

# On Finite-Index Indexed Grammars and Their Restrictions <sup>\*</sup>

Flavio D'Alessandro<sup>1</sup>, Oscar H. Ibarra<sup>2</sup>, and Ian McQuillan<sup>3</sup>

<sup>1</sup> Department of Mathematics  
Sapienza University of Rome  
00185 Rome, Italy  
dalessan@mat.uniroma1.it

and

Department of Mathematics, Boğaziçi University  
34342 Bebek, Istanbul, Turkey

<sup>2</sup> Department of Computer Science  
University of California, Santa Barbara, CA 93106, USA  
ibarra@cs.ucsb.edu

<sup>3</sup> Department of Computer Science, University of Saskatchewan  
Saskatoon, SK S7N 5A9, Canada  
mcquillan@cs.usask.ca

**Abstract.** The family,  $\mathcal{L}(\text{IND}_{\text{LIN}})$ , of languages generated by linear indexed grammars has been studied in the literature. It is known that the Parikh image of every language in  $\mathcal{L}(\text{IND}_{\text{LIN}})$  is semi-linear. However, there are bounded semi-linear languages that are not in  $\mathcal{L}(\text{IND}_{\text{LIN}})$ . Here, we look at larger families of (restricted) indexed languages and study their properties, their relationships, and their decidability properties.

**Keywords:** Indexed Languages, Finite-Index, Full Trios, Semi-linearity, Bounded Languages, ETOL Languages

## 1 Introduction

Indexed grammars [1,2] are a natural generalization of context-free grammars, where variables keep stacks of indices. Despite being all context-sensitive languages, the languages are still quite general as they can generate non-semi-linear languages [1]. Several restrictions have been studied that have desirable computational properties. Linear indexed grammars were first created, restricting the number of variables on the right hand side to be at most one [5]. Other restrictions include another system named exactly linear indexed grammars [6] (see

---

<sup>\*</sup> The research of F. D'Alessandro was supported by a EC-FP7 Marie-Curie/TÜBITAK/Co-Funded Brain Circulation Scheme Project 2236 Fellowship. The research of O. H. Ibarra was supported, in part, by NSF Grant CCF-1117708. The research of I. McQuillan was supported, in part, by a grant from Natural Sciences and Engineering Research Council of Canada.

also [17]), which are different than the first formalisms, although both are sufficiently restricted to only generate semi-linear languages. In this paper, we only examine the first formalism of linear indexed grammars.

We study indexed grammars that are restricted to be finite-index, which is a generalization of linear indexed grammars [5]. Grammar systems that are  $k$ -index are restricted so that, for every word generated by the grammar, there is some successful derivation where at most  $k$  variables (or nonterminals) appear in every sentential form of the derivation [15,13]. A system is finite-index if it is  $k$ -index for some  $k$ . It has been found that when restricting many different types of grammar systems to be finite-index, their languages coincide. This is the case for finite-index ETOL, EDTOL, context-free programmed grammars, ordered grammars, and matrix grammars.

We introduce the family,  $\mathcal{L}(\text{IND}_{\text{FIN}})$ , of languages generated by finite-index indexed grammars and study a sub-family,  $\mathcal{L}(\text{IND}_{\text{UFIN}})$ , of languages generated by uncontrolled finite-index indexed grammars, where every successful derivation has to be finite-index. These have been very recently studied under the name *breadth-bounded grammars*, where it was shown that this family is a semilinear full trio. We also study a special case of the latter, called  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$  that restricts branching productions. We then show the following:

1. All families are semilinear full trios.
2. The following conditions are equivalent for a bounded language  $L$ :
  - $L \in \mathcal{L}(\text{IND}_{\text{UFIN}_1})$ ,
  - $L \in \mathcal{L}(\text{IND}_{\text{UFIN}})$ ,
  - $L$  is bounded semilinear,
  - $L$  can be generated by a finite-index ETOL system,
  - $L$  can be accepted by a DFA augmented with reversal-bounded counters,
3. Every finite-index ETOL language is in  $\mathcal{L}(\text{IND}_{\text{FIN}})$ .
4.  $\mathcal{L}(\text{CFL}) \subset \mathcal{L}(\text{IND}_{\text{LIN}}) \subset \mathcal{L}(\text{IND}_{\text{UFIN}_1}) \subseteq \mathcal{L}(\text{IND}_{\text{UFIN}}) \subset \mathcal{L}(\text{IND}_{\text{FIN}})$ .
5. Containment and equality are decidable for bounded languages in  $\mathcal{L}(\text{IND}_{\text{LIN}})$  and  $\mathcal{L}(\text{IND}_{\text{UFIN}})$ .

## 2 Preliminaries

We assume a basic background in formal languages and automata theory [9].

Let  $\mathbb{N}^k$  be the additive free commutative monoid of non negative integers. If  $B$  is a subset of  $\mathbb{N}^k$ ,  $B^\oplus$  denotes the submonoid of  $\mathbb{N}^k$  generated by  $B$ .

An *alphabet* is a finite set of symbols, and given an alphabet  $A$ ,  $A^*$  is the free monoid generated by  $A$ . An element  $w \in A^*$  is called a *word*, the empty word is denoted by  $\lambda$ , and any  $L \subseteq A^*$  is a *language*. The *length* of a word  $w \in A^*$  is denoted by  $|w|$ , and the number of  $a$ 's,  $a \in A$ , in  $w$  is denoted by  $|w|_a$ , extended to subsets  $X$  of  $A$  by  $|w|_X = \sum_{a \in X} |w|_a$ .

Let  $A = \{a_1, \dots, a_t\}$  be an alphabet of  $t$  letters, and let  $\psi : A^* \rightarrow \mathbb{N}^t$  be the corresponding *Parikh morphism* defined by  $\psi(w) = (|w|_{a_1}, \dots, |w|_{a_t})$  extended to languages  $L \subseteq A^*$ .

A set  $B \subseteq \mathbb{N}^k$  is a *linear set* if there exist vectors  $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_n$  of  $\mathbb{N}^k$  such that  $B = \mathbf{b}_0 + \{\mathbf{b}_1, \dots, \mathbf{b}_n\}^\oplus$ . Further,  $B$  is called a *semi-linear set* if  $B = \bigcup_{i=1}^m B_i, m \geq 1$ , for linear sets  $B_1, \dots, B_m$ . A language  $L \subseteq A^*$  is said to be *semi-linear* if the Parikh morphism applied to  $L$  gives a semi-linear set. A language family is said to be semi-linear if all languages in the family are semi-linear. Many known families are semi-linear, such as the regular languages, context-free (denoted by  $\mathcal{L}(\text{CFL})$ , see [9]), and finite-index ETOL languages ( $\mathcal{L}(\text{ETOL}_{\text{FIN}})$ ), see [14,13]).

A language  $L$  is termed *bounded* if there exist non-empty words  $u_1, \dots, u_k$ , with  $k \geq 1$ , such that  $L \subseteq u_1^* \cdots u_k^*$ . Let  $\varphi : \mathbb{N}^k \rightarrow u_1^* \cdots u_k^*$  be the map defined as: for every tuple  $(\ell_1, \dots, \ell_k) \in \mathbb{N}^k$ ,

$$\varphi(\ell_1, \dots, \ell_k) = u_1^{\ell_1} \cdots u_k^{\ell_k}.$$

The map  $\varphi$  is called the *Ginsburg map*.

**Definition 1.** A bounded language  $L \subseteq u_1^* \cdots u_k^*$  is said to be bounded Ginsburg semi-linear if there exists a semi-linear set  $B$  of  $\mathbb{N}^k$  such that  $\varphi(B) = L$ .

In the literature, bounded Ginsburg semi-linear has also been called just bounded semi-linear, but we will use the terminology bounded Ginsburg semi-linear henceforth in this paper.

A *full trio* is a language family closed under morphism, inverse morphism, and intersection with regular languages [3].

We will also relate our results to the languages accepted by one-way non-deterministic reversal-bounded multicounter machines (denoted by  $\mathcal{L}(\text{NCM})$ ), and to one-way deterministic reversal-bounded multicounter machines (denoted by  $\mathcal{L}(\text{DCM})$ ). These are NFAs (DFAs) augmented by a set of counters that can switch between increasing and decreasing a fixed number of times [10]).

### 3 Restrictions on Indexed Grammars

We first recall the definition of indexed grammar introduced in [1] by following [9], Section 14.3 (see also [4] for a reference book for grammars).

**Definition 2.** An indexed grammar is a 5-tuple  $G = (V, T, I, P, S)$ , where

- $V, T, I$  are finite pairwise disjoint sets: the set of variables, terminals, and indices, respectively;
- $P$  is a finite set of productions of the forms

$$\mathbf{1)} A \rightarrow \nu, \quad \mathbf{2)} A \rightarrow Bf, \quad \text{or} \quad \mathbf{3)} Af \rightarrow \nu,$$

where  $A, B \in V, f \in I$  and  $\nu \in (V \cup T)^*$ ;

- $S \in V$  is the start variable.

