{\tau}.u_1| < m$, then

$$|u| + 4m\#(P) - 2m = |u_1| + 1 + 4m\#(P_1)\,2 - m \geq |t_1| + 1 = |t|$$

  as required. If $|\lambda\overline{x}^{\tau}.u_1| = m$, then

$$\begin{aligned} |u| + 4m\#(P) - 2m &= 0 + 4m(\#(P_1) + 1) - 2m = 4m\#(P_1) + 2m \\ &\geq |t_1| + 2m - |u_1| + 2m > |t_1| \end{aligned}$$

  as required.

– Case $t = t_1 t_2$:
Let $\Phi_m(t_i) = (E_i, u_i, P_i)$. We perform further case analysis.

- Case $|t_i| \geq m$ for each $i \in \{1, 2\}$:
  In this case, we have
  $$P = P_1 \cdot P_2 \qquad u = [\,].$$
  By the induction hypothesis, we have
  $$|u_i| + 4m\#(P_i) - 2m \geq |t_i|$$
  for $i \in \{1, 2\}$. Thus,
  $$\begin{aligned}&|u| + 4m\#(P) - 2m = 0 + 4m(\#(P_1) + \#(P_2)) - 2m\\ &\geq (|t_1| + 2m - |u_1|) + (|t_2| + 2m - |u_2|) - 2m\\ &= |t_1| + |t_2| + (m - |u_1|) + (m - |u_2|)\\ &\geq |t_1| + |t_2| + 2 > |t|\end{aligned}$$
  as required.
- Case $|t_1| \geq m$ and $|u_1 t_2| < m$:
  In this case, we have
  $$P = P_1 \qquad u = u_1 t_2.$$
  By the induction hypothesis, we have
  $$|u_1| + 4m\#(P_1) - 2m \geq |t_1|.$$
  Thus, we have:
  $$\begin{aligned}&|u| + 4m\#(P) - 2m = |u_1| + |t_2| + 1 + 4m\#(P_1) - 2m\\ &\geq |u_1| + |t_2| + 1 + (|t_1| + 2m - |u_1|) - 2m = |t_1| + |t_2| + 1 = |t|\end{aligned}$$
  as required.
- Case $|t_1| \geq m$, $|t_2| < m$ and $|u_1 t_2| \geq m$:
  In this case, we have
  $$P = u_1 t_2 \cdot P_1 \qquad u = [\,].$$
  By the induction hypothesis, we have
  $$|u_1| + 4m\#(P_1) - 2m \geq |t_1|.$$
  Thus, we have
  $$\begin{aligned}&|u| + 4m\#(P) - 2m = 0 + 4m(\#(P_1) + 1) - 2m\\ &\geq |t_1| + 2m - |u_1| + 2m > |t_1| + 4m - m\\ &> |t_1| + |t_2| + 1 = |t|\end{aligned}$$
  as required.

- Case $|t_1| < m$, and $|t_1 u_2| < m$:
  It must be the case $|t_2| \geq m$, because otherwise we have $u_2 = t_2$, which would imply $|t| = |t_1 t_2| = |t_1 u_2| < m$, contradicting the assumption $|t| \geq m$. Thus, the required result can be obtained in a manner similar to the case for $|t_1| \geq m$ and $|u_1 t_2| < m$.
- Case $|t_1| < m$ and $|t_1 u_2| \geq m$: In this case, we have:

$$P = t_1 u_2 \cdot P_2 \qquad u = [\,].$$

  If $|t_2| \geq m$, then the required result can be obtained in a manner similar to the case for $|t_1| \geq m$, $|t_2| < m$ and $|u_1 t_2| \geq m$. If $|t_2| < m$, then $P_2 = \epsilon$ and $u_2 = t_2$. Therefore, we have:

$$|u| + 4m\#(P) - 2m = 0 + 4m - 2m = 2m \geq |t_1| + |t_2| + 1 = |t|$$

  as required. $\square$

E.3 Proof of Item 1 of Lemma 2

For $n > 0$ and $m > 2$, we define an auxiliary set $\widehat{\mathcal{B}}^m_n(\delta, \iota, \xi)$ of second-order contexts as:

$$\widehat{\mathcal{B}}^m_n(\delta, \iota, \xi) \triangleq \{(E, u) \mid (E, u, P) = \Phi_m(\nu([t]_\alpha)) \text{ for some } [t]_\alpha \in \Lambda^\alpha_n(\delta, \iota, \xi)\}.$$

Recall that we have defined $\mathcal{B}^m_n(\delta, \iota, \xi)$ as:

$$\mathcal{B}^m_n(\delta, \iota, \xi) = \{E \mid (E, P) = \widehat{\Phi}_m(\nu([t]_\alpha)) \text{ for some } [t]_\alpha \in \Lambda^\alpha_n(\delta, \iota, \xi)\},$$

and obviously, we have $\mathcal{B}^m_n(\delta, \iota, \xi) = \{E[\![u]\!] \mid (E, u) \in \widehat{\mathcal{B}}^m_n(\delta, \iota, \xi)\}$.

We call a sequence P of contexts *good context sequence for* (E, u, m) if $E : u{\cdot}P$ and $\Phi_m(E[\![u{\cdot}P]\!]) = (E, u, P)$. We write $\mathcal{P}^m_{E,u}(\delta, \iota, \xi)$ for the set of all good context sequences for (E, u, m). We define a subset $\widehat{\mathcal{P}}^m_{E,u}(\delta, \iota, \xi)$ of $\mathcal{P}^m_{E,u}(\delta, \iota, \xi)$ as:

$$\widehat{\mathcal{P}}^m_{E,u}(\delta, \iota, \xi) \triangleq \{P \in \mathcal{P}^m_{E,u}(\delta, \iota, \xi) \mid E[\![u{\cdot}P]\!] = \nu([E[\![u{\cdot}P]\!]]_\alpha),\ [E[\![u{\cdot}P]\!]]_\alpha \in \Lambda^\alpha(\delta, \iota, \xi)\}.$$

In the rest of this subsection, we will show that the following two properties holds:

- $\widehat{\Phi}_m$ induces a bijection between: $\Lambda^\alpha_n(\delta, \iota, \xi)$ and $\coprod_{(E,u) \in \widehat{\mathcal{B}}^m_n(\delta,\iota,\xi)} \widehat{\mathcal{P}}^m_{E,u}(\delta, \iota, \xi)$.
- $\widehat{\mathcal{P}}^m_{E,u}(\delta, \iota, \xi) = \prod_{2 \leq i \leq \mathtt{shn}(E)} \widehat{U}^m_{E.i}(\delta, \iota, \xi)$ for $(E, u) \in \widehat{\mathcal{B}}^m_n(\delta, \iota, \xi)$.

