Acta Informatica manuscript No.
(will be inserted by the editor)

Gilles Geeraerts · **Jean-François Raskin** ·
Laurent Van Begin

Well-Structured Languages

the date of receipt and acceptance should be inserted later

Abstract This paper introduces the notion of *well-structured language*. A well-structured language can be defined by a *labelled well-structured transition system*, equipped with an *upward-closed set* of accepting states. That peculiar class of transition systems has been frequently studied in the field of *computer-aided verification*, where it has direct applications. Petri nets, and their monotonic extensions (like Petri nets with non-blocking arcs or Petri nets with transfer arcs), for instance, are special subclasses of well-structured transition systems.

We show that the class of well-structured languages enjoy several important closure properties. In order to establish these properties, we propose several pumping lemmata that are applicable respectively to the whole class of well-structured languages and to the classes of languages recognized by Petri nets or Petri nets with non-blocking arcs. These pumping lemmata also allow us to strictly separate the expressive power of Petri nets, Petri nets with non-blocking arcs and Petri nets with transfer arcs.

1 Introduction

In this paper, we study the family of languages defined by *well-structured (labelled) transition systems* (WSTS for short). WSTS [6] are transition systems whose state space is infinite but equipped with a well-quasi ordering (wqo for short) and whose transition relation is monotonic w.r.t. this wqo. WSTS have recently attracted a large interest in the community of *model-checking* because they

Laurent Van Begin has been supported by a "First Europe" grant EPH3310300R0012 of the Walloon Region.

This work has been partially supported by a FRFC grant number 2.4530.02 of the belgian National Fund for Scientifi c Research (FNRS).

Université Libre de Bruxelles – Département d'Informatique – CPI 212 – Boulevard du Triomphe, B-1050 Bruxelles BELGIUM. E-mail: {`gigeerae, jraskin, lvbegin`}`@ulb.ac.be`

enjoy nice decidability results and are useful to model important classes of systems (like parametric systems [4] and communication protocols [2]). In particular, the *coverability problem* (a variation of the reachability problem) has been shown decidable for the whole class of WSTS [1,6]. A large number of popular models define WSTS: Petri nets [11], monotonic extensions of Petri nets (e.g., Petri nets with transfer arcs [3]), lossy channel systems [2], broadcast protocols [4].

While the decidability properties of those models have been studied extensively (see, for example [6]), there are few known results about their expressive power in term of *recognized languages*. For example, several extensions of Petri nets have been proposed but their expressive power has not been studied and compared[1] so far.

In a previous paper [5], we have started to study the expressive power of monotonic extensions of Petri nets w.r.t. their ability to define sets of infinite words (omega languages). Unfortunately, the techniques that we had developed in that work were only applicable to omega languages. In the present paper, we generalize those techniques to make them applicable to the study of the expressive power of WSTS measured in term of definable sets of finite words. This classical measure allows us to compare the expressive power of WSTS with other well-studied formalisms like finite automata (defining regular languages), push-down automata (defining context free languages) or Turing machines (defining recursively enumerable languages). We propose proof techniques that intensively use basic properties of wqo. We believe that those proof techniques are interesting on their own.

The main contributions of our paper can be summarized as follows: (i) we define a natural class of languages recognized by WSTS for which the emptiness problem is decidable, (ii) we show that this class has important closure properties and forms an *Abstract Family of Languages* (AFL for short), (iii) to show the limits of the expressive power of WSTS, we introduce a general pumping lemma and show some examples of its possible applications, (iv) we study the relative expressive power of Petri nets and two important monotonic extensions of theirs. This study is made possible by two stronger pumping lemmata for these models.

The rest of this paper is structured as follows. In section 2, we recall some preliminaries about wqo, WSTS and (monotonic extensions of) Petri nets. In section 3, by considering different kinds of accepting conditions, we define three classes of languages recognized by WSTS, and we show that one of them has several interesting properties. That class is called the *well-structured languages* (WSL for short). In section 4, we propose a general pumping lemma applicable to any formalism that defines WSL. Two stronger versions of this lemma are defined and shown applicable to monotonic extensions of Petri nets. In section 5, we use the pumping lemmata to show the limits of WSL, some non-closure properties, and a strict hierarchy of expressive power among the monotonic extensions of Petri nets that we have considered.

[1] Some partial results are know about Petri nets, see for example [11].

2 Preliminaries

In this first section, we recall the main basic results that will be useful in the sequel of the paper. More precisely, we recall the classical notions of *languages*, *Abstract family of Languages* [7,13]. Then, we define *well-quasi orderings* and *well-structured transitions systems* that form the basis of our definition of well-structured languages. We close the section by recalling two classical models of computation. The first one is constituted by the (monotonic extensions of) *Petri nets*, whose languages are actually well-structured. The latter is the *two counter machine* [10], that we use in section 3 to prove undecidable some interesting properties about well-structured languages.

Languages and abstract family of languages Given an (finite) alphabet Σ, a (finite) word on Σ is either the empty word ε or a finite concatenation of symbols in Σ. A language on Σ is a (possibly infinite) set of words on Σ.

Let $\cdot$ denote the word concatenation. As usual $w \cdot \varepsilon = \varepsilon \cdot w = w$. The concatenation of two languages L_1 and L_2 is the language $L_1 \cdot L_2 = \{w_1 \cdot w_2 \mid w_1 \in L_1, w_2 \in L_2\}$. The iteration of a language L is the language $L^+ = \{w_1 \cdot \ldots \cdot w_n \mid n \geq 1 \wedge \forall 1 \leq i \leq n : w_i \in L\}$. Given a finite alphabet Σ, an *homomorphism* is a function $h : \Sigma^* \mapsto \Sigma^*$ s.t. $\forall w_1, w_2 \in \Sigma^* : h(w_1 \cdot w_2) = h(w_1) \cdot h(w_2)$. The inverse of h is the function $h^{-1} : \Sigma^* \mapsto 2^{\Sigma^*}$ such that $h^{-1}(w) = \{w' \mid h(w') = w\}$.

Definition 1 ([7,13]) A full *abstract family of languages* (full AFL for short) is a set of languages closed under (i) union, (ii) concatenation, (iii) intersection with regular languages, (iv) iteration, (v) homomorphism and (vi) inverse homomorphism.

Well-quasi orderings Well-quasi orderings are special cases of quasi orders that are the cornerstone of the definition of WSTS.

Definition 2 A *well quasi ordering* $\leq$ on C (wqo for short) is a *reflexive* and *transitive* relation s.t. for any infinite sequence $c_0, c_1, \ldots$ of elements in C, there are i and j, with $i < j$ and $c_i \leq c_j$.

In the sequel, we note $c_i < c_j$ iff $c_i \leq c_j$ but $c_j \not\leq c_i$. When a set C of elements is equipped with an ordering $\leq$, one can define the notion of *upward-closed set*. That notion will be useful in the sequel to define *accepting conditions* of languages of WSTS.

Definition 3 $\mathscr{U} \subseteq C$ is a $\leq$*-upward-closed set* if and only if: for any $c \in \mathscr{U}$, for any $c' \in C$ such that $c \leq c'$: $c' \in \mathscr{U}$.

Well-structured transition systems The transition systems have the characteristic that their set of configurations is ordered by a wqo $\leq$, and their transition relation is $\leq$-monotonic, as stated by the following definition:

Definition 4 A *(labelled) well-structured transition system* (WSTS for short) is a tuple $\langle C, c_0, \Sigma, \Rightarrow, \leq \rangle$ where:

- C is a (possibly infinite) set of configurations;

- $c_0 \in C$ is the initial configuration;
- Σ is a finite alphabet (that contains ε);
- $\Rightarrow \subseteq C \times \Sigma \times C$ is the transition relation;
- $\leq$ is a wqo for the elements of C.

Moreover, $\Rightarrow$ is *monotonic* w.r.t. to $\leq$, that is, for any c_1, c_2 and c_3 in C: if $(c_1, a, c_2) \in \Rightarrow$ and $c_1 \leq c_3$, then, there exist a finite sequence $\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_k \in C$ (with $k \geq 2$) and $1 \leq \ell < k$, such that:

- $\overline{c}_1 = c_3$;
- for any $1 \leq i < \ell$: $(\overline{c}_i, \varepsilon, \overline{c}_{i+1}) \in \Rightarrow$;
- $(\overline{c}_\ell, a, \overline{c}_{\ell+1}) \in \Rightarrow$;
- for any $\ell + 1 \leq i < k$: $(\overline{c}_i, \varepsilon, \overline{c}_{i+1}) \in \Rightarrow$;

In the sequel we often write $c_1 \stackrel{a}{\Rightarrow} c_2$ instead of $(c_1, a, c_2) \in \Rightarrow$. When the character labelling the transition is not relevant, we might omit it and write $c_1 \Rightarrow c_2$ to mean that there exists $a \in \Sigma$ s.t. $c_1 \stackrel{a}{\Rightarrow} c_2$. We also write $c \stackrel{w}{\Rightarrow} c'$ to mean that there exists a (finite) sequence of configurations $c_1, c_2, \ldots, c_n$ such that (i) $c \stackrel{a_0}{\Rightarrow} c_1 \stackrel{a_1}{\Rightarrow} c_2 \cdots c_n \stackrel{a_n}{\Rightarrow} c'$ and (ii) $w = a_0 \cdot a_1 \cdots a_n$ (thus, some of the a_i's may be ε).

For any configuration $c \in C$, let $\mathsf{PreUp}(c)$ be the set of all configurations whose one-step successors by $\Rightarrow$ are larger (w.r.t. $\leq$) than c i.e., $\mathsf{PreUp}(c) = \{c' \mid c' \Rightarrow c'', c \leq c''\}$. When both $\Rightarrow$ and $\leq$ are decidable, and when we can effectively compute $\mathsf{PreUp}(c)$, for any $c \in C$, the WSTS is called an *effective* WSTS (EWSTS for short) [1].

The following lemma is a direct consequence of the definition of wqo and will be useful in the sequel:

Lemma 1 *Given a set C with the well-quasi ordering $\leq \subseteq C \times C$ and an infinite sequence $S = c_1, c_2, \ldots$ with $\forall i \geq 1 : c_i \in C$, there exists an infinite subsequence $c_{i_1}, c_{i_2}, \ldots$ of S such that $\forall j \geq 1 : c_{i_j} \leq c_{i_{j+1}}$.*

Extended Petri nets In the sequel, we study in particular a subclass of EWSTS defined by Extended Petri Nets. We distinguish three subclasses of Extended Petri nets: the (regular) Petri nets, the Petri nets with non-blocking arcs and the Petri nets with transfer arcs. Those models are classically used to model parameterized systems [15].

A (labelled) *Extended Petri Net* (EPN) $\mathcal{N}$ is a tuple $\langle \mathcal{P}, \mathcal{T}, \Sigma, \mathbf{m}_0 \rangle$, where $\mathcal{P}$ is a finite set $\{p_1, p_2, \ldots, p_n\}$ of places, $\mathcal{T}$ is a finite set of transitions and Σ is a finite alphabet containing ε. A *marking* of the places is a function $\mathbf{m} : \mathcal{P} \mapsto \mathbb{N}$. A marking can also be seen as a vector v such that $v^T = [\mathbf{m}(p_1), \mathbf{m}(p_2), \ldots, \mathbf{m}(p_n)]$. $\mathbf{m}_0 : \mathcal{P} \mapsto \mathbb{N}$ is the initial marking. Each transition is of the form $\langle I, O, s, d, b, \lambda \rangle$, where I and $O : \mathcal{P} \mapsto \mathbb{N}$ are multi-sets of input and output places respectively. By convention, $O(p)$ (resp. $I(p)$) denotes the number of occurrences of p in O (resp. I). $s, d \in \mathcal{P} \cup \{\bot\}$ are the source and the destination places respectively, $b \in \mathbb{N} \cup \{+\infty\}$ is the bound and $\lambda \in \Sigma$ is the label of the transition. Let us divide $\mathcal{T}$ into $\mathcal{T}_r$ and $\mathcal{T}_e$ such that $\mathcal{T} = \mathcal{T}_r \cup \mathcal{T}_e$ and $\mathcal{T}_r \cap \mathcal{T}_e = \emptyset$. Without loss of generality, we assume that for each transition $\langle I, O, s, d, b, \lambda \rangle \in \mathcal{T}$, either $b = 0$ and $s = \bot = d$ (regular Petri transitions, grouped into $\mathcal{T}_r$); or $b > 0$, $s \neq d$, $s \neq \bot$ and $d \neq \bot$ (extended transitions, grouped into $\mathcal{T}_e$). We identify several non-disjoint classes of EPN, depending on $\mathcal{T}_e$:

1. *Petri nets* (PN for short): an EPN is a PN iff $\mathscr{T}_e = \emptyset$;
2. *Petri nets with non-blocking arcs* (PN+NBA): an EPN is a PN+NBA iff $\forall t = \langle I, O, s, d, b, \lambda \rangle \in \mathscr{T}_e : b = 1$;
3. *Petri nets with transfer arcs* (PN+T): an EPN is PN+T iff $\forall t = \langle I, O, s, d, b, \lambda \rangle \in \mathscr{T}_e : b = +\infty$.

Places are graphically depicted by circles; transitions by filled rectangles. For any transition $t = \langle I, O, s, d, b, \lambda \rangle$, we draw an arrow from any place $p \in I$ to transition t and from t to any place $p \in O$. For a PN+NBA (resp. PN+T), we draw a dotted (grey) arrow from s to t and from t to d (provided that $s, d \neq \bot$).

Given an extended Petri net $\mathscr{N} = \langle \mathscr{P}, \mathscr{T}, \Sigma, \mathbf{m}_0 \rangle$, and a marking $\mathbf{m}$ of $\mathscr{N}$, a transition $t = \langle I, O, s, d, b, \lambda \rangle$ is said to be *enabled in* $\mathbf{m}$ (notation: $\mathbf{m} \xrightarrow{t}$) iff $\forall p \in \mathscr{P} : \mathbf{m}(p) \geq I(p)$. An enabled transition $t = \langle I, O, s, d, b, \lambda \rangle$ can *occur*, which deterministically transforms the marking $\mathbf{m}$ into a new marking $\mathbf{m}'$ (we denote this by $\mathbf{m} \xrightarrow{t} \mathbf{m}'$). $\mathbf{m}'$ is computed as follows:

1. First compute $\mathbf{m}_1$ such that: $\forall p \in \mathscr{P} : \mathbf{m}_1(p) = \mathbf{m}(p) - I(p)$.
2. Then compute $\mathbf{m}_2$ as follows. If $s = d = \bot$, then $\mathbf{m}_2 = \mathbf{m}_1$. Otherwise:

$$\mathbf{m}_2(s) = \begin{cases} 0 & \text{if } \mathbf{m}_1(s) \leq b \\ \mathbf{m}_1(s) - b & \text{otherwise} \end{cases} \quad \mathbf{m}_2(d) = \begin{cases} \mathbf{m}_1(d) + \mathbf{m}_1(s) & \text{if } \mathbf{m}_1(s) \leq b \\ \mathbf{m}_1(d) + b & \text{otherwise} \end{cases}$$
$$\forall p \in \mathscr{P} \setminus \{d, s\} : \mathbf{m}_2(p) = \mathbf{m}_1(p)$$

3. Finally, compute $\mathbf{m}'$, such that $\forall p \in O : \mathbf{m}'(p) = \mathbf{m}_2(p) + O(p)$.

Let $\sigma = t_1 t_2 \ldots t_n$ be a sequence of transitions. We write $\mathbf{m} \xrightarrow{\sigma} \mathbf{m}'$ to mean that there exist $\mathbf{m}_1, \ldots, \mathbf{m}_{n-1}$ such that $\mathbf{m} \xrightarrow{t_1} \mathbf{m}_1 \xrightarrow{t_2} \ldots \xrightarrow{t_{n-1}} \mathbf{m}_{n-1} \xrightarrow{t_n} \mathbf{m}'$. Moreover, we let $\Lambda(\sigma) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$, where $\forall 1 \leq i \leq n$: λ_i is the label of t_i. We sometimes write $\mathbf{m} \xrightarrow{*} \mathbf{m}'$ to mean that there exists a sequence of transitions σ such that $\mathbf{m} \xrightarrow{\sigma} \mathbf{m}'$.

An EPN $\langle \mathscr{P}, \mathscr{T}, \Sigma, \mathbf{m}_0 \rangle$, defines a WSTS $S = \langle \mathbb{N}^{|\mathscr{P}|}, \mathbf{m}_0, \Sigma, \Rightarrow, \preccurlyeq \rangle$; where $\Rightarrow$ is such that $\mathbf{m}_1 \overset{a}{\Rightarrow} \mathbf{m}_2$ iff there is a transition $t \in \mathscr{T}$ with label a and $\mathbf{m}_1 \xrightarrow{t} \mathbf{m}_2$.

Example 1 Fig. 1 presents a transition $t = \langle I, O, s, d, +\infty, \mathtt{a} \rangle$ equipped with a transfer arc. I and O are such that : $I(p_1) = I(s) = 1$, $I(p_2) = I(d) = 0$, $O(p_2) = 1$ and $O(p_1) = O(s) = O(d) = 0$.

The successive steps to compute the effect of the firing of t are shown. Namely, (a) presents a marking $\mathbf{m}$ before the firing of t; (b) presents the marking $\mathbf{m}_1$ obtained by removing $I(p)$ tokens in every place p; (c) presents $\mathbf{m}_2$ obtained from $\mathbf{m}_1$ by transferring to d the two tokens present in s; and (d) presents the resulting marking $\mathbf{m}'$ obtained after producing $O(p)$ tokens in every place p.

If t had been equipped with a non-blocking arc (hence $t = \langle I, O, s, d, 1, \mathtt{a} \rangle$), only one token would have been transfered from s to d at step (c). In both cases, t would have been firable even if $\mathbf{m}_1(s)$ had been 0. ◇

Let $\preccurlyeq$ denote the wqo (see [1]) on markings, defined as follows: let $\mathbf{m}$ and $\mathbf{m}'$ be two markings on the set of places $\mathscr{P}$, then $\mathbf{m} \preccurlyeq \mathbf{m}'$ iff $\forall p \in \mathscr{P} : \mathbf{m}(p) \leq \mathbf{m}'(p)$. Since $\preccurlyeq$ is a wqo, we obtain the following property, useful in the sequel:

Lemma 2 *Given an infinite sequence of markings $\mathbf{m}_1, \mathbf{m}_2, \ldots$ we can always extract an infinite sub-sequence $\mathbf{m}_{i_1}, \mathbf{m}_{i_2}, \ldots$ ($\forall j : i_j < i_{j+1}$) s.t. for any place p, either $\mathbf{m}_{i_j}(p) < \mathbf{m}_{i_{j+1}}(p)$ for all $j \geq 1$ or $\mathbf{m}_{i_j}(p) = \mathbf{m}_{i_{j+1}}(p)$ for all $j \geq 1$.*

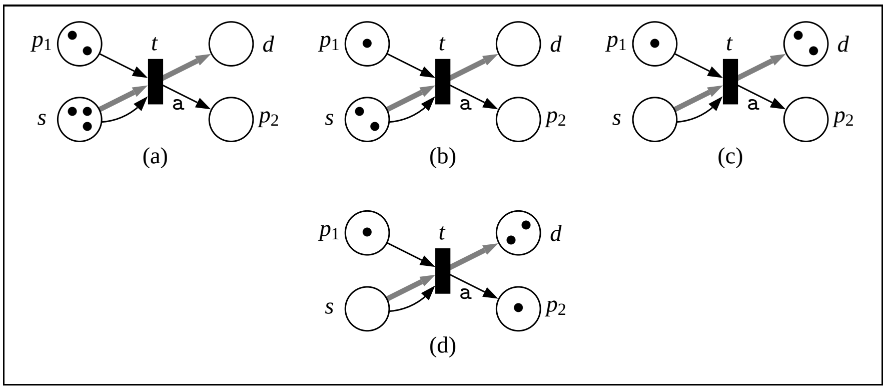

Fig. 1 The four steps to compute the effect of a transfer arc

Two counter machines Another classical model of computation is that of two counter machines. It is well-known to be as expressive as Turing machines are, and is therefore often used to prove the undecidability of relevant problems. Two counter machines have been introduced in [10].

