History-deterministic Vector Addition Systems

Sougata Bose
University of Liverpool, UK

David Purser
University of Liverpool, UK

Patrick Totzke
University of Liverpool, UK

Abstract

We consider history-determinism, a restricted form of non-determinism, for Vector Addition Systems with States (VASS) when used as acceptors to recognise languages of finite words. History-determinism requires that the non-deterministic choices can be resolved on-the-fly; based on the past and without jeopardising acceptance of any possible continuation of the input word.

Our results show that the history-deterministic (HD) VASS sit strictly between deterministic and non-deterministic VASS regardless of the number of counters. We compare the relative expressiveness of HD systems, and closure-properties of the induced language classes, with coverability and reachability semantics, and with and without ε-labelled transitions.

Whereas in dimension 1, inclusion and regularity remain decidable, from dimension two onwards, HD-VASS with suitable resolver strategies, are essentially able to simulate 2-counter Minsky machines, leading to several undecidability results: It is undecidable whether a VASS is history-deterministic, or if a language equivalent history-deterministic VASS exists. Checking language inclusion between history-deterministic 2-VASS is also undecidable.

2012 ACM Subject Classification Theory of computation → Formal languages and automata theory

Keywords and phrases Vector Addition Systems, History-determinism, Good-for Games

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2023.18

Related Version *Full version*: `arXiv:2305.01981` [7]

Funding We gratefully acknowledge support from the EPSRC, project number `EP/V025848/1` and the Royal Society, `IES\R3\203109`.

Acknowledgements The authors thank Ismaël Jecker, Piotr Hofman and Filip Mazowiecki for fruitful discussions.

1 Introduction

Vector addition systems with states (VASSs) are an established model of concurrency with extensive applications in modelling and analysis of hardware, software, chemical, biological and business processes. They are non-deterministic finite automata equipped with a fixed number of integer counters that may be incremented or decremented when changing control state, as long as they remain non-negative.

We explore the notion of *history-determinism* for VASSs when used as acceptors to define languages of finite words. History-determinism is a restricted form of non-determinism. In a nutshell, a non-deterministic automaton is history-deterministic (HD) if there exists a *resolver*, which is a strategy to stepwise produce a run for any input word given one letter at a time, in such a way that if there exist some accepting run on the given word then the run produced by the resolver is also accepting.

The original motivation for HDness comes from formal verification: most modelling formalisms incorporate some form of non-determinism, e.g., to over-approximate deterministic algorithms, to state specifications concisely, or to model system behaviour due to uncontrollable external environments. However, for non-deterministic models, many formal analysis techniques require costly determinisation steps that are often the main barrier to efficient procedures. History-deterministic automata provide a middle ground: they are typically more succinct, or even more expressive, than their deterministic counterparts while preserving some of their good algorithmic properties. They were also called "good-for-games" as they preserve the winner of games under composition and thus allow solving games without determinisation.

Any resolver must always choose language-maximal successors, that is, the language of the chosen successor must also include the language of any alternative successor. When considering languages of finite words, being able to continue making language-maximal choices is even a sufficient condition for being a resolver. Therefore, in this case, resolvers can be assumed to be (configuration) positional: they base their decisions only on the current configuration and not the full history leading to it. Perhaps surprisingly, resolvers for VASSs are not necessarily monotone with respect to counter values, and may require more than just comparing counters to integer thresholds (see full version [7]).

Related Work. VASSs, also known as Petri nets or partially blind counter automata, have been studied intensively since their inception in the 1960s. Early works focussed on modelling capabilities, relative expressiveness and closure properties of their recognised languages [15, 13, 37, 22] but the bulk of research on VASSs concerns decidability and complexity of decision problems [24, 29, 33, 25, 21, 23, 2, 28, 10]. In order to define languages with VASSs, different definitions distinguish between coverability and reachability acceptance conditions, and whether or not silent (ε) transitions are permitted. Checking language emptiness amounts to testing coverability or reachability, which are EXPSPACE [33, 29] and Ackermann-complete [10] respectively. Many other decision problems are undecidable, such as checking language inclusion, bisimulation and related equivalences [20] as well as checking (language) regularity [23]. Universality is undecidable for reachability acceptance [37] and decidable for coverability acceptance, via a well-quasi-order argument but with extremely high complexity (Hyper-Ackermannian in general [21] and still Ackermannian in dimension 1 [19]). These negative results by and large rely on the presence of non-deterministic choice, which motivates restricted forms of non-determinism such as bounded ambiguity (that allows for decidable inclusion [9]) or the notion of history-determinism studied here.

VASS recognisable languages over infinite words are significantly more complex than their finite-word cousins, both topologically and in terms of decision problems: already 1-VASS with (cover) Büchi acceptance can recognise Σ_1^1-complete languages [35, 12] and have an undecidable universality problem [1]. Again, the added complexity is due to non-determinism (languages of deterministic models are Borel, lower in the analytical hierarchy).

History-determinism was introduced independently, with slightly different definitions, by Henzinger and Piterman [17] for solving games without determinisation, by Colcombet [8] for cost-functions, and by Kupferman, Safra, and Vardi [26] for recognising derived tree languages of word automata. These different definitions all coincide for finite automata [3] but not necessarily for more general quantitative automata [4].

Until now, history-determinism has mainly been studied for finite-state systems. In this paper we continue a recent line of work [14, 27, 11, 16, 6, 32] that studies the notion for infinite-state models capable of recognising languages beyond (ω-)regular ones. For infinite-state systems, deterministic models are in general less expressive, not just less succinct, than their non-deterministic counterparts. In some cases they can be determinised, such is the case for quantitative automata [4] and timed automata with safety and reachability acceptance [16]. In contrast, for pushdown automata [14] and Parikh automata (VASS with $\mathbb{Z}$-valued counters; [11]), and timed automata with co-Büchi acceptance, allowing history-determinism strictly increases expressiveness (and adds more closure properties) compared to the deterministic variant. Whenever HD automata are strictly less expressive than fully non-deterministic ones, one can reasonably ask if there exists an equivalent HD automaton for a given non-deterministic one. This language HDness question is undecidable for pushdown and Parikh-automata [14, 11]. In fact, even checking if a given (pushdown or Parikh) automaton is itself HD is undecidable (for Parikh automata this follows for example by the undecidability of 2-dim. robot games [31]). On the other hand, checking HDness for timed automata is decidable [16] and various models of quantitative automata [5].

Most closely related to our work is that of Prakash and Thejaswini [32] who study history-deterministic one counter automata (OCA; PDA with unary stack alphabet) and nets (OCN; 1-dimensional VASSs) with state-based (coverability) acceptance. They show that checking automata HDness and inclusion are undecidable for OCA but remain decidable for OCNs. A useful consequence of their construction is that for any OCN one can construct a language equivalent deterministic OCA (with zero-test), albeit with a doubly exponential blow-up. They do not consider closure properties and leave open whether history-deterministic OCNs can be determinised, are equally expressive as fully non-deterministic OCNs, or fall strictly in between in expressiveness. Our work extends and generalises this paper in several directions.

Our Contributions. We study history-deterministic VASSs on finite words and without restricting the dimension. We consider coverability and reachability acceptance conditions, with and without silent (ε) transitions, and in all cases study the relative expressiveness, closure properties, and related decision problems.

We show that HD VASSs are more expressive than deterministic, but less expressive than non-deterministic ones. The same is true for languages recognised by VASSs of any fixed dimensions k, which answers the open question in [32] for $k = 1$. In particular, we provide examples of 1-dim. HD VASSs for which no equivalent deterministic ones exist in any dimension k, and also demonstrate that HD VASSs are strictly more expressive than finitely sequential ones (another restricted form of non-determinism).

We show that HD VASS languages are closed under inverse homomorphisms and intersections for both coverability and reachability semantics, although sometimes necessarily increasing the dimension. Coverability languages are closed under unions, whereas reachability languages are not. Neither are closed under other standard operations, including complementation, concatenation, homomorphisms, iteration and commutative closures.

We report that HDness is not sufficient for decidability of inclusion checking, even for 2-dimensional VASSs. A direct consequence is the undecidability of checking HDness of a given 2-VASS, contrasting decidability in dimension 1. Further, it is undecidable to check if a given VASS has a HD equivalent, and also if a given HD VASS recognises a regular language.

2 Definitions

Vector-Addition Systems and their recognised languages. A k-dimensional *vector-addition system* (k-VASS) is a non-deterministic finite automaton whose transitions manipulate k non-negative integer counters. It is given by $\mathcal{A} = (\Sigma, Q, \delta, s_0, F)$ consisting of a finite alphabet Σ; a finite set of control states Q; a transition relation $\delta \subseteq Q \times \Sigma \cup \{\varepsilon\} \times \mathbb{Z}^k \times Q$; an initial state s_0; a subset $F \subseteq Q$ of final states.