Item 1 of Lemma 2 is a direct consequence of these two properties.

First we show the coproduct part.

Lemma 15. *For any $\delta, \iota, \xi \geq 0$, $n \geq 1$ and $m \geq 2$, there is a bijection*

$$\begin{aligned} \Lambda^\alpha_n(\delta, \iota, \xi) \quad &\cong \coprod_{(E,u) \in \widehat{\mathcal{B}}^m_n(\delta,\iota,\xi)} \widehat{\mathcal{P}}^m_{E,u}(\delta, \iota, \xi) \\ &\Big(= \biguplus_{(E,u) \in \widehat{\mathcal{B}}^m_n(\delta,\iota,\xi)} \{(E, u)\} \times \widehat{\mathcal{P}}^m_{E,u}(\delta, \iota, \xi) \Big). \end{aligned}$$

Proof. We define a function

$$f : \coprod_{(E,u)\in\widehat{\mathcal{B}}^m_n(\delta,\iota,\xi)} \widehat{\mathcal{P}}^m_{E,u}(\delta,\iota,\xi) \longrightarrow \Lambda^\alpha_n(\delta,\iota,\xi)$$

by $f((E,u),P) \triangleq [E[\![u \cdot P]\!]]_\alpha$, and a function

$$g : \Lambda^\alpha_n(\delta,\iota,\xi) \longrightarrow \coprod_{(E,u)\in\widehat{\mathcal{B}}^m_n(\delta,\iota,\xi)} \widehat{\mathcal{P}}^m_{E,u}(\delta,\iota,\xi)$$

by $g([t]_\alpha) \triangleq ((E,u),P)$ where $(E,u,P) = \Phi_m(\nu([t]_\alpha))$. Then we have

$$f(g([t]_\alpha)) = f((E,u),P) = [E[\![u \cdot P]\!]]_\alpha = [\nu([t]_\alpha)]_\alpha = [t]_\alpha$$

and

$$g(f((E,u),P)) = g([E[\![u \cdot P]\!]]_\alpha) = ((E,u),P)$$

where the last equation holds since

$$\Phi_m(\nu([E[\![u \cdot P]\!]]_\alpha)) = \Phi_m(E[\![u \cdot P]\!]) = (E,u,P)$$

due to $P \in \widehat{\mathcal{P}}^m_{E,u}(\delta,\iota,\xi)$. □

Next we show the product part.

Lemma 16. *For $\delta,\iota,\xi \geq 0$, $n \geq 1$, $m \geq 2$ and $(E,u) \in \mathcal{B}^m_n(\delta,\iota,\xi)$,*

$$\widehat{\mathcal{P}}^m_{E,u}(\delta,\iota,\xi) = \widehat{U}^m_{E\bullet 2}(\delta,\iota,\xi) \times \cdots \times \widehat{U}^m_{E\bullet \mathtt{shn}(E)}(\delta,\iota,\xi).$$

To prove Lemma 16, we take two steps, Lemmas 20 and 23, with some auxiliary lemmas.

For $m \geq 2$, E and $i \in \{1,\ldots,\mathtt{shn}(E)\}$, we define

$$U^m_{E\bullet i} \triangleq \{u \mid u : E\bullet i \text{ and } u \text{ is good for } m\}.$$

Recall that, for $\delta,\iota,\xi \geq 0$, we have defined $\widehat{U}^m_{E\bullet i}(\delta,\iota,\xi)$ as:

$$\begin{aligned}\widehat{U}^m_{E\bullet i}(\delta,\iota,\xi) &= \{u \in \mathcal{U}^\nu(\delta,\iota,\xi) \mid u : E\bullet i \text{ and } u \text{ is good for } m\}\\ &= U^m_{E\bullet i} \cap \mathcal{U}^\nu(\delta,\iota,\xi).\end{aligned}$$

For sets $U_1,\ldots,U_k$ of 0/1-contexts, we use the following notation:

$$U_1 \times \cdots \times U_k \triangleq \{u_1 \cdots u_k \mid u_i \in U_i \text{ for each } i\}.$$

It is easy to show, by induction on C, that if $C \in \mathcal{C}(\theta_1 \cdots \theta_k \Rightarrow \theta)$, $1 \leq i \leq k$ and $C' \in \mathcal{C}(\theta'_1 \cdots \theta'_{k'} \Rightarrow \theta_i)$, then $C[C']_i \in \mathcal{C}(\theta_1 \cdots \theta_{i-1}\theta'_1 \cdots \theta'_{k'}\theta_{i+1} \cdots \theta_k \Rightarrow \theta)$. For κ-context C with its derivation Δ and a subcontext occurrence C' of C, we write $\bar{\kappa}(C' \preceq C : \kappa)$ for the context typing of C' that is determined by the corresponding sub-derivation of Δ.

Thus we have the following two trivial facts; we may use these two lemmas without referring explicitly.

Lemma 17. *For $\kappa = \theta_1 \cdots \theta_n \Rightarrow \theta$, $\kappa_i = \theta_1^i \cdots \theta_{n_i}^i \Rightarrow \theta_i$ $(i = 1, \ldots, n)$, $C \in \mathcal{C}^\nu(\kappa, \delta, \iota, \xi)$ and $C_i \in \mathcal{C}^\nu(\kappa_i, \delta, \iota, \xi)$, let $\kappa' = \theta_1^1 \cdots \theta_{n_1}^1 \cdots \theta_1^n \cdots \theta_{n_n}^n \Rightarrow \theta$; then*

$$C[C_1] \cdots [C_n] \in \mathcal{C}^\nu(\kappa', \delta, \iota, \xi).$$

Lemma 18. *For $C \in \mathcal{C}^\nu(\kappa, \delta, \iota, \xi)$ and $C' \preceq C$, let $\kappa' = \bar{\kappa}(C' \preceq C : \kappa)$. Then $C' \in \mathcal{C}^\nu(\kappa', \delta, \iota, \xi)$.*

Lemma 19. *For any $\delta, \iota, \xi \geq 0$, E, typing θ, and two context sequences P and P' of same length k (i.e., , $\#(P) = \#(P') = k$), if $P' : E$, $P : E$, $E[\![P']\!] \in \mathcal{C}^\nu(\theta, \delta, \iota, \xi)$ and $C_i \in \mathcal{C}^\nu(\bar{\kappa}(E.i), \delta, \iota, \xi)$ for all $i \leq k$, then $E[\![P]\!] \in \mathcal{C}^\nu(\theta, \delta, \iota, \xi)$.*

Proof. Let $P = C_1 \cdots C_k$ and $P' = C'_1 \cdots C'_k$. The proof proceeds by induction on E. Below we use Lemmas 17 and 18.

