

Reachability in Two-Dimensional Vector Addition Systems with States is PSPACE-complete

Michael Blondin^{*†‡}, Alain Finkel^{†§}, Stefan Göller^{†¶||}, Christoph Haase^{†§||} and Pierre McKenzie^{*†**}

^{*}DIRO, Université de Montréal, Canada

{blondimi, mckenzie}@iro.umontreal.ca

[†]Laboratoire Spécification et Vérification (LSV) & CNRS, ENS de Cachan, France

{finkel, goeller, haase}@lsv.ens-cachan.fr

Abstract—Known to be decidable since 1981, there still remains a huge gap between the best known lower and upper bounds for the reachability problem for vector addition systems with states (VASS). Here the problem is shown PSPACE-complete in the two-dimensional case, vastly improving on the doubly exponential time bound established in 1986 by Howell, Rosier, Huynh and Yen. Coverability and boundedness for two-dimensional VASS are also shown PSPACE-complete, and reachability in two-dimensional VASS and in integer VASS under unary encoding are considered.

I. INTRODUCTION

Petri nets have a long history. Since their introduction [19] by Petri in 1962, thousands of papers on Petri nets have been published. Nowadays, Petri nets find a variety of applications, ranging, for instance, from modeling of biological, chemical and business processes to the formal verification of concurrent programs, see *e.g.* [1], [4], [8], [21], [27]. For the analysis of their algorithmic properties, Petri nets are often equivalently viewed as *vector addition systems with states* (VASS), and we will adopt this view throughout this paper. A VASS comprises a finite-state controller with a finite number of counters ranging over the natural numbers. The number of counters is usually referred to as the *dimension* of the VASS, and we write *d*-VASS to denote VASS in dimension *d*. When taking a transition, a VASS can add or subtract an integer from a counter, provided that the resulting counter values are greater than or equal to zero; otherwise the transition is blocked. A configuration of a VASS is a tuple consisting of a control state and an assignment of natural numbers to the counters. The central decision problem for VASS is *reachability*: given two configurations, is there a path connecting them in the infinite graph induced by the VASS?

Resolving decidability of the VASS reachability problem required tremendous effort, extending until 1981. This was achieved by Mayr [18], who built upon an earlier partial proof by Sacerdote and Tenney [23]. Mayr’s argument was

then polished and simplified by Kosaraju [11] in 1982, and Kosaraju’s argument was in turn simplified ten years later by Lambert [12]. More recently, beginning in 2009, Leroux began developing a fundamentally different approach to deciding the VASS reachability problem [15], [16]. Finally, at the time of writing of this paper, Leroux and Schmitz could establish the first explicit upper bound for VASS reachability and show that it can be decided in F_{ω^3} [13].

Milestones in the work on the complexity of the VASS reachability problem include Lipton’s 1976 proof that the problem, regardless of the choice of encoding for numbers but without fixed dimension, is EXPSPACE-hard [17]. Yet our knowledge of the situation for any fixed dimension *d* is vastly lacking. For 1-VASS, reachability under unary encoding is easily seen to be NL-complete: the hardness is inherited from graph reachability and the upper bound follows from a simple pumping argument. Under binary encoding, 1-VASS reachability is known to be NP-complete [5]. As a substantial contribution towards showing decidability of the general problem, Hopcroft and Pansiot in 1979 showed the two-dimensional case decidable [9]. At the core of their proof lies an intricate algorithm that implicitly exploits the fact that the reachability set of a 2-VASS is semi-linear. Exhibiting a 3-VASS with a reachability set that is not semi-linear, Hopcroft and Pansiot could show that their method breaks down for *d*-VASS for any *d* greater than 2. Further complexity aspects were left unanswered in [9]. In 1986, Howell, Rosier, Huynh and Yen [10] observed that Hopcroft and Pansiot’s algorithm runs in nondeterministic doubly-exponential time, under both unary and binary encoding. They then managed to improve this bound from nondeterministic to deterministic doubly-exponential time, and to identify a 2-VASS family on which Hopcroft and Pansiot’s algorithm requires this much. To summarize the state of the art today, 2-VASS reachability in 2-EXPTIME has stood since 1986, with its NL-hardness and NP-hardness depending on number encodings. For any *d* greater than 2, reachability is in F_{ω^3} [13].

The main contribution of this paper is to show that reachability in 2-VASS is PSPACE-complete when numbers are encoded in binary. The PSPACE lower bound follows as an easy consequence of a recent result by Fearnley and Jurdziński who showed PSPACE-completeness of reachability in bounded one-counter automata [3]. Our PSPACE upper

[‡]Supported by the Fonds québécois de la recherche sur la nature et les technologies and by the French Centre national de la recherche scientifique.

[§]Supported by the French Agence nationale de la recherche, REACHARD (grant ANR-11-BS02-001).

[¶]Parts of this work were carried out while the author was at Technische Universität München, Germany.

^{||}Supported by Labex Digicosme, Univ. Paris-Saclay, project VERICONISS

^{**}Supported by the Natural Sciences and Engineering Research Council of Canada and by the “Chaire Digiteo, ENS Cachan — École Polytechnique”.

bound is obtained from showing that the length of a run witnessing reachability can be exponentially bounded in the size of the input, and consequently the existence of such a run can be decided by a PSPACE-algorithm. The difficult and main part of this paper is, of course, to establish the exponential upper bound on the length of witnessing runs. Our starting point is a careful analysis of an argument developed by Leroux and Sutre in [14] for the purpose of showing that reachability relations of 2-VASS can be captured by bounded languages, *i.e.*, speaking in the terminology of [14], 2-VASS can be *flattened*. More precisely, this means that for any 2-VASS there is a finite set S of regular languages over the set of transitions, viewed as an alphabet, each of the form $u_0 v_1^* u_1 \cdots v_k^* u_k$ such that for any two configurations reachable from one another there exists a witnessing run in the language defined by S . The paper of Leroux and Sutre reports that from any 2-VASS it is possible to construct such a bounded language; it has however not appeared as a fully refereed publication and omits some proof details. Thus, while we follow closely the proof strategy presented in [14], we provide a proof that 2-VASS can be flattened by small bounded languages. In doing so we develop new arguments setting the stage for the much deeper analysis of our constructions required for the purpose of establishing a PSPACE upper bound. In summary, we contribute:

- 1) a PSPACE-completeness proof for 2-VASS reachability,
- 2) a proof that 2-VASS can be flattened by bounded languages that have small presentations, and
- 3) remarks that reachability in 2-VASS with numbers encoded in unary is NL-hard and in NP.

