The AC^0-Complexity Of Visibly Pushdown Languages

Stefan Göller, Nathan Grosshans

The AC^0-Complexity Of Visibly Pushdown Languages

Stefan Göller*
stefan.goeller@uni-kassel.de

School of Electrical Engineering and Computer Science, Universität Kassel, Germany

Nathan Grosshans
nathan.grosshans@polytechnique.edu

Independent Scholar, France

Abstract

We study the question of which visibly pushdown languages (VPLs) are in the complexity class AC^0 and how to effectively decide this question. Our contribution is to introduce a particular subclass of one-turn VPLs, called *intermediate VPLs*, for which the raised question is entirely unclear: to the best of our knowledge our research community is unaware of containment or non-containment in AC^0 for *any* intermediate VPL. Our main result states that there is an algorithm that, given a visibly pushdown automaton, correctly outputs either that its language is in AC^0, outputs some $m \geq 2$ such that MOD_m is constant-depth reducible to L (implying that L is not in AC^0), or outputs a finite disjoint union of intermediate VPLs that L is constant-depth equivalent to. In the latter case one can moreover effectively compute $k, l \in \mathbb{N}_{>0}$ with $k \neq l$ such that the concrete intermediate VPL $L(S \to \varepsilon \mid ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2)$ is constant-depth reducible to the language L. Due to their particular nature we conjecture that either all intermediate VPLs are in AC^0 or all are not. As a corollary of our main result we obtain that in case the input language is a visibly counter language our algorithm can effectively determine if it is in AC^0 — hence our main result generalizes a result by Krebs et al. stating that it is decidable if a given visibly counter language is in AC^0 (when restricted to well-matched words).

For our proofs we revisit so-called Ext-algebras (introduced by Czarnetzki et al.), which are closely related to forest algebras (introduced by Bojańczyk and Walukiewicz), and use Green's relations.

Contents

*The first author was supported by the Agence Nationale de la Recherche grant no. ANR-17-CE40-0010.

1 Introduction

This paper studies the circuit complexity of formal word languages. It is well-known that the regular word languages are characterized as the languages recognizable by finite monoids. When restricting the finite monoids to be aperiodic Schützenberger proved that one obtains precisely the star-free regular languages [28]. In terms of logic, these correspond to the languages definable in first-order logic $\mathsf{FO}[<]$ by a result of McNaughton and Papert [29]. The more general class of regular languages expressible in $\mathsf{FO}[\mathsf{arb}]$, i.e. first-order logic with arbitrary numerical predicates, coincides with the regular languages in AC^0 [16, 19]. These can be characterized algebraically as the regular languages whose syntactic morphism is quasi-aperiodic [5]. The latter algebraic characterization also shows that it is decidable if a regular language is in AC^0.

Generalizing regular languages, input-driven languages were introduced by Mehlhorn [25]. They are described by pushdown automata whose input alphabet is partitioned into letters that are either of type call, internal, or return. Rediscovered by Alur and Madhusudan in 2004 [2] under the name

of *visibly pushdown languages (VPLs)*, it was shown that they enjoy many of the desirable effective closure properties of the regular languages. For instance, the visibly pushdown languages form an effective Boolean algebra. Algebraically, VPLs were characterized by Alur et al. [1] by congruences on well-matched words of finite index. Extending upon these, Czarnetzki et al. introduced so-called Ext-algebras [10]; these involve pairs of monoids (R, O) where O is a submonoid of R^R. Being tailored towards recognizing word languages, Ext-algebras are closely connected to forest algebras, introduced by Bojańczyk and Walukiewicz [7]: in [10] it is shown that a language of well-matched words is visibly pushdown if, and only if, its syntactic Ext-algebra is finite. While context-free languages are generally in $\mathsf{LOGCFL} = \mathsf{SAC}^1$, the visibly pushdown languages, as the regular languages, are known to be in NC^1 [11]. By a famous result of Barrington [4], there already exist regular languages that are NC^1-hard. In this paper we study the question of which visibly pushdown languages are in AC^0 and how one can effectively decide this question.

Related work. Visibly pushdown languages (VPLs) were introduced [2] via deterministic visibly pushdown automata (DVPA for short). Inspired by forest algebras [7] the paper [10] introduces Ext-algebras. Unfortunately, the definition of Ext-algebra morphisms in [10] is incorrect in that it provably does not lead to freeness.

The regular languages that are in AC^0 were effectively characterized by Barrington et al. [5]: a regular language is in AC^0 if, and only if, its syntactic morphism is quasi-aperiodic. By an automata-theoretic approach, Krebs et al. [22] effectively characterized the visibly counter languages that are in AC^0. These are particular VPLs that are essentially accepted by visibly pushdown automata that use only one stack symbol. In his PhD thesis [24] Ludwig already considers the question of which VPLs are in AC^0. Yet, his conjectural characterization contains several serious flaws — a detailled discussion of these shortcomings can be found in Section 8.

Our contribution. We reintroduce Ext-algebras, fix the notion of Ext-algebra morphisms and define the languages they recognize. We also reintroduce the syntactic Ext-algebra of languages of well-matched words. We rigorously prove classical results like freeness and minimality of syntactic Ext-algebras with respect to recognition. We prove that a language of well-matched words is a VPL if, and only if, it is recognizable by a finite Ext-algebra. While these results essentially revisit the constructions of [10], we use Ext-algebras as a technical tool for studying the complexity of visibly pushdown languages.

Fix a visibly pushdown alphabet Σ, i.e. Σ is partitioned into Σ_{call} (call letters), Σ_{int} (internal letters), and Σ_{ret} (return letters). Denoting $\Delta(u)$ as the difference between the number of occurrences of call and return letters in $u \in \Sigma^*$ a word $w \in \Sigma^*$ is *well-matched* if $\Delta(w) = 0$ and $\Delta(u) \geq 0$ for all prefixes u of w. A *context* is a pair (u, v) such that uv is well-matched — contexts have a natural composition operation: $(u, v) \circ (u', v') = (uu', v'v)$.

We introduce the following notions: a set of contexts $\mathcal{R}$ is *length-synchronous* if $|u|/|v| = |u'|/|v'|$ for all $(u, v), (u', v') \in \mathcal{R}$ with $\Delta(u), \Delta(u') > 0$ and *weakly length-synchronous* if $u = u'$ implies $|v| = |v'|$ and $v = v'$ implies $|u| = |u'|$ for all $(u, v), (u', v') \in \mathcal{R}$ with $\Delta(u), \Delta(u') > 0$. Any language L of well-matched words induces a congruence $\equiv_L$ on contexts: $(u, v) \equiv_L (u', v')$ if $xuwvy \in L \Leftrightarrow xu'wv'y \in L$ for all contexts (x, y) and all well-matched words w. We introduce the notion of quasi-counterfreeness: a VPL is *quasi-counterfree* if for all contexts $\sigma \in \Sigma^k \times \Sigma^l$ we have $\sigma^n \equiv_L \sigma^{n+1}$ for some $n \in \mathbb{N}$ or all contexts in $\Sigma^k \times \Sigma^l$ are not $\equiv_L$-equivalent to $\sigma \circ \sigma$. Finally, we introduce our central class of *intermediate VPLs*: a VPL is *intermediate* if it is quasi-counterfree and generated by a context-free grammar containing the production $S \to_G \varepsilon$, where S is the start nonterminal and whose other productions are of the form $T \to_G uT'v$ such that uv is well-matched, $u \in (\Sigma_{\text{int}}^* \Sigma_{\text{call}} \Sigma_{\text{int}}^*)^+$ and $v \in (\Sigma_{\text{int}}^* \Sigma_{\text{ret}} \Sigma_{\text{int}}^*)^+$ such that the set of contexts $\{(u, v) \mid S \Rightarrow_G^* uSv\}$ is weakly length-synchronous but not length-synchronous. Note that intermediate VPLs are particular one-turn visibly pushdown

languages, that is, visibly pushdown languages that are subsets of $(\Sigma \setminus \Sigma_{\text{ret}})^*(\Sigma \setminus \Sigma_{\text{call}})^*$. As an example, for all $k, l \geq 1$ with $k \neq l$ a concrete intermediate VPL, denoted by $\mathcal{L}_{k,l}$ is the one that is generated by the context-free grammar $S \to \varepsilon \mid ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2$: here a is a call letter, c is an internal letter and b_1 and b_2 are return letters.

As far as we know our community is unaware of whether *at all* there is some intermediate VPL that is *provably in* AC^0 *or provably not in* AC^0 — analogous remarks apply to ACC^0.

Our main result states that there is an algorithm that, given a DVPA A correctly outputs either $L(A) \in \mathsf{AC}^0$, outputs some $m \geq 2$ such that MOD_m is constant-depth reducible to L (thus witnessing that $L(A) \notin \mathsf{AC}^0$), or outputs a non-empty disjoint finite union of intermediate VPLs that $L(A)$ is constant-depth equivalent to. In the latter case one can moreover effectively compute $k, l \in \mathbb{N}_{>0}$ with $k \neq l$ such that the above-mentioned $\mathcal{L}_{k,l}$ is constant-depth reducible to $L(A)$. We conjecture that either all intermediate VPLs are in AC^0 or all are not: note that together with our main result this conjecture implies the existence of an algorithm that can determine if a given visibly pushdown language is in AC^0. As a corollary of our main result we obtain that in case the input language is a visibly counter language our algorithm can determine if it is in AC^0, hence our main result generalizes a result by Krebs et al. stating that it is decidable if a given visibly counter lanugage is in AC^0 (when restricted to well-matched words).

For our main result we extensively study Ext-algebras, the syntactic morphisms of VPLs, and make use of Green's relations.

Organization. Our paper is organized as follows. We introduce notation and give an overview of our main result in Section 2. In Section 3 we first recall general algebraic concepts and then revisit Ext-algebras and their correspondence to visibly pushdown languages. Section 4 introduces central notions like length-synchronicity and weak length-synchronicity for Ext-algebra morphisms and visibly pushdown languages. The proof of our main result is content of Section 5. In Section 6 we concern ourselves with the computability of the syntactic Ext-algebra as well as decidability of quasi-aperiodicity and (weak) length-synchronicity. We conclude in Section 7.

2 Preliminaries

By $\mathbb{N}$ we denote the non-negative integers and by $\mathbb{N}_{>0}$ the positive integers. For integers $i, j \in \mathbb{Z}$ we denote by $[i, j]$ the set $\{i, \ldots, j\}$. For any function $f : X \to Y$ and any subset $Z \subseteq X$ we denote by $f|_Z : Z \to Y$ the restriction of f to domain Z, i.e. $f|_Z(z) = f(z)$ for all $z \in Z$.

For all words $w = w_1 \cdots w_n$, where $w_i \in \Sigma$ for all $i \in [1, n]$, and for subsets $\Gamma \subseteq \Sigma$, let $|w|_\Gamma = |\{i \in [1, |w|] \mid w_i \in \Gamma\}|$ denote the number of occurrences of letters in Γ. For all $a \in \Gamma$ we write $|w|_a$ to denote $|w|_{\{a\}}$.

We define the languages

$$\text{EQUALITY} = \{w \in \{0,1\}^* : |w|_0 = |w|_1\} \quad \text{and} \quad \text{MOD}_m = \{w \in \{0,1\}^* : m \text{ divides } |w|_1\}$$

for each $m \geq 2$.

A *visibly pushdown alphabet* is a finite alphabet $\Sigma = \Sigma_{call} \cup \Sigma_{int} \cup \Sigma_{ret}$, where the alphabets $\Sigma_{call}, \Sigma_{int}$, and Σ_{ret} are pairwise disjoint.

Definition 2.1. *The set of* well-matched words over a visibly pushdown alphabet Σ*, denoted by* $\Sigma^\triangle$*, is the smallest set satisfying the following:*

- $\varepsilon \in \Sigma^\triangle$ *and* $c \in \Sigma^\triangle$ *for all* $c \in \Sigma_{int}$*,*
- $awb \in \Sigma^\triangle$ *for all* $w \in \Sigma^\triangle$*,* $a \in \Sigma_{call}$ *and* $b \in \Sigma_{ret}$*, and*

- $uv \in \Sigma^{\triangle}$ *for all* $u, v \in \Sigma^{\triangle} \setminus \{\varepsilon\}$.

A well-matched word $w \in \Sigma^{\triangle}$ is *one-turn* if $w \in (\Sigma \setminus \Sigma_{\mathrm{ret}})^*(\Sigma \setminus \Sigma_{\mathrm{call}})^*$. A language $L \subseteq \Sigma^{\triangle}$ is *one-turn* if it contains only one-turn words. Let Σ be a visibly pushdown alphabet. We define $\Delta\colon \Sigma^* \to \mathbb{Z}$ to be the height monoid morphism such that $\Delta(w) = |w|_{\Sigma_{call}} - |w|_{\Sigma_{ret}}$ for all $w \in \Sigma^*$. A *context* is a pair $(u, v) \in \Sigma^* \times \Sigma^*$ such that $uv \in \Sigma^{\triangle}$. The *composition* of two contexts $(u, v), (x, y) \in \mathsf{Con}(\Sigma)$ is defined as $(u, v) \circ (x, y) = (ux, yv)$. For $\sigma \in \mathsf{Con}(\Sigma)$ by σ^k we denote the k-fold composition $\sigma \circ \cdots \circ \sigma$. For any context $(u, v) \in \mathsf{Con}(\Sigma)$ and well-matched word $w \in \Sigma^{\triangle}$ we define $(u, v)w = uwv$. An equivalence relation $\equiv$ on $\mathsf{Con}(\Sigma)$ is a *congruence* if for all $\chi, \chi', \sigma, \tau \in \mathsf{Con}(\Sigma)$ we have that $\sigma \equiv \tau$ implies $\chi \circ \sigma \circ \chi' \equiv \chi \circ \tau \circ \chi'$. Given a congruence $\equiv$ over $\mathsf{Con}(\Sigma)$ we denote by $[\sigma]_{\equiv}$ the equivalence class of σ. Given a language of well-matched words $L \subseteq \Sigma^{\triangle}$ we write $\sigma \equiv_L \tau$ if for all $\chi \in \mathsf{Con}(\Sigma)$ and all $w \in \Sigma^{\triangle}$ we have $(\chi \circ \sigma)w \in L$ if, and only if, $(\chi \circ \tau)w \in L$. Clearly, $\equiv_L$ is a congruence.

Let us briefly introduce context-free grammars. A *context-free grammar* is a tuple $G = (V, \Sigma, P, S)$, where V is a finite set of *nonterminals*, Σ is a non-empty finite *alphabet*, $P \subseteq V \times (V \cup \Sigma)^*$ is a finite set of *productions*, and $S \in V$ is the *start nonterminal*. We write $T \to_G y$ whenever $(T, y) \in P$. The binary relation $\Rightarrow_G$ over $(V \cup \Sigma)^*$ is defined as $u \Rightarrow_G v$ if there exists a production $T \to_G y$ and $x, z \in (V \cup \Sigma)^*$ such that $u = xTz$ and $v = xyz$. By $L(G) = \{w \in \Sigma^* \mid S \Rightarrow_G^* w\}$ we denote the *language* of G where $\Rightarrow_G^*$ is the reflexive transitive closure of $\Rightarrow_G$.

In the following we introduce deterministic visibly pushdown automata, remarking that nondeterministic visibly pushdown automata are determinizable [2].

Definition 2.2. *A* deterministic visibly pushdown automaton *(DVPA) is a tuple* $A = (Q, \Sigma, \Gamma, \delta, q_0, F, \bot)$*, where*

- Q *is a finite set of* states*,*
- Σ *is a visibly pushdown alphabet, the* input alphabet*,*
- Γ *is a finite alphabet, the* stack alphabet*,*
- $q_0 \in Q$ *is the* initial state*,*
- $F \subseteq Q$ *is the set of* final states*,*
- $\bot \in \Gamma$ *is the* bottom-of-stack symbol*, and*
- $\delta\colon Q \times \Sigma \times \Gamma \to Q \times \big(\{\varepsilon\} \cup \Gamma \cup (\Gamma \setminus \{\bot\})\Gamma\big)$ *is the* transition function *such that for all* $q \in Q, a \in \Sigma, \alpha \in \Gamma$*:*
  - *if* $a \in \Sigma_{call}$*, then* $\delta(q, a, \alpha) \in Q \times (\Gamma \setminus \{\bot\})\alpha$*,*
  - *if* $a \in \Sigma_{ret}$*, then* $\delta(q, a, \alpha) \in Q \times \{\varepsilon\}$*, and*
  - *if* $a \in \Sigma_{int}$*, then* $\delta(q, a, \alpha) \in Q \times \{\alpha\}$*.*

We define the *extended transition function* $\widehat{\delta}\colon Q \times \Sigma^* \times \Gamma^* \to Q \times \Gamma^*$ inductively as

- $\widehat{\delta}(q, \varepsilon, \beta) = (q, \beta)$ for all $q \in Q$ and $\beta \in \Gamma^*$,
- $\widehat{\delta}(q, w, \varepsilon) = (q, \varepsilon)$ for all $q \in Q$ and $w \in \Sigma^+$, and
- $\widehat{\delta}(q, aw, \alpha\beta) = \widehat{\delta}(p, w, \gamma\beta)$, where $\delta(q, a, \alpha) = (p, \gamma)$ for all $q \in Q$, $a \in \Sigma$, $w \in \Sigma^*$, $\alpha \in \Gamma$ and $\beta \in \Gamma^*$.

The *language* accepted by A is the language $L(A) = \{w \in \Sigma^* \mid \widehat{\delta}(q_0, w, \bot) \in F \times \{\bot\}\}$. We call

such a language a *visibly pushdown language* (VPL). We remark that visibly pushdown languages are always subsets of $\Sigma^{\triangle}$.

We refer to [17] for further details on formal language theory.

Semi-linear sets. Given $d \in \mathbb{N}_{>0}$, for $\vec{x} = (x_1, \ldots, x_d), \vec{y} = (y_1, \ldots, y_d) \in \mathbb{N}^d$ we define $\vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_d + y_d)$. We define the *norm* of a vector $\vec{x} \in \mathbb{N}^d$ as $\|\vec{x}\| = \max\{x_i \mid i \in [1, d]\}$. For $X, Y \subseteq \mathbb{N}^d$ define $X + Y = \{\vec{x} + \vec{y} \mid \vec{x} \in X, \vec{y} \in Y\}$. For $\vec{x} = (x_1, \ldots, x_d) \in \mathbb{N}^d$ and $n \in \mathbb{N}$ we define $n\vec{x} = (nx_1, \ldots, nx_d)$ and $\mathbb{N}\vec{x} = \{n\vec{x} \mid n \in \mathbb{N}\}$. A set $X \subseteq \mathbb{N}^d$ is *linear* if $X = \vec{y} + \sum_{i=1}^{k} \mathbb{N}\vec{x}_i$ for $k \in \mathbb{N}$ and $y, x_1, \ldots, x_k \in \mathbb{N}^d$ and it is *semilinear* if X is a finite union of linear sets.

2.1 Complexity and logic

We assume familiarity with standard circuit complexity, we refer to [31, 21] for an introduction to the topic. Recall the following Boolean functions: the AND-function, the OR-function, the majority function (that outputs 1 if the majority of its inputs are 1s), and the mod_m function (that outputs 1 if the number of its inputs that are 1s is divisible by m) for all $m \geq 2$.

A circuit family $(C_n)_{n\in\mathbb{N}}$ *decides* a binary language $L \subseteq \{0,1\}^*$ if C_n is a circuit with n inputs such that $L \cap \{0,1\}^n = \{x_1 \ldots x_n \in \{0,1\}^n \mid C_n(x_1, \ldots, x_n) = 1\}$ for all $n \in \mathbb{N}$. In this paper, we will consider circuits deciding languages over arbitrary finite alphabets: to do this, we just consider implicitly that any language over an arbitrary finite alphabet comes with a fixed binary encoding that encodes each letter as a block of bits of fixed size. By $\leq_{\mathrm{cd}}$ we mean *constant-depth truth table reducibility* (or just constant-depth reducibility) as introduced in [8]. Formally for two languages $K \subseteq \Gamma^*$ and $L \subseteq \Sigma^*$ for finite alphabets Σ, Γ, we write $K \leq_{\mathrm{cd}} L$ in case there is a polynomial p, a constant $d \in \mathbb{N}$, and circuit family $(C_n)_{n\in\mathbb{N}}$ deciding L such that each circuit C_n satisfies the following: it is of depth at most d and size at most $p(n)$ and its non-input gates are either AND-labeled, OR-labeled, or so-called oracle gates, labeled by L, that are gates deciding $L \cap \Sigma^m$, where $m \leq p(n)$, such that there is no path from the output of an oracle gate to an input of another oracle gate. We write $K =_{\mathrm{cd}} L$ if $K \leq_{\mathrm{cd}} L$ and $L \leq_{\mathrm{cd}} K$; we also say that K and L are *constant-depth equivalent*. We say a language L is *hard* for a complexity class C (or just C-*hard*) if $L' \leq_{\mathrm{cd}} L$ for all $L' \in \mathsf{C}$. We say L is C-*complete* if L is C-hard and $L \in \mathsf{C}$. The following complexity classes are relevant in this paper:

- AC^0 is the class of all languages decided by circuit families with NOT gates, AND, OR gates of unbounded fan-in, constant depth and polynomial size;
- ACC^0 is the class of all languages decided by circuit families with NOT gates, AND, OR and modular gates (for some fixed m) of unbounded fan-in, constant depth and polynomial size;
- TC^0 is the class of all languages decided by circuit families with NOT gates, AND, OR and majority gates of unbounded fan-in, constant depth and polynomial size;
- NC^1 is the class of all languages decided by circuit families with NOT gates, AND, OR gates of bounded fan-in, logarithmic depth and polynomial size.

We also consider the framework of first order logic over finite words. (See [20, 29] for a proper introduction to the topic.) A *numerical predicate of arity* $r \in \mathbb{N}_{>0}$ is a symbol of arity r associated to a subset of ${\mathbb{N}_{>0}}^r$. Given a class $\mathcal{C}$ of numerical predicates and a finite alphabet Σ, we call $\mathsf{FO}_\Sigma[\mathcal{C}]$-formula a first order formula over finite words using the alphabet Σ and numerical predicates from the class $\mathcal{C}$. On occasions, we might also consider $\mathsf{FO}_{\Sigma,\leftrightsquigarrow}[\mathcal{C}]$-formulas that in comparison to the previous ones can use an additional binary predicate $\leftrightsquigarrow$ and are interpreted on structures (w, M)

with $w \in \Sigma^*$ and $M \subseteq [1, |w|]^2$, where everything is interpreted as for $\mathsf{FO}_\Sigma[\mathcal{C}]$-formulas on w excepted for $\leftrightsquigarrow$ that is interpreted by M. Given a class $\mathcal{C}$ of numerical predicates, by $\mathsf{FO}[\mathcal{C}]$ we denote the class of all languages over any finite alphabet Σ defined by a $\mathsf{FO}_\Sigma[\mathcal{C}]$-sentence. A classical result at the interplay of circuit complexity and logic is that $\mathsf{AC}^0 = \mathsf{FO}[\mathsf{arb}]$, where arb denotes the class of all numerical predicates (see [29, Theorem IX.2.1] or [20, Corollary 5.32]). The other numerical predicates that we will encounter in this paper are $<$, $+$ and MOD_m for all $m \in \mathbb{N}_{>0}$ (gathered together in the set $\mathsf{MOD} = \{\mathsf{MOD}_m \mid m > 0\}$).

2.2 Main result

The notion of length-synchronicity and weak length-synchronicity will be a central notion in our main result. In the following Σ always denotes a visibly pushdown alphabet.

Definition 2.3 ((Weak) Length-Synchronicity)**.** *Let $\mathcal{R} \subseteq \mathsf{Con}(\Sigma)$ be a set of contexts.*

- *$\mathcal{R}$ is* length-synchronous *if $|u|/|v| = |u'|/|v'|$ for all $(u, v), (u', v') \in \mathcal{R}$ with $\Delta(u), \Delta(u') > 0$.*
- *$\mathcal{R}$ is* weakly length-synchronous *if $u = u'$ implies $|v| = |v'|$ and $v = v'$ implies $|u| = |u'|$ for all $(u, v), (u', v') \in \mathcal{R}$ with $\Delta(u), \Delta(u') > 0$.*

Note that a set of contexts $\mathcal{R}$ is weakly length-synchronous if $\mathcal{R}$ is length-synchronous. Indeed, if, say $(u, v), (u, v') \in \mathcal{R}$, where $|v| \neq |v'|$ and $\Delta(u) > 0$, then $|u|, |v|, |v'| > 0$ and so the quotients $\frac{|u|}{|v|}$ and $\frac{|u|}{|v'|}$ are distinct, thus violating length-synchronicity of $\mathcal{R}$.

Definition 2.4 (Quasi-Counterfree)**.** *A VPL $L \subseteq \Sigma^\triangle$ is* quasi-counterfree *if for all $\sigma = (u, v) \in \mathsf{Con}(\Sigma)$ we have $\sigma^n \equiv_L \sigma^{n+1}$ for some $n \in \mathbb{N}$ or for all $\tau \in \Sigma^{|u|} \times \Sigma^{|v|} \cap \mathsf{Con}(\Sigma)$ we have $\tau \not\equiv_L \sigma \circ \sigma$.*

We will later show that quasi-counterfreeness of a VPL $L \subseteq \Sigma^\triangle$ is equivalent to the condition that there is no $k, l \in \mathbb{N}$ such that there is a subset of $\mathsf{Con}(\Sigma) \cap \Sigma^k \times \Sigma^l$ that forms a non-trivial group when considering the associated equivalence classes with respect to $\equiv_L$ (Proposition 4.18).

Example 2.5. Consider the visibly pushdown alphabet Σ, where $\Sigma_{\text{call}} = \{a\}$, $\Sigma_{\text{int}} = \{c\}$ and $\Sigma_{\text{ret}} = \{b_1, b_2\}$. For all $k, l \in \mathbb{N}_{>0}$ satisfying $k \neq l$, consider the language $\mathcal{L}_{k,l}$ generated by the context-free grammar $S \to ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2 \mid \varepsilon$. We have that the set of contexts $\{(u, v) \in \mathsf{Con}(\Sigma) \mid S \Rightarrow_G^* uSv\}$ is weakly length-synchronous since both the relation and its reverse is a partial function — however, it is not length-synchronous. It is also not hard to see that $\mathcal{L}_{k,l}$ is quasi-counterfree.

We say a context-free grammar $G = (V, \Sigma, P, S)$ is *vertically visibly pushdown* if the underlying alphabet Σ is a visibly pushdown alphabet, $S \to_G \varepsilon$, and all other productions of G are of the form $T \to_G uT'v$, where $uv \in \Sigma^\triangle$ is one-turn such that $u \in (\Sigma_{\text{int}}^* \Sigma_{\text{call}} \Sigma_{\text{int}}^*)^+$ and $v \in (\Sigma_{\text{int}}^* \Sigma_{\text{ret}} \Sigma_{\text{int}}^*)^+$. Note that each grammar generating $\mathcal{L}_{k,l}$ in Example 2.5 is vertically visibly pushdown. The following remark is obvious.

Remark 2.6. *The languages generated by vertically visibly pushdown grammars are one-turn VPLs.*

Definition 2.7 (Intermediate VPL)**.** *A VPL L is intermediate if it is quasi-counterfree and $L = L(G)$ for some vertically visibly pushdown grammar G for which $\mathcal{R}(G) = \{(u, v) \in \mathsf{Con}(\Sigma) \mid S \Rightarrow_G^* uSv\}$ is weakly-length synchronous but not length synchronous.*

Thus the languages $\mathcal{L}_{k,l}$ from Example 2.5 are all intermediate VPLs. Loosely speaking, they are the simplest intermediate VPLs. We have the following conjecture.

Conjecture 2.8. *There is no intermediate VPL that is in* ACC^0 *or* TC^0*-hard under constant-depth reductions.*

In fact, the authors are not even aware of any intermediate VPL that is provably not in AC^0. An indication for the inadequacies of known techniques to prove it is that the *robustness* [21] of intermediate VPLs can be proven to be constant. Further techniques, based for instance on the *switching lemma* [18] or on the *polynomial method* [6] also do not seem to be applicable.

Our main result is the following theorem.

Theorem 2.9. *There is an algorithm that, given a DVPA* A*, correctly outputs either*

- $L(A) \in \mathsf{AC}^0$*,*
- $m \geq 2$ *such that* $\mathrm{MOD}_m \leq_{cd} L(A)$ *(hence implying* $L(A) \notin \mathsf{AC}^0$*),*
- *vertically visibly pushdown grammars* $G_1, \ldots, G_m$ *each generating intermediate VPLs such that* $L =_{cd} \biguplus_{i=1}^{m} L(G_i)$*. In this case one can moreover effectively compute* $k, l \in \mathbb{N}$ *with* $k \neq l$ *such that* $\mathcal{L}_{k,l} \leq_{cd} L(A)$*.*

Theorem 2.9 and the following conjecture imply the existence of an algorithm that decides if a given visibly pushdown language is in AC^0.

Conjecture 2.10. *Either all intermediate VPLs are in* AC^0 *or all are not.*

2.3 Corollary for visibly counter languages

A *visibly counter automaton* with threshold m (m-VCA) over a visibly pushdown alphabet Σ is a tuple $A = (Q, \Sigma, q_0, F, \delta_0, \ldots, \delta_m)$, where Q is a finite set of *states*, q_0 is the *initial state*, $F \subseteq Q$ is a set of *final states*, $m \geq 0$ is a *threshold*, and $\delta_i : Q \times \Sigma \to Q$ is a *transition function* for each $i \in [0, m]$.

A *configuration of* A is an element of $Q \times \mathbb{N}$. For any two configurations $(q, n), (q', n')$ and any $x \in \Sigma$ we define $(q, n) \xrightarrow{x}_A (q', n')$ if $q' = \delta_{\min(n,m)}(q, x)$ and $n' = n + \Delta(x)$. The relation $\xrightarrow{x}_A$ is naturally extended to $\xrightarrow{w}_A$ for $w \in \Sigma^*$. By $L(A) = \{w \in \Sigma^{\triangle} \mid \exists q \in F : (q_0, 0) \xrightarrow{w}_A (q, 0)\}$ we denote the *language* (of well-matched words) of A. We remark that the language of any m-VCA is a visibly pushdown language. We also remark that the languages of m-VCAs are defined to be sets of well-matched words as in [3], whereas in [23] the well-matched requirement is not present.

The following corollary implies the main result of [23] when restricted to well-matched words.

Corollary 2.11. *There is an algorithm that, given an* m*-VCA* A*, correctly outputs either that* $L(A)$ *is in* AC^0 *or some* $m \geq 2$ *such that* $\mathrm{MOD}_m \leq_{cd} L(A)$ *(hence implying* $L(A) \notin \mathsf{AC}^0$*).*

For the proof of Corollary 2.11 we refer to Section 5.5.

3 Language-theoretic and algebraic foundations and Ext-Algebras

3.1 Basic algebraic automata theory

For a thorough introduction to algebraic automata theory, we refer the reader to the two classical references of the domain by Eilenberg [12, 13] and Pin [26], but also to the following central reference in automata theory [27, Chapter 1].

A *semigroup* is a pair $(M, \cdot)$, where M is a non-empty set and $\cdot$ is a binary operation on M that is associative, i.e. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ for all $x, y, z \in M$. Usually, when the operation is clear from

the context, we write it multiplicatively and write just M instead of $(M, \cdot)$. The semigroup M is *trivial* if $|M| = 1$, and *non-trivial* otherwise. A *subsemigroup* of M is a semigroup N such that N is a subset of M and the operation of N is the restriction of the operation of M to N. We often just write xy to denote $x \cdot y$. An *idempotent* of a semigroup M is an element $x \in M$ satisfying $x = xx$. The *idempotent power* of a finite semigroup M is the smallest positive integer ω such that x^ω is an idempotent for all $x \in M$. The *zero* of a semigroup M is the unique element $x \in M$ (if it exists) satisfying $xy = yx = x$ for all $y \in M$. A *monoid* is a semigroup M with a neutral element, that is, an element $e \in M$ such that $x \cdot e = e \cdot x = x$ for all $x \in M$. We usually denote the neutral element of a monoid M by 1_M. A *submonoid* of M is a monoid N that is a subsemigroup of M containing 1_M (which is thus also the neutral element of N). Consider some monoid M. A *congruence* on M is an equivalence relation $\sim$ on M that satisfies $vxz \sim vyz$ for all $v, z \in M$ and all $x, y \in M$ with $x \sim y$. We denote by $[x]_\sim$ the equivalence class of $x \in M$. The *quotient* of M with respect to a congruence $\equiv$ is the monoid $M/{\equiv}$ with base set $M/{\equiv} = \{[m]_\equiv \mid m \in M\}$ and operation given by $[x]_\equiv \cdot [y]_\equiv = [xy]_\equiv$ for all $x, y \in M$.

A *group* is a monoid M in which for all $x \in M$ there exists an inverse, that is, an element $x' \in M$ such that $xx' = x'x = 1_M$. Each element in a group M has a unique inverse, so we denote by x^{-1} the unique inverse of an $x \in M$. A *subgroup* of a group M is a submonoid of M that is a group. Given a semigroup M, a set S and a subsemigroup N of M, whenever $N \subseteq S$, N is said to be *contained* in S. A semigroup M is *aperiodic* if it does not contain any non-trivial group. It is well-known that a finite semigroup M is aperiodic if, and only if, given ω the idempotent power of M, it holds that $x^\omega = x^{\omega+1}$ for all $x \in M$ if, and only if, there exists $k \in \mathbb{N}_{>0}$ such that $x^k = x^{k+1}$ for all $x \in M$.

A *morphism* from a monoid M to a monoid N is a mapping $\varphi\colon M \to N$ such that $\varphi(1_M) = 1_N$ and $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in M$. If $M = \Sigma^*$ and $N = \Gamma^*$ where Σ and Γ are finite alphabets, we call φ *length-multiplying* whenever there exists $k \in \mathbb{N}$ such that $\varphi(\Sigma) \subseteq \Gamma^k$. Let $\varphi\colon \Sigma^* \to M$ be a morphism, where Σ is a finite alphabet and M is finite. Then there exists $l \in \mathbb{N}_{>0}$ such that $\varphi(\Sigma^l) = \varphi(\Sigma^{2l})$: this implies that $\varphi(\Sigma^l)$ is a semigroup. The smallest such l is called the *stability index of the morphism* φ. It is easily shown that if $\varphi(\Sigma^n)$ contains a non-trivial group for some $n \in \mathbb{N}$, then so does $\varphi(\Sigma^l)$. We say that h is *quasi-aperiodic* if $\varphi(\Sigma^n)$ does not contain any non-trivial group for all $n \in \mathbb{N}$, which is equivalent to the fact that $\varphi(\Sigma^l)$ is aperiodic. (See [5, 29] for the original definition and [30] for the definition using the stability index, though it has been only formulated for surjective morphisms.)

A language L over a finite alphabet Σ is *recognized* by a monoid M if there is a morphism $\varphi\colon \Sigma^* \to M$ and $F \subseteq M$ such that $L = \varphi^{-1}(F)$. The *syntactic monoid* of a language $L \subseteq \Sigma^*$ is the quotient of Σ^* by the congruence $\sim_L$ (called the *syntactic congruence* of L) defined by $x \sim_L y$ for $x, y \in \Sigma^*$ whenever for all $u, v \in \Sigma^*$, $uxv \in L \Leftrightarrow uyv \in L$. The syntactic monoid of L recognizes L via the *syntactic morphism* of L sending any word $w \in \Sigma^*$ to $[w]_{\sim_L}$. A fundamental and well-known result is that a language L is regular if, and only if, it is recognized by a finite monoid if, and only if, its syntactic monoid is finite.

3.2 Ext-algebras

This section builds on [10], but identifies an inaccuracy in the definition of Ext-algebra morphisms to establish freeness.

Let $(M, \cdot, 1_M)$ be a monoid. For each $m \in M$, we shall respectively denote by left_m and right_m the *left-multiplication map* $x \mapsto m \cdot x$ and the *right-multiplication map* $x \mapsto x \cdot m$.

Definition 3.1. *An* Ext-algebra $(R, O, \cdot, \circ)$ *consists of a monoid* $(R, \cdot, 1_R)$ *and a monoid* $(O, \circ, 1_O)$ *that is a submonoid of* $(R^R, \circ)$ *containing the maps* left_r *and* right_r *for each* $r \in R$.

Definition 3.2. *Let (R, O) and (S, P) be* Ext*-algebras. An* Ext-algebra morphism from (R, O) to (S, P) *is a pair (φ, ψ) of monoid morphisms $\varphi\colon R \to S$ and $\psi\colon O \to P$ such that:*

- *for all $e \in O$ and $r \in R$ we have $\psi(e)(\varphi(r)) = \varphi(e(r))$;*
- *for all $r \in R$ we have $\psi(\mathrm{left}_r) = \mathrm{left}_{\varphi(r)}$ and $\psi(\mathrm{right}_r) = \mathrm{right}_{\varphi(r)}$.*

We write $(\varphi, \psi)\colon (R, O) \to (S, P)$. The morphism (φ, ψ) is called surjective *(respectively* bijective*) if both φ and ψ are* surjective *(respectively* bijective*).*

When it is clear from the context, we shall write *morphism* to mean Ext*-algebra morphism.*

Remark 3.3. *In the above definition, φ is totally determined by ψ, because for each $r \in R$, we have $\varphi(r) = \varphi(\mathrm{left}_r(1_R)) = \psi(\mathrm{left}_r)(\varphi(1_R)) = \psi(\mathrm{left}_r)(1_S)$.*

Definition 3.4. *Let (R, O) and (S, P) be* Ext*-algebras. Then*

- *(R, O) is a* sub-Ext-algebra *of (S, P) whenever R is a submonoid of S and there exists a submonoid O' of P such that $O = \{e|_R \mid e \in O'\}$, so that we may denote O by $O'|_R$.*
- *(R, O) is a* quotient *of (S, P) whenever there exists a surjective morphism from (S, P) to (R, O).*
- *(R, O)* divides *(S, P) whenever (R, O) is a quotient of a sub-*Ext*-algebra of (S, P).*

For the rest of this section, let us fix some visibly pushdown alphabet Σ.

Definition 3.5. *For all $(u, v) \in \mathsf{Con}(\Sigma)$, consider the function $\mathrm{ext}_{u,v}\colon \Sigma^\triangle \to \Sigma^*$ such that $\mathrm{ext}_{u,v}(x) = uxv$ for all $x \in \Sigma^\triangle$. We call*

$$\mathrm{ext}_{u,v} \quad = \quad \mathrm{ext}_{x_1,y_1} \circ \cdots \circ \mathrm{ext}_{x_m,y_m}$$

a factorization of $\mathrm{ext}_{u,v}$. That is, $u = x_1 \dots x_m$, $v = y_m \cdots y_1$.

The following lemma states that each $\mathrm{ext}_{u,v}$ has a unique factorization when restricting the (x_i, y_i) to be from $\Sigma^\triangle \times \Sigma^\triangle$ or from $\Sigma_{\mathrm{call}} \times \Sigma_{\mathrm{ret}}$ and minimizing the number of $(x_i, y_i) \in \Sigma^\triangle \times \Sigma^\triangle$: we obtain its so-called *stair factorization.*

Lemma 3.6. *For all $\mathrm{ext}_{u,v}$ there exists a unique factorization*

$$\mathrm{ext}_{u,v} \quad = \quad \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \quad \cdots \quad \circ \mathrm{ext}_{x_{h-1},y_{h-1}} \circ \mathrm{ext}_{a_{h-1},b_{h-1}} \circ \mathrm{ext}_{x_h,y_h}$$

satisfying $h \geq 1$, $x_i, y_i \in \Sigma^\triangle$ for all $i \in [1, h]$ and $a_i \in \Sigma_{call}$ and $b_i \in \Sigma_{ret}$ for all $i \in [1, h-1]$. In particular, $\mathrm{ext}_{u,v}$ is in fact a function from $\Sigma^\triangle$ to $\Sigma^\triangle$.

Proof. We show additionally that the required factorization must satisfy $h = \Delta(u) + 1$. We prove the statement by induction on $|uv|$. In case $|uv| \leq 1$, then either $\mathrm{ext}_{u,v} = \mathrm{ext}_{\varepsilon,\varepsilon}$, or there is some $c \in \Sigma_{\mathrm{int}}$ such that $\mathrm{ext}_{u,v} = \mathrm{ext}_{\varepsilon,c}$ or $\mathrm{ext}_{u,v} = \mathrm{ext}_{c,\varepsilon}$. In any of these cases, we uniquely factorize $\mathrm{ext}_{u,v}$ as ext_{x_1,y_1} with $x_1 = u$ and $y_1 = v$.

Let us consider the case when $|uv| \geq 2$ and let $h = \Delta(u) + 1$. Note that since $uv \in \Sigma^\triangle$ we have $u \in \Sigma^\triangle$ if, and only if, $v \in \Sigma^\triangle$. In case $u, v \in \Sigma^\triangle$ we have $\Delta(u) = 0$, hence the only factorization of the desired form is indeed $\mathrm{ext}_{u,v} = \mathrm{ext}_{x_1,y_1}$, where $x_1 = u$ and $y_1 = v$. Let us finally treat the case when $u, v \notin \Sigma^\triangle$, thus $\Delta(u) \geq 1$ and hence $h \geq 2$. Let x be the maximal prefix of u satisfying $x \in \Sigma^\triangle$ and let y be the maximal suffix of v satisfying $y \in \Sigma^\triangle$. Due to maximality of x and y there

must exist $a \in \Sigma_{\text{call}}$, $b \in \Sigma_{\text{ret}}$, and $u', v' \in \Sigma^*$ such that $u = xau'$, $v = v'by$ and $u'v' \in \Sigma^{\triangle}$ with $\Delta(u') = \Delta(u) - 1 = h - 2$. Let $\text{ext}_{x_1,y_1} \circ \text{ext}_{a_1,b_1} \circ \cdots \circ \text{ext}_{x_{h-2},y_{h-2}} \circ \text{ext}_{a_{h-2},b_{h-2}} \circ \text{ext}_{x_{h-1},y_{h-1}}$ be the unique factorization of the desired form for $\text{ext}_{u',v'}$ by induction hypothesis. We claim that

$$\text{ext}_{x,y} \circ \text{ext}_{a,b} \circ \text{ext}_{x_1,y_1} \circ \text{ext}_{a_1,b_1} \circ \cdots \circ \text{ext}_{x_{h-2},y_{h-2}} \circ \text{ext}_{a_{h-2},b_{h-2}} \circ \text{ext}_{x_{h-1},y_{h-1}}$$

is the unique factorization of the desired form for $\text{ext}_{u,v}$. Indeed, since $\Delta(u) \geq 1$ any potential factorization of the desired form for $\text{ext}_{u,v}$ must be of the form $\text{ext}_{x',y'} \circ \text{ext}_{a',b'} \circ \pi$, where x' is a prefix of u satisfying $x' \in \Sigma^{\triangle}$, y' is a suffix of v satisfying $y' \in \Sigma^{\triangle}$, $a' \in \Sigma_{\text{call}}$, and $b' \in \Sigma_{\text{ret}}$. In particular x' is a prefix of x and y' is suffix of y. In case $x' = x$ and $y' = y$ it follows $a' = a$ and $b' = b$ and uniqueness follows from induction hypothesis. It remains to consider the case when x' is a strict prefix of x or y' is a strict suffix of y. We only treat the former case. It must hold $x = x'a's$ for some $s \in \Sigma^+$ such that $a's \in \Sigma^{\triangle}$. But then π is a factorization for $\text{ext}_{su',v'z}$ for some $z \in \Sigma^*$ which is a contradiction since $\Delta(s) = -1$ due to $a's \in \Sigma^{\triangle}$. □

In the following we will denote the unique factorization provided by Lemma 3.6 as the *stair factorization* of $\text{ext}_{u,v}$. Consider now the set $\mathcal{O}(\Sigma^{\triangle})$ of all functions $\text{ext}_{u,v}$ for $(u, v) \in \mathsf{Con}(\Sigma)$: it is a subset of $(\Sigma^{\triangle})^{\Sigma^{\triangle}}$ closed under composition. Thus, $(\mathcal{O}(\Sigma^{\triangle}), \circ)$ is a submonoid of $((\Sigma^{\triangle})^{\Sigma^{\triangle}}, \circ)$. Since for all $w \in \Sigma^{\triangle}$ we have $\text{left}_w = \text{ext}_{w,\varepsilon}$ and $\text{right}_w = \text{ext}_{\varepsilon,w}$, the set $\mathcal{O}(\Sigma^{\triangle})$ contains the functions left_w and right_w for all $w \in \Sigma^{\triangle}$. Hence, $(\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle}), \cdot, \circ)$ is an Ext-algebra. The following important proposition establishes freeness of $(\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle}))$.

Proposition 3.7. *Let (R, O) be an* Ext*-algebra and consider two functions $\varphi\colon \Sigma_{int} \to R$ and $\psi\colon \{\text{ext}_{a,b} \mid a \in \Sigma_{call}, b \in \Sigma_{ret}\} \to O$. Then there exists a unique* Ext*-algebra morphism $(\overline{\varphi}, \overline{\psi})$ from $(\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle}))$ to (R, O) satisfying $\overline{\varphi}(c) = \varphi(c)$ for each $c \in \Sigma_{int}$ and $\overline{\psi}(\text{ext}_{a,b}) = \psi(\text{ext}_{a,b})$ for each $a \in \Sigma_{call}, b \in \Sigma_{ret}$.*

Proof. We define $\overline{\varphi}$ based on a refinement of the structural definition of well-matched words. For each $w \in \Sigma^{\triangle}$ we inductively define:

$$\overline{\varphi}(w) = \begin{cases} 1_R & \text{if } w = \varepsilon \text{ (type 1)} \\ \varphi(c) & \text{if } w = c \in \Sigma_{\text{int}} \text{ (type 2)} \\ \psi(\text{ext}_{a,b})(\overline{\varphi}(x)) & \text{if } w = axb \text{ for } a \in \Sigma_{\text{call}},\ b \in \Sigma_{\text{ret}} \text{ and } x \in \Sigma^{\triangle} \text{ (type 3)} \\ \overline{\varphi}(x)\overline{\varphi}(y) & \text{if } w = xy \text{ for } x, y \in \Sigma^{\triangle} \setminus \{\varepsilon\}, \text{ where } |x| \text{ is minimal (type 4)} \end{cases}$$

Observe that the four above types give unique decompositions. For proving that $\overline{\varphi}$ is indeed a monoid morphism one proves that for all $w, v \in \Sigma^{\triangle}$ we have $\overline{\varphi}(wv) = \overline{\varphi}(w)\overline{\varphi}(v)$ by structural induction on w given by the four types. The case $v = \varepsilon$ is direct, we only treat the case $v \in \Sigma^{\triangle} \setminus \{\varepsilon\}$ in the following. If w is of type 1 we have $\overline{\varphi}(wv) = \overline{\varphi}(v) = 1_R \cdot \overline{\varphi}(v) = \overline{\varphi}(w)\overline{\varphi}(v)$. If w is of type 2 or 3, then wv is of type 4 and w is the shortest prefix of wv with $w \in \Sigma^{\triangle} \setminus \{\varepsilon\}$, hence $\overline{\varphi}(wv) = \overline{\varphi}(w)\overline{\varphi}(v)$. If w is of type 4, then $w = xy$ for some $x, y \in \Sigma^{\triangle} \setminus \{\varepsilon\}$, where x is of minimal length. Then wv is of type 4, where $wv = x(yv)$ and x is the shortest prefix of wv with $x \in \Sigma^{\triangle} \setminus \{\varepsilon\}$. Hence $\overline{\varphi}(wv) = \overline{\varphi}(x)\overline{\varphi}(yv) = \overline{\varphi}(x)\overline{\varphi}(y)\overline{\varphi}(v) = \overline{\varphi}(xy)\overline{\varphi}(v) = \overline{\varphi}(w)\overline{\varphi}(v)$, where the first equality follows by definition of $\overline{\varphi}$ and the second and third equality follow from the induction hypothesis. Given any $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle})$ let

$$\text{ext}_{u,v} \quad = \quad \text{ext}_{x_1,y_1} \circ \text{ext}_{a_1,b_1} \circ \quad \cdots \quad \circ \text{ext}_{x_{h-1},y_{h-1}} \circ \text{ext}_{a_{h-1},b_{h-1}} \circ \text{ext}_{x_h,y_h}$$

be the unique stair factorization given by Lemma 3.6. We define

$$\overline{\psi}(\mathrm{ext}_{u,v}) = \bigcirc_{i=1}^{h-1}\big(\mathrm{left}_{\overline{\varphi}(x_i)} \circ \mathrm{right}_{\overline{\varphi}(y_i)} \circ \psi(\mathrm{ext}_{a_i,b_i})\big) \circ \mathrm{left}_{\overline{\varphi}(x_h)} \circ \mathrm{right}_{\overline{\varphi}(y_h)} \ .$$

For showing that $\overline{\psi}$ is indeed a monoid morphism, one proves $\overline{\psi}(\mathrm{ext}_{uu',v'v}) = \overline{\psi}(\mathrm{ext}_{u,v}) \circ \overline{\psi}(\mathrm{ext}_{u',v'})$ for all $\mathrm{ext}_{u,v}, \mathrm{ext}_{u',v'} \in \mathcal{O}(\Sigma^{\triangle})$ by observing simply that the unique stair factorization of $\mathrm{ext}_{uu',v'v}$ is obtained by composing the unique stair factorizations of $\mathrm{ext}_{u,v}$ and $\mathrm{ext}_{u',v'}$.