- Case $E = [\![\,]\!]_{m_1}^{\theta_1 \cdots \theta_\ell \Rightarrow \theta}[E_1] \cdots [E_\ell]$: Let $k_1 = 2$, $k_{i+1} = k_i + \mathtt{shn}(E_i)$ for $1 \leq i < k$, and $k'_i = k_i + \mathtt{shn}(E_i) - 1$ for $1 \leq i \leq k$; thus $(C'_{k_i} \cdots C'_{k'_i})$, $(C_{k_i} \cdots C_{k'_i}) : E_i$ for each $i \leq k$. Since
$$E[\![P']\!] = C'_1[E_1[\![C'_{k_1} \cdots C'_{k'_1}]\!]] \cdots [E_\ell[\![C'_{k_\ell} \cdots C'_{k'_\ell}]\!]] \in \mathcal{C}^\nu(\theta, \delta, \iota, \xi),$$
$E_i[\![C'_{k_i} \cdots C'_{k'_i}]\!] \in \mathcal{C}^\nu(\theta_i, \delta, \iota, \xi)$ for each $i \leq k$. By induction hypothesis, $E_i[\![C_{k_i} \cdots C_{k'_i}]\!] \in \mathcal{C}^\nu(\theta_i, \delta, \iota, \xi)$ for each $i \leq k$, and hence
$$E[\![P]\!] = C_1[E_1[\![C_{k_1} \cdots C_{k'_1}]\!]] \cdots [E_\ell[\![C_{k_\ell} \cdots C_{k'_\ell}]\!]] \in \mathcal{C}^\nu(\theta, \delta, \iota, \xi).$$
- Case $E = x$: Clear.
- Case $E = \lambda \overline{x}^\sigma.E'$: We have $E[\![P']\!] = \lambda \overline{x}^\sigma.E'[\![P']\!] \in \mathcal{C}^\nu(\theta, \delta, \iota, \xi)$; let $\theta' = \bar{\kappa}(E'[\![P']\!] \preceq \lambda \overline{x}^\sigma.E'[\![P']\!] : \theta)$. Then $E'[\![P']\!] \in \mathcal{C}^\nu(\theta', \delta, \iota, \xi)$ and $\lambda \overline{x}^\sigma.[] \in \mathcal{C}^\nu(\theta' \Rightarrow \theta, \delta, \iota, \xi)$. By induction hypothesis, $E'[\![P]\!] \in \mathcal{C}^\nu(\theta', \delta, \iota, \xi)$, and hence $E[\![P]\!] = \lambda \overline{x}^\sigma.E'[\![P]\!] \in \mathcal{C}^\nu(\langle \Gamma; \sigma \to \tau \rangle, \delta, \iota, \xi)$.
- Case $E = E_1 E_2$: Let $k_1 = 1$, $k'_1 = \mathtt{shn}(E_1)$, $k_2 = \mathtt{shn}(E_1) + 1$, and $k'_2 = k$. Then $E[\![P']\!] = E_1[\![C'_{k_1} \cdots C'_{k'_1}]\!] E_2[\![C'_{k_2} \cdots C'_{k'_2}]\!] \in \mathcal{C}^\nu(\theta, \delta, \iota, \xi)$. Let $\theta_i = \bar{\kappa}(E_i[\![C'_{k_i} \cdots C'_{k'_i}]\!] \preceq E[\![P']\!] : \theta)$. We have $E_i[\![C'_{k_i} \cdots C'_{k'_i}]\!] \in \mathcal{C}^\nu(\theta_i, \delta, \iota, \xi)$ and $[][] \in \mathcal{C}^\nu(\theta_1 \theta_2 \Rightarrow \theta, \delta, \iota, \xi)$. By induction hypothesis, $E_i[\![C_{k_i} \cdots C_{k'_i}]\!] \in \mathcal{C}^\nu(\theta_i, \delta, \iota, \xi)$. Hence, $E[\![P]\!] = E_1[\![C_{k_1} \cdots C_{k'_1}]\!] E_2[\![C_{k_2} \cdots C_{k'_2}]\!] \in \mathcal{C}^\nu(\theta, \delta, \iota, \xi)$.

□

Lemma 20. *For $\delta, \iota, \xi \geq 0$, $n \geq 1$, $m \geq 2$ and $(E, u) \in \widehat{\mathcal{B}}_n^m(\delta, \iota, \xi)$, we have*

$$\begin{aligned} \widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi) &= \{P \in \mathcal{P}_{E,u}^m(\delta, \iota, \xi) \mid P.i \in \mathcal{U}^\nu(\delta, \iota, \xi) \text{ for each } i \leq \mathtt{shn}(E) - 1\} \\ &= \mathcal{P}_{E,u}^m(\delta, \iota, \xi) \cap (\mathcal{U}^\nu(\delta, \iota, \xi))^{\mathtt{shn}(E)-1}. \end{aligned}$$

Proof. Let $P \in \widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi)$. By the definition of $\widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi)$, $P \in \mathcal{P}_{E,u}^m(\delta, \iota, \xi)$, $E[\![u \cdot P]\!] = \nu([E[\![u \cdot P]\!]]_\alpha)$ and $[E[\![u \cdot P]\!]]_\alpha \in \Lambda^\alpha(\delta, \iota, \xi)$. Hence $E[\![u \cdot P]\!] \in \mathcal{U}^\nu(\delta, \iota, \xi)$, and by Lemma 18 $P \in (\mathcal{U}^\nu(\delta, \iota, \xi))^{\mathtt{shn}(E)-1}$. Thus $P \in \mathcal{P}_{E,u}^m(\delta, \iota, \xi) \cap (\mathcal{U}^\nu(\delta, \iota, \xi))^{\mathtt{shn}(E)-1}$.