For a transition $t = (s, a, e, s') \in \delta$ we sometimes write $\mathit{label}(t) \overset{\mathit{def}}{=} a$ for the letter from $\Sigma \cup \{\varepsilon\}$ it reads and $\mathit{effect}(t) \overset{\mathit{def}}{=} e$ for its *effect* on the counters. $\|\delta\|$ denotes the largest absolute effect among all transitions on any counter.

A VASS naturally induces an infinite-state labelled transition system in which each *configuration* is a pair $(s, v) \in Q \times \mathbb{N}^k$ comprising a control state and a *non-negative* integer vector. Every transition $t = (s, a, e, s') \in \delta$ gives rise to steps $(s, v) \xrightarrow{t} (s', v')$ for all $v, v' \in \mathbb{N}$ with $v' = v + e$. We will call a path $\rho = (s_0, v_0) \xrightarrow{t_1} (s_1, v_1) \xrightarrow{t_1} \dots \xrightarrow{t_k} (s_k, v_k)$ a *run* of the VASS and say it is a *cycle* if $s_0 = s_k$. Its *effect* is the sum of all transition effects $\mathit{effect}(\rho) \overset{\mathit{def}}{=} \sum_{i=1}^{k} \mathit{effect}(t_i)$. A run ρ as above *reads* the word $\mathit{label}(\rho) = \mathit{label}(t_1)\mathit{label}(t_2) \dots \mathit{label}(t_k) \in \Sigma^*$. It is *accepting* if it ends in a final configuration.

We consider two different definitions for what constitutes a final (also *accepting*) configurations: In the *coverability* semantics, the set of final configurations is $F \times \mathbb{N}$. In the *reachability* semantics, only configurations from $F \times \mathbf{0}$ are final. We define the *language* $\mathcal{L}_{\mathcal{A}}(s, v) \subseteq \Sigma^*$ of a configuration (s, v) to contain exactly all words read by some accepting run starting in (s, v) (we omit the subscript $\mathcal{A}$ if the VASS is clear from context). For notational convenience, we will lift this to sets $S \subseteq Q \times \mathbb{N}^k$ of configurations in the natural way: $\mathcal{L}_{\mathcal{A}}(S) \overset{\mathit{def}}{=} \bigcup_{(s,v)\in S} \mathcal{L}_{\mathcal{A}}(s, v)$ and define the language of $\mathcal{A}$ as that of its initial state with all counters zero: $\mathcal{L}(\mathcal{A}) \overset{\mathit{def}}{=} \mathcal{L}_{\mathcal{A}}(s_0, \mathbf{0})$.

We will sometimes denote languages using short-hand "counting expressions". For instance, we write $a^n b^{\leq n}$ for the language $\{a^n b^m \mid n \geq m\}$ over $\Sigma = \{a, b\}$.

Deterministic and finitely-sequential VASSs. A VASS $\mathcal{A} = (\Sigma, Q, \delta, s_0, F)$ is called ε-free if no transition is labelled by ε. It is *deterministic* if it is ε-free and for every pair $(s, a) \in Q \times \Sigma$ there is at most one transition $t = (s, a, e, s') \in \delta$. A VASS is *finitely sequential* if it is the finite union of deterministic VASSs. That is, all transitions from its initial state s_0 are labelled by ε and lead to an initial state of one of finitely many deterministic VASSs.

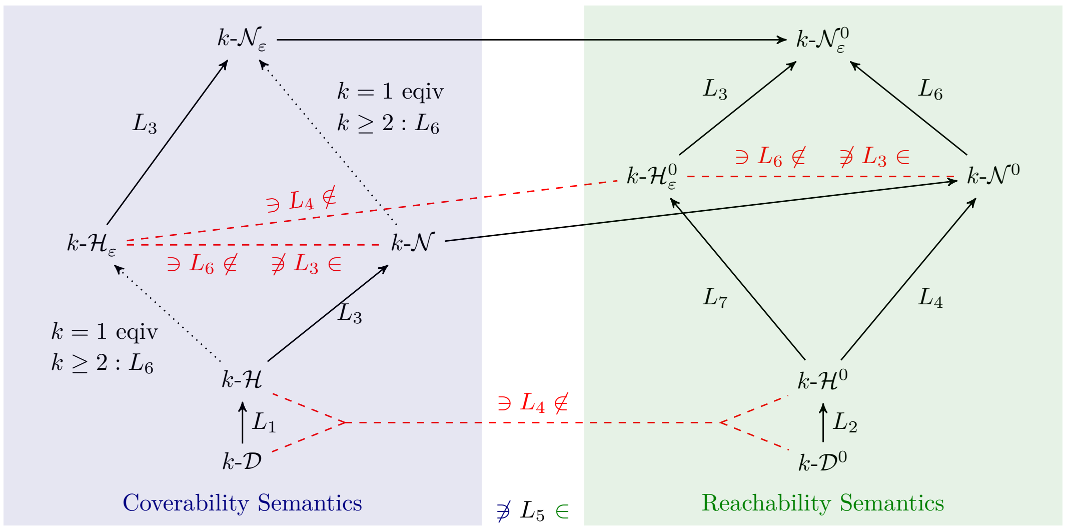
History-deterministic VASSs. A VASS is *history-deterministic* if one can resolve non-deterministic choices on-the-fly. More formally, consider a function $r : (Q \times \mathbb{N}^k \times \delta)^*(Q \times \mathbb{N}) \times \Sigma \to \delta$ that, given a finite run $\rho_i = (s_0, v_0) \xrightarrow{t_1} (s_1, v_1) \xrightarrow{t_2} \dots \xrightarrow{t_i} (s_i, v_i)$ and a next letter $a_{i+1} \in \Sigma$, returns a transition $r(\rho_i, a_{i+1}) = t_{i+1} = (s_i, a_{i+1}, e_{i+1}, s_{i+1}) \in \delta$ with $v_i + e_{i+1} \in \mathbb{N}^k$. This yields, for every word $w = a_1 a_2 \ldots \in \Sigma^*$ and initial configuration (s_0, v_0), a unique run in which the ith step $(s_{i-1}, v_{i-1}) \xrightarrow{t_i} (s_i, v_i)$ results from a transition chosen by r. Such a function is called *resolver* if for any input word $w \in \mathcal{L}_{\mathcal{A}}(s_0, v_0)$ the constructed run ρ from initial configuration (s_0, v_0) is accepting. A k-VASS is *history-deterministic* if such a resolver exists.

Language Classes. We denote by k-$\mathcal{D}$, k-$\mathcal{H}$,and k-$\mathcal{N}$ the classes of languages recognised by k-dimensional ε-free deterministic, history-deterministic, and fully non-deterministic VASSs, in the coverability semantics. Similarly, let k-$\mathcal{D}^0$, k-$\mathcal{H}^0$, and k-$\mathcal{N}^0$ denote the classes of languages recognised by k-dimensional ε-free deterministic, history-deterministic, and fully non-deterministic VASSs, in the reachability semantics. Finally, define k-$\mathcal{H}_\varepsilon$,k-$\mathcal{N}_\varepsilon$ k-$\mathcal{H}^0_\varepsilon$, and k-$\mathcal{N}^0_\varepsilon$, as above but without the restriction to ε-free systems. When dropping the parameter k we refer to the union over all dimensions k. For instance, $\mathcal{H} \overset{\mathit{def}}{=} \bigcup_{k\in\mathbb{N}} k$-$\mathcal{H}$.

3 Expressiveness

We consider the hierarchy of language classes recognised by vector addition systems, varying definitions in three directions: the degree of non-determinism, reachability vs coverability acceptance, and with/without ε-transitions.

The situation is depicted in Figure 1. We start by looking at the classes defined by ε-free systems (in Section 3.1) before discussing the effect of ε-transitions (in Section 3.2) and



Language	Definition	Alphabet	Page
L_1	$a^n b^{\leq n} + a^* b^* c$	$\{a, b, c\}$	5
L_2	$a^n b^{\geq n} \#$	$\{a, b, \#\}$	6
L_3	$(a + b)^* a^n b^{\leq n}$	$\{a, b\}$	6
L_4	$a^n b^{\leq n}$	$\{a, b\}$	7
L_5	$a^n b^n$	$\{a, b\}$	7
L_6	$bin(n)\#0^{\leq n}\#$, where $bin(n)$ is n in binary.	$\{0, 1, \#\}$	8
L_7	$a^n b^{\leq n} \#$	$\{a, b, \#\}$	8

Figure 1 Comparison of expressive power of VASS and H-VASS language classes, with and without silent transitions, in reachability and coverability semantics. A solid arrow A → B indicates strict inclusion A $\subsetneq$ B, with a separating language denoted on the edge. A red/dashed line indicates pair-wise incomparability, with the separating languages denoted. Dotted arrows indicate a special case.

following this up with a comparison with finitely-sequential VASS (in Section 3.3).