Section II below fixes notation. Section III gives an overview of our main results. Section IV proves our main technical result, namely that the global reachability relation of any 2-VASS can be characterized by small *bounded* languages, also known in the literature as *linear path schemes*. Section V proves that 2-VASS reachability is PSPACE-complete and gives further corollaries and implications. Section VI concludes with open problems and directions for future work.

Due to space constraints, the proofs of some lemmas are only sketched, and full proofs can be found in the extended version of this paper¹.

II. PRELIMINARIES

General notation. By $\mathbb{N} \stackrel{\text{def}}{=} \{0, 1, 2, \dots\}$, $-\mathbb{N} \stackrel{\text{def}}{=} \{0, -1, -2, \dots\}$ and \mathbb{Z} we denote the sets of non-negative integers, non-positive integers and integers, respectively. By \mathbb{Q} and $\mathbb{Q}_{\geq 0}$ we denote the set of rationals and non-negative rationals, respectively. For any $i, j \in \mathbb{Z}$, we define $[i, j] \stackrel{\text{def}}{=} \{i, i+1, \dots, j\}$. For each $k \in \mathbb{Z}$ we write $[k, \infty)$ to denote $\{z \in \mathbb{Z} : z \geq k\}$. A *quadrant* is one of the four sets $\mathbb{N}^2, -\mathbb{N} \times \mathbb{N}, \mathbb{N} \times -\mathbb{N}$ and $-\mathbb{N} \times -\mathbb{N}$. Given two vectors $\mathbf{u} = (u_1, \dots, u_d), \mathbf{v} = (v_1, \dots, v_d) \in \mathbb{Z}^d$, we denote by $\mathbf{u} + \mathbf{v} \stackrel{\text{def}}{=} (u_1 + v_1, \dots, u_d + v_d)$ their component-wise sum. Given two sets $U, V \subseteq \mathbb{Z}^d$, we

let $U + V \stackrel{\text{def}}{=} \{\mathbf{u} + \mathbf{v} : \mathbf{u} \in U, \mathbf{v} \in V\}$. The *norm* of a vector $\mathbf{u} = (u_1, \dots, u_d)$ is defined as $\|\mathbf{u}\| \stackrel{\text{def}}{=} \max\{|u_i| : i \in [1, d]\}$. The *norm* of a matrix $A = (a_{ij}) \in \mathbb{Z}^{m \times n}$ is defined as $\|A\| \stackrel{\text{def}}{=} n \cdot \max\{|a_{ij}| : i \in [1, m], j \in [1, n]\}$. For any word $w = a_1 \cdots a_n \in \Sigma^n$ over some alphabet Σ , $w[i, j]$ denotes $a_i a_{i+1} \cdots a_j$ for all $i, j \in [1, n]$.

Graphs, Parikh Images and Linear Path Schemes. For each set Σ , a Σ -labeled directed graph is a pair $G = (U, E)$, where U is a set of *vertices* and $E \subseteq U \times \Sigma \times U$ is a set of *edges*. We say G is *finite* if U and E are finite. Let $\pi = (u_1, a_1, u'_1) \cdots (u_k, a_k, u'_k) \in E^k$. The *Parikh image* Parikh_π of π is the mapping from Σ to \mathbb{N} such that $\text{Parikh}_\pi(a) = |\{i \in [1, k] : a_i = a\}|$ for each $a \in \Sigma$. If $X \subseteq E^*$, then Parikh_X denotes the set of Parikh images of X , *i.e.* $\text{Parikh}_X = \{\text{Parikh}_\pi : \pi \in X\}$. We say π is a *path* (from u_1 to u'_k) if $u'_i = u_{i+1}$ for all $i \in [1, k-1]$. A path π is a *cycle* if $k \geq 1$ and $u_1 = u'_k$, and *cycle-free* if no infix of π is a cycle. A cycle π is called *simple* if π is the only infix of π that is a cycle. A *linear path scheme* (from $u \in U$ to $u' \in U$) is a regular expression (whose language will be referred to implicitly) of the form

$$\rho = \alpha_0 \beta_1^* \alpha_1 \cdots \beta_k^* \alpha_k,$$

where $\alpha_0 \beta_1 \alpha_1 \cdots \beta_k \alpha_k$ is a path (from u to u') and each β_i is a cycle. We define its *length* as $|\rho| \stackrel{\text{def}}{=} |\alpha_0 \beta_1 \alpha_1 \cdots \beta_k \alpha_k|$ and its **-length* as $|\rho|_* \stackrel{\text{def}}{=} k$. We call β_1, \dots, β_k the *cycles* of ρ . Note that every path is a linear path scheme by taking $k = 0$. The general structure of a linear path scheme is illustrated in Fig. 1.

Vector Addition Systems with States. A *vector addition system with states* (VASS) in dimension d (d -VASS for short) is a finite \mathbb{Z}^d -labeled directed graph $V = (Q, T)$, where Q will be referred to as the *states* of V , and where T will be referred to as *transitions* of V . The *size* of V is defined as $|V| \stackrel{\text{def}}{=} |Q| + |T| \cdot d \cdot \lceil \log_2 \|T\| \rceil$, where $\|T\|$ denotes the absolute value of the largest number that appears in T , *i.e.* $\|T\| \stackrel{\text{def}}{=} \max\{\|z\| : (p, z, q) \in T\}$. We say that V is encoded in *binary* when we use this definition of $|V|$, which we will use as standard encoding in this paper. Alternatively, when we set $|V| \stackrel{\text{def}}{=} |Q| + |T| \cdot d \cdot \|T\|$ we say that V is encoded in *unary*.

Subsequently, $Q \times \mathbb{Z}^d$ denotes the set of *configurations* of V . Note that in the literature, the set of configurations is usually $Q \times \mathbb{N}^d$, however in this paper we will often deal with VASS whose counters can take integer values. For the sake of readability, we write configurations $(q, (z_1, \dots, z_d))$ and (q, \mathbf{z}) as $q(z_1, \dots, z_d)$ and $q(\mathbf{z})$, respectively.