Let $P \in \mathcal{P}^m_{E,u}(\delta,\iota,\xi) \cap (\mathcal{U}^\nu(\delta,\iota,\xi))^{\mathtt{shn}(E)-1}$. Since $(E,u) \in \widehat{\mathcal{B}}^m_n(\delta,\iota,\xi)$, there exists $P' \in \widehat{\mathcal{P}}^m_{E,u}(\delta,\iota,\xi)$. Hence $P' \in \mathcal{P}^m_{E,u}(\delta,\iota,\xi)$, $E[\![u \cdot P']\!] = \nu([E[\![u \cdot P']\!]]_\alpha)$ and $[E[\![u \cdot P']\!]]_\alpha \in \Lambda^\alpha(\delta,\iota,\xi)$. Therefore $E[\![u \cdot P']\!] \in \mathcal{C}^\nu(\langle\emptyset;\tau\rangle,\delta,\iota,\xi)$ for some $\tau \in \mathrm{Types}(\delta,\iota)$. By Lemma 19, $E[\![u \cdot P]\!] \in \mathcal{C}^\nu(\langle\emptyset;\tau\rangle,\delta,\iota,\xi)$, and hence $E[\![u \cdot P]\!] = \nu([E[\![u \cdot P]\!]]_\alpha)$ and $[E[\![u \cdot P]\!]]_\alpha \in \Lambda^\alpha(\delta,\iota,\xi)$. Thus, $P \in \widehat{\mathcal{P}}^m_{E,u}(\delta,\iota,\xi)$. □

Lemma 21. *For $m \geq 2$ and $\Phi_m(t) = (E,u,P)$, the following are equivalent:*

- $|t| \geq m$
- $\mathtt{shn}(E) \geq 2$
- *u is a 1-context.*

Proof. Straightforward induction on $|t|$. □

In the next two lemmas, we use the fact: $P \in \mathcal{P}^m_{E,u}(\delta,\iota,\xi)$ if and only if $\Phi_m(t) = (E,u,P)$ for some t. The following lemma states that, for $\Phi_m(u[t]) = (E,u,P)$, P is determined only by t (u does not matter), i.e., decomposition is performed in a bottom-up manner.

Lemma 22. *For $m \geq 2$ and a θ'-term t such that $|t| \geq m$, let $\Phi_m(t) = (E,u,P)$ and $\bar{\kappa}(u \preceq t : \theta') = \theta \Rightarrow \theta'$. Then*

$$\mathcal{P}^m_{E,u}(\delta,\iota,\xi) = \mathcal{P}^m_{E\restriction_0^{\theta\Rightarrow\theta},[\,]}(\delta,\iota,\xi)\ .$$

Proof. The proof proceeds by induction on the size of u. Note that, by the assumption $|t| \geq m$ and Lemma 21, u is a 1-context. The case where $u = [\,]$ is trivial.

- Case $u = \lambda\bar{x}^\sigma.u_1$: By the definition of Φ_m we have $\mathcal{P}^m_{E,u}(\delta,\iota,\xi) = \mathcal{P}^m_{E\restriction u_1,u_1}(\delta,\iota,\xi)$. Then by the induction hypothesis, $\mathcal{P}^m_{E\restriction u_1,u_1}(\delta,\iota,\xi) = \mathcal{P}^m_{E\restriction_0^{\theta\Rightarrow\theta},[\,]}(\delta,\iota,\xi)$.
- Case $u = u_1 t_2$: Since u is a 1-context, it must be the case that $\mathtt{shn}(E) > 1$. Therefore, for every $P_1 \in \mathcal{P}^m_{E\restriction u_1,u_1}(\delta,\iota,\xi)$, $|(E\restriction u_1)[u_1 \cdot P_1]| \geq m$. Thus, we have $P_1 \in \mathcal{P}^m_{E,u_1t_2}(\delta,\iota,\xi)$ by the definition of $\mathcal{P}^m_{E\restriction u_1,u_1}(\delta,\iota,\xi)$. Conversely, if $P_1 \in \mathcal{P}^m_{E,u_1t_2}(\delta,\iota,\xi)$, then (E,u_1t_2,P_1) must have been computed by the second case of $\Phi_m(t_1t_2)$; therefore, we have $P_1 \in \mathcal{P}^m_{E\restriction u_1,u_1}(\delta,\iota,\xi)$. Thus, we have $\mathcal{P}^m_{E,u}(\delta,\iota,\xi) = \mathcal{P}^m_{E\restriction u_1,u_1}(\delta,\iota,\xi) = \mathcal{P}^m_{E\restriction_0^{\theta\Rightarrow\theta},[\,]}(\delta,\iota,\xi)$ as required.
- Case $u = u_2 t_1$: We omit the proof as it is similar to the proof of the above case.

The following lemma shows that the components of a good context sequence are independent of each other.

Lemma 23. *For $n \geq 1$, $m \geq 2$ and (E,u,m),*

$$\mathcal{P}^m_{E,u}(\delta,\iota,\xi) = U^m_{E.2} \times \cdots \times U^m_{E.\mathtt{shn}(E)}.$$

Proof. For readability, we omit the parameters (δ,ι,ξ), and write $\mathcal{P}^m_{E,u}$ for $\mathcal{P}^m_{E,u}(\delta,\iota,\xi)$. Let $k = \mathtt{shn}(E)$. The proof proceeds by induction on $k \geq 1$. When $k = 1$, $U^m_{E.2} \times \cdots \times U^m_{E.k} = \{\epsilon\}$.

When $k \geq 2$, let us pick $P \in \mathcal{P}^m_{E,u}(\delta,\iota,\xi)$; then $(E,u,P) = \Phi_m(E[\![u \cdot P]\!])$. Now we perform a case analysis on how E has been computed. Since $k \geq 2$, u is a 1-context, and hence by Lemma 22, we can assume $u = [\,]$ in this case analysis.

– Case $E = [\![\,]\!][E']$: In this case, E is computed by the second case of $\Phi_m(\lambda \overline{x}'^{\tau'}.t')$, or the third or fifth case of $\Phi_m(t'_1 t'_2)$. By the fact we mentioned above,

$$\begin{aligned}
\mathcal{P}^m_{E,[\,]} &= \{P \mid \Phi_m(t) = (E, [\,], P) \text{ for some } t\} \\
&= (\because \text{ by the definition of } \Phi_m(t)) \\
&\quad \{\lambda \overline{x}^{\tau}.u_1 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E',u_1}, |\lambda \overline{x}^{\tau}.u_1| = m, \lambda \overline{x}^{\tau}.u_1 : E\text{.}2\} \\
&\quad \cup \{u_1 t_2 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E',u_1}, |E'[u_1 \cdot P_1]| \geq m, |t_2| < m, |u_1 t_2| \geq m, u_1 t_2 : E\text{.}2\} \\
&\quad \cup \{t_1 u_2 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E',u_2}, |t_1| < m, |t_1 u_2| \geq m, t_1 u_2 : E\text{.}2\}.
\end{aligned}$$