3.1 Separating determinism, history-determinism and non-determinism

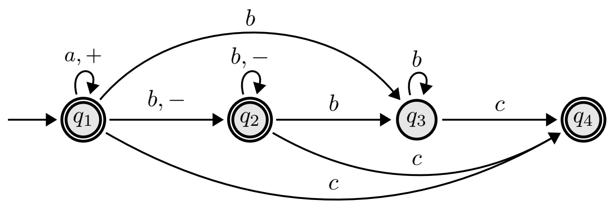
In terms of the classes of languages they define, history-deterministic VASSs are strictly more expressive than deterministic ones, and in turn strictly subsumed by fully non-deterministic ones. The following theorem states this formally. Its proof is split into Lemmas 2–5.

▶ **Theorem 1.** *For all $k \geq 1$, we have k-$\mathcal{D} \subsetneq k$-$\mathcal{H} \subsetneq k$-$\mathcal{N}$ and k-$\mathcal{D}^0 \subsetneq k$-$\mathcal{H}^0 \subsetneq k$-$\mathcal{N}^0$.*

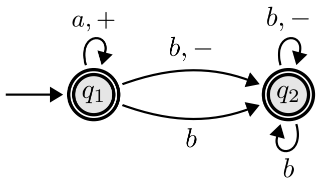
▶ **Lemma 2.** *$L_1 \overset{def}{=} a^n b^{\leq n} + a^* b^* c \in 1$-$\mathcal{H} \setminus \mathcal{D}$.*

Proof. L_1 can be recognised by the 1-H-VASS depicted in Figure 2a. Note that the VASS is HD: the only non-deterministic choice is whether to go to q_2 or q_3 on b, for which the resolver must always choose q_2 if available (if the counter is non-zero). The choice of resolver is unique as going to q_3 unnecessarily is not language maximal.

For a contradiction, suppose L_1 is accepted by a k-D-VASS with n states. Since $w_{n+1} = a^{n+1} b^{n+1} \in L_1$ the run is accepted. There exists $i < j \leq n + 1$ such that $a^{n+1} b^i$ is in state q with counter vector $v \in \mathbb{N}^k$ and $a^{n+1} b^j$ is also in state q with counter vector $v' \in \mathbb{N}^k$. Since $a^{n+1} b^i \in L_1$, we have that state q is accepting.



(a) A 1-H-VASS recognising L_1.



(b) A 1-H-VASS^0 recognising L_2.

Figure 2 Transitions labelled with $+$ increment the counter by 1, and those labelled by $-$ decrement the counter by 1 and otherwise have no effect on the counter.

Suppose $v' - v \geq \mathbf{0}$, then $a^{n+1}b^{i+(j-i)n} \notin L_1$ is accepted. Therefore there exists a dimension such that $v' - v$ is negative. Hence for some ℓ we have $a^{n+1}b^{i+(j-i)\ell}$ is a dead run. Hence it cannot accept $a^{n+1}b^{i+(j-i)\ell}c \in L_1$. ◀

▶ **Lemma 3.** $L_2 \overset{def}{=} a^n b^{\geq n} \in 1\text{-}\mathcal{H}^0 \setminus \mathcal{D}^0$

Proof. L_2 is recognised by the H-VASS^0 depicted in Figure 2b. On b the resolver can choose between decrementing the counter and no effect, the resolver will always choose to decrement whenever the counter is non-zero.

We have $L_2 \notin \mathcal{D}^0$. Suppose a D-VASS^0 with n states exists, consider the run on the word $w_{n+1} = a^{n+1}b^{n+1} \in L_2$. There exists two prefixes of the run in which $a^{n+1}b^i$ and $a^{n+1}b^j$ revisit a state, and so the system cycles through states on extension of $a^{n+1}b^i$ with b^*. Thus, in order to accept $w_{n+1}b^i$ for all i the automaton must visit only accepting states throughout the cycle. Since $a^{n+1}b^i \notin L_2$ the counter must be non-zero, but zero at $u = a^{n+1}b^{i+(j-i)n}$ since $u \in L$, thus the effect of the cycle is decreasing on some counter, there must exist $k > n$ such that the run is dead on $a^{n+1}b^{i+(j-i)k}$. This is a contradiction as $a^{n+1}b^{i+(j-i)k} \in L_2$. ◀

▶ **Lemma 4.** $L_3 \overset{def}{=} \{a,b\}^* a^{n>0} b^{\leq n} \in 1\text{-}\mathcal{N} \setminus \mathcal{H}$.

Proof. L_3 can be accepted a 1-N-VASS, which non-deterministically guesses the start of the last a^*b^* block and accepts if there are fewer b's than a's.

We show that $L_3 \notin \mathcal{H}$. Suppose for contradiction there is a k-H-VASS with $|Q|$ states, $\|\delta\|$ the largest effect on a counter in any transition and a resolver r.

Consider a sequence of accepted words $w_\ell = w_{\ell-1}a^{m_\ell}b^{m_\ell}$, with w_0 the empty word, where m_ℓ is large enough so that there exist $r_{\ell,1} < r_{\ell,2} \leq m_\ell$, such that the run given by the resolver r on $w_{\ell-1}a^{r_{\ell,1}}$ has configuration (q_ℓ, v_ℓ) and $w_{\ell-1}a^{r_{\ell,2}}$ has (q_ℓ, u_ℓ), with $u_\ell \geq v_\ell$. In other words, whilst reading a^{m_ℓ}, the run encounters a cycle on state q_ℓ which does not strictly decrease any counter value. This occurs due to Dickson's lemma and depends on $|Q|, \|\delta\|, k$ and $m_1, \ldots, m_{\ell-1}$. This gives an inductive way to build words w_ℓ consisting of ℓ blocks of as and bs such that each a-block visits a non-decreasing cycle. We consider the word w_n for $n = 2^k + 1$ and the run ρ on w_n given by the resolver.

Given a vector $v \in \mathbb{N}^k$, we define $support(v) = \{i \mid v_i \neq 0\}$. Since there are n blocks of a in w_n, each of which has a non-decreasing cycle (q_ℓ, u_ℓ) and (q_ℓ, v_ℓ), for $\ell \in \{1, \ldots, n\}$. However, there are $2^k + 1$ possible choices for $support(u_\ell - v_\ell)$. Therefore, there exists $\ell < \ell'$ such that $support(u_\ell - v_\ell) = support(u_{\ell'} - v_{\ell'})$. In other words, there are two a-blocks which have a non-decreasing cycle such that the effect of the cycles have the same support. Let $R \in \mathbb{N}$ be such that $R(u_\ell - v_\ell) \geq u_{\ell'} - v_{\ell'}$, which exists since $support(u_\ell - v_\ell) = support(u_{\ell'} - v_{\ell'})$ and $u_\ell - v_\ell \geq \mathbf{0}$ and $u_{\ell'} - v_{\ell'} \geq \mathbf{0}$.

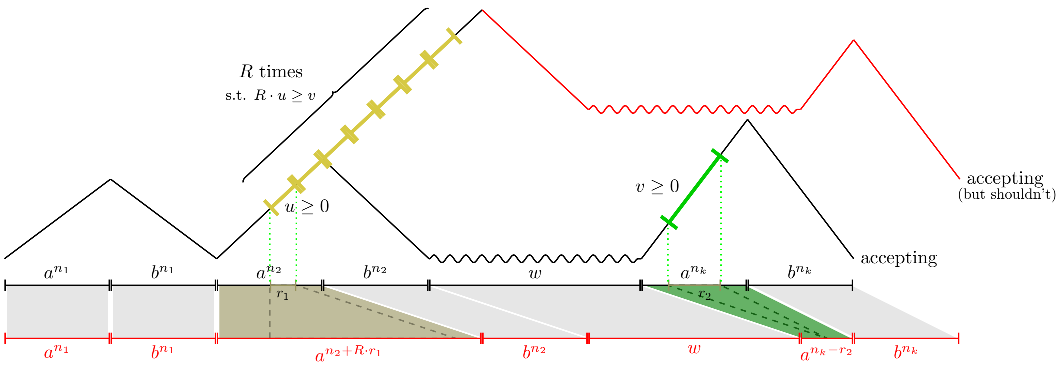


Figure 3 Proof that $L_3 \notin \mathcal{H}$ (Lemma 4). For two cycles of lengths r_1, r_2 chosen in different a^*-blocks with effects $u, v \geq \mathbf{0}$ and $\mathit{support}(u) = \mathit{support}(v)$, repeating the first cycle and removing the second one constructs an accepting run on a word $\notin L_3$.

Let u be the word such that $w_{\ell'-1} = w_\ell u$, i.e, the part between the ℓth b-block and ℓ'th a-block. Consider the word $w' = w_{\ell-1}a^{m_\ell + R(r_{\ell,2}-r_{\ell,1})}b^{m_\ell}ua^{m_{\ell'}-(r_{\ell',2}-r_{\ell',1})}b^{m_{\ell'}}$. The word w' is therefore obtained by adding $R(r_{\ell,2} - r_{\ell,1})$ many a's in the ℓth a-block and removing $(r_{\ell',2} - r_{\ell',1})$ many a's from the ℓ'th a-block. Note that $w' \notin L_3$, since the last block has more b's than a's. We will show that there is an accepting run on w', by modifying the resolver run on w'_ℓ.