For every subset $\mathbb{A} \subseteq \mathbb{Z}^d$, $p(\mathbf{u}), q(\mathbf{v}) \in Q \times \mathbb{A}$ and every transition $t = (p, \mathbf{z}, q)$, we write $p(\mathbf{u}) \xrightarrow{t}_{\mathbb{A}} q(\mathbf{v})$ whenever $\mathbf{v} = \mathbf{u} + \mathbf{z}$. We extend $\xrightarrow{t}_{\mathbb{A}}$ to sequences of transitions $\pi \in T^*$ as follows: $\xrightarrow{\pi}_{\mathbb{A}}$ is the smallest relation satisfying the following conditions for all configurations $p(\mathbf{u}), q(\mathbf{v}), r(\mathbf{w}) \in Q \times \mathbb{A}$ and all $t \in T$,

¹The extended version can be obtained from <http://arxiv.org/abs/1412.4259>.

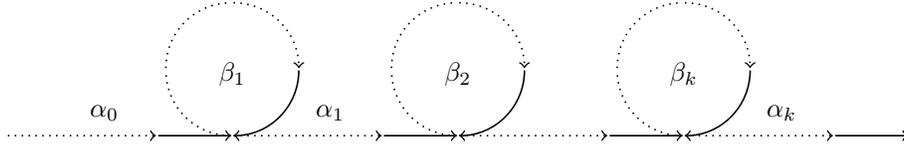


Fig. 1. Illustration of the structure of a linear path scheme $\rho = \alpha_0 \beta_1^* \alpha_1 \cdots \beta_k^* \alpha_k$.

- $p(\mathbf{u}) \xrightarrow{\varepsilon} p(\mathbf{u})$ and
- if $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ and $q(\mathbf{v}) \xrightarrow{t} r(\mathbf{w})$, then $p(\mathbf{u}) \xrightarrow{\pi t} r(\mathbf{w})$.

We extend \xrightarrow{t} to languages $L \subseteq T^*$ in the natural way, $L \xrightarrow{t} \stackrel{\text{def}}{=} \bigcup \{ \xrightarrow{\pi} : \pi \in L \}$. We write $\xrightarrow{*}$ to denote $\xrightarrow{T^*}$. An \mathbb{A} -run from $q_0(\mathbf{v}_0) \in Q \times \mathbb{A}$ to $q_k(\mathbf{v}_k) \in Q \times \mathbb{A}$ that is induced by a path $\pi = t_1 \cdots t_k$ is a sequence of configurations $q_0(\mathbf{v}_0) \xrightarrow{t_1} q_1(\mathbf{v}_1) \cdots \xrightarrow{t_k} q_k(\mathbf{v}_k)$ that we sometimes just abbreviate by $q_0(\mathbf{v}_0) \xrightarrow{\pi} q_k(\mathbf{v}_k)$. When $\mathbb{A} = \mathbb{N}^d$ we also refer to an \mathbb{A} -run as a *run*.

Throughout this paper, we refer to $\xrightarrow{*}$ as the *reachability relation*, and $\xrightarrow{*}$ as the \mathbb{Z} -*reachability relation*. Let $\pi = (p_1, \mathbf{z}_1, p_1) \cdots (p_k, \mathbf{z}_k, p_k) \in T^k$ for some $k \geq 0$. The *displacement of π* is $\delta(\pi) \stackrel{\text{def}}{=} \sum_{i=1}^k \mathbf{z}_i$, and the definition naturally extends to languages $L \subseteq T^*$ as $\delta(L) \stackrel{\text{def}}{=} \{ \delta(\pi) : \pi \in L \}$. Note in particular that if $\text{Parikh}_{\rho'} \subseteq \text{Parikh}_{\rho}$, then $\delta(\rho') \subseteq \delta(\rho)$.

III. MAIN RESULTS

In this paper, our main interest is in the reachability problem for 2-VASS, formally defined as follows:

2-VASS REACHABILITY

- INPUT:** A 2-VASS $V = (Q, T)$ and configurations $p(\mathbf{u})$ and $q(\mathbf{v})$ from $Q \times \mathbb{N}^2$.
- QUESTION:** Is there a run from $p(\mathbf{u})$ to $q(\mathbf{v})$, *i.e.* does $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ hold?

In order to determine the complexity of this problem, we show that the reachability relation of any 2-VASS can be defined by a finite set of linear path schemes. In particular, we are able to show strong bounds on their lengths and $*$ -lengths. For example, consider the 2-VASS V depicted in Fig. 2. Since V contains nested loops, *e.g.* $(t_1 t_3^* t_2)^*$, we cannot directly read off a characterization of its reachability set by a finite set of linear path schemes. However, by carefully unraveling loops we obtain the reachability set from the set of the subsequent

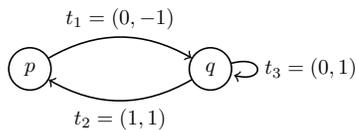


Fig. 2. Example of a 2-VASS.

linear path schemes, and in particular this means that V can be flattened (cf. Fig. 3):

$$\begin{aligned}
 p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v}) &\iff p(\mathbf{u}) \xrightarrow{t_1 t_3^* t_2 (t_1 t_2)^* t_1} q(\mathbf{v}) \\
 p(\mathbf{u}) \xrightarrow{*} p(\mathbf{v}) &\iff p(\mathbf{u}) \xrightarrow{t_1 t_3^* t_2 (t_1 t_2)^* \cup \varepsilon} p(\mathbf{v}) \\
 q(\mathbf{u}) \xrightarrow{*} p(\mathbf{v}) &\iff q(\mathbf{u}) \xrightarrow{(t_2 t_1)^* t_3^* t_2} p(\mathbf{v}) \\
 q(\mathbf{u}) \xrightarrow{*} q(\mathbf{v}) &\iff q(\mathbf{u}) \xrightarrow{(t_2 t_1)^* t_3^*} q(\mathbf{v})
 \end{aligned}$$

We will show that such a flattening exists for any 2-VASS. More precisely, our main technical result states that the global reachability relation of any 2-VASS $V = (Q, T)$ can be defined via a set of linear path schemes whose lengths can be polynomially bounded in $|Q| + \|T\|$, and *a fortiori* are at most exponential in $|V|$, and whose $*$ -lengths are at most quadratic in $|Q|$:

Theorem 1. Let $V = (Q, T)$ be a 2-VASS. There is a finite set S of linear path schemes such that²

- $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ if, and only if, $p(\mathbf{u}) \xrightarrow{S} q(\mathbf{v})$, and
- $|\rho| \leq (|Q| + \|T\|)^{O(1)}$ and $|\rho|_* \leq O(|Q|^2)$ for each $\rho \in S$.

