Weak Bisimulation Finiteness of Pushdown Systems With Deterministic ε-Transitions Is 2-ExpTime-Complete

Stefan Göller[*]   Paweł Parys[†]

Abstract

We consider the problem of deciding whether a given pushdown system all of whose ε-transitions are deterministic is weakly bisimulation finite, that is, whether it is weakly bisimulation equivalent to a finite system. We prove that this problem is 2-ExpTime-complete. This consists of three elements: First, we prove that the smallest finite system that is weakly bisimulation equivalent to a fixed pushdown system, if exists, has size at most doubly exponential in the description size of the pushdown system. Second, we propose a fast algorithm deciding whether a given pushdown system is weakly bisimulation equivalent to a finite system of a given size. Third, we prove 2-ExpTime-hardness of the problem. The problem was known to be decidable, but the previous algorithm had Ackermannian complexity (6-ExpSpace in the easier case of pushdown systems without ε-transitions); concerning lower bounds, only ExpTime-hardness was known.

1 Introduction

An important decision problem in computer science is to decide whether a given infinite system is semantically finite, that is, whether it is semantically equivalent to some finite system. If so, particular techniques and properties of finite systems can be exploited in order to verify such systems. There are different important notions of equivalences that have been studied by the computer science community.

In the area of verification *bisimulation* equivalence [20] can be seen as the central one. It can be seen as a two-player game between Attacker and Defender: given a pair of configurations (c, d) of a system, Attacker chooses a transition $c \rightarrow_a c'$ (resp. $d \rightarrow_a d'$) and Defender must find a reply $d \rightarrow_a d'$ (resp. $c \rightarrow_a c'$), hereby leading the game to a new pair of configurations (c', d') — Attacker wins in case Defender cannot answer, whereas Defender wins every infinite play and every play terminating in a pair of dead ends. It is worth pointing out that trace equivalence coincides with bisimulation equivalence (bisimilarity for short) in case the underlying systems are deterministic. Several central verification logics like modal logic, the modal μ-calculus, CTL*, and Propositional Dynamic Logic can all be characterized as the bisimulation-invariant fragment of well-established logics like first-order logic [1], monadic second-order logic [11], monadic chain logic [15], and weak monadic chain logic [3], respectively. In presence of possible ε-transitions *weak bisimulation equivalence* generalizes bisimulation equivalence [4] in that Attacker can make moves of the form $c \rightarrow_\varepsilon^* \circ \rightarrow_a \circ \rightarrow_\varepsilon^* c'$ (or $c \rightarrow_\varepsilon^* c'$) that can be answered by Defender by moves of the form $d \rightarrow_\varepsilon^* \circ \rightarrow_a \circ \rightarrow_\varepsilon^* d'$ (resp. $d \rightarrow_\varepsilon^* d'$). In the context of (weak) bisimulation equivalence, the semantic finiteness question becomes the *(weak) bisimulation finiteness* problem: given an infinite system, does there exist a *finite* system that is (weakly) bisimilar to it: it is important to emphasize that it is asked whether *there exists* a finite system that is (weakly) bisimilar to the infinite system, so the finite system is not specified in the input to the problem. Indeed, in case both the infinite system and the finite system are part of the input, their equivalence can be reduced to the model checking problem of CTL's fragment EF [14]. The (complexity of the) model checking problem for most classes of infinite systems is well-understood.

To date, it is fair to say that decidability — and in particular the complexity — of this principal and easily-stated problem of (weak) bisimulation finiteness is not well understood. We refer to Srba's survey [19], where it becomes clear that for many classes of infinite state systems decidability is unknown, and if it is known, oftentimes huge complexity gaps exist. A central such class is the class of pushdown systems, that is, systems that can be generated by pushdown automata. Model checking monadic second-order logic is decidable for them [16], whereas

its bisimulation-invariant fragment can be model-checked in ExpTime [22]. Only a few years ago Jančar proved that weak bisimulation finiteness of pushdown systems all of whose ε-transitions are deterministic, is decidable with Ackermannian complexity [7] via a reduction to the Ackermann-complete (weak) bisimulation equivalence problem [9, 23]. Recently, Göller and Parys [5] have shown that the Ackermannian bottleneck can avoided for ε-free pushdown systems: bisimulation finiteness of (ε-free) pushdown systems is in 6-ExpSpace. The problem is known to be ExpTime-hard [13, 18].

Our contribution. We prove that weak bisimulation finiteness of pushdown systems all of whose ε-transitions are deterministic is 2-ExpTime-complete. The problem was known to be decidable [7] with Ackermannian complexity and in 6-ExpSpace in the easier case of pushdown systems without ε-transitions. Concerning lower bounds, only ExpTime-hardness was known. Our upper bound also generalizes Valiant's 2-ExpTime upper bound [21] for the regularity problem of deterministic pushdown automata from 1975: regularity of deterministic pushdown automata is (modulo simple adaptations) the same problem as weak bisimulation finiteness of pushdown systems all of whose transitions (not only ε-transitions) are deterministic. Our contribution consists of three elements. First, we prove that the smallest finite system that is weakly bisimulation equivalent to a fixed pushdown system, if exists, has size at most doubly exponential in the description size of the pushdown system. Second, we propose a fast algorithm deciding whether a given pushdown system is weakly bisimulation equivalent to a finite system of a given size. Third, we prove 2-ExpTime-hardness of the problem.

Related and future work. As mentioned above, (weak) bisimulation finiteness is a problem that is not well understood [19]. Let us mention a few exceptions. Over Petri net's subclass of basic parallel processes bisimulation finiteness is PSpace-complete [17, 12]. Bisimulation finiteness of one-counter systems (which are pushdown systems where there is, apart from the bottom symbol, only one stack symbol) is PTime-complete [2]. For weak bisimulation finiteness a relevant result is undecidability for Petri nets [8], whereas decidability seems to be open for most other central infinite systems [19]. We hope that our results can pave the way to eventually determining the decidability/complexity status of weak bisimulation finiteness of pushdown systems whose ε-transitions are not restricted to be deterministic.

Organization of the paper. We introduce basic notation and state our main result in Section 2. Our 2-ExpTime upper bound is sketched in Section 3. Basics on pushdown systems are subject of Section 4. In Section 5 we discuss decompositions of stacks. In Section 6 we analyze runs in pushdown systems that mainly decrease the stack height. The core arguments of the upper bound proof are content of Section 7. In Section 8 our algorithm running in double exponential time is presented. A matching 2-ExpTime lower bound is given in Section 9. We conclude in Section 10.

2 Preliminaries

If X is a set we denote by 2^X *power set of X*, that is, the set of all subsets of X. By $\mathbb{N} = \{0, 1, 2, \ldots, \}$ we denote the set of non-negative integers. For all finite alphabets Σ we denote by Σ^* the set of finite words over Σ and, for all $n \in \mathbb{N}$, we denote by $\Sigma^{\leq n} = \{w \in \Sigma^* \mid |w| \leq n\}$ the set of finite words in Σ^* of length at most n. The *empty word* is denoted by ε. By Σ_ε we denote the (disjoint) union $\Sigma \cup \{\varepsilon\}$.

A *labeled transition system with ε-transitions (ε-LTS)* is a tuple $\mathcal{L} = (S, \mathbb{A}, (\to_a)_{a \in \mathbb{A}_\varepsilon})$, where S is a (possibly infinite) set of *configurations*, $\mathbb{A}$ is a finite set of *action symbols*, $(\to_a) \subseteq S \times S$ is a binary relation for all $a \in \mathbb{A}_\varepsilon$. We say $\mathcal{L}$ is *finite* if S is finite. We define its *size* as $|\mathcal{L}| = |S|$, thus $|\mathcal{L}| \in \mathbb{N}$ if $\mathcal{L}$ is finite and $|\mathcal{L}| = \omega$ if not (we only consider countable labeled transition systems in this paper). A *pointed ε-LTS* is a pair $(\mathcal{L}, c)$ such that c is a configuration of $\mathcal{L}$. We define the relation $(\overset{\varepsilon}{\Longrightarrow}) = (\to_\varepsilon^*)$. The relation is extended to words in $\mathbb{A}^+$ as follows: for $a \in \mathbb{A}$ and $w \in \mathbb{A}^*$ we define $(\overset{aw}{\Longrightarrow}) = (\to_\varepsilon^*) \circ (\to_a) \circ (\overset{w}{\Longrightarrow})$. Thus, note that $(\overset{a}{\Longrightarrow}) = (\to_\varepsilon^*) \circ (\to_a) \circ (\to_\varepsilon^*)$. We define the binary relation $(\to) = \bigcup_{a \in \mathbb{A}_\varepsilon} (\to_a)$. For all $c, d \in S$ we define $\mathrm{DIST}(c, d) = \min\{|w| \mid c \overset{w}{\Longrightarrow} d\} \in \mathbb{N} \cup \{\omega\}$, the length of the shortest word of action symbols allowing to reach d from c in $\mathcal{L}$.

For an ε-LTS $\mathcal{L}$ we say a binary relation $R_2 \subseteq S \times S$ *weakly covers* a relation $R_1 \subseteq S \times S$ if for all c, d, c' such that $(c, d) \in R_1$ and $c \overset{a}{\Longrightarrow} c'$ with $a \in \mathbb{A}_\varepsilon$ there exists d' such that $d \overset{a}{\Longrightarrow} d'$ and $(c', d') \in R_2$. A *weak bisimulation* is a relation $R \subseteq S \times S$ that is symmetric and covers itself. Observe that the union of two weak bisimulations is again a weak bisimulation. We write $c \approx d$ if $(c, d) \in R$ for some weak bisimulation relation R; note that $(\approx) \subseteq S \times S$ is the largest weak bisimulation on S. If $c \approx d$, we say that c and d are *weakly bisimilar*. For every configuration $c \in S$ we denote by $[c]_\approx = \{d \in S \mid c \approx d\}$ the *weak bisimulation class of c*. The *weak bisimulation quotient* $[\mathcal{L}]$ is the ε-LTS $[\mathcal{L}] = (\{[c]_\approx \mid c \in S\}, \mathbb{A}_\varepsilon, (\to'_a)_{a \in \mathbb{A}_\varepsilon})$, where for $a \in \mathbb{A}_\varepsilon$ we have $\mathcal{C} \to'_a \mathcal{D}$

if $c \xRightarrow{a} d$ for some $c \in \mathcal{C}$, $d \in \mathcal{D}$ (note that then for every $c \in \mathcal{C}$ there exists some $d \in \mathcal{D}$ such that $c \xRightarrow{a} d$). When talking about weak bisimulation equivalence of two configurations, we must not necessarily require that the configurations are from the same ε-LTS: generally we can write $(\mathcal{L}_1, c_1) \approx (\mathcal{L}, c_2)$ if $c_1 \approx c_2$ holds in the disjoint union of $\mathcal{L}_1$ and $\mathcal{L}_2$. We say that $(\mathcal{L}, c)$ (or just c when $\mathcal{L}$ is clear from the context) is *weakly bisimulation finite* if $(\mathcal{L}, c) \approx (\mathcal{L}', c')$ for some finite pointed ε-LTS $(\mathcal{L}', c')$.

We also define relations $(\approx_k) \subseteq S \times S$ for $k \in \mathbb{N}$, by induction: $\approx_0$ is the full relation $S \times S$, and $\approx_{k+1}$ is the largest symmetric relation that is covered by $\approx_k$. For configurations $c, c' \in S$ we define $[c]_k = \{d \in S \mid c \approx_k d\}$, and $\text{EQLEV}(c, c') = \sup\{k \in \mathbb{N} \mid c \approx_k c'\} \in \mathbb{N} \cup \{\omega\}$.

If for every configuration $c \in S$ and every $a \in \mathbb{A}$ there are finitely many configurations $d \in S$ such that $c \xRightarrow{a} d$, then $(\approx) = \bigcap_{k \in \mathbb{N}}(\approx_k)$, and $\text{EQLEV}(c, c') = \omega$ implies $c \approx c'$ (below we restrict ourselves to ε-LTSs that have this property).

A sequence $c_0 \to_{a_1} c_1 \to_{a_2} \cdots \to_{a_n} c_n$ is called a *run* from c_0 to c_n. For such a run ϱ we write $|\varrho|$ for n, and $\varrho(i)$ for c_i. A composition $\varrho \circ \varrho'$ of two runs is defined in the expected way, assuming $\varrho'(0) = \varrho(|\varrho|)$. We say that runs $\varrho = (c_0 \to_{a_1} c_1 \to_{a_2} \cdots \to_{a_m} c_m)$ and $\varrho' = (d_0 \to_{b_1} d_1 \to_{b_2} \cdots \to_{b_n} d_n)$ are *parallel* if we can find indices $0 = i_0 < i_1 < \cdots < i_k = m$ and $0 = i_0' < i_1' < \cdots < i_k' = n$ such that for all $j \in [0, k]$ we have $c_{i_j} \approx d_{i_j'}$ and for all $j \in [1, k]$ the two words $a_{i_{j-1}+1} a_{i_{j-1}+2} \ldots a_{i_j}$ and $b_{i_{j-1}'+1} b_{i_{j-1}'+2} \ldots b_{i_j'}$ (after dropping all ε's) are equal and have length at most 1.

A *pushdown system with deterministic ε-transitions* (ε-PDS for short) is a tuple $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$, where Q is a finite set of *control states*, Γ is a finite *stack alphabet*, $\mathbb{A}$ is finite set of *action symbols*, $\Delta \subseteq Q \times \Gamma \times \mathbb{A}_\varepsilon \times Q \times \Gamma^*$ is a finite *rewrite relation* such that whenever $(p_1, X_1) = (p_2, X_2)$ and $a_1 = \varepsilon$ for two tuples $(p_1, X_1, a_1, q_1, \beta_1), (p_2, X_2, a_2, q_2, \beta_2) \in \Delta$, then $(p_1, X_1, a_1, q_1, \beta_1) = (p_2, X_2, a_2, q_2, \beta_2)$.

In case $(p, X, a, q, \beta) \in \Delta$ we simply write $qX \xhookrightarrow{a} q'\beta$ and refer to it as a *rule*: in case $a \in \mathbb{A}$ we call it a *reading rule* in case $a = \varepsilon$ we call it an *ε-rule*. We say a pair $(p, X) \in Q \times \Gamma$ *is in ε-mode* if for some $(q, \beta) \in Q \times \Gamma^*$ we have $(p, X, \varepsilon, q, \beta) \in \Delta$; we remark that there can be at most one such pair (q, β). If (p, X) is in ε-mode, then there is no $(a, q', \beta') \in \mathbb{A} \times Q \times \Gamma^*$ with $(p, X, a, q', \beta') \in \Delta$. We say that an ε-PDS is *ε-popping* if for all $(p, X, \varepsilon, q, \beta) \in \Delta$ we have $\beta = \varepsilon$.

The *size* of $\mathcal{P}$ is defined as $|\mathcal{P}| = |Q| + |\Gamma| + |\mathbb{A}| + |\Delta|$. A *configuration* of an ε-PDS is an element from $Q\Gamma^*$. An ε-PDS induces a (potentially infinite) ε-LTS $\mathcal{L}(\mathcal{P}) = (S, \mathbb{A}_\varepsilon, (\to_a)_{a \in \mathbb{A}_\varepsilon})$, where $S = Q\Gamma^*$ and for all $a \in \mathbb{A}_\varepsilon$ we have $(\to_a) = \{(pX\gamma, q\beta\gamma) \mid \gamma \in \Gamma^*, (p, X, a, q, \beta) \in \Delta\}$.

In this paper, we are interested in the following decision problem:

WEAK BISIMULATION FINITENESS FOR ε-PDS

INPUT: A ε-PDS $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$ and a configuration $q_{\text{init}} X_{\text{init}} \in Q\Gamma$.
QUESTION: Is $(\mathcal{L}(\mathcal{P}), q_{\text{init}} X_{\text{init}})$ weakly bisimulation finite?

We remark that a seemingly more general problem when having an initial configuration from $Q\Gamma^*$ as input can always be reduced in polynomial time to our restriction, namely where it is from $Q\Gamma$. It is folklore that, given an ε-PDS $\mathcal{P}$ and a configuration $q_{\text{init}} \alpha_{\text{init}} \in Q\Gamma^*$, one can compute in polynomial time an ε-PDS $\mathcal{P}'$ that is ε-popping and a configuration $q_{\text{init}} X_{\text{init}} \in Q\Gamma$ such that $(\mathcal{L}(\mathcal{P}), q_{\text{init}} \alpha_{\text{init}}) \approx (\mathcal{L}(\mathcal{P}'), q_{\text{init}} X_{\text{init}})$. A formal proof of this reduction can be found in [9, Proposition 11]. Additionally, and this is also folklore, one can restrict the resulting $\mathcal{P}'$ in such a way that all rules (p, X, a, q, β) that appear in it satisfy $|\beta| \leq 2$.

Let us state the main result of this paper:

THEOREM 2.1. WEAK BISIMULATION FINITENESS FOR ε-PDS *is 2-EXPTIME-complete.*

3 Overview of the upper-bound proof

In this section we present overall ideas for proving that WEAK BISIMULATION FINITENESS FOR ε-PDS is in 2-EXPTIME. As already mentioned in the introduction, the proof consists of two main parts. First, in Theorem 7.1, we prove that if an ε-PDS $\mathcal{P}$ is weakly bisimulation finite, that is, is weakly bisimilar to some finite ε-LTS (of an arbitrary, unknown size), then it is weakly bisimilar to some finite ε-LTS of at most doubly exponential size. This already gives a 2-EXPSPACE algorithm for deciding weak bisimulation finiteness by a result by Kučera and Mayr [14], who proved that checking whether a given ε-PDS is weakly bisimilar to a given finite ε-LTS is PSPACE-complete: it is enough to enumerate all finite ε-LTSs of at most doubly exponential size, and for each of them check whether it is weakly bisimilar to the given ε-PDS $\mathcal{P}$. We improve this in Theorem 8.1, where we show

how to compute the weak bisimulation quotient of $\mathcal{P}$ (i.e., the smallest ε-LTS weakly bisimilar to $\mathcal{P}$) in time polynomial in its size (and exponential in $|\mathcal{P}|$).

In this description we concentrate on the proof of Theorem 7.1, which is the essential part. The second part (the algorithm) is much shorter, and is described in Section 8.

For simplicity, in this overview we concentrate on the case of an ε-free PDS $\mathcal{P}$ (i.e., an ε-PDS without any ε-transition), which is already non-trivial. We then briefly explain how ideas from the ε-free case can be extended to the general case, when deterministic ε-transitions are present. In particular, when there are no ε-transitions, the notions of bisimulation and of weak bisimulation coincide, so we drop the word "weak" in this description. Moreover, we formulate some statements here in a loose way, neglecting some details. Full details can be found in the following sections.

Let us assume that our initial configuration $q_{\mathsf{init}}X_{\mathsf{init}}$ of our ε-free PDS $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$ is bisimulation finite. Consider an arbitrary configuration $q\delta$ reachable from $q_{\mathsf{init}}X_{\mathsf{init}}$. Our first step is to locate some possibilities of pumping in the stack content δ. Namely, we represent δ as $\alpha\beta\gamma$, so that configurations $q\alpha\beta^i\gamma$, obtained by repeating the β part (we require $\beta \neq \varepsilon$), are "similar" to the original configuration $q\delta$ in the following sense: first, all these configurations are also reachable (from $q_{\mathsf{init}}X_{\mathsf{init}}$) and, second, the set of control states reachable after popping the topmost part $\alpha\beta^j$ is the same for every j. A standard application of the pigeonhole principle allows us to find such α and β of at most exponential size. This gives us at most a doubly exponential number of different pairs α, β. It is thus enough to show that for each fixed pair α, β, configurations of the form $q\alpha\beta\gamma$ constitute only a doubly exponential number of bisimulation classes.

Next, we shift the study of configurations of the form $q\alpha\beta\gamma$ to the study of configurations of the form $r\gamma$, but where we still assume that $r\gamma$ can be reached from reachable configurations of the form $q\alpha\beta^i\gamma$, for all $i \geq 1$. It is enough to prove that the number of classes of these configurations $r\gamma$ is small, because the class of $q\alpha\beta\gamma$ is determined by the classes of configurations $r\gamma$, obtainable by popping the fixed stack prefix $\alpha\beta$ (cf. Lemma 4.5).

We now exploit our assumption that our initial configuration $q_{\mathsf{init}}X_{\mathsf{init}}$ is bisimulation finite; it can reach F classes, for some $F \in \mathbb{N}$ (at this point we do not have yet any bound on F, we only know that it is finite). An interesting property of finite ε-LTSs is that the $\approx_F$ relation coincides with $\approx$. In other words, if two configurations are not bisimilar, then this can be detected in the first F steps, that is, while reading at most F action symbols. In the ε-free case, reading at most F action symbols is equivalent to performing at most F transitions. Bearing in mind $\beta \neq \varepsilon$, it follows that configurations of the form $q\alpha\beta^i\gamma$ for $i \geq F$ are all bisimilar (here q, α, and β are fixed, but the number $i \geq F$ and the stack content γ are arbitrary): during the first F transitions, $\mathcal{P}$ can pop at most F topmost stack symbols, and they are identical in all these configurations. If we take an even larger bound, say $F + k$, then not only $q\alpha\beta^i\gamma \approx q\alpha\beta^{i'}\gamma'$ for all $i, i' \geq F + k$ and $\gamma, \gamma' \in \Gamma^*$, but also after executing a run of length k from $q\alpha\beta^i\gamma$ and an analogous (i.e., performing the same transitions) run from $q\alpha\beta^{i'}\gamma'$, the resulting configurations remain bisimilar. We utilize this observation by introducing stack contents of the form $\alpha\beta^\omega\gamma$. The ω exponent can be formalized in three ways. First, we may assume that ω denotes some "very large" finite number (large enough for purposes of the proof). A second formalization is that the β part of the stack is repeated infinitely many times. In our proofs we have chosen yet another formalization, where we understand β^ω as a formal expression such that after popping β from β^ω we again have β^ω. Nevertheless, no matter which formalization one chooses, it is only important that the finite (although possibly very long) prefix of the stack that may be analyzed by the PDS consists of one α and repetitions of β.

Bounding the number of repetitions. Next, given some stack contents α, β, and γ we are interested in a possibly small $n \in \mathbb{N}$ such that $q\alpha\beta^n\gamma \approx q\alpha\beta^\omega\gamma$. We fix $e \in \mathbb{N}$ to be the smallest number such that $q\alpha\beta^e\gamma$ becomes equivalent to $q\alpha\beta^\omega\gamma$ (strictly speaking, we require that $r\beta^e\gamma \approx r\beta^\omega\gamma$ for all control states r reachable after popping α; this implies that $q\alpha\beta^e\gamma \approx q\alpha\beta^\omega\gamma$, but it is a slightly stronger condition). We remark that Jančar [7] has shown an Ackermannian bound for e. For us, an important step is to prove that e can be at most doubly exponential. To this end, for every $i \in [0, e]$, we consider the control state r_i for which $\mathrm{EQLEV}(r_i\beta^i\gamma, r_i\beta^\omega\gamma)$ is minimal (i.e., the control state from which the stack contents $\beta^i\gamma$ and $\beta^\omega\gamma$ can be distinguished in the least possible number of transitions), and for this control state we take $M_i = \mathrm{EQLEV}(r_i\beta^i\gamma, r_i\beta^\omega\gamma)$. It is rather easy to prove that $M_i < M_{i+1}$ for all $i < e$: the similarity between $\beta^i\gamma$ and $\beta^\omega\gamma$ grows when we increase i, until it reaches level ω for $i = e$. In particular (because $\mathrm{EQLEV}(\cdot, \cdot)$ depends only on the class of a configuration) this means that if $r_i = r_{i'}$ for $i \neq i'$, then the classes $[r_i\beta^i\gamma]_\approx$ and $[r_i\beta^{i'}\gamma]_\approx$ are different.

Given some stack contents α, β, and γ and some $i \in [0, e-1]$, we now want to provide a succinct description of the class of $r_i\beta^i\gamma$. To this end, we consider two runs. The first of them, π, is the shortest possible run from $q\alpha\beta^e\gamma$

to $r_i\beta^i\gamma$. The second of them, π', is created as a run parallel to π that starts in the "much larger" configuration $q\alpha\beta^\omega\gamma$ (it exists because $q\alpha\beta^\omega\gamma \approx q\alpha\beta^e\gamma$). Recall that then π and π' are required to visit bisimilar configurations, but they are *not* required to execute the same transitions (in fact, they cannot execute the same transitions: if π', starting from $q\alpha\beta^\omega$, were to pop $\alpha\beta^{e-i}$ in the same way as π did, then π' would end in $r_i\beta^\omega\gamma$, a configuration that is not bisimilar to $r_i\beta^i\gamma$ by assumption). Depending on the shape of π', we have two cases.

First, it is possible that no suffix of π' pushes D stack symbols, for an appropriately chosen exponential constant D; so, roughly speaking, π' concentrates on popping. We then observe that no matter how much we pop from $q\alpha\beta^\omega\gamma$, we can only pop to a stack content of the form $\beta'\beta^\omega\gamma$, where β' is a suffix of either α or β. Even taking into account the fact that exponentially many symbols may be pushed at the very end, on top of β', in this case we have only doubly exponentially many possibilities for the configuration $\pi'(|\pi'|)$, hence also for its class, which is simultaneously the class of $r_i\beta^i\gamma$. Recalling that for different values of i (having the same state r_i) the classes of $r_i\beta^i\gamma$ are all different, this means that only doubly exponentially many values of i may be handled by this case. We can thus concentrate on the opposite case, which we discuss next.

This opposite case, being significantly more complicated, is that π' has a suffix π'_3 that pushes D stack symbols (note that if more than D symbols are pushed, then we can consider a shorter suffix, pushing exactly D symbols). The suffix π'_3 leads from $t_1 X\theta$ to $t\mu\theta$ for some control states t_1, t, stack contents μ, θ, and stack symbol X, where $|\mu| = \mathsf{D} + 1$. We now use the simple observation that the class of $t\mu\theta$ is determined by the control state t, by the small stack content μ, and by the classes of $u\theta$ for all control states u reachable from $t\mu\theta$ after popping μ (as formalized in Lemma 4.5).

Next, we would somehow like to describe the classes of $u\theta$. To this end, we observe that since $u\theta$ is reachable from $t\mu\theta$, it is also reachable from $t_1 X\theta$ (due to the existence of the run π'_3 from $t_1 X\theta$ to $t\mu\theta$). Moreover, a simple pumping argument (namely, Lemma 4.1) allows us to shorten the run from $t_1 X\theta$ to $u\theta$ (observe that, in total, only one stack symbol needs to be popped) into a very short run; namely, shorter than some constant B (being exponential in $|\mathcal{P}|$). Coming back to the analysis of π, note that the parallel counterpart of $\pi'_3(0) = t_1 X\theta$ in π' is, say, the configuration $q_1\chi\beta^j\gamma$ in π: it is visited while going from $q\alpha\beta^e\gamma$ to $r_i\beta^i\gamma$. The χ part is obtained from β (or α) by popping its prefix as quickly as possible by definition of π. Thus, standard pumping arguments imply that the length of χ is short. Possibly decreasing j and appending a few copies of β into χ, we may artificially redefine $|\chi|$ to be larger than B, but not much larger. If the value of the constant D is appropriate so that π'_3 is long enough, then also the subrun of π from $q_1\chi\beta^j\gamma$ to $r_i\beta^i\gamma$ is long enough to ensure that $j > i$.

We have thus two bisimilar configurations, $q_1\chi\beta^j\gamma$ and $t_1 X\theta$, and we have a run of length B from $t_1 X\theta$ to $u\theta$. Then a run of the same length exists also from $q_1\chi\beta^j\gamma$ to a configuration bisimilar to $u\theta$; this configuration is necessarily of the form $v_u\xi_u\beta^j\gamma$ (we cannot pop the whole χ with $|\chi| > \mathsf{B}$ in only B steps), where the size of ξ_u is again bounded exponentially. We can find such v_u and ξ_u for all considered states u (note that j does not depend on u).

The classes of $v_u\xi_u\beta^j\gamma$ in turn are determined by the control states v_u, by the small stack contents ξ_u, and by the classes $r\beta^j\gamma$ for all control states r reachable from $v_u\xi_u\beta^j\gamma$ after popping ξ_u (hence also reachable from $q\alpha\beta^e\gamma$ after popping $\alpha\beta^{e-j}$).

Summing up this part, we have shown that the class of $r_i\beta^i\gamma$ can be determined by the control state/stack pairs $t\mu$ and $v_u\xi_u$ (indexed by control states u from an appropriate set U) and by the classes of $r\beta^j\gamma$ for some $j > i$. All the stack contents μ and ξ_u have at most exponential size. The number of possibilities for $(t\mu, (v_u\xi_u)_{u\in U})$ is then doubly exponential.

Is it possible that the same tuple $(t\mu, (v_u\xi_u)_{u\in U})$ is assigned to two distinct indices $i, i' < e$, such that moreover $r_i = r_{i'}$? We prove that this is not possible, which immediately implies a doubly exponential upper bound for e, being the number of considered indices $i \in [0, e-1]$. Suppose thus, to the contrary, we had two such indices i, i', where $i < i'$. Then the class of $r_i\beta^i\gamma$ is determined by the classes of $r\beta^j\gamma$, for some $j > i$, in the same way as the class of $r_i\beta^{i'}\gamma$ is determined by the classes of $r\beta^{j'}\gamma$, for some $j' > i' > i$ (due to equality of the descriptions assigned to i and i'). Recalling the inequalities $M_i + 1 \le M_j$ and $M_i + 1 \le M_{i'} < M_{j'}$, and the definition of M_j and $M_{j'}$, we have $r\beta^j\gamma \approx_{M_i+1} r\beta^\omega\gamma \approx_{M_i+1} r\beta^{j'}\gamma$ for all control states r under consideration. It follows easily by the equality of their descriptions, that then also $r_i\beta^i\gamma \approx_{M_i+1} r_i\beta^{i'}\gamma$ (we depend in the same way on $\approx_{M_i+1}$-equivalent configurations, so we remain $\approx_{M_i+1}$-equivalent). Due to $r_i\beta^{i'}\gamma \approx_{M_i+1} r_i\beta^\omega\gamma$ this implies $r_i\beta^i\gamma \approx_{M_i+1} r_i\beta^\omega\gamma$, which contradicts with $\textsc{EqLev}(r_i\beta^i\gamma, r_i\beta^\omega\gamma) = M_i$. This finishes the proof of the doubly exponential bound on e.

Bounding the number of classes. We now come back to our initial goal, namely bounding the number of classes of $r\gamma$, where $r\gamma$ is reachable from the initial configuration via $q\alpha\beta^k\gamma$ for all $k \geq 1$, where α and β are fixed and short. We employ here exactly the same characterization of the classes as in the previous part, where we were bounding e. This time we only take $i = 0$ (i.e., we consider configurations $r\beta^0\gamma = r\gamma$), but we allow an arbitrary control state r in place of r_i, and do not assume that γ is fixed. As previously, we have two possibilities.

First, it may happen that $r\gamma$ is bisimilar to a configuration of the form $t\beta'\beta^\omega\gamma$ that is reachable from $q\alpha\beta^\omega\gamma$ by a run π' that "focused on popping", that is, belonging to the first case we have considered. As β' is small (at most exponential), and since the class of $t\beta'\beta^\omega\gamma$ does not depend on γ (as already explained, $(\approx) = (\approx_F)$, while γ is too deep in the stack to be seen in the first F transitions), we only have a doubly exponential number of classes of $r\gamma$ in this case.

The second case is that the class of $r\gamma$ is determined by a small (exponential) information $(t\mu, (v_u, \xi_u)_{u \in U})$ and by the classes of $r'\beta^j\gamma$ for some $j > i = 0$. The latter classes, in turn, are determined by the stack content β^j (hence by j, because β is fixed) and by the classes of $r''\gamma$. Thus, if we add j to the remembered information, we can say that the class of $r\gamma$ is determined by $(t\mu, (v_u, \xi_u)_{u \in U}, j)$ and by the classes of $r''\gamma$. The previously shown doubly exponential bound on e implies that the number of possibilities for $(t\mu, (v_u, \xi_u)_{u \in U}, j)$ is doubly exponential. A minor detail is that the tuple $(t\mu, (v_u, \xi_u)_{u \in U}, j)$ describes the class of $r\gamma$ for a single control state r; we should rather, given a stack content γ, consider a tuple of such tuples, indexed by states r from an appropriate set R (control states reachable from $q\alpha\beta^e\gamma$ after popping $\alpha\beta^e$), hence describing the classes of $r\gamma$ simultaneously for all $r \in R$. We now have a cyclic situation: the classes of $r\gamma$ are determined by the small a small tuple (indexed by control states) of the above-mentioned information and by the classes of $r\gamma$ themselves. One can show that in such a situation, the small information is enough to determine the classes of $r\gamma$. We thus have a doubly exponential bound on the number of these classes.

Adding ε-transitions. We now briefly discuss on adapting the above proof idea to the general case, where ε-transitions may be present. Recall that every pushing transition reads some action symbol; ε-transitions are allowed only for popping and need to be deterministic. It turns out that if, while popping β, at least one action symbol is necessarily read, then the proof sketch presented above still works, up to adjusting some details. The main difficulty comes with the fact that it may be possible to pop arbitrarily many copies of β without reading any action symbols. In particular, starting from a configuration of the form $q\alpha\beta^k\gamma$, the γ part of the stack may be possibly reached after reading very few action symbols, no matter how large k is.

In order to deal with this difficulty, we proceed as follows: instead of splitting the whole stack content into a single *pumping triple* $\alpha\beta\gamma$, we now rather try to find pumping triples $\alpha\beta\gamma$ with slightly stronger properties, but being only infixes of the whole stack. If some infix of the stack content is represented as a pumping triple $\alpha\beta\gamma$, being located on top of a stack content η, then we still have the previous requirement saying, roughly, that the β infix may be pumped. Beside that, we have a new requirement: if $r\beta^i\gamma\eta$ is reached by an ε-run popping β from some larger configuration $r'\beta^{i+1}\gamma\eta$, then further ε-transitions allow to pop the whole $\beta^i\gamma$, and reach a configuration of the form $s\eta$.

As already said, the previous proof works correctly when the considered runs from $q\alpha\beta^e\gamma\eta$ (and likewise from $q\alpha\beta^\omega\gamma\eta$) read at least one action symbol while popping every copy of β. There are also runs that at some moment pop a copy of β using only ε-transitions. Then these runs (possibly after prolonging them) continue with popping ε-transitions until the stack content η is uncovered. In the proof we need to add special cases for such runs. It turns out, however, that these runs do not introduce many new possibilities. An intuitive reason for this is that if a run, while being in some configuration $r\beta^i\gamma\eta$, starts performing ε-transitions until $s\eta$ is reached, then the number i and the stack content γ are "forgotten"; neither the target configuration nor the action symbols read on the way (we do not read anything) depend on i or on γ.

In Section 5 we introduce decompositions of stacks. They give a way to decompose the whole stack content into many pumping triples $\alpha\beta\gamma$, located possibly in a nested way. In this decomposition, most of the stack content, except for exponentially many symbols, is put into the γ parts of pumping triples. The proof presented above, dealing with a single pumping triple, is made formal and complete in Lemma 7.2, which is our main technical lemma. Then, an appropriate induction gives us the final upper-bound result, Theorem 7.1.