Definition 5 A *two-counter machine* C (2CM for short) is a tuple $\langle c_1, c_2, L, \mathsf{Instr}\rangle$ where:

- c_1, c_2 are two counters taking their values in $\mathbb{N}$;
- $L = \{l_1, l_2, \ldots, l_u\}$ is a finite non-empty set of u locations;
- Instr is a function that labels each location $l \in L$ with an instruction that has one of the three following forms:
  - $l : c_j := c_j + 1;$ `goto` l';, where $j \in \{1,2\}$ and $l' \in L$, this is called an increment;
  - $l : c_j := c_j - 1;$ `goto` l';, where $j \in \{1,2\}$ and $l' \in L$, this is called a decrement;
  - l : `if` $c_j = 0$ `then goto` l' `else goto` l'';, where $j \in \{1,2\}$ and $l', l'' \in L$, this is called a zero-test.

Those instructions have their usual obvious semantics, in particular, decrement can only be done if the value of the counter is strictly greater than zero.

A *configuration* of a 2CM $\langle c_1, c_2, L, \mathsf{Instr}\rangle$ is a tuple $\langle loc, v^1, v^2\rangle$ where $loc \in L$ is the value of the program counter and v^1, respectively v^2, is a natural number that gives the valuation of the counter c_1, respectively c_2. A *computation* γ of a 2CM $\langle c_1, c_2, L, \mathsf{Instr}\rangle$ is a finite sequence of configurations

$$\langle loc_1, v_1^1, v_1^2\rangle, \langle loc_2, v_2^1, v_2^2\rangle, \ldots, \langle loc_r, v_r^1, v_r^2\rangle$$

that respects the two following conditions:

1. "Initialization": $loc_1 = l_1$, $v_1^1 = 0$, and $v_1^2 = 0$, i.e., a computation starts in l_1 and the two counters have the value zero;
2. "Consecution": for each $i \in \mathbb{N}$ such that $1 \le i < r$, $\langle loc_{i+1}, v_{i+1}^1, v_{i+1}^2\rangle$ is the configuration obtained from $\langle loc_i, v_i^1, v_i^2\rangle$ by applying the instruction $\mathsf{Instr}(loc_i)$.

Moreover, we let $\mathsf{final}(\gamma) = \langle loc_r, v_r^1, v_r^2 \rangle$. A configuration $\langle loc, v^1, v^2 \rangle$ is *reachable* in the 2CM $\langle c_1, c_2, L, \mathsf{Instr} \rangle$, if there exists a finite computation γ such that $\mathsf{final}(\gamma) = \langle loc, v^1, v^2 \rangle$.

A natural problem about 2CM is the *boundedness problem*:

Problem 1 Given a 2CM $C = \langle c_1, c_2, L, \mathsf{Instr} \rangle$, the *boundedness problem for* 2CM asks whether there is $c \in \mathbb{N}$ such that for all reachable configuration $\langle loc, v_1, v_2 \rangle$ of C, we have $v_1 + v_2 \leq c$.

It is well-known that this problem cannot be answered completely by an algorithm:

Theorem 1 ([10]) *The boundedness problems is undecidable for* 2CM.

3 Well-structured languages

This section is mainly devoted to the definitions of languages of WSTS (and the motivations of these definitions). In accordance to previous classical works on the expressive power of Petri nets (such as [11], [14] or [9], for instance), we distinguish several classes of languages of WSTS, depending on the form of the set of *accepting states*. Then, we study several properties of these different classes of languages. As we will see, the class one obtains when considering $\leq$-*upward-closed* sets of accepting states enjoys nice properties (the emptiness is decidable, that class forms a full AFL, closed under intersection) that do not hold if we chose, for instance, a finite set of accepting states. This will motivate our choice for the definition of *well-structured languages*. Unfortunately, the universality problem is undecidable for EWSTS. That result is proved by adapting a proof formerly published in [12].

3.1 Languages of WSTS

We first define the notion of language of a WSTS:

Definition 6 Given a WSTS $S = \langle C, c_0, \Sigma, \Rightarrow, \leq \rangle$, and a set $C' \subseteq C$ of *accepting configurations*, the *language* of S, noted $L(S, C')$ is the set of all the words w such that $c_0 \stackrel{w}{\Rightarrow} c$ for some $c \in C'$.

By imposing some well-chosen restrictions about the set of accepting configurations, one can obtain different classes of languages. In the restricted case of PN, this approach has already been followed in classical works of the literature such as [11], [14] or [9]. Namely, if $\mathscr{S}$ is a set of WSTS, then $L^L(\mathscr{S})$, $L^T(\mathscr{S})$ and $L^G(\mathscr{S})$ are the classes of languages defined by a WSTS in $\mathscr{S}$, and where the set of accepting configurations is (resp.) a *finite set* of configurations; the set of every *deadlock* configuration or; a $\leq$-*upward-closed set* of configurations.

Not surprisingly, these different classes of languages enjoy different properties, as shown by the following propositions. Proposition 1 states that $L^L(\mathsf{EWSTS})$ and $L^T(\mathsf{EWSTS})$ are both equal to the set of recursively enumerable languages (R.E.). This proposition stems from the fact that $L^L(\mathsf{PN{+}T}) = R.E.$, as shown in [3]. Hence the emptiness is undecidable on these classes.

Proposition 1 **([3])** $L^L(\mathsf{EWSTS}) = L^T(\mathsf{EWSTS}) = R.E.$

On the other hand, the emptiness is decidable for EWSTS with $\leq$-upward-closed accepting set. That result stems from the fact that the *coverability problem* is decidable on that class:

Problem 2 Given an EWSTS S and an upward-closed set $\mathscr{U}$ of configurations of S, the *coverability problem* asks whether there exists a configuration c that is reachable in S and that belongs to $\mathscr{U}$.

The proof of decidability of that problem can be found, for instance in [6]. From the definition of the problem, it is not difficult to see that, given an EWSTS S and an upward-closed set $\mathscr{U}$ of configurations of S, the language $L^G(S,\mathscr{U}) = \emptyset$ iff the answer to the coverability problem is *negative* on S and $\mathscr{U}$. This provides us with an effective procedure to test the emptiness of the language of an EWSTS when an upward-closed set of accepting configurations is considered. Hence, the Theorem:

Theorem 2 *The emptiness problem is decidable for the class of* EWSTS*, when we consider $\leq$-upward-closed accepting sets.*

As a direct consequence, one obtains:

Corollary 1 $L^G(\mathsf{EWSTS}) \neq$ R.E.

Finally, one can prove that $L^G(\mathsf{WSTS})$ is a full AFL closed under intersection, which is a strong indication that it is a class worth of attention.

Theorem 3 $L^G(\mathsf{WSTS})$ *is a full* AFL*, closed under intersection.*

Proof According to Definition 1, one has to show seven closure properties (the six properties that define an AFL, plus the closure under intersection) in order to establish this result. In the sequel, we assume that $S_1 = \langle C_1, i_1, \Sigma_1, \Rightarrow_1, \leq_1\rangle$ and $S_2 = \langle C_2, i_2, \Sigma_2, \Rightarrow_2, \leq_2\rangle$ are two WSTS, and that $\mathscr{U}_1$ and $\mathscr{U}_2$ are their associated upward-closed sets of accepting states. We also assume that $h : \Sigma_1 \mapsto \Sigma_1^*$ is an homomorphism s.t. $h(\varepsilon) = \varepsilon$, according to the definition from [7,13]. We prove the closure of the seven operations by showing building a WSTS $S = \langle C, i, \Sigma, \Rightarrow, \leq\rangle$ and a set of accepting states $\mathscr{U}$, s.t. $L^G(S,\mathscr{U})$ is the result of the operation in question. We ensure that $L^G(S,\mathscr{U})$ is a WSL by proving that $\leq$ is a wqo, $\Rightarrow$ is $\leq$-monotonic and $\mathscr{U}$ is upward-closed.

Intersection Let us build S and $\mathscr{U}$ s.t. $L^G(S,\mathscr{U}) = L^G(S_1,\mathscr{U}_1) \cap L^G(S_2,\mathscr{U}_2)$. The way we build S is described in the following: $C = C_1 \times C_2$; $i = (i_1, i_2)$; $\Sigma = \Sigma_1 \cap \Sigma_2$; $\Rightarrow = \{\big((c_1,c_2), a, (c_1',c_2')\big) \mid c_1 \overset{a}{\Rightarrow}_1 c_1' \wedge c_2 \overset{a}{\Rightarrow}_2 c_2'\}$; $\leq = \{\big((c_1,c_2),(c_1',c_2')\big) \mid c_1 \leq_1 c_1' \wedge c_2 \leq_2 c_2'\}$; and $\mathscr{U} = \{(c_1,c_2) \mid c_1 \in \mathscr{U}_1 \wedge c_2 \in \mathscr{U}_2\}$.

Clearly, $L^G(S,\mathscr{U}) = L^G(S_1,\mathscr{U}_1) \cap L^G(S_2,\mathscr{U}_2)$. Let us prove that $\leq$, $\Rightarrow$ and $\mathscr{U}$ have the desired properties:

– **$\leq$ is a wqo** Let $\varsigma = (c_1^1, c_1^2), (c_2^1, c_2^2), \ldots, (c_n^1, c_n^2), \ldots$ be an infinite sequence of elements of C. Since $\leq_1$ is a wqo on C_1, following Lemma 1, one can extract from ς a subsequence

$$\varsigma' = (c^1_{\rho(1)}, c^2_{\rho(1)}), (c^1_{\rho(2)}, c^2_{\rho(2)}), \ldots, (c^1_{\rho(n)}, c^2_{\rho(n)}), \ldots$$

such that for any $j \geq 1$: $c^1_{\rho(j)} \leq_1 c^1_{\rho(j+1)}$. Since $\leq_2$ is a wqo on the elements of C_2, there are, in ς', two positions k and ℓ s.t. $k < \ell$ and $c^2_{\rho(k)} \leq_2 c^2_{\rho(\ell)}$. Hence, $(c^1_{\rho(k)}, c^2_{\rho(k)}) \leq (c^1_{\rho(\ell)}, c^2_{\rho(\ell)})$, which proves that $\leq$ is a wqo, according to Definition 2.

- $\Rightarrow$ **is** $\leq$**-monotonic** Let (c^1_1, c^2_1), (c^1_2, c^2_2), and (c^1_3, c^2_3) be three configurations of C s.t. $(c^1_1, c^2_1) \overset{a}{\Rightarrow} (c^1_2, c^2_2)$ and $(c^1_1, c^2_1) \leq (c^1_3, c^2_3)$. By definition of $\Rightarrow$ and $\leq$, this implies that $c^1_1 \overset{a}{\Rightarrow}_1 c^1_2$, $c^2_1 \overset{a}{\Rightarrow}_2 c^2_2$, $c^1_1 \leq_1 c^1_3$ and $c^2_1 \leq_2 c^2_3$. Since $\Rightarrow_1$ and $\Rightarrow_2$ are resp. $\leq_1$- and $\leq_2$-monotonic, there are $c \in C_1$ and $c' \in C_2$ s.t.: $c^1_3 \overset{a}{\Rightarrow}_1 c$, $c^2_3 \overset{a}{\Rightarrow}_1 c'$, $c^1_2 \leq_1 c$ and $c^2_2 \leq_2 c'$. Hence $(c^1_3, c^2_3) \overset{a}{\Rightarrow} (c, c')$ and $(c^1_2, c^2_2) \leq (c, c')$.
- $\mathscr{U}$ **is** $\leq$**-upward-closed** Let (c^1_1, c^2_1) and (c^1_2, c^2_2), both in C, be s.t. $(c^1_1, c^2_1) \leq (c^1_2, c^2_2)$ and $(c^1_1, c^2_1) \in \mathscr{U}$. Let us show that and $(c^1_2, c^2_2) \in \mathscr{U}$ too. Since $(c^1_1, c^2_1) \in \mathscr{U}$, we have $c^1_1 \in \mathscr{U}_1$ and $c^2_1 \in \mathscr{U}_2$, by definition of $\mathscr{U}$. Since $(c^1_1, c^2_1) \leq (c^1_2, c^2_2)$, $c^1_1 \leq_1 c^1_2$ and $c^2_1 \leq_2 c^2_2$, by definition of $\leq$. But $\mathscr{U}_1$ and $\mathscr{U}_2$ are resp. $\leq_1$- and $\leq_2$-upward-closed, which implies that $c^1_2 \in \mathscr{U}_1$ and $c^2_2 \in \mathscr{U}_2$. Hence $(c^1_2, c^2_2) \in \mathscr{U}$.

Union Let us construct S and $\mathscr{U}$ such that $L^G(S, \mathscr{U}) = L^G(S_1, \mathscr{U}_1) \cup L^G(S_2, \mathscr{U}_2)$. We let $C = \{i\} \cup C_1 \cup C_2$ (where i is not in C_1 nor in C_2); $\Sigma = \Sigma_1 \cup \Sigma_2$; $\leq = \leq_1 \cup \leq_2 \cup \{(i,i)\}$; $\Rightarrow = \{(i, \varepsilon, i_1), (i, \varepsilon, i_2)\} \cup \Rightarrow_1 \cup \Rightarrow_2$ and $\mathscr{U} = \mathscr{U}_1 \cup \mathscr{U}_2$.

Clearly, $L^G(S, \mathscr{U}) = L^G(S_1, \mathscr{U}_1) \cup L^G(S_2, \mathscr{U}_2)$. Let us show that we have the right properties. By definition, $\Rightarrow$ is $\leq$-monotonic (remark that i is $\leq$-uncomparable to any other element of C). Thus, it remains to prove that:

- $\leq$ **is a wqo** Let $\varsigma = c_0, c_2, \ldots, c_n, \ldots$ be an infinite sequence of elements of C. Because it is infinite, one can extract, from that sequence, an infinite subsequence $\varsigma' = c_{j_1}, c_{j_2}, c_{j_3}, \ldots$, s.t. either $\forall k \geq 1 : c_{j_k} \in C_1$ or $\forall k \geq 1 : c_{j_k} \in C_2$. Since $\leq_1$ and $\leq_2$ are both wqo, there exist two positions k and ℓ s.t. $k < \ell$ and either $c_{j_k} \leq_1 c_{j_\ell}$ or $c_{j_k} \leq_2 c_{j_\ell}$. Hence $c_{j_k} \leq c_{j_\ell}$, which proves that $\leq$ is a wqo following Definition 2.
- $\mathscr{U}$ **is** $\leq$**-upward-closed** Let c_1, c_2 be two configurations in C s.t. $c_1 \in \mathscr{U}$ and $c_1 \leq c_2$. Let us show that $c_2 \in \mathscr{U}$. We consider two cases: either $c_1 \in \mathscr{U}_1$ or $c_1 \in \mathscr{U}_2$. In the former case, since c_1 and c_2 are $\leq$-comparable, we deduce that $c_2 \in C_1$ and thus, $c_1 \leq_1 c_2$, by definition of $\leq$. Hence, $c_2 \in \mathscr{U}_1$, since $\mathscr{U}_1$ is $\leq_1$-upward-closed. This implies that $c_2 \in \mathscr{U}$. In the latter case, we obtain the same conclusion by a similar reasoning

Concatenation Let us construct S and $\mathscr{U}$ such that $L^G(S, \mathscr{U}) = L^G(S_1, \mathscr{U}_1) \cdot L^G(S_2, \mathscr{U}_2)$. We let $C = C_1 \cup C_2$; $i = i_1$; $\Sigma = \Sigma_1 \cup \Sigma_2$; $\Rightarrow = \{(c, \varepsilon, i_2) \mid c \in \mathscr{U}_1\} \cup \Rightarrow_2 \cup \Rightarrow_1$; $\leq = \leq_1 \cup \leq_2$ and $\mathscr{U} = \mathscr{U}_2$.

Clearly, $L^G(S, \mathscr{U})$ is the concatenation of $L^G(S_1, \mathscr{U}_1)$ and $L^G(S_2, \mathscr{U}_2)$. Moreover, it is trivial to see that $\Rightarrow$ is $\leq$-monotonic (see previous property) and $\mathscr{U}$ is $\leq$-upward-closed. Finally, one can show that $\leq$ is a wqo by reusing the same reasoning as for the union.

Iteration Let us construct S such that $L^G(S, \mathscr{U}) = L^G(S_1, \mathscr{U}_1)^+$. We consider a new configuration $i_0 \notin C_1$ and let $C = C_1 \cup \{i_0\}$; $i = i_0$; $\leq = \leq_1 \cup \{(i_0, i_0)\}$; $\Rightarrow = \{(i_0, \varepsilon, i_1)\} \cup \{(c, \varepsilon, i_0) \mid c \in \mathscr{U}_1\} \cup \Rightarrow_1$ and $\mathscr{U} = \mathscr{U}_1$.

From these definitions, it is trivial to see that $L^G(S, \mathscr{U}) = L^G(S_1, \mathscr{U}_1)^+$ and that $\leq$, $\Rightarrow$ and $\mathscr{U}$ enjoy the desired properties.

Intersection with regular languages It is not difficult to see that any deterministic finite-state automaton is a WSTS, when we chose the equality between states as wqo. Hence, any regular language is a WSL. Since WSL are closed under intersection (see above), the closure with regular languages holds too.
Arbitrary homomorphism Let us construct S and $\mathscr{U}$ such that $L^G(S,\mathscr{U}) = h\big(L^G(S_1,\mathscr{U}_1)\big)$. We extend the set of states C_1 with elements from $C_1 \times \Sigma \times \mathbb{N}$ in the following way: $C = C_1 \cup \{(c,a,j) \mid c \in C_1 \wedge \exists c' : c \stackrel{a}{\Rightarrow}_1 c' \wedge 0 \le j \le |h(a)|\}$. Intuitively, these extra states are the intermediate states that have to appear along the path from c to c' when reading $h(a)$. More precisely, (c,a,j) is the state reached after having read the j first characters of $h(a)$ from c. We also let $i = i_1$; $\le = \le_1 \cup \{\big((c_1,a,j),(c_2,a,j)\big) \mid (c_1,a,j),(c_2,a,j) \in C \wedge c_1 \le_1 c_2\}$. The transition relation is built according to the intuition we have sketched when introducing C:

$$\Rightarrow = \left\{ \begin{array}{l|c} \big(c,\varepsilon,(c,a,0)\big), & \\ \big((c,a,0),w_1,(c,a,1)\big), & c \stackrel{a}{\Rightarrow}_1 c' \\ \quad\vdots & \text{and} \\ \big((c,a,|h(a)|-1),w_{|h(a)|},(c,a,|h(a)|)\big) & h(a) = w_1 w_2 \ldots w_{|h(a)|} \\ \big((c,a,|h(a)|),\varepsilon,c'\big) & \end{array} \right\}$$

Finally, $\mathscr{U} = \mathscr{U}_1$.

By construction, $L^G(S,\mathscr{U}) = h\big(L^G(S_1,\mathscr{U}_1)\big)$, and $\mathscr{U}$ is a $\le$-upward-closed set. It remains to show that:

- **$\le$ is a wqo** Let us suppose it is not the case. Then, there exists a sequence of elements of C: $\varsigma = c_1, c_2, \ldots, c_n, \ldots$ s.t. for any $k \ge 1$, for any $1 \le n < k$: $c_n \not\le c_k$ (each configuration is $\le$-uncomparable to all the previous ones). Remark that, since $\le_1$ is a wqo on the elements of C_1 and since $c \le_1 c'$ implies $c \le c'$ (by definition of $\le$), one cannot extract, from ς, an infinite subsequence of elements from C_1. Thus, there is, in ς, an infinite subsequence $\varsigma' = c_{j_1}, c_{j_2}, \ldots c_{j_n}, \ldots$ s.t. for any $k \ge 1$: (i) $c_{j_k} \notin C_1$ and (ii) for any $1 \le n < k$: $c_n \not\le c_k$.
  By definition of an homomorphism, the value $\ell = \max_{a\in\Sigma}\{|h(a)|\}$ is a finite value. Hence, there exists $0 \le \ell' \le \ell$ and a character a of Σ s.t. the sequence $(c_{j_1},a,\ell'),(c_{j_2},a,\ell'),\ldots,(c_{j_n},a,\ell'),\ldots$ is a subsequence of ς' and $(c_{j_1},a,\ell') \not\le (c_{j_2},a,\ell') \not\le \ldots \not\le (c_{j_n},a,\ell') \not\le \ldots$. But this implies that $c_{j_1} \not\le_1 c_{j_2} \not\le_1 \ldots \not\le_1 \ldots \not\le_1 c_{j_n} \not\le_1 \ldots$, which contradicts the fact that $\le_1$ is a wqo.
- **$\Rightarrow$ is $\le$-monotonic** Let us show that, for any $c_1, c_2, c_3 \in C$, and for any $a \in \Sigma$ s.t. $c_1 \stackrel{a}{\Rightarrow} c_2$ and $c_1 \le c_3$, there exists c_4 s.t. $c_3 \stackrel{a}{\Rightarrow} c_4$ and $c_2 \le c_4$. We consider two cases. (i) Either $c_1 \in C_1$. In that case, by definition of $\Rightarrow$, we have $a = \varepsilon$ and $c_2 = (c_1,b,0)$ for some b. Clearly, $c_4 = (c_3,b,0)$ satisfies the conditions. (ii) Or, $c_1 \notin C_1$. In that case $c_1 = (c',b,i)$ and $c_3 = (c'',b,i)$ with $c' \le_1 c''$, for some b. Again, we have to consider two subcases. (a) Either $i < |h(b)|$. In that case $c_2 = (c',b,i+1)$, and we chose $c_4 = (c'',b,i+1)$, which satisfies the conditions. (b) Or $i = |h(b)|$. In this case, c_2 is a configuration of C_1 such that $c' \stackrel{a}{\Rightarrow}_1 c_2$. We chose c_4 s.t. $c_2 \le_1 c_4$ and $c'' \stackrel{a}{\Rightarrow}_1 c_4$. Such a configuration exists by monotonicity of $\Rightarrow_1$, and satisfies the monotonicity conditions of $\Rightarrow$.