By Lemma 21, $|E'[u_1 \cdot P_1]| \geq m$ if and only if $\mathtt{shn}(E') > 1$, we further have:

$$\begin{aligned}
\mathcal{P}^m_{E,[\,]} =& \{\lambda \overline{x}^{\tau}.u_1 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E',u_1}, |\lambda \overline{x}^{\tau}.u_1| = m, \lambda \overline{x}^{\tau}.u_1 : E\text{.}2\} \\
&\cup \{u_1 t_2 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E',u_1}, \mathtt{shn}(E') > 1, |t_2| < m, |u_1 t_2| \geq m, u_1 t_2 : E\text{.}2\} \\
&\cup \{t_1 u_2 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E',u_2}, |t_1| < m, |t_1 u_2| \geq m, t_1 u_2 : E\text{.}2\}. \qquad (3)
\end{aligned}$$

We perform further case analysis:

- Case $\mathtt{shn}(E') = 1$: In this case, in the right hand side of Equation (3), $\mathtt{shn}(E') = 1$, $P_1 = \epsilon$ and u_1 and u_2 are 0-contexts by Lemma 21. Then we can deduce from Equation (3) that:

$$\begin{aligned}
\mathcal{P}^m_{E,[\,]} &= \{\lambda \overline{x}^{\tau}.u_1 \mid |\lambda \overline{x}^{\tau}.u_1| = m, \lambda \overline{x}^{\tau}.u_1 : E\text{.}2\} \\
&\quad \cup \{t_1 u_2 \mid |t_1| < m, |t_1 u_2| \geq m, t_1 u_2 : E\text{.}2\} \\
&= \{\lambda \overline{x}^{\tau}.u_1 \mid \lambda \overline{x}^{\tau}.u_1 \text{ is good}, \lambda \overline{x}^{\tau}.u_1 : E\text{.}2\} \\
&\quad \cup \{t_1 u_2 \mid t_1 u_2 \text{ is good}, t_1 u_2 : E\text{.}2\} \\
&= U^m_{E\text{.}2}.
\end{aligned}$$

- Case $\mathtt{shn}(E') > 1$: Let $\bar{\kappa}(E'\text{.}2) = \kappa \Rightarrow \langle \Gamma; \tau \rangle$. In this case, in the right hand side of Equation (3), u_1 and u_2 are 1-contexts by Lemma 21. Further, $\lambda \overline{x}^{\tau}.u_1 : E\text{.}2, t_1 u_2 : E\text{.}2$ and $t_1 u_2 : E\text{.}2$ imply that u_1 and u_2 have the same hole typing κ, i.e., $\emptyset \vdash u_1 : \kappa \Rightarrow \langle \Gamma_1; \tau_1 \rangle$ and $\emptyset \vdash u_2 : \kappa \Rightarrow \langle \Gamma_2; \tau_2 \rangle$ for some $\langle \Gamma_1; \tau_1 \rangle, \langle \Gamma_2; \tau_2 \rangle$. Therefore $\mathcal{P}^m_{E',u_i} = \mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]}$ by Lemma 22. Then we can deduce from Equation (3) that:

$$\begin{aligned}
\mathcal{P}^m_{E,[\,]} &= \{\lambda \overline{x}^{\tau}.u_1 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]}, |\lambda \overline{x}^{\tau}.u_1| = m, \lambda \overline{x}^{\tau}.u_1 : E\text{.}2\} \\
&\quad \cup \{u_1 t_2 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]}, \mathtt{shn}(E') > 1, |t_2| < m, |u_1 t_2| \geq m, u_1 t_2 : E\text{.}2\} \\
&\quad \cup \{t_1 u_2 \cdot P_1 \mid P_1 \in \mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]}, |t_1| < m, |t_1 u_2| \geq m, |t_1 u_2| : E\text{.}2\} \\
&= (\{\lambda \overline{x}^{\tau}.u_1 \mid |\lambda \overline{x}^{\tau}.u_1| = m, \lambda \overline{x}^{\tau}.u_1 : E\text{.}2\} \\
&\quad \cup \{u_1 t_2 \mid \mathtt{shn}(E') > 1, |t_2| < m, |u_1 t_2| \geq m, u_1 t_2 : E\text{.}2\} \\
&\quad \cup \{t_1 u_2 \mid |t_1| < m, |t_1 u_2| \geq m, t_1 u_2 : E\text{.}2\}) \times \mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]} \\
&= (\{\lambda \overline{x}^{\tau}.u_1 \mid \lambda \overline{x}^{\tau}.u_1 \text{ is good}, u_1 : E\text{.}2\} \\
&\quad \cup \{t_1 u_2 \mid t_1 u_2 \text{ is good}, u_1 : E\text{.}2\}) \times \mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]} \\
&= U^m_{E\text{.}2} \times \mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]}
\end{aligned}$$

  By the induction hypothesis, $\mathcal{P}^m_{E' \upharpoonright_0^{\kappa \Rightarrow \kappa}, [\,]} = U^m_{E\text{.}3} \times \cdots \times U^m_{E\text{.}k}$, thus we obtain $\mathcal{P}^m_{E,[\,]} = U^m_{E\text{.}2} \times \cdots \times U^m_{E\text{.}k}$.

– Case $E = [\![\,]\!][(E_1[\![u_1]\!])(E_2[\![u_2]\!])]$: In this case, E is computed by the first case of $\Phi_m(t_1't_2')$. By the definition of Φ_m, we have:

$$\begin{aligned}\mathcal{P}^m_{E,[\,]} &= \{P_1 \cdot P_2 \mid P_1 \in \mathcal{P}^m_{E_1,u_1} \text{ and } P_2 \in \mathcal{P}^m_{E_2,u_2}\} \\ &= \mathcal{P}^m_{E_1,u_1} \times \mathcal{P}^m_{E_2,u_2}\end{aligned}$$

Let $k_1 = \mathtt{shn}(E_1)$, By the induction hypothesis, $\mathcal{P}^m_{E_1,u_1} = U^m_{E.\mathbf{2}} \times \cdots \times U^m_{E.k_1}$ and $\mathcal{P}^m_{E_2,u_2} = U^m_{E.k_1+1} \times \cdots \times U^m_{E.k}$, thus we obtain $\mathcal{P}^m_{E,[\,]} = U^m_{E.\mathbf{2}} \times \cdots \times U^m_{E.k}$. □

Combining Lemmas 20 and 23, we obtain Lemma 16 as a corollary.