Let $\rho_{\ell'}$ be the run on $w_{\ell'}$ given by the resolver r. We consider the run ρ' where we take the cycle between (q_ℓ, v_ℓ) and (q_ℓ, u_ℓ) an additional R times in the ℓ-th a-block, but removes the cycle between $(q_{\ell'}, v_{\ell'})$ and $(q_{\ell'}, u_{\ell'})$. We show that ρ' is a run on w'. To see this, we must verify that no counter drops below zero in ρ'. Note that the runs $\rho_{\ell'}$ and ρ' are the same till the prefix $w_{\ell-1}a^{r_{\ell,2}}$ after which it reaches the configuration (q_ℓ, u_ℓ). Then it does R additional cycles which results in the configuration $(q_\ell, u_\ell + R(u_\ell - v_\ell))$. From this point ρ' follows the same sequence of transitions as $\rho_{\ell'}$ till it reads the prefix up to $w_{\ell'-1}a^{r_{\ell',1}}$ ending up in the configuration $(q_{\ell'}, v_{\ell'} + R(u_\ell - v_\ell))$. Since $v_{\ell'} + R(u_\ell - v_\ell) \geq v_{\ell'} + (u_{\ell'} - v_{\ell'}) = u_{\ell'}$, ρ' can follow the suffix of the run $\rho_{\ell'}$ from $(q_{\ell'}, u_{\ell'})$ on $a^{m_{\ell'}-r_{\ell',2}}b^{m_{\ell'}}$, which ends in the same state as $\rho_{\ell'}$ with a non-zero counter value. This is a contradiction as we get a accepting run on $w' \notin L_3$. We conclude that there is no k−H-VASS that recognises the language L_3. ◀

▶ **Lemma 5.** $L_4 \overset{def}{=} a^n b^{\leq n} \in 1\text{-}\mathcal{N}^0 \setminus \mathcal{H}^0$

Proof. In the non-deterministic case reachability semantics can recognise $L_4 \in 1\text{-}\mathcal{N}^0$: On a, non-deterministically choose either to increment by 1 or not, guessing ahead of time how many b's will be seen. On b, the machine moves to a new state and counts down, preventing more b's than the guessed number.

However L_4 cannot be recognised with history-determinism. To see this, observe that since $a^n \in L_4$ all the counters must be zero after reading a^n, then, for n larger than the number of states, the machine cannot distinguish $a^n b^n \in L_4$ and $a^n b^{n+1} \notin L_4$. ◀

3.2 Silent transitions

First observe that $L_5 \overset{def}{=} a^n b^n$ can be recognised with reachability semantics (even $\mathcal{D}^0$), but cannot be recognised under coverability semantics (even $\mathcal{N}_\varepsilon$). On the other hand $L_4 = a^n b^{\leq n}$ can be recognised by coverability semantics (even $\mathcal{D}$), but cannot be recognised by $\mathcal{H}^0_\varepsilon$, thus together L_4 and L_5 show pairwise incomparability between reachability and coverability

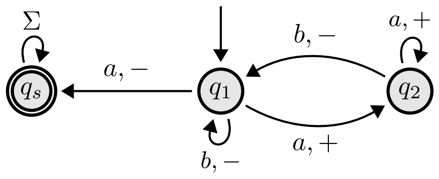


Figure 4 A 1-H-VASS automaton with language $L_8 = \mathcal{L}(q_1, 0)$ that is not finitely sequential. The automaton reads blocks of a's followed by blocks of b's. If some block of a's is followed by fewer b's then the automaton can read anything after the next a. If every block is followed by the same number of a's and b's then it must read another block of the form $a^n b^n$ or $a^n b^{<n}$. The language is thus $L_8 = \bigcup_{k=0}^{\infty} a^{n_0} b^{n_0} \ldots a^{n_{k-1}} b^{n_{k-1}} a^{n_k} b^{<n_k} a\Sigma^*$.

semantics for deterministic and history-deterministic systems. However, if the languages have an end marker then coverability acceptance can be turned into reachability acceptance (with ε-transitions) as ε-transitions can be used to take the counters to zero at the end marker.

The separation between $\mathcal{N}$ and $\mathcal{N}_\varepsilon$ is due to [13] for which $L_6 \overset{def}{=} bin(n)\#0^{\leq n}\# \in \mathcal{N}_\varepsilon \setminus \mathcal{N}$, where $bin(n)$ is the binary representation of $n \in \mathbb{N}, n > 0$ in $1\{0,1\}^*$. This language cannot be recognised without ε transitions (see full version [7] or [13] for details). We observe that the same language separates $\mathcal{H}$ and $\mathcal{H}_\varepsilon$, as the 2-VASS of [13] recognising L_6 is in fact history-deterministic. However, in dimension 1, the two classes collapse:

▶ **Lemma 6.** $1\text{-}\mathcal{H} = 1\text{-}\mathcal{H}_\varepsilon$.

While in coverability semantics, the presence of ε-transitions separates languages recognised by k-H-VASS and k-H-VASS$_\varepsilon$ only for dimensions $k \geq 2$, in reachability semantics the separation occurs already in dimension 1: $L_7 \overset{def}{=} a^n b^{\leq n}\#$ is in $\mathcal{H}^0_\varepsilon$ but not in $\mathcal{H}^0$.

3.3 Comparison with Finitely Sequential VASS

Recall that finitely sequential VASS are the union of finitely many D-VASS. In Lemma 8 we show that language of a finite union of history-deterministic VASS is also history-deterministic. In particular, the deterministic VASSs comprising the finitely sequential VASS are themselves history-deterministic, so any finitely sequential VASS has an equivalent history-deterministic VASS recognising the same language. On the other hand, we show that history-deterministic VASS with coverability acceptance are strictly more powerful:

▶ **Lemma 7.** *There exists a language in* $1\text{-}\mathcal{H}$ *that is not finitely sequential.*

Proof. Consider the language $L_8 \overset{def}{=} \mathcal{L}(q_1, 0)$ of the VASS depicted in Figure 4. Observe that it is history-deterministic: when reading a at state q_1, the resolver goes to q_s if possible. This choice is language-maximal and there is no other non-determinism to resolve.

We show the language is not finitely sequential. Suppose for contradiction the language is accepted by a finitely sequential VASS that is the union of k many D-VASS, each with at most m states. We consider the word $a^{m+1}b^{m+1}$, which can be extended into an accepting word in L_8, and thus is alive in some D-VASS. For D-VASS in which the run is still alive, reading this word goes through a cycle in the run while reading a^{m+1} and similarly also whilst reading b^{m+1}.

Let $c_1, \ldots, c_k$ be the lengths of these cycles while reading a's in each D-VASS respectively, $d_1, \ldots, d_k$ be the lengths of the cycles reading b's, and fix $C = \prod_{i \leq k} c_i$ and $D = \prod_{i \leq k} d$. Observe that, for every x, for each D-VASS the same state is reached after reading a^{m+1+xC}.

Similarly, for any y, the same state is reached after reading $a^{m+1+xC}b^{m+1+yD}$. In particular, fix words $w = a^{m+1+CD}b^{m+1+CD}$ and the words $u = a^{m+1+CD}b^{m+1+(C-1)D}$.

Observe that after reading ua, the system in Figure 4 can be in state q_s and therefore, any extension of ua is accepted. However, the automaton of Figure 4 can only reach state q_2 on wa and so, for any $z \in \mathbb{N}, i \geq 1$, $wa^z b^{z+i} \notin L_8$. Consider this word for $z = m + 1$. Since there is a cycle somewhere while reading b^z, then when reading more b's the automaton visits only states on that cycle. Since $wa^z b^{z+i} \notin L_8$ for $i \geq 1$ either every state on the cycle is non-accepting, or the cycle has a negative effect on at least one counter and therefore becomes unavailable for large enough i.

Recall, in M both wa and ua are in the same control location in each constituent D-VASS, and thus for any $v \in \Sigma^*$ we have wav and uav reach the same control locations (or possibly the run is dead). However, for every z, i, there is some D-VASS in which the word $ua^z b^{z+i}$ is accepting. However, we have argued that for every D-VASS, for sufficiently large i, the run on $wa^z b^{z+i}$ is stuck in a rejecting cycle, or a cycle in which the counter is decreasing. Thus for sufficiently large i, in every D-VASS, either the run on $ua^z b^{z+i}$ is also dead or in a rejecting cycle, which contradicts $ua^z b^{z+i} \in L_8$. ◀

4 Closure Properties

We take a look at closure properties of the classes $\mathcal{H}$ and $\mathcal{H}^0$ recognised by history-deterministic VASSs in coverability and reachability semantics, respectively.

Union closure (of $\mathcal{H}$ and $\mathcal{H}^0$) and closure under intersection (for $\mathcal{H}$) can be shown using a straightforward product construction at the cost of increasing the dimension.