Inverse homomorphism Let us build S and $\mathscr{U}$ s.t. $L^G(S,\mathscr{U}) = h^{-1}\big(L^G(S_1,\mathscr{U}_1)\big)$. We let $C = C_1$; $i = i_1$; $\le = \le_1$; $\Rightarrow = \{(c_1,a,c_2) \mid \exists m \in \Sigma^* : h(a) = m \wedge c_1 \stackrel{m}{\Rightarrow}_1 c_2\}$ and $\mathscr{U} = \mathscr{U}_1$.

Clearly, $L^G(S, \mathscr{U}) = h^{-1}\big(L^G(S_1, \mathscr{U}_1)\big)$. By definition, $\mathscr{U}$ is $\leq$-upward-closed and $\leq$ is a wqo. It remains to show that $\Rightarrow$ is $\leq$-monotonic. Let c_1, c_2, c_3 be three configurations in C s.t. $c_1 \overset{a}{\Rightarrow} c_2$ for some a, and $c_1 \leq c_3$. By definition of $\Rightarrow$, there exists $m \in \Sigma^*$ s.t. $h(a) = m$ and $c_1 \overset{m}{\Rightarrow}_1 c_2$. Moreover, $c_3 \in C_1$ and $c_1 \leq_1 c_3$, by definition. By using an inductive reasoning on the length of $|m|$, one can show that there exists $c_4 \in C_1$ s.t. $c_3 \overset{m}{\Rightarrow}_1 c_4$ and $c_2 \leq_1 c_4$. Hence, $c_4 \in C$ and $c_3 \overset{a}{\Rightarrow} c_4$, by definition of $\Rightarrow$. □

It should now be clear that the class $L^G(\mathsf{WSTS})$ enjoys interesting properties: the emptiness is decidable on this class, under reasonable effectiveness assumptions (Theorem 2), and it forms a full AFL closed under intersection (Theorem 3). Moreover, the transition relation of WSTS is, by definition, $\leq$-monotonic. Thus, $\leq$-upward-closed sets are perfectly suited accepting conditions for these systems. For all these reasons, we will henceforth restrict ourselves to the study of $L^G(\mathsf{WSTS})$. The languages in this class are called *well-structured languages*:

Definition 7 A language L is a *well-structured language* (WSL for short) iff $L \in L^G(\mathsf{WSTS})$.

Remark 1 It is worth recalling that a fourth kind of accepting condition has been routinely studied in the literature. In our context, it is the class $L^P(\mathsf{WSTS})$ of prefix languages one obtains by taking the whole set of configurations as accepting set. By definition, such a set is upward-closed. Since a language that contains no words of length < 2 cannot be in $L^P(\mathsf{WSTS})$, we have: $L^P(\mathsf{WSTS}) \subset L^G(\mathsf{WSTS})$. Most of the results about the classes L^G we are about to present can easily be re-obtained on their corresponding classes L^P.

3.2 Undecidability of universality

Unfortunately, the universality problem is undecidable on EWSTS. This problem is defined as follows:

Problem 3 Given an EWSTS S, an alphabet Σ, and an upward-closed set of accepting markings $\mathscr{U}_f$, the *universality problem* asks whether $L^G(S, \mathscr{U}_f) = \Sigma^*$.

Our proof of undecidability of the universality problem is an adaptation of the proof of Theorem 5.6 in [12], which states the undecidability of the *place-boundedness problem* for PN+NBA. That latter problem is defined as follows:

Problem 4 Given a PN+NBA $\mathscr{N} = \langle \mathscr{P}, \mathscr{T}, \Sigma, \mathbf{m}_0 \rangle$ and a place $p \in \mathscr{P}$ of $\mathscr{N}$, the *place-boundedness problem for* PN+NBA asks whether there exists, for all $i \in \mathbb{N}$, a marking $\mathbf{m}$ such that $\mathbf{m}_0 \xrightarrow{*} \mathbf{m}$ and $\mathbf{m}(p) > i$.

The proof of undecidability of that problem is based on a construction that, given a 2CM, produces a PN+NBA to simulate it. This allows to reduce the boundedness problem of 2CM to place-boundedness of PN+NBA. Since we want to adapt the proof, we first have to recall the construction presented in [12]. We also give two lemmata that state properties of the PN+NBA's obtained thanks to the construction (these lemmata will be exploited when we adapt the proof).

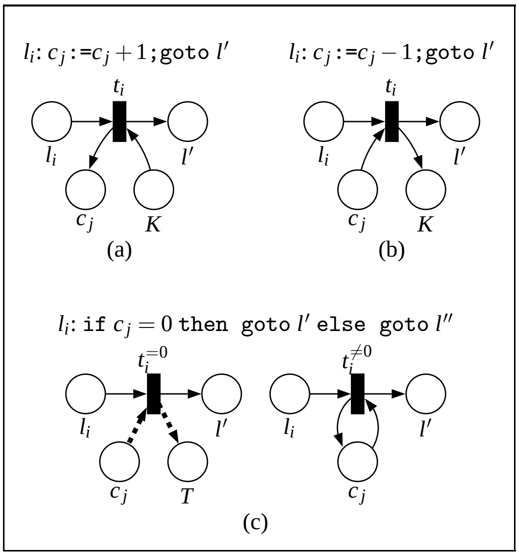

Fig. 2 Simulation of the operations of a 2CM by PN+NBA transitions.

Reduction to place boundedness The construction used in reduction of the boundedness problem for 2CM to the place boundedness problem of PN+NBA works as follows. For any 2CM $C = \langle c_1, c_2, L = \{l_1, l_2, \ldots, l_u\}, \mathsf{Instr}\rangle$, we build a PN+NBA $\mathcal{N}_C = \langle \mathcal{P}, \mathcal{T}, \Sigma, \mathbf{m}_0\rangle$ defined as follows. $\Sigma = \{a, \varepsilon\}$. The set of places $\mathcal{P}$ is equal to $\{c_1, c_2, l_1, l_2, \ldots, l_u, K, T, p_1, p_2\}$. The places c_1 and c_2 will be used to keep track of the values of the two counters of C, $l_1, l_2, \cdots, l_u$ called the *control places* will be used to keep track of the program counter of C, K is called the *capacity place*, T is called the *trash*. Finally, p_1 and p_2 are used to reinitialize the PN+NBA. The set of transitions $\mathcal{T}$ is the smallest set of transitions such that for each $l_i \in L$:

– if $\mathsf{Instr}(l_i)$ is of the form $c_j := c_j + 1; \texttt{goto } l'$, then $\mathcal{T}$ contains the transition $t_i = \langle I, O, \perp, \perp, 0, \varepsilon\rangle$ with $I(l_i) = 1$,$I(K) = 1$ and $\forall p \neq l_i, K : I(p) = 0$, $O(c_j) = 1, O(l') = 1$ and $\forall p \neq c_j, l' : O(p) = 0$;
– if $\mathsf{Instr}(l_i)$ is of the form $c_j := c_j - 1; \texttt{goto } l'$, then $\mathcal{T}$ contains the transition $t_i = \langle I, O, \perp, \perp, 0, \varepsilon\rangle$ with $I(l_i) = 1, I(c_j) = 1$ and $\forall p \neq l_i, c_j : I(p) = 0$, $O(l') = 1, O(K) = 1$ and $\forall p \neq l', K : O(p) = 0$;
– if $\mathsf{Instr}(l_i)$ is of the form $\texttt{if } c_j = 0 \texttt{ then goto } l' \texttt{ else goto } l''$ then $\mathcal{T}$ contains two transitions $t_i^{=0}$ and $t_i^{\neq 0}$ defined as:
  – $t_i^{=0} = \langle I, O, c_j, T, 1, \varepsilon\rangle$ with $I(l_i) = 1$ and $\forall p \neq l_i : I(p) = 0$, $O(l') = 1$ and $\forall p \neq l' : O(l') = 0$;
  – $t_i^{\neq 0} = \langle I, O, \perp, \perp, 0, \varepsilon\rangle$ with $I(l_i) = 1, I(c_j) = 1$ and $\forall p \neq l_i, c_j : I(p) = 0$, $O(c_j) = 1, O(l'') = 1$ and $\forall p \neq c_j, l'' : O(p) = 0$.

Figure 2(a) shows a transition that simulates an increment of c_j by moving one token from the capacity place to c_j. Figure 2(b) shows a transition that simulates a decrement of c_j by moving one token from c_j to the capacity place. Figure

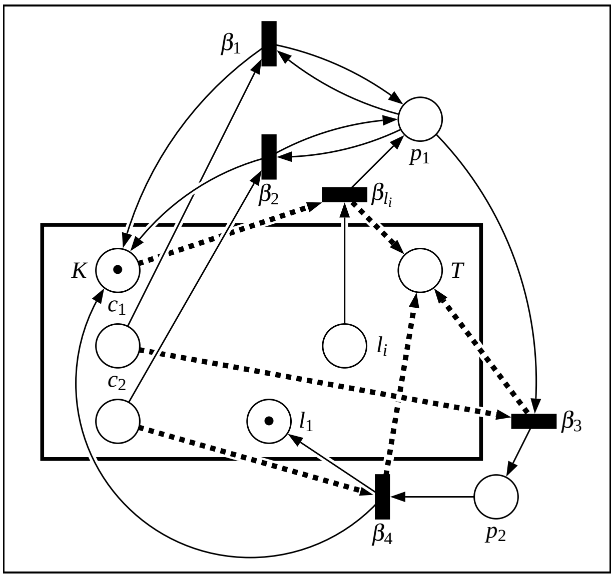

Fig. 3 The PN+NBA $\mathcal{N}_C$.

2(c) shows a transitions that simulates a zero-test on c_j when c_j is equal to zero (transition $t_i^{=0}$) and when c_j is greater than zero.

Finally, we also add to $\mathscr{T}$ the transitions β_{l_i} ($1 \leq i \leq u$), β_1, β_2,β_3 and β_4 that are used to reinitialize the PN+NBA and defined as follows: $\beta_1 = \langle I, O, \perp, \perp, 0, \varepsilon \rangle$ such that $I(p_1) = 1, I(c_1) = 1$ and $\forall p \neq p_1, c_1 : I(p) = 0$, $O(p_1) = 1, O(K) = 1$ and $\forall p \neq p_1, K : O(p) = 0$; $\beta_2 = \langle I, O, \perp, \perp, 0, \varepsilon \rangle$ such that $I(p_1) = 1, I(c_2) = 1$ and $\forall p \neq p_1, c_2 : I(p) = 0$, $O(p_1) = 1, O(K) = 1$ and $\forall p \neq p_1, K : O(p) = 0$; $\beta_3 = \langle I, O, c_1, T, 1, \varepsilon \rangle$ such that $I(p_1) = 1$ and $\forall p \neq p_1 : I(p) = 0$, $O(p_2) = 1$ and $\forall p \neq p_2 : O(p) = 0$; $\beta_4 = \langle I, O, c_2, T, 1, \varepsilon \rangle$ such that $I(p_2) = 1$ and $\forall p \neq p_2 : I(p) = 0$, $O(l_1) = 1$, $O(K) = 1$ and $\forall p \neq l_1, K : O(p) = 0$; for all i such that $1 \leq i \leq u$: $\beta_{l_i} = \langle I, O, K, T, 1, \varepsilon \rangle$ such that $I(l_i) = 1$ and $\forall p \neq l_i : I(p) = 0$, $O(p_1) = 1$ and $\forall p \neq p_1 : O(p) = 0$.

We define the initial marking $\mathbf{m}_0$ as follows: $\mathbf{m}_0(l_1) = 1$, $\mathbf{m}_0(K) = 1$, and $\forall p \neq p_1, K : \mathbf{m}_0(p) = 0$. The construction is depicted in Fig. 3.

From [12], we have the following results:

Lemma 3 ([12]) *For all the reachable markings* $\mathbf{m}$ *of* $\mathcal{N}_C'$*, we have that* $\mathbf{m}(l_1) + \ldots + \mathbf{m}(l_u) + \mathbf{m}(p_1) + \mathbf{m}(p_2) = 1$.

Lemma 4 ([12]) *A* 2CM *C is unbounded if and only if the place K of* $\mathcal{N}_C$ *is unbounded.*

Adaptation of the reduction We are now ready to show how that reduction can be adapted to prove the undecidability of the universality problem for EWSTS. First of all, we have to slightly modify the construction. Let $\mathcal{N}_C'$ be the PN+NBA built

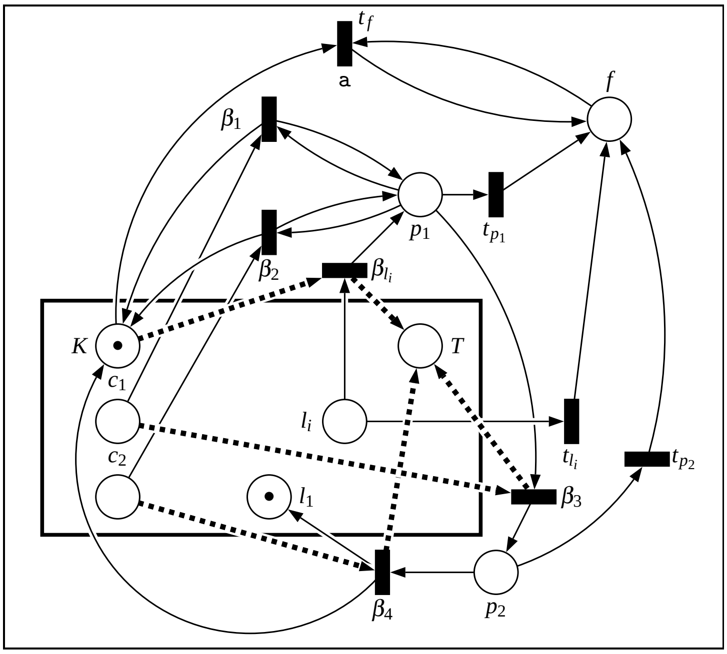

Fig. 4 The PN+NBA $\mathcal{N}'_C$.

from $\mathcal{N}_C$ by adding a place f and transitions $t_{l_1},\ldots,t_{l_u},t_{p_1},t_{p_2}$ and t_f such that for all $1 \le i \le u : t_{l_i} = \langle I,O,\perp,\perp,0,\varepsilon\rangle$ where $I(l_i)=1$ and $\forall p \neq l_i : I(p)=0, O(f)=1$ and $\forall p \neq f : O(p)=0$; for $i=1,2 : t_{p_i} = \langle I,O,\perp,\perp,0,\varepsilon\rangle$ where $I(p_i)=1$ and $\forall p \neq p_i : I(p)=0, O(f)=1$ and $\forall p \neq f : O(p)=0$; $t_f = \langle I,O,\perp,\perp,0,a\rangle$ where $I(f)=1, I(K)=1$ and $\forall p \neq f,K : I(p)=0, O(f)=1$ and $\forall p \neq f : O(p)=0$. Each transition $t=\langle I,O,s,t,b,\varepsilon\rangle$ of $\mathcal{N}_C$ is extended to place f such that $I(f)=0$ and $O(f)=0$.

The initial marking of $\mathcal{N}'_C$ is $\mathbf{m}'_0$ s.t. $\mathbf{m}'_0(f)=0$ and for any place p in the set $\{l_1,\ldots,l_u,c_1,c_2,K,T,p_1,p_2\}$, $\mathbf{m}'_0(p)=\mathbf{m}_0(p)$, .

Thus, $\mathcal{N}'_C$ can stop to simulate C by firing t_p ($p \in \{l_1,\ldots,l_u,p_1,p_2\}$) and then recognizes words of the form a^i by firing t_f (where i is bounded by the number of tokens in K) The construction is shown in Fig. 4.

Let $\mathcal{U}_f$ be the $\preccurlyeq$-upward-closed of markings $\{\mathbf{m} \mid \mathbf{m}(f) \geq 1\}$. Let us first notice that the following lemma holds:

Lemma 5 *if $\mathbf{m}'_0 \xrightarrow{\sigma} \mathbf{m}$ in $\mathcal{N}'_C$, then σ is of the form $t_1\cdots t_n \cdot t_l \cdot t_f^i$ where $n,i \geq 0$; $t_1,\ldots,t_n$ are transitions of $\mathcal{N}_C$ (extended to f) and $l \in \{p_1,p_2,l_1,\ldots,l_u\}$.*

Proof First, it is easy to show that t_f can be only fired if a transition t_l with $l \in \{p_1,p_2,l_1,\ldots,l_u\}$ has been fired before. Indeed, remark that $\mathbf{m}'_0(f)=0$ and the transitions t_l with $l \in \{p_1,p_2,l_1,\ldots,l_u\}$ are the only ones that put one token into place f and t_f need at least one token into f to be fired.

Let us now show that all the transitions different from t_f cannot be fired once a transition t_l with $l \in \{p_1,p_2,l_1,\ldots,l_u\}$ is fired. From Lemma 3, while transitions of $\mathcal{N}_C$ (extended to f) are fired, the set of places $\{p_1, p_2, l_1, \ldots, l_u\}$ contains exactly one token. Suppose that we reach $\mathbf{m}$ by only firing transitions of $\mathcal{N}_C$ (extended to f) from which a transition t_l with $l \in \{p_1, p_2, l_1, \ldots, l_u\}$ is fired. By definition, that transition removes the token in $\{p_1, p_2, l_1, \ldots, l_u\}$ and we reach a marking where there is no token anymore in that set of places. All the transitions except t_f need at least one token in the set of places $\{p_1, p_2, l_1, \ldots, l_u\}$ to be fired and t_f does not add tokens in that set of places. We conclude that all the transitions different from t_f cannot be fired after firing t_l with $l \in \{p_1, p_2, l_1, \ldots, l_u\}$. Hence, the lemma. □

From Lemma 4, we deduce the following lemma:

Lemma 6 *A* 2CM *C is unbounded if and only if $L^G(\mathcal{N}'_C, \mathcal{U}_f) = \{a\}^*$.*

Proof $\boxed{\Rightarrow}$ Suppose that C is unbounded. First notice that $\mathcal{N}'_C$ is constructed from $\mathcal{N}_C$ by adding a place f (not modified by transitions of $\mathcal{N}_C$) and transitions. Hence, from Lemma 3 and Lemma 4 we know that for all $i \geq 0$ we can reach a marking $\mathbf{m}$ from $\mathbf{m}'_0$ by only firing transitions of $\mathcal{N}_C$ (extended to f) such that $\mathbf{m}(K) > i$ and $\mathbf{m}(p) = 1$ with $p \in \{p_1, p_2, l_1, \ldots, l_u\}$. Hence, the sequence of transitions $t_p t_f^i$ may be fired from $\mathbf{m}$ leading to a marking $\mathbf{m}' \in \mathcal{U}_f$. Since t_p and the transitions fired to reach $\mathbf{m}$ are labelled by ε and t_f is labelled by a, we conclude that for all $i \geq 0$, the word a^i is accepted by $\mathcal{N}'_C$.