4 Some basics on pushdown systems

In this section we present some definitions and known facts about pushdown systems, useful in our proofs. For purposes of the whole section let $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$ be an ε-PDS.

Our first lemma says that if there is a run between two similar configurations, then there is a short run between them. Essentially, this boils down to the standard pumping lemma for pushdown automata.

LEMMA 4.1. ([5, LEMMA 3.3]) *There exists a constant*[1] $\mathsf{B} \in 2^{|\mathcal{P}|^{O(1)}}$ *such that whenever* $p\alpha \to^* q\beta$ *for two configurations* $p\alpha, q\beta \in Q\Gamma^*$*, then* $\mathrm{DIST}(p\alpha, q\beta) \leq (|\alpha| + |\beta|) \cdot \mathsf{B}$. $\qquad\square$

For every stack content $\alpha \in \Gamma^*$, we define two sets, describing possible ways of pushing α:

$$\mathsf{up}(\alpha) = \{(pX, qY) \in (Q\Gamma)^2 \mid pX \to^* qY\alpha\} \qquad \text{for } \alpha \in \Gamma^*, \text{ and}$$
$$\mathsf{up}_0(\alpha) = \{(pX, q) \in (Q\Gamma) \times Q \mid pX \to^* q\alpha\} \qquad \text{for } \alpha \in \Gamma^* \setminus \{\varepsilon\}.$$

It is easy to see that for all $\alpha, \beta, \gamma \in \Gamma^*$ with $\alpha \neq \varepsilon$,

$$\mathsf{up}(\beta\gamma) = \mathsf{up}(\gamma) \circ \mathsf{up}(\beta) \qquad \text{and} \qquad \mathsf{up}_0(\alpha\beta) = \mathsf{up}(\beta) \circ \mathsf{up}_0(\alpha),$$

where "$\circ$" denotes the composition of relations, understood in the usual way.

As explained in Section 3, it is convenient to extend the set of configurations into expressions containing ω, which we do as follows. A *generalized stack* is defined by induction: it is a (possibly empty) sequence $\alpha_1\alpha_2\ldots\alpha_k$, where every α_i is either

- a stack symbol from Γ, or
- an expression of the form $(\beta_1\beta_2\ldots\beta_\ell)^\omega$, where $\beta_1\beta_2\ldots\beta_\ell$ is a nonempty (i.e., satisfying $\ell \geq 1$) generalized stack.

The set of generalized stacks is denoted Γ^{gs}. A *generalized configuration* is an element of $Q\Gamma^{\mathsf{gs}}$. From now on, whenever we talk about a configuration, we usually mean a generalized one. Configurations that are not generalized, are called *standard*. Note that Lemma 4.1 and the definition of $\mathsf{up}(\cdot)$ were only given for standard configurations.

Let us emphasize that, formally, generalized stacks are expressions (terms); we do not assign any concrete meaning to the $(\cdot)^\omega$ operation. However, the intuition staying behind is that $(\beta_1\beta_2\ldots\beta_\ell)^\omega$ describes a stack in which the sequence $\beta_1\beta_2\ldots\beta_\ell$ is repeated very many times.

We allow ourselves to write α^e when $e \in \mathbb{N} \cup \{\omega\}$, mixing two different semantics of the same notation: for $e \in \mathbb{N}$ we mean the result of repeating e times the sequence α, while for $e = \omega$ we mean the expression α^ω.

For $a \in \mathbb{A}_\varepsilon$ we define $\to_a$ to be the unique relation on $Q\Gamma^{\mathsf{gs}}$ that holds only in the following cases:

1. if $(p, X, a, q, \beta) \in \Delta$, then $pX\gamma \to_a q\beta\gamma$ for all $\gamma \in \Gamma^{\mathsf{gs}}$,
2. if $p\alpha \to_a q\beta$ and $(p\alpha, p) \notin (\to_\varepsilon^*)$, then $p\alpha^\omega\gamma \to_a q\beta\alpha^\omega\gamma$, and
3. if $p\alpha \to_\varepsilon^* p$, then $p\alpha^\omega\gamma \to_\varepsilon p\gamma$.

Note that in order to know whether $p\alpha^\omega\gamma \to_a q\beta\alpha^\omega\gamma$ or $p\alpha^\omega\gamma \to_\varepsilon p\gamma$, one has to check whether $p\alpha \to_a q\beta$ and whether $p\alpha \to_\varepsilon^* p$. It is important here that the nesting depth of the $(\cdot)^\omega$ operation in α is strictly smaller than in $\alpha^\omega\gamma$, so the whole definition is well-formed.

Based on that, we then define the $\overset{w}{\Longrightarrow}$ relation over generalized configurations, as previously. Note that, when restricted to standard configurations, the above definition coincides with the previous one. Moreover, only standard configurations may be reached from $q_{\mathsf{init}}X_{\mathsf{init}}$.

Item 2 of the above definition treats $p\alpha^\omega\gamma$ almost as $p\alpha\alpha^\omega\gamma$: we can go to $q\beta\alpha^\omega\gamma$ when there is a transition from $p\alpha$ to $q\beta$. There is one exception, though: if $p\alpha \to_\varepsilon^* p$, then we rather pop the whole α^ω, going directly to $p\gamma$ (cf. Item 3). The intuition is that if α may be popped using ε-transitions, then also arbitrarily many copies of α, embodied in the α^ω expression, may be popped using ε-transitions. Transitions from Item 2 are disallowed in this case, so that ε-transitions remain deterministic. In Item 3 it is required that the state p before and after popping α is the same; then it makes sense to say that after popping very many copies of α the state will be again p. If we had $p\alpha \to_\varepsilon^* q$ and $q\alpha \to_\varepsilon^* p$, then it would be unclear in which state (in p or in q) we should finish popping α^ω. For this reason we do not add any special transition in such a case.

The following property is a direct consequence of determinism of ε-transitions:

if $c \to_\varepsilon^* d$, then $c \approx d$ (and $c \approx_k d$ for all $k \in \mathbb{N}$).

---

[1] Constants depending on $\mathcal{P}$ are denoted by capital letters in Sans Serif font; for constants singly exponential in the size of $\mathcal{P}$ we use initial letters of the alphabet ($\mathsf{B}, \mathsf{C}, \mathsf{D}$), while for doubly exponential constants we use letters near the end of the alphabet ($\mathsf{T}, \mathsf{U}, \mathsf{V}, \mathsf{Z}$).

Yet another useful property of generalized configurations is the following:

for all $p \in Q$ and $\alpha, \gamma \in \Gamma^{\mathsf{gs}}$ we have $p\alpha^\omega\gamma \approx p\alpha\alpha^\omega\gamma$.

Indeed, if $p\alpha \to_\varepsilon^* p$ (hence also $p\alpha\alpha^\omega\gamma \to_\varepsilon^* p\alpha^\omega\gamma$), equivalence follows from the previous property; otherwise, exactly the same transitions can be performed from $p\alpha^\omega\gamma$ and from $p\alpha\alpha^\omega\gamma$, leading to exactly the same configurations.

We say that a run ϱ' is a *shift* of a run ϱ if ϱ' is obtained by appending or removing the same suffix to every configuration in ϱ.

For $\alpha \in \Gamma^{\mathsf{gs}}$ and $P \subseteq Q$ we define two functions, saying how states can change while popping α:

$$|\alpha\rangle(P) = \{r \in Q \mid \exists q \in P.\, q\alpha \to^* r\} \qquad \text{and} \qquad |\alpha\rangle_\varepsilon(P) = \{r \in Q \mid \exists q \in P.\, q\alpha \to_\varepsilon^* r\}.$$

For singleton sets we abbreviate $|\alpha\rangle(\{q\})$ as $|\alpha\rangle(q)$, and likewise we abbreviate $|\alpha\rangle_\varepsilon(\{q\})$ as $|\alpha\rangle_\varepsilon(q)$. Observe that

$$|\beta\rangle(|\alpha\rangle(P)) = |\alpha\beta\rangle(P), \qquad |\beta\rangle_\varepsilon(|\alpha\rangle_\varepsilon(P)) = |\alpha\beta\rangle_\varepsilon(P), \qquad \text{and} \qquad |\alpha\rangle_\varepsilon(P) \subseteq |\alpha\rangle(P)$$

for all $\alpha, \beta \in \Gamma^{\mathsf{gs}}$ and $P \subseteq Q$. The following lemma is a direct consequence of the definition:

LEMMA 4.2. *Let $p, q \in Q$ and $\alpha, \beta \in \Gamma^{\mathsf{gs}}$. If $p\alpha \to^* q\beta$, then $|\alpha\rangle(p) \supseteq |\beta\rangle(q)$.* $\qquad\square$

The next property follows from determinism of ε-transitions:

LEMMA 4.3. *Let $q \in Q$ and $\alpha \in \Gamma^{\mathsf{gs}}$. If $|\alpha\rangle_\varepsilon(q) \neq \emptyset$, then $|\alpha\rangle(q) = |\alpha\rangle_\varepsilon(q)$, and this set has size 1.*

Proof. If $q\alpha \to_\varepsilon^* r$, then we cannot have $q\alpha \to^* s$ for any other state $s \neq r$. $\qquad\square$

Recall that DIST counts the number of action symbols needed to reach some configuration. Because ε-transitions are not counted, there may be arbitrarily many configurations in a given distance. We fix this by defining NEAR: for $p\alpha \in Q\Gamma^{\mathsf{gs}}$ and $k \in \mathbb{N}$ let $\mathrm{NEAR}(p\alpha, k)$ be the set of configurations that can be reached from $p\alpha$ by a run reading at most k action symbols and not ending with an ε-transition.

LEMMA 4.4. *For all $p\alpha \in Q\Gamma^{\mathsf{gs}}$ and $k \in \mathbb{N}$ we have*

$$|\{\mathcal{C} \mid \mathrm{DIST}([p\alpha]_\approx, \mathcal{C}) \le k\}| \le |\mathrm{NEAR}(p\alpha, k)| \le |\mathcal{P}|^k.$$

Proof. If $\mathrm{DIST}([p\alpha]_\approx, \mathcal{C}) \le k$, then there is a run ϱ from $p\alpha$ to a configuration $c \in \mathcal{C}$, reading at most k action symbols. Let c' be the first configuration reached by ϱ after reading all these action symbols; after c' we have only ε-transitions. Then $c' \in \mathrm{NEAR}(p\alpha, k)$, and $c' \approx c$, so $c' \in \mathcal{C}$. Thus in every class from the set $\{\mathcal{C} \mid \mathrm{DIST}([p\alpha]_\approx, \mathcal{C}) \le k\}$ there is an element of the set $\mathrm{NEAR}(p\alpha, k)$, so the former set cannot be larger; we obtain the first inequality.

To obtain the second inequality, we observe that every configuration in $\mathrm{NEAR}(p\alpha, k)$ is fully characterized by the list of at most k transitions performed while reading action symbols by a run not ending with an ε-transition. Indeed, from every configuration c we can go by ε-transitions to at most one configuration d such that action symbols can be read from d. When we are in d, we perform a transition from our list. Then again we go by ε-transitions to the unique configuration from which action symbols can be read, and so on. At the very end, we have to stop immediately after performing the last transition from the list. The number of lists of at most k transitions is

$$\sum_{i=0}^{k} |\Delta|^i \le (|\Delta| + 1)^k \le |\mathcal{P}|^k. \qquad\square$$

For a length-1 run ϱ let

$$\mathrm{STACKGROWTH}(\varrho) = \begin{cases} |\beta| - 1 & \text{if } \varrho = (pX\gamma \to_a q\beta\gamma), \\ \mathrm{STACKGROWTH}(p\alpha \to_a q\beta) & \text{if } \varrho = (p\alpha^\omega\gamma \to_a q\beta\alpha^\omega\gamma), \text{ and} \\ -\omega & \text{if } \varrho = (p\alpha^\omega\gamma \to_\varepsilon p\gamma), \end{cases}$$

and for a run $\varrho = \varrho_1 \circ \cdots \circ \varrho_n$ of length n let

$$\textsc{StackGrowth}(\varrho) = \sum_{i=1}^{n} \textsc{StackGrowth}(\varrho_i).$$

Here we assume that $-\omega + n = n - \omega = -\omega - \omega = -\omega$ for every $n \in \mathbb{N}$. Note that if ϱ is a run between standard configurations $p\alpha, q\beta \in Q\Gamma^*$, then $\textsc{StackGrowth}(\varrho) = |\beta| - |\alpha|$.

Finally, we state two simple but important lemmata, which are useful while proving that two configurations are weakly bisimilar (similar lemmata appeared already in prior work [5, 6, 7]):

LEMMA 4.5. *For all $q \in Q$ and $\alpha, \eta, \eta' \in \Gamma^{\mathsf{gs}}$ we have*

$$\textsc{EqLev}(q\alpha\eta, q\alpha\eta') \geq \min\{\textsc{EqLev}(r\eta, r\eta') \mid r \in |\alpha\rangle(q)\}.$$

In other words, if $r\eta \approx r\eta'$ for all $r \in |\alpha\rangle(q)$ then $q\alpha\eta \approx q\alpha\eta'$, and if $r\eta \approx_k r\eta'$ for all $r \in |\alpha\rangle(q)$ then $q\alpha\eta \approx_k q\alpha\eta'$.

Proof. It should be clear that the second part is equivalent to the first part. To see the first part, we consider two cases: If there is some $r \in |\alpha\rangle_\varepsilon(q) \subseteq |\alpha\rangle(q)$, then $q\alpha\eta \to_\varepsilon^* r\eta$ and $q\alpha\eta' \to_\varepsilon^* r\eta'$; we have $q\alpha\eta \approx r\eta$ and $q\alpha\eta' \approx r\eta'$, so $\textsc{EqLev}(q\alpha\eta, q\alpha\eta') = \textsc{EqLev}(r\eta, r\eta')$; the inequality follows. Otherwise, if $|\alpha\rangle_\varepsilon(q) = \emptyset$, we just use Lemma 4.6, stated below. $\square$

LEMMA 4.6. *Let $q \in Q$ and $\alpha, \eta, \eta' \in \Gamma^{\mathsf{gs}}$. If $|\alpha\rangle_\varepsilon(q) = \emptyset$, then*

$$\textsc{EqLev}(q\alpha\eta, q\alpha\eta') \geq 1 + \min\{\textsc{EqLev}(r\eta, r\eta') \mid r \in |\alpha\rangle(q)\}.$$

Proof. We prove for every $k \in \mathbb{N}$ that if $|\alpha\rangle_\varepsilon(q) = \emptyset$ and $r\eta \approx_k r\eta'$ for all $r \in |\alpha\rangle(q)$, then $q\alpha\eta \approx_{k+1} q\alpha\eta'$; the thesis of the lemma follows easily from this statement. The proof is by induction on k. Consider two configurations $q\alpha\eta$ and $q\alpha\eta'$ such that $|\alpha\rangle_\varepsilon(q) = \emptyset$ and $r\eta \approx_k r\eta'$ for all $r \in |\alpha\rangle(q)$; we want to prove that $q\alpha\eta \approx_{k+1} q\alpha\eta'$. Before starting observe that

$$\text{if } q\alpha \to^* s\beta \text{ then } s\beta\eta \approx_k s\beta\eta'. \tag{$\star$}$$

Indeed, if $k = 0$, this holds trivially (all configurations are in the $\approx_0$ relation). If $k \geq 1$ and there is some $r \in |\beta\rangle_\varepsilon(s) \subseteq |\alpha\rangle(q)$ (inclusion by Lemma 4.2), then $s\beta\eta \to_\varepsilon^* r\eta$ and $s\beta\eta' \to_\varepsilon^* r\eta'$ imply that $s\beta\eta \approx r\eta \approx_k r\eta' \approx s\beta\eta'$. If $k \geq 1$ and $|\beta\rangle_\varepsilon(s) = \emptyset$, we observe that $r\eta \approx_{k-1} r\eta'$ for all $r \in |\beta\rangle(s)$ (due to $(\approx_k) \subseteq (\approx_{k-1})$ and $|\beta\rangle(s) \subseteq |\alpha\rangle(q))$, so we have $s\beta\eta \approx_k s\beta\eta'$ by the induction hypothesis. This finishes the proof of Property $(\star)$.

In order to obtain $q\alpha\eta \approx_{k+1} q\alpha\eta'$, it is enough to show that the two-element symmetric relation $\{(q\alpha\eta, q\alpha\eta'), (q\alpha\eta', q\alpha\eta)\}$ is covered by $\approx_k$; the relation will be then contained in $\approx_{k+1}$, the largest symmetric relation covered by $\approx_k$. Expanding the definition of being covered, we should thus prove that if $q\alpha\eta \stackrel{a}{\Longrightarrow} c$ with $a \in \mathbb{A}_\varepsilon$, then there exists c' such that $c \approx_k c'$ and $q\alpha\eta' \stackrel{a}{\Longrightarrow} c'$ (and that if $q\alpha\eta' \stackrel{a}{\Longrightarrow} c'$, then there exists c such that $c' \approx_k c$ and $q\alpha\eta \stackrel{a}{\Longrightarrow} c$; this part is symmetric, so we do not need a separate proof). Suppose thus that $q\alpha\eta \stackrel{a}{\Longrightarrow} c$. If $a = \varepsilon$, we can take $c' = q\alpha\eta'$; then $c \approx q\alpha\eta \approx_k q\alpha\eta' = c'$ by Property $(\star)$ and obviously $q\alpha\eta' \stackrel{\varepsilon}{\Longrightarrow} c'$. If $a \in \mathbb{A}$, then $q\alpha\eta \to_\varepsilon^* c_1 \to_a c_2 \to_\varepsilon^* c$ for some c_1, c_2. Because $|\alpha\rangle_\varepsilon(q) = \emptyset$, no configuration on a run from $q\alpha\eta$ to c_1 may have stack η; we have $c_1 = s_1\beta\eta$ for some $s \in Q$ and $\beta \in \Gamma^{\mathsf{gs}}$ with $\beta \neq \varepsilon$. Then also $q\alpha\eta' \to_\varepsilon^* s_1\beta\eta'$. A single transition cannot pop more than the whole β, so $c_2 = s_2\gamma\eta$ for some $s_2 \in Q$ and $\gamma \in \Gamma^{\mathsf{gs}}$; then also $s_1\beta\eta' \to_a s_2\gamma\eta'$. We take $c' = s_2\gamma\eta'$. We have $q\alpha\eta' \stackrel{a}{\Longrightarrow} c'$, and we obtain $c \approx c_2 = s_2\gamma\eta \approx_k s_2\gamma\eta = c'$ by Property $(\star)$. $\square$

5 Decompositions of stacks

In this section we define decompositions of stacks, prove that they exist, and show their basic properties. In essence, decompositions specify how to split a stack content in a way suitable for pumping. For the whole section we fix an ε-PDS $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$ together with an initial configuration $q_{\mathsf{init}}X_{\mathsf{init}}$.

A *decomposition* $\overline{\delta}$ is defined by induction, and can be in one of three forms:
- a single symbol (base case): $\overline{\delta} = X \in \Gamma$,

- a nonempty sequence of decompositions: $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$, where $\overline{\delta}_1, \ldots, \overline{\delta}_k$ are decompositions and $k \geq 1$, or
- a *pumping triple*: $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$, where $\overline{\alpha}$ and $\overline{\beta}$ are decompositions and $\gamma \in \Gamma^*$ is a stack.

The *height* (resp. *degree*) of a decomposition $\overline{\delta}$ is inductively defined as follows:
- the height and degree of $\overline{\delta} = X \in \Gamma$ is 0,
- the height (resp. degree) of $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$ is 1 plus the maximum of heights of any of the $\overline{\delta}_i$ (resp. the maximum of k and the degree of any of the $\overline{\delta}_i$), and
- the height (resp. degree) of $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ is 1 plus the maximum of the height of $\overline{\alpha}$ and the height of $\overline{\beta}$ (resp. the maximum of 2, the degree of $\overline{\alpha}$, and the degree of $\overline{\beta}$).

In other words, a decomposition can be seen as a tree for which, as expected, the height is the length of some longest path from the root to some leaf and the degree is the maximal number of children of some node in the tree.

For a decomposition $\overline{\delta}$ and for $e \in \mathbb{N} \cup \{\omega\}$ we define a stack $\overline{\delta}^{[e]} \in \Gamma^*$, called the *e-th realization* of $\overline{\delta}$:
- $X^{[e]} = X$;
- $(\overline{\delta}_1 \ldots \overline{\delta}_k)^{[e]} = \overline{\delta}_1^{[e]} \ldots \overline{\delta}_k^{[e]}$;
- $\langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle^{[e]} = \overline{\alpha}^{[e]}(\overline{\beta}^{[e]})^e \gamma$ (i.e., we repeat the middle part e times).

If $\delta = \overline{\delta}^{[1]}$, we say that $\overline{\delta}$ is a *decomposition of* δ. Note that $\overline{\delta}^{[1]}$ is obtained by simply concatenating all stack symbols that literally appear in $\overline{\delta}$, and that $\overline{\delta}^{[e]}$ for greater numbers e is obtained by repeating e times some fragments of $\overline{\delta}^{[1]}$, in a nested way. For $e = \omega$ the intuition is that we repeat infinitely many times appropriate fragments of the stack, but formally we just write ω in the exponent, without repeating anything.

Having the above definition, we can say when a decomposition is *well-formed*: A decomposition consisting of a single stack symbol is always well-formed, a decomposition $\overline{\delta}_1 \ldots \overline{\delta}_k$ is a well-formed when all $\overline{\delta}_i$ are well-formed, and a decomposition $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ is well-formed when both $\overline{\alpha}$ and $\overline{\beta}$ are well-formed and additionally

1. $\mathsf{up}(\gamma) = \mathsf{up}(\overline{\beta}^{[1]}\gamma)$,
2. $|\overline{\alpha}^{[\omega]}\rangle = |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle$, and
3. $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r) = \{r\}$ and $|\gamma\rangle_\varepsilon(r) \neq \emptyset$ for all $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$.

Below, whenever talking about a decomposition, we assume that it is well-formed.

Let us explain the above conditions. Item 1 implies that $\mathsf{up}(\gamma) = \mathsf{up}((\overline{\beta}^{[1]})^e \gamma)$ for every $e \in \mathbb{N}$, so it ensures that when we pump a stack content of a reachable configuration according to a decomposition, then the obtained configuration is again reachable. Item 2 implies that $|\overline{\alpha}^{[\omega]}\rangle = |\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^e\rangle$ for every $e \in \mathbb{N}$: no matter how many copies of $\overline{\beta}^{[\omega]}$ we have, the set of states that can be reached by popping the stack is the same. Item 3 says that there are only two ways of popping $\overline{\beta}^{[\omega]}$: either we read at least one action symbol, or we end in a state from which there is an ε-run popping arbitrarily many copies of $\overline{\beta}^{[\omega]}$ and then γ; in particular, the only way of accessing the γ part of $\overline{\delta}^{[\omega]}$ is by an ε-run popping the whole γ.

There are some similarities between our pumping triples $\langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ and notions from Valiant's paper [21]. Namely, the $\overline{\beta}$ part of such a triple, which can be pumped, corresponds to Valiant's notion of a null-transparent segment, and the γ part corresponds to Valiant's notion of an invisible segment (it is "invisible" after repeating $\overline{\beta}$ many times). There is one important difference, though: in Valiant's proof it is enough to find just one invisible segment and one null-transparent segment; here, we rather need to decompose the whole stack into such segments.

We now give two simple lemmata, being immediate consequences of well-formedness:

LEMMA 5.1. *For every well-formed pumping triple $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ and every set $P \subseteq Q$ we have*
1. *$|(\overline{\beta}^{[\omega]})^\omega\rangle_\varepsilon(P) = |\overline{\beta}^{[\omega]}\rangle_\varepsilon(P)$,*
2. *$|\overline{\delta}^{[\omega]}\rangle_\varepsilon(P) = |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\gamma\rangle_\varepsilon(P)$, and*
3. *$|\overline{\delta}^{[\omega]}\rangle(P) = |\overline{\beta}^{[\omega]}\gamma\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle(P))$ for all $P \subseteq Q$.*

Proof. For the left-to-right inclusion in Item 1, suppose that $r \in |(\overline{\beta}^{[\omega]})^\omega\rangle_\varepsilon(r_0)$ for some $r_0 \in P$, which by definition means that $r_0(\overline{\beta}^{[\omega]})^\omega \rightarrow_\varepsilon^* r$. We have to prove that $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_0)$. The only way of popping $(\overline{\beta}^{[\omega]})^\omega$ is to pop $\overline{\beta}^{[\omega]}$ some number of times, and then to pop the whole $(\overline{\beta}^{[\omega]})^\omega$ in a single step, without changing the state; we thus have states $r_1, \ldots, r_k$ for some $k \in \mathbb{N}$ such that $r_i \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_{i-1})$ for $i \in [1, k]$, and $r_k = r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_k)$.

If $k = 0$, the latter gives us the thesis. If $k \geq 1$, for $i \in [1, k-1]$ we have $r_i \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_{i-1}) \subseteq |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$, so well-formedness of $\overline{\delta}$ implies that $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_i) = \{r_i\}$; because $r_{i+1} \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_i)$, necessarily $r_{i+1} = r_i$. We thus have $r_1 = r_2 = \cdots = r_k = r$, hence $r_1 \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_0)$ gives us the thesis.

Conversely, if $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_0)$ for some $r_0 \in P$, then by well-formedness of $\overline{\delta}$ we also have $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r)$. We thus have $r_0(\overline{\beta}^{[\omega]})^\omega \to_\varepsilon^* r(\overline{\beta}^{[\omega]})^\omega \to_\varepsilon^* r$, that is, $r \in |(\overline{\beta}^{[\omega]})^\omega\rangle_\varepsilon(r_0)$. This finishes the proof of Item 1.

We have $|\overline{\delta}^{[\omega]}\rangle_\varepsilon(P) = |\gamma\rangle_\varepsilon(|(\overline{\beta}^{[\omega]})^\omega\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle_\varepsilon(P)))$ and $|\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\gamma\rangle_\varepsilon(P) = |\gamma\rangle_\varepsilon(|\overline{\beta}^{[\omega]}\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle_\varepsilon(P)))$, so Item 2 is an immediate consequence of Item 1.

In order to prove that $|\overline{\delta}^{[\omega]}\rangle(P) \subseteq |\overline{\beta}^{[\omega]}\gamma\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle(P))$ consider a run that pops $\overline{\delta}^{[\omega]}$. Such a run pops $\overline{\alpha}^{[\omega]}$ and k copies of $\overline{\beta}^{[\omega]}$, for some $k \in \mathbb{N}$, then it pops the whole $(\overline{\beta}^{[\omega]})^\omega$ while staying in some state r, and then pops γ going to some state s. Then $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r) \subseteq |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$ and $s \in |\gamma\rangle(r)$. By well-formedness of $\overline{\delta}$ we then have $r \in |\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^k\rangle(P) = |\overline{\alpha}^{[\omega]}\rangle(P)$ and $|\gamma\rangle_\varepsilon(r) \neq \emptyset$, which implies $s \in |\gamma\rangle_\varepsilon(r)$ by Lemma 4.3. In consequence $s \in |\overline{\beta}^{[\omega]}\gamma\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle(P))$, as needed.

Conversely, consider some $s \in |\overline{\beta}^{[\omega]}\gamma\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle(P))$. Then there is a state $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle(P)) \subseteq |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$ for which $s \in |\gamma\rangle_\varepsilon(r)$. By well-formedness of $\overline{\delta}$ also $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r)$. We can thus pop $\overline{\alpha}^{[\omega]}$ and one copy of $\overline{\beta}^{[\omega]}$ while going to state r, then pop the whole $(\overline{\beta}^{[\omega]})^\omega$ remaining in state r, and then pop γ going to state s; we have $s \in |\overline{\delta}^{[\omega]}\rangle(P)$, as needed. □

LEMMA 5.2. *For every well-formed decomposition $\overline{\delta}$, every set $P \subseteq Q$, and every $e \in \mathbb{N}$ we have*
1. $|\overline{\delta}^{[\omega]}\rangle_\varepsilon(P) = |\overline{\delta}^{[1]}\rangle_\varepsilon(P)$,
2. $|\overline{\delta}^{[\omega]}\rangle(P) \subseteq |\overline{\delta}^{[e]}\rangle(P)$, *and*
3. $\mathsf{up}(\overline{\delta}^{[e]}) = \mathsf{up}(\overline{\delta}^{[1]})$ *and* $\mathsf{up}_0(\overline{\delta}^{[e]}) = \mathsf{up}_0(\overline{\delta}^{[1]})$.

Proof. Induction on the structure of $\overline{\delta}$. The base case of $\overline{\delta} = X \in \Gamma$ is trivial, because $\overline{\delta}^{[\omega]} = \overline{\delta}^{[e]} = \overline{\delta}^{[1]} = X$. In the case of $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\beta}_k$ the thesis follows directly from the induction hypothesis, because we can write $|\overline{\delta}^{[\omega]}\rangle(P) = |\overline{\delta}_k^{[\omega]}\rangle(|\overline{\delta}_{k-1}^{[\omega]}\rangle(\ldots(|\overline{\delta}_1^{[\omega]}\rangle(P))\ldots))$, and likewise for the other functions and relations. For $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ we have, using Lemma 5.1, the induction hypothesis, and well-formedness of $\overline{\delta}$:

$$|\overline{\delta}^{[\omega]}\rangle_\varepsilon(P) = |\gamma\rangle_\varepsilon(|\overline{\beta}^{[\omega]}\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle_\varepsilon(P))) = |\gamma\rangle_\varepsilon(|\overline{\beta}^{[1]}\rangle_\varepsilon(|\overline{\alpha}^{[1]}\rangle_\varepsilon(P))) = |\overline{\delta}^{[1]}\rangle_\varepsilon(P)\,,$$

$$|\overline{\delta}^{[\omega]}\rangle(P) = |\gamma\rangle_\varepsilon(|\overline{\beta}^{[\omega]}\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle(P))) \subseteq |\gamma\rangle(|\overline{\beta}^{[\omega]}\rangle(|\overline{\alpha}^{[\omega]}\rangle(P)))$$

$$= |\gamma\rangle(\underbrace{|\overline{\beta}^{[\omega]}\rangle(\ldots(|\overline{\beta}^{[\omega]}\rangle}_{e}(|\overline{\alpha}^{[\omega]}\rangle(P)))\ldots))$$

$$\subseteq |\gamma\rangle(\underbrace{|\overline{\beta}^{[e]}\rangle(\ldots(|\overline{\beta}^{[e]}\rangle}_{e}(|\overline{\alpha}^{[e]}\rangle(P)))\ldots)) = |\overline{\delta}^{[e]}\rangle(P)\,,$$

$$\mathsf{up}(\overline{\delta}^{[e]}) = \mathsf{up}(\gamma) \circ \underbrace{\mathsf{up}(\overline{\beta}^{[e]}) \circ \cdots \circ \mathsf{up}(\overline{\beta}^{[e]})}_{e} \circ \mathsf{up}(\overline{\alpha}^{[e]})$$

$$= \mathsf{up}(\gamma) \circ \underbrace{\mathsf{up}(\overline{\beta}^{[1]}) \circ \cdots \circ \mathsf{up}(\overline{\beta}^{[1]})}_{e} \circ \mathsf{up}(\overline{\alpha}^{[1]}) = \mathsf{up}(\gamma) \circ \mathsf{up}(\overline{\beta}^{[1]}) \circ \mathsf{up}(\overline{\alpha}^{[1]}) = \mathsf{up}(\overline{\delta}^{[1]})\,,$$

and likewise for $\mathsf{up}_0(\cdot)$. □

The next lemma is very important. It allows us to find a decomposition of bounded degree and height for an arbitrarily large stack content δ. This means that arbitrarily large segments of δ have to fit into the γ parts of pumping triples in the decomposition.

LEMMA 5.3. *Every nonempty stack content $\delta \in \Gamma^*$ has a well-formed decomposition $\overline{\delta}$ of height at most $2|Q| + 2$ and degree at most C, for some $\mathsf{C} \in 2^{|\mathcal{P}|^{O(1)}}$.*

Proof. Take $\mathsf{C} = 2^{|Q|^2(|\Gamma|^2+1)+1} + 1$, and denote $h(\alpha) = |Q \setminus |\alpha\rangle_\varepsilon(Q)|$ for $\alpha \in \Gamma^*$. We strengthen the thesis of the lemma, showing that the height of the resulting decomposition is at most $2h(\delta) + 2$ (which is not greater than $2|Q| + 2$). The proof is by induction on $h(\delta)$.

We represent the stack content δ as $\delta = \delta_1 X_1 \delta_2 X_2 \ldots \delta_n X_n \delta_{n+1}$ (where $\delta_i \in \Gamma^*$ and $X_i \in \Gamma$) in such a way that either $h(\delta_i) < h(\delta)$ or $\delta_i = \varepsilon$ for all $i \in [1, n+1]$, and $h(\delta_i X_i) \geq h(\delta)$ for all $i \in [1, n]$. To obtain such a representation, as δ_1 we take the longest prefix of δ satisfying $h(\delta_1) < h(\delta)$, or just $\delta_1 = \varepsilon$ if there is no such prefix; as X_1 we take the next letter for which by maximality of δ_1 we have $h(\delta_1 X_1) \geq h(\delta)$; then as δ_2 we take the longest prefix of the remaining part of δ satisfying..., and so on.

By the induction hypothesis, for every nonempty fragment δ_i we have a well-formed decomposition $\overline{\delta}_i$ of height at most $2h(\delta_i) + 2 \leq 2h(\delta)$ and degree at most C. Note that when $h(\delta) = 0$, then all δ_i are empty (we cannot have $h(\delta_i) < 0$), hence the induction hypothesis is not needed in this base case.

If $n \leq 2^{|Q|^2(|\Gamma|^2+1)}$, as $\overline{\delta}$ we just take the sequence consisting of the decompositions $\overline{\delta}_i$ for nonempty δ_i, and of single-node decompositions for symbols X_i; this sequence has length at most $2n + 1 \leq \mathsf{C}$. Well-formedness of $\overline{\delta}$ follows immediately from well-formedness of the decompositions $\overline{\delta}_i$.

Suppose that $n > 2^{|Q|^2(|\Gamma|^2+1)}$. For $k \leq \ell \leq n$ let $\overline{\delta}_{k,\ell}$ be the well-formed decomposition of $\delta_k X_k \ldots \delta_\ell X_\ell$, obtained as the sequence consisting of the decompositions $\overline{\delta}_i$ for nonempty δ_i, and of single-node decompositions for symbols X_i, where $i \in [k, \ell]$. To every $i \in [1, n]$ we assign the tuple

$$S_i = \left((|\overline{\delta}_{1,i}^{[\omega]}\rangle(q))_{q \in Q}, \mathsf{up}(\delta_{i+1} X_{i+1} \ldots \delta_n X_n \delta_{n+1})\right).$$

The first $|Q|$ components are subsets of Q, and the last component is a subset of $(Q \times \Gamma)^2$, so there are at most $2^{|Q|^2} \cdot 2^{|Q|^2 \cdot |\Gamma|^2}$ possible tuples S_i. Thus, there are some k, ℓ with $k < \ell \leq 2^{|Q|^2(|\Gamma|^2+1)} + 1 \leq n$ for which $S_k = S_\ell$. Let

$$\alpha = \delta_1 X_1 \ldots \delta_k X_k, \qquad \beta = \delta_{k+1} X_{k+1} \ldots \delta_\ell X_\ell, \qquad \text{and} \qquad \gamma = \delta_{\ell+1} X_{\ell+1} \ldots \delta_n X_n \delta_{n+1}.$$

Then $\overline{\alpha} = \overline{\delta}_{1,k}$ and $\overline{\beta} = \overline{\delta}_{k+1,\ell}$ are well-formed decompositions for α and β, respectively. These decompositions are sequences of length at most $2k$ and $2(\ell - k)$, respectively, which is smaller than C. Moreover, they both have height at most $2h(\delta) + 1$. Then $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ is a decomposition of δ, has degree at most C, and height a most $2h(\delta) + 2$, so it remains to see that $\overline{\delta}$ is well-formed.