Let us now define the derivation relation  $\Rightarrow_G$  of  $G$ . Let  $\nu$  be an arbitrary sentential form of  $G$ ,

$$u_1 A_1 \alpha_1 u_2 A_2 \alpha_2 \cdots u_k A_k \alpha_k u_{k+1},$$

with  $A_i \in V, \alpha_i \in I^*, u_i \in T^*$ . For a sentential form  $\nu' \in (VI^* \cup T)^*$ , we set  $\nu \Rightarrow_G \nu'$  if one of the following three conditions holds:

- 1) In  $P$ , there exists a production of the form (1)  $A \rightarrow w_1 C_1 \cdots w_\ell C_\ell w_{\ell+1}$ ,  $C_j \in V, w_j \in T^*$ , such that in the sentential form  $\nu$ , for some  $i$  with  $1 \leq i \leq k$ , one has  $A_i = A$  and

$$\nu' = u_1 A_1 \alpha_1 \cdots u_i (w_1 C_1 \alpha_i \cdots w_\ell C_\ell \alpha_i w_{\ell+1}) u_{i+1} A_{i+1} \alpha_{i+1} \cdots u_k A_k \alpha_k u_{k+1}.$$

- 2) In  $P$ , there exists a production of the form (2)  $A \rightarrow Bf$  such that in the sentential form  $\nu$ , for some  $i$  with  $1 \leq i \leq k$ , one has  $A_i = A$  and  $\nu' = u_1 A_1 \alpha_1 \cdots u_i (Bf \alpha_i) u_{i+1} A_{i+1} \alpha_{i+1} \cdots u_k A_k \alpha_k u_{k+1}$ .

- 3) In  $P$ , there exists a production of the form (3)  $Af \rightarrow w_1 C_1 \cdots w_\ell C_\ell w_{\ell+1}$ ,  $C_j \in V, w_j \in T^*$ , such that in the sentential form  $\nu$ , for some  $i$  with  $1 \leq i \leq k$ , one has  $A_i = A, \alpha_i = f \alpha'_i, \alpha'_i \in I^*$ , and

$$\nu' = u_1 A_1 \alpha_1 \cdots u_i (w_1 C_1 \alpha'_i \cdots w_\ell C_\ell \alpha'_i w_{\ell+1}) u_{i+1} A_{i+1} \alpha_{i+1} \cdots u_k A_k \alpha_k u_{k+1}.$$

In this case, one says that the index  $f$  is consumed.

For every  $n \in \mathbb{N}$ ,  $\Rightarrow_G^n$  stands for the  $n$ -fold product of  $\Rightarrow_G$  and  $\Rightarrow_G^*$  stands for the reflexive and transitive closure of  $\Rightarrow_G$ . The language  $L(G)$  generated by  $G$  is the set  $L(G) = \{u \in T^* : S \Rightarrow_G^* u\}$ .

**Notation and Convention.** *In the sequel we will adopt the following notation and conventions for an indexed grammar  $G$ .*

- If no ambiguity arises, the relations  $\Rightarrow_G, \Rightarrow_G^n, n \in \mathbb{N}$ , and  $\Rightarrow_G^*$  will be simply denoted by  $\Rightarrow, \Rightarrow^n$ , and  $\Rightarrow^*$ , respectively.
- capital letters as  $A, B, \dots$  etc will denote variables of  $G$ .
- the small letters  $e, f$ , as well as  $f_i$ , will be used to denote indices while  $\alpha, \beta$  and  $\gamma$ , as well as its indexed version (as for instance  $\alpha_i$ ), will denote arbitrary words over  $I$ .
- Small letters as  $a, b, c, \dots$  etc (as well as its indexed version) will denote letters of  $T$  and small letters as  $u, v, w, r, \dots$ , etc (as well as its indexed version) will denote words over  $T$ .
- $\nu$  and  $\mu$ , as well as  $\nu_i$  and  $\mu_i$ , will denote arbitrary sentential forms of  $G$ .
- in order to shorten the notation, according to Definition 2, if  $p$  is a production of  $G$  of the form (1) or (3), we will simply write

$$Af \rightarrow \nu, \quad f \in I \cup \{\lambda\},$$

where it is understood that if  $f = \lambda$ , the production  $p$  has form (1) and if  $f \in I$ , the production  $p$  has form (3).

- If  $p_1 \cdots p_n \in P^*$  is a string of productions of  $G$ , then  $\Rightarrow_{p_1 \cdots p_n}$  denotes a derivation of  $G$  of the form  $\nu_0 \Rightarrow_{p_1} \nu_1 \Rightarrow_{p_2} \cdots \Rightarrow_{p_n} \nu_n$ .

The following set of definitions defines the main objects studied in this draft. Let  $G$  be an indexed grammar and let  $L(G)$  be the language generated by  $G$ . The first definition is from [5].

**Definition 3.** We say that  $G$  is linear if the right side component of every production of  $G$  has at most one variable. A language  $L$  is said to be linear indexed if there exists a linear indexed grammar  $G$  such that  $L = L(G)$ .

**Definition 4.** Given an integer  $k \geq 1$ , a derivation  $\nu_0 \Rightarrow \nu_1 \Rightarrow \cdots \Rightarrow \nu_n$  of  $G = (V, T, I, P, S)$ , is said to be of index- $k$  if  $|\nu_i|_V \leq k$ , for all  $i$ ,  $0 \leq i \leq n$ .

**Definition 5.** Given an integer  $k \geq 1$ ,  $G$  is said to be of index- $k$  if, for every word  $u \in L(G)$ , there exists a derivation of  $u$  in  $G$  of index- $k$ .

A language  $L$  is said to be an indexed language of index- $k$  if there exists an indexed grammar  $G$  of index- $k$  such that  $L = L(G)$ . An indexed language  $L$  is said to be of finite-index if  $L$  is of index- $k$ , for some  $k$ .

**Definition 6.** An indexed grammar  $G$  is said to be uncontrolled index- $k$  if, for every derivation  $\nu_0 \Rightarrow \cdots \Rightarrow \nu_n$  generating  $u \in L(G)$ ,  $|\nu_i|_V \leq k$ , for all  $i$ ,  $0 \leq i \leq n$ .  $G$  is uncontrolled finite-index if  $G$  is uncontrolled index- $k$ , for some  $k$ . A language  $L$  is said to be an uncontrolled finite-index indexed language if there exists an uncontrolled finite-index grammar  $G$  such that  $L = L(G)$ .

*Remark 7.* It is worth noticing that, according to Definition 6, if  $G$  is a grammar of index- $k_1$ , then  $G$  is a grammar of index- $k_2$ , for every integer  $k_1 \leq k_2$ .

*Remark 8.* It is interesting to observe that Definition 6 corresponds, in the case of context-free grammars, to the definition of nonterminal bounded grammar (cf [8], Section 5.7). We recall that nonterminal bounded grammars are equivalent to ultralinear grammars and thus provide a characterisations of the family of languages that are accepted by Finite-Turn pushdown automata.

Finally let us denote by

- $\mathcal{L}(\text{IND}_{\text{LIN}})$  the family of linear indexed languages [5];
- $\mathcal{L}(\text{IND}_{\text{UFIN}})$  the family of uncontrolled finite-index indexed languages;
- $\mathcal{L}(\text{IND}_{\text{FIN}})$  the family of finite-index indexed languages.

A reminder that uncontrolled finite-index corresponds to breadth-bounded indexed grammars [18]. Therefore, the following is implied.

**Theorem 9.** [18]  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  is a semilinear full trio.

The family  $\mathcal{L}(\text{IND}_{\text{LIN}})$  has been introduced in [5] where results of algebraic and combinatorial nature characterize the structure of its languages. Recall that a linear indexed grammar  $G$  is said to be *right linear indexed* if, according to Definition 2, in every production  $p$  of  $G$  of the form (1) or (3), the right hand component  $\nu$  of  $p$  has the form  $\nu = u$ , or  $\nu = uB$ , where  $u \in T^*$ ,  $B \in V$ . In [1] (see also [5]), the following theorem has been proved:

**Theorem 10.** [5] *If  $L$  is an arbitrary language,  $L$  is context-free if and only if there exists a right linear indexed grammar  $G$  such that  $L = L(G)$ .*

From this, the following is evident.

**Proposition 11.**  $\mathcal{L}(\text{CFL}) \subset \mathcal{L}(\text{IND}_{\text{LIN}}) \subset \mathcal{L}(\text{IND}_{\text{UFIN}}) \subseteq \mathcal{L}(\text{IND}_{\text{FIN}})$ .

Indeed Theorem 10 provides the inclusion  $\mathcal{L}(\text{CFL}) \subseteq \mathcal{L}(\text{IND}_{\text{LIN}})$ . The inclusions  $\mathcal{L}(\text{IND}_{\text{LIN}}) \subseteq \mathcal{L}(\text{IND}_{\text{UFIN}}) \subseteq \mathcal{L}(\text{IND}_{\text{FIN}})$  come immediately from the definitions of the corresponding families. Now, for every  $k \geq 1$ , let  $L_k = \{w^k : w \in A^*\}$ . It is easy to construct a linear indexed grammar that generates  $L_2$  so that  $L_2 \in \mathcal{L}(\text{IND}_{\text{LIN}}) \setminus \mathcal{L}(\text{CFL})$  (cf, for instance, [5]). Moreover it is proved that  $L_4 \notin \mathcal{L}(\text{IND}_{\text{LIN}})$  (see [5], Theorem 3.8). On the other hand, it is easily shown that  $L_k \in \mathcal{L}(\text{IND}_{\text{UFIN}})$ ,  $k \geq 0$ .