$\boxed{\Leftarrow}$ Suppose that $L^G(\mathcal{N}'_C, \mathcal{U}_f) = \{a\}^*$. From Lemma 5 and since all the transitions except t_f are labelled by ε, we know that for all $i \geq 0$ there exists $\sigma = t_1 \cdots t_n \cdot t_l \cdot t_f^i$ with $l \in \{p_1, p_2, l_1, \ldots, l_u\}$ such that $\mathbf{m}'_0 \xrightarrow{t_1 \cdots t_n} \mathbf{m} \xrightarrow{t_l \cdot t_f^i} \mathbf{m}'$ and $\mathbf{m}' \in \mathcal{U}_f$. The sequence of transitions $t_l \cdot t_f^i$ removes i tokens from K. Hence, $\mathbf{m}(K) \geq i$ and we conclude that for all $i \geq 0$ there exists a reachable marking $\mathbf{m}$ such that $\mathbf{m}(K) \geq i$.Moreover, $\mathbf{m}$ is reachable by firing transitions of $\mathcal{N}_C$ (extended to f). Hence, the marking $\mathbf{m}'$ such that for all place $p \neq f : \mathbf{m}'(p) = \mathbf{m}(p)$ is reachable in $\mathcal{N}_C$, i.e. K is unbounded in $\mathcal{N}_C$. From Lemma 4, we conclude that C is unbounded. □

This allows us to prove the following theorem:

Theorem 4 *The universality problem is undecidable for* EWSTS*, when we consider $\leq$-upward-closed accepting sets.*

Proof From Theorem 1, we know that the boundedness problem is undecidable for 2CM. Moreover, from Lemma 6 we know that we can reduce the boundedness problem to the universality problem for EWSTS. We conclude that the universality problem is undecidable for EWSTS. □

4 Pumping lemmata

This section presents three lemmata that show the limitations in the expressiveness of WSTS (for the first one), PN (for the second one), and PN+NBA (for the third one). All these lemmata have a similar statement: if a given WSTS (resp. PN, PN+NBA) accepts an infinite set of words $\{w_1, w_2, \ldots\}$ with a given structure, then it must also accept other words that are built upon the words $w_1, w_2, \ldots$ In some sense, these lemmata allow to "inflate" the set of accepted words. For that reason, we have chosen to call them *pumping lemmata*, owing to their similarities to the classical pumping lemmata for regular and context-free languages (see for instance [8]).

The proof techniques rely on properties of infinite sequences of configurations (equipped with a wqo), and monotonicity properties. The usefulness of these pumping lemmata will be demonstrated in Section 5, where we apply them to obtain several results about WSL.

4.1 A pumping lemma for WSL

Our first pumping lemma deals with WSL, and is very easy to prove:

Lemma 7 *Let L be a* WSL, *and let $w_1, w_2, \ldots$ be an infinite sequence of words s.t. $\forall k \geq 1 : w_k \in L$ and $w_k = B_k \cdot E_k$. Then, here exist $i < j$ s.t. $B_j \cdot E_i \in L$.*

Proof Let $S = \langle C, c_0, \Sigma, \Rightarrow, \leq \rangle$ be a WSTS s.t. $L(S, \mathscr{U}) = L$ for some $\leq$-upward-closed set $\mathscr{U}$. For any $k \geq 1$, let $c_k \in C$ be a configuration s.t. $c_0 \stackrel{B_k}{\Rightarrow} c_k \stackrel{E_k}{\Rightarrow} c'_k$, with $c'_k \in \mathscr{U}$. By definition of wqo, there exists $i < j$ s.t. $c_i \leq c_j$. Hence, $c_0 \stackrel{B_j}{\Rightarrow} c_j \stackrel{E_i}{\Rightarrow} c'$, with $c'_i \leq c'$ by monotonicity. Thus, $c' \in \mathscr{U}$ and $B_j \cdot E_i \in L$. □

4.2 A pumping lemma for PN

Our second pumping lemma states properties of languages of Petri nets (more precisely, languages in the class $L^G(\mathsf{PN})$). This lemma will be exploited in section 5.2, to strictly separate the expressive power of PN and PN+NBA.

Lemma 8 *Let $\mathscr{N}$ be a* PN *and $\mathscr{U}$ be an $\preccurlyeq$-upward-closed set of markings of $\mathscr{N}$.* **If** *there exists an infinite sequence of words $w_1, w_2, \ldots$ such that for any $i \geq 1$, there exist two words B_i, E_i with $\{B_i w_i^* E_i\} \subseteq L(\mathscr{N}, \mathscr{U})$,* **then** *there exist $0 < n_1 < n_2 < n_3$ such that for any $K \geq 0$, there exists $K' \geq K$ such that the word $B_{n_3} w_{n_3}^{i_1} w_{n_1}^{K'} w_{n_2}^{i_2} E_{n_2}$ is in $L(\mathscr{N}, \mathscr{U})$.*

The proof of the lemma is quite tedious and technical. However, we believe that the technique at work in this proof is interesting by itself, since it directly exploits the monotonicity and well-quasi ordering properties that are characteristic of WSTS. Before giving the proof, we provide the reader with a sketch that presents the main arguments. By this mean, we hope to make the task of reading the proof easier. Throughout this explanation, we refer to peculiar markings using the same notations as in the proof. The reader is advised to refer to Fig. 5 and 6 to get the intuition of the meaning of these notations.

The proof is constructive. From the fact that the PN accepts the words $B_i w_i^* E_i$ for any $i \geq 1$, we build, by applying Lemma 2, infinite sequences of markings that are ordered (this is the purpose of the two first steps of the proof). Then, at the third step, we exploit these ordering properties, as well as the monotonicity of the

PN, to show that a sequence of transitions with the desirable form is firable, and leads to the accepting $\preccurlyeq$-upward-closed set of markings.

Step 1 For all $i \geq 1$, we build the infinite sequences of runs $\mathscr{M}_i$ where the $j-th$ element of those sequences is a run that accepts the word $B_i w_i^{2^{|\mathscr{P}|}+j} E_i$ (where $\mathscr{P}$ is the set of places of the PN considered). Then, for all $i \geq 1$ we build the sub-sequence $\mathscr{M}_i^{\preccurlyeq}$ of $\mathscr{M}_i$ by applying successively Lemma 2. Those sub-sequences have the property that markings appearing in different runs are $\preccurlyeq$-ordered, as shown on Fig. 5. The increasing sequences appear along the $2^{|\mathscr{P}|}+1$ first "columns", and along the "diagonals" whose first element appears in one of these "columns".

Step 2 The second step consists to select an infinite subset of the $\mathscr{M}_i^{\preccurlyeq}$'s. We do this by building a sequence of runs such that the jth run is the first run appearing in $\mathscr{M}_j^{\preccurlyeq}$. Again, we extract a sub-sequence S where markings appearing in different runs are $\preccurlyeq$-ordered by applying successively Lemma 2. In this case, only markings appearing along the $2^{|\mathscr{P}|}+1$ first "columns" are $\preccurlyeq$-ordered. This is shown at Fig. 6 (a).

Step 3 Finally, we show how to split and combine parts of runs appearing in the $\mathscr{M}_i$'s and S to obtain a run that allows the PN to accept a word of the desired form. This is shown at Fig. 6 (b).

In order to build this sequence, we rely on several variables, namely: c_1, c_2, n, x and y. At the present step of the proof, we present several constraints that relate x, y and n to c_1, c_2 and K. These constraints are meant to produce a sequence of transitions that accepts a word of the desired form. The main (and most technical) part of step 3 consists to show that these constraints are satisfiable.

The first part of the sequence is the prefix of $\mathscr{M}_{\rho(n)}^{\preccurlyeq}(x)$, up to the "column" c_1 (see Fig. 6(b)). At that point, we are guaranteed that the marking we obtain is larger than $\mathscr{M}_{\rho(1)}^{\preccurlyeq}(y,c_1)$. This allows us to continue the sequence with a part of $\mathscr{M}_{\rho(1)}^{\preccurlyeq}(y)$, starting at column c_1 and ending at column c_2+y-1. Again, by exploiting the properties of the sequences built at steps 1 and 2, we are ensured that the marking we have reached is larger than $\mathscr{M}_{\rho(2)}^{\preccurlyeq}(1,c_2)$. This allow us to finish the sequence with the suffix of $\mathscr{M}_{\rho(2)}^{\preccurlyeq}(1)$. The word accepted by this sequence is of the desired form, since we have correctly chosen the values of x, y and n (in particular, y is large enough to ensure that the central part of the word is longer than K times $|w_{n_1}|$).

We are now ready to present the proof of Lemma 8.

Proof Let $\mathscr{N}$ be a PN with set of places $\mathscr{P}$ and initial marking $\mathbf{m}_{init}$, such that $\{B_i w_i^* E_i\} \subseteq L(\mathscr{N},\mathscr{U})$ for all $i \geq 1$.

Step 1 For any $i \geq 1$, let $\mathscr{M}_i$ be the infinite sequence of all the runs accepting the words of the form $B_i w_i^j E_i$, with $j \geq 2^{|\mathscr{P}|}+1$. That is, $\mathscr{M}_i$ is the sequence of runs

$$
\begin{array}{llllllllllllll}
\mathscr{M}_i^{\preccurlyeq}(1) = \mathbf{m}_{init} \xrightarrow{B_i} & \mathscr{M}_i^{\preccurlyeq}(1,1) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(1,2) & \xrightarrow{w_i} \cdots \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(1,2^{|\mathscr{P}|}+1) & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \xrightarrow{E_i} n_{i,1} \in \mathscr{U} \\
& \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} \quad \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & & & & & & \\
\mathscr{M}_i^{\preccurlyeq}(2) = \mathbf{m}_{init} \xrightarrow{B_i} & \mathscr{M}_i^{\preccurlyeq}(2,1) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(2,2) & \xrightarrow{w_i} \cdots \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(2,2^{|\mathscr{P}|}+1) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(2,2^{|\mathscr{P}|}+2) & \cdots & & \cdots & \cdots & \cdots \xrightarrow{E_i} n_{i,2} \in \mathscr{U} \\
& \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} \quad \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & & \rotatebox{-45}{$\preccurlyeq$} & & & & \\
\mathscr{M}_i^{\preccurlyeq}(3) = \mathbf{m}_{init} \xrightarrow{B_i} & \mathscr{M}_i^{\preccurlyeq}(3,1) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(3,2) & \xrightarrow{w_i} \cdots \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(3,2^{|\mathscr{P}|}+1) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(3,2^{|\mathscr{P}|}+2) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(3,2^{|\mathscr{P}|}+3) & \cdots & \cdots & \xrightarrow{E_i} n_{i,3} \in \mathscr{U} \\
& \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} \quad \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & & \rotatebox{-45}{$\preccurlyeq$} & & \rotatebox{-45}{$\preccurlyeq$} & & \\
& \vdots & & \vdots & & \vdots & & \ddots & & \ddots & & \ddots & \\
\mathscr{M}_i^{\preccurlyeq}(k) = \mathbf{m}_{init} \xrightarrow{B_i} & \mathscr{M}_i^{\preccurlyeq}(k,1) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(k,2) & \xrightarrow{w_i} \cdots \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(k,2^{|\mathscr{P}|}+1) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(k,2^{|\mathscr{P}|}+2) & \xrightarrow{w_i} & \mathscr{M}_i^{\preccurlyeq}(k,2^{|\mathscr{P}|}+3) & \xrightarrow{w_i} & \cdots & \xrightarrow{E_i} n_{i,k} \in \mathscr{U} \\
& \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} \quad \rotatebox{-45}{$\preccurlyeq$} & \rotatebox{-90}{$\preccurlyeq$} & \rotatebox{-45}{$\preccurlyeq$} & & \rotatebox{-45}{$\preccurlyeq$} & & \rotatebox{-45}{$\preccurlyeq$} & & \\
& \vdots & & \vdots & & \vdots & & \ddots & & \ddots & & \ddots &
\end{array}
$$

Fig. 5 The sequence of runs $\mathscr{M}_i^{\preccurlyeq}$. The runs have been selected in order to obtain the increasing sequences of markings that are shown on the Figure.

(a)

$$
\begin{array}{llllllll}
S(1) = \mathbf{m}_{init} & \xrightarrow{B_{\rho(1)}} \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,1) & \xrightarrow{w_{\rho(1)}} \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,2) & \xrightarrow{w_{\rho(1)}} \cdots & \xrightarrow{w_{\rho(1)}} \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,2^{|\mathscr{P}|}+1) & \cdots & \xrightarrow{E_{\rho(1)}} n_{\rho(1),1} \in \mathscr{U} \\
& \curlyvee & \curlyvee & & \curlyvee & & \\
S(2) = \mathbf{m}_{init} & \xrightarrow{B_{\rho(2)}} \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1,1) & \xrightarrow{w_{\rho(2)}} \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1,2) & \xrightarrow{w_{\rho(2)}} \cdots & \xrightarrow{w_{\rho(2)}} \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1,2^{|\mathscr{P}|}+1) & \cdots & \xrightarrow{E_{\rho(2)}} n_{\rho(2),1} \in \mathscr{U} \\
& \curlyvee & \curlyvee & & \curlyvee & & \\
S(3) = \mathbf{m}_{init} & \xrightarrow{B_{\rho(3)}} \mathscr{M}^{\preccurlyeq}_{\rho(3)}(1,1) & \xrightarrow{w_{\rho(3)}} \mathscr{M}^{\preccurlyeq}_{\rho(3)}(1,2) & \xrightarrow{w_{\rho(3)}} \cdots & \xrightarrow{w_{\rho(3)}} \mathscr{M}^{\preccurlyeq}_{\rho(3)}(1,2^{|\mathscr{P}|}+1) & \cdots & \xrightarrow{E_{\rho(3)}} n_{\rho(3),1} \in \mathscr{U} \\
& \curlyvee & \curlyvee & & \curlyvee & & \\
& \vdots & \vdots & & \vdots & & \\
S(k) = \mathbf{m}_{init} & \xrightarrow{B_{\rho(k)}} \mathscr{M}^{\preccurlyeq}_{\rho(k)}(1,1) & \xrightarrow{w_{\rho(k)}} \mathscr{M}^{\preccurlyeq}_{\rho(k)}(1,2) & \xrightarrow{w_{\rho(k)}} \cdots & \xrightarrow{w_{\rho(k)}} \mathscr{M}^{\preccurlyeq}_{\rho(k)}(1,2^{|\mathscr{P}|}+1) & \cdots & \xrightarrow{E_{\rho(k)}} n_{\rho(k),1} \in \mathscr{U} \\
& \curlyvee & \curlyvee & & \curlyvee & & \\
& \vdots & \vdots & & \vdots & &
\end{array}
$$

(b)

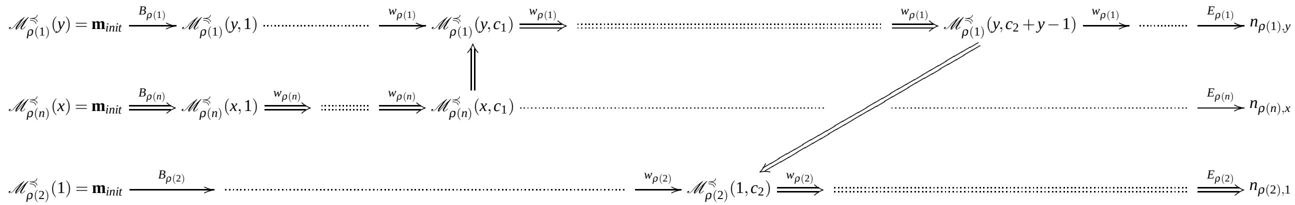

Fig. 6 (a) shows the sequence of runs S. It is built by considering the first runs of every $\mathscr{M}^{\preccurlyeq}_i$, and keeping only those that allow to build the $2^{|\mathscr{P}|}+1$ infinite increasing sequences of markings that are shown. (b) shows the firable sequence (along the $\Rightarrow$'s) that accepts a word of the form $B_{n_3} w_{n_3}^{i_1} w_{n_1}^{K'} w_{n_2}^{i_2} E_{n_2}$.

(where $j \geq 2^{|\mathscr{P}|}+1$):

$$\begin{array}{l}
\mathbf{m}_{init} \xrightarrow{\upsilon_j} \mathbf{m}_j^1 \xrightarrow{\varsigma_j^1} \mathbf{m}_j^2 \xrightarrow{\varsigma_j^2} \cdots \xrightarrow{\varsigma_j^j} \mathbf{m}_j^{j+1} \xrightarrow{\upsilon'_j} \mathbf{n}_j \\
\mathbf{m}_{init} \xrightarrow{\upsilon_{j+1}} \mathbf{m}_{j+1}^1 \xrightarrow{\varsigma_{j+1}^1} \mathbf{m}_{j+1}^2 \xrightarrow{\varsigma_{j+1}^2} \cdots \xrightarrow{\varsigma_{j+1}^{j+1}} \mathbf{m}_{j+1}^{j+2} \xrightarrow{\upsilon'_{j+1}} \mathbf{n}_{j+1} \\
\cdots
\end{array}$$

where for any $\ell \geq j$: $\mathbf{n}_\ell \in \mathscr{U}$, $\Lambda(\upsilon_\ell) = B_i$ and $\Lambda(\upsilon'_\ell) = E_i$. Moreover, $\forall \ell \geq j : \forall 1 \leq k \leq \ell : \Lambda(\varsigma_\ell^k) = w_i$. By applying Lemma 2 successively, we construct, for any $i \geq 1$, an infinite subsequence $\mathscr{M}_i^{\preccurlyeq}$ of $\mathscr{M}_i$:

$$\begin{array}{l}
\mathbf{m}_{init} \xrightarrow{\upsilon_{\pi(1,i)}} \mathbf{m}_{\pi(1,i)}^1 \xrightarrow{\varsigma_{\pi(1,i)}^1} \mathbf{m}_{\pi(1,i)}^2 \xrightarrow{\varsigma_{\pi(1,i)}^2} \cdots \xrightarrow{\varsigma_{\pi(1,i)}^{\pi(1,i)}} \mathbf{m}_{\pi(1,i)}^{\pi(1,i)+1} \xrightarrow{\upsilon'_{\pi(1,i)}} \mathbf{n}_{\pi(1,i)} \\
\mathbf{m}_{init} \xrightarrow{\upsilon_{\pi(2,i)}} \mathbf{m}_{\pi(2,i)}^1 \xrightarrow{\varsigma_{\pi(2,i)}^1} \mathbf{m}_{\pi(2,i)}^2 \xrightarrow{\varsigma_{\pi(2,i)}^2} \cdots \xrightarrow{\varsigma_{\pi(2,i)}^{\pi(2,i)}} \mathbf{m}_{\pi(2,i)}^{\pi(2,i)+1} \xrightarrow{\upsilon'_{\pi(2,i)}} \mathbf{n}_{\pi(2,i)} \\
\cdots
\end{array}$$

where every $\mathscr{M}_i^{\preccurlyeq}$ is such that:

(a) For any $1 \leq j \leq 2^{|\mathscr{P}|}+1$, the sequence of markings $\mathbf{m}_{\pi(1,i)}^j, \mathbf{m}_{\pi(2,i)}^j \ldots$ is increasing: $\forall \ell \geq 1 : \mathbf{m}_{\pi(\ell,i)}^j \preccurlyeq \mathbf{m}_{\pi(\ell+1,i)}^j$;

(b) For any $1 \leq j \leq 2^{|\mathscr{P}|}+1$, there exists a set of places $Places(\mathscr{M}_i^{\preccurlyeq}, j) \subseteq \mathscr{P}$ that strictly increase along the sequence $\mathbf{m}_{\pi(1,i)}^j, \mathbf{m}_{\pi(2,i)}^j \ldots$ The other places stay constant along the sequence: $\forall \ell \geq 1 : \mathbf{m}_{\pi(\ell,i)}^j(p) < \mathbf{m}_{\pi(\ell+1,i)}^j(p)$ if and only if $p \in Places(\mathscr{M}_i^{\preccurlyeq}, j)$;

(c) For any $1 \leq j \leq 2^{|\mathscr{P}|}+1$, for any $\ell \geq 1$ we also have: $\mathbf{m}_{\pi(\ell,i)}^{j+\ell-1} \preccurlyeq \mathbf{m}_{\pi(\ell+1,i)}^{j+\ell}$.