Observe that

$$S_k = \left((|\overline{\alpha}^{[\omega]}\rangle(q))_{q \in Q}, \mathsf{up}(\overline{\beta}^{[1]}\gamma)\right) \qquad \text{and} \qquad S_\ell = \left((|\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle(q))_{q \in Q}, \mathsf{up}(\gamma)\right).$$

Equality of the last components of S_k and S_ℓ gives us the first condition of well-formedness, namely $\mathsf{up}(\gamma) = \mathsf{up}(\overline{\beta}^{[1]}\gamma)$. Equalities $|\overline{\alpha}^{[\omega]}\rangle(q) = |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle(q)$ for all $q \in Q$ give us $|\overline{\alpha}^{[\omega]}\rangle(P) = \bigcup_{q \in P} |\overline{\alpha}^{[\omega]}\rangle(q) = \bigcup_{q \in P} |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle(q) = |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle(P)$ for every set $P \subseteq Q$, which is the second condition of well-formedness.

Recall that $h(\delta_\ell X_\ell) \geq h(\delta)$; equivalently $||\delta_\ell X_\ell\rangle_\varepsilon(Q)| = |Q| - h(\delta_\ell X_\ell) \leq |Q| - h(\delta) = ||\delta\rangle_\varepsilon(Q)|$. Because $|\alpha\beta\rangle_\varepsilon(Q) = |\beta\rangle_\varepsilon(|\alpha\rangle_\varepsilon(Q)) \subseteq |\beta\rangle_\varepsilon(Q) = |\delta_\ell X_\ell\rangle_\varepsilon(|\delta_k X_k \ldots \delta_{\ell-1} X_{\ell-1}\rangle_\varepsilon(Q)) \subseteq |\delta_\ell X_\ell\rangle_\varepsilon(Q)$, this implies $||\alpha\beta\rangle_\varepsilon(Q)| \leq ||\beta\rangle_\varepsilon(Q)| \leq ||\delta\rangle_\varepsilon(Q)|$. Consider now the relation $\mathcal{R} = \{(r, s) \in |\alpha\beta\rangle_\varepsilon(Q) \times |\delta\rangle_\varepsilon(Q) \mid s \in |\gamma\rangle_\varepsilon(r)\}$. By Lemma 4.3 this is a partial function (for every r the set $|\gamma\rangle_\varepsilon(r)$ contains at most one s). On the other hand, for every state $s \in |\delta\rangle_\varepsilon(Q) = |\gamma\rangle_\varepsilon(|\alpha\beta\rangle_\varepsilon(Q))$ there exists $r \in |\alpha\beta\rangle_\varepsilon(Q)$ such that $s \in |\gamma\rangle_\varepsilon(r)$. Together with the cardinality argument $||\alpha\beta\rangle_\varepsilon(Q)| \leq ||\delta\rangle_\varepsilon(Q)|$ this implies that $\mathcal{R}$ is a bijection. Thus, the three sets, $|\alpha\beta\rangle_\varepsilon(Q)$, $|\beta\rangle_\varepsilon(Q)$, and $|\delta\rangle_\varepsilon(Q)$, are of the same size. Recalling that $|\alpha\beta\rangle_\varepsilon(Q)$ is a subset of $|\beta\rangle_\varepsilon(Q)$, this implies that actually $|\alpha\beta\rangle_\varepsilon(Q) = |\beta\rangle_\varepsilon(Q)$. Taking into account Lemma 5.2 (which can be used because $\overline{\alpha}$ and $\overline{\beta}$ are both well-formed), we moreover have $|\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q) = |\alpha\beta\rangle_\varepsilon(Q) = |\beta\rangle_\varepsilon(Q) = |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$.

In order to show the last condition of well-formedness, consider now a state $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$. We need to show that $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r) = \{r\}$ and $|\gamma\rangle_\varepsilon(r) \neq \emptyset$. The latter follows immediately from the fact that $\mathcal{R}$ is a bijection: $r \in |\alpha\beta\rangle_\varepsilon(Q)$, so r has a corresponding element in $|\gamma\rangle_\varepsilon(r)$. In order to show the former, note that, due to $r \in |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$, there have to exist states $q \in Q$ and $r' \in |\overline{\alpha}^{[\omega]}\rangle_\varepsilon(q)$ such that $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r')$. By Lemma 4.3 we then have $|\overline{\alpha}^{[\omega]}\rangle(q) = |\overline{\alpha}^{[\omega]}\rangle_\varepsilon(q) = \{r'\}$, and we have already shown that $|\overline{\alpha}^{[\omega]}\rangle(q) = |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle(q)$ (cf. the second condition of well-formedness). Thus $r \in |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle_\varepsilon(q) \subseteq |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle(q) = \{r'\}$ implies $r = r'$, so we have $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r)$, as required (by Lemma 4.3 the set $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r)$ is a singleton). $\square$

An interesting property of decompositions is described by Lemma 5.4; it says that if we increase the number of copies of $\overline{\beta}^{[\omega]}$ in a configuration, then the configuration becomes more similar to a configuration with ω copies of $\overline{\beta}^{[\omega]}$. We assume here that $1 + \omega = \omega$.

LEMMA 5.4. *Let* $q \in Q$, *let* $\langle\!\langle\overline{\alpha},\overline{\beta},\gamma\rangle\!\rangle$ *be a well-formed pumping triple, let* $\eta \in \Gamma^{\mathsf{gs}}$, *and let* $M_i = \min\{\mathrm{EQLEV}(r(\overline{\beta}^{[\omega]})^i\gamma\eta, r(\overline{\beta}^{[\omega]})^\omega\gamma\eta) \mid r \in |\overline{\alpha}^{[\omega]}\rangle(q)\}$ *for all* $i \in \mathbb{N}$. *Then* $M_{i+1} \geq 1 + M_i$ *for all* $i \in \mathbb{N}$.

Proof. Fix some $i \in \mathbb{N}$. Let $r_{\min} \in |\overline{\alpha}^{[\omega]}\rangle(q)$ be a state for which $M_{i+1} = \mathrm{EQLEV}(r_{\min}(\overline{\beta}^{[\omega]})^{i+1}\gamma\eta, r_{\min}(\overline{\beta}^{[\omega]})^\omega\gamma\eta)$. We distinguish two cases.

Suppose first that $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_{\min}) = \emptyset$. As observed on Page 8, we have $r_{\min}\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \approx r_{\min}(\overline{\beta}^{[\omega]})^\omega\gamma\eta$. Thus by Lemma 4.6 we have

$$M_{i+1} = \mathrm{EQLEV}(r_{\min}\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^i\gamma\eta, r_{\min}\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta)$$
$$\geq 1 + \min\{\mathrm{EQLEV}(r(\overline{\beta}^{[\omega]})^i\gamma\eta, r(\overline{\beta}^{[\omega]})^\omega\gamma\eta) \mid r \in |\overline{\beta}^{[\omega]}\rangle(r_{\min})\}\,.$$

By well-formedness of the pumping triple we have $|\overline{\beta}^{[\omega]}\rangle(r_{\min}) \subseteq |\overline{\beta}^{[\omega]}\rangle(|\overline{\alpha}^{[\omega]}\rangle(q)) = |\overline{\alpha}^{[\omega]}\overline{\beta}^{[\omega]}\rangle(q) = |\overline{\alpha}^{[\omega]}\rangle(q)$, so for all $r \in |\overline{\beta}^{[\omega]}\rangle(r_{\min})$ we have $\mathrm{EQLEV}(r(\overline{\beta}^{[\omega]})^i\gamma\eta, r(\overline{\beta}^{[\omega]})^\omega\gamma\eta) \geq M_i$, which implies $M_{i+1} \geq 1 + M_i$.

Conversely, suppose that there is some $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r_{\min}) \subseteq |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$. By well-formedness of the pumping triple we have $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r) = \{r\}$ and $s \in |\gamma\rangle_\varepsilon(r)$ for some s. This implies that $r_{\min}(\overline{\beta}^{[\omega]})^{i+1}\gamma\eta \to_\varepsilon^* r(\overline{\beta}^{[\omega]})^i\gamma\eta \to_\varepsilon^* s\eta$ and $r_{\min}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \to_\varepsilon^* r(\overline{\beta}^{[\omega]})^\omega\gamma\eta \to_\varepsilon^* s\eta$, that is, $M_{i+1} = \omega \geq 1 + M_i$. $\quad\square$

The next lemma says that if we can pop a stack content of the form $\overline{\delta}^{[\omega]}$, then we can pop it quickly:

LEMMA 5.5. *If* $s \in |\overline{\delta}^{[\omega]}\rangle(q)$ *for a well-formed decomposition* $\overline{\delta}$ *of degree at most* C *and height* ℓ, *then* $\mathrm{DIST}(q\overline{\delta}^{[\omega]}, s) \leq \mathsf{BC}^\ell$.

Proof. Induction on the structure of $\overline{\delta}$. Suppose that $s \in |\overline{\delta}^{[\omega]}\rangle(q)$. If $\overline{\delta} = X = \overline{\delta}^{[\omega]}$, then $s \in |X\rangle(q)$ (i.e., $qX \to^* s$) implies $\mathrm{DIST}(q\overline{\delta}^{[\omega]}, s) \leq \mathsf{B}$ by Lemma 4.1.

Suppose that $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$. Then $\overline{\delta}^{[\omega]} = \overline{\delta}_1^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}$, so there have to exist states $r_0, r_1, \ldots, r_k$ such that $r_0 = q$, $r_k = r$, and $r_i \in |\overline{\delta}_i^{[\omega]}\rangle(r_{i-1})$ for all $i \in [1, k]$. By the induction hypothesis $\mathrm{DIST}(r_{i-1}\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}, r_i\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}) \leq \mathrm{DIST}(r_{i-1}\overline{\delta}_i^{[\omega]}, r_i) \leq \mathsf{BC}^{\ell-1}$. Because $k \leq \mathsf{C}$, we obtain $\mathrm{DIST}(r_0\overline{\delta}^{[\omega]}, r_k) \leq \mathsf{C} \cdot \mathsf{BC}^{\ell-1}$.

Finally, suppose that $\overline{\delta} = \langle\!\langle\overline{\alpha},\overline{\beta},\gamma\rangle\!\rangle$. By Lemma 5.1, $s \in |\overline{\delta}^{[\omega]}\rangle(q)$ implies that there is a state $r \in |\overline{\alpha}^{[\omega]}\rangle(q)$ for which $s \in |\overline{\beta}^{[\omega]}\gamma\rangle_\varepsilon(r) = |(\overline{\beta}^{[\omega]})^\omega\gamma\rangle_\varepsilon(r)$. The latter means that $\mathrm{DIST}(r(\overline{\beta}^{[\omega]})^\omega\gamma, s) = 0$. We conclude using the induction hypothesis:

$$\mathrm{DIST}(q\overline{\delta}^{[\omega]}, s) \leq \mathrm{DIST}(q\overline{\delta}^{[\omega]}, r(\overline{\beta}^{[\omega]})^\omega\gamma) + \mathrm{DIST}(r(\overline{\beta}^{[\omega]})^\omega\gamma, s) \leq \mathrm{DIST}(q\overline{\alpha}^{[\omega]}, r) + 0 \leq \mathsf{BC}^{\ell-1} \leq \mathsf{BC}^\ell\,. \quad\square$$

We now define by induction when two decompositions $\overline{\delta}, \widehat{\delta}$ have the same type, written $\overline{\delta} \overset{\mathsf{tp}}{=} \widehat{\delta}$:
- if $\overline{\delta} = \widehat{\delta}$ then $\overline{\delta} \overset{\mathsf{tp}}{=} \widehat{\delta}$,
- if $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$, and $\widehat{\delta} = \widehat{\delta}_1 \ldots \widehat{\delta}_k$, and $\overline{\delta}_i \overset{\mathsf{tp}}{=} \widehat{\delta}_i$ for all $i \in \{1, \ldots, k\}$, then $\overline{\delta} \overset{\mathsf{tp}}{=} \widehat{\delta}$, and
- if $\overline{\delta} = \langle\!\langle\overline{\alpha},\overline{\beta},\gamma\rangle\!\rangle$, and $\widehat{\delta} = \langle\!\langle\widehat{\alpha},\widehat{\beta},\gamma'\rangle\!\rangle$, and $\overline{\alpha} \overset{\mathsf{tp}}{=} \widehat{\alpha}$, and $\overline{\beta} \overset{\mathsf{tp}}{=} \widehat{\beta}$, and $|\gamma\rangle_\varepsilon = |\gamma'\rangle_\varepsilon$, and $\mathsf{up}(\gamma) = \mathsf{up}(\gamma')$, then $\overline{\delta} \overset{\mathsf{tp}}{=} \widehat{\delta}$.

The next two lemmata say that, on the one hand, the number of different types is small, and, on the other hand, that while considering stacks of the form $\overline{\delta}^{[\omega]}$, only the type of $\overline{\delta}$ is relevant:

LEMMA 5.6. *The number of different types (i.e., equivalence classes of the* $\overset{\mathsf{tp}}{=}$ *relation) for decompositions of height at most* ℓ *and degree at most* C *is at most* $2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^{\ell+1}}$.

Proof. Induction on ℓ. For $\ell = 0$ we have only $|\Gamma|$ decompositions, consisting of single stack symbols, and $|\Gamma| \leq 2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)}$. For $\ell \geq 1$ we have

- $|\Gamma|$ decompositions consisting of single stack symbols;
- decompositions being sequences of at most C decompositions of height at most $\ell - 1$; for each of them we have to know its type, where by the induction hypothesis we have at most $2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^\ell}$ possibilities;
- pumping triples $\langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$, where we have to know the types of $\overline{\alpha}$ and $\overline{\beta}$ (at most $2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^\ell}$ possibilities for each, by the induction hypothesis), and the objects $|\gamma\rangle_\varepsilon$ and $\mathsf{up}(\gamma)$.

In the last case, we notice that $|\gamma\rangle_\varepsilon(P) = \bigcup_{p \in P} |\gamma\rangle_\varepsilon(p)$ for every $P \subseteq Q$, meaning that the function $|\gamma\rangle_\varepsilon$ is determined by the set $\{(p,q) \mid q \in |\gamma\rangle_\varepsilon(p)\} \subseteq Q^2$. Likewise, $\mathsf{up}(\gamma)$ is a subset of $(Q\Gamma)^2$. We thus have at most $2^{|Q|^2 + |Q|^2 |\Gamma|^2} \leq 2^{|\mathcal{P}|^4}$ possibilities for combinations of $|\gamma\rangle_\varepsilon$ and $\mathsf{up}(\gamma)$. Using inequalities $\mathsf{C} \geq 3$ and $\mathsf{C} + 3 \leq 2^{\mathsf{C}} \leq 2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^\ell}$, we obtain that the number of possible types is at most

$$|\Gamma| + \sum\nolimits_{i=0}^{\mathsf{C}} 2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^\ell \cdot i} + 2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^\ell \cdot 2 + |\mathcal{P}|} \leq (\mathsf{C}+3) \cdot 2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^\ell \cdot \mathsf{C}} \leq 2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^{\ell+1}}. \qquad \square$$

LEMMA 5.7. *Let $\overline{\delta}, \widehat{\delta}$ be decompositions. If $\overline{\delta}$ is well-formed and $\overline{\delta} \stackrel{\mathsf{tp}}{=} \widehat{\delta}$, then*
1. $|\overline{\delta}^{[\omega]}\rangle = |\widehat{\delta}^{[\omega]}\rangle$ and $|\overline{\delta}^{[\omega]}\rangle_\varepsilon = |\widehat{\delta}^{[\omega]}\rangle_\varepsilon$,
2. $\mathsf{up}(\overline{\delta}^{[1]}) = \mathsf{up}(\widehat{\delta}^{[1]})$ and $\mathsf{up}_0(\overline{\delta}^{[1]}) = \mathsf{up}_0(\widehat{\delta}^{[1]})$,
3. $\widehat{\delta}$ is well-formed as well, and
4. $q\overline{\delta}^{[\omega]}\eta \approx q\widehat{\delta}^{[\omega]}\eta$ for all $q \in Q$ and $\eta \in \Gamma^{\mathsf{gs}}$.

Proof. We proceed by induction on the structure of $\overline{\delta}$. If $\overline{\delta} = X$, then also $\widehat{\delta} = X$, so the thesis is trivial.

Suppose that $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$. Then $\widehat{\delta} = \widehat{\delta}_1 \ldots \widehat{\delta}_k$ with $\overline{\delta}_i \stackrel{\mathsf{tp}}{=} \widehat{\delta}_i$ for all $i \in [1,k]$. Item 3 follows directly from the induction hypothesis. Items 1 and 2 as well: $|\overline{\delta}^{[\omega]}\rangle(P) = |\overline{\delta}_k^{[\omega]}\rangle(|\overline{\delta}_{k-1}^{[\omega]}\rangle(\ldots(|\overline{\delta}_1^{[\omega]}\rangle(P))\ldots)) = |\widehat{\delta}_k^{[\omega]}\rangle(|\widehat{\delta}_{k-1}^{[\omega]}\rangle(\ldots(|\widehat{\delta}_1^{[\omega]}\rangle(P))\ldots)) = |\widehat{\delta}^{[\omega]}\rangle(P)$ for every $P \subseteq Q$, and likewise for the other equalities. By the induction hypothesis we also have $r\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}\eta \approx r\widehat{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}\eta$ for all $r \in Q$ and $i \in [1,k]$, which implies $q\widehat{\delta}_1^{[\omega]} \ldots \widehat{\delta}_{i-1}^{[\omega]}\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}\eta \approx q\widehat{\delta}_1^{[\omega]} \ldots \widehat{\delta}_{i-1}^{[\omega]}\widehat{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}\eta$ by Lemma 4.5. Having this for all $i \in [1,k]$ we obtain $q\overline{\delta}^{[\omega]}\eta \approx q\widehat{\delta}^{[\omega]}\eta$, as needed for Item 4.

Finally, suppose that $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$. Then $\widehat{\delta} = \langle\!\langle \widehat{\alpha}, \widehat{\beta}, \gamma' \rangle\!\rangle$ with $\overline{\alpha} \stackrel{\mathsf{tp}}{=} \widehat{\alpha}$, $\overline{\beta} \stackrel{\mathsf{tp}}{=} \widehat{\beta}$, $|\gamma\rangle_\varepsilon = |\gamma'\rangle_\varepsilon$, and $\mathsf{up}(\gamma) = \mathsf{up}(\gamma')$. We obtain Item 1 using Lemma 5.1 and the induction hypothesis:

$$|\overline{\delta}^{[\omega]}\rangle(P) = |\gamma\rangle_\varepsilon(|\overline{\beta}^{[\omega]}\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle(P))) = |\gamma'\rangle_\varepsilon(|\widehat{\beta}^{[\omega]}\rangle_\varepsilon(|\widehat{\alpha}^{[\omega]}\rangle(P))) = |\widehat{\delta}^{[\omega]}\rangle(P), \qquad \text{and}$$

$$|\overline{\delta}^{[\omega]}\rangle_\varepsilon(P) = |\gamma\rangle_\varepsilon(|\overline{\beta}^{[\omega]}\rangle_\varepsilon(|\overline{\alpha}^{[\omega]}\rangle_\varepsilon(P))) = |\gamma'\rangle_\varepsilon(|\widehat{\beta}^{[\omega]}\rangle_\varepsilon(|\widehat{\alpha}^{[\omega]}\rangle_\varepsilon(P))) = |\widehat{\delta}^{[\omega]}\rangle_\varepsilon(P) \qquad \text{for all } P \subseteq Q.$$

Item 2 follows directly from the induction hypothesis and the equality $\mathsf{up}(\gamma) = \mathsf{up}(\gamma')$. We observe that the definition of well-formedness for $\overline{\delta}$ uses only the objects $\mathsf{up}(\gamma)$, $\mathsf{up}(\overline{\beta}^{[1]})$, $|\overline{\alpha}^{[\omega]}\rangle$, $|\overline{\beta}^{[\omega]}\rangle$, $|\overline{\beta}^{[\omega]}\rangle_\varepsilon$, and $|\gamma\rangle_\varepsilon$, and for each of them we already have equality with analogous object for the other decomposition; Item 3 follows.

In order to obtain Item 4, we prove that

$$r(\overline{\beta}^{[\omega]})^\omega \gamma\eta \approx_k r(\widehat{\beta}^{[\omega]})^\omega \gamma'\eta \qquad \text{for all } r \in Q \text{ and } k \in \mathbb{N}.$$

We proceed by induction on k. For $k = 0$ this is trivial, as all configurations are related by $\approx_0$. Suppose now that the thesis holds for some k, and we want to prove it for $k+1$. As observed on Page 8, we have $r(\overline{\beta}^{[\omega]})^\omega \gamma\eta \approx r\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega \gamma\eta$ and $r(\widehat{\beta}^{[\omega]})^\omega \gamma'\eta \approx r\widehat{\beta}^{[\omega]}(\widehat{\beta}^{[\omega]})^\omega \gamma'\eta$, and by the induction hypothesis of the external induction we have $r\widehat{\beta}^{[\omega]}(\widehat{\beta}^{[\omega]})^\omega \gamma'\eta \approx r\overline{\beta}^{[\omega]}(\widehat{\beta}^{[\omega]})^\omega \gamma'\eta$. We thus have to prove that

$$r\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega \gamma\eta \approx_{k+1} r\overline{\beta}^{[\omega]}(\widehat{\beta}^{[\omega]})^\omega \gamma'\eta.$$

If $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r) = \emptyset$, this follows by Lemma 4.6 from the induction hypothesis saying that $r'(\overline{\beta}^{[\omega]})^\omega \gamma\eta \approx_k r'(\widehat{\beta}^{[\omega]})^\omega \gamma'\eta$ for all $r' \in |\overline{\beta}^{[\omega]}\rangle(r) \subseteq Q$. Suppose thus that $|\overline{\beta}^{[\omega]}\rangle_\varepsilon(r) \neq \emptyset$, and take some state $r' \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r) \subseteq |\overline{\beta}^{[\omega]}\rangle_\varepsilon(Q)$. By well-formedness of $\overline{\delta}$ we have $r' \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r') = |\widehat{\beta}^{[\omega]}\rangle_\varepsilon(r')$ and $s \in |\gamma\rangle_\varepsilon(r') = |\gamma'\rangle_\varepsilon(r')$ for some s. In consequence

$r\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \to_\varepsilon^* r'(\overline{\beta}^{[\omega]})^\omega\gamma\eta \to_\varepsilon^* r'\gamma\eta \to_\varepsilon^* s\eta$ and $r\overline{\beta}^{[\omega]}(\widehat{\beta}^{[\omega]})^\omega\gamma'\eta \to_\varepsilon^* r'(\widehat{\beta}^{[\omega]})^\omega\gamma'\eta \to_\varepsilon^* r'\gamma'\eta \to_\varepsilon^* s\eta$, so we even have $r\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \approx s\eta \approx \overline{\beta}^{[\omega]}(\widehat{\beta}^{[\omega]})^\omega\gamma'\eta$.

Because $\bigcap_{k\in\mathbb{N}}(\approx_k) = (\approx)$, from the above it follows that $r(\overline{\beta}^{[\omega]})^\omega\gamma\eta \approx r(\widehat{\beta}^{[\omega]})^\omega\gamma'\eta$ for all $r \in Q$, so $q\widehat{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \approx q\widehat{\alpha}^{[\omega]}(\widehat{\beta}^{[\omega]})^\omega\gamma'\eta$ by Lemma 4.5. We conclude using the induction hypothesis, saying that $q\overline{\delta}^{[\omega]}\eta = q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \approx q\widehat{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta$. □

6 Runs going down

In this section we concentrate on runs "going down", that is, runs oriented on popping the stack. Such runs are called D-almost-popping runs, and classes reachable by them are collected in $Below(\cdot)$ sets. We also prove basic properties of these notions. As in previous sections, we assume some fixed ε-PDS $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$.

Let $\mathsf{D} = (\mathsf{B}+3)\cdot\mathsf{BC}^{2|Q|+1}+2$ (it turns out that this value of D is suitable for our proofs presented in the next section). Note that $\mathsf{D} \in 2^{|\mathcal{P}|^{O(1)}}$. We say that a run π is D-*almost-popping* if there is no run π'_D that is parallel to a suffix of π and satisfies $\textsc{StackGrowth}(\pi'_\mathsf{D}) \geq \mathsf{D}$. For a set Ω of classes, we define $Below(\Omega)$ to be the set of classes $\mathcal{C}$ for which there is a D-almost-popping run π from a configuration in a class in Ω to a configuration in $\mathcal{C}$. When Ω is a set of configurations (instead of classes), we abbreviate $Below(\{[c]_\approx \mid c \in \Omega\})$ into $Below(\Omega)$, and for a single configuration c we abbreviate $Below(\{c\})$ into $Below(c)$.

Observe that if π is D-almost-popping, then every run parallel to π is D-almost-popping as well. It follows that if from a configuration c there exists a D-almost-popping run to a configuration in a class $\mathcal{C}$, then such a run exists from every configuration $c' \in [c]_\approx$. Observe also that every suffix of a D-almost-popping run is D-almost-popping as well. The next four lemmata bound the size of the $Below(\cdot)$ sets in specific situations:

LEMMA 6.1. *Let $\overline{\delta}$ be a well-formed decomposition of degree at most C and height ℓ, and let $\eta \in \Gamma^{\mathsf{gs}}$. Then there are at most $\mathsf{C}^\ell \cdot |\mathcal{P}|^{\mathsf{D}+1}$ classes $\mathcal{C}$ for which there is a D-almost-popping run π that starts in $p\overline{\delta}^{[\omega]}\eta$ for some $p \in Q$, leads to a configuration in $\mathcal{C}$, and can be shifted to a run from $p\overline{\delta}^{[\omega]}$. In particular, for every $P \subseteq Q$,*

$$|Below(\{p\overline{\delta}^{[\omega]}\eta \mid p \in P\})| \leq |Below(\{r\eta \mid r \in |\overline{\delta}^{[\omega]}\rangle(P)\})| + \mathsf{C}^\ell \cdot |\mathcal{P}|^{\mathsf{D}+1}.$$

Proof. The second part of the lemma easily follows from the first part. Indeed, consider a class $\mathcal{C} \in Below(\{p\overline{\delta}^{[\omega]}\eta \mid p \in P\})$. By definition, there exists a D-almost-popping run from $p\overline{\delta}^{[\omega]}\eta$ for some $p \in P \subseteq Q$ to a configuration in $\mathcal{C}$. If π can be shifted to a run from $p\overline{\delta}^{[\omega]}$, then $\mathcal{C}$ is among the at most $\mathsf{C}^\ell \cdot |\mathcal{P}|^{\mathsf{D}+1}$ classes counted by the first part of the lemma. If not, then π reaches a configuration with stack η, at some moment. For the first such configuration $r\eta$ we have $r \in |\overline{\delta}^{[\omega]}\rangle(P)$, so the second part of π, from $r\eta$ to a configuration in $\mathcal{C}$, witnesses that $\mathcal{C} \in Below(\{r\eta \mid r \in |\overline{\delta}^{[\omega]}\rangle(P)\})$.

We prove the first part by induction on the structure of $\overline{\delta}$. Suppose first that $\overline{\delta} = X \in \Gamma$, and consider a D-almost-popping run π that starts in $pX\eta$ for some $p \in Q$, and can be shifted to a run from pX. Necessarily $\textsc{StackGrowth}(\pi) < \mathsf{D}$ (because π itself is also parallel to a suffix of π). Thus π leads to a configuration $q\mu\eta$ with $|\mu| = |X| + \textsc{StackGrowth}(\pi) \leq \mathsf{D}$. The number of such configurations $q\mu\eta$, hence also the number of classes of these configurations, is at most $|Q| \cdot \sum_{i=0}^{\mathsf{D}} |\Gamma|^i \leq |Q| \cdot (|\Gamma|+1)^\mathsf{D} \leq |\mathcal{P}|^{\mathsf{D}+1}$ (recall that η is fixed).

Next, suppose that $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$. For every D-almost-popping run that starts in $p\overline{\delta}^{[\omega]}\eta$ and can be shifted to a run from $p\overline{\delta}^{[\omega]}$ we can find the smallest $i \in [1, k]$ such that π can be shifted to a run from $p\overline{\delta}_1^{[\omega]} \ldots \overline{\delta}_i^{[\omega]}$. By minimality of i, the run π crosses a configuration of the form $p'\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}\eta$; then the suffix π' of π starting there is a D-almost-popping run that starts in $p'\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]} \ldots \overline{\delta}_k^{[\omega]}\eta$ and can be shifted to a run from $p'\overline{\delta}_i^{[\omega]}$. By the induction hypothesis, the number of classes reached by such runs π', for every fixed i, is at most $\mathsf{C}^{\ell-1} \cdot |\mathcal{P}|^{\mathsf{D}+1}$. There are C possible values for i, so in total we have at most $\mathsf{C}^\ell \cdot |\mathcal{P}|^{\mathsf{D}+1}$ classes.

Finally, suppose that $\overline{\delta} = \langle\!\langle\overline{\alpha}, \overline{\beta}, \gamma\rangle\!\rangle$. For every D-almost-popping run π that starts in $p\overline{\delta}^{[\omega]}\eta$ and can be shifted to a run from $p\overline{\delta}^{[\omega]}$ we have one of the following possibilities:
- Maybe π can be shifted to a run from $p\overline{\alpha}^{[\omega]}$. By the induction hypothesis, the number of classes reached by such runs π is at most $\mathsf{C}^{\ell-1} \cdot |\mathcal{P}|^{\mathsf{D}+1}$.

- Maybe π cannot be shifted to a run from $p\overline{\alpha}^{[\omega]}$ (hence π visits a configuration with stack $(\overline{\beta}^{[\omega]})^\omega\gamma\eta$), but π never pops $(\overline{\beta}^{[\omega]})^\omega$ going to a configuration with stack $\gamma\eta$. Consider the last moment when π visits a configuration of the form $r(\overline{\beta}^{[\omega]})^\omega\gamma\eta$; let π' be the suffix of π starting at this moment. Let also π'' be the run starting in $r\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta$ and performing exactly the same transitions as π'; after performing the first transition the two runs coincide. We have $r\overline{\beta}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \approx r(\overline{\beta}^{[\omega]})^\omega\gamma\eta$, so π' ans π'' are parallel; in particular π'' is D-almost-popping. Moreover, π'' can be shifted to a run from $r\overline{\beta}^{[\omega]}$, because otherwise π'' and π' would visit a configuration with stack $(\overline{\beta}^{[\omega]})^\omega\gamma\eta$ again, contradicting the maximality of the starting point. Thus, by the induction hypothesis, the number of classes reached by such runs π'' is at most $\mathsf{C}^{\ell-1}\cdot|\mathcal{P}|^{\mathsf{D}+1}$.
- Finally, maybe at some moment π reaches $r(\overline{\beta}^{[\omega]})^\omega\gamma\eta$, then goes to $r\gamma\eta$, and then continues somehow. Because $(\overline{\beta}^{[\omega]})^\omega$ could be popped, we have $r \in |\overline{\beta}^{[\omega]}\rangle_\varepsilon(r)$, hence by well-formedness of $\overline{\delta}$ there exists some $s \in |\gamma\rangle_\varepsilon(r)$; in other words, $r\gamma \to_\varepsilon^* s$ for some state s. Recall that π can be shifted to a run from $p\overline{\delta}^{[\omega]}$, hence its suffix after visiting $r\gamma\eta$ can be shifted to a run ϱ from $r\gamma$. Thus, by determinism of ε-transitions, ϱ consists entirely of ε-transitions. In consequence, the prefix π' of π ending in $r(\overline{\beta}^{[\omega]})^\omega\gamma\eta$ leads to the same class as π; moreover, π' is parallel to π, so it is D-almost-popping as well. But π' satisfies assumptions of the previous case; no new classes can be reached.

Thus, the total number of classes under consideration is at most $2\mathsf{C}^{\ell-1}\cdot|\mathcal{P}|^{\mathsf{D}+1} \leq \mathsf{C}^\ell\cdot|\mathcal{P}|^{\mathsf{D}+1}$.   □

LEMMA 6.2. *Let $\overline{\delta}_1,\ldots,\overline{\delta}_k$ be well-formed decompositions of degree at most C and height ℓ, and let $\eta \in \Gamma^{\mathsf{gs}}$. Then there are at most $k\cdot\mathsf{C}^\ell\cdot|\mathcal{P}|^{\mathsf{D}+1}$ classes $\mathcal{C}$ for which there is a D-almost-popping run π that starts in $p\overline{\delta}_1^{[\omega]}\ldots\overline{\delta}_k^{[\omega]}\eta$ for some $p \in Q$, leads to a configuration in $\mathcal{C}$, and can be shifted to a run from $p\overline{\delta}_1^{[\omega]}\ldots\overline{\delta}_k^{[\omega]}$.*

Proof. As in the proof of Lemma 6.1 (the case of $\overline{\delta} = \overline{\delta}_1\ldots\overline{\delta}_k$) we observe that for every class $\mathcal{C}$ under consideration there is a D-almost-popping run that starts in $p'\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[\omega]}\ldots\overline{\delta}_k^{[\omega]}\eta$, ends in a configuration in $\mathcal{C}$, and can be shifted to a run from $p'\overline{\delta}_i^{[\omega]}$, for some $i \in [1,k]$. By Lemma 6.1, the number of such classes for every fixed i is at most $\mathsf{C}^\ell\cdot|\mathcal{P}|^{\mathsf{D}+1}$. There are k possible values for i, so in total we have at most $k\cdot\mathsf{C}^\ell\cdot|\mathcal{P}|^{\mathsf{D}+1}$ classes.   □

LEMMA 6.3. *Let π be a D-almost-popping run. Then there are at most $|\mathcal{P}|^{2\mathsf{BD}}-1$ classes $\mathcal{C}$ for which there exists a D-almost-popping run π' from $\pi(|\pi|)$ to a configuration in $\mathcal{C}$, but the composition $\pi \circ \pi'$ is not D-almost-popping.*

Proof. We prove that every class $\mathcal{C}$ under consideration satisfies $\mathrm{DIST}([\pi(|\pi|)]_\approx,\mathcal{C}) \leq \mathsf{B}\cdot(2\mathsf{D}-1)$; by Lemma 4.4 this bounds the number of possible classes $\mathcal{C}$ by $|\mathcal{P}|^{\mathsf{B}\cdot(2\mathsf{D}-1)} \leq |\mathcal{P}|^{2\mathsf{BD}}-1$.

