



**HAL**  
open science

# The $AC^0$ -Complexity Of Visibly Pushdown Languages

Stefan Göller, Nathan Grosshans

► **To cite this version:**

Stefan Göller, Nathan Grosshans. The  $AC^0$ -Complexity Of Visibly Pushdown Languages. 2023. hal-04005294v1

**HAL Id: hal-04005294**

**<https://hal.science/hal-04005294v1>**

Preprint submitted on 26 Feb 2023 (v1), last revised 17 Oct 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# 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 concern ourselves with the question which visibly pushdown languages are in the complexity class  $AC^0$ . We provide a conjectural characterization that isolates a stubborn subclass of particular one-turn visibly pushdown languages (that we call intermediate VPLs) all of which our community seems to lack tools for determining containment in  $AC^0$ . Our main result states that there is an algorithm that, given a visibly pushdown automaton, correctly outputs if its language is in  $AC^0$ , some  $m \geq 2$  such that  $MOD_m \leq_{cd} L$  (implying that  $L$  is not in  $AC^0$ ), or a finite disjoint union of intermediate languages  $L$  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 visibly pushdown language is hard for the more concrete intermediate language  $L(S \rightarrow \varepsilon \mid ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2)$ . For our proofs we revisit so-called Ext-algebras, introduced by Czarnetzki, Krebs and Lange [11], which in turn are closely related to forest algebras introduced by Bojańczyk and Walukiewicz [7], and Green's relations.

## Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>2</b>  |
| <b>2</b> | <b>Preliminaries</b>   | <b>4</b>  |
| 2.1      | Complexity and logic . . . . .                                       | 5         |
| 2.2      | Main result . . . . .  | 6         |
| 2.3      | Corollary for visibly counter languages . . . . .                    | 8         |
| <b>3</b> | <b>Language-theoretic and algebraic foundations and Ext-Algebras</b> | <b>8</b>  |
| 3.1      | Basic algebraic automata theory . . . . .                            | 8         |
| 3.2      | Ext-algebras . . . . .   | 9         |
| <b>4</b> | <b>(Weak) length-synchronicity and the nesting depth of VPLs</b>     | <b>21</b> |
| 4.1      | A pumping lemma for Ext-algebra morphisms . . . . .                  | 21        |
| 4.2      | Weak length-synchronicity and length-synchronicity . . . . .         | 22        |
| 4.3      | (Weak) length-synchronicity for one-turn VPLs . . . . .              | 27        |
| 4.4      | The nesting depth of visibly pushdown languages . . . . .            | 29        |

---

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

|          |  |           |
|----------|--|-----------|
| <b>5</b> | <b>Proof of the main theorem</b>   | <b>31</b> |
| 5.1      | Proof outline for Theorem 2.9 . . . . .  | 32        |
| 5.2      | Lower bounds . . . . .   | 32        |
| 5.2.1    | The non-solvable case . . . . .  | 35        |
| 5.3      | In $AC^0$ : Length-synchronous and quasi-aperiodic . . . . .   | 36        |
|          | Approximate matchings. . . . .   | 36        |
|          | Horizontal and vertical evaluation languages. . . . .  | 36        |
| 5.3.1    | Strategy for the proof of Theorem 5.7 . . . . .  | 37        |
| 5.3.2    | $\mathcal{V}_{\varphi_L}$ and $\mathcal{V}_{\psi_L}$ are quasi-aperiodic (Proof of Point 1) . . . . .                            | 38        |
| 5.3.3    | Quasi-aperiodicity of evaluation languages $\mathcal{E}_{\varphi_L,r}$ and $\mathcal{E}_{\psi_L,e}$ (Proof of Point 2) . . . . . | 39        |
| 5.3.4    | Approximate matching relation in $FO[+]$ (Proof of Point 3) . . . . .  | 42        |
|          | Building an approximate matching assuming predicates $\pi_e$ . . . . .   | 42        |
| 5.3.5    | Evaluation in $FO[+]$ via approximate matching and quasi-aperiodic evaluation languages (Proof of Point 4) . . . . .             | 48        |
| 5.4      | The intermediate case . . . . .  | 51        |
| 5.4.1    | Proof strategy . . . . .   | 52        |
| 5.4.2    | Computation of $k, l$ . . . . .  | 60        |
| 5.5      | Proof of Corollary 2.10 . . . . .  | 63        |
| <b>6</b> | <b>Computability and decidability: Proof of Proposition 5.1</b>  | <b>64</b> |
| <b>7</b> | <b>Conclusion</b>  | <b>75</b> |
| <b>8</b> | <b>Some errata in previous work</b>  | <b>77</b> |

## 1 Introduction

It is well-known that the regular word languages are characterized by 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 [29]. In terms of logic, these correspond to the languages definable in first-order logic  $FO[<]$  by a result of McNaughton and Papert [30]. The more general class of regular languages expressible in  $FO[arb]$ , i.e. first-order logic with arbitrary numerical predicates, coincides with the regular languages in  $AC^0$  [17, 20]. The algebraic counterpart of these are the languages whose syntactic morphism is quasi-aperiodic [5]. The regular languages are known to be in  $NC^1$ , hence the previous algebraic characterization of  $AC^0$  yields that it is decidable if a regular language is in  $AC^0$ .

Generalizing regular languages, input-driven languages were introduced by Mehlhorn [26]. 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 (VPL)*, 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 finite congruences on monoids. Extending upon these, Czarnetzki et al. introduced so-called Ext-algebras [11]; 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]. While context-free languages are generally in  $LOGCFL = SAC^1$ , the visibly pushdown languages, as the regular languages, are known to be in  $NC^1$  [12]. By a famous result of Barrington [4] there already exist regular languages that are  $NC^1$ -hard. In this paper we

concern ourselves with the question which visibly pushdown languages are in  $\text{AC}^0$  and how one can effectively decide this question.

**Related work.** Visibly pushdown languages (VPLs) were introduced by Alur and Madhusudan [2] via deterministic visibly pushdown automata (DVPA for short). Inspired by forest algebras [7] the paper [11] introduces Ext-algebras. Unfortunately, the definition of Ext-algebra morphism in [11] is incorrect in that it provably does not lead to freeness.

The regular languages that are in  $\text{AC}^0$  were effectively characterized by Barrington et al. [5]: a regular language is in  $\text{AC}^0$  if, and only if, its syntactic morphism is quasi-aperiodic. By an automata-theoretic approach, Krebs et al. [23] effectively characterized the visibly counter languages that are in  $\text{AC}^0$ . These are particular VPLs that are accepted by visibly pushdown automata that use only one stack symbol. In his PhD thesis [25] Ludwig already considers the question which VPLs are in  $\text{AC}^0$ . Yet, his conjectural characterization contains several serious flaws — a detailed 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 rigorously prove classical results like freeness and closure under division. We prove that a language of well-matched words is a VPL if, and only if, it is recognizable by a finite Ext-algebra; we also reintroduce the syntactic Ext-algebra of languages of well-matched words. While these results improve upon the constructions of [11], we use Ext-algebras as a technical tool for studying the complexity of visibly pushdown languages.

Fix a visibly pushdown alphabet  $\Sigma$ , i.e.  $\Sigma$  is partitioned into  $\Sigma_{\text{call}}$  (call letters),  $\Sigma_{\text{int}}$  (internal letters), and  $\Sigma_{\text{ret}}$  (return letters). Denoting  $\Delta(u)$  as the difference between call and return letters in  $u$ , we introduce the central notions of *length-synchronicity* and *weak length-synchronicity*: a binary relation  $R \subseteq \Sigma^* \times \Sigma^*$  is *length-synchronous* if  $|u|/|v| = |u'|/|v'|$  for all  $(u, v), (u', v') \in R$  with  $\Delta(u), \Delta(u') > 0$  and *weakly length-synchronous* if  $u = u'$  implies  $|v| = |v'|$  and  $v = v'$  implies  $|u| = |u'|$  for all  $(u, v), (u', v') \in R$  with  $\Delta(u), \Delta(u') > 0$ . Every word  $w = (x_1, y_1) \cdots (x_n, y_n)$  over the synchronization alphabet  $\Sigma_{\otimes_2} = (\Sigma_{\text{call}} \times \Sigma_{\text{ret}}) \cup \{(\varepsilon, c), (c, \varepsilon) \mid c \in \Sigma_{\text{int}}\}$  naturally induces the well-matched “one-turn” word  $w^\bowtie = x_1 \cdots x_n y_n \cdots y_1$  over  $\Sigma$  and induces the underlying pair  $R(w) = (x_1 \cdots x_n, y_n \cdots y_1) \in \Sigma^* \times \Sigma^*$ . We introduce the class of *intermediate languages*, which are one-turn visibly pushdown languages of the form  $L^\bowtie$ , where  $L \subseteq \Sigma_{\otimes_2}^*$  is a star-closed (i.e.  $L = L^*$ ) regular language whose syntactic morphism contains no non-trivial group when applied to words whose underlying pairs are in  $\Sigma^k \times \Sigma^l$  for all  $k, l \in \mathbb{N}$  and moreover  $R(L)$  is weakly length-synchronous but not length-synchronous. As far as we know our community is lacking tools to show whether at all there is some intermediate language that is provably in  $\text{AC}^0$  or provably not in  $\text{AC}^0$  — the somewhat simplest example of such an intermediate language is the language generated by the grammar  $S \rightarrow \varepsilon \mid aSb_1 \mid acSb_2$ . Similar remarks apply to  $\text{ACC}^0$ .

Our main result states that there is an algorithm that, given a DVPA  $A$  correctly outputs  $L(A) \in \text{AC}^0$ , some  $m \geq 2$  such that  $\text{MOD}_m$  is constant-depth reducible to  $L$ , or a non-empty disjoint finite union of intermediate languages that  $L(A)$  is constant-depth equivalent to and moreover outputs  $k, l \in \mathbb{N}_{>0}$  with  $k \neq l$  such that  $\mathcal{L}_{k,l} \leq_{\text{cd}} L(A)$ , where  $\mathcal{L}_{k,l}$  is a concrete intermediate language that is generated by the context-free grammar  $S \rightarrow \varepsilon \mid ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2$ .

For our main result we extensively study Ext-algebras and their syntactic morphism 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 notion of quasi-aperiodicity and the proof of our main result are

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  $\{i, \dots, j\}$ . For any function  $f : X \rightarrow Y$  and any subset  $Z \subseteq X$  we denote by  $f|_Z : Z \rightarrow Y$  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  $\Gamma$ . For all  $a \in \Gamma$  we write  $|w|_a$  to denote  $|w|_{\{a\}}$ .

A *deterministic finite automaton (DFA for short)* is a tuple  $A = (Q, \Sigma, \delta, q_0, F)$ , where  $Q$  is a finite set of *states*,  $\Sigma$  is a finite *alphabet*,  $\delta : Q \times \Sigma \rightarrow Q$  is the *transition function*,  $q_0 \in Q$  is the *initial state*, and  $F \subseteq Q$  is the set of *final states*. By  $L(A) \subseteq \Sigma^*$  we denote the *language* of  $A$ . The function  $\delta$  is naturally extended to a function  $\delta^*$  from  $Q \times \Sigma^*$  to  $Q$ . A language  $L$  is *regular* if  $L = L(A)$  for some DFA  $A$ . We refer to [18] for further details on formal language theory. A language  $L$  is *star-closed* if  $L = L^*$ . We define the languages

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

for each  $m \geq 2$ .

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

**Definition 2.1.** *The set of well-matched words over a visibly pushdown alphabet  $\Sigma$ , denoted by  $\Sigma^\Delta$ , is defined as follows:*

- $\varepsilon \in \Sigma^\Delta$  and  $c \in \Sigma^\Delta$  for all  $c \in \Sigma_{\text{int}}$ ,
- $awb \in \Sigma^\Delta$  for all  $w \in \Sigma^\Delta$ ,  $a \in \Sigma_{\text{call}}$  and  $b \in \Sigma_{\text{ret}}$ , and
- $uv \in \Sigma^\Delta$  for all  $u, v \in \Sigma^\Delta \setminus \{\varepsilon\}$ .

A well-matched word  $w \in \Sigma^\Delta$  is *one-turn* if  $w \in (\Sigma \setminus \Sigma_{\text{ret}})^*(\Sigma \setminus \Sigma_{\text{call}})^*$ . A language  $L \subseteq \Sigma^\Delta$  is *one-turn* if it contains only one-turn words. Let  $\Sigma$  be a visibly pushdown alphabet. We define  $\Delta : \Sigma^* \rightarrow \mathbb{Z}$  to be the height monoid morphism such that  $\Delta(w) = |w|_{\Sigma_{\text{call}}} - |w|_{\Sigma_{\text{ret}}}$  for all  $w \in \Sigma^*$ .

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

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

- $Q$  is a finite set of states,
- $\Sigma$  is a visibly pushdown alphabet, the input alphabet,
- $\Gamma$  is a finite alphabet, the stack alphabet,
- $q_0 \in Q$  is the initial state,
- $F \subseteq Q$  is the set of final states,

- $\perp \in \Gamma$  is the bottom-of-stack symbol, and
- $\delta: Q \times \Sigma \times \Gamma \rightarrow Q \times (\{\varepsilon\} \cup \Gamma \cup (\Gamma \setminus \{\perp\})\Gamma)$  is the transition function such that for all  $q \in Q, a \in \Sigma, \alpha \in \Gamma$ :
  - if  $a \in \Sigma_{call}$ , then  $\delta(q, a, \alpha) \in Q \times (\Gamma \setminus \{\perp\})\alpha$ ,
  - 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*  $\widehat{\delta}: Q \times \Sigma^* \times \Gamma^* \rightarrow Q \times \Gamma^*$  inductively as

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

The *language* accepted by  $A$  is the language  $L(A) = \{w \in \Sigma^* \mid \widehat{\delta}(q_0, w, \perp) \in F \times \{\perp\}\}$ . We call such a language a *visibly pushdown language* (VPL). We remark that visibly pushdown languages are always subsets of  $\Sigma^\Delta$ .

**Semi-linear sets.** Given  $d \in \mathbb{N}_{>0}$ , for  $\vec{x} = (x_1, \dots, x_d), \vec{y} = (y_1, \dots, y_d) \in \mathbb{N}^d$  we define  $\vec{x} + \vec{y} = (x_1 + y_1, \dots, x_d + y_d)$ . 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  $\vec{x} = (x_1, \dots, x_d) \in \mathbb{N}^d$  and  $n \in \mathbb{N}$  we define  $n\vec{x} = (nx_1, \dots, nx_d)$  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  $y, x_1, \dots, x_k \in \mathbb{N}^d$  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 [32, 22] 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  $\text{mod}_m$  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 an  $n$ -circuit such that  $L \cap \{0, 1\}^n = \{x_1 \dots x_n \in \{0, 1\}^n \mid C_n(x_1, \dots, x_n) = 1\}$  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  $\leq_{cd}$  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  $\Sigma, \Gamma$  finite alphabets, we write  $K \leq_{cd} L$  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 an oracle gate to an input of another oracle gate. We write  $K =_{cd} L$  if  $K \leq_{cd} L$  and  $L \leq_{cd} K$ ; 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  $L' \leq_{cd} L$  for all  $L' \in C$ . We say  $L$  is *C-complete* if  $L$  is C-hard and  $L \in C$ . The following complexity classes are relevant in this paper:

- $AC^0$  is the class of all languages decided by circuit families with AND, OR gates of unbounded fan-in, constant depth and polynomial size;

- $\text{ACC}^0$  is the class of all languages decided by circuit families with AND, OR and modular gates (for some fixed  $m$ ) of unbounded fan-in, constant depth and polynomial size;
- $\text{TC}^0$  is the class of all languages decided by circuit families with AND, OR and majority gates of unbounded fan-in, constant depth and polynomial size;
- $\text{NC}^1$  is the class of all languages decided by circuit families with 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 [21, 30] 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  $\mathcal{C}$  a class of numerical predicates and  $\Sigma$  a finite alphabet, we call  $\text{FO}_\Sigma[\mathcal{C}]$ -formula a first order formula over finite words using the alphabet  $\Sigma$  and numerical predicates from the class  $\mathcal{C}$ . On occasions, we might also consider  $\text{FO}_{\Sigma, \rightsquigarrow}[\mathcal{C}]$ -formulas that in comparison to the previous ones can use an additional binary predicate  $\rightsquigarrow$  and are interpreted on structures  $(w, M)$  with  $w \in \Sigma^*$  and  $M \subseteq [1, |w|]^2$ , where everything is interpreted as for  $\text{FO}_\Sigma[\mathcal{C}]$ -formulas on  $w$  excepted for  $\rightsquigarrow$  that is interpreted by  $M$ . Given  $\mathcal{C}$  a class of numerical predicates, by  $\text{FO}[\mathcal{C}]$  we denote the class of all languages over any finite alphabet  $\Sigma$  defined by a  $\text{FO}_\Sigma[\mathcal{C}]$ -sentence. A classical result at the interplay of circuit complexity and logic is that  $\text{AC}^0 = \text{FO}[\text{arb}]$ , where  $\text{arb}$  denotes the class of all numerical predicates (see [30, Theorem IX.2.1] or [21, Corollary 5.32]). The other numerical predicates that we will encounter in this paper are  $<$ ,  $+$  and  $\text{MOD}_m$  for all  $m \in \mathbb{N}_{>0}$  (gathered together in the set  $\text{MOD} = \{\text{MOD}_m \mid m > 0\}$ ).

## 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 we fix a visibly pushdown alphabet  $\Sigma$ .

**Definition 2.3** (Length-synchronicity/Weak length-synchronicity). *Let  $\mathcal{R} \subseteq \Sigma^* \times \Sigma^*$  be a binary relation.*

- We call  $\mathcal{R}$  length-synchronous if  $|u|/|v| = |u'|/|v'|$  for all  $(u, v), (u', v') \in \mathcal{R}$  satisfying  $uv, u'v' \in \Sigma^\Delta$  and  $\Delta(u), \Delta(u') > 0$ .
- We call  $\mathcal{R}$  weakly length-synchronous if  $u = u'$  implies  $|v| = |v'|$  and that  $v = v'$  implies  $|u| = |u'|$  for all  $(u, v), (u', v') \in \mathcal{R}$  satisfying  $uv, u'v' \in \Sigma^\Delta$  and  $\Delta(u), \Delta(u') > 0$ .

We remark that that the conditions  $uv, u'v' \in \Sigma^\Delta$  and  $\Delta(u), \Delta(u') > 0$  imply  $\Delta(v) = -\Delta(u) < 0$  and  $\Delta(v') = -\Delta(u') < 0$ . For each visibly pushdown alphabet  $\Sigma$  we define its *synchronization alphabet*

$$\Sigma_{\otimes_2} = (\Sigma_{\text{call}} \times \Sigma_{\text{ret}}) \cup \{(\varepsilon, c), (c, \varepsilon) \mid c \in \Sigma_{\text{int}}\}.$$

For each word  $w = (x_1, y_1) \cdots (x_n, y_n) \in \Sigma_{\otimes_2}^*$  we define  $w^\bowtie = x_1 \cdots x_n y_n \cdots y_1 \in \Sigma^\Delta$  and the pair of words  $\mathcal{R}(w) = (x_1 \cdots x_n, y_n \cdots y_1)$ . Note that every synchronization language  $X \subseteq \Sigma_{\otimes_2}^*$  *accepts the relation* (transduction)  $\mathcal{R}(X) = \{\mathcal{R}(w) \mid w \in X\} \subseteq \Sigma^* \times \Sigma^*$  and *generates the language of well-matched words*  $X^\bowtie = \{w^\bowtie \mid w \in X\}$ . The following remark is straightforward.

**Remark 2.4.**  $X^\bowtie$  is a one-turn VPL for every regular synchronization language  $X \subseteq \Sigma_{\otimes_2}^*$ .

The following proposition, formally following from Proposition 4.9, states that if a one-turn VPL is generated by a star-closed regular synchronization language whose relation is length-synchronous (resp. weakly-length-synchronous), then so is any star-closed regular synchronization language generating it.

**Proposition 2.5.** *Let  $L = X^{\boxtimes}$  for some star-closed regular synchronization language  $X$ . If  $\mathcal{R}(X)$  is a length-synchronous (resp. weakly length-synchronous) relation, then so is  $\mathcal{R}(Y)$  for all star-closed regular synchronization languages  $Y$  satisfying  $L = Y^{\boxtimes}$ .*

**Example 2.6.** Consider the visibly pushdown alphabet  $\Sigma$ , where  $\Sigma = \Sigma_{\text{call}} \uplus \Sigma_{\text{int}} \uplus \Sigma_{\text{ret}}$ ,  $\Sigma_{\text{call}} = \{a\}$ ,  $\Sigma_{\text{int}} = \{c\}$  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  $S \rightarrow ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2 \mid \varepsilon$ . We claim that every language  $\mathcal{L}_{k,l}$  satisfying  $k \neq l$  can be generated by a star-closed regular synchronization language accepting a weakly length-synchronous relation. Indeed, we have that  $\mathcal{L}_{k,l} = X_{k,l}^{\boxtimes}$ , where  $X_{k,l} = ((a, b_1)(c, \varepsilon)^{k-1} + (a, b_2)(c, \varepsilon)^{l-1})^*$ . Note that  $\mathcal{R}(X_{k,l})$  is weakly length-synchronous since both  $\mathcal{R}(X_{k,l})$  and  $\mathcal{R}(X_{k,l})^{-1}$  are partial functions.

On the other hand, we claim that any star-closed regular synchronization language generating  $\mathcal{L}_{k,l}$  cannot be length-synchronous for all  $k, l \in \mathbb{N}_{>0}$  satisfying  $k \neq l$ . To see this, assume some DFA  $A$  such that  $L(A)$  is star-closed and  $L(A)^{\boxtimes} = \mathcal{L}_{k,l}$ . Since  $ac^{l-1}b_2, ac^{k-1}b_1 \in \mathcal{L}_{k,l}$  there must exist words  $u, v \in L(A)$  such that  $u^{\boxtimes} = ac^{k-1}b_1$  and  $v^{\boxtimes} = ac^{l-1}b_2$ . Moreover, it must necessarily hold that  $\mathcal{R}(u) = (ac^{k-1}, b_1)$  and  $\mathcal{R}(v) = (ac^{l-1}, b_2)$ . As  $\frac{|ac^{k-1}|}{|b_1|} = k \neq l = \frac{|ac^{l-1}|}{|b_2|}$ , it follows that  $\mathcal{R}(L(A))$  is not length-synchronous.

We refer to Section 3.1 for algebraic foundations; there the syntactic morphism of a language is also formally introduced.

**Definition 2.7** (Intermediate language). *A VPL  $L \subseteq \Sigma^\Delta$  is intermediate if it is a one-turn VPL such that  $L = X^{\boxtimes}$  for a star-closed regular synchronization language  $X \subseteq \Sigma_{\otimes_2}^*$  whose syntactic morphism  $\varphi : \Sigma_{\otimes_2}^* \rightarrow M$  satisfies that the set  $\varphi(\{w \in \Sigma_{\otimes_2}^* \mid \mathcal{R}(w) = (u, v), |u| = k, |v| = l\})$  does not contain any non-trivial group for all  $k, l \in \mathbb{N}$  and moreover  $\mathcal{R}(X)$  is weakly length-synchronous but not length-synchronous.*

We remark that the languages  $\mathcal{L}_{k,l}$  from Example 2.6 are all intermediate languages. We have the following conjecture that can equivalently be formulated in terms of non-empty finite disjoint unions of intermediate languages.

**Conjecture 2.8.** *There is no intermediate language that is in  $\text{ACC}^0$  or  $\text{TC}^0$ -hard under constant-depth reductions.*

In fact, the authors are not even aware of any intermediate language that is provably not in  $\text{AC}^0$ . An indication for this is that the *robustness* [22] of intermediate languages can be proven to be constant. Further techniques, as for instance the *switching lemma* [19] or 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*

- $L(A) \in \text{AC}^0$ ,
- $m \geq 2$  such that  $\text{MOD}_m \leq_{\text{cd}} L(A)$  (hence implying  $L(A) \notin \text{AC}^0$ ),
- regular synchronization languages  $X_1, \dots, X_m$  witnessing that  $X_1^{\boxtimes}, \dots, X_m^{\boxtimes}$  are intermediate and moreover  $L =_{\text{cd}} \biguplus_{i=1}^m X_i^{\boxtimes}$ . In this case one can moreover effectively compute  $k, l \in \mathbb{N}$  with  $k \neq l$  such that  $\mathcal{L}_{k,l} \leq_{\text{cd}} L(A)$ .



## 2.3 Corollary for visibly counter languages

A *visibly counter automaton* with threshold  $m$  ( $m$ -VCA) over a visibly pushdown alphabet  $\Sigma$  is a tuple  $A = (Q, \Sigma, q_0, F, \delta_0, \dots, \delta_m)$ , 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  $\delta_i : Q \times \Sigma \rightarrow Q$  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  $(q, n) \xrightarrow{x} (q', n')$  if  $q' = \delta_{\min(n, m)}(q, x)$  and  $n' = n + \Delta(x)$ . The relation  $\xrightarrow{x}$  is naturally extended to  $\xrightarrow{w}$  for  $w \in \Sigma^*$ . By  $L(A) = \{w \in \Sigma^\Delta \mid \exists q \in F : (q_0, 0) \xrightarrow{w} (q, 0)\}$  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 the language of  $m$ -VCA are defined to be sets of well-matched words as in [3], whereas in [24] the well-matched requirement is not present.

When restricted to well-matched words, the following corollary implies the main result of [24].

**Corollary 2.10.** *There is an algorithm that, given an  $m$ -VCA  $A$ , correctly outputs that  $L(A)$  is in  $\text{AC}^0$  or some  $m \geq 2$  such that  $\text{MOD}_m \leq_{\text{cd}} L(A)$  (hence implying  $L(A) \notin \text{AC}^0$ ).*

For the proof of Corollary 2.10 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 [13, 14] and Pin [27], but also to the following central reference in automata theory [28, 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  $\omega$  such that  $x^\omega$  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  $vzx \sim vyz$  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  $\omega$  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  $\varphi: M \rightarrow N$  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  $\Sigma$  and  $\Gamma$  are finite alphabets, we call  $\varphi$  *length-multiplying* whenever there exists  $k \in \mathbb{N}$  such that  $\varphi(\Sigma) \subseteq \Gamma^k$ . Let  $\varphi: \Sigma^* \rightarrow M$  be a morphism, where  $\Sigma$  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*  $\varphi$ . (Defined in [9] for surjective morphisms.) 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, 30] for the original definition and [31] for the definition using the stability index, though it has been only formulated for surjective morphisms.)

A language  $L$  over a finite alphabet  $\Sigma$  is *recognized* by a monoid  $M$  if there is a morphism  $\varphi: \Sigma^* \rightarrow M$  and  $F \subseteq M$  such that  $L = \varphi^{-1}(F)$ . The *syntactic monoid* of a language  $L \subseteq \Sigma^*$  is the quotient of  $\Sigma^*$  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 [11], 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  $\text{left}_m$  and  $\text{right}_m$  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  $\text{left}_r$  and  $\text{right}_r$  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  $(\varphi, \psi)$  of monoid morphisms  $\varphi: R \rightarrow S$  and  $\psi: O \rightarrow P$  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  $\psi(\text{left}_r) = \text{left}_{\varphi(r)}$  and  $\psi(\text{right}_r) = \text{right}_{\varphi(r)}$ .

