On the Significance of the Collapse Operation

Paweł Parys
University of Warsaw
ul. Banacha 2, 02-097 Warszawa, Poland
Email: parys@mimuw.edu.pl

***Abstract*—We show that deterministic collapsible pushdown automata of second level can recognize a language which is not recognizable by any deterministic higher order pushdown automaton (without collapse) of any level. This implies that there exists a tree generated by a second level collapsible pushdown system (equivalently: by a recursion scheme of second level), which is not generated by any deterministic higher order pushdown system (without collapse) of any level (equivalently: by any safe recursion scheme of any level). As a side effect, we present a pumping lemma for deterministic higher order pushdown automata, which potentially can be useful for other applications.**

I. Introduction

In verification we often approximate an arbitrary program by a program with variables from a finite domain, remembering only a part of information. Then the outcome of some conditions in the program (e.g. in the *if* or *while* statements) cannot be determined, hence they are replaced by a nondeterministic choice (branching). If the program does not use recursion, the set of its possible control flows is a regular language, and the program itself is (in a sense) a deterministic finite automaton recognizing it. If the program contains recursion, we obtain a deterministic context free language, and from the program one can construct a deterministic pushdown automaton (PDA for short) recognizing this language. In other words, stack can be used to simulate recursion (notice that the same is true for compilers: they convert a recursive program into a program using stack). In verification it is interesting to analyze the possibly infinite tree of all possible control flows of a program. This tree has a decidable MSO theory [1].

A next step is to consider higher order programs, i.e. programs in which procedures can take procedures as parameters. Such programs closely correspond to so-called higher order recursion schemes and to typed λ-terms. They no longer can be simulated by classical PDA. Here higher order PDA come into play. They were originally introduced by Maslov [2]. In automata of level n we have a level n stack of level $n-1$ stacks of ... of level 1 stacks. The idea is that the PDA operates only on the topmost level 1 stack, but additionally it can make a copy of the topmost stack of some level, or can remove the topmost stack of some level. However the correspondence between higher order automata and recursion schemes (programs) is not perfect. Trees generated (in suitable sense) by a PDA of level n coincide with higher order recursion schemes of level n with *safety* restriction [3]. See [4], [5] for another characterizations of the same hierarchy. It is important that these trees have decidable MSO theory [3].

To overcome the safety restriction, a new model of pushdown automata were introduced, called collapsible pushdown automata [6], [7]. These automata are allowed to perform an additional operation called *collapse* (or *panic* in [6]); it allows to remove all stacks on which a copy of the currently topmost stack symbol is present. These automata correspond to all higher order recursion schemes (not only safe ones) [8], and trees generated by them also have decidable MSO theory [9]. It is also worth to mention that verification of some real life higher order programs can be performed in reasonable time [10].

A question arises if these two hierarchies are possibly the same hierarchy? This is an open problem stated in [3] and repeated in other papers concerning higher order PDA [6], [11], [9], [8]. A partial answer to this question was given in [12]: there exists a language recognized by a deterministic collapsible pushdown automaton of second level, which is not recognized by any deterministic higher order pushdown automaton (without collapse) of second level. We prove a stronger property, that the sum of both hierarchies is different, which is our main theorem.

Theorem 1.1. *There exists a language recognized by a deterministic collapsible pushdown automaton of second level, which is not recognized by any deterministic higher order pushdown automaton (without collapse) of any level.*

The result can be also stated as follows (the parts about recursion schemes follow from the equivalences mentioned above).

Corollary 1.2. *There exists a tree generated by a collapsible pushdown system of second level (equivalently: by a recursion scheme of second level), which is not generated by any higher order pushdown system (without collapse) of any level (equivalently: by a safe recursion scheme of any level).*

This confirms that the correspondence between higher order recursion schemes and higher order pushdown systems is not perfect. The language used in Theorem 1.1 (after some adaptations) comes from [3] and from that time was conjectured to be a good example.

As a side effect, in Section VI we present a pumping lemma for deterministic higher order pushdown automata. Although its formulation is not very natural, we believe it may be

Work partially supported by the Polish Ministry of Science grant nr N N206 567840.

useful for some other applications. Our lemma is similar (but incomparable: neither weaker or stronger) to the pumping lemma from [13] (in fact, the lemma from [13] was already stated by Blumensath [14], but there is an irrecoverable error in his proof). Earlier, several pumping lemmas related to the second level of the pushdown hierarchy were proposed [15], [16], [17].

Related work: One may ask a similar question for nondeterministic automata rather than for deterministic ones: is there a language recognized by a nondeterministic collapsible pushdown automaton, which is not recognized by any nondeterministic higher order pushdown automaton (without collapse). This is an independent problem. The answer is known only for level 2 and is opposite: one can see that for level 2 the collapse operation can be simulated by nondeterminism, hence normal and collapsible nondeterministic level 2 PDA recognize the same languages [11]. However it seems that in context of verification considering deterministic automata is a more natural choice, for the following reasons. First, most problems for nondeterministic PDA are not decidable: even the very basic problem of universality for level 1 PDA is undecidable. Second, we want to verify deterministic programs (possibly with some not deterministic input). A nondeterministic program is something rather strange: it has an oracle which says what to do in order to accept. Normally, when a program is going to make some not deterministic choice, we want to analyze all possibilities, not only these which are leading to some „acceptance" (hence we have branching, not nondeterminism).

We know [8] that there is a collapsible pushdown graph of level 2, which has undecidable MSO theory, hence which is not a pushdown graph of any level (as they all have decidable MSO theory).

In [18] we simultaneously prove that the hierarchy of collapsible pushdown trees (and also graphs) is infinite, i.e. that for each level there exists a tree generated by a collapsible pushdown system of that level which is not generated by any collapsible pushdown system of a lower level.

II. Preliminaries

An n-th level *deterministic higher order pushdown automaton* (n-HOPDA for short) is a tuple $(A, \Gamma, \gamma_I, Q, q_I, F, \delta)$ where A is an input alphabet, Γ is a stack alphabet, $\gamma_I \in \Gamma$ is an initial stack symbol, Q is a set of states, $q_I \in Q$ is an initial state, $F \subseteq Q$ is a set of accepting states, and δ is a transition function which maps every element of $Q \times \Gamma$ into one of the following operations:

- $\mathsf{read}(f)$, where $f : A \to Q$ is an injective function,
- $\mathsf{pop}^k(q)$, where $1 \le k \le n$ and $q \in Q$, or
- $\mathsf{push}^k(t^0, q)$, where $1 \le k \le n$, and $t^0 \in \Gamma$, and $q \in Q$.

The letter n is used exclusively for the level of pushdown automata.

For any alphabet Γ (of stack symbols) we define a *k-th level pushdown store* (k-pds for short) as an element of the following set Γ^k_*:

$$\Gamma^0_* = \Gamma,$$
$$\Gamma^k_* = (\Gamma^{k-1}_*)^* \qquad \text{for } 1 \le k \le n.$$

In other words, a 0-pds is just a single symbol, and a k-pds for $1 \le k \le n$ is a (possibly empty) sequence of $(k-1)$-pds's. The last element of a k-pds is also called the *topmost* one. For any $s^k \in \Gamma^k_*$ and $s^{k-1} \in \Gamma^{k-1}_*$ we write $s^k : s^{k-1}$ for the k-pds obtained from s^k by placing s^{k-1} at its end. The operator ":" is assumed to be right associative, i.e. $s^2 : s^1 : s^0 = s^2 : (s^1 : s^0)$. We say for $k \ge 1$ that a k-pds is *proper* if it is nonempty and every $(k-1)$-pds in it is proper; a 0-pds is always proper.

A *configuration* of $\mathcal{A}$ consists of a state and of a proper n-pds, i.e. is an element of $Q \times \Gamma^n_*$ in which the n-pds is proper. The *initial* configuration consists of the initial state q_I and of the n-pds containing only one 0-pds, which is the initial stack symbol γ_I. For a configuration c, its state is denoted $state(c)$, and its n-pds is denoted $\pi(c)$.

Next, we define when a configuration d is a *successor* of a configuration c. Let $p = state(c)$, and let s^0 be the topmost 0-pds of $\pi(c)$. We have three cases depending on $\delta(p, s^0)$:

- if $\delta(p, s^0) = \mathsf{read}(f)$ then $state(d) = f(a)$ for some $a \in A$, and $\pi(d) = \pi(c)$,
- if $\delta(p, s^0) = \mathsf{pop}^k(q)$ then $state(d) = q$, and $\pi(d)$ is obtained from $\pi(c)$ by replacing its topmost k-pds $s^k : s^{k-1}$ by s^k (i.e. we remove the topmost $(k-1)$-pds; in particular the topmost k-pds of $\pi(c)$ has to contain at least two $(k-1)$-pds's),
- if $\delta(p, s^0) = \mathsf{push}^k(t^0, q)$ then $state(d) = q$, and $\pi(d)$ is obtained from $\pi(c)$ by replacing its topmost k-pds $s^k : s^{k-1}$ by $(s^k : s^{k-1}) : s^{k-1}$, and then by replacing its topmost 0-pds by t^0 (i.e. we copy the topmost $(k-1)$-pds, and then we change the topmost symbol in the copy[1]).

Notice that most configurations have exactly one successor. However when the operation is read, there are several successors. It is also possible that there are no successors: when the operation is pop^k but there is only one $(k-1)$-pds on the topmost k-pds.

A *run* is a function R from numbers $0, 1, \dots, l$ (for some $l \ge 0$) to configurations such that $R(i)$ is a successor of $R(i-1)$ for $1 \le i \le l$. The number l is called the *length* of R, and denoted $|R|$. We say that R is a run from $R(0)$ to $R(|R|)$. For $0 \le x \le y \le |R|$ we can consider the *subrun of R from x to y*; this is the run of length $y - x$ which maps i to $R(i + x)$. For two runs R, S such that $S(0) = R(|R|)$ we can consider their *composition*; this is the run of length $|R| + |S|$ which maps $i \le |R|$ to $R(i)$, and $i > |R|$ to $S(i - |R|)$. We also consider infinite runs.

The *word read by a run* is a word over the input alphabet A. For a run from a configuration c to its successor d, it is the empty word if the operation between them is pop or

[1] In the classical definition the topmost symbol can be changed only when $k = 1$ (for $k \ge 2$ it has to be $s^0 = t^0$). We make this (not important) extension to have an unified definition of push^k for every k.

TABLE I
STACK CONTENTS AND SETS pre FOR THE EXAMPLE RUN

i	$\pi(R(i))$	$pre^0_R(i)$	$pre^1_R(i)$	$pre^2_R(i)$	$pre^3_R(i)$
0	$[[[ab]]]$	$\{0\}$	$\{0\}$	$\{0\}$	$\{0\}$
1	$[[[ab][ac]]]$	$\{0,1\}$	$\{0,1\}$	$\{0,1\}$	$\{0,1\}$
2	$[[[ab][ac]][[ab][ad]]]$	$\{0,1,2\}$	$\{0,1,2\}$	$\{0,1,2\}$	$\{0,1,2\}$
3	$[[[ab][ac]][[ab][a]]]$	$\{3\}$	$\{0,1,2,3\}$	$\{0,1,2,3\}$	$\{0,1,2,3\}$
4	$[[[ab][ac]][[ab][a]][[ab][e]]]$	$\{3,4\}$	$\{0,1,2,3,4\}$	$\{0,1,2,3,4\}$	$\{0,1,2,3,4\}$
5	$[[[ab][ac]][[ab][a]][[ab]]]$	$\{0,5\}$	$\{0,5\}$	$\{0,1,2,3,4,5\}$	$\{0,1,2,3,4,5\}$
6	$[[[ab][ac]][[ab][a]]]$	$\{3,6\}$	$\{0,1,2,3,6\}$	$\{0,1,2,3,6\}$	$\{0,1,2,3,4,5,6\}$

push. If the operation is $\mathsf{read}(f)$, this is the one-letter word consisting on the letter a for which $state(d) = f(a)$ (this letter is determined uniquely, as f is injective). For a longer run R this is defined as the concatenation of the words read by every subrun of R from $i-1$ to i (for $1 \leq i \leq k$). A word w is accepted by $\mathcal{A}$ if it is read by some run from the initial configuration to a configuration having an accepting state. The *language* recognized by $\mathcal{A}$ is the set of words accepted by $\mathcal{A}$.

A *position* is a vector $x = (x_n, x_{n-1}, \dots, x_1)$ of n positive integers. The symbol on position x in configuration c (which is an element of Γ) is defined in the natural way (we take the x_n-th $(n-1)$-pds of $\pi(c)$, then its x_{n-1}-th $(n-2)$-pds, and so on; elements of pds's are numbered from bottom to top). We say that x is a position of c, if at position x there is a symbol in c.

For $0 \leq k \leq n$, by $top^k(c)$ we denote the position of the bottommost symbol of the topmost k-pds of c. In particular $top^0(c)$ is the topmost position of c.

For any run R, indices $0 \leq a \leq b \leq |R|$, and a position y of $R(b)$, we define a position $hist_R(b,y)(a)$. It is y when $b = a$. It is y also when $b = a+1$, and the operation between $R(a)$ and $R(b)$ is read or pop, as well as when the operation is push^k and y is not in the topmost $(k-1)$-pds of $R(b)$. If the operation between $R(a)$ and $R(b)$ is push^k and y is in the topmost $(k-1)$-pds of $R(b)$, then $hist_R(b,y)(a)$ is the position of $R(a)$ from which a symbol was copied to y (i.e. this is y with the $(n-k+1)$-th coordinate decreased by 1). When $b > a+1$, $hist_R(b,y)(a)$ is defined (by induction) as $hist_R(a+1, hist_R(b,y)(a+1))(a)$. In other words, $hist_R(b,y)(a)$ is the (unique) position of $R(a)$, from which the symbol was copied to y in $R(b)$.

For $0 \leq k \leq n$, a run R, and an index $0 \leq b \leq |R|$ we define a set $pre^k_R(b)$ consisting of these indices a ($0 \leq a \leq b$) for which $hist_R(b, top^k(R(b)))(a) = top^k(R(a))$. We also denote $pre^{-1}_R(b) = \{b\}$. Intuitively, $a \in pre^k_R(b)$ means that the topmost k-pds of $R(b)$ "comes from" the topmost k-pds of $R(a)$, in the sense that the topmost k-pds of $R(b)$ is a copy of the topmost k-pds of $R(a)$, but possibly some changes were done to it. Notice that $pre^k_R(b) \subseteq pre^l_R(b)$ for $k \leq l$, and $pre^k_R(a) = pre^k_R(b) \cap \{0,1,\dots,a\}$ for $a \in pre^k_R(b)$.

Example 2.1. Consider a PDS of level 3. Below, brackets are used in descriptions of pds's as follows: symbols taken in brackets form one 1-pds, 1-pds's taken in brackets form one 2-pds, and 2-pds's taken in brackets form one 3-pds. Consider a run R of length 6 in which $\pi(R(0)) = [[[ab]]]$ and the operations between consecutive configurations are (we omit the state):

$$\mathsf{push}^2(c),\ \mathsf{push}^3(d),\ \mathsf{pop}^1,\ \mathsf{push}^3(e),\ \mathsf{pop}^2,\ \mathsf{pop}^3.$$

The contents of the 3-pds's of the configurations in the run, and the pre sets, are presented in Table I. In configuration $R(0)$ symbol a is on position $(1,1,1)$ and symbol b is on position $(1,1,2)$. We have

$$hist_R(2,(2,2,1))(1) = (1,2,1), \quad \text{and}$$
$$hist_R(2,(2,2,1))(0) = (1,1,1).$$

Notice that positions y in $R(b)$ and $hist_R(b,y)(a)$ in $R(a)$ not necessarily contain the same symbol, for example on position $(1,2,2)$ in $R(1)$ we have c, and on position $(1,1,2)$ in $R(0)$ we have b, but $hist_R(1,(1,2,2))(0) = (1,1,2)$.

Collapsible 2-HOPDA: In Section VII we also use deterministic collapsible pushdown automata of second level (2-CPDA for short). Such automaton is defined like 2-HOPDA, with the following differences. A 0-pds contains now two parts: a symbol from Γ, and a natural number, but still only the symbol (together with a state) is used to determine which transition is performed from a configuration. The push^1 operation sets the number in the topmost 0-pds to the current size of the 2-pds (the number of 1-pds's). We have a new operation $\mathsf{collapse}(q)$. When it is performed between configurations c and d, then $state(d) = q$, and $\pi(d)$ is obtained from $\pi(c)$ by removing its topmost 1-pds's, so that only $k-1$ of them is left, where k is the number stored in the topmost 0-pds of c (intuitively, we remove all 1-pds's on which the topmost 0-pds is present).

III. BASIC PROPERTIES OF RUNS

In this section we present four propositions, which follow easily from definitions.

Proposition 3.1. *Let R be a run of an n-HOPDA, let $0 \leq k \leq n$, and let $i \leq j$ be such that $pre^k_R(j) \cap \{i, i+1, \dots, j\} = \{i,j\}$. Then*

- *the topmost k-pds of $R(i)$ and $R(j)$ is the same; additionally for every position x in the topmost k-pds of $R(j)$, $hist_R(j,x)(i)$ is the corresponding position in $R(i)$, or*
- *$j = i+1$ and the operation between $R(i)$ and $R(j)$ is pop^r for $r \leq k$, or push^r for $r \leq k$.*

Proof: For $j = i$ and for $j = i+1$ we immediately get one of the possibilities. Otherwise, we look at the history of the topmost k-pds of $R(j)$. It is covered by the first operation after $R(i)$, and then it is not the topmost k-pds until $R(j)$. Thus it remains unchanged (we get the first possibility). ■

Proposition 3.2. *Let R be a run of an n-HOPDA, let $1 \leq k \leq n$, and let i, j be such that $i \in pre_R^k(j)$. Then $i \in pre_R^{k-1}(j)$ if and only if the size (the number of $(k-1)$-pds's) of the topmost k-pds of $R(i)$ is not greater than the size of the topmost k-pds of $R(l)$ for every $l \in pre_R^k(j) \cap \{i, i+1, \dots, j\}$.*

Proof: Let r_l be the size of the topmost k-pds of $R(l)$ (for each l), and let $y_l(r)$ be the bottommost symbol of the r-th $(k-1)$-pds of the topmost k-pds of $R(j)$ (for $1 \leq r \leq r_l$). Let r_{min} be the smallest among r_l for $l \in pre_R^k(j)$, $l \geq i$. We will prove that for $1 \leq r \leq r_j$,

$$hist_R(j, y_j(r))(i) = y_i(\min(r, r_{min})). \tag{1}$$

Then the thesis of the proposition follows immediately, as $i \in pre_R^{k-1}(j)$ if and only if $hist_R(j, y_j(r_j))(i) = y_i(r_i)$. To prove (1) we make an induction on j. For $j = i$ it is true. Otherwise, let $l < j$ be the greatest index being in $pre_R^k(j)$. We have some of the two cases described by Proposition 3.1 (where l is taken as i), and for both of them we see that for $1 \leq r \leq r_j$,

$$hist_R(j, y_j(r))(l) = y_l(\min(r, r_l)).$$

Together with the induction assumption for l, this implies (1). ■

Proposition 3.3. *Let $1 \leq k \leq n$, let R be a run of an n-HOPDA, and let $0 \leq i \leq j \leq l \leq |R|$. Assume that $i \notin pre_R^{k-1}(j)$ and $j \in pre_R^k(l)$. Then $i \notin pre_R^{k-1}(l)$.*

Proof: If $i \notin pre_R^k(l)$ then as well $i \notin pre_R^{k-1}(l)$. So assume that $i \in pre_R^k(l)$; then also $i \in pre_R^k(j)$. By Proposition 3.2 we know that for some $a \in pre_R^k(j) \cap \{i, i+1, \dots, j\}$ the size of the topmost k-pds of $R(a)$ is smaller than the size of the topmost k-pds of $R(i)$. Again Proposition 3.2 (now for l used as j) gives us that $i \notin pre_R^{k-1}(l)$. ■

Proposition 3.4. *Let $1 \leq k \leq n$. Let R be a run of an n-HOPDA, and let $1 \leq k \leq n$. Assume that $0 \in pre_R^{k-1}(|R|-1)$ and $0 \notin pre_R^{k-1}(|R|)$ and $|R| - 1 \in pre_R^k(|R|)$. Then the topmost k-pds of $R(0)$ after removing its topmost $(k-1)$-pds is equal to the topmost k-pds of $R(|R|)$.*

Proof: We also have $0 \in pre_R^k(|R| - 1)$ and $0 \in pre_R^k(|R|)$. Let r be the size of the topmost k-pds of the topmost k-pds of $R(0)$. From Proposition 3.2 (used for 0 as i and $|R| - 1$ as j) we know that the size of the topmost k-pds of $R(l)$ is at least r, for every $l \in pre_R^k(|R| - 1)$. Moreover from Proposition 3.1 we know that between these configuration $R(l)$, we can modify the topmost k-pds only using a single operation; thus the first $r - 1$ $(k-1)$-pds's of the topmost k-pds of $R(0)$ and of $R(|R| - 1)$ are the same. From Proposition 3.2, now used for 0 as i and $|R|$ as j, we get that the size of the topmost k-pds of $R(|R|)$ is smaller than r. Thus it has to be $r - 1$ (the last operation has to be pop^k), so we get the thesis. ■

IV. Types and sequence equivalence

Let $\mathcal{A}$ be an n-HOPDA over input alphabet A which contains a distinguished symbol $\sharp$, and over stack alphabet Γ. Let $\varphi\colon A^* \to M$ be a morphism into a finite monoid. In the appendix we define the following objects:

- a finite set $\mathcal{T}_{\mathcal{A},\varphi}$,
- a function $type_{\mathcal{A},\varphi}$ which assigns to every configuration of $\mathcal{A}$ an element of $\mathcal{T}_{\mathcal{A},\varphi}$,
- a partial order $\leq$ over $\mathcal{T}_{\mathcal{A},\varphi}$,
- an equivalence relation over infinite sequences of configurations of $\mathcal{A}$, called $(\mathcal{A}, \varphi)$-sequence equivalence, which has finitely many equivalence classes.

Basing on $type_{\mathcal{A},\varphi}$, for each $0 \leq k \leq n$ we define a function $type_{\mathcal{A},\varphi}^k$ which assigns to every configuration c of $\mathcal{A}$ a pair from $\mathcal{T}_{\mathcal{A},\varphi} \times \Gamma_*^k$, which is $type_{\mathcal{A},\varphi}(c)$, and the topmost k-pds of c. We extend partial order $\leq$ to $\mathcal{T}_{\mathcal{A},\varphi} \times \Gamma_*^k$:

$$(t_1, s_1^k) \leq (t_2, s_2^k) \iff t_1 \leq t_2 \text{ and } s_1^k = s_2^k.$$

The important properties of the $type_{\mathcal{A},\varphi}$ function and the $(\mathcal{A}, \varphi)$-sequence equivalence are described by the following two theorems. Theorem 4.1 is already present in [13], but Theorem 4.3 is new.

Theorem 4.1. *Let $\mathcal{A}$ be an n-HOPDA with an input alphabet A, let $\varphi\colon A^* \to M$ be a morphism into a finite monoid M, and let $0 \leq k \leq n$. Let R be a run of $\mathcal{A}$ such that $0 \in pre_R^k(|R|)$, and let c be a configuration such that $type_{\mathcal{A},\varphi}^k(R(0)) \leq type_{\mathcal{A},\varphi}^k(c)$. Then there exists a run S from c such that*

1) *if $|R| > 0$ then $|S| > 0$, and*
2) *$0 \in pre_S^k(|S|)$, and*
3) *the words read by R and by S evaluate to the same under φ, and*
4) *$type_{\mathcal{A},\varphi}^k(R(|R|)) \leq type_{\mathcal{A},\varphi}^k(S(|S|))$.*

Remark 4.2. We explain the intuition why the above theorem is true. Assume $n = 1$ and $k = 0$. Then it is enough if $type_{\mathcal{A},\varphi}^0(c)$ returns the topmost stack symbol of c, and the state of c (and the relation $\leq$ is trivial). Notice that the assumption $0 \in pre_R^0(|R|)$ means that the topmost stack symbol of $R(0)$ is not removed in the whole run. Thus if the topmost stack symbol and the state of c is the same as of $R(0)$, we can perform from c the same sequence of operations as in R.

Assume now that $n = 2$ and $k = 0$. Although again the topmost stack symbol of $R(0)$ is not removed directly, now the run can depend on all the symbols of the topmost 1-pds, after making a copy of it. But the assumption $0 \in pre_R^0(|R|)$ guaranties that if the topmost symbol of $R(0)$ is removed in some copy of the topmost 1-pds, then the copy is later removed. Thus for the topmost 1-pds only the following information is needed for each state q, and each monoid element m: if we start in state q in which states the topmost 1-pds can be removed, if the word read between these two states evaluates to m.