Consider thus a D-almost-popping run π' that starts in $\pi(|\pi|)$, ends in a class $\mathcal{C}$, and is such that $\pi \circ \pi'$ is not D-almost-popping. By definition we then have a run ϱ_D that is parallel to a suffix of $\pi \circ \pi'$ and satisfies $\mathrm{STACKGROWTH}(\varrho_\mathsf{D}) \geq \mathsf{D}$. Because π' is D-almost-popping, ϱ_D cannot be parallel to a suffix of π'. We can thus represent ϱ_D as $\varrho_1 \circ \varrho_\mathsf{D}'$, where ϱ_1 is parallel to a suffix of π, and ϱ_D' is parallel to π'. Let also ϱ_3 be a suffix of ϱ_D' for which $\mathrm{STACKGROWTH}(\varrho_3)$ is maximal; let us write $\varrho_\mathsf{D}' = \varrho_2 \circ \varrho_3$. Denote $k_i = \mathrm{STACKGROWTH}(\varrho_i)$ for $i \in \{1,2,3\}$. Because π and π' are D-almost-popping, necessarily $k_1 \leq \mathsf{D}-1$ and $k_3 \leq \mathsf{D}-1$ (because ϱ_1 and ϱ_3 are parallel to suffixes of π and π', respectively). Recall also that $\mathrm{STACKGROWTH}(\varrho_\mathsf{D}) = k_1 + k_2 + k_3 \geq \mathsf{D}$, so $k_1 + k_2 \geq 1$, implying that on the top of the stack of $\varrho_3(0)$ we have a real stack symbol, pushed there by $\varrho_1 \circ \varrho_2$ (i.e., not an expression of the form γ^ω). Maximality of $\mathrm{STACKGROWTH}(\varrho_3)$ means that we can then write $\varrho_2(0) = p\alpha\eta$, $\varrho_3(0) = qX\eta$, and $\varrho_3(|\varrho_3|) = r\beta\eta$ (for $\alpha,\beta \in \Gamma^*$), where $|X|-|\alpha| = k_2$ and $|\beta|-|X| = k_3$, and $\varrho_2 \circ \varrho_3$ can be shifted to a run from $p\alpha$ to $r\beta$ (intuitively: the "stack height" during $\varrho_2 \circ \varrho_3$ is minimal at $\varrho_3(0)$). Due to $k_3 \leq \mathsf{D}-1$ we have $|\beta| = k_3 + 1 \leq \mathsf{D}$, and due to $k_1 \leq \mathsf{D}-1$ and $k_1 + k_2 \geq 1$ we have $|\alpha| = 1 - k_2 \leq k_1 \leq \mathsf{D}-1$. Thus

$$\mathrm{DIST}([\pi(|\pi|)]_\approx,\mathcal{C}) \leq \mathrm{DIST}(\varrho_2(0),\varrho_3(|\varrho_3|)) \leq \mathrm{DIST}(p\alpha,r\beta) \leq (|\alpha|+|\beta|)\cdot\mathsf{B} \leq \mathsf{B}\cdot(2\mathsf{D}-1),$$

as we wanted; the third inequality is by Lemma 4.1.   □

LEMMA 6.4. *For every set Ω of classes,*

$$|Below(Below(\Omega))| \leq |Below(\Omega)|\cdot|\mathcal{P}|^{2\mathsf{BD}}.$$

Proof. For every class $\mathcal{C} \in Below(\Omega)$ we fix some D-almost-popping run $\pi_{\mathcal{C}}$ from a configuration in a class in Ω to a configuration in $\mathcal{C}$. For every class $\mathcal{D} \in Below(Below(\Omega)) \setminus Below(\Omega)$ we fix a class $\mathcal{C}(\mathcal{D}) \in Below(\Omega)$ such that $\mathcal{D} \in Below(\mathcal{C}(\mathcal{D}))$, and we fix a D-almost-popping run $\pi'_{\mathcal{D}}$ from $\pi_{\mathcal{C}(\mathcal{D})}(|\pi_{\mathcal{C}(\mathcal{D})}|)$ to a configuration in $\mathcal{D}$ (recall that a D-almost-popping run to a configuration in $\mathcal{D}$ exists not only from some configuration in $\mathcal{C}(\mathcal{D})$, but actually from every configuration in $\mathcal{C}(\mathcal{D})$, in particular from $\pi_{\mathcal{C}(\mathcal{D})}(|\pi_{\mathcal{C}(\mathcal{D})}|)$). Because $\mathcal{D} \notin Below(\Omega)$, the composition $\pi_{\mathcal{C}(\mathcal{D})} \circ \pi'_{\mathcal{D}}$ is not D-almost-popping. Thus, for every class $\mathcal{C}_0$, the number of classes $\mathcal{D} \in Below(Below(\Omega)) \setminus Below(\Omega)$ with $\mathcal{C}(\mathcal{D}) = \mathcal{C}_0$ is at most $|\mathcal{P}|^{2\mathsf{BD}} - 1$, by Lemma 6.3. We then have at most $|Below(\Omega)|$ choices for the class $\mathcal{C}_0$. Additionally, $Below(Below(\Omega))$ may contain at most $|Below(\Omega)|$ classes from $Below(\Omega)$. We thus have $|Below(Below(\Omega))| \le |Below(\Omega)| \cdot (|\mathcal{P}|^{2\mathsf{BD}} - 1) + |Below(\Omega)|$, as required. $\square$

7 The core of the upper bound proof

Let us fix an ε-PDS $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$ together with an initial configuration $q_{\mathsf{init}} X_{\mathsf{init}} \in Q\Gamma$. Assuming that $(\mathcal{L}(\mathcal{P}), q_{\mathsf{init}} X_{\mathsf{init}})$ is weakly bisimulation finite, we want to bound the number of its classes reachable from $[q_{\mathsf{init}} X_{\mathsf{init}}]_{\approx}$, showing that it is in $2^{2^{|\mathcal{P}|^{O(1)}}}$.

Our first lemma says that if the ε-PDS is weakly bisimulation finite, then generalized configurations with ω repetitions of some stack fragment are equivalent to configurations with a large enough number $e \in \mathbb{N}$ of repetitions:

LEMMA 7.1. *Let $q \in Q$, let $\langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ be a well-formed pumping triple, and let $\eta \in \Gamma^*$. If $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\alpha}^{[1]}\overline{\beta}^{[1]}\gamma\eta$, and if $(\mathcal{L}(\mathcal{P}), q_{\mathsf{init}} X_{\mathsf{init}})$ is weakly bisimulation finite, then for some $e \in \mathbb{N}$ we have $r(\overline{\beta}^{[\omega]})^e \gamma\eta \approx r(\overline{\beta}^{[\omega]})^\omega \gamma\eta$ for all $r \in |\overline{\alpha}^{[\omega]}\rangle(q)$.*

Proof. We assume that $(\mathcal{L}(\mathcal{P}), q_{\mathsf{init}} X_{\mathsf{init}})$ is weakly bisimulation finite, and we denote the number of its classes by F. We prove the lemma for $e = F^{|Q|}$. Let us formulate a version of the lemma suitable for induction: for every $q \in Q$, every well-formed decomposition $\overline{\delta}$, and every $\eta \in \Gamma^*$, if $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\delta}^{[e]}\eta$, then $q\overline{\delta}^{[e]}\eta \approx q\overline{\delta}^{[\omega]}\eta$. We prove this statement by induction on the structure of $\overline{\delta}$, and in the case of $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$ we also prove that $r(\overline{\beta}^{[\omega]})^e \gamma\eta \approx r(\overline{\beta}^{[\omega]})^\omega \gamma\eta$ for all $r \in |\overline{\alpha}^{[\omega]}\rangle(q)$. This gives us the actual statement of the lemma; note that the assumption of the lemma, $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\delta}^{[1]}\eta$ (which can be reformulated as $(q_{\mathsf{init}} X_{\mathsf{init}}, q) \in \mathsf{up}_0(\overline{\delta}^{[1]})$), is equivalent to the assumption of our inductive statement, $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\delta}^{[e]}\eta$ (which can be reformulated as $(q_{\mathsf{init}} X_{\mathsf{init}}, q) \in \mathsf{up}_0(\overline{\delta}^{[e]})$), by Lemma 5.2.

If $\overline{\delta} = X \in \Gamma$, the thesis is trivial because $\overline{\delta}^{[\omega]} = \overline{\delta}^{[e]} = X$.

Suppose that $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$, and consider some $i \in [1, k]$. For every $r \in |\overline{\delta}_1^{[\omega]} \ldots \overline{\delta}_{i-1}^{[\omega]}\rangle(q)$ we have $r \in |\overline{\delta}_1^{[e]} \ldots \overline{\delta}_{i-1}^{[e]}\rangle(q)$ by Lemma 5.2, so $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\delta}^{[e]}\eta \to^* r\overline{\delta}_i^{[e]}\overline{\delta}_{i+1}^{[e]} \ldots \overline{\delta}_k^{[e]}\eta$; by the induction hypothesis we thus have $r\overline{\delta}_i^{[e]}\overline{\delta}_{i+1}^{[e]} \ldots \overline{\delta}_k^{[e]}\eta \approx r\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[e]} \ldots \overline{\delta}_k^{[e]}\eta$. By Lemma 4.5 this implies $q\overline{\delta}_1^{[\omega]} \ldots \overline{\delta}_{i-1}^{[\omega]}\overline{\delta}_i^{[e]}\overline{\delta}_{i+1}^{[e]} \ldots \overline{\delta}_k^{[e]}\eta \approx q\overline{\delta}_1^{[\omega]} \ldots \overline{\delta}_{i-1}^{[\omega]}\overline{\delta}_i^{[\omega]}\overline{\delta}_{i+1}^{[e]} \ldots \overline{\delta}_k^{[e]}\eta$. Having this for all $i \in [1, k]$, we obtain $q\overline{\delta}^{[e]}\eta \approx q\overline{\delta}^{[\omega]}\eta$, as needed.

Finally, suppose that $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$. Let $R = |\overline{\alpha}^{[\omega]}\rangle(q)$. By well-formedness of $\overline{\delta}$ we have $R = |\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^i\rangle(q)$ for all $i \ge 0$. By applying the previous case to the decomposition $\widehat{\delta} = \overline{\alpha} \underbrace{\overline{\beta} \ldots \overline{\beta}}_{e}$ we obtain that

$$q\overline{\delta}^{[e]}\eta = q\overline{\alpha}^{[e]}(\overline{\beta}^{[e]})^e \gamma\eta \approx q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^e \gamma\eta.$$

The former configuration is reachable from $q_{\mathsf{init}} X_{\mathsf{init}}$, so the class of $q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^e \gamma\eta$ is reachable from $[q_{\mathsf{init}} X_{\mathsf{init}}]_{\approx}$. Moreover, $q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^e \gamma\eta \to^* r(\overline{\beta}^{[\omega]})^i \gamma\eta$ for all $i \in [0, e]$ and $r \in R$, so the classes of $r(\overline{\beta}^{[\omega]})^i \gamma\eta$ are reachable as well; they are among the F classes of our pointed ε-LTS. For every $i \in [0, e]$ let

$$\boldsymbol{\sigma}_i = ([r(\overline{\beta}^{[\omega]})^i \gamma\eta]_{\approx})_{r \in R} \qquad \text{and} \qquad M_i = \min\{\mathrm{EqLev}(r(\overline{\beta}^{[\omega]})^i \gamma\eta, r(\overline{\beta}^{[\omega]})^\omega \gamma\eta) \mid r \in R\}.$$

There exist at most $F^{|Q|} = e$ distinct tuples of $|R|$ classes, so necessarily $\boldsymbol{\sigma}_i = \boldsymbol{\sigma}_j$ for some i, j with $0 \le i < j \le e$, and thus also $M_i = M_j$. On the other hand $M_j \ge (j - i) + M_i$ by Lemma 5.4 applied to indices $i, i+1, \ldots, j-1$.

Thus necessarily $M_i = \omega$, meaning that $r(\overline{\beta}^{[\omega]})^i\gamma\eta \approx r(\overline{\beta}^{[\omega]})^\omega\gamma\eta$ for all $r \in R$. By Lemma 4.5 we also have $r(\overline{\beta}^{[\omega]})^e\gamma\eta \approx r(\overline{\beta}^{[\omega]})^{e-i}(\overline{\beta}^{[\omega]})^\omega\gamma\eta \approx r(\overline{\beta}^{[\omega]})^\omega\gamma\eta$ for all $r \in R$, and $q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^e\gamma\eta \approx q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta$. Altogether we obtain

$$q\overline{\delta}^{[e]}\eta \approx q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^e\gamma\eta \approx q\overline{\alpha}^{[\omega]}(\overline{\beta}^{[\omega]})^\omega\gamma\eta = q\overline{\delta}^{[\omega]}\eta,$$

as needed. □

We now have Lemma 7.2, our main technical lemma. It says that the number of possible classes of $r\gamma\eta$ is small, assuming that the classes $\mathcal{C}_s$ of $s\eta$ are fixed, and that γ belongs to a pumping triple $\langle\!\langle\cdot,\cdot,\gamma\rangle\!\rangle$ of a fixed type. This is very powerful, because infinitely many stack contents γ may be handled this way (intuitively, γ is almost arbitrary). The assumption that γ belongs to a pumping triple of a fixed type is very mild; we easily deal with it later (using the fact that every stack content has a decomposition whose type comes from a small set).

LEMMA 7.2. *Let $q \in Q$, let $\overline{\delta}_0 = \langle\!\langle\overline{\alpha}, \overline{\beta}, \gamma_0\rangle\!\rangle$ be a well-formed decomposition of degree at most C and height at most $2|Q| + 2$, let $(\mathcal{C}_s)_{s\in|\overline{\delta}_0^{[\omega]}\rangle(q)}$ be a tuple of classes, and let $K = |Below(\{\mathcal{C}_s \mid s \in |\overline{\delta}_0^{[\omega]}\rangle(q)\})|$. Let also Ω be the set of pairs $(\gamma, \eta) \in \Gamma^* \times \Gamma^*$ for which*
- *$\langle\!\langle\overline{\alpha}, \overline{\beta}, \gamma\rangle\!\rangle \overset{\mathsf{tp}}{=} \overline{\delta}_0$,*
- *$q_{\mathsf{init}}X_{\mathsf{init}} \rightarrow^* q\overline{\alpha}^{[1]}\overline{\beta}^{[1]}\gamma\eta$, and*
- *$s\eta \in \mathcal{C}_s$ for all $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$.*

If $(\mathcal{L}(\mathcal{P}), q_{\mathsf{init}}X_{\mathsf{init}})$ is weakly bisimulation finite, then the set of tuples of classes

$$\Theta = \{([r\gamma\eta]_\approx)_{r\in|\overline{\alpha}^{[\omega]}\rangle(q)} \mid (\gamma, \eta) \in \Omega\}$$

has at most $(K + \mathsf{T})^{3|Q|}$ elements, for some $\mathsf{T} \in 2^{2^{|\mathcal{P}|^{O(1)}}}$. Moreover, for every tuple $(\mathcal{C}'_r)_{r\in|\overline{\alpha}^{[\omega]}\rangle(q)} \in \Theta$ we have $|Below(\{\mathcal{C}'_r \mid r \in |\overline{\alpha}^{[\omega]}\rangle(q)\})| \leq (K + 1) \cdot \mathsf{U}$ for some $\mathsf{U} \in 2^{2^{|\mathcal{P}|^{O(1)}}}$.

Proof. If Ω is empty, the thesis holds trivially. Assuming that Ω is nonempty, let us fix, for the duration of the whole proof, a stack content $\eta_0 \in \Gamma^*$ that occurs on the second coordinate of some pair from Ω. Then $s\eta_0 \in \mathcal{C}_s$ for all $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$, and thus the condition $s\eta \in \mathcal{C}_s$ can be reformulated as $s\eta \approx s\eta_0$. We can also write $K = |Below(\{s\eta_0 \mid s \in |\overline{\delta}_0^{[\omega]}\rangle(q)\})|$.

Denote $\widetilde{\alpha} = \overline{\alpha}^{[\omega]}$ and $\widetilde{\beta} = \overline{\beta}^{[\omega]}$. Note that $\widetilde{\beta}$ is obtained from $\overline{\beta}$ by adding ω exponents in all pumping triples inside $\overline{\beta}$, but the whole $\overline{\beta}$ remains repeated only once (in particular, if $\overline{\beta}$ is just a sequence of symbols, then $\widetilde{\beta} = \overline{\beta}$). Let also $R = |\widetilde{\alpha}\rangle(q)$. By well-formedness of $\overline{\delta}_0$ we have $R = |\widetilde{\alpha}\widetilde{\beta}^i\rangle(q)$ for all $i \geq 0$, as well as $|\widetilde{\beta}\rangle(R) = R$.

For $(\gamma, \eta) \in \Omega$ let $e_{\gamma,\eta}$ be the smallest number such that $([r\widetilde{\beta}^{e_{\gamma,\eta}}\gamma\eta]_\approx)_{r\in R} = ([r\widetilde{\beta}^\omega\gamma\eta]_\approx)_{r\in R}$. By Lemma 7.1 we have $e_{\gamma,\eta} \in \mathbb{N}$. We define also the constant $\mathsf{D}' = (\mathsf{B} + 2) \cdot \mathsf{BC}^{2|Q|+1} + 1$ and remark that $\mathsf{D}' < \mathsf{D}$, recalling that $\mathsf{D} = (\mathsf{B} + 3) \cdot \mathsf{BC}^{2|Q|+1} + 2$.

Next, for all $r \in R$, all $(\gamma, \eta) \in \Omega$, and all $i \in [0, e_{\gamma,\eta} - 1]$ we fix some $j_{r,\gamma,\eta,i} \in [i + 1, e_{\gamma,\eta}]$, $\theta_{r,\gamma,\eta,i} \in \Gamma^{\mathsf{gs}}$, and some finite information $\boldsymbol{\tau}_{r,\gamma,\eta,i}$ (whose goal is to describe the class of $r\widetilde{\beta}^i\gamma\eta$) in such a way that either

(A) $\boldsymbol{\tau}_{r,\gamma,\eta,i} = [r\widetilde{\beta}^i\gamma\eta]_\approx \in Below(q\overline{\delta}_0^{[\omega]}\eta_0)$, or

(B) $\boldsymbol{\tau}_{r,\gamma,\eta,i} = (r', t, \mu, (\tau_u)_{u\in|\mu\rangle(t)})$ with $r' \in R$, $t \in Q$, $\mu \in \Gamma^{\mathsf{D}'+1}$, where $r\widetilde{\beta}^i\gamma\eta \approx t\mu\theta_{r,\gamma,\eta,i}$, and for all $u \in |\mu\rangle(t)$ we have either (where the goal of τ_u is to describe the class of $u\theta_{r,\gamma,\eta,i}$)

(B1) $\tau_u = (v_u, \xi_u) \in Q \times \Gamma^{\mathsf{gs}}$, where $u\theta_{r,\gamma,\eta,i} \approx v_u\xi_u\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta$ and $v_u\xi_u \in \text{NEAR}(r'\widetilde{\beta}^{\mathsf{B}+2}, \mathsf{BC}^{2|Q|+1} + \mathsf{B})$, or

(B2) $\tau_u = [u\theta_{r,\gamma,\eta,i}]_\approx$ and $\text{DIST}(\mathcal{C}_{s_u}, [u\theta_{r,\gamma,\eta,i}]_\approx) \leq \mathsf{B}$ for some $s_u \in |\overline{\delta}_0^{[\omega]}\rangle(q)$.

If there are multiple choices for some r, γ, η, i, we just take any of them. In the following claim we prove that there is always at least one choice.

CLAIM 7.1. *For all $r \in R$, all $(\gamma, \eta) \in \Omega$, and all $i \in [0, e_{\gamma,\eta} - 1]$ there exist $j_{r,\gamma,\eta,i}$, $\theta_{r,\gamma,\eta,i}$, and $\boldsymbol{\tau}_{r,\gamma,\eta,i}$ satisfying the above conditions.*

Proof. Let $r \in R$, $(\gamma, \eta) \in \Omega$, and $i \in [0, e_{\gamma,\eta} - 1]$. To simplify the notation, let us write e for $e_{\gamma,\eta}$, for the chosen γ, η.

If $[r\widetilde{\beta}^i\gamma\eta]_\approx \in Below(q\overline{\delta}_0^{[\omega]}\eta_0)$, we can take $\boldsymbol{\tau}_{r,\gamma,\eta,i} = [r\widetilde{\beta}^i\gamma\eta]_\approx$, obtaining Item (A). Suppose thus that $[r\widetilde{\beta}^i\gamma\eta]_\approx \notin Below(q\overline{\delta}_0^{[\omega]}\eta_0)$, in which case we prove Item (B). Note that $q\overline{\delta}_0^{[\omega]}\eta_0 \approx q\overline{\delta}_0^{[\omega]}\eta \approx q\widetilde{\alpha}\widetilde{\beta}^\omega\gamma\eta \approx q\widetilde{\alpha}\widetilde{\beta}^e\gamma\eta$, where the first equivalence holds by Lemma 4.5 because $s\eta \approx s\eta_0$ for all $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$, the second equivalence holds by Lemma 5.7 because $\langle\!\langle\overline{\alpha}, \overline{\beta}, \gamma\rangle\!\rangle \overset{\mathsf{tp}}{=} \overline{\delta}_0$, and the third equivalence holds by Lemma 4.5 because $r'\widetilde{\beta}^\omega\gamma\eta \approx r'\widetilde{\beta}^e\gamma\eta$ for all $r' \in |\widetilde{\alpha}\rangle(q)$ (cf. the definition of e). We can thus write $[r\widetilde{\beta}^i\gamma\eta]_\approx \notin Below(q\widetilde{\alpha}\widetilde{\beta}^e\gamma\eta)$.

Recall that $r \in R = |\widetilde{\alpha}\widetilde{\beta}^{e-i}\rangle(q)$, which implies that there exists a sequence of states $r_e, r_{e-1}, \ldots, r_i \in R$ such that $r_e \in |\widetilde{\alpha}\rangle(q)$, and $r_j \in |\widetilde{\beta}\rangle(r_{j+1})$ for all $j \in [i, e-1]$, and $r_i = r$. It follows that there is a run from $q\widetilde{\alpha}$ to r_e and runs from $r_{j+1}\widetilde{\beta}$ to r_j for all $j \in [i, e-1]$, each of them reading at most $\mathsf{BC}^{2|Q|+1}$ action symbols (cf. Lemma 5.5). We compose these runs shifted appropriately, so that we obtain a run π from $q\widetilde{\alpha}\widetilde{\beta}^e\gamma\eta$ to $r\widetilde{\beta}^i\gamma\eta$ (going through $r_e\widetilde{\beta}^e\gamma\eta, r_{e-1}\widetilde{\beta}^{e-1}\gamma\eta, \ldots, r_{i+1}\widetilde{\beta}^{i+1}\gamma\eta$).

Because $[\pi(|\pi|)]_\approx \notin Below(\pi(0))$, the run π is not D-almost-popping, which by definition means that there exists a run π'_D that is parallel to a suffix of π and satisfies $\textsc{StackGrowth}(\pi'_\mathsf{D}) \geq \mathsf{D}$. Let π'_3 be the shortest suffix of π'_D satisfying $\textsc{StackGrowth}(\pi'_3) = \mathsf{D}'$; note that the stack growth of π'_D changes by at most one in each step, and $\mathsf{D}' < \mathsf{D}$, so such a suffix exists. For some $t_1, t \in Q$, $X \in \Gamma$, $\mu \in \Gamma^{\mathsf{D}'+1}$, and $\theta_{r,\gamma,\eta,i} \in \Gamma^{\mathsf{gs}}$ we can write

$$t_1 X \theta_{r,\gamma,\eta,i} \xrightarrow{\pi'_3} t\mu\theta_{r,\gamma,\eta,i}$$

(in particular, on the top of the stack of $\pi'_3(0)$ there is a standard symbol, pushed earlier by π'_D). Note that π'_3 can be shifted to a run from $t_1 X$ to $t\mu$, and that π'_3 is parallel to a suffix π_3 of π.

Because every non-popping transition reads an action symbol, π'_D reads at least $\mathsf{D} - \mathsf{D}' = \mathsf{BC}^{2|Q|+1} + 1$ action symbols before π'_3 starts. Simultaneously $r_e\widetilde{\beta}^e\gamma\eta$ is reached by π before $\mathsf{BC}^{2|Q|+1} + 1$ action symbols are read, so before π_3 starts. Let h be the smallest number in $[i+1, e]$ such that $r_h\widetilde{\beta}^h\gamma\eta$ is reached by π before π_3 starts. We take $j_{r,\gamma,\eta,i} = h - \mathsf{B} - 2$ and $r' = r_h$. Let us prove that $j_{r,\gamma,\eta,i} > i$. Clearly $j_{r,\gamma,\eta,i} + \mathsf{B} + 2 > i$, by definition. The configuration $\pi_3(0)$ can be written as $q_1\chi\widetilde{\beta}^{j_{r,\gamma,\eta,i}+\mathsf{B}+1}\gamma\eta$ for some $q_1 \in Q$ and $\chi \in \Gamma^{\mathsf{gs}}$. Let π_1 and π_2 be the fragments of π before $r_h\widetilde{\beta}^h\gamma\eta$, and between $r_h\widetilde{\beta}^h\gamma\eta$ and $\pi_3(0)$, respectively; we have

$$q\widetilde{\alpha}\widetilde{\beta}^e\gamma\eta \xrightarrow{\pi_1} r'\widetilde{\beta}^{j_{r,\gamma,\eta,i}+\mathsf{B}+2}\gamma\eta \xrightarrow{\pi_2} q_1\chi\widetilde{\beta}^{j_{r,\gamma,\eta,i}+\mathsf{B}+1}\gamma\eta \xrightarrow{\pi_3} r\widetilde{\beta}^i\gamma\eta .$$

Recalling how π was constructed, we see that π_3 reads at most $(j_{r,\gamma,\eta,i} + \mathsf{B} + 2 - i) \cdot \mathsf{BC}^{2|Q|+1}$ action symbols (it is a suffix of a run popping $j_{r,\gamma,\eta,i} + \mathsf{B} + 2 - i$ copies of $\widetilde{\beta}$); likewise π_2 reads at most $\mathsf{BC}^{2|Q|+1}$ action symbols. On the other hand, every non-popping transition reads some action symbol, so π'_3 (and π_3 as well) reads at least $|\mu| - 1 = \mathsf{D}' > (\mathsf{B} + 2) \cdot \mathsf{BC}^{2|Q|+1}$ of them. This implies that $(j_{r,\gamma,\eta,i} + \mathsf{B} + 2 - i) \cdot \mathsf{BC}^{2|Q|+1} > (\mathsf{B} + 2) \cdot \mathsf{BC}^{2|Q|+1}$, that is, $j_{r,\gamma,\eta,i} > i$. By construction we also have $r' \in R$, $j_{r,\gamma,\eta,i} \leq e$, and $r\widetilde{\beta}^i\gamma\eta \approx t\mu\theta_{r,\gamma,\eta,i}$.

It remains to define τ_u for $u \in |\mu\rangle(t)$, satisfying Items (B1) or (B2). Take some $u \in |\mu\rangle(t)$. Because $t_1 X \to^* t\mu$, we also have $t_1 X \to^* u$, so $\textsc{Dist}(t_1 X, u) \leq \mathsf{B}$ by Lemma 4.1. On the other hand $\textsc{Dist}(t_1 X, u) > 0$, because ε-transitions are deterministic and cannot push, while there is a run from $t_1 X$ to u going through $t\mu$ (a configuration with a larger stack). Since $q_1\chi\widetilde{\beta}^{j_{r,\gamma,\eta,i}+\mathsf{B}+1}\gamma\eta \approx t_1 X\theta_{r,\gamma,\eta,i}$, there is a run ϱ from $q_1\chi\widetilde{\beta}^{j_{r,\gamma,\eta,i}+\mathsf{B}+1}\gamma\eta$ to a configuration c in $[u\theta_{r,\gamma,\eta,i}]_\approx$, reading at most B action symbols. Let us choose c so that no ε-transitions are performed at the end of this run (this is possible because ε-transitions do not change the class). Because ϱ reads some action symbols, also the composition $\pi_2 \circ \varrho$ does not end with an ε-transition. We have two cases:

- Suppose first that $\pi_2 \circ \varrho$ may be shifted to a run from $r'\widetilde{\beta}^{\mathsf{B}+2}$. This means that its final configuration c may be written as $v_u\xi_u\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta$, and the considered shift of the run $\pi_2 \circ \varrho$ witnesses that $v_u\xi_u \in \textsc{Near}(r'\widetilde{\beta}^{\mathsf{B}+2}, \mathsf{BC}^{2|Q|+1} + \mathsf{B})$; taking $\tau_u = (v_u, \xi_u)$ we obtain Item (B1).
- Suppose conversely: $\pi_2 \circ \varrho$ cannot be shifted to a run from $r'\widetilde{\beta}^{\mathsf{B}+2}$. In other words, $\pi_2 \circ \varrho$ (before its end) reaches a configuration with stack $\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta$. Recall that the stack in the π_2 part always contains $\widetilde{\beta}^{j_{r,\gamma,\eta,i}+\mathsf{B}+1}\gamma\eta$ as a suffix. Thus, ϱ pops the $\mathsf{B}+1$ copies of $\widetilde{\beta}$ while reading at most B action symbols; some copy of $\widetilde{\beta}$ is popped using only ε-transitions. After popping this copy, we reach a configuration of the form $r''\widetilde{\beta}^{j_{r,\gamma,\eta,i}+k}\gamma\eta$, where $k \geq 0$ and $r'' \in |\widetilde{\beta}\rangle_\varepsilon(Q)$. By well-formedness of $\overline{\delta}_0$ this implies that $|\widetilde{\beta}\rangle_\varepsilon(r'') = \{r''\}$

and $|\gamma\rangle_\varepsilon(r'') = |\gamma_0\rangle_\varepsilon(r'') \neq \emptyset$. Recall that ε-transitions are deterministic and that ϱ does not end with an ε-transition. It follows that ϱ, after visiting $r''\widetilde{\beta}^{j_{r,\gamma,\eta,i}+k}\gamma\eta$, necessarily pops all the remaining copies of $\widetilde{\beta}$, then pops γ, and then continues from $s_u\eta$, for some $s_u \in Q$. Because ϱ reads at most B action symbols (and ends in $c \in [u\theta_{r,\gamma,\eta,i}]_\approx$), we obtain $\mathrm{DIST}([s_u\eta]_\approx, [u\theta_{r,\gamma,\eta,i}]_\approx) \leq \mathsf{B}$. A prefix of $\pi_1 \circ \pi_2 \circ \varrho$ leading to $r''\widetilde{\beta}^{j_{r,\gamma,\eta,i}+k}\gamma\eta$ can be shifted to a run from $q\widetilde{\alpha}\widetilde{\beta}^{e-j_{r,\gamma,\eta,i}-k}$ to r'', witnessing that $r'' \in |\widetilde{\alpha}\widetilde{\beta}^{e-j_{r,\gamma,\eta,i}-k}\rangle(q) = |\widetilde{\alpha}\rangle(q)$. Together with $r'' \in |\widetilde{\beta}\rangle_\varepsilon(r'')$ and $s_u \in |\gamma\rangle_\varepsilon(r'') = |\gamma_0\rangle_\varepsilon(r'')$ this gives us $s_u \in |\overline{\delta}_0^{[\omega]}\rangle(q)$ (cf. Lemma 5.1). Thus, by assumptions of the lemma, we have $s_u\eta \in \mathcal{C}_{s_u}$, which gives us $\mathrm{DIST}(\mathcal{C}_{s_u}, [u\theta_{r,\gamma,\eta,i}]_\approx) \leq \mathsf{B}$, as needed for Item (B2), where we take $\tau_u = [u\theta_{r,\gamma,\eta,i}]_\approx$. □

Observe that there is a number $\mathsf{T} \in 2^{2^{|\mathcal{P}|^{O(1)}}}$ such that

$$|\{\boldsymbol{\tau}_{r,\gamma,\eta,i} \mid r \in R, (\gamma,\eta) \in \Omega, i \in [0, e_{\gamma,\eta}-1]\}| \leq K + \mathsf{T}.$$

Indeed, we have $|Below(q\overline{\delta}_0^{[\omega]}\eta_0)|$ possible values of $\boldsymbol{\tau}_{r,\gamma,\eta,i}$ conforming with Item (A) of the definition, which is at most $K + \mathsf{C}^{2|Q|+2} \cdot |\mathcal{P}|^{\mathsf{D}+1}$ by Lemma 6.1. For Item (B) we have $|Q|^2 \cdot |\Gamma|^{\mathsf{D}'+1}$ possibilities for r', t, μ, and then for every of at most $|Q|$ states u we have

- at most $|\mathrm{NEAR}(r'\widetilde{\beta}^{\mathsf{B}+2}, \mathsf{BC}^{2|Q|+1} + \mathsf{B})|$ possibilities for $v_u\xi_u$ in Item (B1), which is at most $|\mathcal{P}|^{\mathsf{BC}^{2|Q|+1}+\mathsf{B}}$ by Lemma 4.4, and
- at most $|Q|$ possibilities for a state s_u, and then at most $|\{\mathcal{C} \mid \mathrm{DIST}(\mathcal{C}_{s_u}, \mathcal{C}) \leq \mathsf{B}\}|$ possibilities for $[u\theta_{r,\gamma,\eta,i}]_\approx$ in Item (B2), which is at most $|\mathcal{P}|^{\mathsf{B}}$ by Lemma 4.4.