▶ **Lemma 8.** *Let $L \in k$-$\mathcal{H}$ and $L' \in k'$-$\mathcal{H}$. Then $L \cup L' \in (k+k')$-$\mathcal{H}$ and $L \cap L' \in (k+k')$-$\mathcal{H}$. Let $L \in k$-$\mathcal{H}^0$ and $L' \in k'$-$\mathcal{H}^0$. Then $L \cap L' \in (k+k')$-$\mathcal{H}^0$.*

A naïve product of the two systems recognising L and L' does *not* work for showing the union closure of $\mathcal{H}^0$ because here, acceptance requires all counters to be zero even for inputs that are only in one of the two languages (note the absence of ε-transitions). Indeed, $\mathcal{H}^0$ are not closed under union, as witnessed by $L_9 \overset{def}{=} a^n b^n \cup a^n b^{2n}$ not being in $\mathcal{H}^0$ (see full version [7]).

Taking a direct product yields a H-VASS that may not be optimal in terms of the number of counters and in general, increasing the dimension is not avoidable. For instance, the languages $L_{10} \overset{def}{=} a^n b^{\leq n} c^* \cup a^n b^* c^{\leq n}$ and $L_{11} \overset{def}{=} a^n b^{\leq n} c^* \cap a^n b^* c^{\leq n}$ are not in 1-$\mathcal{H}$, while the individual component languages are. Similarly, the language $L_{12} \overset{def}{=} a^n b^n c^* \cap a^n b^* c^n = a^n b^n c^n$ witnesses non-closure of 1-$\mathcal{H}^0$ under intersection.

The theorems below summarise our findings regarding closure properties of history-deterministic classes. Full proofs are in the full version [7].

▶ **Theorem 9.** *$\mathcal{H}$ is closed under union, intersection and inverse homomorphisms. It is not closed under complementation, concatenation, homomorphisms, iteration, nor commutative closure.*

▶ **Theorem 10.** *$\mathcal{H}^0$ is closed under intersection and inverse homomorphisms. It is not closed under union, complementation, concatenation, homomorphisms, iteration, nor commutative closure.*

5 Decision Problems

In this section we consider decision problems related to history-determinism: checking if a given N-VASS is history-deterministic, HD definability (as well as regularity) of its recognised language, and language inclusion between HD VASSs.

Prakash and Thejaswini [32] showed that in dimension 1 (and for coverability semantics), checking HDness and inclusion is decidable in PSPACE by reduction to simulation preorder [18]. This can be generalised slightly as follows.

▶ **Theorem 11.** *Language inclusion $\mathcal{L}(\mathcal{B}) \subseteq \mathcal{L}(\mathcal{A})$ is decidable for any* 1*-H-VASS $\mathcal{A}$ and for any N-VASS $\mathcal{B}$.*

Proof. By Theorem 19 in [32], for any 1-H-VASS, one can effectively construct a language equivalent deterministic one-counter automaton (DOCA; a 1-VASS with zero-testing transitions). DOCA can be complemented [36] and so the inclusion question is equivalent to the emptiness (reachability) of $\overline{\mathcal{A}} \times \mathcal{B}$, a VASS with one zero-testable counter, which is decidable [34]. Note this result is independent of the number of counters of B. ◀

We continue to show that in higher dimensions, these questions are undecidable. Throughout this section, when we show undecidability for ε-free VASS, the result naturally also applies for superclass with silent transitions. Our constructions proving this are similar, yet require subtle differences, and are all based on weakly simulating two-counter machines [30]. Let us recall these in a suitable syntax first.

▶ **Definition 12.** *A two-counter Minsky machine (2CM) $M = (Q, q_0, q_h, \delta)$ consists of a finite set of states Q, including a distinguished starting and final state q_0, q_h, respectively, as well as a finite set of transitions $\delta \subseteq Q \times \Gamma \times Q$, where $\Gamma = \{inc_1, inc_2, dec_1, dec_2, ztest_1, ztest_2\}$ are the operations on the counters*[1]*.*

A configuration of M is an element of $Q \times \mathbb{N}^2$, comprising the current state and the value of the two counters. For every state q either:

1. *There is only one transition of the form (q, inc_i, q'). This allows to move from state q to q', increment counter i by one and leaves the other counter untouched; or*
2. *There are exactly two transitions from q, of the form $(q, ztest_i, q')$ and (q, dec_i, q''). The former allows to move to q' without changing the counters, but only if counter i has value* 0*. The latter allows to move from q to q'' and decrease counter i, and leaves the other counter unchanged.*

Notice that from any configuration there is exactly one possible successor configuration. We can therefore speak of *the* run of M, and its sequence of counter operations, from the initial configuration $(q_0, 0, 0)$. We say that M *terminates* if its run visits the final state q_h. W.l.o.g., we can assume that both counters have value 0 whenever M terminates.

Deciding whether a given 2CM terminates is undecidable [30]. An easy consequence, and the basis for our construction for regularity, is the undecidability of checking finiteness of the reachability set for a given 2CM.

▶ **Lemma 13.** *It is undecidable to check, for given 2CM M, if its run visits infinitely many different configurations.*

[1] Readers may be more familiar with an instruction of the form if $C_i = 0$ goto q_ℓ else goto q_k, this can be simulated by a ztest$_i$ to q_ℓ and a decrement followed by an increment to q_k.

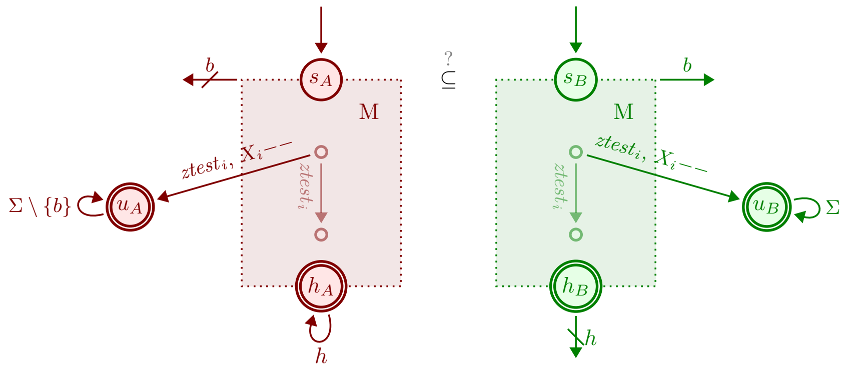


Figure 5 The 2-VASSs A (in red) and B (in green) both include a copy of, and weakly simulate, a given 2CM M. For any zero-testing operation ztest_i in M both can go to a sink state if counter i is in fact non-zero, reading the letter ztest_i and decreasing the VASS counter i, as indicated by the effect vector $X_i{-}{-}$. The extra letter b ensures that $\mathcal{L}(B) \not\subseteq \mathcal{L}(A)$; Only A can accept words that consist of valid sequences of 2CM operations and that end in the letter h.

5.1 Checking HDness and Inclusion

We focus on the questions of whether a given VASS is history-deterministic, and whether language inclusion holds for two languages given by H-VASS. For languages of finite words these two decision problems are intrinsically linked due to the connection with language maximal resolvers.

▶ **Lemma 14.** *For a given 2CM M one can construct two history-deterministic 2-VASSs with initial states s_A and s_B, respectively, so that $\mathcal{L}(s_A, 0) \subseteq \mathcal{L}(s_B, 0)$ if, and only if, the unique valid run of M never reaches a halting state.*

Proof. Suppose we are given 2CM M with designated initial and halting states s and h, respectively, and let Γ denote the set of counter operations. W.l.o.g., there is exactly one valid sequence of counter operations that is either infinite or finite. We define two 2-VASSs A and B over the alphabet $\Sigma = \Gamma \uplus \{b, h\}$. These are just copies of, and just weakly simulate the machine M: For every state q of M, there are states q_A and q_B; For every transition $q \xrightarrow{\gamma} q'$ of M, there are corresponding edges $q_A \xrightarrow{\gamma} q'_A$ and $q_B \xrightarrow{\gamma} q'_B$ that read the letter γ and manipulates the counter accordingly: if $\gamma = \mathrm{inc}_i$ (or dec_i) then counter i is incremented (or decremented, respectively). If $\gamma = \mathrm{ztest}_i$ then counter i remains as is. The only accepting states so far are h_A and h_B, corresponding to the designated halting state of M.

Additionally, for every zero-testing transition $q \xrightarrow{\mathrm{ztest}_i} q'$ in M, both A and B have a transition from state q that decreases counter i and goes to a new, accepting, sink state u with language $\supseteq (\Gamma \cup \{h\})^*$. This way, both systems will accept any word that prescribes a run of M that contains a "counter cheat", meaning that the word contains operation ztest_i but the run of M so far ends in a configuration where counter i is not zero.