F Proof of Lemma 3

Proof. First, by the definition of $\overline{2}_k$, we have:

$$|\overline{2}_k| = 7, \quad \vdash \overline{2}_k : \tau(k+2), \quad \mathtt{ord}(\overline{2}_k) = k, \quad \mathtt{iar}(\overline{2}_k) = k.$$

Item 1 is clear.

On Item 2, since we have: $|t \uparrow^m s| = m|t| + |s| + m$, $|Dup| = 8$ and $|Id| = 2$,

$$\begin{aligned}|Expl^k_m| &= 1 + (7m + 7 + m) + 1 + 7 + 1 + \cdots + 7 + 1 + 8 + 1 + 2 + 1 + 0 \\ &= 1 + (7m + 7 + m) + 1 + (k-3)(7+1) + 8 + 1 + 2 + 1 = 8m + 8k - 3.\end{aligned}$$

On Item 3, $\#(\mathbf{V}(Expl^k_m)) = 2$ is clear. To calculate $\mathtt{ord}(-)$ and $\mathtt{iar}(-)$ inductively, we show the following by induction:

$$\mathtt{ord}\left(\frac{}{\Gamma \vdash x : \tau}\right) = \mathtt{ord}(\tau) \qquad \mathtt{ord}\left(\frac{\Delta}{\Gamma \vdash \lambda \overline{x}^{\tau'}.t : \tau}\right) = \max\{\mathtt{ord}(\tau), \mathtt{ord}(\Delta)\}$$

$$\mathtt{ord}\left(\frac{\Delta_1 \quad \Delta_2}{\Gamma \vdash ts : \tau}\right) = \max\{\mathtt{ord}(\Delta_1), \mathtt{ord}(\Delta_2)\}$$

$$\mathtt{iar}\left(\frac{}{\Gamma \vdash x : \tau}\right) = \mathtt{iar}(\tau) \qquad \mathtt{iar}\left(\frac{\Delta}{\Gamma \vdash \lambda \overline{x}^{\tau'}.t : \tau}\right) = \max\{\mathtt{iar}(\tau), \mathtt{iar}(\Delta)\}$$

$$\mathtt{iar}\left(\frac{\Delta_1 \quad \Delta_2}{\Gamma \vdash ts : \tau}\right) = \max\{\mathtt{iar}(\Delta_1), \mathtt{iar}(\Delta_2)\}$$

where note that we have excluded $\mathtt{ord}(\tau)$ and $\mathtt{iar}(\tau)$ in the right hand sides of the two application cases. Also, we have

$$\mathtt{ord}(Dup) = 1 \qquad \mathtt{ord}(Id) = 1 \qquad \mathtt{iar}(Dup) = 2 \qquad \mathtt{iar}(Id) = 1.$$

Hence,

$$\begin{aligned}\mathtt{ord}(Expl^k_m[x]) &= \max\{\mathtt{ord}(\mathtt{o}{\to}\mathtt{o}), k, \ldots, 2, \mathtt{ord}(Dup), \mathtt{ord}(Id), 1\} = k \\ \mathtt{iar}(Expl^k_m[x]) &= \max\{\mathtt{iar}(\mathtt{o}{\to}\mathtt{o}), k, \ldots, 2, \mathtt{iar}(Dup), \mathtt{iar}(Id), 1\} = k.\end{aligned}$$

Item 4 follows from Items 1 and 3.

The proof for Item 5 is the same as that in [1]. □

G Proof of Lemma 4

Proof. We shall construct u' by embedding $Expl^k_{m'}$ in an appropriate context. We use Corollary 1 to adjust the goodness, typing and size of u'.

For given $\delta, \iota, \xi \geq 2$, let $d_0 \geq 1$ be the constant in Corollary 1. We choose b and c so that the following condition holds for any $m' \geq b$:

$$(cm' - 1)/2 \geq |Expl^k_{m'}| + 2d_0. \tag{4}$$

For example, we can set $b = 8k - 2 + 2d_0$ and $c = 18$; note that $|Expl^k_{m'}| = 8m' + 8k - 3$ by Lemma 3.

Suppose that n, m', (E, u) and i in the statement are given. By the definition, $\widehat{U}^{cm'}_{E\bullet i}(\delta, \iota, \xi)$ is non-empty. Pick an arbitrary element $u'' \in \widehat{U}^{cm'}_{E\bullet i}(\delta, \iota, \xi)$. We construct u' by finding a subcontext u_0 of u'' such that $|Expl^k_{m'}| + 2d_0 \leq |u_0| < cm'$, and replacing it with a 0/1-context of the form $S^\circ[Expl^k_{m'}[u^\circ]]$ such that (i) $|S^\circ| = d_0$, (ii) $|u^\circ| = |u_0| - |Expl^k_{m'}| - d_0 \geq d_0$, (iii) S°, $u^\circ \in \mathcal{U}^\nu(\delta, \iota, \xi)$, and (iv) the size and typing of $S^\circ[Expl^k_{m'}[u^\circ]]$ are the same as those of u_0. The existence of such S° and u° are guaranteed by Corollary 1. Since u'' is good, so is u' (note that $|u_0| < cm'$).

It remains to show how to find the subcontext u_0. We perform a case analysis:

- Case where u'' is of the form $\lambda\overline{x}^\sigma.u'_0$. In this case, since $\lambda\overline{x}^\sigma.u'_0$ is good for cm', we have $|E\bullet i| = |\lambda\overline{x}^\sigma.u'_0| = cm'$. Thus, we can let $u_0 = u'_0$. Note that the required condition $|Expl^k_{m'}| + 2d_0 \leq |u_0|$ follows from the inequality (4).
- Otherwise, u'' must be of the form $u_1 u_2$. Since $u_1 u_2$ is good for cm', we have $|u_1|, |u_2| < cm'$, and $|u_1 u_2| = |u_1| + |u_2| + 1 \geq cm'$. Let u_0 be u_i such that $|u_i| \geq |u_{3-i}|$. By the conditions above and the inequality (4), we have $|Expl^k_{m'}| + 2d_0 \leq |u_0| < cm'$ as required.

□

H Proof of the Main Theorem

The following lemma will be used in the main proof.

Lemma 24. *Let $\delta, \iota, \xi \geq 2$ be integers. There exists a constant d such that*

$$\#(\mathcal{U}^\nu_n(\delta, \iota, \xi)) \leq \#\big(\mathcal{U}^\nu_{n+d+n'}(\delta, \iota, \xi)\big)$$

holds for any $n, n' \geq 0$.