Theorem 4.3. *Let $\mathcal{A}$ be an n-HOPDA with an input alphabet A containing a $\sharp$ symbol, let $\varphi\colon A^* \to M$ be a morphism into* *a finite monoid M, and let $0 \leq k \leq n$. Let $c_1, c_2, c_3, \ldots$ and $d_1, d_2, d_3, \ldots$ be infinite sequences of configurations which are $(\mathcal{A}, \varphi)$-sequence equivalent. Let also R be a run, and $0 = l_0 \leq l_1 \leq \cdots \leq l_r = |R|$ indices such that $l_{j-1} \in pre^k_R(l_j)$ for $1 \leq j \leq r-1$ (not for $j = r$), and in R there is no push^n operation and there is only one pop^n operation, which is the last operation of R. Assume that $type^k_{\mathcal{A},\varphi}(R(0)) = type^k_{\mathcal{A},\varphi}(c_i) = type^k_{\mathcal{A},\varphi}(d_i)$ for each i. Then for each i there exist runs S_i from c_i, and T_i from d_i such that*

1) *the last operation of S_i and of T_i is pop^n, and*
2) *the word read by S_i decomposes as $w_1 w_2 \ldots w_r$, where w_j evaluates under φ to the same as the word read by the subrun of R from l_{j-1} to l_j (for $1 \leq j \leq r$); the same for T_i, and*
3) *let x_i and y_i be the number of the $\sharp$ symbols read by the run S_i and T_i, respectively; then either the sequences $x_1, x_2, x_3, \ldots$ and $y_1, y_2, y_3, \ldots$ are both bounded, or both unbounded.*

Remark 4.4. We explain the intuition why the above theorem is true. Assume $n = 2$ and $k = 0$. To characterize a sequence $c_1, c_2, c_3, \ldots$, for each state q and monoid element m, it is enough to know if it is possible that starting in configuration $(q, \pi(c_i))$ and finishing by a pop^2 operation, we read a word evaluating to m with a bounded number of the $\sharp$ symbols. Then, from every c_i we can perform the same operations as in the initial part of R, as long as the topmost symbol of $R(0)$ is not removed (this is true only for $n = 2$); this part reads a bounded (constant) number of $\sharp$ symbols. It ends in a configuration $(q, \pi(c_i))$ for some state q (which depends only on R). By assumptions, this can not happen before $R(l_{r-1})$. Thus, from $(q, \pi(c_i))$ we need to have a run which just reads a word evaluating to the same as the rest of R. Whether such run can read a bounded number of the $\sharp$ symbols is determined by the equivalence class of the sequence.

V. Milestone configurations

In this section we define so-called milestone configurations and we show their basic properties. The idea of considering milestone configurations comes from [19], but our definition is slightly different. Starting from this section, we often assume that the input alphabet contains a distinguished symbol denoted $\star$ (star).

Definition 5.1. Let $\mathcal{A}$ be an n-HOPDA with input alphabet containing a $\star$ symbol, and let c be a configuration of $\mathcal{A}$. We say that c is a *milestone* (or a milestone configuration) if there exists an infinite run R from c reading only stars, and an infinite set I of indices such that $0 \in I$, and $i \in pre^0_R(j)$ for each $i, j \in I$, $i \leq j$.

Example 5.2. Consider a PDS of level 3. Assume there is a run which begins in a pds $[[[aa]]]$ (the notation is the same as in the previous example), and performs forever the following sequence of operations, in a loop:

$$\mathsf{push}^2(a),\ \mathsf{push}^3(a),\ \mathsf{pop}^1,\ \mathsf{push}^3(a),\ \mathsf{pop}^2,\ \mathsf{push}^3(a).$$

Then the topmost 2-pds is alternatively: $[[aa]]$ or $[[aa][aa]]$ or $[[aa][a]]$. This run does not read any symbols, so it is a degenerate case of an infinite run which reads only stars. Configurations with topmost 2-pds $[[aa]]$ are milestones (and no other configurations in this run). To obtain a less degenerate case, we may consider a loop of operations as above, but containing additionally a read operation; when a star is read, the loop continues (we do not care what happens when any other symbol is read). Then again configurations with topmost 2-pds $[[aa]]$ are milestones.

Lemma 5.3. *Let $\mathcal{A}$ be an n-HOPDA with input alphabet containing a $\star$ symbol. Let R be an infinite run of $\mathcal{A}$ reading only stars. Then, for infinitely many i the configuration $R(i)$ is a milestone.*

Proof: Let $I^n = \mathbb{N}$. For $k = n-1, n-2, \ldots, 0$ we define

$$I^k = \{i \in I^{k+1} | \forall_{j \geq i}(j \in I^{k+1} \Rightarrow i \in pre^k_R(j))\}.$$

It is enough to show that set I^0 is infinite. Then, by definition, I^0 contains only milestone configurations.

We prove that I^k is infinite by induction on k, from $k = n$ down to $k = 0$. The induction basis for $k = n$ is true, because $I^n = \mathbb{N}$. Let now $k \leq n-1$; assume that I^{k+1} is infinite. For each index l, we want to find an index $i \geq l$ which is in I^k. By s_j denote the size of the topmost $(k+1)$-pds of $R(j)$. We can choose an index $i \in I^{k+1}$ such that s_i is minimal among all s_j for $j \in I^{k+1} \cap \{l, l+1, l+2, \ldots\}$. By Proposition 3.2 (used for $k+1$ as k) we see that $i \in I^k$. ■

If c is a milestone, R the (unique) infinite run from c reading only stars, and I a set like in the definition of a milestone, then also $R(i)$ is a milestone for each $i \in I$. The following lemma shows that in fact the set I can contain all i for which $R(i)$ is a milestone.

Lemma 5.4. *Let $\mathcal{A}$ be an n-HOPDA with input alphabet containing a $\star$ symbol. Let R be a run reading only stars, which begins and ends in a milestone. Then $0 \in pre^0_R(|R|)$.*

Proof: Consider the infinite run S from $R(0)$ reading only stars (since $R(0)$ is a milestone, the run is really infinite); R is its prefix. We use the sets I^k from the proof of Lemma 5.3 (for run S). We will show that if $S(i)$ is a milestone, then $i \in I^0$. It will mean that both 0 and $|R|$ are in I^0; it follows that $0 \in pre^0_R(|R|)$.

We prove by induction on k, from $k = n$ down to $k = 0$, that if $S(i)$ is a milestone, then $i \in I^k$. We have $i \in I^n$ for each i. Let $k \leq n-1$. Assume that $S(i)$ is a milestone, and that for each milestone $S(j)$ we have $j \in I^{k+1}$. Choose any $j \in I^{k+1}$, $j \geq i$. We need to prove that $i \in pre^k_S(j)$. By definition of a milestone, we have arbitrarily large l (in particular $l \geq j$) such that $i \in pre^0_S(l)$ and $S(l)$ is a milestone. It follows that $i \in pre^k_S(l)$. We also know that $j, l \in I^{k+1}$ (by induction assumption), so $j \in pre^{k+1}_S(l)$. We use Proposition 3.3 for i, j, l, and $k+1$ (as k). We get that $i \in pre^k_S(j)$, as wanted. ■

We also prove a finitary version of Lemma 5.3, which is used in the proof of the pumping lemma in the next section.

Lemma 5.5. *Let $\mathcal{A}$ be an n-HOPDA with input alphabet containing a $\star$ symbol and stack alphabet Γ, and let $1 \leq k \leq n$. There exists a function $b\colon \Gamma_*^k \to \mathbb{N}$, assigning a number to a k-pds, having the following properties. Let R be a run which reads only stars, and let y be a position of $R(|R|)$. Let s^k be the k-pds of $R(0)$ containing $hist_R(|R|, y)(0)$. Assume that there exist at least $b(s^k)$ indices i such that position $hist_R(|R|, y)(i)$ is in the topmost k-pds of $R(i)$. Then for some i, configuration $R(i)$ is a milestone and position $hist_R(|R|, y)(i)$ is in the topmost k-pds of $R(i)$.*

In order to do prove the above, we need an auxiliary lemma. The intuition behind it is, by analogy to the proof of Lemma 5.3, that we can choose an arbitrarily big set I^k having a big enough set I^{k+1}.

Lemma 5.6. *Let $\mathcal{A}$ be an n-HOPDA with input alphabet containing a $\star$ symbol, and let $0 \leq k \leq n-1$. There exists a function $f^k\colon \mathbb{N}\times\mathbb{N} \to \mathbb{N}$, having the following properties. Let N be a natural number, let R be a run reading only stars, and let $0 \leq x \leq |R|$. Assume that $|pre_R^{k+1}(x)| \geq f^k(r, N)$, where r is the size of the topmost $(k+1)$-pds of $R(\min(pre_R^{k+1}(x)))$. Then there exists an index $y \in pre_R^{k+1}(x)$ such that*

1) *$|pre_R^k(y)| \geq N$, and*
2) *the topmost k-pds of $R(\min(pre_R^k(y)))$ is one of k-pds's in the topmost $(k+1)$-pds of $R(\min(pre_R^{k+1}(x)))$.*

Proof: We prove the lemma by induction on N. For $N = 1$ we can take $f^k(r, 1) = 1$, and then $y = \min(pre_R^{k+1}(x))$. Let now $N \geq 2$. We take

$$f^k(r, N) = 1 + \sum_{m=1}^{r} f^k(m+1, N-1).$$

Fix some R and x satisfying the assumptions. Let $a = \min(pre_R^{k+1}(x))$. By r_i we denote the size of the topmost $(k+1)$-pds of $R(i)$ (for each $i \in pre_R^{k+1}(x)$). Then $r = r_a$.

For each $j \in pre_R^{k+1}(x)$ denote

$$m_j = \min\{r_i : i \in pre_R^{k+1}(x) \wedge i \leq j\}.$$

Notice that $1 \leq m_j \leq r$ (because $r_a = r$) and that $m_j \geq m_{j'}$ for $j \leq j'$. From the formula for $f^k(r, N)$ we see that for some m we have at least $f(m+1, N-1)+1$ indexes $j \in pre_R^{k+1}(x)$ such that $m_j = m$. Choose some such m; let b be the first index such that $m_b = m$, and e the last such index. We have $m = r_b$.

Let b' be the next index after b which is in $pre_R^{k+1}(x)$. Notice that $r_{b'} \leq r_b + 1 = m+1$; this follows from Proposition 3.1 used for R, $k+1$ (as k), b (as i), and b' (as j). Thus we have

$$|pre_R^{k+1}(e) \cap \{b', b'+1, \dots, e\}| \geq f(m+1, N-1) \geq \\ \geq f(r_{b'}, N-1).$$

We use the induction assumption for the subrun of R from b' to e (as R), and $e - b'$ (as x). We obtain an index $y \in pre_R^k(e) \subseteq pre_R^{k+1}(x)$ such that $|pre_R^k(y) \cap \{b', b'+1, \dots, e\}| \geq N-1$. Recall that $r_l \geq r_b$ for every $l \in pre_R^{k+1}(e) \cap \{b, b+1, \dots, e\}$. From Proposition 3.2, used for R, $k+1$ (as k), b (as i), and y (as j) we get that $b \in pre_R^k(y)$. This implies that $|pre_R^k(y)| \geq N$.

Finally we show condition 2. Assume that $pre_R^k(b) \cap \{a, a+1, \dots, b-1\} \neq \emptyset$; let i be any its element. Then $i \in pre_R^{k+1}(x)$, so $r_i > r_b$. This contradicts with Proposition 3.2, used for R, $k+1$ (as k), i, and b (as j). It means that $pre_R^k(b) \cap \{a, a+1, \dots, b-1\} = \emptyset$. On the other hand $a \in pre_R^{k+1}(b)$. Thus the topmost k-pds of $R(b)$ is one of the k-pds's in the topmost $(k+1)$-pds of $R(a)$. ■

Proof of Lemma 5.5: Let $\varphi\colon A^* \to M$ be a morphism into a finite monoid, which checks if a word consists of only $\star$ symbols. Fix some k-pds s^k. Let $N_0 = |\mathcal{T}_{\mathcal{A},\varphi}| \cdot |\Gamma| + 1$, i.e. this is the number of possible values of $type^0_{\mathcal{A},\varphi}$, plus one. For $0 \leq l \leq k-1$ let $N_{l+1} = f^l(r_l, N_l)$, where r_l is the maximal size of an $(l+1)$-pds which appears in s^k, and f^l is the function from Lemma 5.6. We define $b(s^k) = N_k$.

Now take a run R and a position y in $R(|R|)$, such that the assumptions of the lemma are satisfied. First, for each $0 \leq l \leq k$ we want to show that for some m_l ($0 \leq m_l \leq |R|$) we have

a) $|pre_R^l(m_l)| \geq N_l$, and
b) $hist_R(|R|, y)(m_l)$ is in the topmost k-pds of $R(m_l)$, and
c) the topmost l-pds of $R(\min(pre_R^l(m_l)))$ is one of the l-pds's of s^k.

We show this by induction on l, from $l = k$ down to $l = 0$. Consider first $l = k$. As m_k we take the greatest index for which position $hist_R(|R|, y)(m_k)$ is in the topmost k-pds of $R(m_k)$. Then for every index i such that position $hist_R(|R|, y)(i)$ is in the topmost k-pds of $R(i)$, we have $i \in pre_R^k(m_k)$, so by assumption of the lemma we have $|pre_R^k(m_k)| \geq b(s^k) = N_k$. Moreover, the topmost k-pds of $R(\min(pre_R^k(m_k)))$ was not modified from the beginning of the run (as it was not the topmost k-pds). So this k-pds is the same as the k-pds of $R(0)$ containing $hist_R(|R|, y)(0)$, which is s^k.

Let now $l < k$, and assume that we have an index m_{l+1} satisfying a)-c). From condition c) we know that the size of the topmost $(l+1)$-pds of $R(\min(pre_R^{l+1}(m_{l+1})))$ is at most r_l. Thus Lemma 5.6 can be used for l (as k), N_l (as N), R, and m_{l+1} (as x); it requires that $|pre_R^{l+1}(m_{l+1})| \geq f^k(r_l, N_l)$, which is guaranteed by condition a). As m_l we take the index y from the lemma. Observe that such m_l satisfies conditions a)-c): condition a) is the same as condition 1 from the lemma; condition b) is true because $m_l \in pre_R^{l+1}(m_{l+1})$ implies $m_l \in pre_R^k(m_{l+1})$, and we have b) for m_{l+1}; condition c) follows from condition 2 from the lemma, and c) for m_{l+1}.

Next, let S be the unique maximal run from $R(0)$ which reads only stars (it is either infinite, or ends in a configuration with no successor); R is its prefix. Observe that in $pre_S^0(m_0)$ we have two indices $l_0 < l_1$ such that $type^0_{\mathcal{A},\varphi}(S(l_0)) = type^0_{\mathcal{A},\varphi}(S(l_1))$ (because we have more indices than possible values of $type^0_{\mathcal{A},\varphi}$). We also have $l_0 \in pre_S^0(l_1)$. Now, assume that (for some $r \geq 1$) we have a sequence of indices

$l_0 < l_1 < \dots < l_r$ such that for each $0 \le i < r$ it holds $l_i \in pre^0_S(l_{i+1})$ and $type^0_{\mathcal{A},\varphi}(S(l_i)) \le type^0_{\mathcal{A},\varphi}(S(l_{i+1}))$; initially we have such sequence for $r = 1$. We use Theorem 4.1 for 0 (as k), the subrun of S from l_{r-1} to l_r (as R), and $S(l_r)$ (as c). We obtain a run from $S(l_r)$ which reads only stars; it is necessarily a subrun of S from l_r to some l_{r+1}. It holds $l_{r+1} \ne l_r$ (condition 1), $l_r \in pre^0_S(l_{r+1})$ (condition 2), and $type^0_{\mathcal{A},\varphi}(S(l_r)) \le type^0_{\mathcal{A},\varphi}(S(l_{r+1}))$ (condition 4). So we obtain a sequence satisfying the same conditions, but having one more element. We can continue like that for infinity: we obtain an infinite sequence $l_0 < l_1 < l_2 < \dots$ such that for each i it holds $l_i \in pre^0_S(l_{i+1})$ (in particular S is infinite). By transitivity of pre we also have $l_i \in pre^0_S(l_j)$ for each $i < j$. This shows that $S(l_0)$ is a milestone (to the set I required in the definition we take all l_i). Moreover recall that l_0 is from $pre^0_R(m_0)$. Thus, by condition b), position $hist_R(|R|, y)(l_0)$ is in the topmost k-pds of $R(l_0)$. ■

VI. Pumping lemma

In this section we present a pumping lemma which can be used to prove that a language cannot be recognized by an n-HOPDA.

Definition 6.1. Let $\mathcal{A}$ be an n-HOPDA with input alphabet A containing a $\star$ symbol. Let $\varphi\colon A^* \to M$ be a morphism into a finite monoid. We say that a run R' is a *pumping witness* for a run R with respect to φ, if $R(0) = R'(0)$, and the words read by R and by R' evaluate to the same under φ, and for each $0 \le k \le n$,

- $type^k_{\mathcal{A},\varphi}(R(|R|)) \le type^k_{\mathcal{A},\varphi}(R'(|R'|))$, or
- $0 \in pre^{k-1}_R(|R|)$.

We say that run R can be *pumped* with respect to φ if for each r' there exists a pumping witness R' such that the word read by R' begins with at least r' stars.

Intuitively, a run can be pumped, if we can change the number of stars at its beginning in such way that the final configuration does not change too much. In the definition we describe the behavior of the topmost k-pds (for each k). The second option, $0 \in pre^{k-1}_R(|R|)$, corresponds to a situation when a part of the topmost k-pds was created while reading the initial stars; then we do not say anything about the topmost k-pds of $R'(|R'|)$. Otherwise, the topmost k-pds of $R(|R|)$, in some sense, was not touched while reading the initial stars; then we guarantee that $type^k_{\mathcal{A},\varphi}(R(|R|)) \le type^k_{\mathcal{A},\varphi}(R'(|R'|))$ (in particular the topmost k-pds of $R(|R|)$ and of $R'(|R'|)$ are the same).

Theorem 6.2 (Pumping lemma). *Let $\mathcal{A}$ be an n-HOPDA with input alphabet A containing a $\star$ symbol. Let $\varphi\colon A^* \to M$ be a morphism into a finite monoid M, and c a milestone configuration of $\mathcal{A}$. There exists $r > 0$ such that each run from c, which reads a word beginning with (at least) r stars, can be pumped with respect to φ.*

The rest of this section is devoted to a proof of this theorem.

Lemma 6.3. *Let $\mathcal{A}$ be an n-HOPDA with input alphabet A containing a $\star$ symbol, and let $1 \le k \le n$. Let c be a milestone configuration of $\mathcal{A}$. Then there exists a finite set $\mathcal{S}^k$ of k-pds's having the following property. Let R be a run from c reading only stars. Let x be the bottommost position of the topmost $(k-1)$-pds in some k-pds of $R(|R|)$. Assume that $hist_R(|R|, x)(0) \ne top^{k-1}(c)$. Then the k-pds of $R(|R|)$ containing x is in $\mathcal{S}^k$.*

Proof: Let $\mathcal{X}$ be the set containing all k-pds's of c, and additionally the topmost k-pds of c with its topmost $(k-1)$-pds removed. Let $\mathcal{S}^k$ contain all k-pds which can be obtained from a k-pds $s^k \in \mathcal{X}$ by applying at most $b(s^k)$ of push and pop operations,[2] where b is the function from Lemma 5.5.

Fix a run R from c which reads only stars. We say that a k-pds of some configuration $R(i)$ is *c-clear*, if $hist_R(i, x)(0) \ne top^{k-1}(c)$ for x being the bottommost position of the topmost $(k-1)$-pds in the considered k-pds of $R(i)$. Fix some c-clear k-pds of $R(|R|)$, let y be its bottommost position. Our goal is to show that this k-pds is in $\mathcal{S}^k$.

Let i be the smallest index such that the k-pds of $R(i)$ containing $hist_R(|R|, y)(i)$ is c-clear (such i exists, as $|R|$ is a good candidate). We claim that the k-pds of $R(i)$ containing $hist_R(|R|, y)(i)$ is in $\mathcal{X}$. Indeed, one possibility is that $i = 0$, then we are done. Otherwise, the k-pds of $R(i-1)$ containing $hist_R(|R|, y)(i-1)$ is not c-clear, but in the next configuration it becomes c-clear; so necessarily this is the topmost k-pds and $i-1 \in pre^k_R(i)$ (the operation between $R(i-1)$ and $R(i)$ has to be pop^k). We have $0 \in pre^{k-1}_R(i-1)$ (the topmost k-pds of $R(i-1)$ is not c-clear) and $0 \notin pre^{k-1}_R(i)$ (the topmost k-pds of d is c-clear). So we can use Proposition 3.4 for the subrun of R from 0 to i. It says that the topmost k-pds of $R(i)$ can be obtained from the topmost k-pds of c by removing its topmost $(k-1)$-pds; thus it is in $\mathcal{X}$.

Now observe that if $hist_R(|R|, y)(j) = top^k(R(j))$ for some $j \ge i$ (where i as above), then $0 \notin pre^0_R(j)$. Indeed, one possibility is that $i = 0$; then $hist_R(|R|, y)(0) \ne top^k(c)$ (since the k-pds of c containing $hist_R(|R|, y)(0)$ is c-clear), so $0 \notin pre^k_R(j)$, which implies $0 \notin pre^0_R(j)$. Otherwise, as already observed, $0 \notin pre^{k-1}_R(i)$ and $hist_R(|R|, y)(i) = top^k(R(i))$, which implies $i \in pre^k_R(j)$. By Proposition 3.3, applied for 0 (as i), i (as j), and j (as l), we get $0 \notin pre^{k-1}_R(j)$, which implies $0 \notin pre^0_R(j)$.

Let s^k be the k-pds of $R(i)$ containing $hist_R(|R|, y)(i)$ (where i as above). We have two cases. Assume first that there exist at least $b(s^k)$ indices j such that $i \le j \le |R|$ and position $hist_R(|R|, y)(j)$ is in the topmost k-pds of $R(j)$. Then we can use Lemma 5.5 for the subrun of R from i to $|R|$ (and for y). Then for some j, configuration $R(j)$ is a milestone and position $hist_R(|R|, y)(j)$ is in the topmost k-pds of $R(j)$. As both c and $R(j)$ are milestones, we can use Lemma 5.4. We get that $0 \in pre^0_R(j)$. But from the above paragraph we know that

[2] Although we have not defined operations on k-pds's, only on configurations, the definition is natural: the pop^r operation (for $1 \le r \le k$) removes the topmost $(r-1)$-pds; the push^r operation (for $1 \le r \le k$) copies the topmost $(r-1)$-pds and changes the topmost 0-pds.

$0 \notin pre^0_R(j)$; a contradiction. So there exist less than $b(s^k)$ indices j such that $i \leq j \leq |R|$ and position $hist_R(|R|, y)(j)$ is in the topmost k-pds of $R(j)$. Observe that a k-pds can be changed only if it is the topmost k-pds. So the k-pds of $R(|R|)$ containing y can be obtained from the k-pds of $R(i)$ containing $hist_R(|R|, y)(i)$ (i.e. from s^k, which is in $\mathcal{X}$) by applying at most $b(s^k)$ of push and pop operations. Thus it is in $\mathcal{S}^k$. ∎

Corollary 6.4. *Let $\mathcal{A}$ be an n-HOPDA with input alphabet A containing a $\star$ symbol. Let c be a milestone configuration of $\mathcal{A}$, and let $\varphi\colon A^* \to M$ be a morphism into a finite monoid M. Then there exists a finite set $\mathcal{S}$ of configurations having the following property. Let $0 \leq k \leq n$, let R be a run from c, and let $0 \leq r \leq |R|$ be such that the subrun of R from 0 to r reads only stars. Assume that $0 \notin pre^{k-1}_R(|R|)$, but $0 \in pre^{k-1}_R(i)$ for each $i \in pre^k_R(|R|) \cap \{r, r+1, \dots, |R|-1\}$. Then for some configuration $d \in \mathcal{S}$, we have $type^k_{\mathcal{A},\varphi}(R(|R|)) = type^k_{\mathcal{A},\varphi}(d)$.*

Proof: Recall that $type^k_{\mathcal{A},\varphi}(d)$ returns an element of $\mathcal{T}_{\mathcal{A},\varphi}$, and the topmost k-pds of d. We have only finitely elements of $\mathcal{T}_{\mathcal{A},\varphi}$. So it is enough to show, for each k, that there are only finitely many possible topmost k-pds's over all configurations $R(|R|)$ satisfying the assumptions. For $k = 0$ this is trivial as 0-pds contains just one symbol. So let $1 \leq k \leq n$. We have two cases.

First assume that $pre^k_R(|R|) \cap \{r, r+1, \dots, |R|-1\} \neq \emptyset$. Let i be the greatest index in this set. Is it possible that $i < |R|-1$? Then the first case of Proposition 3.1 (used for $|R|$ as j) holds, so we could not have simultaneously $0 \in pre^{k-1}_R(i)$ and $0 \notin pre^{k-1}_R(|R|)$, which are our assumptions. Thus $i = |R| - 1$. As we have $0 \in pre^{k-1}_R(|R| - 1)$ and $0 \notin pre^{k-1}_R(|R|)$ and $|R| - 1 \in pre^k_R(|R|)$, we can use Proposition 3.4. It says that the topmost k-pds of $R(|R|)$ can be obtained from the topmost k-pds of c by removing the topmost $(k-1)$-pds. Thus the content of this k-pds is fixed.