We now modify the systems A and B so that they differ in two ways:

1. the halting state h_A of A admits a h-labelled step (to itself) but s_B does not.
2. All states in B have b-labelled steps (to the accepting sink u_B) but none of As states do.

See Figure 5 for a depiction of the constructed 2-VASS.

Notice that $\mathcal{L}(B) \not\subseteq \mathcal{L}(A)$ by design, because no word containing the letter b can be accepted by A. Notice that both A and B are indeed history-deterministic: the only choices to be resolved are upon reading a zero-testing letter ztest_i from a configuration where the corresponding counter i is not zero. In any such case, moving to the sink is language maximal.

It remains to argue that $\mathcal{L}(s_A) \not\subseteq \mathcal{L}(s_B)$ if, and only if, M has a finite run from an initial configuration to its final state. Indeed, if M terminates via a sequence $\rho = e_0 e_1, \ldots e_k$, then $\mathcal{L}(A)$ contains the word $\rho \cdot h$. Since this run does not contain "cheats" nor letters b, the system B cannot possibly reach the winning sink u_B and therefore not accept. Conversely, if M does not terminate, then any word $\rho \in \Gamma^* \cdot h$ accepted by A must prescribe a run of M that contains a cheat. Say $\rho = \rho_1 \cdot \mathrm{ztest}_i \cdot \rho_1 \cdot h$. But then, B will be able to reach the sink u_B after reading the prefix $\rho_1 \cdot \mathrm{ztest}_i$ and thus accept. ◀

The construction in the previous lemma works both in coverability and reachability semantics (note that we assume that a 2CM terminates with counters at 0). The next two theorems are direct consequences and again hold for coverability and reachability semantics.

▶ **Theorem 15.** *Checking language inclusion is undecidable for* 2*-HD VASSs.*

▶ **Theorem 16.** *It is undecidable to check if a given* 2*-VASS is history-deterministic.*

Proof. By reduction from 2CM termination: Construct the two systems A and B as given by Lemma 14 and add one new initial state s that, upon reading some letter b can move to the initial state s_A of A or s_B of B. The so-constructed system is HD iff $\mathcal{L}(s_A) \subseteq \mathcal{L}(s_B)$, which is true iff M does not terminate. ◀

5.2 Checking HDness of VASS Languages

We turn to showing undecidability of *language* history-determinism, i.e., the question if for a given VASS there exists an equivalent history-deterministic VASS. We start with the more interesting and involved case, for the coverability semantics (Theorem 17) and present an easier construction for reachability (Theorem 18) afterwards.

We give a proof by reduction from the 2CM halting problem, combining the constructions to show the non-HDness of $L_3 = (a,b)^* a^n b^{\leq n}$, (Lemma 4) and the proof of [23] that checking regularity for N-VASS languages is undecidable.

▶ **Theorem 17.** *It is undecidable to check if $\mathcal{L}(\mathcal{A}) \in \mathcal{H}$ holds for a given N-VASS $\mathcal{A}$.*

Proof. By reduction from the 2CM halting problem. For a given 2CM M with states Q_M and counter operations $\Gamma = \{\mathrm{inc}_1, \mathrm{inc}_2, \mathrm{dec}_1, \mathrm{dec}_2, \mathrm{ztest}_1, \mathrm{ztest}_2\}$ we construct a 3-VASS $\mathcal{A} = (\Sigma, Q, \delta, s_0, F)$ so that $\mathcal{L}(\mathcal{A})$ is history-deterministic iff the faithful run of M is finite.

We refer to the three counters as X_1, X_2, X_3 and write $X_i{-}{-}$ and $X_i{+}{+}$ for the effects of (VASS) transitions that decrement/increment counter i only.

The construction. $\mathcal{A}$ uses the alphabet $\Sigma = \Gamma \cup \{a, b\}$, consisting of counter operations of M and two fresh symbols. The control states of $\mathcal{A}$ mimic those of M, except that in between any simulated step of M, $\mathcal{A}$ can read a word in a^+b^+: For every state $q \in Q_M$ we introduce states q_{in}, q_{out} and q_{step}. In addition, we add three other states sink, r_1, r_2. We make sink universal by adding self-loops $(s, a, \mathbf{0}, s)$ for every letter $a \in \Sigma$. First we consider the simulation of M.

For every step $q \xrightarrow{\gamma} p$ of M, $\mathcal{A}$ has a transition $t = (q_{out}, \gamma, e, p_{in})$ from q_{out} to p_{in} that reads the letter $\mathit{label}(t) = \gamma$ and manipulates the counter accordingly: if $\gamma = \mathrm{inc}_i$ then $e = X_i{+}{+}$; if $\gamma = \mathrm{dec}_i$ then $e = X_i{-}{-}$; if $\gamma = \mathrm{ztest}_i$ then $e = \mathbf{0}$. In addition, for zero-testing steps $q \xrightarrow{\mathrm{ztest}}_i p$, $\mathcal{A}$ in M, $\mathcal{A}$ contains a decreasing transition $t = (q_{out}, \mathrm{ztest}_i, X_i{-}{-}, \mathit{sink})$ to the universal sink state. From a state q_{in}. There are two possible continuations:

1. Reading a word in a^+b^+ and moving to q_{out}, via transitions $q_{in} \xrightarrow{a,\mathbf{0}} q_{step}$, $q_{step} \xrightarrow{a,\mathbf{0}} q_{step}$, $q_{step} \xrightarrow{b,\mathbf{0}} q_{out}$ and $q_{out} \xrightarrow{b,\mathbf{0}} q_{out}$.
2. Reading a word in $a^n b^{\leq n}$ and stopping. For this, there are transitions $q_{in} \xrightarrow{a,X_3{+}{+}} r_1$, $r_1 \xrightarrow{a,X_3{+}{+}} r_1$, $r_1 \xrightarrow{b,X_3{-}{-}} r_2$ and $r_2 \xrightarrow{b,X_3{-}{-}} r_2$.

The accepting states of $\mathcal{A}$ are $F = \{r_2, \mathit{sink}\}$. Its initial state is $s_0 = q_{out}$, where $q \in Q_M$ is the initial state of M.

The recognised language. The language of the constructed 3-VASS $\mathcal{A}$ contains sequences of instructions of M interspersed with blocks of the form a^+b^+. Let's call a sequence $\gamma_1\gamma_2 \ldots \gamma_k \in \Gamma^*$ of operations in M *faithful* if for all $i \leq k$, γ_i is the ith instruction in the run of M from its initial configuration $(q, 0, 0)$. Clearly, for any k less or equal to the length of the run of M, there is a unique faithful sequence ρ_k of length k. Define $\mathsf{Correct}_k \overset{\text{def}}{=} \gamma_1(a^+b^+)\gamma_2(a^+b^+)\gamma_3 \ldots (a^+b^+)\gamma_k$ where $\gamma_1\gamma_2 \ldots \gamma_k = \rho_k$. Let $\mathsf{Incorrect}_k \subseteq \Sigma^*$ contain exactly all words $w\gamma \in \Sigma^* \setminus \mathsf{Correct}_k$ where $w \in \mathsf{Correct}_{k-1}$ and $\gamma \in \{\text{ztest}_1, \text{ztest}_2\}$. That is, words whose projection into the operations of M is faithful up to step $k-1$ but that contain an incorrect zero-test at step k.

Observe that if the faithful sequence of length k takes M to (q, C_1, C_2) then $\mathcal{A}$ can read any word in $\mathsf{Correct}_k$ and every run on such a word leads to the configuration $(q_{in}, C_1, C_2, 0)$. Such a run of $\mathcal{A}$ can be extended in two ways to reach an accepting state. Either by reading a word in $a^n b^{\leq n}$ to reach r_2, or by continuing on the run of M and eventually erroneously reading a ztest_i to reach sink. We can therefore write the language of $\mathcal{A}$ as

$$\mathcal{L}(\mathcal{A}) \quad = \quad \bigcup_{k \geq 0} \mathsf{Correct}_k \cdot (a^n b^{\leq n}) \quad \cup \quad \bigcup_{k \geq 0} \mathsf{Incorrect}_k \cdot \Sigma^*$$

HDness. We show that if M terminates, meaning its run has some length $k \in \mathbb{N}$, then $\mathcal{L}(\mathcal{A})$ is history-deterministic. Observe that for every $0 \leq i \leq k$, both languages $\mathsf{Correct}_i$ and $\mathsf{Incorrect}_i$ are regular. We can concatenate a DFA recognising the former with a 1-H-VASS for $a^n b^{\leq n}$ to construct an 1-H-VASS recognising $\mathsf{Correct}_i \cdot (a^n b^{\leq n})$. Observe that $\mathsf{Incorrect}_i$, $i > k+1$ is empty. Now, $\mathcal{L}(\mathcal{A})$ is the finite union of k many 1-H-VASS languages (and a regular language) and therefore recognisable by a k-dimensional H-VASS.