We now show that $(\overline{\varphi}, \overline{\psi})$ is in fact an Ext-algebra morphism. The discussion above first shows that both $\overline{\varphi}\colon \Sigma^{\triangle} \to R$ and $\overline{\psi}\colon \mathcal{O}(\Sigma^{\triangle}) \to O$ are monoid morphisms. Next, let us prove that for all $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle})$ and $w \in \Sigma^{\triangle}$ we have $\overline{\psi}(\mathrm{ext}_{u,v})(\overline{\varphi}(w)) = \overline{\varphi}(\mathrm{ext}_{u,v}(w))$. Let

$$\mathrm{ext}_{u,v} \quad = \quad \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \quad \cdots \quad \circ \mathrm{ext}_{x_{h-1},y_{h-1}} \circ \mathrm{ext}_{a_{h-1},b_{h-1}} \circ \mathrm{ext}_{x_h,y_h}$$

be the unique stair factorization of $\mathrm{ext}_{u,v}$ provided by Lemma 3.6. If $h = 1$, then

$$\overline{\psi}(\mathrm{ext}_{u,v})(\overline{\varphi}(w)) = \mathrm{left}_{\overline{\varphi}(x_h)} \circ \mathrm{right}_{\overline{\varphi}(y_h)}(\overline{\varphi}(w)) = \overline{\varphi}(x_h w y_h) = \overline{\varphi}(\mathrm{ext}_{u,v}(w)) \ .$$

Otherwise, we have

$$\begin{aligned}
&\overline{\psi}(\mathrm{ext}_{u,v})(\overline{\varphi}(w)) \\
=& \bigcirc_{i=1}^{h-1} \big(\mathrm{left}_{\overline{\varphi}(x_i)} \circ \mathrm{right}_{\overline{\varphi}(y_i)} \circ \psi(\mathrm{ext}_{a_i,b_i})\big) \circ \mathrm{left}_{\overline{\varphi}(x_h)} \circ \mathrm{right}_{\overline{\varphi}(y_h)}(\overline{\varphi}(w)) \\
=& \bigcirc_{i=1}^{h-1} \big(\mathrm{left}_{\overline{\varphi}(x_i)} \circ \mathrm{right}_{\overline{\varphi}(y_i)} \circ \psi(\mathrm{ext}_{a_i,b_i})\big)\big(\overline{\varphi}(x_h w y_h)\big) \\
=& \bigcirc_{i=1}^{h-2} \big(\mathrm{left}_{\overline{\varphi}(x_i)} \circ \mathrm{right}_{\overline{\varphi}(y_i)} \circ \psi(\mathrm{ext}_{a_i,b_i})\big) \circ \\
& \mathrm{left}_{\overline{\varphi}(x_{h-1})} \circ \mathrm{right}_{\overline{\varphi}(y_{h-1})} \circ \psi(\mathrm{ext}_{a_{h-1},b_{h-1}})\big(\overline{\varphi}(x_h w y_h)\big) \\
=& \bigcirc_{i=1}^{h-2} \big(\mathrm{left}_{\overline{\varphi}(x_i)} \circ \mathrm{right}_{\overline{\varphi}(y_i)} \circ \psi(\mathrm{ext}_{a_i,b_i})\big) \circ \\
& \mathrm{left}_{\overline{\varphi}(x_{h-1})} \circ \mathrm{right}_{\overline{\varphi}(y_{h-1})}\big(\overline{\varphi}(a_{h-1} x_h w y_h b_{h-1})\big) \\
=& \bigcirc_{i=1}^{h-2} \big(\mathrm{left}_{\overline{\varphi}(x_i)} \circ \mathrm{right}_{\overline{\varphi}(y_i)} \circ \psi(\mathrm{ext}_{a_i,b_i})\big)\big(\overline{\varphi}(x_{h-1} a_{h-1} x_h w y_h b_{h-1} y_{h-1})\big) \\
=& \cdots \\
=& \overline{\varphi}(x_1 a_1 \cdots x_{h-1} a_{h-1} x_h w y_h b_{h-1} y_{h-1} \cdots b_1 y_1) \\
=& \overline{\varphi}(\mathrm{ext}_{u,v}(w)) \ .
\end{aligned}$$

Let us prove that for all $w \in \Sigma^{\triangle}$ we have $\overline{\psi}(\mathrm{left}_w) = \mathrm{left}_{\overline{\varphi}(w)}$. Noting that the unique stair factorization of left_w is $\mathrm{ext}_{w,\varepsilon}$ we obtain

$$\overline{\psi}(\mathrm{left}_w) = \overline{\psi}(\mathrm{ext}_{w,\varepsilon}) = \mathrm{left}_{\overline{\varphi}(w)} \circ \mathrm{right}_{\overline{\varphi}(\varepsilon)} = \mathrm{left}_{\overline{\varphi}(w)} \circ \mathrm{right}_{1_R} = \mathrm{left}_{\overline{\varphi}(w)} \circ 1_O = \mathrm{left}_{\overline{\varphi}(w)} \ .$$

One proves $\overline{\psi}(\mathrm{right}_w) = \mathrm{right}_{\overline{\varphi}(w)}$ for all $w \in \Sigma^{\triangle}$ analogously.

Therefore, $(\overline{\varphi}, \overline{\psi})$ is an Ext-algebra morphism and it is the unique one satisfying $\overline{\varphi}(c) = \varphi(c)$ for each $c \in \Sigma_{\mathrm{int}}$ and $\overline{\psi}(\mathrm{ext}_{a,b}) = \psi(\mathrm{ext}_{a,b})$ for each $a \in \Sigma_{\mathrm{call}}, b \in \Sigma_{\mathrm{ret}}$. Take indeed any such Ext-algebra morphism (φ', ψ'): using the properties of Ext-algebra morphisms, it is straightforward to prove that then $\overline{\varphi}(w) = \varphi'(w)$ for all $w \in \Sigma^{\triangle}$ by structural induction on w and then to prove that $\overline{\psi}(\mathrm{ext}_{u,v}) = \psi'(\mathrm{ext}_{u,v})$ for all $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle})$ by using the unique stair factorization of $\mathrm{ext}_{u,v}$ provided by Lemma 3.6. □

Remark 3.8. *The second condition in Definition 3.2, i.e. for all $r \in R$ we have $\psi(\mathrm{left}_r) = \mathrm{left}_{\varphi(r)}$ and $\psi(\mathrm{right}_r) = \mathrm{right}_{\varphi(r)}$, does not appear in the definition of* Ext*-algebra morphisms given in [10].*

But this is actually problematic, because then Proposition 3.7 would not hold in general.

Indeed, consider for instance the visibly pushdown alphabet Γ *where* $\Gamma_{call} = \{a\}$, $\Gamma_{int} = \emptyset$ *and* $\Gamma_{ret} = \{b\}$, *where* R *the is semi-lattice on two elements* $\{0,1\}$ *such that* $1 \cdot 1 = 1$ *and* $0 \cdot 1 = 1 \cdot 0 = 0 \cdot 0 = 0$*; and moreover* O *is defined as* $\{\mathrm{id}, \mathbf{0}, \mathbf{1}\}$ *with* $\mathbf{0}(0) = \mathbf{0}(1) = 0$ *and* $\mathbf{1}(0) = \mathbf{1}(1) = 1$. *Then* (R, O) *is an* Ext*-algebra. Let us define the function* $\varphi\colon \Gamma^{\triangle} \to R$ *by* $\varphi(w) = 1$ *for all* $w \in \Gamma^{\triangle}$ *and the function* $\psi\colon \mathcal{O}(\Gamma^{\triangle}) \to O$ *by* $\psi(\mathrm{ext}_{a^n,b^n}) = \mathrm{id}$ *for all* $n \in \mathbb{N}$ *and* $\psi(\mathrm{ext}_{u,v}) = \mathbf{1}$ *for all* $u, v \in \Gamma^*$ *with* $uv \in \Gamma^{\triangle}$ *and* $(u \in a\Gamma^*b\Gamma^*$ *or* $v \in \Gamma^*a\Gamma^*b)$. *The pair* (φ, ψ) *forms a morphism from* $(\Gamma^{\triangle}, \mathcal{O}(\Gamma^{\triangle}))$ *to* (R, O)*, but it is not the only one sending* $\mathrm{ext}_{a,b}$ *to* id*, because we could also take* ψ *to send all elements of* $\mathcal{O}(\Gamma^{\triangle})$ *to* id.

Definition 3.9. *A language* $L \subseteq \Sigma^{\triangle}$ *is* recognized *by an* Ext*-algebra* (R, O) *whenever there exists a morphism* $(\varphi, \psi)\colon (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R, O)$ *such that* $L = \varphi^{-1}(F)$ *for some* $F \subseteq R$.

Example 3.10. Consider the language $\mathcal{L}_{1,2} = L(S \to aSb_1 \mid acSb_2 \mid \varepsilon)$ from Example 2.5 over the visibly pushdown alphabet Γ, where $\Gamma_{\mathrm{int}} = \{c\}$, $\Gamma_{\mathrm{call}} = \{a\}$ and $\Gamma_{\mathrm{ret}} = \{b_1, b_2\}$. Consider the Ext-algebra (R, O) defined as follows. We set $R = \{acb_1, \varepsilon, c, cab_1, ab_1\}$ with multiplication given by the following table:

$\cdot$	acb_1	ε	c	cab_1	ab_1
acb_1	acb_1	acb_1	acb_1	acb_1	acb_1
ε	acb_1	ε	c	cab_1	ab_1
c	acb_1	c	acb_1	acb_1	cab_1
cab_1	acb_1	cab_1	acb_1	acb_1	acb_1
ab_1	acb_1	ab_1	acb_1	acb_1	acb_1

Thus, observe that $\varepsilon = 1_R$ and acb_1 is the zero of R. Omitting its multiplication table, we set the monoid O to be the following

$$\begin{aligned} O \;=\; & \{(acb_1, \varepsilon), (\varepsilon, \varepsilon), (c, \varepsilon), (\varepsilon, c), (ab_1, \varepsilon), (\varepsilon, ab_1), (cab_1, \varepsilon)\} \;\cup \\ & \{(a, b_2), (ca, b_2), (ca, ab_1b_2), (ca, b_1), (a, ab_1b_2), (a, b_1)\}, \end{aligned}$$

where the elements in the first set comprise $\{\mathrm{left}_r, \mathrm{right}_r \mid r \in R\}$, more precisely

- $(acb_1, \varepsilon) = \mathrm{left}_{acb_1} = \mathrm{right}_{acb_1}$,
- $(\varepsilon, \varepsilon) = \mathrm{left}_{\varepsilon} = \mathrm{right}_{\varepsilon} = 1_O$,
- $(c, \varepsilon) = \mathrm{left}_c$,
- $(\varepsilon, c) = \mathrm{right}_c$,
- $(ab_1, \varepsilon) = \mathrm{left}_{ab_1}$,
- $(\varepsilon, ab_1) = \mathrm{right}_{ab_1}$,
- $(cab_1, \varepsilon) = \mathrm{left}_{cab_1} = \mathrm{right}_{cab_1}$,

and where the elements from the second set are the following functions from R to R, respectively:

- (a, b_2):

r	acb_1	ε	c	cab_1	ab_1
$(a, b_2)(r)$	acb_1	acb_1	ab_1	ab_1	acb_1

- (ca, b_2):

r	acb_1	ε	c	cab_1	ab_1
$(ca, b_2)(r)$	acb_1	acb_1	cab_1	cab_1	acb_1

- (ca, ab_1b_2):

r	acb_1	ε	c	cab_1	ab_1
$(ca, ab_1b_2)(r)$	acb_1	acb_1	cab_1	acb_1	acb_1

- (ca, b_1):

r	acb_1	ε	c	cab_1	ab_1
$(ca, b_1)(r)$	acb_1	cab_1	acb_1	acb_1	cab_1

- (a, ab_1b_2):

r	acb_1	ε	c	cab_1	ab_1
$(a, ab_1b_2)(r)$	acb_1	acb_1	ab_1	acb_1	acb_1

- (a, b_1):

r	acb_1	ε	c	cab_1	ab_1
$(a, b_1)(r)$	acb_1	ab_1	acb_1	acb_1	ab_1

.

Consider the unique morphism $(\varphi, \psi)\colon (\Gamma^{\triangle}, \mathcal{O}(\Gamma^{\triangle})) \to (R, O)$ that (thanks to Proposition 3.7) satisfies $\varphi(c) = c$, $\psi(\mathrm{ext}_{a,b_1}) = (a, b_1)$ and $\psi(\mathrm{ext}_{a,b_2}) = (a, b_2)$. We have $L = \varphi^{-1}(\{\varepsilon, ab_1\})$.

Definition 3.11. *Let (R, O) be an* Ext*-algebra. An* equivalence relation *on (R, O) is an equivalence relation $\sim$ on R. We say an equivalence relation $\sim$ is a* congruence *on (R, O) whenever for all $e \in O$ and for all $x, y \in R$ we have that $x \sim y$ implies $e(x) \sim e(y)$. In case $\sim$ is a congruence we denote by $(R, O)/\sim$ the pair $(R/\sim, O')$, where*

$$O' = \{e' \in (R/\sim)^{R/\sim} \mid \exists e \in O\ \forall x \in R\colon e'([x]_\sim) = [e(x)]_\sim\}.$$

The following lemma actually shows that $(R, O)/\sim$ is again an Ext-algebra, that we call *the quotient of (R, O) by $\sim$*.

Lemma 3.12. *Let (R, O) be an* Ext*-algebra and $\sim$ be a congruence on (R, O). Then $(R/\sim, O')$, with*

$$O' = \{e' \in (R/\sim)^{R/\sim} \mid \exists e \in O\ \forall x \in R\colon e'([x]_\sim) = [e(x)]_\sim\}$$

a submonoid of $(R/\sim)^{R/\sim}$, is an Ext*-algebra and the pair (φ, ψ) of functions $\varphi\colon R \to R/\sim$ and $\psi\colon O \to O'$ satisfying $\varphi(r) = [r]_\sim$ for all $r \in R$ and $\psi(e)([r]_\sim) = [e(r)]_\sim$ for all $e \in O$ and $r \in R$ is a surjective morphism from (R, O) to $(R/\sim, O')$.*

Proof. Let $u, v \in R$ such that $u \sim v$. Take any $x, y \in R$: we have that

$$xuy = \mathrm{right}_y \circ \mathrm{left}_x(u) \sim \mathrm{right}_y \circ \mathrm{left}_x(v) = xvy$$

by definition of congruence. Thus, $\sim$ is a congruence on R. This implies that $R/\sim$ is a monoid.

Let $e', f' \in O'$: this means there exist $e, f \in O$ such that $e'([r]_\sim) = [e(r)]_\sim$ and $f'([r]_\sim) = [f(r)]_\sim$ for all $r \in R$. Given any $r \in R$, we thus have

$$e' \circ f'([r]_\sim) = e'([f(r)]_\sim) = [e(f(r))]_\sim = [e \circ f(r)]_\sim\ ,$$

so that $e' \circ f' \in O'$. Therefore, O' is a submonoid of $(R/\sim)^{R/\sim}$ that contains the functions $\mathrm{left}_{[r]_\sim}$ and $\mathrm{right}_{[r]_\sim}$ for all $[r]_\sim \in R/\sim$. Thus, $(R/\sim, O')$ is an Ext-algebra.

Now define the functions $\varphi\colon R \to R/\sim$ and $\psi\colon O \to O'$ by respectively $\varphi(r) = [r]_\sim$ for all $r \in R$ and $\psi(e) = e'$ with $e' \in O'$ such that $e'([r]_\sim) = [e(r)]_\sim$ for all $r \in R$: this is well-defined because

$\sim$ is a congruence on (R, O). Since $\sim$ is a congruence on R, φ is a surjective monoid morphism. Further, let $e, f \in O$. We have

$$\begin{aligned}\psi(e) \circ \psi(f)([r]_\sim) &= \psi(e)([f(r)]_\sim) \\ &= [e(f(r))]_\sim \\ &= [e \circ f(r)]_\sim \\ &= \psi(e \circ f)([r]_\sim)\end{aligned}$$

for all $r \in R$, so that $\psi(e) \circ \psi(f) = \psi(e \circ f)$. Therefore, as $\psi(1_O)([r]_\sim) = [1_O(r)]_\sim = [r]_\sim$ for all $r \in R$, it follows that ψ is also a monoid morphism, that is obviously surjective. By construction, we do of course have that

$$\psi(e)(\varphi(r)) = \psi(e)([r]_\sim) = [e(r)]_\sim = \varphi(e(r))$$

for all $e \in O$ and $r \in R$. Moreover, for all $r \in R$, it holds that

$$\psi(\mathrm{left}_r)([x]_\sim) = [\mathrm{left}_r(x)]_\sim = [rx]_\sim = [r]_\sim[x]_\sim = \mathrm{left}_{\varphi(r)}([x]_\sim)$$

for all $x \in R$, so that $\psi(\mathrm{left}_r) = \mathrm{left}_{\varphi(r)}$. Similarly, we can prove that $\psi(\mathrm{right}_r) = \mathrm{right}_{\varphi(r)}$ for all $r \in R$. Thus, (φ, ψ) is a surjective morphism from (R, O) to $(R/\sim, O')$. □

The lemma also proves that the pair (φ, ψ) of functions $\varphi\colon R \to R/\sim$ and $\psi\colon O \to O'$ satisfying $\varphi(r) = [r]_\sim$ for all $r \in R$ and $\psi(e)([r]_\sim) = [e(r)]_\sim$ for all $e \in O$ and $r \in R$ is a surjective morphism from (R, O) to $(R, O)/\sim$. We also call this pair (φ, ψ) the *morphism associated to the congruence* $\sim$.

Definition 3.13. *The* syntactic congruence of a language $L \subseteq \Sigma^\triangle$ *is the congruence* $\sim_L$ *on* $(\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle))$ *defined by* $u \sim_L v$ *for* $u, v \in \Sigma^\triangle$ *whenever* $e(u) \in L \Leftrightarrow e(v) \in L$ *for all* $e \in \mathcal{O}(\Sigma^\triangle)$. *We define the* syntactic Ext-algebra of L *to be* $(R_L, O_L) = (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle))/\sim_L$ *and the* syntactic morphism of L *to be the morphism* (φ_L, ψ_L) *associated to* $\sim_L$.

Note that the syntactic Ext-algebra (R_L, O_L) of L recognizes L via the syntactic morphism (φ_L, ψ_L). Indeed, for all $u, v \in \Sigma^\triangle$, we have that if $u \sim_L v$, then $u \in L \Leftrightarrow v \in L$. This implies that $L = \varphi_L^{-1}(\varphi_L(L))$. For instance, it can be proven that the Ext-algebra recognizing the language $\mathcal{L}_{1,2}$ in Example 3.10 is in fact a certain presentation of the syntactic Ext-algebra of $\mathcal{L}_{1,2}$.

The next lemma states that all languages recognized by an Ext-algebra are also recognized by the Ext-algebras it divides.

Lemma 3.14. *Let* (R, O) *and* (S, P) *be two* Ext*-algebras such that* (R, O) *divides* (S, P). *Then any language* $L \subseteq \Sigma^\triangle$ *recognized by* (R, O) *is also recognized by* (S, P).

Proof. Let $L \subseteq \Sigma^\triangle$ be a language recognized by (R, O). This means that there exists a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ such that $L = \varphi^{-1}(F)$ for some $F \subseteq R$. We will prove the lemma by combining the following two points:

(1) if (R, O) is a sub-Ext-algebra of (S, P), then so does (S, P) recognize L, and

(2) if (R, O) is a quotient of (S, P), then so does (S, P) recognize L.

For Point (1), assume that (R, O) is a sub-Ext-algebra of (S, P). This means that R is a submonoid of S and that there exists a submonoid O' of P satisfying $O = O'|_R$. Take an arbitrary

function $\sigma\colon O \to P$ such that $\sigma(e)|_R = e$ for all $e \in O$. Let us consider the unique morphism $(\varphi', \psi')\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (S, P)$ such that $\varphi'(c) = \varphi(c)$ for all $c \in \Sigma_{\text{int}}$ and $\psi'(\text{ext}_{a,b}) = \sigma(\psi(\text{ext}_{a,b}))$ for all $a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}$, given to us by Proposition 3.7. We can prove by induction on w that $\varphi'(w) = \varphi(w)$ for all $w \in \Sigma^\triangle$:

- $w = \varepsilon$. Then $\varphi'(w) = 1_S = 1_R = \varphi(w)$.
- $w = c$ for some $c \in \Sigma_{\text{int}}$. Then $\varphi'(w) = \varphi'(c) = \varphi(c) = \varphi(w)$.
- $w = aw'b$ for some $a \in \Sigma_{\text{call}}$, $b \in \Sigma_{\text{ret}}$ and $w' \in \Sigma^\triangle$. Then

$$\begin{aligned}
\varphi'(w) &= \varphi'(\text{ext}_{a,b}(w')) \\
&= \psi'(\text{ext}_{a,b})(\varphi'(w')) \\
&\overset{\text{IH}}{=} \sigma(\psi(\text{ext}_{a,b}))(\varphi(w')) \\
&= \sigma(\psi(\text{ext}_{a,b}))|_R(\varphi(w')) \\
&= \psi(\text{ext}_{a,b})(\varphi(w')) \\
&= \varphi(\text{ext}_{a,b}(w')) \\
&= \varphi(w) \ .
\end{aligned}$$

- $w = uv$ for some $u, v \in \Sigma^\triangle \setminus \{\varepsilon\}$. Then

$$\varphi'(w) = \varphi'(u)\varphi'(v) \overset{\text{IH}}{=} \varphi(u)\varphi(v) = \varphi(w) \ .$$

Thus, $\varphi'^{-1}(F) = L$, which implies that (S, P) recognizes L.

For Point (2), assume that (R, O) is a quotient of (S, P). This means that there exists a surjective morphism $(\alpha, \beta)\colon (S, P) \to (R, O)$. Let us define an arbitrary function $\rho\colon \Sigma_{\text{int}} \to S$ such that $\rho(c) \in \alpha^{-1}(\varphi(c))$ for all $c \in \Sigma_{\text{int}}$ as well as an arbitrary function $\sigma\colon \{\text{ext}_{a,b} \mid a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}\} \to P$ such that $\sigma(\text{ext}_{a,b}) \in \beta^{-1}(\psi(\text{ext}_{a,b}))$ for all $a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}$. Now, take the unique morphism $(\varphi', \psi')\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (S, P)$ given by Proposition 3.7 for ρ and σ: we claim that it is such that $\alpha(\varphi'(w)) = \varphi(w)$ for all $w \in \Sigma^\triangle$. We can prove it by induction on w:

- $w = \varepsilon$. Then $\alpha(\varphi'(w)) = \alpha(1_S) = 1_R = \varphi(w)$.
- $w = c$ for some $c \in \Sigma_{\text{int}}$. Then $\alpha(\varphi'(w)) = \alpha(\rho(c)) = \varphi(c) = \varphi(w)$.
- $w = aw'b$ for some $a \in \Sigma_{\text{call}}$, $b \in \Sigma_{\text{ret}}$ and $w' \in \Sigma^\triangle$. Then

$$\begin{aligned}
\alpha(\varphi'(w)) &= \alpha\big(\varphi'(\text{ext}_{a,b}(w'))\big) \\
&= \alpha\big(\psi'(\text{ext}_{a,b})(\varphi'(w'))\big) \\
&= \beta(\psi'(\text{ext}_{a,b}))\big(\alpha(\varphi'(w'))\big) \\
&\overset{\text{IH}}{=} \beta(\sigma(\text{ext}_{a,b}))(\varphi(w')) \\
&= \psi(\text{ext}_{a,b})(\varphi(w')) \\
&= \varphi(\text{ext}_{a,b}(w')) \\
&= \varphi(w) \ .
\end{aligned}$$

- $w = uv$ for some $u, v \in \Sigma^\triangle \setminus \{\varepsilon\}$. Then

$$\alpha(\varphi'(w)) = \alpha(\varphi'(u))\alpha(\varphi'(v)) \overset{\text{IH}}{=} \varphi(u)\varphi(v) = \varphi(w) \ .$$

Therefore, $\varphi'^{-1}(\alpha^{-1}(F)) = L$, which implies that (S, P) recognizes L. □

Next, we show that any language recognized by an Ext-algebra is also recognized by one of its sub-Ext-algebras via a surjective morphism.

Lemma 3.15. *Let $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ be a morphism and let $L \subseteq \Sigma^\triangle$ be a language recognized by (R, O) via (φ, ψ). Then $\big(\varphi(\Sigma^\triangle), \psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}\big)$ is a sub-*Ext*-algebra of (R, O) recognizing L via the surjective morphism (φ, ψ') where $\psi'(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u,v})|_{\varphi(\Sigma^\triangle)}$ for all $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle)$.*

Proof. Since (R, O) recognizes L via (φ, ψ), this means that there exists $F \subseteq R$ such that $\varphi^{-1}(F) = L$. We have that $\varphi(\Sigma^\triangle)$ is a submonoid of R and $\psi(\mathcal{O}(\Sigma^\triangle))$ is a submonoid of O. Observe that for all $e \in \psi(\mathcal{O}(\Sigma^\triangle))$ and $r \in \varphi(\Sigma^\triangle)$, we have

$$e(r) = \psi(\mathrm{ext}_{u,v})(\varphi(w)) = \varphi(uwv) \in \varphi(\Sigma^\triangle)$$

because $\psi(\mathrm{ext}_{u,v}) = e$ for $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle)$ and $r = \varphi(w)$ for $w \in \Sigma^\triangle$. Moreover, for all $e, f \in \psi(\mathcal{O}(\Sigma^\triangle))$, it holds that $e|_{\varphi(\Sigma^\triangle)} \circ f|_{\varphi(\Sigma^\triangle)} = (e \circ f)|_{\varphi(\Sigma^\triangle)}$. Therefore, $\psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}$ is a submonoid of $\varphi(\Sigma^\triangle)^{\varphi(\Sigma^\triangle)}$. In addition, for each $r \in \varphi(\Sigma^\triangle)$, we have that $r = \varphi(w)$ for some $w \in \Sigma^\triangle$ and thus that

$$\mathrm{left}_r = \mathrm{left}_{\varphi(w)} = \psi(\mathrm{left}_w)|_{\varphi(\Sigma^\triangle)} = \psi(\mathrm{ext}_{w,\varepsilon})|_{\varphi(\Sigma^\triangle)}$$

as well as $\mathrm{right}_r = \psi(\mathrm{ext}_{\varepsilon,w})|_{\varphi(\Sigma^\triangle)}$. Thus, $\big(\varphi(\Sigma^\triangle), \psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}\big)$ is a sub-Ext-algebra of (R, O).

It is clear that φ is a surjective monoid morphism from $\Sigma^\triangle$ to $\varphi(\Sigma^\triangle)$. Further,

$$\begin{aligned}
\psi'(\mathrm{ext}_{u,v}) \circ \psi'(\mathrm{ext}_{u',v'}) &= \psi(\mathrm{ext}_{u,v})|_{\varphi(\Sigma^\triangle)} \circ \psi(\mathrm{ext}_{u',v'})|_{\varphi(\Sigma^\triangle)} \\
&= \big(\psi(\mathrm{ext}_{u,v}) \circ \psi(\mathrm{ext}_{u',v'})\big)\big|_{\varphi(\Sigma^\triangle)} \\
&= \psi(\mathrm{ext}_{u,v} \circ \mathrm{ext}_{u',v'})|_{\varphi(\Sigma^\triangle)} \\
&= \psi'(\mathrm{ext}_{u,v} \circ \mathrm{ext}_{u',v'})
\end{aligned}$$

for all $\mathrm{ext}_{u,v}, \mathrm{ext}_{u',v'} \in \mathcal{O}(\Sigma^\triangle)$, hence since $\psi'(\mathrm{ext}_{\varepsilon,\varepsilon}) = 1_O|_{\varphi(\Sigma^\triangle)}$, it follows that ψ' is a surjective monoid morphism from $\mathcal{O}(\Sigma^\triangle)$ to $\psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}$. Moreover, we have

- $\psi'(\mathrm{ext}_{u,v})(\varphi(w)) = \psi(\mathrm{ext}_{u,v})|_{\varphi(\Sigma^\triangle)}(\varphi(w)) = \psi(\mathrm{ext}_{u,v})(\varphi(w)) = \varphi(\mathrm{ext}_{u,v}(w))$ for all $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle)$ and $w \in \Sigma^\triangle$;
- $\psi'(\mathrm{left}_w) = \psi(\mathrm{ext}_{w,\varepsilon})|_{\varphi(\Sigma^\triangle)} = \mathrm{left}_{\varphi(w)}$ and $\psi'(\mathrm{right}_w) = \mathrm{right}_{\varphi(w)}$ for all $w \in \Sigma^\triangle$.

Therefore, (φ, ψ') is a surjective morphism recognizing L. □

The following lemma states that a language is recognized by an Ext-algebra via a surjective morphism if, and only if, the syntactic morphism of the language factors through the former morphism.

Lemma 3.16. *Let $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ be a surjective morphism, let $L \subseteq \Sigma^\triangle$ and let $(\varphi_L, \psi_L)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R_L, O_L)$ be the syntactic morphism of L. Then (R, O) recognizes L via (φ, ψ) if and only if there is a surjective morphism $(\alpha, \beta)\colon (R, O) \to (R_L, O_L)$ such that $\varphi_L = \alpha \circ \varphi$ (we say that (φ_L, ψ_L)* factors through *(φ, ψ)).*

Proof. Assume first that there is a surjective morphism $(\alpha, \beta)\colon (R, O) \to (R_L, O_L)$ such that $\varphi_L = \alpha \circ \varphi$. Then

$$\varphi^{-1}\big(\alpha^{-1}(\varphi(L))\big) = (\alpha \circ \varphi)^{-1}\big(\varphi_L(L)\big) = \varphi_L^{-1}(\varphi_L(L)) = L \ ,$$

hence (R, O) recognizes L via (φ, ψ).

Assume now that (R, O) recognizes L via (φ, ψ). This means that there exists $F \subseteq R$ satisfying $\varphi^{-1}(F) = L$. Take $w, w' \in \Sigma^\triangle$ such that $\varphi(w) = \varphi(w')$. Then, given any $e \in \mathcal{O}(\Sigma^\triangle)$, we have that

$$\varphi(e(w)) = \psi(e)(\varphi(w)) = \psi(e)(\varphi(w')) = \varphi(e(w')) \ .$$

Therefore, since $\varphi^{-1}(F) = L$, it holds that $w \sim_L w'$, that is, $\varphi_L(w) = \varphi_L(w')$.

Take $\mathrm{ext}_{u,v}, \mathrm{ext}_{u',v'} \in \mathcal{O}(\Sigma^\triangle)$ such that $\psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u',v'})$. Then, for each $w \in \Sigma^\triangle$, we have that

$$\varphi(\mathrm{ext}_{u,v}(w)) = \psi(\mathrm{ext}_{u,v})(\varphi(w)) = \psi(\mathrm{ext}_{u',v'})(\varphi(w)) = \varphi(\mathrm{ext}_{u',v'}(w)) \ ,$$

that is, $\mathrm{ext}_{u,v}(w) \sim_L \mathrm{ext}_{u',v'}(w)$. Hence, $\psi_L(\mathrm{ext}_{u,v}) = \psi_L(\mathrm{ext}_{u',v'})$.

We can now define the functions $\alpha\colon R \to R_L$ and $\beta\colon O \to O_L$ such that $\alpha(\varphi(w)) = \varphi_L(w)$ for all $w \in \Sigma^\triangle$ and $\beta(\psi(\mathrm{ext}_{u,v})) = \psi_L(\mathrm{ext}_{u,v})$ for all $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle)$: those are well-defined by surjectivity of (φ, ψ) and what we have proven just above. Since (φ_L, ψ_L) is a surjective morphism from $(\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle))$ to (R_L, O_L), we can easily prove that (α, β) is a surjective morphism from (R, O) to (R_L, O_L) that does of course satisfy $\varphi_L = \alpha \circ \varphi$. □

The following proposition shows that the syntactic Ext-algebra of a given language of well-matched words is the least Ext-algebra recognizing this language.

Proposition 3.17. *An* Ext*-algebra (R, O) recognizes a language $L \subseteq \Sigma^\triangle$ if, and only if, its syntactic* Ext*-algebra (R_L, O_L) divides (R, O).*

Proof. Let (R, O) be an Ext-algebra and let $L \subseteq \Sigma^\triangle$ be a language. Consider also its syntactic Ext-algebra (R_L, O_L) and its syntactic morphism (φ_L, ψ_L).

Implication from right to left. Assume that the syntactic Ext-algebra (R_L, O_L) of L divides (R, O). We have that (R_L, O_L) recognizes L and we then use Lemma 3.14 to conclude that (R, O) does also recognize L.

Implication from left to right. Assume that (R, O) recognizes L through a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$. By Lemma 3.15, we have that $\big(\varphi(\Sigma^\triangle), \psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}\big) = (R', O')$ is a sub-Ext-algebra of (R, O) recognizing L via the surjective morphism (φ, ψ') where $\psi'(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u,v})|_{\varphi(\Sigma^\triangle)}$ for all $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle)$. Then, by Lemma 3.16, there exists a surjective morphism $(\alpha, \beta)\colon (R', O') \to (R_L, O_L)$ such that $\varphi_L = \alpha \circ \varphi$. Thus, we have that (R_L, O_L) divides (R, O). □

We say that an Ext-algebra (R, O) is *finite* whenever R is finite (which is the case if and only if O is finite). The following theorem establishes the equivalence between visibly pushdown languages and languages recognizable by finite Ext-algebras. Its proof provides effective translations from DVPAs to Ext-algebras and vice versa.

Theorem 3.18. *A language $L \subseteq \Sigma^\triangle$ is a VPL if, and only if, it is recognized by a finite* Ext*-algebra.*

Proof. Let $L \subseteq \Sigma^\triangle$ be a language. Before we prove the theorem we have the following claim, which can be easily proven by induction on $|u|$ and structural induction on w, respectively.

Claim. Let $A = (Q, \Sigma, \Gamma, \delta, q_0, F, \bot)$ be a DVPA. We denote by π_Q the projection of $Q \times \Gamma^*$ on Q and by π_{Γ^*} the projection of $Q \times \Gamma^*$ on Γ^*. It holds that $L(A) \subseteq \Sigma^\triangle$ and additionally we have that

$$\widehat{\delta}(q, uv, \sigma) = \widehat{\delta}\big(\pi_Q(\widehat{\delta}(q, u, \sigma)), v, \pi_{\Gamma^*}(\widehat{\delta}(q, u, \sigma))\big)$$

and

$$\widehat{\delta}(q, w, \alpha\sigma) = \big(\pi_Q(\widehat{\delta}(q, w, \alpha)), \alpha\sigma\big)$$

for all $q \in Q$, $u, v \in \Sigma^*$, $\sigma \in \Gamma^*$, $w \in \Sigma^\triangle$ and $\alpha \in \Gamma$.

Implication from left to right. Assume that L is a VPL. This means there exists a DVPA $A = (Q, \Sigma, \Gamma, \delta, q_0, F, \bot)$ such that $L(A) = L$. Consider the operation $*$ on $R = Q^{Q\times\Gamma}$ defined so that for all $f, g \in R$, we have $f * g(q, \alpha) = g(f(q, \alpha), \alpha)$ for all $q \in Q$ and $\alpha \in \Gamma$. Observe that for all $f, g, h \in Q^{Q\times\Gamma}$, we have

$$(f * g) * h(q, \alpha) = h(f * g(q, \alpha), \alpha) = h(g(f(q, \alpha), \alpha), \alpha) = g * h(f(q, \alpha), \alpha) = f * (g * h)(q, \alpha)$$

for all $q \in Q$ and $\alpha \in \Gamma$. Thus $*$ is associative and since it also has $i \in R$ with $i(q, \alpha) = q$ for all $q \in Q$ and $\alpha \in \Gamma$ as an identity, we have that $R = Q^{Q\times\Gamma}$ with operation $*$ forms a monoid. Take O to be the monoid R^R (for composition). Since O clearly contains the functions left_r and right_r for all $r \in R$, it follows that (R, O) is a finite Ext-algebra. We now prove that it recognizes L. For each $w \in \Sigma^\triangle$, define $f_w \in R$ by $f_w(q, \alpha) = \pi_Q(\widehat{\delta}(q, w, \alpha))$ for all $q \in Q$ and $\alpha \in \Gamma$. Let us consider the unique morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$, given by Proposition 3.7, such that for each $c \in \Sigma_{\mathrm{int}}$, we have $\varphi(c) = f_c$ and for each $a \in \Sigma_{\mathrm{call}}, b \in \Sigma_{\mathrm{ret}}$, we have that $\psi(\mathrm{ext}_{a,b})$ sends any $f \in R$ to $g \in R$ satisfying that $g(q, \alpha) = \pi_Q\big(\delta(f(p, \beta), b, \beta)\big)$ with $\delta(q, a, \alpha) = (p, \beta\alpha)$ for all $q \in Q$ and $\alpha \in \Gamma$. We claim that for all $w \in \Sigma^\triangle$, we have that $\varphi(w) = f_w$. We prove it by induction on w.

- $w = \varepsilon$. Then $\varphi(w) = i = f_w$.
- $w = c$ for some $c \in \Sigma_{\mathrm{int}}$. Then $\varphi(w) = f_c = f_w$.
- $w = aw'b$ for some $a \in \Sigma_{\mathrm{call}}$, $b \in \Sigma_{\mathrm{ret}}$ and $w' \in \Sigma^\triangle$. Then

$$\varphi(w) = \varphi(\mathrm{ext}_{a,b}(w')) = \psi(\mathrm{ext}_{a,b})(\varphi(w')) \overset{\mathrm{IH}}{=} \psi(\mathrm{ext}_{a,b})(f_{w'}) \ .$$

  So $\varphi(w) = g$ such that for all $q \in Q$ and $\alpha \in \Gamma$, if we set $\delta(q, a, \alpha) = (p, \beta\alpha)$, we have, recalling that $\widehat{\delta}$ extends δ,

$$\begin{aligned}
g(q, \alpha) &= \pi_Q\big(\delta(f_{w'}(p, \beta), b, \beta)\big) \\
&= \pi_Q\Big(\delta\big(\pi_Q(\widehat{\delta}(p, w', \beta)), b, \beta\big)\Big) \\
&= \pi_Q\Big(\widehat{\delta}\big(\pi_Q(\widehat{\delta}(p, w', \beta)), b, \beta\big)\Big) \\
&= \pi_Q\Big(\widehat{\delta}\big(\pi_Q(\widehat{\delta}(q, aw', \alpha)), b, \pi_{\Gamma^*}(\widehat{\delta}(q, aw', \alpha))\big)\Big) \\
&= \pi_Q(\widehat{\delta}(q, aw'b, \alpha)) \\
&= f_{aw'b}(q, \alpha) \\
&= f_w(q, \alpha) \ .
\end{aligned}$$

  Thus $\varphi(w) = f_w$.

- $w = uv$ for some $u, v \in \Sigma^\triangle \setminus \{\varepsilon\}$. Then $\varphi(w) = \varphi(u) * \varphi(v) \overset{\text{IH}}{=} f_u * f_v$. But

$$\begin{aligned} f_u * f_v(q,\alpha) &= f_v(f_u(q,\alpha),\alpha) \\ &= f_v\big(\pi_Q(\widehat{\delta}(q,u,\alpha)),\alpha\big) \\ &= \pi_Q\Big(\widehat{\delta}\big(\pi_Q(\widehat{\delta}(q,u,\alpha)),v,\alpha\big)\Big) \\ &= \pi_Q\Big(\widehat{\delta}\big(\pi_Q(\widehat{\delta}(q,u,\alpha)),v,\pi_{\Gamma^*}(\widehat{\delta}(q,u,\alpha))\big)\Big) \\ &= \pi_Q(\widehat{\delta}(q,uv,\alpha)) \\ &= f_{uv}(q,\alpha) \end{aligned}$$

  for all $q \in Q$ and $\alpha \in \Gamma$. Therefore $\varphi(w) = f_w$.

Finally, set $P = \{f \in R \mid f(q_0,\bot) \in F\}$. It holds that

$$\begin{aligned} \varphi^{-1}(P) &= \{w \in \Sigma^\triangle \mid f_w(q_0,\bot) \in F\} \\ &= \{w \in \Sigma^\triangle \mid \pi_Q(\widehat{\delta}(q_0,w,\bot)) \in F\} \\ &= \{w \in \Sigma^\triangle \mid \widehat{\delta}(q_0,w,\bot) \in F \times \{\bot\}\} \\ &= L(A) \\ &= L\ . \end{aligned}$$

Therefore, (R,O) recognizes L.

Implication from right to left. Assume there exists a finite Ext-algebra (R,O) that recognizes L. This means that there exists a morphism $(\varphi,\psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R,O)$ such that $L = \varphi^{-1}(F)$ for some $F \subseteq R$. Let us now define the DVPA

$$A = (Q, \Sigma, \Gamma, \delta, 1, F, \bot),$$

where $Q = R$, $1 = 1_R$, $\Gamma = R \times \Sigma_{\text{call}} \cup \{\bot\}$, and

$$\delta(r,a,\alpha) = \begin{cases} (1,(r,a)\alpha) & \text{if } a \in \Sigma_{\text{call}} \\ (s\psi(\text{ext}_{b,a})(r),\varepsilon) & \text{if } a \in \Sigma_{\text{ret}} \text{ and } \alpha = (s,b) \in R \times \Sigma_{\text{call}} \\ (r,\varepsilon) & \text{if } a \in \Sigma_{\text{ret}} \text{ and } \alpha = \bot \\ (r\varphi(c),\alpha) & \text{if } a \in \Sigma_{\text{int}} \end{cases}$$

for all $r \in R$, $a \in \Sigma$ and $\alpha \in \Gamma$. We prove that $\widehat{\delta}(r,w,\sigma) = (r\varphi(w),\sigma)$ for all $r \in R$, $w \in \Sigma^\triangle$ and $\sigma \in \Gamma^*\bot$ by induction on w.

- $w = \varepsilon$. Then $\widehat{\delta}(r,w,\sigma) = (r,\sigma) = (r\varphi(w),\sigma)$.
- $w = c$ for some $c \in \Sigma_{\text{int}}$. Then $\widehat{\delta}(r,w,\sigma) = (r\varphi(c),\sigma) = (r\varphi(w),\sigma)$.

- $w = aw'b$ for some $a \in \Sigma_{\mathrm{call}}$, $b \in \Sigma_{\mathrm{ret}}$ and $w' \in \Sigma^{\triangle}$. Then

$$\begin{aligned}\widehat{\delta}(r, w, \sigma) &= \widehat{\delta}(1, w'b, (r, a)\sigma) \\ &= \widehat{\delta}\big(\pi_Q(1, w', (r, a)\sigma), b, \pi_{\Gamma^*}(1, w', (r, a)\sigma)\big) \\ &\overset{\mathrm{IH}}{=} \widehat{\delta}(\varphi(w'), b, (r, a)\sigma) \\ &= \big(r\psi(\mathrm{ext}_{a,b})(\varphi(w')), \sigma\big) \\ &= (r\varphi(w), \sigma)\ .\end{aligned}$$

- $w = uv$ for some $u, v \in \Sigma^{\triangle} \setminus \{\varepsilon\}$. Then

$$\begin{aligned}\widehat{\delta}(r, w, \sigma) &= \widehat{\delta}\big(\pi_Q(\widehat{\delta}(r, u, \sigma)), v, \pi_{\Gamma^*}(\widehat{\delta}(r, u, \sigma))\big) \\ &\overset{\mathrm{IH}}{=} \widehat{\delta}(r\varphi(u), v, \sigma) \\ &\overset{\mathrm{IH}}{=} (r\varphi(u)\varphi(v), \sigma) \\ &= (r\varphi(w), \sigma)\ .\end{aligned}$$

Hence,

$$\begin{aligned}L(A) &= \{w \in \Sigma^{\triangle} \mid \widehat{\delta}(1, w, \bot) \in F \times \{\bot\}\} \\ &= \{w \in \Sigma^{\triangle} \mid \pi_Q(\widehat{\delta}(1, w, \bot)) \in F\} \\ &= \{w \in \Sigma^{\triangle} \mid \varphi(w) \in F\} \\ &= \varphi^{-1}(F) = L\ .\end{aligned}$$

Therefore, L is a VPL. □

4 (Weak) length-synchronicity, nesting depth, and quasi-aperiodicity

For the rest of this section let us fix a visibly pushdown alphabet Σ, a finite Ext-algebra (R, O) and consider a morphism $(\varphi, \psi)\colon (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R, O)$. Suitably adjusting the pumping lemma for context-free language we introduce a pumping lemma for Ext-algebra morphisms in Section 4.1. In Section 4.2 we extend the notions of weak length-synchronicity and length-synchronicity to Ext-algebras morphisms and to visibly pushdown languages. It is shown that for languages generated by vertically visibly pushdown grammars, (weak) length-synchronicity of the relation of the generating grammar coincides with (weak) length-synchronicity of language. We concern ourselves with the nesting depth of visibly pushdown languages in Section 4.3. Finally in Section 4.4 we introduce quasi-aperiodicity of Ext-algebra morphisms and prove that a VPL is quasi-counterfree if, and only if, its syntactic morphism is quasi-aperiodic.

4.1 A pumping lemma for Ext-algebra morphisms

The following is an adaption of the pumping lemma for context-free languages to Ext-algebra morphisms. It states that if $uv \in \Sigma^{\triangle}$ and u (resp. v) contains a well-matched factor that is sufficiently long, we can pump certain infixes of u (resp. v): thus, one can find longer and longer words $u_1, u_2, \ldots$ (resp. $v_1, v_2, \ldots$) such that $u_1 v, u_2 v, \ldots \in \Sigma^{\triangle}$ (resp. $uv_1, uv_2, \ldots \in \Sigma^{\triangle}$) and the morphism ψ sends $\mathrm{ext}_{u,v}$ to the same element in O as $\mathrm{ext}_{u_i,v}$ (resp. as ext_{u,v_i}).

Lemma 4.1 (Pumping Lemma). *There exists $n \in \mathbb{N}_{>0}$ such that for all $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle})$ we have:*

- *If there exists a factor $w \in \Sigma^{\triangle}$ of u satisfying $|w| \geq n$, then $u = sxzyt$ with $s, x, z, y, t \in \Sigma^*$ such that $|xy| \geq 1$, $|xzy| \leq n$ and for all $i \in \mathbb{N}$, $sx^izy^itv \in \Sigma^{\triangle}$ and $\psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{sx^izy^it,v})$.*
- *If there exists a factor $w \in \Sigma^{\triangle}$ of v satisfying $|w| \geq n$, then $v = sxzyt$ with $s, x, z, y, t \in \Sigma^*$ such that $|xy| \geq 1$, $|xzy| \leq n$ and for all $i \in \mathbb{N}$, $usx^izy^it \in \Sigma^{\triangle}$ and $\psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u,sx^izy^it})$.*

Proof. For each $r \in R$, let $n_r \in \mathbb{N}_{>0}$ be the pumping constant for the context-free language $\varphi^{-1}(r)$: it is a VPL and hence a context-free language by Theorem 3.18. We set $n = \max_{r \in R} n_r$. Let $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle})$ be such that there exists a factor $w \in \Sigma^{\triangle}$ of u satisfying $|w| \geq n$. Let

$$\mathrm{ext}_{u,v} \quad = \quad \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \quad \cdots \quad \circ \mathrm{ext}_{x_{h-1},y_{h-1}} \circ \mathrm{ext}_{a_{h-1},b_{h-1}} \circ \mathrm{ext}_{x_h,y_h}$$

be the stair factorization of $\mathrm{ext}_{u,v}$ provided by Lemma 3.6. Since no factor of u spanning one of the a_j's in the factorization can be well-matched, there must exist some $j \in [1, h]$ satisfying $|x_j| \geq n$, so that if we set $u' = x_1a_1 \cdots x_{j-1}a_{j-1}$, $v' = b_{j-1}y_{j-1} \cdots b_1y_1$, $u'' = a_jx_{j+1} \cdots a_{h-1}x_h$ and $v'' = y_hb_{h-1} \cdots y_{j+1}b_jy_j$, we have $u'v', u''v'' \in \Sigma^{\triangle}$ and $\mathrm{ext}_{u,v} = \mathrm{ext}_{u',v'} \circ \mathrm{ext}_{x_j,\varepsilon} \circ \mathrm{ext}_{u'',v''}$. By the pumping lemma for context-free languages we have $x_j = x'xzyy'$ with $x', x, z, y, y' \in \Sigma^*$ such that $|xy| \geq 1$, $|xzy| \leq n$ and for all $i \in \mathbb{N}$, $x'x^izy^iy' \in \Sigma^{\triangle}$ and $\varphi(x_j) = \varphi(x'x^izy^iy')$. This implies that if we set $s = u'x'$ and $t = y'u''$, then for all $i \in \mathbb{N}$, we have $sx^izy^itv = \mathrm{ext}_{u',v'} \circ \mathrm{ext}_{x'x^izy^iy',\varepsilon} \circ \mathrm{ext}_{u'',v''}(\varepsilon) \in \Sigma^{\triangle}$ and

$$\begin{aligned}
\psi(\mathrm{ext}_{u,v}) &= \psi(\mathrm{ext}_{u',v'}) \circ \mathrm{left}_{\varphi(x_j)} \circ \psi(\mathrm{ext}_{u'',v''}) \\
&= \psi(\mathrm{ext}_{u',v'}) \circ \mathrm{left}_{\varphi(x'x^izy^iy')} \circ \psi(\mathrm{ext}_{u'',v''}) \\
&= \psi(\mathrm{ext}_{u',v'}) \circ \psi(\mathrm{ext}_{x'x^izy^iy',\varepsilon}) \circ \psi(\mathrm{ext}_{u'',v''}) \\
&= \psi(\mathrm{ext}_{sx^izy^it,v}) \ .
\end{aligned}$$

We handle the case where for $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle})$ there exists a factor $w \in \Sigma^{\triangle}$ of v such that $|w| \geq n$ symmetrically. □

4.2 Weak length-synchronicity and length-synchronicity

In this section we introduce the notions of weak length-synchronicity and length-synchronicity for Ext-algebra morphisms and visibly pushdown languages. Before we do that, let us give some motivation how TC^0-hardness can be proven if the syntactic morphism maps certain $\mathrm{ext}_{u,v}, \mathrm{ext}_{u',v}$ with $|u| \neq |u'|$ to particular idempotents. For these we require the following notion of reachability.

For $F \subseteq R$ we say that an element $r \in R$ is *F-reachable* if $e(r) \in F$ for some $e \in O$. We say $e \in O$ is *F-reachable* if $e(r)$ is F-reachable for some $r \in R$. Although we will mainly study F-reachable elements over finite Ext-algebras we remark that the notion of F-reachability is defined over any Ext-algebra, in particular over $(\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle}))$. Fix any VPL L, its syntactic Ext-algebra (R_L, O_L) along with its syntactic morphism (φ_L, ψ_L). Assume some idempotent $e \in O_L$ that is $\varphi(L)$-reachable.

We claim that if $\psi_L(\mathrm{ext}_{u,v}) = \psi_L(\mathrm{ext}_{u',v}) = e$ and $\Delta(u), \Delta(u') > 0$ for some $\mathrm{ext}_{u,v}, \mathrm{ext}_{u',v} \in \mathcal{O}(\Sigma^{\triangle})$ with $|u| \neq |u'|$, then L is TC^0-hard. We remark that we must have $\Delta(u) = \Delta(u')$. Exploiting the fact that $|u| \neq |u'|$ we give a constant-depth reduction from the TC^0-complete language EQUALITY to L.

Since $\psi_L(\mathrm{ext}_{u,v})$ is $\varphi_L(L)$-reachable, we can fix some $x, y, z \in \Sigma^*$ such that $xuyvz \in L$. Given a word $w \in \{0,1\}^*$ of length $2n$ (binary words of odd length can directly be rejected) we map it to

$xh(w)zv^{n\cdot(|u|+|u'|)}y$, where $h\colon \{0,1\}^* \to \Sigma^*$ is the length-multiplying morphism satisfiying $h(0) = u^{|u'|}$ and $h(1) = u'^{|u|}$: one can prove that $w \in \text{EQUALITY}$ if, and only if, $h(w)v^{n\cdot(|u|+|u'|)} \in \Sigma^{\triangle}$ if, and only if, $xh(w)zv^{n\cdot(|u|+|u'|)}y \in L$.