Let us now introduce some notations. We denote by $\mathscr{M}_i^{\preccurlyeq}(\ell)$ the ℓ-th run of $\mathscr{M}_i^{\preccurlyeq}$. We denote by $\mathscr{M}_i^{\preccurlyeq}(\ell, j)$ the marking number j in $\mathscr{M}_i^{\preccurlyeq}(\ell)$, i.e., $\mathbf{m}_{\pi(\ell,i)}^j$ in $\mathscr{M}_i^{\preccurlyeq}(\ell)$. We denote by $\sigma_i^\ell(k_1, k_2)$ the sequence of transitions of $\mathscr{M}_i^{\preccurlyeq}(\ell)$ one fires from $\mathscr{M}_i^{\preccurlyeq}(\ell, k_1)$ to reach $\mathscr{M}_i^{\preccurlyeq}(\ell, k_2)$. That is, $\sigma_i^\ell(k_1, k_2) = \varsigma_{\pi(\ell,i)}^{k_1} \cdot \varsigma_{\pi(\ell,i)}^{k_1+1} \cdot \ldots \cdot \varsigma_{\pi(\ell,i)}^{k_2-1}$ in $\mathscr{M}_i^{\preccurlyeq}(\ell)$. We also denote by $\sigma_i^\ell(\cdot, k)$ the sequence $\upsilon_{\pi(\ell,i)} \cdot \varsigma_{\pi(\ell,i)}^1 \cdots \varsigma_{\pi(\ell,i)}^{k-1}$; and by $\sigma_i^\ell(k, \cdot)$ the sequence $\varsigma_{\pi(\ell,i)}^k \cdot \ldots \cdot \varsigma_{\pi(\ell,i)}^{\pi(\ell,i)} \cdot \upsilon'_{\pi(\ell,i)}$ both in $\mathscr{M}_i^{\preccurlyeq}(\ell)$.

Step 2 To finish with the construction, we consider the infinite sequence of runs $\mathscr{M}_1^{\preccurlyeq}(1), \mathscr{M}_2^{\preccurlyeq}(1), \ldots$ made up of the first runs of all $\mathscr{M}_i^{\preccurlyeq}$. From this sequence, we extract the infinite subsequence $S = \mathscr{M}_{\rho(1)}^{\preccurlyeq}(1), \mathscr{M}_{\rho(2)}^{\preccurlyeq}(1), \ldots$ by successively applying Lemma 2 again. We construct S such that:

(d) For any $1 \leq j \leq 2^{|\mathscr{P}|}+1$ the sequence $\mathscr{M}_{\rho(1)}^{\preccurlyeq}(1, j), \mathscr{M}_{\rho(2)}^{\preccurlyeq}(1, j), \ldots$ is increasing:

$$\forall k \geq 1 : \mathscr{M}_{\rho(k)}^{\preccurlyeq}(1, j) \preccurlyeq \mathscr{M}_{\rho(k+1)}^{\preccurlyeq}(1, j) \tag{1}$$

(e) For any $1 \leq j \leq 2^{|\mathscr{P}|}+1$, there exists a set of places $Places(S,j) \subseteq \mathscr{P}$ that strictly increase along the sequence $\mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,j), \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1,j), \ldots$ All the other places stay constant along the sequence:

$$\forall k \geq 1 : \mathscr{M}^{\preccurlyeq}_{\rho(k)}(1,j)(p) < \mathscr{M}^{\preccurlyeq}_{\rho(k+1)}(1,j)(p) \text{ iff } p \in Places(S,j) \tag{2}$$

Let c_1 and c_2 be s.t. $1 \leq c_1 < c_2 \leq 2^{|\mathscr{P}|}+1$ and $Places(S,c_1) = Places(S,c_2)$. Remark that c_1 and c_2 always exist because there are $2^{|\mathscr{P}|}$ subsets of $\mathscr{P}$.

(f) The sets of strictly increasing places of the selected $\mathscr{M}^{\preccurlyeq}_i$ are equal: for any $1 \leq j \leq 2^{|\mathscr{P}|}+1$, for any $k \geq 1 : Places(\mathscr{M}^{\preccurlyeq}_{\rho(k)}, j) = Places(\mathscr{M}^{\preccurlyeq}_{\rho(k+1)}, j)$. This is possible because there is a finite number of subsets of $\mathscr{P}$.

One can now summarize points (a) to (c) of the construction by the two following conditions[2], which hold for any $k \geq 1$, for any $1 \leq j \leq 2^{|\mathscr{P}|}+1$ and for any $\ell \geq 1$:

$$\forall \mu \geq \nu \geq 0 : \mathscr{M}^{\preccurlyeq}_{\rho(k)}(\ell,j) \preccurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(k)}(\ell+\mu, j+\nu) \tag{3}$$

$$\mathscr{M}^{\preccurlyeq}_{\rho(k)}(\ell,j)(p) < \mathscr{M}^{\preccurlyeq}_{\rho(k)}(\ell+1,j)(p) \text{ iff } p \in Places(\mathscr{M}^{\preccurlyeq}_{\rho(k)}, j) \tag{4}$$

Step 3 The rest of the proof consists in showing that there exist $i_1 \geq 0, i_2 \geq 0$ and $0 < n_1 < n_2 < n_3$ such that for any $K \in \mathbb{N}$, one can devise a word $w = B_{n_3} w^{i_1}_{n_3} w^{K'}_{n_1} w^{i_2}_{n_2} E_{n_2}$ with $K' \geq K$ that is accepted by $\mathscr{N}$. The accepting sequence of transitions (called σ) is built as follows: $\sigma = \sigma^x_{\rho(n)}(\cdot, c_1) \cdot \sigma^y_{\rho(1)}(c_1, c_2+y-1) \cdot \sigma^1_{\rho(2)}(c_2, \cdot)$ for well-chosen values of n, x and y. We now explain how to compute those values for any K.

Choice of n Let $\mathbf{m}_n$ be the marking such that $\mathscr{M}^{\preccurlyeq}_{\rho(n)}(1,c_1) \xrightarrow{\sigma^1_{\rho(1)}(c_1,c_2)} \mathbf{m}_n$. Remark that, since we are dealing with Petri nets, the sequence $\sigma^1_{\rho(1)}(c_1,c_2)$ has a constant effect (i.e., characterized by a vector of natural constants) equal to $\mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_2) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)$. Thus $\mathbf{m}_n = \mathscr{M}^{\preccurlyeq}_{\rho(n)}(1,c_1) + \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_2) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)$. We choose $n > 2$ such that:

$$\mathbf{m}_n = \mathscr{M}^{\preccurlyeq}_{\rho(n)}(1,c_1) + \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_2) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_1) \succcurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1,c_2) \tag{5}$$

Let us show that such a n exists. First notice that $\sigma^1_{\rho(1)}(c_1,c_2)$ is firable from $\mathscr{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)$ for all $n > 2$, because $\mathscr{M}^{\preccurlyeq}_{\rho(n)}(1,c_1) \succcurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)$ following (1). Then, recall that $Places(S,c_1) = Places(S,c_2)$. Since, for any $p \in Places(S,c_1)$, the sequence $\mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p), \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1,c_1)(p), \ldots$ is strictly growing by (2), we have $\forall p \in Places(S,c_1) : \forall n \geq 1 : \mathscr{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) \geq n-1$. Thus there exists $n \geq 1$ s.t. $\forall p \in Places(S,c_1) : \mathscr{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) \geq \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1,c_2)(p) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1,c_2)(p) +$

[2] More precisely, 3 is a consequence of (a) and (c), and 4 stems from (b).

$\mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p)$. This is equivalent to $\forall p \in Places(S,c_1): \mathbf{m}_n(p) \geq \mathcal{M}^{\preccurlyeq}_{\rho(2)}(1,c_2)(p)$, by definition of $\mathbf{m}_n$. On the other hand, for any $p \in \mathscr{P} \setminus Places(S,c_1)$, we have: $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) = \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p)$ and $\mathcal{M}^{\preccurlyeq}_{\rho(2)}(1,c_2)(p) = \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_2)(p)$, by (2) again. From these two equalities, we obtain:

$$\mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) - \mathcal{M}^{\preccurlyeq}_{\rho(2)}(1,c_2)(p) = \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_2)(p)$$

and thus:

$$\mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) + \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_2)(p) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p) = \mathbf{m}_n(p) = \mathcal{M}^{\preccurlyeq}_{\rho(2)}(1,c_2)(p)$$

Hence, we conclude that, for any place $p \in \mathscr{P} : \mathbf{m}_n(p) \geq \mathcal{M}^{\preccurlyeq}_{\rho(2)}(1,c_2)(p)$.

Choice of y We choose $y > 0$ such that:

$$y > K + c_1 - c_2 + 1 \tag{6}$$

Choice of x Finally, we choose $x > y$ such that:

$$\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1) \succcurlyeq \mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1) + \mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1) \tag{7}$$

One can prove that such a x always exists by a the same reasoning as in the choice of n, and by the fact that $Places(\mathcal{M}^{\preccurlyeq}_{\rho(n)},c_1) = Places(\mathcal{M}^{\preccurlyeq}_{\rho(1)},c_1)$ (Point (f) above). Indeed, $\forall p \in Places(\mathcal{M}^{\preccurlyeq}_{\rho(1)},c_1)$, the sequence $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p), \mathcal{M}^{\preccurlyeq}_{\rho(n)}(2,c_1)(p), \ldots$ is strictly increasing by (4) and (f), and we can thus choose x large enough to have $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1)(p) \geq \mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) + \mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1)(p) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p)$, for any place p in the set $Places(\mathcal{M}^{\preccurlyeq}_{\rho(1)},c_1)$. On the other hand, for any $p \in \mathscr{P} \setminus Places(\mathcal{M}^{\preccurlyeq}_{\rho(1)},c_1)$, we know, by the points (b) and (f) of the construction, that: $\mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1)(p) = \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p)$. Thus, $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1)(p) \geq \mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) + \mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1)(p) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p)$ iff $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1)(p) \geq \mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p)$. This latter point is true by (b). We conclude that for any $p \in \mathscr{P} : \mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1)(p) \geq \mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1)(p) + \mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1)(p) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)(p)$.

The next step amounts to showing that the sequence σ is firable. From $\mathbf{m}_{init}$, we fire $\sigma^x_{\rho(n)}(\cdot,c_1)$ and reach $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1)$. From that marking, we can fire the sequence $\sigma^y_{\rho(1)}(c_1,c_2+y-1)$. This is possible because $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1) \succcurlyeq \mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1)$. Indeed, by (7): $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1) \succcurlyeq \mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1) + \big(\mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)\big)$. However, we know that $\big(\mathcal{M}^{\preccurlyeq}_{\rho(n)}(1,c_1) - \mathcal{M}^{\preccurlyeq}_{\rho(1)}(1,c_1)\big) \succcurlyeq \langle 0,\ldots,0\rangle$, by (1). This implies that $\mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1) \succcurlyeq \mathcal{M}^{\preccurlyeq}_{\rho(1)}(y,c_1)$ and we have:

$$\mathbf{m}_{init} \xrightarrow{\sigma^x_{\rho(n)}(0,c_1)} \mathcal{M}^{\preccurlyeq}_{\rho(n)}(x,c_1) \xrightarrow{\sigma^y_{\rho(1)}(c_1,c_2+y-1)} \mathbf{m}$$

To finish the sequence, we have to show that $\mathbf{m} \succcurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1, c_2)$. Since the effect of $\sigma^{y}_{\rho(1)}(c_1, c_2 + y - 1)$ is constant and equal to $\mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_2 + y - 1) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_1)$, we have:

$$\begin{aligned}
\mathbf{m} &= \mathscr{M}^{\preccurlyeq}_{\rho(n)}(x, c_1) + \mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_2 + y - 1) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_1) \\
\Rightarrow \mathbf{m} &\succcurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(n)}(1, c_1) + \mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_1) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1, c_1) \\
&\quad + \mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_2 + y - 1) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_1) && \text{by (7)} \\
\Rightarrow \mathbf{m} &\succcurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(n)}(1, c_1) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1, c_1) + \mathscr{M}^{\preccurlyeq}_{\rho(1)}(y, c_2 + y - 1) \\
\Rightarrow \mathbf{m} &\succcurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(n)}(1, c_1) - \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1, c_1) + \mathscr{M}^{\preccurlyeq}_{\rho(1)}(1, c_2) && \text{by (3)} \\
\Rightarrow \mathbf{m} &\succcurlyeq \mathscr{M}^{\preccurlyeq}_{\rho(2)}(1, c_2) && \text{by (5)}
\end{aligned}$$

We can thus fire $\sigma^{1}_{\rho(2)}(c_2, \cdot)$ from $\mathbf{m}$ and obtain $\mathbf{m}'$ such that $\mathbf{m}' \succcurlyeq \mathbf{n}_{\rho(2)}$ (by monotonicity), which implies that $\mathbf{m}' \in \mathscr{U}$. Thus, $\mathscr{N}$ accepts $\Lambda(\sigma)$, which is of the form $B_{n_3} w_{n_3}^{i_1} w_{n_1}^{K'} w_{n_2}^{i_2} E_{n_2}$ with *(i)* $n_1 = \rho(1)$, $n_2 = \rho(2)$ and $n_3 = \rho(n)$, hence $0 < n_1 < n_2 < n_3$; *(ii)* $i_1 \geq 0$, $i_2 \geq 0$ and *(iii)* $K' \geq K$. □

4.3 A pumping lemma for PN+NBA

Let us turn our attention to the third pumping lemma. Its proof relies on the following auxiliary lemma:

Lemma 9 *Let $\mathscr{N} = \langle \mathscr{P}, \mathscr{T}, \Sigma, \mathbf{m}_0 \rangle$ be a* PN+NBA, *and let σ be a finite sequence of transitions of $\mathscr{N}$ that contains n occurrences of transitions in $\mathscr{T}_e$. Let $\mathbf{m}_1$, $\mathbf{m}'_1$, $\mathbf{m}_2$ and $\mathbf{m}'_2$ be four makings such that (i) $\mathbf{m}_1 \xrightarrow{\sigma} \mathbf{m}'_1$, (ii) $\mathbf{m}_2 \xrightarrow{\sigma} \mathbf{m}'_2$ and (iii) $\mathbf{m}_2 \succcurlyeq \mathbf{m}_1$. Then, for every place $p \in \mathscr{P}$: $\mathbf{m}'_2(p) - \mathbf{m}'_1(p) \geq \mathbf{m}_2(p) - \mathbf{m}_1(p) - n$.*

Proof Let us consider a place $p \in \mathscr{P}$. First, we remark that when we fire σ from $\mathbf{m}_2$ instead of $\mathbf{m}_1$, its Petri net arcs will have the same effect on p. On the other hand, since we want to find a lower bound on $\mathbf{m}'_2(p) - \mathbf{m}'_1(p)$, we consider the situation where no non-blocking arcs affect p when σ is fired from $\mathbf{m}_1$, but they all remove one token from p when σ is fired from $\mathbf{m}_2$. In the latter case, the effect of σ on p is $\mathbf{m}'_1(p) - \mathbf{m}_1(p) - n$. We obtain thus: $\mathbf{m}'_2(p) \geq \max\{\mathbf{m}_2(p) + \mathbf{m}'_1(p) - \mathbf{m}_1(p) - n, 0\}$. Hence $\mathbf{m}'_2(p) \geq \mathbf{m}'_1(p) + \mathbf{m}_2(p) - \mathbf{m}_1(p) - n$, and thus: $\mathbf{m}'_2(p) - \mathbf{m}'_1(p) \geq \mathbf{m}_2(p) - \mathbf{m}_1(p) - n$. □

We can now state our pumping lemma for PN+NBA:

Lemma 10 *Let $\mathscr{N}$ be a* PN+NBA *and $\mathscr{U}$ be an $\preccurlyeq$-upward-closed set of markings of $\mathscr{N}$.* **If** *there exists an infinite sequence of words $w_1, w_2, \ldots$ such that for any $i \geq 1$, there exist two words B_i, E_i with $\{B_i w_i^* E_i\} \subseteq L(\mathscr{N}, \mathscr{U})$,* **then** *there exist $i_1 \geq 0$, $i_2 > 0$, $i_3 \geq 0$ and $0 < n_1 < n_2 < n_3$ such that the word $B_{n_3} w_{n_3}^{i_1} w_{n_1}^{i_2} w_{n_2}^{i_3} E_{n_2}$ is in $L(\mathscr{N}, \mathscr{U})$.*

Once again, since the proof of Lemma 10 is rather technical, we first sketch it informally. The proof may be decomposed into two steps:

Step 1 This step is similar to Step 1 of Lemma 8. More precisely, for all $i \geq 2^{|\mathscr{P}|}+1$ (where $\mathscr{P}$ is the set of places of the PN+NBA considered), we build the infinite sequences of runs where the $j-th$ element of those sequences is a run that accepts the word $B_i w_i^{i+j-1} E_i$. Then, for all $i \geq 2^{|\mathscr{P}|}+1$ we build a subsequence of runs by applying successively Lemma 2. Those sub-sequences have the property that markings appearing in different runs are $\preccurlyeq$-ordered. The increasing sequences appear along the $2^{|\mathscr{P}|}+1$ first "columns".

Step 2 Finally, we show how to split and combine parts of runs appearing in the $\mathscr{M}_i$'s to obtain a run that allows the PN+NBA to accept a word of the desired form.
In order to build this sequence, we rely on several variables, namely: c_1, c_2 and n. At the present step of the proof, we present several constraints on c_1, c_2 and n. These constraints are meant to produce a sequence of transitions that accepts a word of the desired form. The main (and most technical) part of step 2 consists to show that these constraints are satisfiable.

Proof Let $\mathscr{N}$ be a PN+NBA with set of places $\mathscr{P}$ and initial marking $\mathbf{m}_{init}$ such that $\{B_i w_i^* E_i\} \subseteq L(\mathscr{N}, \mathscr{U})$.

Step 1 Since $\{B_i w_i^* E_i\} \subseteq L(\mathscr{N}, \mathscr{U})$, all the words of the form $B_i w_i^j E_i (j \geq 0)$ are accepted by $\mathscr{N}$. Let us consider the infinite sequence of the runs that accept all these words, for $i \geq 2^{|\mathscr{P}|}+1$:

$$
\begin{array}{l}
\mathbf{m}_{init} \xrightarrow{\upsilon_i} \mathbf{m}_i^1 \xrightarrow{\varsigma_i^1} \mathbf{m}_i^2 \xrightarrow{\varsigma_i^2} \cdots \xrightarrow{\varsigma_i^i} \mathbf{m}_i^{i+1} \xrightarrow{\upsilon_i'} \mathbf{n}_i \\
\mathbf{m}_{init} \xrightarrow{\upsilon_{i+1}} \mathbf{m}_{i+1}^1 \xrightarrow{\varsigma_{i+1}^1} \mathbf{m}_{i+1}^2 \xrightarrow{\varsigma_{i+1}^2} \cdots \xrightarrow{\varsigma_{i+1}^{i+1}} \mathbf{m}_{i+1}^{i+2} \xrightarrow{\upsilon_{i+1}'} \mathbf{n}_{i+1} \\
\cdots
\end{array}
$$

where for any $i \geq 2^{|\mathscr{P}|}+1$: $\Lambda(\upsilon_i) = B_i$, $\Lambda(\upsilon_i') = E_i$, $\mathbf{n}_i \in \mathscr{U}$ and for any $1 \leq j \leq i$: $\Lambda(\varsigma_i^j) = w_i$.

By applying Lemma 2 successively, we can construct an infinite subsequence of that sequence:

$$
\begin{array}{l}
\mathbf{m}_{init} \xrightarrow{\upsilon_{\rho(1)}} \mathbf{m}_{\rho(1)}^1 \xrightarrow{\varsigma_{\rho(1)}^1} \mathbf{m}_{\rho(1)}^2 \xrightarrow{\varsigma_{\rho(1)}^2} \cdots \xrightarrow{\varsigma_{\rho(1)}^{\rho(1)}} \mathbf{m}_{\rho(1)}^{\rho(1)+1} \xrightarrow{\upsilon_{\rho(1)}'} \mathbf{n}_{\rho(1)} \\
\mathbf{m}_{init} \xrightarrow{\upsilon_{\rho(2)}} \mathbf{m}_{\rho(2)}^1 \xrightarrow{\varsigma_{\rho(2)}^1} \mathbf{m}_{\rho(2)}^2 \xrightarrow{\varsigma_{\rho(2)}^2} \cdots \xrightarrow{\varsigma_{\rho(2)}^{\rho(2)}} \mathbf{m}_{\rho(2)}^{\rho(2)+1} \xrightarrow{\upsilon_{\rho(2)}'} \mathbf{n}_{\rho(2)} \\
\cdots
\end{array}
$$

such that, for any $1 \leq j \leq 2^{|\mathscr{P}|}+1$, the sequence $\mathbf{m}_{\rho(1)}^j \mathbf{m}_{\rho(2)}^j \ldots$ is increasing:

$$\forall 1 \leq j \leq 2^{|\mathscr{P}|}+1 : \forall k \geq 1 : \mathbf{m}_{\rho(k)}^j \preccurlyeq \mathbf{m}_{\rho(k+1)}^j \tag{8}$$

and, for any $1 \leq j \leq 2^{|\mathscr{P}|}+1$ there exists a set of places, noted $Places(j)$ that strictly increase along the sequence $\mathbf{m}_{\rho(1)}^j \mathbf{m}_{\rho(2)}^j \ldots$ while the other places stay constant:

$$\forall 1 \leq j \leq 2^{|\mathscr{P}|}+1 : \forall k \geq 1 : \mathbf{m}_{\rho(k)}^j(p) < \mathbf{m}_{\rho(k+1)}^j(p) \text{ iff } p \in Places(j) \tag{9}$$

Since, there are $2^{|\mathscr{P}|}$ subsets of $\mathscr{P}$, there exist $1 \leq c_1 < c_2 \leq 2^{|\mathscr{P}|}+1$ such that $Places(c_1) = Places(c_2)$.