Having established Theorem 1, we can show that proving the existence of a run between two reachable configurations in a 2-VASS reduces to checking the existence of a solution for suitably constructed systems of linear Diophantine inequalities that depend on S and the properties listed in Theorem 1. The absence of nested cycles in linear path schemes in S is crucial to this reduction. By application of standard bounds from integer linear programming, this in turn enables us to bound the length of paths witnessing reachability, and to prove the upper bound of the the main theorem of this paper in Section V:

Theorem 2. 2-VASS REACHABILITY is PSPACE-complete.

IV. PROOF OF THEOREM 1

In this section, we prove Theorem 1 and show that runs of a 2-VASS $V = (Q, T)$ are captured by a finite set of linear path schemes each of which has length at most $(|Q| + \|T\|)^{O(1)}$

²The expanded technical meaning of this statement is that there are constants c_1 and c_2 such that for every 2-VASS $V = (Q, T)$ there exists a finite set S of linear path schemes, each of length $\leq (|Q| + \|T\|)^{c_1}$ and of $*$ -length $\leq c_2 |Q|^2$, with the property that for every $p(\mathbf{u}), q(\mathbf{v}) \in Q \times \mathbb{N}^2$, $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ if, and only if, $p(\mathbf{u}) \xrightarrow{S} q(\mathbf{v})$. The more familiar statements of this theorem and of lemmas of a similar nature in the rest of the paper were chosen to avoid clutter and to downplay the role of the precise constants.

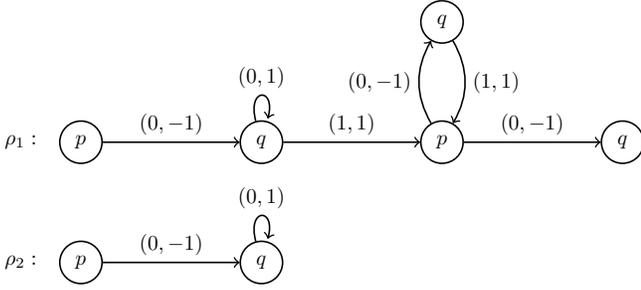


Fig. 3. Illustration of a set $S = \{\rho_1, \rho_2\}$ of linear path schemes defining the reachability relation from p to q of the 2-VASS V depicted in Fig. 2. Here, $\rho_1 = t_1 t_3^* t_2 (t_1 t_2)^* t_1$, $\rho_2 = t_1 t_3^*$, and $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ if, and only if, $p(\mathbf{u}) \xrightarrow{S} q(\mathbf{v})$.

and $*$ -length at most $O(|Q|^2)$. In order to construct this finite set of linear path schemes, we consider the following three types of runs $p(u_1, u_2) \xrightarrow{\pi} q(v_1, v_2)$, depicted in Fig. 4:

- 1) Both counter values of $p(u_1, u_2)$ and of $q(v_1, v_2)$ are sufficiently large and $p = q$, but intermediate configurations on the run $p(u_1, u_2) \xrightarrow{\pi} q(v_1, v_2)$ may have arbitrarily small counter values.
- 2) For all configurations of the run $p(u_1, u_2) \xrightarrow{\pi} q(v_1, v_2)$ both counter values are sufficiently large.
- 3) For all configurations of the run $p(u_1, u_2) \xrightarrow{\pi} q(v_1, v_2)$ at least one counter value is small.

In Subsections IV-A, IV-B and IV-C, we will show how to construct linear path schemes for these three types of runs. Then, in Subsection IV-D, we prove Theorem 1 by showing that any run can be decomposed as finitely many runs of these types.

In some more detail, the first step is to show in Section IV-A that Parikh images of finite labeled graphs can be captured by linear path schemes of polynomial size. This will allow us to prove that \mathbb{Z} -reachability, *i.e.* runs in which counter values may drop below zero, can be captured by linear path schemes of polynomial size. We then give in Section IV-B an effective decomposition of certain linear sets in dimension two into semi-linear sets with special properties, and use this decomposition in order to derive together with the results in Section IV-A linear path schemes of size $(|Q| + \|T\|)^{O(1)}$ and constant $*$ -length for runs of type (1). Linear path schemes for runs of type (2) will then be seen to follow from the type (1) case.

For runs of type (3), in Section IV-C we construct linear path schemes for 1-VASS and show that runs of a 2-VASS that stay within an “L-shaped band” are, essentially, runs of a 1-VASS. Our analysis of such runs of type (3) is a simple consequence of certain normal forms of shortest runs in one-counter automata, which 1-VASS are a subclass of, by Valiant and Paterson [26].

Similarities and differences in comparison with [14]. Our proof strategy of considering the three kinds of runs described above shares some similarities with [14]. There, the bounds on what we referred to above as “large” and “small” in the

runs of type (1), (2) and (3) are not explicitly calculated. Our proofs for obtaining rather tight bounds require new insights. We capture runs of type (1) by linear path schemes of size $(|Q| + \|T\|)^{O(1)}$, whereas in [14] the linear path schemes were of size at least exponential in $|Q|$. To prove the former, we establish a new upper bound on the presentation size of Parikh images of finite automata in Lemma 4 below, which is a result of independent interest. The difference between our runs of type (2) and the ones analyzed in [14] is that our runs have to stay in the “outside region” entirely, whereas in [14] the set of displacements of paths from q to q' is analyzed. Runs of type (3) are treated as special cases of their runs of type (2) in [14], whereas we invoke a result by Valiant and Paterson on normal forms of minimal runs in one-counter automata. Our final proof of Theorem 1 shows that each run can be factorized into segments of runs of types (1), (2) and (3) and requires a more careful treatment than in [14]. At every step, we have to ensure that the $*$ -length of the linear path schemes we construct stays polynomial in the number of control states. This aspect is neglected in [14] as it is of no interest for the goal of [14], however, for us it is by far the technically most challenging part and one of the cornerstones of our PSPACE upper bound.