The other case is that $pre^k_R(|R|) \cap \{r, r+1, \dots, |R|-1\} = \emptyset$. This means that for $r \leq i < |R|$, the k-pds of $R(i)$ containing $hist_R(|R|, top^k(R(|R|)))(i)$ is not the topmost k-pds, so it remains unchanged, and is the same as the topmost k-pds of $R(|R|)$. Thus the topmost k-pds of $R(|R|)$ is the same as some k-pds of $R(r)$ (the one containing $hist_R(|R|, top^k(R(|R|)))(r)$). Let x be the bottommost position of the topmost $(k-1)$-pds of this k-pds in $R(r)$. We have $hist_R(|R|, top^{k-1}(R(|R|)))(0) = hist_R(r, x)(0)$. Because $0 \notin pre^{k-1}_R(|R|)$, it follows that $hist_R(r, x)(0) \neq top^{k-1}(R(0))$. So x and the subrun of R from 0 to r satisfy the assumptions of Lemma 6.3. It follows that our k-pds comes from a finite set $\mathcal{S}^k$. ∎

Proof of Theorem 6.2: Consider the infinite run S starting at the milestone configuration c and reading only stars. Consider first the degenerate case when in S only finitely many stars are read. As r we take their number, plus one. Then the thesis is satisfied trivially, as there is no run from c which reads a word beginning with r stars. So for the rest of the proof assume that S reads infinitely many stars.

Let $\mathcal{S}$ be the set from Corollary 6.4 (used for c and φ). For each $i \geq 1$ we define the set $T_i \subseteq \{0, 1, \dots, n\} \times \mathcal{S} \times M$ as follows. A triple (k, d, m) belongs to T_i if and only if there exists a run R from c such that the word read by R begins with (at least) i stars, evaluates to m under φ, and $type^k_{\mathcal{A},\varphi}(R(|R|)) = type^k_{\mathcal{A},\varphi}(d)$. By definition $T_i \subseteq T_{i+1}$ (for each i), and there are only finitely many possible sets, so from some moment every T_i is the same. As the required number of stars (in the statement of the pumping lemma) we take such $r > 0$ that $T_i = T_r$ for each $i \geq r$.

Consider now any number r' (we may assume that $r' \geq r$) and any run R from c which reads a word beginning with at least r stars. Our goal is to show a pumping witness R' for R such that the word read by R' begins with at least r' stars. Let k be the greatest number ($0 \leq k \leq n$) such that $0 \notin pre^{k-1}_R(|R|)$. Such k exists, as $k = 0$ is always good (recall that by definition $pre^{-1}_R(|R|) = \{|R|\}$ and $|R| > 0$). Let i be an index such that the subrun of R from 0 to i reads exactly r stars. Let $j \geq i$ be the smallest index such that $j \in pre^k_R(|R|)$, and $0 \notin pre^{k-1}_R(j)$. Such j exists, as $j = |R|$ is always good.

We use Corollary 6.4 for k, for the subrun of R from 0 to j (as R), and for i (as r). Its assumptions are satisfied by minimality of j. So we get that $type^k_{\mathcal{A},\varphi}(R(j)) = type^k_{\mathcal{A},\varphi}(d)$, for some $d \in \mathcal{S}$. It means that $(k, d, m) \in T_r$, where m is the image under φ of the word read by the subrun of R from 0 to j. Because $T_r = T_{r'}$, there exists a run U from c such that the word read by U begins with (at least) r' stars, evaluates to m under φ, and $type^k_{\mathcal{A},\varphi}(U(|U|)) = type^k_{\mathcal{A},\varphi}(R(j))$.

Finally, we use Theorem 4.1 for the subrun of R from j to $|R|$ (as R), and for $U(|U|)$ (as c). We have assumed that $j \in pre^k_R(|R|)$, which is the assumption of the theorem. We obtain a run U' from $U(|U|)$ such that $type^k_{\mathcal{A},\varphi}(R(|R|)) \leq type^k_{\mathcal{A},\varphi}(U'(|U'|))$, and that the words read by the subrun of R from j to R and by U' evaluate to the same.

Notice that the composition R' of U and U' is a pumping witness for R, and the word read by R' begins with at least r' stars. Indeed, the words read by R and by R' evaluate to the same under φ. For $0 \leq l \leq k$, we have $type^l_{\mathcal{A},\varphi}(R(|R|)) \leq type^l_{\mathcal{A},\varphi}(R'(|R'|))$. For $k < l \leq n$, by the maximality of k, we have $0 \in pre^{l-1}_R(|R|)$. ∎

VII. The separating language

In this section we define a language U which can be recognized by a 2-CPDA, but not by any n-HOPDA, for any n. It is a language over the alphabet $A = \{[,], \star, \sharp\}$. For a word $w \in \{[,], \star\}^*$ we define $stars(w)$. Whenever in some prefix of w there are more closing brackets than opening brackets, $stars(w) = 0$. Also when in the whole w we have the same number of opening and closing brackets, $stars(w) = 0$. Otherwise, let $stars(w)$ be the number of stars in w before the last opening bracket which is not closed. Let U be the set of words $w\sharp^{stars(w)+1}$, for any $w \in \{[,], \star\}^*$ (i.e. these are words w consisting of brackets and stars, followed by $stars(w) + 1$ sharp symbols).

It is known that languages similar to U can be recognized by a 2-CPDA (e.g. [11]), but for completeness we show it below. The collapsible 2-HOPDA will use three stack symbols: X (used to mark the bottom of 1-pds's), Y (used to count brackets), Z (used to mark the bottommost1-pds). The initial symbol is X. The automaton first pushes Z, makes a copy of the 1-pds (i.e. push^2), and pops Z (hence the first 1-pds is marked with Z, unlike any other 1-pds used later). Then, for an opening bracket we push Y, for a closing bracket we pop Y, and for a star we make push^2. Hence for each star we have an 1-pds and on the last 1-pds we have as many Y symbols as the number of currently open brackets. If for a closing bracket the topmost symbol is X, it means that in the word read so far we have more closing brackets than opening brackets; in this case we should accept suffixes of the form $\{[,],*\}^*\sharp$, which is easy.

Finally the $\sharp$ symbol is read. If the topmost symbol is X, we have read as many opening brackets as closing brackets, hence we should accept one $\sharp$ symbol. Otherwise, the topmost Y symbol corresponds to the last opening bracket which is not closed. We do the collapse operation. It leaves the 1-pds's created by the stars read before this bracket, except one (plus the first 1-pds). Thus the number of 1-pds's is precisely equal to $stars(w)$. Now we should read as many $\sharp$ symbols as we have 1-pds's, plus one (after each $\sharp$ symbol we make pop^2), and then accept.

VIII. Why U cannot be recognized?

In this section we prove that language U cannot be recognized by a deterministic higher order pushdown automaton of any level. Assume oppositely, that for some n we have an $(n-1)$-HOPDA recognizing U. We construct an n-HOPDA $\mathcal{A}$ which works as follows. First it makes a push^n operation. Then it simulates the $(n-1)$-HOPDA (not using the push^n and pop^n operations). When the $(n-1)$-HOPDA is going to accept, $\mathcal{A}$ makes the pop^n operation and afterwards accepts. Clearly, $\mathcal{A}$ recognizes U as well.

Fix a morphism $\lambda\colon A^*\to M$ into a finite monoid M, which checks if a word is of the form $\sharp^*$ (some number of the $\sharp$ symbols), or of the form $\star^*]\star^*$ (a closing bracket surrounded by some number of stars), or of none of these two forms. This means that for words u,v being of different form we have $\lambda(u)\neq\lambda(v)$. Let N be the number of equivalence classes of the $(\mathcal{A},\lambda)$-sequence equivalence relation, times $|\mathcal{T}_{\mathcal{A},\lambda}|$, plus one. Consider the following words:

$$w_0=[$$
$$w_{k+1}=w_k^N]^N[\qquad\text{for } 0\leq k\leq n-1,$$

where the number in the superscript (in this case N) denotes the number of repetitions of a word. For a word w, its *pattern* is a word obtained from w by removing its letters other than brackets (leaving only brackets). Fix a morphism $\varphi\colon A^*\to M$ such that from its value $\varphi(w)$ we can deduce

- if word w contains the $\sharp$ symbol, and
- if the pattern of w is longer than $|w_n|$, and
- the exact value of the pattern of w, assuming that the pattern is not longer than $|w_n|$.

We fix a run R, and an index $z(w)$ for each prefix w of w_n, such that the following holds. Run R begins in the initial configuration. Between $R(0)$ and $R(z(\varepsilon))$ only stars are read. For each prefix w of w_n, configuration $R(z(w))$ is a milestone. Just after $z(w)$ run R reads r stars, where r is the constant from Theorem 6.2 used for $\mathcal{A}$, φ, and $R(z(w))$ (as c). If $w=va$ (where a is a single letter), the word read by R between $R(z(w))$ and $R(z(wa))$ consists of a surrounded by some number of stars. Of course such run R exists: we read stars until we reach a milestone (succeeds thanks to Lemma 5.3), then we read as many stars as required by the pumping lemma, then we read next letter of w_n, and so on (because $\mathcal{A}$ accepts U, it will never block).

By construction of $\mathcal{A}$, for every prefix v of w_n we have $z(v)\in pre_R^{n-1}(z(w_n))$ (as we never make a pop^n operation before reading $\sharp$). This contradicts with the following key lemma (taken for $k=n-1$ and $u=\varepsilon$).

Lemma 8.1. *Let $-1\leq k\leq n-1$, and let u be a word such that uw_{k+1} is a prefix of w_n. Then there exist a prefix v of w_{k+1} such that $z(uv)\notin pre_R^k(z(uw_{k+1}))$.*

Proof: The proof is by induction on k. For $k=-1$ this is obvious, we simply take ε as v. By definition $pre_R^{-1}(z(u[))=\{z(u[)\}$, and $z(u)\neq z(u[)$.

Let now $k\geq 0$. Assume that the thesis of the lemma does not hold. Then for each prefix v of w_{k+1} we have $z(uv)\in pre_R^k(z(uw_{k+1}))$. From this we get the following property $\heartsuit$.

Let v' be a prefix of w_{k+1}, and v a prefix of v'. Then $z(uv)\in pre_R^k(z(uv'))$.

From the induction assumption (where uw_k^{i-1} is taken as u), for each $1\leq i\leq N$ there exist a prefix v_i of w_k such that $z(uw_k^{i-1}v_i)\notin pre_R^{k-1}(z(uw_k^i))$. We use Proposition 3.3 for k, R, and $z(uw_k^{i-1}v_i)$ (as i), $z(uw_k^i)$ (as j), $z(uw_k^N)$ (as l). Its second assumption follows from $\heartsuit$. It follows that $z(uw_k^{i-1}v_i)\notin pre_R^{k-1}(z(uw_k^N))$.

Now we are ready to use the pumping lemma (Theorem 6.2). For each $1\leq i\leq N$ we use it for the subrun of R from $z(uw_k^{i-1}v_i)$ to $z(uw_k^N)$. Recall from the definition of R that the word read by this subrun begins with such number of stars that the pumping lemma can be used. So this subrun can be pumped. For each number l we obtain a pumping witness $S_{i,l}$ which reads a word beginning with at least l stars; let $d_{i,l}=S_{i,l}(|S_{i,l}|)$. From the definition of a pumping witness (Definition 6.1), we have a run from the initial configuration to $d_{i,l}$ (namely, the subrun of R from 0 to $z(uw_k^{i-1}v_i)$ composed with $S_{i,l}$) which reads a word having pattern uw_w^N. Moreover, because $z(uw_k^{i-1}v_i)\notin pre_R^{k-1}(z(uw_k^N))$, we get that $type^k_{\mathcal{A},\varphi}(R(z(uw_k^N)))\leq type^k_{\mathcal{A},\varphi}(d_{i,l})$.

Because there are only finitely many possible values of $type_{\mathcal{A},\lambda}$, we can assume that $type_{\mathcal{A},\lambda}(d_{i,l})=type_{\mathcal{A},\lambda}(d_{i,j})$ for $1\leq i\leq N$ and each l and j. Indeed, we can choose (for each i separately) some value of $type_{\mathcal{A},\lambda}(d_{i,l})$ which appears infinitely often, and then we take the subsequence of only these $d_{i,l}$ which give this value.

Since there are more possible indices $i\in\{1,2,\dots,N\}$ than the number of classes of the $(\mathcal{A},\lambda)$-sequence equivalence relation, times $|\mathcal{T}_{\mathcal{A},\lambda}|$, there have to exist two indices $1 \leq x < y \leq N$ such that $type_{\mathcal{A},\lambda}(d_{x,1}) = type_{\mathcal{A},\lambda}(d_{y,1})$, and the sequences $d_{x,1}, d_{x,2}, d_{x,3}, \dots$ and $d_{y,1}, d_{y,2}, d_{y,3}, \dots$ are $(\mathcal{A}, \lambda)$-sequence equivalent. From now we fix these two indices x, y. Furthermore, because $type^k_{\mathcal{A},\varphi}(R(z(uw_k^N))) \leq type^k_{\mathcal{A},\varphi}(d_{i,l})$ for each $1 \leq i \leq N$ and each l, we know that the topmost k-pds's of all $d_{x,l}$ and $d_{y,l}$ are the same. Thus $type^k_{\mathcal{A},\lambda}(d_{x,l}) = type^k_{\mathcal{A},\lambda}(d_{y,j})$ for each l and j.

Let $r = N - x + 1$. We will construct a run R' from $d_{x,1}$, and indices $0 = l_0 \leq l_1 \leq \dots \leq l_r = |R'|$ such that for $1 \leq i \leq r-1$ the word read by R' between l_{i-1} and l_i is of the form $\star^*]\star^*$ (a closing bracket surrounded by some number of stars), and $l_{i-1} \in pre^k_{R'}(l_i)$, and R' between l_{r-1} and l_r reads only $\sharp$ symbols, and the last operation of R' is pop^n. Additionally, we will have $type^k_{\mathcal{A},\varphi}(R(z(uw_k^N]^i))) \leq type^k_{\mathcal{A},\varphi}(R'(l_i))$ for $0 \leq i \leq r-1$. Indeed, as $R'(0)$ we take $d_{x,1}$; we know that $type^k_{\mathcal{A},\varphi}(R(z(uw_k^N))) \leq type^k_{\mathcal{A},\varphi}(d_{x,1})$ Then consecutively for $1 \leq i \leq r-1$ we use Theorem 4.1 for $\mathcal{A}$, φ, the subrun of R from $z(uw_k^N]^{i-1})$ to $z(uw_k^N]^i)$ (as R), and $R'(l_{i-1})$ (as c). From $\heartsuit$ we know that $z(uw_k^N]^{i-1}) \in pre^k_R(z(uw_k^N]^i)$. We obtain a next fragment of R' (from l_{i-1} to some l_i), satisfying the above. Finally, from $R'(l_{r-1})$ we start reading $\sharp$ symbols until we reach an accepting configuration; this gives us the last fragment of R', from l_{r-1} to some l_r. Because $\mathcal{A}$ recognizes U, it will finally reach such configuration. By construction of $\mathcal{A}$, the last operation of R' has to be pop^n. So we obtain a run R' as declared.

Finally we use Theorem 4.3 for $\mathcal{A}$, λ (as φ), k, sequences $d_{x,1}, d_{x,2}, d_{x,3}, \dots$ (as $c_1, c_2, c_3, \dots$) and $d_{y,1}, d_{y,2}, d_{y,3}, \dots$ (as $d_1, d_2, d_3, \dots$), for run R' (as R), and for sequence $l_0, l_1, \dots, l_r$. We have $l_{i-1} \in pre^k_{R'}(l_i)$ for $1 \leq i \leq r-1$. By construction of $\mathcal{A}$, no push^n or pop^n operations are done in R', except the last operation in R', which is pop^n. As noticed above (in particular because $R'(0) = d_{x,1}$) we have $type^k_{\mathcal{A},\lambda}(e_0) = type^k_{\mathcal{A},\lambda}(d_{x,l}) = type^k_{\mathcal{A},\lambda}(d_{y,l})$ for each l. Thus the assumptions of the theorem are satisfied. For each l, we obtain runs S_l (from $d_{x,l}$) and T_l (from $d_{y,l}$). From condition 1 of the theorem we see that, for each l, the word read by S_l (and by T_l) contains $r-1 = N-x$ closing brackets with some number of stars around them, and after them some number of the $\sharp$ symbols.

For each l, let x_l and y_l be the number of the $\sharp$ symbols read by S_l and T_l, respectively. The pattern of the words read between the initial configuration and $S_l(|S_l|)$ (by the subrun of R from 0 to $z(uw_k^{x-1}v_x)$ composed with $S_{x,l}$ and composed with S_l), for each l, is $uw_k^N]^{N-x}$; the same for T_l. In this pattern the last opening bracket which is not closed is the last bracket of the x-th w_k. Recall that configurations $d_{x,l}$ were obtained by pumping inside the x-th w_k, so before this bracket; for $l \to \infty$ the number of stars inserted there is unbounded. From the definition of language U it follows that the sequence $x_1, x_2, x_3, \dots$ has to be unbounded. On the other hand, configurations $d_{y,l}$ were obtained by pumping inside the y-th w_k, so after the last opening bracket which was is not closed (as $y > x$). For each l the number of stars before this bracket is the same. From the definition of language U it follows that the sequence $y_1, y_2, y_3, \dots$ has to be constant, hence bounded. This contradicts with condition 2 of Theorem 4.3, which says that either both these sequences are bounded or both unbounded. ■

References

[1] B. Courcelle, "The monadic second-order logic of graphs IX: Machines and their behaviours," *Theor. Comput. Sci.*, vol. 151, no. 1, pp. 125–162, 1995.

[2] A. N. Maslov, "The hierarchy of indexed languages of an arbitrary level," *Soviet Math. Dokl.*, vol. 15, pp. 1170–1174, 1974.

[3] T. Knapik, D. Niwinski, and P. Urzyczyn, "Higher-order pushdown trees are easy," in *FoSSaCS*, ser. Lecture Notes in Computer Science, M. Nielsen and U. Engberg, Eds., vol. 2303. Springer, 2002, pp. 205–222.

[4] D. Caucal, "On infinite terms having a decidable monadic theory," in *MFCS*, ser. Lecture Notes in Computer Science, K. Diks and W. Rytter, Eds., vol. 2420. Springer, 2002, pp. 165–176.

[5] B. Courcelle and T. Knapik, "The evaluation of first-order substitution is monadic second-order compatible," *Theor. Comput. Sci.*, vol. 281, no. 1-2, pp. 177–206, 2002.

[6] T. Knapik, D. Niwinski, P. Urzyczyn, and I. Walukiewicz, "Unsafe grammars and panic automata," in *ICALP*, ser. Lecture Notes in Computer Science, L. Caires, G. F. Italiano, L. Monteiro, C. Palamidessi, and M. Yung, Eds., vol. 3580. Springer, 2005, pp. 1450–1461.

[7] K. Aehlig, J. G. de Miranda, and C.-H. L. Ong, "The monadic second order theory of trees given by arbitrary level-two recursion schemes is decidable," in *TLCA*, ser. Lecture Notes in Computer Science, P. Urzyczyn, Ed., vol. 3461. Springer, 2005, pp. 39–54.

[8] M. Hague, A. S. Murawski, C.-H. L. Ong, and O. Serre, "Collapsible pushdown automata and recursion schemes," in *LICS*. IEEE Computer Society, 2008, pp. 452–461.

[9] C.-H. L. Ong, "On model-checking trees generated by higher-order recursion schemes," in *LICS*. IEEE Computer Society, 2006, pp. 81–90.

[10] N. Kobayashi, "Model-checking higher-order functions," in *PPDP*, A. Porto and F. J. López-Fraguas, Eds. ACM, 2009, pp. 25–36.

[11] K. Aehlig, J. G. de Miranda, and C.-H. L. Ong, "Safety is not a restriction at level 2 for string languages," in *FoSSaCS*, ser. Lecture Notes in Computer Science, V. Sassone, Ed., vol. 3441. Springer, 2005, pp. 490–504.

[12] P. Parys, "Collapse operation increases expressive power of deterministic higher order pushdown automata," in *STACS*, ser. LIPIcs, T. Schwentick and C. Dürr, Eds., vol. 9. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011, pp. 603–614.

[13] P. Parys, "A pumping lemma for pushdown graphs of any level," 2011, accepted to STACS 2012.

[14] A. Blumensath, "On the structure of graphs in the Caucal hierarchy," *Theor. Comput. Sci.*, vol. 400, no. 1-3, pp. 19–45, 2008.

[15] T. Hayashi, "On derivation trees of indexed grammars," *Publ. RIMS, Kyoto Univ.*, vol. 9, pp. 61–92, 1973.

[16] R. H. Gilman, "A shrinking lemma for indexed languages," *Theor. Comput. Sci.*, vol. 163, no. 1&2, pp. 277–281, 1996.

[17] A. Kartzow, "A pumping lemma for collapsible pushdown graphs of level 2," in *CSL*, ser. LIPIcs, M. Bezem, Ed., vol. 12. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011, pp. 322–336.

[18] A. Kartzow and P. Parys, "Strictness of the collapsible pushdown hierarchy," 2012, submited to LICS 2012.

[19] A. Kartzow, "Collapsible pushdown graphs of level 2 are tree-automatic," in *STACS*, ser. LIPIcs, J.-Y. Marion and T. Schwentick, Eds., vol. 5. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2010, pp. 501–512.

Appendix A
Pushdown automata as tree generators

In this section we describe how pushdown automata can be used to generate trees, and we show how Corollary 1.2 follows from Theorem 1.1. We do not describe recursion schemes in this paper; it is known that a tree is generated by a recursion scheme of second level if and only if it is generated by a collapsible pushdown system of second level [8], and that a tree is generated by a safe recursion scheme of any level if and only if it is generated by a higher order pushdown system (without collapse) of the same level [3], so we get the "equivalently" parts of Corollary 1.2.

We consider ranked, potentially infinite trees. Beside of the input alphabet A we have a function $rank\colon A \to \mathbb{N}$; a tree node labelled by some $a \in A$ has always $rank(a)$ children.

Automata used to generate trees are called *higher order pushdown systems*, or *collapsible pushdown systems* if they use the collapse operation. They are defined like HOPDA, with the difference that they do not have the set of accepting states, and that instead of the $\mathsf{read}(f)$ operations, there are $\mathsf{branch}(a, q_1, q_2, \dots, q_{rank(a)})$ operations, for $a \in A$, and for pairwise distinct states $q_1, q_2, \dots, q_{rank(a)} \in Q$. We use a shorthand $\delta(c)$ for a configuration c to denote $\delta(state(c), s^0)$, where s^0 is the topmost symbol of c. If from $\delta(c) = \mathsf{branch}(a, q_1, q_2, \dots, q_{rank(a)})$, in a successor d of c we have $\pi(d) = \pi(c)$ and $state(d) = q_i$ for some $1 \leq i \leq rank(a)$ (in particular c has no successors if $rank(a) = 0$). Let $T(\mathcal{A})$ be the set of all configurations c of $\mathcal{A}$ reachable from the initial one, such that a branch operation should be performed from c. Additionally we require that from each configuration of $\mathcal{A}$ reachable from the initial one, there exists a run to a configuration from $T(\mathcal{A})$ (this assumption is needed to ensure that the generated tree has a proper form).

A *tree generated by a pushdown system* has runs from the initial configuration to a configuration from $T(\mathcal{A})$ as its nodes. A node R is labelled by $a \in A$ such that $\delta(R(|R|)) = \mathsf{branch}(a, q_1, q_2, \dots, q_{rank(a)})$. A node S is its i-th child ($1 \leq i \leq rank(a)$), if S is the composition of R and a run S' which uses a branch operation only in its first transition. Notice that the graph obtained this way is really an A-labelled ranked tree.

Proof of Corollary 1.2: Let $L \subseteq A^*$ be the language recognized by a deterministic collapsible pushdown automaton $\mathcal{A}$ of second level, which is not recognized by any deterministic higher order pushdown automaton (without collapse) of any level (L exists by Theorem 1.1). We can transform $\mathcal{A}$ to a deterministic collapsible pushdown automaton $\mathcal{B}$ of second level, also recognizing L, such that each configuration of $\mathcal{B}$ reachable from the initial one has a successor. Indeed, the only reason why in $\mathcal{A}$ there are configurations with no successors is that it wants to empty a pds using a pop operation. To avoid such situations, $\mathcal{B}$ should have some bottom-of-stack marker on the bottom of each pds of each level (thus at the beginning $\mathcal{B}$ performs a push operation of each level). Then whenever $\mathcal{A}$ blocks because it wants to empty a pds, in $\mathcal{B}$ the bottom-of-stack marker is uncovered; in such situation $\mathcal{B}$ starts some loop with no accepting state. There is also a technical detail, that a pop operation which would block $\mathcal{A}$, in $\mathcal{B}$ can enter an accepting state; to overcome this problem, every pop operation ending in an accepting state should first end in some auxiliary, not accepting state, from which (if the bottom-of-stack marker is not seen) an accepting state is reached.