We write  $(\varphi, \psi): (R, O) \rightarrow (S, P)$ . The morphism  $(\varphi, \psi)$  is called *surjective* (respectively *bijective*) if both  $\varphi$  and  $\psi$  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,  $\varphi$  is totally determined by  $\psi$ , because for each  $r \in R$ , we have  $\varphi(r) = \varphi(\text{left}_r(1_R)) = \psi(\text{left}_r)(\varphi(1_R)) = \psi(\text{left}_r)(1_S)$ .

**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  $\Sigma$ .

**Definition 3.5.** For all  $(u, v) \in \Sigma^* \times \Sigma^*$  with  $uv \in \Sigma^\Delta$ , consider the function  $\text{ext}_{u,v}: \Sigma^\Delta \rightarrow \Sigma^*$  such that  $\text{ext}_{u,v}(x) = uxv$  for all  $x \in \Sigma^\Delta$ . We call

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

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

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

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

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

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

*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  $\text{ext}_{u,v} = \text{ext}_{\varepsilon, \varepsilon}$ , or there is some  $c \in \Sigma_{\text{int}}$  such that  $\text{ext}_{u,v} = \text{ext}_{\varepsilon, c}$  or  $\text{ext}_{u,v} = \text{ext}_{c, \varepsilon}$ . In any of these cases, we uniquely factorize  $\text{ext}_{u,v}$  as  $\text{ext}_{x_1, y_1}$  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  $uv \in \Sigma^\Delta$  we have  $u \in \Sigma^\Delta$  if, and only if,  $v \in \Sigma^\Delta$ . In case  $u, v \in \Sigma^\Delta$  we have  $\Delta(u) = 0$ , hence the only factorization of the desired form is indeed  $\text{ext}_{u,v} = \text{ext}_{x_1, y_1}$ , where  $x_1 = u$  and  $y_1 = v$ . Let us finally treat the case when  $u, v \notin \Sigma^\Delta$ , thus  $\Delta(u) \geq 1$  and hence  $h \geq 2$ . Let  $x$  be the maximal prefix of  $u$  satisfying  $x \in \Sigma^\Delta$  and let  $y$  be the maximal suffix of  $v$  satisfying  $y \in \Sigma^\Delta$ . 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  $u'v' \in \Sigma^\Delta$  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  $x' \in \Sigma^\Delta$ ,  $y'$  is a suffix of  $v$  satisfying  $y' \in \Sigma^\Delta$ ,  $a' \in \Sigma_{\text{call}}$ , 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  $a's \in \Sigma^\Delta$ . But then  $\pi$  is a factorization for  $\text{ext}_{su', v'z}$  for some  $z \in \Sigma^*$  which is a contradiction since  $\Delta(s) = -1$  due to  $a's \in \Sigma^\Delta$ .  $\square$

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  $\mathcal{O}(\Sigma^\Delta)$  of all functions  $\text{ext}_{u,v}$  for  $(u, v) \in \Sigma^* \times \Sigma^*$  with  $uv \in \Sigma^\Delta$ : it is a subset of  $(\Sigma^\Delta)^{\Sigma^\Delta}$  closed under composition. Thus,  $(\mathcal{O}(\Sigma^\Delta), \circ)$  is a submonoid of  $((\Sigma^\Delta)^{\Sigma^\Delta}, \circ)$ . Since for all  $w \in \Sigma^\Delta$  we have  $\text{left}_w = \text{ext}_{w, \varepsilon}$  and  $\text{right}_w = \text{ext}_{\varepsilon, w}$ , the set  $\mathcal{O}(\Sigma^\Delta)$  contains the functions  $\text{left}_w$  and  $\text{right}_w$  for all  $w \in \Sigma^\Delta$ . Hence,  $(\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta), \cdot, \circ)$  is an Ext-algebra. The following important proposition establishes freeness of  $(\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta))$ .

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

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

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

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

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

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

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

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

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

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

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

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

Otherwise, we have

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

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

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

One proves  $\bar{\psi}(\text{right}_w) = \text{right}_{\bar{\varphi}(w)}$  for all  $w \in \Sigma^\Delta$  analogously.

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

**Remark 3.8.** *The second condition in Definition 3.2, i.e. for all  $r \in R$  we have  $\psi(\text{left}_r) = \text{left}_{\varphi(r)}$  and  $\psi(\text{right}_r) = \text{right}_{\varphi(r)}$ , does not appear in the definition of Ext-algebra morphisms given in [11]. But this is actually problematic, because then Proposition 3.7 would not hold in general.*

*Indeed, consider for instance the visibly pushdown alphabet  $\Gamma$  where  $\Gamma_{\text{call}} = \{a\}$ ,  $\Gamma_{\text{int}} = \emptyset$  and  $\Gamma_{\text{ret}} = \{b\}$ , where  $R$  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  $\{\text{id}, \mathbf{0}, \mathbf{1}\}$  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  $\varphi: \Gamma^\Delta \rightarrow R$  by  $\varphi(w) = 1$  for all  $w \in \Gamma^\Delta$  and the function  $\psi: \mathcal{O}(\Gamma^\Delta) \rightarrow O$  by  $\psi(\text{ext}_{a^n, b^n}) = \text{id}$  for all  $n \in \mathbb{N}$  and  $\psi(\text{ext}_{u,v}) = \mathbf{1}$  for all  $u, v \in \Gamma^*$  with  $uv \in \Gamma^\Delta$  and  $(u \in a\Gamma^*b\Gamma^*$  or  $v \in \Gamma^*a\Gamma^*b)$ . The pair  $(\varphi, \psi)$  forms a morphism from  $(\Gamma^\Delta, \mathcal{O}(\Gamma^\Delta))$  to  $(R, O)$ , but it is not the only one sending  $\text{ext}_{a,b}$  to  $\text{id}$ , because we could also take  $\psi$  to send all elements of  $\mathcal{O}(\Gamma^\Delta)$  to  $\text{id}$ .*

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

**Example 3.10.** Consider the language  $\mathcal{L}_{1,2} = L(S \rightarrow aSb_1 \mid acSb_2 \mid \varepsilon)$  from Example 2.6 over the visibly pushdown alphabet  $\Gamma$ , where  $\Gamma_{\text{int}} = \{c\}$ ,  $\Gamma_{\text{call}} = \{a\}$  and  $\Gamma_{\text{ret}} = \{b_1, b_2\}$ . 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:

|               |         |               |         |         |         |
|---------------|---------|---------------|---------|---------|---------|
| $\cdot$       | $acb_1$ | $\varepsilon$ | $c$     | $cab_1$ | $ab_1$  |
| $acb_1$       | $acb_1$ | $acb_1$       | $acb_1$ | $acb_1$ | $acb_1$ |
| $\varepsilon$ | $acb_1$ | $\varepsilon$ | $c$     | $cab_1$ | $ab_1$  |
| $c$           | $acb_1$ | $c$           | $acb_1$ | $acb_1$ | $cab_1$ |
| $cab_1$       | $acb_1$ | $cab_1$       | $acb_1$ | $acb_1$ | $acb_1$ |
| $ab_1$        | $acb_1$ | $ab_1$        | $acb_1$ | $acb_1$ | $acb_1$ |

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

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

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

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

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

- $(a, b_2)$ :

$$\frac{r}{(a, b_2)(r)} \parallel \begin{array}{c|c|c|c|c|c} acb_1 & \varepsilon & c & cab_1 & ab_1 \\ \hline acb_1 & acb_1 & ab_1 & ab_1 & acb_1 \end{array}$$

- $(ca, b_2)$ :

$$\frac{r}{(ca, b_2)(r)} \parallel \begin{array}{c|c|c|c|c|c} acb_1 & \varepsilon & c & cab_1 & ab_1 \\ \hline acb_1 & acb_1 & cab_1 & cab_1 & acb_1 \end{array}$$

- $(ca, ab_1b_2)$ :

$$\frac{r}{(ca, ab_1b_2)(r)} \parallel \begin{array}{c|c|c|c|c|c} acb_1 & \varepsilon & c & cab_1 & ab_1 \\ \hline acb_1 & acb_1 & cab_1 & acb_1 & acb_1 \end{array}$$

- $(ca, b_1)$ :

$$\frac{r}{(ca, b_1)(r)} \parallel \begin{array}{c|c|c|c|c|c} acb_1 & \varepsilon & c & cab_1 & ab_1 \\ \hline acb_1 & cab_1 & acb_1 & acb_1 & cab_1 \end{array}$$

- $(a, ab_1b_2)$ :

$$\frac{r}{(a, ab_1b_2)(r)} \parallel \begin{array}{c|c|c|c|c|c} acb_1 & \varepsilon & c & cab_1 & ab_1 \\ \hline acb_1 & acb_1 & ab_1 & acb_1 & acb_1 \end{array}$$

- $(a, b_1)$ :

$$\frac{r}{(a, b_1)(r)} \left\| \begin{array}{c|c|c|c|c|c} acb_1 & \varepsilon & c & cab_1 & ab_1 & \\ \hline acb_1 & ab_1 & acb_1 & acb_1 & ab_1 & \end{array} \right\|.$$

Consider the unique morphism  $(\varphi, \psi): (\Gamma^\Delta, \mathcal{O}(\Gamma^\Delta)) \rightarrow (R, O)$  that (thanks to Proposition 3.7) satisfies  $\varphi(c) = c$ ,  $\psi(\text{ext}_{a,b_1}) = (a, b_1)$  and  $\psi(\text{ext}_{a,b_2}) = (a, b_2)$ . 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  $\sim$  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

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

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

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

a submonoid of  $(R/\sim)^{R/\sim}$ , is an Ext-algebra and the pair  $(\varphi, \psi)$  of functions  $\varphi: R \rightarrow R/\sim$  and  $\psi: O \rightarrow O'$  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

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

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  $\text{left}_{[r]_\sim}$  and  $\text{right}_{[r]_\sim}$  for all  $[r]_\sim \in R/\sim$ . Thus,  $(R/\sim, O')$  is an Ext-algebra.

Now define the functions  $\varphi: R \rightarrow R/\sim$  and  $\psi: O \rightarrow O'$  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  $R$ ,  $\varphi$  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  $\psi$  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

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

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

The lemma also proves that the pair  $(\varphi, \psi)$  of functions  $\varphi: R \rightarrow R/\sim$  and  $\psi: O \rightarrow O'$  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  $(\varphi, \psi)$  the *morphism associated to the congruence*  $\sim$ .

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

Note that the syntactic Ext-algebra  $(R_L, O_L)$  of  $L$  recognizes  $L$  via the syntactic morphism  $(\varphi_L, \psi_L)$ . Indeed, for all  $u, v \in \Sigma^{\Delta}$ , 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  $L \subseteq \Sigma^{\Delta}$  recognized by  $(R, O)$  is also recognized by  $(S, P)$ .*

*Proof.* Let  $L \subseteq \Sigma^{\Delta}$  be a language recognized by  $(R, O)$ . This means that there exists a morphism  $(\varphi, \psi): (\Sigma^{\Delta}, \mathcal{O}(\Sigma^{\Delta})) \rightarrow (R, O)$  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  $\sigma: O \rightarrow P$  such that  $\sigma(e)|_R = e$  for all  $e \in O$ . Let us consider the unique morphism  $(\varphi', \psi'): (\Sigma^{\Delta}, \mathcal{O}(\Sigma^{\Delta})) \rightarrow (S, P)$  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  $w \in \Sigma^{\Delta}$ :

- $w = \varepsilon$ . 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  $w' \in \Sigma^\Delta$ . Then

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

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

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

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  $(\alpha, \beta): (S, P) \rightarrow (R, O)$ . Let us define an arbitrary function  $\rho: \Sigma_{\text{int}} \rightarrow S$  such that  $\rho(c) \in \alpha^{-1}(\varphi(c))$  for all  $c \in \Sigma_{\text{int}}$  as well as an arbitrary function  $\sigma: \{\text{ext}_{a,b} \mid a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}\} \rightarrow P$  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  $(\varphi', \psi'): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (S, P)$  given by Proposition 3.7 for  $\rho$  and  $\sigma$ : we claim that it is such that  $\alpha(\varphi'(w)) = \varphi(w)$  for all  $w \in \Sigma^\Delta$ . 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  $w' \in \Sigma^\Delta$ . Then

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

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

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

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

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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  be a morphism and let  $L \subseteq \Sigma^\Delta$  be a language recognized by  $(R, O)$  via  $(\varphi, \psi)$ . Then  $(\varphi(\Sigma^\Delta), \psi(\mathcal{O}(\Sigma^\Delta))|_{\varphi(\Sigma^\Delta)})$  is a sub-Ext-algebra of  $(R, O)$  recognizing  $L$  via the surjective morphism  $(\varphi, \psi')$  where  $\psi'(\text{ext}_{u,v}) = \psi(\text{ext}_{u,v})|_{\varphi(\Sigma^\Delta)}$  for all  $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^\Delta)$ .*

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

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

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

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

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

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

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

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

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

Therefore,  $(\varphi, \psi')$  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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  be a surjective morphism, let  $L \subseteq \Sigma^\Delta$  and let  $(\varphi_L, \psi_L): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R_L, O_L)$  be the syntactic morphism of  $L$ . Then  $(R, O)$  recognizes  $L$  via  $(\varphi, \psi)$  if and only if there is a surjective morphism  $(\alpha, \beta): (R, O) \rightarrow (R_L, O_L)$  such that  $\varphi_L = \alpha \circ \varphi$  (we say that  $(\varphi_L, \psi_L)$  factors through  $(\varphi, \psi)$ ).*

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

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

hence  $(R, O)$  recognizes  $L$  via  $(\varphi, \psi)$ .

Assume now that  $(R, O)$  recognizes  $L$  via  $(\varphi, \psi)$ . This means that there exists  $F \subseteq R$  satisfying  $\varphi^{-1}(F) = L$ . Take  $w, w' \in \Sigma^\Delta$  such that  $\varphi(w) = \varphi(w')$ . Then, given any  $e \in \mathcal{O}(\Sigma^\Delta)$ , 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  $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^\Delta)$  such that  $\psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v'})$ . Then, for each  $w \in \Sigma^\Delta$ , we have that

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

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

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

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  $L \subseteq \Sigma^\Delta$  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  $L \subseteq \Sigma^\Delta$  be a language. Consider also its syntactic Ext-algebra  $(R_L, O_L)$  and its syntactic morphism  $(\varphi_L, \psi_L)$ .

**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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ . By Lemma 3.15, we have that  $(\varphi(\Sigma^\Delta), \psi(\mathcal{O}(\Sigma^\Delta)))|_{\varphi(\Sigma^\Delta)} = (R', O')$  is a sub-Ext-algebra of  $(R, O)$  recognizing  $L$  via the surjective morphism  $(\varphi, \psi')$  where  $\psi'(\text{ext}_{u,v}) = \psi(\text{ext}_{u,v})|_{\varphi(\Sigma^\Delta)}$  for all  $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^\Delta)$ . Then, by Lemma 3.16, there exists a surjective morphism  $(\alpha, \beta): (R', O') \rightarrow (R_L, O_L)$  such that  $\varphi_L = \alpha \circ \varphi$ . Thus, we have that  $(R_L, O_L)$  divides  $(R, O)$ .  $\square$

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 DVPA's to Ext-algebras and vice versa.

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

*Proof.* Let  $L \subseteq \Sigma^\Delta$  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  $A = (Q, \Sigma, \Gamma, \delta, q_0, F, \perp)$  be a DVPA. We denote by  $\pi_Q$  the projection of  $Q \times \Gamma^*$  on  $Q$  and by  $\pi_{\Gamma^*}$  the projection of  $Q \times \Gamma^*$  on  $\Gamma^*$ . It holds that  $L(A) \subseteq \Sigma^\Delta$  and additionally we have that

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

and

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

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

**Implication from left to right.** Assume that  $L$  is a VPL. This means there exists a DVPA  $A = (Q, \Sigma, \Gamma, \delta, q_0, F, \perp)$  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  $\text{left}_r$  and  $\text{right}_r$  for all  $r \in R$ , it follows that  $(R, O)$  is a finite Ext-algebra. We now prove that it recognizes  $L$ . For each  $w \in \Sigma^\Delta$ , 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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ , given by Proposition 3.7, such that for each  $c \in \Sigma_{\text{int}}$ , we have  $\varphi(c) = f_c$  and for each  $a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}$ , we have that  $\psi(\text{ext}_{a,b})$  sends any  $f \in R$  to  $g \in R$  satisfying that  $g(q, \alpha) = \pi_Q(\delta(f(p, \beta), b, \beta))$  with  $\delta(q, a, \alpha) = (p, \beta\alpha)$  for all  $q \in Q$  and  $\alpha \in \Gamma$ . We claim that for all  $w \in \Sigma^\Delta$ , 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  $c \in \Sigma_{\text{int}}$ . Then  $\varphi(w) = f_c = f_w$ .
- $w = aw'b$  for some  $a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}$  and  $w' \in \Sigma^\Delta$ . Then

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

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  $\delta$ ,

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

Thus  $\varphi(w) = f_w$ .

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

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

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

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

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

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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  such that  $L = \varphi^{-1}(F)$  for some  $F \subseteq R$ . Let us now define the DVPA

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

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

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

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  $r \in R$ ,  $w \in \Sigma^\Delta$  and  $\sigma \in \Gamma^*\perp$  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  $a \in \Sigma_{\text{call}}$ ,  $b \in \Sigma_{\text{ret}}$  and  $w' \in \Sigma^\Delta$ . Then

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

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

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

Hence,

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

Therefore,  $L$  is a VPL. □

## 4 (Weak) length-synchronicity and the nesting depth of VPLs

For the rest of this section let us fix a visibly pushdown alphabet  $\Sigma$ , a finite Ext-algebra  $(R, O)$  and consider a morphism  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ . Suitably adjusting the pumping lemma for context-free language we introduce an 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. In Section 4.3 we prove a connection between the (weak) length-synchronicity of those one-turn VPLS generated by star-closed regular synchronization languages and the (weak) length-synchronicity of the word relations they generate. Finally, we concern ourselves with the nesting depth of visibly pushdown languages in Section 4.4.

### 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  $uv \in \Sigma^\Delta$  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  $u_1, u_2, \dots$  (resp.  $v_1, v_2, \dots$ ) such that  $u_1v, u_2v, \dots \in \Sigma^\Delta$  (resp.  $uv_1, uv_2, \dots \in \Sigma^\Delta$ ) and the morphism  $\psi$  sends  $\text{ext}_{u,v}$  to the same element in  $O$  as  $\text{ext}_{u_i,v}$  (resp. as  $\text{ext}_{u,v_i}$ ).

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

- *If there exists a factor  $w \in \Sigma^\Delta$  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  $i \in \mathbb{N}$ ,  $sx^i zy^i tv \in \Sigma^\Delta$  and  $\psi(\text{ext}_{u,v}) = \psi(\text{ext}_{sx^i zy^i tv})$ .*
- *If there exists a factor  $w \in \Sigma^\Delta$  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  $i \in \mathbb{N}$ ,  $usx^i zy^i t \in \Sigma^\Delta$  and  $\psi(\text{ext}_{u,v}) = \psi(\text{ext}_{usx^i zy^i t})$ .*

*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  $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^\Delta)$  be such that there exists a factor  $w \in \Sigma^\Delta$  of  $u$  satisfying  $|w| \geq n$ . Let

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

be the stair factorization of  $\text{ext}_{u,v}$  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  $u' = x_1 a_1 \cdots x_{j-1} a_{j-1}$ ,  $v' = b_{j-1} y_{j-1} \cdots b_1 y_1$ ,  $u'' = a_j x_{j+1} \cdots a_{h-1} x_h$  and  $v'' = y_h b_{h-1} \cdots y_{j+1} b_j y_j$ , we have  $u'v', u''v'' \in \Sigma^\Delta$  and  $\text{ext}_{u,v} = \text{ext}_{u',v'} \circ \text{ext}_{x_j, \varepsilon} \circ \text{ext}_{u'',v''}$ . 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  $i \in \mathbb{N}$ ,  $x'x^i zy^i y' \in \Sigma^\Delta$  and  $\varphi(x_j) = \varphi(x'x^i zy^i y')$ . This implies that if we set  $s = u'x'$  and  $t = y'u''$ , then for all  $i \in \mathbb{N}$ , we have  $sx^i zy^i tv = \text{ext}_{u',v'} \circ \text{ext}_{x'x^i zy^i y', \varepsilon} \circ \text{ext}_{u'',v''}(\varepsilon) \in \Sigma^\Delta$  and

$$\begin{aligned} \psi(\text{ext}_{u,v}) &= \psi(\text{ext}_{u',v'}) \circ \text{left}_{\varphi(x_j)} \circ \psi(\text{ext}_{u'',v''}) \\ &= \psi(\text{ext}_{u',v'}) \circ \text{left}_{\varphi(x'x^i zy^i y')} \circ \psi(\text{ext}_{u'',v''}) \\ &= \psi(\text{ext}_{u',v'}) \circ \psi(\text{ext}_{x'x^i zy^i y', \varepsilon}) \circ \psi(\text{ext}_{u'',v''}) \\ &= \psi(\text{ext}_{sx^i zy^i tv}). \end{aligned}$$

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

## 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  $\text{TC}^0$ -hardness can be proven if the syntactic morphism maps certain  $\text{ext}_{u,v}, \text{ext}_{u',v}$  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  $(\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta))$ . Fix any VPL  $L$ , its syntactic Ext-algebra  $(R_L, O_L)$  along with its syntactic morphism  $(\varphi_L, \psi_L)$ . Assume some idempotent  $e \in O_L$  that is  $\varphi(L)$ -reachable.

We claim that if  $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v}) = e$  and  $\Delta(u), \Delta(u') > 0$  for some  $\text{ext}_{u,v}, \text{ext}_{u',v} \in \mathcal{O}(\Sigma^\Delta)$  with  $|u| \neq |u'|$ , then  $L$  is  $\text{TC}^0$ -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  $\text{TC}^0$ -complete language EQUALITY to  $L$ .

Since  $\text{ext}_{u,v}$  is  $\varphi(L)$ -reachable, we can fix some  $x, y, z \in \Sigma^*$  such that  $\varphi(xuyvz) \in L$ . 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  $h : \{0, 1\}^* \rightarrow \Sigma^*$  is the length-multiplying morphism satisfying  $h(0) = u^{|u'|}$  and  $h(1) = u'^{|u|}$ : one can prove that  $w \in \text{EQUALITY}$  if, and only if,  $h(w)v^{n \cdot (|u|+|u'|)} \in \Sigma^\Delta$  if, and only if,  $xh(w)zv^{n \cdot (|u|+|u'|)}y \in L$ .

Dually, one can show that  $L$  is  $\text{TC}^0$ -hard in case  $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u,v'}) = e$  and  $\Delta(v), \Delta(v') < 0$  for some  $\text{ext}_{u,v}, \text{ext}_{u,v'} \in \mathcal{O}(\Sigma^\Delta)$  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 binary word relations, given in Definition 2.3, to morphisms and VPLs, respectively. Recall that  $\mathcal{R} \subseteq \Sigma^* \times \Sigma^*$  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  $uv \in \Sigma^\Delta$ ,  $u'v' \in \Sigma^\Delta$ , and  $\Delta(u), \Delta(u') > 0$ .

**Definition 4.2** (Weak Length-Synchronicity). *The morphism  $(\varphi, \psi) : (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  is  $F$ -weakly-length-synchronous (where  $F \subseteq R$ ) if for all  $F$ -reachable idempotents  $e \in O$  the relation*

$$\mathcal{U}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\Delta, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = e\}$$

*is weakly length-synchronous. We call  $L \subseteq \Sigma^\Delta$  weakly length-synchronous if its syntactic morphism  $(\varphi_L, \psi_L)$  is  $\varphi_L(L)$ -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 to  $F$ -reachable elements that are not necessarily idempotent.

**Proposition 4.3.** *For all  $F \subseteq R$  we have that  $(\varphi, \psi)$  is  $F$ -weakly-length-synchronous if, and only if, for all  $F$ -reachable  $e \in O$  the relation  $\mathcal{R}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\Delta, \Delta(u) > 0, e \circ \psi(\text{ext}_{u,v}) = e\}$  is weakly length-synchronous.*

*Proof.* Let  $F \subseteq R$ .

If  $\mathcal{R}_e$  is weakly length-synchronous for all  $F$ -reachable  $e \in O$ , then in particular the relation  $\mathcal{U}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\Delta, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = e\}$  is weakly length-synchronous for all  $F$ -reachable idempotents  $e \in O$ .

Conversely, assume that  $(\varphi, \psi)$  is  $F$ -weakly-length-synchronous. Fix any  $F$ -reachable  $e \in O$ . We need to prove that  $\mathcal{R}_e$  is weakly length-synchronous. For this we consider, without of generality, some  $\text{ext}_{u,v}, \text{ext}_{u',v} \in \mathcal{O}(\Sigma^\Delta)$  satisfying  $e \circ \psi(\text{ext}_{u,v}) = e \circ \psi(\text{ext}_{u',v}) = e$ ,  $\Delta(v) < 0$ , and  $|u| \neq |u'|$ . We remark that  $\Delta(u) = \Delta(u') > 0$ . We need to prove  $|u| = |u'|$ . Let  $\omega \in \mathbb{N}_{>0}$  be the idempotent power of  $O$  and consider

$$\text{ext}_{x,y} = \text{ext}_{(u^{2 \cdot \omega} u'^{\omega})_{\omega}, v^{3 \cdot \omega^2}} = \left( \text{ext}_{u^{2 \cdot \omega} u'^{\omega}, v^{3 \cdot \omega}} \right)^\omega$$

and

$$\text{ext}_{x',y} = \text{ext}_{(u^\omega u'^{2 \cdot \omega})_{\omega}, v^{3 \cdot \omega^2}} = \left( \text{ext}_{u^\omega u'^{2 \cdot \omega}, v^{3 \cdot \omega}} \right)^\omega .$$

By definition there exists an idempotent  $e' \in O$  such that  $\psi(\text{ext}_{x,y}) = \psi(\text{ext}_{x',y}) = e'$ . Using our assumption that  $e$  is  $F$ -reachable we claim that  $e'$  is  $F$ -reachable as well. Indeed, by iterated application of  $e \circ \psi(\text{ext}_{u,v}) = e \circ \psi(\text{ext}_{u',v}) = e$ , we get  $e \circ \psi(\text{ext}_{x,y}) = e \circ \psi(\text{ext}_{x',y}) = e$ . So as  $(\varphi, \psi)$  is  $F$ -weakly-length-synchronous, it must be that  $\mathcal{U}_{e'}$  is length-synchronous, thus  $|x| = |x'|$ .



Hence, as required, we obtain

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

In the same way, one can prove that for all  $\text{ext}_{u,v}, \text{ext}_{u,v'} \in \mathcal{O}(\Sigma^\Delta)$  satisfying  $e \circ \psi(\text{ext}_{u,v}) = e \circ \psi(\text{ext}_{u,v'}) = e$ , and  $\Delta(u) > 0$ , we have  $|v| = |v'|$ .  $\square$

Using Lemma 4.1 and Proposition 4.3 and the following proposition follows immediately.