Dually, one can show that L is TC^0-hard in case $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u,v'}) = e$ and $\Delta(u) > 0$ for some $\text{ext}_{u,v}, \text{ext}_{u,v'} \in \mathcal{O}(\Sigma^{\triangle})$ with $|v| \neq |v'|$.

The following definition of weak length-synchronicity captures the situation when such idempotents do not exist — it adapts the notion of weak length-synchronicity of sets of contexts, given in Definition 2.3, to morphisms and VPLs, respectively. Recall that $\mathcal{R} \subseteq \mathsf{Con}(\Sigma)$ is defined to be weakly length-synchronous if $u = u'$ implies $|v| = |v'|$ and $v = v'$ implies $|u| = |u'|$ for all $(u,v),(u',v') \in \mathcal{R}$ satisfying $\Delta(u), \Delta(u') > 0$.

Definition 4.2. *For all $e \in O$ define the sets of contexts $\mathcal{R}_e$ and $\mathcal{U}_e$ as follows:*

$$\mathcal{R}_e = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi(\text{ext}_{u,v}) = e\} \quad \text{and} \quad \mathcal{U}_e = \{(u,v) \in \mathsf{Con}(\Sigma) \mid e \circ \psi(\text{ext}_{u,v}) = e\}$$

Definition 4.3 (Weak Length-Synchronicity)**.** *The morphism $(\varphi,\psi)\colon (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R,O)$ is F-weakly-length-synchronous (where $F \subseteq R$) if for all F-reachable idempotents $e \in O$ the set of contexts $\mathcal{R}_e$ is weakly length-synchronous. We call a VPL $L \subseteq \Sigma^{\triangle}$* weakly length-synchronous *if its syntactic morphism (φ_L, ψ_L) is $\varphi_L(L)$-weakly-length-synchronous.*

In fact, F-weak-length-synchronicity actually implies weak length-synchronicity of the set of contexts associated to any subsemigroup of F-reachable elements.

Lemma 4.4. *For all $F \subseteq R$ and subsemigroup P of O, if all elements in P are F-reachable and (φ,ψ) is F-weakly-length-synchronous, then $\bigcup_{e\in P} \mathcal{R}_e$ is weakly length-synchronous.*

Proof. Let $F \subseteq R$ and P be a subsemigroup of O. Assume all elements in P are F-reachable and (φ,ψ) is F-weakly-length-synchronous.

Let $(u,v),(u',v) \in \mathsf{Con}(\Sigma)$ be such that $\Delta(u), \Delta(u') > 0$ and $\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v}) \in P$. Set $e = \psi(\text{ext}_{u,v})$ and $f = \psi(\text{ext}_{u',v})$. By hypothesis, given $\omega \in \mathbb{N}_{>0}$ the idempotent power of O, we have $(e^\omega f^\omega)^\omega \in P$, hence $(e^\omega f^\omega)^\omega$ is an F-reachable idempotent and thus $\mathcal{R}_{(e^\omega f^\omega)^\omega}$ is weakly length-synchronous. But

$$\psi(\text{ext}_{(u^{2\cdot\omega}u'^\omega)^\omega, v^{3\cdot\omega^2}}) = (e^\omega f^\omega)^\omega = \psi(\text{ext}_{(u^\omega u'^{2\cdot\omega})^\omega, v^{3\cdot\omega^2}})$$

so since $\Delta((u^{2\cdot\omega}u'^\omega)^\omega) = \Delta((u^\omega u'^{2\cdot\omega})^\omega) > 0$, we obtain

$$\begin{aligned} & \omega \cdot (2\cdot\omega\cdot |u| + \omega\cdot|u'|) = \omega\cdot(\omega\cdot|u| + 2\cdot\omega\cdot|u'|) \\ \iff \quad & 2\cdot|u| + |u'| = |u| + 2\cdot|u'| \\ \iff \quad & |u| = |u'| \ . \end{aligned}$$

In the same way, one can prove that for all $(u,v),(u,v') \in \mathsf{Con}(\Sigma)$ such that $\Delta(u) > 0$ and $\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u,v'}) \in P$, we have $|v| = |v'|$.

Therefore, $\bigcup_{e\in P} \mathcal{R}_e$ is weakly length-synchronous. □

Instead of considering those pairs (u,v) such that $\text{ext}_{u,v}$ is being mapped to an F-reachable idempotent, the following characterization of weak length-synchronicity consider pairs (u,v) such that $\text{ext}_{u,v}$ is being mapped to an element that behaves neutrally with respect to right multiplication with an F-reachable element that is not necessarily idempotent.

Proposition 4.5. *For all $F \subseteq R$ we have that (φ, ψ) is F-weakly-length-synchronous if, and only if, for all F-reachable $e \in O$ the set of contexts $\mathcal{U}_e$ is weakly length-synchrononous.*

Proof. Let $F \subseteq R$.

If $\mathcal{U}_e = \{(u, v) \in \mathsf{Con}(\Sigma) \mid e \circ \psi(\mathrm{ext}_{u,v}) = e\}$ is weakly length-synchronous for all F-reachable $e \in O$, then in particular the set of contexts $\mathcal{R}_e = \{(u, v) \in \mathsf{Con}(\Sigma) \mid \psi(\mathrm{ext}_{u,v}) = e\}$ is weakly length-synchronous for all F-reachable idempotents $e \in O$.

Conversely, assume that (φ, ψ) is F-weakly-length-synchronous. Fix any F-reachable $e \in O$. We need to prove that $\mathcal{U}_e$ is weakly length-synchronous. It is clear that $P = \{f \in O \mid e \circ f = e\}$ forms a subsemigroup of O whose elements are all F-reachable, by F-reachability of e. Therefore, by Lemma 4.4, $\bigcup_{f \in P} \mathcal{R}_f = \mathcal{U}_e$ is weakly length-synchronous. □

Using Lemma 4.1 and Proposition 4.5, the following proposition follows immediately.

Proposition 4.6. *Let n be the pumping constant from Lemma 4.1, let $F \subseteq R$, let $e \in O$ be F-reachable, and let $\mathrm{ext}_{u,v}$ be such that $\Delta(u) > 0$ and $e \circ \psi(\mathrm{ext}_{u,v}) = e$. If (φ, ψ) is F-weakly-length-synchronous, then the stair factorization*

$$\mathrm{ext}_{u,v} = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_{h-1},y_{h-1}} \circ \mathrm{ext}_{a_{h-1},b_{h-1}} \circ \mathrm{ext}_{x_h,y_h}$$

satisfies $|x_i|, |y_i| \leq n$ for all $i \in [1, h]$,

As above, the following definition adapts the notion of length-synchronicity of sets of contexts, given in Definition 2.3, to Ext-algebra morphisms and VPLs, respectively.

Definition 4.7 (Length-Synchronicity)**.** *The morphism $(\varphi, \psi) \colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ is F-length-synchronous (where $F \subseteq R$) if for all F-reachable idempotents $e \in O$ the set of contexts $\mathcal{R}_e$ is length-synchronous. We call a VPL $L \subseteq \Sigma^\triangle$* length-synchronous *if its syntactic morphism (φ_L, ψ_L) is $\varphi_L(L)$-length-synchronous.*

Example 4.8. Consider our running example $\mathcal{L}_{1,2} = L(S \to aSb_1 \mid acSb_2 \mid \varepsilon)$. Recall that the monoid $O_{\mathcal{L}_{1,2}}$ of the syntactic Ext-algebra $(R_{\mathcal{L}_{1,2}}, O_{\mathcal{L}_{1,2}})$ and syntactic morphism $(\varphi_{\mathcal{L}_{1,2}}, \psi_{\mathcal{L}_{1,2}})$ of $\mathcal{L}_{1,2}$, given in Example 3.10, has the idempotents $(\varepsilon, \varepsilon)$, (acb_1, ε) and (a, b_1). Also recall that $\varphi_{\mathcal{L}_{1,2}}(\mathcal{L}_{1,2}) = \{\varepsilon, ab_1\}$. Since $\psi^{-1}_{\mathcal{L}_{1,2}}((\varepsilon, \varepsilon)) = \{\mathrm{ext}_{\varepsilon,\varepsilon}\}$ and (acb_1, ε) is a zero we have that $O_{\mathcal{L}_{1,2}}$'s only idempotent that is $\{\varepsilon, ab_1\}$-reachable and whose pre-image under $\psi_{\mathcal{L}_{1,2}}$ contains at least one $\mathrm{ext}_{u,v}$ with $\Delta(u) > 0$ is the idempotent (a, b_1). However, both ext_{a,b_1} and ext_{ac,b_2}, where $\Delta(a) = \Delta(ac) = 1 > 0$, are sent to the idempotent $(a, b_1) = (a, b_2) \circ (c, \varepsilon)$. Since $|a|/|b_1| = 1 \neq 2 = |ac|/|b_2|$, we have that $\mathcal{L}_{1,2}$ is not length-synchronous. On the other hand, note that if any $\mathrm{ext}_{u,v}$ and $\mathrm{ext}_{u',v}$ (resp. $\mathrm{ext}_{u,v}$ and $\mathrm{ext}_{u,v'}$) are sent to (a, b_1) then $u = u'$ and thus $|u| = |u'|$ (resp. $v = v'$ and thus $|v| = |v'|$). Hence, $\mathcal{L}_{1,2}$ is weakly length-synchronous.

As above, F-length-synchronicity actually implies length-synchronicity of the set of contexts associated to any subsemigroup of F-reachable elements.

Lemma 4.9. *For all $F \subseteq R$ and subsemigroup P of O, if all elements in P are F-reachable and (φ, ψ) is F-length-synchronous, then $\bigcup_{e \in P} \mathcal{R}_e$ is length-synchronous.*

Proof. Let $F \subseteq R$ and P be a subsemigroup of O. Assume all elements in P are F-reachable and (φ, ψ) is F-length-synchronous.

Let $(u, v), (u', v') \in \mathsf{Con}(\Sigma)$ be such that $\Delta(u), \Delta(u') > 0$ and $\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v'}) \in P$. Set $e = \psi(\mathrm{ext}_{u,v})$ and $f = \psi(\mathrm{ext}_{u',v'})$. By hypothesis, given $\omega \in \mathbb{N}_{>0}$ the idempotent power of O,

we have $(e^\omega f^\omega)^\omega \in P$, hence $(e^\omega f^\omega)^\omega$ is an F-reachable idempotent and thus $\mathcal{R}_{(e^\omega f^\omega)^\omega}$ is length-synchronous. But

$$\psi(\mathrm{ext}_{(u^{2\cdot\omega}u'^\omega)^\omega,(v'^\omega v^{2\cdot\omega})^\omega}) = (e^\omega f^\omega)^\omega = \psi(\mathrm{ext}_{(u^\omega u'^{2\cdot\omega})^\omega,(v'^{2\cdot\omega}v^\omega)^\omega})$$

so since $\Delta((u^{2\cdot\omega}u'^\omega)^\omega), \Delta((u^\omega u'^{2\cdot\omega})^\omega) > 0$, setting $(x,y) = ((u^{2\cdot\omega}u'^\omega)^\omega, (v'^\omega v^{2\cdot\omega})^\omega)$ and $(x',y') = ((u^\omega u'^{2\cdot\omega})^\omega, (v'^{2\cdot\omega}v^\omega)^\omega)$, we obtain (using that for $a,b,c,d>0$ we have that $\frac{a}{b} = \frac{c}{d}$ implies $\frac{a}{b} = \frac{c}{d} = \frac{a+c}{b+d}$ and, if additionally $a > c$, it implies $\frac{a}{b} = \frac{c}{d} = \frac{a-c}{b-d}$)

$$\begin{aligned}
\frac{|x|}{|y|} = \frac{|x'|}{|y'|} \quad &\Longrightarrow \quad \frac{|x|}{|y|} = \frac{|x'|}{|y'|} = \frac{|x|+|x'|}{|y|+|y'|} = \frac{\omega^2\cdot(|u|+|u'|)}{\omega^2\cdot(|v|+|v'|)} \\
&\Longrightarrow \quad \frac{|x|-\omega^2\cdot(|u|+|u'|)}{|y|-\omega^2\cdot(|v|+|v'|)} = \frac{|x'|-\omega^2\cdot(|u|+|u'|)}{|y'|-\omega^2\cdot(|v|+|v'|)} \qquad (1) \\
&\Longrightarrow \quad \frac{|u|}{|v|} = \frac{\omega^2\cdot|u|}{\omega^2\cdot|v|} \overset{(1)}{=} \frac{\omega^2\cdot|u'|}{\omega^2\cdot|v'|} = \frac{|u'|}{|v'|}\ .
\end{aligned}$$

Therefore, $\bigcup_{e\in P}\mathcal{R}_e$ is length-synchronous. □

The two following propositions characterize length-synchronicity of Ext-algebra morphisms, which will be of particular importance when approximating the matching relation of a length-synchronous VPL in terms of $\mathsf{FO}[+]$. This will be an important ingredient to proving that VPLs that both are length-synchronous and have a quasi-aperiodic syntactic morphism (a notion to be defined in Subsection 4.4) are in $\mathsf{FO}[+]$ and thus in AC^0.

Proposition 4.10. *For all $F \subseteq R$, we have that (φ,ψ) is F-length-synchronous if, and only if, for all F-reachable $e \in O$ the set of contexts $\mathcal{U}_e$ is length-synchronous.*

Proof. Let $F \subseteq R$.

If $\mathcal{U}_e = \{(u,v) \in \mathsf{Con}(\Sigma) \mid e \circ \psi(\mathrm{ext}_{u,v}) = e\}$ is length-synchronous for all F-reachable $e \in O$, then in particular the set of contexts $\mathcal{R}_e = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi(\mathrm{ext}_{u,v}) = e\}$ is length-synchronous for all F-reachable idempotents $e \in O$.

Conversely, assume that (φ,ψ) is F-length-synchronous. Fix any F-reachable $e \in O$. We need to prove that $\mathcal{U}_e$ is length-synchronous. It is clear that $P = \{f \in O \mid e \circ f = e\}$ forms a subsemigroup of O whose elements are all F-reachable, by F-reachability of e. Therefore, by Lemma 4.9, $\bigcup_{f\in P}\mathcal{R}_f = \mathcal{U}_e$ is length-synchronous. □

Proposition 4.11. *Let $F \subseteq R$ and assume (φ,ψ) is F-weakly-length-synchronous. Then for all F-reachable $e \in O$ the following two statements are equivalent.*

1. *The set of contexts $\mathcal{U}_e$ is length-synchronous.*
2. *There exist $\alpha \in \mathbb{Q}_{>0}$, $\beta \in \mathbb{N}$, $\gamma \in \mathbb{N}_{>0}$ such that for all $(u,v) \in \mathcal{U}_e$ with $\Delta(u) > 0$ we have:*
   (a) *$\frac{|u|}{|v|} = \alpha$.*
   (b) *For all $u', v' \in \Sigma^+$ with u' prefix of u and v' suffix of v such that $\frac{|u'|}{|v'|} = \alpha$, we have that $-\Delta(v') - \beta \le \Delta(u') \le -\Delta(v') + \beta$.*
   (c) *For all factors $u' \in \Sigma^*$ of u such that $|u'| = \gamma$, we have $\Delta(u') \ge 1$.*
   (d) *For all factors $v' \in \Sigma^*$ of v such that $|v'| = \gamma$, we have $\Delta(v') \le -1$.*

Proof. The implication from Point 2 to Point 1 is trivial since Point 2 (a) implies Point 1.

Let us now prove that Point 1 implies Point 2. Fix any $e \in O$ that is F-reachable and assume that $\mathcal{U}_e$ is length-synchronous. Point 2 (a) follows immediately from length-synchronicity of $\mathcal{U}_e$. We can hence write $\alpha = \frac{A}{B}$ for some $A, B \in \mathbb{N}_{>0}$.

For proving Point 2 (b), we define $\beta = (n+1) \cdot (|O| + \max(A,B) + 1)$, where n is the constant taken from Lemma 4.1. Let $(u,v) \in \mathcal{U}_e$ with $\Delta(u) > 0$ and let

$$\mathrm{ext}_{u,v} = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_{h-1},y_{h-1}} \circ \mathrm{ext}_{a_{h-1},b_{h-1}} \circ \mathrm{ext}_{x_h,y_h}$$

be the stair factorization of $\mathrm{ext}_{u,v}$ according to Lemma 3.6. Since our morphism (φ,ψ) is F-weakly-length-synchronous by assumption, we have $|x_i|, |y_i| \leq n$ by Lemma 4.6. Let $u' \in \Sigma^*$ be a prefix of u and v' be a suffix of v such that $\frac{|u'|}{|v'|} = \alpha$. If $(u',v') = (u,v)$ we are done since then $\Delta(u') = -\Delta(v')$. Thus, it remains to consider the case when u' is a strict prefix of u and v' is a strict suffix of v: indeed, due to $\frac{|u|}{|v|} = \frac{|u'|}{|v'|} = \alpha$ we have that u' is a strict prefix of u if, and only if, v' is a strict suffix of v.

Let $j \in [1,h]$ be maximal such that $x_1 \cdots a_{j-1}x_j$ is a prefix of u' and $y_j b_{j-1} \cdots y_1$ is a suffix of v'. If $j = 1$ we are done, since then $\min\{|u'|,|v'|\} \leq n$, so that $|\Delta(u') + \Delta(v')| \leq |u'| + |v'| \leq n + n \cdot \max(A,B) \leq \beta$. So assume now that $j > 1$, which implies that $\Delta(u'') > 0$. Note that $j < h$ since $(u',v') \neq (u,v)$. Hence there exist unique words $s,t \in \Sigma^*$ such that $u' = u''s$ and $v' = tv''$, where $u'' = x_1 \cdots a_{j-1}x_j$ and $v'' = y_j b_{j-1} \cdots y_1$. By maximality of j we have $\min\{|s|,|t|\} \leq n$. Setting $f = \psi(\mathrm{ext}_{u'',v''})$ and $g = \psi(\mathrm{ext}_{a_j x_{j+1} \cdots a_{h-1} x_h, y_h b_{h-1} \cdots y_{j+1} b_j})$ we have $\psi(\mathrm{ext}_{u,v}) = f \circ g$. We claim that there exist $\mathrm{ext}_{x_g,y_g} \in \mathcal{O}(\Sigma^\triangle)$ such that $\psi(\mathrm{ext}_{x_g,y_g}) = g$ and $|x_g|, |y_g| \leq |O| \cdot (n+1)$: indeed, by the pigeonhole principle and Lemma 4.6 (as $e \circ f \circ g = e$ and $\Delta(u'') > 0$), any $\mathrm{ext}_{x,y} \in \mathcal{O}(\Sigma^\triangle)$ such that $\psi(\mathrm{ext}_{x,y}) = g$ and $\max(|x|,|y|) > |O| \cdot (n+1)$ must have a stair factorization according to Lemma 3.6 with an $h > |O|$ and can thus be factorized as $\mathrm{ext}_{x,y} = \mathrm{ext}_{x',y'} \circ \mathrm{ext}_{x'',y''} \circ \mathrm{ext}_{x''',y'''}$ such that $\psi(\mathrm{ext}_{x,y}) = \psi(\mathrm{ext}_{x',y'}) \circ \psi(\mathrm{ext}_{x''',y'''})$, where moreover $(x'',y'') \in \Sigma^+ \times \Sigma^+$. Thus, $\psi(\mathrm{ext}_{u''x_g,y_g v''}) = \psi(\mathrm{ext}_{u,v})$ and therefore $(u''x_g, y_g v'') \in \mathcal{U}_e$ with $\Delta(u''x_g) > 0$. It follows $\alpha = \frac{|u''x_g|}{|y_g v''|} = \frac{|u'|-|s|+|x_g|}{|y_g|+|v'|-|t|}$, or equivalently, using $\frac{|u'|}{|v'|} = \alpha$:

$$|s| = |u'| + |x_g| + \alpha(|t| - |y_g| - |v'|) = |x_g| + \alpha(|t| - |y_g|) \tag{2}$$

$$|t| = \frac{|s| - |u'| - |x_g|}{\alpha} + |y_g| + |v'| = \frac{|s| - |x_g|}{\alpha} + |y_g| \; . \tag{3}$$

Finally, we obtain

$$\begin{aligned}
|\Delta(u') + \Delta(v')| &= |\Delta(u''s) + \Delta(tv'')| \\
&\overset{\Delta(u'') = -\Delta(v'')}{=} |\Delta(s) + \Delta(t)| \\
&\leq |s| + |t| \\
&= \min(|s|,|t|) + \max(|s|,|t|) \\
&\leq n + \max(|s|,|t|) \\
&\overset{(2),(3)}{\leq} n + \max\left(|x_g| + \alpha(n - |y_g|), \frac{n - |x_g|}{\alpha} + |y_g|\right) \\
&\leq n + |O| \cdot (n+1) + n \cdot \max(A,B) \\
&\leq (n+1) \cdot (|O| + \max(A,B) + 1) \\
&= \beta \; .
\end{aligned}$$

This proves Point 2 (b).

For Point 2 (c) and Point 2 (d) we set $\gamma = (\lceil \frac{n}{2} \rceil + 1) \cdot (n+1) + n$ and remark that γ does not depend on u nor v. We only prove Point 2 (c), the proof of Point 2 (d) is analogous. As above, let

$$\mathrm{ext}_{u,v} = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_{h-1},y_{h-1}} \circ \mathrm{ext}_{a_{h-1},b_{h-1}} \circ \mathrm{ext}_{x_h,y_h}$$

be the stair factorization of $\mathrm{ext}_{u,v}$ according to Lemma 3.6. Let u' with $|u'| \geq \gamma$ be a factor of u and hence of $x_1 a_1 x_2 \cdots x_{h-1} a_{h-1} x_h$. By definition of stair factorization we have $\Delta(x_i) = 0$ for all $i \in [1,h]$ and $\Delta(a_i) = 1$ for all $i \in [1, h-1]$. Let w be the longest prefix of u' such that $\Delta(w) = \min\{\Delta(x) \mid x \text{ is a prefix of } u'\}$. Since $|x_1|, |y_1|, \ldots, |x_h|, |y_h| \leq n$, it immediately follows $\Delta(w) \geq -\frac{n}{2}$ and $|w| \leq n$. By the same reason, every prefix of the form ws of u' satisfies $\Delta(ws) \geq \Delta(w) + \frac{|s|}{n+1}$. Thus we have

$$\Delta(u') \geq \Delta(w) + \frac{|u'| - |w|}{n+1} \geq -\frac{n}{2} + \frac{((\lceil \frac{n}{2} \rceil + 1) \cdot (n+1) + n) - n}{n+1} \geq 1 \ . \qquad \square$$

The following proposition relates, for languages L generated by vertically visibly pushdown grammars G, (weak) length-synchronicity of L with (weak) length-synchronicity of $\mathcal{R}(G)$.

Proposition 4.12. *Let $L = L(G)$ for some vertically visibly pushdown grammar $G = (V, \Sigma, P, S)$. Moreover, let $\mathcal{R}(G) = \{(u,v) \in \mathsf{Con}(\Sigma) \mid S \Rightarrow_G^* uSv\}$. Then the following equivalences hold:*

1. *$L(G)$ is length-synchronous if, and only if, $\mathcal{R}(G)$ is length-synchronous.*
2. *$L(G)$ is weakly length-synchronous if, and only if, $\mathcal{R}(G)$ is weakly length-synchronous.*

Proof. Let $L = L(G)$ for a vertically visibly pushdown grammar $G = (V, \Sigma, P, S)$. Moreover, let $(\varphi_L, \psi_L) : (\mathcal{O}(\Sigma), \Sigma^{\triangle}) \to (R_L, O_L)$ be the syntactic morphism. For all $e \in O_L$ recall the set of contexts $\mathcal{R}_e = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi_L(\mathrm{ext}_{u,v}) = e\}$.

Let $F = \{e \in O_L \mid \exists\, \mathrm{ext}_{u,v} \in \psi_L^{-1}(e) : (u,v) \in \mathcal{R}(G)\}$. Observe that F is a submonoid of O_L all of whose elements are $\varphi_L(L)$-reachable since G is vertically visibly pushdown. Also observe that $\mathcal{R}(G) = \bigcup_{f \in F} \mathcal{R}_f$. We claim that since G is vertically visibly pushdown, there exists a constant $C > 0$ such that $\frac{1}{C} \leq \frac{|x|}{|y|} \leq C$ for all $(x,y) \in \mathcal{R}(G) \setminus \{(\varepsilon,\varepsilon)\}$: indeed, one can take $C = \frac{A}{B}$ where $A = \max\big\{\max\{|u|,|v|\} \;\big|\; T \to_G uT'v\big\}$ and $B = \min\big\{\min\{|u|,|v|\} \;\big|\; T \to_G uT'v\big\}$ since whenever $T \Rightarrow_G^* xT'y$ thanks to a derivation comprising $k \in \mathbb{N}_{>0}$ steps, we have

$$\frac{1}{C} = \frac{k \cdot B}{k \cdot A} \leq \frac{|x|}{|y|} \leq \frac{k \cdot A}{k \cdot B} = C \ .$$

Next, we prove that for all $\varphi_L(L)$-reachable idempotents $e \in O_L$ there exist $g, h \in O_L$ such that $g \circ e \circ h \in F$. Fix any such $\varphi_L(L)$-reachable idempotent $e \in O_L$. Without loss of generality let us assume that e is not the identity in O_L (indeed, if e is the identity in O_L, then we are done since we can then choose $g = h = e \in F$). Thus there exist $g \in O_L$ and $r \in R_L$ such that $(g \circ e)(r) \in \varphi_L(L)$. Moreover, let $\mathrm{ext}_{u',v'} \in \psi_L^{-1}(g), \mathrm{ext}_{u,v} \in \psi_L^{-1}(e)$, and $w \in \varphi_L^{-1}(r)$. Observe that we must have $(u,v) \neq (\varepsilon,\varepsilon)$ since $\psi_L(\mathrm{ext}_{\varepsilon,\varepsilon})$ is the identity in O_L. Since e is an idempotent we have that $u'u^n wv^n v' \in L$ for all $n \geq 1$. Fix a sufficiently large $N \geq 1$ such that $\frac{|u'|+|u|}{(N-1)|u|+|w|+N|v|+|v'|} \leq \frac{1}{C}$ and $\frac{|u'|+N|u|+|w|+(N-1)|v|}{|v|+|v'|} \geq C$, which exists due to $(u,v) \neq (\varepsilon,\varepsilon)$. Since $u'u^N wv^N v' \in L$ there exists $(x,y) \in \mathcal{R}(G) \setminus \{(\varepsilon,\varepsilon)\}$ such that $xy = u'u^N wv^N v'$, $S \Rightarrow_G^* xSy$, $|x| \geq |u'u|$, and $|y| \geq |vv'|$. Let $(x',y') \in \Sigma^* \times \Sigma^*$ be such that $(x,y) = (u'ux', y'vv')$. As $(u',v'), (u,v), (x,y) \in \mathsf{Con}(\Sigma)$,

$x' \in (\Sigma \setminus \Sigma_{\text{ret}})^*$, and $y' \in (\Sigma \setminus \Sigma_{\text{call}})^*$, we can conclude $(x', y') \in \mathsf{Con}(\Sigma)$. That is, $(x, y) = (u', v') \circ (u, v) \circ (x', y') \in \mathcal{R}(G)$. Hence,

$$g \circ e \circ \psi_L(\text{ext}_{x',y'}) = \psi_L(\text{ext}_{u'ux',y'vv'}) = \psi_L(\text{ext}_{x,y}) \in F.$$

We are now ready to prove Point 1. For the first direction, let us assume that $L(G)$ is length-synchronous. Recalling that F is a submonoid of O_L all of whose elements are $\varphi_L(L)$-reachable, we obtain that $R(G) = \bigcup_{f \in F} \mathcal{R}_f$ is length-synchronous by Lemma 4.9.

Conversely, let us assume that $\mathcal{R}(G)$ is length-synchronous. Assume by contradiction that $L(G)$ is not length-synchronous. Hence $\mathcal{R}_e$ is not length-synchronous for some $\varphi_L(L)$-reachable idempotent $e \in O_L$, i.e. $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'}) = e$ for some $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^{\triangle})$ such that $\Delta(u), \Delta(u') > 0$ and $\frac{|u|}{|v|} \neq \frac{|u'|}{|v'|}$. Without loss of generality we may assume that $|v| = |v'|$ (indeed, if $|v| \neq |v'|$, then $\text{ext}_{u^{|v'|},v^{|v'|}}, \text{ext}_{(u')^{|v|},(v')^{|v|}}$ satisfies the desired property). As a consequence we have $|u| \neq |u'|$, say $|u| < |u'|$ without loss of generality. Since e is a $\varphi_L(L)$-reachable idempotent, as argued above, there exist $g, h \in O_L$ such that $g \circ e \circ h = f'$ for some $f' \in F$. Let us fix $\text{ext}_{x_h,y_h} \in \psi_L^{-1}(h)$ and $\text{ext}_{x_g,y_g} \in \psi_L^{-1}(g)$. We have $\psi_L(\text{ext}_{x_gux_h,y_hvy_g}) = f' = \psi_L(\text{ext}_{x_gu'x_h,y_hv'y_g})$. Since $|x_gux_h| < |x_gu'x_h|$ and $|y_hvy_g| = |y_hv'y_g|$ it follows that $\mathcal{R}_{f'}$ is not length-synchronous, contradicting our assumption that $R(G) = \bigcup_{f \in F} \mathcal{R}_f$ is length-synchronous.

Let us next prove Point 2.

Let us first assume that $L(G)$ is weakly length-synchronous. Again, since F is a submonoid of O_L all of whose elements are $\varphi_L(L)$-reachable, we obtain that $R(G) = \bigcup_{f \in F} \mathcal{R}_f$ is weakly length-synchronous by Lemma 4.4.

Conversely, assume $\mathcal{R}(G)$ is weakly length-synchronous. Assume by contradiction that $\mathcal{R}_e$ is not weakly length-synchronous for some $\varphi_L(L)$-reachable idempotent $e \in O_L$. Thus, $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'}) = e$ for some $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^{\triangle})$ such that $\Delta(u), \Delta(u') > 0$ and moreover either $u = u'$ and $|v| \neq |v'|$ or $v = v'$ and $|u| \neq |u'|$. Without loss of generality let us assume that $u = u'$ and $|v| \neq |v'|$. As mentioned above, there exist $g, h \in O_L$ such that $g \circ e \circ h = f'$ for some $f' \in F$. Fix some $\text{ext}_{x_g,y_g} \in \psi_L^{-1}(g)$ and some $\text{ext}_{x_h,y_h} \in \psi_L^{-1}(h)$. Analogously, as argued above, we have $\psi_L(\text{ext}_{x_gux_h,y_hvy_g}) = \psi_L(\text{ext}_{x_gux_h,y_hv'y_g}) = f'$, $\Delta(x_gux_h) > 0$, and $|y_hvy_g| \neq |y_hv'y_g|$, implying that $\mathcal{R}_{f'}$ is not weakly length-synchronous, a contradiction to our assumption that $\mathcal{R}(G) = \bigcup_{f \in F} \mathcal{R}_f$ is weakly length-synchronous. □

4.3 The nesting depth of visibly pushdown languages

Another central notion is the nesting depth of well-matched words, which is the Horton-Strahler number [15] of the underlying trees.

Definition 4.13. *The* nesting depth *of well-matched words is given by the function* $\text{nd}\colon \Sigma^{\triangle} \to \mathbb{N}$ *defined inductively as follows:*

- $\text{nd}(\varepsilon) = 0$*;*
- $\text{nd}(c) = 0$ *for all* $c \in \Sigma_{int}$*;*
- $\text{nd}(uv) = \max\{\text{nd}(u), \text{nd}(v)\}$ *for all* $u \in \Sigma_{call}\Sigma^{\triangle}\Sigma_{ret} \cup \Sigma_{int}$ *and* $v \in \Sigma^{\triangle} \setminus \{\varepsilon\}$*;*
- $\text{nd}(awb) = \begin{cases} \text{nd}(w) + 1 & \text{if } w = uv \text{ with } u, v \in \Sigma^{\triangle} \text{ and } \text{nd}(w) = \text{nd}(u) = \text{nd}(v) \\ \text{nd}(w) & \text{otherwise} \end{cases}$ *for all* $a \in \Sigma_{call}$*,* $b \in \Sigma_{ret}$ *and* $w \in \Sigma^{\triangle}$*.*

An important property of weakly length-synchronous VPLs is that their words have bounded nesting depth.

Proposition 4.14. *For each weakly length-synchronous VPL $L \subseteq \Sigma^\triangle$ there exists a constant $d \in \mathbb{N}$ such that $L \subseteq \{w \in \Sigma^\triangle \mid \mathrm{nd}(w) \leq d\}$.*

Proposition 4.14 is proved in several steps. For these we introduce a factorization that can be seen as one that witnesses the nesting depth of a word.

Definition 4.15. *A* nesting-maximal stair factorization *of $w \in \Sigma^\triangle$ with $\mathrm{nd}(w) \geq 1$ is a factorization of w as*

$$w = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(w')$$

such that $k \geq 0$, $x_i, y_i \in \Sigma^\triangle$, $a_i \in \Sigma_{call}$, and $b_i \in \Sigma_{ret}$ for all $i \in [1,l]$, and $w' \in \Sigma^*_{int}$ satisfying that for all $i \in [1,k]$ we have

$$\mathrm{nd}(\mathrm{ext}_{x_i,y_i}(w_i)) = \mathrm{nd}(w_i),$$

where $w_i = \mathrm{ext}_{a_i,b_i} \circ \mathrm{ext}_{x_{i+1},y_{i+1}} \circ \cdots \circ \mathrm{ext}_{a_k,b_k}(w')$.

Lemma 4.16. *All words $w \in \Sigma^\triangle$ have a nesting-maximal stair factorization.*

Proof. The proof goes by structural induction on w.

- $w = \varepsilon$. Then we are done because w contains only internal letters.
- $w = c$ for a $c \in \Sigma_{\mathrm{int}}$. Then we are again done because w contains only internal letters.
- $w = aw'b$ for $a \in \Sigma_{\mathrm{call}}$, $b \in \Sigma_{\mathrm{ret}}$ and $w' \in \Sigma^\triangle$. By using the inductive hypothesis, w' has a nesting-maximal stair factorization $\mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(w'')$. It directly follows that $\mathrm{ext}_{a,b} \circ \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(w'')$ is a nesting-maximal stair factorization of w.
- $w = uv$ for $u, v \in \Sigma^\triangle \setminus \{\varepsilon\}$. Then w can be decomposed as $z_1 \cdots z_m$ with $z_1, \ldots, z_m \in \Sigma_{\mathrm{call}}\Sigma^\triangle\Sigma_{\mathrm{ret}} \cup \Sigma_{\mathrm{int}}$ and $m \in \mathbb{N}, m \geq 2$. In this case, either $z_i \in \Sigma_{\mathrm{int}}$ for all $i \in [1,m]$ and thus we are done because w contains only internal letters, or there exists some $i \in [1,m]$ such that $z_i \in \Sigma_{\mathrm{call}}\Sigma^\triangle\Sigma_{\mathrm{ret}}$ and has maximal nesting depth, i.e. $\mathrm{nd}(w) = \mathrm{nd}(z_i)$. In this second subcase, we have that $z_i = az'_ib$ with $a \in \Sigma_{\mathrm{call}}$, $b \in \Sigma_{\mathrm{ret}}$ and $z'_i \in \Sigma^\triangle$. By using the inductive hypothesis, z'_i has a nesting-maximal stair factorization $\mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(w'')$. Therefore,

  $$\mathrm{ext}_{z_1\cdots z_{i-1},z_{i+1}\cdots z_k} \circ \mathrm{ext}_{a,b} \circ \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \ldots \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(w'')$$

  is a nesting-maximal stair factorization of w. □

The following lemma will be a useful tool for proofs by induction on the nesting depth of well-matched words.

Lemma 4.17. *Let $u = a_1vb_1 \in \Sigma^\triangle$ for some $a_1 \in \Sigma_{call}$, $b_1 \in \Sigma_{ret}$, and $v \in \Sigma^\triangle$ such that $\mathrm{nd}(u) = d > 0$. Moreover, let $u = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(u')$ be a nesting-maximal stair factorization of u (i.e. $x_1 = y_1 = \varepsilon$). Then there exists $h \in [1,k]$ such that, setting $u_i = \mathrm{ext}_{a_i,b_i} \circ \mathrm{ext}_{x_{i+1},y_{i+1}} \circ \cdots \circ \mathrm{ext}_{a_k,b_k}(u')$ for all $i \in [1,k]$ and $u_{k+1} = u'$, we have*

1. $\mathrm{nd}(u) = \mathrm{nd}(u_h) = d$,
2. $\mathrm{nd}(u_{h+1}) = d - 1$, *and*
3. $\mathrm{nd}(x_1), \mathrm{nd}(y_1), \ldots, \mathrm{nd}(x_h), \mathrm{nd}(y_h) < d$.

Proof. Let $u_j = \mathrm{ext}_{a_j,b_j} \circ \mathrm{ext}_{x_{j+1},y_{j+1}} \circ \cdots \circ \mathrm{ext}_{a_k,b_k}(u')$ for all $j \in [1,k]$. Note that we have $\mathrm{nd}(u) = \mathrm{nd}(u_1) = d > 0$ by assumption. Moreover, $\mathrm{nd}(u_j) \geq \mathrm{nd}(u_{j+1})$ for all $j \in [1,k-1]$ by definition of nesting depth. Thus, since $\mathrm{nd}(u_k) = 1 > 0 = \mathrm{nd}(u_{k+1})$, it follows that

$$h = \min\{j \in [1,k] \mid \mathrm{nd}(u_j) > \mathrm{nd}(u_{j+1})\}$$

is well-defined and $\mathrm{nd}(u) = \mathrm{nd}(u_1) = \mathrm{nd}(u_h) = d$, thus showing Point 1. Since

$$d = \mathrm{nd}(u_h) \leq \mathrm{nd}(u_{h+1}) + 1$$

and $\mathrm{nd}(u_{h+1}) < \mathrm{nd}(u_h) = d$ it follows $\mathrm{nd}(u_{h+1}) = d-1$, thus showing Point 2. To prove Point 3, assume by contradiction that $\mathrm{nd}(x_j) \geq d$ or $\mathrm{nd}(y_j) \geq d$ for some $j \in [1,h]$. Without loss of generality assume $\mathrm{nd}(x_j) \geq d$. Since $x_1 = y_1 = \varepsilon$ and $d > 0$ we must have $j \in [2,h]$. It follows

$$\mathrm{nd}(u) \geq \mathrm{nd}(u_{j-1}) = \mathrm{nd}(a_{j-1}x_j u_j y_j b_{j-1}) \geq \min(\mathrm{nd}(x_j), \mathrm{nd}(u_j)) + 1 \geq d+1 > d = \mathrm{nd}(u)\ ,$$

which is a contradiction. □

We are now ready to prove Proposition 4.14.

Proof of Proposition 4.14. Let $L \subseteq \Sigma^\triangle$ be a weakly length-synchronous VPL. We claim that $\mathrm{nd}(L) \leq n+1$, where n is the pumping constant from Lemma 4.1. Assume by contradiction that $\mathrm{nd}(u) = d$ for some $u \in L$ and some $d > n+1$. Let $u = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(u')$ be a nesting-maximal stair factorization of u according to Lemma 4.16. According to Lemma 4.17 there exists $i \in [1,k]$ such that, setting $u_j = \mathrm{ext}_{a_j,b_j} \circ \mathrm{ext}_{x_{j+1},y_{j+1}} \circ \cdots \circ \mathrm{ext}_{a_k,b_k}(u')$ for all $j \in [1,k]$ and $u_{k+1} = u'$, we have $\mathrm{nd}(u) = \mathrm{nd}(u_i) = d$ and $\mathrm{nd}(u_{i+1}) = d-1$. Since $d-1 > n > 0$, we must have $i+1 \leq k$, so that $u_i = a_i x_{i+1} u_{i+1} y_{i+1} b_i$ with $u_{i+1} \in \Sigma_{\mathrm{call}}\Sigma^\triangle\Sigma_{\mathrm{ret}}$ and $\mathrm{nd}(x_{i+1}u_{i+1}y_{i+1}) = \mathrm{nd}(u_{i+1}) = d-1$. Hence it follows that $\mathrm{nd}(x_{i+1}) = d-1$ or $\mathrm{nd}(y_{i+1}) = d-1$. Without loss of generality let us assume $\mathrm{nd}(y_{i+1}) = d-1 > n$. A simple induction shows that $|x| \geq 2^{\mathrm{nd}(x)} - 1 \geq \mathrm{nd}(x)$ for all $x \in \Sigma^\triangle$. Thus, we have $|y_{i+1}| \geq \mathrm{nd}(y_{i+1}) > n$, contradicting Proposition 4.6. □

4.4 Quasi-aperiodicity and its correspondence with quasi-counterfreeness

Let us revisit the notion of quasi-aperiodicity. It has already been defined for visibly pushdown languages in [24]. Let us define $\mathcal{O}(\Sigma^\triangle)^{k,l} = \{\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle) : |u| = k, |v| = l\}$ for all $k,l \in \mathbb{N}$. We say the morphism $(\varphi,\psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R,O)$ is *quasi-aperiodic* if all semigroups contained in the set $\psi(\mathcal{O}(\Sigma^\triangle)^{k,l})$ are aperiodic for all $k,l \in \mathbb{N}$.

The following proposition relates quasi-counterfreeness of a VPL with quasi-aperiodicity of its syntactic morphism.

Proposition 4.18. *A VPL $L \subseteq \Sigma^\triangle$ is quasi-counterfree if, and only if, its syntactic morphism is quasi-aperiodic.*

Proof. Recall that $(R_L, O_L) = (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle))/\!\sim_L$ by Definition 3.13. Hence for all $(u,v),(u',v') \in \mathsf{Con}(\Sigma)$ we have $(u,v) \equiv_L (u',v')$ if, and only if, $\psi_L(\mathrm{ext}_{u,v}) = \psi_L(\mathrm{ext}_{u',v'})$.

First, let us assume that $(\varphi_L,\psi_L)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R_L,O_L)$ is quasi-aperiodic. Assume by contradiction that L is not quasi-counterfree. Thus, there exists some $\sigma = (u,v) \in \mathsf{Con}(\Sigma)$ such that $\sigma^n \not\equiv_L \sigma^{n+1}$ for all $n \in \mathbb{N}$ and $\tau \equiv_L \sigma \circ \sigma$ for some $\tau = (x,y) \in \mathsf{Con}(\Sigma) \cap \Sigma^{|u|} \times \Sigma^{|v|}$. The latter can equivalently be rephrased as $\psi_L(\mathrm{ext}_{x,y}) = \psi_L(\mathrm{ext}_{u^2,v^2})$. By choice of σ and the fact that $\equiv_L$ has finite index there exist $s \geq 1$ and $t \geq 2$ such that $\sigma^{s+i} \equiv_L \sigma^{s+i+t}$ for all $i \geq 0$ and $\sigma^{s+i} \not\equiv_L \sigma^{s+j}$ for all $i,j \in [0,t-1]$ with $i \neq j$. It follows that the set $G = \psi_L(\{\mathrm{ext}_{u^{s+i},v^{s+i}} \mid i \in [0,t-1]\})$ is a non-trivial

group with identity $g_0 = \psi_L(\mathrm{ext}_{u^{s+k},v^{s+k}}) \in \psi_L(\mathcal{O}(\Sigma^\triangle)^{(s+k)|u|,(s+k)|v|})$, where k is the unique integer in $[0, t-1]$ such that $s+k$ is divisible by t. Also observe that G is a cyclic group that is generated by $g_1 = \psi_L(\mathrm{ext}_{u^{s+k+1},v^{s+k+1}})$. That is $G = \{g_0, g_1, \ldots, g_{t-1}\}$, where $g_i = g_1^i$ for all $i \in [2, t-1]$. But note that due to $\psi_L(\mathrm{ext}_{x,y}) = \psi_L(\mathrm{ext}_{u^2,v^2})$ we obtain $g_1 = \psi_L(\mathrm{ext}_{u^{s+k+1},v^{s+k+1}}) = \psi_L(\mathrm{ext}_{xu^{s+k-1},v^{s+k-1}y})$ and thus $g_0, g_1 \in \psi_L(\mathcal{O}(\Sigma^\triangle)^{(s+k)|u|,(s+k)|v|})$ due to $|x| = |u|$ and $|y| = |v|$. Since $g_i = g_0^{t-1-i} g_1^i$ for all $i \in [0, t-1]$ we obtain that $G = \{g_0, g_1, \ldots, g_{t-1}\}$ is contained in $\psi_L(\mathcal{O}(\Sigma^\triangle)^{(t-1)(s+k)|u|,(t-1)(s+k)|v|})$, hereby contradicting quasi-aperiodicity of (φ_L, ψ_L).

For the converse direction, let us assume that L is quasi-counterfree. Assume by contradiction that (φ_L, ψ_L) is not quasi-aperiodic. That is, for some $k, l \in \mathbb{N}$ the set $\psi_L(\mathcal{O}(\Sigma^\triangle)^{k,l})$ contains a non-trivial group G. Let $g_0 \in G$ be the identity of G. Fix some $g \in G$ with $g \neq g_0$ and some $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle)^{k,l}$ such that $\psi_L(\mathrm{ext}_{u,v}) = g$. Let $\sigma = (u, v) \in \mathsf{Con}(\Sigma) \cap \Sigma^k \times \Sigma^l$. Since g is not the identity in G we have that $\psi_L(\mathrm{ext}_{u^n,u^n}) \neq \psi_L(\mathrm{ext}_{u^{n+1},v^{n+1}})$ for all $n \in \mathbb{N}$, equivalently $\sigma^n \not\equiv_L \sigma^{n+1}$ for all $n \in \mathbb{N}$. But since moreover $\psi_L(\mathrm{ext}_{u^n,v^n})$ is in G and thus in $\psi_L(\mathcal{O}(\Sigma^\triangle)^{k,l})$ for all $n \in \mathbb{N}$ it follows that $\sigma \circ \sigma \equiv_L \tau$ for some $\tau \in \mathsf{Con}(\Sigma) \cap \Sigma^k \times \Sigma^l$. We thus obtain a contradiction to our assumption that L is quasi-counterfree. □

5 Proof of the main theorem

Before giving an overview of the proof of Theorem 2.9 we will state a proposition saying that the syntactic Ext-algebra and the syntactic morphism of a given visibly pushdown language L is computable and that it is decidable if L is quasi-aperiodic, length-synchronous, and weakly length-synchronous, respectively. Its proof is subject of Section 6.

Proposition 5.1. *The following computability and decidability results hold:*

1. *Given a DVPA A, one can effectively compute the syntactic* Ext*-algebra of $L = L(A)$, its syntactic morphism (φ_L, ψ_L) and $\varphi_L(L)$.*
2. *Given a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ for a visibly pushdown alphabet Σ and a finite* Ext*-algebra (R, O), all of the following are decidable for (φ, ψ):*
   (a) *Quasi-aperiodicity. In case (φ, ψ) is not quasi-aperiodic, one can effectively compute $k, l \in \mathbb{N}$ such that $\psi(\mathcal{O}(\Sigma^\triangle)^{k,l})$ is not aperiodic.*
   (b) *F-length-synchronicity for a given $F \subseteq R$. In case (φ, ψ) is not F-length-synchronous, one can effectively compute a quadruple $(k, l, k', l') \in \mathbb{N}^4_{>0}$ such that there exist $uv, u'v' \in \Sigma^\triangle$ and some F-reachable idempotent $e \in O$ such that $\psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u',v'}) = e$, $\Delta(u) > 0$, $\Delta(u') > 0$, $k = |u|, l = |v|, k' = |u'|$, $l' = |v'|$, and $\frac{k}{l} \neq \frac{k'}{l'}$.*
   (c) *F-weakly-length-synchronicity for a given $F \subseteq R$.*

5.1 Proof outline for Theorem 2.9

Towards proving our main result (Theorem 2.9), given a DVPA A, where $L = L(A)$ is a VPL over a visibly pushdown alphabet Σ, we apply Proposition 5.1 and compute its syntactic Ext-algebra (R_L, O_L) along with its syntactic morphism (φ_L, ψ_L) and the subset $\varphi_L(L)$. Then we make the following effective case distinction which immediately implies Theorem 2.9.

1. If L is not weakly length-synchronous, then L is TC^0-hard and hence not in AC^0 (Proposition 5.4 in Section 5.2). Thus, we can output any $m > 1$ since $\mathrm{MOD}_m \leq_{\mathrm{cd}} \mathrm{EQUALITY} \leq_{\mathrm{cd}} L$ for any $m > 1$.

2. If L is not quasi-aperiodic, then one can effectively compute some $m \geq 2$ such that $\mathrm{MOD}_m \leq_{\mathrm{cd}} L$ (Proposition 5.5 in Section 5.2).

3. If L is length-synchronous and (φ_L, ψ_L) is quasi-aperiodic, then $L \in \mathsf{AC}^0$ (Theorem 5.7 in Section 5.3).

4. If L is weakly length-synchronous but not length-synchronous, and its syntactic morphism (φ_L, ψ_L) is quasi-aperiodic, one can effectively compute vertically visibly pushdown grammars $G_1, \ldots, G_m$ generating intermediate VPLs such that $L =_{\mathrm{cd}} \biguplus_{i=1}^{m} L(G_i)$ (Theorem 5.17 in Section 5.4). Moreover, already if a VPL L is weakly length-synchronous but not length-synchronous, one can effectively compute $k, l \in \mathbb{N}_{>0}$ with $k \neq l$ such that $\mathcal{L}_{k,l} \leq_{\mathrm{cd}} L$ (Proposition 5.24 in Section 5.4).

We refer to Section 5.5 for the proof of Corollary 2.11.

5.2 Lower bounds

The following visibly pushdown languages are helpful for proving lower bounds.

Definition 5.2. *Let $L \subseteq \Sigma^{\triangle}$ be a VPL. For each $e \in O_L$ and for $\# \notin \Sigma$ a fresh internal letter we define*

$$L_e = \{u \# v \mid (u, v) \in \mathsf{Con}(\Sigma) : \psi_L(\mathrm{ext}_{u,v}) = e\}$$

and

$$L_{e\uparrow} = \{u \# v \mid (u, v) \in \mathsf{Con}(\Sigma) : \Delta(u) > 0, \psi(\mathrm{ext}_{u,v}) = e\} = L_e \cap \{u \# v \mid (u, v) \in \mathsf{Con}(\Sigma) : \Delta(u) > 0\}.$$

The next lemma shows that both $\varphi_L^{-1}(r)$ and L_e are constant-depth reducible to L in case $r \in R_L$ and $e \in O_L$ are $\varphi_L(L)$-reachable, respectively.