Next, we create a collapsible pushdown system $\mathcal{C}$ of second level, which generates a tree over alphabet $B = \{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$, where $rank(\mathsf{X}) = |A|$ and $rank(\mathsf{Y}) = rank(\mathsf{Z}) = 1$. It is obtained from $\mathcal{B}$ in two steps. First, we replace every operation $\mathsf{read}(f)$ of $\mathcal{B}$ by operation $\mathsf{branch}(\mathsf{X}, f(a_1), f(a_2), \dots, f(a_{|A|}))$, where $A = \{a_1, a_2, \dots, a_{|A|}\}$. Then, in every operation we replace the resulting state q by its auxiliary copy $\overline{q}$, and from $\overline{q}$ (for any topmost stack symbol) we perform operation $\mathsf{branch}(\mathsf{Y}, q)$ if q was accepting, or operation $\mathsf{branch}(\mathsf{Z}, q)$ if q was not accepting (this way, after every operation of the original automaton, we perform operation $\mathsf{branch}(\mathsf{Y}, \cdot)$ or $\mathsf{branch}(\mathsf{Z}, \cdot)$). Notice that from each configuration of $\mathcal{C}$ reachable from the initial one, there exists a run to a configuration from $T(\mathcal{C})$, as required by the definition of a (collapsible) pushdown system. Let $t_{\mathcal{C}}$ be the tree generated by $\mathcal{C}$.

Finally, assume that $t_{\mathcal{C}}$ can be generated also by some higher order pushdown system $\mathcal{D}$ (without collapse). From $\mathcal{D}$ we create a deterministic higher order pushdown automaton $\mathcal{E}$ (without collapse) of the same level. We replace every operation $\mathsf{branch}(\mathsf{X}, q_1, q_2, \dots, q_{|A|})$ of $\mathcal{D}$ by operation $\mathsf{read}(f)$, where $f(a_i) = q_i$. We replace every operation $\mathsf{branch}(\mathsf{Y}, q)$ of $\mathcal{D}$ by operation $\mathsf{push}^1(s^0, p)$ for a fresh accepting state p and some stack symbol s^0; from (p, s^0) we perform operation $\mathsf{pop}^1(q)$ (thus we replace $\mathsf{branch}(\mathsf{Y}, q)$ by a pass through an accepting state). The same for a $\mathsf{branch}(\mathsf{Z}, q)$ operation, but the fresh state p is not accepting.

Notice that $\mathcal{E}$ recognizes L; this contradicts our assumptions about L, so $t_{\mathcal{C}}$ is not generated by any higher order pushdown system (without collapse). Indeed, take any word $w \in L$. We have an accepting run of $\mathcal{B}$ which reads w. This run corresponds to a run of $\mathcal{C}$, so to a path p in $t_{\mathcal{C}}$, from the root to a Y-labelled node. Letters of w say which child p chooses in X-labelled nodes: if i-th letter of w is a_j, then from the i-th X-labelled node of p, the path continues to the j-th child. This path p corresponds also to a run of $\mathcal{D}$, so to a run of $\mathcal{E}$. This run ends in an accepting state, and reads w; thus $\mathcal{E}$ accepts w. Similarly, every word accepted by $\mathcal{E}$ is also accepted by $\mathcal{B}$. ∎

APPENDIX B
TYPES AND SEQUENCE EQUIVALENCE

In this appendix we define objects used in the statements of Theorems 4.1 and 4.3, and we prove these theorems.

For the whole Appendix B fix an n-HOPDA $\mathcal{A}$. Let A be its input alphabet, Γ its stack alphabet, and Q its set of states. Assume A contains a distinguished symbol $\sharp$. Fix also a morphism $\varphi\colon A^* \to M$ into a finite monoid M.

A. Returns

It is important to distinguish special runs, called returns.

Definition B.1. A run R is called an *r-return* (where $1 \leq r \leq n$) if

- the topmost r-pds of $R(0)$ contains at least two $(r-1)$-pds's, and
- $hist_R(|R|, top^{r-1}(R(|R|)))(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$, and
- $pre_R^{r-1}(|R|) = \{|R|\}$.

In other words, R is an r-return when the topmost r-pds of $R(|R|)$ is obtained from the topmost r-pds of $R(0)$ by removing its topmost $(r-1)$-pds (but not only in the sense of contents, but we require that really it was obtained this way). In particular we have the following proposition.

Proposition B.2. *Let R be an r-return. Then the topmost r-pds of $R(0)$ after removing its topmost $(r-1)$-pds is equal to the topmost r-pds of $R(|R|)$. Additionally, for every position y in the topmost r-pds of $R(|R|)$, $hist_R(|R|, y)(0)$ is the corresponding position in the topmost r-pds of $R(0)$.*

Proof: For each index $i \in pre_R^r(|R|)$ let s_i be the size of the topmost r-pds. Choose the index $i \in pre_R^r(|R|)$, $i < |R|$, for which s_i is smallest. Because $i \notin pre_R^{r-1}(|R|)$, by Proposition 3.2 (used for this i, and for $|R|$ as j), we know that $s_i \geq s_{|R|}+1$. Thus also $s_i \geq s_{|R|}+1$ for each index $i \in pre_R^r(|R|)$, $i < |R|$. For all indices $i \leq j$ such that $pre_R^r(|R|) \cap \{i, i+1, \dots, j\} = \{i, j\}$, by Proposition 3.1 we know that the first $s_{|R|}$ $(r-1)$-pds's of the topmost r-pds of $R(j)$ and of $R(i)$ are the same, and that for every position y in such $(r-1)$-pds of $R(j)$, $hist_R(j, y)(i)$ is the corresponding position in the topmost r-pds of $R(i)$. From the second point in the definition of an r-return it follows that $0 \in pre_R^r(|R|)$. Composing this property for each such pair i, j, we get that the topmost r-pds of $R(|R|)$ is equal to the first $s_{|R|}$ $(r-1)$-pds's of the topmost r-pds of $R(0)$, and that for every position y of the topmost r-pds $R(|R|)$, $hist_R(|R|, y)(0)$ is the corresponding position in the topmost r-pds of $R(0)$. By the second point of the definition of an r-return, we get that $s_0 = s_{|R|}+1$, which gives us the thesis. ■

Example: Consider a PDS of level 2, and a run R of length 6 in which $\pi(R(0)) = [[ab][cd]]$, and the operations between consecutive configurations are:

$$\mathsf{push}^2(e),\ \mathsf{pop}^1,\ \mathsf{pop}^2,\ \mathsf{pop}^1,\ \mathsf{push}^1(d),\ \mathsf{pop}^1.$$

The contents of the 2-pds's of the configurations in the run are presented in the table below.

i	0	1	2	3	4	5	6
$\pi(R(i))$	$[[ab][cd]]$	$[[ab][cd][ce]]$	$[[ab][cd][c]]$	$[[ab][cd]]$	$[[ab][c]]$	$[[ab][cd]]$	$[[ab][c]]$

The subruns of R from 0 to 2, from 0 to 4, from 1 to 2, from 3 to 4, and from 5 to 6 are 1-returns; the subruns of R from 1 to 3, and from 2 to 3 are 2-returns. These are the only subruns of R being returns, in particular R is not an 1-return because $4 \in pre_R^0(6)$.

The lemma below shows possible forms of runs described by Theorems 4.1 and 4.3 (i.e. such that $0 \in pre_R^k(|R|)$).

Lemma B.3. *Let $0 \leq k \leq n$, and let R be a run. Then we have $0 \in pre_R^k(|R|)$ if and only if*

1) *$|R| = 0$, or*
2) *$|R| = 1$, and the operation performed in R is* read, *or* push^r *for any r, or* pop^r *for $r \leq k$, or*
3) *R is a one-step run performing a* push^r *operation composed with an r-return, where $r \geq k+1$, or*
4) *R is a composition of runs S and T such that $0 \in pre_S^k(|S|)$ and $0 \in pre_T^k(|T|)$, and $|S| \neq 0 \neq |T|$.*

Proof: Concentrate first on the left-to-right implication. If $|R| = 0$, we have case 1; assume that $|R| \geq 1$. Notice that the first operation, between $R(0)$ and $R(1)$, cannot be pop^r for $r \geq k+1$, as such operation removes the topmost k-pds of $R(0)$, which contradicts with the assumption that $0 \in pre_R^k(|R|)$. Thus, if $|R| = 1$, we have case 2; assume that $|R| \geq 2$. If the first operation is read or pop^r for $r \leq k$, or push^r for $r \leq k$, then $0 \in pre_R^k(|R|)$ implies that $1 \in pre_R^k(|R|)$; we have case 4 (for subruns from 0 to 1 and from 1 to $|R|$). We can do the same if the operation is push^r for $r \geq k+1$ and $1 \in pre_R^k(|R|)$.

The remaining case is when the first operation is push^r for $r \geq k+1$ and $1 \notin pre_R^k(|R|)$. Notice that $hist_R(1, y)(0) = top^k(R(0))$ holds only for $y = top^k(R(1))$ and $y = top^k(R(0))$. So, because $0 \in pre_R^k(|R|)$ (which by definition means $hist_R(|R|, top^k(R(|R|)))(0) = top^k(R(0))$) and $1 \notin pre_R^k(|R|)$, it has to be $hist_R(|R|, top^k(R(|R|)))(1) = top^k(R(0))$. Thus also $hist_R(|R|, top^{r-1}(R(|R|)))(1) = top^{r-1}(R(0))$, which is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(1)$. Let $j \geq 1$ be the smallest positive index for which $j \in pre_R^{r-1}(|R|)$; such j exists since $|R|$ always can be taken as j. Then $hist_R(j, top^{r-1}(R(j)))(1) = top^{r-1}(R(0))$ and $pre_R^{r-1}(j) = \{0, j\}$, thus the subrun of R from 1 to j is an r-return. If $j = |R|$, we have case 3; assume that $j < |R|$. By Proposition B.2, the only position y in the topmost $(r-1)$-pds (even in the topmost r-pds) of $R(j)$ for which $hist_R(j, y)(0) = top^{r-1}(R(0))$ is $y = top^{r-1}(R(j))$. Thus $0 \in pre_R^k(|R|)$ and $j \in pre_R^{r-1}(|R|)$ implies that $j \in pre_R^k(|R|)$; we have case 4 (for subruns from 0 to j and from j to $|R|$).

The right-to-left implication is almost immediate; in case 3 we use Proposition B.2. ■

The next lemma shows possible forms of returns.

Lemma B.4. *Let $1 \leq r \leq n$. A run R is an r-return if and only if*

1) *R is a one-step run performing a* read *operation composed with an r-return, or*
2a) *R is a one-step run performing a* pop^r *operation, or*
2b) *R is a one-step run performing a* pop^k *operation composed with an r-return, where $k \leq r-1$, or*
3a) *R is a one-step run performing a* push^k *operation composed with an r-return, where $k \neq r$, or*
3b) *R is a one-step run performing a* push^k *operation composed with a k-return composed with an r-return, where $k \geq r$.*

Proof: Concentrate first on the left-to-right implication. Of course $|R| \geq 1$. Notice that the first operation, between $R(0)$ and $R(1)$, cannot be pop^k for $k \geq r+1$, as such operation removes the topmost r-pds of $R(0)$, which contradicts with the assumption that $hist_R(|R|, top^{r-1}(R(|R|))(0)$ is in the topmost r-pds of $R(0)$.

Assume that the first operation is read, or pop^k for $k \leq r-1$, or push^k for $k \leq r-1$. Then $hist_R(1, y)(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$ only if y is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(1)$. Thus, because R is an r-return, $hist_R(|R|, top^{r-1}(R(|R|)))(1)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(1)$, so the subrun of R from 1 to $|R|$ is an r-return; we have case 1 or 2b or 3a.

Next, assume that the first operation of R is pop^r. Then $hist_R(1, y)(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$ only if $y = top^{r-1}(R(1))$. Thus, because R is an r-return, $hist_R(|R|, top^{r-1}(R(|R|)))(1) = top^{r-1}(R(1))$, which means $1 \in pre_R^{r-1}(|R|) = \{|R|\}$. So $|R| = 1$; we have case 2a.

Finally, assume that the first operation of R is push^k for $k \geq r$. Let s be the size of the topmost k-pds of $R(0)$. For each i from 1 to $|R|$ we look at the size of the k-pds containing $hist_R(|R|, top^{r-1}(R(|R|)))(i)$. Recall that $hist_R(|R|, top^{r-1}(R(|R|)))(0)$ is in the topmost k-pds of $R(0)$, so for $i = 1$ this is also the topmost k-pds and its size is $s+1$. Assume first that this size is at least $s+1$ for each i. Then $1 \in pre_R^{k-1}(|R|)$ (Proposition 3.2). Because R is an r-return, we know that $1 \notin pre_R^{r-1}(|R|)$ (of course $|R| \neq 1$), so $k \neq r$ ($k > r$). As $1 \in pre_R^{k-1}(|R|)$, we know that $hist_R(|R|, top^{r-1}(R(|R|)))(1)$ is in the topmost $(k-1)$-pds of $R(1)$, so it is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(1)$. It follows that the subrun of R from 1 to $|R|$ is an r-return, and $k \neq r$ (case 3a). The opposite possibility is that for some i ($1 \leq i \leq |R|$), the size of the k-pds containing $hist_R(|R|, top^{r-1}(R(|R|)))(i)$ becomes s. Fix the first such i. Then $pre_R^{k-1}(i) = \{0, i\}$ (Proposition 3.2), and $hist_R(i, top^{k-1}(R(i)))(1)$ is the bottommost position of the $(k-1)$-pds just below the topmost $(k-1)$-pds of $R(1)$, thus the subrun of R from 1 to i is a k-return. By Proposition B.2, if $hist_R(i, y)(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$, and if y is in the topmost k-pds, then it y has to be the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(i)$. Because the size of the k-pds containing $hist_R(|R|, top^{r-1}(R(|R|)))(i)$ has changed size between $R(i-1)$ and $R(i)$, it has to be the topmost k-pds, so necessarily $hist_R(|R|, top^{r-1}(R(|R|)))(i)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(i)$. Thus the subrun of R from i to $|R|$ is an r-return.

Next we concentrate on the right-to-left implication. Case 2a is trivial. In cases 1, 2b, 3a we observe that for y being the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(1)$, also $hist_R(1, y)(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$ (it is important that $k \neq r$ in case 3a). Thus $hist_R(|R|, top^{r-1}(R(|R|)))(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$. Then of course $0 \notin pre_R^{r-1}(|R|)$. So R is an r-return. In case 3c, let i be the length of the first return, plus one (so the k-return ends in $R(i)$). Recall that $k \geq r$. By Proposition B.2, for y being the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(i)$, also $hist_R(i, y)(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$. Thus $hist_R(|R|, top^{r-1}(R(|R|)))(0)$ is the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(0)$. Then of course $0 \notin pre_R^{r-1}(|R|)$. If $j \in pre_R^{r-1}(|R|)$ for some $1 \leq j \leq i$, we have $hist_R(i, y)(j) = top^{r-1}(R(j))$ for y being the bottommost position of the $(r-1)$-pds just below the topmost $(r-1)$-pds of $R(i)$. This implies that $j \in pre_R^{k-1}(i)$ (both for $k > r$ and $k = r$), which is impossible, as the subrun of R from 1 to i is a k-return. We get that R is an r-return. ■

B. Types of pds's

We are going to define a type of a k-pds for each $0 \leq k \leq n$. A set of possible level k types (types of k-pds's) will be denoted $\mathcal{T}^k$.

Definition B.5. We define $\mathcal{T}^k$ by induction on k, going down from $k = n$ to $k = 0$. Let $0 \le k \le n$. Assume we have already defined sets $\mathcal{T}^i$ for $k+1 \le i \le n$. We take

$$\mathcal{T}^k = \{(\mathsf{ne},\mathsf{tr})\} \cup \big(\mathcal{P}(\mathcal{T}^n) \times \mathcal{P}(\mathcal{T}^{n-1}) \cdots \times \mathcal{P}(\mathcal{T}^{k+1}) \times Q \times \mathcal{D}^k \times \{\mathsf{tr},\mathsf{nt}\}\big), \quad \text{where}$$
$$\mathcal{D}^k = \bigcup_{r=k+1}^{n} M \times \{r\} \times \mathcal{P}(\mathcal{T}^n) \times \mathcal{P}(\mathcal{T}^{n-1}) \times \cdots \times \mathcal{P}(\mathcal{T}^{r+1}) \times Q,$$

and by $\mathcal{P}(X)$ we denote the power set of X (the set of all subsets of X). In particular $\mathcal{D}^n = \emptyset$ and $\mathcal{T}^n = \{(\mathsf{ne},\mathsf{tr})\}$. By $\mathcal{T}_{\mathsf{tr}}$ and $\mathcal{T}_{\mathsf{nt}}$ we denote the subset of $\bigcup_{0\le k\le n} \mathcal{T}^k$ having respectively tr or nt on the last coordinate.

Before we define types, we present their intended meaning. This is described by the following definition and lemma, which connect types with existence of some returns.

Definition B.6. Let $0 \le l \le n$, let $\widehat{\sigma} = (m, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q) \in \mathcal{T}^l$, and let R be a run. Decompose $\pi(R(|R|)) = t^n : t^{n-1} : \cdots : t^r$. We say that run R *agrees with* $\widehat{\sigma}$ if

- the word read by R evaluates to m under φ, and
- R is an r-return, and
- $\xi^i \subseteq \mathit{type}(t^i)$ for $r+1 \le i \le n$, and
- $q = \mathit{state}(R(|R|))$.

Lemma B.7. *Let $0 \le l \le n$ and let $\widehat{\rho} = (m, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q) \in \mathcal{D}^l$. Assume there exists a run from a configuration c which agrees with $\widehat{\rho}$. Decompose $\pi(c) = s^n : s^{n-1} : \cdots : s^l$. Then in $\mathit{type}(s^l)$ we have a tuple $(\eta^n, \eta^{n-1}, \dots, \eta^{l+1}, \mathit{state}(c), \widehat{\rho}, \cdot)$ such that $\eta^i \subseteq \mathit{type}(s^i)$ for $l+1 \le i \le n$.*

We will also have implication in the opposite direction, saying that having a tuple in type implies existence of appropriate run. However we will be talking not only about the existence of a run, but also about the number of the $\sharp$ characters read. Then the last coordinate (tr/nt) of types will be used (notice that it is not used in the above definition and lemma, including its proof). Its intended meaning is, roughly speaking, that using a tuple with nt increases the number of the $\sharp$ characters read, and a tuple with tr does not change it.

Now we come to the definition of types. We first define a composer, which is then (in Definition B.10) used to compose types of smaller pds's into types of greater pds's.

Definition B.8. We define when $(\beta^k, \beta^{k-1} \dots, \beta^l; \alpha^k; \mathit{fl})$ is a *composer*, where $0 \le l < k \le n$, $\beta^i \subseteq \mathcal{T}^i$ for $l \le i \le k$, $\alpha^k \subseteq \mathcal{T}^k$, and $\mathit{fl} \in \{\mathsf{tr},\mathsf{nt}\}$.

1) Let $\sigma^l = (\gamma^n, \gamma^{n-1}, \dots, \gamma^{l+1}, p, \widehat{\sigma}, \cdot) \in \mathcal{T}^l$, and let $\sigma^k = (\gamma^n, \gamma^{n-1}, \dots, \gamma^{k+1}, p, \widehat{\sigma}, \mathit{fl}) \in \mathcal{T}^k$ (in particular $\widehat{\sigma} \in \mathcal{D}^k$), where $\mathit{fl} = \mathsf{tr}$ if and only if $\{\sigma^l\} \cup \gamma^{l+1} \cup \gamma^{l+2} \cup \cdots \cup \gamma^k \subseteq \mathcal{T}_{\mathsf{tr}}$. Then we say that $(\gamma^k, \gamma^{k-1}, \dots, \gamma^{l+1}, \{\sigma^l\}; \{\sigma^k\}; \mathsf{tr})$ is a composer.
2) Tuple $(\emptyset, \emptyset, \dots, \emptyset; \{(\mathsf{ne};\mathsf{tr})\}; \mathsf{tr})$ is a composer.
3) Let $\alpha^k \subseteq \mathcal{T}^k$; Assume for each $\sigma \in \alpha^k$ we have a composer $(\beta^k_\sigma, \beta^{k-1}_\sigma, \dots, \beta^l_\sigma; \{\sigma\}; \mathsf{tr})$. Let $\beta^i = \bigcup_{\sigma\in\alpha^k} \beta^i_\sigma$. Then we say that $(\beta^k, \beta^{k-1}, \dots, \beta^l; \alpha^k; \mathit{fl})$ is a composer, where $\mathit{fl} = \mathsf{tr}$ if and only if $\beta^i_\sigma \cap \beta^i_\tau \subseteq \mathcal{T}_{\mathsf{tr}}$ for each $\sigma, \tau \in \alpha^k$, $\sigma \ne \tau$, $l \le i \le k$.

Now we will define the set of types of a 0-pds as a fixpoint. First, for each $z \in \mathbb{N}$ and each 0-pds s^0 we define $\mathit{type}_z(s^0) \subseteq \mathcal{T}^0$. The cases in the definition below correspond to the cases of Lemma B.4.

Definition B.9. Let $z \in \mathbb{N}$ and let $s^0 \in \Gamma$. If $z = 0$ we put $\mathit{type}_z(s^0) = \emptyset$. For $z > 0$ we proceed by induction: assume that type_{z-1} are already defined. We define $\mathit{type}_z(s^0)$ as the smallest set satisfying the following conditions. Let p be a state.

1) Assume that $\delta(s^0, p) = \mathsf{read}(f)$ and in $\mathit{type}_{z-1}(s^0)$ we have a tuple
$$\sigma = \big(\alpha^n, \alpha^{n-1}, \dots, \alpha^1, f(a), (m, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q_2), \cdot\big)$$
for some $a \in A$. Then $\big(\alpha^n, \alpha^{n-1}, \dots, \alpha^1, p, (\varphi(a)m, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q_2), \mathit{fl}\big) \in \mathit{type}_z(s^0)$, where $\mathit{fl} = \mathsf{tr}$ if and only if $\sigma \in \mathcal{T}_{\mathsf{tr}}$ and $a \ne \sharp$.
2) Assume that $\delta(s^0, p) = \mathsf{pop}^k(q_1)$ (where $1 \le k \le n$). We have two subcases:
   a) Let $\alpha^i \subseteq \mathcal{T}^i$ for $k+1 \le i \le n$. Then
   $$\big(\alpha^n, \alpha^{n-1}, \dots, \alpha^{k+1}, \{(\mathsf{ne},\mathsf{tr})\}, \emptyset, \emptyset, \dots, \emptyset, p, (\varphi(\varepsilon), k, \alpha^n, \alpha^{n-1}, \dots, \alpha^{k+1}, q_1), \mathsf{tr}\big) \in \mathit{type}_z(s^0).$$
   b) Let $\sigma^k = (\alpha^n, \alpha^{n-1}, \dots, \alpha^{k+1}, q_1, \widehat{\sigma}, \cdot) \in \mathcal{T}^k$. Then
   $$(\alpha^n, \alpha^{n-1}, \dots, \alpha^{k+1}, \{\sigma^k\}, \emptyset, \emptyset, \dots, \emptyset, p, \widehat{\sigma}, \mathsf{tr}) \in \mathit{type}_z(s^0).$$

3) Assume that $\delta(s^0, p) = \mathsf{push}^k(t^0, q_1)$ (where $1 \leq k \leq n$) and in $type_{z-1}(t^0)$ we have a tuple

$$\sigma = (\alpha^n, \alpha^{n-1}, \dots, \alpha^1, q_1, \widehat{\sigma}, \cdot), \quad \text{where } \widehat{\sigma} = (m_1, r_1, \zeta^n, \zeta^{n-1}, \dots, \zeta^{r_1+1}, q_2).$$

Assume also that $(\beta^k, \beta^{k-1}, \dots, \beta^0; \alpha^k; fl)$ is a composer such that $\beta^0 \subseteq type_{z-1}(s^0)$. We have two subcases:

a) Assume that $r_1 \neq k$. Define

$$\eta^i = \begin{cases} \alpha^i & \text{for } k+1 \leq i \leq n, \\ \beta^i & \text{for } i = k, \\ \alpha^i \cup \beta^i & \text{for } 1 \leq i \leq k-1. \end{cases}$$

Then $(\eta^n, \eta^{n-1}, \dots, \eta^1, p, \widehat{\sigma}, fl') \in type_z(s^0)$, where $fl' = \mathsf{tr}$ if and only if $fl = \mathsf{tr}$ and $\beta^0 \cup \{\sigma\} \subseteq \mathcal{T}_{\mathsf{tr}}$ and $\alpha^i \cap \beta^i \subseteq \mathcal{T}_{\mathsf{tr}}$ for $1 \leq i \leq k-1$.

b) Assume that $r_1 = k$ and in $type_{z-1}(s^0)$ we have a tuple

$$\tau = \left(\zeta^n, \zeta^{n-1} \dots, \zeta^{k+1}, \gamma^k, \gamma^{k-1}, \dots, \gamma^1, q_2, (m_2, r_2, \xi^n, \xi^{n-1}, \dots, \xi^{r_2+1}, q_3), \cdot\right),$$

where $r_2 \leq k$. Define

$$\eta^i = \begin{cases} \alpha^i & \text{for } k+1 \leq i \leq n, \\ \beta^i \cup \gamma^i & \text{for } i = k, \\ \alpha^i \cup \beta^i \cup \gamma^i & \text{for } 1 \leq i \leq k-1. \end{cases}$$

Then $\left(\eta^n, \eta^{n-1}, \dots, \eta^1, p, (m_1 m_2, r_2, \xi^n, \xi^{n-1}, \dots, \xi^{r_2+1}, q_3), fl'\right) \in type_z(s^0)$, where $fl' = \mathsf{tr}$ if and only if $fl = \mathsf{tr}$, and $\beta^0 \cup \{\sigma, \tau\} \subseteq \mathcal{T}_{\mathsf{tr}}$ and $\beta^i \cap \gamma^i \subseteq \mathcal{T}_{\mathsf{tr}}$ for $1 \leq i \leq k$, and $\alpha^i \cap (\beta^i \cup \gamma^i) \subseteq \mathcal{T}_{\mathsf{tr}}$ for $1 \leq i \leq k-1$ (the last two conditions say that only elements of $\mathcal{T}_{\mathsf{tr}}$ may appear in more than one component of the union defining η^i).