Proof. Let $d = d_0$ where d_0 is the constant in Corollary 1. We show that there is a injection $f : \mathcal{U}^\nu_n(\delta, \iota, \xi) \to \mathcal{U}^\nu_{n+d+n'}(\delta, \iota, \xi)$. For a 0/1-$\kappa$-context $u \in \mathcal{U}^\nu_n(\delta, \iota, \xi)$, we define

$$f(u) \triangleq S_\kappa[u]$$

where S_κ is a 1-context determined by the context typing κ of u such that $S_\kappa[u] \in \mathcal{U}^\nu_{n+d+n'}(\delta, \iota, \xi)$, i.e., $|S_\kappa| = d + n'$, and $S_\kappa[u]$ is a κ-context. The existence of such S_κ is assured by Corollary 1, and it is clear by the definition that f is an injection. □

Now, we give a complete proof of the main theorem.

Proof (of Theorem 1). Given $\delta, \iota, \xi, k \geq 2$ in the assumption, let $k = \max(\delta, \iota)$. We take the constants b and c in Lemma 4. Let $n \geq 2^{2^b}$, $m' \triangleq \lceil \log^{(2)}(n) \rceil$ and $m \triangleq cm'$; then $m' \geq \log^{(2)}(n) = b$. By Lemma 15, there is a bijection

$$g : \Lambda_n^{\alpha}(\delta, \iota, \xi) \quad \cong \coprod_{(E,u) \in \widehat{\mathcal{B}}_n^m(\delta,\iota,\xi)} \widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi).$$

For each $(E, u) \in \widehat{\mathcal{B}}_n^m(\delta, \iota, \xi)$, we have $\widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi) = \widehat{U}_2^{\delta,\iota,\xi,m,E}(\delta, \iota, \xi) \times \cdots \times \widehat{U}_{\mathtt{shn}(E)}^{\delta,\iota,\xi,m,E}(\delta, \iota, \xi)$ by Lemma 16. Let $\overline{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi) \triangleq \{u_2 \cdots u_{\mathtt{shn}(E)} \in \widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi) \mid \mathit{Expl}_{m'}^k \not\preceq u_i$ for all $2 \leq i \leq \mathtt{shn}(E)\}$. Then, by Lemma 3 (Item 5), we can restrict g to an injection:

$$g' : \{[t]_\alpha \in \Lambda_n^{\alpha}(\delta, \iota, \xi) \mid \beta(t) < \mathbf{exp}_{k-2}(n)\} \;\rightarrowtail \coprod_{(E,u) \in \widehat{\mathcal{B}}_n^m(\delta,\iota,\xi)} \overline{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi)$$

By Lemma 4, every $\widehat{U}_{E \centerdot i}^m(\delta, \iota, \xi)$ contains some u_i' such that $\mathit{Expl}_{m'}^k \preceq u_i'$, so $\overline{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi) \subseteq (\widehat{U}_{E \centerdot 2}^m(\delta, \iota, \xi) \setminus \{u_2'\}) \times \cdots \times (\widehat{U}_{E \centerdot \mathtt{shn}(E)}^m(\delta, \iota, \xi) \setminus \{u_{\mathtt{shn}(E)}'\})$. Then we have

$$\frac{\#\left(\overline{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi)\right)}{\#\left(\widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi)\right)} \leq \frac{\prod_{i=2}^{\mathtt{shn}(E)} (\#\left(\widehat{U}_{E \centerdot i}^m(\delta, \iota, \xi)\right) - 1)}{\prod_{i=2}^{\mathtt{shn}(E)} \#\left(\widehat{U}_{E \centerdot i}^m(\delta, \iota, \xi)\right)} = \prod_{i=2}^{\mathtt{shn}(E)} \left(1 - \frac{1}{\#\left(\widehat{U}_{E \centerdot i}^m(\delta, \iota, \xi)\right)}\right).$$

By Lemma 24, there exists a constant d such that for any $n', n'' \geq 0$ we have $\#(\mathcal{U}_{n'}^{\nu}(\delta, \iota, \xi)) \leq \#\big(\mathcal{U}_{n'+d+n''}^{\nu}(\delta, \iota, \xi)\big)$. Hence we can deduce

$$\begin{aligned} \frac{\#\left(\overline{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi)\right)}{\#\left(\widehat{\mathcal{P}}_{E,u}^m(\delta, \iota, \xi)\right)} &\leq \prod_{i=2}^{\mathtt{shn}(E)} \left(1 - \frac{1}{\#\left(\widehat{U}_{E \centerdot i}^m(\delta, \iota, \xi)\right)}\right) \leq \prod_{i=2}^{\mathtt{shn}(E)} \left(1 - \frac{1}{\#\left(\mathcal{U}_{|E \centerdot i|}^{\nu}(\delta, \iota, \xi)\right)}\right) \\ &= \left(1 - \frac{1}{\#\big(\mathcal{U}_{2m+d}^{\nu}(\delta, \iota, \xi)\big)}\right)^{\mathtt{shn}(E)-1} \leq \left(1 - \frac{1}{\#\big(\mathcal{U}_{2m+d}^{\nu}(\delta, \iota, \xi)\big)}\right)^{\frac{n}{4m}} \end{aligned}$$

where note that $m \leq |E \centerdot i| < 2m$ for each $2 \leq i \leq \mathtt{shn}(E)$ by Lemma 13 and $\mathtt{shn}(E) - 1 \geq n/4m$ by Lemma 14.

Now, for sufficiently large n, we have

$$
\begin{aligned}
&\frac{\#\bigl(\{[t]_\alpha \in \Lambda_n^\alpha(\delta,\iota,\xi) \mid \beta(t) < \mathbf{exp}_{k-2}(n)\}\bigr)}{\#(\Lambda_n^\alpha(\delta,\iota,\xi))} \\
\leq\ & \frac{\sum_{(E,u)\in\widehat{\mathcal{B}}_n^m(\delta,\iota,\xi)} \#\Bigl(\overline{\mathcal{P}}_{E,u}^m(\delta,\iota,\xi)\Bigr)}{\sum_{(E,u)\in\widehat{\mathcal{B}}_n^m(\delta,\iota,\xi)} \#\Bigl(\widehat{\mathcal{P}}_{E,u}^m(\delta,\iota,\xi)\Bigr)} \\
\leq\ & \frac{\sum_{(E,u)\in\widehat{\mathcal{B}}_n^m(\delta,\iota,\xi)} \left(\left(1-\frac{1}{\#(\mathcal{U}_{2m+d}^\nu(\delta,\iota,\xi))}\right)^{\frac{n}{4m}} \times \#\Bigl(\widehat{\mathcal{P}}_{E,u}^m(\delta,\iota,\xi)\Bigr)\right)}{\sum_{(E,u)\in\widehat{\mathcal{B}}_n^m(\delta,\iota,\xi)} \#\Bigl(\widehat{\mathcal{P}}_{E,u}^m(\delta,\iota,\xi)\Bigr)} \\
=\ & \left(1-\frac{1}{\#\bigl(\mathcal{U}_{2m+d}^\nu(\delta,\iota,\xi)\bigr)}\right)^{\frac{n}{4m}} \\
\leq\ & \left(1-\frac{1}{c'\overline{\gamma}(\delta,\iota,\xi)^{2m}}\right)^{\frac{n}{4m}} \qquad (\because \text{ for sufficiently large } n \text{ by Lemma 1}) \\
=\ & \left(1-\frac{1}{c'\overline{\gamma}(\delta,\iota,\xi)^{2c\lceil\log^{(2)}(n)\rceil}}\right)^{\frac{n}{4c\lceil\log^{(2)}(n)\rceil}} \longrightarrow 0 \quad (\text{as } n \longrightarrow \infty)
\end{aligned}
$$