Lemma 5.3. *Let $L \subseteq \Sigma^{\triangle}$ be a VPL. Then*

- *$\varphi_L^{-1}(r) \leq_{cd} L$ for all $\varphi_L(L)$-reachable $r \in R_L$, and*
- *$L_e \leq_{cd} L$ for all $\varphi_L(L)$-reachable $e \in O_L$.*

Proof. To show the first point, let us fix some $\varphi_L(L)$-reachable $r \in R_L$. Thus, there exist $w_r \in \Sigma^{\triangle}$ and $(u_r, v_r) \in \mathsf{Con}(\Sigma)$ such that $\varphi_L(w_r) = r$ and $\varphi_L(u_r w_r v_r) \in \varphi_L(L)$. By definition of the syntactic morphism of L (Definition 3.13) for all $r_1, r_2 \in R_L$ with $r_1 \neq r_2$ there exists some $e_{r_1,r_2} \in O_L$ such that $e_{r_1,r_2}(r_1) \in \varphi_L(L) \Leftrightarrow e_{r_1,r_2}(r_2) \notin \varphi_L(L)$. For each such $e_{r_1,r_2} \in O_L$ fix $(u_{r_1,r_2}, v_{r_1,r_2}) \in \mathsf{Con}(\Sigma)$ with $\psi_L(\mathrm{ext}_{u_{r_1,r_2}, v_{r_1,r_2}}) = e_{r_1,r_2}$.

Hence, for all $w \in \Sigma^*$ we have

$$w \in \varphi_L^{-1}(r) \quad \Longleftrightarrow \quad u_r w v_r \in L \ \wedge \bigwedge_{\substack{r' \in R_L \\ r \neq r'}} u_{r,r'} w v_{r,r'} \in L \leftrightarrow u_{r,r'} w_r v_{r,r'} \in L \ ,$$

thus showing $\varphi_L^{-1}(r) \leq_{\mathrm{cd}} L$.

For the second point, let us fix some $\varphi_L(L)$-reachable $e \in O_L$. Fix some $(u_e, v_e) \in \mathsf{Con}(\Sigma)$ such that $\psi_L(\mathrm{ext}_{u_e,v_e}) = e$. Thus, again, there exist $w_e \in \Sigma^{\triangle}$ and $(u'_e, v'_e) \in \mathsf{Con}(\Sigma)$ such that $\varphi_L(u'_e u_e w_e v_e v'_e) \in \varphi_L(L)$. Again by definition of the syntactic morphism of L (Definition 3.13), for all $e_1, e_2 \in O_L$ with $e_1 \neq e_2$ there exist some $f_{e_1,e_2} \in O_L$ and some $r_{e_1,e_2} \in R_L$ such that

$f_{e_1,e_2}(e_1(r_{e_1,e_2})) \in \varphi_L(L) \Leftrightarrow f_{e_1,e_2}(e_2(r_{e_1,e_2})) \notin \varphi_L(L)$. For each such f_{e_1,e_2} and r_{e_1,e_2} fix, respectively, $(u_{e_1,e_2}, v_{e_1,e_2}) \in \mathsf{Con}(\Sigma)$ and $w_{e_1,e_2} \in \Sigma^{\triangle}$ such that $\psi_L(\mathrm{ext}_{u_{e_1,e_2},v_{e_1,e_2}}) = f_{e_1,e_2}$ and $\varphi_L(w_{e_1,e_2}) = r_{e_1,e_2}$, respectively. Hence, for all $u\#v \in \Sigma^*\#\Sigma^*$ we have

$$u\#v \in L_e \quad \Longleftrightarrow \quad u'_e u w_e v v'_e \in L \ \wedge \bigwedge_{\substack{e' \in O_L \\ e \neq e'}} u_{e,e'} u w_{e,e'} v v_{e,e'} \in L \leftrightarrow u_{e,e'} u_e w_{e,e'} v_e v_{e,e'} \in L \ ,$$

thus showing $L_e \leq_{\mathrm{cd}} L$. □

The following lower bound has already been sketched in Section 4.

Proposition 5.4. *If L is not weakly length-synchronous, then L is TC^0-hard.*

Proof. Recall that (R_L, O_L) is the syntactic Ext-algebra of L and $(\varphi_L, \psi_L) \colon (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R_L, O_L)$ is its syntactic morphism. Assume that (φ_L, ψ_L) is not $\varphi_L(L)$-weakly-length-synchronous.

Assume first there exist $\mathrm{ext}_{u,v}, \mathrm{ext}_{u',v} \in \mathcal{O}(\Sigma^{\triangle})$ satisfying that $\psi_L(\mathrm{ext}_{u,v}) = \psi_L(\mathrm{ext}_{u',v})$ that is a $\varphi_L(L)$-reachable idempotent such that $\Delta(u)$, $\Delta(u') > 0$, but $|u| \neq |u'|$. We exploit the fact that $|u| \neq |u'|$ to reduce EQUALITY $= \{w \in \{0,1\}^* : |w|_0 = |w|_1\}$ to $L_{\psi_L(\mathrm{ext}_{u,v})}$. The constant-depth reduction works as follows on input $w \in \{0,1\}^*$:

1. Check if $|w| = 2n$ for some $n \in \mathbb{N}$, reject if it is not the case.
2. Compute $w' = \alpha(w)$, where $\alpha\alpha\{0,1\}^* \to \Sigma^*$ is the length-multiplying morphism satisfying $\alpha(1) = u^{|u'|}$ and $\alpha(0) = u'^{|u|}$.
3. Accept whenever $w'\#v^{n(|u|+|u'|)} \in L_{\psi(\mathrm{ext}_{u,v})}$.

Bearing in mind that $0 < \Delta(u) = -\Delta(v) = \Delta(u')$, the latter forms a valid reduction, because given a word $w \in \{0,1\}^*$ of length $2n$ for an $n \in \mathbb{N}$ that contains $k \in [0, 2n]$ many 1's, for $w'\#v^{n(|u|+|u'|)}$ to be in $L_{\psi_L(\mathrm{ext}_{u,v})}$, it is in particular required that $w'v^{n(|u|+|u'|)}$ is well-matched, so it is necessary and sufficient that

$$\begin{aligned}
& k \cdot \Delta(u) \cdot |u'| + (2n-k) \cdot \Delta(u') \cdot |u| &&= -n \cdot \Delta(v) \cdot (|u| + |u'|) \\
\Longleftrightarrow \quad & (k-n) \cdot \Delta(u) \cdot |u'| + (n-k) \cdot \Delta(u') \cdot |u| &&= 0 \\
\Longleftrightarrow \quad & (k-n) \cdot \Delta(u) \cdot (|u'| - |u|) &&= 0 \\
\Longleftrightarrow \quad & k &&= n \ .
\end{aligned}$$

Additionally applying Lemma 5.3 we obtain EQUALITY $\leq_{\mathrm{cd}} L_{\psi_L(\mathrm{ext}_{u,v})} \leq_{\mathrm{cd}} L$. Assume now there exist $\mathrm{ext}_{u,v}, \mathrm{ext}_{u,v'} \in \mathcal{O}(\Sigma^{\triangle})$ satisfying that $\psi_L(\mathrm{ext}_{u,v}) = \psi_L(\mathrm{ext}_{u,v'})$ is an $\varphi_L(L)$-reachable idempotent such that $\Delta(u) > 0$ but $|v| \neq |v'|$. Symmetrically, one can prove that we also have EQUALITY $\leq_{\mathrm{cd}} L$ in this case.

In conclusion, as EQUALITY is TC^0-complete under constant-depth reductions, it follows that L is TC^0-hard under constant-depth reductions. □

The following proposition has essentially already been shown in [24, Proposition 135], yet with some inaccuracies (we refer to Section 8) that we fix here.

Proposition 5.5. *If L is not quasi-aperiodic, then one can effectively compute some $m \geq 2$ such that* $\mathrm{MOD}_m \leq_{cd} L$.

Proof. Since L is not quasi-aperiodic, by Point 2 (a) Proposition 5.1 one can effectively compute $k, l \in \mathbb{N}$ such that $\psi_L(\mathcal{O}(\Sigma^\triangle)^{k,l})$ is not aperiodic. Thus, one can compute $m \geq 2$ such that $\psi_L(\mathcal{O}(\Sigma^\triangle)^{k,l})$ contains the additive group $G = ([0, m-1], +, 0)$ of $\mathbb{Z}/m\mathbb{Z}$ for some prime number m. Moreover, there exist $\mathrm{ext}_{u_0,v_0}, \mathrm{ext}_{u_1,v_1} \in \mathcal{O}(\Sigma^\triangle)^{k,l}$ such that $\psi_L(\mathrm{ext}_{u_0,v_0}) = 0_G$ and $\psi_L(\mathrm{ext}_{u_1,v_1}) = 1_G$. Since G is a group both $\psi_L(\mathrm{ext}_{u_0,v_0})$ and $\psi_L(\mathrm{ext}_{u_1,v_1})$ are $\varphi_L(L)$-reachable. Moreover there exist $xy, z \in \Sigma^\triangle$ such that $xu_0zv_0y \in L$ if, and only if, $xu_1zv_1y \notin L$. Let us assume without loss of generality that $xu_0zv_0y \notin L$ and $xu_1zv_1y \in L$ (the case when $xu_0zv_0y \in L$ and $xu_1zv_1y \notin L$ can be proven analogously). Let $h_\uparrow, h_\downarrow : \{0,1\}^* \to \Sigma^*$ be the length-multiplying morphisms satisfying $h_\uparrow(i) = u_i$ and $h_\downarrow(i) = v_i$ for all $i \in \{0,1\}$. We claim that

$$w \in \mathrm{MOD}_m \iff \bigwedge_{i=1}^{m-1} xh_\uparrow(w)^{i^{m-2}} zh_\downarrow(w^R)^{i^{m-2}} y \notin L.$$

Let $w_i = xh_\uparrow(w)^{i^{m-2}} zh_\downarrow(w^R)^{i^{m-2}} y$ for all $i \in [1, m-1]$. Observe that $w_i \in \Sigma^\triangle$ for all $i \in [1, m-1]$ directly by definition of the morphisms $h_\uparrow$ and $h_\downarrow$.

To show the above equivalence, let us first assume that $|w|_1$ is divisible by m. Then we have $\psi_L(\mathrm{ext}_{h^\uparrow(w),h^\downarrow(w^R)}) = \psi_L(\mathrm{ext}_{u_0,v_0}) = 0_G$, and consequently $\psi_L(\mathrm{ext}_{h^\uparrow(w)^{i^{m-2}},h^\downarrow(w)^{i^{m-2}}}) = 0_G$ for all $i \in [1, m-1]$. It follows $w_i \notin L$ for all $i \in [1, m-1]$, as desired. Conversely, assume that $|w|_1$ is not divisible by m, i.e. $|w|_1 \equiv i \bmod m$ for some $i \in [1, m-1]$. Hence $\psi_L(\mathrm{ext}_{h^\uparrow(w),h^\uparrow(w^R)}) = i_G \neq 0_G$ and thus $\psi_L(\mathrm{ext}_{h^\uparrow(w)^{i^{m-2}},h^\uparrow(w^R)^{i^{m-2}}}) = (i^{m-1} \bmod m)_G = 1_G$ by Fermat's Little Theorem. Hence $w_i \in L$ as required.

Altogether we obtain $\mathrm{MOD}_m \leq_{\mathrm{cd}} L$. □

5.2.1 The non-solvable case

In this additional section we prove a stronger lower bound, namely when the syntactic morphism not only is not quasi-aperiodic but the syntactic Ext-algebra not solvable. For this we revisit solvable groups and introduce solvable Ext-algebras.

Let G be a finite group. The *word problem for* G is the question, given a word $w_1 \cdots w_n$ over G, to decide if their product $w_1 \cdots w_n$ in G evaluates to 1_G. The *commutator* of $g, h \in G$ is $ghg^{-1}h^{-1} \in G$, denoted by $[g, h]$. The *commutator subgroup* $[G, G]$ of G is the subgroup of G that is generated by the commutators of G. We say that G is *perfect* if $G = [G, G]$. We say that G is *solvable* if in the series of commutator subgroups (a.k.a. derived series) $G^{(0)}, G^{(1)}, \ldots$ a trivial group is contained, where $G^{(0)} = G$ and $G^{(i+1)} = [G^{(i)}, G^{(i)}]$ for all $i \in \mathbb{N}$. Thus, note that any non-solvable finite group contains a perfect subgroup.

We say the Ext-algebra (R, O) is *solvable* if all subsets of R or O that are groups (under the multiplication of R, resp. of O) are solvable. It is worth mentioning that one can prove that if $(\varphi, \psi) : (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ is quasi-aperiodic, then (R, O) is solvable. In fact, one can prove that if (φ, ψ) is quasi-aperiodic, then (R, O) must contain only Abelian groups.

Our proof that L is NC^1-hard (and thus TC^0-hard) when (R_L, O_L) is not solvable can be reduced to the case for words [4], by showing that already $\psi_L(\mathcal{O}(\Sigma^\triangle)^{k,l})$ contains such a non-solvable group for some fixed $k, l \geq 0$.

Proposition 5.6. *If (R_L, O_L) is not solvable, then L is NC^1-hard and thus not in AC^0.*

Before we prove the proposition we remark that not every subset $G \subseteq R_L$ (resp. $G \subseteq O_L$) that is a group is necessarily a submonoid of R_L (resp. O_L); in particular the neutral element of G need not necessarily be the neutral element of O_L. Indeed, for instance assume $R_L = \{1, a, b\}$ where

$1 \cdot r = 1 \cdot r = r$ for all $r \in R$ and where $a \cdot b = b \cdot a = b$ and $a \cdot a = b \cdot b = a$; the subset $\{a, b\}$ forms the additive group of $\mathbb{Z}/2\mathbb{Z}$ with neutral element a. It is also worth mentioning that since R is (isomorphic to) a submonoid of O we could have equivalently defined an Ext-algebra to be solvable if all subsets of O that are groups are solvable.

Proof of Proposition 5.6. Assume (R_L, O_L) is not solvable. Then there exists a subset $G \subseteq O_L$, where G is a non-trivial perfect group, i.e. $G = [G, G]$. Let ω be the idempotent power of G. For all $g, h \in G$ there exist $\mathrm{ext}_{u_g, v_g}, \mathrm{ext}_{u_h, v_h} \in \mathcal{O}(\Sigma^{\triangle})$ such that

$$[g, h] = ghg^{-1}h^{-1} = ghg^{\omega-1}h^{\omega-1} = \psi_L\left(\mathrm{ext}_{u_g u_h u_g^{\omega-1} u_h^{\omega-1}, v_h^{\omega-1} v_g^{\omega-1} v_h v_g}\right)$$

and $1_G = g^\omega h^\omega = \psi_L(\mathrm{ext}_{u_g^\omega u_h^\omega, v_h^\omega v_g^\omega})$. Therefore, for all $g, h \in G$ we have

$$[g, h] = \psi_L\left(\mathrm{ext}_{u_g u_h u_g^{\omega-1} u_h^{\omega-1}, v_h^{\omega-1} v_g^{\omega-1} v_h v_g} \circ \bigcirc_{\substack{(g',h') \in G^2 \\ (g',h') \neq (g,h)}} \mathrm{ext}_{u_{g'}^\omega u_{h'}^\omega, v_{h'}^\omega, v_{g'}^\omega}\right) .$$

Hence, $\{[g, h] \mid g, h \in G\} \subseteq \mathcal{O}(\Sigma^{\triangle})^{k,l}$ for

$$k = \sum_{(g,h) \in G^2} (|u_g| + |u_h|) \cdot \omega \quad \text{and} \quad l = \sum_{(g,h) \in G^2} (|v_h| + |v_g|) \cdot \omega \quad .$$

Since $G = [G, G]$ every element of G can be written as the product of at most $|G|$ elements in $\{[g, h] \mid g, h \in G\}$ and, in fact, even as the product of exactly $|G|$ elements in $\{[g, h] \mid g, h \in G\}$, since it contains the identity 1_G. Thus, we can conclude that $G \subseteq \psi_L(\mathcal{O}(\Sigma^{\triangle})^{k \cdot |G|, l \cdot |G|}$. Since the word problem of any non-solvable finite group is NC^1-hard by [4] and $G \subseteq \psi_L(\mathcal{O}(\Sigma^{\triangle})^{k \cdot |G|, l \cdot |G|}$, it follows that the word problem for G is constant-depth reducible to L. Hence L is NC^1-hard and in particular TC^0-hard. □

5.3 In AC^0: Length-synchronous and quasi-aperiodic

This section is devoted to the following theorem.

Theorem 5.7. *If L is length-synchronous and (φ_L, ψ_L) is quasi-aperiodic, then L is in $\mathsf{FO}[+]$ and thus in AC^0.*

For the rest of this section let us fix a VPL L, its syntactic Ext-algebra (R_L, O_L), and its syntactic morphism $(\varphi_L, \psi_L) : (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R_L, O_L)$.

Before we explain our proof strategy we introduce approximate matchings and horizontal and vertical evaluation languages. Approximate matchings generalize the classical matching relation on well-matched words with respect to our VPL L in the sense that they are subsets of the matching relation but must equal the matching relation on all those words that are in L. Approximate matchings in the context of visibly pushdown languages were introduced by Ludwig [24]. We then introduce suitably padded word languages mimicking the evaluation problem of the horizontal monoid R_L and the vertical monoid O_L, respectively.

Approximate matchings. For any word $w \in \Sigma^*$, we say that two positions $i, j \in [1, |w|]$ in w are *matched* whenever $i < j$, $w_i \in \Sigma_{\mathrm{call}}$, $w_j \in \Sigma_{\mathrm{ret}}$ and $w_{i+1} \cdots w_{j-1} \in \Sigma^{\triangle}$; we also say that i is

matched to j in w. Observe that a word w over Σ is well-matched if and only if for each position $i \in [1, |w|]$,

- if $i \in \Sigma_{\text{call}}$, then there exists a position $j \in [1, |w|]$ such that i is matched to j in w;
- if $i \in \Sigma_{\text{ret}}$, then there exists a position $j \in [1, |w|]$ such that j is matched to i in w.

Given a word $w \in \Sigma^{\triangle}$, we denote by $M^{\triangle}(w)$ its *matching relation* (or *matching*), that is the relation $\{(i,j) \in [1,|w|]^2 \mid i \text{ is matched to } j \text{ in } w\}$. An *approximate matching relative to* $L \subseteq \Sigma^{\triangle}$ is a function $M\colon \Sigma^* \to {\mathbb{N}_{>0}}^2$ such that $M(w) = M^{\triangle}(w)$ for all $w \in L$ and $M(w) \subseteq M^{\triangle}(w)$ for all $w \in \Sigma^* \setminus L$.

Horizontal and vertical evaluation languages. For all $k \in \mathbb{N}$, we define

$$\mathcal{O}(\Sigma^{\triangle})^{k,*} = \{\text{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle}) : |u| = k\} \quad \text{and} \quad \mathcal{O}(\Sigma^{\triangle})^{*,k} = \{\text{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle}) : |v| = k\} \ .$$

We also define $\mathcal{O}(\Sigma^{\triangle})_{\uparrow} = \{\text{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle}) \mid \Delta(u) > 0\}$ and finally for all $k \in \mathbb{N}$, we define

$$\mathcal{O}(\Sigma^{\triangle})_{\uparrow}^{k,*} = \mathcal{O}(\Sigma^{\triangle})^{k,*} \cap \mathcal{O}(\Sigma^{\triangle})_{\uparrow} \quad \text{and} \quad \mathcal{O}(\Sigma^{\triangle})_{\uparrow}^{*,k} = \mathcal{O}(\Sigma^{\triangle})^{*,k} \cap \mathcal{O}(\Sigma^{\triangle})_{\uparrow} \ .$$

Consider the alphabets $\Gamma_{\varphi_L} = \varphi_L(\Sigma^{\triangle} \setminus \{\varepsilon\}) \cup \{\$\}$ and $\Gamma_{\psi_L} = \psi_L\big(\mathcal{O}(\Sigma^{\triangle})_{\uparrow}\big) \cup \{\$\}$ for a letter $\$ \notin R_L \cup O_L$. We also define

$$\mathcal{V}_{\varphi_L} = \{\$^k s \mid k \in \mathbb{N}, s \in \varphi_L(\Sigma^{k+1})\}^* \text{ and } \mathcal{V}_{\psi_L} = \big\{\$^k f \;\big|\; k \in \mathbb{N}, f \in \psi_L\big(\mathcal{O}(\Sigma^{\triangle})_{\uparrow}^{k+1,*}\big)\big\}^* \ .$$

Define the φ_L-evaluation morphism $\text{eval}_{\varphi_L}\colon \Gamma_{\varphi_L}^* \to R_L$ by $\text{eval}_{\varphi_L}(s) = s$ for all $s \in \varphi_L(\Sigma^{\triangle} \setminus \{\varepsilon\})$ and $\text{eval}_{\varphi_L}(\$) = 1_R$. Similarly, define the ψ_L-evaluation morphism $\text{eval}_{\psi_L}\colon \Gamma_{\psi_L}^* \to O_L$ by $\text{eval}_{\psi_L}(f) = f$ for all $f \in \psi_L\big(\mathcal{O}(\Sigma^{\triangle})_{\uparrow}\big)$ and $\text{eval}_{\psi_L}(\$) = 1_{O_L}$. Finally, for all $r \in R_L$, we set

$$\mathcal{E}_{\varphi_L,r} = \mathcal{V}_{\varphi_L} \cap \text{eval}_{\varphi_L}^{-1}(r)$$

and for all $e \in O_L$, we set

$$\mathcal{E}_{\psi_L,e} = \mathcal{V}_{\psi_L} \cap \text{eval}_{\psi_L}^{-1}(e) \ .$$

5.3.1 Strategy for the proof of Theorem 5.7

We are now ready to give the proof strategy for Theorem 5.7. The proof consists of the following steps.

1. Lemma 5.9: $\mathcal{V}_{\varphi_L}$ and $\mathcal{V}_{\psi_L}$ are regular languages whose syntactic morphisms are quasi-aperiodic.
2. Proposition 5.10: Let L be a VPL whose syntactic morphism (φ_L, ψ_L) is quasi-aperiodic.
   - For all $r \in R_L$, the language $\mathcal{E}_{\varphi_L,r}$ is regular and its syntactic morphism is quasi-aperiodic.
   - For all $e \in O_L$, if for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ such that there exist $g, h \in O_L$ satisfying $e = g \circ f \circ h$ we have that $\mathcal{R}_f = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi_L(\text{ext}_{u,v}) = f\}$ is length-synchronous, then $\mathcal{E}_{\psi_L,e}$ is a regular language whose syntactic morphism is quasi-aperiodic.
3. Proposition 5.12: If $L \subseteq \Sigma^{\triangle}$ is length-synchronous, then there exists an $\mathsf{FO}_{\Sigma}[+]$-formula $\mu(x,y)$ such that $M\colon \Sigma^* \to {\mathbb{N}_{>0}}^2$ defined by $M(w) = \{(i,j) \in [1,|w|]^2 \mid w \models \mu(i,j)\}$ for all $w \in \Sigma^*$ is an approximate matching relative to L.

4. Proposition 5.16: Assume a VPL L has bounded nesting depth and

   - $\mathcal{E}_{\varphi_L,r}$ is a regular language whose syntactic morphism is quasi-aperiodic for all $\varphi_L(L)$-reachable $r \in R_L$, and
   - $\mathcal{E}_{\psi_L,e}$ is a regular language whose syntactic morphism is quasi-aperiodic for all $\varphi_L(L)$-reachable $e \in O_L$.

   Then there exists an $\mathsf{FO}_{\Sigma,\rightsquigarrow}[+]$-sentence η such that for any approximate matching M relative to L, we have $w \in L$ if, and only if, $(w, M(w)) \models \eta$ for all $w \in \Sigma^*$.

Let us argue that Points 2, 3 and 4 indeed imply Theorem 5.7 (Point 1 will be used in the proof of Point 2). Length-synchronicity of L implies weak length-synchronicity of L and thus bounded nesting depth of L (Proposition 4.14). Point 2 implies the other assumptions of Point 4: length-synchronicity of L means by definition that for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ we have that $\mathcal{R}_f = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi_L(\mathrm{ext}_{u,v}) = f\}$ is length-synchronous, so Point 2 implies that $\mathcal{E}_{\varphi_L,r}$ and $\mathcal{E}_{\psi_L,e}$ are quasi-aperiodic for all $r \in R_L$ and all $e \in O_L$, respectively. Finally, combining the $\mathsf{FO}_{\Sigma,\rightsquigarrow}[+]$-sentence of Point 4 with the $\mathsf{FO}_\Sigma[+]$-formula given by Point 3 that defines an approximate matching relative to L yields an $\mathsf{FO}_\Sigma[+]$-sentence defining L, thus proving Theorem 5.7.

5.3.2 $\mathcal{V}_{\varphi_L}$ and $\mathcal{V}_{\psi_L}$ are quasi-aperiodic (Proof of Point 1)

Before proving Point 1 in the proof strategy for Theorem 5.7 we require the following auxiliary lemma. It provides an important periodicity property of Ext-algebra morphisms.

Lemma 5.8. *The following periodicity holds:*

1. *There exist $t \in \mathbb{N}$ and $p \in \mathbb{N}_{>0}$ such that $\varphi_L(\Sigma^\triangle \cap \Sigma^i) = \varphi_L(\Sigma^\triangle \cap \Sigma^j)$ for all $i, j \in \mathbb{N}$ satisfying $i, j \geq t$ and $i \equiv j \pmod p$.*
2. *There exist $t \in \mathbb{N}$ and $p \in \mathbb{N}_{>0}$ such that $\psi_L\big(\mathcal{O}(\Sigma^\triangle)^{i,*}_\uparrow\big) = \psi_L\big(\mathcal{O}(\Sigma^\triangle)^{j,*}_\uparrow\big)$ for all $i, j \in \mathbb{N}$ satisfying $i, j \geq t$ and $i \equiv j \pmod p$.*
3. *There exist $t \in \mathbb{N}$ and $p \in \mathbb{N}_{>0}$ such that $\psi_L\big(\mathcal{O}(\Sigma^\triangle)^{*,i}_\uparrow\big) = \psi_L\big(\mathcal{O}(\Sigma^\triangle)^{*,j}_\uparrow\big)$ for all $i, j \in \mathbb{N}$ satisfying $i, j \geq t$ and $i \equiv j \pmod p$.*

Proof. To prove Point 1 recall that $\varphi_L^{-1}(r)$ is a VPL and hence a context-free language for all $r \in R_L$. By Parikh's Theorem [14, Section 3] it follows that $S_r = \{|w| : w \in \Sigma^\triangle, \varphi_L(w) = r\} \subseteq \mathbb{N}$ is a semilinear set for all $r \in R_L$. It follows that for all $U \subseteq R_L$ the set $S_U = \{|w| : w \in \Sigma^\triangle, \varphi_L(w) \in U\} \subseteq \mathbb{N}$ is semilinear since semilinear sets are closed under union. Point 1 follows immediately from this observation.

Next we prove Point 2, Point 3 can be proven analogously. According to Lemma 6.4 in Section 6 for $\# \notin \Sigma$ the language $L_e = \{u\#v \mid uv \in \Sigma^\triangle : \psi_L(\mathrm{ext}_{u,v}) = e\}$ is a VPL for all $e \in O_L$. As the language $K = \{u\#v \mid u, v \in \Sigma^\triangle\}$ is obviously a VPL, it follows that for all $e \in O_L$ the language

$$L_{e\uparrow} = L_e \setminus K = \{u\#v \mid uv \in \Sigma^\triangle : \psi_L(\mathrm{ext}_{u,v}) = e, \Delta(u) > 0\} \subseteq L_e$$

is a VPL as well. By Lemma 6.5 in Section 6 the set

$$S_e = \{(k,l) \in \mathbb{N} \times \mathbb{N} \mid \exists u \in \Sigma^k, v \in^l\colon u\#v \in L_{e\uparrow}\}$$

is semilinear as well for all $e \in O_L$. As a consequence we obtain that for all $Y \subseteq O_L$ the set

$$S_Y = \{(k,l) \in \mathbb{N} \times \mathbb{N} \mid \exists u \in \Sigma^k, v \in \Sigma^l, e \in Y : u\#v \in L_{e\uparrow}\} \subseteq \mathbb{N} \times \mathbb{N}$$

is semilinear as well since semilinear sets are closed under union. Since for all $Y \subseteq O_L$ the set $\{k \in \mathbb{N} \mid \psi_L(\mathcal{O}(\Sigma^{\triangle})^{k,*}_{\uparrow} = Y\}$ is nothing but the projection of S_Y onto the first component and semilinear sets are closed under projection, Point 2 follows. □

The following lemma holds irrespective of whether the syntactic morphism (φ_L, ψ_L) of L is quasi-aperiodic or not.

Lemma 5.9. *$\mathcal{V}_{\varphi_L}$, $\mathcal{V}_{\psi_L}$ are regular languages whose syntactic morphisms are quasi-aperiodic.*

Proof. Take $t \in \mathbb{N}$ and $p \in \mathbb{N}_{>0}$ given by Lemma 5.8 such that $\psi_L\big(\mathcal{O}(\Sigma^{\triangle})^{i,*}_{\uparrow}\big) = \psi_L\big(\mathcal{O}(\Sigma^{\triangle})^{j,*}_{\uparrow}\big)$ for all $i, j \in \mathbb{N}$ satisfying $i, j \geq t$ and $i \equiv j \pmod p$. Define $\theta_{t,p} \colon \mathbb{N} \to \mathbb{N}$ as

$$\theta_{t,p}(n) = \begin{cases} n & \text{if } n < t \\ \min\{n' \in \mathbb{N} \mid n' \geq t \wedge n' \equiv n \pmod p\} & \text{otherwise} \end{cases}$$

for all $n \in \mathbb{N}$. Take M to be the syntactic monoid of $\mathcal{V}_{\psi_L}$ and $h \colon \Gamma^*_{\psi_L} \to M$ to be its syntactic morphism.

If there exists $f \in \psi_L\big(\mathcal{O}(\Sigma^{\triangle})_{\uparrow}\big)$ and $k \in \mathbb{N}$ such that $\$^k f \notin \mathcal{V}_{\psi_L}$, let us fix some $f_\perp$ and $k_\perp$ that satisfy this. Observe that for all $k \in \mathbb{N}$, we have that $h(\$^k) = h(\$^{\theta_{t,p}(k)})$. Further, for all $n \in \mathbb{N}_{>0}$, $k_1, \ldots, k_{n+1} \in \mathbb{N}$ and $f_1, \ldots, f_n \in \psi_L\big(\mathcal{O}(\Sigma^{\triangle})_{\uparrow}\big)$, we have $h(\$^{k_1} f_1 \cdots \$^{k_n} f_n \$^{k_{n+1}}) = h(\$^{\theta_{t,p}(k_1)} f_1 \alpha \$^{\theta_{t,p}(k_{n+1})})$, where

$$\alpha = \begin{cases} \varepsilon & \text{if } \$^{k_2} f_2 \cdots \$^{k_n} f_n \in \mathcal{V}_{\psi_L} \\ \$^{k_\perp} f_\perp & \text{otherwise .} \end{cases}$$

Therefore, M is finite and thus $\mathcal{V}_{\psi_L}$ is regular. Let $l \in \mathbb{N}_{>0}$ be the stability index of h and take $q \in \mathbb{N}_{>0}$ such that $q \cdot l \geq t$ and $q \cdot l \equiv 0 \pmod p$. By definition, we have $h(\Gamma^l_{\psi_L}) = h(\Gamma^{q \cdot l}_{\psi_L})$. Thus, to show that h is quasi-aperiodic it is sufficient to prove that for all $m \in h(\Gamma^{q \cdot l}_{\psi_L})$, we have $m^2 = m^3$. Indeed, given $m \in h(\Gamma^{q \cdot l}_{\psi_L})$, only the following three cases can occur.

1. $m = h(\$^{q \cdot l})$. In this case, we have

$$m^2 = h(\$^{2 \cdot q \cdot l}) = h(\$^{\theta_{t,p}(2 \cdot q \cdot l)}) = h(\$^{\theta_{t,p}(q \cdot l)}) = h(\$^{q \cdot l}) = m \ ,$$

   where the third equality follows from $\theta_{t,p}(2 \cdot q \cdot l) = \theta_{t,p}(q \cdot l)$.

2. $m = h(\$^{k_1} f \$^{k_\perp} f_\perp \$^{k_2})$ for $f \in \psi_L\big(\mathcal{O}(\Sigma^{\triangle})_{\uparrow}\big)$ and $k_1, k_2 \in \mathbb{N}$ satisfying $\theta_{t,p}(k_1) = k_1$ and $\theta_{t,p}(k_2) = k_2$. In this case, we have

$$m^2 = h(\$^{k_1} f \$^{k_\perp} f_\perp \$^{k_1 + k_2} f \$^{k_\perp} f_\perp \$^{k_2}) = h(\$^{k_1} f \$^{k_\perp} f_\perp \$^{k_2}) = m \ ,$$

   where the second equality follows from $\$^{k_\perp} f_\perp \$^{k_1 + k_2} f \$^{k_\perp} f_\perp \notin \mathcal{V}_{\psi_L}$.

3. $m = h(\$^{k_1} f \$^{k_2})$ for $f \in \psi_L\big(\mathcal{O}(\Sigma^{\triangle})_{\uparrow}\big)$ and $k_1, k_2 \in \mathbb{N}$ satisfying $\theta_{t,p}(k_1) = k_1$ and $\theta_{t,p}(k_2) = k_2$. If $f \in \psi_L\big(\mathcal{O}(\Sigma^{\triangle})^{k_1 + k_2 + 1,*}_{\uparrow}\big)$, then

$$m^2 = h(\$^{k_1} f \$^{k_1 + k_2} f \$^{k_2}) = h(\$^{k_1} f \$^{k_2}) = m$$

   because $\$^{k_1 + k_2} f \in \mathcal{V}_{\psi_L}$. Otherwise, $f \notin \psi_L\big(\mathcal{O}(\Sigma^{\triangle})^{k_1 + k_2 + 1,*}_{\uparrow}\big)$ and then

$$m^2 = h(\$^{k_1} f \$^{k_1 + k_2} f \$^{k_2}) = h(\$^{k_1} f \$^{k_\perp} f_\perp \$^{k_2})$$

because $\$^{k_1+k_2} f \notin \mathcal{V}_{\psi_L}$, so

$$m^3 = h(\$^{k_1} f \$^{k_\perp} f_\perp \$^{k_1+k_2} f \$^{k_2}) = h(\$^{k_1} f \$^{k_\perp} f_\perp \$^{k_2}) = m^2$$

because $\$^{k_\perp} f_\perp \$^{k_1+k_2} f \notin \mathcal{V}_{\psi_L}$.

Therefore, $\mathcal{V}_{\psi_L}$ is a regular language whose syntactic morphism is quasi-aperiodic. □

5.3.3 Quasi-aperiodicity of evaluation languages $\mathcal{E}_{\varphi_L,r}$ and $\mathcal{E}_{\psi_L,e}$ (Proof of Point 2)

One important consequence of Lemma 5.9 is that for all $r \in R_L$ and $e \in O_L$, the languages $\mathcal{E}_{\varphi_L,r}$ and $\mathcal{E}_{\psi_L,e}$ are in fact regular languages. The following proposition gives conditions under which those languages have moreover quasi-aperiodic syntactic morphisms when the syntactic morphism of L itself is.

Proposition 5.10. *Let L be a VPL whose syntactic morphism (φ_L, ψ_L) is quasi-aperiodic.*

- *For all $r \in R_L$, the language $\mathcal{E}_{\varphi_L,r}$ is regular and its syntactic morphism is quasi-aperiodic.*
- *For all $e \in O_L$, if for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ such that there exist $g, h \in O_L$ satisfying $e = g \circ f \circ h$[1] we have that $\mathcal{R}_f = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi_L(\mathrm{ext}_{u,v}) = f\}$ is length-synchronous, then $\mathcal{E}_{\psi_L,e}$ is a regular language whose syntactic morphism is quasi-aperiodic.*

Proof. We already know that for all $r \in R_L$ and $e \in O_L$, the languages $\mathcal{E}_{\varphi_L,r}$ and $\mathcal{E}_{\psi_L,e}$ are regular. To prove the lemma, we then just have to prove that

- if there exists $r \in R_L$ such that the syntactic morphism of $\mathcal{E}_{\varphi_L,r}$ is not quasi-aperiodic, then there exists $k \in \mathbb{N}$ such that $\varphi(\Sigma^\triangle \cap \Sigma^k)$ contains a semigroup that is not aperiodic;
- if there exists $e \in O_L$ such that the syntactic morphism of $\mathcal{E}_{\psi_L,e}$ is not quasi-aperiodic and for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ such that there exist $f', f'' \in O_L$ satisfying $e = f' \circ f \circ f''$ we have that $\mathcal{R}_f = \{(u,v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\triangle, \Delta(u) > 0, \psi_L(\mathrm{ext}_{u,v}) = f\}$ is length-synchronous, then there exist $k, l \in \mathbb{N}$ such that $\psi_L(\mathcal{O}(\Sigma^\triangle)^{k,l})$ contains a semigroup that is not aperiodic.

Indeed, the first point allows to conclude that (φ_L, ψ_L) is not quasi-aperiodic, since if there exists a non-aperiodic semigroup S contained in $\varphi_L(\Sigma^\triangle \cap \Sigma^k)$, then $\{\mathrm{left}_s \mid s \in S\}$ is a semigroup contained in $\psi_L(\mathcal{O}(\Sigma^\triangle)^{k,0})$ (because for each $s \in S$, there exists $w \in \Sigma^\triangle \cap \Sigma^k$ satisfying $\varphi_L(w) = s$, so that $\psi_L(\mathrm{ext}_{w,\varepsilon}) = \mathrm{left}_{\varphi_L(w)} = \mathrm{left}_s$). But this semigroup is non-aperiodic as well, since as S is non-aperiodic, it must be that for all $i \in \mathbb{N}_{>0}$, there exists $s \in S$ such that $s^i \neq s^{i+1}$, so that ${\mathrm{left}_s}^i \neq {\mathrm{left}_s}^{i+1}$.

We only prove the second point, the first point can be proved in a similar way by leaving out the last paragraph of the following proof, that is the sole place where we need, given an $e \in O_L$, length-synchronicity of $\mathcal{R}_f$ for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ such that there exist $f', f'' \in O_L$ satisfying $e = f' \circ f \circ f''$.

Take $t \in \mathbb{N}$ and $p \in \mathbb{N}_{>0}$ given by Lemma 5.8 such that $\psi_L\big(\mathcal{O}(\Sigma^\triangle)^{i,*}_\uparrow\big) = \psi_L\big(\mathcal{O}(\Sigma^\triangle)^{j,*}_\uparrow\big)$ for all $i, j \in \mathbb{N}$ satisfying $i, j \geq t$ and $i \equiv j \pmod p$.

Assume there exists $e \in O_L$ such that the syntactic morphism of $\mathcal{E}_{\psi_L,e}$ is not quasi-aperiodic. Take M to be the syntactic monoid of $\mathcal{E}_{\psi_L,e}$ and $h \colon \Gamma^*_{\psi_L} \to M$ to be its syntactic morphism. Let

[1] In the terminology of Green's relations, to be introduced in the proof of Proposition 5.24, this is equivalent to the fact that $e \leq_\mathfrak{J} f$.

$s \in \mathbb{N}_{>0}$ be the stability index of h and let $\omega \geq 2$ be a multiple both of the idempotent power of M and the idempotent power of O_L. Non-quasi-aperiodicity of h implies that there exists $g \in h(\Gamma_{\psi_L}^s)$ satisfying $g^\omega \neq g^{\omega+1}$.

By definition of the stability index, there exists $w \in \Gamma_{\psi_L}^{q \cdot s}$ for $q \in \mathbb{N}_{>0}$ such that $q \cdot s \geq t$ and $q \cdot s \equiv 0 \pmod{p}$ satisfying $h(w) = g$. Since $t \leq q \cdot s \leq q \cdot s \cdot \omega \leq q \cdot s \cdot (\omega+1)$ and $q \cdot s \equiv q \cdot s \cdot \omega \equiv q \cdot s \cdot (\omega+1) \pmod{p}$, we cannot have $w = \$^{q \cdot s}$, for otherwise we would have $g^\omega = h(\$^{q \cdot s \cdot \omega}) = h(\$^{q \cdot s \cdot (\omega+1)}) = g^{\omega+1}$ because $\psi_L\big(\mathcal{O}(\Sigma^\triangle)_\uparrow^{q \cdot s \cdot \omega + k + 1, *}\big) = \psi_L\big(\mathcal{O}(\Sigma^\triangle)_\uparrow^{q \cdot s \cdot (\omega+1) + k + 1, *}\big)$ for all $k \in \mathbb{N}$. Therefore, we have $w = \$^{k_1} f_1 \cdots \$^{k_n} f_n \$^{k_{n+1}}$ for $n \in \mathbb{N}_{>0}$, $k_1, \ldots, k_{n+1} \in \mathbb{N}$ and $f_1, \ldots, f_n \in \psi_L\big(\mathcal{O}(\Sigma^\triangle)_\uparrow\big)$. Since $g^\omega \neq g^{\omega+1}$, there exist $x, y \in \Gamma_{\psi_L}^*$ such that either $xw^\omega y \in \mathcal{E}_{\psi_L, e}$ and $xw^{\omega+1}y \notin \mathcal{E}_{\psi_L, e}$, or $xw^\omega y \notin \mathcal{E}_{\psi_L, e}$ and $xw^{\omega+1}y \in \mathcal{E}_{\psi_L, e}$. Assume the first case holds. Then we have $x = x'\$^{k_x}$ and $y = \$^{k_y} y'$ with $k_x, k_y \in \mathbb{N}$ and $x', y' \in \Gamma_{\psi_L}^*$ satisfying $x', \$^{k_x + k_1} f_1, \$^{k_2} f_2, \ldots, \$^{k_n} f_n, \$^{k_{n+1} + k_1} f_1, \$^{k_{n+1} + k_y} y' \in \mathcal{V}_{\psi_L}$ and $\mathrm{eval}_{\psi_L}(x') \circ (f_1 \circ \cdots \circ f_n)^\omega \circ \mathrm{eval}_{\psi_L}(y') = e$. Therefore, we also have $xw^{\omega+1}y \in \mathcal{V}_{\psi_L}$, hence since $xw^{\omega+1}y \notin \mathcal{E}_{\psi_L, e}$ we necessarily have

$$\begin{aligned} e &= \mathrm{eval}_{\psi_L}(x') \circ (f_1 \circ \cdots \circ f_n)^\omega \circ \mathrm{eval}_{\psi_L}(y') \\ &\neq \mathrm{eval}_{\psi_L}(xw^{\omega+1}y) = \mathrm{eval}_{\psi_L}(x') \circ (f_1 \circ \cdots \circ f_n)^{\omega+1} \circ \mathrm{eval}_{\psi_L}(y') \ . \end{aligned}$$

Thus we have $(f_1 \circ \cdots \circ f_n)^\omega \neq (f_1 \circ \cdots \circ f_n)^{\omega+1}$ and $\$^{k_{n+1}+k_1} f_1 \$^{k_2} f_2 \cdots \$^{k_n} f_n \in \mathcal{V}_{\psi_L}$. This is also true for the case when $xw^\omega y \notin \mathcal{E}_{\psi_L, e}$ and $xw^{\omega+1}y \in \mathcal{E}_{\psi_L, e}$.

Therefore, we have $(f_1 \circ \cdots \circ f_n)^\omega \neq (f_1 \circ \cdots \circ f_n)^{\omega+1}$ with $(f_1 \circ \cdots \circ f_n)^i \in \psi_L\big(\mathcal{O}(\Sigma^\triangle)_\uparrow^{q \cdot s \cdot i, *}\big) = \psi_L\big(\mathcal{O}(\Sigma^\triangle)_\uparrow^{q \cdot s, *}\big)$ for each $i \in \mathbb{N}_{>0}$ because $k_{n+1} + k_1 + \cdots + k_n + n = q \cdot s \geq t$ and $k_{n+1} + k_1 + \cdots + k_n + n = q \cdot s \equiv 0 \pmod{p}$. But given ω' the idempotent power of $\{(f_1 \circ \cdots \circ f_n)^i \mid i \in \mathbb{N}_{>0}\}$, we have that $(f_1 \circ \cdots \circ f_n)^\omega = (f_1 \circ \cdots \circ f_n)^{\omega\omega'} = (f_1 \circ \cdots \circ f_n)^{\omega'}$, so that $(f_1 \circ \cdots \circ f_n)^{\omega'} \neq (f_1 \circ \cdots \circ f_n)^{\omega'+1}$, hence $\{(f_1 \circ \cdots \circ f_n)^i \mid i \in \mathbb{N}_{>0}\}$ is not aperiodic.

Assume additionally that for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ such that there exist $f', f'' \in O_L$ satisfying $e = f' \circ f \circ f''$ we have that $\mathcal{R}_f = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\triangle, \Delta(u) > 0, \psi_L(\mathrm{ext}_{u,v}) = f\}$ is length-synchronous. For each $i \in \mathbb{N}_{>0}$, let $\mathrm{ext}_{u_i, v_i} \in \psi_L\big(\mathcal{O}(\Sigma^\triangle)_\uparrow^{q \cdot s, *}\big)$ such that $\psi_L(\mathrm{ext}_{u_i, v_i}) = (f_1 \circ \cdots \circ f_n)^i$. If $(f_1 \circ \cdots \circ f_n)^\omega$ were not $\varphi_L(L)$-reachable, then it would imply that $\mathrm{ext}_{u_\omega, v_\omega}$ is not L-reachable. This would in turn entail that for all $z \in \Sigma^\triangle$ and $\mathrm{ext}_{\alpha,\beta} \in \mathcal{O}(\Sigma^\triangle)$ we have

$$\mathrm{ext}_{\alpha,\beta}(\mathrm{ext}_{u_\omega, v_\omega}(z)) \notin L \wedge \mathrm{ext}_{\alpha,\beta}(\mathrm{ext}_{u_1 u_\omega, v_\omega v_1}(z)) = \mathrm{ext}_{\alpha u_1, v_1 \beta}(\mathrm{ext}_{u_\omega, v_\omega}(z)) \notin L \ ,$$

so that it would follow that $(f_1 \circ \cdots \circ f_n)^\omega = \psi_L(\mathrm{ext}_{u_\omega, v_\omega}) = \psi_L(\mathrm{ext}_{u_1 u_\omega, v_\omega v_1}) = (f_1 \circ \cdots \circ f_n)^{\omega+1}$, a contradiction. Hence, since $(f_1 \circ \cdots \circ f_n)^\omega$ is a $\varphi_L(L)$-reachable idempotent and

$$e = \begin{cases} \mathrm{eval}_{\psi_L}(x') \circ (f_1 \circ \cdots \circ f_n)^\omega \circ \mathrm{eval}_{\psi_L}(y') & \text{if } xw^\omega y \in \mathcal{E}_{\psi_L, e} \text{ and } xw^{\omega+1}y \notin \mathcal{E}_{\psi_L, e} \\ \mathrm{eval}_{\psi_L}(x') \circ (f_1 \circ \cdots \circ f_n)^{\omega+1} \circ \mathrm{eval}_{\psi_L}(y') & \text{otherwise } (xw^\omega y \notin \mathcal{E}_{\psi_L, e} \text{ and } xw^{\omega+1}y \in \mathcal{E}_{\psi_L, e}) \end{cases} \ ,$$

it follows that $\mathcal{R}_{(f_1 \circ \cdots \circ f_n)^\omega}$ is length-synchronous. So for all $i \in \mathbb{N}, i \geq 2$, since $\psi_L(\mathrm{ext}_{u_1^\omega, v_1^\omega}) = \psi_L(\mathrm{ext}_{u_i^\omega, v_i^\omega}) = (f_1 \circ \cdots \circ f_n)^\omega$ with $\Delta(u_1^\omega) > 0$ and $\Delta(u_i^\omega) > 0$, since $|u_1| = |u_i|$, we have

$$\frac{|u_1^\omega|}{|v_1^\omega|} = \frac{|u_i^\omega|}{|v_i^\omega|} \Rightarrow \frac{|u_1|}{|v_1|} = \frac{|u_i|}{|v_i|} \Rightarrow |v_1| = |v_i| \ .$$

To conclude, we obtain that the non-aperiodic semigroup $\{(f_1 \circ \cdots \circ f_n)^i \mid i \in \mathbb{N}_{>0}\}$ is contained in $\psi_L(\mathcal{O}(\Sigma^\triangle)^{q \cdot s, |v_1|})$. □

The following remark states that the length-synchronicity precondition in the second point of Proposition 5.10 is important. In fact it shows that weak length-synchronicity is not sufficient.

Remark 5.11. *For the second point of Proposition 5.10 it is generally not sufficient to assume, given e, that for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ such that there exist $g, h \in O_L$ satisfying $e = g \circ f \circ h$ we have that $\mathcal{R}_f = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\triangle, \Delta(u) > 0, \psi_L(\mathrm{ext}_{u,v}) = f\}$ is weakly length-synchronous. Indeed, the VPL K generated by the grammar with rules*

$$S \to aSb_1 \mid acTb_2 \mid \varepsilon$$
$$T \to aTb_1 \mid acSb_2 \ .$$

using S as start symbol is not length-synchronous, but weakly length-synchronous, and has a quasi-aperiodic syntactic morphism. However, for the syntactic Ext*-algebra (R_K, O_K) and the syntactic morphism (φ_K, ψ_K) of K, we claim that there exists a $\varphi_K(K)$-reachable $e \in O_K$ such that $\mathcal{E}_{\psi_K,e}$ is a regular language whose syntactic morphism is not quasi-aperiodic while, as K is weakly length-synchronous, for each $\varphi_L(L)$-reachable idempotent $f \in O_L$ we have that $\mathcal{R}_f = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\triangle, \Delta(u) > 0, \psi_L(\mathrm{ext}_{u,v}) = f\}$ is weakly length-synchronous.*

Let Γ be the visibly pushdown alphabet of K. Note that we have $K \subset \mathcal{L}_{1,2}$, where $\mathcal{L}_{1,2} = L(S \to aSb_1|acSb_2|\varepsilon)$ is the VPL initially introduced in Example 2.5. For all $uv, u'v' \in \mathcal{L}_{1,2}$ with $u, u' \in \{a, c\}^+$, $v, v' \in \{b_1, b_2\}^+$, $|u|_c \equiv |u'|_c \pmod 2$ we have $xuzvy \in K \Leftrightarrow xu'zv'y \in K$ for all $xy, z \in \Gamma^\triangle$. This implies that if we set $e_0 = \psi_K(\mathrm{ext}_{a,b_1})$ and $e_1 = \psi_K(\mathrm{ext}_{ac,b_2})$, we have that for all $uv \in \mathcal{L}_{1,2}$ with $u \in \{a, c\}^+$, $v \in \{b_1, b_2\}^+$, it holds that $\psi_K(\mathrm{ext}_{u,v}) = e_{|u|_c \bmod 2}$. Therefore, while $e_0 \neq e_1$, we have $e_0 \circ e_1 = e_1 \circ e_0 = e_1$ and $e_0 \circ e_0 = e_1 \circ e_1 = e_0$.