Notice that the sequence $type_z$ is monotone: $type_z(s^0) \subseteq type_{z+1}(s^0)$ for each $z \in \mathbb{N}$ and each 0-pds s^0. Because we have only finitely many 0-pds's, the sequence stabilizes for each 0-pds after a finite number of steps. This fixpoint is denoted as $type$ (without an index).

Definition B.10. We define $type(s^k)$ for any $1 \leq k \leq n$ and any k-pds s^k by induction on k and on the size of s^k. If s^k is empty, $type(s^k) = \emptyset$. Otherwise we decompose $s^k = t^k : t^{k-1}$ ($type$ for t^k and t^{k-1} are already defined). As $type(s^k)$ we take the set containing all elements $\sigma \in \mathcal{T}^k$ such that there exists a composer $(\beta^k, \beta^{k-1}; \{\sigma\}; \cdot)$ for which $\beta^k \subseteq type(t^k)$ and $\beta^{k-1} \subseteq type(t^{k-1})$.

The following two observations follows immediately from Definitions B.8 and B.10.

Proposition B.11. *Let $0 \leq l < k \leq n$, let $s^k : s^{k-1} : \dots : s^l$ be a k-pds, and let $\alpha^k \subseteq \mathcal{T}^k$. Then $\alpha^k \subseteq type(s^k : s^{k-1} \dots : s^l)$ if and only if there exists a composer $(\beta^k, \beta^{k-1}, \dots, \beta^l; \alpha^k; \cdot)$ such that $\beta^i \subseteq type(s^i)$ for each $l \leq i \leq k$.*

Proposition B.12. *Let $1 \leq k \leq n$, and let s^k be a k-pds. Then $(\mathsf{ne}, \mathsf{tr}) \in type(s^k)$ if and only if s^k is not empty.*

Proof of Lemma B.7: We make an external induction on the length of the run and an internal induction on l. Fix some tuple $\widehat{\rho} = (m, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q) \in \mathcal{D}^l$. Assume we have a run R from c which agrees with $\widehat{\rho}$. Decompose $\pi(c) = s^n : s^{n-1} : \dots : s^l$. Consider first the case $l = 0$. Lemma B.4 gives us five possible forms of R; we analyze these cases.

Assume that R is a one-step run performing a $\mathsf{read}(f)$ operation composed with an r-return S (case 1). Let a be the symbol read by the first operation of R. Then the word read by S evaluates to m' such that $\varphi(a)m' = m$. Thus S agrees with $\widehat{\sigma} = (m', r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q)$. From the induction assumption (for run S) we know that in $type(s^0)$ we have a tuple $(\alpha^n, \alpha^{n-1}, \dots, \alpha^1, f(a), \widehat{\sigma}, \cdot)$ such that $\alpha^i \subseteq type(s^i)$ for $1 \leq i \leq n$. From Definition B.9 (point 1) we know that in $type(s^0)$ there is a tuple $(\alpha^n, \alpha^{n-1}, \dots, \alpha^1, p, \widehat{\rho}, \cdot)$.

Assume that R is a one-step run performing a $\mathsf{pop}^r(q)$ operation (case 2a). Then $\pi(R(1)) = s^n : s^{n-1} : \dots : s^r$, so $\xi^i \subseteq type(s^i)$ for $r+1 \leq i \leq n$. Of course s^r is not empty, so $(\mathsf{ne}, \mathsf{tr}) \in type(s^r)$. We also have $m = \varphi(\varepsilon)$. From Definition B.9 (point 2a) we know that in $type(s^0)$ there is a tuple $(\xi^n, \xi^{n-1}, \dots, \xi^{r+1}, \{(\mathsf{ne}, \mathsf{tr})\}, \emptyset, \emptyset, \dots, \emptyset, p, \widehat{\rho}, \cdot)$.

Assume that R is a one-step run performing a $\mathsf{pop}^k(q_1)$ operation composed with an r-return S, where $k \leq r-1$ (case 2b). Then $\pi(S(0)) = s^n : s^{n-1} : \dots : s^k$, and S also agrees with $\widehat{\rho}$. From the induction assumption (for run S and for k as l) we get that in $type(s^k)$ there is a tuple $\rho^k = (\alpha^n, \alpha^{n-1}, \dots, \alpha^{k+1}, q_1, \widehat{\rho}, \cdot)$ such that $\alpha^i \subseteq type(s^i)$ for $k+1 \leq i \leq n$. From Definition B.9 (point 2b) we know that in $type(s^0)$ there is a tuple $(\alpha^n, \alpha^{n-1}, \dots, \alpha^{k+1}, \{\rho^k\}, \emptyset, \emptyset, \dots, \emptyset, p, \widehat{\rho}, \cdot)$.

Assume that R is a one-step run performing a $\mathsf{push}^k(t^0, q_1)$ operation composed with an r-return S, where $k \neq r$ (case 3a). We have

$$\pi(S(0)) = s^n : s^{n-1} : \dots : s^{k+1} : (s^k : s^{k-1} : \dots : s^0) : s^{k-1} : s^{k-2} : \dots : s^1 : t^0,$$

and S also agrees with $\widehat{\rho}$. From the induction assumption (for shorter run S) we get that in $type(t^0)$ there is a tuple $(\alpha^n, \alpha^{n-1}, \dots, \alpha^1, q_1, \widehat{\rho}, q, \cdot)$ such that $\alpha^i \subseteq type(s^i)$ for $1 \leq i \leq k-1$ and for $k+1 \leq i \leq n$, and $\alpha^k \subseteq type(s^k : s^{k-1} : \dots : s^0)$. From Proposition B.11 we get that there exists a composer $(\beta^k, \beta^{k-1}, \dots, \beta^0; \alpha^k; \cdot)$ such that $\beta^i \subseteq type(s^i)$ for $0 \leq i \leq k$. Let η^i be defined like in Definition B.9, point 3a; we see that $\eta^i \subseteq type(s^i)$ for $1 \leq i \leq n$. We get that in $type(s^0)$ there is a tuple $(\eta^n, \eta^{n-1}, \dots, \eta^1, p, \widehat{\rho}, \cdot)$.

Assume that R is a one-step run performing a $\mathsf{push}^k(t^0, q_1)$ operation composed with a k-return S composed with an r-return T, where $k \geq r$ (case 3b) Let m_1 and m_2 be the images under φ of the words read by S and by T, respectively; we have $m_1 m_2 = m$. Notice that the topmost k-pds's of $R(0)$ and of $T(0)$ are the same (Proposition B.2). Decompose $\pi(T(0)) = u^n : u^{n-1} : \dots : u^{k+1} : s^k : s^{k-1} : \dots : s^0$. Run T agrees with $\widehat{\tau} = (m_2, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q)$. From the induction assumption for T we get that in $type(s^0)$ there is a tuple $(\zeta^n, \zeta^{n-1}, \dots, \zeta^{k+1}, \gamma^k, \gamma^{k-1}, \dots, \gamma^1, state(T(0)), \widehat{\tau}, \cdot)$ such that $\zeta^i \subseteq type(u^i)$ for $k+1 \leq i \leq n$ and $\gamma^i \subseteq type(s^i)$ for $1 \leq i \leq k$. Run S agrees with $\widehat{\sigma} = (m_1, k, \zeta^n, \zeta^{n-1}, \dots, \zeta^{k+1}, state(T(0)))$. From the induction assumption for S we get that in $type(t^0)$ there is a tuple $(\alpha^n, \alpha^{n-1}, \dots, \alpha^1, q_1, \widehat{\sigma}, \cdot)$ such that $\alpha^i \subseteq type(s^i)$ for $1 \leq i \leq k-1$ and for $k+1 \leq i \leq n$, and $\alpha^k \subseteq type(s^k : s^{k-1} : \dots : s^0)$. From Proposition B.11 we get that there exists a composer $(\beta^k, \beta^{k-1}, \dots, \beta^0; \alpha^k; \cdot)$ such that $\beta^i \subseteq type(s^i)$ for $0 \leq i \leq k$. Let η^i be defined like in Definition B.9, point 3b; we see that $\eta^i \subseteq type(s^i)$ for $1 \leq i \leq n$. We get that in $type(s^0)$ there is a tuple $(\eta^n, \eta^{n-1}, \dots, \eta^1, p, \widehat{\rho}, \cdot)$.

Consider now the case $l > 0$. Decompose $s^l = t^l : t^{l-1}$. From the induction assumption for $l-1$ we get that in $type(t^{l-1})$ there is a tuple $\rho^{l-1} = (\alpha^n, \alpha^{n-1}, \dots, \alpha^l, p, \widehat{\rho}, \cdot)$ such that $\alpha^i \subseteq type(s^i)$ for $l+1 \leq i \leq n$, and $\alpha^l \subseteq type(t^l)$. From Definition B.8, point 1 we get that $(\alpha^l, \{\rho^{l-1}\}; \{\rho^l\}; \{\mathsf{tr}\})$ is a composer, where $\rho^l = (\alpha^n, \alpha^{n-1}, \dots, \alpha^{l+1}, p, \widehat{\rho}, fl)$ (for some fl). From Proposition B.11 we get that $\rho^l \in type(s^l)$. ■

Next, we are going to prove the opposite: that types witnesses existence of some runs. Simultaneously we want to bound the number of the $\sharp$ symbols read by these runs. To handle this, we need to define a decomposition.

Definition B.13. Let $0 \leq k \leq n$, let s^k be a k-pds, and let $\alpha^k \subseteq type(s^k)$. We define when ω is a *decomposition* of (s^k, α^k). The definition is by induction on k and on the size of s^k. If $|\alpha^k| \neq 1$, ω consists of decompositions ω_σ of $(s^k, \{\sigma\})$ for each $\sigma \in \alpha^k$, as defined below. If $\alpha^k = \{\sigma\}$ and $k = 0$ we have exactly one decomposition $\omega = (s^0, \sigma)$. If $\alpha^k = \{\sigma\}$ and $s^k = t^k : t^{k-1}$, a decomposition ω of (s^k, α^k) consists of

- s^k and σ, and
- a composer $(\beta^k, \beta^{k-1}; \alpha^k; \cdot)$ such that $\beta^k \subseteq type(t^k)$ and $\beta^{k-1} \subseteq type(t^{k-1})$, and
- a decomposition ω^k of (t^k, β^k), and
- a decomposition ω^{k-1} of (t^{k-1}, β^{k-1}).

Proposition B.14. *Let $0 \leq k \leq n$, let s^k be a k-pds, and let $\alpha^k \subseteq type(s^k)$. Then there exists a decomposition of (s^k, α^k).*

To each decomposition ω we are going to assign two natural numbers $low(\omega)$ and $high(\omega)$.

Definition B.15. For positive natural numbers $m_1, \dots, m_k$ we define $pow(m_1, \dots, m_k)$ by induction on k:

$$pow() = 1, \qquad \text{and} \qquad pow(m_1, m_2, \dots, m_k) = (1 + m_1)^{pow(m_2, \dots, m_k)} - 1.$$

Proposition B.16. *The following is true for any positive natural numbers:*

$$pow(a_1, a_2, \dots, a_k, pow(b_1, b_2, \dots, b_l)) = pow(a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_l), \tag{1}$$

$$\begin{aligned} &pow(a_1, a_2, \dots, a_k, b_1 \cdot pow(c_1, c_2, \dots, c_l), b_2, b_3, \dots, b_l) \leq \\ &\qquad \leq pow(a_1, a_2, \dots, a_k, b_1 c_1, b_2 c_2, \dots, b_l c_l), \end{aligned} \tag{2}$$

$$pow(a_1, \dots, a_{i-1}, a_i^x, a_{i+1}, \dots, a_k) \leq pow(a_1, a_2 \dots, a_{k-1}, x a_k) \qquad \textit{for } 1 \leq i \leq k-1, \tag{3}$$

$$pow(a_1, a_2, \dots, a_k) + 1 \leq pow(a_1, a_2, \dots, a_{k-1}, a_k + 1), \tag{4}$$

$$pow(a_1, a_2, \dots, a_k) \cdot pow(b_1, b_2, \dots, b_k) \leq pow(a_1 b_1, a_2 b_2, \dots, a_k b_k). \tag{5}$$

Definition B.17. Let $0 \leq k \leq n$, and let ω be a decomposition of (s^k, α^k) for some k-pds s^k and some $\alpha^k \subseteq type(s^k)$. We define natural numbers $low(\omega)$ and $high(\omega)$ by induction on k and on the size of s^k.

- If $k = 0$ and $\alpha^0 = \{\sigma\} \subseteq \mathcal{T}_{\mathsf{tr}}$, we take $low(\omega) = 0$ and $high(\omega) = 1$.
- If $k = 0$ and $\alpha^0 = \{\sigma\} \subseteq \mathcal{T}_{\mathsf{nt}}$, we take $low(\omega) = 1$ and $high(\omega) = C_z$ where z is the smallest number such that $\sigma \in type_z(s^0)$ and C_z is defined inductively:

$$C_0 = 1, \qquad \text{and} \qquad C_z = (3|\mathcal{T}^0|)^n \cdot (C_{z-1})^{|\mathcal{T}^0|^2 + 2}.$$

- If $s^k = t^k : t^{k-1}$ and $\alpha^k = \{\sigma\}$, the decomposition consists of a composer $(\beta^k, \beta^{k-1}; \alpha^k; \cdot)$ and decompositions ω^k and ω^{k-1} of (t^k, β^k) and (t^{k-1}, β^{k-1}), respectively. We take

$$low(\omega) = low(\omega^k) + low(\omega^{k-1}), \qquad \text{and} \qquad high(\omega) = pow(high(\omega^k), high(\omega^{k-1})).$$

- In general (for α^k of size different than 1), ω consists of decompositions ω_σ of $(s^k, \{\sigma\})$, for each $\sigma \in \alpha^k$. We take

$$low(\omega) = \sum_{\sigma\in\alpha^k} low(\omega_\sigma), \qquad \text{and} \qquad high(\omega) = \prod_{\sigma\in\alpha^k} high(\omega_\sigma).$$

Proposition B.18. *Let $0 \le k \le n$ and let ω be a decomposition of (s^k, α^k) for some k-pds s^k and some $\alpha^k \subseteq type(s^k)$. If $\alpha^k \subseteq \mathcal{T}_{\mathsf{tr}}$ then $low(\omega) = 0$ and $high(\omega) = 1$; otherwise $low(\omega) \ge 1$ and $high(\omega) \ge 2$.*

Proof: By induction on k and on the size of s^k. We analyze Definition B.17. In the third case observe (see Definition B.8, point 1) that $\alpha^k \subseteq \mathcal{T}_{\mathsf{tr}}$ if and only if $\beta^k \subseteq \mathcal{T}_{\mathsf{tr}}$ and $\beta^{k-1} \subseteq \mathcal{T}_{\mathsf{tr}}$; the other cases are immediate. ■

Proposition B.19. *For each $L \in \mathbb{N}$ there exists $H_L \in \mathbb{N}$ such that for each decomposition ω such that $low(\omega) \le L$ we have $high(\omega) \le H_L$.*

Proof: We prove this by induction on L. For $L = 0$ we can take $H_0 = 1$ (by the above proposition, $low(\omega) = 0$ implies $high(\omega) = 1$). Fix some $L \ge 1$. Let $N_{L,0}$ be the maximum of $high(\omega)$ for decompositions ω of (s^0, α^0) over every 0-pds s^0 and every $\alpha^0 \subseteq type(s^0)$. There are only finitely many of them, so this maximum exists. For $1 \le k \le n$ we define by induction

$$N_{L,k} = \max(pow(1, N_{L,k-1}), pow(H_{L-1}, H_{L-1}), (H_{L-1})^{|\mathcal{T}^k|}),$$

and we take $H_L = N_{L,n}$. We will show that if ω is a decomposition of (s^k, α^k) for some k-pds s^k and some $\alpha^k \subseteq type(s^k)$ such that $low(\omega) \le L$, then $high(\omega) \le N_{L,k}$. This gives the thesis as $N_{L,k} \le H_L$.

We prove this by induction on k and on the size of the k-pds s^k. For $k = 0$ this is true by definition. Let $1 \le k \le n$, and let first $|\alpha^k| = 1$. Then s^k is nonempty, $s^k = t^k : t^{k-1}$. The decomposition ω contains a composer $(\beta^k, \beta^{k-1}; \alpha^k; \cdot)$, and a decomposition ω^k of (t^k, β^k), and a decomposition ω^{k-1} of (t^{k-1}, β^{k-1}). By definition

$$low(\omega) = low(\omega^k) + low(\omega^{k-1}), \qquad \text{and} \qquad high(\omega) = pow(high(\omega^k), high(\omega^{k-1})).$$

We have three cases. The first case is that $low(\omega^k) \le L-1$ and $low(\omega^{k-1}) \le L-1$. Then, by the "big" induction assumption, $high(\omega^k) \le H_{L-1}$ and $high(\omega^{k-1}) \le H_{L-1}$. Thus

$$high(\omega) \le pow(H_{L-1}, H_{L-1}) \le N_{L,k}.$$

The second case is that $low(\omega^k) = L$ and $low(\omega^{k-1}) = 0$. By the induction assumption for a smaller k-pds we get that $high(\omega^k) \le N_{L,k}$. Since $high(\omega^{k-1}) = 1$, we have

$$high(\omega) \le pow(N_{L,k}, 1) = N_{L,k}.$$

The third case is that $low(\omega^k) = 0$ and $low(\omega^{k-1}) = L$. By the induction assumption for $k-1$ we get that $high(\omega^{k-1}) \le N_{L,k-1}$. Since $high(\omega^k) = 1$, we have

$$high(\omega) \le pow(1, N_{L,k-1}) \le N_{L,k}.$$

Let now $|\alpha^k| \ne 1$. Then ω contains decompositions ω_σ of $(s^k, \{\sigma\})$ for every $\sigma \in \alpha^k$. By definition

$$low(\omega) = \sum_{\sigma\in\alpha^k} low(\omega_\sigma), \qquad \text{and} \qquad high(\omega) = \prod_{\sigma\in\alpha^k} high(\omega_\sigma).$$

One possibility is that $low(\omega_\sigma) \le L-1$ for all $\sigma \in \alpha^k$. Then, by the "big" induction assumption, $high(\omega_\sigma) \le H_{L-1}$ for all $\sigma \in \alpha^k$, so

$$high(\omega) \le (H_{L-1})^{|\alpha^k|} \le (H_{L-1})^{|\mathcal{T}^k|} \le N_{L,k}.$$

Otherwise for some $\sigma \in \alpha^k$ we have $low(\omega_\sigma) = L$, and $low(\omega_\tau) = 0$ for all $\tau \in \alpha^k$, $\tau \ne \sigma$. Then $high(\omega_\sigma) \le N_{l,k}$ (by the analysis for $|\alpha^k| = 1$), and $high(\omega_\tau) = 1$ for all $\tau \in \alpha^k$, $\tau \ne \sigma$, so

$$high(\omega) = high(\omega_\sigma) \le N_{L,k}.$$

■

Definition B.20. We simultaneously define what does it mean that a decomposition ω^k has a witness, and what does it mean that a run agrees with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \dots, \omega^k)$, where $\widehat{\sigma} \in \mathcal{D}^k$ and $\omega^n, \omega^{n-1}, \dots, \omega^k$ are decompositions. This definition is by induction on k (going down). Let $0 \le k \le n$. Assume that these two notions are already defined for levels greater than k.

1) Let $r \ge k+1$ and let R be a run. Decompose $\pi(R(0)) = s^n : s^{n-1} : \cdots : s^k$ and $\pi(R(|R|)) = t^n : t^{n-1} : \cdots : t^r$. Let $\widehat{\sigma} = (m, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q)$ and, for $k \le i \le n$, let ω^i be a decomposition of (s^i, α^i) for some $\alpha^i \subseteq type(s^i)$. Let x be the number of the $\sharp$ symbols read by R. We say that run R *agrees with* $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \dots, \omega^k)$ if R agrees with $\widehat{\sigma}$,

and for each $r+1 \leq i \leq n$ there exists a decomposition χ^i of (t^i, ξ^i) which has a witness, such that for each positive integer K it holds

$$\sum_{i=k}^{n} low(\omega^i) \leq x + \sum_{i=r+1}^{n} low(\chi^i), \qquad \text{and}$$
$$pow\big(high(\omega^n), \dots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \dots, high(\omega^k)\big) \geq$$
$$\geq x + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r+1}), K\big)\,.$$

2) Let ω^k be a decomposition of (s^k, α^k) for some k-pds s^k and some $\alpha^k \subseteq type(s^k)$. If $|\alpha^k| \neq 1$, ω^k consists of a decomposition ω^k_σ for each $\sigma \in \alpha^k$; we say that ω^k *has a witness* if every ω^k_σ has a witness, as defined below. If $\alpha^k = \{(\mathsf{ne}, \mathsf{tr})\}$, we always say that ω^k *has a witness*. If $\alpha^k = \{\sigma\}$ and $\sigma = (\alpha^n, \alpha^{n-1}, \dots, \alpha^{k+1}, p, \widehat{\sigma}, \cdot)$, we say that ω^k *has a witness* if the following implication is true.

   For $k+1 \leq i \leq n$, let ω^i be a decomposition of (s^i, α^i) which has a witness, for some i-pds s^i such that $\alpha^i \subseteq type(s^i)$. Then there exists a run from $(p, s^n : s^{n-1} : \dots : s^k)$ which agrees with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \dots, \omega^k)$.

In the last part of the above definition we mean that for all appropriate $\omega^{k+1}, \omega^{k+2}, \dots, \omega^n$ the run exists. Our goal now is to prove that every decomposition has a witness (Corollary B.23). Heading toward this, we first show how having a witness interplays with composers.

Proposition B.21. *Let $0 \leq l < k \leq n$. For each $l \leq i \leq k$ let ω^i be a decomposition of (s^i, β^i) which has a witness, for some i-pds s^i and some $\beta^i \subseteq type(s^i)$. Let $(\beta^k, \beta^{k-1}, \dots, \beta^l; \alpha^k; fl)$ be a composer. Then there exists a decomposition ω of $(s^k : s^{k-1} : \dots : s^l, \alpha^k)$ which has a witness, such that*

$$\sum_{i=l}^{k} low(\omega^i) \leq low(\omega),$$
$$\sum_{i=l}^{k} low(\omega^i) < low(\omega) \qquad \text{if } fl = \mathsf{nt},$$
$$pow\Big(high(\omega^k), high(\omega^{k-1}), \dots, high(\omega^{l+1}), \left|\mathcal{T}^0\right|^n \cdot high(\omega^l)^{|\mathcal{T}^0|}\Big) \geq high(\omega),$$
$$pow\big(high(\omega^k), high(\omega^{k-1}), \dots, high(\omega^{l+1}), high(\omega^l)\big) \geq high(\omega) \qquad \text{if } fl = \mathsf{tr}.$$

Proof: Assume first that $l = k-1$ and $\alpha^k = \{\sigma^k\}$ has one element, which is not $(\mathsf{ne}, \mathsf{tr})$. In such case also $\beta^{k-1} = \{\sigma^{k-1}\}$ has one element. Let $\sigma^{k-1} = (\gamma^n, \gamma^{n-1}, \dots, \gamma^k, p, \widehat{\sigma}, \cdot)$ and $\widehat{\sigma} = (m, r, \xi^n, \xi^{n-1}, \xi^{r+1}, q)$. Then $\beta^k = \gamma^k$ and $\sigma^k = (\gamma^n, \gamma^{n-1}, \dots, \gamma^{k+1}, p, \widehat{\sigma}, \cdot)$. To the decomposition ω we take composer $(\beta^k, \beta^{k-1}; \alpha^k; fl)$, and decompositions ω^k and ω^{k-1}. By Definition B.17 we have

$$low(\omega) = low(\omega^k) + low(\omega^{k-1}), \qquad \text{and} \tag{6}$$
$$high(\omega) = pow\big(high(\omega^k), high(\omega^{k-1})\big)\,. \tag{7}$$

This immediately gives the required inequalities (in particular $fl = \mathsf{tr}$, so the second inequality is not needed). To show that ω has a witness, for $k+1 \leq i \leq n$ take a decomposition ω^i of (s^i, γ^i) which has a witness, for some i-pds s^i such that $\gamma^i \in type(s^i)$. Recall also that $\gamma^k \in type(s^k)$ and ω^k is a decomposition of (s^k, γ^k) which has a witness. By expanding the definition of witness for ω^{k-1} we obtain a run R from $(p, s^n : s^{n-1} : \dots : s^{k-1})$ which agrees with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \dots, \omega^{k-1})$. It is enough to show that R agrees also with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \dots, \omega^{k+1}, \omega)$. We already know that R agrees with $\widehat{\sigma}$. To check the rest, decompose $\pi(R(|R|)) = t^n : t^{n-1} : \dots : t^r$. Let x be the number of the $\sharp$ symbols read by R. Because R agrees with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \dots, \omega^{k-1})$, for each $r+1 \leq i \leq n$ we obtain a decomposition χ^i of (t^i, ξ^i) which has a witness, such that for each positive integer K it holds

$$\sum_{i=k-1}^{n} low(\omega^i) \leq x + \sum_{i=r+1}^{n} low(\chi^i), \qquad \text{and}$$
$$pow\big(high(\omega^n), \dots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \dots, high(\omega^{k-1})\big) \geq$$
$$\geq x + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r+1}), K\big)\,.$$

By substituting (6) and (7) on the left side of the above inequalities, and using (1), we get that (as $\widehat{\sigma} \in \mathcal{D}^k$, we have $r \geq k+1$, so the K factor does not interfere)

$$low(\omega) + \sum_{i=k+1}^{n} low(\omega^i) \le x + \sum_{i=r+1}^{n} low(\chi^i), \qquad \text{and}$$
$$pow\big(high(\omega^n), \dots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \dots, high(\omega^{k+1}), high(\omega)\big) \ge$$
$$\ge x + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r+1}), K\big)\,,$$

which are the inequalities required for the run to agree with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \dots, \omega^{k+1}, \omega)$.