In the following, we denote by $\sigma_{\rho(j)}(k_1,k_2)$ with $k_1 < k_2$ the sequence $\varsigma^{k_1}_{\rho(j)} \cdot \ldots \cdot \varsigma^{k_2-1}_{\rho(j)}$. We also denote by $\sigma_{\rho(j)}(\cdot,k)$, the sequence $\upsilon_{\rho(j)} \cdot \varsigma^{1}_{\rho(j)} \cdot \ldots \cdot \varsigma^{k-1}_{\rho(j)}$; and by $\sigma_{\rho(j)}(k,\cdot)$ the sequence $\varsigma^{k}_{\rho(j)} \cdot \ldots \cdot \varsigma^{\rho(j)}_{\rho(j)} \cdot \upsilon'_{\rho(j)}$

Step 2 The rest of the proof consists in devising a word of $L(\mathscr{N},\mathscr{U})$ that is of the form $B_{n_3} w^{i_1}_{n_3} w^{i_2}_{n_1} w^{i_3}_{n_2} E_{n_2}$, with $i_1 \geq 0$, $i_2 > 0$, $i_3 \geq 0$ and $0 < n_1 < n_2 < n_3$. The sequence of transitions that accepts this word (called σ) is built as follows:

$$\sigma = \sigma_{\rho(n)}(\cdot,c_1) \cdot \sigma_{\rho(1)}(c_1,c_2) \cdot \sigma_{\rho(2)}(c_2,\cdot)$$

for a well-chosen value of n. We next explain how to compute this value.

We choose $n > 2$ such that, when firing $\sigma_{\rho(1)}(c_1,c_2)$ from $\mathbf{m}^{c_1}_{\rho(n)}$, we reach a marking $\mathbf{m} \succcurlyeq \mathbf{m}^{c_2}_{\rho(2)}$. Let us show that such a n always exists. First, remark that for any $n > 2$: $\sigma_{\rho(1)}(c_1,c_2)$ is firable from $\mathbf{m}^{c_1}_{\rho(n)}$ since, by (8), $\mathbf{m}^{c_1}_{\rho(n)} \succcurlyeq \mathbf{m}^{c_1}_{\rho(1)}$. Let k be the number of non-blocking arcs in $\sigma_{\rho(1)}(c_1,c_2)$. By Lemma 9, we have that

$$\forall p \in \mathscr{P} : \mathbf{m}(p) \geq \mathbf{m}^{c_1}_{\rho(n)}(p) + \mathbf{m}^{c_2}_{\rho(1)}(p) - \mathbf{m}^{c_1}_{\rho(1)}(p) - k \tag{10}$$

But, since $Places(c_1) = Places(c_2)$, we can state the following. For any place $p \in Places(c_1)$ and for any $n \geq 1$: $\mathbf{m}^{c_1}_{\rho(n)}(p) \geq n-1$, since by (9) the sequence $\mathbf{m}^{c_2}_{\rho(1)}(p), \mathbf{m}^{c_2}_{\rho(2)}(p), \ldots$ is strictly increasing. In particular, if we choose n such that

$$n > \max_{p \in Places(c_1)} \left(\mathbf{m}^{c_2}_{\rho(2)}(p) - \mathbf{m}^{c_2}_{\rho(1)}(p) + \mathbf{m}^{c_1}_{\rho(1)}(p) \right) + k$$

we have $\forall p \in Places(c_1) : \mathbf{m}^{c_1}_{\rho(n)}(p) \geq \mathbf{m}^{c_2}_{\rho(2)}(p) - \mathbf{m}^{c_2}_{\rho(1)}(p) + \mathbf{m}^{c_1}_{\rho(1)}(p) + k$ and thus:

$$\forall p \in Places(c_1) : \mathbf{m}^{c_1}_{\rho(n)}(p) + \mathbf{m}^{c_2}_{\rho(1)}(p) - \mathbf{m}^{c_1}_{\rho(1)}(p) - k \geq \mathbf{m}^{c_2}_{\rho(2)}(p) \tag{11}$$

By (10) and (11), we obtain:

$$\forall p \in Places(c_1) : \mathbf{m}(p) \geq \mathbf{m}^{c_2}_{\rho(2)}(p) \tag{12}$$

On the other hand, for any place p, the monotonicity property of PN+NBA implies that $\mathbf{m}(p) \geq \mathbf{m}^{c_2}_{\rho(1)}(p)$. And since, by (9): $\forall p \in \mathscr{P} \setminus Places(c_1) : \mathbf{m}^{c_2}_{\rho(1)}(p) = \mathbf{m}^{c_2}_{\rho(2)}(p)$, we obtain:

$$\forall p \in \mathscr{P} \setminus Places(c_1) : \mathbf{m}(p) \geq \mathbf{m}^{c_2}_{\rho(2)}(p) \tag{13}$$

By (12) and (13), we conclude that $\mathbf{m} \succcurlyeq \mathbf{m}^{c_2}_{\rho(2)}$.

Thus, the sequence of transitions $\sigma = \sigma_{\rho(n)}(\cdot,c_1) \cdot \sigma_{\rho(1)}(c_1,c_2) \cdot \sigma_{\rho(2)}(c_2,\cdot)$ is firable from $\mathbf{m}_{init}$ (with n computed as explained above) and leads to a marking $\mathbf{m}'$, i.e $\mathbf{m}_{init} \xrightarrow{\sigma} \mathbf{m}'$. Since $\mathbf{m} \succcurlyeq \mathbf{m}^{c_2}_{\rho(2)}$, we also have that $\mathbf{m}' \succcurlyeq \mathbf{n}_{\rho(2)}$, by monotonicity.

Hence $\mathbf{m}' \in \mathscr{U}$, and the word $\Lambda(\sigma) \in L(\mathscr{N}, \mathscr{U})$. It is not difficult to see that by the previous construction this word is of the form: $B_{n_3} w_{n_3}^{i_1} w_{n_1}^{i_2} w_{n_2}^{i_3} E_{n_2}$ with *(i)* $n_3 = \rho(n)$, $n_1 = \rho(1)$ and $n_2 = \rho(2)$, hence $0 < n_1 < n_2 < n_3$, and *(ii)* $i_1 \geq 0$, $i_2 = c_2 - c_1 > 0$, $i_3 \geq 0$. □

5 Properties of WSL

In this section, we apply the pumping lemmata of the previous section to obtain several results about WSL and languages of EPN. Section 5.1 presents properties of WSL that can be proved thanks to Lemma 7. Then, the pumping lemmata on PN and PN+NBA are exploited in sections 5.2 and 5.3 to prove a strict hierarchy among the languages of PN, PN+NBA and PN+T; as well as in section 5.4, to obtain closure properties of languages of EPN.

5.1 Consequences of Lemma 7

We first study several classical languages and show that they are not well-structured. These languages are: the set of all words of the form $\mathtt{a}^n\mathtt{b}^n$, the set of all words of the form $\mathtt{a}^n\mathtt{b}^m$ with $m \geq n$, and the set of all palindromes.

- $\mathscr{L} = \{\mathtt{a}^n\mathtt{b}^n | n \geq 1\} \notin L^G(\mathsf{WSTS})$. Suppose that $\mathscr{L} \in L^G(\mathsf{WSTS})$. Since, $\forall k \geq 1 : \mathtt{a}^k\mathtt{b}^k \in \mathscr{L}$, we can apply Lemma 7 (letting $B_k = \mathtt{a}^k$ and $E_k = \mathtt{b}^k$, for any $k \geq 1$). We conclude that there is $i < j$ s.t. $\mathtt{a}^j\mathtt{b}^i \in \mathscr{L}$, which is a contradiction. Notice that this results is also a consequence of Theorem 3 and Theorem 2, following the reasoning given in [11, pages 175–176].
- $\mathscr{L}^{\geq} = \{\mathtt{a}^n\mathtt{b}^m | m \geq n\} \notin L^G(\mathsf{WSTS})$. The proof is similar to the previous one.
- $\mathscr{L}^R = \{w \cdot w^R\} \notin L^G(\mathsf{WSTS})$. Let Σ be an alphabet and $w = \mathtt{a}_1 \cdot \ldots \cdot \mathtt{a}_n \in \Sigma^*$, we define the mirror of w, as the word $w^R = \mathtt{a}_n \cdot \ldots \cdot \mathtt{a}_1$. Let us suppose $\mathscr{L}^R \in L^G(\mathsf{WSTS})$. Since $\{\mathtt{a}^n\mathtt{bba}^n \mid n \geq 0\} \subseteq \mathscr{L}^R$, we can apply Lemma 7 (letting $B_k = \mathtt{a}^k\mathtt{b}$ and $E_k = \mathtt{ba}^k$, for $k \geq 1$). We conclude that there exist $i < j$ such that $\mathtt{a}^j\mathtt{bba}^i \in \mathscr{L}^R$, which is a contradiction. Hence $\mathscr{L}^R \notin L^G(\mathsf{WSTS})$.

These results allow us to show that neither the class of WSL, nor $L^G(\mathsf{PN})$, nor $L^G(\mathsf{PN{+}NBA})$, nor $L^G(\mathsf{PN{+}T})$ are not closed under complement.

Proposition 2 *$L^G(\mathsf{WSTS})$, $L^G(\mathsf{PN})$, $L^G(\mathsf{PN{+}NBA})$ and $L^G(\mathsf{PN{+}T})$ are not closed under complement.*

Proof It is not difficult to devise a PN $\mathscr{N}$ and an $\preccurlyeq$-upward-closed set $\mathscr{U}$ such that $L(\mathscr{N}, \mathscr{U}) = \{\mathtt{a}^n\mathtt{b}^m \mid m < n\}$. It is well-known [11] that $L^G(\mathsf{PN})$ is closed under union and that the regular languages are all in $L^G(\mathsf{PN})$. Hence, $\{\mathtt{a}^n\mathtt{b}^m \mid m < n\} \cup \big((\mathtt{a}+\mathtt{b})^* \setminus \mathtt{a}^*\mathtt{b}^*\big)$ is in $L^G(\mathsf{PN})$, but also in PN+NBA and in PN+T, since PN is a syntactic subclass of theirs. However, its complement is $\mathscr{L}^{\geq} = \{\mathtt{a}^n\mathtt{b}^m \mid m \geq n\}$, which is not a WSL. □

Finally, we can also exploit the previous results to show that the class of WSL is incomparable to the class of Context Free Languages (C.F.L., for short).

Proposition 3 *The class $L^G(\mathsf{WSTS})$ is incomparable to the class of context-free languages.*

Proof C.F.L. $\not\subseteq L^G(\mathsf{WSTS})$ stems from the fact that $\mathscr{L}$, which is well-known to be a C.F.L., is not in $L(\mathsf{WSTS})$. We prove that $L^G(\mathsf{WSTS}) \not\subseteq$ C.F.L. thanks to $\mathscr{L}_1 = \{\mathtt{a}^i\mathtt{b}^j\mathtt{c}^k \mid i \geq j \geq k \geq 0\}$. It is not difficult to devise a PN that accepts $\mathscr{L}_1$ for some $\preccurlyeq$-upward-closed set. On the other hand, we prove that $\mathscr{L}_1$ is not a C.F.L. thanks to the classical pumping lemma for C.F.L.

For that purpose, we have to devise, for any constant $n \in \mathbb{N}$, a word $\omega_n \in \mathscr{L}_1$ such that $|\omega_n| \geq n$ and, for any words u, v, w, x and y respecting (i) $\omega = u\cdot v\cdot w\cdot x\cdot y$, (ii) $|v\cdot w\cdot x| \leq n$ and (iii) $|v\cdot x| > 0$, we can find $i \geq 0$ s.t. $u\cdot v^i\cdot w\cdot x^i\cdot y \notin \mathscr{L}_1$.

For any $n \geq 0$, we let $\omega_n = \mathtt{a}^n\mathtt{b}^n\mathtt{c}^n$. Clearly $\omega_n \in \mathscr{L}_1$ and $|\omega_n| \geq n$, for any n. Let us consider all the possible values of $u, v, \ldots, y$ that respect the three conditions above, and let us show that, for all these values, there exists a $i \geq 0$ such that $u\cdot v^i\cdot w\cdot x^i\cdot y \notin \mathscr{L}_1$.

- *If either v or x contain at least two different characters*, the word $u\cdot v^2\cdot w\cdot x^2\cdot y$ is clearly not a word of $\mathscr{L}_1$.
- *If $v \in \mathtt{a}^*$*, then, since $|v\cdot w\cdot x| \leq n$, there are two possibilities. Either $x \in \mathtt{a}^*$. In that case, we choose $i = 0$ and the word $u\cdot v^0\cdot w\cdot x^0\cdot y$ is of the form $\mathtt{a}^{n-|v\cdot x|}\mathtt{b}^n\mathtt{c}^n$, and is clearly not in $\mathscr{L}_1$, since $|v\cdot x| > 0$. Otherwise, $x \in \mathtt{b}^*$. In that case, we choose $i = 0$ again and we obtain a word of the form $\mathtt{a}^{n-|v|}\mathtt{b}^{n-|x|}\mathtt{c}^n$, which is not in $\mathscr{L}_1$ because $|v\cdot x| > 0$.
- *If $v \in \mathtt{b}^*$*, there are two possibilities again. Either $x \in \mathtt{b}^*$, and by choosing $i = 0$ we obtain $\mathtt{a}^n\mathtt{b}^{n-|v\cdot x|}\mathtt{c}^n \notin \mathscr{L}_1$. Or, $x \in \mathtt{c}^*$, we choose $i = 2$ and obtain $\mathtt{a}^n\mathtt{b}^{n+|v|}\mathtt{c}^{n+|x|} \notin \mathscr{L}_1$.
- *If $v \in \mathtt{c}^*$*, then, $x \in \mathtt{c}^*$, and, by choosing $i = 2$, we obtain $\mathtt{a}^n\mathtt{b}^n\mathtt{c}^{n+|v\cdot x|} \notin \mathscr{L}_1$. □

5.2 PN+NBA are more expressive than PN

In this section we prove that the class of languages accepted by PN+NBA strictly contains the class of languages accepted by PN (when the acceptance condition is an $\preccurlyeq$-upward-closed set). Since the class of PN form a syntactic subclass of PN+NBA, we obtain this result by showing that there is a language accepted by a PN+NBA that cannot be accepted by any PN.

Separation of PN+NBA *and* PN The strategy adopted in the proof is as follows. We look into the PN+NBA $\mathscr{N}_1$ of Fig. 7 with initial marking $\mathbf{m}_0$ such that $\mathbf{m}_0(p_1) = 1$ and $\mathbf{m}_0(p) = 0$ for $p \in \{p_2, p_3, p_4, p_5, p_6\}$, and prove it accepts every word of the form $\mathtt{i}^k\mathtt{s}\left(\mathtt{a}^k\mathtt{c}\mathtt{b}^k\mathtt{d}\right)^j$, for $k \geq 0$ and $j \geq 0$ (Lemma 11), but not those of the form $\mathtt{i}^{n_3}\mathtt{s}\mathtt{a}^{n_3}\mathtt{c}(\mathtt{b}^{n_3}\mathtt{d}\mathtt{a}^{n_3}\mathtt{c})^{i_1}\left(\mathtt{b}^{n_1}\mathtt{d}\mathtt{a}^{n_1}\mathtt{c}\right)^k\left(\mathtt{b}^{n_2}\mathtt{d}\mathtt{a}^{n_2}\mathtt{c}\right)^{i_2}\mathtt{b}^{n_2}\mathtt{d}$, for k big enough, and $0 < n_1 < n_2 < n_3$ (Lemma 12). Then we invoke Lemma 8 (pumping lemma on PN) to prove that every PN accepting the words of the first form also accepts words of the latter, which implies that no PN accepts $L(\mathscr{N}_1, \mathbb{N}^6)$.

Lemma 11 *For any $k \geq 0$, for any $j \geq 0$, the word $\mathtt{i}^k\mathtt{s}\left(\mathtt{a}^k\mathtt{c}\mathtt{b}^k\mathtt{d}\right)^j$ is in $L(\mathscr{N}_1, \mathbb{N}^6)$.*

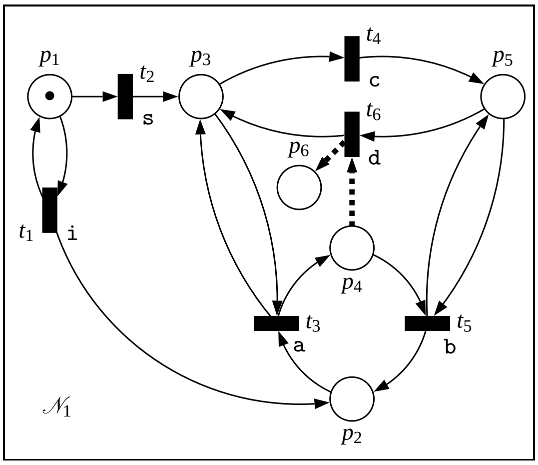

Fig. 7 The PN+NBA used in the proof of Theorem 5.

Proof Remark that, since the $\preccurlyeq$-upward-closed set considered here is $\mathbb{N}^6$, we just need to show that a sequence of transitions labelled by $\mathtt{i}^k\mathtt{s}\big(\mathtt{a}^k\mathtt{c}\mathtt{b}^k\mathtt{d}\big)^j$ is firable in $\mathcal{N}_1$ to get the Lemma.