CLAIM 7.2. *If $r \in R$, $(\gamma,\eta), (\gamma',\eta') \in \Omega$, $i \in [0, e_{\gamma,\eta}-1]$, $i' \in [0, e_{\gamma',\eta'}-1]$ are such that $\boldsymbol{\tau}_{r,\gamma,\eta,i} = \boldsymbol{\tau}_{r,\gamma',\eta',i'}$, then*

$$\mathrm{EQLEV}(r\widetilde{\beta}^i\gamma\eta, r\widetilde{\beta}^{i'}\gamma'\eta') \geq \min\{\mathrm{EQLEV}(r''\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta, r''\widetilde{\beta}^{j_{r,\gamma',\eta',i'}}\gamma'\eta') \mid r'' \in R\}.$$

Proof. If $\boldsymbol{\tau}_{r,\gamma,\eta,i}$ (which equals $\boldsymbol{\tau}_{r,\gamma',\eta',i'}$) is defined according to Item (A), we have $[r\widetilde{\beta}^i\gamma\eta]_\approx = \boldsymbol{\tau}_{r,\gamma,\eta,i} = [r\widetilde{\beta}^{i'}\gamma'\eta']_\approx$, that is, $\mathrm{EQLEV}(r\widetilde{\beta}^i\gamma\eta, r\widetilde{\beta}^{i'}\gamma'\eta') = \omega$. Suppose that $\boldsymbol{\tau}_{r,\gamma,\eta,i}$ is defined according to Item (B): we have $\boldsymbol{\tau}_{r,\gamma,\eta,i} = (r', t, \mu, (\tau_u)_{u\in|\mu\rangle(t)})$, where $r\widetilde{\beta}^i\gamma\eta \approx t\mu\theta_{r,\gamma,\eta,i}$ and $r\widetilde{\beta}^{i'}\gamma'\eta' \approx t\mu\theta_{r,\gamma',\eta',i'}$. By Lemma 4.5 we thus have

$$\mathrm{EQLEV}(r\widetilde{\beta}^i\gamma\eta, r\widetilde{\beta}^{i'}\gamma'\eta') \geq \min\{\mathrm{EQLEV}(u\theta_{r,\gamma,\eta,i}, u\theta_{r,\gamma',\eta',i'}) \mid u \in |\mu\rangle(t)\}.$$

For each $u \in |\mu\rangle(t)$ we then consider the component τ_u:

- If it is defined according to Item (B2), we simply have $[u\theta_{r,\gamma,\eta,i}]_\approx = \tau_u = [u\theta_{r,\gamma',\eta',i'}]_\approx$, that is, $\mathrm{EQLEV}(u\theta_{r,\gamma,\eta,i}, u\theta_{r,\gamma',\eta',i'}) = \omega$.
- Suppose that τ_u is defined according to Item (B1): we have $\tau_u = (v_u, \xi_u)$, where $u\theta_{r,\gamma,\eta,i} \approx v_u\xi_u\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta$ and $u\theta_{r,\gamma',\eta',i'} \approx v_u\xi_u\widetilde{\beta}^{j_{r,\gamma',\eta',i'}}\gamma'\eta'$; moreover $v_u\xi_u \in \mathrm{NEAR}(r'\widetilde{\beta}^{\mathsf{B}+2}, \mathsf{BC}^{2|Q|+1} + \mathsf{B})$, meaning that $r'\widetilde{\beta}^{\mathsf{B}+2} \rightarrow^* v_u\xi_u$, which implies that $|\xi_u\rangle(v_u) \subseteq |\widetilde{\beta}^{\mathsf{B}+2}\rangle(r') \subseteq |\widetilde{\beta}^{\mathsf{B}+2}\rangle(R) = R$. By Lemma 4.5 we have

$$\begin{aligned}
\mathrm{EQLEV}(u\theta_{r,\gamma,\eta,i}, u\theta_{r,\gamma',\eta',i'}) &= \mathrm{EQLEV}(v_u\xi_u\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta, v_u\xi_u\widetilde{\beta}^{j_{r,\gamma',\eta',i'}}\gamma'\eta') \\
&\geq \min\{\mathrm{EQLEV}(r''\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta, r''\widetilde{\beta}^{j_{r,\gamma',\eta',i'}}\gamma'\eta') \mid r'' \in |\xi_u\rangle(v_u)\} \\
&\geq \min\{\mathrm{EQLEV}(r''\widetilde{\beta}^{j_{r,\gamma,\eta,i}}\gamma\eta, r''\widetilde{\beta}^{j_{r,\gamma',\eta',i'}}\gamma'\eta') \mid r'' \in R\}.
\end{aligned}$$

By combining the inequalities obtained above we get the thesis of the claim. □

CLAIM 7.3. *For all $(\gamma,\eta) \in \Omega$ we have $e_{\gamma,\eta} \leq |Q| \cdot (K + \mathsf{T})$.*

Proof. Fix some $(\gamma,\eta) \in \Omega$. For every $i \in \mathbb{N}$ let $M_i = \min\{\mathrm{EQLEV}(r\widetilde{\beta}^i\gamma\eta, r\widetilde{\beta}^\omega\gamma\eta) \mid r \in R\}$, and let $r_i \in R$ be a state such that $\mathrm{EQLEV}(r_i\widetilde{\beta}^i\gamma\eta, r_i\widetilde{\beta}^\omega\gamma\eta) = M_i$. To shorten the notation denote $e = e_{\gamma,\eta}$, $\boldsymbol{\tau}_i = \boldsymbol{\tau}_{r_i,\gamma,\eta,i}$, and $j_i = j_{r_i,\gamma,\eta,i}$ for $i \in [0, e-1]$.

By Lemma 5.4 we have $M_{i+1} \geq 1 + M_i$ for all $i \in \mathbb{N}$. Note that $M_i = \omega$ implies $([r\widetilde{\beta}^i\gamma\eta]_\approx)_{r\in R} = ([r\widetilde{\beta}^\omega\gamma\eta]_\approx)_{r\in R}$, which is impossible for $i < e$. Thus for all $i < e$ we have $M_{i+1} > M_i$ with $M_i \in \mathbb{N}$.

It remains to prove that $(r_i, \boldsymbol{\tau}_i) \neq (r_{i'}, \boldsymbol{\tau}_{i'})$ whenever $0 \leq i < i' < e$: because there are at most $|Q|$ possibilities for r_i and at most $K + \mathsf{T}$ possibilities for $\boldsymbol{\tau}_i$, this implies that $e \leq |Q| \cdot (K + \mathsf{T})$. Suppose to the contrary that $(r_i, \boldsymbol{\tau}_i) = (r_{i'}, \boldsymbol{\tau}_{i'})$ for some i, i' with $0 \leq i < i' < e$. Recall that $M_0 < M_1 < \cdots < M_e$, that is, $M_i < M_{i'}$. Because $\mathrm{E}_Q\mathrm{LEV}(r_i \widetilde{\beta}^i \gamma \eta, r_i \widetilde{\beta}^\omega \gamma \eta) = M_i$ and because $\mathrm{E}_Q\mathrm{LEV}(r_i \widetilde{\beta}^\omega \gamma \eta, r_{i'} \widetilde{\beta}^{i'} \gamma \eta) = M_{i'} > M_i$ (recall that $r_i = r_{i'}$), we have $\mathrm{E}_Q\mathrm{LEV}(r_i \widetilde{\beta}^i \gamma \eta, r_{i'} \widetilde{\beta}^{i'} \gamma \eta) = M_i$. Using Claim 7.2 we obtain

$$M_i = \mathrm{E}_Q\mathrm{LEV}(r_i \widetilde{\beta}^i \gamma \eta, r_{i'} \widetilde{\beta}^{i'} \gamma \eta) \geq \min\{\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^{j_i} \gamma \eta, r'' \widetilde{\beta}^{j_{i'}} \gamma \eta) \mid r'' \in R\}.$$

For all $r'' \in R$, because $\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^{j_i} \gamma \eta, r'' \widetilde{\beta}^\omega \gamma \eta) \geq M_{j_i} > M_i$ and $\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^\omega \gamma \eta, r'' \widetilde{\beta}^{j_{i'}} \gamma \eta) \geq M_{j_{i'}} > M_{i'} > M_i$, we also have $\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^{j_i} \gamma \eta, r'' \widetilde{\beta}^{j_{i'}} \gamma \eta) > M_i$. We thus obtain $M_i > M_i$, a contradiction. $\quad\square$

To every $(\gamma, \eta) \in \Omega$ we now assign the tuple $((\boldsymbol{\tau}_{r,\gamma,\eta,0}, j_{r,\gamma,\eta,0}))_{r \in R}$. The number of possibilities for $\boldsymbol{\tau}_{r,\gamma,\eta,0}$ is at most $K + \mathsf{T}$, and the number of possibilities for $j_{r,\gamma,\eta,0} \in [1, e_{\gamma,\eta}]$ is at most $|Q| \cdot (K + \mathsf{T})$. We can safely assume that $|Q| \leq \mathsf{T} \leq K + \mathsf{T}$, so the number of possible tuples is at most $(K + \mathsf{T})^{3|Q|}$.

To prove the first part of the lemma, it remains to see that if the same tuple is assigned to two pairs $(\gamma, \eta), (\gamma', \eta') \in \Omega$ then $r\gamma\eta \approx r\gamma'\eta'$ for all $r \in R$. Consider thus two such pairs $(\gamma, \eta), (\gamma', \eta') \in \Omega$. Let $M = \min\{\mathrm{E}_Q\mathrm{LEV}(r\gamma\eta, r\gamma'\eta') \mid r \in R\}$, and let $r_{\min} \in R$ be a state such that $\mathrm{E}_Q\mathrm{LEV}(r_{\min}\gamma\eta, r_{\min}\gamma'\eta') = M$. By Claim 7.2 (used with $i = i' = 0$) we have

$$M = \mathrm{E}_Q\mathrm{LEV}(r_{\min}\gamma\eta, r_{\min}\gamma'\eta') \geq \min\{\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^j \gamma \eta, r'' \widetilde{\beta}^j \gamma' \eta') \mid r'' \in R\},$$

where $j = j_{r,\gamma,\eta,0} = j_{r,\gamma',\eta',0}$. Next, for every $r'' \in R$ we prove that $\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^j \gamma \eta, r'' \widetilde{\beta}^j \gamma' \eta') \geq 1 + M$, considering two cases:

- Suppose first that $|\widetilde{\beta}^j\rangle_\varepsilon(r'') = \emptyset$. Then, by Lemma 4.6 we have

$$\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^j \gamma \eta, r'' \widetilde{\beta}^j \gamma' \eta') \geq 1 + \min\{\mathrm{E}_Q\mathrm{LEV}(r\gamma\eta, r\gamma'\eta') \mid r \in |\widetilde{\beta}^j\rangle(r'')\}.$$

  We can conclude recalling that $|\widetilde{\beta}^j\rangle(r'') \subseteq |\widetilde{\beta}^j\rangle(R) = R$ and that $\mathrm{E}_Q\mathrm{LEV}(r\gamma\eta, r\gamma'\eta') \geq M$ for all $r \in R$.
- Conversely, suppose that there is some $r \in |\widetilde{\beta}^j\rangle_\varepsilon(r'') = |\widetilde{\beta}\rangle_\varepsilon(|\widetilde{\beta}^{j-1}\rangle_\varepsilon(r'')) \subseteq |\widetilde{\beta}\rangle_\varepsilon(Q)$ (recall that $j \geq 1$). By well-formedness of $\overline{\delta}_0$ we have $r \in |\widetilde{\beta}\rangle_\varepsilon(r)$ and $s \in |\gamma_0\rangle_\varepsilon(r) = |\gamma\rangle_\varepsilon(r) = |\gamma'\rangle_\varepsilon(r)$ for some s. This implies that $r'' \widetilde{\beta}^j \gamma \eta \rightarrow^*_\varepsilon r\gamma\eta \rightarrow^*_\varepsilon s\eta$ and $r'' \widetilde{\beta}^j \gamma' \eta' \rightarrow^*_\varepsilon r\gamma'\eta' \rightarrow^*_\varepsilon s\eta'$. Recalling that $r \in R = |\widetilde{\alpha}\rangle(q)$ we also have $s \in |\widetilde{\beta}\gamma\rangle_\varepsilon(|\widetilde{\alpha}\rangle(q)) = |\overline{\delta}_0^{[\omega]}\rangle(q)$ (cf. Lemma 5.1). By assumptions of the lemma we then have $s\eta \approx s\eta'$, implying that $\mathrm{E}_Q\mathrm{LEV}(r'' \widetilde{\beta}^j \gamma \eta, r'' \widetilde{\beta}^j \gamma' \eta') = \omega \geq 1 + M$.

It follows that $M \geq 1 + M$, which is only possible for $M = \omega$; thus indeed $r\gamma\eta \approx r\gamma'\eta'$ for all $r \in R$. This finishes the proof of the first part of the thesis.

Recall that every tuple in Θ is of the form $([r\gamma\eta]_\approx)_{r \in R}$ for some $(\gamma, \eta) \in \Omega$. Thus, in order to obtain the second part of the thesis, we should bound the size of the set $Below(\{r\gamma\eta \mid r \in R\})$ for every pair $(\gamma, \eta) \in \Omega$. Fix some $(\gamma, \eta) \in \Omega$. We classify classes $\mathcal{C} \in Below(\{r\gamma\eta \mid r \in R\})$ as follows:

1. Suppose that $\mathcal{C} \in Below(r\gamma\eta)$ for a state $r \in R$ such that $[r\gamma\eta]_\approx \in Below(q\overline{\delta}_0^{[\omega]}\eta_0)$. Then $\mathcal{C} \in Below(Below(q\overline{\delta}_0^{[\omega]}\eta_0))$, and the size of this set is at most $(K + \mathsf{C}^{2|Q|+2} \cdot |\mathcal{P}|^{\mathsf{D}+1}) \cdot |\mathcal{P}|^{2\mathsf{BD}}$ by Lemmata 6.1 and 6.4.
2. Suppose that $\mathcal{C} \in Below(r\gamma\eta)$ for a state $r \in R \cap |\widetilde{\beta}\rangle_\varepsilon(Q)$. By well-formedness of $\overline{\delta}_0$ we have $r \in |\widetilde{\beta}\rangle_\varepsilon(r)$ and $s \in |\gamma_0\rangle_\varepsilon(r) = |\gamma\rangle_\varepsilon(r)$ for some s, which implies $r\gamma\eta \rightarrow^*_\varepsilon s\eta$ and $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$. Then $r\gamma\eta \approx s\eta \approx s\eta_0$, so $\mathcal{C}$ belongs to $Below(\{s\eta_0 \mid s \in |\overline{\delta}_0^{[\omega]}\rangle(q)\})$, being a set of size K.
3. Suppose that $\mathcal{C} \in Below(\mathcal{C}')$ for a class $\mathcal{C}'$ such that $\mathrm{DIST}(\mathcal{C}_s, \mathcal{C}') \leq \mathsf{B}$ for some $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$. We then have a run from $s\eta_0 \in \mathcal{C}_s$ to a configuration in $\mathcal{C}'$, reading at most B action symbols. While reading only B action symbols one cannot push $\mathsf{D} > \mathsf{B}$ stack symbols (only transitions reading a symbol can push), so this run is necessarily D-almost-popping; we have $\mathcal{C}' \in Below(\{\mathcal{C}_s \mid s \in |\overline{\delta}_0^{[\omega]}\rangle(q)\})$. This means that $\mathcal{C} \in Below(Below(\{\mathcal{C}_s \mid s \in |\overline{\delta}_0^{[\omega]}\rangle(q)\}))$; by Lemma 6.4 there are at most $K \cdot |\mathcal{P}|^{2\mathsf{BD}}$ such classes $\mathcal{C}$.
4. Let us assume that none of the three above cases occurs. Among D-almost-popping runs that start from a configuration in $\{r\gamma\eta \mid r \in R\}$ and lead to a configuration in $\mathcal{C}$, we choose a run $\pi_\mathcal{C}$ reading the minimal

number of action symbols. Let us fix the state r for which $\pi_{\mathcal{C}}(0) = r\gamma\eta$; we have at most $|Q|$ possibilities for that. Let us abbreviate $j = j_{r,\gamma,\eta,0}$ and $\theta = \theta_{r,\gamma,\eta,0}$. Because the first case above does not apply, necessarily $\boldsymbol{\tau}_{r,\gamma,\eta,0}$ is of the form $(r', t, \mu, (\tau_u)_{u \in |\mu\rangle(t)})$. We have $r\gamma\eta \approx t\mu\theta$, so there is also a D-almost-popping run π'_{C} that starts in $t\mu\theta$, leads a configuration in $\mathcal{C}$, and reads the same number of action symbols as $\pi_{\mathcal{C}}$. The stack $\mu \in \Gamma^{\mathsf{D}'+1}$ can be seen as a sequence of $\mathsf{D}' + 1$ well-formed decompositions of height 0 (consisting of single stack symbols). Thus, by Lemma 6.2, there are at most $(\mathsf{D}' + 1) \cdot |\mathcal{P}|^{\mathsf{D}+1}$ classes for which the run $\pi'_{\mathcal{C}}$ can be shifted to a run from $t\mu$. For the remaining classes $\mathcal{C}$ the run $\pi'_{\mathcal{C}}$ crosses a configuration with stack θ, that is, a configuration $u\theta$ for some $u \in |\mu\rangle(t)$. Let us now also fix this state u; we have at most $|Q|$ possibilities for that.

Of course every suffix of a D-almost-popping run is again D-almost-popping; in particular this is the case for the suffix of $\pi'_{\mathcal{C}}$ starting in $u\theta$, so $\mathcal{C} \in Below(u\theta)$. Because the previous case does not apply, we cannot have $\mathrm{DIST}(\mathcal{C}_s, [u\theta]_\approx) \leq \mathsf{B}$ for any $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$. Thus τ_u is necessarily of the form (v_u, ξ_u), where $u\theta \approx v_u\xi_u\widetilde{\beta}^j\gamma\eta$. Let $\pi''_{\mathcal{C}}$ be a D-almost-popping run from $v_u\xi_u\widetilde{\beta}^j\gamma\eta$ to a configuration in $\mathcal{C}$, parallel to $\pi'_{\mathcal{C}}$. Because $v_u\xi_u \in \mathrm{NEAR}(r'\overline{\beta}^{\mathsf{B}+2}, \mathsf{BC}^{2|Q|+1} + \mathsf{B})$, there is a run ϱ from $r'\overline{\beta}^{j+\mathsf{B}+2}\gamma\eta$ to $v_u\xi_u\widetilde{\beta}^j\gamma\eta$ reading at most $\mathsf{BC}^{2|Q|+1} + \mathsf{B} < \mathsf{D}$ action symbols. A run reading so few action symbols cannot push D stack symbols, so ϱ is D-almost-popping. By Lemma 6.3 there are at most $|\mathcal{P}|^{2\mathsf{BD}} - 1$ classes $\mathcal{C}$ for which the composition $\varrho \circ \pi''_{\mathcal{C}}$ is not D-almost-popping. For remaining classes $\mathcal{C}$ the run $\varrho \circ \pi''_{\mathcal{C}}$ is D-almost-popping.

Note that $j + \mathsf{B} + 2 \leq e_{\gamma,\eta} + \mathsf{B} + 2 \leq |Q| \cdot (K + \mathsf{T}) + \mathsf{B} + 2$. By Lemma 6.2 (applied to the sequence of $j + \mathsf{B} + 2$ copies of $\overline{\beta}$) there are at most $(|Q| \cdot (K + \mathsf{T}) + \mathsf{B} + 2) \cdot \mathsf{C}^{2|Q|+1} \cdot |\mathcal{P}|^{\mathsf{D}+1}$ classes $\mathcal{C}$ for which the run $\varrho \circ \pi''_{\mathcal{C}}$ can be shifted to a run from $r'\overline{\beta}^{j+\mathsf{B}+2}$. For remaining classes $\mathcal{C}$ the run $\pi''_{\mathcal{C}}$ crosses a configuration with stack $\gamma\eta$, that is, a configuration $r''\gamma\eta$ for some $r'' \in |\widetilde{\beta}^{j+\mathsf{B}+2}\rangle(r') \subseteq R$. The suffix of $\pi''_{\mathcal{C}}$ starting in $r''\gamma\eta$ is again D-almost-popping, that is, $\mathcal{C} \in Below(r''\gamma\eta)$. Because the second case above does not apply, we have $r'' \notin |\widetilde{\beta}\rangle_\varepsilon(Q)$, so the part of $\pi''_{\mathcal{C}}$ that pops the last copy of $\widetilde{\beta}$ necessarily reads some action symbol; the suffix of $\pi''_{\mathcal{C}}$ starting in $r''\gamma\eta$ reads less action symbols than $\pi_{\mathcal{C}}$. However this contradicts the minimality of $\pi_{\mathcal{C}}$, so there are no such classes $\mathcal{C}$.

By summing up the numbers obtained above we get a number $\mathsf{U} \in 2^{2^{|\mathcal{P}|^{O(1)}}}$ such that we have $|Below(\{r\gamma\eta \mid r \in R\})| \leq (K + 1) \cdot \mathsf{U}$. $\quad\square$

Let us reformulate Lemma 7.2 so that it uses a set of states P in place of a single state q:

LEMMA 7.3. *Let $P \subseteq Q$, let $\overline{\delta}_0 = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma_0 \rangle\!\rangle$ be a well-formed decomposition of degree at most C and height at most $2|Q| + 2$, let $(\mathcal{C}_s)_{s \in |\overline{\delta}_0^{[\omega]}\rangle(P)}$ be a tuple of classes, and let $K = |Below(\{\mathcal{C}_s \mid s \in |\overline{\delta}_0^{[\omega]}\rangle(P)\})|$. Let also Ω be the set of pairs $(\gamma, \eta) \in \Gamma^* \times \Gamma^*$ for which*

- $\langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle \overset{\mathsf{tp}}{=} \overline{\delta}_0$,
- $q_{\mathsf{init}}X_{\mathsf{init}} \to^* q\overline{\alpha}^{[1]}\overline{\beta}^{[1]}\gamma\eta$ *for all $q \in P$, and*
- $s\eta \in \mathcal{C}_s$ *for all $s \in |\overline{\delta}_0^{[\omega]}\rangle(P)$.*

If $(\mathcal{L}(\mathcal{P}), q_{\mathsf{init}}X_{\mathsf{init}})$ is weakly bisimulation finite, then the set of tuples of classes

$$\Theta = \{([r\gamma\eta]_\approx)_{r \in |\overline{\alpha}^{[\omega]}\rangle(P)} \mid (\gamma, \eta) \in \Omega\}$$

has at most $(K + \mathsf{T})^{3|Q|^2}$ elements. Moreover, for every tuple $(\mathcal{C}'_r)_{r \in |\overline{\alpha}^{[\omega]}\rangle(P)} \in \Theta$ we have $|Below(\{\mathcal{C}'_r \mid r \in |\overline{\alpha}^{[\omega]}\rangle(P)\})| \leq (K + 1) \cdot |Q| \cdot \mathsf{U}$.

Proof. For every $q \in P$ let Ω_q be the set of pairs $(\gamma, \eta) \in \Gamma^* \times \Gamma^*$ for which $\langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle \overset{\mathsf{tp}}{=} \overline{\delta}_0$, $q_{\mathsf{init}}X_{\mathsf{init}} \to^* q\overline{\alpha}^{[1]}\overline{\beta}^{[1]}\gamma\eta$, and $s\eta \in \mathcal{C}_s$ for all $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$; let also $\Theta_q = \{([r\gamma\eta]_\approx)_{r \in |\overline{\alpha}^{[\omega]}\rangle(q)} \mid (\gamma, \eta) \in \Omega_q\}$. Lemma 7.2 implies that $|\Theta_q| \leq (K + \mathsf{T})^{3|Q|}$ and that $|Below(\{\mathcal{C}'_r \mid r \in |\overline{\alpha}^{[\omega]}\rangle(q)\})| \leq (K + 1) \cdot \mathsf{U}$ for every tuple $(\mathcal{C}'_r)_{r \in |\overline{\alpha}^{[\omega]}\rangle(q)} \in \Theta_q$.

Observe that $\Omega = \bigcap_{q \in P} \Omega_q$. In consequence, for every tuple $(\mathcal{C}'_r)_{r \in |\overline{\alpha}^{[\omega]}\rangle(P)} \in \Theta$ and for every $q \in P$, the sub-tuple $(\mathcal{C}'_r)_{r \in |\overline{\alpha}^{[\omega]}\rangle(q)}$ belongs to Θ_q. Because $|\overline{\alpha}^{[\omega]}\rangle(P) = \bigcup_{q \in P} |\overline{\alpha}^{[\omega]}\rangle(q)$ (i.e., every full tuple is completely determined by the sub-tuples for all $q \in P$), we thus have $|\Theta| \leq \Pi_{q \in P}|\Theta_q| \leq ((K + \mathsf{T})^{3|Q|})^{|Q|}$. Moreover, for every tuple $(\mathcal{C}'_r)_{r \in |\overline{\alpha}^{[\omega]}\rangle(P)} \in \Theta$ we have $Below(\{\mathcal{C}'_r \mid r \in |\overline{\alpha}^{[\omega]}\rangle(P)\}) = \bigcup_{q \in P} Below(\{\mathcal{C}'_r \mid r \in |\overline{\alpha}^{[\omega]}\rangle(q)\})$, which implies $|Below(\{\mathcal{C}'_r \mid r \in |\overline{\alpha}^{[\omega]}\rangle(P)\})| \leq |Q| \cdot (K + 1) \cdot \mathsf{U}$, as required. $\quad\square$

While in Lemma 7.3 we describe the situation after appending γ to a stack η (where γ is the last component of some pumping triple), in Lemma 7.4 we describe the situation after appending the whole $\overline{\delta}^{[1]}$ for a decomposition $\overline{\delta}$. The decomposition $\overline{\delta}$ may contain multiple pumping triples (also in a nested way); Lemma 7.4 uses Lemma 7.3 for each of these triples.

LEMMA 7.4. *Let $P \subseteq Q$, let $\overline{\delta}_0$ be a well-formed decomposition of degree at most C and height $\ell \leq 2|Q| + 2$, let $(\mathcal{C}_s)_{s \in |\overline{\delta}_0^{[\omega]}\rangle(P)}$ be a tuple of classes, and let $K = |Below(\{\mathcal{C}_s \mid s \in |\overline{\delta}_0^{[\omega]}\rangle(P)\})|$. Let also Ω be the set of pairs $(\overline{\delta}, \eta)$ (of a decomposition and a standard stack content) for which*

- $\overline{\delta} \stackrel{\mathsf{tp}}{=} \overline{\delta}_0$,
- $q_{\mathsf{init}} X_{\mathsf{init}} \rightarrow^* q\overline{\delta}^{[1]}\eta$ *for all $q \in P$, and*
- $s\eta \in \mathcal{C}_s$ *for all $s \in |\overline{\delta}_0^{[\omega]}\rangle(P)$.*

If $(\mathcal{L}(\mathcal{P}), q_{\mathsf{init}} X_{\mathsf{init}})$ is weakly bisimulation finite, then the set of tuples of classes

$$\Theta = \{([q\overline{\delta}^{[1]}\eta]_{\approx})_{q \in P} \mid (\overline{\delta}, \eta) \in \Omega\}$$

has at most $((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell}}$ elements. Moreover, for every tuple $(\mathcal{C}'_q)_{q \in P} \in \Theta$ we have $|Below(\{\mathcal{C}'_q \mid q \in P\})| \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell}} - 1$.

Proof. We proceed by induction on the structure of $\overline{\delta}_0$. Suppose first that $\overline{\delta}_0 = X \in \Gamma$ (i.e., $\ell = 0$). For all $(\overline{\delta}, \eta), (\widehat{\delta}, \eta') \in \Omega$ we have $\overline{\delta}^{[1]} = X = \widehat{\delta}^{[1]}$, and $qX\eta \approx qX\eta'$ for all $q \in P$ by Lemma 4.5 (because $s\eta \approx s\eta'$ for all $s \in |X\rangle(q) \subseteq |X\rangle(P)$ by assumption). It follows that $|\Theta| \leq 1$. Moreover the only tuple in Θ (if exists) is of the form $([qX\eta]_{\approx})_{q \in P}$ for $(X, \eta) \in \Omega$. By Lemma 6.1 we have $|Below(\{qX\eta \mid q \in P\})| \leq K + |\mathcal{P}|^{\mathsf{D}+1}$. Checking the definition of U in the proof of Lemma 7.2 we see that $|\mathcal{P}|^{\mathsf{D}+1} \leq \mathsf{U}$ (already the component of U needed to handle Case 1 is greater than $|\mathcal{P}|^{\mathsf{D}+1}$), which gives us the desired inequality $|Below(\{qX\eta \mid q \in P\})| \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1) - 1$.

Next, suppose that $\overline{\delta}_0 = \overline{\delta}_{0,1} \ldots \overline{\delta}_{0,k}$. Let $P_0 = P$ and $P_i = |\overline{\delta}_{0,i}^{[\omega]}\rangle(P_{i-1})$ for $i \in [1, k]$. Note that $P_k = |\overline{\delta}_0^{[\omega]}\rangle(P)$. For each $i \in [0, k]$ and each tuple of classes $\boldsymbol{\sigma} = (\mathcal{D}_s)_{s \in P_i}$ let $K_{i,\boldsymbol{\sigma}} = |Below(\{\mathcal{D}_s \mid s \in P_i\})|$; when $i \in [1, k]$ let $\Omega_{i,\boldsymbol{\sigma}}$ be the set of pairs $(\overline{\delta}_i, \eta_i)$ for which $\overline{\delta}_i \stackrel{\mathsf{tp}}{=} \overline{\delta}_{0,i}$, $q_{\mathsf{init}} X_{\mathsf{init}} \rightarrow^* q\overline{\delta}_i^{[1]}\eta_i$ for all $q \in P_{i-1}$, and $s\eta_i \in \mathcal{D}_s$ for all $s \in P_i$; let also $\Theta_{i,\boldsymbol{\sigma}} = \{([q\overline{\delta}_i^{[1]}\eta_i]_{\approx})_{q \in P_{i-1}} \mid (\overline{\delta}_i, \eta_i) \in \Omega_{i,\boldsymbol{\sigma}}\}$. The induction hypothesis (where as P we take P_{i-1}, as $\overline{\delta}_0$ we take $\overline{\delta}_{0,i}$, and as $(\mathcal{C}_s)_{s \in P_i}$ we take $\boldsymbol{\sigma}$) says that

$$(7.1) \qquad |\Theta_{i,\boldsymbol{\sigma}}| \leq ((K_{i,\boldsymbol{\sigma}} + 1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}} \qquad \text{and}$$

$$(7.2) \qquad K_{i-1,\boldsymbol{\sigma}'} \leq (K_{i,\boldsymbol{\sigma}} + 1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} - 1 \qquad \text{for all } \boldsymbol{\sigma}' \in \Theta_{i,\boldsymbol{\sigma}}.$$

Consider now some tuple $(\mathcal{C}'_q)_{q \in P} \in \Theta$. It is of the form $([q\overline{\delta}^{[1]}\eta]_{\approx})_{q \in P}$ for some $(\overline{\delta}, \eta) \in \Omega$. Because $\overline{\delta} \stackrel{\mathsf{tp}}{=} \overline{\delta}_0$, we have $\overline{\delta} = \overline{\delta}_1 \ldots \overline{\delta}_k$, where $\overline{\delta}_i \stackrel{\mathsf{tp}}{=} \overline{\delta}_{0,i}$ for all $i \in [1, k]$. For $i \in [0, k]$ let $\eta_i = \overline{\delta}_{i+1}^{[1]} \ldots \overline{\delta}_k^{[1]}\eta$, and $\boldsymbol{\sigma}_i = ([s\eta_i]_{\approx})_{s \in P_i}$. By definition of Ω we have $q_{\mathsf{init}} X_{\mathsf{init}} \rightarrow^* q\overline{\delta}^{[1]}\eta$ for all $q \in P$. By Lemmata 5.2 and 5.7 we have $P_{i-1} = |\overline{\delta}_{0,1}^{[\omega]} \ldots \overline{\delta}_{0,i-1}^{[\omega]}\rangle(P) = |\overline{\delta}_1^{[\omega]} \ldots \overline{\delta}_{i-1}^{[\omega]}\rangle(P) \subseteq |\overline{\delta}_1^{[1]} \ldots \overline{\delta}_{i-1}^{[1]}\rangle(P)$, which implies $q_{\mathsf{init}} X_{\mathsf{init}} \rightarrow^* q\overline{\delta}_i^{[1]}\eta_i$ for all $q \in P_{i-1}$ and $i \in [1, k]$ (we can first push the whole stack content $q\overline{\delta}^{[1]}\eta$, and then pop its topmost part). It follows that $(\overline{\delta}_i, \eta_i) \in \Omega_{i,\boldsymbol{\sigma}_i}$, hence also $\boldsymbol{\sigma}_{i-1} = ([q\overline{\delta}_i^{[1]}\eta_i]_{\approx})_{q \in P_{i-1}} \in \Theta_{i,\boldsymbol{\sigma}_i}$ for all $i \in [1, k]$. Inequality (7.2) used for the tuple $\boldsymbol{\sigma}_{i-1} \in \Theta_{i,\boldsymbol{\sigma}_i}$ gives us

$$K_{i-1,\boldsymbol{\sigma}_{i-1}} \leq (K_{i,\boldsymbol{\sigma}_i} + 1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} - 1 \qquad \text{for all } i \in [1, k].$$

Notice that $\boldsymbol{\sigma}_k = ([s\eta]_{\approx})_{s \in P_k} = (\mathcal{C}_s)_{s \in P_k}$ and $\boldsymbol{\sigma}_0 = ([q\overline{\delta}^{[1]}\eta]_{\approx})_{q \in P} = (\mathcal{C}'_q)_{q \in P}$; in particular $K_{k,\boldsymbol{\sigma}_k} = K$ and $Below(\{\mathcal{C}'_q \mid q \in P\}) = K_{0,\boldsymbol{\sigma}_0}$. Thus, the above inequalities imply that

$$K_{i,\boldsymbol{\sigma}_i} \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{(k-i) \cdot \mathsf{C}^{\ell-1}} - 1 \qquad \text{for all } i \in [0, k].$$

In particular, because $k \leq \mathsf{C}$, we have

(7.3) $\qquad Below(\{\mathcal{C}'_q \mid q \in P\}) = K_{0,\boldsymbol{\sigma}_0} \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C} \cdot \mathsf{C}^{\ell-1}} - 1$ $\qquad\qquad$ and

(7.4) $\qquad\qquad\qquad K_{i,\boldsymbol{\sigma}_i} \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{(\mathsf{C}-1) \cdot \mathsf{C}^{\ell-1}} - 1$ $\qquad\qquad$ for all $i \in [1, k]$.

Inequality (7.3) already gives us the second part of the thesis of the lemma. Inequality (7.4) can be substituted in Inequality (7.1), giving us

(7.5) $\qquad\qquad\qquad |\Theta_{i,\boldsymbol{\sigma}_i}| \leq ((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C} \cdot \mathsf{C}^{\ell-1}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}}.$

From the above it follows that for every tuple $\boldsymbol{\sigma}_0 \in \Theta$ we have found a sequence of tuples $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_k$ such that $\boldsymbol{\sigma}_{i-1} \in \Theta_{i,\boldsymbol{\sigma}_i}$ for all $i \in [1, k]$, where $\boldsymbol{\sigma}_k = (\mathcal{C}_s)_{s \in P_k}$ is fixed (does not depend on $\boldsymbol{\sigma}_0$). Notice that there are at most $|\Theta_{k,\boldsymbol{\sigma}_k}|$ choices for $\boldsymbol{\sigma}_{k-1}$, then at most $|\Theta_{k-1,\boldsymbol{\sigma}_{k-1}}|$ choices for $\boldsymbol{\sigma}_{k-2}$, and so on. Thus, by Inequality (7.5), we obtain the desired inequality

$$|\Theta| \leq \left(((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^\ell} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}}\right)^k \leq ((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^\ell} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^\ell}.$$

$\qquad$ Finally, suppose that $\overline{\delta}_0 = \langle\!\langle \overline{\alpha}_0, \overline{\beta}_0, \gamma_0 \rangle\!\rangle$. Let $R = |\overline{\alpha}_0^{[\omega]}\rangle(P)$ and $S = |\overline{\delta}_0^{[\omega]}\rangle(P)$. Notice that $|\overline{\beta}_0^{[\omega]}\rangle(R) = R$, by well-formedness of $\overline{\delta}_0$. Let Ω_3 be the set of pairs (γ, η) for which $\langle\!\langle \overline{\alpha}_0, \overline{\beta}_0, \gamma \rangle\!\rangle \overset{\mathsf{tp}}{=} \overline{\delta}_0$, $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\alpha}_0^{[1]}\overline{\beta}_0^{[1]}\gamma\eta$ for all $q \in P$, and $s\eta \in \mathcal{C}_s$ for all $s \in S$; let also $\Theta_3 = \{([r\gamma\eta]_{\approx})_{r \in R} \mid (\gamma, \eta) \in \Omega_3\}$. Going further, for every tuple $\boldsymbol{\sigma} = (\mathcal{D}_r)_{r \in R}$ let $K_{\boldsymbol{\sigma}} = |Below(\{\mathcal{D}_r \mid r \in R\})|$, let $\Omega_{2,\boldsymbol{\sigma}}$ be the set of pairs $(\overline{\beta}, \eta')$ for which $\overline{\beta} \overset{\mathsf{tp}}{=} \overline{\beta}_0$, $q_{\mathsf{init}} X_{\mathsf{init}} \to^* r\overline{\beta}^{[1]}\eta'$ for all $r \in R$, and $r\eta' \in \mathcal{D}_r$ for all $r \in R$; let also $\Theta_{2,\boldsymbol{\sigma}} = \{([r\overline{\beta}^{[1]}\eta']_{\approx})_{r \in R} \mid (\overline{\beta}, \eta') \in \Omega_{2,\boldsymbol{\sigma}}\}$. Likewise, again for every tuple $\boldsymbol{\sigma} = (\mathcal{D}_r)_{r \in R}$, let $\Omega_{1,\boldsymbol{\sigma}}$ be the set of pairs $(\overline{\alpha}, \eta'')$ for which $\overline{\alpha} \overset{\mathsf{tp}}{=} \overline{\alpha}_0$, $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\alpha}^{[1]}\eta''$ for all $q \in P$, and $r\eta'' \in \mathcal{D}_r$ for all $r \in R$; let also $\Theta_{1,\boldsymbol{\sigma}} = \{([q\overline{\alpha}^{[1]}\eta'']_{\approx})_{q \in Q} \mid (\overline{\beta}, \eta'') \in \Omega_{1,\boldsymbol{\sigma}}\}$. Finally, for every tuple $\boldsymbol{\sigma} = (\mathcal{C}'_q)_{q \in P}$ let $K'_{\boldsymbol{\sigma}} = |Below(\{\mathcal{C}'_q \mid q \in P\})|$. Lemma 7.3 gives us inequalities

(7.6) $\qquad |\Theta_3| \leq (K+\mathsf{T})^{3|Q|^2} \leq ((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^\ell} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}}$ $\qquad\qquad$ and

(7.7) $\qquad K_{\boldsymbol{\sigma}} \leq (K+1) \cdot |Q| \cdot \mathsf{U} \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} - 1$ $\qquad\qquad$ for all $\boldsymbol{\sigma} \in \Theta_3$.