It remains to show that if the run of M is infinite, then $\mathcal{L}(\mathcal{A})$ is not in k-$\mathcal{H}$, for any k. Our proof mirrors the proof of Lemma 4, except that we interleave $\{a, b\}$-blocks with the faithful operations of M. Suppose towards a contradiction that there exists a k-H-VASS $\mathcal{B}$ with states $Q_{\mathcal{B}}$ and let $\rho = \gamma_1\gamma_2, \dots \in \Gamma^\omega$ denote the infinite run of M. That is, every length-i prefix ρ_i is faithful. Consider a sequence $(w_n)_{n \geq 0}$ of words in $\mathcal{L}(B)$ such that $w_0 = \varepsilon$ and otherwise $w_\ell = w_{\ell-1}\gamma_\ell a^{m_\ell} b^{m_\ell}$ with m_ℓ large enough so that the resolved run on w_ℓ contains a non-decreasing cycle while reading the last a-block. Say,

$$(s_0, \mathbf{0}) \xrightarrow{w_{\ell-1}\gamma_\ell a^{r_{\ell,1}}} (q_\ell, u_\ell) \xrightarrow{a^{r_{\ell,2}}} (q_\ell, v_\ell)$$

with $u_\ell \leq v_\ell$. This is well-defined by Dickson's Lemma.

Setting $n = |Q_{\mathcal{B}}|2^k + 1$ is sufficiently high so that there must be $\ell < \ell'$ with $q_\ell = q_{\ell'}$ and $\mathit{support}(u_\ell - v_\ell) = \mathit{support}(u_{\ell'} - v_{\ell'})$. Take R be such that $R(u_\ell - v_\ell) \geq (u_{\ell'} - v_{\ell'})$ and let u be the word such that $w_{\ell'-1} = w_\ell u$. Now consider the word

$$w' = w_{\ell-1}\gamma_\ell a^{m_\ell + R(r_{2,\ell})} b^{m_\ell} u \gamma_{\ell'} a^{m_{\ell'} - r_{2,\ell'}} b^{m_{\ell'}}$$

that results from w_n by removing one iteration of the loop in block ℓ' and making up for it by inserting R iterations of the loop in block ℓ. Notice that w' is accepted by the run that follows the resolved run on w_n and repeats the designated loops on the extra letters. However, $w' \notin \mathcal{L}(\mathcal{A})$ because its last $\{a,b\}$-block contains more b's than a's. ◀

Notice that if the given 2CM terminates then our construction produces a history-deterministic VASS where the number of counters corresponds to the length of the terminating run. Therefore it remains open whether the language k-HDness problem is decidable, which ask whether there is an equivalent k-HDVASS for the given language.

The analogous statement for reachability is simpler to prove; by adapting the construction for the regularity problem of Parikh-automata [11], we reduce from the universality problem, which is undecidable in reachability semantics [37, Theorem 10].

▶ **Theorem 18.** *It is undecidable to check if $\mathcal{L}(\mathcal{A}) \in \mathcal{H}^0$ holds for a given N-VASS $\mathcal{A}$.*

Proof. We reduce from the undecidable universality problem for VASS languages in reachability semantics [37, Theorem 10]. The construction is the same as for the regularity problem of Parikh-automata, recently presented in [11]. For an alphabet Σ let $\Sigma_\$ = \Sigma \uplus \{\$\}$ for some fresh symbol $\$ \notin \Sigma$. For two words u, v let $u \otimes v$ be the word $w = (a_1, b_1)(a_2, b_2) \dots (a_k, b_k)$ so that either $u = a_1 a_2 \dots a_k$ and $b_1 b_2 \dots b_k \in v\$^*$ or $v = b_1 b_2 \dots b_k$ and $a_1 a_2 \dots a_k \in u\$^*$.

For two languages $L, L' \subseteq \Sigma^*$ define their *cross-union* $L \varowedge L \subseteq (\Sigma_\$^2)^*$ to be the langugae of words $u \otimes v$ such that $u \in L$ or $v \in L'$. That is, for any word $w \in L \varowedge L$, either the projection into the first components is $\pi_1(w) \in L$ or that into the second components $\pi_2(w) \in L'$.

Recall the language $L_4 = a^n b^{\leq n} \in \mathcal{N}^0 \setminus \mathcal{H}^0$, which is not HD recognisable. To show our claim, let L be some given N-VASS language and consider the language $L_{14} \overset{def}{=} \$ \cdot (L \varowedge \emptyset) \cup \$ \cdot (\emptyset \varowedge L_4)$. This is clearly in $\mathcal{N}^0$. Now, if $L = \Sigma^*$ is universal then $L \varowedge \emptyset$ is universal over $\Sigma_\$^2$ and so $L_{14} = \$(\Sigma_\$^2)^* \in \mathcal{H}^0$ (even, regular). If conversely, suppose L is not universal as witnessed by $w \notin L$, then L_{14} cannot be recognised by any H-VASS^0 for the same reason as $a^n b^{\geq n} \notin \mathcal{H}^0$: suppose it is accepted by some k-H-VASS^0 run on n states and consider run of the resolver on the word $u = \$(w \otimes a^{|w|+n+1}) \in L_{14}$, thus must end with counter $\mathbf{0}$. The extension of u by $(\$, b)^{n+1}$ is also accepting, it must remain at $\mathbf{0}$ and cycle on accepting states. Hence $u(\$, b)^{|w|+n+1} \in L_{14}$ cannot be distinguished from $u(\$, b)^{|w|+n+2} \notin L_{14}$. ◀

5.3 Regularity

We turn to the decision problem of whether a given VASS recognises a regular language. This regularity question is undecidable for general N-VASS [23]. It again turns out that for history-deterministic VASSs, the decidability status of regularity depends on the dimension. For 1-H-VASS, one can effectively construct a language equivalent DOCA [32], for which checking regularity remains decidable [2, 36].

▶ **Theorem 19.** *Given a 1-H-VASS $\mathcal{A}$, checking if $\mathcal{L}(\mathcal{A})$ is regular is decidable in* EXPSPACE.

Although checking regularity of DOCA is NL-complete, the added complexity here is due to the doubly exponentially large DOCA produced in the reduction. Since 1-H-VASS$_\varepsilon$ can be transformed into 1-H-VASS by Lemma 6, the theorem also holds for 1-H-VASS$_\varepsilon$. We now show undecidability already for dimension 2.

▶ **Theorem 20.** *Given a 2-H-VASS $\mathcal{A}$, it is undecidable if $\mathcal{L}(\mathcal{A})$ is regular.*

Proof. By reduction from the finiteness problem for 2CM (Lemma 13). For a given 2CM M we construct a 2-H-VASS whose language will be regular iff M's run visits only finitely many configurations. We make the argument for coverability semantics first.

Let $\rho = \gamma_1\gamma_2\ldots$ be the faithful run of M and $|\rho| \in \mathbb{N} \cup \{\infty\}$ for its length. Write $\mathit{correct}_k$ for its length-k prefixes and let x_k be 1 plus the sum of both counter-values in the configuration M reaches after reading $\mathit{correct}_k$. Further, wherever $\mathit{correct}_k = \mathit{correct}_{k-1}\mathrm{dec}_i$, define $\mathit{incorrect}_k$ as $\mathit{correct}_{k-1}\mathrm{ztest}_i$.

Consider the language $L = G \uplus B$ over the alphabet $\Sigma = \Gamma \uplus \{a\}$,

$$G \overset{def}{=} \bigcup_{k\geq 0} \left(\rho_k \cdot a^{\leq x_k}\right) \qquad \text{and} \qquad B \overset{def}{=} \bigcup_{k\geq 0} \left(\mathit{incorrect}_k \cdot \Sigma^*\right)$$

G consists of words that describe some length-k prefix of M's run followed by x_k or fewer symbols a; B contains all words describing the run of M up to length-k, followed by an incorrect zero-test, and then anything.

We claim that this language L is recognised by a 2-H-VASS. To see this, again build a VASS that weakly simulates M as done before, for example in the proof of Theorem 17. This will simulate increment and decrement operations faithfully, reading letters inc_i or dec_i, respectively. For any step $q \xrightarrow{\mathrm{ztest}_i} q'$ in M, the VASS $\mathcal{A}$ will have a transition $(q, \mathrm{ztest}_i, \mathbf{0}, q')$ as well as one that reads ztest_i, decreases counter i and leads to a universal state. This allows to accept exactly all words in B. In addition, from any state q of M, $\mathcal{A}$ can move to a new countdown phase: there is a transition $q \xrightarrow{a,\mathbf{0}} c$ to a new, final, control state that can continue to read $a's$ while at least one of the counters remains non-zero. This allows to accept exactly all words in G. Note that the only non-determinism is for letters ztest_i when M's ith counter after reading ρ_i is not zero. In this case, the only language-maximal choice is to move to the universal state. The constructed system is therefore history-deterministic.

To conclude the proof, we argue that L is regular iff ρ visits only finitely many configurations. Indeed, if so, then G is finite because all x_i, $i \leq k$ are bounded, and B is regular because at most k many words $\mathit{incorrect}_k$ exist. So L is the finite union of regular languages and thus regular.