Consider the length-multiplying monoid morphism $\beta\colon \{0,1\}^* \to \Gamma_{\psi_K}^*$ such that $\beta(0) = e_0e_0$ and $\beta(1) = \$e_1$. Then $\mathrm{MOD}_2 = \beta^{-1}(\mathcal{E}_{\psi_K,e_0})$, so $\mathcal{E}_{\psi_K,e_0}$ cannot have a quasi-aperiodic syntactic morphism, for otherwise, by closure of the class of regular languages whose syntactic morphism is quasi-aperiodic under inverses of length-multiplying morphisms (see [30]), we would have that MOD_2 *has a quasi-aperiodic syntactic morphism.*

5.3.4 Approximate matching relation in FO[+] (Proof of Point 3)

The following proposition states that there is a $\mathsf{FO}_\Sigma[+]$-definable approximate matching relative to any length-synchronous visibly pushdown language.

Proposition 5.12. *If $L \subseteq \Sigma^\triangle$ is length-synchronous, then there exists an $\mathsf{FO}_\Sigma[+]$-formula $\eta(x, y)$ such that $M\colon \Sigma^* \to {\mathbb{N}_{>0}}^2$ defined by $M(w) = \{(i, j) \in [1, |w|]^2 \mid w \models \eta(i, j)\}$ for all $w \in \Sigma^*$ is an approximate matching relative to L.*

The technical heart of the proof is the following lemma whose proof is postponed and will take most part of this subsubsection. This lemma realizes the characterization of length-synchronicity given by Proposition 4.11 via an $\mathsf{FO}_\Sigma[+]$-formula.

Lemma 5.13. *Assume that (φ_L, ψ_L) is weakly length-synchronous. Let $e \in O_L$ be $\varphi_L(L)$-reachable and assume that $\mathcal{U}_e = \{(u, v) \in \mathsf{Con}(\Sigma) \mid e \circ \psi_L(\mathrm{ext}_{u,v}) = e\}$ is length-synchronous. Then there exists an $\mathsf{FO}_\Sigma[+]$-formula $\pi_e(x, x', y', y)$ such that for all $w \in \Sigma^+$ and $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$ the following holds,*

- *if $w \models \pi_e(i, i', j', j)$, then $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\triangle$ and*
- *if $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\triangle$ and $(w_i \ldots w_{i'}, w_{j'} \ldots w_j) \in \mathcal{U}_e$, then $w \models \pi_e(i, i', j', j)$.*

Building an approximate matching assuming predicates π_e. Let us prove Proposition 5.12 by making use of Lemma 5.13.

Proof of Proposition 5.12. By assumption (φ_L, ψ_L) is $\varphi_L(L)$-length-synchronous. Thus, the set of contexts $\mathcal{U}_e$ is length-synchronous for all $\varphi_L(L)$-reachable $e \in O_L$ by Proposition 4.10. Moreover, there exists $d_L \in \mathbb{N}$ bounding the nesting depth of the words in L Proposition 4.14. For defining our desired formula μ, we will construct $\mathsf{FO}_\Sigma[+]$ formulas μ_d and $\mu_d^\uparrow$ for all $0 \leq d \leq d_L$ with the following properties: for all $w \in \Sigma^+$ and for all $i, j \in [1, |w|]$, we have

- if $w \models \mu_d^\uparrow(i,j)$ or $w \models \mu_d(i,j)$, then $w_i \cdots w_j \in \Sigma^\triangle$,
- if $w \in L$, $w_i \dots w_j \in \Sigma^\triangle$, $\mathrm{nd}(w_i \dots w_j) \leq d$ and i is matched to j in w, then $w \models \mu_d^\uparrow(i,j)$, and
- if $w \in L$, $\mathrm{nd}(w_i \dots w_j) \leq d$ and $w_i \dots w_j \in \Sigma^\triangle$, then $w \models \mu_d(i,j)$.

We therefore define $\mu = \mu_{d_L}$. The construction of $\mu_d^\uparrow$ and μ_d is by induction on d. We set

$$\mu_0(i,j) = \bot \text{ and } \mu_0(i,j) = \forall z \big(x \leq z \leq y \to \Sigma_{\mathrm{int}}(z) \big).$$

Let us assume $d > 0$. The formula μ_d is easily defined assuming $\mu_d^\uparrow$. We define

$$\mu_d(x,y) \;=\; \forall z \Big[x \leq z \leq y \to \Big(\Sigma_{\mathrm{int}}(z) \vee \exists z' \Big((\Sigma_{\mathrm{call}}(z) \wedge \Sigma_{\mathrm{ret}}(z') \wedge \mu_d^\uparrow(z,z')) \vee (\Sigma_{\mathrm{call}}(z') \wedge \Sigma_{\mathrm{ret}}(z) \wedge \mu_d^\uparrow(z',z)) \Big) \Big) \Big].$$

It remains to define $\mu_d^\uparrow$. Let us assume $u = w_i \dots w_j \in \Sigma^\triangle$, that i is matched to j in w and that $\mathrm{nd}(u) = d > 0$. Hence, $u = a_1 v b_1 \in \Sigma^\triangle$ for some $a_1 \in \Sigma_{\mathrm{call}}$, $b_1 \in \Sigma_{\mathrm{ret}}$, and $v \in \Sigma^\triangle$. We then apply Lemma 4.17 which states that u has a nesting-maximal stair factorization $u = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \dots \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(u')$ such that for some $h \in [1,k]$, setting $u_\ell = \mathrm{ext}_{a_\ell,b_\ell} \circ \mathrm{ext}_{x_{\ell+1},y_{\ell+1}} \circ \cdots \circ \mathrm{ext}_{a_k,b_k}(u')$ for all $\ell \in [1,k]$ and $u_{k+1} = u'$, we have

1. $\mathrm{nd}(u) = \mathrm{nd}(u_h) = d$,
2. $\mathrm{nd}(u_{h+1}) = d - 1$, and
3. $\mathrm{nd}(x_1), \mathrm{nd}(y_1), \dots, \mathrm{nd}(x_h), \mathrm{nd}(y_h) < d$.

We remark that $x_1 = y_1 = \varepsilon$. Let $i = i_1 < \cdots < i_h$ and $j_h < \cdots < j_1 = j$ be the positions that correspond to the positions of the letters $a_1, \dots, a_h \in \Sigma_{\mathrm{call}}$ and $b_1, \dots, b_h \in \Sigma_{\mathrm{ret}}$ of u in w, respectively: more precisely $i_\ell = i + |x_1 \cdots a_{\ell-1} x_\ell|$ and $j_\ell = |x_1 a_1 \cdots x_k a_k u' b_k y_k \cdots b_{\ell+1} y_{\ell+1}| + 1$ for all $\ell \in [1,h]$. The formula $\eta_d^\uparrow$ could guess the positions $i = i_1 < \cdots < i_h$ and $j_h < \cdots < j_1 = j$ and verify the following (recalling that $x_1 = y_1 = \varepsilon$):

(a) the infix $w_{i_h+1} \cdots w_{j_h-1} = \mathrm{ext}_{x_{h+1},y_{h+1}}(u_{h+1})$ is well-matched, and

(b) the word $w_{i_1} \cdots w_{i_h} w_{j_h} \cdots w_{j_1}$ is well-matched.

Point (a) can be realized via the formula μ_{d-1} by making use of Point 2 from above, whereas Point (b) can be realized by the following ad-hoc formula, this time making use of Point 3 and from above:

$$\begin{aligned}\kappa_h(x,x',y',y) \;=\; & \exists x_1 \cdots x_h \exists y_1 \cdots y_h \Bigg(x = x_1 \wedge x' = x_h \wedge y' = y_h \wedge y = y_h \wedge \\ & \bigwedge_{t=1}^{h} \Sigma_{\mathrm{call}}(x_t) \wedge \Sigma_{\mathrm{ret}}(y_t) \wedge \bigwedge_{t=1}^{h-1} \mu_{d-1}(x_t+1, x_{t+1}-1) \wedge \mu_{d-1}(y_{t+1}+1, y_t-1) \Bigg)\end{aligned}$$

The problem with this approach is that the size of the formula depends on the size of w. For instance, for $a \in \Sigma_{\mathrm{call}}$, $b \in \Sigma_{\mathrm{ret}}$, $c \in \Sigma_{\mathrm{int}}$, and $u = a^n c b^n$ we have $\mathrm{nd}(u) = \mathrm{nd}(acb) = 1$ for all $n \geq 1$. Hence we would have $h = n-1$, so h would depend on u which is problematic. Therefore, towards expressing Point (b) by a formula whose size only depends on $|O_L|$, let us define, for all $\ell, \ell' \in [1,h]$, the product

$$e_{\ell,\ell'} = \psi_L(\mathrm{ext}_{x_\ell, y_\ell} \circ \mathrm{ext}_{a_\ell, b_\ell} \cdots \mathrm{ext}_{x_{\ell'}, y_{\ell'}} \circ \mathrm{ext}_{a_{\ell'}, b_{\ell'}}) \quad \text{and} \quad e_\ell = e_{1,\ell}.$$

We remark that all $e_{\ell,\ell'}$ are $\varphi_L(L)$-reachable since w is assumed to be in L. For $e \in O_L$ we say an interval $I = [s,t] \subseteq [1,h]$ is *e-repetitive* if $s < t$ and $e_s = e_t$. We say $[s,t] \subseteq [1,h]$ is *repetitive* if it is e-repetitive for some $e \in O_L$.

Claim 5.14. *There exist indices $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_q < t_q \leq s_{q+1} = h$ such that $[s_1,t_1], \ldots, [s_q,t_q]$ are all repetitive and for $D_0 = [t_0,s_1], D_1 = [t_1,s_2], \ldots, D_q = [t_q, s_{q+1}]$ we have $q + \sum_{p=0}^{q} |D_p| \leq 3|O_L|$.*

Proof of the Claim. For all $z \in [1,h]$ let $\lambda(z) = \max\{\ell \in [1,h] \mid e_\ell = e_z\}$. Observe $\lambda(z) \geq z$ for all $z \in [1,h]$ and that $|\lambda([1,h])| \leq |O_L|$. We define $t_0 = 1$. Let $p > 0$ and assume that we have already defined t_{p-1}. In case $t_{p-1} = h$ we are done and define $q = p-1$ and $s_{q+1} = h$. So let us assume $t_{p-1} < h$. In case there exists $z \in [t_{p-1}, h]$ such that $z < \lambda(z)$ we define $s_p = \min\{z \in [t_{p-1},h] \mid z < \lambda(z)\}$ and $t_p = \lambda(s_p)$, otherwise (i.e. in case $z = \lambda(z)$ for all $z \in [t_{p-1},h]$) we are done and define $q = p-1$ and $s_{q+1} = h$. Immediately by definition we have $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_q < t_q \leq s_{q+1} = h$ (because if we had $t_{p-1} = s_p$ for a $p \in [2,q]$, we would have $e_{s_{p-1}} = e_{t_{p-1}} = e_{s_p}$, so $\lambda(s_p) = \lambda(s_{p-1}) = t_{p-1} = s_p < \lambda(s_p)$, a contradiction) and $e_{s_p} = e_{t_p}$ for all $p \in [1,q]$. Moreover, the intervals $[s_1,t_1], \ldots, [s_q,t_q]$ are indeed all repetitive. Since moreover $t_p \in \lambda([1,h])$ for all $p \in [1,q]$ and $|\lambda([1,h])| \leq |O_L|$ we must have $q \leq |O_L|$. Now let $D_0 = [t_0,s_1], D_1 = [t_1,s_2], \ldots, D_q = [t_q, s_{q+1}]$. Clearly, these sets are pairwise disjoint. Moreover, by construction, the only elements $z \in \bigcup_{p=0}^{q} D_p$ such that $z < \lambda(z)$ are those in $X = \{s_1, \ldots, s_q\}$, so that all elements $z \in (\bigcup_{p=0}^{q} D_p) \setminus X$ satisfy $z = \lambda(z)$, i.e. are elements from $\lambda([1,h])$. Thus, we obtain $q + \sum_{p=0}^{q} |D_p| = q + \left|\bigcup_{p=0}^{q} D_p\right| \leq |O_L| + |X| + |\lambda([1,h])| = 3\,|O_L|$. □

Let $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_q < t_q \leq s_{q+1} = h$ be the indices satisfying Claim 5.14 along with $D_0 = [t_0,s_1], D_1 = [t_1,s_2], \ldots, D_q = [t_q, s_{q+1}]$. Let $d_p = |D_p|$ for all $p \in [0,q]$. Since, for all $p \in [1,q]$, the non-empty interval $[s_p,t_p]$ is repetitive, we have $e_{s_p} = e_{t_p}$ and thus obtain

$$e_{s_p} = e_{t_p} = e_{s_p} \circ \psi_L(\mathrm{ext}_{x_{s_p+1} \cdots a_{t_p}, b_{t_p} \cdots y_{s_p+1}}).$$

Hence, we have $w \models \pi_{e_{s_p}}(i_{s_p}+1, i_{t_p}, j_{t_p}, j_{s_p}-1)$ where $\pi_{e_{s_p}}$ is the formula given by Lemma 5.13 (recall that e_{s_p} is $\varphi_L(L)$-reachable and that $\mathcal{U}_{e_{s_p}}$ is length-synchronous). We can therefore use the

formula $\pi_{e_{s_p}}$ to witness that $w_{i_{s_p}+1} \cdots w_{i_{t_p}} w_{j_{t_p}} \cdots w_{j_{s_p}-1}$ is indeed a well-matched word. It will thus remain to verify that $w_{i_{t_p}} \cdots w_{i_{s_{p+1}}} w_{j_{s_{p+1}}} \cdots w_{i_{t_p}}$ is well-matched for all $p \in [0,q]$: this can be guaranteed by evaluating $\kappa_{d_p}(i_{t_p}, i_{s_{p+1}}, j_{s_{p+1}}, i_{t_p})$. We can now define our final formula $\mu_d^{\uparrow}$:

$$\begin{aligned}
\mu_d^{\uparrow}(x,y) \quad = \quad & \bigvee_{\substack{q \in [0,|O_L|] \\ d_0,\ldots,d_q \geq 1: \\ q+d_0+\cdots+d_q \leq 3|0|}} \exists x_1 \ldots x_{q+1} \exists x'_0 \ldots x'_q \exists y_1 \ldots y_{q+1} \exists y'_1 \ldots y'_0 \\
& \Bigg[x_0 \leq x_1 < x'_1 < x_2 < \cdots < x'_q < y'_q < y_q < \cdots < y'_1 < y_1 \leq y'_0 \wedge \\
& \quad x'_0 = x \wedge y'_0 = y \wedge \mu_{d-1}(x_{q+1}+1, y_{q+1}-1) \wedge \\
& \quad \bigwedge_{p=1}^{q} \left(\bigvee_{e \in O_L \ \varphi_L(L)\text{-reachable}} \pi_e(x_p+1, x'_p, y'_p, y_p-1) \right) \wedge \bigwedge_{p=0}^{q} \kappa_{d_p}(x'_p, x_{p+1}, y_{p+1}, y'_p) \Bigg] \ . \quad \square
\end{aligned}$$

The following remark is obvious but will be important in Section 5.4.

Remark 5.15. *When constructing our predicate $\mu_d^{\uparrow}$, we could have replaced any subset of the predicates π_e, where e is $\varphi_L(L)$-reachable from above, by the predicate π_e^{exact} expressing that for all $w \in \Sigma^+$ and $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$ it holds:*

$$\begin{aligned}
w \models \pi_e^{\mathsf{exact}}(i,i',j',j) \quad \iff \quad & w_i \ldots w_{i'} w_{j'} \ldots w_j \in \Sigma^{\triangle}, e \circ \psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e, \text{ and} \\
& \Delta(w_i \cdots w_{i'}) > 0
\end{aligned}$$

It remains to prove Lemma 5.13.

Proof of Lemma 5.13

In essence, our proof is inspired by the approach taken in [24, Proof of Proposition 126], which is itself a flawed adaptation (we refer to Section 8 for more details) of the approach taken in [23, Proof of Lemma 15].

Let $\alpha_e \in \mathbb{Q}_{>0}$, $\beta_e \in \mathbb{N}$ and $\gamma_e \in \mathbb{N}_{>0}$ given by Proposition 4.11 for e. There exist unique $n_e, d_e \in \mathbb{N}_{>0}$ that are relatively prime such that $\alpha_e = \frac{n_e}{d_e}$. We are going to build an $\mathsf{FO}[+]$-formula $\pi_e(x,x',y',y)$ such that for all $w \in \Sigma^+$ and $i, i', j', j \in [1,|w|], i \leq i' < j' \leq j$, we have that $w \models \pi_e(i,i',j',j)$ if, and only if, all of the following conditions are satisfied:

(i) $\frac{i'-i+1}{j-j'+1} = \frac{n_e}{d_e}$;

(ii) $-\beta_e \leq \Delta(w_i \cdots w_{i+k \cdot n_e - 1} w_{j-k \cdot d_e+1} \cdots w_j) \leq \beta_e$ for all $k \in \mathbb{N}_{>0}$ such that $k \leq (j-j'+1)/d_e$ and $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_j) = 0$;

(iii) $\Delta(w_{i+(q-1)\cdot\gamma_e} \cdots w_{i+q\cdot\gamma_e-1}) \geq 1$ for all $q \in \mathbb{N}_{>0}$ with $q \cdot \gamma_e \leq i'-i+1$ and $\Delta(w_i \cdots w_{i+p-1}) \geq 0$ for all $p \in [1, i'-i+1]$;

(iv) $\Delta(w_{j-q\cdot\gamma_e+1} \cdots w_{j-(q-1)\cdot\gamma_e}) \leq -1$ for all $q \in \mathbb{N}_{>0}$ with $q \cdot \gamma_e \leq j-j'+1$ and $\Delta(w_{j-p+1} \cdots w_j) \leq 0$ for all $p \in [1, j-j'+1]$.

Let us first prove that these four conditions whose conjunction the $\mathsf{FO}[+]$-formula $\pi_e(x,x',y',y)$ will express, indeed imply the two conditions of the lemma.

If conditions (i) to (iv) are satisfied for a $w \in \Sigma^+$ and $i, i', j', j \in [1,|w|], i \leq i' < j' \leq j$, we actually have that $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}$. Indeed, condition (ii) ensures that $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_j) =$

0. Conditions (iii) and (iv) then additionally imply that $\Delta(w_i \cdots w_{i+p-1}) \geq 0$ for all $p \in [1, i'-i+1]$ and $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_{j'+p-1}) \geq 0$ for all $p \in [1, j-j'+1]$. This is because if there were a $p \in [1, j-j'+1]$ such that $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_{j'+p-1}) < 0$, then it should be that $\Delta(w_{j'+p} \cdots w_j) > 0$ with $p \leq j - j'$ as we already know that $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_j) = 0$: this would be a contradiction to condition (iv).

Conversely, let us fix some $w \in \Sigma^+$ and indices $i, i', j', j \in [1, |w|]$ such that $i \leq i' < j' \leq j$, $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}$, $\Delta(w_i \cdots w_{i'}) > 0$ and $e \circ \psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$. In the terminology of Proposition 4.11, for $F = \varphi_L(L)$, we have $(w_i \cdots w_{i'}, w_{j'} \cdots w_j) \in \mathcal{U}_e$. We claim that Points (i) to (iv) are actually satisfied. Indeed, recalling that L is length-synchronous by assumption, 2(a) of Proposition 4.11 for e in fact states that that Point (i) is satisfied. Next, since for all $k \in \mathbb{N}_{>0}$ such that $k \leq \frac{j-j'+1}{d_e} = \frac{i'-i+1}{n_e}$, the word $w_i \cdots w_{i+k \cdot n_e - 1}$ is a prefix of $w_i \cdots w_{i'}$ and the word $w_{j-k \cdot d_e+1} \cdots w_j$ is a suffix of $w_{j'} \cdots w_j$ such that $\frac{|w_i \cdots w_{i+k \cdot n_e-1}|}{|w_{j-k \cdot d_e+1} \cdots w_j|} = \frac{k \cdot n_e}{k \cdot d_e} = \alpha_e$, it must hold that $-\beta_e \leq \Delta(w_i \cdots w_{i+k \cdot n_e-1} w_{j-k \cdot d_e+1} \cdots w_j) \leq \beta_e$ by Point 2(b) of Proposition 4.11. We have that $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_j) = 0$ immediately follows from our assumption $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}$, thus Point (ii) holds. Another consequence of our assumption $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}$ is that $\Delta(w_i \cdots w_{i+p-1}) \geq 0$ for all $p \in [1, i'-i+1]$ and $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_{j'+p-1}) \geq 0$ for all $p \in [1, j-j'+1]$. This implies that $\Delta(w_{j-p+1} \cdots w_j) \leq 0$ for all $p \in [1, j-j'+1]$, as already argued above. Since $w_{i+(q-1) \cdot \gamma_e} \cdots w_{i+q \cdot \gamma_e - 1}$ is a factor of $w_i \cdots w_{i'}$ of length γ_e for all $q \in \mathbb{N}_{>0}$ such that $q \cdot \gamma_e \leq i' - i + 1$ and $w_{j-q \cdot \gamma_e+1} \cdots w_{j-(q-1) \cdot \gamma_e}$ is a factor of $w_{j'} \cdots w_j$ of length γ_e for all $q \in \mathbb{N}_{>0}$ such that $q \cdot \gamma_e \leq j - j' + 1$, by Points 2(c) and 2(d) of Proposition 4.11, we finally have that conditions (iii) and (iv) are also satisfied.

It now remains to construct the formula $\pi_e(x, x', y', y)$. We set

$$\begin{aligned}\pi_e(x, x', y', y) = &(x' - x + 1) \cdot d_e = (y - y' + 1) \cdot n_e \wedge \\ &\mu_{n_e, d_e, \beta_e}(x, x', y', y) \wedge \\ &\nu^+_{\gamma_e}(x, x') \wedge \nu^-_{\gamma_e}(y', y),\end{aligned}$$

where the first line checks condition (i), the $\mathsf{FO}[+]$-formula $\mu_{n_e, d_e, \beta_e}(x, x', y', y)$ will check condition (ii) under the assumption condition (i) is satisfied and the $\mathsf{FO}[+]$-formulas $\nu^+_{\gamma_e}(x, x')$ and $\nu^-_{\gamma_e}(y', y)$ respectively will check conditions (iii) and (iv). We now explain how to build those formulas.

Helper formulas. For all $k \in \mathbb{N}_{>0}$ and $h \in \mathbb{Z}$ such that $-k \leq h \leq k$, we let

$$H^h_k(x) = \bigvee_{\substack{I, J \subseteq [1,k] \\ I \cap J = \emptyset \\ |I| - |J| = h}} \Big(\bigwedge_{p \in I} \Sigma_{\mathrm{call}}(x + p - 1) \wedge \bigwedge_{p \in J} \Sigma_{\mathrm{ret}}(x + p - 1) \wedge \bigwedge_{p \in [1,k] \setminus (I \cup J)} \Sigma_{\mathrm{int}}(x + p - 1) \Big)$$

such that for all $w \in \Sigma^+$ and $i \in [1, |w|]$ such that $i \leq |w| - k + 1$, we have $w \models H^h_k(i)$ if, and only if, $\Delta(w_i \cdots w_{i+k-1}) = h$.

For all $n, d \in \mathbb{N}_{>0}$ relatively prime and $h \in \mathbb{Z}$, $-n - d \leq h \leq n + d$, we define

$$D^h_{n,d}(x, y, z) = \bigvee_{\substack{-n \leq h_1 \leq n \\ -d \leq h_2 \leq d \\ h_1 + h_2 = h}} \Big(H^{h_1}_n\big(x + (z-1) \cdot n\big) \wedge H^{h_2}_d\big(y - z \cdot d + 1\big) \Big) ,$$

such that for all $w \in \Sigma^+$ and $i, j, k \in [1, |w|]$ with $i + k \cdot n - 1 \leq |w|$ and $j - k \cdot d + 1 \geq 1$, we have

$w \models D^h_{n,d}(i,j,k)$ if, and only if,

$$\Delta(w_{i+(k-1)\cdot n} \cdots w_{i+k\cdot n-1} w_{j-k\cdot d+1} \cdots w_{j-(k-1)\cdot d}) = h \ .$$

Formula $\mu_{n,d,q}(x,x',y',y)$**.** For each $p \in \mathbb{N}$ let $\Gamma_p = \{a_{-p}, \ldots, a_{-1}, a_0, a_1, \ldots, a_p\}$ and define $\Delta_p \colon \Gamma_p^* \to \mathbb{Z}$ to be the p-height monoid morphism satisfying $\Delta_p(a_h) = h$ for all $a_h \in \Gamma_p$. Consider the language

$$L_{p,q} = \{w \in \Gamma_p^* \mid \Delta_p(w) = 0 \wedge \forall i \in [1,|w|], -q \leq \Delta_p(w_1 \cdots w_i) \leq q\} \ .$$

We claim that this language is recognized by a finite aperiodic monoid. This implies, by a theorem by McNaughton and Papert (see [29, Theorem VI.1.1]), that there exists an $\mathsf{FO}_{\Gamma_{n+d}}[<]$-sentence $\tilde{\mu}_{p,q}$ defining $L_{p,q}$.

Let now $n,d \in \mathbb{N}_{>0}$ relatively prime and $q \in \mathbb{N}$. Consider $w \in \Sigma^+$ and $i,i',j',j \in [1,|w|]$ such that $i \leq i' < j' \leq j$ and $\frac{i'-i+1}{j-j'+1} = \frac{n}{d}$. We want to check whether we have

$$-q \leq \Delta(w_i \cdots w_{i+k\cdot n-1} w_{j-k\cdot d+1} \cdots w_j) \leq q$$

for all $k \in \mathbb{N}_{>0}$ such that $k \leq (j-j'+1)/d$ and moreover $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_j) = 0$. Since n and d are relatively prime, this means that there exists $l \in [1,|w|]$ such that $i'-i+1 = l \cdot n$ and $j-j'+1 = l \cdot d$. We can hence decompose $w_i \cdots w_{i'}$ as $u_1 \cdots u_l$ with $u_1, \ldots, u_l \in \Sigma^n$ and $w_{j'} \cdots w_j$ as $v_l \cdots v_1$ with $v_1, \ldots, v_l \in \Sigma^d$. Observe that $\Delta(u_i v_i) \in [-n-d, n+d]$ for all $i \in [1,l]$. Using this decomposition, we now need to check whether $-q \leq \Delta(u_1 v_1) + \cdots + \Delta(u_k v_k) \leq q$ for all $k \in [1,l]$ and $\Delta(u_1 v_1) + \cdots + \Delta(u_l v_l) = 0$. This is equivalent to checking whether the word $\tilde{w} = a_{\Delta(u_1 v_1)} \cdots a_{\Delta(u_l v_l)}$ in Γ_{n+d}^* belongs to $L_{n+d,q}$.

We thus transform the $\mathsf{FO}_{\Gamma_{n+d}}[<]$-sentence $\tilde{\mu}_{n+d,q}$ into an $\mathsf{FO}_\Sigma[+]$-formula $\mu_{n,d,q}(x,x',y',y)$ by

- replacing any quantification $\exists z \rho(z)$ by $\exists z\big(z \leq (y-y'+1)/d \wedge \rho(z)\big)$;
- replacing any quantification $\forall z \rho(z)$ by $\forall z\big(z \leq (y-y'+1)/d \to \rho(z)\big)$;
- replacing any atomic formula of the form $a_h(z)$ for $a_h \in \Gamma_{n+d}$ by $D^h_{n,d}(x,y,z)$.

By this translation for all $w \in \Sigma^+$ and $i,i',j',j \in [1,|w|]$ with $i \leq i' < j' \leq j$ and $\frac{i'-i+1}{j-j'+1} = \frac{n}{d}$ we have $w \models \mu_{n,d,q}(i,i',j',j)$ if, and only if, $-q \leq \Delta(w_i \cdots w_{i+k\cdot n-1} w_{j-k\cdot d+1} \cdots w_j) \leq q$ for all $k \in \mathbb{N}_{>0}, k \leq (j-j'+1)/d$ and $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_j) = 0$.

It remains to show that $L_{p,q}$ is recognized by a finite aperiodic monoid for all $p,q \in \mathbb{N}$. Set $Q_q = \{-q, \ldots, -1, 0, 1, \ldots, q, \bot\}$ and consider the monoid ${Q_q}^{Q_q}$ with function composition from left to right. For each $a_h \in \Gamma_p$, we define the function $f_{a_h} \colon Q_q \to Q_q$ to be such that

$$f_{a_h}(h') = \begin{cases} h'+h & \text{if } h' \neq \bot \text{ and } -q \leq h'+h \leq q \\ \bot & \text{otherwise} \end{cases}$$

for all $h' \in Q_q$. We take $M_{p,q}$ to be the submonoid of ${Q_q}^{Q_q}$ generated by $\{f_{a_h} \mid a_h \in \Gamma_p\}$ and define $\varphi_{p,q} \colon \Gamma_p^* \to M_{p,q}$ as the unique monoid morphism such that $\varphi_{p,q}(a_h) = f_{a_h}$ for all $a_h \in \Gamma_p$.

It is straightforward to show, by induction on the length of w, that for all $w \in \Gamma_p^*$ and all $h \in Q_q$, we have

$$\varphi_{p,q}(w)(h) = \begin{cases} h + \Delta_p(w) & \text{if } h \neq \bot \text{ and } -q \leq h + \Delta_p(w_1 \cdots w_i) \leq q \text{ for all } i \in [1,|w|] \\ \bot & \text{otherwise.} \end{cases}$$

Thus $L_{p,q} = \varphi_{p,q}^{-1}(\{f \in M_{p,q} \mid f(0) = 0\})$. We claim that the monoid $M_{p,q}$ is aperiodic. Indeed, take $f \in M_{p,q}$; we claim that $f^{2q+1} = f^{2q+2}$. Since $M_{p,q}$ is generated by $\{f_{a_h} \mid a_h \in \Gamma_p\}$, there exists $w \in \Gamma_p^*$ satisfying $\varphi_{p,q}(w) = f$. There are three subcases to consider.

- If $\Delta_p(w) = 0$, then since $h + \Delta_p(w^{n-1}w_1 \cdots w_i) = h + \Delta_p(w_1 \cdots w_i)$ for all $h \in \mathbb{Z}, -q \leq h \leq q$, for all $n \in \mathbb{N}_{>0}$ and $i \in [1, |w|]$, we have that $f^n = f$ for all $n \in \mathbb{N}_{>0}$.
- If $\Delta_p(w) > 0$, then since $q < h + \Delta_p(w^{2q+1}) \leq h + \Delta_p(w^{2q+2})$ for all $h \in \mathbb{Z}, -q \leq h \leq q$, both f^{2q+1} and f^{2q+2} must be equal to the function sending every element to $\perp$.
- If $\Delta_p(w) < 0$, then since $h + \Delta_p(w^{2q+2}) \leq h + \Delta_p(w^{2q+1}) < -q$ for all $h \in \mathbb{Z}, -q \leq h \leq q$, both f^{2q+1} and f^{2q+2} must be equal to the function sending every element to $\perp$.

Formula $\nu_l^+(x, x')$. For all $l \in \mathbb{N}_{>0}$, we let

$$\nu_l^+(x, x') = \bigwedge_{p=1}^{l^2} \Big(x' - x + 1 \geq p \rightarrow \bigwedge_{k=1}^{p} \bigvee_{h=0}^{k} H_k^h(x) \Big) \wedge$$
$$\forall z \Big(z \cdot l \leq x' - x + 1 \rightarrow \bigvee_{h=1}^{l} H_l^h\big(x + (z-1) \cdot l\big) \Big) .$$

Fix any $w \in \Sigma^+$ and $i, i' \in [1, |w|]$ such that $i \leq i'$. We have $w \models \nu_l^+(i, i')$ if, and only if, $\Delta(w_i \cdots w_{i+p-1}) \geq 0$ for all $p \in [1, \min\{l^2, i'-i+1\}]$ and $\Delta(w_{i+(q-1)\cdot l} \cdots w_{i+q\cdot l-1}) \geq 1$ for all $q \in \mathbb{N}_{>0}$ such that $q \cdot l \leq i' - i + 1$. The latter is clearly equivalent to having $\Delta(w_{i+(q-1)\cdot l} \cdots w_{i+q\cdot l-1}) \geq 1$ for all $q \in \mathbb{N}_{>0}, q \cdot l \leq i' - i + 1$ and $\Delta(w_i \cdots w_{i+p-1}) \geq 0$ for all $p \in [1, i' - i + 1]$, as required.

Formula $\nu_l^-(y', y)$. For all $l \in \mathbb{N}_{>0}$, we let

$$\nu_l^-(y', y) = \bigwedge_{p=1}^{l^2} \Big(y - y' + 1 \geq p \rightarrow \bigwedge_{k=1}^{p} \bigvee_{h=0}^{k} H_k^{-h}(y - k + 1) \Big) \wedge$$
$$\forall z \Big(z \cdot l \leq y - y' + 1 \rightarrow \bigvee_{h=1}^{l} H_l^{-h}\big(y - z \cdot l + 1\big) \Big) .$$

Therefore, analogously as for $\nu_l^+(x, x')$, for all $w \in \Sigma^+$ and $j', j \in [1, |w|]$ such that $j' \leq j$, we have $w \models \nu_l^-(j', j)$ if, and only if, $\Delta(w_{j-q\cdot l+1} \cdots w_{j-(q-1)\cdot l}) \leq -1$ for all $q \in \mathbb{N}_{>0}$ such that $q \cdot l \leq j - j' + 1$ and $\Delta(w_{j-p+1} \cdots w_j) \leq 0$ for all $p \in [1, j - j' + 1]$.

5.3.5 Evaluation in FO[+] (Proof of Point 4)

The following proposition states that every VPL L that has bounded nesting depth and for which the horizontal and vertical evaluation languages $\mathcal{E}_{\varphi_L, r}$ and $\mathcal{E}_{\psi_L, e}$ are quasi-aperiodic for all $\varphi_L(L)$-reachable r and e, respectively, is definable by an $\mathsf{FO}_{\Sigma, \leftrightsquigarrow}[+]$-sentence in case an approximate matching is present as built-in predicate.

Proposition 5.16. *Assume a VPL L has bounded nesting depth and*

- *$\mathcal{E}_{\varphi_L,r}$ is a regular language whose syntactic morphism is quasi-aperiodic for all $\varphi_L(L)$-reachable $r \in R_L$, and*
- *$\mathcal{E}_{\psi_L,e}$ is a regular language whose syntactic morphism is quasi-aperiodic for all $\varphi_L(L)$-reachable $e \in O_L$.*

Then there exists an $\mathsf{FO}_{\Sigma,\leftrightsquigarrow}[+]$-sentence η such that for any approximate matching M relative to L, we have $w \in L$ if, and only if, $(w, M(w)) \models \eta$ for all $w \in \Sigma^*$.

Proof. By hypothesis, there exists $d_L \in \mathbb{N}$ bounding the nesting depth of the words in L.

By hypothesis also, for each $\varphi_L(L)$-reachable $r \in R_L$, the language $\mathcal{E}_{\varphi_L,r}$ is regular and its syntactic morphism is quasi-aperiodic. This implies, by [29, Theorem VI.4.1], that for each $\varphi_L(L)$-reachable $r \in R_L$, there exists an $\mathsf{FO}_{\Gamma_{\varphi_L}}[<, \mathsf{MOD}]$-sentence $\nu_{\varphi_L,r}$ defining $\mathcal{E}_{\varphi_L,r}$.

Finally, by hypothesis, for each $\varphi_L(L)$-reachable $e \in O_L$, the language $\mathcal{E}_{\psi_L,e}$ is regular and its syntactic morphism is quasi-aperiodic. Again, by [29, Theorem VI.4.1], for each $\varphi_L(L)$-reachable $e \in O_L$, there exists an $\mathsf{FO}_{\Gamma_{\psi_L}}[<, \mathsf{MOD}]$-sentence $\nu_{\psi_L,e}$ defining $\mathcal{E}_{\psi_L,e}$.

Auxiliary formulas. We introduce a few auxiliary formulas that all assume access to the full matching relation $M^{\triangle}(w)$, represented by the relational symbol $\leftrightsquigarrow$.

First let us define a formula A such that for all $w \in \Sigma^{\triangle}$ and $i, j, k \in [1, |w|]$ satisfying $w_i \cdots w_j \in \Sigma^{\triangle}$, we have that $(w, M^{\triangle}(w)) \models A(i, j, k)$ if, and only if, $i \leq k < j$ and $\Delta(w_i \cdots w_k) > 0$. We let

$$A(x, y, z) = \exists x' \exists y' (x \leq x' \leq z < y' \leq y \wedge x' \leftrightsquigarrow y') \ .$$

Next, we define a formula U such that for all $w \in \Sigma^{\triangle}$ and $i, i', k \in [1, |w|]$, we have that $(w, M^{\triangle}(w)) \models U(i, i', k)$ if, and only if, $i \leq k \leq i'$ and k is matched to some position larger than i' in w. We let

$$U(x, x', z) = x \leq z \leq x' \wedge \exists t (z \leftrightsquigarrow t \wedge x' < t) \ .$$

The last formulas we introduce are N_d which express that the infix $w_i \cdots w_j \in \Sigma^{\triangle}$ of $w \in \Sigma^{\triangle}$ has nesting depth at least $d \geq 0$. More precisely, for all $d \in \mathbb{N}$, we introduce auxiliary formulas N_d such that for all $w \in \Sigma^{\triangle}$ and $i, j \in [1, |w|]$ satisfying $w_i \cdots w_j \in \Sigma^{\triangle}$, we have that $(w, M^{\triangle}(w)) \models N_d(i, j)$ if, and only if, $\mathrm{nd}(w_i \cdots w_j) \geq d$. The case $d = 0$ is trivial since we can set $N_0(i, j) = \top$.

Take $w \in \Sigma^{\triangle}$ such that $\mathrm{nd}(w) = d' \geq 1$. Note that then w can be factorized as $w = w_1 u w_2$ such that $w_1, w_2 \in \Sigma^{\triangle}$, $u \in \Sigma_{\mathrm{call}} \Sigma^{\triangle} \Sigma_{\mathrm{ret}}$ and $\mathrm{nd}(w) = \mathrm{nd}(u) = d'$. This means that $u = a_1 v b_1$ for $a_1 \in \Sigma_{\mathrm{call}}$, $b_1 \in \Sigma_{\mathrm{ret}}$ and $v \in \Sigma^{\triangle}$. We then apply Lemma 4.16 and Lemma 4.17 implying that u has a nesting-maximal stair factorization

$$u = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}(u')$$

for which there exists $h \in [1, k]$ such that, setting $u_i = \mathrm{ext}_{a_i,b_i} \circ \mathrm{ext}_{x_{i+1},y_{i+1}} \circ \cdots \circ \ \mathrm{ext}_{a_k,b_k}(u')$ for all $i \in [1, k]$ and $u_{k+1} = u'$, we have $\mathrm{nd}(u) = \mathrm{nd}(u_h) = d'$ and $\mathrm{nd}(\mathrm{ext}_{x_{h+1},y_{h+1}} \circ \mathrm{ext}_{a_{h+1},b_{h+1}} \circ \cdots \circ \mathrm{ext}_{a_k,b_k}(u')) = \mathrm{nd}(u_{h+1}) = d' - 1$. Thus, by definition of the nesting depth of a well-matched word, $u_h = a_h z_1 z_2 b_h$ for $z_1, z_2 \in \Sigma^{\triangle}$ satisfying $\mathrm{nd}(z_1) = \mathrm{nd}(z_2) = d' - 1$.

Hence, we set

$$N_1(x, y) = \exists x' \exists y' (x \leq x' < y' \leq y \wedge x' \leftrightsquigarrow y')$$

and for $d \geq 2$ we set

$$N_d(x,y) = \exists x' \exists y' \exists z \big(x \leq x' < z < y' \leq y \wedge x' \leftrightsquigarrow y' \wedge \neg A(x'+1, y'-1, z) \wedge \\ N_{d-1}(x'+1, z) \wedge N_{d-1}(z+1, y'-1) \big) \ .$$

Main construction. To build the $\mathsf{FO}_{\Sigma,\leftrightsquigarrow}[+]$-sentence η, we build $\mathsf{FO}_{\Sigma,\leftrightsquigarrow}[+]$-formulas

- $\eta^{\uparrow}_{d,r}(x,y)$ for all $d \in \mathbb{N}$ and all $\varphi_L(L)$-reachable $r \in R_L$ and
- $\eta_{d,r}(x,y)$ for all $d \in \mathbb{N}$ and all $\varphi_L(L)$-reachable $r \in R_L$

that also assume access to the full matching relation $M^{\Delta}(w)$. They will have the following properties for all $w \in \Sigma^{\triangle}$ and all $i, j \in [1, |w|]$:

- if i is matched to j in w, then $(w, M^{\triangle}(w)) \models \eta^{\uparrow}_{d,r}(i,j)$ if, and only if, $\mathrm{nd}(w_i \cdots w_j) \leq d$ and $\varphi_L(w_i \cdots w_j) = r$ and
- if $w_i \cdots w_j \in \Sigma^{\triangle}$, then $(w, M^{\triangle}(w)) \models \eta_{d,r}(i,j)$ if, and only if, $\mathrm{nd}(w_i \cdots w_j) \leq d$ and $\varphi_L(w_i \cdots w_j) = r$.

Let the formula E be defined as $\forall x(x \neq x)$ if $\varepsilon \in L$ and $\bot = \exists x(x \neq x)$ otherwise. Our final formula η will then be defined as

$$\eta = \forall z \exists t \big((\Sigma_{\mathrm{call}}(z) \to z \leftrightsquigarrow t) \wedge (\Sigma_{\mathrm{ret}}(z) \to t \leftrightsquigarrow z) \big) \wedge \\ \Big(E \vee \exists x \exists y \big(\neg \exists x'(x' < x) \wedge \neg \exists y'(y < y') \wedge \bigvee_{r \in \varphi_L(L)} \eta_{d_L,r}(x,y) \big) \Big).$$

It now remains to build $\eta^{\uparrow}_{d,r}(x,y)$ and $\eta_{d,r}(x,y)$ for all $d \in \mathbb{N}$ and $\varphi_L(L)$-reachable $r \in R_L$. The construction is by induction on d. Let $r \in R_L$ that is $\varphi_L(L)$-reachable. We define $\eta^{\uparrow}_{0,r}(x,y) = \bot$. We define $\eta_{0,r}$ as

$$\eta_{0,r}(x,y) = \neg N_1(x,y) \wedge \tau_0(\nu_{\varphi_L,r}) \ ,$$

where the translation τ_0 is inductively defined as follows:

- $\tau_0(z < z') = z < z'$;
- $\tau_0(s(z)) = \bigvee_{c \in \varphi_L^{-1}(s) \cap \Sigma_{\mathrm{int}}} c(z)$ for all $s \in \varphi_L(\Sigma^{\triangle} \setminus \{\varepsilon\})$;
- $\tau_0(\mathsf{MOD}_m(z)) = \exists t \big((z = x \to 1 = t \cdot m) \wedge (z \neq x \to z - x + 1 = t \cdot m) \big)$ for all $m \in \mathbb{N}_{>0}$;
- $\tau_0(\$(z)) = \bot$;
- $\tau_0(\rho_1(\boldsymbol{z}_1) \wedge \rho_2(\boldsymbol{z}_2)) = \tau_0(\rho_1(\boldsymbol{z}_1)) \wedge \tau_0(\rho_1(\boldsymbol{z}_2))$;
- $\tau_0(\neg \rho(\boldsymbol{z})) = \neg \tau_0(\rho(\boldsymbol{z}))$;
- $\tau_0(\exists z \rho(z, \boldsymbol{z})) = \exists z \big(x \leq z \leq y \wedge \tau_0(\rho(z, \boldsymbol{z})) \big)$.

Now let $d > 0$. Let us first define $\eta_{d,r}$ when assuming that we have already defined $\eta^{\uparrow}_{d,r}$. Given $w \in \Sigma^{\triangle} \setminus \{\varepsilon\}$ and $i, j \in [1, |w|]$ such that $w_i \cdots w_j \in \Sigma^{\triangle} \setminus \{\varepsilon\}$, note that in case $\mathrm{nd}(w_i \cdots w_j) \leq d$, then one can factorize $w_i \cdots w_j$ as $w_i \cdots w_j = u_1 \cdots u_m$ such that $u_\ell \in \Sigma_{\mathrm{int}} \cup \Sigma_{\mathrm{call}} \Sigma^{\triangle} \Sigma_{\mathrm{ret}}$ and

$\mathrm{nd}(u_\ell) \leq d$ for all $\ell \in [1,m]$. Note also that if $\varphi_L(w_i \cdots w_j) = r$, as r is $\varphi_L(L)$-reachable, then $\varphi_L(u_\ell)$ is $\varphi_L(L)$-reachable for all $\ell \in [1,m]$. Using these observations we define

$$\eta_{d,r}(x,y) = \neg N_{d+1}(x,y) \wedge \forall z\Big(x \leq z \leq y \to \big(\tau_1(\$(z)) \vee \bigvee_{\substack{s \in \varphi_L(\Sigma^\triangle \setminus \{\varepsilon\}) \\ \varphi_L(L)\text{-reachable}}} \tau_1(s(z))\big)\Big) \wedge \tau_1(\nu_{\varphi_L,r}) \ ,$$

where the translation τ_1 agrees with the above translation τ_0 (where, as expected, occurrences of τ_0 are replaced by τ_1) except for the following kinds of subformulas:

- $\tau_1(\$(z)) = A(x,y,z)$;
- $\tau_1(s(z)) = \neg A(x,y,z) \wedge \Big(\bigvee_{c \in \varphi_L^{-1}(s) \cap \Sigma_{\mathrm{int}}} c(z) \vee \exists t\big(x \leq t \leq y \wedge t \leftrightsquigarrow z \wedge \eta^{\uparrow}_{d,s}(t,z)\big)\Big)$ if s is $\varphi_L(L)$-reachable, $\tau_1(s(z)) = \bot$ otherwise.

It remains to define $\eta^{\uparrow}_{d,r}$.

We first construct for all $\varphi_L(L)$-reachable $e \in O_L$ a formula $\chi_{d,e}(x,x',y',y)$ such that for all $w \in \Sigma^\triangle$ and $i,i',j',j \in [1,|w|], i \leq i' < j' \leq j$ we have that if $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\triangle$ and i' is matched to j' in w, then given

$$\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j} = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k}$$

the stair factorization of $\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}$ provided by Lemma 3.6, we have $(w, M^\triangle(w)) \models \chi_{d,e}(i,i',j',j)$ if, and only if, $\mathrm{nd}(x_\ell), \mathrm{nd}(y_\ell) < d$ for all $\ell \in [1,k]$ and $\psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$. Given w,i,i',j',j and the associated stair factorization as above, note that if $\psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$, as e is $\varphi_L(L)$-reachable, then $\varphi_L(x_\ell)$ and $\varphi_L(y_\ell)$ are $\varphi_L(L)$-reachable for all $\ell \in [1,k]$. If additionally $\mathrm{nd}(x_\ell), \mathrm{nd}(y_\ell) < d$ for all $\ell \in [1,k]$, we can inductively make use of the formulas $\{\eta_{d-1,r'} \mid r' \in R_L \ \varphi_L(L)\text{-reachable}\}$ in order to evaluate $\varphi_L(x_1), \varphi_L(y_1), \ldots, \varphi_L(x_k), \varphi_L(y_k)$. Let $p_1, \ldots, p_k \in [i,i']$ be the positions in $w_i \cdots w_{i'}$ where, respectively, $a_1, \ldots, a_k$ in the above factorization of $\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}$ appear: the formula $\chi_{d,e}$ will verify if $\sigma_i \cdots \sigma_{i'} \in \mathcal{E}_{\psi_L,e}$, where

$$\sigma_q = \begin{cases} \mathrm{left}_{\varphi_L(x_m)} \circ \mathrm{right}_{\varphi_L(y_m)} \circ \psi_L(\mathrm{ext}_{a_m,b_m}) & \text{if } q = p_m \text{ for } m \in [1,k] \\ \$ & \text{otherwise} \end{cases}$$

for all $q \in [i,i']$. Hence we set

$$\chi_{d,e}(x,x',y',y) = \forall z\Big(x \leq z \leq x' \to \big(\tau_2(\$(z)) \vee \bigvee_{\substack{f \in \psi_L\left(\mathcal{O}(\Sigma^\triangle)_\uparrow\right) \\ \varphi_L(L)\text{-reachable}}} \tau_2(f(z))\big)\Big) \wedge \tau_2(\nu_{\psi_L,e}) \ ,$$

where the translation τ_2 agrees with translation τ_0 (where, as expected, occurrences of τ_0 are replaced by τ_2) with the following exceptions:

- $\tau_2(\exists z \rho(z, \boldsymbol{z})) = \exists z\big(x \leq z \leq x' \wedge \tau_2(\rho(z, \boldsymbol{z}))\big)$
- $\tau_2(\$(z)) = \neg U(x,x',z)$
- $\tau_2(f(z)) = \begin{cases} \exists t\Big(U(x,x',z) \wedge z \leftrightsquigarrow t \wedge \big(\iota_{d,f}(x,x',y',y,z,t) \vee \zeta_{d,f}(x,x',z,t)\big)\Big) & \text{if } f \text{ is } \varphi_L(L)\text{-reachable} \\ \bot & \text{otherwise} \end{cases}$

for all $f \in \psi_L\big(\mathcal{O}(\Sigma^\triangle)_\uparrow\big)$, where

$$\iota_{d,f}(x,x',y',y,z,t) = \neg\exists z'\big(U(x,x',z') \wedge z' < z\big) \wedge \bigvee_{\substack{a\in\Sigma_{\mathrm{call}}, b\in\Sigma_{\mathrm{ret}}, r',r''\in R_L \\ f=\mathrm{left}_{r'}\circ\mathrm{right}_{r''}\circ\psi_L(\mathrm{ext}_{a,b})}} \Big(a(z)\wedge b(t)\wedge \Big(\big(x<z\wedge\eta_{d-1,r'}(x,z-1)\big)\vee F_{r'}(x,z)\Big)\wedge \Big(\big(t<y\wedge\eta_{d-1,r''}(t+1,y)\big)\vee F_{r''}(y,t)\Big)\Big)$$

with $F_s(x,z) = \begin{cases} x=z & \text{if } s = 1_{R_L} \\ \bot & \text{otherwise} \end{cases}$ for all $\varphi_L(L)$-reachable $s \in R_L$ and

$$\zeta_{d,f}(x,x',z,t) = \exists z'\exists t'\Big(U(x,x',z')\wedge z'<z\wedge\neg\exists z''\big(z'<z''<z\wedge U(x,x',z'')\big)\wedge z' \leftrightsquigarrow t' \wedge \bigvee_{\substack{a\in\Sigma_{\mathrm{call}}, b\in\Sigma_{\mathrm{ret}}, r',r''\in R_L \\ f=\mathrm{left}_{r'}\circ\mathrm{right}_{r''}\circ\psi_L(\mathrm{ext}_{a,b})}} \Big(a(z)\wedge b(t)\wedge \eta_{d-1,r'}(z'+1,z-1)\wedge\eta_{d-1,r''}(t+1,t'-1)\Big)\Big)\,.$$

We now construct $\eta^{\uparrow}_{d,r}$ itself. Given $w \in \Sigma^\triangle \setminus \{\varepsilon\}$ and $i,j \in [1,|w|]$ such that i is matched to j in w, observe first that the infix $w_i\cdots w_j$ is of the form $w_i\cdots w_j = a_1 v b_1 \in \Sigma^\triangle$ for some $a_1 \in \Sigma_{\mathrm{call}}$, $b_1 \in \Sigma_{\mathrm{ret}}$, and $v \in \Sigma^\triangle$. As above, we can directly express $\mathrm{nd}(w_i\cdots w_j) \leq d$ via the formula $\neg N_{d+1}$. Assuming this holds, towards expressing that $\varphi_L(w_i\cdots w_j) = r$, we make use of Lemma 4.16 and Lemma 4.17: for the infix $w_i\cdots w_j$ there is a nesting-maximal stair factorization

$$w_i\cdots w_j = \mathrm{ext}_{x_1,y_1}\circ\mathrm{ext}_{a_1,b_1}\circ\cdots\circ\mathrm{ext}_{x_k,y_k}\circ\mathrm{ext}_{a_k,b_k}(u')$$

such that we have

1. $\mathrm{nd}(x_1),\mathrm{nd}(y_1),\ldots,\mathrm{nd}(x_k),\mathrm{nd}(y_k) < d$; and
2. if $\varphi_L(w_i\cdots w_j) = r$, as r is $\varphi_L(L)$-reachable, then $\psi_L(\mathrm{ext}_{x_1,y_1}\circ\mathrm{ext}_{a_1,b_1}\circ\cdots\circ\mathrm{ext}_{x_k,y_k}\circ\mathrm{ext}_{a_k,b_k})$ and $\varphi_L(u')$ are $\varphi_L(L)$-reachable.