**Proposition 4.4.** *Let  $n$  be the pumping constant from Lemma 4.1, let  $F \subseteq R$ , let  $e \in O$  be  $F$ -reachable, and let  $\text{ext}_{u,v}$  be such that  $\Delta(u) > 0$  and  $e \circ \psi(\text{ext}_{u,v}) = e$ . If  $(\varphi, \psi)$  is  $F$ -weakly-length-synchronous, then the stair factorization*

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

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 word relations, given in Definition 2.3, to Ext-algebra morphisms and VPLs, respectively.

**Definition 4.5** (Length-Synchronicity). *The morphism  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  is  $F$ -weakly-length-synchronous (where  $F \subseteq R$ ) if for all  $F$ -reachable idempotents  $e \in O$  the relation*

$$\mathcal{U}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\Delta, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = e\}$$

is length-synchronous. We call a the VPL  $L \subseteq \Sigma^\Delta$  length-synchronous if its syntactic morphism  $(\varphi_L, \psi_L)$  is  $\varphi_L(L)$ -length-synchronous.

**Example 4.6.** Consider our running example  $\mathcal{L}_{1,2} = L(S \rightarrow aSb_1 \mid acSb_2 \mid \varepsilon)$ . 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  $(\varepsilon, \varepsilon)$ ,  $(acb_1, \varepsilon)$  and  $(a, b_1)$ . Also recall that  $\varphi_{\mathcal{L}_{1,2}}(\mathcal{L}_{1,2}) = \{\varepsilon, ab_1\}$ . Since  $\psi_{\mathcal{L}_{1,2}}^{-1}((\varepsilon, \varepsilon)) = \{\text{ext}_{\varepsilon, \varepsilon}\}$  and  $(acb_1, \varepsilon)$  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  $\text{ext}_{u,v}$  with  $\Delta(u) > 0$  is the idempotent  $(a, b_1)$ . However, both  $\text{ext}_{a, b_1}$  and  $\text{ext}_{ac, b_2}$ , 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  $\text{ext}_{u,v}$  and  $\text{ext}_{u',v}$  (resp.  $\text{ext}_{u,v}$  and  $\text{ext}_{u',v'}$ ) 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.

The following proposition characterizes length-synchronicity of Ext-morphisms, which will be of particular importance when approximating the matching relation of a length-synchronous VPL in terms of  $\text{FO}[+]$ . This will be an important ingredient to proving that VPLs that are both length-synchronous and quasi-aperiodic (a notion to be defined in the next Section 4.3) are in  $\text{FO}[+]$  and thus in  $\text{AC}^0$ .

**Proposition 4.7.** *The morphism  $(\varphi, \psi)$  is  $F$ -length-synchronous if, and only if, for all  $F$ -reachable  $e \in O$  the relation  $\mathcal{R}_e = \{(u, v) \mid uv \in \Sigma^\Delta, \Delta(u) > 0, e \circ \psi(\text{ext}_{u,v}) = e\}$  is length-synchronous.*

*Proof.* Assume that for all  $F$ -reachable  $e \in O$  the relation  $\mathcal{R}_e$  is length-synchronous. Then in particular  $\mathcal{R}_e$  is length-synchronous for all  $F$ -reachable idempotents  $e \in O$ . Hence  $(\varphi, \psi)$  is length-synchronous.

Conversely, assume  $(\varphi, \psi)$  is  $F$ -length-synchronous. Then the relation  $\mathcal{U}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\Delta, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = e\}$  is length-synchronous for all  $F$ -reachable idempotents  $e \in O$ . We need to prove that  $\mathcal{R}_e$  is length-synchronous for all  $F$ -reachable  $e \in O$ .

Fix any  $F$ -reachable  $e \in O$ . Moreover, fix any  $(u, v), (u', v') \in \mathcal{R}_e$ , i.e.  $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^\Delta)$ ,  $e \circ \psi(\text{ext}_{u,v}) = e \circ \psi(\text{ext}_{u',v'}) = e$ , and  $\Delta(u), \Delta(u') > 0$ . We have to prove  $\frac{|u|}{|v|} = \frac{|u'|}{|v'|}$ . In analogy to the proof of Lemma 4.3, consider

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

and

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

We have  $\psi(\text{ext}_{x,y}) = \psi(\text{ext}_{x',y'}) = e'$  for some  $F$ -reachable idempotent  $e' \in O$ . Since  $\mathcal{U}_{e'}$  is length-synchronous we have  $\frac{|x|}{|y|} = \frac{|x'|}{|y'|}$ . Hence (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  $\frac{a}{b} = \frac{c}{d} = \frac{a-c}{b-d}$ ) we obtain

$$\begin{aligned} \frac{|x|}{|y|} = \frac{|x'|}{|y'|} &\implies \frac{|x|}{|y|} = \frac{|x'|}{|y'|} = \frac{|x| + |x'|}{|y| + |y'|} = \frac{\omega^2 \cdot (|u| + |u'|)}{\omega^2 \cdot (|v| + |v'|)} \\ &\implies \frac{|x| - \omega^2 \cdot (|u| + |u'|)}{|y| - \omega^2 \cdot (|v| + |v'|)} = \frac{|x'| - \omega^2 \cdot (|u| + |u'|)}{|y'| - \omega^2 \cdot (|v| + |v'|)} \\ &\implies \frac{|u|}{|v|} = \frac{\omega^2 \cdot |u|}{\omega^2 \cdot |v|} \stackrel{(1)}{=} \frac{\omega^2 \cdot |u'|}{\omega^2 \cdot |v'|} = \frac{|u'|}{|v'|} \end{aligned} \tag{1}$$

as required.  $\square$

**Proposition 4.8.** *Let  $F \subseteq R$  and assume  $(\varphi, \psi)$  is  $F$ -weakly-length-synchronous. Then for all  $F$ -reachable  $e \in O$  the following two statements are equivalent.*

1. *The relation  $\mathcal{R}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^\Delta, \Delta(u) > 0, e \circ \psi(\text{ext}_{u,v}) = e\}$  is length-synchronous.*
2. *There exist  $\alpha \in \mathbb{Q}_{>0}$ ,  $\beta \in \mathbb{N}$ ,  $\gamma \in \mathbb{N}_{>0}$  such that for all  $(u, v) \in \mathcal{R}_e$  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  $-\Delta(v') - \beta \leq \Delta(u') \leq -\Delta(v') + \beta$ .*

(c) *For all factors  $u' \in \Sigma^*$  of  $u$  such that  $|u'| = \gamma$ , we have  $\Delta(u') \geq 1$ .*

(d) *For all factors  $v' \in \Sigma^*$  of  $v$  such that  $|v'| = \gamma$ , we have  $\Delta(v') \leq -1$ .*

*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  $\mathcal{R}_e$  is length-synchronous. Point 2 (a) follows immediately from length-synchronicity of  $\mathcal{R}_e$ . 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  $(u, v) \in \mathcal{R}_e$  and let

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

be the stair factorization of  $\text{ext}_{u,v}$  according to Lemma 3.6. Since our morphism  $(\varphi, \psi)$  is  $F$ -weakly-length-synchronous by assumption, we have  $|x_i|, |y_i| \leq n$  by Lemma 4.4. 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  $x_1 \dots a_{j-1} x_j$  is a prefix of  $u'$  and  $y_j b_{j-1} \dots y_1$  is a suffix of  $v'$ . 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  $u'' = x_1 \dots a_{j-1} x_j$  and  $v'' = y_j b_{j-1} \dots y_1$ . By maximality of  $j$  we have  $\min\{|s|, |t|\} \leq n$ . Setting  $f = \psi(\text{ext}_{u'', v''})$  and  $g = \psi(\text{ext}_{a_j x_{j+1} \dots a_{h-1} x_h, y_h b_{h-1} \dots y_{j+1} b_j})$  we have  $\psi(\text{ext}_{u,v}) = f \circ g$ . We claim that there exist  $\text{ext}_{x_g, y_g} \in \mathcal{O}(\Sigma^\Delta)$  such that  $\psi(\text{ext}_{x_g, y_g}) = g$  and  $|x_g|, |y_g| \leq |O| \cdot (n+1)$ : indeed, by the pigeonhole principle and Lemma 4.4 any  $\text{ext}_{x,y} \in \mathcal{O}(\Sigma^\Delta)$  such that  $\psi(\text{ext}_{x,y}) = g$  and  $\min(|x|, |y|) > |O| \cdot (n+1)$  must satisfy  $\Delta(x) > |O|$  and can thus be factorized as  $\text{ext}_{x,y} = \text{ext}_{x', y'} \circ \text{ext}_{x'', y''} \circ \text{ext}_{x''', y'''}$  such that  $\psi(\text{ext}_{x,y}) = \psi(\text{ext}_{x', y'}) \circ \psi(\text{ext}_{x'', y''}) \circ \psi(\text{ext}_{x''', y'''})$ , where moreover  $(x'', y'') \in \Sigma^+ \times \Sigma^+$ . Thus,  $\psi(\text{ext}_{u'' x_g, y_g v''}) = \psi(\text{ext}_{u,v})$  and therefore  $(u'' x_g, y_g v'') \in \mathcal{R}_e$ . It follows  $\alpha = \frac{|u'' x_g|}{|y_g v''|} = \frac{|u''| - |s| + |x_g|}{|y_g| + |v''| - |t|}$ , or equivalently, using  $\frac{|u'|}{|v'|} = \alpha$ :

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

$$|t| = \frac{|s| - |u'| - |x_g|}{\alpha} + |y_g| + |v'| = \frac{|s| - |x_g|}{\alpha} + |y_g| \quad (3)$$

Finally, we obtain

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

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  $\gamma$  does not depend on  $u$  nor  $v$ . We only prove Point 2 (c), the proof of Point 2 (d) is analogous. As above, let

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

be the stair factorization of  $\text{ext}_{u,v}$  according to Lemma 3.6. Let  $u'$  with  $|u'| \geq \gamma$  be a factor

of  $u$  and hence of  $x_1 a_1 x_2 \dots x_{h-1} a_{h-1} x_h$ . 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  $|x_1|, |y_1|, \dots, |x_h|, |y_h| \leq n$ , 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

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

□

### 4.3 (Weak) length-synchronicity for one-turn VPLs

In this section we relate (weak) length-synchronicity of relations generated by regular synchronization languages with (weak) length-synchronicity of the one-turn visibly pushdown languages they generate.

We recall that our running example language  $\mathcal{L}_{1,2}$  is generated by a star-closed regular synchronization language whose underlying relation is also weakly length-synchronous and not length-synchronous. The latter is not a coincidence: the following proposition implies that a one-turn VPL that is generated by a star-closed regular synchronization language is weakly length-synchronous (resp. length-synchronous) if, and only if, the relation of *any* star-closed regular synchronization language generating the language is weakly length-synchronous (resp. length-synchronous).

**Proposition 4.9.** *Let  $L = X^\boxtimes$  for some star-closed regular synchronization language  $X$ . Then the following two equivalences hold:*

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

*Proof.* Let  $L = X^\boxtimes$  for some star-closed regular synchronization language  $X$ . Hence,  $L$  is a one-turn VPL. Let us fix some DFA  $A = (Q, \Sigma_{\otimes_2}, \delta, q_0, F)$  such that  $L(A) = X$ . Let us first prove Point 1.

“Only-if”: Assume by contradiction that  $\mathcal{R}(X)$  is a weakly length-synchronous relation and that  $L = X^\boxtimes$  is not weakly length-synchronous, i.e.  $(\varphi_L, \psi_L)$  is not  $\varphi_L(L)$ -weakly-length synchronous. Then there exist  $u, v, u', v' \in \Sigma^\Delta$  such that  $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'})$  is a  $\varphi_L(L)$ -reachable idempotent and moreover either

1.  $v = v'$  and  $\Delta(u), \Delta(u') > 0$  and  $|u| \neq |u'|$ , or
2.  $u = u'$  and  $\Delta(v), \Delta(v') < 0$  and  $|v| \neq |v'|$ .

We only treat the first case, the second case can be shown analogously. Without loss of generality assume  $|u| > |u'|$ . Fix some  $\text{ext}_{x,y} \in \mathcal{O}(\Sigma^\Delta)$  and some  $w \in \Sigma^\Delta$  such that  $\varphi_L(\text{ext}_{xu,vy}(w)), \varphi_L(\text{ext}_{xu',vy}(w)) \in L$ . We remark that  $u, u' \in (\Sigma \setminus \Sigma_{\text{ret}})^+$  and  $v \in (\Sigma \setminus \Sigma_{\text{call}})^+$  since  $L$  is one-turn. Moreover,  $\Delta(u) = \Delta(u') > 0$  and  $\Delta(v) < 0$  since  $v = v'$  by our case. There are unique factorizations  $u' = \alpha_1 \alpha_2$  and  $v = \beta_2 \beta_1$ , where  $\alpha_1 \in (\Sigma \setminus \Sigma_{\text{ret}})^* \Sigma_{\text{call}}, \beta_1 \in \Sigma_{\text{ret}} (\Sigma \setminus \Sigma_{\text{call}})^*$ , and  $\alpha_2, \beta_2 \in \Sigma_{\text{int}}^*$ . Hence, for all  $i, j \geq 1$  one can factorize  $xu^i (u')^j w v^{i+j} y \in L$  as

$$xu^i (u')^j w v^{i+j} y = xu^i (u')^{j-1} \alpha_1 \alpha_2 w \beta_2 \beta_1 v^{i+j-1} y.$$

For every  $i, j \geq 1$  there exists some  $\sigma_{i,j} \tau_{i,j} \in X$  such that  $\mathcal{R}(\sigma_{i,j}) = (xu^i (u')^{j-1} \alpha_1, \beta_1 v^{i+j-1} y)$  and  $\tau_{i,j}^\boxtimes = \alpha_2 w \beta_2 \in \Sigma^\Delta$ . There are only finitely many distinct words  $\tau \in \Sigma_{\otimes_2}^*$  with  $\tau^\boxtimes = \alpha_2 w \beta_2$ , say

$N \in \mathbb{N}$  many such words. Thus, by the pigeonhole principle there exist  $(i_1, j_1), (i_2, j_2) \in \mathbb{N}_{>0} \times \mathbb{N}_{>0}$  such that

1.  $\tau_{i_1, j_1} = \tau_{i_2, j_2}$ ,
2.  $\delta^*(q_0, \sigma_{i_1, j_1}) = \delta^*(q_0, \sigma_{i_2, j_2})$ , and
3.  $(i_1, j_1) \neq (i_2, j_2)$  and  $i_1 + j_1 = i_2 + j_2 = N \cdot |Q| + 1$ .

Without loss of generality let us assume  $i_1 > i_2$ . Let  $\tau_{i_1, j_1} = \tau_{i_2, j_2} = (s_1, t_1) \cdots (s_k, t_k) \in \Sigma_{\otimes 2}^*$ , i.e.  $s_1 \cdots s_k t_k \cdots t_1 = \alpha_2 w \beta_2$ . By Point 2 we have  $\sigma_{i_1, j_1} \tau_{i_1, j_1}, \sigma_{i_2, j_2} \tau_{i_2, j_2} \in X$  and hence both

$$(xu^{i_1}(u')^{j_1-1}\alpha_1 s_1 \cdots s_k, t_k \cdots t_1 \beta_1 v^{i_1+j_1-1}y)$$

and

$$(xu^{i_2}(u')^{j_2-1}\alpha_1 s_1 \cdots s_k, t_k \cdots t_1 \beta_1 v^{i_2+j_2-1}y)$$

are in  $\mathcal{R}(L(A)) = \mathcal{R}(X)$ . We claim that the two pairs contradict our assumption that  $\mathcal{R}(X)$  is weakly length-synchronous. Indeed, the right entries are identical and the left entries have different lengths: recalling  $|u| > |u'|$  and  $i_1 + j_1 = i_2 + j_2$ , we have

$$\begin{aligned} & (i_1 - i_2)|u| > (j_2 - j_1)|u'| \\ \implies & i_1|u| + (j_1 - 1)|u'| > i_2|u| + (j_2 - 1)|u'| \\ \implies & |xu^{i_1}(u')^{j_1-1}\alpha_1 s_1 \cdots s_k| > |xu^{i_2}(u')^{j_2-1}\alpha_1 s_1 \cdots s_k| \quad . \end{aligned}$$

“If”: Assume by contradiction that  $L$  is weakly length-synchronous but  $\mathcal{R}(X)$  is not. Then there exist  $uv, u'v' \in \Sigma^\Delta$  satisfying  $(u, v), (u', v') \in \mathcal{R}(X)$ ,  $\Delta(u) > 0, \Delta(u') > 0$  such that either

1.  $u = u'$  and  $|v| \neq |v'|$  or
2.  $v = v'$  and  $|u| \neq |u'|$ .

We only treat the first case, the second case can be proven analogously. Recall that  $(R_L, O_L)$  denotes syntactic Ext-algebra and  $(\varphi_L, \psi_L)$  the syntactic morphism of  $L$ . Since  $X$  is star-closed and  $L = X^\boxtimes$  we have that  $\psi_L(\text{ext}_{u^{i+j}, v^i(v')^j})^k$  is  $\varphi_L(L)$ -reachable for all  $i, j, k \in \mathbb{N}$ . Let  $\omega$  be the idempotent power of  $O_L$ . Hence, there exists an  $\varphi_L(L)$ -reachable idempotent  $e \in O_L$  such that  $e = \psi_L(\text{ext}_{u^{\omega \cdot (i+j)}, (v^{\omega \cdot i}(v')^{\omega \cdot j})})$  for all  $i, j \geq 1$ . Considering the two cases  $(i, j) = (1, 2)$  and  $(i, j) = (2, 1)$  the entries of the left-hand sides are identical and the lengths of the right-hand sides are distinct: indeed, due to  $|v'| \neq |v|$  we have

$$\begin{aligned} |(v^\omega(v')^{\omega \cdot 2})| \cdot \omega &= \omega^2 \cdot (|v| + 2 \cdot |v'|) \\ &\neq \omega^2 \cdot (2 \cdot |v| + |v'|) \\ &= |(v^{\omega \cdot 2}(v')^\omega)| \cdot \omega, \end{aligned}$$

thus contradicting that  $L$  is weakly length-synchronous. This concludes the proof of Point 1 of the Lemma. Let us now prove Point 2.

“Only-if”: Assume by contradiction that  $\mathcal{R}(X)$  is a length-synchronous relation and that  $L$  is not length-synchronous. Then there exist  $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^\Delta)$  such that  $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'})$  is a  $\varphi_L(L)$ -reachable idempotent,  $\Delta(u) > 0, \Delta(u') > 0$ , and  $\frac{|u|}{|v|} \neq \frac{|u'|}{|v'|}$ . Fix some  $\text{ext}_{x,y} \in \mathcal{O}(\Sigma)$  and some  $w \in \Sigma^\Delta$  such that  $\psi_L(\text{ext}_{xu,vy}(w)) = \psi_L(\text{ext}_{xu',v'y}(w)) \in L$ . We have

$$\varphi_L(\text{ext}_{xu^i, v^i y}(w)), \varphi_L(\text{ext}_{x(u')^i, (v')^i y}(w)) \in L$$

for all  $i \geq 1$  since  $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v'})$  is an idempotent. As above, there exist unique factorizations  $u = \alpha_1\alpha_2$ ,  $v = \beta_2\beta_1$ , where  $\alpha_1 \in (\Sigma \setminus \Sigma_{\text{ret}})^*\Sigma_{\text{call}}$ ,  $\beta_1 \in \Sigma_{\text{ret}}(\Sigma \setminus \Sigma_{\text{call}})^*$  and  $\alpha_2, \beta_2 \in \Sigma_{\text{int}}^*$ .

By the pigeonhole principle there exist  $i, j, k \geq 1$  and words  $\varrho, \sigma, \tau \in \Sigma_{\otimes_2}^*$  and  $w_1, w_2 \in \Sigma^*$  such that

1.  $\alpha_2 w \beta_2 = w_1 w_2$  and
2.  $\varrho \sigma \tau \in X$  and  $\mathcal{R}(\varrho \sigma \tau) = (x u^{i+j+k-1} \alpha_1 w_1, w_2 \beta_1 v^{i+j+k-1} y)$ ,
3.  $\mathcal{R}(\varrho \sigma) = (x u^{i+j} \alpha_1, \beta_1 v^{i+j} y)$  and thus  $\mathcal{R}(\tau) = (\alpha_2 u^{k-1} \alpha_1 w_1, w_2 \beta_1 u^{k-1} \beta_2)$ ,
4.  $\mathcal{R}(\varrho) = (x u^{i-1} \alpha_1, \beta_1 v^{i-1} y)$  and thus  $\mathcal{R}(\sigma) = (\alpha_2 u^{j-1} \alpha_1, \beta_1 v^{j-1} \beta_2)$ , and
5.  $\delta^*(q_0, \varrho) = \delta^*(q_0, \varrho \sigma)$  and thus  $\delta^*(q_0, \varrho) = \delta^*(q_0, \varrho \sigma^N)$  for all  $N \in \mathbb{N}$ .

It follows  $\varrho \sigma^N \tau \in X$  for all  $N \in \mathbb{N}$ . Moreover, setting  $M = i + k - 1$  we have

$$(x u^{M+j \cdot N} \alpha_1 w_1, w_2 \beta_1 v^{M+j \cdot N} y) \in \mathcal{R}(L(A)) = \mathcal{R}(X) \quad \text{for all } N \in \mathbb{N} \quad . \quad (4)$$

Analogously one can show that there exists a factorization  $\alpha'_2 w \beta'_2 = w'_1 w'_2$ ,  $j' \geq 1$  and  $M' \in \mathbb{N}$  such that

$$(x (u')^{M'+j' \cdot N} \alpha'_1 w'_1, w'_2 \beta'_1 (v')^{M'+j' \cdot N} y) \in \mathcal{R}(L(A)) = \mathcal{R}(X) \quad \text{for all } N \in \mathbb{N} \quad . \quad (5)$$

Combining (4) and (5) and

$$\lim_{N \rightarrow \infty} \frac{|x u^{M+j \cdot N} \alpha_1 w_1|}{|w_2 \beta_1 v^{M+j \cdot N} y|} = \frac{|u|}{|v|} \neq \frac{|u'|}{|v'|} = \lim_{N \rightarrow \infty} \frac{|x (u')^{M'+j' \cdot N} \alpha'_1 w'_1|}{|w'_2 \beta'_1 (v')^{M'+j' \cdot N} y|}$$

it follows that  $\mathcal{R}(X)$  is not length-synchronous, a contradiction.

“If”: Assume by contradiction that  $L$  is length-synchronous but  $\mathcal{R}(X)$  is not. Then there exist  $uv, u'v' \in L = X^{\boxtimes}$  satisfying  $(u, v), (u', v') \in \mathcal{R}(X)$ ,  $\Delta(u) > 0$ ,  $\Delta(u') > 0$  such that  $\frac{|u|}{|v|} \neq \frac{|u'|}{|v'|}$ . Let  $\omega$  be the idempotent power of  $O_L$ . Thus, there exists a  $\varphi_L(L)$ -reachable idempotent  $e \in O_L$  such that  $e = \psi_L(\text{ext}_{u^\omega \cdot i, u'^\omega \cdot j, v^\omega \cdot i, v'^\omega \cdot j})^\omega$  for all  $i, j \geq 1$ . However, fixing  $j = 1$  and letting  $i$  grow to infinity and conversely, fixing  $i = 1$  and letting  $j$  grow to infinity, yield different length ratios, namely

$$\lim_{i \rightarrow \infty} \frac{|(u^{\omega \cdot i} (u')^\omega)^\omega|}{|((v')^\omega v^{\omega \cdot i})^\omega|} = \frac{|u|}{|v|} \neq \frac{|u'|}{|v'|} = \lim_{j \rightarrow \infty} \frac{|(u^\omega (u')^{\omega \cdot j})^\omega|}{|((v')^{\omega \cdot j} v^\omega)^\omega|},$$

contradicting that  $L$  is length-synchronous. □

#### 4.4 The nesting depth of visibly pushdown languages

Another central notion is the nesting depth of well-matched words, which is the Horton-Strahler number [16] of the underlying trees.

**Definition 4.10.** *The nesting depth of well-matched words is given by the function  $\text{nd}: \Sigma^\Delta \rightarrow \mathbb{N}$  defined inductively as follows:*

- $\text{nd}(\varepsilon) = 0$ ;
- $\text{nd}(c) = 0$  for all  $c \in \Sigma_{\text{int}}$ ;

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

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

**Proposition 4.11.** *For each weakly length-synchronous VPL  $L \subseteq \Sigma^{\Delta}$  there exists a constant  $d \in \mathbb{N}$  such that  $L \subseteq \{w \in \Sigma^{\Delta} \mid \text{nd}(w) \leq d\}$ .*

Proposition 4.11 is proved in several steps. We first give an equivalent characterization of weak length-synchronicity.

Next, we introduce a factorization that can be seen as a factorization that witnesses the nesting depth of a word.

**Definition 4.12.** *A nesting-maximal stair factorization of  $w \in \Sigma^{\Delta}$  with  $\text{nd}(w) \geq 1$  is a factorization of  $w$  as*

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

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

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

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

**Lemma 4.13.** *All words  $w \in \Sigma^{\Delta}$  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  $c \in \Sigma_{\text{int}}$ . Then we are again done because  $w$  contains only internal letters.
- $w = aw'b$  for  $a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}$  and  $w' \in \Sigma^{\Delta}$ . By using the inductive hypothesis,  $w'$  has a nesting-maximal stair factorization  $\text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \cdots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}(w'')$ . It directly follows that  $\text{ext}_{a, b} \circ \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \cdots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}(w'')$  is a nesting-maximal stair factorization of  $w$ .
- $w = uv$  for  $u, v \in \Sigma^{\Delta} \setminus \{\varepsilon\}$ . Then  $w$  can be decomposed as  $z_1 \cdots z_m$  with  $z_1, \dots, z_m \in \Sigma_{\text{call}}\Sigma^{\Delta}\Sigma_{\text{ret}} \cup \Sigma_{\text{int}}$  and  $m \in \mathbb{N}, m \geq 2$ . In this case, either  $z_i \in \Sigma_{\text{int}}$  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  $z_i \in \Sigma_{\text{call}}\Sigma^{\Delta}\Sigma_{\text{ret}}$  and has maximal nesting depth, i.e.  $\text{nd}(w) = \text{nd}(z_i)$ . In this second subcase, we have that  $z_i = az'_ib$  with  $a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}$  and  $z'_i \in \Sigma^{\Delta}$ . By using the inductive hypothesis,  $z'_i$  has a nesting-maximal stair factorization  $\text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \cdots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}(w'')$ . Therefore,

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

is a nesting-maximal stair factorization of  $w$ . □

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

**Lemma 4.14.** *Let  $u = a_1 v b_1 \in \Sigma^\Delta$  for some  $a_1 \in \Sigma_{\text{call}}$ ,  $b_1 \in \Sigma_{\text{ret}}$ , and  $v \in \Sigma^\Delta$  such that  $\text{nd}(u) = d > 0$ . Moreover, let  $u = \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \cdots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}(u')$  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  $u_i = \text{ext}_{a_i, b_i} \circ \text{ext}_{x_{i+1}, y_{i+1}} \circ \cdots \circ \text{ext}_{a_k, b_k}(u')$  for all  $i \in [1, k]$  and  $u_{k+1} = u'$ , we have*

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

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

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

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

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

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

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

which is a contradiction.  $\square$

We are now ready to prove Proposition 4.11.