A. Parikh images of finite directed graphs and \mathbb{Z} -reachability of d -VASS

The purpose of this subsection is to prove the following proposition.

Proposition 3. Let $V = (Q, T)$ be a d -VASS. There exists a finite set S of linear path schemes such that

- (i) $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ if, and only if, $p(\mathbf{u}) \xrightarrow{S} q(\mathbf{v})$, and
- (ii) $|\rho| \leq 2 \cdot |Q| \cdot |T|$ and $|\rho|_* \leq |T|$ for each $\rho \in S$.

In order to prove Proposition 3, we will prove suitable bounds on the representation size of the Parikh images of paths of a Σ -labeled finite graphs (or equivalently, nondeterministic finite automata) in terms of linear path schemes. Even though estimations on this size have been made in the literature (*e.g.* in [14] or [25, Prop. 7.3.4]), we are not aware of any in which the $*$ -length is linear in the number of edges (in the aforementioned references, the $*$ -length may be exponential in the number of control states).

Lemma 4. Let $G = (U, E)$ be a finite Σ -labeled graph. There exists a finite set S of linear path schemes such that

- (i) $\{\text{Parikh}_\pi : \pi \text{ is a path in } G\} = \{\text{Parikh}_\rho : \rho \in S\}$, and
- (ii) $|\rho| \leq 2 \cdot |U| \cdot |E|$ and $|\rho|_* \leq |E|$ for each $\rho \in S$.

Proof. We first provide some additional definitions. Let $\sigma, \sigma' : E \rightarrow \mathbb{N}$ be mappings and let X be a set of such mappings. We define $\sigma + \sigma' \in \mathbb{N}^E$ as $(\sigma + \sigma')(e) \stackrel{\text{def}}{=} \sigma(e) + \sigma'(e)$ for each $e \in E$ and $X + \sigma \stackrel{\text{def}}{=} \{\tau + \sigma : \tau \in X\}$. For each $u \in U$, let $\text{in}(u) \stackrel{\text{def}}{=} \{(u', a, u'') \in E : u'' = u\}$ and $\text{out}(u) \stackrel{\text{def}}{=} \{(u', a, u'') \in E : u' = u\}$ denote the set of incoming and outgoing edges of

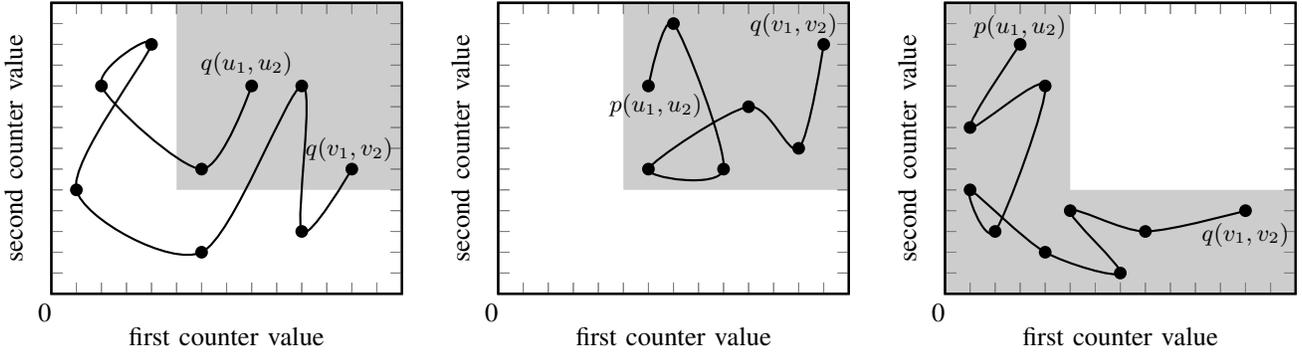


Fig. 4. Example of the three types of runs. The region depicted in each case is the positive quadrant in the Cartesian plane. (1) left: run from q to q starting and ending sufficiently high; (2) middle: run staying sufficiently high; (3) right: run within an L-shaped band, *i.e.*, running high on at most one component at a time.

u , respectively. We say that σ is *flow-preserving* if for every $u \in U$ we have

$$\sum_{e \in \text{in}(u)} \sigma(e) = \sum_{e \in \text{out}(u)} \sigma(e).$$

We will show the following claim:

Claim. Let $\pi \in E^*$ be a path. There exists some $h \geq 1$, a sequence of linear path schemes $\rho_1, \dots, \rho_h \subseteq E^*$, and a sequence $\sigma_1, \dots, \sigma_h \in \mathbb{N}^E$ such that

- (a) ρ_1 is a path of length at most $|U| \cdot |E|$ that visits each vertex of π at least once,
 - (b) σ_1 is flow-preserving, and
 - (c) $\text{Parikh}_\pi = \text{Parikh}_{\rho_1} + \sigma_1$,
- and for every $1 < i \leq h$,
- (1) ρ_i is a linear path scheme that can be obtained from ρ_1 by inserting $i - 1$ simple cycles (in the form β^*),
 - (2) σ_i is flow-preserving,
 - (3) $\text{Parikh}_{\rho_{i-1}} + \sigma_{i-1} \subseteq \text{Parikh}_{\rho_i} + \sigma_i$,
 - (4) $\sigma_{i-1}(e) \geq \sigma_i(e)$ for all $e \in E$ and there exists some $e \in E$ s.t. $\sigma_{i-1}(e) > \sigma_i(e) = 0$, and
 - (5) $\sigma_h(e) = 0$ for all $e \in E$.

First observe that due to (4) we have $h \leq |E|$, and due to (1) we have $|\rho_i| \leq |\rho_1| + |U| \cdot (i-1)$. Thus, $|\rho_h| \leq |\rho_1| + |U| \cdot |E| \leq 2 \cdot |U| \cdot |E|$, where the last inequality is due to (a). Moreover $|\rho_h|_* \leq |E|$ due to (1) and $h \leq |E|$.