By these points, we can use the formulas $\{\chi_{d,e}(x,x',y',y) \mid e \in O_L\ \varphi_L(L)\text{-reachable}\}$ to evaluate $\psi_L(\mathrm{ext}_{x_1,y_1}\circ\mathrm{ext}_{a_1,b_1}\circ\cdots\circ\mathrm{ext}_{x_k,y_k}\circ\mathrm{ext}_{a_k,b_k})$ and the formulas $\{\eta_{0,r'} \mid r' \in R_L\ \varphi_L(L)\text{-reachable}\}$ to evaluate $\varphi_L(u')$. We are now ready to give the formula $\eta^{\uparrow}_{d,r}$. We set

$$\eta^{\uparrow}_{d,r}(x,y) = \neg N_{d+1}(x,y)\wedge\exists x'\exists y'\Big(x\leq x'<y'\leq y\wedge x' \leftrightsquigarrow y' \wedge \bigvee_{\substack{r'\in R_L, e\in O_L \\ e(r')=r}} \big(\chi_{d,e}(x,x',y',y)\wedge\eta_{0,r'}(x'+1,y'-1)\big)\Big)\,.$$

□

5.4 The intermediate case

The following theorem effectively characterizes the remaining case, namely those VPLs that are weakly length-synchronous but not length-synchronous and whose syntactic morphism is quasi-aperiodic: such VPLs are shown to be constant-depth equivalent to a non-empty disjoint union of

intermediate VPLs. The computability of $k, l \in \mathbb{N}$ with $k \neq l$ such that $\mathcal{L}_{k,l} \leq_{\mathrm{cd}} L$ is subject of Section 5.4.2.

Theorem 5.17. *If L is weakly length-synchronous but not length-synchronous, and its syntactic morphism (φ_L, ψ_L) is quasi-aperiodic, one can effectively compute vertically visibly pushdown grammars $G_1, \ldots, G_m$ generating intermediate VPLs such that $L =_{cd} \biguplus_{i=1}^{m} L(G_i)$.*

Before we give the proof of the theorem we need a bit of notation. Let $L \subseteq \Sigma^{\triangle}$ be a VPL that is weakly length-synchronous, not length-synchronous, and whose syntactic morphism (φ_L, ψ_L) is quasi-aperiodic. By Proposition 5.1 one can effectively compute its syntactic Ext-algebra (R_L, O_L), (φ_L, ψ_L) and $\varphi_L(L)$ from (a given DVPA for) L.

For all $e \in O_L$ recall that

$$\mathcal{U}_e = \{(u, v) \in \mathsf{Con}(\Sigma) \mid e \circ \psi_L(\mathrm{ext}_{u,v}) = e\}.$$

For all $\varphi_L(L)$-reachable $e \in O_L$ and some fresh internal letter $\# \notin \Sigma$ let

$$M_e = \{u\#v \mid (u, v) \in \mathsf{Con}(\Sigma), \Delta(u) > 0, e \circ \psi_L(\mathrm{ext}_{u,v}) = e\}.$$

Note that since L is assumed to be weakly length-synchronous, by Proposition 4.5, $\mathcal{U}_e$ is weakly length-synchronous for all $\varphi_L(L)$-reachable $e \in O_L$.

Also note that since $M_e = \bigcup\{L_f \mid f \in O_L \colon e \circ f = e\} \cap \{u\#v \mid uv \in \Sigma^{\triangle}, \Delta(u) > 0\}$, since for all languages L_f one can effectively compute Ext-algebras recognizing them by Lemma 6.4, and since one can effectively computable an Ext-algebra recognizing the language $\{u\#v \mid uv \in \Sigma^{\triangle}, \Delta(u) > 0\}$, we obtain that one can effectively compute an Ext-algebra recognizing $M_e \subseteq \Sigma^*\#\Sigma^*$. The set $S_e = \{(k, l) \in \mathbb{N}^2 \mid \exists (u, v) \in \Sigma^k \times \Sigma^l : u\#v \in M_e\} \subseteq \mathbb{N}^2_{>0}$ is hence effectively semilinear by Lemma 6.5. Note that the word relation $\mathcal{U}_e$ is length-synchronous if, and only if, there exists some $\alpha \in \mathbb{Q}_{>0}$ such that $\frac{k}{l} = \alpha$ for all $(k, l) \in S_e$. Lemma 6.7 implies that the latter condition is decidable. As a consequence one can effectively compute the set

$$\mathcal{Z} = \{e \in O_L \mid e \text{ is } \varphi_L(L)\text{-reachable and } \mathcal{U}_e \text{ is not length-synchronous}\}.$$

Observe that since L is not length-synchronous by assumption, we have $\mathcal{Z} \neq \emptyset$ (Proposition 4.10).

Let us introduce two fresh copies $\widetilde{\Sigma} = \{\widetilde{\sigma} \mid \sigma \in \Sigma\}$ and $\overline{\Sigma} = \{\overline{\sigma} \mid \sigma \in \Sigma\}$ of our alphabet Σ. Let $\widetilde{\vartheta} : (\widetilde{\Sigma} \cup \Sigma)^* \to \widetilde{\Sigma}^*$ and $\overline{\vartheta} : (\overline{\Sigma} \cup \Sigma)^* \to \overline{\Sigma}^*$ be the (letter-to-letter and hence length-multiplying) morphisms satisfying $\widetilde{\vartheta}(\sigma) = \widetilde{\vartheta}(\widetilde{\sigma}) = \widetilde{\sigma}$ and $\overline{\vartheta}(\sigma) = \overline{\vartheta}(\overline{\sigma}) = \overline{\sigma}$ for all $\sigma \in \Sigma$. Conversely, let $\widetilde{\vartheta}^{-1} : (\widetilde{\Sigma} \cup \Sigma)^* \to \Sigma^*$ and $\overline{\vartheta}^{-1} : (\overline{\Sigma} \cup \Sigma)^* \to \Sigma^*$ be the morphisms satisfying $\widetilde{\vartheta}^{-1}(\widetilde{\sigma}) = \widetilde{\vartheta}^{-1}(\sigma) = \sigma$ and $\overline{\vartheta}^{-1}(\overline{\sigma}) = \overline{\vartheta}^{-1}(\sigma) = \sigma$ for all $\sigma \in \Sigma$.

We define a new visibly pushdown alphabet $\Upsilon = \Upsilon_{\mathrm{call}} \cup \Upsilon_{\mathrm{int}} \cup \Upsilon_{\mathrm{ret}}$ where $\Upsilon_{\mathrm{call}} = \Sigma_{\mathrm{call}}$, $\Upsilon_{\mathrm{int}} = \Sigma_{\mathrm{int}} \cup \widetilde{\Sigma} \cup \overline{\Sigma} \cup \{\#\}$, and $\Upsilon_{\mathrm{ret}} = \Sigma_{\mathrm{ret}}$.

For every word $u\#v \in M_e$ consider the unique factorization

$$u\#v = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k} \circ \mathrm{ext}_{x_{k+1},y_{k+1}}(\#)$$

where $k \geq 1$, $x_1, \ldots, x_{k+1}, y_1, \ldots, y_{k+1} \in \Sigma^{\triangle}$, $a_1, \ldots, a_k \in \Sigma_{\mathrm{call}}$, and $b_1, \ldots, b_k \in \Sigma_{\mathrm{ret}}$. For these we define

$$(u\#v)^{\ddagger} = \mathrm{ext}_{\widetilde{\vartheta}(x_1),\overline{\vartheta}(y_1)} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{\widetilde{\vartheta}(x_k),\overline{\vartheta}(y_k)} \circ \mathrm{ext}_{a_k,b_k} \circ \mathrm{ext}_{\widetilde{\vartheta}(x_{k+1}),\overline{\vartheta}(y_{k+1})}(\varepsilon) \in \Upsilon^{\triangle} .$$

Notably, $(u\#v)^\ddagger$ does not contain the letter $\#$. Finally, for all $e \in \mathcal{Z}$ we define the language

$$N_e = \{(u\#v)^\ddagger \in \Upsilon^* \mid u\#v \in M_e\} \cup \{\varepsilon\} \subseteq \Upsilon^\triangle \ .$$

Remark 5.18. *Let be $n \in \mathbb{N}$ be the constant from Lemma 4.6 for L and let $e \in \mathcal{Z}$. When setting $F = \varphi_L(L)$ and $(\varphi, \psi) = (\varphi_L, \psi_L)$, Lemma 4.6 states that the factorization*

$$u\#v = \mathrm{ext}_{x_1,y_1} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{x_k,y_k} \circ \mathrm{ext}_{a_k,b_k} \circ \mathrm{ext}_{x_{k+1},y_{k+1}}(\#)$$

of every word $u\#v \in M_e$ satisfies $|x_1|, \ldots, |x_{k+1}|, |y_1|, \ldots, |y_{k+1}| \leq n$. As a consequence, for the corresponding factorization

$$(u\#v)^\ddagger = \mathrm{ext}_{\widetilde{\vartheta}(x_1),\overline{\vartheta}(y_1)} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{\widetilde{\vartheta}(x_k),\overline{\vartheta}(y_k)} \circ \mathrm{ext}_{a_k,b_k} \circ \mathrm{ext}_{\widetilde{\vartheta}(x_{k+1}),\overline{\vartheta}(y_{k+1})}(\varepsilon)$$

of every word $(u\#v)^\ddagger \in N_e \setminus \{\varepsilon\}$ we have $\left|\widetilde{\vartheta}(x_1)\right|, \ldots, \left|\widetilde{\vartheta}(x_{k+1})\right|, \left|\overline{\vartheta}(y_1)\right|, \ldots, \left|\overline{\vartheta}(y_{k+1})\right| \leq n$.

5.4.1 Proof strategy

We are now ready to give the proof strategy for Theorem 5.17. The proof consists of the following steps.

1. N_e is an intermediate VPL for all $e \in \mathcal{Z}$. Moreover, one can effectively compute a vertically visibly pushdown grammar G_e witnessing that N_e is indeed intermediate (Lemma 5.19).
2. $N_e \leq_{\mathrm{cd}} M_e$ for all $e \in \mathcal{Z}$ (Lemma 5.20).
3. $M_e \leq_{\mathrm{cd}} N_e$ for all $e \in \mathcal{Z}$ (Lemma 5.21).
4. $M_e \leq_{\mathrm{cd}} \biguplus_{f \in O_L \text{ is } \varphi_L(L)\text{-reachable}} L_f$ for all $e \in \mathcal{Z}$ (Lemma 5.22).
5. $L \leq_{\mathrm{cd}} \biguplus_{e \in \mathcal{Z}} M_e$ (Lemma 5.23).

Let us argue that Theorem 5.17 follows from the above steps. By Point 1 for all $e \in \mathcal{Z}$ we have that N_e is an intermediate VPL, for which moreover one can effectively compute a vertically visibly pushdown grammar G_e witnessing that N_e is indeed intermediate. Recalling that $\mathcal{Z} \neq \emptyset$, it remains to argue that $L =_{\mathrm{cd}} \biguplus\{N_e \mid e \in \mathcal{Z}\}$. Before we prove this let us recall some basics of constant-depth reductions. For this, let $K, L_1, \ldots, L_n, K_1, \ldots, K_n$ be languages. Firstly, observe that if $L_i \leq_{\mathrm{cd}} K_i$ for all $i \in [1, n]$, then $\biguplus_{i \in [1,n]} L_i \leq_{\mathrm{cd}} \biguplus_{i \in [1,n]} K_i$. Secondly, if $L_i \leq_{\mathrm{cd}} K$ for all $i \in [1, n]$, then $\biguplus_{i \in [1,n]} \leq_{\mathrm{cd}} K$.

Hence we obtain the following sequence of reductions showing that $L =_{\mathrm{cd}} \biguplus\{N_e \mid e \in \mathcal{Z}\}$.

$$
\begin{aligned}
L \quad &\overset{\text{Point 5}}{\leq_{\mathrm{cd}}} \quad \biguplus_{e\in\mathcal{Z}} M_e \\
&\overset{\text{Point 3}}{\leq_{\mathrm{cd}}} \quad \biguplus_{e\in\mathcal{Z}} N_e \\
&\overset{\text{Point 2}}{\leq_{\mathrm{cd}}} \quad \biguplus_{e\in\mathcal{Z}} M_e \\
&\overset{\text{Point 4}}{\leq_{\mathrm{cd}}} \quad \biguplus_{f \in O_L \text{ is } \varphi_L(L)\text{-reachable}} L_f \\
&\overset{\text{Lemma 5.3}}{\leq_{\mathrm{cd}}} \quad L \ .
\end{aligned}
$$

Lemma 5.19. *N_e is an intermediate language for all $e \in \mathcal{Z}$. Moreover, one can effectively compute a vertically visibly pushdown grammar witnessing that N_e is indeed an intermediate VPL.*

Proof. Let $e \in \mathcal{Z}$. For showing that N_e is an intermediate VPL we first show that N_e is quasi-counterfree. To this end we show that $(\varphi_{N_e}, \psi_{N_e})$ of N_e is quasi-aperiodic, which is equivalent by Proposition 4.18.

Assume by contradiction that $(\varphi_{N_e}, \psi_{N_e})$ is not quasi-aperiodic. Then there exist $k, l \in \mathbb{N}$ such that $\psi_{N_e}(\mathcal{O}(\Upsilon^{\triangle})^{k,l})$ contains a non-trivial group, say $G \subseteq O_{N_e}$. Let g_0 be the identity of G and let $g_1 \in G$ be such that $g_1 \neq g_0$. Thus, we have $g_1^{i+1} \neq g_1^i$ for all $\mathbb{N}_{>0}$. We claim that all $g \in G$ are $\varphi_{N_e}(N_e)$ reachable: indeed, if $g \in G$ were not $\varphi_{N_e}(N_e)$-reachable, then the same would hold for all $g' \in G$ since $g' = g'g^{-1}g$, hence implying that g' is the (one and only) zero of O_{N_e}, contradicting that G is non-trivial. Fix some $\mathrm{ext}_{u_0,v_0}, \mathrm{ext}_{u_1,v_1} \in \mathcal{O}(\Upsilon^{\triangle})^{k,l}$ such that $\psi_{N_e}(\mathrm{ext}_{u_0,v_0}) = g_0$ and $\psi_{N_e}(\mathrm{ext}_{u_1,v_1}) = g_1$. Note that we must have $k > 0$ or $l > 0$ for otherwise we would have $u_0 = u_1 = \varepsilon$ and $v_0 = v_1 = \varepsilon$, a contradiction to $g_0 \neq g_1$. We moreover claim that $u_0, u_1 \notin \widetilde{\Sigma}^*$ and hence $v_0, v_1 \notin \overline{\Sigma}^*$: indeed, a consequence of Remark 5.18 is that $\mathrm{ext}_{u,v} \in \psi_{N_e}^{-1}(f) \cap \widetilde{\Sigma}^s \times \overline{\Sigma}^t$ implies $s, t \leq n$ for all $\varphi_{N_e}(N_e)$-reachable $f \in O_{N_e}$, yet the fact that $\psi_{N_e}(\mathrm{ext}_{u^i,v^i})$ is in G and thus $\varphi_{N_e}(N_e)$-reachable for all $i \geq 1$ contradicts this (recall that $|u_1| > 0$ or $|v_1| > 0$). In other words, both u_0 and u_1 contain at least one letter from Σ_{call} and both v_0 and v_1 contain a letter from Σ_{ret}.

Next we claim that neither u_0 nor u_1 contain any letter from Σ_{ret}: indeed, without loss of generality if u_1 were to contain a letter from Σ_{ret} then $\psi_{N_e}(\mathrm{ext}_{u_1^2,v_1^2})$ would be the zero of O_{N_e}, contradicting that $\psi_{N_e}(\mathrm{ext}_{u_1^2,v_1^2})$ is in G. It follows that $u_0, u_1 \in (\widetilde{\Sigma}^*\Sigma_{\mathrm{call}}\widetilde{\Sigma}^*)^+$ and hence that $v_0, v_1 \in (\overline{\Sigma}^*\Sigma_{\mathrm{ret}}\overline{\Sigma}^*)^+$.

Let $u_0' = \widetilde{\vartheta}^{-1}(u_0)$, $u_1' = \widetilde{\vartheta}^{-1}(u_1)$, $v_0' = \overline{\vartheta}^{-1}(v_0)$, $v_1' = \overline{\vartheta}^{-1}(v_1)$ and note that $\mathrm{ext}_{u_0',v_0'}, \mathrm{ext}_{u_1',v_1'} \in \mathcal{O}(\Sigma^{k,l})$. Since $\mathfrak{P}(O_L)$ forms a monoid there exists $p \in \mathbb{N}_{>0}$ such that

$$\psi_L(\{\mathrm{ext}_{u_0',v_0'}, \mathrm{ext}_{u_1',v_1'}\})^p = \psi_L(\{\mathrm{ext}_{u_0',v_0'}, \mathrm{ext}_{u_1',v_1'}\})^{2p} \ .$$

This implies that for all $i \in \mathbb{N}_{>0}$, we have

$$\psi_L(\mathrm{ext}_{u_0'^{p-1}u_1',v_1'v_0'^{p-1}})^i \in \psi_L(\{\mathrm{ext}_{u_0',v_0'}, \mathrm{ext}_{u_1',v_1'}\})^{ip} = \psi_L(\{\mathrm{ext}_{u_0',v_0'}, \mathrm{ext}_{u_1',v_1'}\})^p \ .$$

Hence, as $|u_0| = |u_1| = |u_0'| = |u_1'| = k$ and $|v_0| = |v_1| = |v_0'| = |v_1'| = l$, the semigroup

$\{\psi_L(\mathrm{ext}_{u_0'^{p-1}u_1',v_1'v_0'^{p-1}})^i \mid i \in \mathbb{N}_{>0}\}$ is contained in $\psi_L(\mathcal{O}(\Sigma^\triangle)^{kp,lp})$. We show that the latter semigroup is not aperiodic by showing $\psi_L(\mathrm{ext}_{u_0'^{p-1}u_1',v_1'v_0'^{p-1}})^i \neq \psi_L(\mathrm{ext}_{u_0'^{p-1}u_1',v_1'v_0'^{p-1}})^{i+1}$ for all $i \in \mathbb{N}_{>0}$. Indeed, since

$$\psi_{N_e}\left(\left(\mathrm{ext}_{u_0^{p-1}u_1,v_1v_0^{p-1}}\right)^{i+1}\right) = g_1^{i+1} \neq g_1^{i+2} = \psi_{N_e}\left(\left(\mathrm{ext}_{u_0^{p-1}u_1,v_1v_0^{p-1}}\right)^{i+2}\right)$$

there exist $\mathrm{ext}_{x,y} \in \mathcal{O}(\Upsilon^\triangle)$ and $w \in \Upsilon^\triangle$ such that, without loss of generality

$$\left(\mathrm{ext}_{x,y} \circ \left(\mathrm{ext}_{u_0^{p-1}u_1,v_1v_0^{p-1}}\right)^{i+1}\right)(w) \in N_e \quad \text{and} \quad \left(\mathrm{ext}_{x,y} \circ \left(\mathrm{ext}_{u_0^{p-1}u_1,v_1v_0^{p-1}}\right)^{i+2}\right)(w) \notin N_e.$$

Now since $u_0, u_1 \in (\widetilde{\Sigma}^*\Sigma_{\mathrm{call}}\widetilde{\Sigma}^*)^+$ there exists a factorization $u_0^{p-1}u_1wv_1v_0^{p-1} = x'y'$ such that

$$\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x),\overline{\vartheta}^{-1}(y)} \circ \left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(u_0^{p-1}u_1),\overline{\vartheta}^{-1}(v_1v_0^{p-1})}\right)^i\right)(\widetilde{\vartheta}^{-1}(x')\#\overline{\vartheta}^{-1}(y')) \in M_e$$

and

$$\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x),\overline{\vartheta}^{-1}(y)} \circ \left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(u_0^{p-1}u_1),\overline{\vartheta}^{-1}(v_1v_0^{p-1})}\right)^{i+1}\right)(\widetilde{\vartheta}^{-1}(x')\#\overline{\vartheta}^{-1}(y')) \notin M_e$$

By definition of M_e we obtain

$$\begin{aligned} e &= e \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x(u_0^{p-1}u_1)^ix',\overline{\vartheta}^{-1}(y'(v_1v_0^{p-1})^iy)}\right) \\ &= e \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x),\overline{\vartheta}^{-1}(y)}\right) \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(u_0^{p-1}u_1),\overline{\vartheta}^{-1}(v_1v_0^{p-1})}\right)^i \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x'),\overline{\vartheta}^{-1}(y')}\right) \end{aligned}$$

and

$$\begin{aligned} e &\neq e \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x(u_0^{p-1}u_1)^{i+1}x',\overline{\vartheta}^{-1}(y'(v_1v_0^{p-1})^{i+1}y)}\right) \\ &= e \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x),\overline{\vartheta}^{-1}(y)}\right) \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(u_0^{p-1}u_1),\overline{\vartheta}^{-1}(v_1v_0^{p-1})}\right)^{i+1} \circ \psi_L\left(\mathrm{ext}_{\widetilde{\vartheta}^{-1}(x'),\overline{\vartheta}^{-1}(y')}\right). \end{aligned}$$

Thus, we must have $\psi_L(\mathrm{ext}_{u_0'^{p-1}u_1',v_1'v_0'^{p-1}})^i \neq \psi_L(\mathrm{ext}_{u_0'^{p-1}u_1',v_1'v_0'^{p-1}})^{i+1}$, as required. As this is true for each $i \in \mathbb{N}_{>0}$, the semigroup $\{\psi_L(\mathrm{ext}_{u_0'^{p-1}u_1',v_1'v_0'^{p-1}})^i \mid i \in \mathbb{N}_{>0}\}$ that is contained in $\psi_L(\mathcal{O}(\Sigma^\triangle)^{kp,lp})$ is not aperiodic, contradicting quasi-aperiodicity of (φ_L, ψ_L).

It remains to show that one can compute a vertically visibly pushdown grammar G_e with $L(G_e) = N_e$ such that $\mathcal{R}(G_e)$ is weakly length-synchronous but not length-synchronous. By Remark 5.18 each non-empty word in N_e is of the form

$$(u\#v)^\ddagger = \mathrm{ext}_{\widetilde{\vartheta}(x_1),\overline{\vartheta}(y_1)} \circ \mathrm{ext}_{a_1,b_1} \circ \cdots \circ \mathrm{ext}_{\widetilde{\vartheta}(x_k),\overline{\vartheta}(y_k)} \circ \mathrm{ext}_{a_k,b_k} \circ \mathrm{ext}_{\widetilde{\vartheta}(x_{k+1}),\overline{\vartheta}(y_{k+1})}(\varepsilon)$$

for some $k \geq 1$ and some words $x_1, \ldots, x_{k+1}, y_1, \ldots, y_{k+1} \in \Sigma^\triangle$ all of which have length at most n such that moreover $e \circ \psi_L(\mathrm{ext}_{u,v}) = e$. We construct the grammar $G_e = (V, \Upsilon, P, S_e)$ as follows. The set of nonterminals is $V = \{S_f \mid f \in O_L\}$, $S_e \in V$ is the start nonterminal, the set of productions

consists of the union of $\{S_e \to_{G_e} \varepsilon\}$ and

$$\Big\{ S_f \to_{G_e} \widetilde{\vartheta}(x_1) a \widetilde{\vartheta}(x_2) S_g \overline{\vartheta}(y_2) b \overline{\vartheta}(y_1) \;\Big|\; f, g \in O_L, x_1, x_2, y_1, y_2 \in \Sigma^{\triangle}, |x_1|, |x_2|, |y_1|, |y_2| \leq n, \\ a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}, f \circ \psi_L(\text{ext}_{x_1 a x_2, y_2 b y_1}) = g \Big\}.$$

As a consequence we obtain

$$\mathcal{R}(G_e) = \Big\{ \Big(\widetilde{\vartheta}(x_1) \dots a_k \widetilde{\vartheta}(x_{x+1}), \overline{\vartheta}(y_{k+1}) b_k \dots \overline{\vartheta}(y_1) \Big) \;\Big|\; x_1, \dots, x_{k+1}, x_1, \dots, y_{k+1} \in \Sigma^{\triangle}, \\ a_1, \dots, a_k \in \Sigma_{\text{call}}, b_1, \dots, b_k \in \Sigma_{\text{ret}}, \\ e \circ \psi_L(\text{ext}_{a_1 \dots a_k x_{k+1}, y_{k+1} b_k \dots y_1}) = e \Big\}$$

$$= \Big\{ \Big(\widetilde{\vartheta}(x_1) \dots a_k \widetilde{\vartheta}(x_{x+1}), \overline{\vartheta}(y_{k+1}) b_k \dots \overline{\vartheta}(y_1) \Big) \;\Big|\; x_1, \dots, x_{k+1}, x_1, \dots, y_{k+1} \in \Sigma^{\triangle}, \\ a_1, \dots, a_k \in \Sigma_{\text{call}}, b_1, \dots, b_k \in \Sigma_{\text{ret}}, \\ (x_1 \dots a_k x_{k+1}, y_{k+1} b_k \dots y_1) \in \mathcal{U}_e \Big\}.$$

It follows that $\mathcal{R}(G_e)$ is weakly length-synchronous since $\mathcal{U}_e$ is: indeed, if $\mathcal{R}(G_e)$ were not weakly length-synchronous and without loss of generality there were to exist $(u, v), (u', v') \in \mathcal{R}(G_e)$ with $u = u'$ and $|v| \neq |v'|$, then both $(\widetilde{\vartheta}^{-1}(u), \overline{\vartheta}^{-1}(v))$ and $(\widetilde{\vartheta}^{-1}(u'), \overline{\vartheta}^{-1}(v'))$ would be in $\mathcal{U}_e$ by definition. Yet $\widetilde{\vartheta}^{-1}(u) = \widetilde{\vartheta}^{-1}(u')$ and $|\overline{\vartheta}^{-1}(v)| = |v| \neq |v'| = \overline{\vartheta}^{-1}(v')$, so this would contradict that $\mathcal{U}_e$ is indeed weakly length-synchronous. Analogously it follows that $\mathcal{R}(G_e)$ is not length-synchronous since $\mathcal{U}_e$ is not length-synchronous by assumption. □

Lemma 5.20. *$N_e \leq_{cd} M_e$ for all $e \in \mathcal{Z}$.*

Proof. Assume we are given $w \in \Upsilon^*$. To decide if $w \in N_e$ using an oracle to M_e we do the following constant-depth computation:

1. Accept if $w = \varepsilon$, otherwise continue.
2. Check if $w = uv$ for some $u \in (\widetilde{\Sigma} \cup \Sigma_{\text{call}})^*$ and some $v \in (\overline{\Sigma} \cup \Sigma_{\text{ret}})^*$, reject if this is not the case.
3. Check whether u can be factorized as $u = x_1 a_1 \cdots x_k a_k x_{k+1}$, where $k \geq 1$, $x_1, \dots, x_{k+1} \in \{x \in \widetilde{\Sigma}^* \mid |x| \leq n \wedge \widetilde{\vartheta}^{-1}(x) \in \Sigma^{\triangle}\}$ and $a_1, \dots, a_k \in \Sigma_{\text{call}}$ and whether v can be factorized as $v = y_{l+1} b_l y_l \cdots a_1 y_1$, where $l \geq 1$, $y_1, \dots, y_{l+1} \in \{y \in \overline{\Sigma}^* \mid |y| \leq n \wedge \overline{\vartheta}^{-1}(y) \in \Sigma^{\triangle}\}$ and $b_1, \dots, b_l \in \Sigma_{\text{ret}}$. Reject if it is not possible. (Observe that this is doable by a constant depth and polynomial size circuit family since we test membership in finite sets that do not depend on the input.)
4. Finally accept if, and only if, the word

$$\widetilde{\vartheta}^{-1}(x_1) a_1 \cdots \widetilde{\vartheta}^{-1}(x_{k-1}) a_k \widetilde{\vartheta}^{-1}(x_{k+1}) \# \overline{\vartheta}^{-1}(y_{l+1}) b_l \overline{\vartheta}^{-1}(y_{l+1}) \cdots b_1 \overline{\vartheta}^{-1}(y_1)$$

is in M_e. □

Lemma 5.21. *$M_e \leq N_e$ for all $e \in \mathcal{Z}$.*

Proof. Assume we are given $w \in (\Sigma \cup \{\#\})^*$, where $w = w_1 \cdots w_m$ and where $w_i \in \Sigma \cup \{\#\}$ for all $i \in [1, m]$. To decide if $w \in M_e$ using an oracle to N_e we do the following constant-depth computation:

1. Check if $w = u\#v$ for some $u \in \Sigma^+$ and some $v \in \Sigma^+$, reject otherwise.
2. For all return letters $b \in \Sigma_{\text{ret}}$ and all positions j within u at which b appears, check whether there exists a position i within u such that $1 \leq j - i \leq n - 1$ and the infix $w_i \cdots w_j$ is in $\Sigma^{\triangle}$. (As above, this is doable by a constant depth and polynomial size circuit family since we check well-matchedness of at most a fixed number of words that does not depend on the input.) Reject if it is not the case.
3. For all call letters $a \in \Sigma_{\text{call}}$ and all positions i within v at which a appears, check whether there exists a position j within v such that $1 \leq j - i \leq n - 1$ and the infix $w_i \cdots w_j$ is in $\Sigma^{\triangle}$. Reject if it is not the case.
4. For each position i within u, compute $P_{\text{call}}(i)$ where P_{call} is the unary predicate defined by $w \models P_{\text{call}}(i)$ if, and only if, i is a position within u, $w_i \in \Sigma_{\text{call}}$, and there does not exist any position j within u such that $1 \leq j - i \leq n - 1$ and the infix $w_i \cdots w_j$ is in $\Sigma^{\triangle}$.
5. For each position j within v, compute $P_{\text{ret}}(j)$ where P_{ret} is the unary predicate defined by $w \models P_{\text{ret}}(j)$ if, and only if, j is a position within v, $w_j \in \Sigma_{\text{ret}}$, and there does not exist any position i within v such that $1 \leq j - i \leq n - 1$ and the infix $w_i \cdots w_j$ is in $\Sigma^{\triangle}$.
6. Let $1 \leq i_1 < i_2 \cdots < i_k \leq |u|$ be an enumeration of $\{i \in [1, |u|] \mid w \models P_{\text{call}}(i)\}$ and let $|u| + 2 \leq j_l < j_{l-1} \cdots < j_1 \leq m$ be an enumeration of $\{j \in [|u| + 2, m] \mid w \models P_{\text{ret}}(j)\}$. Build
$$u' = \widetilde{\vartheta}(w_1 \cdots w_{i_1 - 1}) w_{i_1} \cdots \widetilde{\vartheta}(w_{i_{k-1}+1} \cdots w_{i_k - 1}) w_{i_k} \widetilde{\vartheta}(w_{i_{k+1}+1} \cdots w_{|u|})$$
and
$$v' = \overline{\vartheta}(w_{|u|+2} \cdots w_{j_{l+1}-1}) w_{j_l} \cdots \overline{\vartheta}(w_{j_2} + 1 \cdots w_{j_1 - 1}) w_{j_1} \overline{\vartheta}(w_{j_1} + 1 \cdots w_m).$$
7. Accept if, and only if, the word $u'v'$ is in N_e. □

Lemma 5.22. *$M_e \leq_{cd} \biguplus_{f \in O_L \text{ is } \varphi_L(L)\text{-reachable}} L_f$ for all $e \in \mathcal{Z}$.*

Proof. Note that the following equivalence holds:

$$u\#v \in M_e \iff \exists f \in O_L \text{ that is } \varphi_L(L)\text{-reachable} : e \circ f = e \wedge u\#v \in L_f \wedge \Delta(u) > 0 \ .$$

This holds because for $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^{\triangle})$ satisfying $e \circ \psi_L(\text{ext}_{u,v}) = e$, as e is $\varphi_L(L)$-reachable, $\psi_L(\text{ext}_{u,v})$ must also be $\varphi_L(L)$-reachable. Assume we are given $w \in (\Sigma \cup \{\#\})^*$. To decide if $w \in M_e$ we do the following constant-depth computation using oracles to $\biguplus_{f \in O_L \text{ is } \varphi_L(L)\text{-reachable}} L_f$:

1. Check if $w = u\#v$ for some $u, v \in \Sigma^*$, reject otherwise.
2. Check if $u\#v \in L_f$ for some $\varphi_L(L)$-reachable $f \in O_L$ satisfying $e \circ f = e$, reject otherwise.
3. Finally, accept if, and only if for all $\varphi_L(L)$-reachable $f \in O_L$ we have $u\# \notin L_f$.

If the second check is successful, then $\psi_L(\text{ext}_{u,v})$ is necessarily $\varphi_L(L)$-reachable, so in that case when $\Delta(u) = 0$ it holds that $u \in \Sigma^{\triangle}$ and $\psi_L(\text{ext}_{u,\varepsilon})$ is $\varphi_L(L)$-reachable. Hence, in combination with the second check, the third check is successful if, and only if $\Delta(u) > 0$. □

Lemma 5.23. $L \leq_{cd} \biguplus_{e \in \mathcal{Z}} M_e$.

Proof. By assumption L is weakly length-synchronous but not length-synchronous, and its syntactic morphism (φ_L, ψ_L) is quasi-aperiodic. There is a constant d_L such that all words in L have nesting depth at most d_L by Proposition 4.14.

By the first point of Proposition 5.10 we may assume that the evaluation language $\mathcal{E}_{\varphi_L,r}$ is regular and its syntactic morphism is quasi-aperiodic for all $\varphi_L(L)$-reachable $r \in R_L$. This implies, by [29, Theorem VI.4.1], that for each $\varphi_L(L)$-reachable $r \in R_L$, there exists an $\mathsf{FO}_{\Gamma_{\varphi_L}}[<, \mathsf{MOD}]$-sentence $\nu_{\varphi_L,r}$ defining $\mathcal{E}_{\varphi_L,r}$.

As L is not length-synchronous we cannot assume analogous sentences for the evaluation languages $\mathcal{E}_{\psi_L,e}$ for all $\varphi_L(L)$-reachable $e \in O_L$. Indeed, Remark 5.11 provides an example of a weakly length-synchronous but non-length-synchronous VPL whose syntactic morphism is quasi-aperiodic but for which some evaluation language $\mathcal{E}_{\psi_L,e}$ for $e \in O_L$ that is $\varphi_L(L)$-reachable has a non-quasi-aperiodic syntactic morphism.

However, let $e \in O_L$ be such that $\mathcal{R}_e = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi_L(\mathrm{ext}_{u,v}) = e\}$ is length-synchronous. Take any $\varphi_L(L)$-reachable idempotent $f \in O_L$ such that there exist $g, h \in O_L$ satisfying $e = g \circ f \circ h$. There exist $\mathrm{ext}_{x_g,y_g}, \mathrm{ext}_{x_h,y_h} \in \mathcal{O}(\Sigma^{\triangle})$ such that $\psi_L(\mathrm{ext}_{x_g,y_g}) = g$ and $\psi_L(\mathrm{ext}_{x_h,y_h}) = h$. Let $(u,v), (u',v') \in \mathsf{Con}(\Sigma)$ such that $\Delta(u), \Delta(u') > 0$ and $\psi_L(\mathrm{ext}_{u,v}) = \psi_L(\mathrm{ext}_{u',v'}) = f$. Because f is idempotent, we have that $\psi_L(\mathrm{ext}_{u^{|v'|},v^{|v'|}}) = \psi_L(\mathrm{ext}_{u'^{|v|},v'^{|v|}}) = f$, thus $\psi_L(\mathrm{ext}_{x_g u^{|v'|} x_h, y_h v^{|v'|} y_g}) = \psi_L(\mathrm{ext}_{x_g u'^{|v|} x_h, y_h v'^{|v|} y_g}) = e$. Therefore, because of length-synchronicity of $\mathcal{R}_e$, it follows that

$$\begin{aligned}
\frac{\left|x_g u^{|v'|} x_h\right|}{\left|y_h v^{|v'|} y_g\right|} &= \frac{\left|x_g u'^{|v|} x_h\right|}{\left|y_h v'^{|v|} y_g\right|} \\
\frac{|x_g| + |v'| \cdot |u| + |x_h|}{|y_h| + |v'| \cdot |v| + |y_g|} &= \frac{|x_g| + |v| \cdot |u'| + |x_h|}{|y_h| + |v| \cdot |v'| + |y_g|} \\
\left|v'\right| \cdot |u| &= |v| \cdot \left|u'\right| \\
\frac{|u|}{|v|} &= \frac{|u'|}{|v'|} \,.
\end{aligned}$$

So we can conclude that $\mathcal{R}_f = \{(u,v) \in \mathsf{Con}(\Sigma) \mid \psi_L(\mathrm{ext}_{u,v}) = f\}$ is length-synchronous. Thus, by the second point of Proposition 5.10 we may assume that the evaluation language $\mathcal{E}_{\psi_L,e}$ is regular and its syntactic morphism is quasi-aperiodic for all $\varphi_L(L)$-reachable $e \in O_L$ with $\mathcal{R}_e$ length-synchronous. This implies again, by [29, Theorem VI.4.1], that for each $\varphi_L(L)$-reachable $e \in O_L$ with $\mathcal{R}_e$ length-synchronous, there exists an $\mathsf{FO}_{\Gamma_{\psi_L}}[<, \mathsf{MOD}]$-sentence $\nu_{\psi_L,e}$ defining $\mathcal{E}_{\psi_L,e}$.

For proving $L \leq_{\mathrm{cd}} \biguplus_{e \in \mathcal{Z}} M_e$ we must thus make use of the oracles to $\biguplus_{e \in \mathcal{Z}} M_e$. All of the following predicates can be computed by a circuit family of constant depth and polynomial size with access to these oracles. More concretely, by accessing oracles to $\biguplus_{e \in \mathcal{Z}} M_e$, for all $e \in \mathcal{Z}$ we may assume that we have a predicate π_e^{exact} such that for all $w \in \Sigma^+$ and $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$ the following holds:

$$w \models \pi_e^{\mathsf{exact}}(i,i',j',j) \quad \Longleftrightarrow \quad \begin{aligned} &w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}, e \circ \psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e \text{ and} \\ &\Delta(w_i \cdots w_{i'}) > 0 \end{aligned} \tag{4}$$

For all $\varphi_L(L)$-reachable $e \in O_L$ that are *not in* $\mathcal{Z}$ we may assume, by Lemma 5.13, that we have the $\mathsf{FO}[+]$-definable (and hence constant-depth computable) predicate π_e at hand. It has the following properties: for all $w \in \Sigma^+$ and $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$:

- if $w \models \pi_e(i,i',j',j)$, then $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}$ and

- if $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}$, $\Delta(w_i \cdots w_{i'}) > 0$ and $e \circ \psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$, then $w \models \pi_e(i, i', j', j)$.

We can first build an approximate matching μ relative to L. This is done totally analogously as done in Section 5.3.4 by replacing the there appearing π_e for each $e \in \mathcal{Z}$ by our predicate π_e^{exact}: indeed, Remark 5.15 states that the predicates π_e from of Lemma 5.13 could have been replaced by the predicate π_e^{exact}.

Thus, as in the proof of Proposition 5.16 we may assume that we have full access to the matching relation $M^{\Delta}(w)$ of our input word w.

For verifying if a given word $w \in \Sigma^{\triangle}$ is in L we follow the same approach as the main construction in Section 5.3.5. It is however important to stress that this time we cannot assume quasi-aperiodicity of the syntactic morphisms of the evaluation languages $\mathcal{E}_{\psi_L, e}$ for all $\varphi_L(L)$-reachable $e \in O_L$. Still, we build formulas

- $\overline{\eta}^{\uparrow}_{d,r}(x, y)$ for all $d \in [0, d_L]$ and all $\varphi_L(L)$-reachable $r \in R_L$ and
- $\overline{\eta}_{d,r}(x, y)$ for all $d \in [0, d_L]$ and all $\varphi_L(L)$-reachable $r \in R_L$

that will have the following properties (as $\eta_{d,r}$ and $\eta^{\uparrow}_{d,r}$) for all $w \in \Sigma^{\triangle}$ and all $i, j \in [1, |w|]$:

- if i is matched to j in w, then $(w, M^{\triangle}(w)) \models \overline{\eta}^{\uparrow}_{d,r}(i, j)$ if, and only if, $\mathrm{nd}(w_i \cdots w_j) \leq d$ and $\varphi_L(w_i \cdots w_j) = r$;
- if $w_i \cdots w_j \in \Sigma^{\triangle}$, then $(w, M^{\triangle}(w)) \models \overline{\eta}_{d,r}(i, j)$ if, and only if, $\mathrm{nd}(w_i \cdots w_j) \leq d$ and $\varphi_L(w_i \cdots w_j) = r$.

It remains to define the formulas $\overline{\eta}_{d,r}$ and $\overline{\eta}^{\uparrow}_{d,r}$ for all $d \in [0, d_L]$ and all $\varphi_L(L)$-reachable $r \in R_L$. For the definition of the $\overline{\eta}^{\uparrow}_{0,r}$ and the $\overline{\eta}_{0,r}$ we can simply reuse $\eta_{0,r}$ and $\eta^{\uparrow}_{0,r}$ as in the proof of Proposition 5.16 respectively ($\eta_{0,r}$ will make use of our sentence $\nu_{\varphi_L, r}$). So let us assume $d > 0$.

We first construct for all $\varphi_L(L)$-reachable $e \in O_L$ a formula $\overline{\chi}_{d,e}(x, x', y', y)$ such that for all $w \in \Sigma^{\triangle}$ and $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$ we have that if $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\triangle}$ and i' is matched to j' in w, then given

$$\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j} = \mathrm{ext}_{x_1, y_1} \circ \mathrm{ext}_{a_1, b_1} \circ \cdots \circ \mathrm{ext}_{x_k, y_k} \circ \mathrm{ext}_{a_k, b_k}$$

the stair factorization of $\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}$ provided by Lemma 3.6, we have $(w, M^{\triangle}(w)) \models \overline{\chi}_{d,e}(i, i', j', j)$ if, and only if, $\mathrm{nd}(x_\ell), \mathrm{nd}(y_\ell) < d$ for all $\ell \in [1, k]$ and $\psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$. Given w, i, i', j', j and the associated stair factorization as above, note that if $\psi_L(\mathrm{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$, as e is $\varphi_L(L)$-reachable, then $\varphi_L(x_\ell)$ and $\varphi_L(y_\ell)$ are $\varphi_L(L)$-reachable for all $\ell \in [1, k]$. If additionally $\mathrm{nd}(x_\ell), \mathrm{nd}(y_\ell) < d$ for all $\ell \in [1, k]$, we can inductively make use of the formulas $\{\overline{\eta}_{d-1,r'} \mid r' \in R_L \ \varphi_L(L)\text{-reachable}\}$ in order to evaluate $\varphi_L(x_1), \varphi_L(y_1), \ldots, \varphi_L(x_k), \varphi_L(y_k)$.

As expected, the problems are, firstly, that we cannot access our evaluation languages $\mathcal{E}_{\psi_L, e}$ and, secondly, that we have to build a formula that may not depend on k. As in Section 5.3.4 we define the product

$$e_{\ell, \ell'} = \psi_L(\mathrm{ext}_{x_\ell, y_\ell} \circ \mathrm{ext}_{a_\ell, b_\ell} \circ \cdots \circ \mathrm{ext}_{x_{\ell'}, y_{\ell'}} \circ \mathrm{ext}_{a_{\ell'}, b_{\ell'}}) \quad \text{and} \quad e_\ell = e_{1,\ell}$$

for all $\ell, \ell' \in [1, k]$. For $e \in O_L$ we say an interval $I = [s, t] \subseteq [1, k]$ is *e-repetitive* if $s < t$ and $e_s = e_t$. We say $[s, t] \subseteq [1, k]$ is *repetitive* if it is e-repetitive for some $e \in O_L$.

By Claim 5.14 there exist indices $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_q < t_q \leq s_{q+1} = k$ such that $[s_1, t_1], \ldots, [s_q, t_q]$ are all repetitive and for $D_0 = [t_0, s_1], D_1 = [t_1, s_2], \ldots, D_q = [t_q, s_{q+1}]$

we have $q + \sum_{p=0}^{q} |D_p| \leq 3\,|O_L|$. Let $i = i_1 < \cdots < i_k$ and $j_k < \cdots < j_1 = j$ be the positions that correspond to the positions of the letters $a_1, \ldots, a_k \in \Sigma_{\text{call}}$ and $b_k, \ldots, b_1 \in \Sigma_{\text{ret}}$ of the factorization of $\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}$ in w, respectively: more precisely $i_\ell = i + |x_1 \cdots a_{\ell-1} x_\ell|$ and $j_\ell = \left|x_1 a_1 \cdots x_k a_k w_{i'+1} \cdots w_{j'-1} b_k y_k \cdots b_{\ell+1} y_{\ell+1}\right| + 1$ for all $\ell \in [1, k]$. Since the non-empty interval $[s_p, t_p]$ is repetitive for all $p \in [1, q]$, we have $e_{s_p} = e_{t_p}$ and thus obtain

$$e_{s_p} = e_{t_p} = e_{s_p} \circ \psi(\text{ext}_{x_{s_p+1} \cdots a_{t_p}, b_{t_p} \cdots y_{s_p+1}}) \,.$$

Given $p \in [1, q]$, if $e_{s_p} \in \mathcal{Z}$, we can use the predicate $\pi^{\mathsf{exact}}_{e_{s_p}}$ to check the above equality; we set $\theta_{d, e_{s_p}}(x, x', y', y) = \pi^{\mathsf{exact}}_{e_{s_p}}(x, x', y', y)$. If $e_{s_p} \notin \mathcal{Z}$ and is $\varphi_L(L)$-reachable, then $\mathcal{U}_{e_{s_p}}$ is length-synchronous, so for all $e' \in O_L$ such that $e_{s_p} \circ e' = e_{s_p}$, we have that e' is $\varphi_L(L)$-reachable and $\mathcal{R}_{e'} = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\triangle, \Delta(u) > 0, \psi_L(\text{ext}_{u,v}) = e'\}$ is length-synchronous. So to check the above equality, we can use the formula $\theta_{d, e_{s_p}}(x, x', y', y)$ built by taking the disjunction over all $e' \in O_L$ such that $e_{s_p} \circ e' = e_{s_p}$ of the formulas defined inductively completely analogously as $\chi_{d,e'}(x, x', y', x')$ in Section 5.3.5 (using the sentence $\nu_{\psi_L, e'}$ defining $\mathcal{E}_{\psi_L, e'}$): we simply replace every occurrence of $\eta_{d-1,r}$ by $\overline{\eta}_{d-1,r}$.