Also, in [5], it is shown that for an alphabet  $T$ ,  $\$ \notin T$ , and  $A, B \subseteq T^*$ , if  $L = A\$B$  is a linear indexed language, then  $A$  or  $B$  is a context-free language. Then, let  $T = \{a, b, c\}$ ,  $A = \{a^n b^n c^n : n > 0\}$ , and  $B = \{a^n b^n c^n : n > 0\}$ . Then  $L = \{a^n b^n c^n \$ a^m b^m c^m : n, m > 0\}$ . But since both  $A$  and  $B$  are not context-free, then  $L$  must not be linear indexed.

Next, closure under union is addressed with a straightforward adaptation of the first part of the proof of Theorem 6.1 of [9].

**Lemma 12.** *The families  $\mathcal{L}(\text{IND}_{\text{FIN}})$  and  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  are closed under union.*

*Proof.* Let us prove the claim for the family  $\mathcal{L}(\text{IND}_{\text{FIN}})$ , the proof for  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  being similar. Let  $L_1$  and  $L_2$  be indexed languages of indices  $k_1$  and  $k_2$  respectively, and let  $G_1$  and  $G_2$  be grammars

$$G_1 = (V_1, T_1, I_1, P_1, S_1), \quad G_2 = (V_2, T_2, I_2, P_2, S_2),$$

such that  $L_1 = L(G_1)$  and  $L_2 = L(G_2)$ . Since we may rename variables and indices without changing the language generated, we assume that  $V_1 \cap V_2 = I_1 \cap I_2 = \emptyset$ . Moreover let  $S$  be a new variable not in  $V_1 \cup V_2$ .

Construct a new grammar  $G = (V, T, I, P, S)$ , where  $V = V_1 \cup V_2 \cup \{S\}$ ,  $I = I_1 \cup I_2$ , and  $P$  is equal to  $P = P_1 \cup P_2$ , plus the two productions  $S \rightarrow S_1$  and  $S \rightarrow S_2$ .

It is easily checked that  $L_1 \cup L_2 = L(G)$  and  $G$  is of index  $\max\{k_1, k_2\}$ .  $\square$

Next, we show that  $\mathcal{L}(\text{IND}_{\text{FIN}})$  is a full trio, and the result also holds for  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  as well (shown in [18]). We will prove the more general fact that they are closed under rational transductions. The proof is structured using a chain of lemmas.

**Lemma 13.**  *$\mathcal{L}(\text{IND}_{\text{UFIN}})$  and  $\mathcal{L}(\text{IND}_{\text{FIN}})$  are closed under morphisms.*

*Proof.* We will demonstrate the proof for  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  with the proof for  $\mathcal{L}(\text{IND}_{\text{FIN}})$  following similarly.

Let  $L \in \mathcal{L}(\text{IND}_{\text{UFIN}})$  and let  $G = (V, T, I, P, S)$  be an uncontrolled  $k$ -index indexed grammar such that  $L = L(G)$ . Let  $\varphi : T^* \rightarrow (T')^*$  be a morphism

where  $T$  and  $T'$  are two alphabets. Construct a new grammar  $G'$  by replacing each production of  $G$  of the form

$$Xf \rightarrow u_1 X_1 \cdots u_\ell X_\ell u_{\ell+1},$$

where  $f \in I \cup \{\lambda\}$ ,  $u_i \in T^*$ ,  $X, X_i \in V$ , by the production

$$Xf \rightarrow \varphi(u_1)X_1 \cdots \varphi(u_\ell)X_\ell \varphi(u_{\ell+1}).$$

It is easily verified that the resulting grammar  $G'$  satisfies  $\varphi(L) = L(G')$  and  $G'$  is an uncontrolled grammar.  $\square$

**Lemma 14.**  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  and  $\mathcal{L}(\text{IND}_{\text{FIN}})$  are closed under intersection with regular languages.

*Proof.* We will show the result for uncontrolled grammars, and the other result follows similarly.

In order to prove that  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  is closed under intersection with regular sets, the following Claim is needed.

**Claim.** Let  $G = (V, T, I, P, S)$  be a finite-index (resp. uncontrolled finite-index) indexed grammar and let  $L = L(G)$ . Then there exists a finite-index (resp. uncontrolled finite-index) indexed grammar  $G' = (V', T, I', P', S')$  generating  $L$  such that  $I' = I$  and the productions of  $P'$  are of the form:

$$1) A \rightarrow \nu, \quad 2) A \rightarrow Bf, \quad \text{or} \quad 3) Af \rightarrow \nu,$$

where  $A, B \in V'$ ,  $f \in I'$  and  $\nu \in (V' \cup T)^*$  is a word of the form

$$\nu = u, \quad \text{or} \quad \nu = uXZ, \quad \text{or} \quad \nu = uXv, \quad X, Z \in V', \quad u, v \in T^*.$$

**Proof of the Claim.** Let first assume that  $G$  has a sole production  $p$  of the form

$$A \rightarrow \nu = u_1 X_1 u_2 X_2 \cdots u_k X_k u_{k+1}, \quad k \geq 2, \quad A, X_i \in V, u_i \in T^*. \quad (1)$$

Define the following list of productions:

- i.  $A \rightarrow u_1 X_1 Z_1$
- ii. For every  $j = 1, \dots, k-2$ ,  $Z_j \rightarrow u_{j+1} X_{j+1} Z_{j+1}$
- iii.  $Z_{k-1} \rightarrow u_k X_k u_{k+1}$ ,

where  $Z_j$ , ( $j = 1, \dots, k-1$ ), are new variables not in  $V$ .

Remove the production (1) from  $P$ , add to  $P$  the list of productions defined at (i)-(ii)-(iii) above, and add to  $V$  the corresponding list of variables  $Z_j$ 's. Finally call  $G'$  the grammar obtained from  $G$  by using the previous transformation. We now observe that the unique derivation of  $G'$  where the productions defined above appear is the one that simulates  $p$ :

$$A \Rightarrow_{G'} u_1 X_1 Z_1 \Rightarrow_{G'} u_1 X_1 u_2 X_2 Z_2 \Rightarrow_{G'} \cdots \Rightarrow_{G'} u_1 X_1 u_2 \cdots u_{k-1} X_{k-1} Z_k \Rightarrow_{G'} \nu.$$

Moreover such derivation has index not larger than that of  $G$ . >From the latter remark, it is easily checked, by induction on the length of the derivations of  $G'$ , that  $G'$  has the same index of  $G$  and that  $L = L(G')$ .

The case of productions  $Af \rightarrow \nu, f \in I$  is similarly treated. If  $G$  has two or more productions of the form previously considered, the claim is obtained by iterating the previous argument.  $\diamond$

Let  $G = (V, I, T, P, S)$  be an uncontrolled finite-index indexed grammar in the form given by the previous Claim. Let  $\mathcal{A} = (Q, T, \lambda, q_0, K)$  be a finite deterministic automaton accepting  $R$ , where  $Q$  is the set of states of  $\mathcal{A}$ ,  $\lambda : Q \times T \rightarrow Q$  is its transition function,  $q_0 \in Q$  is its unique initial state while  $K$  is the set of final states of  $\mathcal{A}$ . In the sequel, for the sake of simplicity, the extension of the function  $\lambda$  to the set  $Q \times T^*$  will be still denoted by  $\lambda$ .

We proceed to construct a new uncontrolled finite-index grammar  $G'$  such that  $G' = (V', I', T, P', S')$  and  $L(G') = L \cap R$ .

The set  $V'$  of variables of  $G'$  will be of the form  $\langle p, X, q \rangle$ , where  $p$  and  $q$  are in  $Q$  and  $X$  is in  $V$ , together with a new symbol  $S'$ , denoting the start variable of  $G'$ .

The set  $I'$  of indices of  $G'$  is a copy of  $I$  disjoint with it. For every index  $f$  of  $I$ , we will denote by  $f'$  the corresponding copy of  $f$  in  $I'$ .

The set  $P'$  of productions of  $G'$  is defined as follows.

1. If  $Af \rightarrow u$  is in  $P$ , where  $f \in I \cup \{\lambda\}$ ,  $u \in T^*$ , and  $\lambda(p, u) = q$ , then  $P'$  contains the set of productions  $\langle p, A, q \rangle f' \rightarrow u$ , for all  $p, q \in Q$ .
2. If  $A \rightarrow Bf$  is in  $P$ , where  $f \in I$ , then  $P'$  contains the set of productions

$$\langle p, A, q \rangle \rightarrow \langle p, B, q \rangle f',$$

where  $p, q$  are two arbitrary states of  $Q$ .

3. If  $Af \rightarrow vDw$  is in  $P$ , where  $f \in I \cup \{\lambda\}$ ,  $A, D \in V, v, w \in T^*$ , then  $P'$  contains, for all  $p, q, r, s \in Q$ , the set of productions

$$\langle p, A, q \rangle f' \rightarrow v \langle r, D, s \rangle w,$$

provided that  $\lambda(p, v) = r$ , and  $\lambda(s, w) = q$ .