Before proving the claim, let us first see how it proves the lemma. We define

$$S \stackrel{\text{def}}{=} \left\{ \rho : \rho \text{ is a linear path scheme, } |\rho| \leq 2 \cdot |U| \cdot |E| \text{ and } |\rho|_* \leq |E| \right\}.$$

Trivially, (ii) is satisfied. To establish (i), let us fix an arbitrary path π and obtain a linear path scheme $\rho_h \in S$ for π from the above claim. We have $\text{Parikh}_\pi \stackrel{(c)}{=} \text{Parikh}(\rho_1) + \sigma_1 \stackrel{(3)}{\subseteq} \text{Parikh}(\rho_2) + \sigma_2 \stackrel{(3)}{\subseteq} \dots \stackrel{(3)}{\subseteq} \text{Parikh}_{\rho_h} + \sigma_h \stackrel{(5)}{=} \text{Parikh}_{\rho_h}$ as required.

We now prove the claim. Let π be a path and let us first define ρ_1 and σ_1 such that (a), (b) and (c) are satisfied. The

path π can be decomposed as $\pi = e_1 \pi_1 \dots e_k \pi_k$ where $k \leq |U|$ and each $e_j = (u, a, u')$ is the first transition such that u or u' appears in π . We define ρ_1 and σ_1 as the result of the following iterative process: We initially set ρ_1 to π and set $\sigma_1(e) = 0$ for all $e \in E$; then we successively remove a simple cycle β from some π_j , and add Parikh_β to σ_1 . We repeat this process until no longer possible. The resulting ρ_1 is a path of length at most $|U| \cdot |E|$. Moreover, σ_1 is flow-preserving since we successively removed cycles only, and clearly $\text{Parikh}_\pi = \text{Parikh}_{\rho_1} + \sigma_1$, by construction. Thus (a), (b) and (c) hold.

Let us prove (1) to (5) by induction on $1 < i \leq h$. We only prove the induction step, the base case can be proven analogously. Let $E' \stackrel{\text{def}}{=} \{e \in E : \sigma_{i-1}(e) > 0\}$. If $E' = \emptyset$, then (5) holds and we are done. Thus, we assume that $E' \neq \emptyset$. Let us fix a choice function $\chi : E' \rightarrow E'$ satisfying

$$\chi(u_1, a, u_2) = (u'_1, a, u'_2) \implies u_2 = u'_1.$$

Note that χ exists since σ_{i-1} is flow-preserving by induction hypothesis. By the pigeonhole principle, there exists some $e \in E'$ and some $\ell \geq 0$ such that $\beta \stackrel{\text{def}}{=}} e \chi(e) \chi^2(e) \dots \chi^\ell(e)$ is a simple cycle. Without loss of generality let us assume that $c \stackrel{\text{def}}{=} \sigma_{i-1}(e) = \min\{\sigma_{i-1}(\chi^h(e)) : h \in [0, \ell]\}$, hence $\sigma_{i-1}(e)$ is minimal among all edges that lie on the simple cycle β . We define $\sigma_i \stackrel{\text{def}}{=} \sigma_{i-1} - \text{Parikh}_{\beta^c}$ and observe that σ_i is flow-preserving because β is a cycle and $\sigma_i \in \mathbb{N}^E$ due to minimality of c ; thus (2) and (4) are shown. Let $e = (u, a, u')$, hence β is a simple cycle from u to u . By (1) of induction hypothesis the linear path scheme ρ_{i-1} can be obtained from ρ_1 by inserting $(i-2)$ simple cycles and can hence be factorized as $\rho_{i-1} = \alpha \gamma$, where α is a linear path scheme from some state to u . We set $\rho_i \stackrel{\text{def}}{=} \alpha \beta^* \gamma$ and hence (1) holds. Furthermore, (3) holds due to $\text{Parikh}_{\rho_{i-1}} + \sigma_{i-1} = \text{Parikh}_{\rho_{i-1}} + \text{Parikh}_{\beta^c} + \sigma_i \subseteq \text{Parikh}_{\rho_i} + \sigma_i$. \square

Proof of Proposition 3. We have $T \subseteq Q \times \Sigma \times Q$ for some finite subset $\Sigma \subseteq \mathbb{Z}^d$. Let S be the finite set of linear path schemes from Lemma 4, then (ii) of Proposition 3 is clear. Let us now prove (i). We have $p(u) \xrightarrow{*}_{\mathbb{Z}^d} q(v)$ if, and only

if, there exists a path π from p to q in V such that $\mathbf{v} - \mathbf{u} = \sum_{\mathbf{z} \in \Sigma} \text{Parikh}_\pi(\mathbf{z}) \cdot \mathbf{z}$. By Lemma 4 (i), this is equivalent to the existence of some $\rho \in S$ from p to q , and some $f \in \text{Parikh}_\rho$ such that $\mathbf{v} - \mathbf{u} \in \sum_{\mathbf{z} \in \Sigma} f(\mathbf{z}) \cdot \mathbf{z}$. Now the latter existence of f is equivalent to $\mathbf{v} - \mathbf{u} \in \delta(\rho)$. This shows that $p(\mathbf{u}) \xrightarrow{*}_{\mathbb{Z}^d} q(\mathbf{v})$ if, and only if, $p(\mathbf{u}) \xrightarrow{S}_{\mathbb{Z}^d} q(\mathbf{v})$. \square

B. Starting and ending in “sufficiently large” configurations

The goal of this subsection is to prove that, given a 2-VASS V , there exists a sufficiently small bound D such that the reachability relation between any two configurations $q(u_1, v_1)$ and $q(u_2, v_2)$ with $u_1, u_2, v_1, v_2 \geq D$ can be captured by a finite set of small linear path schemes (in the sense of Theorem 1). In [14], this property is referred to as *ultimately flat*. As a consequence of this result, we can show that the reachability relation between arbitrary configurations for which there exists a run on which both counter values on all configurations stay above D can be captured by a finite set of small linear path schemes as well.

Proposition 5. Let $V = (Q, T)$ be a 2-VASS. There exist $D \leq (|Q| + \|T\|)^{O(1)}$ and finite sets of linear path schemes R, X such that for $\mathbb{O} \stackrel{\text{def}}{=} [D, \infty)^2$, $p, q \in Q$ and $\mathbf{u}, \mathbf{v} \in \mathbb{O}$,

- (a) • $q(\mathbf{u}) \xrightarrow{*}_{\mathbb{N}^2} q(\mathbf{v})$ if, and only if, $q(\mathbf{u}) \xrightarrow{R}_{\mathbb{N}^2} q(\mathbf{v})$, and
 - $|\rho| \leq (|Q| + \|T\|)^{O(1)}$ and $|\rho|_* \leq 2$ for every $\rho \in R$.
- (b) • $p(\mathbf{u}) \xrightarrow{*}_{\mathbb{O}} q(\mathbf{v})$ implies $p(\mathbf{u}) \xrightarrow{X}_{\mathbb{N}^2} q(\mathbf{v})$, and
 - $|\rho| \leq (|Q| + \|T\|)^{O(1)}$ and $|\rho|_* \leq 2 \cdot |Q|$ for every $\rho \in X$.