Next, for all $m > 0$ and all $\varphi_L(L)$-reachable $f \in O_L$ we will construct a formula $\alpha_{d,m,f}(x, x', y', y)$ such that for all $w \in \Sigma^\triangle$ and $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$ we have that if $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\triangle$ and i' is matched to j' in w, then given

$$\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j} = \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \cdots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}$$

the stair factorization of $\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}$ provided by Lemma 3.6, we have $(w, M^\triangle(w)) \models \alpha_{d,m,f}(i, i', j', j)$ if, and only if, $\text{nd}(x_\ell), \text{nd}(y_\ell) < d$ for all $\ell \in [1, k]$, $\Delta(w_i \cdots w_{i'}) = -\Delta(w_{j'} \cdots w_j) = m$ and $\psi_L(\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$. For $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_m) \in \Sigma^m_{\text{call}}$, $\boldsymbol{\xi} = (\boldsymbol{\xi}_1 \ldots, \boldsymbol{\xi}_m) \in \Sigma^m_{\text{ret}}$, $\boldsymbol{r} = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m) \in {R_L}^m$, and $\boldsymbol{r}^\dagger = (\boldsymbol{r}^\dagger_1, \ldots, \boldsymbol{r}^\dagger_m) \in {R_L}^m$ we define

$$\prod(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{r}, \boldsymbol{r}^\dagger) = \bigcirc_{g=1}^{m} \text{left}_{\boldsymbol{r}_g} \circ \text{right}_{\boldsymbol{r}^\dagger_g} \circ \psi_L(\text{ext}_{\boldsymbol{\sigma}_g, \boldsymbol{\xi}_g}).$$

The formula $\alpha_{d,m,f}$ can be expressed as follows:

$$\begin{aligned}
\alpha_{d,m,f}(x, x', y', y) = \bigvee_{\substack{\boldsymbol{\sigma} \in \Sigma^m_{\text{call}}, \boldsymbol{\xi} \in \Sigma^m_{\text{ret}} \\ \boldsymbol{r}, \boldsymbol{r}^\dagger \in {R_L}^m : f = \prod(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{r}, \boldsymbol{r}^\dagger)}} & \exists x_1, \ldots, x_m \exists y_1, \ldots, y_m \\
& \Big(x' = x_m \wedge y' = y_m \wedge x \leq x_1 < \cdots < x_m < y_m < \cdots < y_1 \leq y \wedge \\
& \bigwedge_{g=1}^{m} \big(\boldsymbol{\sigma}_g(x_g) \wedge \boldsymbol{\xi}_g(y_g) \wedge x_g \leftrightsquigarrow y_g\big) \wedge \\
& \forall z \big((x \leq z \leq x' \wedge \bigwedge_{g=1}^{m} z \neq x_g) \to \neg U(x, x', z) \big) \wedge \\
& \Big(\big(x < x_1 \wedge \overline{\eta}_{d-1, \boldsymbol{r}_1}(x, x_1 - 1)\big) \vee F_{\boldsymbol{r}_1}(x, x_1) \Big) \wedge \\
& \Big(\big(y_1 < y \wedge \overline{\eta}_{d-1, \boldsymbol{r}^\dagger_1}(y_1 + 1, y)\big) \vee F_{\boldsymbol{r}^\dagger_1}(y_1, y) \Big) \wedge \\
& \bigwedge_{g=2}^{m} \overline{\eta}_{d-1, \boldsymbol{r}_g}(x_{g-1} + 1, x_g - 1) \wedge \bigwedge_{g=2}^{m} \overline{\eta}_{d-1, \boldsymbol{r}^\dagger_g}(y_g + 1, y_{g-1} - 1) \Big)
\end{aligned}$$

with $F_s(x,z) = \begin{cases} x = z & \text{if } s = 1_{R_L} \\ \bot & \text{otherwise} \end{cases}$ for all $\varphi_L(L)$-reachable $s \in R_L$.

We are now ready to define the formula $\overline{\chi}_{d,e}$:

$$\begin{aligned}
\overline{\chi}_{d,e}(x,x',y',x) \quad = \quad & \bigvee_{\substack{q\in[0,|O|] \\ d_0,\dots,d_q\geq 1: \\ q+d_0+\cdots+d_q\leq 3|0|}} \bigvee_{\substack{e_0,f_0,e_1,\dots,f_{q-1},e_q,f_q\in O \\ e_0=1_O\wedge\forall j\in[1,q]: e_j=e_{j-1}\circ f_{j-1} \\ e=e_q\circ f_q}} \\
& \exists x_1\dots x_{q+1}\exists x'_0\dots x'_q\exists y_1\dots y_{q+1}\exists y'_1\dots y'_0 \\
& \Bigg(\quad x'_0 \leq x_1 < x'_1 < x_2 < \cdots < x'_q < y'_q < y_q < \cdots < y'_1 < y_1 \leq y'_0 \wedge \\
& \quad x'_0 = x \wedge y'_0 = y \wedge x'_q = x' \wedge y'_q = y' \wedge \\
& \quad \bigwedge_{p=1}^{q} \big(\theta_{e_p}(x_p+1, x'_p, y'_p, y_p-1)\big) \wedge \bigwedge_{p=0}^{q} \alpha_{d_p,f_p}(x'_p+1, x_{p+1}, y_{p+1}, y'_p-1) \Bigg) \ .
\end{aligned}$$

The inductive definition of $\overline{\eta}^{\uparrow}_{d,r}$ is completely analogous to the definition of $\eta_{d,r}$ in Section 5.3.5: we simply replace every occurrence of $\eta_{d-1,r}$ by $\overline{\eta}_{d-1,r}$ and every occurrence of $\chi_{d,e}$ by $\overline{\chi}_{d,e}$

The inductive definition of $\overline{\eta}_{d,r}$ is completely analogous to the definition of $\eta_{d,r}$ in Section 5.3.5: we access the horizontal evaluation languages $\mathcal{E}_{\varphi_L,r}$ for all $\varphi_L(L)$-reachable $r \in R_L$ by making use of the sentence $\nu_{\varphi,r}$ and the already defined $\overline{\eta}^{\uparrow}_{d,r}$. □

5.4.2 Computation of k, l

The following proposition implies the computability of $k, l \in \mathbb{N}$ such that $\mathcal{L}_{k,l} \leq_{\text{cd}} L$ already when VPL L is weakly length-synchronous but not length-synchronous.

Proposition 5.24. *If a VPL L is weakly length-synchronous but not length-synchronous, one can effectively compute $k, l \in \mathbb{N}_{>0}$ with $k \neq l$ such that $\mathcal{L}_{k,l} \leq_{cd} L$.*

Proof. Let $L \subseteq \Sigma^{\triangle}$ be a weakly length-synchronous VPL that is not length-synchronous. According to Point 2 (b) of Proposition 5.1 one can effectively compute a quadruple $(k_0, l_0, k'_0, l'_0) \in \mathbb{N}^4_{>0}$ for which there exist $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^{\triangle})$ such that

- $|u| = k_0$, $|v| = l_0$, $|u'| = k'_0$, $|v'| = l'_0$,
- $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'})$ is a $\varphi_L(L)$-reachable idempotent,
- $\Delta(u), \Delta(u') > 0$, and
- $\frac{k_0}{l_0} = \frac{|u|}{|v|} \neq \frac{|u'|}{|v'|} = \frac{k'_0}{l'_0}$.

We can explicitly compute such $\text{ext}_{u,v}$ and $\text{ext}_{u',v'}$ by just doing an exhaustive search. This enables us to assume without loss of generality while maintaining effective computability that $\Delta(u) = \Delta(u')$: indeed, in case $\Delta(u) \neq \Delta(u')$, we can consider $\text{ext}_{u_1,v_1} = \text{ext}_{u^{\Delta(u')},v^{\Delta(u')}}$ and $\text{ext}_{u_2,v_2} = \text{ext}_{(u')^{\Delta(u)},(v')^{\Delta(u)}}$ satisfying the desired properties.

Let us now define Green's relations on O_L (see [26, Chapter 3, Section 1]). Let us consider two elements x, y of O_L.

- We write $x \leq_{\mathfrak{J}} y$ whenever there are elements e, f of O_L such that $x = e \circ y \circ f$. We write $x \mathfrak{J} y$ if $x \leq_{\mathfrak{J}} y$ and $y \leq_{\mathfrak{J}} x$. We finally write $x <_{\mathfrak{J}} y$ if $x \leq_{\mathfrak{J}} y$ and $x \not\mathfrak{J} y$.
- We write $x \leq_{\mathfrak{R}} y$ whenever there is an element e of O_L such that $x = y \circ e$. We write $x \mathfrak{R} y$ if $x \leq_{\mathfrak{R}} y$ and $y \leq_{\mathfrak{R}} x$.
- We write $x \leq_{\mathfrak{L}} y$ whenever there is an element e of O_L such that $x = e \circ y$. We write $x \mathfrak{L} y$ if $x \leq_{\mathfrak{L}} y$ and $y \leq_{\mathfrak{L}} x$.
- We write $x \mathfrak{H} y$ if $x \mathfrak{R} y$ and $x \mathfrak{L} y$.

Observe that because $\Delta(u) = \Delta(u')$, we have that $uv' \in \Sigma^{\triangle}$ and $u'v \in \Sigma^{\triangle}$, so that we can consider the elements $\mathrm{ext}_{uuu,vv'v} = \mathrm{ext}_{u,v} \circ \mathrm{ext}_{u,v'} \circ \mathrm{ext}_{u,v}$ and $\mathrm{ext}_{uu'u,vvv} = \mathrm{ext}_{u,v} \circ \mathrm{ext}_{u',v} \circ \mathrm{ext}_{u,v}$ in $\mathcal{O}(\Sigma^{\triangle})$. These elements satisfy $\psi_L(\mathrm{ext}_{uuu,vv'v}) \leq_{\mathfrak{J}} \psi_L(\mathrm{ext}_{u,v})$ and $\psi_L(\mathrm{ext}_{uu'u,vvv}) \leq_{\mathfrak{J}} \psi_L(\mathrm{ext}_{u,v})$. We claim that we actually have $\psi_L(\mathrm{ext}_{uuu,vv'v}) <_{\mathfrak{J}} \psi_L(\mathrm{ext}_{u,v})$ and $\psi_L(\mathrm{ext}_{uu'u,vvv}) <_{\mathfrak{J}} \psi_L(\mathrm{ext}_{u,v})$. Indeed, assume we would have $\psi_L(\mathrm{ext}_{uu'u,vvv}) \mathfrak{J} \psi_L(\mathrm{ext}_{u,v})$. Set $x = \psi_L(\mathrm{ext}_{u,v})$ and $y = \psi_L(\mathrm{ext}_{u',v})$. By a classical property of Green's relations (see [26, Chapter 3, Proposition 1.4]), since it would hold that $x \circ y \circ x \leq_{\mathfrak{R}} x$ and $x \circ y \circ x \mathfrak{J} x$, we would have $x \circ y \circ x \mathfrak{R} x$ and dually, since it would hold that $x \circ y \circ x \leq_{\mathfrak{L}} x$ and $x \circ y \circ x \mathfrak{J} x$, we would have $x \circ y \circ x \mathfrak{L} x$. Therefore, we would have $x \circ y \circ x \mathfrak{H} x$. By another classical result on Green's relations [26, Chapter 3, Corollary 1.7], as x is an idempotent, its $\mathfrak{H}$-class is a group, hence for $\omega \in \mathbb{N}_{>0}$ the idempotent power of O_L, we would have $(x \circ y \circ x)^\omega = x^\omega = x$ (as the only idempotent element in a group is the identity). This would finally entail that $\psi_L(\mathrm{ext}_{(uu'u)^\omega,(vvv)^\omega}) = \psi_L(\mathrm{ext}_{(uuu)^\omega,(vvv)^\omega})$ is a $\varphi_L(L)$-reachable idempotent and $\Delta((uu'u)^\omega) = \Delta((uuu)^\omega) > 0$ but $|(uu'u)^\omega| \neq |(uuu)^\omega|$, a contradiction to the fact that (φ_L, ψ_L) is $\varphi_L(L)$-weakly-length-synchronous. Symmetrically, we can prove that if we had $\psi_L(\mathrm{ext}_{uuu,vv'v}) \mathfrak{J} \psi_L(\mathrm{ext}_{u,v})$, this would contradict the fact that (φ_L, ψ_L) is $\varphi_L(L)$-weakly-length-synchronous.

We distinguish three cases. In each of these we prove that there exist $k, l \in \mathbb{N}_{>0}, k \neq l$ such that $\mathcal{L}_{k,l} \leq_{\mathrm{cd}} L_{\psi_L(\mathrm{ext}_{u,v})}$, so that since $L_{\psi_L(\mathrm{ext}_{u,v})} \leq_{\mathrm{cd}} L$ (by Lemma 5.3) and by transitivity of $\leq_{\mathrm{cd}}$ we have $\mathcal{L}_{k,l} \leq_{\mathrm{cd}} L$.

Case $|v| = |v'|$. In that case, we necessarily have $|u| \neq |u'|$. Then, we can exploit the fact that matching u^3 with $vv'v$ or $uu'u$ with v^3 makes us fall down to a smaller $\mathfrak{J}$-class to reduce $\mathcal{L}_{3|u|,2|u|+|u'|}$ to $L_{\psi_L(\mathrm{ext}_{u,v})}$. The constant-depth reduction works as follows on input $w \in \Sigma^*$:

1. check if $w = xy$ with $x \in (ac^{3|u|-1} + ac^{2|u|+|u'|-1})^*$ and $y \in (b_1 + b_2)^*$, reject if it's not the case;
2. build x' by sending $ac^{3|u|-1}$ to u^3, $ac^{2|u|+|u'|-1}$ to $uu'u$ and y' by sending b_1 to v^3 and b_2 to $vv'v$;
3. accept whenever $x'\#y' \in L_{\psi_L(\mathrm{ext}_{u,v})}$.

This forms a valid reduction. Indeed, take a word $w = xy$ with $x \in (ac^{3|u|-1} + ac^{2|u|+|u'|-1})^n$ for $n \in \mathbb{N}$ and $y \in (b_1 + b_2)^m$ for $m \in \mathbb{N}$ and consider $x'\#y'$ produced by the reduction with $x' \in (u^3 + uu'u)^n$ and $y' \in (v^3 + vv'v)^m$. If $w \in \mathcal{L}_{3|u|,2|u|+|u'|}$, then it easily follows that $x'\#y' \in L_{\psi_L(\mathrm{ext}_{u,v})}$. Otherwise, if $w \notin \mathcal{L}_{3|u|,2|u|+|u'|}$, then either $n \neq m$ and thus $x'y'$ is not well-matched because $\Delta(x') = n \cdot 3 \cdot \Delta(u)$ and $\Delta(y') = m \cdot 3 \cdot \Delta(v)$, or $n = m$ and thus $x'y'$ is well-matched, so $\mathrm{ext}_{x',y'} = \mathrm{ext}_{z'_1,t'_1} \circ \cdots \circ \mathrm{ext}_{z'_n,t'_n}$

with $z'_1, \ldots, z'_n \in \{u^3, uu'u\}$ and $t'_1, \ldots, t'_n \in \{v^3, vv'v\}$ such that there exists $i \in [1, n]$ satisfying $\mathrm{ext}_{z'_i,t'_i} \in \{\mathrm{ext}_{u^3,vv'v}, \mathrm{ext}_{uu'u,v^3}\}$, thereby implying

$$\psi_L(\mathrm{ext}_{x',y'}) \leq_{\mathfrak{J}} \psi_L(\mathrm{ext}_{z'_i,t'_i}) <_{\mathfrak{J}} \psi_L(\mathrm{ext}_{u,v}) \ .$$

Our algorithm therefore outputs the pair $(k, l) = (3k_0, 2k_0 + k'_0)$.

Case $|u| = |u'|$. This case is symmetric to the previous case. Our algorithm outputs the pair $(k, l) = (2l_0 + l'_0, 3l_0)$.

Case $|u| \neq |u'|$ and $|v| \neq |v'|$. Then, we can again exploit the fact that matching u^3 with $vv'v$ or $uu'u$ with v^3 makes us fall down to a smaller $\mathfrak{J}$-class to reduce $\mathcal{L}_{A\cdot B', A'\cdot B}$ where $A = 3|u| = 3k_0$, $A' = 2|u| + |u'| = 2k_0 + k'_0$, $B = 3|v| = 3l_0$ and $B' = 2|v| + |v'| = 2l_0 + l'_0$ to $L_{\psi_L(\mathrm{ext}_{u,v})}$. Indeed, we have $A \cdot B' \neq A' \cdot B$ because otherwise we would have

$$\begin{aligned} 3|u| \cdot (2|v| + |v'|) &= (2|u| + |u'|) \cdot 3|v| \\ 6|u||v| + 3|u||v'| &= 6|u||v| + 3|u'||v| \\ |u||v'| &= |u'||v| \ . \end{aligned}$$

The constant-depth reduction works as follows on input $w \in \Sigma^*$:

1. check if $w = xy$ with $x \in (ac^{A\cdot B'-1} + ac^{A'\cdot B-1})^*$ and $y \in (b_1 + b_2)^*$, reject if it is not the case;
2. build x' by sending $ac^{A\cdot B'-1}$ to $(u^3)^{B'}$, $ac^{A'\cdot B-1}$ to $(uu'u)^B$ and y' by sending b_1 to $(v^3)^{B'}$ and b_2 to $(vv'v)^B$;
3. accept whenever $x'\#y' \in L_{\psi_L(\mathrm{ext}_{u,v})}$.

This forms a valid reduction. Indeed, take a word $w = xy$ with $x = z_1 \cdots z_n$ where $n \in \mathbb{N}$ and $z_1, \ldots, z_n \in \{ac^{A\cdot B'-1}, ac^{A'\cdot B-1}\}$ and $y = t_1 \cdots t_m$ where $m \in \mathbb{N}$ and $t_1, \ldots, t_m \in \{b_1, b_2\}$. Consider $x'\#y'$ produced by the reduction with $x' = z'_1 \cdots z'_n$ where $z'_1, \ldots, z'_n \in \{(u^3)^{B'}, (uu'u)^B\}$ and $y' = t'_1 \cdots t'_m$ where $t'_1, \ldots, t'_m \in \{(v^3)^{B'}, (vv'v)^B\}$. If $w \in \mathcal{L}_{A\cdot B', A'\cdot B}$, then it easily follows that $x'\#y' \in L_{\psi_L(\mathrm{ext}_{u,v})}$. Otherwise, if $w \notin \mathcal{L}_{A\cdot B', A'\cdot B}$, three situations can occur.

- There exists $i \in [1, \min\{n, m\}]$ such that $z_1 \cdots z_{i-1} t_{i-1} \cdots t_1 \in \mathcal{L}_{A\cdot B', A'\cdot B}$ but it holds that $(z_i, t_i) \in \{(ac^{A\cdot B'-1}, b_2), (ac^{A'\cdot B-1}, b_1)\}$. Assume first $(z_i, t_i) = (ac^{A'\cdot B-1}, b_1)$. In this case, let $\tilde{x}' = (uu'u)^{B-1} z'_{i+1} \cdots z'_n$ and $\tilde{y}' = t'_n \cdots t'_{i+1} (v^3)^{B'-1}$. If $\Delta(\tilde{x}'\tilde{y}') \neq 0$, then

  $$\Delta(x'y') = \Delta(z'_1 \cdots z'_{i-1}(uu'u)v^3 t'_{i-1} \cdots t'_1) + \Delta(\tilde{x}'\tilde{y}') = \Delta(\tilde{x}'\tilde{y}') \neq 0 \ ,$$

  thus $x'y'$ is not well-matched. Otherwise, if $\Delta(\tilde{x}'\tilde{y}') = 0$, we can show that $\tilde{x}'\tilde{y}'$ is well-matched. Indeed, since $uv \in \Sigma^{\triangle}$, for all $j \in [1, |u|]$, we have $\Delta(u_1 \cdots u_j) \geq 0$ and for all $j \in [1, |v|]$, we have $\Delta(v_j \cdots v_{|v|}) = -\Delta(uv_1 \cdots v_{j-1}) \leq 0$. Similarly, since $u'v' \in \Sigma^{\triangle}$, for all $j \in [1, |u'|]$, we have $\Delta(u'_1 \cdots u'_j) \geq 0$ and for all $j \in [1, |v'|]$, we have $\Delta(v'_j \cdots v'_{|v'|}) = -\Delta(u'v'_1 \cdots v'_{j-1}) \leq 0$. This implies that for all $j \in [1, |\tilde{x}'|]$, we have $\Delta(\tilde{x}'_1 \cdots \tilde{x}'_j) \geq 0$ and for all $j \in [1, |\tilde{y}'|]$, we have $\Delta(\tilde{x}'\tilde{y}'_1 \cdots \tilde{y}'_{j-1}) = -\Delta(\tilde{y}'_j \cdots \tilde{y}'_{|\tilde{y}'|}) \geq 0$. Therefore, $\tilde{x}'\tilde{y}' \in \Sigma^{\triangle}$. Hence $x'y'$ is well-matched and

  $$\mathrm{ext}_{x',y'} = \mathrm{ext}_{z'_1 \cdots z'_{i-1}, t'_{i-1} \cdots t'_1} \circ \mathrm{ext}_{z'_i,t'_i} \circ \mathrm{ext}_{\tilde{x}',\tilde{y}'} \ ,$$

  so that

  $$\psi_L(\mathrm{ext}_{x',y'}) \leq_{\mathfrak{J}} \psi_L(\mathrm{ext}_{z'_i,t'_i}) <_{\mathfrak{J}} \psi_L(\mathrm{ext}_{u,v}) \ .$$

If we assume that $(z_i, t_i) = (ac^{A\cdot B'-1}, b_2)$, then we prove in the same way that either $x'y'$ is not well-matched or it is well-matched and $\psi_L(\mathrm{ext}_{x',y'}) <_{\mathfrak{J}} \psi_L(\mathrm{ext}_{u,v})$.

- It holds that $n < m$ and $z_1 \cdots z_n t_n \cdots t_1 \in \mathcal{L}_{A\cdot B', A'\cdot B}$. This entails that $\Delta(x' t'_n \cdots t'_1) = \Delta(z'_1 \cdots z'_n t'_n \cdots t'_1) = 0$, so that

$$\Delta(x'y') = \Delta(x' t'_n \cdots t'_1) + \Delta(t'_m \cdots t'_{n+1}) = \Delta(t'_m \cdots t'_{n+1}) < 0$$

  because $m > n$ and $\Delta(v) < 0$ as well as $\Delta(v') < 0$. Therefore, $x'y'$ is not well-matched.

- It holds that $n > m$ and $z_1 \cdots z_m t_m \cdots t_1 \in \mathcal{L}_{A\cdot B', A'\cdot B}$. Symmetrically to the previous case, we can also show that then, $x'y'$ is not well-matched.

Hence, our algorithm outputs the pair $(k, l) = (A \cdot B', A' \cdot B) = (3k_0(2l_0 + l'_0), (2k_0 + k'_0)3l_0)$ in this last case. □

5.5 Proof of Corollary 2.11

Let $A = (Q, \Sigma, q_0, F, \delta_0, \ldots, \delta_m)$ be a m-VCA and let $L = L(A)$. One easily computes from A' a DVPA such that $L(A') = L$. Details of this standard translation are omitted. It will be sufficient to prove that L is weakly length-synchronous if, and only if, L is length-synchronous: indeed, one can simply perform the case distinction of Section 5.1 and observe that, under the assumption that weak length-synchronicity and length synchronicity coincide, the algorithm for Theorem 2.9 will either output that L is in AC^0 or some $m \geq 2$ such that $\mathrm{MOD}_m \leq_{\mathrm{cd}} L$.

It thus suffices to prove that if L is not length-synchronous, then L is not weakly length-synchronous. Let (R_L, O_L) be the syntactic Ext-algebra of L along with with its syntactic morphism $(\varphi_L, \psi_L) : (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R_L, O_L)$.

Let $\mathcal{M} = [0, m] \times (Q^Q)^{[0,m]} \times (Q^Q)^{[0,m]}$. The behavior of the m-VCA can be described as follows. To each $\mathrm{ext}_{u,v} \in \mathcal{O}(\Sigma^\triangle)$ we assign the triple $\zeta(\mathrm{ext}_{u,v}) = (j, (f_i)_{i\in[0,m]}, (g_i)_{i\in[0,m]}) \in \mathcal{M}$, where

- $j = \min(\Delta(u), m) \in [0, m]$,
- $f_i(q) = q'$ where $q' \in Q$ is the unique state such that $q(i) \xrightarrow{u}_A q'(i + j)$ for all $q \in Q$ and all $i \in [0, m]$, and
- $g_i(q) = q'$, where $q' \in Q$ is the unique state such that $q'(i + j) \xrightarrow{v}_A q'(i)$ for all $q \in Q$ and all $i \in [0, m]$.

Over $\mathcal{M}$ we define the product

$$(j, (f_i)_{i\in[0,m]}, (g_i)_{i\in[0,m]}) \odot_{\mathcal{M}} (j', (f'_i)_{i\in[0,m]}, (g'_i)_{i\in[0,m]})$$

as

$$(\min(j + j', m), (f'_{\min(i+j,m)} \circ f_i)_{i\in[0,m]}, (g_{\min(i+j,m)} \circ g'_{\min(i+j+j',m)})_{i\in[0,m]}).$$

We claim $\odot_{\mathcal{M}}$ is associative. For this, let us fix

- $m = (j, (f_i)_{i\in[0,m]}, (g_i)_{i\in[0,m]}) \in \mathcal{M}$,
- $m' = (j', (f'_i)_{i\in[0,m]}, (g'_i)_{i\in[0,m]}) \in \mathcal{M}$, and
- $m'' = (j'', (f''_i)_{i\in[0,m]}, (g''_i)_{i\in[0,m]}) \in \mathcal{M}$.

Let $\widetilde{m} = m \odot_{\mathcal{M}} m' = (\widetilde{j}, (\widetilde{f}_i)_{i\in[0,m]}, (\widetilde{g}_i)_{i\in[0,m]})$ and $\overline{m} = m' \odot_{\mathcal{M}} m'' = (\overline{j}, (\overline{f}_i)_{i\in[0,m]}, (\overline{g}_i)_{i\in[0,m]})$. We need to prove $\widetilde{m} \odot_{\mathcal{M}} m'' = m \odot_{\mathcal{M}} \overline{m}$.

Since $j, j', j'', m \geq 0$ we have $\min(\widetilde{j}+j'', m) = \min(\min(j+j', m)+j'', m) = \min(j+j'+j'', m) = \min(j + \min(j' + j'', m), m)) = \min(j + \overline{j}, m)$ associativity holds on the first component. For each $i \in [0, m]$ the i^{th} component of the second component of $\widetilde{m} \odot_{\mathcal{M}} m''$ is

$$\begin{aligned}
f''_{\min(i+\widetilde{j},m)} \circ \widetilde{f}_i &= f''_{\min(i+\widetilde{j},m)} \circ \left(f'_{\min(i+j,m)} \circ f_i\right) \\
&= f''_{\min(i+\min(j+j',m),m)} \circ \left(f'_{\min(i+j,m)} \circ f_i\right) \\
&= f''_{\min(i+j+j',m)} \circ \left(f'_{\min(i+j,m)} \circ f_i\right) \\
&= \left(f''_{\min(i+j'+j'',m)} \circ f'_{\min(i+j,m)}\right) \circ f_i \\
&= \left(f''_{\min(\min(i+j,m)+j',m)} \circ f'_{\min(i+j,m)}\right) \circ f_i \\
&= \overline{f}_{\min(i+j,m)} \circ f_i
\end{aligned}$$

the latter of which is the i^{th} component of the second component of $m \odot_{\mathcal{M}} \overline{m}$, as required. For each $i \in [0, m]$ the i^{th} component of the third component of $m \odot_{\mathcal{M}} \overline{m}$ is

$$\begin{aligned}
g_{\min(i+j,m)} \circ \overline{g}_{\min(i+j+\overline{j},m))} &= g_{\min(i+j,m)} \circ \overline{g}_{\min(i+j+j'+j'',m)} \\
&= g_{\min(i+j,m)} \circ \overline{g}_{\min(\min(i+j,m)+j'+j'',m)} \\
&= g_{\min(i+j,m)} \circ \left(g'_{\min(\min(i+j,m)+j')} \circ g''_{\min(\min(i+j,m),j'+j'',m)}\right) \\
&= \left(g_{\min(i+j,m)} \circ g'_{\min(\min(i+j,m)+j')}\right) \circ g''_{\min(\min(i+j,m),j'+j'',m)} \\
&= \left(g_{\min(i+j,m)} \circ g'_{\min(i+j+j',m)}\right) \circ g''_{\min(i+\min(j+j',m)+j'',m)} \\
&= \widetilde{g}_i \circ g''_{\min(i+\widetilde{j}+j'',m)}
\end{aligned}$$

the latter of which is the i^{th} component of the third component of $\widetilde{m} \odot_{\mathcal{M}} m''$, as required. Clearly $(0, (\text{id}_Q)^{[0,m]}, (\text{id}_Q)^{[0,m]})$ is the identity of $\mathcal{M}$ with respect to $\odot_{\mathcal{M}}$, hence $(\mathcal{M}, \odot_{\mathcal{M}})$ is a monoid. The following points can easily be verified.

1. The function $\zeta : \mathcal{O}(\Sigma^{\triangle}) \to \mathcal{M}$ is a monoid morphism.
2. For all $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^{\triangle})$ with $\Delta(u) = \Delta(u')$ we have
$$\zeta(\text{ext}_{u,v}) = \zeta(\text{ext}_{u',v'}) \implies \zeta(\text{ext}_{u,v}) = \zeta(\text{ext}_{u',v}) = \zeta(\text{ext}_{u,v'}) = \zeta(\text{ext}_{u',v'}).$$
3. For all $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^{\triangle})$ we have
$$\zeta(\text{ext}_{u,v}) = \zeta(\text{ext}_{u',v'}) \implies \psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'}).$$

Now assume that L is not length-synchronous. We will prove that L is not length-synchronous. By assumption there exist a $\varphi_L(L)$-reachable idempotent $e \in O_L$ and $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^{\triangle})$ such that $\Delta(u), \Delta(u') > 0$, $\frac{|u|}{|v|} \neq \frac{|u'|}{|v'|}$, and $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'}) = e$. Without loss of generality we may assume $\Delta(u) = \Delta(u')$. Let ω denote the idempotent power of $\mathcal{M}$. Consider the elements

$$\text{ext}_{x,y} = (\text{ext}^{2\omega}_{u,v} \circ \text{ext}^{\omega}_{u',v'})^{\omega} \text{ and } \text{ext}_{x',y'} = (\text{ext}^{\omega}_{u,v} \circ \text{ext}^{2\omega}_{u',v'})^{\omega}.$$

By definition we have $\zeta(\mathrm{ext}_{x,y}) = \zeta(\mathrm{ext}_{x',y'})$, and since $\Delta(x) = \Delta(x')$ we obtain $\zeta(\mathrm{ext}_{x,y}) = \zeta(\mathrm{ext}_{x',y}) = \zeta(\mathrm{ext}_{x,y'}) = \zeta(\mathrm{ext}_{x',y'})$ by Point 2. Hence,

$$\psi_L(\mathrm{ext}_{x,y}) = \psi_L(\mathrm{ext}_{x',y}) = \psi_L(\mathrm{ext}_{x,y'}) = \psi_L(\mathrm{ext}_{x',y'}) = e$$

by Point 3.

We finally make a case distinction on whether $|u| = |u'|$ or not.

First, assume $|u| \neq |u'|$. Then $|x| \neq |x'|$ by construction. Since $\psi_L(\mathrm{ext}_{x,y}) = \psi_L(\mathrm{ext}_{x',y}) = e$, which is an idempotent of O_L, we obtain that L is not weakly length-synchronous.

Now assume $|u| = |u'|$. Since $\frac{|u|}{|v|} \neq \frac{|u'|}{|v'|}$ by assumption, we conclude that $|v| \neq |v'|$. By construction, the latter implies $|y| \neq |y'|$. Since $\psi_L(\mathrm{ext}_{x,y}) = \psi_L(\mathrm{ext}_{x,y'}) = e$, again we obtain that L is not weakly length-synchronous.

6 Proof of Proposition 5.1

We will prove the different statements appearing in Proposition 5.1 in the following subsections.

Computability of the syntactic Ext**-algebra.** This paragraph will be devoted to proving Point 1 of Proposition 5.1, rephrased in the following proposition.

Proposition 6.1. *Given a DVPA A with $L = L(A)$, one can compute the syntactic* Ext*-algebra (R_L, O_L) of L, its syntactic morphism (φ_L, ψ_L) and $\varphi_L(L)$.*

We require a bit of notation. For each visibly pushdown alphabet Σ and each finite Ext-algebra (R, O) it follows from Proposition 3.7 that each morphism $(\varphi, \psi)\colon (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R, O)$ has a unique finite presentation: it is given by the tuples

$$(\varphi(c))_{c \in \Sigma_{\mathrm{int}}} \quad \text{and} \quad (\psi(\mathrm{ext}_{a,b}))_{(a,b) \in \Sigma_{\mathrm{call}} \times \Sigma_{\mathrm{ret}}} \quad .$$

The syntactic Ext-algebra (R_L, O_L) of a VPL L over a visibly pushdown alphabet Σ can be represented by any Ext-algebra (R, O) such that R has $[1, |R_L|]$ as base set and such that there exists a bijective morphism $(\alpha, \beta)\colon (R, O) \to (R_L, O_L)$. Indeed, in that case we have

1. $xy = z \Leftrightarrow \alpha(x)\alpha(y) = \alpha(z)$ for all $x, y, z \in R$;
2. $x'y' = z' \Leftrightarrow \alpha^{-1}(x')\alpha^{-1}(y') = \alpha^{-1}(z')$ for all $x', y', z' \in R_L$;
3. $f(x) = y \Leftrightarrow \beta(f)(\alpha(x)) = \alpha(y)$ for all $f \in O$ and all $x, y \in R$; and
4. $f'(x') = y' \Leftrightarrow \beta^{-1}(f')(\alpha^{-1}(x')) = \alpha^{-1}(y')$ for all $f' \in O_L$ and all $x', y' \in R_L$.

For the following claim we avoid the tedious standard algebraic constructions on Ext-algebras to show decidability of the equivalence problem, since the latter decidability has already been established in [2].

Claim 6.2. *There is an algorithm that decides, given two morphisms into finite* Ext*-algebras $(\varphi_1, \psi_1)\colon (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R_1, O_1)$ and $(\varphi_2, \psi_2)\colon (\Sigma^{\triangle}, \mathcal{O}(\Sigma^{\triangle})) \to (R_1, O_1)$ for Σ a visibly pushdown alphabet and subsets $F_1 \subseteq R_1$ and $F_2 \subseteq R_2$, whether $\varphi_1^{-1}(F_1) = \varphi_2^{-1}(F_2)$.*

Proof of the Claim. The proof of Theorem 3.18 shows that one can effectively compute DVPAs A_1 and A_2 such that $L(A_1) = \varphi_1^{-1}(F_1)$ and $L(A_2) = \varphi_2^{-1}(F_2)$. By [2] one can effectively decide if $L(A_1) = L(A_2)$ by deciding $L(A_1) \subseteq L(A_2)$ and $L(A_2) \subseteq L(A_1)$. □

Proof of Proposition 6.1. By Theorem 3.18 we first compute from our DVPA A on the visibly push-down alphabet Σ an Ext-algebra (R_A, O_A), a morphism $(\varphi_A, \psi_A)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R_A, O_A)$, and a subset $F_A \subseteq R_A$ such that $L(A) = \varphi_A^{-1}(F_A)$. For an Ext-algebra (R, O) define $\#(R, O) = (|R|, |O|)$. Let $\prec \subseteq (\mathbb{N} \times \mathbb{N})^2$ be the lexicographic order on $\mathbb{N} \times \mathbb{N}$, i.e. $(i, j) \prec (k, l)$ if, and only if either $i < k$, or $i = k$ and $j < l$.

Observe that since (R_A, O_A) recognizes L, we have that the syntactic Ext-algebra (R_L, O_L) of L divides (R_A, O_A) by Proposition 3.17, so that $\#(R_L, O_L) \leq \#(R_A, O_A)$. In fact, we have that any Ext-algebra (R, O) having $[1, i]$ for $i \in [1, |R_L|]$ as base set, satisfying $\#(R, O) \leq \#(R_L, O_L)$ and recognizing L via a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ is a presentation of (R_L, O_L) with (φ, ψ) and F presentations of, respectively, (φ_L, ψ_L) and $\varphi_L(L)$. Indeed, since such an Ext-algebra recognizes L, by Proposition 3.17 it is divided by (R_L, O_L): this implies that $\#(R_L, O_L) \leq \#(R, O)$, but as also $\#(R, O) \leq \#(R_L, O_L)$, we have $\#(R, O) = \#(R_L, O_L)$. The morphism (φ, ψ) must be surjective, otherwise, by Lemma 3.15, $\big(\varphi(\Sigma^\triangle), \psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}\big)$ would be a sub-Ext-algebra of (R, O) recognizing L such that $\#\big(\varphi(\Sigma^\triangle), \psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}\big) < \#(R, O) = \#(R_L, O_L)$ while (R_L, O_L) divides $\big(\varphi(\Sigma^\triangle), \psi(\mathcal{O}(\Sigma^\triangle))|_{\varphi(\Sigma^\triangle)}\big)$, which is contradictory. Therefore, by Lemma 3.16, there is a surjective morphism $(\alpha, \beta)\colon (R, O) \to (R_L, O_L)$, that must be bijective, such that $\varphi_L = \alpha \circ \varphi$, so that (R, O) is a presentation of (R_L, O_L) with (φ, ψ) and F presentations of, respectively, (φ_L, ψ_L) and $\varphi_L(L)$.

Under the assumption that such an Ext-algebra exists, we compute (R_L, O_L), (φ_L, ψ_L) and $\varphi_L(L)$ by enumerating all the finitely many triples made of a finite Ext-algebra (R, O), a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ and a subset $F \subseteq R$ such that R has $[1, i]$ for $i \in [1, |R_A|]$ as base set and $\#(R, O) \leq \#(R_A, O_A)$. For each of these we test whether $\varphi^{-1}(F) = \varphi_A^{-1}(F_A)$, which is possible by the above claim and take (R, O), (φ, ψ) and F from a triple validating this test with $\#(R, O)$ minimal with respect to $\prec$.

It remains to prove that an Ext-algebra (R, O) having $[1, i]$ for $i \in [1, |R_L|]$ as base set, satisfying $\#(R, O) \leq \#(R_L, O_L)$ and recognizing L exists. Take any bijection $\alpha\colon R_L \to [1, |R_L|]$. We define R to be the monoid with base set $[1, |R_L|]$ and operation defined by $x \cdot y = \alpha\big(\alpha^{-1}(x)\alpha^{-1}(y)\big)$ for all $x, y \in [1, |R_L|]$. This is a monoid because

- $x \cdot \alpha(1_R) = \alpha\big(\alpha^{-1}(x)\alpha^{-1}(\alpha(1_R))\big) = \alpha(\alpha^{-1}(x)) = \alpha\big(\alpha^{-1}(\alpha(1_R))\alpha^{-1}(x)\big) = \alpha(1_R) \cdot x$ for all $x \in R$; and
- for all $x, y, z \in R$, we have

$$\begin{aligned} x \cdot (y \cdot z) &= \alpha\Big(\alpha^{-1}(x)\alpha^{-1}\Big(\alpha\big(\alpha^{-1}(y)\alpha^{-1}(z)\big)\Big)\Big) \\ &= \alpha\big(\alpha^{-1}(x)\alpha^{-1}(y)\alpha^{-1}(z)\big) \\ &= \alpha\Big(\alpha^{-1}\Big(\alpha\big(\alpha^{-1}(x)\alpha^{-1}(y)\big)\Big)\alpha^{-1}(z)\Big) \\ &= (x \cdot y) \cdot z\ . \end{aligned}$$

Define the function $\beta\colon O_L \to R^R$ by $\beta(f')(x) = \alpha\big(f'(\alpha^{-1}(x))\big)$ for all $f' \in O_L$ and $x \in R$. Set O to be the monoid with base set $\beta(O_L)$ and with composition as operation. This is a monoid because

- $\beta(1_O)(x) = \alpha\big(1_O(\alpha^{-1}(x))\big) = x = \mathrm{id}_R(x)$ for all $x \in R$; and

- for all $f', g' \in O_L$,

$$\beta(f') \circ \beta(g')(x) = \alpha\Big(f'\Big(\alpha^{-1}\Big(\alpha\big(g'(\alpha^{-1}(x))\big)\Big)\Big)\Big) = \alpha\big(f' \circ g'(\alpha^{-1}(x))\big) = \beta(f' \circ g')(x)$$

  for all $x \in R$, so that $\beta(f') \circ \beta(g') \in O$.

Then (R, O) is an Ext-algebra, because for all $r' \in R_L$, we have

$$\beta(\mathrm{left}_{r'})(x) = \alpha\big(\mathrm{left}_{r'}(\alpha^{-1}(x))\big) = \alpha\big(\alpha^{-1}(\alpha(r'))\alpha^{-1}(x)\big) = \alpha(r') \cdot x = \mathrm{left}_{\alpha(r')}(x)$$

for all $x \in R$ and $\beta(\mathrm{right}_{r'})(x) = \mathrm{right}_{\alpha(r')}(x)$ for all $x \in R$, so that $\mathrm{left}_r, \mathrm{right}_r \in O$ for all $r \in R$ by surjectivity of α.

We now define $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ as the unique morphism satisfying $\varphi(c) = \alpha(\varphi_L(c))$ for all $c \in \Sigma_{\mathrm{int}}$ and $\psi(\mathrm{ext}_{a,b}) = \beta(\psi_L(\mathrm{ext}_{a,b}))$ for all $a \in \Sigma_{\mathrm{call}}$, $b \in \Sigma_{\mathrm{ret}}$ given by Proposition 3.7. It is easy to show that then, $\varphi(w) = \alpha(\varphi_L(w))$ for all $w \in \Sigma^\triangle$ by structural induction on w. Hence, by injectivity of α, we have

$$\begin{aligned}\varphi^{-1}\big(\alpha(\varphi_L(L))\big) &= \{w \in \Sigma^\triangle \mid \alpha(\varphi_L(w)) \in \alpha(\varphi_L(L))\} \\ &= \{w \in \Sigma^\triangle \mid \varphi_L(w) \in \varphi_L(L)\} = \varphi_L^{-1}(\varphi_L(L)) = L\ ,\end{aligned}$$

thus (R, O) recognizes L. □

Decidability of quasi-aperiodicity. This paragraph is devoted to proving Point 2 (a) of Proposition 5.1, rephrased in the following proposition.

Proposition 6.3. *Given a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$ for Σ a visibly pushdown alphabet and (R, O) a finite* Ext*-algebra, it is decidable if (φ, ψ) is quasi-aperiodic. If (φ, ψ) is not quasi-aperiodic, one can effectively compute $k, l \in \mathbb{N}$ such that $\psi(\mathcal{O}(\Sigma^\triangle)^{k,l})$ is not aperiodic.*

For the rest of this paragraph, let us fix a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$, where Σ is a visibly pushdown alphabet and (R, O) is some finite Ext-algebra that is the input to our problem. We first have the following lemma.

Lemma 6.4. *For all $e \in O$ one can effectively compute a finite* Ext*-algebra recognizing $L_e = \{u\#v \mid uv \in \Sigma^\triangle : \psi(\mathrm{ext}_{u,v}) = e\}$, where $\#$ is a fresh internal letter that does not appear in Σ, along with an associated morphism and subset.*

Proof. Let Σ' be the alphabet that emerges from Σ by additionally declaring $\#$ as an internal letter. We will construct an Ext-algebra (R', O') and a morphism $(\varphi', \psi')\colon (\Sigma'^\triangle, \mathcal{O}(\Sigma')^\triangle) \to (R', O')$ such that for some element $r' \in R'$ we have $L_e = \varphi'^{-1}(r')$.

We define $R' = R \cup O \cup \{\bot\}$, for some fresh zero $\bot$, where multiplication between two elements in R' is defined as follows:

- multiplication between two elements in R is inherited from the monoid R;
- $r \cdot f = \mathrm{left}_r \circ f$ and $f \cdot r = \mathrm{right}_r \circ f$ for all $r \in R$ and all $f \in O$;
- $\bot$ acts as a zero, i.e. $\bot \cdot r' = r' \cdot \bot = \bot$ for all $r' \in R'$;
- $f \cdot g = \bot$ for all $f, g \in O$.

Clearly the identity of R is the identity of R'. Associativity is immediate except for products of the form $r_1 \cdot f \cdot r_2, r_1 \cdot r_2 \cdot f$, and $f \cdot r_1 \cdot r_2$, where $f \in O$ and $r_1, r_2 \in R$. In the first case we have

$$\begin{aligned}(r_1 \cdot f) \cdot r_2 = (\mathrm{left}_{r_1} \circ f) \cdot r_2 = \mathrm{right}_{r_2} \circ (\mathrm{left}_{r_1} \circ f) = (\mathrm{right}_{r_2} \circ \mathrm{left}_{r_1}) \circ f \\ = (\mathrm{left}_{r_1} \circ \mathrm{right}_{r_2}) \circ f = r_1 \cdot (\mathrm{right}_{r_2} \circ f) = r_1 \cdot (f \cdot r_2)\ .\end{aligned}$$

In the second case we have

$$(r_1 \cdot r_2) \cdot f = \mathrm{left}_{r_1 r_2} \circ f = (\mathrm{left}_{r_1} \circ \mathrm{left}_{r_2}) \circ f = \mathrm{left}_{r_1} \circ (\mathrm{left}_{r_2} \circ f) = r_1 \cdot (r_2 \cdot f)$$

and in the third case we have

$$f \cdot (r_1 \cdot r_2) = \mathrm{right}_{r_1 r_2} \circ f = (\mathrm{right}_{r_2} \circ \mathrm{right}_{r_1}) \circ f = \mathrm{right}_{r_2} \circ (\mathrm{right}_{r_1} \circ f) = (f \cdot r_1) \cdot r_2\ .$$

We define $O' = (R')^{R'}$ which is clearly a monoid for composition and thus directly get that (R', O') is an Ext-algebra. Applying Proposition 3.7, we define the morphism $(\varphi', \psi')\colon (\Sigma'^{\triangle}, \mathcal{O}(\Sigma')^{\triangle}) \to (R', O')$ as the unique one satisfying $\varphi'(c) = \varphi(c)$ for all $c \in \Sigma_{\mathrm{int}}$, $\varphi'(\#) = \mathrm{id}_O$ and where for all $a \in \Sigma_{\mathrm{call}}$, $b \in \Sigma_{\mathrm{ret}}$, we have

$$\psi'(\mathrm{ext}_{a,b})(x) = \begin{cases} \psi(\mathrm{ext}_{a,b})(x) & \text{if } x \in R \\ \psi(\mathrm{ext}_{a,b}) \circ x & \text{if } x \in O \\ \bot & \text{otherwise (i.e. if } x = \bot) \end{cases}$$

for all $x \in R'$. It suffices to prove the following claim, which directly implies the desired equality $\varphi'^{-1}(e) = \{u\#v \mid uv \in \Sigma^{\triangle} \text{ s.t. } \psi(\mathrm{ext}_{u,v}) = e\}$. For all $w \in \Sigma'^{\triangle}$ we have

$$\varphi'(w) = \begin{cases} \varphi(w) & \text{if } w \in \Sigma^{\triangle} \\ \psi(\mathrm{ext}_{u,v}) & \text{if } w = u\#v \text{ for some } uv \in \Sigma^{\triangle} \\ \bot & \text{otherwise.} \end{cases}$$

We prove it by structural induction on w. The cases when $w = \varepsilon$ or $w = c \in \Sigma_{\mathrm{int}}$ follow immediately from the definition of φ'. In case $w = \# = \varepsilon\#\varepsilon$, we have $\varphi'(w) = \mathrm{id}_O = \psi(\mathrm{ext}_{\varepsilon,\varepsilon})$.

For the inductive step first assume $w = aw'b$ for some $w' \in \Sigma'^{\triangle}$. If w' is neither in $\Sigma^{\triangle}$ nor of the form $u\#v$ with $uv \in \Sigma^{\triangle}$, then $\varphi'(w') = \bot$ by induction hypothesis and thus $\varphi'(w) = \psi'(\mathrm{ext}_{a,b})(\varphi'(w')) = \psi'(\mathrm{ext}_{a,b})(\bot) = \bot$ as required. If $w' \in \Sigma^{\triangle}$, then $\varphi'(w') = \varphi(w') \in R$ by induction hypothesis, and hence $\varphi'(w) = \psi'(\mathrm{ext}_{a,b})(\varphi'(w')) = \psi'(\mathrm{ext}_{a,b})(\varphi(w')) = \psi(\mathrm{ext}_{a,b})(\varphi(w')) = \varphi(w)$ as required. If $w' = u\#v$ with $uv \in \Sigma^{\triangle}$, i.e. $w = au\#vb$, then $\varphi'(w') = \psi(\mathrm{ext}_{u,v}) \in O$ by induction hypothesis. Hence, we have $\varphi'(w) = \psi'(\mathrm{ext}_{a,b})(\varphi'(w')) = \psi'(\mathrm{ext}_{a,b})(\psi(\mathrm{ext}_{u,v})) = \psi(\mathrm{ext}_{a,b}) \circ \psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{au,vb})$.

Finally assume $w = xy$ for some $x, y \in \Sigma'^{\triangle} \setminus \{\varepsilon\}$. The case when x or y is neither in $\Sigma^{\triangle}$ nor of the form $u\#v$ with $uv \in \Sigma^{\triangle}$ is easily handled by applying the induction hypothesis and observing that $\bot$ is a zero in R'. Two other immediate cases are when both x and y are in $\Sigma^{\triangle}$ and when both x and y are of the form $u\#v$ with $uv \in \Sigma^{\triangle}$. Consider the case when $x \in \Sigma^{\triangle} \setminus \{\varepsilon\}$ and $y = u\#v$ with $uv \in \Sigma^{\triangle}$, hence $w = xu\#v$. The induction hypothesis yields $\varphi'(x) = \varphi(x) \in R$ and $\varphi'(y) = \psi(\mathrm{ext}_{u,v}) \in O$. We obtain

$$\varphi'(xy) = \varphi'(x) \cdot \varphi'(y) = \varphi(x) \cdot \psi(\mathrm{ext}_{u,v}) = \mathrm{left}_{\varphi(x)} \circ \psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{xu,v})$$

as required. Finally, let us treat the case when $x = u\#v$ with $uv \in \Sigma^{\triangle}$ and $y \in \Sigma^{\triangle} \setminus \{\varepsilon\}$, i.e. $w = u\#vy$. The induction hypothesis yields $\varphi'(x) = \psi(\mathrm{ext}_{u,v}) \in O$ and $\varphi'(y) = \varphi(y) \in R$. We obtain

$$\varphi'(xy) = \varphi'(x) \cdot \varphi'(y) = \psi(\mathrm{ext}_{u,v}) \cdot \varphi(y) = \mathrm{right}_{\varphi(y)} \circ \psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u,vy})$$

as required. □

The next goal will be to prove that the set of pairs of word lengths $(|u|, |v|)$ of words $u\#v \in L_e$ is effectively semilinear for each $e \in O$.

A *(realtime) pushdown automaton (PDA for short)* is a tuple $A = (Q, \Sigma, \Gamma, \Omega, q_0, F, \bot)$, where Q is a finite set of *states*, Σ is a finite *input alphabet*, Γ is a finite *stack alphabet*, $q_0 \in Q$ is an *initial state*, $F \subseteq Q$ is the set of *final states*, $\bot \in \Gamma \setminus \Sigma$ is the *bottom-of-stack symbol*, and $\Omega \subseteq Q \times \Sigma \times \Gamma \times Q \times \Gamma^*$ is a finite *transition relation* such that for all $(p, a, X, q, \alpha) \in \Omega$ we have $\alpha \in \Gamma^*\bot$ if $X = \bot$ and $\alpha \in (\Gamma \setminus \{\bot\})^*$ otherwise. The relation Ω is naturally extended to the relation $\Omega^* \subseteq Q \times \Sigma^* \times \Gamma^*\bot \times Q \times \Gamma^*\bot$, namely as the smallest relation containing the set $\{(p, \varepsilon, \alpha, p, \alpha) \mid p \in Q, \alpha \in \Gamma^*\bot\}$ and such that moreover, if $(p, a, X, q, \alpha) \in \Omega$ and $(q, w, \alpha\beta, r, \gamma) \in \Omega^*$, then $(p, aw, X\beta, r, \gamma) \in \Omega^*$. The *language* of A is $L(A) = \{w \in \Sigma^* \mid \exists \alpha \in \Gamma^*\bot, \exists q \in F : (q_0, w, \bot, q, \alpha) \in \Omega^*\}$. Hence it is clear that one can compute a PDA A' such that $L(A') = L(A)$.