4. If  $Af \rightarrow uBC$  is in  $P$ , where  $f \in I \cup \{\lambda\}$ ,  $A, B, C \in V, u \in T^*$ , then  $P'$  contains, for all  $p, q, r', r'' \in Q$ , the set of productions

$$\langle p, A, q \rangle f' \rightarrow u \langle r', B, r'' \rangle \langle r'', C, q \rangle,$$

provided that  $\lambda(p, u) = r'$ .

5. Finally  $P'$  contains the production  $S' \rightarrow \langle s_0, X, p \rangle$ , for all  $p \in K$ .

No other productions different from the form specified in the list above is in  $P'$ .

The first task is to show that  $L \cap R = L(G')$ . For this purpose, we first show that:  $\langle p, A, q \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^* u$ , with  $i \geq 0$ ,  $u \in T^*$ , if and only if  $Af_1 \cdots f_i \Rightarrow_G^* u$  and  $\lambda(p, u) = q$ . Indeed, from this statement, we get  $S' \Rightarrow_{G'} \langle s_0, S, q \rangle \Rightarrow_{G'}^* u$ ,



for some  $q \in K$ , if and only if  $S \Rightarrow_G^* u$ , and  $\lambda(s_0, u) = q$ , which is sufficient to complete the proof.

Let us first prove that:

(\*) If  $\langle p, A, q \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^\ell u$  is a derivation of  $G'$  of length  $\ell \geq 0$  then  $Af_1 \cdots f_i \Rightarrow_G^* u$  and  $\lambda(p, u) = q$ .

(\*) is easily checked to be true for derivations of length 1. Now suppose that (\*) is true for all  $m < \ell$  with  $m \geq 1$  and let  $\langle p, A, q \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^\ell u$  be a derivation of  $G'$  of length  $\ell$ . Such a derivation can be of one of the following forms.

(i)  $\langle p, A, q \rangle f'_1 \cdots f'_i \Rightarrow_{G'} \langle p, B, q \rangle f' f'_1 \cdots f'_i \Rightarrow_{G'}^{\ell-1} u$ ,

that is, the first production of the derivation has the form (2). By the inductive hypothesis, we then have  $Bf f_1 \cdots f_i \Rightarrow_G^* u$  and  $\lambda(p, u) = q$ , which yields  $Af_1 \cdots f_i \Rightarrow_G Bf f_1 \cdots f_i \Rightarrow_G^* u$  and  $\lambda(p, u) = q$ .

(ii)  $\langle p, A, q \rangle f' f'_1 \cdots f'_i \Rightarrow_{G'} v \langle r, D, s \rangle f'_1 \cdots f'_i w \Rightarrow_{G'}^{\ell-1} u$ ,  $f' \in I' \cup \{\lambda\}$ ,

that is, the first production of the derivation has the form (3). Set  $u = vu'w$ . From the latter, we get  $\langle r, D, s \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^{\ell-1} u'$  so that, by the inductive hypothesis,  $Df_1 \cdots f_i \Rightarrow_G^* u'$  and  $\lambda(r, u') = s$ . On the other hand, we know that

$$Af \Rightarrow_G vDw, \quad \lambda(p, v) = r, \quad \lambda(s, w) = q,$$

thus yielding  $Af f_1 \cdots f_i \Rightarrow_G vDf_1 \cdots f_i w \Rightarrow_G^* vu'w = u$ . Furthermore,  $\lambda(p, v) = r$ ,  $\lambda(s, w) = q$  which gives  $\lambda(p, u) = q$ .

(iii)  $\langle p, A, q \rangle f' f'_1 \cdots f'_i \Rightarrow_{G'} v \langle r', B, r'' \rangle f'_1 \cdots f'_i \langle r'', C, q \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^{\ell-1} u$ ,  $f' \in I' \cup \{\lambda\}$ ,  $r' = \lambda(p, v)$ ,

that is, the first production of the derivation has the form (4). Set  $u = vu'$ , with  $u' \in A^*$ . From the second sentential form, we get

$$\langle r', B, r'' \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^{\ell_1} u'_1, \quad \langle r'', C, q \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^{\ell_2} u'_2,$$

where  $u' = u'_1 u'_2$ , with  $u_1, u_2 \in A^*$ ,  $l_1 < l$ ,  $l_2 < l$ . By the inductive hypothesis, we have

$$Bf_1 \cdots f_i \Rightarrow_G^* u'_1, \quad Cf_1 \cdots f_i \Rightarrow_G^* u'_2,$$

together with

$$\lambda(r', u'_1) = r'', \quad \lambda(r'', u'_2) = q, \tag{2}$$

thus yielding

$$Af f_1 \cdots f_i \Rightarrow_G vBf_1 \cdots f_i Cf_1 \cdots f_i \Rightarrow_G^*$$

$$vu'_1 Cf_1 \cdots f_i \Rightarrow_G^* vu'_1 u'_2 = vu' = u.$$

Finally, from (2) and  $\lambda(p, v) = r'$ , we get  $\lambda(p, u) = q$ .

(iv)  $\langle p, A, q \rangle f' f'_1 \cdots f'_i \Rightarrow_{G'} wf'_1 \cdots f'_i \Rightarrow_{G'}^{\ell-1} u$ ,

that is, the first production of the derivation has the form (1). In this case,  $f'_1 = \cdots = f'_i = 1$ , and  $\ell = 1$  so that the claim is trivially proved.

Since the latter cases represent all the possible ways an arbitrary derivation can start, (\*) is proved. Similarly, one proves by induction on the length of a derivation in  $G$  that if  $Af_1 \cdots f_i \Rightarrow_G^\ell u$  is a derivation of  $G$  of length  $\ell \geq 0$  and  $\lambda(p, u) = q$  then  $\langle p, A, q \rangle f'_1 \cdots f'_i \Rightarrow_{G'}^* u$ . By the previous remark, this implies that  $L(G') = L(G) \cap R$ .

Finally it is checked that the grammars  $G$  and  $G'$  have the same index so that  $L(G')$  belongs to the family  $\mathcal{L}(\text{IND}_{\text{UFIN}})$ . This concludes the proof.  $\square$

Next, we show closure under a inverse morphisms of a specific type.

Let  $T$  and  $T'$  be two alphabets with  $T \subseteq T'$  and let  $\widehat{\pi}_T : (T')^* \rightarrow T^*$  be the projection of  $(T')^*$  onto  $T^*$ , that is the epi-morphism from  $(T')^*$  onto  $T^*$  generated by the mapping  $\pi_T : T' \rightarrow T \cup \{\lambda\}$

$$\forall \sigma \in T', \pi_T(\sigma) = \begin{cases} \lambda & \text{if } \sigma \notin T, \\ \sigma & \text{if } \sigma \in T. \end{cases}$$

In the sequel, for the sake of simplicity, we denote the projection  $\widehat{\pi}_T$  by  $\pi_T$ . It is useful to remark that, for every  $w \in T^*$  and  $w' \in (T')^*$ , with  $w = a_1 \cdots a_n$ ,  $n \geq 0$ ,  $a_i \in T$ ,

$$w' \in \pi_T^{-1}(w) \Leftrightarrow w' = w_1 a_1 \cdots w_n a_n w_{n+1}, \quad w_i \in (T' \setminus T)^*. \quad (3)$$

**Lemma 15.** *If  $L \in \mathcal{L}(\text{IND}_{\text{UFIN}})$  (resp.  $\mathcal{L}(\text{IND}_{\text{FIN}})$ ) with  $L \subseteq T^*$ , then  $\pi_T^{-1}(L) \in \mathcal{L}(\text{IND}_{\text{UFIN}})$  (resp.  $\mathcal{L}(\text{IND}_{\text{FIN}})$ ).*

*Proof.* Let  $G = (V, T, I, P, S)$  be an uncontrolled finite-index indexed grammar generating  $L$ . We construct a new uncontrolled finite-index grammar  $G'$  generating  $\pi_T^{-1}(L)$ .

For this purpose, let  $p = Xf \rightarrow \nu$ , with  $X \in V, f \in I \cup \{\lambda\}$ , and  $\nu \in (VI^* \cup T^*)^*$ , be a production of  $G$  of the form (1) or (3) (according to Definition 2). Then  $p$  has the form

$$Xf \rightarrow \nu = u_1 X_1 \cdots u_k X_k u_{k+1}, \quad u_i \in T^*,$$

where  $X, X_i \in V$ , with  $i = 1, \dots, k$ , and, for every  $i = 1, \dots, k+1$ ,

$$u_i = a_{i,1} \cdots a_{i,n_i}, \quad n_i \geq 0, \quad a_{i,j} \in T.$$

Let us associate with  $p$ , the following set of productions:

- $Xf \rightarrow Y_{1,0} \cdots Y_{k,0} Y_{k+1,0}$ ,
- $\forall i = 1, \dots, k+1, \forall j = 0, \dots, n_i - 1, \quad Y_{i,j} \rightarrow c Y_{i,j}, \quad c \in T' \setminus T$ ,
- $\forall i = 1, \dots, k+1, \forall j = 0, \dots, n_i - 1, \quad Y_{i,j} \rightarrow a_{i,j+1} Y_{i,j+1}$ ,
- $Y_{k+1,n_{k+1}} \rightarrow Y_{k+1,n_{k+1}} c, \quad c \in T' \setminus T$ ,
- $Y_{k+1,n_{k+1}} \rightarrow c, \quad c \in T' \setminus T$ ,
- $\forall i = 1, \dots, k, \quad Y_{i,n_i} \rightarrow X_i$ ,

where, for all  $i, j$ , with  $1 \leq i \leq k + 1$ , and  $0 \leq j \leq n_i$ ,  $Y_{i,j}$  are new variables not in  $V$ .