For every tuple $\boldsymbol{\sigma}$ (indexed by elements of R), by the induction hypothesis for $\overline{\beta}_0$ we obtain

(7.8) $\qquad |\Theta_{2,\boldsymbol{\sigma}}| \leq ((K_{\boldsymbol{\sigma}}+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}}$ $\qquad\qquad$ and

(7.9) $\qquad K_{\boldsymbol{\sigma}'} \leq (K_{\boldsymbol{\sigma}}+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} - 1$ $\qquad\qquad$ for all $\boldsymbol{\sigma}' \in \Theta_{2,\boldsymbol{\sigma}}$,

and by the induction hypothesis for $\overline{\alpha}_0$ we obtain

(7.10) $\qquad |\Theta_{1,\boldsymbol{\sigma}}| \leq ((K_{\boldsymbol{\sigma}}+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}}$ $\qquad\qquad$ and

(7.11) $\qquad K'_{\boldsymbol{\sigma}'} \leq (K_{\boldsymbol{\sigma}}+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell-1}} - 1$ $\qquad\qquad$ for all $\boldsymbol{\sigma}' \in \Theta_{1,\boldsymbol{\sigma}}$.

$\qquad$ Consider now some tuple $\boldsymbol{\sigma}_0 = (\mathcal{C}'_q)_{q \in P} \in \Theta$. It is of the form $([q\overline{\delta}^{[1]}\eta]_{\approx})_{q \in P}$ for some $(\overline{\delta}, \eta) \in \Omega$. Because $\overline{\delta} \overset{\mathsf{tp}}{=} \overline{\delta}_0$, we have $\overline{\delta} = \langle\!\langle \overline{\alpha}, \overline{\beta}, \gamma \rangle\!\rangle$, where $\overline{\alpha} \overset{\mathsf{tp}}{=} \overline{\alpha}_0$, $\overline{\beta} \overset{\mathsf{tp}}{=} \overline{\beta}_0$, $|\gamma\rangle_\varepsilon = |\gamma_0\rangle_\varepsilon$, and $\mathsf{up}(\gamma) = \mathsf{up}(\gamma_0)$. Let $\eta' = \gamma\eta$ and $\eta'' = \overline{\beta}^{[1]}\gamma\eta$. Let also $\boldsymbol{\sigma}_1 = ([r\eta'']_{\approx})_{r \in R}$ and $\boldsymbol{\sigma}_2 = ([r\eta']_{\approx})_{r \in R}$. By definition of Ω we have $q_{\mathsf{init}} X_{\mathsf{init}} \to^* q\overline{\delta}^{[1]}\eta = q\overline{\alpha}^{[1]}\eta''$ for all $q \in P$, so $(\overline{\alpha}, \eta'') \in \Omega_{1,\boldsymbol{\sigma}_1}$. We also have $\langle\!\langle \overline{\alpha}_0, \overline{\beta}_0, \gamma \rangle\!\rangle \overset{\mathsf{tp}}{=} \overline{\delta}_0$, so $(\gamma, \eta) \in \Omega_3$. By Lemmata 5.2 and 5.7 we have $R = |\overline{\alpha}_0^{[\omega]}\rangle(P) = |\overline{\alpha}^{[\omega]}\rangle(P) \subseteq |\overline{\alpha}^{[1]}\rangle(P)$, which implies $q_{\mathsf{init}} X_{\mathsf{init}} \to^* r\overline{\beta}^{[1]}\eta'$ for all $r \in R$ (we can first push the whole stack content $q\overline{\delta}^{[1]}\eta$, and then pop its topmost part), and thus $(\overline{\beta}, \eta') \in \Omega_{2,\boldsymbol{\sigma}_2}$. In consequence $\boldsymbol{\sigma}_0 = ([q\overline{\alpha}^{[1]}\eta'']_{\approx})_{q \in P} \in \Theta_{1,\boldsymbol{\sigma}_1}$, $\boldsymbol{\sigma}_1 = ([r\overline{\beta}^{[1]}\eta']_{\approx})_{r \in R} \in \Theta_{2,\boldsymbol{\sigma}_2}$, and $\boldsymbol{\sigma}_2 = ([r\gamma\eta]_{\approx})_{r \in R} \in \Theta_3$. Inequalities (7.7), (7.9) and (7.11) give us

$$K'_{\boldsymbol{\sigma}_0} \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{3\mathsf{C}^{\ell-1}} - 1 \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^\ell} - 1,$$

which is the second part of the thesis of the lemma. For $i \in \{1,2\}$ we rather need

$$K_{\boldsymbol{\sigma}_i} \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{2\mathsf{C}^{\ell-1}} - 1 \leq (K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{(\mathsf{C}-1)\mathsf{C}^{\ell-1}} - 1 \,,$$

Substituting this to Inequalities (7.8) and (7.10) we obtain that

$$(7.12) \qquad \max(|\Theta_{1,\boldsymbol{\sigma}_i}|, |\Theta_{2,\boldsymbol{\sigma}_2}|) \leq ((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}\cdot\mathsf{C}^{\ell-1}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}} \,.$$

From the above it follows that for every tuple $\boldsymbol{\sigma}_0 \in \Theta$ we have found tuples $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$ such that $\boldsymbol{\sigma}_0 \in \Theta_{1,\boldsymbol{\sigma}_1}$, $\boldsymbol{\sigma}_1 \in \Theta_{2,\boldsymbol{\sigma}_2}$, and $\boldsymbol{\sigma}_2 \in \Theta_3$. There are at most $|\Theta_3|$ choices for $\boldsymbol{\sigma}_2$, then at most $|\Theta_{2,\boldsymbol{\sigma}_2}|$ choices for $\boldsymbol{\sigma}_1$, and then at most $|\Theta_{1,\boldsymbol{\sigma}_1}|$ choices for $\boldsymbol{\sigma}_0$. Thus, by Inequalities (7.6) and (7.12), we obtain the desired inequality

$$|\Theta| \leq \big(((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell-1}}\big)^3 \leq ((K+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{\ell}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{\ell}} \,. \qquad \square$$

THEOREM 7.1. *If $(\mathcal{L}(\mathcal{P}), q_{\mathsf{init}}X_{\mathsf{init}})$ is weakly bisimulation finite, then it has at most Z classes for some $\mathsf{Z} \in 2^{2^{|\mathcal{P}|^{O(1)}}}$.*

Proof. Let $\mathcal{C}_\perp = [s]_\approx$ for all $s \in Q$, which is the class containing configurations with no successors. Clearly $Below(\mathcal{C}_\perp) = 1$. For every type of a well-formed decomposition (i.e., equivalence class of the $\overset{\mathsf{tp}}{=}$ relation) fix some decomposition $\overline{\delta}_0$ having this type. For such a decomposition $\overline{\delta}_0$ and for a state $q \in Q$ let $\Omega_{q,\overline{\delta}_0}$ be the set of pairs $(\overline{\delta}, \eta)$ for which $\overline{\delta} \overset{\mathsf{tp}}{=} \overline{\delta}_0$, $q_{\mathsf{init}}X_{\mathsf{init}} \rightarrow^* q\overline{\delta}^{[1]}\eta$, and $s\eta \in \mathcal{C}_\perp$ for all $s \in |\overline{\delta}_0^{[\omega]}\rangle(q)$. Let also $\Theta_{q,\overline{\delta}_0} = \{[q\overline{\delta}^{[1]}\eta]_\approx \mid (\overline{\delta}, \eta) \in \Omega_{q,\overline{\delta}_0}\}$. Assuming that $\overline{\delta}_0$ has degree at most C and height at most $2|Q| + 2$, by Lemma 7.4 we have

$$|\Theta_{q,\overline{\delta}_0}| \leq \mathsf{V} = ((1+1) \cdot (|Q| \cdot \mathsf{U} + 1)^{\mathsf{C}^{2|Q|+2}} + \mathsf{T})^{3|Q|^2 \cdot \mathsf{C}^{2|Q|+2}} \,.$$

Consider now an arbitrary configuration $q\delta$ reachable from $q_{\mathsf{init}}X_{\mathsf{init}}$. Configurations with empty stack belong to $\mathcal{C}_\perp$. If δ is nonempty, by Lemma 5.3 we obtain a well-formed decomposition $\overline{\delta}$ of height at most $2|Q| + 2$ and degree at most C. Let $\overline{\delta}_0$ be the fixed representative of the type of $\overline{\delta}$. We see that $(\overline{\delta}, \varepsilon) \in \Omega_{q,\overline{\delta}_0}$, so $[q\delta]_\approx \in \Theta_{q,\overline{\delta}_0}$. By Lemma 5.6 there are at most $2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^{2|Q|+3}}$ choices for the type of $\overline{\delta}$, hence for the decomposition $\overline{\delta}_0$. Once q and $\overline{\delta}_0$ is fixed, we have only V choices for the class $[q\delta]_\approx \in \Theta_{q,\overline{\delta}_0}$. Thus, in total we have at most

$$\mathsf{Z} = 1 + |Q| \cdot 2^{|\mathcal{P}|^4 \cdot (\mathsf{C}+1)^{2|Q|+3}} \cdot \mathsf{V} \in 2^{2^{|\mathcal{P}|^{O(1)}}} \text{ possible classes.} \qquad \square$$

8 The Algorithm

In this section we prove the following theorem:

THEOREM 8.1. *Given an ε-PDS $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$, its initial configuration $q_{\mathsf{init}}X_{\mathsf{init}} \in Q\Gamma$, and a number $Z \in \mathbb{N}$, one can decide in time $O(Z^{|\mathcal{P}|^{O(1)}})$ whether the number of classes of $\mathcal{P}$ reachable from $[q_{\mathsf{init}}X_{\mathsf{init}}]_\approx$ is at most Z.*

Having in mind results from the previous section (in particular, Theorem 7.1), it follows that the weak bisimulation finiteness problem for ε-PDS can be solved in 2-ExpTime.

We remark that the algorithm provided by Theorem 8.1 not only answers Yes or No, but in the case of a positive answer, it actually computes a description of the weak bisimulation quotient of $\mathcal{P}$, which is an ε-LTS that is weakly bisimilar to $\mathcal{P}$ and has at most Z configurations.

Let us fix the input to our problem: an ε-PDS $\mathcal{P}$, an initial configuration $q_{\mathsf{init}}X_{\mathsf{init}}$, and a bound $Z \in \mathbb{N}$. The algorithm will compute relations $\approx_k$ over the set of reachable configurations, for consecutive $k = 0, 1, 2, \ldots$. By definition we have $(\approx) \subseteq (\approx_{k+1}) \subseteq (\approx_k)$ for all $k \in \mathbb{N}$, and it is easy to see that if $(\approx_{k+1}) = (\approx_k)$ for some k, then necessarily $(\approx) = (\approx_k)$. It follows that either every $\approx_{k+1}$ has more classes than $\approx_k$ until the number of classes starts exceeding Z, or at some moment, necessarily for $k < Z$, the number of classes stops growing at quantity at most Z with $(\approx) = (\approx_{k+1}) = (\approx_k)$. It is thus enough to examine the first Z relations $\approx_k$, checking whether any of them has more than Z classes.

In order to avoid special treatment of configurations with empty stack, we assume that such configurations can never be reached. This is without loss of generality: we can add a new initial configuration $q'_{\mathsf{init}}X'_{\mathsf{init}}$ together

with a transition $q'_{\mathsf{init}} X'_{\mathsf{init}} \xrightarrow{\mathsf{init}} q_{\mathsf{init}} X_{\mathsf{init}} X'_{\mathsf{init}}$ (for a fresh action symbol init). After performing this transition, the system behaves as previously, but has additionally the X'_{init} symbol on the bottom on the stack; such a change adds exactly one new weak bisimulation class, containing the new initial configuration.

Let us now see how we can represent the relations $\approx_k$ in a finite, succinct way. First, note that $\approx_0$ always has exactly one class, so we need no representation for it. Consider now some $k \geq 1$. Note that a class of $\approx_k$ is uniquely described by a tuple of classes of $\approx_{k-1}$ reachable after reading a single action symbol. More formally, for a configuration c we define

$$\mathrm{DESC}_k(c) = \{(a, [d]_{k-1}) \mid c \xRightarrow{a} d, a \in \mathbb{A}\};$$

then we have $\mathrm{DESC}_k(c) = \mathrm{DESC}_k(c') \Leftrightarrow c \approx_k c'$. We can thus take $\mathrm{DESC}_k(c)$ as a finite representation of the class $[c]_k$. By Lemma 4.4 $\mathrm{DESC}_k(c)$, as a set, has size at most $|\mathcal{P}|$, and while storing it in memory, we can remember every class $[d]_{k-1}$ as a number of this class on the list of all classes of $\approx_{k-1}$. The memory size needed for storing $\mathrm{DESC}_k(c)$ is thus negligible compared to the desired complexity of our algorithm.

In our algorithm, beside of a list of classes, we also compute a partial function CONS_k saying how a class changes when a stack grows. It is a function that assigns a class of $\approx_k$ to some tuples $(q, X, (\mathcal{C}_r)_{r \in |X\rangle(q)})$ with $q \in Q$, $X \in \Gamma$, and $\mathcal{C}_r$ being classes of $\approx_k$, defined by taking $\mathrm{CONS}_k(q, X, ([r\alpha]_k)_{r \in |X\rangle(q)}) = [qX\alpha]_k$ for all reachable configurations $qX\alpha$. Note that $([r\alpha]_k)_{r \in |X\rangle(q)} = ([r\alpha']_k)_{r \in |X\rangle(q)}$ implies $qX\alpha \approx_k qX\alpha'$ by Lemma 4.5, meaning that $[qX\alpha]_k$ depends only on the classes $[r\alpha]_k$ (without necessarily knowing the stack content α). While storing CONS_k, we represent both the result and the classes $\mathcal{C}_r$ as their numbers on the list of all reachable classes.

Moreover, for every reachable class $\mathcal{C}$ of $\approx_k$ let $\mathrm{UP}_k(\mathcal{C})$ be the class of $\approx_{k-1}$ containing $\mathcal{C}$ (recall that $c \approx_k c'$ implies $c \approx_{k-1} c'$).

We have already said that our algorithm computes the relations $\approx_k$ for consecutive $k = 0, 1, 2, \ldots$. Now we can be more precise: for every k, it computes a list of reachable classes of $\approx_k$, together with the functions CONS_k and UP_k (except for UP_0, which makes no sense). This can be done easily for $k = 0$, because $\approx_0$ is a trivial relation consisting of a single class.

When we start the computation for some $k \geq 1$, we assume that the objects mentioned above are already computed for $k - 1$. Note first that whenever we have some new reachable class $\mathcal{C}$ of $\approx_k$, which means that we know $\mathrm{DESC}_k(c)$ for $c \in \mathcal{C}$, then we can easily compute $\mathrm{UP}_k(\mathcal{C})$. Indeed, if $k = 1$, then $\mathrm{UP}_k(\mathcal{C})$ is always the only class of $\approx_0$, and if $k \geq 2$, recall that $\mathrm{UP}_k(\mathcal{C})$ is a class described by $\mathrm{DESC}_{k-1}(c)$ for configurations $c \in \mathcal{C}$, and we have

$$\mathrm{DESC}_{k-1}(c) = \{(a, \mathrm{UP}_{k-1}(\mathcal{D})) \mid (a, \mathcal{D}) \in \mathrm{DESC}_k(c)\}.$$

The function CONS_k is computed by considering configurations with stack height n, for consecutive $n = 1, 2, 3, \ldots$. Formally, we define $\mathrm{CONS}_{k,n}$ to be CONS_k with domain restricted to tuples $(q, X, ([r\alpha]_k)_{r \in |X\rangle(q)})$ for reachable configurations $qX\alpha$ satisfying $|X\alpha| \leq n$. Note that the domain of $\mathrm{CONS}_{k,n}$ only becomes larger when n increases. For $n = 0$ the domain of this function is empty (we cannot have $|X\alpha| \leq 0$). For $n \geq 1$ we assume that $\mathrm{CONS}_{k,n-1}$ is already computed. We first compute the domain of $\mathrm{CONS}_{k,n}$ using the following lemma:

LEMMA 8.1. *The domain of* $\mathrm{CONS}_{k,n}$ *is the set of*
- *tuples* $(q, X, ()_{r \in \emptyset})$ *with* $q_{\mathsf{init}} X_{\mathsf{init}} \to^* qX$, *and*
- *tuples* $(q, X, (\mathrm{CONS}_{k,n-1}(r, Y, (\mathcal{C}_s)_{s \in |Y\rangle(r)}))_{r \in |X\rangle(q)})$ *such that in the domain of* $\mathrm{CONS}_{k,n-1}$ *there is a tuple* $(p', Y', (\mathcal{C}_s)_{s \in |Y'\rangle(p')})$, *and there is a transition* $p'Y' \xhookrightarrow{a} pX'Y$, *and* $pX' \to^* qX$.

Proof. By definition, the domain of $\mathrm{CONS}_{k,n}$ contains tuples $(q, X, ([r\alpha]_k)_{r \in |X\rangle(q)})$ for reachable configurations $qX\alpha$ with $|X\alpha| \leq n$. Consider such a configuration. If $|X\alpha| = 1$ (i.e., $\alpha = \varepsilon$), we have $q_{\mathsf{init}} X_{\mathsf{init}} \to^* qX$; the tuple is added by the first item of the lemma. Note that we have $|X\rangle(q) = \emptyset$ thanks to our assumption that no configurations with empty stack are reachable.

Conversely, suppose that $2 \leq |X\alpha| \leq n$, consider a run from $q_{\mathsf{init}} X_{\mathsf{init}}$ to $qX\alpha$, and consider the last configuration with stack of height $|\alpha|$ on this run. It is of the form $p'Y'\alpha'$, and $\alpha = Y\alpha'$ for some Y. Then we have a transition $p'Y' \xhookrightarrow{a} pX'Y$ leading to $pX'\alpha$. The remaining part of the run may be shifted to a run from pX', so we have $pX' \to^* qX$. The configuration $p'Y'\alpha'$ is reachable, and $|Y'\alpha'| \leq n - 1$, so the tuple $(p', Y', (\mathcal{C}_s)_{s \in |Y'\rangle(p')})$ is in the domain of $\mathrm{CONS}_{k,n-1}$. Thus our original tuple fulfils the conditions of the second item of the lemma.

The proof of the opposite inclusion (i.e., that every tuple specified in the lemma belongs to the domain of $\text{Cons}_{k,n}$) is very similar, and thus it is left to the reader. $\square$

Using standard means we can compute in time polynomial in $|\mathcal{P}|$:
- the set of pairs pX, qY of configurations with stack of height 1 such that $pX \to^* qY$, and
- the set $|X\rangle(p)$ for every configuration pX with stack of height 1.

Thus, Lemma 8.1 allows us to easily find the domain of $\text{Cons}_{k,n}$ based on the previously computed function $\text{Cons}_{k,n-1}$.

Next, we compute values of this function. Take some tuple $(q, X, (\mathcal{C}_r)_{r \in |X\rangle(q)})$ in its domain. Imagine also a reachable configuration $qX\alpha$ such that $\mathcal{C}_r = [r\alpha]_k$ for $r \in |X\rangle(q)$ and $|X\alpha| \leq n$ (the algorithm described below does not depend on α, only its correctness proof does). Notice that the height of the stack content α is at most $n - 1$, so $\text{Up}_k(\mathcal{C}_r)$ is already known. We have two cases:

- If (q, X) is in ε-mode, then for some r we have a popping transition $qX \overset{\varepsilon}{\hookrightarrow} r$ (and $|X\rangle(q) = \{r\}$); we have $qX\alpha \approx_k r\alpha$, so we should take $\mathcal{C}_r$ as the value of $\text{Cons}_{k,n}$.
- Otherwise (q, X) is not in ε-mode. We compute the description $\text{Desc}_k(qX\alpha)$ of the class $[qX\alpha]_k$ from definition, by listing all possible classes of $\approx_{k-1}$ that can be reached from $qX\alpha$ after reading a single letter. To this end, we consider all transitions starting in (q, X):
  - for transitions of the form $qX \overset{a}{\hookrightarrow} r$ we have $qX\alpha \overset{a}{\Longrightarrow} r\alpha$ and $[r\alpha]_{k-1} = \text{Up}_k(\mathcal{C}_r)$;
  - for transitions of the form $qX \overset{a}{\hookrightarrow} sY$ we have $qX\alpha \overset{a}{\Longrightarrow} sY\alpha$ and

$$[sY\alpha]_{k-1} = \text{Cons}_{k-1}(s, Y, (\text{Up}_k(\mathcal{C}_r))_{r \in |Y\rangle(s)})$$

  (where $|Y\rangle(s) \subseteq |X\rangle(q)$);
  - for transitions of the form $qX \overset{a}{\hookrightarrow} tZY$ we have $qX\alpha \overset{a}{\Longrightarrow} tZY\alpha$ and

$$[tZY\alpha]_{k-1} = \text{Cons}_{k-1}(t, Z, (\text{Cons}_{k-1}(s, Y, (\text{Up}_k(\mathcal{C}_r))_{r \in |Y\rangle(s)}))_{s \in |Z\rangle(t)})$$

  (where $|Y\rangle(s) \subseteq |X\rangle(q)$ for $s \in |Z\rangle(t)$).

This finishes the description of how to compute $\text{Cons}_{k,n}$ knowing $\text{Cons}_{k,n-1}$ and Cons_{k-1}. Of course the values of $\text{Cons}_{k,n}$ do not depend on n, only the domain of this function grows. We can also see (cf. Lemma 8.1) that if $\text{Cons}_{k,n}$ has the same domain as $\text{Cons}_{k,n-1}$, then the domain will not grow any more. For such n we have $\text{Cons}_k = \text{Cons}_{k,n}$; we can stop the computation. Recall that we also stop the computation whenever we see more than Z classes of $\approx_k$. Thus, in the domain of $\text{Cons}_{k,n}$ we may have at most $|Q| \cdot |\Gamma| \cdot Z^{|Q|}$ tuples, and this is simultaneously a bound for n. As already said, the considered values of k are bounded by Z. It follows that the running time is in $O(Z^{|\mathcal{P}|^{O(1)}})$, as declared.

9 Lower bound

Our lower bound is shown via a reduction from the acceptance problem of exponentially space-bounded alternating Turing machines. These are being introduced in Section 9.1. Section 9.2 recalls Defender's Forcing gadgets and introduces a gadget for checking if a certain prefix of the stack is in a given regular language. Encodings of numbers and configurations are subject of Section 9.3 and Section 9.4, respectively. A gadget for verifying if a certain prefix of the stack consists of two consecutive configurations of an exponentially space-bounded alternating Turing machine is introduced in Section 9.5, whereas Section 9.6 provides a gadget for pushing a successor configuration on top of a configuration that can be found as the prefix of the stack. The actual simulation of an exponentially space-bounded alternating Turing machine is given in Section 9.7. Building upon these gadgets, Section 9.8 provides the final reduction from the acceptance problem of exponentially space-bounded alternating Turing machines to the weak bisimulation finiteness problem for ε-PDS.

9.1 Alternating Turing machines. An *alternating Turing machine* is a tuple $\mathcal{M} = (Q, \Upsilon, \Sigma, T, q_0, \square)$, where
- $Q = Q_\forall \uplus Q_\exists$ is a finite set of *states* that is partitioned into *universal states* $Q_\forall$ and existential states $Q_\exists$,
- Υ is a finite *tape alphabet*,
- $\Sigma \subseteq \Upsilon$ is an *input alphabet*,
- $T \subseteq Q \times \Upsilon \times Q \times \Upsilon \times \{-1, 1\}$ is a set of *transitions*,

- $q_0 \in Q$ is an initial state, and
- $\square \in \Upsilon \setminus \Sigma$ is a *blank symbol*.

A *configuration* of $\mathcal{M}$ is a word of the form $u(p, U)v$, where $u, v \in \Upsilon^*$ and $(p, U) \in Q \times \Upsilon$. It is *universal* if $p \in Q_\forall$ and *existential* if $p \in Q_\exists$. For all $(p, U) \in Q \times \Upsilon$ we define $T_{(p,U)} = \{(p, U, q, Y, d) \in T \mid q \in Q_\mathcal{M}, Y \in \Upsilon, d \in \{-1, 1\}\}$. By $\mathsf{Conf}_\mathcal{M}$ we denote the set of configurations of $\mathcal{M}$. For each pair $\mu = (t, Z) \in T \times \Upsilon$, where $t = (p, U, q, Y, d)$ we define the relation $\vdash_\mu \subseteq \mathsf{Conf}_\mathcal{M} \times \mathsf{Conf}_\mathcal{M}$, where $w, w' \in \mathsf{Conf}_\mathcal{M}$:

$$w \vdash_\mu w' \iff \exists u, v \in \Upsilon^*. \begin{cases} w = uZ(p, U)v \text{ and } w' = u(q, Z)Yv & \text{if } d = -1, \\ w = u(p, U)Zv \text{ and } w' = uY(q, Z)v & \text{if } d = 1. \end{cases}$$

Note that if $w \vdash_\mu w$ and $w \vdash_\mu w''$, then $w' = w''$. We say that w' a *successor configuration of* w if $w \vdash_\mu w'$ for some $\mu \in T \times \Upsilon$. Since we use them for showing 2-ExpTime-hardness, we assume without loss of generality that our alternating Turing machines do not contain any cyclic computations, that is, computations of the form $w_{(0)} \vdash w_{(1)} \cdots \vdash w_{(n)}$, where $n \geq 1$ and $w_{(0)} = w_{(n)}$. This implies that we can do induction on the length of the longest computation starting in a given configuration. We define the set of *accepting configurations* as the smallest set $C \subseteq \mathsf{Conf}_\mathcal{M}$ that satisfies the following properties:
- For every existential configuration $w \in \Upsilon^*(p, U)\Upsilon^*$ for which there exists some $t \in T_{(p,U)}$, some $Z \in \Upsilon$ and some configuration $w' \in C$ with $w \vdash_{(t,Z)} w'$, we have $w \in C$.
- For every universal configuration $w \in \Upsilon^*(p, U)\Upsilon^*$ such that for all $t \in T_{(p,U)}$ there exists some $Z \in \Upsilon$ and some configuration $w' \in C$ with $w \vdash_{(t,Z)} w'$, we have $w \in C$.

We note that a configuration $w \in \Upsilon^*(p, U)\Upsilon^*$ with $T_{(p,U)} = \emptyset$ is accepting if, and only if, it is universal.

Given a function $f \colon \mathbb{N} \to \mathbb{N}$ with $f(n) \geq n$ for all $n \in \mathbb{N}$, the language of an f-space-bounded alternating Turing machine $\mathcal{M}$ is given by

$$L(\mathcal{M}) = \{x_0 \ldots x_{n-1} \in \Sigma^n \mid n \geq 1, (q_0, x_0)x_1 \ldots x_{n-1}\square^{f(n)-n} \text{ is accepting}\}.$$

THEOREM 9.1. WEAK BISIMULATION FINITENESS FOR ε-PDS *is 2-ExpTime-hard under polynomial time reductions.*

Fix any language L in 2-ExpTime. Fix a $2^{\ell(n)}$-space-bounded alternating Turing machine $\mathcal{M} = (Q_\mathcal{M}, \Upsilon, \Sigma_\mathcal{M}, T, q_0, \square)$ such that $L = L(\mathcal{M})$, where ℓ is a polynomial (it is folklore that such a machine exists for every language in 2-ExpTime). Moreover fix some input $x = x_0 \ldots x_{n-1} \in \Sigma_\mathcal{M}^n$. We construct in time polynomial in n an ε-PDS $\mathcal{P} = (Q, \Gamma, \mathbb{A}, \Delta)$ and a configuration $q_\updownarrow\# \in Q\Gamma$ of $\mathcal{P}$ such that $x \in L(\mathcal{M})$ if, and only if, $(\mathcal{L}(\mathcal{P}), q_\updownarrow\#)$ is *not* weakly bisimulation finite. This is sufficient since 2-ExpTime is closed under complement.

In the following let us identify ℓ with $\ell(n)$ and let $N = 2^\ell$. We assume without loss of generality that $\ell \geq 2$ and hence $N - 1 > n$. We set $\Gamma = \{\#, 0, 1\} \cup \Upsilon \cup (Q_\mathcal{M} \times \Upsilon) \cup \{\#_i \mid i \in [0, \ell - 1]\}$.

Note that configurations of $\mathcal{M}$ needed for accepting inputs of length n are words from the language $\bigcup_{k=0}^{N-1} \Upsilon^k(Q_\mathcal{M} \times \Upsilon)\Upsilon^{N-1-k}$. For every such configuration $w = w_0 \ldots w_{N-1}$ we write $\mathrm{Pos}(w)$ to denote the unique $k \in [0, N - 1]$ such that $w_k \in Q_\mathcal{M} \times \Upsilon$, hereby denoting the position of the read/write head of the configuration w.

The following lemma is an immediate consequence of the definition of $\vdash_\mu$:

LEMMA 9.1. *For all $\mu = (t, Z) \in T \times \Upsilon$ with $t = (p, U, q, Y, d)$ and all configurations w and w' of $\mathcal{M}$ we have $w \vdash_\mu w'$ if, and only if, $\mathrm{Pos}(w') = \mathrm{Pos}(w) + d$ and $h_\mu(w) = h'_\mu(w')$, where $h_\mu, h'_\mu \colon (\Upsilon \cup (Q_\mathcal{M} \times \Upsilon))^* \to (\Upsilon \cup (Q_\mathcal{M} \times \Upsilon))^*$ are the letter-to-letter morphisms such that for all $X \in \Upsilon \cup (Q_\mathcal{M} \times \Upsilon)$ we have*

$$h_\mu(X) = \begin{cases} Y & \text{if } X = (p, U), \\ X & \text{otherwise,} \end{cases} \qquad \text{and} \qquad h'_\mu(X) = \begin{cases} Z & \text{if } X = (q, Z), \\ X & \text{otherwise.} \end{cases} \qquad \square$$

Throughout Section 9 we constantly add fresh rules; they implicitly come with fresh control states and action symbols. Whenever we introduce such rules, we prove that certain configurations, whose control states appear in those rules, have certain properties. These properties however are not influenced by rules that are introduced later. In particular, if we describe rules starting in some pair $(q, x) \in Q \times \Gamma$, then the reader may assume that no other rules starting in this pair are introduced later in the text. Before we introduce these rules we will introduce some gadgets in the next section.

9.2 Gadgets for the lower bound construction. Inspired by the "Defender's forcing" technique of Jančar and Srba [10], given a finite set of control states $\{\textsc{Source}, \textsc{Source}'\} \cup \{\textsc{Target}_i, \textsc{Target}'_i \mid i \in I\}$ and a stack symbol X, we introduce a notation

$$\langle \textsc{Source}, \textsc{Source}' \rangle X \xleftarrow{\textsc{Def}} \{\langle \textsc{Target}_i, \textsc{Target}'_i \rangle X \mid i \in I\}$$

to denote the following set of rules, where all action symbols and control states that are not in $\{\textsc{Source}, \textsc{Source}'\}$ $\cup \{\textsc{Target}_i, \textsc{Target}'_i \mid i \in I\}$ are newly introduced and where $i, j \in I$:

$$\textsc{Source}X \xhookrightarrow{a} \textsc{Wait}X\,, \qquad\qquad \textsc{Source}'X \xhookrightarrow{a} \textsc{Choose}'_i X\,,$$

$$\textsc{Source}X \xhookrightarrow{a} \textsc{Choose}'_i X\,,$$

$$\textsc{Wait}X \xhookrightarrow{a_i} \textsc{Target}_i X\,, \qquad\qquad \textsc{Choose}'_i X \xhookrightarrow{a_i} \textsc{Target}'_i X\,,$$

$$\textsc{Choose}'_i X \xhookrightarrow{a_j} \textsc{Target}_j X \quad \text{if } i \neq j.$$

The wit of the gadget, justified in details in Jančar and Srba [10, Section 4], is that for all stack contents $\gamma \in \Gamma^*$ we have

$$\textsc{Source}X\gamma \approx \textsc{Source}'X\gamma \quad \Longleftrightarrow \quad \exists i \in I.\, \textsc{Target}_i X\gamma \approx \textsc{Target}'_i X\gamma\,.$$

Analogously (by adding appropriate push rules at the end), given a set of control states $\{\textsc{Source}, \textsc{Source}'\} \cup$ $\{\textsc{Target}, \textsc{Target}'\}$, a set of stack symbols $\{X\} \cup \{Y_i \in \Gamma \mid i \in I\}$, one can construct a gadget

$$\langle \textsc{Source}, \textsc{Source}' \rangle X \xleftarrow{\textsc{Def}} \langle \textsc{Target}, \textsc{Target}' \rangle \ \{Y_i X \mid i \in I\}$$

such that for all stack contents $\gamma \in \Gamma^*$ we have $\textsc{Source}X\gamma \approx \textsc{Source}'X\gamma$ if, and only if, $\textsc{Target}Y_i X\gamma \approx$ $\textsc{Target}'Y_i X\gamma$ for some $i \in I$. In any case, both gadgets have size $O(|I|)$.

Going further, given a finite language $L \subseteq \Gamma^*$ of the form $L = \Sigma_1 \ldots \Sigma_m$, where $\Sigma_j \subseteq \Gamma$ for all $j \in [1, m]$, by suitably cascading the above gadget m times, one can construct a gadget

$$\langle \textsc{Source}, \textsc{Source}' \rangle X \xleftarrow{\textsc{Def}} \langle \textsc{Target}, \textsc{Target}' \rangle \ L$$

such that for all stack contents $\gamma \in \Gamma^*$ we have $\textsc{Source}X\gamma \approx \textsc{Source}'X\gamma$ if, and only if, there exists some $w \in L$ such that $\textsc{Target}w\gamma \approx \textsc{Target}'w\gamma$.