Lemma 6.5. *Let A be a DVPA that accepts a language over a visibly pushdown alphabet Σ' such that $L(A) \subseteq \{u\#v \mid uv \in \Sigma'^{\triangle} \setminus \Sigma'^*\#\Sigma'^*\}$ and $\# \in \Sigma'_{int}$. Then the set*

$$P(L(A)) = \{(k, l) \in \mathbb{N} \times \mathbb{N} \mid \exists u \in (\Sigma' \setminus \{\#\})^k, v \in (\Sigma' \setminus \{\#\})^l : u\#v \in L(A)\}$$

is effectively semilinear.

Proof. We first compute a PDA A' accepting the same language as A, i.e. $L(A') = L(A)$. Let us assume without loss of generality that $0, 1 \notin \Sigma'$. We claim that from $A' = (Q, \Sigma', \Gamma, \Omega, q_0, F, \bot)$ one can compute a PDA A'' such that

$$L(A'') = \{0^{|u|}\#1^{|v|} \mid u\#v \in L(A')\} \ .$$

Indeed, the PDA A'' can simply be computed as follows: we set

$$A'' = (Q \times \{0, 1\}, \{0, 1, \#\}, \Gamma, \Omega', \langle q_0, 0\rangle, F \times \{1\}, \bot) \ ,$$

where Ω' is the union of $\{(\langle p, i\rangle, i, X, \langle q, i\rangle, \alpha) \mid i \in \{0, 1\}, \exists c \in \Sigma' \setminus \{\#\} : (p, c, X, q, \alpha) \in \Omega\}$ and $\{(\langle p, 0\rangle, \#, X, \langle q, 1\rangle, \alpha) \mid (p, \#, X, q, \alpha) \in \Omega\}$. Finally, we apply Parikh's Theorem, cf. [14, Section 3], which implies that the set $\{(m, n) \in \mathbb{N} \times \mathbb{N} \mid 0^m\#1^n \in L(A'')\} = P(L(A))$ is effectively semilinear. □

We are now ready to prove Proposition 6.3.

Proof of Proposition 6.3. Let $e \in O$. By Lemma 6.4 we first compute a finite Ext-algebra recoginizing L_e, along with an associated morphism and subset. From the latter we can compute (by Theorem 3.18) a DVPA A_e accepting L_e. We then use Lemma 6.5 to conclude that the set

$$P(L_e) = \{(k, l) \in \mathbb{N} \times \mathbb{N} \mid \exists u \in \Sigma^k, v \in \Sigma^l : uv \in \Sigma^{\triangle}, \psi(\mathrm{ext}_{u,v}) = e\}$$

is effectively semilinear, and this holds for all $e \in O$.

We make use of the folklore fact that semilinear sets are effectively closed under Boolean operations, cf. [9] for a recent study. To decide whether (φ, ψ) is quasi-aperiodic, we go through all

possible subsets $U \subseteq O$: if it is a subsemigroup of O that is a non-trivial group, we compute the set $\bigcap_{e\in U} P(L_e)$ and reject if it is non-empty (which is easy to check given a semilinear presentation of the set), otherwise we continue. If we were able to go through all those subsets without rejecting, we accept.

Thus, if (φ,ψ) is not quasi-aperiodic we can find a subset $U \subseteq O$ that contains a non-trivial group and output a pair $(k,l) \in \bigcap_{e\in U} P(L_e)$; it witnesses that $\psi(\mathcal{O}(\Sigma^{\triangle})^{k,l})$ is not aperiodic. □

Decidability of length-synchronicity. This paragraph is devoted to proving Point 2 (b) of Proposition 5.1, rephrased in the following proposition.

Proposition 6.6. *Given a morphism $(\varphi,\psi)\colon (\Sigma^{\triangle},\mathcal{O}(\Sigma^{\triangle})) \to (R,O)$, for Σ a visibly pushdown alphabet and (R,O) a finite* Ext*-algebra, and some $F \subseteq R$, it is decidable if (φ,ψ) is F-length-synchronous. If (φ,ψ) is not length-synchronous, one can effectively compute a tuple $(k,l,k',l') \in \mathbb{N}^4_{>0}$ such that that there exist $uv, u'v' \in \Sigma^{\triangle}$ and some F-reachable idempotent $e \in O$ such that $\psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u',v'}) = e$, $k = |u|$, $l = |v|$, $k' = |u'|$, $l' = |v'|$ and $\frac{k}{l} \neq \frac{k'}{l'}$.*

Before proving the proposition we need a technical lemma characterizing when a two-dimensional semilinear set contains only vectors with the same slope. We say two vectors $\vec{x},\vec{y} \in \mathbb{N}^2$ are *collinear* if $\vec{y} = \alpha\cdot\vec{x}$ for some $\alpha \in \mathbb{Q}_{>0}$

Lemma 6.7. *Let $S = \bigcup_{i\in I}\left(\vec{x}_{i,0} + \sum_{j=1}^{t_i} \mathbb{N}\vec{x}_{i,j}\right) \subseteq \mathbb{N}^2_{>0}$ be a non-empty semilinear set, where $\vec{x}_{i,j} \neq (0,0)$ for all $i \in I$ and all $j \in [0,t_i]$. Then,*

$$\left|\left\{\frac{k}{l}\middle|(k,l)\in S\right\}\right| = 1 \quad\iff\quad \forall i,i' \in I\ \forall j \in [0,t_i]\ \forall j' \in [0,t_{i'}] : \vec{x}_{i,j} \text{ and } \vec{x}_{i',j'} \text{ are collinear.}$$

Proof. First assume that $\vec{x}_{i,j}$ and $\vec{x}_{i',j'}$ are collinear for all $i,i' \in I$, $j \in [0,t_i]$, and $j' \in [0,t_{i'}]$. Let $(k,l),(k',l') \in S$. That is, $(k,l) = \vec{x}_{i,0}+n_1\vec{x}_{i,1}+\cdots+n_{t_i}\vec{x}_{i,t_i}$ and $(k',l') = \vec{x}_{i',0}+n'_1\vec{x}_{i',1}+\cdots+n'_{t_{i'}}\vec{x}_{i',t_{i'}}$ for some $i,i' \in I$ and some $n_1,\ldots,n_{t_i},n'_1,\ldots,n'_{t_{i'}} \in \mathbb{N}$. But due to pairwise collinearity there exist $\alpha,\alpha' \in \mathbb{Q}_{>0}$ such that $(k,l) = \alpha\vec{x}_{i,0}$ and $(k',l') = \alpha'\vec{x}_{i,0}$, thus implying $\frac{k}{l} = \frac{k'}{l'}$.

Conversely assume that there exist two vectors $(k,l) = \vec{x}_{i,j}$ and $(k',l') = \vec{x}_{i',j'}$ that are not collinear. In case this is possible when $i \neq i'$ and $j = j' = 0$ we are done, since then $(k,l),(k',l') \in S$ and thus $\frac{k}{l} \neq \frac{k'}{l'}$. Otherwise $\vec{x}_{i,0}$ and $\vec{x}_{i',0}$ are collinear for all $i,i' \in I$, so there must exist $i \in I$ and $j \in [0,t_i]$ such that $\vec{x}_{i,0}$ and $\vec{x}_{i,j}$ are not collinear. Then $\vec{x}_{i,0}$ and $\vec{x}_{i,0}+\vec{x}_{i,j}$ are in S but also not collinear: indeed, if $\alpha\vec{x}_{i,0} = \vec{x}_{i,0}+\vec{x}_{i,j}$ for some $\alpha \in \mathbb{Q}_{>0}$, then $\vec{x}_{i,j} = (\alpha-1)\vec{x}_{i,0}$ with $\alpha - 1 > 0$ due to $\vec{x}_{i,0},\vec{x}_{i,j} \in \mathbb{N}^2 \setminus \{(0,0)\}$, a contradiction. Hence there exist $(k,l),(k',l') \in S$ that are not collinear, and therefore $\frac{k}{l} \neq \frac{k'}{l'}$. □

Proof of Proposition 6.6. Let us fix the Ext-algebra morphism $(\varphi,\psi)\colon (\Sigma^{\triangle},\mathcal{O}(\Sigma^{\triangle})) \to (R,O)$, where (R,O) is a finite Ext-algebra and where $F \subseteq R$.

Recall that over the alphabet Σ', obtained from Σ by declaring a fresh letter # as internal, the language

$$L_{e\uparrow} = \{u\#v \mid uv \in \Sigma^{\triangle} : \Delta(u) > 0, \psi(\mathrm{ext}_{u,v}) = e\} = L_e \cap \{u\#v \mid uv \in \Sigma^{\triangle} : \Delta(u) > 0\}$$

is given for all $e \in O$. The language $\{u\#v \mid uv \in \Sigma^{\triangle} : \Delta(u) > 0\}$ is a clearly a VPL. Thus, for all $e \in O$, we have that the set

$$P(L_{e\uparrow}) = \left\{(k,l) \in \mathbb{N}\times\mathbb{N} \mid \exists u \in \Sigma^k, v \in \Sigma^l : uv \in \Sigma^{\triangle}, \Delta(u) > 0, \psi(\mathrm{ext}_{u,v}) = e\right\}$$

is effectively semilinear: indeed, given $e \in O$, using Lemma 6.4 and Theorem 3.18, we can as in the proof of Point 2 (a) of Proposition 5.1 compute a DVPA A_e accepting L_e; we then compute a DVPA $A_{e\uparrow}$ accepting $L_{e\uparrow} = L(A_e) \cap L(A)$ by using the effective construction given in [2] and finally use Lemma 6.5 to conclude.

Observe that (φ, ψ) is F-length-synchronous if, and only if, for each F-reachable idempotent $e \in O$ for which $P(L_{e\uparrow})$ is non-empty we have $|\{\frac{k}{l} \mid (k,l) \in P(L_{e\uparrow})\}| = 1$. The latter condition is easily seen to be decidable by the characterization provided in Lemma 6.7. Hence, for deciding if (φ, ψ) is length-synchronous our algorithm verifies if for all F-reachable $e \in O$ for which $P(L_{e\uparrow})$ is non-empty we have $|\{\frac{k}{l} \mid (k,l) \in P(L_{e\uparrow})\}| = 1$. On the other hand, if this verification fails, i.e. in case (φ, ψ) is not F-length-synchronous, our algorithm outputs, again using the characterization of Lemma 6.7, a quadruple $(k, l, k', l') \in \mathbb{N}^4_{>0}$ such that for some F-reachable idempotent $e \in O$ we have $(k,l), (k',l') \in P(L_{e\uparrow})$ and $\frac{k}{l} \neq \frac{k'}{l'}$. □

Decidability of weak length-synchronicity. This paragraph is devoted to proving Point 2 (c) of Proposition 5.1, rephrased in the following proposition.

Proposition 6.8. *Given a morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$, for Σ a visibly pushdown alphabet and (R, O) a finite* Ext*-algebra, and some $F \subseteq R$, it is decidable if (φ, ψ) is F-weakly-length-synchronous.*

Let us fix the morphism $(\varphi, \psi)\colon (\Sigma^\triangle, \mathcal{O}(\Sigma^\triangle)) \to (R, O)$, for Σ a visibly pushdown alphabet and (R, O) a finite Ext-algebra, and some $F \subseteq R$.

Define the new visibly pushdown alphabet $\overline{\Sigma}$ by $\overline{\Sigma}_{\text{call}} = \{\overline{b} \mid b \in \Sigma_{\text{ret}}\}$, $\overline{\Sigma}_{\text{int}} = \{\overline{c} \mid c \in \Sigma_{\text{int}}\}$ and $\overline{\Sigma}_{\text{ret}} = \{\overline{a} \mid a \in \Sigma_{\text{call}}\}$. For all $w \in \Sigma^*$, we define

$$\overline{w} = \begin{cases} \varepsilon & \text{if } w = \varepsilon \\ \overline{w_n} \cdots \overline{w_1} & \text{if } w = w_1 \cdots w_n \text{ for } n \in \mathbb{N}_{>0} \text{ and } w_1, \ldots w_n \in \Sigma\ . \end{cases}$$

We have the following lemma, that we prove later on.

Lemma 6.9. *For all $e \in O$ one can effectively compute a finite* Ext*-algebra recognizing the language of well-matched words $K_e = \{u\#\overline{u'} \mid u, u' \in \Sigma^*, \exists v \in \Sigma^* : uv \in \Sigma^\triangle, u'v \in \Sigma^\triangle, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e\}$, where $\#$ is a fresh internal letter that does not appear in $\Sigma \cup \overline{\Sigma}$, along with an associated morphism and subset.*

Over the alphabet Σ' obtained from $\Sigma \cup \overline{\Sigma}$ by declaring the fresh letter $\#$ as internal, we define

$$\begin{aligned} K_{e\uparrow} &= \left\{ u\#\overline{u'} \;\middle|\; \begin{array}{l} u, u' \in \Sigma^*, \exists v \in \Sigma^* : \\ uv \in \Sigma^\triangle, u'v \in \Sigma^\triangle, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e \end{array} \right\} \\ &= K_e \cap \{u\#u' \mid uu' \in (\Sigma \cup \overline{\Sigma})^\triangle : \Delta(u) > 0\} \end{aligned}$$

for all $e \in O$. As in the proof of Point (2) of the second statement of Proposition 5.1, we can prove that the language $\{u\#u' \mid uu' \in (\Sigma \cup \overline{\Sigma})^\triangle : \Delta(u) > 0\}$ is a VPL and thus conclude that for all $e \in O$, the set

$$P(K_{e\uparrow}) = \left\{ (k,l) \in \mathbb{N} \times \mathbb{N} \;\middle|\; \begin{array}{l} \exists u \in \Sigma^k, u' \in \Sigma^l, v \in \Sigma^* : \\ uv \in \Sigma^\triangle, u'v \in \Sigma^\triangle, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e \end{array} \right\}$$

is effectively semilinear.

It is clear that (φ, ψ) is F-weakly-length-synchronous if and only if for each idempotent $e \in O$ that is F-reachable, there does not exist any $(x_1, x_2) \in P(K_{e\uparrow})$ such that $x_1 \neq x_2$. Therefore, to decide whether (φ, ψ) is F-weakly-length-synchronous, we go through all $e \in O$: if e is an idempotent that is F-reachable, we compute the set $P(K_{e\uparrow})$ and reject if it contains a vector (x_1, x_2) such that $x_1 \neq x_2$ (which is easy to check given a semilinear presentation of the set), otherwise we continue. Finally, if we were able to go through all those elements without rejecting, we accept.

Proof of Lemma 6.9. Let Σ' be the alphabet that emerges from $\Sigma \cup \overline{\Sigma}$ by additionally declaring $\#$ as an internal letter. We will construct an Ext-algebra (R', O') and a morphism (φ', ψ') from $(\Sigma'^{\triangle}, \mathcal{O}(\Sigma')^{\triangle})$ to (R', O') such that for some subset $F \subseteq R'$ we have $K_e = \varphi'^{-1}(F)$.

Let $\overline{R} = \{\overline{r} \mid r \in R\}$. We define $R' = R \cup \overline{R} \cup \mathfrak{P}(O^2) \setminus \emptyset \cup \{\bot, 1\}$, for some fresh zero $\bot$ and identity 1, where multiplication between two elements in R' is defined as follows:

- for all $r_1, r_2 \in R$,

$$\begin{aligned} r_1 \cdot r_2 &= r_1 r_2 & \qquad \overline{r_1} \cdot r_2 &= \bot \\ \overline{r_1} \cdot \overline{r_2} &= \overline{r_2 r_1} & \qquad r_1 \cdot \overline{r_2} &= \bot \ ; \end{aligned}$$

- for all $r \in R$ and $E \in \mathfrak{P}(O^2) \setminus \emptyset$,

$$\begin{aligned} r \cdot E &= \{(\mathrm{left}_r \circ e_1, e_2) \mid (e_1, e_2) \in E\} & \qquad E \cdot r &= \bot \\ E \cdot \overline{r} &= \{(e_1, \mathrm{left}_r \circ e_2) \mid (e_1, e_2) \in E\} & \qquad \overline{r} \cdot E &= \bot \ ; \end{aligned}$$

- for all $E_1, E_2 \in \mathfrak{P}(O^2) \setminus \emptyset$, we have $E_1 \cdot E_2 = \bot$;
- $\bot$ acts as a zero, i.e. $\bot \cdot r' = r' \cdot \bot = \bot$ for all $r' \in R'$;
- 1 acts as an identity, i.e. $1 \cdot r' = r' \cdot 1 = r'$ for all $r' \in R'$.

Associativity is immediate except for products of the form $\overline{r_1} \cdot \overline{r_2} \cdot \overline{r_3}$, $r_1 \cdot E \cdot \overline{r_2}$, $r_1 \cdot r_2 \cdot E$ and $E \cdot \overline{r_1} \cdot \overline{r_2}$, where $E \in \mathfrak{P}(O^2) \setminus \emptyset$ and $r_1, r_2, r_3 \in R$. In the first case we have

$$(\overline{r_1} \cdot \overline{r_2}) \cdot \overline{r_3} = \overline{r_2 r_1} \cdot \overline{r_3} = \overline{r_3 r_2 r_1} = \overline{r_1} \cdot \overline{r_3 r_2} = \overline{r_1} \cdot (\overline{r_2} \cdot \overline{r_3}) \ .$$

In the second case we have

$$\begin{aligned} (r_1 \cdot E) \cdot \overline{r_2} &= \{(\mathrm{left}_{r_1} \circ e_1, e_2) \mid (e_1, e_2) \in E\} \cdot \overline{r_2} \\ &= \{(\mathrm{left}_{r_1} \circ e_1, \mathrm{left}_{r_2} \circ e_2) \mid (e_1, e_2) \in E\} \\ &= r_1 \cdot \{(e_1, \mathrm{left}_{r_2} \circ e_2) \mid (e_1, e_2) \in E\} = r_1 \cdot (E \cdot \overline{r_2}) \ . \end{aligned}$$

In the third case we have

$$\begin{aligned} (r_1 \cdot r_2) \cdot E &= \{(\mathrm{left}_{r_1 r_2} \circ e_1, e_2) \mid (e_1, e_2) \in E\} \\ &= \{(\mathrm{left}_{r_1} \circ \mathrm{left}_{r_2} \circ e_1, e_2) \mid (e_1, e_2) \in E\} \\ &= r_1 \cdot \{(\mathrm{left}_{r_2} \circ e_1, e_2) \mid (e_1, e_2) \in E\} = r_1 \cdot (r_2 \cdot E) \end{aligned}$$

and in the fourth case we have

$$\begin{aligned} E \cdot (\overline{r_1} \cdot \overline{r_2}) &= \{(e_1, \mathrm{left}_{r_2 r_1} \circ e_2) \mid (e_1, e_2) \in E\} \\ &= \{(e_1, \mathrm{left}_{r_2} \circ \mathrm{left}_{r_1} \circ e_2) \mid (e_1, e_2) \in E\} \\ &= \{(e_1, \mathrm{right}_{r_1} \circ e_2) \mid (e_1, e_2) \in E\} \cdot \overline{r_2} = (E \cdot \overline{r_1}) \cdot \overline{r_2} \ . \end{aligned}$$

We define $O' = (R')^{R'}$ which is clearly a monoid for composition and thus directly get that (R', O') is a finite Ext-algebra. Applying Proposition 3.7, we define the morphism $(\varphi', \psi') \colon (\Sigma'^{\triangle}, \mathcal{O}(\Sigma')^{\triangle}) \to (R', O')$ as the unique one satisfying $\varphi'(c) = \varphi(c)$ and $\varphi'(\overline{c}) = \overline{\varphi(c)}$ for all $c \in \Sigma_{\mathrm{int}}$, $\varphi'(\#) = \{(\psi(\mathrm{ext}_{\varepsilon,v}), \psi(\mathrm{ext}_{\varepsilon,v})) \mid v \in \Sigma^{\triangle}\}$ and where for all $a, a' \in \Sigma_{\mathrm{call}}$, $b, b' \in \Sigma_{\mathrm{ret}}$, we have

$$\psi'(\mathrm{ext}_{a,b})(x) = \begin{cases} \psi(\mathrm{ext}_{a,b})(1_R) & \text{if } x = 1 \\ \psi(\mathrm{ext}_{a,b})(x) & \text{if } x \in R \\ \bot & \text{otherwise} \end{cases}$$

$$\psi'(\mathrm{ext}_{\overline{b'},\overline{a'}})(x) = \begin{cases} \overline{\psi(\mathrm{ext}_{a',b'})(1_R)} & \text{if } x = 1 \\ \overline{\psi(\mathrm{ext}_{a',b'})(x')} & \text{if } x = \overline{x'} \text{ for } x' \in R \\ \bot & \text{otherwise} \end{cases}$$

$$\psi'(\mathrm{ext}_{a,\overline{a'}})(x) = \begin{cases} \bigcup\limits_{\substack{b \in \Sigma_{\mathrm{ret}} \\ z \in \Sigma^{\triangle} \\ (e_1,e_2) \in x}} \{(\psi(\mathrm{ext}_{a,bz}) \circ e_1, \psi(\mathrm{ext}_{a',bz}) \circ e_2)\} & \text{if } x \in \mathfrak{P}(O^2) \setminus \emptyset \\ \bot & \text{otherwise} \end{cases}$$

$$\psi'(\mathrm{ext}_{\overline{b'},b})(x) = \bot$$

for all $x \in R'$. Note that (φ', ψ') is computable because

$$\begin{aligned} \{(\psi(\mathrm{ext}_{\varepsilon,v}), \psi(\mathrm{ext}_{\varepsilon,v})) \mid v \in \Sigma^{\triangle}\} &= \{(\mathrm{right}_{\varphi(v)}, \mathrm{right}_{\varphi(v)}) \mid v \in \Sigma^{\triangle}\} \\ &= \{(\mathrm{right}_r, \mathrm{right}_r) \mid r \in R\} \end{aligned}$$

and

$$\begin{aligned} &\bigcup_{\substack{b \in \Sigma_{\mathrm{ret}} \\ z \in \Sigma^{\triangle} \\ (e_1,e_2) \in x}} \{(\psi(\mathrm{ext}_{a,bz}) \circ e_1, \psi(\mathrm{ext}_{a',bz}) \circ e_2)\} \\ = &\bigcup_{\substack{b \in \Sigma_{\mathrm{ret}} \\ z \in \Sigma^{\triangle} \\ (e_1,e_2) \in x}} \{(\mathrm{right}_{\varphi(z)} \circ \psi(\mathrm{ext}_{a,b}) \circ e_1, \mathrm{right}_{\varphi(z)} \circ \psi(\mathrm{ext}_{a',b}) \circ e_2)\} \\ = &\bigcup_{\substack{b \in \Sigma_{\mathrm{ret}} \\ r \in R \\ (e_1,e_2) \in x}} \{(\mathrm{right}_r \circ \psi(\mathrm{ext}_{a,b}) \circ e_1, \mathrm{right}_r \circ \psi(\mathrm{ext}_{a',b}) \circ e_2)\} \end{aligned}$$

for all $x \in \mathfrak{P}(O^2) \setminus \emptyset$ and $a, a' \in \Sigma_{\mathrm{call}}$.

Now define the set of pairs $P = \{(u, u') \in \Sigma^* \times \Sigma^* \mid \exists v \in \Sigma^* : uv \in \Sigma^{\triangle}, u'v \in \Sigma^{\triangle}\}$; it is not difficult to check that for all $w \in \Sigma'^*$, $w \in \Sigma'^{\triangle} \cap \Sigma^* \# \overline{\Sigma}^*$ if and only if $w = u \# \overline{u'}$ for $(u, u') \in P$. It

suffices to prove the following claim, which directly implies the desired equality

$$\begin{aligned}
&\varphi'^{-1}(\{E \in \mathfrak{P}(O^2) \setminus \emptyset \mid (e,e) \in E\}) \\
=&\{u\#\overline{u'} \mid (u,u') \in P, \exists v \in \Sigma^* : uv \in \Sigma^\triangle, u'v \in \Sigma^\triangle, \psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u',v}) = e\} \\
=&\{u\#\overline{u'} \mid u,u' \in \Sigma^*, \exists v \in \Sigma^* : uv \in \Sigma^\triangle, u'v \in \Sigma^\triangle, \psi(\mathrm{ext}_{u,v}) = \psi(\mathrm{ext}_{u',v}) = e\}\ .
\end{aligned}$$

For all $w \in \Sigma'^\triangle$ we have

$$\varphi'(w) = \begin{cases}
1 & \text{if } w = \varepsilon \\
\varphi(w) & \text{if } w \in \Sigma^\triangle \setminus \{\varepsilon\} \\
\overline{\varphi(w')} & \text{if } w = \overline{w'} \text{ for } w' \in \Sigma^\triangle \setminus \{\varepsilon\} \\
\{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} & \text{if } w = u\#\overline{u'} \text{ for } (u,u') \in P \\
\bot & \text{otherwise.}
\end{cases}$$

We prove it by structural induction on w. The cases when $w = \varepsilon$ or $w = c \in \Sigma_{\mathrm{int}}$ or $w = \overline{c} \in \overline{\Sigma}_{\mathrm{int}}$ follow immediately from the definition of φ'. In case $w = \# = \varepsilon\#\varepsilon$, we have

$$\varphi'(w) = \{(\psi(\mathrm{ext}_{\varepsilon,v}), \psi(\mathrm{ext}_{\varepsilon,v})) \mid v \in \Sigma^\triangle\} = \{(\psi(\mathrm{ext}_{\varepsilon,v}), \psi(\mathrm{ext}_{\varepsilon,v})) \mid v \in \Sigma^*, \varepsilon v \in \Sigma^\triangle\}$$

as required.

For the inductive step first assume $w = \alpha w' \beta$ for $w' \in \Sigma'^\triangle$, $\alpha \in \Sigma_{\mathrm{call}} \cup \overline{\Sigma}_{\mathrm{call}}$ and $\beta \in \Sigma_{\mathrm{ret}} \cup \overline{\Sigma}_{\mathrm{ret}}$. If w' is neither in $\Sigma^\triangle \cup \overline{\Sigma}^\triangle$ nor of the form $u\#\overline{u'}$ with $(u,u') \in P$, then $\varphi'(w') = \bot$ by induction hypothesis and thus $\varphi'(w) = \psi'(\mathrm{ext}_{\alpha,\beta})(\varphi'(w')) = \psi'(\mathrm{ext}_{\alpha,\beta})(\bot) = \bot$ as required, since w is also neither in $\Sigma^\triangle \cup \overline{\Sigma}^\triangle$ nor of the form $u\#\overline{u'}$ with $(u,u') \in P$. If $w' = \varepsilon$, then $\varphi'(w') = 1$ and hence

$$\begin{aligned}
\varphi'(w) &= \psi'(\mathrm{ext}_{\alpha,\beta})(\varphi'(w')) \\
&= \psi'(\mathrm{ext}_{\alpha,\beta})(1) \\
&= \begin{cases}
\psi(\mathrm{ext}_{\alpha,\beta})(1_R) = \varphi(w) & \text{if } \alpha \in \Sigma_{\mathrm{call}} \text{ and } \beta \in \Sigma_{\mathrm{ret}} \\
\overline{\psi(\mathrm{ext}_{a,b})(1_R)} = \overline{\varphi(ab)} & \text{if } \alpha = \overline{b} \in \overline{\Sigma}_{\mathrm{call}} \text{ and } \beta = \overline{a} \in \overline{\Sigma}_{\mathrm{ret}} \\
\bot & \text{otherwise}
\end{cases}
\end{aligned}$$

as required, because $w = \overline{b}\,\overline{a} = \overline{ab}$ in the second case. If $w' \in \Sigma^\triangle \setminus \{\varepsilon\}$, then $\varphi'(w') = \varphi(w') \in R$ by induction hypothesis, and hence

$$\begin{aligned}
\varphi'(w) = \psi'(\mathrm{ext}_{\alpha,\beta})(\varphi'(w')) &= \psi'(\mathrm{ext}_{\alpha,\beta})(\varphi(w')) \\
&= \begin{cases}
\psi(\mathrm{ext}_{\alpha,\beta})(\varphi(w')) = \varphi(w) & \text{if } \alpha \in \Sigma_{\mathrm{call}} \text{ and } \beta \in \Sigma_{\mathrm{ret}} \\
\bot & \text{otherwise}
\end{cases}
\end{aligned}$$

as required. If $w' \in \overline{\Sigma}^\triangle \setminus \{\varepsilon\}$, then $w' = \overline{w''}$ for $w'' \in \Sigma^\triangle \setminus \{\varepsilon\}$, so $\varphi'(w') = \overline{\varphi(w'')} \in \overline{R}$ by induction

hypothesis, and hence

$$\begin{aligned}\varphi'(w) &= \psi'(\mathrm{ext}_{\alpha,\beta})(\varphi'(w')) \\ &= \psi'(\mathrm{ext}_{\alpha,\beta})(\overline{\varphi(w'')}) \\ &= \begin{cases} \overline{\psi(\mathrm{ext}_{a,b})(\varphi(w''))} = \overline{\varphi(aw''b)} & \text{if } \alpha = \overline{b} \in \overline{\Sigma}_{\mathrm{call}} \text{ and } \beta = \overline{a} \in \overline{\Sigma}_{\mathrm{ret}} \\ \bot & \text{otherwise} \end{cases}\end{aligned}$$

as required, because $w = \overline{b}\,\overline{w''}\,\overline{a} = \overline{aw''b}$ in the first case. If $w' = u\#u'$ with $(u, u') \in P$, i.e. $w = \alpha u \# u' \beta$, then

$$\varphi'(w') = \{(\psi(\mathrm{ext}_{u,v'}), \psi(\mathrm{ext}_{u',v'})) \mid v' \in \Sigma^*, uv', u'v' \in \Sigma^\triangle\} \in \mathfrak{P}(O^2) \setminus \emptyset$$

by induction hypothesis. Hence, we have

$$\begin{aligned}\varphi'(w) &= \psi'(\mathrm{ext}_{\alpha,\beta})(\varphi'(w')) \\ &= \psi'(\mathrm{ext}_{\alpha,\beta})(\{(\psi(\mathrm{ext}_{u,v'}), \psi(\mathrm{ext}_{u',v'})) \mid v' \in \Sigma^*, uv', u'v' \in \Sigma^\triangle\}) \\ &= \bigcup_{\substack{b \in \Sigma_{\mathrm{ret}} \\ z \in \Sigma^\triangle}} \{(\psi(\mathrm{ext}_{au,v'bz}), \psi(\mathrm{ext}_{a'u',v'bz})) \mid v' \in \Sigma^*, uv', u'v' \in \Sigma^\triangle\} \\ &= \{(\psi(\mathrm{ext}_{au,v}), \psi(\mathrm{ext}_{a'u',v})) \mid v \in \Sigma^*, auv, a'u'v \in \Sigma^\triangle\}\end{aligned}$$

if $\alpha = a \in \Sigma_{\mathrm{call}}$ and $\beta = \overline{a'} \in \overline{\Sigma}_{\mathrm{ret}}$ (where the last inclusion from right to left follows by considering the unique stair factorizations given by Lemma 3.6 for the elements of each pair) and $\varphi'(w) = \bot$ otherwise, as required.

Finally assume $w = xy$ for some $x, y \in \Sigma'^\triangle \setminus \{\varepsilon\}$. The case when x or y is neither in $\Sigma^\triangle \cup \overline{\Sigma}^\triangle$ nor of the form $u\#\overline{u'}$ with $(u, u') \in P$ is easily handled by applying the induction hypothesis and observing that $\bot$ is a zero in R'. Four other immediate cases are when both x and y are in $\Sigma^\triangle$, when x is in $\Sigma^\triangle$ and y in $\overline{\Sigma}^\triangle$, when x is in $\overline{\Sigma}^\triangle$ and y in $\Sigma^\triangle$ and when both x and y are of the form $u\#u'$ with $(u, u') \in P$. For the case when both x and y are in $\overline{\Sigma}^\triangle$, we have that $x = \overline{x'}$ and $y = \overline{y'}$ for $x', y' \in \Sigma^\triangle$, so that $\varphi'(x) = \overline{\varphi(x')} \in \overline{R}$ and $\varphi'(y) = \overline{\varphi(y')} \in \overline{R}$ by induction hypothesis, hence

$$\varphi'(xy) = \varphi'(x) \cdot \varphi'(y) = \overline{\varphi(x')} \cdot \overline{\varphi(y')} = \overline{\varphi(y')\varphi(x')} = \overline{\varphi(y'x')}$$

as required, because $xy = \overline{x'}\,\overline{y'} = \overline{y'x'}$. Consider the case when $x \in (\Sigma^\triangle \cup \overline{\Sigma}^\triangle) \setminus \{\varepsilon\}$ and $y = u\#\overline{u'}$ with $(u, u') \in P$, hence $w = xu\#\overline{u'}$. The induction hypothesis yields

$$\varphi'(x) = \begin{cases} \varphi(x) \in R & \text{if } x \in \Sigma^\triangle \setminus \{\varepsilon\} \\ \overline{\varphi(x')} \in \overline{R} & \text{if } x = \overline{x'} \text{ for } x' \in \Sigma^\triangle \setminus \{\varepsilon\} \end{cases}$$

and $\varphi'(y) = \{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} \in \mathfrak{P}(O^2) \setminus \emptyset$. We obtain

$$\begin{aligned}\varphi'(w) &= \varphi'(x) \cdot \varphi'(y) \\ &= \varphi(x) \cdot \{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} \\ &= \{(\mathrm{left}_{\varphi(x)} \circ \psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} \\ &= \{(\psi(\mathrm{ext}_{xu,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, xuv, u'v \in \Sigma^\triangle\}\end{aligned}$$

if $x \in \Sigma^\triangle \setminus \{\varepsilon\}$ and $\varphi'(w) = \overline{\varphi(x')} \cdot \{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} = \bot$ if $x = \overline{x'}$ for $x' \in \Sigma^\triangle \setminus \{\varepsilon\}$, as required. Eventually, let us treat the case when $x = u\#u'$ with $(u, u') \in P$ and $y \in (\Sigma^\triangle \cup \overline{\Sigma}^\triangle) \setminus \{\varepsilon\}$, hence $w = u\#\overline{u'}y$. The induction hypothesis yields $\varphi'(x) = \{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} \in \mathfrak{P}(O^2) \setminus \emptyset$ and $\varphi'(y) = \begin{cases} \varphi(y) \in R & \text{if } y \in \Sigma^\triangle \setminus \{\varepsilon\} \\ \overline{\varphi(y')} \in \overline{R} & \text{if } y = \overline{y'} \text{ for } y' \in \Sigma^\triangle \setminus \{\varepsilon\} \end{cases}$. We obtain

$$\begin{aligned}
\varphi'(w) &= \varphi'(x) \cdot \varphi'(y) \\
&= \{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} \cdot \overline{\varphi(y')} \\
&= \{(\psi(\mathrm{ext}_{u,v}), \mathrm{left}_{\varphi(y')} \circ \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} \\
&= \{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{y'u',v})) \mid v \in \Sigma^*, uv, y'u'v \in \Sigma^\triangle\}
\end{aligned}$$

if $y = \overline{y'}$ for $y' \in \Sigma^\triangle \setminus \{\varepsilon\}$ and $\varphi'(w) = \{(\psi(\mathrm{ext}_{u,v}), \psi(\mathrm{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\triangle\} \cdot \varphi(y) = \bot$ if $x \in \Sigma^\triangle \setminus \{\varepsilon\}$, as required, because $\overline{y'u'} = \overline{u'}\,\overline{y'}$ in the first case. □

7 Conclusion

In this paper we have studied the question which visibly pushdown languages lie in the complexity class AC^0.

We have introduced the notions of length-synchronicity, weak length-synchronicity and quasi-counterfreeness. We have introduced intermediate VPLs: these are quasi-counterfree VPLs generated by context-free grammars G involving the production $S \to_G \varepsilon$ for the start nonterminal S and whose further productions are all of the form $T \to_G uT'v$, where uv is well-matched, $u \in (\Sigma^*_{\mathrm{int}}\Sigma_{\mathrm{call}}\Sigma^*_{\mathrm{int}})^+$, $v \in (\Sigma^*_{\mathrm{int}}\Sigma_{\mathrm{ret}}\Sigma^*_{\mathrm{int}})^+$, and the set of contexts $\{(u, v) \in \mathsf{Con}(\Sigma) \mid S \Rightarrow^*_G uSv\}$ weakly-length synchronous but not length-synchronous. To the best of our knowledge our community is unaware of whether at all there is an intermediate VPL that is provably in AC^0 (even in ACC^0) or provably not in AC^0. We conjecture that none of the intermediate VPLs are in ACC^0 nor TC^0-hard.

Our main result states that there is an algorithm that, given a visibly pushdown language L, outputs if L surely lies in AC^0, surely does not lie in AC^0 (by providing some $m > 1$ such that MOD_m is constant-depth reducible to L), or outputs a disjoint finite union of intermediate VPLs that L is constant-depth equivalant to. In the latter case one can moreover compute distinct $k, l \in \mathbb{N}_{>0}$ such that already $\mathcal{L}_{k,l} = L(S \to \varepsilon \mid ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2)$ is constant-depth reducible to L.

We conjecture that due to the particular nature of intermediate VPLs, either all of them are in AC^0 or all are not: this conjecture together with our main result indeed implies that there is an algorithm that decides if a given visibly pushdown language is in AC^0.

As main tools we carefully revisited Ext-algebras, introduced by Czarnetzki et al. [10], being closely related to forest algebras, introduced by Bojańczyk and Walukiewicz [7]. For the reductions from $\mathcal{L}_{k,l}$ we made use of Green's relations.

Natural questions arise. Is there any *concrete* intermediate VPL that is provably in ACC^0, provably not in AC^0, or hard for TC^0? Another exciting question is whether one can effectively compute those visibly pushdown languages that lie in the complexity class TC^0. Is there is a $\mathsf{TC}^0/\mathsf{NC}^1$ complexity dichotomy? For these questions new techniques seem to be necessary. In this context it is already interesting to mention there is an NC^1-complete visibly pushdown language whose syntactic Ext-algebra is aperiodic. Another exciting question is to give an algebraic characterization of the visibly counter languages.

8 Some errata in previous work

The following section summarizes some crucial errata in [24].

1. On page 176 line 8 it is written

   > if a VPL has SHB there exist unique rationals $\Delta_m^{\uparrow}$ and $\Delta_m^{\downarrow}$ such that for $(u, v) \in \eta_L^{-1}(m)$ we unambiguously have $\frac{\Delta(u)}{|u|} = \Delta_m^{\uparrow}$ and $\frac{\Delta(v)}{|v|} = \Delta_m^{\downarrow}$

   The following language is a counter-example to this claim: consider the VPL generated by the context-free grammar $S \to aSb \mid a'cSb' \mid \varepsilon$, where a, a' are call letters, b, b' are return letters and c is an internal letter. This rest of the section makes use of the above.

2. The previous point leads to problems, for instance the last two sentences on page 176 are problematic. By definition, it does not follow that for each m there is a unique slope γ such that for all $(u, v) \in \eta^{-1}(m)$ we have $\gamma = \frac{\Delta(u)}{|u|}$.

3. The reduction in Proposition 135 has some problems. Firstly, one cannot assume that $\alpha u \beta v \gamma$ is necessarily in L. It can be assumed without loss of generality though. Secondly, if $p > 2$, then $w \mapsto \alpha\phi(w)\beta\psi(w^R)\gamma$ could possibly be mapped to an element $i \in \mathbb{Z}_p$, where $i \notin \{0, 1\}$: in this case it is not clear if $\alpha\phi(w)\beta\psi(w^R)\gamma$ is in L or not.

4. Top of page 182: The quotient $n_v^{\uparrow}/d_v^{\uparrow} \in \mathbb{Q}$. As mentioned in Point 1 its existence does not follow from the definition of bounded corridor. The construction of the approximate matching (proof of Proposition 126 relies on this).

5. Page 184: The relation $\rightsquigarrow_L$ is not well-defined. Proposition 126 essentially states a property that $\rightsquigarrow_L$ should satisfy, but the relation $\rightsquigarrow_L$ is defined by the formula appearing in Proposition 126. Yet, the formulas appearing already rely on the wrong observation that unique slopes exist (Point 1 from above). This has consequences for Lemma 127, Conjecture 128, Corollary 129, Conjecture 130, Conjecture 132, and Proposition 137.

6. Conjecture 128: If one were to interpret $\rightsquigarrow_L$ it as "the matching relation", then the Conjecture 128 is easily seen to be wrong. The VPL generated by the grammar $S \to acbc \mid aSb \mid \varepsilon$ does not satisfy SHB, but its matching relation is definable in $\mathsf{FO}[\mathsf{arb}]$.

7. Page 177, line -6. It is written

   > If such an m exists, we also find such an element that is idempotent.

   The language generated by the grammar $S \to aSb \mid a_1cb_1 \mid a_2ccb_2$ is a counter-example.

8. The statement of Proposition 131 is wrong. The language $\{a^n b^n \mid n \geq 0\}^*$ is a counter-example.

9. Proposition 131: the proof has problems since the morphism is not length-multiplying.

10. Page 181: In the characterization the first bullet point is incorrect.

11. The statement of Lemma 125 is wrong. Counter-example: $L = \{a^n b^n \mid n \geq 0\}^*$. Clearly, $\text{cancel}^{|H_L|}(w) = \Sigma_{\text{int}}^*$ for all $w \in L$ but L does not have the WSHB.

12. The statement of Proposition 144 is wrong. Consider the language generated by the grammar $S \to aSb \mid a_1cb_1 \mid a_2ccb_2$ which is a visibly counter language that does not have the SHB property. However, it is in AC^0.

13. Corollary 145 is wrong due to the previous point.

14. In the proof on page 192 in line 3 one cannot assume that an idempotent $m' \in V$ exists for which $\eta_L^{-1}(m')$ is also a witness.

15. Statement of Lemma 146 is unclear since $\rightsquigarrow_L$ is not clearly defined.

16. Lemma 147 is unclear since $\rightsquigarrow_L$ is not clearly defined. There is no proof given.

References

[1] Rajeev Alur, Viraj Kumar, P. Madhusudan, and Mahesh Viswanathan. Congruences for visibly pushdown languages. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, *Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings*, volume 3580 of *Lecture Notes in Computer Science*, pages 1102–1114. Springer, 2005. doi:10.1007/11523468_89.

[2] Rajeev Alur and P. Madhusudan. Visibly pushdown languages. In László Babai, editor, *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 202–211. ACM, 2004. doi:10.1145/1007352.1007390.

[3] Vince Bárány, Christof Löding, and Olivier Serre. Regularity problems for visibly pushdown languages. In Bruno Durand and Wolfgang Thomas, editors, *STACS 2006, 23rd Annual Symposium on Theoretical Aspects of Computer Science, Marseille, France, February 23-25, 2006, Proceedings*, volume 3884 of *Lecture Notes in Computer Science*, pages 420–431. Springer, 2006. URL: https://doi.org/10.1007/11672142_34.

[4] David A. Mix Barrington. Bounded-Width Polynomial-Size Branching Programs Recognize Exactly Those Languages in NC^1. *J. Comput. Syst. Sci.*, 38(1):150–164, 1989. doi:10.1016/0022-0000(89)90037-8.

[5] David A. Mix Barrington, Kevin J. Compton, Howard Straubing, and Denis Thérien. Regular languages in nc^1. *J. Comput. Syst. Sci.*, 44(3):478–499, 1992. doi:10.1016/0022-0000(92)90014-A.

[6] Richard Beigel. The polynomial method in circuit complexity. In *Proceedings of the Eigth Annual Structure in Complexity Theory Conference, San Diego, CA, USA, May 18-21, 1993*, pages 82–95. IEEE Computer Society, 1993. doi:10.1109/SCT.1993.336538.

[7] Mikolaj Bojanczyk and Igor Walukiewicz. Forest algebras. In Jörg Flum, Erich Grädel, and Thomas Wilke, editors, *Logic and Automata: History and Perspectives [in Honor of Wolfgang Thomas]*, volume 2 of *Texts in Logic and Games*, pages 107–132. Amsterdam University Press, 2008.

[8] Ashok K. Chandra, Larry J. Stockmeyer, and Uzi Vishkin. Constant depth reducibility. *SIAM J. Comput.*, 13(2):423–439, 1984. doi:10.1137/0213028.

[9] Dmitry Chistikov and Christoph Haase. The taming of the semi-linear set. In *Automata, Languages, and Programming, ICALP*, volume 55 of *LIPIcs*, pages 128:1–128:13. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016. doi:10.4230/LIPIcs.ICALP.2016.128.

[10] Silke Czarnetzki, Andreas Krebs, and Klaus-Jörn Lange. Visibly pushdown languages and free profinite algebras. *CoRR*, abs/1810.12731, 2018. URL: http://arxiv.org/abs/1810.12731, arXiv:1810.12731.

[11] Patrick W. Dymond. Input-driven languages are in log n depth. *Inf. Process. Lett.*, 26(5):247–250, 1988. doi:10.1016/0020-0190(88)90148-2.

[12] Samuel Eilenberg. *Automata, Languages, and Machines. Volume A.* Pure and applied mathematics. Academic Press, New York, 1974. URL: https://www.sciencedirect.com/bookseries/pure-and-applied-mathematics/vol/59/part/PA.

[13] Samuel Eilenberg. *Automata, Languages, and Machines. Volume B.* Pure and applied mathematics. Academic Press, New York, 1976. URL: https://www.sciencedirect.com/bookseries/pure-and-applied-mathematics/vol/59/part/PB.

[14] Javier Esparza, Pierre Ganty, Stefan Kiefer, and Michael Luttenberger. Parikh's theorem: A simple and direct automaton construction. *Inf. Process. Lett.*, 111(12):614–619, 2011. doi:10.1016/j.ipl.2011.03.019.

[15] Javier Esparza, Michael Luttenberger, and Maximilian Schlund. A brief history of strahler numbers. In Adrian-Horia Dediu, Carlos Martín-Vide, José Luis Sierra-Rodríguez, and Bianca Truthe, editors, *Language and Automata Theory and Applications - 8th International Conference, LATA 2014, Madrid, Spain, March 10-14, 2014. Proceedings*, volume 8370 of *Lecture Notes in Computer Science*, pages 1–13. Springer, 2014. doi:10.1007/978-3-319-04921-2_1.

[16] Yuri Gurevich and Harry R. Lewis. A logic for constant-depth circuits. *Inf. Control.*, 61(1):65–74, 1984. doi:10.1016/S0019-9958(84)80062-5.

[17] Michael A. Harrison. *Introduction to Formal Language Theory.* Addison-Wesley, 1978.

[18] Johan Håstad. Almost optimal lower bounds for small depth circuits. In Juris Hartmanis, editor, *Proceedings of the 18th Annual ACM Symposium on Theory of Computing, May 28-30, 1986, Berkeley, California, USA*, pages 6–20. ACM, 1986. doi:10.1145/12130.12132.

[19] Neil Immerman. Languages that capture complexity classes. *SIAM J. Comput.*, 16(4):760–778, 1987. doi:10.1137/0216051.

[20] Neil Immerman. *Descriptive complexity.* Graduate texts in computer science. Springer, 1999. doi:10.1007/978-1-4612-0539-5.

[21] Stasys Jukna. *Boolean Function Complexity - Advances and Frontiers*, volume 27 of *Algorithms and combinatorics.* Springer, 2012. doi:10.1007/978-3-642-24508-4.

[22] Andreas Krebs, Klaus-Jörn Lange, and Michael Ludwig. Visibly counter languages and constant depth circuits. In Ernst W. Mayr and Nicolas Ollinger, editors, *32nd International Symposium on Theoretical Aspects of Computer Science, STACS 2015, March 4-7, 2015, Garching, Germany*, volume 30 of *LIPIcs*, pages 594–607. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015. doi:10.4230/LIPIcs.STACS.2015.594.

[23] Andreas Krebs, Klaus-Jörn Lange, and Michael Ludwig. Visibly counter languages and constant depth circuits. In Ernst W. Mayr and Nicolas Ollinger, editors, *32nd International Symposium on Theoretical Aspects of Computer Science, STACS 2015, March 4-7, 2015, Garching, Germany*, volume 30 of *LIPIcs*, pages 594–607. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015. doi:10.4230/LIPIcs.STACS.2015.594.

[24] Michael Ludwig. *Tree-Structured Problems and Parallel Computation.* PhD thesis, University of Tübingen, Germany, 2019. URL: https://publikationen.uni-tuebingen.de/xmlui/handle/10900/85960/.

[25] Kurt Mehlhorn. Pebbling Moutain Ranges and its Application of DCFL-Recognition. In J. W. de Bakker and Jan van Leeuwen, editors, *Automata, Languages and Programming, 7th Colloquium, Noordweijkerhout, The Netherlands, July 14-18, 1980, Proceedings*, volume 85 of *Lecture Notes in Computer Science*, pages 422–435. Springer, 1980. doi:10.1007/3-540-10003-2_89.

[26] Jean-Éric Pin. *Varieties of Formal Languages.* North Oxford, London and Plenum, New-York, 1986. (Traduction de Variétés de langages formels).

[27] Jean-Éric Pin, editor. *Handbook of Automata Theory.* European Mathematical Society Publishing House, Zürich, Switzerland, 2021. doi:10.4171/Automata.

[28] Marcel Paul Schützenberger. On finite monoids having only trivial subgroups. *Inf. Control.*, 8(2):190–194, 1965. doi:10.1016/S0019-9958(65)90108-7.

[29] Howard Straubing. *Finite Automata, Formal Logic, and Circuit Complexity.* Birkhäuser Boston, 1994. doi:10.1007/978-1-4612-0289-9.

[30] Howard Straubing. On logical descriptions of regular languages. In Sergio Rajsbaum, editor, *LATIN 2002: Theoretical Informatics, 5th Latin American Symposium, Cancun, Mexico, April 3-6, 2002, Proceedings*, volume 2286 of *Lecture Notes in Computer Science*, pages 528–538. Springer, 2002. doi:10.1007/3-540-45995-2_46.

[31] Heribert Vollmer. *Introduction to Circuit Complexity - A Uniform Approach.* Texts in Theoretical Computer Science. An EATCS Series. Springer, 1999. doi:10.1007/978-3-662-03927-4.