The proof of this proposition requires two intermediate steps. First, in Lemma 6 below we prove an effective decomposition of certain linear sets in dimension two into semi-linear sets with nice properties. Similar decompositions have been the cornerstone of the results by Hopcroft and Pansiot [9] and Leroux and Sutre [14]. The contribution of Lemma 6 is to establish a new proof from which we can obtain sufficiently small bounds on this decomposition. Next, in Lemma 7 we show how this decomposition can be applied in order to capture reachability instances by linear path schemes of $*$ -length two whose displacements all point into the same quadrant. This in turn enables us to prove Part (a) of Proposition 5, from which we can then prove Part (b).

Let us recall some definitions concerning semi-linear sets. Let $P = \{\mathbf{p}_1, \dots, \mathbf{p}_n\} \subseteq \mathbb{Z}^m$ and $\mathbb{D} \subseteq \mathbb{Q}_{\geq 0}$. The \mathbb{D} -cone generated by P is defined as

$$\text{cone}_{\mathbb{D}}(P) \stackrel{\text{def}}{=} \left\{ \sum_{i \in [1, n]} \lambda_i \cdot \mathbf{p}_i : \lambda_i \in \mathbb{D} \right\}.$$

A linear set $L(\mathbf{b}; P)$ is determined by a base vector $\mathbf{b} \in \mathbb{Z}^d$ and a finite set of period vectors $P \subseteq \mathbb{Z}^d$, where $L(\mathbf{b}; P) \stackrel{\text{def}}{=} \mathbf{b} + \text{cone}_{\mathbb{N}}(P)$. A semi-linear set is a finite union of linear sets. The norm $\|P\|$ of a finite set $P \subseteq \mathbb{Z}^d$ is defined as $\|P\| \stackrel{\text{def}}{=} \max\{\|\mathbf{p}\| : \mathbf{p} \in P\}$. Recall that $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^d$ are linearly dependent if $\mathbf{0} = \lambda_1 \cdot \mathbf{u} + \lambda_2 \cdot \mathbf{v}$ for some $\lambda_1, \lambda_2 \in \mathbb{Q} \setminus \{0\}$, and linearly independent otherwise.

We now show the following decomposition of linear sets.

Lemma 6. Let $\mathbf{b} \in \mathbb{Z}^2$, let $P \subseteq \mathbb{Z}^2$ be finite with $\mathbf{b} \in P$ and let Z be a quadrant. Then there exists a finite set of indices I such that $L(\mathbf{b}; P) \cap Z \subseteq \bigcup_{i \in I} L(\mathbf{c}_i; P_i)$, and for each $i \in I$ we have

- $|P_i| \leq 2$,
- $P_i \subseteq (P \cup L(\mathbf{b}; P)) \cap Z$, and
- there exists $e \leq \|P\|^{O(1)}$ such that $\{\mathbf{c}_i\} \cup (P_i \cap L(\mathbf{b}; P)) \subseteq \mathbf{b} + \text{cone}_{[0, e]}(P)$.

Proof sketch. It is sufficient to show the statement for $Z = \mathbb{N}^2$. Let $P = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ and let \mathbf{r} be a point in $L(\mathbf{b}; P) \cap \mathbb{N}^2$. By definition,

$$\mathbf{r} = \mathbf{b} + \lambda_1 \mathbf{p}_1 + \dots + \lambda_n \mathbf{p}_n$$

for some $\lambda_i \in \mathbb{N}$, $i \in [1, n]$. First, it is not difficult to show that we may assume that all but two of the λ_i do not exceed $O(\|P\|^2)$. Hence, with no loss of generality we may assume that $\mathbf{r} \in L(\mathbf{c}; P')$, where

$$\mathbf{c} \in \mathbf{b} + \text{cone}_{[0, e]}(P)$$

for some $e \leq \|P\|^{O(1)}$ and $P' = \{\mathbf{u}, \mathbf{v}\} \subseteq P$. The most interesting case is when \mathbf{u} and \mathbf{v} are linearly independent, which entails making a case distinction in which quadrant \mathbf{u} and \mathbf{v} lie. If both \mathbf{u} and \mathbf{v} lie in \mathbb{N}^2 then we are done.

If $\mathbf{u} \in \mathbb{N}^2$ and $\mathbf{v} \notin \mathbb{N}^2$ then we can show that there exists some natural number $\alpha \leq \|P\|^{O(1)}$ such that either

$$\begin{aligned} \begin{pmatrix} \alpha \\ 0 \end{pmatrix} &\in L(\mathbf{b}; P) \cap \text{cone}_{\mathbb{N}}(P') \text{ or} \\ \begin{pmatrix} 0 \\ \alpha \end{pmatrix} &\in L(\mathbf{b}; P) \cap \text{cone}_{\mathbb{N}}(P'), \end{aligned} \quad (1)$$

depending on the relative angle between \mathbf{u} and \mathbf{v} . In the following, assume $(0, \alpha) \in L(\mathbf{b}; P) \cap \text{cone}_{\mathbb{N}}(P')$, the other case follows symmetrically. A careful analysis allows to conclude that \mathbf{r} can equivalently be obtained as

$$\mathbf{r} \in \mathbf{c} + \omega \cdot \mathbf{v} + \text{cone}_{\mathbb{N}}(P'')$$

for some $\omega \leq \|P\|^{O(1)}$ and $P'' = \{\mathbf{u}, (0, \alpha)\}$, i.e., $\mathbf{r} \in L(\mathbf{c} + \omega \cdot \mathbf{v}; P'')$, which fulfills the requirements of the lemma.