A *deterministic finite automaton (DFA)* is a tuple $A = (S, \Sigma, \delta_A, s_0, F)$, where S is a finite set of *states*, Σ is a finite *alphabet*, $\delta_A : S \times \Sigma \to S$ is the *transition function*, $s_0 \in S$ is the *initial state*, and $F \subseteq S$ is a set of *final states*. The function δ_A is naturally extended to a function from $S \times \Sigma^*$ inductively via $\delta_A(s, \varepsilon) = s$ and $\delta_A(s, aw) = \delta_A(\delta_A(s, a), w)$ for all $a \in \Sigma$ and all $w \in \Sigma^*$. A language $L \subseteq \Sigma^*$ is *regular* if $L = L(A) = \{w \in \Sigma^* \mid \delta_A(s_0, w) \in F\}$ for some DFA $A = (S, \Sigma, \delta_A, s_0, F)$. Finally, given a regular language $L \subseteq \Sigma^*$, where $\Sigma \subseteq \Gamma \setminus \{\#\}$, and two control states p and p' we introduce a gadget

$$\langle p, p' \rangle \# \xleftarrow{\textsc{Prefix-Check}\#L\#}$$

such that for all stack contents of the form $\gamma = \#\theta\#\beta$, where $\theta \in \Sigma^*$ and $\beta \in \Gamma^*$, we have $p\gamma \approx p'\gamma$ if, and only if, $\theta \in L$. Assuming some DFA $A = (S, \Sigma, \delta_A, s_0, F)$ such that $L(A) = L$, we construct the gadget as follows, where all symbols and control states (except for p and p') are freshly introduced, where $s \in S$, where s' is a copy of every such $s \in S$, and where $a \in \Sigma$:

$$p\# \xhookrightarrow{1} s_0\,, \qquad\qquad\qquad p'\# \xhookrightarrow{1} s'_0\,,$$

$$sa \xhookrightarrow{\varepsilon} \delta_A(s, a)\,, \qquad\qquad s'a \xhookrightarrow{\varepsilon} \delta_A(s, a)'\,,$$

$$s\# \xhookrightarrow{2} t \quad \text{if } s \in F, \qquad\qquad s'\# \xhookrightarrow{2} t' \quad \text{if } s \in F,$$

$$s\# \xhookrightarrow{3} t \quad \text{if } s \notin F.$$

Since the above rules only allow to execute runs reading at most two action symbols, note that, by construction, both $\{[p\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[p'\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite. For all of the prefix checking rules that we will introduce below where, say, L is an involved regular language, it holds implicitly that one can compute in polynomial time in $n = |x|$ a DFA A with $L(A) = L$.

9.3 Encoding numbers. The *binary presentation* of a number $k \in [0, N-1]$ is defined as

$$\alpha_k = \alpha_{k,\ell-1} \ldots \alpha_{k,0}, \text{ where } \alpha_{k,i} \in \{0,1\} \text{ are such that } k = \sum_{i=0}^{\ell-1} 2^i \cdot \alpha_{k,i}.$$

In particular, note that $\alpha_0 = 0^\ell$ and $\alpha_{N-1} = \alpha_{2^\ell-1} = 1^\ell$. Conversely, for each $u \in \{0,1\}^\ell$ let $\overline{u} \in \mathbb{N}$ be the unique $k \in [0, N-1]$ such that $\alpha(k) = u$. For any two $k, k' \in [0, N-1]$ with $k \neq k'$, we define $\mathsf{MSB}(k, k') = \max\{i \in [0, \ell-1] \mid \alpha_{k,i} \neq \alpha_{k',i}\}$ to be the most significant bit in which the binary presentations of k and k' differ.

In this subsection we present rules allowing us to verify whether one number is a successor of another number. We start by adding the following set of rules, where $i, j \in [0, \ell-1]$ and $b \in \{0,1\}$:

$$\mathrm{INC}_i b \xrightarrow{s} \mathrm{INC}_{i,\ell-1} b, \qquad\qquad\qquad \widetilde{\mathrm{INC}} b \xrightarrow{s} \widetilde{\mathrm{INC}}_{\ell-1} b,$$

$$\mathrm{INC}_{i,j} b \xrightarrow{b} \mathrm{INC}_{i,j-1} \quad \text{if } j > i,$$

$$\mathrm{INC}_{i,j} 0 \xrightarrow{1} \mathrm{INC}_{i,j-1} \quad \text{if } j = i, \qquad\qquad \widetilde{\mathrm{INC}}_j b \xrightarrow{b} \widetilde{\mathrm{INC}}_{j-1}.$$

$$\mathrm{INC}_{i,j} 1 \xrightarrow{0} \mathrm{INC}_{i,j-1} \quad \text{if } j < i,$$

Observe that there is no rule involving $\mathrm{INC}_{i,-1}$ nor $\widetilde{\mathrm{INC}}_{-1}$ on its left-hand side. We have the following characterization:

LEMMA 9.2. *For all $i \in [0, \ell-1]$, all $k, k' \in [0, N-1]$, and all stack contents $\gamma, \gamma' \in \Gamma^*$ we have $\mathrm{INC}_i \alpha_k \gamma \approx \widetilde{\mathrm{INC}} \alpha_{k'} \gamma'$ if, and only if, $k' = k+1$ and $i = \mathsf{MSB}(k, k')$. Moreover, both $\{[\mathrm{INC}_i \gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\widetilde{\mathrm{INC}} \gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite.*

Proof. The first statement of the lemma follows immediately from inspection of the above rules and the equivalence of the following two statements:
- $k' = k + 1$ and $i = \mathsf{MSB}(k, k+1)$;
- the following three conditions hold for all $j \in [0, \ell-1]$:
    - $\alpha_{k,j} = \alpha_{k',j}$ if $j > i$,
    - ($\alpha_{k',j} = 1$ and $\alpha_{k,j} = 0$) if $j = i$, and
    - ($\alpha_{k,j} = 1$ and $\alpha_{k',j} = 0$) if $j < i$.

The second statement follows from inspection of the above rules, which imply that the weak bisimulation class both of $\mathrm{INC}\gamma$ and of $\widetilde{\mathrm{INC}}\gamma'$ is determined by the first ℓ letters of γ. □

We now add a few more rules. Assuming that the top of the stack is of the form $\alpha_k \#_i X \alpha_{k'}$ for some $X \in \Upsilon \cup (Q_{\mathcal{M}} \times \Upsilon)$, these rules allow us to verify, starting from two control states INC and INC', whether it holds that $k' = k+1$ and i is the most significant bit position in which k and k' differ. In the following rules we have $X \in \Upsilon \cup (Q_{\mathcal{M}} \times \Upsilon)$, $i \in [0, \ell-1]$, and $b \in \{0,1\}$:

$$\mathrm{INC} b \xrightarrow{a} \widetilde{\mathrm{INC}} b, \qquad\qquad\qquad \mathrm{INC}' b \xrightarrow{a} \mathrm{TEMP},$$

$$\mathrm{TEMP} b \xrightarrow{\varepsilon} \mathrm{TEMP},$$

$$\mathrm{TEMP} \#_i \xrightarrow{\varepsilon} \mathrm{TEMP}_i,$$

$$\mathrm{TEMP}_i X \xrightarrow{\varepsilon} \mathrm{INC}_i.$$

LEMMA 9.3. *For all $k, k' \in [0, N-1]$, all $i \in [0, \ell-1]$, all $X \in \Upsilon \cup (Q_{\mathcal{M}} \times \Upsilon)$, and all $\gamma \in \Gamma^*$ we have $\mathrm{INC}\alpha_k \#_i X \alpha_{k'} \gamma \approx \mathrm{INC}'\alpha_k \#_i X \alpha_{k'} \gamma$ if, and only if, $k' = k + 1$ and $\mathsf{MSB}(k, k') = i$. Moreover, both $\{[\mathrm{INC}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\mathrm{INC}'\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite.*

Proof. Looking at the first rules what will be applied (yielding an a-labeled transition, possibly followed by a sequence of ε-transitions) we see that $\mathrm{INC}\alpha_k \#_i X \alpha_{k'} \gamma \approx \mathrm{INC}'\alpha_k \#_i X \alpha_{k'} \gamma$ if, and only if, $\widetilde{\mathrm{INC}}\alpha_k \#_i X \alpha_{k'} \gamma \approx \mathrm{INC}_i \alpha_{k'} \gamma$, which by Lemma 9.2 holds if, and only if, $k' = k+1$ and $i = \mathsf{MSB}(k, k')$. That the sets $\{[\mathrm{INC}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\mathrm{INC}'\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are both finite sets of classes all of which are weakly bisimulation finite follows immediately from inspection of the above rules and Lemma 9.2. □

9.4 Encoding configurations. The encoding of a configuration $w = w_0 \ldots w_{N-1}$ of $\mathcal{M}$ is defined to be the following word $\beta_w \in \Gamma^*$:

$$\beta_w = w_0 \alpha_0 \left(\prod_{i=1}^{N-1} \#_{\mathsf{MSB}(i-1,i)} w_i \alpha_i \right).$$

Example. For $\ell = 3$ (hence for $N = 2^\ell = 8$) and $w = ab(q,c)abcbb$ we have

$$\beta_w = a000\#_0 b001\#_1(q,c)010\#_0 a011\#_2 b100\#_0 c101\#_1 b110\#_0 b111.$$

□

We remark that β_w contains precisely $2^{\ell-1-i}$ occurrences of $\#_i$, for every $i \in [0, \ell-1]$.

DEFINITION 9.1. *For all configurations $w = w_0 \ldots w_{N-1}$ of $\mathcal{M}$ and for all $u \in \{0,1\}^{\leq \ell}$ let us define, by induction on $|u|$, the infix $\beta_w^{(u)}$ of β_w:*
- *$\beta_w^{(\varepsilon)} = \beta_w$, and*
- *if $u \in \{0,1\}^i$ for some $i \in [0, \ell-1]$, then $\beta_w^{(u0)}$ and $\beta^{(u1)}$ are the unique infixes of $\beta^{(u)}$ satisfying $\beta_w^{(u)} = \beta_w^{(u0)} \#_{\ell-1-i} \beta_w^{(u1)}$.*

Recalling that $\overline{u}$ denotes the unique number that the binary string $u \in \{0,1\}^\ell$ encodes, for all $u \in \{0,1\}^\ell$ we have $\beta_w^{(u)} = w_{\overline{u}} \alpha_{\overline{u}}$. A simple induction yields that for all $i \in [0, \ell]$ and all $u \in \{0,1\}^i$, the infix $\beta_w^{(u)}$ appears exactly once in β_w and, for all $j \in [0, \ell-1]$ we have that $\beta_w^{(u)}$ contains precisely $\lfloor 2^{\ell-1-i-j} \rfloor$ occurrences of $\#_j$.

9.5 Gadget for checking the successor relation among configurations. In this subsection we show how to check whether the successor relation holds between two configurations written on the top of the stack. First, we construct a gadget verifying whether head positions agree; more precisely, given $d \in \{-1,1\}$, and assuming that the stack starts with $\#\beta_{w'}\#\beta_w$, the gadget allows us to test whether $\mathrm{Pos}(w') = \mathrm{Pos}(w) + d$. To this end, for all $d \in \{-1,1\}$, all $i \in [0, \ell-1]$, and all $X \in \Gamma$ we add the following rules:

$$(9.13) \qquad\qquad \langle \mathrm{Pos}_d, \mathrm{Pos}'_d \rangle \# \xrightarrow{\mathrm{DEF}} \{ \langle \mathrm{Pos}_{d,j}, \mathrm{Pos}'_{d,j} \rangle \# \mid j \in [0, \ell-1] \},$$

$$(9.14) \qquad \mathrm{Pos}_{d,i} X \xrightarrow{\varepsilon} \mathrm{Pos}_{d,i} \quad \text{if } X \notin Q_\mathcal{M} \times \Upsilon, \qquad\qquad \mathrm{Pos}'_{d,i} X \xrightarrow{\varepsilon} \mathrm{Pos}'_{d,i} \quad \text{if } X \notin Q_\mathcal{M} \times \Upsilon,$$

$$(9.15) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \mathrm{Pos}'_{d,i} X \xrightarrow{\varepsilon} \mathrm{Pos}''_{d,i} \quad \text{if } X \in Q_\mathcal{M} \times \Upsilon,$$

$$(9.16) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \mathrm{Pos}''_{d,i} X \xrightarrow{\varepsilon} \mathrm{Pos}''_{d,i} \quad \text{if } X \notin Q_\mathcal{M} \times \Upsilon,$$

$$(9.17) \qquad \mathrm{Pos}_{d,i} X \xrightarrow{\varepsilon} \mathrm{INC}_i \quad \text{if } X \in Q_\mathcal{M} \times \Upsilon, d = -1, \qquad \mathrm{Pos}''_{d,i} X \xrightarrow{\varepsilon} \widetilde{\mathrm{INC}} \quad \text{if } X \in Q_\mathcal{M} \times \Upsilon, d = -1,$$

$$(9.18) \qquad \mathrm{Pos}_{d,i} X \xrightarrow{\varepsilon} \widetilde{\mathrm{INC}} \quad \text{if } X \in Q_\mathcal{M} \times \Upsilon, d = 1, \qquad\quad \mathrm{Pos}''_{d,i} X \xrightarrow{\varepsilon} \mathrm{INC}_i \quad \text{if } X \in Q_\mathcal{M} \times \Upsilon, d = 1.$$

LEMMA 9.4. *For all $d \in \{-1,1\}$, all length-N configurations w, w' of $\mathcal{M}$, and all $\gamma \in \Gamma^*$ we have*

$$\mathrm{Pos}_d \# \beta_{w'} \# \beta_w \gamma \approx \mathrm{Pos}'_d \# \beta_{w'} \# \beta_w \gamma \quad \Longleftrightarrow \quad \mathrm{Pos}(w') = \mathrm{Pos}(w) + d.$$

Moreover, both $\{ [\mathrm{Pos}_d \gamma]_\approx \mid \gamma \in \Gamma^* \}$ and $\{ [\mathrm{Pos}'_d \gamma]_\approx \mid \gamma \in \Gamma^* \}$ are finite sets of classes all of which are weakly bisimulation finite.

Proof. Let $d \in \{-1,1\}$, let $w' = w'_0 \ldots w'_{N-1}$, $w = w_0 \ldots w_{N-1}$, $k' = \mathrm{Pos}(w')$, and $k = \mathrm{Pos}(w)$. We then have $w'_{k'}, w_k \in Q_\mathcal{M} \times \Upsilon$. By the definition on an encoding we have

$$\beta_{w'} = w'_0 \alpha_0 \left(\prod_{i=1}^{N-1} \#_{\mathsf{MSB}(i-1,i)} w'_i \alpha_i \right) \quad \text{and} \quad \beta_w = w_0 \alpha_0 \left(\prod_{i=1}^{N-1} \#_{\mathsf{MSB}(i-1,i)} w_i \alpha_i \right).$$

Let $\gamma \in \Gamma^*$ and let us define $\delta = \#\beta_{w'}\#\beta_w \gamma$. We provide a proof only in the case of $d = -1$; the case of $d = 1$ is completely analogous. By inspecting rules in Lines (9.14)–(9.18), for all $i \in [0, \ell-1]$ we have $\mathrm{Pos}_{d,i} \delta \gamma \approx \mathrm{Pos}'_{d,i} \delta \gamma$ if, and only if,

$$\mathrm{INC}_i \alpha_{k'} \#_{\mathsf{MSB}(k',k'+1)} \ldots w'_{N-1} \alpha_{N-1} \# \beta_w \gamma \approx \widetilde{\mathrm{INC}} \alpha_k \#_{\mathsf{MSB}(k,k+1)} \ldots w_{N-1} \alpha_{N-1} \gamma.$$

By Lemma 9.2 the latter equivalence holds if, and only if, $k' = k + 1$ and $i = \mathsf{MSB}(k, k')$. By properties of the Defender's forcing gadget used in Line (9.13) we have $\mathrm{Pos}_d\delta\gamma \approx \mathrm{Pos}'_d\delta\gamma$ if, and only if, there exists some $i \in [0, \ell-1]$ such that $\mathrm{Pos}_{d,i}\delta\gamma \approx \mathrm{Pos}'_{d,i}\delta\gamma$. Thus, $\mathrm{Pos}_d\delta\gamma \approx \mathrm{Pos}'_d\delta\gamma$ if, and only if, there exists some $i \in [0, \ell-1]$ such that $k' = k + 1$ and $i = \mathsf{MSB}(k, k')$, that is, if, and only if, $\mathrm{Pos}(w') = \mathrm{Pos}(w) + d$, as required.

Let us finally prove that $\{[\mathrm{Pos}_d\gamma]_\approx \mid \gamma \in \Gamma^*\}$ is a finite set of classes all of which are weakly bisimulation finite; the same can analogously be proven for $\{[\mathrm{Pos}'_d\gamma]_\approx \mid \gamma \in \Gamma^*\}$. By inspecting the above rules, there is a constant $c \in \mathbb{N}$ (essentially depending on the Defender's forcing gadget) such that for all stack contents $\gamma \in \Gamma^*$, when applying the rules from a configuration of the form $\mathrm{Pos}_d\gamma$ one can only execute a run reading at most c action symbols and for some $\gamma' \in \Gamma^*$ either get stuck in a dead end of the form $\mathrm{Pos}_{d,i}\gamma'$, or reach a configuration of the form $\mathrm{Inc}_i\gamma'$ or $\widetilde{\mathrm{Inc}}\gamma'$. In the former case, finiteness is clear, in the latter two cases finiteness immediately follows from the fact that $\{[\mathrm{Inc}_i\gamma]_\approx \mid \gamma \in \Gamma^*\}$ is a finite set of classes all of which are weakly bisimulation finite for all $i \in [0, \ell-1]$ and the same holds for the set $\{[\widetilde{\mathrm{Inc}}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ according to Lemma 9.2. $\qquad\square$

Once head positions are verified, the remaining part of the successor relation may be checked with a help of letter-to-letter morphisms, as explained in Lemma 9.1. This is realized by our next gadget, which for each $\mu \in T \times \Upsilon$ allows to verify, assuming that the top of the stack is $\#\beta_{w'}\#\beta_w$, whether $w \vdash_\mu w'$ holds. In order to construct this gadget, for all $\mu = (t, Z) \in T \times \Gamma$ with $t = (p, U, q, Y, d)$, all $X \in \Gamma$, and all $i \in [0, \ell-1]$ we add the following rules, where rules involving the control states Pos_d and Pos'_d have already been introduced above and where the morphisms h_μ and h'_μ are defined in Lemma 9.1:

$$(9.19) \qquad \mathrm{Succ}_\mu\# \xrightarrow{0} \mathrm{Pos}_d\#\,, \qquad\qquad\qquad \mathrm{Succ}'_\mu\# \xrightarrow{0} \mathrm{Pos}'_d\#\,,$$

$$(9.20) \qquad \mathrm{Succ}_\mu\# \xrightarrow{1} \mathrm{Desc}_{\mu,\ell}\,, \qquad\qquad\quad \mathrm{Succ}'_\mu\# \xrightarrow{1} \mathrm{Desc}'_{\mu,\ell-1}\,,$$

$$(9.21) \qquad \mathrm{Desc}_{\mu,\ell}X \xrightarrow{\varepsilon} \mathrm{Desc}_{\mu,\ell} \quad \text{if } X \neq \#,$$

$$(9.22) \qquad \mathrm{Desc}_{\mu,\ell}\# \xrightarrow{\varepsilon} \mathrm{Desc}_{\mu,\ell-1}\,,$$

$$(9.23) \qquad \mathrm{Desc}_{\mu,i}X \xrightarrow{0} \mathrm{Desc}_{\mu,i-1}X\,, \qquad\qquad \mathrm{Desc}'_{\mu,i}X \xrightarrow{0} \mathrm{Desc}'_{\mu,i-1}X\,,$$

$$(9.24) \qquad \mathrm{Desc}_{\mu,i}X \xrightarrow{1} \mathrm{Pop}_{\mu,i}X\,, \qquad\qquad\quad \mathrm{Desc}'_{\mu,i}X \xrightarrow{1} \mathrm{Pop}'_{\mu,i}X\,,$$

$$(9.25) \qquad \mathrm{Pop}_{\mu,i}X \xrightarrow{\varepsilon} \mathrm{Pop}_{\mu,i} \quad \text{if } X \neq \#_i, \qquad \mathrm{Pop}'_{\mu,i}X \xrightarrow{\varepsilon} \mathrm{Pop}'_{\mu,i} \quad \text{if } X \neq \#_i,$$

$$(9.26) \qquad \mathrm{Pop}_{\mu,i}\#_i \xrightarrow{\varepsilon} \mathrm{Desc}_{\mu,i-1}\,, \qquad\qquad \mathrm{Pop}'_{\mu,i}\#_i \xrightarrow{\varepsilon} \mathrm{Desc}'_{\mu,i-1}\,,$$

$$(9.27) \qquad \mathrm{Desc}_{\mu,-1}X \xrightarrow{h_\mu(X)} \mathrm{Final}_\mu X\,, \qquad\quad \mathrm{Desc}'_{\mu,-1}X \xrightarrow{h'_\mu(X)} \mathrm{Final}'_\mu X\,.$$

LEMMA 9.5. *For all length-N configurations w, w' of $\mathcal{M}$, all $\mu \in T \times \Upsilon$, and all $\gamma \in \Gamma^*$ we have*

$$\mathrm{Succ}_\mu\#\beta_{w'}\#\beta_w\gamma \approx \mathrm{Succ}'_\mu\#\beta_{w'}\#\beta_w\gamma \quad \Longleftrightarrow \quad w \vdash_\mu w'\,.$$

Moreover, both $\{[\mathrm{Succ}_\mu\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\mathrm{Succ}'_\mu\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite.

Proof. Let us first prove the second statement of the lemma. Consider the following (smallest) partial order $\succ$ relating the above control states as follows:

$$\mathrm{Succ}_\mu \nearrow^{\mathrm{Desc}_{\mu,\ell}} \searrow_{\mathrm{Desc}_{\mu,\ell-1} \succ \mathrm{Pop}_{\mu,\ell-1} \succ \mathrm{Desc}_{\mu,\ell-2} \succ \mathrm{Pop}_{\mu,\ell-2} \succ \cdots \succ \mathrm{Desc}_{\mu,-1} \succ \mathrm{Final}_\mu\,,}_{\mathrm{Pos}_\mu}$$

$$\mathrm{Succ}'_\mu \succ \mathrm{Desc}'_{\mu,\ell-1} \succ \mathrm{Pop}'_{\mu,\ell-1} \succ \mathrm{Desc}'_{\mu,\ell-2} \succ \mathrm{Pop}'_{\mu,\ell-2} \succ \cdots \succ \mathrm{Desc}'_{\mu,-1} \succ \mathrm{Final}'_\mu\,.$$
$$\searrow_{\mathrm{Pos}'_\mu}$$

For all rules that appear in Lines (9.19)–(9.27) in which, say, p_1 is the control state on the left-hand side of the rule and p_2 is the control state on the right-hand side, observe that either $p_1 = p_2$ or $p_1 \succ p_2$. Furthermore, if the

rule is a reading rule (i.e., is not an ε-rule) we have $\ell \succ r$. Since moreover any configuration having control state FINAL_μ or FINAL'_μ is a dead end, and recalling that both $\{[\text{POS}_\mu \gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\text{POS}'_\mu \gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite by Lemma 9.4, it follows that both $\{[\text{SUCC}_\mu \gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\text{SUCC}'_\mu \gamma]_\approx \mid \gamma \in \Gamma^*\}$ are indeed finite sets of classes all of which are weakly bisimulation finite.

Let us now prove the first statement of the lemma. Let $\mu = (t, Z)$ with $t = (p, U, q, Y, d)$ and let us introduce the abbreviation $\delta = \#\beta_{w'}\#\beta_w$. By inspecting Lines (9.19) and (9.20) we have $\text{SUCC}_\mu \delta \gamma \approx \text{SUCC}'_\mu \delta \gamma$ if, and only if, $\text{POS}_d \delta \gamma \approx \text{POS}'_d \delta \gamma$ and $\text{DESC}_{\mu,\ell}\beta_{w'}\#\beta_w\gamma \approx \text{DESC}'_{\mu,\ell-1}\beta_{w'}\#\beta_w\gamma$. Since, on the one hand, $\text{POS}_d \delta \gamma \approx \text{POS}'_d \delta \gamma$ if, and only if $\text{POS}(w') = \text{POS}(w) + d$ by Lemma 9.4 and, on the other hand, $w \vdash_\mu w'$ if, and only if, $\text{POS}(w') = \text{POS}(w) + d$ and $h_\mu(w) = h'_\mu(w')$ by Lemma 9.1, for proving the first statement of the lemma it is enough to prove that $\text{DESC}_{\mu,\ell}\beta_{w'}\#\beta_w\gamma \approx \text{DESC}'_{\mu,\ell-1}\beta_{w'}\#\beta_w\gamma$ if, and only if, $h_\mu(w) = h'_\mu(w')$. Hence, by inspection of Lines (9.21) and (9.22), it is sufficient to prove that

$$(9.28) \qquad \text{DESC}_{\mu,\ell-1}\beta_w\gamma \approx \text{DESC}'_{\mu,\ell-1}\beta_{w'}\#\beta_w\gamma \quad \Longleftrightarrow \quad h_\mu(w) = h'_\mu(w') \,.$$

Let $w = w_0 \ldots w_{N-1}$ and $w' = w'_0 \ldots w'_{N-1}$. For all $u \in \{0,1\}^{\leq \ell}$ we have defined an infix $\beta_w^{(u)}$ of β_w and an infix $\beta_{w'}^{(u)}$ of $\beta_{w'}$ (cf. Definition 9.1); recall that $\beta_w^{(u)}$ (and $\beta_{w'}^{(u)}$) appears precisely once in β_w (in $\beta_{w'}$, respectively). Let $\delta_w^{(u)}$ be the unique suffix of β_w starting just after $\beta_w^{(u)}$ (then $\beta_w^{(u)}\delta_w^{(u)}$ is a suffix of β_w), and analogously, let $\delta_{w'}^{(u)}$ be the unique suffix of $\beta_{w'}$ starting just after $\beta_{w'}^{(u)}$, (then $\beta_{w'}^{(u)}\delta_{w'}^{(u)}$ is a suffix of $\beta_{w'}$). For all $i \in [0, \ell-1]$ and all $u \in \{0,1\}^i$ we have that both $\beta_w^{(u)}$ and $\beta_{w'}^{(u)}$ contain precisely one occurrence of $\#_{\ell-1-i}$; we have $\beta_w^{(u)} = \beta_w^{(u0)}\#_{\ell-1-i}\beta_w^{(u1)}$, and analogously $\beta_{w'}^{(u)} = \beta_{w'}^{(u0)}\#_{\ell-1-i}\beta_{w'}^{(u1)}$ by Definition 9.1. As a consequence, by inspection of the rules from Lines (9.23)–(9.26), for all $i \in [0, \ell-1]$, all $u \in \{0,1\}^i$, and all $b \in \{0,1\}$ we have $\text{DESC}_{\mu,\ell-1-i}\beta_w^{(u)}\delta_w^{(u)}\gamma \xrightarrow{b} \text{DESC}_{\mu,\ell-2-i}\beta_w^{(ub)}\delta_w^{(ub)}\gamma$ and $\text{DESC}'_{\mu,\ell-1-i}\beta_{w'}^{(u)}\delta_{w'}^{(u)}\#\beta_w\gamma \xrightarrow{b} \text{DESC}'_{\mu,\ell-2-i}\beta_{w'}^{(ub)}\delta_{w'}^{(ub)}\#\beta_w\gamma$. We also have $\beta_w = \beta_w^{(\varepsilon)}\delta_w^{(\varepsilon)}$ and $\beta_{w'} = \beta_{w'}^{(\varepsilon)}\delta_{w'}^{(\varepsilon)}$. By a direct induction on $i \in [0, \ell]$ this implies that for all $u \in \{0,1\}^i$ we have $\text{DESC}_{\mu,\ell-1}\beta_w\gamma \xRightarrow{u} \text{DESC}_{\mu,\ell-1-i}\beta_w^{(u)}\delta_w^{(u)}\gamma$ and $\text{DESC}'_{\mu,\ell-1}\beta_{w'}\#\beta_w\gamma \xRightarrow{u} \text{DESC}'_{\mu,\ell-1-i}\beta_{w'}^{(u)}\delta_{w'}^{(u)}\#\beta_w\gamma$.

For all $u \in \{0,1\}^\ell$ we have $\beta_w^{(u)} = w_{\overline{u}}\alpha_{\overline{u}}$ and $\beta_{w'}^{(u)} = w'_{\overline{u}}\alpha_{\overline{u}}$ by Definition 9.1. As a consequence, by inspection of the rules in Line (9.27), for all $u \in \{0,1\}^\ell$ we have

$$\text{DESC}_{\mu,\ell-1}\beta_w\gamma \xRightarrow{u} \text{DESC}_{\mu,-1}\beta_w^{(u)}\delta_w^{(u)}\gamma \xRightarrow{h_\mu(w_{\overline{u}})} \text{FINAL}_\mu\beta_w^{(u)}\delta_w^{(u)}\gamma$$

and analogously

$$\text{DESC}'_{\mu,\ell-1}\beta_{w'}\#\beta_w\gamma \xRightarrow{u} \text{DESC}'_{\mu,-1}\beta_{w'}^{(u)}\delta_{w'}^{(u)}\#\beta_w\gamma \xRightarrow{h'_\mu(w'_{\overline{u}})} \text{FINAL}'_\mu\beta_{w'}^{(u)}\delta_{w'}^{(u)}\#\beta_w\gamma \,.$$

Inspecting the rules once again, we see that they are all deterministic, and that only symbols from $\{0,1\}$ can be read, until the state becomes $\text{DESC}_{\mu,-1}$ or $\text{DESC}'_{\mu,-1}$. After reaching such a state, it is only possible to read the symbol $h_\mu(X)$ or $h'_\mu(X)$ (depending on whether the state is $\text{DESC}_{\mu,-1}$ or $\text{DESC}'_{\mu,-1}$), where X is the topmost stack symbol. No further transitions are possible after reading this symbol (and reaching the state FINAL_μ or FINAL'_μ). It follows that all runs that one can execute from the configuration $\text{DESC}_{\mu,\ell-1}\beta_w\gamma$ are deterministic and the maximal such runs are precisely the runs reading a word from the set $\{uh_\mu(w_{\overline{u}}) \mid u \in \{0,1\}^\ell\}$. Analogously, all runs that one can execute from the configuration $\text{DESC}'_{\mu,\ell-1}\beta_{w'}\#\beta_w\gamma$ are deterministic and the maximal such runs are precisely the runs reading a word from the set $\{uh'_\mu(w'_{\overline{u}}) \mid u \in \{0,1\}^\ell\}$. Thus, Equivalence (9.28) and hence the lemma hold by the following equivalences:

$$\text{DESC}_{\mu,\ell-1}\beta_w\gamma \approx \text{DESC}'_{\mu,\ell-1}\beta_{w'}\#\beta_w\gamma$$
$$\Longleftrightarrow \quad \{uh_\mu(w_{\overline{u}}) \mid u \in \{0,1\}^\ell\} = \{uh'_\mu(w'_{\overline{u}}) \mid u \in \{0,1\}^\ell\}$$
$$\Longleftrightarrow \quad \forall u \in \{0,1\}^\ell.\, h_\mu(w_{\overline{u}}) = h'_\mu(w'_{\overline{u}})$$
$$\Longleftrightarrow \quad h_\mu(w) = h'_\mu(w') \,. \quad \square$$

9.6 A gadget for pushing successor configurations. Before listing quite involved rules, the following lemma states that one can design a gadget that allows to push a successor configuration of a configuration whose encoding is assumed to be on the top of the stack:

LEMMA 9.6. *For all $t = (p, U, q, Y, d) \in T$, when adding the rules Lines (9.29)–(9.40) below, for all length-N configurations w of $\mathcal{M}$ and all $\gamma \in \Gamma^*$ we have*

$$\textsc{Push}_t \# \beta_w \gamma \approx \textsc{Push}'_t \# \beta_w \gamma \quad \Longleftrightarrow \quad \begin{array}{c} \exists Z \in \Upsilon, w' \in \Upsilon^*(q, Z)\Upsilon^*. \, w \vdash_{(t,Z)} w' \text{ and} \\ \textsc{Play}_{(q,Z)} \# \beta_{w'} \# \beta_w \gamma \approx \textsc{Play}'_{(q,Z)} \# \beta_{w'} \# \beta_w \gamma, \end{array}$$

For all $t = (p, U, q, Y, d) \in T$ and all $X \in \Gamma$ we add the rule

$$(9.29) \qquad \langle \textsc{Push}_t, \textsc{Push}'_t \rangle X \xrightarrow{\textsc{Def}} \{ \langle \textsc{Push}_{(t,Z)}, \textsc{Push}_{(t,Z)} \rangle X \mid Z \in \Upsilon \},$$

and for all $\mu = (t, Z) \in \{t\} \times \Upsilon$ we add the following rules, where $\Theta = \Upsilon \cup (Q_{\mathcal{M}} \times \Upsilon)$ and where $\Omega = \{\#_j \mid j \in [0, \ell - 1]\}$:

$$(9.30) \qquad \langle \textsc{Push}_\mu, \textsc{Push}'_\mu \rangle X \xrightarrow{\textsc{Def}} \langle \textsc{Check}_\mu, \textsc{Check}'_\mu \rangle \, \Theta \{0,1\}^\ell \#_0 \Theta 1^\ell X,$$

$$(9.31) \qquad \langle \textsc{Ctd}_\mu, \textsc{Ctd}'_\mu \rangle X \xrightarrow{\textsc{Def}} \{ \langle \textsc{Next}_\mu, \textsc{Next}'_\mu \rangle X, \langle \textsc{Done}_\mu, \textsc{Done}'_\mu \rangle X \},$$

$$(9.32) \qquad \langle \textsc{Next}_\mu, \textsc{Next}'_\mu \rangle X \xrightarrow{\textsc{Def}} \langle \textsc{Check}_\mu, \textsc{Check}'_\mu \rangle \, \Theta \{0,1\}^\ell \Omega X,$$

$$(9.33) \qquad \textsc{Check}_\mu X \xrightarrow{1} \textsc{Inc}, \qquad\qquad\qquad \textsc{Check}'_\mu X \xrightarrow{1} \textsc{Inc}',$$

$$(9.34) \qquad \textsc{Check}_\mu X \xrightarrow{2} \textsc{Ctd}_\mu X, \qquad\qquad\qquad \textsc{Check}'_\mu X \xrightarrow{2} \textsc{Ctd}'_\mu X,$$

$$(9.35) \qquad \textsc{Done}_\mu X \xrightarrow{1} \textsc{HeadChk}_\mu \# X, \qquad\qquad \textsc{Done}'_\mu X \xrightarrow{1} \textsc{HeadChk}'_\mu \# X,$$

$$(9.36) \qquad \langle \textsc{HeadChk}_\mu, \textsc{HeadChk}'_\mu \rangle \# \xrightarrow{\textsc{Prefix-Check} \# L^*(Q_{\mathcal{M}} \times \Upsilon) L^* \#}, \quad \text{where } L = \Upsilon \cup \Omega \cup \{0,1\},$$

$$(9.37) \qquad \textsc{Done}_\mu X \xrightarrow{2} \textsc{FinChk}_\mu \# X, \qquad\qquad \textsc{Done}'_\mu X \xrightarrow{2} \textsc{FinChk}'_\mu \# X,$$

$$(9.38) \qquad \langle \textsc{FinChk}_\mu, \textsc{FinChk}'_\mu \rangle \# \xrightarrow{\textsc{Prefix-Check} \# \Theta 0^\ell (\Gamma \setminus \{\#\})^* \#},$$

$$(9.39) \qquad \textsc{Done}_\mu X \xrightarrow{3} \textsc{Succ}_\mu \# X, \qquad\qquad\qquad \textsc{Done}'_\mu X \xrightarrow{3} \textsc{Succ}'_\mu \# X,$$

$$(9.40) \qquad \textsc{Done}_\mu X \xrightarrow{4} \textsc{Play}_{(q,Z)} \# X, \qquad\qquad \textsc{Done}'_\mu X \xrightarrow{4} \textsc{Play}'_{(q,Z)} \# X.$$

We remark that the final control states of the form $\textsc{Play}_{(q,Z)}$ and $\textsc{Play}'_{(q,Z)}$ will be connected to a gadget further below.