Now assume only that $\alpha^k = \{\sigma^k\}$ has one element, which is not $(\mathsf{ne}, \mathsf{tr})$. We make an induction on $k - l$. For $l = k - 1$ we have the above case. Otherwise we have composers $(\beta^k, \{\sigma^{k-1}\}; \{\sigma^k\}; \mathsf{tr})$ and $(\beta^{k-1}, \beta^{k-2}, \dots, \beta^l; \{\sigma^{k-1}\}; \mathsf{tr})$ for some $\sigma^{k-1} \ne (\mathsf{ne}, \mathsf{tr})$ (follows easily from Definition B.8). By the induction assumption (used for the second of the composers) we have a decomposition χ of $(s^{k-1} : s^{k-2} : \dots : s^l, \{\sigma^{k-1}\})$ which has a witness, and satisfies the inequalities. So we may use the base case for the first composer, which gives a decomposition ω of $(s^k : s^{k-1} : \dots : s^l, \{\sigma^k\})$ which has a witness, and satisfies the inequalities. It remains to deduce the required inequalities from those which we get from the induction assumption. The first inequality follows immediately by summing the first inequalities from the induction assumption. The second one is not needed, as we always have $fl = \mathsf{tr}$. The fourth inequality we get from the induction assumption and then using (1):

$$high(\omega) \le pow\big(high(\omega^k), pow\big(high(\omega^{k-1}), high(\omega^{k-2}), \dots, high(\omega^l)\big)\big) = pow\big(high(\omega^k), high(\omega^{k-1}), \dots, high(\omega^l)\big)\,.$$

The third inequality is also true, as it is weaker.

Next, assume that $\alpha^k = \{(\mathsf{ne}, \mathsf{tr})\}$. Then, by definition, ω has a witness. Moreover $\beta^i = \emptyset$, so $low(\omega^i) = 0$ and $high(\omega^i) = 1$, for $l \le i \le k$, and $fl = \mathsf{tr}$. This trivially gives us the required inequalities.

Finally, consider an arbitrary situation. By Definition B.8, for each $\sigma \in \alpha^k$ we have a composer $(\beta^k_\sigma, \beta^{k-1}_\sigma, \dots, \beta^l_\sigma; \{\sigma\}; \mathsf{tr})$ such that $\beta^i = \bigcup_{\sigma \in \alpha^k} \beta^i_\sigma$ for $l \le i \le k$. Recall that decomposition ω^i (for $l \le i \le k$) consists of decompositions ω^i_τ of $(s^i, \{\tau\})$ for each $\tau \in \beta^i$; each of them has a witness. We can create from them a decomposition χ^i_σ of (s^i, β^i_σ) for each $\sigma \in \alpha^k$. By the above cases, we have a decomposition ω_σ of $(s^k : s^{k-1} : \dots : s^l, \{\sigma\})$ which has a witness, for each $\sigma \in \alpha^k$. This gives us the required decomposition ω of $(s^k : s^{k-1} : \dots : s^l, \alpha^k)$. To get what we need, we only have to check the inequalities. The first of them is get by summing the first inequality for each $\sigma \in \alpha^k$ and by observing that each element of β^i is in some β^i_σ:

$$low(\omega) = \sum_{\sigma \in \alpha^k} low(\omega_\sigma) \ge \sum_{\sigma \in \alpha^k} \sum_{i=l}^{k} low(\chi^i_\sigma) = \sum_{\sigma \in \alpha^k} \sum_{i=l}^{k} \sum_{\tau \in \beta^i_\sigma} low(\omega^i_\tau) \ge \sum_{i=l}^{k} \sum_{\tau \in \beta^i} low(\omega^i_\tau) = \sum_{i=l}^{k} low(\omega^i).$$

For the second inequality observe that if $fl = \mathsf{nt}$, then some $\tau \in \mathcal{T}_{\mathsf{nt}}$ appears in two β^i_σ, which means that some positive component $low(\omega^i_\tau)$ appears in two sums $\sum_{\tau \in \beta^i_\sigma} low(\omega^i_\tau)$, so the inequality becomes strict. Now we multiply the fourth inequality for each $\sigma \in \alpha^k$, we use (5), and we get

$$\begin{aligned} high(\omega) = \prod_{\sigma \in \alpha^k} high(\omega_\sigma) &\le \prod_{\sigma \in \alpha^k} pow\big(high(\chi^k_\sigma), high(\chi^{k-1}_\sigma), \dots, high(\chi^l_\sigma)\big) \le \\ &\le pow\Big(\prod_{\sigma \in \alpha^k} high(\chi^k_\sigma), \prod_{\sigma \in \alpha^k} high(\chi^{k-1}_\sigma), \dots, \prod_{\sigma \in \alpha^k} high(\chi^l_\sigma)\Big). \end{aligned} \tag{8}$$

Now observe for $l \le i \le k$ that

$$\prod_{\sigma \in \alpha^k} high(\chi^i_\sigma) = \prod_{\sigma \in \alpha^k} \prod_{\tau \in \beta^i_\sigma} high(\omega^i_\tau) \le \Big(\prod_{\tau \in \beta^i} high(\omega^i_\tau)\Big)^{|\alpha^k|} \le (high(\omega^i))^{|\mathcal{T}^0|}.$$

The last inequality is true, because $|\alpha^k| \le |\mathcal{T}^k| \le |\mathcal{T}^0|$. Using (3) we move the $|\mathcal{T}^0|$ exponents (there is at most n of them) into the last argument of pow and we get the third inequality.

$$\begin{aligned} high(\omega) &\le pow\Big((high(\omega^k))^{|\mathcal{T}^0|}, (high(\omega^{k-1}))^{|\mathcal{T}^0|}, \dots, (high(\omega^l))^{|\mathcal{T}^0|}\Big) \le \\ &\le pow\Big(high(\omega^k), high(\omega^{k-1}), \dots, high(\omega^{l+1}), |\mathcal{T}^0|^n \cdot (high(\omega^l))^{|\mathcal{T}^0|}\Big) \end{aligned}$$

Now assume that $fl = \mathsf{tr}$. It implies that each $\tau \in \beta^i \cap \mathcal{T}_{\mathsf{nt}}$ belongs to only one set β^i_σ, so all the common factors are equal to 1:

$$\prod_{\sigma \in \alpha^k} high(\chi^i_\sigma) = \prod_{\sigma \in \alpha^k} \prod_{\tau \in \beta^i_\sigma} high(\omega^i_\tau) = \prod_{\tau \in \beta^i} high(\omega^i_\tau) = high(\omega^i).$$

By substituting this to (8) we get the fourth inequality. ∎

Lemma B.22. *Let s^0 be a 0-pds and let $\rho \in type(s^0)$. Then the (unique) decomposition of $(s^0, \{\rho\})$ has a witness.*

Proof: We make an induction on $z \in \mathbb{N}$ for which $\rho \in type_z(s^0)$. For $z = 0$ this is trivial, as $type_0(s^0)$ is empty. Let $z > 0$. Take some 0-pds s^0 and $\rho = (\eta^n, \eta^{n-1}, \ldots, \eta^1, p, \widehat{\rho}, \cdot) \in type_z(s^0)$, where $\widehat{\rho} = (m, r, \xi^n, \xi^{n-1}, \ldots, \xi^{r+1}, q)$. We can assume that $\rho \notin type_{z-1}(s^0)$, as otherwise the thesis follows immediately from the induction assumption. Let ω^0 be the unique decomposition of $(s^0, \{\rho\})$. For $1 \leq i \leq n$, let ω^i be a decomposition of (s^i, η^i) which has a witness, for some i-pds s^i such that $\eta^i \subseteq type(s^i)$. Let $c = (p, s^n : s^{n-1} : \cdots : s^0)$. To show that ω^0 has a witness, we have to show that there exists a run from c which agrees with $(\widehat{\rho}, \omega^n, \omega^{n-1}, \ldots, \omega^0)$. Choose some positive integer K and denote

$$L = \sum_{i=0}^{n} low(\omega^i), \qquad \text{and}$$
$$H = pow\big(high(\omega^n), \ldots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \ldots, high(\omega^0)\big).$$

The element ρ is in $type_z(s^0)$ because it was added there by some of the points of Definition B.9; we analyze each of these points separately.

Case 1: Assume ρ was added to $type_z(s^0)$ by point 1 of Definition B.9. Then $\delta(s^0, p) = \mathsf{read}(f)$, and for some $a \in A$ in $type_{z-1}(s^0)$ we have a tuple $\sigma = (\eta^n, \eta^{n-1}, \ldots, \eta^1, f(a), \widehat{\sigma}, \cdot)$, where $\widehat{\sigma} = (m', r, \xi^n, \xi^{n-1}, \ldots, \xi^{r+1}, q)$ and $\varphi(a)m' = m$. Let d be the configuration obtained in one step from c, by reading the a symbol. We have $\pi(d) = \pi(c)$. From the induction assumption we know that the unique decomposition ω_σ of (s^0, σ) has a witness. By expanding the definition of having a witness we obtain a run S from d which agrees with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \ldots, \omega^1, \omega_\sigma)$. Let R be the composition of the one-step run from c to d with run S. We claim that R agrees with $(\widehat{\rho}, \omega^n, \omega^{n-1}, \ldots, \omega^0)$. It is easy to see that R agrees with $\widehat{\rho}$ (in particular Lemma B.4 implies that R is an r-return). Let $\pi(R(|R|)) = t^n : t^{n-1} : \cdots : t^r$. Let x and x' be the number of the $\sharp$ symbols read by R, and by S, respectively. Because S agrees with $(\widehat{\sigma}, \omega^n, \omega^{n-1}, \ldots, \omega^1, \omega_\sigma)$, for $r+1 \leq i \leq n$ we have a decomposition χ^i of (t^i, ξ^i) which has a witness, such that

$$low(\omega_\sigma) + \sum_{i=1}^{n} low(\omega^i) \leq x' + \sum_{i=r+1}^{n} low(\chi^i), \qquad \text{and} \tag{9}$$
$$pow\big(high(\omega^n), \ldots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \ldots, high(\omega^1), high(\omega_\sigma)\big) \geq$$
$$\geq x' + pow\big(high(\chi^n), high(\chi^{n-1}), \ldots, high(\chi^{r+1}), K\big). \tag{10}$$

Notice that $low(\omega^0) - low(\omega_\sigma) \leq x - x'$. Indeed, either $low(\omega^0) = low(\omega_\sigma)$ and $x \geq x'$, or $low(\omega^0) = 1$ and $low(\omega_\sigma) = 0$ and $x = x' + 1$. Together with (9) it gives us that

$$L = \Big(low(\omega_\sigma) + \sum_{i=1}^{n} low(\omega^i)\Big) + \big(low(\omega^0) - low(\omega_\sigma)\big) \leq \Big(x' + \sum_{i=r+1}^{n} low(\chi^i)\Big) + (x - x') = x + \sum_{i=r+1}^{n} low(\chi^i),$$

which is exactly what we need. If $\rho \in \mathcal{T}_{\mathsf{tr}}$, then also $\sigma \in \mathcal{T}_{\mathsf{tr}}$ and $x = x'$, so $high(\omega^0) = high(\omega_\sigma)$, and (10) gives us that

$$H \geq x + pow\big(high(\chi^n), high(\chi^{n-1}), \ldots, high(\chi^{r+1}), K\big).$$

If $\rho \notin \mathcal{T}_{\mathsf{tr}}$, then $high(\omega^0) \geq 1 + high(\omega_\sigma)$, because $\sigma \in type_{z-1}(s^0)$ but $\rho \notin type_{z-1}(s^0)$. Using (4), (10), and $x' + 1 \geq x$ we get the required inequality:

$$H \geq pow\big(high(\omega^n), \ldots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \ldots, high(\omega^1), 1 + high(\omega_\sigma)\big) \geq$$
$$\geq pow\big(high(\omega^n), \ldots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \ldots, high(\omega^1), high(\omega_\sigma)\big) + 1 \geq$$
$$\geq x + pow\big(high(\chi^n), high(\chi^{n-1}), \ldots, high(\chi^{r+1}), K\big).$$

Case 2a: Assume ρ was added to $type_z(s^0)$ by point 2a of Definition B.9. Then $\delta(s^0, p) = \mathsf{pop}^r(q)$, and $m = \varphi(\varepsilon)$, and $\rho \in \mathcal{T}_{\mathsf{tr}}$, and $\eta^i = \emptyset$ for $1 \leq i \leq r-1$, and $\eta^r = \{(\mathsf{ne}, \mathsf{tr})\}$, and $\eta^i = \xi^i$ for $r+1 \leq i \leq n$. Because $\eta^r \subseteq type(s^r)$, s^r is not empty, so it is possible to execute the pop^r operation from c; let d be the configuration obtained in one step from c. We have $\pi(d) = s^n : s^{n-1} : \cdots : s^r$. Let R be the one-step run from c to d. We see that R is an r-return (see Lemma B.4), and $q = state(d)$, and the word read by R is empty, so it evaluates to m under φ, and $\xi^i = \eta^i \subseteq type(s^i)$ for $r+1 \leq i \leq n$. Thus R agrees with $\widehat{\rho}$. The number of the $\sharp$ symbols read by R is 0. Because $\rho \in \mathcal{T}_{\mathsf{tr}}$ and $\eta^i = \emptyset$ for $1 \leq i \leq r-1$ and $\eta^r = \{(\mathsf{ne}, \mathsf{tr})\}$, we have $low(\omega^i) = 0$ and $high(\omega^i) = 1$ for $0 \leq i \leq r$. We see that

$$L = \sum_{i=r+1}^{n} low(\omega^i), \qquad \text{and}$$
$$H = pow\big(high(\omega^n), high(\omega^{n-1}), \ldots, high(\omega^{r+1}), K\big),$$

which says that R agrees with $(\widehat{\rho}, \omega^n, \omega^{n-1}, \dots, \omega^0)$.

Case 2b: Assume ρ was added to $type_z(s^0)$ by point 2b of Definition B.9. Then $\delta(s^0, p) = \mathsf{pop}^k(q_1)$, and $\rho \in \mathcal{T}_{\mathsf{tr}}$, and $\eta^i = \emptyset$ for $1 \le i \le k-1$, and $\eta^k = \{\sigma\}$, where $\sigma = (\eta^n, \eta^{n-1}, \dots, \eta^{k+1}, q_1, \widehat{\rho}, \cdot)$. In particular $k \le r-1$ (which is required by $\widehat{\rho} \in \mathcal{D}^k$). Because $\sigma \in type(s^k) \neq \emptyset$, we know that s^k is nonempty, so it is possible to execute the pop^k operation from c. Let d be the configuration obtained in one step from c. We have $\pi(d) = s^n : s^{n-1} : \dots : s^k$. We expand the definition of having a witness (for ω^k): we obtain a run S from d which agrees with $(\widehat{\rho}, \omega^n, \omega^{n-1}, \dots, \omega^k)$. Let $\pi(S(|S|)) = t^n : t^{n-1} : \dots : t^r$. Let R be the composition of the one-step run from c to d with run S. We get that R is an r-return (see Lemma B.4), as well as $q = state(R(|R|))$; the word read by R is the same as by S, so it evaluates to m under φ; additionally $\xi^i \subseteq type(t^i)$ for $r+1 \le i \le n$. Thus R agrees with $\widehat{\rho}$. The number of the $\sharp$ symbols read by R and by S is the same; denote it x. Because S agrees with $(\widehat{\rho}, \omega^n, \omega^{n-1}, \dots, \omega^k)$, for $r+1 \le i \le n$ we have a decomposition χ^i of (t^i, ξ^i) which has a witness, such that

$$\sum_{i=k}^{n} low(\omega^i) \le x + \sum_{i=r+1}^{n} low(\chi^i), \qquad \text{and}$$
$$pow\big(high(\omega^n), \dots, high(\omega^{r+1}), K \cdot high(\omega^r), high(\omega^{r-1}), \dots, high(\omega^k)\big) \ge$$
$$\ge x + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r+1}), K\big)\,.$$

Because $\rho \in \mathcal{T}_{\mathsf{tr}}$ and $\eta^i = \emptyset$ for $1 \le i \le k-1$, we have $low(\omega^i) = 0$ and $high(\omega^i) = 1$ for $0 \le i \le k-1$; we get that

$$L \le x + \sum_{i=r+1}^{n} low(\chi^i), \qquad \text{and}$$
$$H \ge x + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r+1}), K\big)\,,$$

which says that R agrees with $(\widehat{\rho}, \omega^n, \omega^{n-1}, \dots, \omega^0)$.

Case 3: Assume ρ was added to $type_z(s^0)$ by point 3a or 3b of Definition B.9. Then $\delta(s^0, p) = \mathsf{push}^k(t^0, q_1)$. In $type_{z-1}(t^0)$ we have a tuple $\sigma = (\alpha^n, \alpha^{n-1}, \dots, \alpha^1, q_1, \widehat{\sigma}, \cdot)$, where $\widehat{\sigma} = (m_1, r_1, \zeta^n, \zeta^{n-1}, \dots, \zeta^{r_1+1}, q_2)$ such that $\alpha^i \subseteq \eta^i \subseteq type(s^i)$ for $1 \le i \le k-1$ and for $k+1 \le i \le n$. We also have a composer $(\beta^k, \beta^{k-1}, \dots, \beta^0; \alpha^k; fl)$ such that $\beta^i \subseteq \eta^i \subseteq type(s^i)$ for $1 \le i \le k$, and $\beta^0 \subseteq type_{z-1}(s^0)$. As a part of decomposition ω^i we obtain a decomposition ω_{α^i} of (s^i, α^i) which has a witness, for $1 \le i \le k-1$ and for $k+1 \le i \le n$, and a decomposition ω_{β^i} of (s^i, β^i) which has a witness, for $1 \le i \le k$. By the induction assumption the unique decomposition ω_σ of $(t^0, \{\sigma\})$ and the unique decomposition ω_{β^0} of (s^0, β^0) have a witness. From Proposition B.11 we know that $\alpha^k \subseteq type(s^k : s^{k-1} : \dots : s^0)$. Using Proposition B.21 we obtain a decomposition ω_{α^k} of $(s^k : s^{k-1} : \dots : s^0, \alpha^k)$ which has a witness. Denote

$$\begin{aligned} a_i &= high(\omega_{\alpha^i}) && \text{for } 1 \le i \le n, \\ b_i &= high(\omega_{\beta^i}) && \text{for } 0 \le i \le k, \\ K_i &= K && \text{for } i = r, \\ K_i &= 1 && \text{for } i \neq r. \end{aligned}$$

From Proposition B.21 we get the following inequalities; the first of them is strict if $fl = \mathsf{nt}$:

$$\sum_{i=0}^{k} low(\omega_{\beta^i}) \le low(\omega_{\alpha^k}), \tag{11}$$
$$pow(b_k, b_{k-1}, \dots, b_1, b_0) \ge a_k \qquad \text{if } fl = \mathsf{tr}, \tag{12}$$
$$pow\Big(b_k, b_{k-1}, \dots, b_1, |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|}\Big) \ge a_k. \tag{13}$$

Let d be the configuration obtained in one step from c. We have

$$\pi(d) = s^n : s^{n-1} : \dots : s^{k+1} : (s^k : s^{k-1} : \dots : s^0) : s^{k-1} : s^{k-2} : \dots : s^1 : t^0.$$

Because ω_σ has a witness, we have a run S from d which agrees with $(\widehat{\sigma}, \omega_{\alpha^n}, \omega_{\alpha^{n-1}}, \dots, \omega_{\alpha^1}, \omega_\sigma)$. Let $\pi(S(|S|)) = u^n : u^{n-1} : \dots : u^{r_1}$. Let x_1 be the number of the $\sharp$ symbols read by S. Because S agrees with $(\widehat{\sigma}, \omega_{\alpha^n}, \omega_{\alpha^{n-1}}, \dots, \omega_{\alpha^1}, \omega_\sigma)$, for $r_1+1 \le i \le n$ we have a decomposition χ^i of (u^i, ζ^i) which has a witness, such that for each positive integer K' it holds

$$low(\omega_\sigma) + \sum_{i=1}^{n} low(\omega_{\alpha^i}) \le x_1 + \sum_{i=r_1+1}^{n} low(\chi^i), \qquad \text{and} \tag{14}$$
$$pow(a_n, \dots, a_{r+1}, K' \cdot a_r, a_{r-1}, \dots, a_1, high(\omega_\sigma)) \ge x_1 + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r_1+1}), K'\big)\,. \tag{15}$$

Case 3a: Assume ρ was added to $type_z(s^0)$ by point 3a of Definition B.9. Then $r_1 = r \neq k$, and $m_1 = m$, and $q_2 = q$, and $\xi^i = \zeta^i$ for $r+1 \leq i \leq n$. Let R be the composition of the one-step run from c to d with run S. We see that R is an r-return (see Lemma B.4), as well as $q = state(R(|R|))$; the word read by R is the same as by S, so it evaluates to m under φ; additionally $\xi^i \subseteq type(u^i)$ for $r+1 \leq i \leq n$. Thus R agrees with $\widehat{\rho}$. The number of the $\sharp$ symbols read by R is x_1. Recall that, for $1 \leq i \leq k-1$, the decomposition ω^i of (s^i, η^i) consists of decompositions of $(s^i, \{\upsilon\})$ for every $\upsilon \in \eta^i$, and $low(\omega^i)$ is the sum of low for these decompositions. Simultaneously, decomposition ω_{α^i} of (s^i, α^i) consists of some of these decompositions, as $\alpha^i \subseteq \eta^i$, so in $low(\omega_{\alpha^i})$ we sum only some of the components summed in $low(\omega^i)$; similarly for ω_{β^i}. However, for $1 \leq i \leq k-1$, every element of η^i is in some α^i or in some β^i. Additionally $\eta^i = \alpha^i$ for $k+1 \leq i \leq n$, and $\eta^k = \beta^k$. Thus

$$\begin{aligned} &low(\omega^i) = low(\omega_{\alpha^i}) && \text{for } k+1 \leq i \leq n, \\ &low(\omega^i) = low(\omega_{\beta^i}) && \text{for } i = k. \\ &low(\omega^i) \leq low(\omega_{\alpha^i}) + low(\omega_{\beta^i}) && \text{for } 1 \leq i \leq k-1. \end{aligned}$$

Notice that if for some $1 \leq i \leq k-1$ we have $\alpha^i \cap \beta^i \not\subseteq \mathcal{T}_{\mathsf{tr}}$, then the appropriate inequality is strict (as the positive component corresponding to $\upsilon \in \alpha^i \cap \beta^i \cap \mathcal{T}_{\mathsf{nt}}$ appears in both low's on the right side, and only once on the left side). We sum these inequalities together; next we substitute (11), which is strict if $fl = \mathsf{nt}$, and then (14); we get

$$\begin{aligned} L &\leq low(\omega^0) + \sum_{i=1}^{k-1} low(\omega_{\alpha^i}) + \sum_{i=k+1}^{n} low(\omega_{\alpha^i}) + \sum_{i=1}^{k} low(\omega_{\beta^i}) \leq low(\omega^0) - low(\omega_{\beta^0}) + \sum_{i=1}^{n} low(\omega_{\alpha^i}) \leq \\ &\leq low(\omega^0) - low(\omega_{\beta^0}) - low(\omega_\sigma) + x_1 + \sum_{i=r+1}^{n} low(\chi^i) \leq low(\omega^0) + x_1 + \sum_{i=r+1}^{n} low(\chi^i). \end{aligned}$$

If $\beta^0 \cup \{\sigma\} \not\subseteq \mathcal{T}_{\mathsf{tr}}$, the last inequality is strict, as we have removed negative components. Because $low(\omega^0) \leq 1$, if some on the above inequalities was strict, we can remove $low(\omega^0)$, and we get what is needed:

$$L \leq x_1 + \sum_{i=r+1}^{n} low(\chi^i).$$

On the other hand, if none of these inequalities was strict, we have $\alpha^i \cap \beta^i \subseteq \mathcal{T}_{\mathsf{tr}}$ for each $1 \leq i \leq k-1$, and $fl = \mathsf{tr}$, and $\beta^0 \cup \{\sigma\} \subseteq \mathcal{T}_{\mathsf{tr}}$; from Definition B.9 it follows that $\rho \in \mathcal{T}_{\mathsf{tr}}$, so $low(\omega^0) = 0$ and we also get the above inequality.