*Proof of Proposition 4.11.* Let  $L \subseteq \Sigma^\Delta$  be a weakly length-synchronous VPL. We claim that  $\text{nd}(L) \leq n + 1$ , where  $n$  is the pumping constant from Lemma 4.1. Assume by contradiction that  $\text{nd}(u) = d$  for some  $u \in L$  and some  $d > n + 1$ . Let  $u = \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \cdots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}(u')$  be a nesting-maximal stair factorization of  $u$  according to Lemma 4.13. According to Lemma 4.14 there exists  $i \in [1, k]$  such that, setting  $u_j = \text{ext}_{a_j, b_j} \circ \text{ext}_{x_{j+1}, y_{j+1}} \circ \cdots \circ \text{ext}_{a_k, b_k}(u')$  for all  $j \in [1, k]$  and  $u_{k+1} = u'$ , we have  $\text{nd}(u) = \text{nd}(u_i) = d$  and  $\text{nd}(u_{i+1}) = d - 1$ . 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  $u_{i+1} \in \Sigma_{\text{call}} \Sigma^\Delta \Sigma_{\text{ret}}$  and  $\text{nd}(x_{i+1} u_{i+1} y_{i+1}) = \text{nd}(u_{i+1}) = d - 1$ . Hence it follows that  $\text{nd}(x_{i+1}) = d - 1$  or  $\text{nd}(y_{i+1}) = d - 1$ . Without loss of generality let us assume  $\text{nd}(y_{i+1}) = d - 1 > n$ . A simple induction shows that  $|x| \geq 2^{\text{nd}(x)} - 1 \geq \text{nd}(x)$  for all  $x \in \Sigma^\Delta$ . Thus, we have  $|y_{i+1}| \geq \text{nd}(y_{i+1}) > n$ , contradicting Proposition 4.4.  $\square$

## 5 Proof of the main theorem

Before giving an overview of the proof of Theorem 2.9 we revisit the notion of quasi-aperiodicity (a notion that has already been defined for visibly pushdown languages in [25]).

Let  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, \mathcal{O})$  for a visibly pushdown alphabet  $\Sigma$  and a finite Ext-algebra  $(R, \mathcal{O})$ . Let us define  $\mathcal{O}(\Sigma^\Delta)^{k,l} = \{\text{ext}_{u,v} \in \mathcal{O}(\Sigma^\Delta) \mid |u| = k, |v| = l\}$  for all  $k, l \in \mathbb{N}$ . We say  $(\varphi, \psi)$  is *quasi-aperiodic* if all semigroups contained in the set  $\psi(\mathcal{O}(\Sigma^\Delta)^{k,l})$  are aperiodic for all  $k, l \in \mathbb{N}$ .



The following proposition implies 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  $(\varphi_L, \psi_L)$  and  $\varphi_L(L)$ .*
2. *Given a morphism  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  for a visibly pushdown alphabet  $\Sigma$  and a finite Ext-algebra  $(R, O)$ , all of the following are decidable for  $(\varphi, \psi)$ :*
  - (a) *Quasi-aperiodicity. In case  $(\varphi, \psi)$  is not quasi-aperiodic, one can effectively compute  $k, l \in \mathbb{N}$  such that  $\psi(\mathcal{O}(\Sigma^\Delta)^{k,l})$  is not aperiodic.*
  - (b)  *$F$ -length-synchronicity for a given  $F \subseteq R$ . In case  $(\varphi, \psi)$  is not  $F$ -length-synchronous, one can effectively compute a quadruple  $(k, l, k', l') \in \mathbb{N}_{>0}^4$  such that there exist  $uv, u'v' \in \Sigma^\Delta$  and some  $F$ -reachable idempotent  $e \in O$  such that  $\psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v'}) = e$ ,  $\Delta(u) > 0$ ,  $\Delta(u') > 0$ ,  $k = |u|$ ,  $l = |v|$ ,  $k' = |u'|$ ,  $l' = |v'|$ , 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  $\Sigma$ , we apply Proposition 5.1 and compute its syntactic Ext-algebra  $(R_L, O_L)$  along with its syntactic morphism  $(\varphi_L, \psi_L)$  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  $\text{TC}^0$ -hard and hence not in  $\text{AC}^0$  (Proposition 5.4 in Section 5.2).
2. If  $L$  is not quasi-aperiodic, then one can effectively compute some  $m \geq 2$  such that  $\text{MOD}_m \leq_{\text{cd}} L$  (Proposition 5.5 in Section 5.2).
3. If  $L$  is length-synchronous and  $(\varphi_L, \psi_L)$  is quasi-aperiodic, then  $L \in \text{AC}^0$  (Theorem 5.7 in Section 5.3).
4. If a VPL  $L$  that is weakly length-synchronous but not length-synchronous, and whose syntactic morphism  $(\varphi_L, \psi_L)$  is quasi-aperiodic, one can effectively compute regular synchronization languages  $X_1, \dots, X_m$  witnessing that  $X_1^{\boxtimes}, \dots, X_m^{\boxtimes}$  are intermediate languages and moreover  $L =_{\text{cd}} \boxplus_{i=1}^m X_i^{\boxtimes}$  (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  $\mathcal{L}_{k,l} \leq_{\text{cd}} L$  (Proposition 5.24 in Section 5.4).

We refer to Section 5.5 for the proof of Corollary 2.10.

## 5.2 Lower bounds

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

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

$$L_e = \{u\#v \mid uv \in \Sigma^\Delta : \psi_L(\text{ext}_{u,v}) = e\}$$

and

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

The next lemma shows that both  $\varphi_L^{-1}(r)$  and  $L_e$  and 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  $L \subseteq \Sigma^\Delta$  a VPL. Then

- $\varphi_L^{-1}(r) \leq_{\text{cd}} L$  for all  $\varphi_L(L)$ -reachable  $r \in R_L$ , and
- $L_e \leq_{\text{cd}} L$  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  $w_r \in \Sigma^\Delta$  and  $(u_r, v_r) \in \Sigma^* \times \Sigma^*$  with  $u_r v_r \in \Sigma^\Delta$  such that  $\varphi_L(w_r) = r$  and  $\varphi_L(u_r w_r v_r) \in L$ . By definition of the syntactic morphism (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 a pair of words  $(u_{r_1, r_2}, v_{r_1, r_2}) \in \Sigma^* \times \Sigma^*$  with  $u_{r_1, r_2} v_{r_1, r_2} \in \Sigma^\Delta$  and  $\psi_L(\text{ext}_{u_{r_1, r_2}, v_{r_1, r_2}}) = e_{r_1, r_2}$ .

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

$$w \in \varphi_L^{-1}(r) \iff u_r w v_r \in L \wedge \bigwedge_{\substack{r_1, r_2 \in R_L \\ r_1 \neq r_2}} u_{r_1, r_2} w v_{r_1, r_2} \in L \leftrightarrow u_{r_1, r_2} w_r v_{r_1, r_2} \in L,$$

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

For the second point, let us fix some  $\varphi_L(L)$ -reachable  $e \in O_L$ . Fix some  $(u_e, v_e) \in \Sigma^* \times \Sigma^*$  such that  $u_e v_e \in \Sigma^\Delta$  and  $\psi_L(\text{ext}_{u_e, v_e}) = e$ . Again by definition of the syntactic morphism (Definition 3.13), for all  $e_1, e_2 \in O_L$  with  $e_1 \neq e_2$  there exists some  $r_{e_1, e_2} \in \Sigma^\Delta$  such that  $e_1(r_{e_1, e_2}) \in \psi_L(L) \Leftrightarrow e_2(r_{e_1, e_2}) \notin \psi_L(L)$ . For each such  $r_{e_1, e_2}$  fix some word  $w_{e_1, e_2} \in \Sigma^\Delta$  such that  $\varphi_L(w_{e_1, e_2}) = r_{e_1, e_2}$ . Hence, for all  $u\#v \in \Sigma^* \# \Sigma^*$  we have

$$u\#v \in L_e \iff \bigvee_{r \in R_L} uv \in \varphi_L^{-1}(r) \wedge \bigwedge_{\substack{e_1, e_2 \in O_L \\ e_1 \neq e_2}} uw_{e_1, e_2} v \in L \leftrightarrow u_e w_{e_1, e_2} v_e \in L,$$

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

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

**Proposition 5.4.** If  $L$  is not weakly length-synchronous, then  $L$  is  $\text{TC}^0$ -hard.

*Proof.* Recall that  $(R_L, O_L)$  is the syntactic Ext-algebra of  $L$  and  $(\varphi_L, \psi_L): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R_L, O_L)$  is its syntactic morphism. By assumption we have that  $(\varphi_L, \psi_L)$  is not  $\varphi_L(L)$ -weakly-length-synchronous.

Assume first there exist  $\text{ext}_{u,v}, \text{ext}_{u',v} \in \mathcal{O}(\Sigma^\Delta)$  satisfying that  $\psi_L(\text{ext}_{u,v}) = \psi_L(\text{ext}_{u',v})$  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  $L_{\psi_L(\text{ext}_{u,v})}$ . 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  $\alpha : \{0, 1\}^* \rightarrow \Sigma^*$  is the length-multiplying morphism satisfying  $\alpha(1) = u^{|u'|}$  and  $\alpha(0) = u'^{|u|}$ .
3. Accept whenever  $w' \# v^{n(|u|+|u'|)} \in L_{\psi(\text{ext}_{u,v})}$ .

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  $k \in [0, 2n]$  1's, for  $w' \# v^{n(|u|+|u'|)}$  to be in  $L_{\psi_L(\text{ext}_{u,v})}$ , it is in particular required that  $w'v^{n(|u|+|u'|)}$  is well-matched, so it is necessary and sufficient that

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

Additionally applying Lemma 5.3 we obtain  $\text{EQUALITY} \leq_{\text{cd}} L_{\psi_L(\text{ext}_{u,v})} \leq_{\text{cd}} L$ , and thus that  $\text{EQUALITY} \leq_{\text{cd}} L$  by transitivity of  $\leq_{\text{cd}}$ .

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

In conclusion, as  $\text{EQUALITY}$  is  $\text{TC}^0$ -complete under constant-depth reductions, it follows that  $L$  is  $\text{TC}^0$ -hard under constant-depth reductions.  $\square$

The following proposition has essentially already been shown in [25, 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  $\text{MOD}_m \leq_{\text{cd}} L$ .*

*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  $\psi_L(\mathcal{O}(\Sigma^\Delta)^{k,l})$  is not aperiodic. Thus, one can compute  $m \geq 2$  such that  $\psi_L(\mathcal{O}(\Sigma^\Delta)^{k,l})$  contains the additive group  $G = ([0, m-1], +, 0)$  of  $\mathbb{Z}/m\mathbb{Z}$  for some prime number  $m$ . Moreover, there exist  $\text{ext}_{u_0,v_0}, \text{ext}_{u_1,v_1} \in \mathcal{O}(\Sigma^\Delta)^{k,l}$  such that  $\psi_L(\text{ext}_{u_0,v_0}) = 0_G$  and  $\psi_L(\text{ext}_{u_1,v_1}) = 1_G$ . Since  $G$  is a group both  $\psi_L(\text{ext}_{u_0,v_0})$  and  $\psi_L(\text{ext}_{u_1,v_1})$  are  $\varphi_L(L)$ -reachable. Moreover there exist  $xy, z \in \Sigma^\Delta$  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  $h_\uparrow, h_\downarrow : \{0, 1\}^* \rightarrow \Sigma^*$  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

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

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  $w_i \in \Sigma^\Delta$  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  $\psi_L(\text{ext}_{h_\uparrow(w), h_\downarrow(w^R)}) = \psi_L(\text{ext}_{u_0,v_0}) = 0_G$ , and consequently  $\psi_L(\text{ext}_{h_\uparrow(w)^{i^{m-2}}, h_\downarrow(w^R)^{i^{m-2}}}) = 0_G$  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.  $|w|_1 \equiv i \pmod m$  for some  $i \in [1, m-1]$ . Hence  $\psi_L(\text{ext}_{h_\uparrow(w), h_\uparrow(w^R)}) = i_G \neq 0_G$

and thus  $\psi_L(\text{ext}_{h^\uparrow(w)^{i^{m-2}}, h^\uparrow(w^R)^{i^{m-2}}}) = (i^{m-1} \bmod m)_G = 1_G$  by Fermat's Little Theorem. Hence  $w_i \in L$  as required.

Altogether we obtain  $\text{MOD}_m \leq_{\text{cd}} L$ .  $\square$

### 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)  $G^{(0)}, G^{(1)}, \dots$  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  $(\varphi, \psi) : (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  is quasi-aperiodic, then  $(R, O)$  is solvable. In fact, one can prove that if  $(\varphi, \psi)$  is quasi-aperiodic, then  $(R, O)$  must contain only Abelian groups.

Our proof that  $L$  is  $\text{NC}^1$ -hard (and thus  $\text{TC}^0$ -hard) when  $(R_L, O_L)$  is not solvable can be reduced to the case for words [4], by showing that already  $\psi_L(\mathcal{O}(\Sigma^\Delta)^{k,l})$  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  $\text{NC}^1$ -hard and thus not in  $\text{AC}^0$ .*

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  $\omega$  be the idempotent power of  $G$ . For all  $g, h \in G$  there exist  $\text{ext}_{u_g, v_g}, \text{ext}_{u_h, v_h} \in \mathcal{O}(\Sigma^\Delta)$  such that

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

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

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

Hence,  $\{[g, h] \mid g, h \in G\} \subseteq \mathcal{O}(\Sigma^\Delta)^{k,l}$  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  $G \subseteq \psi_L(\mathcal{O}(\Sigma^\Delta)^{k \cdot |G|, l \cdot |G|})$ . Since the word problem of any non-solvable finite group is  $\text{NC}^1$ -hard by [4] and  $G \subseteq \psi_L(\mathcal{O}(\Sigma^\Delta)^{k \cdot |G|, l \cdot |G|})$ , it follows that the word problem for  $G$  is constant-depth reducible to  $L$ . Hence  $L$  is  $\text{NC}^1$ -hard and in particular  $\text{TC}^0$ -hard.  $\square$

### 5.3 In $\text{AC}^0$ : Length-synchronous and quasi-aperiodic

This section is devoted to the following theorem.

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

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

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 [25]. 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  $i < j$ ,  $w_i \in \Sigma_{\text{call}}$ ,  $w_j \in \Sigma_{\text{ret}}$  and  $w_{i+1} \cdots w_{j-1} \in \Sigma^\Delta$ ; we also say that  $i$  is *matched to*  $j$  in  $w$ . Observe that a word  $w$  over  $\Sigma$  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  $w \in \Sigma^\Delta$ , we denote by  $M^\Delta(w)$  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  $L \subseteq \Sigma^\Delta$*  is a function  $M : \Sigma^* \rightarrow \mathbb{N}_{>0}^2$  such that  $M(w) = M^\Delta(w)$  for all  $w \in L$  and  $M(w) \subseteq M^\Delta(w)$  for all  $w \in \Sigma^* \setminus L$ .

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

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

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

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

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

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

Define the  $\varphi_L$ -evaluation morphism  $\text{eval}_{\varphi_L} : \Gamma_{\varphi_L}^* \rightarrow R_L$  by  $\text{eval}_{\varphi_L}(s) = s$  for all  $s \in \varphi_L(\Sigma^\Delta \setminus \{\varepsilon\})$  and  $\text{eval}_{\varphi_L}(\$) = 1_{R_L}$ . Similarly, define the  $\psi_L$ -evaluation morphism  $\text{eval}_{\psi_L} : \Gamma_{\psi_L}^* \rightarrow O_L$  by  $\text{eval}_{\psi_L}(f) = f$  for all  $f \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow)$  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  $(\varphi_L, \psi_L)$  is quasi-aperiodic.
  - Then  $\mathcal{E}_{\varphi_L, r}$  is a regular language whose syntactic morphism is quasi-aperiodic for all  $r \in R_L$ .
  - If  $L$  is length-synchronous, then  $\mathcal{E}_{\psi_L, e}$  is a regular language whose syntactic morphism is quasi-aperiodic for all  $e \in O_L$ .
3. Proposition 5.12: If  $L \subseteq \Sigma^\Delta$  is length-synchronous, then there exists an  $\text{FO}_\Sigma[+]$ -formula  $\mu(x, y)$  such that  $M : \Sigma^* \rightarrow \mathbb{N}_{>0}^2$  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  $r \in R_L$ , and
  - $\mathcal{E}_{\psi_L, e}$  is a regular language whose syntactic morphism is quasi-aperiodic for all  $e \in O_L$ .

Then there exists an  $\text{FO}_{\Sigma, \rightsquigarrow}[+]$ -sentence  $\eta$  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). Points Point 2 and 3 together imply the precondition of Point 4: recalling that length-synchronicity implies weak length-synchronicity 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, whereas Point 3 provides a first-order definable approximate matching relation relative to  $L$ , in turn being a predicate assumed by Point 4. Finally, Point 4 implies 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  $\varphi_L(\Sigma^\Delta \cap \Sigma^i) = \varphi_L(\Sigma^\Delta \cap \Sigma^j)$  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  $\psi_L(\mathcal{O}(\Sigma^\Delta)_\dagger^{i,*}) = \psi_L(\mathcal{O}(\Sigma^\Delta)_\dagger^{j,*})$  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  $\psi_L(\mathcal{O}(\Sigma^\Delta)_\dagger^{*,i}) = \psi_L(\mathcal{O}(\Sigma^\Delta)_\dagger^{*,j})$  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 [15, Section 3] it follows that  $S_r = \{|w| : w \in \Sigma^\Delta, \varphi_L(w) = r\} \subseteq \mathbb{N}$  is a semilinear set for all  $r \in R_L$ . It follows that for all  $U \subseteq R_L$  the set  $S_U = \{|w| : \varphi_L(w) \in U\} \subseteq \mathbb{N}$  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  $L_e = \{u\#v \mid uv \in \Sigma^\Delta : \psi_L(\text{ext}_{u,v}) = e\}$  is a VPL for all  $e \in O_L$ . As the language  $K = \{u\#v \mid u, v \in \Sigma^\Delta\}$  is obviously a VPL, it follows that for all  $e \in O_L$  the language

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

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

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

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

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

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

The following lemma holds irrespective of whether the syntactic morphism  $(\varphi_L, \psi_L)$  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  $\psi_L(\mathcal{O}(\Sigma^\Delta)_\dagger^{i,*}) = \psi_L(\mathcal{O}(\Sigma^\Delta)_\dagger^{j,*})$  for all  $i, j \in \mathbb{N}$  satisfying  $i, j \geq t$  and  $i \equiv j \pmod{p}$ . Define  $\theta_{t,p}: \mathbb{N} \rightarrow \mathbb{N}$  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  $h: \Gamma_{\psi_L}^* \rightarrow M$  to be its syntactic morphism.

If there exists  $f \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow)$  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  $n \in \mathbb{N}_{>0}$ ,  $k_1, \dots, k_{n+1} \in \mathbb{N}$  and  $f_1, \dots, f_n \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow)$ , we have  $h(\$^{k_1} f_1 \dots \$^{k_n} f_n \$^{k_{n+1}}) = h(\$^{\theta_{t,p}(k_1)} f_1 \alpha \$^{\theta_{t,p}(k_{n+1})})$ , where

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

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  $h(\Gamma_{\psi_L}^l) = h(\Gamma_{\psi_L}^{q \cdot l})$ . Thus, to show that  $h$  is quasi-aperiodic it is sufficient to prove that for all  $m \in h(\Gamma_{\psi_L}^{q \cdot l})$ , we have  $m^2 = m^3$ . Indeed, given  $m \in h(\Gamma_{\psi_L}^{q \cdot l})$ , 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  $f \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow)$  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  $f \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow)$  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  $f \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow^{k_1+k_2+1,*})$ , 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,  $f \notin \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow^{k_1+k_2+1,*})$  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.  $\square$

### 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 states that the respective evaluation languages  $\mathcal{E}_{\varphi_L, r}$  and  $\mathcal{E}_{\psi_L, e}$  are all quasi-aperiodic if the syntactic morphism  $(\varphi_L, \psi_L)$  of  $L$  is and the latter is additionally length-synchronous.



**Proposition 5.10.** *Let  $L$  be a VPL whose syntactic morphism  $(\varphi_L, \psi_L)$  is quasi-aperiodic.*

- *Then  $\mathcal{E}_{\varphi_L, r}$  is a regular language whose syntactic morphism is quasi-aperiodic for all  $r \in R_L$ .*
- *If  $L$  is length-synchronous, then  $\mathcal{E}_{\psi_L, e}$  is a regular language whose syntactic morphism is quasi-aperiodic for all  $e \in O_L$ .*

*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  $\varphi(\Sigma^\Delta \cap \Sigma^k)$  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  $L$  is length-synchronous, then there exist  $k, l \in \mathbb{N}$  such that  $\psi_L(\mathcal{O}(\Sigma^\Delta)^{k, l})$  contains a semigroup that is not aperiodic.

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

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 length-synchronicity of  $L$ .

Take  $t \in \mathbb{N}$  and  $p \in \mathbb{N}_{>0}$  given by Lemma 5.8 such that  $\psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow^{i, *}) = \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow^{j, *})$  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  $h: \Gamma_{\psi_L}^* \rightarrow M$  to be its syntactic morphism. Let  $s \in \mathbb{N}_{>0}$  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  $\psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow^{q \cdot s \cdot \omega + k + 1, *}) = \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow^{q \cdot s \cdot (\omega + 1) + k + 1, *})$  for all  $k \in \mathbb{N}$ . Therefore, we have  $w = \$^{k_1} f_1 \dots \$^{k_n} f_n \$^{k_{n+1}}$  for  $n \in \mathbb{N}_{>0}$ ,  $k_1, \dots, k_{n+1} \in \mathbb{N}$  and  $f_1, \dots, f_n \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow)$ . 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  $x', \$^{k_x + k_1} f_1, \$^{k_2} f_2, \dots, \$^{k_n} f_n, \$^{k_{n+1} + k_1} f_1, \$^{k_{n+1} + k_y} y' \in \mathcal{V}_{\psi_L}$  and  $\text{eval}_{\psi_L}(x') \circ (f_1 \circ \dots \circ f_n)^\omega \circ \text{eval}_{\psi_L}(y') = e$ . 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

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

Thus we have  $(f_1 \circ \dots \circ f_n)^\omega \neq (f_1 \circ \dots \circ f_n)^{\omega+1}$  and  $\$^{k_{n+1} + k_1} f_1 \$^{k_2} f_2 \dots \$^{k_n} f_n \in \mathcal{V}_{\psi_L}$ . 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  $(f_1 \circ \dots \circ f_n)^\omega \neq (f_1 \circ \dots \circ f_n)^{\omega+1}$  with  $(f_1 \circ \dots \circ f_n)^i \in \psi_L(\mathcal{O}(\Sigma^\Delta)_{\dagger}^{q \cdot s \cdot i, *}) = \psi_L(\mathcal{O}(\Sigma^\Delta)_{\dagger}^{q \cdot s, *})$  for each  $i \in \mathbb{N}_{>0}$  because  $k_{n+1} + k_1 + \dots + k_n + n = q \cdot s \geq t$  and  $k_{n+1} + k_1 + \dots + k_n + n = q \cdot s \equiv 0 \pmod{p}$ . But given  $\omega'$  the idempotent power of  $\{(f_1 \circ \dots \circ f_n)^i \mid i \in \mathbb{N}_{>0}\}$ , we have that  $(f_1 \circ \dots \circ f_n)^\omega = (f_1 \circ \dots \circ f_n)^{\omega\omega'} = (f_1 \circ \dots \circ f_n)^{\omega'}$ , so that  $(f_1 \circ \dots \circ f_n)^{\omega'} \neq (f_1 \circ \dots \circ f_n)^{\omega'+1}$ , hence  $\{(f_1 \circ \dots \circ f_n)^i \mid i \in \mathbb{N}_{>0}\}$  is not aperiodic.

Assume  $L$  is length-synchronous. For each  $i \in \mathbb{N}_{>0}$ , let  $\text{ext}_{u_i, v_i} \in \psi_L(\mathcal{O}(\Sigma^\Delta)_{\dagger}^{q \cdot s, *})$  such that  $\psi_L(\text{ext}_{u_i, v_i}) = (f_1 \circ \dots \circ f_n)^i$ . If  $(f_1 \circ \dots \circ f_n)^\omega$  were not  $\varphi_L(L)$ -reachable, then it would imply that  $\text{ext}_{u_\omega, v_\omega}$  is not  $L$ -reachable. This would in turn entail that for all  $w \in \Sigma^\Delta$  and  $\text{ext}_{x, y} \in \mathcal{O}(\Sigma^\Delta)$  we have

$$\text{ext}_{x, y}(\text{ext}_{u_\omega, v_\omega}(w)) \notin L \wedge \text{ext}_{x, y}(\text{ext}_{u_1 u_\omega, v_\omega v_1}(w)) = \text{ext}_{x u_1, v_1 y}(\text{ext}_{u_\omega, v_\omega}(w)) \notin L,$$

so that it would follow that  $(f_1 \circ \dots \circ f_n)^\omega = \psi_L(\text{ext}_{u_\omega, v_\omega}) = \psi_L(\text{ext}_{u_1 u_\omega, v_\omega v_1}) = (f_1 \circ \dots \circ f_n)^{\omega+1}$ , a contradiction. Hence, since  $(f_1 \circ \dots \circ f_n)^\omega$  is a  $\varphi_L(L)$ -reachable idempotent and  $(\varphi_L, \psi_L)$  is  $\varphi_L(L)$ -length-synchronous, it follows that for all  $i \in \mathbb{N}$ ,  $i \geq 2$ , since  $\psi_L(\text{ext}_{u_1^\omega, v_1^\omega}) = \psi_L(\text{ext}_{u_i^\omega, v_i^\omega}) = (f_1 \circ \dots \circ f_n)^\omega$  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  $\{(f_1 \circ \dots \circ f_n)^i \mid i \in \mathbb{N}_{>0}\}$  is contained in  $\psi_L(\mathcal{O}(\Sigma^\Delta)^{q \cdot s, |v_1|})$ .  $\square$

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 that  $L$  is weakly length-synchronous. Indeed, the VPL  $K$  generated by the grammar with rules*

$$\begin{aligned} S &\rightarrow aSb_1 \mid aTb_2 \mid \varepsilon \\ T &\rightarrow aTb_1 \mid acSb_2. \end{aligned}$$

*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  $(\varphi_K, \psi_K)$  of  $K$ , we claim that there exists  $e \in O_K$  such that  $\mathcal{E}_{\psi_K, e}$  is a regular language whose syntactic morphism is not quasi-aperiodic.*

*Let  $\Gamma$  the visibly pushdown alphabet of  $K$ . Note that we have we have  $K \subset \mathcal{L}_{1,2}$ , where  $\mathcal{L}_{1,2} = L(S \rightarrow aSb_1 \mid acSb_2 \mid \varepsilon)$  is the VPL initially introduced in Example 2.6. 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  $xy, z \in \Gamma^\Delta$ . This implies that if we set  $e_0 = \psi_K(\text{ext}_{a, b_1})$  and  $e_1 = \psi_K(\text{ext}_{ac, b_2})$ , 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  $\psi_K(\text{ext}_{u, v}) = e_{|u|_c \pmod{2}}$ . 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  $\beta: \{0, 1\}^* \rightarrow \Gamma_{\psi_K}^*$  such that  $\beta(0) = e_0 e_0$  and  $\beta(1) = e_1$ . Then  $\text{MOD}_2 = \beta^{-1}(\mathcal{E}_{\psi_K, e_0})$ , 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 [31]), we would have that  $\text{MOD}_2$  has a quasi-aperiodic syntactic morphism.*

### 5.3.4 Approximate matching relation in $\text{FO}[+]$ (Proof of Point 3)

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

**Proposition 5.12.** *If  $L \subseteq \Sigma^{\Delta}$  is length-synchronous, then there exists an  $\text{FO}_{\Sigma}[+]$ -formula  $\eta(x, y)$  such that  $M: \Sigma^* \rightarrow \mathbb{N}_{>0}^2$  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 for the the following lemma whose proof is postponed and will most part of this section. The following lemma realizes the characterization of length-synchronicity given by Proposition 4.8 via an  $\text{FO}_{\Sigma}[+]$ -formula.