Conversely, suppose that M's run ρ visits infinitely many different configurations. Then in particular, there are infinitely many faithful prefixes ρ_k. Let us assume towards contradiction that L is regular and recognised by a DFA with d many states. We pick a prefix ρ_k so that $x_k > d$ and consider the word $\rho_k a^{x_k} \in L$. While reading the suffix a^m, our DFA must repeat some cycle of length $c \leq d$. But then it must also accept $\rho_k a^{x_k+c} \notin L$ by going through that cycle twice.

The same proof goes through for the reachability semantics if we set $G \overset{def}{=} \bigcup_{k\geq 0} (\rho_k \cdot a^{=x_k})$ and $B \overset{def}{=} \bigcup_{k\geq 0} \left(\mathit{incorrect}_k \cdot \Sigma^{\geq x_k - 1}\right)$. Then again, if the run of M visits finitely many configurations then both G and B are regular. Otherwise G is not regular. The extra symbols at the end (of words in G and B) allow a run of the VASS $\mathcal{A}$ to decrease the counters to 0 and accept (and therefore to conclude that language $L = G \uplus B$ is in 2-$\mathcal{H}^0$). ◀

References

1 Stanislav Böhm, Stefan Göller, Simon Halfon, and Piotr Hofman. On büchi one-counter automata. In *International Symposium on Theoretical Aspects of Computer Science*, volume 66 of *LIPIcs*, pages 14:1–14:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.STACS.2017.14.

2 Stanislav Böhm, Stefan Göller, and Petr Jancar. Bisimulation equivalence and regularity for real-time one-counter automata. *Journal of Computer and System Sciences*, 80(4):720–743, 2014. doi:10.1016/j.jcss.2013.11.003.

3 Udi Boker and Karoliina Lehtinen. Good for games automata: From nondeterminism to alternation. In *International Conference on Concurrency Theory*, volume 140 of *LIPIcs*, pages 19:1–19:16, 2019.

4 Udi Boker and Karoliina Lehtinen. History Determinism vs. Good for Gameness in Quantitative Automata. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, volume 213 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 38:1–38:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.FSTTCS.2021.38.

5 Udi Boker and Karoliina Lehtinen. Token games and history-deterministic quantitative automata. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, pages 120–139. Springer International Publishing, 2022.

6 Sougata Bose, Thomas A. Henzinger, Karoliina Lehtinen, Sven Schewe, and Patrick Totzke. History-deterministic timed automata are not determinizable. In *International Workshop on Reachability Problems*, 2022.

7 Sougata Bose, David Purser, and Patrick Totzke. History-deterministic vector addition systems. *CoRR*, abs/2305.01981, 2023.

8 Thomas Colcombet. The theory of stabilisation monoids and regular cost functions. In *International Colloquium on Automata, Languages and Programming*, pages 139–150, 2009.

9 Wojciech Czerwinski and Piotr Hofman. Language inclusion for boundedly-ambiguous vector addition systems is decidable. In *International Conference on Concurrency Theory*, volume 243 of *LIPIcs*, pages 16:1–16:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.CONCUR.2022.16.

10 Wojciech Czerwinski and Lukasz Orlikowski. Reachability in vector addition systems is ackermann-complete. In *Annual Symposium on Foundations of Computer Science*, pages 1229–1240. IEEE, 2021. doi:10.1109/FOCS52979.2021.00120.

11 Enzo Erlich, Shibashis Guha, Ismaël Jecker, Karoliina Lehtinen, and Martin Zimmermann. History-deterministic parikh automata. In *International Conference on Concurrency Theory*, LIPIcs. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023.

12 Olivier Finkel and Michal Skrzypczak. On the topological complexity of ω-languages of non-deterministic Petri nets. *Information Processing Letters*, 114(5):229–233, 2014. doi:10.1016/j.ipl.2013.12.007.

13 Sheila A. Greibach. Remarks on blind and partially blind one-way multicounter machines. *Theoretical Computer Science*, 7:311–324, 1978. doi:10.1016/0304-3975(78)90020-8.

14 Shibashis Guha, Ismaël Jecker, Karoliina Lehtinen, and Martin Zimmermann. A Bit of Nondeterminism Makes Pushdown Automata Expressive and Succinct. In *International Symposium on Mathematical Foundations of Computer Science*, volume 202 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 53:1–53:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.MFCS.2021.53.

15 Michel Henri Theódore Hack. Petri net language. 1976.

16 Thomas A. Henzinger, Karoliina Lehtinen, and Patrick Totzke. History-deterministic timed automata. In *International Conference on Concurrency Theory*, 2022.

17 Thomas A. Henzinger and Nir Piterman. Solving games without determinization. In *EACSL Annual Conference on Computer Science Logic*, pages 395–410. Springer Berlin Heidelberg, 2006.

18 Piotr Hofman, Slawomir Lasota, Richard Mayr, and Patrick Totzke. Simulation problems over one-counter nets. *Logical Methods in Computer Science*, 12(1), 2016. `doi:10.2168/LMCS-12(1:6)2016`.

19 Piotr Hofman and Patrick Totzke. Trace inclusion for one-counter nets revisited. *Theoretical Computer Science*, 11174, 2017. `doi:10.1016/j.tcs.2017.05.009`.

20 Petr Jancar. Nonprimitive recursive complexity and undecidability for Petri net equivalences. *Theoretical Computer Science*, 256(1-2):23–30, 2001. `doi:10.1016/S0304-3975(00)00100-6`.

21 Petr Jancar, Javier Esparza, and Faron Moller. Petri nets and regular processes. *Journal of Computer and System Sciences*, 59(3):476–503, 1999. `doi:10.1006/jcss.1999.1643`.

22 Matthias Jantzen. Language theory of Petri nets. In *Petri Nets: Central Models and Their Properties, Advances in Petri Nets 1986, Part I, Proceedings of an Advanced Course, Bad Honnef, Germany, 8-19 September 1986*, volume 254 of *LNCS*, pages 397–412. Springer, 1986. `doi:10.1007/BFb0046847`.

23 Petr Jančar and Faron Moller. Checking regular properties of Petri nets. In *International Conference on Concurrency Theory*, pages 348–362, 1995.

24 Richard M. Karp and Raymond E. Miller. Parallel program schemata. *Journal of Computer and System Sciences*, 3(2):147–195, 1969. `doi:10.1016/S0022-0000(69)80011-5`.

25 S. Rao Kosaraju. Decidability of reachability in vector addition systems. In *STOC'82*, pages 267–281, 1982. `doi:10.1145/800070.802201`.

26 Orna Kupferman, Shmuel Safra, and Moshe Y Vardi. Relating word and tree automata. *Annals of Pure and Applied Logic*, 138(1-3):126–146, 2006.

27 Karoliina Lehtinen and Martin Zimmermann. Good-for-games ω-pushdown automata. *Logical Methods in Computer Science*, 18, 2022.

28 Jérôme Leroux. The reachability problem for Petri nets is not primitive recursive. In *Annual Symposium on Foundations of Computer Science*, pages 1241–1252. IEEE, 2021. `doi:10.1109/FOCS52979.2021.00121`.

29 Richard J. Lipton. The reachability problem requires exponential space. Technical Report 62, Yale University, 1976.

30 Marvin L. Minsky. *Computation: finite and infinite machines*. Prentice-Hall, Inc., 1967.

31 Reino Niskanen, Igor Potapov, and Julien Reichert. Undecidability of two-dimensional robot games. In *International Symposium on Mathematical Foundations of Computer Science*, volume 58 of *LIPIcs*, pages 73:1–73:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. `doi:10.4230/LIPIcs.MFCS.2016.73`.

32 Aditya Prakash and K. S. Thejaswini. On history-deterministic one-counter nets. In *International Conference on Foundations of Software Science and Computational Structures*. Springer, 2023. `doi:10.1007/978-3-031-30829-1_11`.

33 Charles Rackoff. The covering and boundedness problems for vector addition systems. *Theoretical Computer Science*, pages 223–231, 1978. `doi:10.1016/0304-3975(78)90036-1`.

34 Klaus Reinhardt. Reachability in Petri nets with inhibitor arcs. *Electronic Notes in Theoretical Computer Science*, 223:239–264, 2008. `doi:10.1016/j.entcs.2008.12.042`.

35 Michal Skrzypczak. Büchi VASS recognise w-languages that are sigma^1_1 - complete. *CoRR*, abs/1708.09658, 2017.

36 Leslie G. Valiant and Mike Paterson. Deterministic one-counter automata. *Journal of Computer and System Sciences*, 10(3):340–350, 1975. `doi:10.1016/S0022-0000(75)80005-5`.

37 Rüdiger Valk and Guy Vidal-Naquet. Petri nets and regular languages. *Journal of Computer and System Sciences*, 23(3):299–325, 1981. `doi:10.1016/0022-0000(81)90067-2`.