Proof of Lemma 9.6. Let $t = (p, U, q, Y, d) \in T$, let $w \in \Upsilon^*(Q_{\mathcal{M}} \times \Upsilon)\Upsilon^*$ be some configuration of length N, and let $\gamma \in \Gamma^*$ be some stack content. For any stack content $\delta \in \Gamma^*$ and control states r and r' we introduce the notation $\langle r \approx r' \rangle \delta$ and $\langle r \not\approx r' \rangle \delta$ as an abbreviation for $r\delta \approx r'\delta$ and $r\delta \not\approx r'\delta$, respectively. For all $i \in [0, N-1]$ and all $Y_i, \ldots, Y_{N-1} \in \Theta$, let us also introduce the notation

$$\delta(Y_i, \ldots, Y_{N-1}) = Y_i \alpha_i \#_{\textsf{MSB}(i,i+1)} Y_{i+1} \alpha_{i+1} \ldots \#_{\textsf{MSB}(N-2,N-1)} Y_{N-1} \alpha_{N-1} \# \beta_w \gamma.$$

We have the following claim:

CLAIM 9.1. *For all $\mu \in \{t\} \times \Upsilon$, all $i \in [1, N-2]$, and all $Y_i, \ldots, Y_{N-1} \in \Theta$ we have*

$$\langle \textsc{Ctd}_\mu \approx \textsc{Ctd}'_\mu \rangle \delta(Y_i, \ldots, Y_{N-1}) \quad \Longleftrightarrow \quad \exists Y_{i-1} \in \Theta. \, \langle \textsc{Ctd}_\mu \approx \textsc{Ctd}'_\mu \rangle \delta(Y_{i-1}, Y_i \ldots, Y_{N-1}).$$

Proof of Claim 9.1. First we claim that $\langle \textsc{Done}_\mu \not\approx \textsc{Done}'_\mu \rangle \delta(Y_i, \ldots, Y_{N-1})$: indeed, the prefix

$$\# Y_i \alpha_i \#_{\textsf{MSB}(i,i+1)} Y_{i+1} \alpha_{i+1} \ldots \#_{\textsf{MSB}(N-2,N-1)} Y_{N-1} \alpha_{N-1} \#$$

of $\#\delta(Y_i,\ldots,Y_{N-1})$ is not in the regular language $\#\Theta 0^\ell(\Gamma\setminus\{\#\})^*\#$, simply because $\alpha_i\neq 0^\ell$ due to $i\neq 0$; thus by the rules in Lines (9.37) and (9.38) it follows that $\langle\mathrm{DONE}_\mu\not\approx\mathrm{DONE}'_\mu\rangle\delta(Y_i,\ldots,Y_{N-1})$.

Let us fix any $i\in[1,N-2]$. The claim follows from the following equivalences, where the second one follows from the just proven $\langle\mathrm{DONE}_\mu\not\approx\mathrm{DONE}'_\mu\rangle\delta(Y_i,\ldots,Y_{N-1})$ and the penultimate equivalence follows from the equivalence $\langle\mathrm{INC}\approx\mathrm{INC}'\rangle y\#_j\delta(Y_i,\ldots,Y_{N-1})\Longleftrightarrow y=\alpha_{i-1}$ and $j=\mathsf{MSB}(i-1,i)$, which holds for $y\in\{0,1\}^\ell$ by Lemma 9.3:

$$\langle\mathrm{CTD}_\mu\approx\mathrm{CTD}'_\mu\rangle\delta(Y_i,\ldots,Y_{N-1})$$

$$\overset{(9.31),(9.32)}{\Longleftrightarrow}\big(\exists Y_{i-1}\in\Theta,y\in\{0,1\}^\ell,\#_j\in\Omega.\,\langle\mathrm{CHECK}_\mu\approx\mathrm{CHECK}'_\mu\rangle Y_{i-1}y\#_j\delta(Y_i,\ldots,Y_{N-1})\big)\ \text{or}$$
$$\langle\mathrm{DONE}_\mu\approx\mathrm{DONE}'_\mu\rangle\delta(Y_i,\ldots,Y_{N-1})$$

$$\Longleftrightarrow\quad\exists Y_{i-1}\in\Theta,y\in\{0,1\}^\ell,\#_j\in\Omega.\,\langle\mathrm{CHECK}_\mu\approx\mathrm{CHECK}'_\mu\rangle Y_{i-1}y\#_j\delta(Y_i,\ldots,Y_{N-1})$$

$$\overset{(9.33),(9.34)}{\Longleftrightarrow}\exists Y_{i-1}\in\Theta,y\in\{0,1\}^\ell,\#_j\in\Omega.$$
$$\langle\mathrm{INC}\approx\mathrm{INC}'\rangle y\#_j\delta(Y_i,\ldots,Y_{N-1})\ \text{and}\ \langle\mathrm{CTD}_\mu\approx\mathrm{CTD}'_\mu\rangle Y_{i-1}y\#_j\delta(Y_i,\ldots,Y_{N-1})$$

$$\Longleftrightarrow\quad\exists Y_{i-1}\in\Theta.\,\langle\mathrm{CTD}_\mu\approx\mathrm{CTD}'_\mu\rangle Y_{i-1}\alpha_{i-1}\#_{\mathsf{MSB}(i-1,i)}\delta(Y_i,\ldots,Y_{N-1})$$

$$\Longleftrightarrow\quad\exists Y_{i-1}\in\Theta.\,\langle\mathrm{CTD}_\mu\approx\mathrm{CTD}'_\mu\rangle\delta(Y_{i-1},Y\ldots,Y_{N-1}).$$

This completes the proof of Claim 9.1. $\square$

Next we have the following claim, where we recall that $t=(p,U,q,Y,d)$ and $\mu=(t,Z)$:

CLAIM 9.2. *For all $\mu=(t,Z)\in\{t\}\times\Upsilon$ and all $w'=Y_0\ldots Y_{N-1}\in\Theta^N$ we have $\langle\mathrm{CTD}_\mu\approx\mathrm{CTD}'_\mu\rangle\delta(Y_0,\ldots,Y_{N-1})$ if, and only if, $w\vdash_\mu w'$ and $\langle\mathrm{PLAY}_{(q,Z)}\approx\mathrm{PLAY}'_{(q,Z)}\rangle\#\beta_{w'}\#\beta_w\gamma$.*

Proof of Claim 9.2. Let $\mu=(t,Z)\in\{t\}\times\Upsilon$ and let $w'=Y_0\ldots Y_{N-1}\in\Theta^N$. Recall that $\delta(Y_0,\ldots,Y_{N-1})=\beta_{w'}\#\beta_w\gamma$. Firstly, we claim that for all $Y\in\Theta$, all $y\in\{0,1\}^\ell$, and all $\#_j\in\Omega$ we have

$$(9.41)\qquad\qquad\qquad\qquad\langle\mathrm{CHECK}_\mu\not\approx\mathrm{CHECK}'_\mu\rangle Yy\#_j\beta_{w'}\#\beta_w\gamma.$$

Indeed, $\langle\mathrm{INC}\not\approx\mathrm{INC}'\rangle y\#_j\beta_{w'}\#\beta_w\gamma$ follows directly from Lemma 9.3, hence by the rules in Line (9.33) we obtain Property (9.41) as a consequence.

Secondly, we remark that the presence of the prefix checking rule from Line (9.38) (reachable due to the rules in Line (9.37)) does not impact the equivalence $\langle\mathrm{DONE}_\mu\approx\mathrm{DONE}'_\mu\rangle\beta_{w'}\#\beta_w\gamma$, simply as the unique prefix of $\#\beta_{w'}\#\beta_w\gamma$ lying in $\#(\Gamma\setminus\{\#\})^*\#$, namely $\#\beta_{w'}\#$, clearly lies in the regular language $\#\Theta 0^\ell(\Gamma\setminus\{\#\})^*\#$ appearing in the rule. Hence, it follows that the relevant equivalence $\langle\mathrm{DONE}_\mu\approx\mathrm{DONE}'_\mu\rangle\beta_{w'}\#\beta_w\gamma$ only depends on the application of the rules appearing in Lines (9.35), (9.36), (9.39) and (9.40). The claim now follows from the following equivalences:

$$\langle\mathrm{CTD}_\mu,\mathrm{CTD}'_\mu\rangle\delta(Y_0,\ldots,Y_{N-1})$$

$$\Longleftrightarrow\quad\langle\mathrm{CTD}_\mu,\mathrm{CTD}'_\mu\rangle\beta_{w'}\#\beta_w\gamma$$

$$\overset{(9.31),(9.32)}{\Longleftrightarrow}\quad\exists Y\in\Theta,y\in\{0,1\}^\ell,\#_j\in\Omega:\langle\mathrm{CHECK}_\mu\approx\mathrm{CHECK}'_\mu\rangle Yy\#_j\beta_{w'}\#\beta_w\gamma$$
$$\text{or }\langle\mathrm{DONE}_\mu\approx\mathrm{DONE}'_\mu\rangle\beta_{w'}\#\beta_w\gamma$$

$$\overset{(9.41)}{\Longleftrightarrow}\quad\langle\mathrm{DONE}_\mu\approx\mathrm{DONE}'_\mu\rangle\beta_{w'}\#\beta_w\gamma$$

$$\overset{(9.35),(9.36),(9.39),(9.40)}{\Longleftrightarrow}\quad w'\in\Upsilon^*(Q_\mathcal{M}\times\Upsilon)\Upsilon^*,\,\langle\mathrm{SUCC}_\mu\approx\mathrm{SUCC}'_\mu\rangle\#\beta_{w'}\#\beta_w\gamma,$$
$$\text{and }\langle\mathrm{PLAY}_{(q,Z)}\approx\mathrm{PLAY}'_{(q,Z)}\rangle\#\beta_{w'}\#\beta_w\gamma$$

$$\Longleftrightarrow\quad w'\text{ is a configuration, }\langle\mathrm{SUCC}_\mu\approx\mathrm{SUCC}'_\mu\rangle\#\beta_{w'}\#\beta_w\gamma,$$
$$\text{and }\langle\mathrm{PLAY}_{(q,Z)}\approx\mathrm{PLAY}'_{(q,Z)}\rangle\#\beta_{w'}\#\beta_w\gamma)$$

$$\overset{\text{Lemma 9.5}}{\Longleftrightarrow}\quad w'\text{ is a configuration, }w\vdash_\mu w',\text{ and }\langle\mathrm{PLAY}_{(q,Z)}\approx\mathrm{PLAY}'_{(q,Z)}\rangle\#\beta_{w'}\#\beta_w\gamma$$

$$\Longleftrightarrow\quad w\vdash_\mu w'\text{ and }\langle\mathrm{PLAY}_{(q,Z)}\approx\mathrm{PLAY}'_{(q,Z)}\rangle\#\beta_{w'}\#\beta_w\gamma.\quad\square$$

The lemma now follows from the following equivalences:

$$\langle \text{PUSH}_t, \text{PUSH}'_t\rangle \# \beta_w \gamma$$

$$\overset{(9.29)}{\Longleftrightarrow} \quad \exists \mu \in \{t\} \times \Upsilon : \langle \text{PUSH}_\mu \approx \text{PUSH}'_\mu\rangle \# \beta_w \gamma$$

$$\overset{(9.30)}{\Longleftrightarrow} \quad \exists \mu \in \{t\} \times \Upsilon, Y_{N-2}, Y_{N-1} \in \Theta, y \in \{0,1\}^\ell.$$
$$\langle \text{CHECK}_\mu \approx \text{CHECK}'_\mu\rangle Y_{N-2}y\#_0 Y_{N-1}1^\ell \# \beta_w \gamma$$

$$\overset{(9.33),(9.34)}{\Longleftrightarrow} \quad \exists \mu \in \{t\} \times \Upsilon, Y_{N-2}, Y_{N-1} \in \Theta, y \in \{0,1\}^\ell.$$
$$\left(\langle \text{INC} \approx \text{INC}'\rangle y\#_0 Y_{N-1}1^\ell \# \beta_w \gamma \text{ and } \langle \text{CTD}_\mu \approx \text{CTD}'_\mu\rangle Y_{N-2}y\#_0 Y_{N-1}1^\ell \# \beta_w \gamma \right)$$

$$\overset{\text{Lemma } 9.3}{\Longleftrightarrow} \quad \exists \mu \in \{t\} \times \Upsilon, Y_{N-2}, Y_{N-1} \in \Theta, \langle \text{CTD}_\mu \approx \text{CTD}'_\mu\rangle Y_{N-2}\alpha_{N-2}\#_0 Y_{N-1}\alpha_{N-1}\# \beta_w \gamma$$

$$\Longleftrightarrow \quad \exists \mu \in \{t\} \times \Upsilon, Y_{N-2}, Y_{N-1} \in \Theta. \langle \text{CTD}_\mu \approx \text{CTD}'_\mu\rangle \delta(Y_{N-2}, Y_{N-1})$$

$$\overset{\text{Claim } 9.1}{\Longleftrightarrow} \quad \exists \mu \in \{t\} \times \Upsilon, Y_{N-3}, Y_{N-2}, Y_{N-1} \in \Theta. \langle \text{CTD}_\mu \approx \text{CTD}'_\mu\rangle \delta(Y_{N-3}, Y_{N-2}, Y_{N-1})$$

$$\cdots$$

$$\overset{\text{Claim } 9.1}{\Longleftrightarrow} \quad \exists \mu \in \{t\} \times \Upsilon, Y_0, \ldots, Y_{N-1} \in \Theta. \langle \text{CTD}_\mu \approx \text{CTD}'_\mu\rangle \delta(Y_0, \ldots, Y_{N-1})$$

$$\overset{\text{Claim } 9.2}{\Longleftrightarrow} \quad \exists \mu = (t, Z) \in \{t\} \times \Upsilon, w' \in \Theta^N. w \vdash_\mu w' \text{ and } \langle \text{PLAY}_{(q,Z)} \approx \text{PLAY}'_{(q,Z)}\rangle \# \beta_{w'} \# \beta_w \gamma$$

$$\Longleftrightarrow \quad \exists Z \in \Upsilon, w' \in \Upsilon^*(q, Z)\Upsilon^*. w \vdash_\mu w' \text{ and } \langle \text{PLAY}_{(q,Z)} \approx \text{PLAY}'_{(q,Z)}\rangle \# \beta_{w'} \# \beta_w \gamma. \qquad \square$$

9.7 Simulating $\mathcal{M}$. The next gadget allows us to decide, assuming that the topmost stack content is of the form $\#\alpha_w$, whether the configuration w is indeed accepting. For all $(p, U) \in Q_\mathcal{M} \times \Upsilon$ and all $t \in T_{(p,U)}$ we add the following rules:

$$(9.42) \qquad \text{PLAY}_{(p,U)}\# \overset{t}{\hookrightarrow} \text{PUSH}_t\# \quad \text{if } p \in Q_\exists, \qquad \qquad \text{PLAY}'_{(p,U)}\# \overset{t}{\hookrightarrow} \text{PUSH}'_t\# \quad \text{if } p \in Q_\exists,$$

$$(9.43) \qquad \langle \text{PLAY}_{(p,U)}, \text{PLAY}'_{(p,U)}\rangle\# \overset{\text{DEF}}{\longleftrightarrow} \{\langle \text{PUSH}_{t'}, \text{PUSH}'_{t'}\rangle\# \mid t' \in T_{(p,U)}\} \quad \text{if } p \in Q_\forall.$$

We have the following accompanying lemma:

LEMMA 9.7. *For all $(p, U) \in Q_\mathcal{M} \times \Upsilon$, all length-$N$ configurations $w \in \Upsilon^*(p, U)\Upsilon^*$ of $\mathcal{M}$, and all $\gamma \in \Gamma^*$ we have*

$$\text{PLAY}_{(p,U)}\#\beta_w\gamma \not\approx \text{PLAY}'_{(p,U)}\#\beta_w\gamma \quad \Longleftrightarrow \quad w \text{ is accepting}.$$

Moreover, both $\{[\text{PLAY}_{(p,U)}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\text{PLAY}'_{(p,U)}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite.

Proof. Let us first prove the second statement of the lemma. Since the rules in Lines (9.42) and (9.43) mutually depend on the rules Lines (9.29)–(9.40) we analyze them together. First of all, observe that, immediately by the definition of the gadget, neither of the prefix checking rules in Lines (9.36) and (9.38) can contribute to an infinity of classes. An important consequence of Lemma 4.5 is that the class of every configuration $qX\eta$ is determined by the tuple $(q, X, ([r\eta]_\approx)_{r\in|X\rangle(q)})$.

By inspection of the rules in Lines (9.29)–(9.40), (9.42), and (9.43) one realizes that the only rule that *decreases* the stack height are the rules in Line (9.33), leading to the control states INC and INC', respectively. Hence the only such classes of the above form $[r\eta]_\approx$ are the classes of the form $[\text{INC}\gamma]_\approx$ or $[\text{INC}'\gamma]_\approx$. But the sets $\{[\text{INC}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\text{INC}'\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite by Lemma 9.3. Thus, it follows that $\{[\text{PLAY}_{(p,U)}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\text{PLAY}'_{(p,U)}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are finite sets of classes all of which are weakly bisimulation finite.

Let us now prove the first statement of the lemma. Let $(p, U) \in Q_\mathcal{M} \times \Upsilon$, let $w = \Upsilon^*(p, U)\Upsilon^*$ be a configuration of $\mathcal{M}$ and let $\gamma \in \Gamma^*$ be any stack content. We prove the statement by induction on the length of the longest computation starting in w. We make a case distinction whether w is universal or existential.

If w is existential we have the following equivalences:

$$w \text{ is accepting}$$
$$\Longleftrightarrow \quad \exists t = (p, U, q, Y, d) \in T_{(p,U)}, Z \in \Upsilon, w' \in \Upsilon^*(q, Z)\Upsilon^*. \, w \vdash_{(t,Z)} w' \text{ and } w' \text{ is accepting}$$
$$\stackrel{\text{IH}}{\Longleftrightarrow} \quad \exists t = (p, U, q, Y, d) \in T_{(p,U)}, Z \in \Upsilon, w' \in \Upsilon^*(q, Z)\Upsilon^*. \, w \vdash_{(t,Z)} w' \text{ and}$$
$$\text{PLAY}_{(q,Z)}\#\beta_{w'}\#\beta_w\gamma \not\approx \text{PLAY}'_{(q,Z)}\#\beta_{w'}\#\beta_w\gamma$$
$$\stackrel{\text{Lemma 9.6}}{\Longleftrightarrow} \quad \exists t = (p, U, q, Y, d) \in T_{(p,U)}. \, \text{PUSH}_t\#\beta_w\gamma \not\approx \text{PUSH}_t\#\beta_w\gamma$$
$$\stackrel{p\in Q_\exists,(9.42)}{\Longleftrightarrow} \quad \text{PLAY}_{(p,U)}\#\beta_w\gamma \not\approx \text{PLAY}'_{(p,U)}\#\beta_w\gamma \,.$$

If w is universal we have the following equivalences:

$$w \text{ is accepting}$$
$$\Longleftrightarrow \quad \forall t = (p, U, q, Y, d) \in T_{(p,U)}. \, \exists Z \in \Upsilon, w' \in \Upsilon^*(q, Z)\Upsilon^*. \, w \vdash_{(t,Z)} w' \text{ and } w' \text{ is accepting}$$
$$\stackrel{\text{IH}}{\Longleftrightarrow} \quad \forall t = (p, U, q, Y, d) \in T_{(p,U)}. \, \exists Z \in \Upsilon, w' \in \Upsilon^*(q, Z)\Upsilon^*. \, w \vdash_{(t,Z)} w' \text{ and}$$
$$\text{PLAY}_{(q,Z)}\#\beta_{w'}\#\beta_w\gamma \not\approx \text{PLAY}'_{(q,Z)}\#\beta_{w'}\#\beta_w\gamma$$
$$\stackrel{\text{Lemma 9.6}}{\Longleftrightarrow} \quad \forall t = (p, U, q, Y, d) \in T_{(p,U)}. \, \text{PUSH}_t\#\beta_w\gamma \not\approx \text{PUSH}_t\#\beta_w\gamma$$
$$\stackrel{p\in Q_\forall,(9.43)}{\Longleftrightarrow} \quad \text{PLAY}_{(p,U)}\#\beta_w\gamma \not\approx \text{PLAY}'_{(p,U)}\#\beta_w\gamma \,. \qquad \square$$

9.8   Setting up the initial configuration. Recall that α_k denotes the binary encoding for every $k \in [0, 2^\ell - 1]$, that $x = x_0 \ldots x_{n-1}$ is the input to $\mathcal{M}$ and that $q_0 \in Q$ is the initial state of $\mathcal{M}$. We add the following rules, where $\Theta_0 = \Upsilon \cup \{(q_0, x_0)\}$ and $\Omega = \{\#_j \mid j \in [0, \ell - 1]\}$:

$$(9.44) \qquad \langle \text{INIT}\#, \text{INIT}'\rangle\# \stackrel{\text{DEF}}{\longleftrightarrow} \langle \text{CHECK}, \text{CHECK}'\rangle \, \Box\alpha_{N-2}\#_0\Box\alpha_{N-1}\# \,,$$

$$(9.45) \qquad \langle \text{CTD}, \text{CTD}'\rangle X \stackrel{\text{DEF}}{\longleftrightarrow} \{\langle \text{NEXT}, \text{NEXT}'\rangle X, \langle \text{DONE}, \text{DONE}'\rangle X\} \,,$$

$$(9.46) \qquad \langle \text{NEXT}, \text{NEXT}'\rangle X \stackrel{\text{DEF}}{\longleftrightarrow} \langle \text{CHECK}, \text{CHECK}'\rangle \, \Theta_0\{0,1\}^\ell\Omega X \,,$$

$$(9.47) \qquad \text{CHECK}X \stackrel{1}{\hookrightarrow} \text{INC} \,, \qquad\qquad\qquad\qquad \text{CHECK}'X \stackrel{1}{\hookrightarrow} \text{INC}' \,,$$

$$(9.48) \qquad \text{CHECK}X \stackrel{2}{\hookrightarrow} \text{CTD}X \,, \qquad\qquad\qquad\qquad \text{CHECK}'X \stackrel{2}{\hookrightarrow} \text{CTD}'X \,,$$

$$(9.49) \qquad \text{DONE}X \stackrel{1}{\hookrightarrow} \text{INICHK}\#X \,, \qquad\qquad\qquad \text{DONE}'X \stackrel{1}{\hookrightarrow} \text{INICHK}'\#X \,.$$

$$(9.50) \qquad \langle \text{INICHK}, \text{INICHK}'\rangle\# \stackrel{\text{PREFIX-CHECK}\#(q_0,x_0)\alpha_0\#_0x_1\alpha_1\ldots\#_{\text{MSB}(n-2,n-1)}x_{n-1}\alpha_{n-1}(\Omega\Box\{0,1\}^\ell)^*\#}{\longleftrightarrow} \,,$$

$$(9.51) \qquad \text{DONE}X \stackrel{1}{\hookrightarrow} \text{PLAY}_{(q_0,x_0)}\#X \,, \qquad\qquad\qquad \text{DONE}'X \stackrel{1}{\hookrightarrow} \text{PLAY}'_{(q_0,x_0)}\#X \,.$$

LEMMA 9.8. *For all $\gamma \in \Gamma^*$ we have $\text{INIT}\#\gamma \not\approx \text{INIT}'\#\gamma$ if, and only if, $x \in L(\mathcal{M})$. Moreover, $\{[\text{INIT}\gamma]_\approx \mid \gamma \in \Gamma^*\}$ and $\{[\text{INIT}'\gamma]_\approx \mid \gamma \in \Gamma^*\}$ are both finite sets of classes all of which are weakly bisimulation finite.*

Proof. Let us first prove the first statement of the lemma. Since the proof is very closely related to the proof of Lemma 9.6, so we only sketch it here.

Let $w_x = (q_0, x_0)x_1 \ldots x_{n-1}\Box^{N-n}$. We have the following equivalences, where the first equivalence can be proven analogously as Lemma 9.6:

$$\text{INIT}\#\gamma \not\approx \text{INIT}'\#\gamma \qquad \Longleftrightarrow \qquad \text{PLAY}_{(q_0,x_0)}\#\beta_{w_x}\#\gamma \not\approx \text{PLAY}'_{(q_0,x_0)}\#\beta_{w_x}\#\gamma$$
$$\stackrel{\text{Lemma 9.7}}{\Longleftrightarrow} \quad w_x \text{ is accepting}$$
$$\Longleftrightarrow \quad x \in L(\mathcal{M}) \,.$$

The second statement of the lemma can be shown in the same way as the analogous statement of Lemma 9.7.

$\square$

We conclude our construction by adding the following initial rules:

$$(9.52) \qquad q_\updownarrow \# \xrightarrow{a} q_\updownarrow A \#,$$

$$(9.53) \qquad q_\updownarrow A \xrightarrow{a} q_\updownarrow AA,$$

$$(9.54) \qquad q_\updownarrow A \xrightarrow{a} q_\updownarrow,$$

$$(9.55) \qquad q_\updownarrow A \xrightarrow{\$} q_\varepsilon,$$

$$(9.56) \qquad q_\varepsilon A \xrightarrow{\varepsilon} q_\varepsilon,$$

$$(9.57) \qquad q_\varepsilon \# \xrightarrow{\varepsilon} \textsc{Init}\#,$$

$$(9.58) \qquad q_\updownarrow \# \xrightarrow{\$} \textsc{Init}'\#.$$

The following lemma provides the final desired reduction.

LEMMA 9.9. *We have $x \in L(\mathcal{M})$ if, and only if, $(\mathcal{L}(\mathcal{P}), q_\updownarrow \#)$ is not weakly bisimulation finite.*

Proof. Let us first assume $x \in L(\mathcal{M})$. By Lemma 9.8 we have $\textsc{Init}\# \not\approx \textsc{Init}'\#$. For $n \geq 1$ we have $q_\updownarrow A^n \# \xRightarrow{\$} \textsc{Init}\#$, while reading $\$$ from $q_\updownarrow \#$ necessarily leads to $\textsc{Init}'\# \not\approx \textsc{Init}'\#$. It follows that $q_\updownarrow A^n \# \not\approx q_\updownarrow \#$ for all $n \geq 1$. In consequence $q_\updownarrow \#$ is the only reachable configuration in its class (once we enter the configuration $\textsc{Init}\#$ or $\textsc{Init}'\#$, we cannot read $\$$ any more). Observe now that if $q_\updownarrow A^n \# \xRightarrow{w} q_\updownarrow \#$ for some $n \in \mathbb{N}$, then $|w| \geq n$; on the other hand $q_\updownarrow A^n \# \xRightarrow{a^n} q_\updownarrow \#$. Then $\textsc{Dist}([q_\updownarrow A^n \#]_\approx, [q_\updownarrow \#]_\approx) = n$, which implies $q_\updownarrow A^n \# \not\approx q_\updownarrow A^m \#$ for all $n, m \in \mathbb{N}$ with $n \neq m$. Moreover, we have $q_\updownarrow \# \to^* q_\updownarrow A^n \#$ for all $n \in \mathbb{N}$. Hence, $q_\updownarrow \#$ is not weakly bisimulation finite since infinitely many configurations, that are pairwise not weakly bisimilar, are reachable from it.

Conversely, assume $x \notin L(\mathcal{M})$. Then $\textsc{Init}\# \approx \textsc{Init}'\#$ by Lemma 9.8. Moreover, $\textsc{Init}\#$ is weakly bisimulation finite by Lemma 9.8. From this and by inspection of the rules in Lines (9.52)–(9.58) one easily sees that $q_\updownarrow A^n \# \approx q_\updownarrow A^m \#$ for all $n, m \in \mathbb{N}$. In fact, $q_\updownarrow \#$ is weakly bisimilar to a finite pointed ε-LTS $(\mathcal{L}, c)$, where $\mathcal{L}$ has the following shape: the configuration c has an a-loop plus a $\$$-labeled transition to the weakly bisimulation finite $[\textsc{Init}\#]_\approx = [\textsc{Init}'\#]_\approx$. $\qquad \square$

10 Conclusion

In this paper we have shown that weak bisimulation finiteness is 2-EXPTIME-complete for pushdown systems with deterministic ε-transitions. This improves a previously known ACKERMANN upper bound of the problem and improves the previously best known 6-EXPSPACE upper bound when ε-transitions are not present. It also generalizes the 2-EXPTIME upper bound for regularity of deterministic pushdown automata and tightens a previously known EXPTIME lower bound for the problem. The more general case with unrestricted ε-transitions is indeed challenging, since our upper bound proof heavily relied on the fact that the underlying transition system is finitely branching.

References

[1] J. van Benthem. *Modal Correspondence Theory.* PhD thesis, University of Amsterdam, 1976.

[2] S. Böhm, S. Göller, and P. Jančar. Bisimulation equivalence and regularity for real-time one-counter automata. *J. Comput. Syst. Sci.*, 80(4):720–743, 2014.

[3] F. Carreiro. PDL is the bisimulation-invariant fragment of weak chain logic. In *30th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2015, Kyoto, Japan, July 6-10, 2015*, pages 341–352. IEEE Computer Society, 2015.

[4] R. J. van Glabbeek. The linear time-branching time spectrum (extended abstract). In J. C. M. Baeten and J. W. Klop, editors, *CONCUR '90, Theories of Concurrency: Unification and Extension, Amsterdam, The Netherlands, August 27-30, 1990, Proceedings*, volume 458 of *Lecture Notes in Computer Science*, pages 278–297. Springer, 1990.

[5] S. Göller and P. Parys. Bisimulation finiteness of pushdown systems is elementary. In H. Hermanns, L. Zhang, N. Kobayashi, and D. Miller, editors, *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 521–534. ACM, 2020.

[6] P. Jančar. Finiteness up to bisimilarity is decidable for pushdown processes. *CoRR*, abs/1305.0516v1, 2013.

[7] P. Jančar. Deciding semantic finiteness of pushdown processes and first-order grammars w.r.t. bisimulation equivalence. *J. Comput. Syst. Sci.*, 109:22–44, 2020.

[8] P. Jančar and J. Esparza. Deciding finiteness of Petri nets up to bisimulation. In F. M. auf der Heide and B. Monien, editors, *Automata, Languages and Programming, 23rd International Colloquium, ICALP96, Paderborn, Germany, 8-12 July 1996, Proceedings*, volume 1099 of *Lecture Notes in Computer Science*, pages 478–489. Springer, 1996.

[9] P. Jančar and S. Schmitz. Bisimulation equivalence of first-order grammars is ACKERMANN-complete. In *34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019, Vancouver, BC, Canada, June 24-27, 2019*, pages 1–12. IEEE, 2019.

[10] P. Jančar and J. Srba. Undecidability of bisimilarity by Defender's forcing. *J. ACM*, 55(1):5:1–5:26, 2008.

[11] D. Janin and I. Walukiewicz. On the expressive completeness of the propositional mu-calculus with respect to monadic second order logic. In U. Montanari and V. Sassone, editors, *CONCUR '96, Concurrency Theory, 7th International Conference, Pisa, Italy, August 26-29, 1996, Proceedings*, volume 1119 of *Lecture Notes in Computer Science*, pages 263–277. Springer, 1996.

[12] M. Kot. Regularity of BPP is PSPACE-complete. In *Proceedings of the 3rd Ph.D. Workshop of Faculty of Electrical En gineering and Computer Science, WOFEX*, pages 393–398. VŠB-TUO FEECS, 2005.

[13] A. Kučera and R. Mayr. On the complexity of semantic equivalences for pushdown automata and BPA. In K. Diks and W. Rytter, editors, *Mathematical Foundations of Computer Science 2002, 27th International Symposium, MFCS 2002, Warsaw, Poland, August 26-30, 2002, Proceedings*, volume 2420 of *Lecture Notes in Computer Science*, pages 433–445. Springer, 2002.

[14] A. Kučera and R. Mayr. On the complexity of checking semantic equivalences between pushdown processes and finite-state processes. *Inf. Comput.*, 208(7):772–796, 2010.

[15] F. Moller and A. M. Rabinovich. Counting on CTL*: On the expressive power of monadic path logic. *Inf. Comput.*, 184(1):147–159, 2003.

[16] D. E. Muller and P. E. Schupp. The theory of ends, pushdown automata, and second-order logic. *Theor. Comput. Sci.*, 37:51–75, 1985.

[17] J. Srba. Strong bisimilarity and regularity of basic parallel processes is pspace-hard. In H. Alt and A. Ferreira, editors, *STACS 2002, 19th Annual Symposium on Theoretical Aspects of Computer Science, Antibes - Juan les Pins, France, March 14-16, 2002, Proceedings*, volume 2285 of *Lecture Notes in Computer Science*, pages 535–546. Springer, 2002.

[18] J. Srba. Strong bisimilarity and regularity of basic process algebra is PSPACE-hard. In P. Widmayer, F. T. Ruiz, R. M. Bueno, M. Hennessy, S. J. Eidenbenz, and R. Conejo, editors, *Automata, Languages and Programming, 29th International Colloquium, ICALP 2002, Malaga, Spain, July 8-13, 2002, Proceedings*, volume 2380 of *Lecture Notes in Computer Science*, pages 716–727. Springer, 2002.

[19] J. Srba. Roadmap of infinite results. *Current Trends In Theoretical Computer Science, The Challenge of the New Century*, Vol 2: Formal Models and Semantics:337–350, 2004.

[20] C. Stirling. The joys of bisimulation. In L. Brim, J. Gruska, and J. Zlatuška, editors, *Mathematical Foundations of Computer Science 1998, 23rd International Symposium, MFCS'98, Brno, Czech Republic, August 24-28, 1998, Proceedings*, volume 1450 of *Lecture Notes in Computer Science*, pages 142–151. Springer, 1998.

[21] L. G. Valiant. Regularity and related problems for deterministic pushdown automata. *J. ACM*, 22(1):1–10, 1975.

[22] I. Walukiewicz. Pushdown processes: Games and model-checking. *Inf. Comput.*, 164(2):234–263, 2001.

[23] W. Zhang, Q. Yin, H. Long, and X. Xu. Bisimulation equivalence of pushdown automata is Ackermann-complete. In A. Czumaj, A. Dawar, and E. Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 141:1–141:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.