**Lemma 5.13.** *Assume that  $(\varphi_L, \psi_L)$  is weakly length-synchronous. Let  $e \in O_L$  be  $\varphi_L(L)$ -reachable and let  $\mathcal{R}_e = \{(u, v) \mid uv \in \Sigma^{\Delta}, \Delta(u) > 0, e \circ \psi_L(\text{ext}_{u,v}) = e\}$  be length-synchronous. Then there exists an  $\text{FO}_{\Sigma}[+]$ -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  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\Delta}$  and
- if  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^{\Delta}$  and  $(w_i \dots w_{i'}, w_{j'} \dots w_j) \in \mathcal{R}_e$ , then  $w \models \pi_e(i, i', j', j)$ .

**Building an approximate matching assuming predicates  $\pi_e$ .** Let us prove Proposition 5.12 by making use of Lemma 5.13.

*Proof of Proposition 5.12.* By assumption  $(\varphi_L, \psi_L)$  is  $\varphi_L(L)$ -length-synchronous. Thus, the relation  $\mathcal{R}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid uv \in \Sigma^{\Delta}, \Delta(u) > 0, e \circ \psi_L(\text{ext}_{u,v}) = e\}$  is length-synchronous for all  $\varphi_L(L)$ -reachable  $e \in O_L$  by Proposition 4.7. Moreover, there exists  $d_L \in \mathbb{N}$  bounding the nesting depth of the words in  $L$  Proposition 4.11. For defining our desired formula  $\mu$ , we will construct  $\text{FO}_{\Sigma}[+]$  formulas  $\mu_d$  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  $w_i \cdots w_j \in \Sigma^{\Delta}$ ,
- if  $w \in L, w_i \dots w_j \in \Sigma^{\Delta}, \text{nd}(w_i \dots w_j) \leq d$  and  $i$  is matched to  $j$  in  $w$ , then  $w \models \mu_d^{\uparrow}(i, j)$ , and
- if  $w \in L, \text{nd}(w_i \dots w_j) \leq d$  and  $w_i \dots w_j \in \Sigma^{\Delta}$ , then  $w \models \mu_d(i, j)$ .

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

$$\mu_0(i, j) = \perp \text{ and } \mu_0(i, j) = \forall z (x \leq z \leq y \rightarrow \Sigma_{\text{int}}(z)).$$

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

$$\mu_d(x, y) = \forall z \left[ x \leq z \leq y \rightarrow \left( \Sigma_{\text{int}}(z) \vee \exists z' \left( (\Sigma_{\text{call}}(z) \wedge \Sigma_{\text{ret}}(z') \wedge \mu_d^{\uparrow}(z, z')) \vee (\Sigma_{\text{call}}(z') \wedge \Sigma_{\text{ret}}(z) \wedge \mu_d^{\uparrow}(z', z)) \right) \right) \right].$$

It remains to define  $\mu_d^{\uparrow}$ . Let us assume  $u = w_i \dots w_j \in \Sigma^{\Delta}$ , that  $i$  is matched to  $j$  in  $w$  and that  $\text{nd}(u) = d > 0$ . Hence,  $u = a_1 v b_1 \in \Sigma^{\Delta}$  for some  $a_1 \in \Sigma_{\text{call}}, b_1 \in \Sigma_{\text{ret}}$ , and  $v \in$

$\Sigma^\Delta$ . We then apply Lemma 4.14 which states that  $u$  has a nesting-maximal stair factorization  $u = \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \dots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}(u')$  such that for some  $h \in [1, k]$ , setting  $u_\ell = \text{ext}_{a_\ell, b_\ell} \circ \text{ext}_{x_{\ell+1}, y_{\ell+1}} \circ \dots \circ \text{ext}_{a_k, b_k}(u')$  for all  $\ell \in [1, k]$  and  $u_{k+1} = u'$ , we have

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

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

- (a) the infix  $w_{i_{h+1}} \dots w_{j_{h-1}} = \text{ext}_{x_{h+1}, y_{h+1}}(u_{h+1})$  is well-matched, and
- (b) the word  $w_{i_1} \dots w_{i_h} w_{j_h} \dots w_{j_1}$  is well-matched.

Point (a) can be realized via the formula  $\mu_{d-1}$  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:

$$\begin{aligned} \kappa_h(x, x', y', y) &= \exists x_1 \dots x_h \exists y_1 \dots y_h \left( x = x_1 \wedge x' = x_h \wedge y' = y_h \wedge y = y_h \wedge \right. \\ &\quad \left. \bigwedge_{t=1}^h \Sigma_{\text{call}}(x_t) \wedge \Sigma_{\text{ret}}(y_t) \wedge \bigwedge_{t=1}^{h-1} \mu_{d-1}(x_{t+1} + 1, x_t - 1) \wedge \mu_{d-1}(y_{t+1} + 1, y_t - 1) \right) \end{aligned}$$

The problem with this approach is that the size of the formula depends on the size of  $w$ . For instance, for  $a \in \Sigma_{\text{call}}$ ,  $b \in \Sigma_{\text{ret}}$ ,  $c \in \Sigma_{\text{int}}$ , and  $u = a^n c b^n$  we have  $\text{nd}(u) = \text{nd}(acb) = 1$  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

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

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  $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \dots < s_q < t_q \leq s_{q+1} = h$  such that  $[s_1, t_1], \dots, [s_q, t_q]$  are all repetitive and for  $D_0 = [t_0, s_1], D_1 = [t_1, s_2], \dots, D_q = [t_q, s_{q+1}]$  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  $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \dots < s_q < t_q \leq s_{q+1} = h$  (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  $[s_1, t_1], \dots, [s_q, t_q]$  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  $D_0 = [t_0, s_1], D_1 = [t_1, s_2], \dots, D_q = [t_q, s_{q+1}]$ . 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  $X = \{s_1, \dots, s_q\}$ , 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  $q + \sum_{p=0}^q |D_p| = q + \left| \bigcup_{p=0}^q D_p \right| \leq |O_L| + |X| + |\lambda([1, h])| = 3|O_L|$ .  $\square$

Let  $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \dots < s_q < t_q \leq s_{q+1} = h$  be the indices satisfying Claim 5.14 along with  $D_0 = [t_0, s_1], D_1 = [t_1, s_2], \dots, D_q = [t_q, s_{q+1}]$ . 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

$$e_{s_p} = e_{t_p} = e_{s_p} \circ \psi_L(\text{ext}_{x_{s_{p+1}} \dots a_{t_p}, b_{t_p} \dots y_{s_{p+1}}}).$$

Hence, we have  $w \models \pi_{e_{s_p}}(i_{s_p} + 1, i_{t_p}, j_{t_p}, j_{s_p} - 1)$  where  $\pi_e$  is the formula given by Lemma 5.13 (recall that  $\mathcal{R}_{e_{s_p}}$  is length-synchronous). We can therefore use the formula  $\pi_{e_{s_p}}$  to witness that  $w_{i_{s_{p+1}}} \dots w_{i_{t_p}} w_{j_{t_p}} \dots w_{j_{s_{p-1}}}$  is indeed a well-matched word. It will thus remain to verify that  $w_{i_{t_p}} \dots w_{i_{s_{p+1}}} w_{j_{s_{p+1}}} \dots w_{i_{t_p}}$  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  $\mu_d^\uparrow$ :

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

$\square$

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  $\pi_e$ , where  $e$  is  $\varphi_L(L)$ -reachable from above, by the predicate  $\pi_e^{\text{exact}}$  expressing that for all  $w \in \Sigma^+$  and  $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$  it holds:

$$w \models \pi_e^{\text{exact}}(i, i', j', j) \iff w_i \dots w_{i'} w_{j'} \dots w_j \in \Sigma^\Delta, e \circ \psi_L(\text{ext}_{w_i \dots w_{i'}, w_{j'} \dots w_j}) = e, \text{ and } \Delta(w_i \dots w_{i'}) > 0$$

It remains to prove Lemma 5.13.

### Proof of Lemma 5.13

In essence, our proof is inspired by the approach taken in [25, Proof of Proposition 126], which is itself a flawed adaptation (we refer to Section 8 for more details) of the approach taken in [24, 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 4.8 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 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 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  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\Delta$ . 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  $i \leq i' < j' \leq j$ ,  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\Delta$ ,  $\Delta(w_i \cdots w_{i'}) > 0$  and  $e \circ \psi_L(\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$ . In the terminology of Proposition 4.8, for  $F = \varphi_L(L)$ , we have  $(w_i \cdots w_{i'}, w_{j'} \cdots w_j) \in \mathcal{R}_e$ . We claim that Points (i) to (iv) are actually satisfied. Indeed, recalling that  $L$  is length-synchronous by assumption, 2(a) of Proposition 4.8 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 4.8. We have that  $\Delta(w_i \cdots w_{i'} w_{j'} \cdots w_j) = 0$  immediately follows from our assumption  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\Delta$ , thus Point (ii) holds. Another consequence of our assumption  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\Delta$  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  $\gamma_e$  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  $\gamma_e$  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 4.8, 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

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

where the first line checks condition (i), the 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 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

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

such that for all  $w \in \Sigma^+$  and  $i \in [1, |w|]$  such that  $i \leq |w| - k + 1$ , we have  $w \models H_k^h(i)$  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

$$D_{n,d}^h(x, y, z) = \bigvee_{\substack{-n \leq h_1 \leq n \\ -d \leq h_2 \leq d \\ h_1 + h_2 = h}} \left( H_n^{h_1}(x + (z - 1) \cdot n) \wedge H_d^{h_2}(y - z \cdot d + 1) \right),$$

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  $w \models D_{n,d}^h(i, j, k)$  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  $\Gamma_p = \{a_{-p}, \dots, a_{-1}, a_0, a_1, \dots, a_p\}$  and define  $\Delta_p: \Gamma_p^* \rightarrow \mathbb{Z}$  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 [30, Theorem VI.1.1]), that there exists an  $\text{FO}_{\Gamma_{n+d}[\langle \rangle]}$ -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  $u_1, \dots, u_l \in \Sigma^n$  and  $w_{j'} \cdots w_j$  as  $v_1 \cdots v_l$  with  $v_1, \dots, v_l \in \Sigma^d$ . 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  $-q \leq \Delta(u_1 v_1) + \cdots + \Delta(u_l v_l) \leq q$  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  $\Gamma_{n+d}^*$  belongs to  $L_{n+d,q}$ .

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

- replacing any quantification  $\exists z \rho(z)$  by  $\exists z (z \leq (y - y' + 1)/d \wedge \rho(z))$ ;
- replacing any quantification  $\forall z \rho(z)$  by  $\forall z (z \leq (y - y' + 1)/d \rightarrow \rho(z))$ ;

- replacing any atomic formula of the form  $a_h(z)$  for  $a_h \in \Gamma_{n+d}$  by  $D_{n,d}^h(x, y, z)$ .

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  $Q_q = \{-q, \dots, -1, 0, 1, \dots, q, \perp\}$  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  $f_{a_h}: Q_q \rightarrow Q_q$  to be such that

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

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  $\varphi_{p,q}: \Gamma_p^* \rightarrow M_{p,q}$  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

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

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

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

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

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

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[+] via approximate matching and quasi-aperiodic evaluation languages (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 all quasi-aperiodic, is definable by FO $_{\Sigma, \rightsquigarrow}$ [+] sentences 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  $r \in R_L$ , and
- $\mathcal{E}_{\psi_L, e}$  is a regular language whose syntactic morphism is quasi-aperiodic for all  $e \in O_L$ .

*Then there exists an FO $_{\Sigma, \rightsquigarrow}$ [+]-sentence  $\eta$  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  $r \in R_L$ , the language  $\mathcal{E}_{\varphi_L, r}$  is regular and its syntactic morphism is quasi-aperiodic. This implies, by [30, Theorem VI.4.1], that for each  $r \in R_L$ , there exists an FO $_{\Gamma_{\varphi_L}}[<, \text{MOD}]$ -sentence  $\nu_{\varphi_L, r}$  defining  $\mathcal{E}_{\varphi_L, r}$ .

Finally, by hypothesis, for each  $e \in O_L$ , the language  $\mathcal{E}_{\psi_L, e}$  is regular and its syntactic morphism is quasi-aperiodic. Again, by [30, Theorem VI.4.1], for each  $e \in O_L$ , there exists an FO $_{\Gamma_{\psi_L}}[<, \text{MOD}]$ -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  $M^\Delta(w)$ , represented by the relational symbol  $\rightsquigarrow$ . The first ones are  $N_d$  which express that the infix  $w_i \dots w_j \in \Sigma^\Delta$  of  $w \in \Sigma^\Delta$  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  $w \in \Sigma^\Delta$  and  $i, j \in [1, |w|]$  satisfying  $w_i \cdots w_j \in \Sigma^\Delta$ , we have that  $(w, M^\Delta(w)) \models N_d(i, j)$  if, and only if,  $\text{nd}(w_i \cdots w_j) \geq d$ . The case  $d = 0$  is trivial since we can set

$$N_0(i, j) = \top.$$

Note that if  $\text{nd}(w) = d \geq 1$ , then  $w$  can be factorized as  $w = w_1 u w_2$  such that  $u \in \Sigma_{\text{call}} \Sigma^\Delta \Sigma_{\text{ret}}$  and  $\text{nd}(w) = \text{nd}(u) \geq d$ . This means that  $u = a_1 v b_1 \in$  for some  $a_1 \in \Sigma_{\text{call}}$ ,  $b_1 \in \Sigma_{\text{ret}}$  and  $v \in \Sigma^\Delta$ . We then apply Lemma 4.13 and Lemma 4.14 implying that  $u$  has a nesting-maximal stair factorization

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

for which there exists  $h \in [1, k]$  such that, setting  $u_i = \text{ext}_{a_i, b_i} \circ \text{ext}_{x_{i+1}, y_{i+1}} \circ \dots \circ \text{ext}_{a_k, b_k}(u')$  for all  $i \in [1, k]$  and  $u_{k+1} = u'$ , we have  $\text{nd}(u) = \text{nd}(u_h) \geq d$  and  $\text{nd}(u_{h+1}) \geq d - 1$ . Thus,  $u_h = a_h z_1 z_2 b_h$  for some  $z_1, z_2 \in \Sigma^\Delta$  satisfying  $\text{nd}(z_1), \text{nd}(z_2) \geq d - 1$ .

Hence for  $d \geq 1$  we set

$$N_d(x, y) = \exists x' \exists y' \exists x'' \exists y'' \exists s (x \leq x' < x'' < y'' < y' \leq y \wedge x' \rightsquigarrow y' \wedge x'' \rightsquigarrow y'' \wedge N_{d-1}(x'' + 1, s) \wedge N_{d-1}(s, y'' - 1)) .$$

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

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

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

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

**Main construction.** To build the  $\text{FO}_{\Sigma, \rightsquigarrow}[+]$ -sentence  $\eta$ , we build  $\text{FO}_{\Sigma, \rightsquigarrow}[+]$ -formulas

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

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

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

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

$$\eta = \forall z \exists t ((\Sigma_{\text{call}}(z) \rightarrow z \rightsquigarrow t) \wedge (\Sigma_{\text{ret}}(z) \rightarrow t \rightsquigarrow z)) \wedge \left( E \vee \exists x \exists y (\neg \exists x' (x' < x) \wedge \neg \exists y' (y < y')) \wedge \bigvee_{r \in \varphi_L(L)} \eta_{d,r}(x, y) \right) .$$

It now remains to build  $\eta_{d,r}^\uparrow(x, y)$  and  $\eta_{d,r}(x, y)$  for all  $d \in \mathbb{N}$  and  $r \in R_L$ . The construction is by induction on  $d$ . Let  $r \in R_L$ . We define  $\eta_{0,r}^\uparrow(x, y) = \perp$ . 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  $\tau_0$  is inductively defined as follows:

- $\tau_0(z < z') = z < z'$
- $\tau_0(s(z)) = \bigvee_{c \in \varphi_L^{-1}(s) \cap \Sigma_{\text{int}}} c(z)$  for all  $s \in \varphi_L(\Sigma^\Delta \setminus \{\varepsilon\})$

- $\tau_0(\text{MOD}_m(z)) = \exists t(z - x + 1 = t \cdot m)$  for all  $m \in \mathbb{N}_{>0}$
- $\tau_0(\$ (z)) = \perp$
- $\tau_0(\rho_1(\mathbf{z}_1) \wedge \rho_2(\mathbf{z}_2)) = \tau_0(\rho_1(\mathbf{z}_1)) \wedge \tau_0(\rho_2(\mathbf{z}_2))$
- $\tau_0(\neg \rho(z)) = \neg \tau_0(\rho(z))$
- $\tau_0(\exists z \rho(z, \mathbf{z})) = \exists z(x \leq z \leq y \wedge \tau_0(\rho(z, \mathbf{z})))$

Now let  $d > 0$ . Let us first define  $\eta_{d,r}$  when assuming that we have already define  $\eta_{d,r}^\uparrow$ . Note that in case  $\text{nd}(u) \leq d$ , then one can factorize  $u$  as  $u = u_1 \dots u_m$  such that  $u_i \in \Sigma_{\text{int}}^+ \cup \Sigma_{\text{call}} \Sigma_{\text{ret}}^\Delta$  and  $\text{nd}(u_i) \leq d$  for all  $i \in [1, m]$ . Using this observation we define

$$\eta_{d,r}(x, y) = \neg N_{d+1}(x, y) \wedge \tau_1(\nu_{\varphi_L, r}),$$

where the translation  $\tau_1$  agrees with the above translation  $\tau_0$  (where, as expected, occurrences of  $\tau_0$  are replaced by  $\tau_1$ ) except for the following kinds of subformulas:

- $\tau_1(\$ (z)) = A(x, y, z)$
- $\tau_1(s(z)) = \neg A(x, y, z) \wedge \left( \bigvee_{c \in \varphi_L^{-1}(s) \cap \Sigma_{\text{int}}} c(z) \vee \exists t(x \leq t \leq y \wedge t \leftrightarrow z \wedge \eta_{d,s}^\uparrow(t, z)) \right)$

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

First observe that any infix  $u = w_i \dots w_j$  of  $w$  where  $i$  is matched to  $j$  in  $w$  is of the form  $u = a_1 v b_1 \in \Sigma^\Delta$  for some  $a_1 \in \Sigma_{\text{call}}$ ,  $b_1 \in \Sigma_{\text{ret}}$ , and  $v \in \Sigma^\Delta$ . As above, we can directly express  $\text{nd}(u) \leq d$  via the formula  $\neg N_{d+1}$ . Towards expressing that  $\varphi_L(w_i \dots w_j) = \varphi_L(u) = r$ , we make use of Lemma 4.14: for the infix  $u$  there is a nesting-maximal stair factorization  $u = \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \circ \dots \circ \text{ext}_{x_k, y_k} \circ \text{ext}_{a_k, b_k}(u')$  and some  $h \in [1, k]$  such that, setting  $u_\ell = \text{ext}_{a_\ell, b_\ell} \circ \text{ext}_{x_{\ell+1}, y_{\ell+1}} \circ \dots \circ \text{ext}_{a_k, b_k}(u')$  for all  $\ell \in [1, k]$  and  $u_{k+1} = u'$ , we have

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

We first construct for all  $\varphi_L(L)$ -reachable  $e \in O_L$  a formula  $\chi_{d,e}(i, i', j', j)$  that verifies whether for

$$\text{ext}_{w_i \dots w_{i'}, w_{j'} \dots w_j} = \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \dots \text{ext}_{x_{h-1}, y_{h-1}} \circ \text{ext}_{a_h, b_h}$$

we have  $\psi_L(\text{ext}_{w_i \dots w_{i'}, w_{j'} \dots w_j}) = e$ . Exploiting Point 3 from above we can inductively make use of the formulas  $\{\eta_{d-1, r'} \mid r' \in R_L\}$  in order to evaluate  $x_1, y_1, \dots, y_h, y_h$ . The formula  $\chi_{d,e}$  will verify if  $\sigma_i \dots \sigma_{i'} \in \mathcal{E}_{\psi_L, e}$ , where  $\sigma_m = \text{left}_{\varphi_L(x_m)} \circ \text{right}_{\varphi_L(y_m)} \circ \psi_L(\text{ext}_{a_m, b_m})$  for all  $m \in [i, i']$ .

Hence we set  $\chi_{d,e}(x, x', y', y) = \tau_2(\nu_{\psi_L, e})$ , where the translation  $\tau_2$  agrees with translation  $\tau_0$  (where, as expected, occurrences of  $\tau_0$  are replaced by  $\tau_2$ ) with the following exceptions:

- $\tau_2(\exists z \rho(z, \mathbf{z})) = \exists z(x \leq z \leq x' \wedge \tau_2(\rho(z, \mathbf{z})))$
- $\tau_2(\$ (z)) = \neg U(x, x', z)$

- $\tau_2(f(z)) = \iota_{d,f}(x, x', y', y, z) \vee \zeta_{d,f}(x, x', y', y, z)$  for all  $f \in \psi_L(\mathcal{O}(\Sigma^\Delta)_\uparrow)$ , where (using that  $x_1 = y_1 = \varepsilon$ )

$$\iota_{d,f}(x, x', y', y, z) = \left( x = z \wedge \bigvee_{\substack{a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}} \\ \psi_L(\text{ext}_{a,b})=f}} (a(x) \wedge b(y)) \right)$$

and

$$\begin{aligned} \zeta_{d,f}(x, x', y', y, z) = & \exists t \exists z' \exists t' \left( U(x, x', z) \wedge z \rightsquigarrow t \wedge \right. \\ & U(x, x', z') \wedge z' < z \wedge \neg \exists z'' (z' < z'' < z \wedge U(x, x', z'')) \wedge z' \rightsquigarrow t' \wedge \\ & \bigvee_{\substack{a \in \Sigma_{\text{call}}, b \in \Sigma_{\text{ret}}, r', r'' \in R_L \\ f = \text{left}_{r'} \circ \text{right}_{r''} \circ \psi_L(\text{ext}_{a,b})}} \left. \left( a(z) \wedge b(t) \wedge \right. \right. \\ & \left. \left. \eta_{d-1, r'}(z' + 1, z - 1) \wedge \eta_{d-1, r''}(t + 1, t' - 1) \right) \right) \end{aligned}$$

Making use of Point 2 above, we can evaluate remaining part  $\varphi_L(\text{ext}_{x_{h+1}, y_{h+1}}(u_{h+1}))$  via the formulas  $\{\eta_{d-1, r'} \mid r' \in R_L\}$ . We are now ready to give the formula  $\eta_{d,r}^\uparrow$ . We set

$$\begin{aligned} \eta_{d,r}^\uparrow(x, y) = & \neg N_{d+1}(x, y) \wedge x \rightsquigarrow y \wedge \exists x' \exists y' \left( x \leq x' < y' \leq y \wedge x' \rightsquigarrow y' \wedge \right. \\ & \bigwedge_{d'=0}^d (N_{d'}(x, y) \leftrightarrow N_{d'}(x', y')) \wedge \\ & \left. \bigvee_{\substack{r' \in R_L, e \in O_L \\ e(r')=r}} (\chi_{d,e}(x, x', y', y) \wedge \eta_{d-1, r'}(x' + 1, y' - 1)) \right). \end{aligned}$$

□

## 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 languages. The computability of  $k, l \in \mathbb{N}$  with  $k \neq l$  such that  $\mathcal{L}_{k,l} \leq_{\text{cd}} L$  are subject of Section 5.4.2.

**Theorem 5.17.** *If a VPL  $L$  that is weakly length-synchronous but not length-synchronous, and whose syntactic morphism  $(\varphi_L, \psi_L)$  is quasi-aperiodic, one can effectively compute regular synchronization languages  $X_1, \dots, X_m$  witnessing that  $X_1^\boxtimes, \dots, X_m^\boxtimes$  are intermediate languages and moreover  $L =_{\text{cd}} \bigsqcup_{i=1}^m X_i^\boxtimes$ .*

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

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

$$M_e = \{u\#v \mid uv \in \Sigma^\Delta, \Delta(u) > 0, e \circ \psi_L(\text{ext}_{u,v}) = e\} \text{ and } \mathcal{R}_e = \{(u, v) \in \Sigma^* \times \Sigma^* \mid u\#v \in M_e\}.$$

Note that since  $L$  is assumed to be weakly length-synchronous, by Proposition 4.3,  $\mathcal{R}_e$  is weakly length-synchronous for all  $\varphi_L(L)$ -reachable  $e \in O_L$ .

Also note that since  $M_e = \bigcup_{f \in O_L: e \circ f = e} L_f \cap \{u\#v \mid uv \in \Sigma^\Delta, \Delta(u) > 0\}$ , since all languages  $L_f$  are effectively computable VPLs by Lemma 6.4, and since the language  $\{u\#v \mid uv \in \Sigma^\Delta, \Delta(u) > 0\}$  is obviously an effectively computable VPL, we obtain that  $M_e \subseteq \Sigma^* \# \Sigma^*$  is an effectively computable VPL. The set  $S_e = \{(k, l) \mid \exists (u, v) \in \Sigma^k \times \Sigma^l : u\#v \in M_e\} \subseteq \mathbb{N}_{>0}^2$  is hence effectively semilinear by Lemma 6.5. Note that the word relation  $\mathcal{R}_e$  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 when  $S_e$  is non-empty, and that condition is trivially true when  $S_e$  is empty. As a consequence one can effectively compute the set

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

Observe that since  $L$  is not length-synchronous by assumption, we have  $\mathcal{Z} \neq \emptyset$  (Proposition 4.7).

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

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

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

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

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

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

Finally, for all  $e \in \mathcal{Z}$  we define the language

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

**Remark 5.18.** Let be  $n \in \mathbb{N}$  be the constant from Lemma 4.4 for  $L$  and let  $e \in \mathcal{Z}$ . When setting  $F = \varphi_L(L)$  and  $(\varphi, \psi) = (\varphi_L, \psi_L)$ , Lemma 4.4 states that the factorization

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

of every word  $u\#v \in M_e$  satisfies  $|x_1|, \dots, |x_k|, |y_1|, \dots, |y_k| \leq n$ . As a consequence, for the corresponding factorization

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

of every word  $(u\#v)^\ddagger \in N_e \setminus \{\varepsilon\}$  we have  $|\widetilde{x_1}|, \dots, |\widetilde{x_k}|, |\overline{y_1}|, \dots, |\overline{y_k}| \leq n$ .