The following holds for any $k \geq 0$. After firing the transitions $t_1^k t_2$ from the initial marking of $\mathcal{N}_1$, we reach the marking $\mathbf{m}_1$ such that $\mathbf{m}_1(p_2) = k$, $\mathbf{m}_1(p_3) = 1$, and $\mathbf{m}_1(p_j) = 0$ for $j \in \{1,4,5,6\}$. Then, we can fire $t_3^k t_4$ from $\mathbf{m}_1$. This leads to the marking $\mathbf{m}_2$ such that $\mathbf{m}_2(p_4) = k$, $\mathbf{m}_2(p_5) = 1$, and $\mathbf{m}_2(p_j) = 0$ for $j \in \{1,2,3,6\}$. From $\mathbf{m}_2$, t_5^k can be fired. This sequence of transitions moves the k tokens from p_4 to p_2. Then, from the resulting marking, t_6 can be fired. Since, p_4 is now empty, the effect of t_6 only consists in moving the token from p_5 to p_3 (its non-blocking arc has no effect) and we reach $\mathbf{m}_1$ again. Thus, the sequence of transitions $t_3^k t_4 t_5^k t_6$, labelled by $\mathtt{a}^k\mathtt{c}\mathtt{b}^k\mathtt{d}$, can be fired arbitrarily often from $\mathbf{m}_1$, and reaches the same marking. Hence the word $\mathtt{i}^k\mathtt{s}\big(\mathtt{a}^k\mathtt{c}\mathtt{b}^k\mathtt{d}\big)^j$ is in $L(\mathcal{N}_1, \mathbb{N}^6)$, for any $k \geq 0$, any $j \geq 0$. □

Lemma 12 *Let* n_1, n_2 *and* n_3 *be three natural numbers such that* $0 < n_1 < n_2 < n_3$. *The words*

$$\mathtt{i}^{n_3}\mathtt{s}\mathtt{a}^{n_3}\mathtt{c}(\mathtt{b}^{n_3}\mathtt{d}\mathtt{a}^{n_3}\mathtt{c})^{i_1}\big(\mathtt{b}^{n_1}\mathtt{d}\mathtt{a}^{n_1}\mathtt{c}\big)^k\big(\mathtt{b}^{n_2}\mathtt{d}\mathtt{a}^{n_2}\mathtt{c}\big)^{i_2}\mathtt{b}^{n_2}\mathtt{d}$$

are not in $L(\mathcal{N}_1, \mathbb{N}^6)$*, for all* $i_1 \geq 0$*,* $k \geq n_3 - n_1$ *and* $i_2 \geq 0$*.*

Proof In this proof, we will identify a sequence of transitions with the word it accepts (all the transitions have different labels). Clearly (see the proof of Lemma 11), for any $n_3 \geq 0$, $m \geq 0$, the firing of $\mathtt{i}^{n_3}\mathtt{s}\big(\mathtt{a}^{n_3}\mathtt{c}\mathtt{b}^{n_3}\mathtt{d}\big)^m$ from $\mathbf{m}_0$ leads to a marking $\mathbf{m}_1$ such that $\mathbf{m}_1(p_2) = n_3$, $\mathbf{m}_1(p_3) = 1$, and $\forall i \in \{1,4,5,6\} : \mathbf{m}_1(p_i) = 0$ (the non-blocking arc of t_6 hasn't consumed any token in p_4). By firing $\mathtt{a}^{n_3}\mathtt{c}\mathtt{b}^{n_1}\mathtt{d}$ from $\mathbf{m}_1$, we now have n_1 tokens in p_2, $n_3 - n_1 - 1$ tokens in p_4 and one token in p_6 (this time the non-blocking arc has moved one token since $n_1 < n_3$). Clearly, at each subsequent firing of $\mathtt{a}^{n_1}\mathtt{c}\mathtt{b}^{n_1}\mathtt{d}$, the non-blocking arc of t_6 will remove one token from p_4 and the marking of this place will strictly decrease until p_4 becomes empty. Let $\ell = n_3 - n_1 - 1$. It is easy to see that that firing $\mathtt{a}^{n_3}\mathtt{cb}^{n_1}\mathtt{d}\left(\mathtt{a}^{n_1}\mathtt{cb}^{n_1}\mathtt{d}\right)^{\ell}$ from $\mathbf{m}_1$ leads to a marking $\mathbf{m}_2$ with $\mathbf{m}_2(p_2) = n_1$, $\mathbf{m}_2(p_3) = 1$, $\mathbf{m}_2(p_6) = n_3 - n_1$ and $\forall j \in \{1,4,5\} : \mathbf{m}_2(p_j) = 0$. This characterization also implies that we can fire $\mathtt{a}^{n_1}\mathtt{cb}^{n_1}\mathtt{d}$ an arbitrary number of times from $\mathbf{m}_2$ because $\mathbf{m}_2 \xrightarrow{\mathtt{a}^{n_1}\mathtt{cb}^{n_1}\mathtt{d}} \mathbf{m}_2$. On the other hand, it is not possible to fire $\mathtt{a}^{n_1}\mathtt{cb}^{n_2}\mathtt{d}$, with $n_2 > n_1$, from $\mathbf{m}_2$. Indeed $\mathbf{m}_2 \xrightarrow{\mathtt{a}^{n_1}\mathtt{cb}^{n_1}} \mathbf{m}_3$, with $\mathbf{m}_3(p_5) = 1$, $\mathbf{m}_3(p_2) = n_1$, $\mathbf{m}_3(p_6) = n_3 - n_1$ and $\forall j \in \{1,3,4\} : \mathbf{m}_3(p_j) = 0$, which does not allow to fire the $\mathtt{b}$-labelled transition t_5 anymore. We conclude that, $\forall k \geq n_3 - n_1$, a sequence labelled by $\mathtt{i}^{n_3}\mathtt{s}\left(\mathtt{a}^{n_3}\mathtt{cb}^{n_3}\mathtt{d}\right)^m\mathtt{a}^{n_3}\mathtt{c}\left(\mathtt{b}^{n_1}\mathtt{da}^{n_1}\mathtt{c}\right)^k\mathtt{b}^{n_2}\mathtt{da}^{n_2}\mathtt{c}$, is not firable in $\mathcal{N}_1$. Thus, we will not find in $L(\mathcal{N}_1, \mathbb{N}^6)$ any word with this prefix, hence the Lemma. □

Thanks to these lemmata, we can prove Proposition 4.

Proposition 4 *There is no* PN $\mathcal{N}$ *with an* $\preccurlyeq$*-upward-closed set* $\mathcal{U}$ *such that* $L(\mathcal{N}, \mathcal{U}) = L(\mathcal{N}_1, \mathbb{N}^6)$.

Proof By Lemma 11, any PN $\mathcal{N}$ s.t. $L(\mathcal{N}, \mathcal{U}) = L(\mathcal{N}_1, \mathbb{N}^6)$ for some $\preccurlyeq$-upward-closed set of accepting markings $\mathcal{U}$, must accept $\mathtt{i}^k\mathtt{s}\left(\mathtt{a}^k\mathtt{cb}^k\mathtt{d}\right)^j$, for any $k \geq 1$ and $j \geq 0$. Hence, we can apply Lemma 8, by letting $B_k = \mathtt{i}^k\mathtt{sa}^k\mathtt{c}$, $E_k = \mathtt{b}^k\mathtt{d}$ and $w_k = \mathtt{b}^k\mathtt{da}^k\mathtt{c}$, for any $k \geq 1$. We conclude that $\mathcal{N}$ also accepts a word of the form: $\mathtt{i}^{n_3}\mathtt{sa}^{n_3}\mathtt{c}\left(\mathtt{b}^{n_3}\mathtt{da}^{n_3}\mathtt{c}\right)^{i_1}\left(\mathtt{b}^{n_1}\mathtt{da}^{n_1}\mathtt{c}\right)^{L'}\left(\mathtt{b}^{n_2}\mathtt{da}^{n_2}\mathtt{c}\right)^{i_2}\mathtt{b}^{n_2}\mathtt{d}$ such that $0 < n_1 < n_2 < n_3$ and $L' \geq n_3 - n_1$. Since it is not in $L(\mathcal{N}_1, \mathcal{U})$, by Lemma 12, there is no PN $\mathcal{N}$ and no $\preccurlyeq$-upward-closed set $\mathcal{U}$ s.t. $L(\mathcal{N}, \mathcal{U}) = L(\mathcal{N}_1, \mathbb{N}^6)$. □

Thus, we conclude that:

Theorem 5 $L^G(\mathsf{PN}) \subset L^G(\mathsf{PN{+}NBA})$.

Proof $L^G(\mathsf{PN}) \subseteq L^G(\mathsf{PN{+}NBA})$ is trivial since PN is a syntactic subclass of PN+NBA. The strictness of the inclusion is given by Proposition 4. □

5.3 PN+T are more expressive than PN+NBA

Let us now prove a similar result about the classes PN+NBA and PN+T: the class of languages that can be accepted by PN+T strictly contains the class of languages accepted by PN+NBA. For this purpose, we first show that a PN+T can always *simulate* a PN+NBA, hence $L^G(\mathsf{PN{+}NBA}) \subseteq L^G(\mathsf{PN{+}T})$. Then, we prove, thanks to Lemma 10, that there is a language that can be recognized by a PN+T, but not by a PN+NBA, which implies the strictness of the inclusion.

Simulation of a PN+NBA *by a* PN+T Lemma 13 below states that any PN+NBA can be simulated by a PN+T. The proof of this lemma is based on the following construction. Let us consider a PN+NBA $\mathcal{N} = \langle \mathcal{P}, \mathcal{T}, \Sigma, \mathbf{m}_0 \rangle$, and an $\preccurlyeq$-upward-closed set $\mathcal{U}$ of markings, and let us show how to transform them into a PN+T $\mathcal{N}'$ and an $\preccurlyeq$-upward-closed set $\mathcal{U}'$ such that $L^G(\mathcal{N}, \mathcal{U}) = L^G(\mathcal{N}', \mathcal{U}')$.

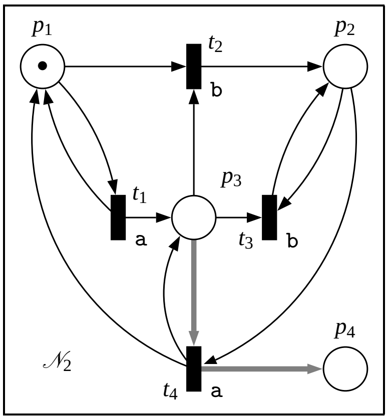

Fig. 8 The PN+T used in the proof of Theorem 6.

Let us consider the partition of $\mathscr{T}$ into $\mathscr{T}_e$ and $\mathscr{T}_r$ as defined in Section 2, and a new place p_{Tr} (the trash place). We show now how to build $\mathscr{N}' = \langle \mathscr{P}', \mathscr{T}', \Sigma, \mathbf{m}'_0 \rangle$ and $\mathscr{U}'$. First, $\mathscr{P}' = \mathscr{P} \cup \{p_{Tr}\}$. For each transition $t = \langle I, O, s, d, 1, \lambda \rangle$ in $\mathscr{T}_e$, we put in $\mathscr{T}'$: $t_l = \langle I, O, s, p_{Tr}, +\infty, \lambda \rangle$ and $t_e = \langle I_e, O_e, \perp, \perp, 0, \lambda \rangle$, two new transitions, such that: $\forall p \in \mathscr{P} : \big(p \neq s \Rightarrow I_e(p) = I(p) \wedge p \neq d \Rightarrow O_e(p) = O(p)\big)$, $I_e(s) = I(s) + 1$ and $O_e(d) = O(d) + 1$. We also add into $\mathscr{T}'$ all the transitions of $\mathscr{T}_r$ (extended to p_{Tr} such that they have no guard and no effect on p_{Tr}). Finally, $\forall p \in \mathscr{P} = \mathbf{m}'_0(p) = \mathbf{m}_0(p)$, $\mathbf{m}'_0(p_{Tr}) = 0$ and $\mathscr{U}' = \{\mathbf{m} \mid \exists \mathbf{m}' \in \mathscr{U} : \forall p \in \mathscr{P} : \mathbf{m}(p) = \mathbf{m}'(p)\}$.

Example 2 Fig. 9 illustrates the above construction. $\diamond$

Lemma 13 *For any* PN+NBA $\mathscr{N}$ *with an* $\preccurlyeq$*-upward-closed set* $\mathscr{U}$*, it is possible to construct a* PN+T $\mathscr{N}'$ *and an* $\preccurlyeq$*-upward closed set* $\mathscr{U}'$ *s.t.:* $L(\mathscr{N}, \mathscr{U}) = L(\mathscr{N}', \mathscr{U}')$.

Proof Let us consider the previous construction and let us prove that $L(\mathscr{N}, \mathscr{U}) = L(\mathscr{N}', \mathscr{U}')$.

$\boxed{L(\mathscr{N}, \mathscr{U}) \subseteq L(\mathscr{N}', \mathscr{U}')}$ We show that, for every sequence of transitions σ of $\mathscr{N}$ that leads into a marking $\mathbf{m} \in \mathscr{U}$, we can find a sequence of transitions σ' of $\mathscr{N}'$ that leads into a marking $\mathbf{m}' \in \mathscr{U}'$ such that $\Lambda(\sigma) = \Lambda(\sigma')$.

Let us define the function $f : \mathscr{T} \times \mathbb{N}^{|\mathscr{P}|} \to \mathscr{T}'$ such that $\forall t \in \mathscr{T}_r : f(t, \mathbf{m}) = t$ and $\forall t = \langle O, I, s, d, 1, \lambda \rangle \in \mathscr{T}_e : f(t, \mathbf{m}) = t_e$, if $\mathbf{m}(s) > I(s)$ (the non-blocking arc still has an effect after the firing of the Petri part of the transition); and $f(t, \mathbf{m}) = t_l$, otherwise.

Let $\sigma = \mathbf{m}_0 \xrightarrow{t_1} \mathbf{m}_1 \xrightarrow{t_2} \dots \xrightarrow{t_n} \mathbf{m}_n \xrightarrow{t_{n+1}} \mathbf{m}_{n+1}$ be a sequence of $\mathscr{N}$ such that $\mathbf{m}_{n+1} \in \mathscr{U}$. Then we may see that $\sigma' = \mathbf{m}'_0 \xrightarrow{f(t_1, \mathbf{m}'_0)} \mathbf{m}'_1 \xrightarrow{f(t_2, \mathbf{m}'_1)} \dots \xrightarrow{f(t_n, \mathbf{m}'_{n-1})} \mathbf{m}'_n \xrightarrow{f(t_{n+1}, \mathbf{m}'_n)} \mathbf{m}'_{n+1}$ is a sequence of $\mathscr{N}'$, where $\forall 1 \leq i \leq n+1 : \mathbf{m}'_i$ is such that $\mathbf{m}'_i(p) = \mathbf{m}_i(p)$ for all $p \in \mathscr{P}$ and $\mathbf{m}'_i(p_{Tr}) = 0$. Hence, $\mathbf{m}'_{n+1} \in \mathscr{U}'$ and $\Lambda(\sigma')$ is accepted. Since we have $\forall 1 \leq i \leq n+1 : \Lambda(t_i) = \Lambda(f(t_i, \mathbf{m}_{i-1}))$, we conclude that $\Lambda(\sigma) = \Lambda(\sigma')$, hence $L(\mathcal{N}, \mathcal{U}) \subseteq L(\mathcal{N}', \mathcal{U}')$.

$\boxed{L(\mathcal{N}', \mathcal{U}') \subseteq L(\mathcal{N}, \mathcal{U})}$ We show that, for every sequence of transitions σ' of $\mathcal{N}'$ that leads into a marking $\mathbf{m}' \in \mathcal{U}'$, we can find a sequence of transitions σ of $\mathcal{N}$ that leads into a marking $\mathbf{m} \in \mathcal{U}$ such that $\Lambda(\sigma') = \Lambda(\sigma)$.

We define the function $g : \mathcal{T}' \to \mathcal{T}$ such that for all $t \in \mathcal{T}_r$: $g(t) = t$ and for all $t \in \mathcal{T}_e : g(t_e) = g(t_l) = t$. Moreover, we define the relation $\preccurlyeq_{\mathcal{P}}$ that compares two markings only on the places that are in $\mathcal{P}$. Thus, if $\mathbf{m}$ is defined on $\mathcal{P}$ and $\mathbf{m}'$ on $\mathcal{P}'$ (remember that $\mathcal{P} \subseteq \mathcal{P}'$), $\mathbf{m}' \preccurlyeq_{\mathcal{P}} \mathbf{m}$ iff $\forall p \in \mathcal{P} : \mathbf{m}'(p) \leq \mathbf{m}(p)$.

Let $\sigma' = \mathbf{m}'_0 \xrightarrow{t_1} \mathbf{m}'_1 \xrightarrow{t_2} \dots \xrightarrow{t_n} \mathbf{m}'_n \xrightarrow{t_{n+1}} \mathbf{m}'_{n+1}$ be a sequence of $\mathcal{N}'$ such that $\mathbf{m}'_{n+1} \in \mathcal{U}'$. Then, there exist $\mathbf{m}_1, \mathbf{m}_2, \dots \mathbf{m}_{n+1}$ in $\mathcal{N}$ such that we have $\mathbf{m}_0 \xrightarrow{g(t_1)} \mathbf{m}_1 \xrightarrow{g(t_2)} \dots \xrightarrow{g(t_n)} \mathbf{m}_n \xrightarrow{g(t_{n+1})} \mathbf{m}_{n+1}$ and $\mathbf{m}_{n+1} \in \mathcal{U}$. To prove that the sequence of markings exists, we show by induction on the indexes, that $\mathbf{m}'_i \preccurlyeq_{\mathcal{P}} \mathbf{m}_i$ for all i such that $0 \leq i \leq n+1$. That implies that $\forall 1 \leq i \leq n+1 : g(t_i)$ is firable from $\mathbf{m}_{i-1}$ because $g(t_i)$ consumes no more tokens in any place p than t_i does.

Base case: $j = 0$. The base case is trivially verified.

Induction step: $j = k$. By induction hypothesis, we have: $\forall 0 \leq j \leq k-1 : \mathbf{m}'_j \preccurlyeq_{\mathcal{P}} \mathbf{m}_j$. In the case where $t_k = \langle I, O, s, d, b, \lambda \rangle$ (from $\mathbf{m}'_{k-1}$) has the same effect on $\mathcal{P}$ than $g(t_k)$ (from $\mathbf{m}_{k-1}$), we directly have that $\mathbf{m}'_k \preccurlyeq_{\mathcal{P}} \mathbf{m}_k$. This happens if t_k is a regular Petri transition or if $\mathbf{m}_{k-1}(s) = \mathbf{m}'_{k-1}(s) = I(s)$.

Otherwise t_k has a transfer arc and we must consider two cases:

- The transfer of t_k has no effect and the non-blocking arc of $g(t_k)$ moves one token from the source s to the target d, hence $I(s) = \mathbf{m}'_{k-1}(s) < \mathbf{m}_{k-1}(s)$. Since t_k and $g(t_k)$ have the same effect except that $g(t_k)$ removes one more token from s and adds one more token in d, and since $\mathbf{m}'_{k-1} \preccurlyeq_{\mathcal{P}} \mathbf{m}_{k-1}$ with $\mathbf{m}'_{k-1}(s) < \mathbf{m}_{k-1}(s)$, we conclude that $\mathbf{m}'_k \preccurlyeq_{\mathcal{P}} \mathbf{m}_k$.
- The transfer of t_k moves at least one token from the source s to p_{Tr} and the non-blocking arc of $g(t_k)$ moves one token from s to d. Since t_k and $g(t_k)$ have the same effect on the places in $\mathcal{P}$ except that $g(t_k)$ adds one more token in d and t_k may remove more tokens from s, and since $\mathbf{m}'_{k-1} \preccurlyeq_{\mathcal{P}} \mathbf{m}_{k-1}$, we conclude that $\mathbf{m}'_k \preccurlyeq_{\mathcal{P}} \mathbf{m}_k$.

Thus, there are $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{n+1}$ s.t. $\mathbf{m}_0 \xrightarrow{g(t_1)} \mathbf{m}_1 \xrightarrow{g(t_2)} \dots \xrightarrow{g(t_n)} \mathbf{m}_n \xrightarrow{g(t_{n+1})} \mathbf{m}_{n+1}$ in $\mathcal{N}$ and $\forall 1 \leq i \leq n+1 : \mathbf{m}'_i \preccurlyeq_{\mathcal{P}} \mathbf{m}_i$. Thus, $\mathbf{m}_{n+1} \in \mathcal{U}$. Since $\Lambda(t_i) = \Lambda(g(t_i))$ for all $1 \leq i \leq n+1$, we conclude that $\Lambda(\sigma') = \Lambda(\sigma)$, hence $L(\mathcal{N}', \mathcal{U}') \subseteq L(\mathcal{N}, \mathcal{U})$. □

Separation of PN+T *and* PN+NBA Let us now prove that $L^G(\text{PN+NBA})$ is *strictly* included in $L^G(\text{PN+T})$. We consider the PN+T $\mathcal{N}_2$ presented in Fig.8 with the initial marking $\mathbf{m}_0(p_1) = 1$ and $\mathbf{m}_0(p) = 0$ for $p \in \{p_2, p_3, p_4\}$. The two following Lemmata allow us to better understand the behaviour of $\mathcal{N}_2$.

Lemma 14 *For any $k \geq 1$, for any $j \geq 0$, the word $\left(\mathtt{a}^k \mathtt{b}^k\right)^j$ is in $L(\mathcal{N}_2, \mathbb{N}^4)$.*

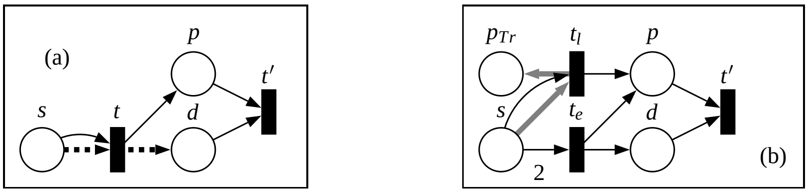

Fig. 9 A PN+NBA $\mathcal{N}$ (a) and the corresponding PN+T $\mathcal{N}'$ (b)

Proof Remark that, since the $\preccurlyeq$-upward-closed set considered here is $\mathbb{N}^4$, we just need to show that a sequence of transitions labelled by $\left(\mathtt{a}^k\mathtt{b}^k\right)^j$ ($j \geq 0$) is firable in $\mathcal{N}_2$ to get the lemma.