Next, we have to show the inequality for the $high$ part. Notice that

$$\begin{aligned} &high(\omega^i) = a_i && \text{for } k+1 \leq i \leq n, \text{ and} \\ &high(\omega^k) = b_k. \end{aligned}$$

Assume first that $\rho \in \mathcal{T}_{\mathsf{tr}}$. Then $\beta^0 \cup \{\sigma\} \subseteq \mathcal{T}_{\mathsf{tr}}$; we have $high(\omega^0) = b_0 = high(\omega_\sigma) = 1$. Moreover $\alpha^i \cap \beta^i \subseteq \mathcal{T}_{tr}$ for each $1 \leq i \leq k-1$; we have

$$high(\omega^i) = a_i b_i \qquad \text{for } 1 \leq i \leq k-1.$$

Using (2) we get

$$\begin{aligned} H &= pow(K_n a_n, K_{n-1} a_{n-1}, \dots, K_{k+1} a_{k+1}, b_k, K_{k-1} a_{k-1} b_{k-1}, K_{k-2} a_{k-2} b_{k-2}, \dots, K_1 a_1 b_1, 1) \geq \\ &\geq pow(K_n a_n, K_{n-1} a_{n-1}, \dots, K_{k+1} a_{k+1}, pow(b_k, b_{k-1}, \dots, b_1, 1), K_{k-1} a_{k-1}, K_{k-2} a_{k-2}, \dots, K_1 a_1, 1). \end{aligned}$$

Now we use (12) and then (15) for $K' = K$ (recall that $r \neq k$, so $K_k = 1$), and we get the required inequality

$$H \geq pow(K_n a_n, K_{n-1} a_{n-1}, \dots, K_1 a_1, 1) \geq x_1 + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r+1}), K\big).$$

Next, assume that $\rho \in \mathcal{T}_{\mathsf{nt}}$. Because $\sigma \in type_{z-1}(t^0)$ and $\beta^0 \subseteq type_{z-1}(s^0)$, but $\rho \notin type_{z-1}(s^0)$, we have

$$high(\omega^0) = C_z = (3|\mathcal{T}^0|)^n \cdot (C_{z-1})^{|\mathcal{T}^0|^2+2} \geq 2^{k-1} \cdot |\mathcal{T}^0|^n \cdot (C_{z-1})^{|\beta^0| \cdot |\mathcal{T}^0|+1} \geq 2^{k-1} \cdot |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|} \cdot high(\omega_\sigma).$$

Using (3) we replace 2^{k-1} in the last argument of pow by 2 in the $k-1$ previous arguments; then we observe that for each $1 \leq i \leq k-1$ we have $(high(\omega^i))^2 \geq a_i b_i$; and then we use (2):

$$H \geq pow\Big(K_n high(\omega^n), K_{n-1} high(\omega^{n-1}), \dots, K_k high(\omega^k),$$
$$K_{k-1}(high(\omega^{k-1}))^2, K_{k-2}(high(\omega^{k-2}))^2, \dots, K_1(high(\omega^1))^2, |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|} \cdot high(\omega_\sigma)\Big) \geq$$
$$\geq pow\Big(K_n a_n, K_{n-1}a_{n-1}, \dots, K_{k+1}a_{k+1}, b_k, K_{k-1}a_{k-1}b_{k-1}, K_{k-2}a_{k-2}b_{k-2}, \dots, K_1 a_1 b_1, |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|} \cdot high(\omega_\sigma)\Big) \geq$$
$$\geq pow\Big(K_n a_n, K_{n-1}a_{n-1}, \dots, K_{k+1}a_{k+1}, pow\Big(b_k, b_{k-1}, \dots, b_1, |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|}\Big), K_{k-1}a_{k-1}, K_{k-2}a_{k-2}, \dots, K_1 a_1,$$
$$high(\omega_\sigma)\Big).$$

Next we use (13) and then (15) for $K' = K$ (recall that $r \neq k$, so $K_k = 1$), and we get the required inequality

$$H \geq pow(K_n a_n, K_{n-1}a_{n-1}, \dots, K_1 a_1, high(\omega_\sigma)) \geq x_1 + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{r+1}), K\big).$$

Case 3b: Finally we come to the case that ρ was added to $type_z(s^0)$ by point 3b of Definition B.9. Then in $type_{z-1}(s^0)$ we have a tuple $\tau = (\zeta^n, \zeta^{n-1}, \dots, \zeta^{k+1}, \gamma^k, \gamma^{k-1}, \dots, \gamma^1, q_2, \widehat{\tau}, \cdot)$, where $\widehat{\tau} = (m_2, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, q)$, such that $\gamma^i \subseteq \eta^i \subseteq type(s^i)$ for $1 \leq i \leq k$, and $m_1 m_2 = m$. We have $r_1 = k$ and $r \leq k$. Because S is an k-return, we have $u^k = s^k : s^{k-1} : \dots : s^0$ (Proposition B.2), so $\pi(S(|S|)) = u^n : u^{n-1} : \dots : u^{k+1} : s^k : s^{k-1} : \dots : s^0$. Recall that for $k+1 \leq i \leq n$ we have a decomposition χ^i of (u^i, ζ^i) which has a witness. For $1 \leq i \leq k$, as a part of decomposition ω^i we obtain a decomposition ω_{γ^i} of (s^i, γ^i) which has a witness; denote $c_i = high(\omega_{\gamma^i})$. By the induction assumption the unique decomposition ω_τ of $(s^0, \{\tau\})$ has a witness. This by definition means that we have a run T from $S(|S|)$ which agrees with $(\widehat{\tau}, \chi^n, \chi^{n-1}, \dots, \chi^{k+1}, \omega_{\gamma^k}, \omega_{\gamma^{k-1}}, \dots, \omega_{\gamma^1}, \omega_\tau)$. Let $\pi(T(|T|)) = v^n : v^{n-1} : \dots : v^r$. Let R be the composition of the one-step run from c to d with run S and with run T. We see that R is an r-return (see Lemma B.4), as well as $q = state(R(|R|))$; the word read by R evaluates to $m_1 m_2 = m$ under φ; additionally $\xi^i \subseteq type(v^i)$ for $r+1 \leq i \leq n$. Thus R agrees with $\widehat{\rho}$.

Finally, we have to check the inequalities; this is very similar to the previous case, but now a part corresponding to τ appears. Let x_2 be the number of the $\sharp$ symbols read by T; the number of the $\sharp$ symbols read by R is $x_1 + x_2$. Because T agrees with $(\widehat{\tau}, \chi^n, \chi^{n-1}, \dots, \chi^{k+1}, \omega_{\gamma^k}, \omega_{\gamma^{k-1}}, \dots, \omega_{\gamma^1}, \omega_\tau)$ we have decompositions χ'^i of (v^i, ξ^i) for $r+1 \leq i \leq n$ such that

$$low(\omega_\tau) + \sum_{i=1}^{k} low(\omega_{\gamma^i}) + \sum_{i=k+1}^{n} low(\chi^i) \leq x_2 + \sum_{i=r+1}^{n} low(\chi'^i), \qquad \text{and} \tag{16}$$
$$pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{k+1}), K_k c_k, K_{k-1}c_{k-1}, \dots, K_1 c_1, high(\omega_\tau)\big) \geq$$
$$\geq x_2 + pow\big(high(\chi'^n), high(\chi'^{n-1}), \dots, high(\chi'^{r+1}), K\big). \tag{17}$$

Because $\eta^i = \alpha^i$ for $k+1 \leq i \leq n$, and $\eta^k = \beta^k \cup \gamma^k$, and $\eta^i = \alpha^i \cup \beta^i \cup \gamma^i$ for $1 \leq i \leq k-1$, we get

$$L \leq low(\omega^0) + \sum_{i=1}^{k-1} low(\omega_{\alpha^i}) + \sum_{i=k+1}^{n} low(\omega_{\alpha^i}) + \sum_{i=1}^{k} low(\omega_{\beta^i}) + \sum_{i=1}^{k} low(\omega_{\gamma^i}).$$

To this inequality we substitute (11), (14), and (16); we get

$$L \leq low(\omega^0) - low(\omega_{\beta^0}) - low(\omega_\sigma) - low(\omega_\tau) + x_1 + x_2 + \sum_{i=r+1}^{n} low(\chi'^i) \leq low(\omega^0) + x_1 + x_2 + \sum_{i=r+1}^{n} low(\chi'^i).$$

An analysis like in the previous case shows that either $low(\omega^0) = 0$, or $low(\omega^0) = 1$ and the above inequality is strict; in both cases we can remove $low(\omega^0)$ on the right side and we get what is needed:

$$L \leq x_1 + x_2 + \sum_{i=r+1}^{n} low(\chi'^i).$$

Next, we have to show the inequality for the $high$ part. Assume first that $\rho \in \mathcal{T}_{\mathsf{tr}}$. Then we have $high(\omega^0) = b_0 = high(\omega_\sigma) = high(\omega_\tau) = 1$. Moreover

$$\begin{array}{ll} high(\omega^i) = a_i & \text{for } k+1 \leq i \leq n, \\ high(\omega^i) = b_i c_i & \text{for } i = k, \\ high(\omega^i) = a_i b_i c_i & \text{for } 1 \leq i \leq k-1. \end{array}$$

Using (2) we get

$$H = pow(a_n, a_{n-1}, \dots, a_{k+1}, K_k b_k c_k, K_{k-1} a_{k-1} b_{k-1} c_{k-1}, K_{k-2} a_{k-2} b_{k-2} c_{k-2}, \dots, K_1 a_1 b_1 c_1, 1) \geq$$
$$\geq pow(a_n, a_{n-1}, \dots, a_{k+1}, pow(b_k, b_{k-1}, \dots, b_1, 1) \cdot pow(K_k c_k, K_{k-1} c_{k-1}, \dots, K_1 c_1, 1), a_{k-1}, a_{k-2}, \dots, a_1, 1)\,.$$

Now we use (12), then (15) for $K' = pow(K_k c_k, K_{k-1} c_{k-1}, \dots, K_1 c_1, 1)$, and then (1) and (17), and we get the required inequality

$$H \geq pow(a_n, a_{n-1}, \dots, a_{k+1}, a_k \cdot pow(K_k c_k, K_{k-1} c_{k-1}, \dots, K_1 c_1, 1), a_{k-1}, a_{k-2}, \dots, a_1, 1) \geq$$
$$\geq x_1 + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{k+1}), pow(K_k c_k, K_{k-1} c_{k-1}, \dots, K_1 c_1, 1)\big) \geq$$
$$\geq x_1 + x_2 + pow\big(high(\chi'^n), high(\chi'^{n-1}), \dots, high(\chi'^{r+1}), K\big)\,.$$

Next, assume that $\rho \in \mathcal{T}_{\mathsf{nt}}$. Because $\sigma \in type_{z-1}(t^0)$ and $\beta^0 \cup \{\tau\} \subseteq type_{z-1}(s^0)$, but $\rho \notin type_{z-1}(s^0)$, we have

$$high(\omega^0) = C_z = (3|\mathcal{T}^0|)^n \cdot (C_{z-1})^{|\mathcal{T}^0|^2+2} \geq 2 \cdot 3^{k-1} \cdot |\mathcal{T}^0|^n \cdot (C_{z-1})^{|\beta^0| \cdot |\mathcal{T}^0|+2} \geq$$
$$\geq 2 \cdot 3^{k-1} \cdot |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|} \cdot high(\omega_\sigma) \cdot high(\omega_\tau).$$

Using (3) we replace $2 \cdot 3^{k-1}$ in the last argument of *pow* by 2 or 3 in the k previous arguments; then we observe that for each $1 \leq i \leq k-1$ we have $(high(\omega^i))^3 \geq a_i b_i c_i$, and $(high(\omega^k))^2 \geq b_k c_k$; and then we use (2):

$$H \geq pow\Big(high(\omega^n), high(\omega^{n-1}), \dots, high(\omega^{k+1}), K_k (high(\omega^k))^2,$$
$$K_{k-1}(high(\omega^{k-1}))^3, K_{k-2}(high(\omega^{k-2}))^3, \dots, K_1(high(\omega^1))^3, |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|} \cdot high(\omega_\sigma) \cdot high(\omega_\tau)\Big) \geq$$
$$\geq pow\Big(a_n, a_{n-1}, \dots, a_{k+1}, K_k b_k c_k, K_{k-1} a_{k-1} b_{k-1} c_{k-1}, K_{k-2} a_{k-2} b_{k-2} c_{k-2}, \dots, K_1 a_1 b_1 c_1,$$
$$|\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|} \cdot high(\omega_\sigma) \cdot high(\omega_\tau)\Big) \geq$$
$$\geq pow\Big(a_n, a_{n-1}, \dots, a_{k+1}, pow\big(b_k, b_{k-1}, \dots, b_1, |\mathcal{T}^0|^n \cdot (b_0)^{|\mathcal{T}^0|}\big) \cdot pow\big(K_k c_k, K_{k-1} c_{k-1}, \dots, K_1 c_1, high(\omega_\tau)\big),$$
$$a_{k-1}, a_{k-2}, \dots, a_1, high(\omega_\sigma)\Big).$$

Next we use (13), then (15) for $K' = pow(K_k c_k, K_{k-1} c_{k-1}, \dots, K_1 c_1, high(\omega_\tau))$, and then (1) and (17), we get the required inequality

$$H \geq x_1 + x_2 + pow\big(high(\chi'^n), high(\chi'^{n-1}), \dots, high(\chi'^{r+1}), K\big)\,.$$

■

Corollary B.23. *Every decomposition has a witness.*

Proof: Let ω be a decomposition of (s^k, α^k) for some k-pds s^k and some $\alpha^k \subseteq type(s^k)$, where $0 \leq k \leq n$. We make an induction on k and on the size of a s^k. If $|\alpha^k| \neq 1$, then ω consists of decompositions of $(s^k, \{\sigma\})$ for each $\sigma \in \alpha^k$; we have to show that each of them has a witness. Thus it remains to consider the case $|\alpha^k| = 1$. For $k = 0$ this is Lemma B.22. For empty s^k the thesis is trivial as $type(s^k)$ is empty. Otherwise $s^k = t^k : t^{k-1}$. The decomposition ω contains a composer $(\beta^k, \beta^{k-1}; \alpha^k; \cdot)$ such that $\beta^k \subseteq type(t^k)$ and $\beta^{k-1} \subseteq type(t^{k-1})$. It contains also a decomposition ω^k of (t^k, β^k) and a decomposition ω^{k-1} of (t^{k-1}, β^{k-1}). By induction assumption both ω^k and ω^{k-1} have a witness. Thus, by Proposition B.21, also ω has a witness. ■

C. Types of configurations

Definition B.24. For a configuration c and for $1 \leq k \leq n$ we define $pds^k(c) \in \Gamma^k_*$ as the topmost k-pds of c with its topmost $(k-1)$-pds removed. Additionally $pds^0(c)$ is the topmost 0-pds of c.

In other words, we always have $\pi(c) = pds^n(c) : pds^{n-1}(c) : \cdots : pds^0(c)$.

Definition B.25.
1) Let $\mathcal{T}_{\mathcal{A},\varphi} = \mathcal{P}(\mathcal{T}^n) \times \mathcal{P}(\mathcal{T}^{n-1}) \times \cdots \times \mathcal{P}(\mathcal{T}^0) \times Q$.
2) For a configuration c, let

$$type_{\mathcal{A},\varphi}(c) = \big(type(pds^n(c)), type(pds^{n-1}(c)), \dots, type(pds^0(c)), state(c)\big)\,.$$

3) We say that $(\alpha^n, \alpha^{n-1}, \dots, \alpha^0, p) \leq (\beta^n, \beta^{n-1}, \dots, \beta^0, q)$ if and only if $p = q$ and $\alpha^i \subseteq \beta^i$ for each $0 \leq i \leq n$.

Basing on $type_{\mathcal{A},\varphi}$, for each $0 \leq k \leq n$ we define a function $type^k_{\mathcal{A},\varphi}$ which assigns to every configuration c of $\mathcal{A}$ a pair from $\mathcal{T}_{\mathcal{A},\varphi} \times \Gamma^k_*$, which is $type_{\mathcal{A},\varphi}(c)$, and the topmost k-pds of c. We extend partial order $\leq$ to $\mathcal{T}_{\mathcal{A},\varphi} \times \Gamma^k_*$:

$$(t_1, s^k_1) \leq (t_2, s^k_2) \qquad \Longleftrightarrow \qquad t_1 \leq t_2 \text{ and } s^k_1 = s^k_2.$$

Lemma B.26. *Let $0 \leq k \leq n$. Let R be a run such that $0 \in pre^k_R(|R|)$. Let also $\xi^i \subseteq type(pds^i(R(|R|)))$, for $k+1 \leq i \leq n$. Then there exist $\alpha^i \subseteq type(pds^i(R(0)))$ for $k+1 \leq i \leq n$, and a function $f_R\colon \mathbb{N} \to \mathbb{N}$ such that the following holds. Let c be a configuration such that $state(R(0)) = state(c)$ and the topmost k-pds of $R(0)$ and of c are the same, and $\alpha^i \subseteq type(pds^i(c))$ for $k+1 \leq i \leq n$. Let ω^i be a decomposition of $(pds^i(c), \alpha^i)$, for $k+1 \leq i \leq n$. Then there exists a run S from c such that*

1) *if $|R| > 0$ then $|S| > 0$, and*
2) *$0 \in pre^k_S(|S|)$, and*
3) *the words read by R and by S evaluate to the same under φ, and*
4) *$state(R(|R|)) = state(S(|S|))$, and $\xi^i \subseteq type(pds^i(S(|S|)))$ for $k+1 \leq i \leq n$, and the topmost k-pds's of $R(|R|)$ and of $S(|S|)$ are the same, and*
5) *there exist decompositions χ^i of $(pds^i(S(|S|)), \xi^i)$, for $k+1 \leq i \leq n$, such that*

$$\sum_{i=k+1}^{n} low(\omega^i) \leq x + \sum_{i=k+1}^{n} low(\chi^i), \qquad \textit{and}$$
$$f_R\big(pow\big(high(\omega^n), high(\omega^{n-1}), \dots, high(\omega^{k+1})\big)\big) \geq x + pow\big(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{k+1})\big)\,,$$

where x is the number of the $\sharp$ symbols read by S.

Theorem 4.1 follows immediately from the above lemma. It is enough to take, for $k+1 \leq i \leq n$, $\xi^i = type(pds^i(R(|R|)))$, and any decomposition ω^i. Then $type^k_{\mathcal{A},\varphi}(R(0)) \leq type^k_{\mathcal{A},\varphi}(c)$ implies that $state(R(0)) = state(c)$ and the topmost k-pds of $R(0)$ and of c are the same, and $\alpha^i \subseteq type(pds^i(c))$ for $k+1 \leq i \leq n$. On the other hand condition 3 implies that $type^k_{\mathcal{A},\varphi}(R(|R|)) \leq type^k_{\mathcal{A},\varphi}(S(|S|))$.

Proof: At the beginning notice that the first operation done from $R(0)$ and from c is the same, since $state(R(0)) = state(c)$ and the topmost symbol (and the whole topmost k-pds) of $R(0)$ and c are the same. We make an induction on the length of R. Lemma B.3 gives us four possible forms of R; we analyze these cases.

Case 1: Assume first that $|R| = 0$. We take $\alpha^i = \xi^i$ for $k+1 \leq i \leq n$, and identity function f_R. Given a configuration c, and decompositions ω^i for $k+1 \leq i \leq n$ as S we take the empty run from c, and we take $\chi^i = \omega^i$ for $k+1 \leq i \leq n$. The thesis is satisfied trivially

Case 2a: Next, assume that $|R| = 1$ and the operation performed in R is read, or push^r for $r \leq k$, or pop^r for $r \leq k$. We take $\alpha^i = \xi^i$ (for $k+1 \leq i \leq n$), and $f_R(N) = N+1$. Given a configuration c, and decompositions ω^i for $k+1 \leq i \leq n$, as S we take the one-step run from c (if the operation is read, we read the same letter as in R). It is important that between $R(0)$ and $R(1)$, and between $S(0)$ and $S(1)$ only the topmost k-pds is modified. In particular the topmost k-pds of $R(1)$ and of $S(1)$ are the same. The first four conditions are immediate. In the last condition we take $\chi^i = \omega^i$; notice that $x \leq 1$.

Case 2b: Next, assume that $|R| = 1$ and the operation performed in R is push^r for $r \geq k+1$. Observe that $pds^i(R(1)) = pds^i(R(0))$ for $1 \leq i \leq r-1$ and for $r+1 \leq i \leq n$, and

$$pds^r(R(1)) = pds^r(R(0)) : pds^{r-1}(R(0)) : \dots : pds^0(R(0)).$$

Choose some composer $(\beta^r, \beta^{r-1}, \dots, \beta^0; \xi^r; \cdot)$ such that $\beta^i \subseteq type(pds^i(R(0)))$ for $0 \leq i \leq r$; it exists by Proposition B.11. We take

$$\alpha^i = \begin{cases} \xi^i & \text{for } r+1 \leq i \leq n, \\ \beta^i & \text{for } i = r, \\ \xi^i \cup \beta^i & \text{for } k+1 \leq i \leq r-1. \end{cases}$$

For $0 \leq i \leq k$ fix some decomposition ω^i of $(pds^i(R(0)), type(pds^i(R(0))))$; of course it exists. As f_R we take a function such that for every positive integers $a_{k+1}, a_{k+2}, \dots, a_n$ we have

$$f_R(pow(a_n, a_{n-1}, \dots, a_{k+1})) \geq pow\Big(a_n^2, a_{n-1}^2, \dots, a_{k+1}^2, high(\omega^k), high(\omega^{k-1}), \dots, high(\omega^1), |\mathcal{T}^0|^n \cdot high(\omega^0)^{|\mathcal{T}^0|}\Big)\,.$$

Such function exists, as for each N there are only finitely many combinations of arguments $a_{k+1}, a_{k+2}, \dots, a_n$ such that $pow(a_n, a_{n-1}, \dots, a_{k+1}) = N$ (in particular no argument can be greater than N).

To show the thesis take any c such that $state(R(0)) = state(c)$ and the topmost k-pds of $R(0)$ and of c are the same, and $\alpha^i \subseteq type(pds^i(c))$ for $k+1 \leq i \leq n$. Take also some decomposition ω^i of $(pds^i(c), \alpha^i)$, for $k+1 \leq i \leq n$. As S we take the one-step run from c. Like for R, we have $pds^i(S(1)) = pds^i(c)$ for $1 \leq i \leq r-1$ and for $r+1 \leq i \leq n$, and

$$pds^r(S(1)) = pds^r(c) : pds^{r-1}(c) : \dots : pds^0(c).$$

The first three conditions are immediate. In condition 4 the only nontrivial part is that $\xi^r \subseteq \mathit{type}(\mathit{pds}^r(S(1)))$. But $\beta^i \subseteq \mathit{type}(\mathit{pds}^i(c))$ for $0 \le i \le r$ (for $0 \le i \le k$ this is true because $\beta^i \subseteq \mathit{type}(\mathit{pds}^i(R(0)))$ and $\mathit{pds}^i(R(0)) = \mathit{pds}^i(c)$, and for $k+1 \le i \le r$ because $\beta^i \subseteq \alpha^i \subseteq \mathit{type}(\mathit{pds}^i(c))$). Using Proposition B.11 for composer $(\beta^r, \beta^{r-1}, \dots, \beta^0; \xi^r; \cdot)$ we get that $\xi^r \subseteq \mathit{type}(\mathit{pds}^r(S(1)))$.