#### 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 language for all  $e \in \mathcal{Z}$ . Moreover, one can effectively compute a regular synchronization language  $X_e$  witnessing that  $N_e = X_e^{\boxtimes}$  is indeed intermediate (Lemma 5.19).
2.  $N_e \leq_{\text{cd}} M_e$  for all  $e \in \mathcal{Z}$  (Lemma 5.20).
3.  $M_e \leq_{\text{cd}} N_e$  for all  $e \in \mathcal{Z}$  (Lemma 5.21).
4.  $M_e \leq_{\text{cd}} \biguplus_{f \in O_L \text{ is } \varphi_L(L)\text{-reachable}} L_f$  for all  $e \in \mathcal{Z}$  (Lemma 5.22).
5.  $L \leq_{\text{cd}} \biguplus_{e \in \mathcal{Z}} M_e$  (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 language, for which moreover one can effectively compute a regular synchronization language  $X_e$  witnessing that  $N_e = X_e^{\boxtimes}$  is indeed intermediate. Recalling that  $\mathcal{Z} \neq \emptyset$ , it remains to argue that  $L =_{\text{cd}} \biguplus\{N_e \mid e \in \mathcal{Z}\}$ . We remark that if a language is constant-depth reducible to at least of the languages  $L_1, \dots, L_n$ , then so is the language constant-depth reducible to the disjoint union  $\biguplus_{i=1}^n L_i$ . Conversely, if all languages  $L_1, \dots, L_n$  are constant-depth reducible to a language, then so is  $\biguplus_{i=1}^n L_i$  constant-depth reducible to the language.

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

**Lemma 5.19.**  $N_e$  is an intermediate language for all  $e \in \mathcal{Z}$ . Moreover, one can effectively compute a regular synchronization language  $X_e$  witnessing that  $N_e = X_e^{\boxtimes}$  is indeed intermediate.

*Proof.* Let  $e \in \mathcal{Z}$ . Recall that  $\tilde{\Sigma} \cup \bar{\Sigma} \subseteq \Upsilon_{\text{int}}$ . Our synchronization language  $X_e$  will be defined over the alphabet

$$\Upsilon_{\otimes_2} = \{(a, b) \mid a \in \Upsilon_{\text{call}}, b \in \Upsilon_{\text{ret}}\} \cup \{(\varepsilon, c), (c, \varepsilon) \mid c \in \Upsilon_{\text{int}}\}.$$

For every  $w = w_1 \cdots w_k \in \Sigma^\Delta \cap \Sigma^{\leq n}$  with  $w_i \in \Sigma$  for all  $i \in [1, k]$  let

$$\langle \tilde{w}, \varepsilon \rangle = \begin{cases} \varepsilon & \text{if } w = \varepsilon \\ (\tilde{w}_1, \varepsilon) \dots (\tilde{w}_k, \varepsilon) & \text{if } w \neq \varepsilon \end{cases} \quad \text{and} \quad \langle \varepsilon, \bar{w} \rangle = \begin{cases} \varepsilon & \text{if } w = \varepsilon \\ (\varepsilon, \bar{w}_1) \dots (\varepsilon, \bar{w}_k) & \text{if } w \neq \varepsilon. \end{cases}$$

Let  $S = (\{\langle \tilde{w}, \varepsilon \rangle \mid w \in \Sigma^\Delta \cap \Sigma^{\leq n}\} \cup \{\langle \varepsilon, \bar{w} \rangle \mid w \in \Sigma^\Delta \cap \Sigma^{\leq n}\})^*$  and  $H = \Sigma_{\text{call}} \times \Sigma_{\text{ret}}$ . Observe that for all  $w \in (SH)^*S$ , setting  $\mathcal{R}(w) = (u, v)$  we have that  $u \in (\tilde{\Sigma} \cup \Sigma)^*$ ,  $v \in (\bar{\Sigma} \cup \Sigma)^*$  and  $\tilde{\vartheta}^{-1}(u)\bar{\vartheta}^{-1}(v) \in \Sigma^\Delta$ . We define

$$X_e = \{w \in (SH)^+S \cup \{\varepsilon\} \mid \mathcal{R}(w) = (u, v) \wedge e \circ \psi_L(\text{ext}_{\tilde{\vartheta}^{-1}(u), \bar{\vartheta}^{-1}(v)}) = e\}.$$

The equality  $X_e^{\boxtimes} = N_e$  is immediate from definition. Moreover is  $X_e$  star-closed by definition. It is also clear that  $X_e$  is regular: one can construct a finite automaton whose states involve the monoid  $O_L$ , whose initial and only final state is  $e$  and which simulates multiplication in  $O_L$  by guessing factors of the form  $\langle \tilde{w}, \varepsilon \rangle$  or  $\langle \varepsilon, \bar{w} \rangle$  for  $w \in \Sigma^\Delta \cap \Sigma^{\leq n}$ , or reading letters  $(a, b) \in H$ . By construction we have

$$\mathcal{R}_e = \{(\tilde{\vartheta}^{-1}(u), \bar{\vartheta}^{-1}(v)) \mid (u, v) \in \mathcal{R}(X_e) \setminus \{(\varepsilon, \varepsilon)\}\} .$$

We claim that  $\mathcal{R}(X_e)$  is not length-synchronous: indeed, if  $\mathcal{R}(X_e)$  were length-synchronous, then by the previous equivalence  $\mathcal{R}_e$  would also be length-synchronous since  $\tilde{\vartheta}^{-1}$  and  $\bar{\vartheta}^{-1}$  preserve lengths and  $\Delta(\tilde{\vartheta}^{-1}(u)) = \Delta(u) > 0$  for all  $(u, v) \in \mathcal{R}(X_e) \setminus \{(\varepsilon, \varepsilon)\}$ . We claim that  $\mathcal{R}(X_e)$  is weakly length-synchronous. Assume by contradiction that  $\mathcal{R}(X_e)$  were not weakly length-synchronous. Then, without loss of generality, there would exist  $(u, v), (u', v') \in \mathcal{R}(X_e)$  such that  $|u| \neq |u'|$  and  $\Delta(u), \Delta(u') > 0$ . Then  $(\tilde{\vartheta}^{-1}(u), \bar{\vartheta}^{-1}(v)), (\tilde{\vartheta}^{-1}(u'), \bar{\vartheta}^{-1}(v')) \in \mathcal{R}_e$ . Again, since  $\tilde{\vartheta}^{-1}$  and  $\bar{\vartheta}^{-1}$  preserve lengths and  $\Delta(\tilde{\vartheta}^{-1}(u)) = \Delta(u) > 0$  as well as  $\Delta(\tilde{\vartheta}^{-1}(u')) = \Delta(u') > 0$ , it follows that  $\mathcal{R}_e$  is not weakly length-synchronous as well, a contradiction.

Finally, let  $\varphi: \Upsilon_{\otimes_2}^* \rightarrow M$  denote the syntactic morphism of  $X_e$ . Assume by contradiction that there exist  $k, l \in \mathbb{N}$  such that  $\varphi(\{w \in \Upsilon_{\otimes_2}^* \mid \mathcal{R}(w) \in \Upsilon^k \times \Upsilon^l\})$  contains a non-trivial group  $G$ : we will derive a contradiction by proving that  $(\varphi_L, \psi_L)$  is not quasi-aperiodic.

We make a first observation. Let  $w \in \varphi^{-1}(G)$  and assume that for all  $x, y \in \Upsilon_{\otimes_2}^*$  we had  $xwy \notin X_e$ : then, for all  $w' \in \varphi^{-1}(G)$  and  $x, y \in \Upsilon_{\otimes_2}^*$ , we would have  $xw'wy \notin X_e$  and  $xww'y \notin X_e$ . This would imply that  $\varphi(w')\varphi(w) = \varphi(w)\varphi(w') = \varphi(w)$  for all  $w' \in \varphi^{-1}(G)$ , that is, as  $G \subseteq \Upsilon_{\otimes_2}^*$ , we would have that  $\varphi(w)$  is a zero in  $G$ , a contradiction to the folklore fact that non-trivial groups have no zero. Hence, for all  $w \in \varphi^{-1}(G)$  there exist  $x, y \in \Upsilon_{\otimes_2}^*$  satisfying  $xwy \in X_e$ .

Let  $g_0$  be the identity of  $G$  and  $g_1$  any other element. By assumption, there exist  $w_0, w_1 \in \Upsilon_{\otimes_2}^*$  such that  $\varphi(w_0) = g_0$  and  $\varphi(w_1) = g_1$ . We claim that  $w_0$  belongs to  $\Upsilon_{\otimes_2}^* H \Upsilon_{\otimes_2}^*$ . Assume it were not the case. Then, given any  $x, y \in \Upsilon_{\otimes_2}^*$ , since  $|w_0| \geq 1$  (because  $\mathcal{R}(w_0)$  and  $\mathcal{R}(w_1)$  both belong to  $\Upsilon^k \times \Upsilon^l$  while not being equal), then either  $xw_0^{2n+1}y \notin (SH)^+S$  or  $xw_0^{2n+1}y \in (SH)^+S$  and then setting  $\mathcal{R}(xw_0^{2n+1}y) = (u, v)$ , at least one of  $u$  or  $v$  must contain  $z \in \Upsilon^\Delta$  of length at least  $n+1$ , showing  $uv \notin N_e$  by Remark 5.18: in any case, we would have  $xw_0^{2n+1}y \notin X_e$  for any  $x, y \in \Upsilon_{\otimes_2}^*$ , a contradiction to what we have proven in the last paragraph.

Now let  $r$  be the shortest prefix of  $w_0$  such that  $r \in \Upsilon_{\otimes_2}^* H$  (which exists because  $w_0 \in \Upsilon_{\otimes_2}^* H \Upsilon_{\otimes_2}^*$ ) and  $t$  be the longest suffix of  $w_0$  such that  $t \in (\Upsilon_{\otimes_2} \setminus H)^*$ ; we then write  $w_0 = rst$  with  $s \in (\Upsilon_{\otimes_2}^* H)^*$ . Because there exists  $x, y \in \Upsilon_{\otimes_2}^*$  satisfying  $xw_0w_0y \in X_e \subseteq (SH)^+S$  and  $w_0w_0 = rstrst$ , we must have that  $tr \in SH$  and  $s \in (SH)^*$ . As there exist  $x, y \in \Upsilon_{\otimes_2}^*$  such that  $xrs(tw_0rs)ty = xw_0w_0y \in X_e \subseteq (SH)^+S$ , since  $xrs$  and  $tw_0rs$  both end with a letter in  $H$ , it follows that  $tw_0rs \in (SH)^+$ . Similarly, we have  $tw_1rs \in (SH)^+$ . Set  $\mathcal{R}(tw_0rs) = (u_0, v_0)$  and  $\mathcal{R}(tw_1rs) = (u_1, v_1)$ . Since  $\mathcal{R}(w_0), \mathcal{R}(w_1) \in \Upsilon^k \times \Upsilon^l$ , we must have  $|u_0| = |u_1|$  and  $|v_0| = |v_1|$  and we will denote those lengths by  $k'$  and  $l'$ , respectively. Let  $u'_0 = \tilde{\vartheta}^{-1}(u_0)$ ,  $v'_0 = \bar{\vartheta}^{-1}(v_0)$ ,  $u'_1 = \tilde{\vartheta}^{-1}(u_1)$  and  $v'_1 = \bar{\vartheta}^{-1}(v_1)$ : because  $tw_0rs, tw_1rs \in (SH)^+$ , it holds that  $u'_0v'_0, u'_1v'_1 \in \Sigma^\Delta$ .

Since  $\mathfrak{B}(O_L)$  forms a monoid there exists  $p \in \mathbb{N}_{>0}$  such that

$$\psi_L(\{\text{ext}_{u'_0, v'_0}, \text{ext}_{u'_1, v'_1}\})^p = \psi_L(\{\text{ext}_{u'_0, v'_0}, \text{ext}_{u'_1, v'_1}\})^{2p} .$$

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

$$\psi_L(\text{ext}_{u_0^{p-1}u'_1, v_0^{p-1}v'_1})^i \in \psi_L(\{\text{ext}_{u'_0, v'_0}, \text{ext}_{u'_1, v'_1}\})^{ip} = \psi_L(\{\text{ext}_{u'_0, v'_0}, \text{ext}_{u'_1, v'_1}\})^{2p} .$$

Hence, as  $|u'_0| = |u'_1| = k'$  and  $|v'_0| = |v'_1| = l'$ , the semigroup  $\{\psi_L(\text{ext}_{u_0^{p-1}u'_1, v_1 v_0^{p-1}})^i \mid i \in \mathbb{N}_{>0}\}$  is contained in  $\psi_L(\mathcal{O}(\Sigma^\Delta)^{k', l'})$ .

Let  $i \in \mathbb{N}_{>0}$ . We have that

$$\varphi((w_0^{2p-1}w_1)^i w_0) = \varphi(w_1)^i \neq \varphi(w_1)^{i+1} = \varphi((w_0^{2p-1}w_1)^{i+1}w_0)$$

because  $g_1^i \neq g_1^{i+1}$ , otherwise we would have  $g_0 = g_1$  by the existence of an inverse of  $g_1^i$  in  $G$ . This means that there exist  $x, y \in \Upsilon_{\otimes_2}^*$  satisfying either  $x(w_0^{2p-1}w_1)^i w_0 y \in X_e$  and  $x(w_0^{2p-1}w_1)^{i+1} w_0 y \notin X_e$ , or  $x(w_0^{2p-1}w_1)^i w_0 y \notin X_e$  and  $x(w_0^{2p-1}w_1)^{i+1} w_0 y \in X_e$ . Assume the first case holds, the second case is handled symmetrically. Observe that

$$\begin{aligned} x(w_0^{2p-1}w_1)^i w_0 y &= x(rst w_0^{2p-2} w_1)^i rst y \\ &= xrs(tw_0^{2p-2} w_1 rs)^i ty \\ &= xrs(t(w_0 rst)^{p-1} w_1 rs)^i ty \\ &= xrs((tw_0 rs)^{p-1} (tw_1 rs))^i ty \end{aligned}$$

and, analogously,

$$x(w_0^{2p-1}w_1)^{i+1} w_0 y = xrs((tw_0 rs)^{p-1} (tw_1 rs))^{i+1} ty .$$

So as in particular  $xrs((tw_0 rs)^{p-1} (tw_1 rs))^i ty \in (SH)^+ S$ , since  $xrs$  ends with a letter in  $H$ , it follows that  $xrs \in (SH)^+$  and since  $tw_1 rs \in (SH)^+$ , it follows that  $ty \in (SH)^* S$ . Therefore, because  $tw_0 rs, tw_1 rs \in (SH)^+$ , we also have  $xrs((tw_0 rs)^{p-1} (tw_1 rs))^{i+1} ty \in (SH)^+ S$ . Take  $u_x, v_x, u_y, v_y \in \Upsilon^*$  such that  $\mathcal{R}(xrs) = (u_x, v_x)$  and  $\mathcal{R}(ty) = (u_y, v_y)$ . Since

$$\mathcal{R}\left(xrs((tw_0 rs)^{p-1} (tw_1 rs))^i ty\right) = (u_x(u_0^{p-1}u_1)^i u_y, v_y(v_1 v_0^{p-1})^i v_x)$$

and

$$\mathcal{R}\left(xrs((tw_0 rs)^{p-1} (tw_1 rs))^{i+1} ty\right) = (u_x(u_0^{p-1}u_1)^{i+1} u_y, v_y(v_1 v_0^{p-1})^{i+1} v_x) ,$$

we must have

$$\begin{aligned} e &= e \circ \psi_L \left( \text{ext}_{\tilde{\vartheta}^{-1}(u_x(u_0^{p-1}u_1)^i u_y), \bar{\vartheta}^{-1}(v_y(v_1 v_0^{p-1})^i v_x)} \right) \\ &= e \circ \psi_L \left( \text{ext}_{\tilde{\vartheta}^{-1}(u_x), \bar{\vartheta}^{-1}(v_x)} \right) \circ \psi_L \left( \text{ext}_{u_0^{p-1}u'_1, v_1 v_0^{p-1}} \right)^i \circ \psi_L \left( \text{ext}_{\tilde{\vartheta}^{-1}(u_y), \bar{\vartheta}^{-1}(v_y)} \right) \end{aligned}$$

and

$$\begin{aligned} e &\neq e \circ \psi_L \left( \text{ext}_{\tilde{\vartheta}^{-1}(u_x(u_0^{p-1}u_1)^{i+1} u_y), \bar{\vartheta}^{-1}(v_y(v_1 v_0^{p-1})^{i+1} v_x)} \right) \\ &= e \circ \psi_L \left( \text{ext}_{\tilde{\vartheta}^{-1}(u_x), \bar{\vartheta}^{-1}(v_x)} \right) \circ \psi_L \left( \text{ext}_{u_0^{p-1}u'_1, v_1 v_0^{p-1}} \right)^{i+1} \circ \psi_L \left( \text{ext}_{\tilde{\vartheta}^{-1}(u_y), \bar{\vartheta}^{-1}(v_y)} \right) . \end{aligned}$$

Thus, it must be that  $\psi_L(\text{ext}_{u_0^{p-1}u'_1, v_1 v_0^{p-1}})^i \neq \psi_L(\text{ext}_{u_0^{p-1}u'_1, v_1 v_0^{p-1}})^{i+1}$ .

As this is true for each  $i \in \mathbb{N}_{>0}$ , the semigroup  $\{\psi_L(\text{ext}_{u_0^{p-1}u'_1, v_1 v_0^{p-1}})^i \mid i \in \mathbb{N}_{>0}\}$  that is contained in  $\psi_L(\mathcal{O}(\Sigma^\Delta)^{k', l'})$  is not aperiodic, contradicting the quasi-aperiodicity of  $(\varphi_L, \psi_L)$ . Therefore, there does not exist  $k, l \in \mathbb{N}$  such that  $\varphi(\{w \in \Upsilon_{\otimes_2}^* \mid \mathcal{R}(w) \in \Upsilon^k \times \Upsilon^l\})$  contains a non-trivial group  $G$ .



As a consequence we obtain that  $N_e = X_e^{\boxtimes}$  is an intermediate language and that the synchronization language  $X_e$  witnesses this.  $\square$

**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  $u \in (\tilde{\Sigma} \cup \Sigma_{\text{call}})^*$  and some  $v \in (\bar{\Sigma} \cup \Sigma_{\text{ret}})^*$ , 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  $k \geq 1$ ,  $x_1, \dots, x_{k+1} \in \{x \in \tilde{\Sigma}^* \mid |x| \leq n \wedge \tilde{\vartheta}^{-1}(x) \in \Sigma^\Delta\}$  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  $l \geq 1$ ,  $y_1, \dots, y_{l+1} \in \{y \in \bar{\Sigma}^* \mid |y| \leq n \wedge \bar{\vartheta}^{-1}(y) \in \Sigma^\Delta\}$  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

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

is in  $M_e$ .  $\square$

**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  $\Sigma^\Delta$ . (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  $\Sigma^\Delta$ . Reject if it is not the case.
4. For each position  $i$  within  $u$ , compute  $P_{\text{call}}(i)$  where  $P_{\text{call}}$  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  $\Sigma^\Delta$ .
5. For each position  $j$  within  $v$ , compute  $P_{\text{ret}}(j)$  where  $P_{\text{ret}}$  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  $\Sigma^\Delta$ .

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

$$u' = \tilde{\vartheta}(w_1 \cdots w_{i_1-1})w_{i_1} \cdots \tilde{\vartheta}(w_{i_{k-1}+1} \cdots w_{i_k-1})w_{i_k} \tilde{\vartheta}(w_{i_{k+1}+1} \cdots w_{|u|})$$

and

$$v' = \bar{\vartheta}(w_{|u|+2} \cdots w_{j_{l+1}-1})w_{j_l} \cdots \bar{\vartheta}(w_{j_2+1} \cdots w_{j_1-1})w_{j_1} \bar{\vartheta}(w_{j_1+1} \cdots w_m).$$

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  $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^\Delta)$  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  $u \in \Sigma^\Delta$  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  $(\varphi_L, \psi_L)$  is quasi-aperiodic. There is a constant  $d_L$  such that all words in  $L$  have nesting depth at most  $d_L$  by Proposition 4.11.

By the first point of Proposition 5.10 we may assume that the evaluation language  $\mathcal{E}_{\varphi_L, r}$  is quasi-aperiodic for all  $r \in R_L$ . This implies, by [30, Theorem VI.4.1], that for each  $r \in R_L$ , there exists an  $\text{FO}_{\Gamma_{\varphi_L}}[<, \text{MOD}]$ -sentence  $\nu_{\varphi_L, r}$  defining  $\mathcal{E}_{\varphi_L, r}$ .

As  $L$  is not length-synchronous we cannot assume analogous formulas 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  $\mathcal{E}_{\psi_L, e}$  is not quasi-aperiodic.

For proving  $L \leq_{cd} \biguplus_{e \in \mathcal{Z}} M_e$  we must rather make use of the oracles from  $\biguplus_{e \in \mathcal{Z}} M_e$ . 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  $M_e$ , for all  $e \in \mathcal{Z}$  we may assume that we have a predicate  $\pi_e^{\text{exact}}$  such that for all  $w \in \Sigma^+$  and  $i, i', j', j \in [1, |w|], i \leq i' < j' \leq j$  the following holds:

$$w \models \pi_e^{\text{exact}}(i, i', j', j) \iff w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\Delta, e \circ \psi(\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e \text{ and } \Delta(w_i \cdots w_{i'}) > 0 \quad (6)$$

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 FO[+]-definable (and hence constant-depth computable) predicate  $\pi_e$  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  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\Delta$  and
- if  $w_i \cdots w_{i'} w_{j'} \cdots w_j \in \Sigma^\Delta$ ,  $\Delta(w_i \cdots w_{i'}) > 0$  and  $e \circ \psi_L(\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j}) = e$ , then  $w \models \pi_e(i, i', j', j)$ .

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

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  $w \in \Sigma^\Delta$  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  $\mathcal{E}_{\psi_L, e}$ . Still, we build predicates

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

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

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

It remains to define the predicates  $\bar{\eta}_{d,r}$  and  $\bar{\eta}_{d,r}^\uparrow$  for all  $d \in [0, d_L]$  and all  $r \in R_L$ . For the definition of the  $\bar{\eta}_{0,r}^\uparrow$  and the  $\bar{\eta}_{0,r}$  we can simply reuse  $\eta_{0,r}$  and  $\eta_{0,r}^\uparrow$  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  $d > 0$ .

Towards expressing the predicate  $\bar{\eta}_{d,r}^\uparrow$  we make use of Lemma 4.14: for any infix  $u = w_i \cdots w_j$  of  $w$  of nesting depth  $d$ , where  $i$  is matched to  $j$ , there is a nesting-maximal stair factorization

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

and some  $h \in [1, k]$  such that, setting  $u_\ell = \text{ext}_{a_\ell, b_\ell} \circ \text{ext}_{x_{\ell+1}, y_{\ell+1}} \circ \cdots \circ \text{ext}_{a_k, b_k}(u')$  for all  $\ell \in [1, k]$  and  $u_{k+1} = u'$ , we have

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

For all  $\varphi_L(L)$ -reachable  $e \in O_L$  we build a predicate  $\bar{\chi}_{d,e}(i, i', j', j)$  that verifies whether for

$$\text{ext}_{w_i \cdots w_{i'}, w_{j'} \cdots w_j} = \text{ext}_{x_1, y_1} \circ \text{ext}_{a_1, b_1} \cdots \text{ext}_{x_h, y_h} \circ \text{ext}_{a_h, b_h}$$

we have  $\psi_L(\text{ext}_{w_i \dots w_{i'}, w_{j'} \dots w_j}) = e$ . By Points 2 and 3 from above we can inductively make use of the formulas  $\{\bar{\eta}_{d-1, r'} \mid r' \in R\}$  to evaluate  $x_1, y_1, \dots, y_h, y_h$  and

$$w_{i'+1} \dots w_{j'-1} = \begin{cases} x_{h+1} u_{h+1} y_{h+1} & \text{if } h+1 \leq k \\ u' & \text{otherwise} \end{cases}.$$

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  $h$ . As in Section 5.3.4 we define the product

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

for all  $\ell, \ell' \in [1, h]$ . 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  $e \in O_L$ .

By Claim 5.14 there exist indices  $1 = t_0 \leq s_1 < t_1 < s_2 < t_2 < \dots < s_q < t_q \leq s_{q+1} = h$  such that  $[s_1, t_1], \dots, [s_q, t_q]$  are all repetitive and for  $D_0 = [t_0, s_1], D_1 = [t_1, s_2], \dots, D_q = [t_q, s_{q+1}]$  we have  $q + \sum_{p=0}^q |D_p| \leq 3|O_L|$ . Let  $i = i_1 < \dots < i_h$  and  $j_h < \dots < j_1 = j$  be the positions that correspond to the positions of the letters  $a_1, \dots, a_h \in \Sigma_{\text{call}}$  and  $b_h, \dots, b_1 \in \Sigma_{\text{ret}}$  of  $u$  in  $w$ , respectively: more precisely  $i_\ell = i + |x_1 \dots a_{\ell-1} x_\ell|$  and  $j_\ell = |x_1 a_1 \dots x_k a_k u' b_k y_k \dots b_{\ell+1} y_{\ell+1}| + 1$  for all  $\ell \in [1, h]$ . 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

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

Hence,  $w \models \pi_{e_{s_p}}^{\text{exact}}(i_{s_p} + 1, i_{t_p}, j_{t_p}, j_{s_p} - 1)$  or  $w \models \pi_{e_{s_p}}(i_{s_p} + 1, i_{t_p}, j_{t_p}, j_{s_p} - 1)$ , depending on whether  $e_{s_p} \in \mathcal{Z}$  or not. We can therefore use the predicate  $\pi_{e_{s_p}}^{\text{exact}}$  or  $\pi_{e_{s_p}}$  to witness the above equalities, depending on whether  $e_{s_p} \in \mathcal{Z}$  or not. Next, for all  $m > 0$  and all  $\varphi_L(L)$ -reachable  $f \in O_L$  we will construct a predicate  $\alpha_{m, f}(x, x', y', y)$  such that the following holds:

$$w \models \alpha_{m, f}(i, i', j', j) \iff w_i \dots w_{i'} w_{j'} \dots w_j \in \Sigma^\Delta \wedge \Delta(w_i \dots w_{i'}) = -\Delta(w_{j'} \dots w_j) = m - 1 \\ \psi_L(\text{ext}_{w_i \dots w_{i'}, w_{j'} \dots w_j}) = f$$

For  $\sigma = (\sigma_1, \dots, \sigma_m) \in \Sigma_{\text{call}}^m$ ,  $\xi = (\xi_1, \dots, \xi_m) \in \Sigma_{\text{ret}}^m$ ,  $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_m) \in R^m$ , and  $\mathbf{r}^\dagger = (\mathbf{r}_1^\dagger, \dots, \mathbf{r}_m^\dagger) \in R^m$  we define

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

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

$$\begin{aligned}
\alpha_{m,f}(x, x', y', y) = & \bigvee_{\substack{\sigma \in \Sigma_{\text{call}}^{m-1}, \xi \in \Sigma_{\text{ret}}^{m-1} \\ \mathbf{r}, \mathbf{r}^\dagger \in R^{m-1} \\ f = \Pi(\sigma, \xi, \mathbf{r}, \mathbf{r}^\dagger)}} \exists x_1, \dots, x_m \exists y_1, \dots, y_m \\
& \left( x = x_1 \wedge x' = x_m \wedge y' = y_m \wedge y = y_1 \wedge \right. \\
& x_1 < x_2 < \dots < x_m < y_m < \dots < y_1 \wedge \\
& \bigwedge_{g=2}^m \sigma_{g-1}(x_g) \wedge \xi_{g-1}(y_g) \wedge x_g \leftrightarrow y_g \wedge \\
& \left. \bigwedge_{g=1}^{m-1} \bar{\eta}_{d-1, \mathbf{r}_g}(x_g + 1, x_{g+1}) \wedge \bigwedge_{g=1}^{m-1} \bar{\eta}_{d-1, \mathbf{r}_g^\dagger}(y_{g+1}, y_g + 1) \right)
\end{aligned}$$

We are now ready to define the predicate  $\bar{\chi}_{d,e}$ , where we set  $\theta_e = \pi_e^{\text{exact}}$  if  $e \in \mathcal{Z}$  and  $\theta_e = \pi_e$  if  $e \notin \mathcal{Z}$ :

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

The inductive definition of  $\bar{\eta}_{d,r}^\uparrow$  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  $\bar{\eta}_{d-1,r}$  and every occurrence of  $\chi_{d,e}$  by  $\bar{\chi}_{d,e}$ .