The following holds for any $k \geq 1$. From the initial marking $\mathbf{m}_0$ of $\mathcal{N}_2$, we can fire $t_1^k t_2 t_3^{k-1}$ (which is labelled by $\mathtt{a}^k\mathtt{b}^k$), and obtain the marking $\mathbf{m}_1$ such that $\mathbf{m}_1(p_2) = 1$ and $\forall p \in \{p_1, p_3, p_4\} : \mathbf{m}_1(p) = 0$. Thus, t_4 is firable from $\mathbf{m}_1$ and does not transfer any token, but produces a token in p_3 and moves the token from p_2 to p_1. It is thus not difficult to see that $t_4 t_1^{k-1} t_2 t_3^{k-1}$, labelled by $\mathtt{a}^k\mathtt{b}^k$, can be fired from $\mathbf{m}_1$. The marking one obtains is $\mathbf{m}_1$ again. Hence, we can fire a sequence labelled by $\mathtt{a}^k\mathtt{b}^k$ arbitrarily often from $\mathbf{m}_1$. Thus, any word of the form $\left(\mathtt{a}^k\mathtt{b}^k\right)^j$ is in $L(\mathcal{N}_2, \mathbb{N}^4)$. □

Lemma 15 *Let n_1, n_2, n_3 be three natural numbers such that $0 < n_1 < n_2 < n_3$. For any $i_1 \geq 0$, $i_2 > 0$ and $i_3 \geq 0$, the words of the form:*

$$\mathtt{a}^{n_3}(\mathtt{b}^{n_3}\mathtt{a}^{n_3})^{i_1}(\mathtt{b}^{n_1}\mathtt{a}^{n_1})^{i_2}(\mathtt{b}^{n_2}\mathtt{a}^{n_2})^{i_3}\mathtt{b}^{n_2}$$

are not in $L(\mathcal{N}_2, \mathbb{N}^4)$.

Proof The following holds for any n_1, n_2, n_3 with $0 < n_1 < n_2 < n_3$. From the initial marking of $\mathcal{N}_2$, the only sequence of transitions labelled by $\mathtt{a}^{n_3}$ is $t_1^{n_3}$. Firing this sequence leads to the marking $\mathbf{m}_1$ such that $\mathbf{m}_1(p_1) = 1, \mathbf{m}_1(p_3) = n_3$ and $\mathbf{m}_1(p) = 0$ if $p \in \{p_2, p_4\}$. From $\mathbf{m}_1$ the only firable sequence of transitions labelled by $\mathtt{b}^{n_3}$ is $t_2 t_3^{n_3-1}$. This leads to the marking $\mathbf{m}_2$ such that $\mathbf{m}_2(p_2) = 1$ and $\mathbf{m}_2(p) = 0$ if $p \neq p_2$. The only sequence of transitions firable from $\mathbf{m}_2$ and labelled by $\mathtt{a}^{n_3}$ is $t_4 t_1^{n_3-1}$. Since $\mathbf{m}_2(p_3) = 0$, the transfer of t_4 has no effect when fired from $\mathbf{m}_2$. Hence, we reach $\mathbf{m}_1$ again after firing $t_4 t_1^{n_3-1}$. By repeating the reasoning, we conclude that the only sequence of transitions firable from the initial marking and labelled by $(\mathtt{a}^{n_3}\mathtt{b}^{n_3})^{i_1}\mathtt{a}^{n_3}$ (when $i_1 > 0$) is $t_1^{n_3} t_2 t_3^{n_3-1}(t_4 t_1^{n_3-1} t_2 t_3^{n_3-1})^{i_1-1} t_4 t_1^{n_3-1}$ and leads to $\mathbf{m}_1$. In the case where $i_1 = 0$, the sequence $t_1^{n_3}$ is firable and leads to $\mathbf{m}_1$ too. From $\mathbf{m}_1$, the only firable sequence of transitions labelled by $\mathtt{b}^{n_1}$ is $t_2 t_3^{n_1-1}$. This leads to a marking similar to $\mathbf{m}_2$, noted $\mathbf{m}_2'$, except that p_3 contains $n_3 - n_1$ tokens. Then, the only firable sequence of transitions labelled by $\mathtt{a}^{n_1}$ is $t_4 t_1^{n_1-1}$. In this case, the transfer of t_4 moves the $n_3 - n_1$ tokens from p_3 to p_4 and we reach a marking similar to $\mathbf{m}_1$, noted $\mathbf{m}_1'$, except that p_4 contains $n_3 - n_1$ tokens and p_3 contains n_1 tokens. From $\mathbf{m}_1'$, the only firable sequence of transitions labelled by $\mathtt{b}^{n_1}\mathtt{a}^{n_1}$ is $t_2 t_3^{n_1-1} t_4 t_1^{n_1-1}$ and leads to $\mathbf{m}_1'$. Hence, the sequence $(t_2 t_3^{n_1-1} t_4 t_1^{n_1-1})^{i_2}$ is firable from $\mathbf{m}_1'$.

However, after firing $t_2 t_3^{n_1-1}$ from $\mathbf{m}_1'$, we reach a marking $\mathbf{m}_2''$ similar to $\mathbf{m}_2$ except that p_4 contains $n_3 - n_1$ tokens and from which no transition labelled by

$\mathtt{b}$ is firable. Since $n_2 > n_1$, we conclude that there is no sequence of transitions labelled by $\mathtt{b}^{n_2}$ that is firable from $\mathbf{m}_1'$, hence $\mathtt{a}^{n_3}(\mathtt{b}^{n_3}\mathtt{a}^{n_3})^{i_1}(\mathtt{b}^{n_1}\mathtt{a}^{n_1})^{i_2}(\mathtt{b}^{n_2}\mathtt{a}^{n_2})^{i_3}\mathtt{a}^{n_2}$ with $i_1 \geq 0, i_2 > 0, i_3 \geq 0$ is not in $L(\mathscr{N}_2, \mathbb{N}^4)$. □

Thanks to these two lemmata, and thanks to Lemma 10, we can now prove Proposition 5, that states that no PN+NBA can accept the language of $\mathscr{N}_2$.

Proposition 5 *There is no* PN+NBA *with an* $\preccurlyeq$*-upward-closed set* $\mathscr{U}$ *such that* $L(\mathscr{N}, \mathscr{U}) = L(\mathscr{N}_2, \mathbb{N}^4)$.

Proof By Lemma 14, any PN+NBA $\mathscr{N}$ s.t. $L(\mathscr{N}, \mathscr{U}) = L(\mathscr{N}_2, \mathbb{N}^4)$ for some $\preccurlyeq$-upward-closed set $\mathscr{U}$, accepts $\left(\mathtt{a}^j\mathtt{b}^j\right)^k$, for any $j \geq 1, k \geq 1$. Thus, we can apply Lemma 10, by letting $B_i = \mathtt{a}^i$, $E_i = \mathtt{b}^i$ and $w_i = \mathtt{b}^i\mathtt{a}^i$, for all $i \geq 1$, and obtain that $\mathscr{N}$ accepts a word of the form: $\mathtt{a}^{n_3}(\mathtt{b}^{n_3}\mathtt{a}^{n_3})^{i_1}(\mathtt{b}^{n_1}\mathtt{a}^{n_1})^{i_2}(\mathtt{b}^{n_2}\mathtt{a}^{n_2})^{i_3}\mathtt{b}^{n_2}$ with $0 < n_1 < n_2 < n_3$ and $i_2 > 0$. Since, by Lemma 15, this word is not in $L^G(\mathscr{N}_2, \mathbb{N}^4)$, there can be no PN+NBA $\mathscr{N}$ and no $\preccurlyeq$-upward-closed-set $\mathscr{U}$ s.t.: $L(\mathscr{N}, \mathscr{U}) = L(\mathscr{N}_2, \mathbb{N}^4)$. □

The two last propositions allow us to conclude that:

Theorem 6 $L^G(\mathsf{PN{+}NBA}) \subset L^G(\mathsf{PN{+}T})$

Proof $L^G(\mathsf{PN{+}NBA}) \subseteq L^G(\mathsf{PN{+}T})$ is given by Lemma 13. The strictness of the inclusion is given by Proposition 5. □

5.4 Closure Properties of EPN

The pumping lemmata on PN and PN+NBA can also be used to show that neither $L^G(\mathsf{PN})$ nor $L^G(\mathsf{PN{+}NBA})$ are closed under iteration.

Theorem 7 $L^G(\mathsf{PN})$ *and* $L^G(\mathsf{PN{+}NBA})$ *are not closed under iteration.*

Proof It is easy to show that $L = \{\mathtt{a}^n\mathtt{b}^m | n \geq m\} \in L^G(\mathsf{PN})$ (hence, L is also in $L^G(\mathsf{PN{+}NBA})$). Let us show, by contradiction, that $L^+ \notin L^G(\mathsf{PN})$. Suppose that there is a PN $\mathscr{N}$ and an upward-closed set $\mathscr{U}$ s.t. $L^G(\mathscr{N}, \mathscr{U}) = L^+$. Let $B_i = \mathtt{a}^i$, $w_i = \mathtt{b}^i\mathtt{a}^i$ and $E_i = \mathtt{b}_i$ for all $i \geq 1$. Thanks to Lemma 8, we obtain that $L^G(\mathscr{N}, \mathscr{U})$ contains a word of the form:

$$\mathtt{a}^{n_3}(\mathtt{b}^{n_3}\mathtt{a}^{n_3})^{i_1}(\mathtt{b}^{n_1}\mathtt{a}^{n_1})^{K}(\mathtt{b}^{n_2}\mathtt{a}^{n_2})^{i_2}\mathtt{b}^{n_2}$$

with $n_1 < n_2 < n_3$, $K \geq 1$, which is not in L^*. Hence the contradiction. A similar proof for PN+NBA invokes Lemma 10. □

Following Definition 1, we immediately deduce that:

Corollary 2 $L^G(\mathsf{PN})$ *and* $L^G(\mathsf{PN{+}NBA})$ *are not full* AFL.

On the other hand, it is easy to show that:

Theorem 8 $L^G(\mathsf{PN{+}T})$ *is a full* AFL, *closed under intersection.*

Proof The proof is quite immediate. Hence, we only report the main ideas. In the following, we consider two PN+T $\mathcal{N}_1$ and $\mathcal{N}_2$ and two upward-closed set $\mathcal{U}_1$ and $\mathcal{U}_2$, and we assume that the set of places and transitions of this two nets are disjoint. For each property to prove we show how to build a PN+T that accepts the desired language.

Union $L(\mathcal{N}_1,\mathcal{U}_1) \cup L(\mathcal{N}_2,\mathcal{U}_2) \in L^G(\mathsf{PN{+}T})$. We build a PN+T such that its set of transitions is the union of the sets of transitions of $\mathcal{N}_1$ and $\mathcal{N}_2$ and its set of places is the union of the sets of places of the two nets. Moreover, we add a place *init* and two transitions t_1 and t_2 such that t_1 consumes a tokens from *init* and adds a number of tokens in the places of $\mathcal{N}_1$ corresponding to the number of tokens assigned by the initial marking of $\mathcal{N}_1$ to those places. t_2 consumes a tokens from *init* and adds a number of tokens in the places of $\mathcal{N}_2$ corresponding to the number of tokens assigned by the initial marking of $\mathcal{N}_2$ to those places. The initial marking of the PN+T contains only one token in *init*. The accepting $\preccurlyeq$-upward-closed set is the Cartesian product of $\mathcal{U}_1$ and $\mathcal{U}_2$.

Concatenation $L(\mathcal{N}_1,\mathcal{U}_1) \cdot L(\mathcal{N}_2,\mathcal{U}_2) \in L^G(\mathsf{PN{+}T})$. We build a PN+T such that its set of transitions is the union of the sets of transitions of $\mathcal{N}_1$ and $\mathcal{N}_2$ and its set of places is the union of the sets of places of the two nets. Moreover, we add transitions that test if the current marking is in $\mathcal{U}_1$, removes all the tokens from places of $\mathcal{N}_1$ by using transfer arcs with the target place *Trash* (added into the set of places), and finally add a number of tokens into the places of $\mathcal{N}_2$ corresponding to the number of tokens assigned by the initial marking of $\mathcal{N}_2$ to those places. Notice that it is easy to define a mechanism that ensures that those transitions are fired sequentially (some new places must by use). The initial marking corresponds to the initial marking of $\mathcal{N}_1$. More precisely, the places of $\mathcal{N}_2$ contain zero tokens into the initial marking. The accepting $\preccurlyeq$-upward-closed set corresponds to $\mathcal{U}_2$ where markings are extended to places of $\mathcal{N}_1$ that may contain any number of tokens.

Intersection $L(\mathcal{N}_1,\mathcal{U}_1) \cap L(\mathcal{N}_2,\mathcal{U}_2) \in L^G(\mathsf{PN{+}T})$. We build a PN+T such that its set of places is the union of the sets of places of the two nets. To each pair of transitions t_1 and t_2, respectively of $\mathcal{N}_1$ and $\mathcal{N}_2$, labelled by the same symbol (different from ε) we have a set of transitions that remove tokens from the input places of t_1 and t_2, apply the transfers of the two transitions and finally add tokens into the output places of the transitions t_1 and t_2. As previously noticed, it is easy to define a mechanism that ensure that those transitions are fired sequentially. Finally, the PN+T also contains all the transitions of the two nets labelled by ε. The initial marking of the PN+T corresponds to the initial marking of $\mathcal{N}_1$ on the places of $\mathcal{N}_1$ and corresponds to the initial marking of $\mathcal{N}_2$ on the places of $\mathcal{N}_2$. The accepting $\preccurlyeq$-upward-closed set is $\mathcal{U}_1 \cap \mathcal{U}_2$ (extended to places used to ensure that transitions are fired sequentially and may contain any number of tokens).

Iteration $L^+(\mathcal{N}_1,\mathcal{U}_1) \in L^G(\mathsf{PN{+}T})$. The idea is similar to the construction for the concatenation.

Arbitrary homomorphism $h(L(\mathcal{N}_1,\mathcal{U}_1)) \in L^G(\mathsf{PN{+}T})$. We replace in $\mathcal{N}_1$ each transition labelled by a by a sequence of transitions labelled by $h(a)$. As noticed previously it is easy to ensure that those transitions are fired sequentially. The initial marking corresponds to the initial marking of $\mathcal{N}_1$ (extended to the extra places and those places are empty). The accepting $\preccurlyeq$-upward-closed set corresponds to $\mathcal{U}_1$ (extended to the extra places).

Inverse homomorphism $h^{-1}(L(\mathcal{N}_1,\mathcal{U}_1)) \in L^G(\mathsf{PN{+}T})$. The PN+T that accepts the language is built as follows. We first define a finite observer that recognizes the sequences $h^{-1}(a)$ for each symbol a. We also ensure that when the observer has recognized $h^{-1}(a)$ then it fires a transition labelled by ε leading to a terminal state. Those transitions are called the terminal transitions of $h^{-1}(a)$. Then, We compute the PN+T by applying the construction presented in the intersection section on the observer and $\mathcal{N}_1$. Finally, for each symbol a the label of the terminal transition of $h^{-1}(a)$ is replaced by a and the label of the other transitions are replaced by ε. The initial marking corresponds to the initial marking of $\mathcal{N}_1$ extended to places of the observer. The accepting $\preccurlyeq$-upward-closed set corresponds to $\mathcal{N}_1$ extended to places of the observer and where at least one terminal place of the observer is non-empty. □

5.5 Some remarks about the pumping lemmata

It is interesting to compare Lemma 7, that provides a very general property holding for any WSL, with Lemma 8 and Lemma 10, which both apply to more restricted classes of languages (namely, $L^G(\mathsf{PN})$ and $L^G(\mathsf{PN{+}NBA})$, respectively), but state more precise properties of these classes of languages.

When we restrict ourselves to the class PN (resp. PN+NBA), Lemma 8 (resp. 10) is more general than Lemma 7. Indeed, we obtain Lemma 7, by letting $w_i = \varepsilon$ in Lemma 8 (resp. 10) for any $i \geq 1$.

As a consequence the following results can be proved thanks to Lemma 8 (see section 5.1 for the definition of these languages): $\mathscr{L} \notin L^G(\mathsf{PN})$; $\mathscr{L}^{\geq} \notin L^G(\mathsf{PN})$; $\mathscr{L}^R \notin L^G(\mathsf{PN})$; $L^G(\mathsf{PN})$ is not closed under complement. Similar results can be obtained about PN+NBA thanks to Lemma 10. Finally, since $\mathscr{L} \in L^L(\mathsf{PN})$, but $\mathscr{L} \notin L^G(\mathsf{PN})$, we have $L^G(\mathsf{PN}) \subset L^L(\mathsf{PN})$ (and $L^G(\mathsf{PN{+}NBA}) \subset L^L(\mathsf{PN{+}NBA}) =$ R.E. by the same reasoning).

6 Conclusion

The (labelled) well-structured transition systems are a well-known class of infinite-state transition systems, that enjoy monotonicity properties and whose set of states is well-quasi ordered. In the present work, we have studied several properties of the classes of languages that can be recognized by WSTS, and some of their sub-classes, such as the EPN. We have proved three pumping lemmata by exploiting specific properties of the WSTS (which is, to the best of our knowledge, original in this context). These lemmata have allowed us mainly to strictly separate the expressiveness of three important classes of EPN: the PN, the PN+NBA, and the PN+T. This last result demonstrates the meaningfulness of the different communication procedures present in these three models.

References

1. P. A. Abdulla, K. Cerans, B. Jonsson, and Y.-K. Tsay. General Decidability Theorems for Infinite-state Systems. In *Proceedings of the 11th Annual Symposium on Logic in Comuter Science (LICS'96)*, pages 313–321. IEEE Computer Society Press, 1996.

2. P.A. Abdulla, A Bouajjani, and B Jonsson. On-the-Fly Analysis of Systems with Unbounded, Lossy FIFO Channels. In *Proceedings of the 10th International Conference on Computer Aided Verification (CAV'98)*, volume 1427 of *LNCS*, pages 305–318. Springer, 1998.
3. G. Ciardo. Petri nets with marking-dependent arc multiplicity: properties and analysis. In *Proceedings of the 15th International Conference on Applications and Theory of Petri Nets (ICATPN 94)*, volume 815 of *LNCS*, pages 179–198. Springer, 1994.
4. E. A. Emerson and K. S. Namjoshi. On Model Checking for Non-deterministic Infinite-state Systems. In *Proceedings of the 13th Annual Symposium on Logic in Computer Science (LICS '98)*, pages 70–80. IEEE Computer Society Press, 1998.
5. A. Finkel, G. Geeraerts, J.-F. Raskin, and L. Van Begin. On the omega-language expressive power of extended Petri nets. In *Proceedings of EXPRESS'04, 11th International Workshop on Expressiveness in Concurrency, London, Great Britain*, volume 128(2) of *Electronic Notes in Theoretical Computer Science*, pages 87–101. Elsevier Publishing, 2004.
6. A. Finkel and P. Schnoebelen. Well-structured transition systems everywhere! *Theoretical Computer Science*, 256(1-2):63–92, 2001.
7. Seymour Ginsburg. *Algebraic and Automata-Theoretic Properties of Formal Languages*. Elsevier Science Inc., 1975.
8. John Hopcroft, Rajeev Motwani, and Jeffrey Ullman. *Introduction to Automata Theory, Languages, and Computation, second edition*. Addison-Wesley, 2001.
9. Matthias Jantzen. Language theory of petri nets. In Wilfried Brauer, Wolfgang Reisig, and Grzegorz Rozenberg, editors, *Proc. of Advances in Petri Nets 1986*, volume 254 of *Lecture Notes in Computer Science*, pages 397–412. Springer, 1986.
10. N.M. Minsky. *Finite and Infinite Machines*. Englewood Cliffs, N.J., Prentice-Hall, 1967.
11. J. L. Peterson. *Petri Net Theory and the Modeling of Systems*. Prentice Hall, 1981.
12. J.-F. Raskin and L. Van Begin. Petri Nets with Non-blocking Arcs are Difficult to Analyse. In *Proceedings of the 5th International Workshop on Verification of Infinite-state Systems (INFINITY 2003)*, volume 96(1) of *ENTCS*. Elsevier, 2003.
13. Arto Salomaa. *Formal Languages*. Academic Press, 1973.
14. Arto Salomaa. *Computation and Automata*, volume 25 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1985.
15. L. Van Begin. *Efficient Verification of Counting Abstractions for Parametric systems*. PhD thesis, Université Libre de Bruxelles, Belgium, 2003.