Finally, we prove condition 5. For $k+1 \le i \le r-1$, as a part of ω^i we can take a decomposition χ^i of $(\mathit{pds}^i(c), \xi^i)$, and a decomposition ω_{β^i} of $(\mathit{pds}^i(c), \beta^i)$; they satisfy the inequalities

$$\mathit{low}(\omega^i) \le \mathit{low}(\chi^i) + \mathit{low}(\omega_{\beta^i}), \quad \text{and} \quad \mathit{high}(\omega^i) \ge \mathit{high}(\chi^i), \quad \text{and} \quad \mathit{high}(\omega^i) \ge \mathit{high}(\omega_{\beta^i}).$$

Denote also $\omega_{\beta^r} = \omega^r$, and $\chi^i = \omega^i$ for $r+1 \le i \le n$. Moreover for $0 \le i \le k$ as a part of the fixed decomposition ω^i we obtain a decomposition ω_{β^i} of $(\mathit{pds}^i(c), \beta^i)$ (recall that $\mathit{pds}^i(R(0)) = \mathit{pds}^i(c)$ for $0 \le i \le k$), such that $\mathit{high}(\omega^i) \ge \mathit{high}(\omega_{\beta^i})$. By Proposition B.21 we obtain a decomposition χ^r of $(\mathit{pds}^r(S(1)), \xi^r)$ such that

$$\sum_{i=0}^{r} \mathit{low}(\omega_{\beta^i}) \le \mathit{low}(\chi^r),$$
$$\mathit{pow}\Big(\mathit{high}(\omega_{\beta^r}), \mathit{high}(\omega_{\beta^{r-1}}), \dots, \mathit{high}(\omega_{\beta^1}), |\mathcal{T}^0|^n \cdot \mathit{high}(\omega_{\beta^0})^{|\mathcal{T}^0|}\Big) \ge \mathit{high}(\chi^r). \tag{18}$$

We get the required inequality for low by composing the inequalities mentioned above:

$$\sum_{i=k+1}^{n} \mathit{low}(\omega^i) \le \sum_{i=k+1}^{r-1} (\mathit{low}(\chi^i) + \mathit{low}(\omega_{\beta^i})) + \mathit{low}(\omega_{\beta^r}) + \sum_{i=r+1}^{n} \mathit{low}(\chi^i) \le$$
$$\le \sum_{i=k+1}^{r-1} \mathit{low}(\chi^i) + \sum_{i=r+1}^{n} \mathit{low}(\chi^i) + \sum_{i=0}^{r} \mathit{low}(\omega_{\beta^i}) \le \sum_{i=k+1}^{n} \mathit{low}(\chi^i) \le x + \sum_{i=k+1}^{n} \mathit{low}(\chi^i).$$

For the second inequality, we first use the definition of f_R and (2):

$$f_R\big(\mathit{pow}\big(\mathit{high}(\omega^n), \mathit{high}(\omega^{n-1}), \dots, \mathit{high}(\omega^{k+1})\big)\big) \ge$$
$$\ge \mathit{pow}\Big((\mathit{high}(\omega^n))^2, (\mathit{high}(\omega^{n-1}))^2, \dots, (\mathit{high}(\omega^{k+1}))^2, \mathit{high}(\omega^k), \mathit{high}(\omega^{k-1}), \dots, \mathit{high}(\omega^1), |\mathcal{T}^0|^n \cdot \mathit{high}(\omega^0)^{|\mathcal{T}^0|}\Big) \ge$$
$$\ge \mathit{pow}\Big(\mathit{high}(\omega^n), \mathit{high}(\omega^{n-1}), \dots, \mathit{high}(\omega^{r+1}), \mathit{pow}\Big(\mathit{high}(\omega^r), \mathit{high}(\omega^{r-1}), \dots, \mathit{high}(\omega^1), |\mathcal{T}^0|^n \cdot \mathit{high}(\omega^0)^{|\mathcal{T}^0|}\Big),$$
$$\mathit{high}(\omega^{r-1}), \mathit{high}(\omega^{r-2}), \dots, \mathit{high}(\omega^{k+1}), 1, 1, \dots, 1\Big).$$

Next, in the internal pow we replace $\mathit{high}(\omega^i)$ by $\mathit{high}(\omega_{\beta^i})$ (which is not greater), and we use (18). In the external pow we replace $\mathit{high}(\omega^i)$ by $\mathit{high}(\chi^i)$ (which is not greater). We can also remove the final ones, as they do not change the result of pow. Finally, we observe that x (the number of the $\sharp$ symbols read) is 0. We get the required inequality

$$f_R\big(\mathit{pow}\big(\mathit{high}(\omega^n), \mathit{high}(\omega^{n-1}), \dots, \mathit{high}(\omega^{k+1})\big)\big) \ge x + \mathit{pow}\big(\mathit{high}(\chi^n), \mathit{high}(\chi^{n-1}), \dots, \mathit{high}(\chi^{k+1})\big).$$

Case 3: Next, assume that the first operation in R is push^r, and the subrun R' of R from 1 to $|R|$ is an r-return, where $r \ge k+1$. From Lemma B.7 we get that in $\mathit{type}(\mathit{pds}^0(R(1)))$ we have a tuple $\sigma = (\gamma^n, \gamma^{n-1}, \dots, \gamma^1, \mathit{state}(R(1)), \widehat{\sigma}, \cdot)$, where $\widehat{\sigma} = (m, r, \xi^n, \xi^{n-1}, \dots, \xi^{r+1}, \mathit{state}(R(|R|)))$ such that $\gamma^i \subseteq \mathit{type}(\mathit{pds}^i(R(1)))$ for $1 \le i \le n$, and m is the image under φ of the word read by R (equivalently: by R'). Observe that $\mathit{pds}^i(R(1)) = \mathit{pds}^i(R(0))$ for $1 \le i \le r-1$ and for $r+1 \le i \le n$, and

$$\mathit{pds}^r(R(1)) = \mathit{pds}^r(R(0)) : \mathit{pds}^{r-1}(R(0)) : \dots : \mathit{pds}^0(R(0)).$$

Choose some composer $(\beta^r, \beta^{r-1}, \dots, \beta^0; \gamma^r; \cdot)$ such that $\beta^i \subseteq \mathit{type}(\mathit{pds}^i(R(0)))$ for $0 \le i \le r$; it exists by Proposition B.11. We take

$$\alpha^i = \begin{cases} \gamma^i & \text{for } r+1 \le i \le n, \\ \beta^i \cup \xi^i & \text{for } i = r, \\ \gamma^i \cup \beta^i \cup \xi^i & \text{for } k+1 \le i \le r-1. \end{cases}$$

For $0 \le i \le k$ fix some decomposition ω^i of $(\mathit{pds}^i(R(0)), \mathit{type}(\mathit{pds}^i(R(0))))$; of course it exists. Let also ω_σ be the unique decomposition of $(\mathit{pds}^0(R(1)), \{\sigma\})$. As f_R we take a function such that for every positive integers $a_{k+1}, a_{k+2}, \dots, a_n$ we have (as previously, such function exists)

$$f_R(\mathit{pow}(a_n, a_{n-1}, \dots, a_{k+1})) \ge \mathit{pow}\Big(a_n^3, a_{n-1}^3, \dots, a_{k+1}^3, (\mathit{high}(\omega^k))^2, (\mathit{high}(\omega^{k-1}))^2, \dots, (\mathit{high}(\omega^1))^2,$$
$$|\mathcal{T}^0|^n \cdot \mathit{high}(\omega^0)^{|\mathcal{T}^0|} \cdot \mathit{high}(\omega_\sigma)\Big).$$

To show the thesis take any c such that $state(R(0)) = state(c)$ and the topmost k-pds of $R(0)$ and of c are the same, and $\alpha^i \subseteq type(pds^i(c))$ for $k+1 \leq i \leq n$. Take also some decomposition ω^i of $(pds^i(c), \alpha^i)$, for $k+1 \leq i \leq n$. Let d be the configuration one step after c. Like for $R(0)$ and $R(1)$, we have $pds^i(d) = pds^i(c)$ for $1 \leq i \leq r-1$ and for $r+1 \leq i \leq n$, and

$$pds^r(d) = pds^r(c) : pds^{r-1}(c) : \cdots : pds^0(c).$$

For $k+1 \leq i \leq r-1$, as a part of ω^i we can take a decomposition ω_{γ^i} of $(pds^i(c), \gamma^i)$; and for $k+1 \leq i \leq r$ a decomposition ω_{β^i} of $(pds^i(c), \beta^i)$ and a decomposition χ^i of $(pds^i(c), \xi^i)$. Denote also $\omega_{\gamma^i} = \omega^i$ for $r+1 \leq i \leq n$. We have

$$\begin{aligned}
&low(\omega^i) \leq low(\omega_{\gamma^i}) + low(\omega_{\beta^i}) + low(\chi^i) \qquad \text{for } k+1 \leq i \leq r-1,\\
&low(\omega^r) \leq low(\omega_{\beta^r}) + low(\chi^r),\\
&high(\omega^i) \geq high(\omega_{\gamma^i}) \qquad \text{for } k+1 \leq i \leq r-1 \text{ and for } r+1 \leq i \leq n,\\
&high(\omega^i) \geq high(\omega_{\beta^i}) \quad \text{and} \quad high(\omega^i) \geq high(\chi^i) \qquad \text{for } k+1 \leq i \leq r.
\end{aligned}$$

Moreover for $0 \leq i \leq k$ as a part of the fixed decomposition ω^i we obtain a decomposition ω_{β^i} of $(pds^i(c), \beta^i)$ and a decomposition ω_{γ^i} of $(pds^i(c), \gamma^i)$ (recall that $pds^i(c) = pds^i(R(0))$ for $0 \leq i \leq k$). By Proposition B.21 we obtain a decomposition ω_{γ^r} of $(pds^r(d), \gamma^r)$ such that

$$\begin{aligned}
&\sum_{i=0}^{r} low(\omega_{\beta^i}) \leq low(\omega_{\gamma^r}),\\
&pow\Big(high(\omega_{\beta^r}), high(\omega_{\beta^{r-1}}), \dots, high(\omega_{\beta^1}), |\mathcal{T}^0|^n \cdot high(\omega_{\beta^0})^{|\mathcal{T}^0|}\Big) \geq high(\omega_{\gamma^r}). \qquad (19)
\end{aligned}$$

From Corollary B.23 we know that every decomposition has a witness. By the definition of witness, there exists a run S' from d which agrees with $(\widehat{\sigma}, \omega_{\gamma^n}, \omega_{\gamma^{n-1}}, \dots, \omega_{\gamma^1}, \omega_\sigma)$. Let S be the composition of the one-step run from c to d with run S'. Run S' agrees with $\widehat{\sigma}$, which means that $state(R(|R|)) = state(S(|S|))$, and $\xi^i \subseteq type(pds^i(R(|R|)))$ for $r+1 \leq i \leq n$, and m is the image under φ of the word read by S (by S'). It also means that S' is an r-return, so $0 \in pre^k_S(|S|)$ (Lemma B.3). Additionally (Proposition B.2), the topmost r-pds of c and of $S(|S|)$ is the same; similarly the topmost r-pds of $R(0)$ and of $R(|R|)$ is the same. It follows that the topmost k-pds of $R(|R|)$ and $S(|S|)$ are the same, and $\xi^i \subseteq \alpha^i \subseteq type(pds^i(c)) = type(pds^i(S(|S|)))$ for $k+1 \leq i \leq r$. This way we get the first four conditions.

Finally, we prove condition 5. We compose the inequalities mentioned above (and we add some new non-negative components):

$$\begin{aligned}
\sum_{i=k+1}^{n} low(\omega^i) &\leq \sum_{i=k+1}^{r-1} \Big(low(\omega_{\gamma^i}) + low(\omega_{\beta^i}) + low(\chi^i)\Big) + \Big(low(\omega_{\beta^r}) + low(\chi^r)\Big) + \sum_{i=r+1}^{n} low(\omega_{\gamma^i}) \leq\\
&\leq \sum_{i=1}^{r-1} low(\omega_{\gamma^i}) + \sum_{i=r+1}^{n} low(\omega_{\gamma^i}) + \sum_{i=0}^{r} low(\omega_{\beta^i}) + \sum_{i=k+1}^{r} low(\chi^i) \leq low(\omega_\sigma) + \sum_{i=1}^{n} low(\omega_{\gamma^i}) + \sum_{i=k+1}^{r} low(\chi^i).
\end{aligned}$$

Because S' agrees with $(\widehat{\sigma}, \omega_{\gamma^n}, \omega_{\gamma^{n-1}}, \dots, \omega_{\gamma^1}, \omega_\sigma)$, there exist decompositions χ^i of $(pds^i(S(|S|)), \xi^i)$ for $r+1 \leq i \leq n$ such that we have

$$low(\omega_\sigma) + \sum_{i=1}^{n} low(\omega_{\gamma^i}) \leq x + \sum_{i=r+1}^{n} low(\chi^i).$$

By combining the last two inequalities, we get the required inequality about low.

For the second inequality, we first use the definition of f_R and and twice (2):

$$\begin{aligned}
&f_R\big(pow\big(high(\omega^n), high(\omega^{n-1}), \dots, high(\omega^{k+1})\big)\big) \geq\\
&\quad\geq pow\Big((high(\omega^n))^3, (high(\omega^{n-1}))^3, \dots, (high(\omega^{k+1}))^3,\\
&\qquad (high(\omega^k))^2, (high(\omega^{k-1}))^2, \dots, (high(\omega^1))^2, |\mathcal{T}^0|^n \cdot high(\omega^0)^{|\mathcal{T}^0|} \cdot high(\omega_\sigma)\Big) \geq\\
&\quad\geq pow\Big(high(\omega^n), high(\omega^{n-1}), \dots, high(\omega^{r+1}), pow\Big(high(\omega^r), high(\omega^{r-1}), \dots, high(\omega^1), |\mathcal{T}^0|^n \cdot high(\omega^0)^{|\mathcal{T}^0|}\Big) \cdot\\
&\qquad \cdot pow\Big(high(\omega^r), high(\omega^{r-1}), \dots, high(\omega^{k+1}), 1, 1, \dots, 1\Big), high(\omega^{r-1}), high(\omega^{r-2}), \dots, high(\omega^1), high(\omega_\sigma)\Big).
\end{aligned}$$

Next, in the first internal pow we replace $high(\omega^i)$ by $high(\omega_{\beta^i})$ (which is not greater), and we use (19). In the second internal pow we replace $high(\omega^i)$ by $high(\chi^i)$ (which is not greater). We can also remove the final ones, as they do not change the result of pow. In the external pow we replace $high(\omega^i)$ by $high(\omega_{\gamma^i})$ (which is not greater). We get

$$f_R\big(pow\big(high(\omega^n), high(\omega^{n-1}), \dots, high(\omega^{k+1})\big)\big) \geq$$
$$\geq pow\Big(high(\omega_{\gamma^n}), high(\omega_{\gamma^{n-1}}), \dots, high(\omega_{\gamma^{r+1}}), high(\omega_{\gamma^r}) \cdot pow\big(high(\chi^r), high(\chi^{r-1}), \dots, high(\chi^{k+1})\big),$$
$$high(\omega_{\gamma^{r-1}}), high(\omega_{\gamma^{r-2}}), \dots, high(\omega_{\gamma^1}), high(\omega_\sigma)\Big).$$

Finally, we use the inequality from Definition B.20 for $K = pow\big(high(\chi^r), high(\chi^r), \dots, high(\chi^{k+1})\big)$, and (1), and we get the required inequality.

Case 4: Assume that R is a composition of shorter runs R_1 and R_2 such that $0 \in pre^k_{R_1}(|R_1|)$ and $0 \in pre^k_{R_2}(|R_2|)$. We use the induction assumption for R_2 (as R), and for $\xi^{k+1}, \xi^{k+2}, \dots, \xi^n$. We get $\gamma^i \subseteq type(pds^i(R_2(0)))$ for $k+1 \leq i \leq n$, and a function f_2. Then we use the induction assumption for R_1 (as R), and for $\gamma^{k+1}, \gamma^{k+2}, \dots, \gamma^n$ (as $\xi^{k+1}, \xi^{k+2}, \dots, \xi^n$). We get $\alpha^i \subseteq type(pds^i(R(0)))$ for $k+1 \leq i \leq n$, and a function f_1. We define

$$f_R(N) = f_1(N) + \max_{1 \leq M \leq f_1(N)} f_2(M).$$

To show the thesis take any c such that $state(R(0)) = state(c)$ and the topmost k-pds of $R(0)$ and of c are the same, and $\alpha^i \subseteq type(pds^i(c))$ for $k+1 \leq i \leq n$. Take also some decomposition ω^i of $(pds^i(c), \alpha^i)$, for $k+1 \leq i \leq n$. From the induction assumption for R_1 we obtain a run S_1 from c, and (from condition 5) decompositions ω_{γ^i} of $(pds^i(S_1(|S_1|)), \gamma^i)$ for $k+1 \leq i \leq n$. By condition 4, $S_1(|S_1|)$ can be used as c in the induction assumption for R_2. We obtain a run S_2 from $S_1(|S_1|)$ and (from condition 5) decompositions χ^i of $(pds^i(S_2(|S_2|)), \xi^i)$ for $k+1 \leq i \leq n$. As S we take the composition of S_1 and S_2. The first four conditions of the induction hypothesis follow trivially, as well as the inequality about low in condition 5. To see the second inequality, denote

$$A = pow(high(\omega^n), high(\omega^{n-1}), \dots, high(\omega^{k+1})),$$
$$B = pow(high(\omega_{\gamma^n}), high(\omega_{\gamma^{n-1}}), \dots, high(\omega_{\gamma^{k+1}})),$$
$$C = pow(high(\chi^n), high(\chi^{n-1}), \dots, high(\chi^{k+1})).$$

Let x_1 and x_2 be the number of the $\sharp$ symbols read by S_1, and by S_2, respectively. The induction assumptions give us the inequalities

$$f_1(A) \geq x_1 + B \qquad \text{and} \qquad f_2(B) \geq x_2 + C.$$

We get the required inequality:

$$f_R(A) \geq x_1 + B + \max_{1 \leq M \leq x_1 + B} f_2(M) \geq x_1 + f_2(B) \geq x_1 + x_2 + C.$$

■

D. Sequence equivalence

Definition B.27. Let $0 \leq k \leq n$, and let $(s^k_i)_{i=1}^\infty$ be a sequence of k-pds's. We define $stype\big((s^k_i)_{i=1}^\infty\big) \subseteq \mathcal{T}^k$ to be the set of those $\sigma \in \mathcal{T}^k$ for which $\sigma \in type(s^k_i)$ for each i and there exists a decomposition ω_i of $(s^k_i, \{\sigma\})$ such that the sequence $(low(\omega_i))_{i=1}^\infty$ is bounded.

Definition B.28. Let $(c_i)_{i=1}^\infty$ and $(d_i)_{i=1}^\infty$ be sequences of configurations. We say that these sequences are $(\mathcal{A}, \varphi)$-sequence equivalent if for each $0 \leq k \leq n$ we have

$$stype\big((pds^k(c_i))_{i=1}^\infty\big) = stype\big((pds^k(d_i))_{i=1}^\infty\big).$$

Proof of Theorem 4.3: Because between $R(l_{r-1})$ and $R(|R|)$ there are no push^n operations and there is one pop^n operation, which is the last one, we know that the subrun of R from l_{r-1} to $|R|$ is an n-return. From Lemma B.7 we get that in $type(pds^0(R(l_{r-1})))$ we have a tuple $\sigma = (\alpha^n_{r-1}, \alpha^{n-1}_{r-1}, \dots, \alpha^1_{r-1}, state(R(l_{r-1})), \widehat{\sigma}, \cdot)$ where $\widehat{\sigma} = (m, n-1, state(R(|R|)),$ such that $\alpha^s_{r-1} \subseteq type(pds^s(R(l_{r-1})))$ for $1 \leq s \leq n$, and m is the image under φ of the word read between $R(l_{r-1})$ and $R(|R|)$. Denote also $\alpha^0_{r-1} = \{\sigma\}$. For $j = r-1, r-2, \dots, 1$, let $\alpha^{k+1}_{j-1}, \alpha^{k+2}_{j-1}, \dots, \alpha^n_{j-1}$ be the sets $\alpha^{k+1}, \alpha^{k+2}, \dots, \alpha^n$ obtained from Lemma B.26 used for the subrun of R from l_{j-1} to l_j (as R), and $\alpha^{k+1}_j, \alpha^{k+2}_j, \dots, \alpha^n_j$ (as $\xi^{k+1}, \xi^{k+2}, \dots, \xi^n$); let also f_j be the function from this lemma.

Case 1: Assume first that $\alpha_0^s \subseteq stype((pds^s(c_i))_{i=0}^\infty)$ for $k+1 \le s \le n$. Then we have decompositions $\omega_{0,i}^s$ of $(pds^s(c_i), \alpha_0^s)$ for $k+1 \le s \le n$ and each i, such that $(low(\omega_{0,i}^s))_{i=0}^\infty$ is bounded (recall that $\omega_{0,i}^s$ is a collection of decompositions of $(pds^s(c_i), \{\tau\})$ for each $\tau \in \alpha_0^s$, and $low(\omega_{0,i}^s)$ is the sum of low for these decompositions). From Proposition B.19 we get that also $(high(\omega_{0,i}^s))_{i=0}^\infty$ is bounded.

For each i, from Lemma B.26 we obtain, for $j = 1, 2, \ldots, r-1$, a run $S_{j,i}$, and decompositions $\omega_{j,i}^s$ of $\big(pds^s(S_{j,i}(|S_{j,i}|), \alpha_j^s\big)$, for $k+1 \le s \le n$; run $S_{1,i}$ is from c_i, and runs $S_{j,i}$ for $j \ge 2$ are from $S_{j-1,i}(|S_{j-1,i}|)$. Notice that the topmost k-pds of $S_{r-1,i}(|S_{r-1,i}|)$ and of $R(l_{r-1})$ are the same. Fix some decomposition ω_{r-1}^s of $\big(pds^s(R(l_{r-1})), \alpha_{r-1}^s\big)$ for $0 \le s \le k$, and denote $\omega_{r-1,i}^s = \omega_{r-1}^s$ for each i. Because every decomposition has a witness (Corollary B.23), we have a run $S_{r,i}$ from $S_{r-1,i}(|S_{r-1,i}|)$ which agrees with $(\widehat{\sigma}, \omega_{r-1,i}^n, \omega_{r-1,i}^{n-1}, \ldots, \omega_{r-1,i}^0)$. As S_i we take the composition of runs $S_{j,i}$ for $j = 1, 2, \ldots, r$. Condition 1 follows trivially, since $S_{r,i}$ is an n-return. We see that the words read by $S_{j,i}$ and by the subrun of R from l_{j-1} to l_j evaluate to the same under φ, for $1 \le j \le r$ and each i (we get condition 2).

Let $x_{j,i}$ (for $1 \le j \le r$ and each i) be the number of the $\sharp$ symbols read by $S_{j,i}$. From condition 5 of Lemma B.26 we have that

$$f_j\big(pow\big(high(\omega_{j-1,i}^n), high(\omega_{j-1,i}^{n-1}), \ldots, high(\omega_{j-1,i}^{k+1})\big)\big) \ge x_{j,i} + pow\big(high(\omega_{j,i}^n), high(\omega_{j,i}^{n-1}), \ldots, high(\omega_{j,i}^{k+1})\big)\,.$$

By assumption $(high(\omega_{0,i}^s))_{i=0}^\infty$ is bounded for $k+1 \le s \le n$. By induction on j (for $1 \le j \le r-1$), since $(high(\omega_{j-1,i}^s))_{i=0}^\infty$ is bounded for $k+1 \le s \le n$, then also $(x_{j,i})_{i=0}^\infty$ and $(high(\omega_{j,i}^s))_{i=0}^\infty$ for $k+1 \le s \le n$ are bounded. From Definition B.20 (we take $K = 1$) we have inequality

$$pow\big(high(\omega_{r-1,i}^n), high(\omega_{r-1,i}^{n-1}), \ldots, high(\omega_{r-1,i}^0)\big) \ge x_{r,i} + 1.$$

Recall that for $0 \le s \le k$ the value of $high(\omega_{r-1,i}^s)$ is the same for each i. Thus also $(x_{r,i})_{i=0}^\infty$ is bounded. As $x_i = \sum_{j=1}^r x_{j,i}$, also $(x_i)_{i=0}^\infty$ is bounded.

Exactly the same we can do for $(d_i)_{i=0}^\infty$, obtaining runs such that $(y_i)_{i=0}^\infty$ is bounded (recall that $stype((pds^s(c_i))_{i=0}^\infty) = stype((pds^s(d_i))_{i=0}^\infty)$ for $k+1 \le s \le n$).

Case 2: This is the opposite case: we assume that for some $k+1 \le s \le n$ we have $\alpha_0^s \not\subseteq stype((pds^s(c_i))_{i=0}^\infty)$. Take any decompositions $\omega_{0,i}^s$ of $(pds^s(c_i), \alpha_0^s)$ for $k+1 \le s \le n$ and each i. Notice that if $(low(\omega_{0,i}^s))_{i=0}^\infty$ is bounded then $\alpha_0^s \subseteq stype\left((pds^s(c_i))_{i=0}^\infty\right)$; thus $(\sum_{s=k+1}^n low(\omega_{0,i}^s))_{i=0}^\infty$ is unbounded. We construct the runs exactly in the same way as previously, but now we use the opposite inequalities. From condition 5 of Lemma B.26 we have that (for $1 \le j \le r-1$ and each i)

$$\sum_{s=k+1}^{n} low(\omega_{j-1,i}^s) \le x_{j,i} + \sum_{s=k+1}^{n} low(\omega_{j,i}^s).$$

From Definition B.20 we have inequality (for each i)

$$\sum_{s=k+1}^{n} low(\omega_{r-1,i}^s) \le \sum_{s=0}^{n} low(\omega_{r-1,i}^s) \le x_{r,i}.$$

By summing theses inequalities we get that for each i,

$$\sum_{s=k+1}^{n} low(\omega_{0,i}^s) \le x_i,$$

thus $(x_i)_{i=0}^\infty$ is unbounded. In the same way we create runs from $(d_i)_{i=0}^\infty$, for which we have that $(y_i)_{i=0}^\infty$ is unbounded. ■