Now remove the production  $p$  from  $P$  and add respectively to  $P$  and  $V$  the productions defined above together with the corresponding set of new variables.

By applying the previous argument to every production  $p$  of the latter form, we will get a new grammar  $G' = (V', T', I', P', S')$ , where  $I' = I$ ,  $S' = S$  and the sets  $V'$  and  $P'$  are obtained from  $V$  and  $P$  respectively by iterating the latter combinatorial transformation.

It is useful now to remark that, in correspondence of every production  $Xf \rightarrow u_1 X_1 \cdots u_k X_k u_{k+1}$ , of  $G$  of the form (1) or (3), there exists a derivation of  $G'$  such that

$$Xf \Rightarrow_{G'}^* w_1 X_1 w_2 X_2 \cdots w_k X_k w_{k+1},$$

where, for all  $i = 0, \dots, k + 1$ ,  $w_i \in (T')^*$  and  $w_i \in \pi_{T'}^{-1}(u_i)$ , for all  $i$ .

Taking into account the latter argument and Eq. (3), by induction on the length of the derivations of  $G$  and  $G'$  respectively, one proves the following two claims:

- for every  $w' \in T'^*$ ,  $S' \Rightarrow_{G'}^* w'$  if and only if there exists a derivation of  $G$  such that  $S \Rightarrow_G^* w$ , with  $w \in T^*$ , and  $w' \in \pi_{T'}^{-1}(w)$ .
- if a non negative integer bounds the index of an arbitrary derivation of  $G$  the same does for  $G'$ . This implies that  $G'$  is an uncontrolled finite-index grammar.

This concludes the proof.  $\square$

Next, it is possible to show closure under rational transductions.

**Lemma 16.** *Let  $T$  and  $T'$  be two alphabets. Let  $\tau : T^* \rightarrow (T')^*$  be a rational transduction from  $T^*$  into  $(T')^*$ . If  $L$  is a language of  $T^*$  in the family  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  (resp.  $\mathcal{L}(\text{IND}_{\text{FIN}})$ ), then  $\tau(L) \in \mathcal{L}(\text{IND}_{\text{UFIN}})$  (resp.  $\mathcal{L}(\text{IND}_{\text{FIN}})$ ).*

*Proof.* We will show it for  $\mathcal{L}(\text{IND}_{\text{UFIN}})$ . Let us first assume that  $T \cap T' = \emptyset$ . By a well known theorem for the representation of rational transductions (see [3], Ch. III, Thm 4.1, see also [7]), there exists a regular set  $R$  over the alphabet  $(T \cup T')$  such that

$$\tau = \{(\pi_T(u), \pi_{T'}(u)) : u \in R\},$$

where  $\pi_T$  and  $\pi_{T'}$  are the projections of  $(T \cup T')^*$  onto  $T^*$  and  $T'^*$  respectively.

From the latter, one has that, for every  $u \in T^*$ ,  $\tau(u) = \pi_{T'}(\pi_T^{-1}(u) \cap R)$ , so that

$$\tau(L) = \bigcup_{u \in L} \tau(u) = \pi_{T'}(\pi_T^{-1}(L) \cap R). \quad (4)$$

Since, by hypothesis,  $L \in \mathcal{L}(\text{IND}_{\text{UFIN}})$ , the claim follows from (4), by applying Lemma 13, 14, and 15.

Let us finally treat the case where  $T$  and  $T'$  are not disjoint. Let  $T''$  be a copy of  $T$  with  $T'' \cap T' = \emptyset$  and let  $c_{T''} : (T')^* \rightarrow (T'')^*$  be the corresponding iso-morphism from  $(T')^*$  onto  $(T'')^*$ . By applying the latter argument to

the rational transduction  $c_{T''}\tau : T^* \rightarrow (T'')^*$ , one has  $(c_{T''}\tau)(L) \in \mathcal{L}(\text{IND}_{\text{UFIN}})$ . Since  $c_{T''}^{-1}(c_{T''}\tau)(L) = \tau(L)$ , then the claim follows from the latter by applying Lemma 13.  $\square$

Then the following is immediate:

**Corollary 17.**  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  and  $\mathcal{L}(\text{IND}_{\text{FIN}})$  are closed under inverse morphisms.

By Lemma 13, Lemma 14, and Corollary 17, we obtain:

**Theorem 18.** The families  $\mathcal{L}(\text{IND}_{\text{FIN}})$  and  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  are full trios.

We now prove a result which extends the semi-linearity of a family of languages to a bigger family. If  $\mathcal{C}$  is a full trio of semilinear languages and  $\mathcal{L}$  is the family of languages accepted by NCMs, let  $\mathcal{CL} = \{L_1 \cap L_2 : L_1 \in \mathcal{C}, L_2 \in \mathcal{L}\}$ .

**Proposition 19.** Every language in  $\mathcal{CL}$  has a semilinear Parikh map.

*Proof.* Let  $A$  and  $B$  be disjoint alphabets. Define a homomorphism

$$h : (A \cup B)^* \rightarrow A^*$$

by  $h(a) = a$  for each  $a \in A$ , and  $h(b) = \lambda$  for each  $b \in B$ . If  $L$  is a language over  $A^*$ , then  $h^{-1}(L) = \{x : x \in (A \cup B)^*, h(x) \in L\}$ .

Let  $A = \{a_1, \dots, a_n\}$  and  $L_1 \subseteq A^*$  be in  $\mathcal{C}$ . Then  $h^{-1}(L_1)$  is also in  $\mathcal{C}$ , since  $\mathcal{C}$  is closed under inverse homomorphism. Note that the Parikh map of  $L_1$ ,  $\psi(L_1)$ , is semi-linear since  $\mathcal{C}$  is a semi-linear family.

Now let  $L_2 \subseteq A^*$  be a language accepted by an NCM. Clearly, any NCM can be simulated by an NCM  $M_2$  whose counters are 1-reversal. We may assume that a string is accepted by  $M_2$  if and only if it enters a unique halting state  $f$  with all counters zero.

Let  $M_2$  have  $k$  1-reversal counters. Let  $B = \{p_1, q_1, \dots, p_k, q_k\}$  be new symbols disjoint from  $A$ . Construct an NFA  $M_3$  which when given a string  $w$  in  $(A \cup B)^*$  simulates  $M_2$ , but whenever counter  $c_i$  increments,  $M_3$  reads  $p_i$ . When  $M_2$  decrements counter  $c_i$ ,  $M_3$  reads  $q_i$ . (Note that after the first  $q_i$  is read, no  $p_i$  should appear on the remaining input symbols.)  $M_3$  guesses when each counter  $c_i$  becomes zero (this may be different time for each  $i$ ), after which,  $M_3$  should no longer read  $q_i$ . At some point,  $M_3$  guesses that all counters are zero. It continues the simulation and when  $M_2$  accepts in state  $f$ ,  $M_3$  accepts. Clearly, a string  $x$  in  $A^*$  is accepted by  $M_2$  if and only if there is a string  $w$  in  $(A \cup B)^*$  accepted by  $M_3$  such that:

- (1)  $h(w) = x$ ,
- (2)  $|w|_{p_i} = |w|_{q_i}$  for each  $1 \leq i \leq k$ .

Let  $R_3$  be the regular set accepted by  $M_3$ . Since  $\mathcal{C}$  is a full trio:

$$h^{-1}(L_1) \in \mathcal{C}, \quad L_4 = (h^{-1}(L_1) \cap R_3) \in \mathcal{C}.$$

Hence the Parikh map of  $L_4, \psi(L_4)$ , is a semi-linear set  $Q_4$ .

Now  $A = \{a_1, \dots, a_n\}$  and  $B = \{p_1, q_1, \dots, p_n, q_n\}$ . Define the semi-linear set  $Q_5 = \{(s_1, \dots, s_n, t_1, t_1, \dots, t_k, t_k) : s_i, t_i \geq 0\}$ . (Note that the first  $n$  coordinates refer to the counts corresponding to symbols  $a_1, \dots, a_n$ , and the last  $2k$  coordinates refer to the counts corresponding to symbols  $(p_1, q_1, \dots, p_k, q_k)$ .)

Then  $Q_6 = Q_4 \cap Q_5$  is semi-linear, since semi-linear sets are closed under intersection. Now  $\psi(L_1 \cap L_2)$  coincides with the projection of  $Q_6$  on the first  $n$  coordinates. Hence  $\psi(L_1 \cap L_2)$  is semi-linear, since semi-linear sets are closed under projections.  $\square$

Note that the above proposition does not depend on how the languages in  $\mathcal{C}$  are specified. It extends the semi-linearity of languages in  $\mathcal{C}$  to a bigger family that can do some “counting”. The proposition applies to all well-known full trios of semilinear languages, in particular, to  $\mathcal{C} = \mathcal{L}(\text{IND}_{\text{UFIN}})$ .