The inductive definition of  $\bar{\eta}_{d,r}$  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  $r \in R_L$  by making use of the sentence  $\nu_{\varphi, r}$  and the already defined  $\bar{\eta}_{d,r}^\uparrow$ .  $\square$

## 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  $\mathcal{L}_{k,l} \leq_{\text{cd}} L$ .*

*Proof.* Let  $L \subseteq \Sigma^\Delta$  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  $(k_0, l_0, k'_0, l'_0) \in \mathbb{N}_{>0}^4$  for which there exist  $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^\Delta)$  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 [27, 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  $uv' \in \Sigma^\Delta$  and  $u'v \in \Sigma^\Delta$ , so that we can consider the elements  $\text{ext}_{uuu,vv'v} = \text{ext}_{u,v} \circ \text{ext}_{u,v'} \circ \text{ext}_{u,v}$  and  $\text{ext}_{uu'u,vvv} = \text{ext}_{u,v} \circ \text{ext}_{u',v} \circ \text{ext}_{u,v}$  in  $\mathcal{O}(\Sigma^\Delta)$ . These elements satisfy  $\psi_L(\text{ext}_{uuu,vv'v}) \leq_{\mathfrak{J}} \psi_L(\text{ext}_{u,v})$  and  $\psi_L(\text{ext}_{uu'u,vvv}) \leq_{\mathfrak{J}} \psi_L(\text{ext}_{u,v})$ . We claim that we actually have  $\psi_L(\text{ext}_{uuu,vv'v}) <_{\mathfrak{J}} \psi_L(\text{ext}_{u,v})$  and  $\psi_L(\text{ext}_{uu'u,vvv}) <_{\mathfrak{J}} \psi_L(\text{ext}_{u,v})$ . Indeed, assume we would have  $\psi_L(\text{ext}_{uu'u,vvv}) \mathfrak{J} \psi_L(\text{ext}_{u,v})$ . Set  $x = \psi_L(\text{ext}_{u,v})$  and  $y = \psi_L(\text{ext}_{u',v})$ . By a classical property of Green's relations (see [27, 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 [27, 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  $\psi_L(\text{ext}_{(uu'u)^\omega,(vvv)^\omega}) = \psi_L(\text{ext}_{(uuu)^\omega,(vvv)^\omega})$  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  $(\varphi_L, \psi_L)$  is  $\varphi_L(L)$ -weakly-length-synchronous. Symmetrically, we can prove that if we had  $\psi_L(\text{ext}_{uuu,vv'v}) \mathfrak{J} \psi_L(\text{ext}_{u,v})$ , this would contradict the fact that  $(\varphi_L, \psi_L)$  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  $\mathcal{L}_{k,l} \leq_{\text{cd}} L_{\psi_L(\text{ext}_{u,v})}$ , so that since  $L_{\psi_L(\text{ext}_{u,v})} \leq_{\text{cd}} L$  (by Lemma 5.3) and by transitivity of  $\leq_{\text{cd}}$  we have  $\mathcal{L}_{k,l} \leq_{\text{cd}} L$ .

**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  $L_{\psi_L(\text{ext}_{u,v})}$ . 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  $x'\#y' \in L_{\psi_L(\text{ext}_{u,v})}$ .

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  $x'\#y' \in L_{\psi_L(\text{ext}_{u,v})}$ . 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  $\text{ext}_{x',y'} = \text{ext}_{z'_1, t'_1} \circ \dots \circ \text{ext}_{z'_n, t'_n}$  with  $z'_1, \dots, z'_n \in \{u^3, uu'u\}$  and  $t'_1, \dots, t'_n \in \{v^3, vv'v\}$  such that there exists  $i \in [1, n]$  satisfying  $\text{ext}_{z'_i, t'_i} \in \{\text{ext}_{u^3, vv'v}, \text{ext}_{uu'u, v^3}\}$ , thereby implying

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

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  $L_{\psi_L(\text{ext}_{u,v})}$ . 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  $x'\#y' \in L_{\psi_L(\text{ext}_{u,v})}$ .

This forms a valid reduction. Indeed, take a word  $w = xy$  with  $x = z_1 \dots z_n$  where  $n \in \mathbb{N}$  and  $z_1, \dots, z_n \in \{ac^{A \cdot B' - 1}, ac^{A' \cdot B - 1}\}$  and  $y = t_1 \dots t_m$  where  $m \in \mathbb{N}$  and  $t_1, \dots, t_m \in \{b_1, b_2\}$ . Consider  $x'\#y'$  produced by the reduction with  $x' = z'_1 \dots z'_n$  where  $z'_1, \dots, z'_n \in \{(u^3)^{B'}, (uu'u)^B\}$  and  $y' = t'_1 \dots t'_m$  where  $t'_1, \dots, t'_m \in \{(v^3)^{B'}, (vv'v)^B\}$ . If  $w \in \mathcal{L}_{A \cdot B', A' \cdot B}$ , then it easily follows that  $x'\#y' \in L_{\psi_L(\text{ext}_{u,v})}$ . 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  $z_1 \dots z_{i-1} t_{i-1} \dots t_1 \in \mathcal{L}_{A \cdot B', A' \cdot B}$  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  $\tilde{x}' = (uu'u)^{B-1} z'_{i+1} \dots z'_n$  and  $\tilde{y}' = t'_n \dots t'_{i+1} (v^3)^{B'-1}$ . If  $\Delta(\tilde{x}'\tilde{y}') \neq 0$ , then

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

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  $uv \in \Sigma^\Delta$ , 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  $u'v' \in \Sigma^\Delta$ , 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,  $\tilde{x}'\tilde{y}' \in \Sigma^\Delta$ . Hence  $x'y'$  is well-matched and

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

so that

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

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  $\psi_L(\text{ext}_{x',y'}) <_{\mathfrak{J}} \psi_L(\text{ext}_{u,v})$ .

- 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.  $\square$

## 5.5 Proof of Corollary 2.10

Let  $A = (Q, \Sigma, q_0, \delta_0, \dots, \delta_m)$  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  $\text{AC}^0$  or some  $m \geq 2$  such that  $\text{MOD}_m \leq_{\text{cd}} L$ .

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  $(\varphi_L, \psi_L) : (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R_L, O_L)$ .

The behavior of the  $m$ -VCA can be described as follows. To each  $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^\Delta)$  one assigns an  $(m+1)$ -tuple  $(f_i, g_i)_{i \in [0, m]}$  of pairs of functions from  $Q$  to  $Q$ .

Consider the monoid  $M = (Q^Q \times Q^Q, \odot, 1_M)$ , where  $\odot$  is reverse function composition on the first component and function composition on the second component, i.e. for all  $(f, g), (f', g') \in Q^Q \times Q^Q$  we have  $(f, g) \odot (f', g') = (f' \circ f, g \circ g')$ . Recall that  $A$  is an  $m$ -VCA. Consider the monoid  $\mathcal{M} = (M^{m+1}, \odot_{m+1}, 1_{\mathcal{M}})$ , where  $\odot_{m+1}$  is given by componentwise multiplication in  $M$ .

Consider the function  $\zeta : \mathcal{O}(\Sigma^\Delta) \rightarrow \mathcal{M}$ , where for all  $\text{ext}_{u,v} \in \mathcal{O}(\Sigma^\Delta)$  with  $\Delta(u) = j$  (and thus  $\Delta(v) = -j$ ) we set  $\zeta(\text{ext}_{u,v}) = (f_i, g_i)_{i \in [0, m]} \in M^{m+1}$  such that for all  $i \in [0, m]$  and all  $q \in Q$  we have  $f_i(q) = q'$  for the unique  $q' \in Q$  satisfying  $(q, i) \xrightarrow{u} (q', i+j)$  and  $g_i(q) = q''$  for the unique  $q'' \in Q$  satisfying  $(q, i+j) \xrightarrow{v} (q'', i)$ . The following points follow from definition.



1. The function  $\zeta : \mathcal{O}(\Sigma^\Delta) \rightarrow \mathcal{M}$  is a monoid morphism.
2. For all  $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^\Delta)$  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  $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^\Delta)$  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. Then there exist a  $\varphi_L(L)$ -reachable idempotent  $e \in O_L$  and  $\text{ext}_{u,v}, \text{ext}_{u',v'} \in \mathcal{O}(\Sigma^\Delta)$  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  $\omega$  denote the idempotent power of  $\mathcal{M}$ . Consider the elements

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

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

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

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  $\psi_L(\text{ext}_{x,y}) = \psi_L(\text{ext}_{x',y'}) = e$ , 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  $\psi_L(\text{ext}_{x,y}) = \psi_L(\text{ext}_{x',y'}) = e$ , again we obtain that  $L$  is not weakly length-synchronous.

## 6 Computability and decidability: 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  $(\varphi_L, \psi_L)$  and  $\varphi_L(L)$ .*

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

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

The syntactic Ext-algebra  $(R_L, O_L)$  of a VPL  $L$  over a visibly pushdown alphabet  $\Sigma$  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  $(\alpha, \beta): (R, O) \rightarrow (R_L, O_L)$ . 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  $(\varphi_1, \psi_1): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R_1, O_1)$  and  $(\varphi_2, \psi_2): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R_2, O_2)$  for  $\Sigma$  a visibly push-down 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 DVPA's  $A_1$  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  $L(A_2) \subseteq L(A_1)$ .  $\square$

*Proof of Proposition 6.1.* By Theorem 3.18 we first compute from our DVPA  $A$  on the visibly push-down alphabet  $\Sigma$  an Ext-algebra  $(R_A, O_A)$ , a morphism  $(\varphi_A, \psi_A): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R_A, O_A)$ , 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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  is a presentation of  $(R_L, O_L)$  with  $(\varphi, \psi)$  and  $F$  presentations of, respectively,  $(\varphi_L, \psi_L)$  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  $(\varphi, \psi)$  must be surjective, otherwise, by Lemma 3.15,  $(\varphi(\Sigma^\Delta), \psi(\mathcal{O}(\Sigma^\Delta)))|_{\varphi(\Sigma^\Delta)}$  would be a sub-Ext-algebra of  $(R, O)$  recognizing  $L$  such that  $\#(\varphi(\Sigma^\Delta), \psi(\mathcal{O}(\Sigma^\Delta)))|_{\varphi(\Sigma^\Delta)} < \#(R, O) = \#(R_L, O_L)$  while  $(R_L, O_L)$  divides  $(\varphi(\Sigma^\Delta), \psi(\mathcal{O}(\Sigma^\Delta)))|_{\varphi(\Sigma^\Delta)}$ , which is contradictory. Therefore, by Lemma 3.16, there is a surjective morphism  $(\alpha, \beta): (R, O) \rightarrow (R_L, O_L)$ , that must be bijective, such that  $\varphi_L = \alpha \circ \varphi$ , so that  $(R, O)$  is a presentation of  $(R_L, O_L)$  with  $(\varphi, \psi)$  and  $F$  presentations of, respectively,  $(\varphi_L, \psi_L)$  and  $\varphi_L(L)$ .

Under the assumption that such an Ext-algebra exists, we compute  $(R_L, O_L)$ ,  $(\varphi_L, \psi_L)$  and  $\varphi_L(L)$  by enumerating all the finitely many triples made of a finite Ext-algebra  $(R, O)$ , a morphism  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  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  $(R, O)$ ,  $(\varphi, \psi)$  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  $\alpha: R_L \rightarrow [1, |R_L|]$ . We define  $R$  to be the monoid with base set  $[1, |R_L|]$  and operation defined by  $x \cdot y = \alpha(\alpha^{-1}(x)\alpha^{-1}(y))$  for all  $x, y \in [1, |R_L|]$ . This is a monoid because

- $x \cdot \alpha(1_R) = \alpha(\alpha^{-1}(x)\alpha^{-1}(\alpha(1_R))) = \alpha(\alpha^{-1}(x)) = \alpha(\alpha^{-1}(\alpha(1_R))\alpha^{-1}(x)) = \alpha(1_R) \cdot x$  for all  $x \in R$ ; and

- for all  $x, y, z \in R$ , we have

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

Define the function  $\beta: O_L \rightarrow R^R$  by  $\beta(f')(x) = \alpha(f'(\alpha^{-1}(x)))$  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

- $\beta(1_O)(x) = \alpha(1_O(\alpha^{-1}(x))) = x = \text{id}_R(x)$  for all  $x \in R$ ; and
- for all  $f', g' \in O_L$ ,

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

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

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

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

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

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

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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$  for  $\Sigma$  a visibly pushdown alphabet and  $(R, O)$  a finite Ext-algebra, it is decidable if  $(\varphi, \psi)$  is quasi-aperiodic. If  $(\varphi, \psi)$  is not quasi-aperiodic, one can effectively compute  $k, l \in \mathbb{N}$  such that  $\psi(\mathcal{O}(\Sigma^\Delta)^{k,l})$  is not aperiodic.*

For the rest of this paragraph, let us fix a morphism  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ , where  $\Sigma$  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  $L_e = \{u\#v \mid uv \in \Sigma^\Delta \text{ s.t. } \psi(\text{ext}_{u,v}) = e\}$ , where  $\#$  is a fresh internal letter that does not appear in  $\Sigma$ , along with an associated morphism and subset.*

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

We define  $R' = R \cup O \cup \{\perp\}$ , for some fresh zero  $\perp$ , where multiplication between two elements in  $R'$  is defined as follows:

- multiplication between two elements in  $R$  is inherited from the monoid  $R$ ;
- $r \cdot f = \text{left}_r \circ f$  and  $f \cdot r = \text{right}_r \circ f$  for all  $r \in R$  and all  $f \in O$ ;
- $\perp$  acts as a zero, i.e.  $\perp \cdot r' = r' \cdot \perp = \perp$  for all  $r' \in R'$ ;
- $f \cdot g = \perp$  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

$$\begin{aligned} (r_1 \cdot f) \cdot r_2 &= (\text{left}_{r_1} \circ f) \cdot r_2 = \text{right}_{r_2} \circ (\text{left}_{r_1} \circ f) = (\text{right}_{r_2} \circ \text{left}_{r_1}) \circ f \\ &= (\text{left}_{r_1} \circ \text{right}_{r_2}) \circ f = r_1 \cdot (\text{right}_{r_2} \circ f) = r_1 \cdot (f \cdot r_2) . \end{aligned}$$

In the second case we have

$$(r_1 \cdot r_2) \cdot f = \text{left}_{r_1 r_2} \circ f = (\text{left}_{r_1} \circ \text{left}_{r_2}) \circ f = \text{left}_{r_1} \circ (\text{left}_{r_2} \circ f) = r_1 \cdot (r_2 \cdot f)$$

and in the third case we have

$$f \cdot (r_1 \cdot r_2) = \text{right}_{r_1 r_2} \circ f = (\text{right}_{r_2} \circ \text{right}_{r_1}) \circ f = \text{right}_{r_2} \circ (\text{right}_{r_1} \circ f) = (f \cdot r_1) \cdot r_2 .$$

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  $(\varphi', \psi') : (\Sigma'^{\Delta}, \mathcal{O}(\Sigma')^{\Delta}) \rightarrow (R', O')$  as the unique one satisfying  $\varphi'(c) = \varphi(c)$  for all  $c \in \Sigma_{\text{int}}$ ,  $\varphi'(\#) = \text{id}_O$  and where for all  $a \in \Sigma_{\text{call}}$ ,  $b \in \Sigma_{\text{ret}}$ , we have

$$\psi'(\text{ext}_{a,b})(x) = \begin{cases} \psi(\text{ext}_{a,b})(x) & \text{if } x \in R \\ \psi(\text{ext}_{a,b}) \circ x & \text{if } x \in O \\ \perp & \text{otherwise (i.e. if } x = \perp) \end{cases}$$

for all  $x \in R'$ . It suffices to prove the following claim, which directly implies the desired equality  $\varphi'^{-1}(e) = \{u\#v \mid uv \in \Sigma^{\Delta} \text{ s.t. } \psi(\text{ext}_{u,v}) = e\}$ . For all  $w \in \Sigma'^{\Delta}$  we have

$$\varphi'(w) = \begin{cases} \varphi(w) & \text{if } w \in \Sigma^{\Delta} \\ \psi(\text{ext}_{u,v}) & \text{if } w = u\#v \text{ for some } uv \in \Sigma^{\Delta} \\ \perp & \text{otherwise.} \end{cases}$$

We prove it by structural induction on  $w$ . The cases when  $w = \varepsilon$  or  $w = c \in \Sigma_{\text{int}}$  follow immediately from the definition of  $\varphi'$ . In case  $w = \# = \varepsilon\#\varepsilon$ , we have  $\varphi'(w) = \text{id}_O = \psi(\text{ext}_{\varepsilon,\varepsilon})$ .

For the inductive step first assume  $w = aw'b$  for some  $w' \in \Sigma'^{\Delta}$ . If  $w'$  is neither in  $\Sigma^{\Delta}$  nor of the form  $u\#v$  with  $uv \in \Sigma^{\Delta}$ , then  $\varphi'(w') = \perp$  by induction hypothesis and thus  $\varphi'(w) = \psi'(\text{ext}_{a,b})(\varphi'(w')) = \psi'(\text{ext}_{a,b})(\perp) = \perp$  as required. If  $w' \in \Sigma^{\Delta}$ , then  $\varphi'(w') = \varphi(w') \in R$  by induction hypothesis, and hence  $\varphi'(w) = \psi'(\text{ext}_{a,b})(\varphi'(w')) = \psi'(\text{ext}_{a,b})(\varphi(w')) = \psi(\text{ext}_{a,b})(\varphi(w')) =$

$\varphi(w)$  as required. If  $w' = u\#v$  with  $uv \in \Sigma^\Delta$ , i.e.  $w = au\#vb$ , then  $\varphi'(w') = \psi(\text{ext}_{u,v}) \in O$  by induction hypothesis. Hence, we have  $\varphi'(w) = \psi'(\text{ext}_{a,b})(\varphi'(w')) = \psi'(\text{ext}_{a,b})(\psi(\text{ext}_{u,v})) = \psi(\text{ext}_{a,b}) \circ \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{au,vb})$ .

Finally assume  $w = xy$  for some  $x, y \in \Sigma'^\Delta \setminus \{\varepsilon\}$ . The case when  $x$  or  $y$  is neither in  $\Sigma^\Delta$  nor of the form  $u\#v$  with  $uv \in \Sigma^\Delta$  is easily handled by applying the induction hypothesis and observing that  $\perp$  is a zero in  $R'$ . Two other immediate cases are when both  $x$  and  $y$  are in  $\Sigma^\Delta$  and when both  $x$  and  $y$  are of the form  $u\#v$  with  $uv \in \Sigma^\Delta$ . Consider the case when  $x \in \Sigma^\Delta \setminus \{\varepsilon\}$  and  $y = u\#v$  with  $uv \in \Sigma^\Delta$ , hence  $w = xu\#v$ . The induction hypothesis yields  $\varphi'(x) = \varphi(x) \in R$  and  $\varphi'(y) = \psi(\text{ext}_{u,v}) \in O$ . We obtain

$$\varphi'(xy) = \varphi'(x) \cdot \varphi'(y) = \varphi(x) \cdot \psi(\text{ext}_{u,v}) = \text{left}_{\varphi(x)} \circ \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{xu,v})$$

as required. Finally, let us treat the case when  $x = u\#v$  with  $uv \in \Sigma^\Delta$  and  $y \in \Sigma^\Delta \setminus \{\varepsilon\}$ , i.e.  $w = u\#vy$ . The induction hypothesis yields  $\varphi'(x) = \psi(\text{ext}_{u,v}) \in O$  and  $\varphi'(y) = \varphi(y) \in R$ . We obtain

$$\varphi'(xy) = \varphi'(x) \cdot \varphi'(y) = \psi(\text{ext}_{u,v}) \cdot \varphi(y) = \text{right}_{\varphi(y)} \circ \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u,vy})$$

as required.  $\square$

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  $A = (Q, \Sigma, \Gamma, \Omega, q_0, F, \perp)$ , where  $Q$  is a finite set of states,  $\Sigma$  is a finite input alphabet,  $\Gamma$  is a finite stack alphabet,  $q_0 \in Q$  is an initial state,  $F \subseteq Q$  is the set of final states,  $\perp \in \Gamma \setminus \Sigma$  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  $\alpha \in \Gamma^* \perp$  if  $X = \perp$  and  $\alpha \in (\Gamma \setminus \{\perp\})^*$  otherwise. The relation  $\Omega$  is naturally extended to the relation  $\Omega^* \subseteq Q \times \Sigma^* \times \Gamma^* \perp \times Q \times \Gamma^* \perp$ , namely as the smallest relation containing the set  $\{(p, \varepsilon, \alpha, p, \alpha) \mid p \in Q, \alpha \in \Gamma^* \perp\}$  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  $L(A) = \{w \in \Sigma^* \mid \exists \alpha \in \Gamma^* \perp, \exists q \in F : (q_0, w, \perp, q, \alpha) \in \Omega^*\}$ . 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  $\Sigma'$  such that  $L(A) \subseteq \{u\#v \mid uv \in \Sigma'^\Delta \setminus \Sigma'^* \# \Sigma'^*\}$  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  $A' = (Q, \Sigma', \Gamma, \Omega, q_0, F, \perp)$  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

$$A'' = (Q \times \{0, 1\}, \{0, 1, \#\}, \Gamma, \Omega', \langle q_0, 0 \rangle, F \times \{1\}, \perp),$$

where  $\Omega'$  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. [15,

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.  $\square$

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 recognizing  $L_e$ , 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

$$P(L_e) = \{(k, l) \in \mathbb{N} \times \mathbb{N} \mid \exists u \in \Sigma^k, v \in \Sigma^l : uv \in \Sigma^\Delta, \psi(\text{ext}_{u,v}) = e\}$$

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. [10] for a recent study. To decide whether  $(\varphi, \psi)$  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  $(\varphi, \psi)$  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  $\psi(\mathcal{O}(\Sigma^\Delta)^{k,l})$  is not aperiodic.  $\square$

**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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ , for  $\Sigma$  a visibly pushdown alphabet and  $(R, O)$  a finite Ext-algebra, and some  $F \subseteq R$ , it is decidable if  $(\varphi, \psi)$  is  $F$ -length-synchronous. If  $(\varphi, \psi)$  is not length-synchronous, one can effectively compute a tuple  $(k, l, k', l') \in \mathbb{N}_{>0}^4$  such that there exist  $uv, u'v' \in \Sigma^\Delta$  and some  $F$ -reachable idempotent  $e \in O$  such that  $\psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v'}) = e$ ,  $k = |u|$ ,  $l = |v|$ ,  $k' = |u'|$ ,  $l' = |v'|$  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  $S = \bigcup_{i \in I} (\vec{x}_{i,0} + \sum_{j=1}^{t_i} \mathbb{N} \vec{x}_{i,j}) \subseteq \mathbb{N}_{>0}^2$  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,*

$$\left| \left\{ \frac{k}{l} \mid (k, l) \in S \right\} \right| = 1 \iff \forall i, i' \in I \forall j \in [0, t_i] \forall j' \in [0, t_{i'}] : \vec{x}_{i,j} \text{ and } \vec{x}_{i',j'} \text{ are collinear.}$$

*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,  $(k, l) = \vec{x}_{i,0} + n_1 \vec{x}_{i,1} + \dots + n_{t_i} \vec{x}_{i,t_i}$  and  $(k', l') = \vec{x}_{i',0} + n'_1 \vec{x}_{i',1} + \dots + n'_{t_{i'}} \vec{x}_{i',t_{i'}}$  for some  $i, i' \in I$  and some  $n_1, \dots, n_{t_i}, n'_1, \dots, n'_{t_{i'}} \in \mathbb{N}$ . 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  $\frac{k}{l} \neq \frac{k'}{l'}$ .  $\square$

*Proof of Proposition 6.6.* Let us fix the Ext-algebra morphism  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ , where  $(R, O)$  is a finite Ext-algebra and where  $F \subseteq R$ .

Recall that over the alphabet  $\Sigma'$ , obtained from  $\Sigma$  by declaring a fresh letter  $\#$  as internal, the language

$$L_{e\uparrow} = \{u\#v \mid uv \in \Sigma^\Delta : \Delta(u) > 0, \psi(\text{ext}_{u,v}) = e\} = L_e \cap \{u\#v \mid uv \in \Sigma^\Delta : \Delta(u) > 0\}$$

is given for all  $e \in O$ . The language  $\{u\#v \mid uv \in \Sigma^\Delta : \Delta(u) > 0\}$  is a clearly a VPL. Thus, for all  $e \in O$ , we have that the set

$$P(L_{e\uparrow}) = \left\{ (k, l) \in \mathbb{N} \times \mathbb{N} \mid \exists u \in \Sigma^k, v \in \Sigma^l : uv \in \Sigma^\Delta, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = e \right\}$$

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  $(\varphi, \psi)$  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  $(\varphi, \psi)$  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  $(\varphi, \psi)$  is not  $F$ -length-synchronous, our algorithm outputs, again using the characterization of Lemma 6.7, a quadruple  $(k, l, k', l') \in \mathbb{N}_{>0}^4$  such that for some  $F$ -reachable idempotent  $e \in O$  we have  $(k, l), (k', l') \in P(L_{e\uparrow})$  and  $\frac{k}{l} \neq \frac{k'}{l'}$ .  $\square$

**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  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ , for  $\Sigma$  a visibly pushdown alphabet and  $(R, O)$  a finite Ext-algebra, and some  $F \subseteq R$ , it is decidable if  $(\varphi, \psi)$  is  $F$ -weakly-length-synchronous.*

Let us fix the morphism  $(\varphi, \psi): (\Sigma^\Delta, \mathcal{O}(\Sigma^\Delta)) \rightarrow (R, O)$ , for  $\Sigma$  a visibly pushdown alphabet and  $(R, O)$  a finite Ext-algebra, and some  $F \subseteq R$ .