where the last convergence is shown in Appendix I. Thus, we have proved Theorem 1. □

I Proof of the Convergences

Proof. First note that, on the convergence in the proof of Proposition 3, we can assume that $\#(A) \geq 2$ and we have

$$
\begin{aligned}
&\lim_{n\to\infty}\left(1-\frac{1}{\#(A)^{\lceil\log^{(2)}(n)\rceil}}\right)^{\left\lfloor\frac{n}{\lceil\log^{(2)}(n)\rceil}\right\rfloor} \\
\leq\ & \lim_{n\to\infty}\left(1-\frac{1}{\#(A)^{2\log^{(2)}(n)}}\right)^{\frac{1}{2}\frac{n}{2\log^{(2)}(n)}} \\
=\ & \left(\lim_{n\to\infty}\left(1-\frac{1}{\left(\#(A)^2\right)^{\log^{(2)}(n)}}\right)^{\frac{n}{\log^{(2)}(n)}}\right)^{\frac{1}{4}}.
\end{aligned}
$$

Also, on the convergence in the proof of Theorem 1, we have

$$
\begin{aligned}
&\lim_{n\to\infty}\left(1-\frac{1}{c'\overline{\gamma}(\delta,\iota,\xi)^{2c\lceil\log^{(2)}(n)\rceil}}\right)^{\frac{n}{4c\lceil\log^{(2)}(n)\rceil}}\\
\leq\ &\lim_{n\to\infty}\left(1-\frac{1}{\overline{\gamma}(\delta,\iota,\xi)^{\log^{(2)}(n)}\overline{\gamma}(\delta,\iota,\xi)^{4c(\log^{(2)}(n))}}\right)^{\frac{n}{8c(\log^{(2)}(n))}}\\
\leq\ &\left(\lim_{n\to\infty}\left(1-\frac{1}{\left(\overline{\gamma}(\delta,\iota,\xi)^{4c+1}\right)^{\log^{(2)}(n)}}\right)^{\frac{n}{\log^{(2)}(n)}}\right)^{\frac{1}{8c}}.
\end{aligned}
$$

Hence, it is enough to show

$$\lim_{n\to\infty}\left(1-\frac{1}{\gamma^{\log^{(2)}(n)}}\right)^{\frac{n}{\log^{(2)}(n)}}=0$$

for any $\gamma>1$, where $\gamma=\#(A)^2$ for Proposition 3 and $\gamma=\overline{\gamma}(\delta,\iota,\xi)^{4c+1}$ for Theorem 1.

Now we transform the variable n by $\gamma^{\log^{(2)}(n)}=x$, i.e., by $n=2^{\left(x^{\frac{1}{\log\gamma}}\right)}$; then

$$
\begin{aligned}
\lim_{n\to\infty}\left(1-\frac{1}{\gamma^{\log^{(2)}(n)}}\right)^{\frac{n}{\log^{(2)}(n)}}&=\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^{\left(\frac{\log\gamma}{\log x}2^{\left(x^{\frac{1}{\log\gamma}}\right)}\right)}\\
&=\left(\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^{\left(\frac{1}{\log x}2^{\left(x^{\frac{1}{\log\gamma}}\right)}\right)}\right)^{\log\gamma}
\end{aligned}
$$

where $\log\gamma>0$. Then

$$
\begin{aligned}
&\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^{\left(\frac{1}{\log x}2^{\left(x^{\frac{1}{\log\gamma}}\right)}\right)}\\
=\ &\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^{\left(2^{\left(x^{\frac{1}{\log\gamma}}-\log^{(2)}(x)\right)}\right)}\\
\leq\ &\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^{\left(2^{\left(\frac{1}{2}x^{\frac{1}{\log\gamma}}\right)}\right)} \qquad \left(\text{since } \frac{1}{2}x^{\frac{1}{\log\gamma}}\geq\log^{(2)}(x) \text{ for } x>1\right)\\
=\ &\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^{\left(\sqrt{2}^{\left(x^{\frac{1}{\log\gamma}}\right)}\right)}\\
=\ &0
\end{aligned}
$$

where the last equation can be calculated by applying log(−) to the left-hand-side:

$$\begin{aligned}
&\lim_{x\to\infty} \log\left(\left(1-\frac{1}{x}\right)^{\left(\sqrt{2}^{\left(x^{\frac{1}{\log\gamma}}\right)}\right)}\right)\\
=\;&\lim_{x\to\infty} \frac{\log\left(1-\frac{1}{x}\right)}{\sqrt{2}^{\left(-x^{\frac{1}{\log\gamma}}\right)}}\\
=\;&\lim_{x\to\infty} \frac{\frac{x^{-2}}{(\log_e 2)(1-\frac{1}{x})}}{\left(\log_e\sqrt{2}\right)\sqrt{2}^{\left(-x^{\frac{1}{\log\gamma}}\right)}\left(-\frac{1}{\log\gamma}x^{\left(\frac{1}{\log\gamma}-1\right)}\right)} \qquad \text{(by the L'Hôpital's Rule)}\\
=\;&\lim_{x\to\infty} \frac{-(\log\gamma)\sqrt{2}^{\left(x^{\frac{1}{\log\gamma}}\right)}}{\left(\log_e\sqrt{2}\right)\left(\log_e 2\right)x^{\left(\frac{1}{\log\gamma}+1\right)}} \lim_{x\to\infty}\frac{1}{1-\frac{1}{x}}\\
=\;&-\infty.
\end{aligned}$$

□