**Corollary 20.** *Let  $\mathcal{C}$  be a full trio whose closures under homomorphism, inverse homomorphism and intersection with regular sets are effective. Moreover, assume that for each  $L$  in  $\mathcal{C}$ ,  $\psi(L)$  can effectively be constructed. Then  $\mathcal{C}\mathcal{L}$  has a decidable emptiness problem.*

Note that  $\mathcal{L}$  is also a full trio of semilinear languages. It is easy to see that the proposition is not true if  $\mathcal{L}$  is an arbitrary full trio of semilinear languages. For example suppose  $\mathcal{C} = \mathcal{L}$  is the family of languages accepted by 1-reversal NPDAs (= linear context-free languages). Let

$$L_1 = \{a^{n_1}\#\dots\#a^{n_k}\$a^{n_k}\#\dots\#a^{n_1} : k \geq 4, n_i \geq 1\},$$

$$L_2 = \{a^{n_1}\#\dots\#a^{n_k}\$a^{m_k}\#\dots\#a^{m_1} : k \geq 4, n_i, m_i \geq 1, m_j = n_{j+1}, 1 \leq j < k\}.$$

Clearly,  $L_1$  and  $L_2$  can be accepted by 1-reversal NPDAs. But  $\psi(L_1 \cap L_2)$  is  $\{(a^n\#)^{k-1}a^n\$(a^n\#)^{k-1}a^n : n \geq 1, k \geq 4\}$  and it is not semilinear.

Similarly, it is easy to show that the proposition does not hold when  $\mathcal{C} = \mathcal{L}$  is the family of languages accepted by NFAs with one unrestricted counter (i.e., NPDAs with a unary stack alphabet in addition to a distinct bottom of the stack symbol which is never altered).

Finally, let  $\mathcal{C}_1$  and  $\mathcal{C}_2$  be any full trios of semilinear languages. It is clear that  $\mathcal{C}_1 + \mathcal{C}_2 = \{L_1 \cup L_2 : L_1 \in \mathcal{C}_1, L_2 \in \mathcal{C}_2\}$  is a semilinear family. One can also show that  $\mathcal{C}_1 \cdot \mathcal{C}_2 = \{L_1 L_2 : L_1 \in \mathcal{C}_1, L_2 \in \mathcal{C}_2\}$  is a semilinear family.

## 4 Bounded Languages and Hierarchy Results

The purpose of this section is to demonstrate that all bounded Ginsburg semi-linear languages are in  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  (thus implying they are in  $\mathcal{L}(\text{IND}_{\text{FIN}})$  as well), but not  $\mathcal{L}(\text{IND}_{\text{LIN}})$ .

Notice that the language  $L$  from Proposition 11 is a bounded Ginsburg semi-linear language. Thus, the following is true:

**Proposition 21.** *There are bounded Ginsburg semi-linear languages that are not in  $\mathcal{L}(\text{IND}_{\text{LIN}})$ .*

Furthermore, it has been shown that in every semi-linear full trio, all bounded languages in the family are bounded Ginsburg semi-linear [12]. Further,  $\mathcal{L}(\text{IND}_{\text{LIN}})$  is a semi-linear full trio [5]. Therefore, the bounded languages in  $\mathcal{L}(\text{IND}_{\text{LIN}})$  are strictly contained in the bounded languages contained in any family containing all bounded Ginsburg semi-linear languages. We only mention here three of the many such families mentioned in [12].

**Corollary 22.** *The bounded languages in  $\mathcal{L}(\text{IND}_{\text{LIN}})$  are strictly contained in the bounded languages from  $\mathcal{L}(\text{NCM}), \mathcal{L}(\text{DCM}), \mathcal{L}(\text{ETOL}_{\text{FIN}})$ .*

**Proposition 23.**  *$\mathcal{L}(\text{IND}_{\text{UFIN}})$  contains all bounded Ginsburg semi-linear languages.*

*Proof.* We now prove that if  $L$  is a bounded Ginsburg semi-linear language, with  $L \subseteq u_1^* \cdots u_k^*$ , then  $L \in \mathcal{L}(\text{IND}_{\text{UFIN}})$ . Since  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  is closed under union by Lemma 12, it is enough to show it for a linear set  $B$ . Let  $B$  be a set of the form  $B = \{\mathbf{b}_0 + x_1 \mathbf{b}_1 + \cdots + x_\ell \mathbf{b}_\ell : x_1, \dots, x_\ell \in \mathbb{N}\}$ , where  $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_\ell$ , are vectors of  $\mathbb{N}^k$ . By denoting the arbitrary vector  $\mathbf{b}_i$  as  $(b_{i1}, \dots, b_{ik})$ , we write  $B$  as

$$\{(\mathbf{b}_{01} + x_1 \mathbf{b}_{11} + \cdots + x_\ell \mathbf{b}_{\ell 1}, \dots, \mathbf{b}_{0k} + x_1 \mathbf{b}_{1k} + \cdots + x_\ell \mathbf{b}_{\ell k}), : x_1, \dots, x_\ell \in \mathbb{N}\},$$

so that the language  $L = \varphi(B)$  becomes

$$u_1^{\mathbf{b}_{01} + x_1 \mathbf{b}_{11} + \cdots + x_\ell \mathbf{b}_{\ell 1}} u_2^{\mathbf{b}_{02} + x_1 \mathbf{b}_{12} + \cdots + x_\ell \mathbf{b}_{\ell 2}} \dots u_k^{\mathbf{b}_{0k} + x_1 \mathbf{b}_{1k} + \cdots + x_\ell \mathbf{b}_{\ell k}}, \quad (5)$$

where  $x_1, \dots, x_\ell \in \mathbb{N}$ . Let us now define an indexed grammar  $G$  such that  $L = L(G)$ . Let  $G = (V, T, I, P, S)$ , where

$$V = \{S, Y, X_1, \dots, X_k\}, \quad T = A, \quad I = \{e, f_1, f_2, \dots, f_\ell\},$$

and the set  $P$  of productions is the following:

1.  $P_{\text{start}} = (S \rightarrow Ye)$
2. For every  $j = 1, \dots, \ell$ ,  $P_j = (Y \rightarrow Yf_j)$
3.  $Q = (Y \rightarrow X_1 X_2 \cdots X_k)$
4. For every  $i = 1, \dots, k$  and for every  $j = 1, \dots, \ell$ ,

$$R_{i0} = (X_i e \rightarrow u_i^{\mathbf{b}_{0i}}), \quad R_{ij} = (X_i f_j \rightarrow u_i^{\mathbf{b}_{ji}} X_i).$$

Let us finally prove that  $L = L(G)$  and  $G$  is an uncontrolled grammar. Let us first show that  $L \subseteq L(G)$ . Let  $w \in L$ . By (5), there exist  $x_1, \dots, x_\ell \in \mathbb{N}$  such that

$$w = u_1^{\mathbf{b}_{01} + x_1 \mathbf{b}_{11} + \cdots + x_\ell \mathbf{b}_{\ell 1}} u_2^{\mathbf{b}_{02} + x_1 \mathbf{b}_{12} + \cdots + x_\ell \mathbf{b}_{\ell 2}} \dots u_k^{\mathbf{b}_{0k} + x_1 \mathbf{b}_{1k} + \cdots + x_\ell \mathbf{b}_{\ell k}}.$$

Consider the derivation defined by the word over the alphabet  $P$ :

$$\mathcal{P} = P_{\text{start}} P_1^{x_1} P_2^{x_2} \cdots P_\ell^{x_\ell} Q Q_1 \cdots Q_k,$$

where, for every  $i = 1, \dots, k$ ,  $Q_i = R_{i\ell}^{x_\ell} \cdots R_{i2}^{x_2} R_{i1}^{x_1} R_{i0}$ . It is easily checked that  $S \Rightarrow_{\mathcal{P}} w$ . Indeed,

$$\begin{aligned}
 S &\Rightarrow_{P_{start}} Y e \Rightarrow_{P_1^{x_1} P_2^{x_2} \dots P_\ell^{x_\ell}} Y f_\ell^{x_\ell} \dots f_1^{x_1} e \Rightarrow_Q \\
 X_1 f_\ell^{x_\ell} \dots f_1^{x_1} e \dots X_k f_\ell^{x_\ell} \dots f_1^{x_1} e &\Rightarrow_{Q_1} u_1^{\mathbf{b}_{01} + x_1 \mathbf{b}_{11} + \dots + x_\ell \mathbf{b}_{\ell 1}} X_2 \dots X_k f_\ell^{x_\ell} \dots f_1^{x_1} e \\
 &\Rightarrow_{Q_2} u_1^{\mathbf{b}_{01} + x_1 \mathbf{b}_{11} + \dots + x_\ell \mathbf{b}_{\ell 1}} u_2^{\mathbf{b}_{02} + x_1 \mathbf{b}_{12} + \dots + x_\ell \mathbf{b}_{\ell 2}} X_3 \dots X_k f_\ell^{x_\ell} \dots f_1^{x_1} e \Rightarrow_{Q_3 \dots Q_k} w,
 \end{aligned}$$

so that  $w \in L(G)$ . Similarly, it can be shown that  $L(G) \subseteq L$ .  $\square$

Since it is known that in any semi-linear full trio, all bounded languages in the family are bounded Ginsburg semi-linear, the bounded languages in  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  coincide with several other families, including a deterministic machine model [12].