Define the new visibly pushdown alphabet  $\bar{\Sigma}$  by  $\bar{\Sigma}_{\text{call}} = \{\bar{b} \mid b \in \Sigma_{\text{ret}}\}$ ,  $\bar{\Sigma}_{\text{int}} = \{\bar{c} \mid c \in \Sigma_{\text{int}}\}$  and  $\bar{\Sigma}_{\text{ret}} = \{\bar{a} \mid a \in \Sigma_{\text{call}}\}$ . For all  $w \in \Sigma^*$ , we define

$$\bar{w} = \begin{cases} \varepsilon & \text{if } w = \varepsilon \\ \bar{w}_n \cdots \bar{w}_1 & \text{if } w = w_1 \cdots w_n \text{ for } n \in \mathbb{N}_{>0} \text{ and } w_1, \dots, w_n \in \Sigma \end{cases}$$

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  $K_e = \{u\#\bar{u}' \mid u, u' \in \Sigma^*, \exists v \in \Sigma^* : uv \in \Sigma^\Delta, u'v \in \Sigma^\Delta, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e\}$ , where  $\#$  is a fresh internal letter that does not appear in  $\Sigma \cup \bar{\Sigma}$ , along with an associated morphism and subset.*

Over the alphabet  $\Sigma'$  obtained from  $\Sigma \cup \bar{\Sigma}$  by declaring the fresh letter  $\#$  as internal, we define

$$\begin{aligned} K_{e\uparrow} &= \left\{ u\#\bar{u}' \mid \begin{array}{l} u, u' \in \Sigma^*, \exists v \in \Sigma^* : \\ uv \in \Sigma^\Delta, u'v \in \Sigma^\Delta, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e \end{array} \right\} \\ &= K_e \cap \{u\#\bar{u}' \mid uu' \in (\Sigma \cup \bar{\Sigma})^\Delta : \Delta(u) > 0\} \end{aligned}$$

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  $\{u\#\bar{u}' \mid uu' \in (\Sigma \cup \bar{\Sigma})^\Delta : \Delta(u) > 0\}$  is a VPL and thus conclude that for all  $e \in O$ , the set

$$P(K_{e\uparrow}) = \left\{ (k, l) \in \mathbb{N} \times \mathbb{N} \mid \begin{array}{l} \exists u \in \Sigma^k, u' \in \Sigma^l, v \in \Sigma^* : \\ uv \in \Sigma^\Delta, u'v \in \Sigma^\Delta, \Delta(u) > 0, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e \end{array} \right\}$$

is effectively semilinear.

It is clear that  $(\varphi, \psi)$  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  $(\varphi, \psi)$  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  $\Sigma'$  be the alphabet that emerges from  $\Sigma \cup \bar{\Sigma}$  by additionally declaring  $\#$  as an internal letter. We will construct an Ext-algebra  $(R', O')$  and a morphism  $(\varphi', \psi')$  from  $(\Sigma'^\Delta, \mathcal{O}(\Sigma')^\Delta)$  to  $(R', O')$  such that for some subset  $F \subseteq R'$  we have  $K_e = \varphi'^{-1}(F)$ .

Let  $\bar{R} = \{\bar{r} \mid r \in R\}$ . We define  $R' = R \cup \bar{R} \cup \mathfrak{P}(O^2) \setminus \emptyset \cup \{\perp, 1\}$ , for some fresh zero  $\perp$  and identity 1, where multiplication between two elements in  $R'$  is defined as follows:

- for all  $r_1, r_2 \in R$ ,

$$\begin{array}{ll} r_1 \cdot r_2 = r_1 r_2 & \bar{r}_1 \cdot r_2 = \perp \\ \bar{r}_1 \cdot \bar{r}_2 = \bar{r}_2 \bar{r}_1 & r_1 \cdot \bar{r}_2 = \perp ; \end{array}$$

- for all  $r \in R$  and  $E \in \mathfrak{P}(O^2) \setminus \emptyset$ ,

$$\begin{array}{ll} r \cdot E = \{(\text{left}_r \circ e_1, e_2) \mid (e_1, e_2) \in E\} & E \cdot r = \perp \\ E \cdot \bar{r} = \{(e_1, \text{left}_r \circ e_2) \mid (e_1, e_2) \in E\} & \bar{r} \cdot E = \perp ; \end{array}$$

- for all  $E_1, E_2 \in \mathfrak{P}(O^2) \setminus \emptyset$ , we have  $E_1 \cdot E_2 = \perp$ ;
- $\perp$  acts as a zero, i.e.  $\perp \cdot r' = r' \cdot \perp = \perp$  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  $\bar{r}_1 \cdot \bar{r}_2 \cdot \bar{r}_3$ ,  $r_1 \cdot E \cdot \bar{r}_2$ ,  $r_1 \cdot r_2 \cdot E$  and  $E \cdot \bar{r}_1 \cdot \bar{r}_2$ , where  $E \in \mathfrak{P}(O^2) \setminus \emptyset$  and  $r_1, r_2, r_3 \in R$ . In the first case we have

$$(\bar{r}_1 \cdot \bar{r}_2) \cdot \bar{r}_3 = \bar{r}_2 \bar{r}_1 \cdot \bar{r}_3 = \bar{r}_3 \bar{r}_2 \bar{r}_1 = \bar{r}_1 \cdot \bar{r}_3 \bar{r}_2 = \bar{r}_1 \cdot (\bar{r}_2 \cdot \bar{r}_3) .$$



In the second case we have

$$\begin{aligned}
(r_1 \cdot E) \cdot \bar{r}_2 &= \{(\text{left}_{r_1} \circ e_1, e_2) \mid (e_1, e_2) \in E\} \cdot \bar{r}_2 \\
&= \{(\text{left}_{r_1} \circ e_1, \text{left}_{r_2} \circ e_2) \mid (e_1, e_2) \in E\} \\
&= r_1 \cdot \{(e_1, \text{left}_{r_2} \circ e_2) \mid (e_1, e_2) \in E\} = r_1 \cdot (E \cdot \bar{r}_2) .
\end{aligned}$$

In the third case we have

$$\begin{aligned}
(r_1 \cdot r_2) \cdot E &= \{(\text{left}_{r_1 r_2} \circ e_1, e_2) \mid (e_1, e_2) \in E\} \\
&= \{(\text{left}_{r_1} \circ \text{left}_{r_2} \circ e_1, e_2) \mid (e_1, e_2) \in E\} \\
&= r_1 \cdot \{(\text{left}_{r_2} \circ e_1, e_2) \mid (e_1, e_2) \in E\} = r_1 \cdot (r_2 \cdot E)
\end{aligned}$$

and in the fourth case we have

$$\begin{aligned}
E \cdot (\bar{r}_1 \cdot \bar{r}_2) &= \{(e_1, \text{left}_{r_2 r_1} \circ e_2) \mid (e_1, e_2) \in E\} \\
&= \{(e_1, \text{left}_{r_2} \circ \text{left}_{r_1} \circ e_2) \mid (e_1, e_2) \in E\} \\
&= \{(e_1, \text{right}_{r_1} \circ e_2) \mid (e_1, e_2) \in E\} \cdot \bar{r}_2 = (E \cdot \bar{r}_1) \cdot \bar{r}_2 .
\end{aligned}$$

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  $(\varphi', \psi') : (\Sigma'^{\Delta}, \mathcal{O}(\Sigma')^{\Delta}) \rightarrow (R', O')$  as the unique one satisfying  $\varphi'(c) = \varphi(c)$  and  $\varphi'(\bar{c}) = \overline{\varphi(c)}$  for all  $c \in \Sigma_{\text{int}}$ ,  $\varphi'(\#) = \{(\psi(\text{ext}_{\varepsilon, v}), \psi(\text{ext}_{\varepsilon, v})) \mid v \in \Sigma^{\Delta}\}$  and where for all  $a, a' \in \Sigma_{\text{call}}$ ,  $b, b' \in \Sigma_{\text{ret}}$ , we have

$$\begin{aligned}
\psi'(\text{ext}_{a,b})(x) &= \begin{cases} \psi(\text{ext}_{a,b})(1_R) & \text{if } x = 1 \\ \psi(\text{ext}_{a,b})(x) & \text{if } x \in R \\ \perp & \text{otherwise} \end{cases} \\
\psi'(\text{ext}_{\bar{a}', \bar{a}})(x) &= \begin{cases} \overline{\psi(\text{ext}_{a', b'})(1_R)} & \text{if } x = 1 \\ \overline{\psi(\text{ext}_{a', b'})(x')} & \text{if } x = \bar{x}' \text{ for } x' \in R \\ \perp & \text{otherwise} \end{cases} \\
\psi'(\text{ext}_{a, \bar{a}'})(x) &= \begin{cases} \bigcup_{\substack{b \in \Sigma_{\text{ret}} \\ z \in \Sigma^{\Delta} \\ (e_1, e_2) \in x}} \{(\psi(\text{ext}_{a, bz}) \circ e_1, \psi(\text{ext}_{a', bz}) \circ e_2)\} & \text{if } x \in \mathfrak{P}(O^2) \setminus \emptyset \\ \perp & \text{otherwise} \end{cases} \\
\psi'(\text{ext}_{\bar{b}', b})(x) &= \perp
\end{aligned}$$

for all  $x \in R'$ . Note that  $(\varphi', \psi')$  is computable because

$$\begin{aligned}
\{(\psi(\text{ext}_{\varepsilon, v}), \psi(\text{ext}_{\varepsilon, v})) \mid v \in \Sigma^{\Delta}\} &= \{(\text{right}_{\varphi(v)}, \text{right}_{\varphi(v)}) \mid v \in \Sigma^{\Delta}\} \\
&= \{(\text{right}_r, \text{right}_r) \mid r \in R\}
\end{aligned}$$

and

$$\begin{aligned}
& \bigcup_{\substack{b \in \Sigma_{\text{ret}} \\ z \in \Sigma^\Delta \\ (e_1, e_2) \in x}} \{(\psi(\text{ext}_{a,bz}) \circ e_1, \psi(\text{ext}_{a',bz}) \circ e_2)\} \\
&= \bigcup_{\substack{b \in \Sigma_{\text{ret}} \\ z \in \Sigma^\Delta \\ (e_1, e_2) \in x}} \{(\text{right}_{\varphi(z)} \circ \psi(\text{ext}_{a,b}) \circ e_1, \text{right}_{\varphi(z)} \circ \psi(\text{ext}_{a',b}) \circ e_2)\} \\
&= \bigcup_{\substack{b \in \Sigma_{\text{ret}} \\ r \in R \\ (e_1, e_2) \in x}} \{(\text{right}_r \circ \psi(\text{ext}_{a,b}) \circ e_1, \text{right}_r \circ \psi(\text{ext}_{a',b}) \circ e_2)\}
\end{aligned}$$

for all  $x \in \mathfrak{P}(O^2) \setminus \emptyset$  and  $a, a' \in \Sigma_{\text{call}}$ .

Now define the set of pairs  $P = \{(u, u') \in \Sigma^* \times \Sigma^* \mid \exists v \in \Sigma^* : uv \in \Sigma^\Delta, u'v \in \Sigma^\Delta\}$ ; it is not difficult to check that for all  $w \in \Sigma'^*$ ,  $w \in \Sigma'^\Delta \cap \Sigma^* \# \overline{\Sigma^*}$  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

$$\begin{aligned}
& \varphi'^{-1}(\{E \in \mathfrak{P}(O^2) \setminus \emptyset \mid (e, e) \in E\}) \\
&= \{u \# \overline{u'} \mid (u, u') \in P, \exists v \in \Sigma^* : uv \in \Sigma^\Delta, u'v \in \Sigma^\Delta, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e\} \\
&= \{u \# \overline{u'} \mid u, u' \in \Sigma^*, \exists v \in \Sigma^* : uv \in \Sigma^\Delta, u'v \in \Sigma^\Delta, \psi(\text{ext}_{u,v}) = \psi(\text{ext}_{u',v}) = e\}.
\end{aligned}$$

For all  $w \in \Sigma'^\Delta$  we have

$$\varphi'(w) = \begin{cases} 1 & \text{if } w = \varepsilon \\ \varphi(w) & \text{if } w \in \Sigma^\Delta \setminus \{\varepsilon\} \\ \frac{\varphi(w)}{\varphi(w')} & \text{if } w = \overline{w'} \text{ for } w' \in \Sigma^\Delta \setminus \{\varepsilon\} \\ \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} & \text{if } w = u \# \overline{u'} \text{ for } (u, u') \in P \\ \perp & \text{otherwise.} \end{cases}$$

We prove it by structural induction on  $w$ . The cases when  $w = \varepsilon$  or  $w = c \in \Sigma_{\text{int}}$  or  $w = \bar{c} \in \overline{\Sigma}_{\text{int}}$  follow immediately from the definition of  $\varphi'$ . In case  $w = \# = \varepsilon \# \varepsilon$ , we have

$$\varphi'(w) = \{(\psi(\text{ext}_{\varepsilon,v}), \psi(\text{ext}_{\varepsilon,v})) \mid v \in \Sigma^\Delta\} = \{(\psi(\text{ext}_{\varepsilon,v}), \psi(\text{ext}_{\varepsilon,v})) \mid v \in \Sigma^*, \varepsilon v \in \Sigma^\Delta\}$$

as required.

For the inductive step first assume  $w = \alpha w' \beta$  for  $w' \in \Sigma'^\Delta$ ,  $\alpha \in \Sigma_{\text{call}} \cup \overline{\Sigma}_{\text{call}}$  and  $\beta \in \Sigma_{\text{ret}} \cup \overline{\Sigma}_{\text{ret}}$ . If  $w'$  is neither in  $\Sigma^\Delta \cup \overline{\Sigma}^\Delta$  nor of the form  $u \# \overline{u'}$  with  $(u, u') \in P$ , then  $\varphi'(w') = \perp$  by induction hypothesis and thus  $\varphi'(w) = \psi'(\text{ext}_{\alpha,\beta})(\varphi'(w')) = \psi'(\text{ext}_{\alpha,\beta})(\perp) = \perp$  as required, since  $w$  is also neither in  $\Sigma^\Delta \cup \overline{\Sigma}^\Delta$  nor of the form  $u \# \overline{u'}$  with  $(u, u') \in P$ . If  $w' = \varepsilon$ , then  $\varphi'(w') = 1$  and hence

$$\begin{aligned}
\varphi'(w) &= \psi'(\text{ext}_{\alpha,\beta})(\varphi'(w')) \\
&= \psi'(\text{ext}_{\alpha,\beta})(1) \\
&= \begin{cases} \psi(\text{ext}_{\alpha,\beta})(1_R) = \varphi(w) & \text{if } \alpha \in \Sigma_{\text{call}} \text{ and } \beta \in \Sigma_{\text{ret}} \\ \psi(\text{ext}_{a,b})(1_R) = \varphi(ab) & \text{if } \alpha = \bar{b} \in \overline{\Sigma}_{\text{call}} \text{ and } \beta = \bar{a} \in \overline{\Sigma}_{\text{ret}} \\ \perp & \text{otherwise} \end{cases}
\end{aligned}$$

as required, because  $w = \bar{b}\bar{a} = \overline{ab}$  in the second case. If  $w' \in \Sigma^\Delta \setminus \{\varepsilon\}$ , then  $\varphi'(w') = \varphi(w') \in R$  by induction hypothesis, and hence

$$\begin{aligned}\varphi'(w) &= \psi'(\text{ext}_{\alpha,\beta})(\varphi'(w')) = \psi'(\text{ext}_{\alpha,\beta})(\varphi(w')) \\ &= \begin{cases} \psi(\text{ext}_{\alpha,\beta})(\varphi(w')) = \varphi(w) & \text{if } \alpha \in \Sigma_{\text{call}} \text{ and } \beta \in \Sigma_{\text{ret}} \\ \perp & \text{otherwise} \end{cases}\end{aligned}$$

as required. If  $w' \in \overline{\Sigma}^\Delta \setminus \{\varepsilon\}$ , then  $w' = \overline{w''}$  for  $w'' \in \Sigma^\Delta \setminus \{\varepsilon\}$ , so  $\varphi'(w') = \overline{\varphi(w'')} \in \overline{R}$  by induction hypothesis, and hence

$$\begin{aligned}\varphi'(w) &= \psi'(\text{ext}_{\alpha,\beta})(\varphi'(w')) \\ &= \psi'(\text{ext}_{\alpha,\beta})(\overline{\varphi(w'')}) \\ &= \begin{cases} \overline{\psi(\text{ext}_{a,b})(\varphi(w''))} = \overline{\varphi(aw''b)} & \text{if } \alpha = \bar{b} \in \overline{\Sigma}_{\text{call}} \text{ and } \beta = \bar{a} \in \overline{\Sigma}_{\text{ret}} \\ \perp & \text{otherwise} \end{cases}\end{aligned}$$

as required, because  $w = \bar{b}\overline{w''}\bar{a} = \overline{aw''b}$  in the first case. If  $w' = u\#u'$  with  $(u, u') \in P$ , i.e.  $w = \alpha u\#u'\beta$ , then

$$\varphi'(w') = \{(\psi(\text{ext}_{u,v'}), \psi(\text{ext}_{u',v'})) \mid v' \in \Sigma^*, uv', u'v' \in \Sigma^\Delta\} \in \mathfrak{P}(O^2) \setminus \emptyset$$

by induction hypothesis. Hence, we have

$$\begin{aligned}\varphi'(w) &= \psi'(\text{ext}_{\alpha,\beta})(\varphi'(w')) \\ &= \psi'(\text{ext}_{\alpha,\beta})(\{(\psi(\text{ext}_{u,v'}), \psi(\text{ext}_{u',v'})) \mid v' \in \Sigma^*, uv', u'v' \in \Sigma^\Delta\}) \\ &= \bigcup_{\substack{b \in \Sigma_{\text{ret}} \\ z \in \Sigma^\Delta}} \{(\psi(\text{ext}_{au,v'bz}), \psi(\text{ext}_{a'u',v'bz})) \mid v' \in \Sigma^*, uv', u'v' \in \Sigma^\Delta\} \\ &= \{(\psi(\text{ext}_{au,v}), \psi(\text{ext}_{a'u',v})) \mid v \in \Sigma^*, auv, a'u'v \in \Sigma^\Delta\}\end{aligned}$$

if  $\alpha = a \in \Sigma_{\text{call}}$  and  $\beta = \bar{a}' \in \overline{\Sigma}_{\text{ret}}$  (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  $\varphi'(w) = \perp$  otherwise, as required.

Finally assume  $w = xy$  for some  $x, y \in \Sigma'^\Delta \setminus \{\varepsilon\}$ . The case when  $x$  or  $y$  is neither in  $\Sigma^\Delta \cup \overline{\Sigma}^\Delta$  nor of the form  $u\#\bar{u}'$  with  $(u, u') \in P$  is easily handled by applying the induction hypothesis and observing that  $\perp$  is a zero in  $R'$ . Four other immediate cases are when both  $x$  and  $y$  are in  $\Sigma^\Delta$ , when  $x$  is in  $\Sigma^\Delta$  and  $y$  in  $\overline{\Sigma}^\Delta$ , when  $x$  is in  $\overline{\Sigma}^\Delta$  and  $y$  in  $\Sigma^\Delta$  and when both  $x$  and  $y$  are of the form  $u\#\bar{u}'$  with  $(u, u') \in P$ . For the case when both  $x$  and  $y$  are in  $\overline{\Sigma}^\Delta$ , we have that  $x = \overline{x'}$  and  $y = \overline{y'}$  for  $x', y' \in \Sigma^\Delta$ , 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  $xy = \overline{x'y'} = \overline{y'x'}$ . Consider the case when  $x \in (\Sigma^\Delta \cup \overline{\Sigma}^\Delta) \setminus \{\varepsilon\}$  and  $y = u\#\bar{u}'$  with  $(u, u') \in P$ , hence  $w = xu\#\bar{u}'$ . The induction hypothesis yields

$$\varphi'(x) = \begin{cases} \varphi(x) \in R & \text{if } x \in \Sigma^\Delta \setminus \{\varepsilon\} \\ \overline{\varphi(x')} \in \overline{R} & \text{if } x = \overline{x'} \text{ for } x' \in \Sigma^\Delta \setminus \{\varepsilon\} \end{cases}$$

and  $\varphi'(y) = \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} \in \mathfrak{P}(O^2) \setminus \emptyset$ . We obtain

$$\begin{aligned}\varphi'(w) &= \varphi'(x) \cdot \varphi'(y) \\ &= \varphi(x) \cdot \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} \\ &= \{(\text{left}_{\varphi(x)} \circ \psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} \\ &= \{(\psi(\text{ext}_{xu,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, xuv, u'v \in \Sigma^\Delta\}\end{aligned}$$

if  $x \in \Sigma^\Delta \setminus \{\varepsilon\}$  and  $\varphi'(w) = \overline{\varphi(x')} \cdot \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} = \perp$  if  $x = \overline{x'}$  for  $x' \in \Sigma^\Delta \setminus \{\varepsilon\}$ , as required. Eventually, let us treat the case when  $x = u\#u'$  with  $(u, u') \in P$  and  $y \in (\Sigma^\Delta \cup \overline{\Sigma^\Delta}) \setminus \{\varepsilon\}$ , hence  $w = u\#\overline{u'}y$ . The induction hypothesis yields  $\varphi'(x) = \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} \in \mathfrak{P}(O^2) \setminus \emptyset$  and  $\varphi'(y) = \begin{cases} \varphi(y) \in R & \text{if } y \in \Sigma^\Delta \setminus \{\varepsilon\} \\ \overline{\varphi(y')} \in \overline{R} & \text{if } y = \overline{y'} \text{ for } y' \in \Sigma^\Delta \setminus \{\varepsilon\} \end{cases}$ . We obtain

$$\begin{aligned}\varphi'(w) &= \varphi'(x) \cdot \varphi'(y) \\ &= \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} \cdot \overline{\varphi(y')} \\ &= \{(\psi(\text{ext}_{u,v}), \text{left}_{\varphi(y')} \circ \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} \\ &= \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{y'u',v})) \mid v \in \Sigma^*, uv, y'u'v \in \Sigma^\Delta\}\end{aligned}$$

if  $y = \overline{y'}$  for  $y' \in \Sigma^\Delta \setminus \{\varepsilon\}$  and  $\varphi'(w) = \{(\psi(\text{ext}_{u,v}), \psi(\text{ext}_{u',v})) \mid v \in \Sigma^*, uv, u'v \in \Sigma^\Delta\} \cdot \varphi(y) = \perp$  if  $x \in \Sigma^\Delta \setminus \{\varepsilon\}$ , as required, because  $\overline{y'u'} = \overline{u'y'}$  in the first case.  $\square$

## 7 Conclusion

In this paper we have studied the question which visibly pushdown languages lie in the complexity class  $\text{AC}^0$ .

We have introduced the notions of length-synchronicity and weak length-synchronicity. We have introduced intermediate languages: these are particular one-turn visibly pushdown languages generated by star-closed regular synchronization languages describing weakly length-synchronous but not length synchronous word relations and whose syntactic morphism does not contain any group when applied (convolutions of) words from any set of the form  $\Sigma^k \times \Sigma^l$ . To the best of our knowledge our community is lacking tools to determine if there is at all any intermediate language that is in  $\text{AC}^0$  (even in  $\text{ACC}^0$ ) or provably not in  $\text{AC}^0$ . We conjecture that none of the intermediate languages are in  $\text{ACC}^0$  nor  $\text{TC}^0$ -hard.

Our main result states that there is an algorithm that, given a visibly pushdown language  $L$ , outputs if  $L$  surely lies in  $\text{AC}^0$ , surely does not lie in  $\text{AC}^0$ , or outputs a disjoint finite union of intermediate languages that  $L$  is constant-depth equivalent to and moreover distinct  $k, l \in \mathbb{N}_{>0}$  such that  $L$  is hard for a concrete intermediate language of the form  $\mathcal{L}_{k,l} = L(S \rightarrow \varepsilon \mid ac^{k-1}Sb_1 \mid ac^{l-1}Sb_2)$ .

As main tools we carefully revisited Ext-algebras, introduced by Czarnetzki, Lange and Krebs [11], being closely related to forest algebras, introduced by Bojańczyk and Walukiewicz [7]. For the reduction from  $\mathcal{L}_{k,l}$  we made use of Green's relations.

Natural questions arise. Is there any *concrete* intermediate language that is provably in  $\text{ACC}^0$ , provably not in  $\text{AC}^0$ , or hard for  $\text{TC}^0$ ? Another exciting question is whether one can effectively compute those visibly pushdown languages that lie in the complexity class  $\text{TC}^0$ . Is there a  $\text{TC}^0/\text{NC}^1$  complexity dichotomy? For these questions new techniques seem to be necessary. In this context it

is already interesting to mention there is an  $\text{NC}^1$ -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 [25].

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  $S \rightarrow aSb \mid a'cSb' \mid \varepsilon$ , 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  $\gamma$  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  $S \rightarrow acbc \mid aSb \mid \varepsilon$  does not satisfy SHB, but its matching relation is definable in FO[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  $S \rightarrow aSb \mid a_1cb_1 \mid a_2ccb_2$  is a counter-example.

8. The statement of Proposition 131 is wrong. The language  $\{a^n b^n \mid n \geq 0\}^*$  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,  $\text{cancel}^{|HL|}(w) = \Sigma_{\text{int}}^*$  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  $S \rightarrow aSb \mid a_1cb_1 \mid a_2ccb_2$  which is a visibly counter language that does not have the SHB property. However, it is in  $\text{AC}^0$ .

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](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 Eighth 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] Laura Chaubard, Jean-Éric Pin, and Howard Straubing. First order formulas with modular predicates. In *21th IEEE Symposium on Logic in Computer Science (LICS 2006), 12-15 August 2006, Seattle, WA, USA, Proceedings*, pages 211–220. IEEE Computer Society, 2006. doi:10.1109/LICS.2006.24.

- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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>.
- [14] 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>.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] Michael A. Harrison. *Introduction to Formal Language Theory*. Addison-Wesley, 1978.
- [19] 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.
- [20] Neil Immerman. Languages that capture complexity classes. *SIAM J. Comput.*, 16(4):760–778, 1987. doi:10.1137/0216051.
- [21] Neil Immerman. *Descriptive complexity*. Graduate texts in computer science. Springer, 1999. doi:10.1007/978-1-4612-0539-5.
- [22] Stasys Jukna. *Boolean Function Complexity - Advances and Frontiers*, volume 27 of *Algorithms and combinatorics*. Springer, 2012. doi:10.1007/978-3-642-24508-4.
- [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] 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.
- [25] 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/>.
- [26] Kurt Mehlhorn. Pebbling Mountain Ranges and its Application of DCFL-Recognition. In J. W. de Bakker and Jan van Leeuwen, editors, *Automata, Languages and Programming, 7th Colloquium, Noordwijkerhout, 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.
- [27] Jean-Éric Pin. *Varieties of Formal Languages*. North Oxford, London and Plenum, New-York, 1986. (Traduction de Variétés de langages formels).
- [28] Jean-Éric Pin, editor. *Handbook of Automata Theory*. European Mathematical Society Publishing House, Zürich, Switzerland, 2021. doi:10.4171/Automata.
- [29] 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.
- [30] Howard Straubing. *Finite Automata, Formal Logic, and Circuit Complexity*. Birkhäuser Boston, 1994. doi:10.1007/978-1-4612-0289-9.
- [31] 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.
- [32] 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.