**Corollary 24.** *The bounded languages in  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  coincides with the bounded Ginsburg semi-linear languages, which coincides with the bounded languages in  $\mathcal{L}(\text{NCM}), \mathcal{L}(\text{DCM}), \mathcal{L}(\text{ETOL}_{\text{FIN}})$  (and several other families listed in [12]).*

Also, since  $\mathcal{L}(\text{IND}_{\text{LIN}})$  does not contain all bounded Ginsburg semi-linear languages by Proposition 21, but  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  does, the following is immediate:

**Corollary 25.**  $\mathcal{L}(\text{IND}_{\text{LIN}}) \subset \mathcal{L}(\text{IND}_{\text{UFIN}})$ .

Next, a restriction of  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  is studied and compared to the other families. And indeed, this family is quite general as it contains all bounded Ginsburg semilinear languages in addition to some languages that are not in  $\mathcal{L}(\text{ETOL}_{\text{FIN}})$ .

Now let  $p = (Af \rightarrow \nu) \in P$ , with  $f \in I \cup \{\lambda\}$ , be a production. Then  $p$  is called *special* if the number of occurrences of variables of  $V$  in  $\nu$  is at least 2, and *linear*, otherwise. Denote by  $P_S$  and  $P_L$  the sets of special and linear productions of  $P$  respectively. By Definition 6, a grammar  $G$  is uncontrolled finite-index if and only if the number of times special productions appear in every successful derivation of  $G$  is upper bounded by a given fixed integer.

Next, we will deal with uncontrolled grammars such that in every successful derivation of  $G$ , at most one special production occurs. The languages generated by such grammars form a family denoted  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$ . It is worth noticing that a careful rereading of the proof of Theorem 18 and Lemma 12 shows that they hold for  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$  as well. Further, it is clear that only one special production is used in every derivation of a word in the proof of Proposition 23. Therefore, the following holds:

**Proposition 26.** *The family  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$  is a union-closed full trio and it contains all bounded Ginsburg semilinear languages.*

It is immediate from the definitions that  $\mathcal{L}(\text{IND}_{\text{LIN}}) \subseteq \mathcal{L}(\text{IND}_{\text{UFIN}_1}) \subseteq \mathcal{L}(\text{IND}_{\text{UFIN}})$ . Further, since  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$  contains all bounded Ginsburg semilinear languages by Proposition 26, but the linear indexed languages do not, by Proposition 21, the following holds:

**Proposition 27.**  $\mathcal{L}(\text{IND}_{\text{LIN}}) \subset \mathcal{L}(\text{IND}_{\text{UFIN}_1}) \subseteq \mathcal{L}(\text{IND}_{\text{UFIN}})$ .

Then the following is true from [12].

**Corollary 28.**  *$\mathcal{L}(\text{IND}_{\text{UFIN}_1})$  is a semilinear full trio containing all bounded Ginsburg semilinear languages. Further, the bounded languages in  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$  coincides with the bounded languages in  $\mathcal{L}(\text{IND}_{\text{UFIN}}), \mathcal{L}(\text{NCM}), \mathcal{L}(\text{DCM}), \mathcal{L}(\text{ETOL}_{\text{FIN}})$ , and several others listed in [12].*

## 5 Some Examples, Separation, and Decidability Results

We start this section by giving an example that clarifies previous results.

*Example 29.* Let  $L = \{a^n b^n c^n \$ a^n b^n c^n : n \in \mathbb{N}\}$ . If  $\varphi : \mathbb{N}^7 \rightarrow a^* b^* c^* \$ a^* b^* c^*$ , then  $L = \varphi(B)$ , where  $B = \{\mathbf{b}_0 + n\mathbf{b}_1 : n \in \mathbb{N}\}$ , with  $\mathbf{b}_0 = (0, 0, 0, 1, 0, 0, 0)$  and  $\mathbf{b}_1 = (1, 1, 1, 0, 1, 1, 1)$ . It is worth noticing that, by the discussion preceding Proposition 21,  $L$  is not a linear indexed grammar. We define an uncontrolled finite-index grammar  $G = (V, T, I, P, S)$  where  $V = \{S, Y, X_1, X_2, X_3, X_4, X_5, X_6, X_7\}$ ,  $T = \{a, b, c, \$\}$ ,  $I = \{e, f\}$ , and the set  $P$  of productions is:

$$P_{start} = S \rightarrow Ye, P = Y \rightarrow Yf, Q = Y \rightarrow X_1 X_2 \cdots X_7$$

$$\begin{aligned} X_1 f &\rightarrow aX_1 & X_2 f &\rightarrow bX_2 & X_3 f &\rightarrow cX_3 & X_4 f &\rightarrow X_4 & X_5 f &\rightarrow aX_5 & X_6 f &\rightarrow bX_6 \\ X_7 f &\rightarrow cX_7 & X_1 e &\rightarrow \varepsilon & X_2 e &\rightarrow \varepsilon & X_3 e &\rightarrow \varepsilon & X_4 e &\rightarrow \$ & X_5 e &\rightarrow \varepsilon \\ X_6 e &\rightarrow \varepsilon & X_7 e &\rightarrow \varepsilon. \end{aligned}$$

In general

$$\begin{aligned} S &\Rightarrow Ye \Rightarrow^n Yf^n e \Rightarrow X_1 X_2 X_3 X_4 X_5 X_6 X_7 f^n e = \\ &= X_1 f^n e X_2 f^n e X_3 f^n e X_4 f^n e X_5 f^n e X_6 f^n e X_7 f^n e \Rightarrow^* a^n b^n c^n \$ a^n b^n c^n. \end{aligned}$$

As the only freedom in derivations of  $G$  consists of how many times the rule  $P$  is applied and of trivial variations in order to perform the rules  $X_i f \rightarrow \sigma X_i, \sigma \in T \cup \{\varepsilon\}$ , it should be clear that  $L = L(G)$ .

It is known that decidability of several properties holds for semilinear trios where the properties are effective [11]. This is the case for  $\mathcal{L}(\text{IND}_{\text{UFIN}})$ , and also for  $\mathcal{L}(\text{IND}_{\text{LIN}})$  [5].

**Corollary 30.** *Containment, equality, membership, and emptiness are decidable for bounded languages in  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  and  $\mathcal{L}(\text{IND}_{\text{LIN}})$ .*

Lastly, it is known that  $\mathcal{L}(\text{ETOL}_{\text{FIN}})$  cannot generate some context-free languages [16], but all context-free languages can be generated by indexed linear grammars by Theorem 10, which are all in  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$ .

**Corollary 31.** *There are languages in  $\mathcal{L}(\text{IND}_{\text{UFIN}_1})$  and  $\mathcal{L}(\text{IND}_{\text{LIN}})$  that are not in  $\mathcal{L}(\text{ETOL}_{\text{FIN}})$ .*

We provide an example of language in  $\mathcal{L}(\text{IND}_{\text{FIN}})$  whose Parikh image is not a semi-linear set.

*Example 32.* We construct a grammar of index 3, which is not uncontrolled, that generates the language  $L = \{aba^2b \cdots a^n ba^{n+1} : n \geq 1\}$ . Let  $G = (V, T, I, P, S)$  be the grammar where  $V = \{S, A, B, X, X', X''\}$ ,  $T = \{a, b\}$ ,  $I = \{e, f, g\}$ , and the set of productions of  $G$  are defined as:

$$\begin{aligned} - p_0 &= S \rightarrow Xe, & p_1 &= X \rightarrow ABX'f, & p_2 &= X' \rightarrow X, & p_3 &= X' \rightarrow X'', \\ - p_4 &= X''f \rightarrow aX'', & p_5 &= X''e \rightarrow a, & p_6 &= Af \rightarrow aA, & p_7 &= Ae \rightarrow \varepsilon, \\ - p_8 &= Bf \rightarrow B, & p_9 &= Be \rightarrow \varepsilon. \end{aligned}$$

One can check that  $G$  is not uncontrolled and  $L = L(G)$ .



**Corollary 33.** *There are languages in  $\mathcal{L}(\text{IND}_{\text{FIN}})$  that are not semilinear. Furthermore, there are bounded (and unary) languages in  $\mathcal{L}(\text{IND}_{\text{FIN}})$  that are not bounded Ginsburg semilinear.*

This allows for the separation of  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  (which only contains semi-linear languages) and  $\mathcal{L}(\text{IND}_{\text{FIN}})$ .

**Corollary 34.**  $\mathcal{L}(\text{CFL}) \subset \mathcal{L}(\text{IND}_{\text{LIN}}) \subset \mathcal{L}(\text{IND}_{\text{UFIN}_1}) \subseteq \mathcal{L}(\text{IND}_{\text{UFIN}}) \subset \mathcal{L}(\text{IND}_{\text{FIN}})$ .

Finally, we show that all finite-index ETOL languages are finite-index.

**Proposition 35.**  $\mathcal{L}(\text{ETOL}_{\text{FIN}}) \subset \mathcal{L}(\text{IND}_{\text{FIN}})$ .

*Proof.* Strictness follows since  $\mathcal{L}(\text{IND}_{\text{FIN}})$  contains non-semilinear languages by Corollary 33, but  $\mathcal{L}(\text{ETOL}_{\text{FIN}})$  only contains semi-linear languages [14].

We refer to [14] for the formal definitions of ETOL systems and finite-index ETOL systems, which we will omit.

Let  $G = (V, \mathcal{P}, S, T)$  be a  $k$ -index ETOL system. We can assume without loss of generality that  $G$  is in so-called active-normal form, so that the set of active symbols of  $V$  (those that can be changed by some production table) is equal to  $V \setminus T$ . Let  $\mathcal{P} = \{f_1, \dots, f_r\}$  be the set of production tables. Then create an indexed grammar  $G' = (V', T, I, P, S')$  where  $V' = (V \setminus T) \cup \{S'\}$ ,  $S'$  is a new variable,  $I = \{f_1, \dots, f_r\}$ , and  $P$  contains the following productions:

1.  $S' \rightarrow S' f_i, \forall i, 1 \leq i \leq r$ ,
2.  $S' \rightarrow S$ ,
3.  $B f_i \rightarrow \nu, \forall (B \rightarrow \nu) \in f_i, B \in V \setminus T$ .

Let  $w \in L(G)$ . Then  $w_0 \Rightarrow_{f_{j_1}} w_1 \Rightarrow \dots \Rightarrow_{f_{j_l}} w_l, w_0 = S, w_l = w$ . Let  $w'_i$  be obtained from  $w_i$  by placing  $f_{j_{i+1}} \dots f_{j_l}$  after each variable of  $w_i$ .

We will show by induction on  $i, 0 \leq i \leq l$ , that  $S' \Rightarrow_{G'}^* w'_i$ . Indeed,  $S' \Rightarrow_{G'}^* S f_{j_1} \dots f_{j_l} = w'_0$ , by using productions of type 1 followed by 2. Assume the inductive hypothesis is true for some  $i, 0 \leq i < l$ . Then  $S' \Rightarrow_{G'}^* w'_i$ . Then the next index on every variable of  $w'_i$  is  $f_{j_{i+1}}$ . Applying the corresponding productions used in the derivation  $w_i \Rightarrow_{f_{j_{i+1}}} w_{i+1}$  in table  $f_{j_{i+1}}$  on each variable of  $w'_i$  one at a time from left-to-right created in step 3,  $w'_{i+1}$  is obtained. It is also clear that if the original derivation is of index- $k$ , then the resulting derivation is of index- $2k$  (since the derivation of the indexed grammar proceeds sequentially instead of in parallel, the number of variables of the indexed grammar could potentially be more than  $k$ , but it is always less than the number of variables in the sentential form of the ETOL system plus the next sentential form).

Let  $w \in L(G')$ . Thus,  $w_0 \Rightarrow_{p_1} w_1 \Rightarrow_{p_2} \dots \Rightarrow_{p_l} w_l$ , where  $S' = w_0$  and  $w_l \in T^*$ . It should also be clear that we can assume without loss of generality that this derivation proceeds by rewriting variables in a “sweeping left-to-right” manner. That is, if  $w_i = w'_i B w''_i$  derives  $w_{i+1}$  by rewriting variable  $B$ , then  $w_{i+1}$  derives  $w_{i+2}$  by rewriting the first variable of  $w''_i$  if it exists, and if not, the first variable of  $w_{i+1}$ . Then one “sweep” of the variables by rewriting each variable is similar to one rewriting step of an ETOL system.

By the construction, there exists  $\alpha > 0$  such that  $p_1, \dots, p_\alpha$  are productions created in step 1,  $p_{\alpha+1}$  is created in step 2, and  $p_{\alpha+2}, \dots, p_l$  are created in step 3. Let  $\beta_1, \dots, \beta_q$  be such that  $\beta_1 = \alpha + 2$ , and the derivation from  $w_{\beta_i}$  is the start of the  $i$ th “sweep” from left-to-right, and let  $\beta_{q+1} = l$ . For  $1 \leq i \leq q + 1$ , let  $u_i$  be obtained from  $w_{\beta_i}$  by removing all indices (so  $u_{q+1} = w_l$ ).

We will show by induction that for all  $i$ ,  $1 \leq i \leq q + 1$ , then  $S \Rightarrow_G^* u_i$ , and all variables in  $w_{\beta_i}$  are followed by the same index sequence. Indeed,  $w_{\beta_1} = w_{\alpha+2} = S\gamma$  for some  $\gamma \in I^*$ ,  $u_1 = S$ , and  $S \Rightarrow_G^* u_1 = S$ . Assume that the inductive hypothesis holds for some  $i$ ,  $1 \leq i \leq q$ . Then in  $w_{\beta_i}$ , all variables are followed by the same index sequence. Let  $f$  be the first index following every variable. Then in the subderivation  $w_{\beta_i} \Rightarrow_{p_{\beta_i}} \dots \Rightarrow_{p_{\beta_{i+1}}} w_{\beta_{i+1}}$ , because all productions applied were created in step 3, they must all pop the first index, and since they all start with the same index, they must all have been created from productions in the same table  $f$ . It is clear that  $u_i \Rightarrow_G u_{i+1}$  using production table  $f$ . It is also immediate that all variables in  $w_{\beta_{i+1}}$  are followed by the same sequence of indices. The proof follows.  $\square$

It is an open question though as to how  $\mathcal{L}(\text{ETOL}_{\text{FIN}})$  compares to  $\mathcal{L}(\text{IND}_{\text{UFIN}})$ . For finite-index ETOL, uncontrolled systems, defined similarly to our definition of uncontrolled, does not restrict languages accepted. Furthermore, it is known that  $\mathcal{L}(\text{ETOL}_{\text{FIN}})$  is closed under Kleene-\* [14] and therefore contains  $\{a^n b^n c^n : n > 0\}^*$ . But we conjecture that this language is not in  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  despite being in  $\mathcal{L}(\text{IND}_{\text{FIN}})$  by the proposition above. This would imply that  $\mathcal{L}(\text{IND}_{\text{UFIN}})$  is incomparable with  $\mathcal{L}(\text{ETOL}_{\text{FIN}})$  by Corollary 31.

## References

1. A. V. Aho, Indexed grammars—an extension of context-free grammars, J. ACM, 15 (4), 647–671, (1968).
2. A. V. Aho, Nested stack automata, J. ACM, 16, 383–406, (1969).
3. J. Berstel, *Transductions and Context-Free Languages*, B.B. Teubner, Stuttgart, 1979.
4. J. Dassow, Gh. Păun, *Regulated Rewriting in Formal Language Theory*. EATCS Monographs on Theoretical Computer Science, 18. Springer-Verlag, Berlin, 1989.
5. J. Duske, R. Parchmann, Linear indexed grammars, Theoret. Comput. Sci. 32, 47–60, (1984).
6. G. Gazdar, Applicability of Indexed Grammars to Natural Languages, pp. 69–94, Springer Netherlands, Dordrecht, (1988).
7. S. Ginsburg, *The Mathematical Theory of Context-free Languages*, Mc Graw- Hill, New York, 1966.
8. M. A. Harrison, *Introduction to Formal Language Theory*, Addison-Wesley Publishing Co., Reading, Mass., 1978.
9. J. E. Hopcroft, J. D. Ullman, *Introduction to Automata Theory, Languages and Computation*, Addison-Wesley Publishing Co., Reading, Mass., 1979.
10. O. H. Ibarra, Reversal-bounded multicounter machines and their decision problems, J. ACM, 1 (25), 116–133, (1978).

11. O.H. Ibarra and I. McQuillan, On bounded semilinear languages, counter machines, and finite-index ETOL. In: Y. Han and K. Salomaa (eds.), *Lecture Notes in Computer Science*, 21st International Conference on Implementation and Application of Automata, CIAA 2016, Seoul, South Korea, vol. 9705, pp. 138–149, (2016).
12. O.H. Ibarra and I. McQuillan, On families of full trios containing counter machine languages. In: S. Brlek and C. Reutenauer (eds.), *Lecture Notes in Computer Science*, 20th International Conference on Developments in Language Theory, DLT 2016, Montreal, Canada, vol. 9840, pp. 216–228, (2016).
13. G. Rozenberg and A. Salomaa, *The Mathematical Theory of L Systems*, Academic Press, Inc., New York, 1980.
14. G. Rozenberg and D. Vermeir, On ETOL systems of finite index, *Information and Control*, 38, 103–133, (1978).
15. G. Rozenberg and D. Vermeir, On the effect of the finite index restriction on several families of grammars, *Information and Control*, 39, 284–302, (1978).
16. B. Rozoy, The Dyck language  $D_1^*$  is not generated by any matrix grammar of finite index, *Information and Computation* 74 (1), 64–89, (1987).
17. K. Vijay-Shanker, and D. J. Weir, The equivalence of four extensions of context-free grammars, *Mathematical Systems Theory*, 27 (6), 511–546, (1994).
18. G. Zetsche, An Approach to Computing Downward Closures. In: M.M. Halldórsson and K. Iwama and N. Kobayashi and B. Speckmann (eds.), *Automata, Languages, and Programming: 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part II*, pp. 440–451, (2015).