Finally, if both $\mathbf{u} \notin \mathbb{N}^2$ and $\mathbf{v} \notin \mathbb{N}^2$ then in the non-trivial case both $(\alpha, 0)$ and $(0, \alpha)$ can be obtained as in (1). Applying similar reasoning as above, it is then possible to show that

$$\mathbf{r} \in \mathbf{c} + \lambda \cdot \mathbf{u} + \gamma \cdot \mathbf{v} + \text{cone}_{\mathbb{N}}(P''),$$

for some $\lambda, \gamma \leq \|P\|^{O(1)}$ and $P'' = \{(\alpha, 0), (0, \alpha)\}$, which again fulfills the requirements of the lemma. \square

Let us give an intuitive idea of how we can prove Proposition 5 (a) by an application of Lemma 6. Suppose we are given a run starting in $q(u_1, u_2)$ and ending in $q(v_1, v_2)$ such that w.l.o.g. $u_1 \leq v_1$ and $u_2 \leq v_2$. From Proposition 3 we know that the \mathbb{Z} -reachability relation can be captured by a finite set of linear path schemes. Since we start and end in the same state one can show that any (slight modification of) such

a linear path scheme describes a set of displacements equal to a linear set $L(\mathbf{b}; P)$ such that $\mathbf{b} \in P$. An application of Lemma 6 then allows us to decompose such a linear set into a semi-linear set whose period vectors all point into the same \mathbb{N}^2 direction. The crucial point is that any linear set in this semi-linear set can again be translated back into a linear path scheme of $*$ -length at most two whose displacements point to \mathbb{N}^2 . Consequently, any path obtained from such a linear path scheme does not, informally speaking, drift away too much, and if u_1 and u_2 are sufficiently large then \mathbb{N} -reachability and \mathbb{Z} -reachability coincide.

Consequently, the first step is to interpret Lemma 6 in terms of linear path schemes. As in [14], subsequently we say that a linear path scheme $\alpha_0\beta_1^*\alpha_1\cdots\beta_k^*\alpha_k$ is *zigzag-free* if $\{\delta(\beta_1), \dots, \delta(\beta_k)\} \subseteq Z$ for some quadrant Z .

Lemma 7. Let $q \in Q$. For every linear path scheme ρ from q to q , there exists a finite set R_ρ of zigzag-free linear path schemes such that

- (i) $\delta(\rho) \subseteq \delta(R_\rho)$, and
- (ii) $|\sigma| \leq (|\rho| + \|T\|)^{O(1)}$ and $|\sigma|_* \leq 2$ for each $\sigma \in R_\rho$.

Proof sketch. Let $\rho = \alpha_0\beta_1^*\alpha_1\cdots\beta_k^*\alpha_k$ be a linear path scheme. We have that $\alpha_0\cdots\alpha_k$ and β_1, \dots, β_k are cycles. The idea is to view the displacements of those cycles as linear sets, *i.e.*, define $\mathbf{b} \stackrel{\text{def}}{=} \delta(\alpha_0\cdots\alpha_k)$, $\mathbf{p}_i = \delta(\beta_i)$ for $i \in [1, k]$, and $P \stackrel{\text{def}}{=} \{\mathbf{b}, \mathbf{p}_1, \dots, \mathbf{p}_k\}$. It is not difficult to verify that

$$\delta(\rho) \subseteq L(\mathbf{b}; P) = \bigcup_{Z \text{ is a quadrant}} L(\mathbf{b}; P) \cap Z.$$

For a fixed quadrant Z , by Lemma 6 we have $\delta(\mathbf{b}; P) \cap Z \subseteq \bigcup_{i \in I} L(\mathbf{c}_i; P_i)$ such that for every $i \in I$, $|P_i| \leq 2$, $P_i \subseteq (P \cup L(\mathbf{b}; P)) \cap Z$, and $\mathbf{u} \in \mathbf{b} + \text{cone}_{[0, e]}(P)$ for some $e \leq \|P\|^{O(1)}$ for every $\mathbf{u} \in \{\mathbf{c}_i\} \cup (P_i \cap L(\mathbf{b}; P))$. The latter property allows for translating every such \mathbf{u} into a path $\pi_{\mathbf{u}} \in \rho$ such that $\mathbf{u} = \delta(\pi_{\mathbf{u}})$. Hence, for every $i \in I$ we can obtain from ρ some ρ_i such that $\delta(\rho_i) = L(\mathbf{c}_i; P_i)$, the length of ρ_i does not increase too much, and ρ_i only has the two cycles from P_i which point into Z . \square

We are now fully prepared to give a proof of Proposition 5.

Proof of Proposition 5. Let us fix a 2-VASS $V = (Q, T)$.

Proof of (a): Let S be the finite set of linear path schemes from Proposition 3 such that

- $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ if, and only if, $p(\mathbf{u}) \xrightarrow{S} q(\mathbf{v})$, and
- $|\rho| \leq 2 \cdot |Q| \cdot |T|$ and $|\rho|_* \leq |T|$ for each $\rho \in S$.

For each $\rho \in S$, let R_ρ be the set of zigzag-free linear path schemes from Lemma 7, and define $R \stackrel{\text{def}}{=} \bigcup_{\rho \in S} R_\rho$. Hence, for each $\sigma \in R$ we have $|\sigma| \leq (2 \cdot |Q| \cdot |T| + \|T\|)^{O(1)} = (|Q| + \|T\|)^{O(1)}$ by (ii) of Lemma 7. We set D required in Proposition 5 to $D \stackrel{\text{def}}{=} \max\{|\sigma| : \sigma \in R\} \cdot \|T\| \leq (|Q| + \|T\|)^{O(1)}$. The monotonicity of zigzag-free linear path schemes now provides the key ingredient for proving Proposition 5 (a). For the rest of the proof let us fix $\mathbf{u}, \mathbf{v} \in [D, \infty)^2$

and some zigzag-free linear path scheme $\sigma = \alpha_0\beta_1^*\alpha_1\beta_2^*\alpha_2 \in R$. Suppose $q(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ for some $\pi = \alpha_0\beta_1^{e_1}\alpha_1\beta_2^{e_2}\alpha_2$, then by definition of D and the fact that σ is zigzag-free it is clear that for every $i \in [0, |\pi|]$,

$$\mathbf{0} \leq \mathbf{u} + \delta(\pi[1, i]) \quad (2)$$

It remains to prove $q(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ if, and only if, $q(\mathbf{u}) \xrightarrow{\sigma} q(\mathbf{v})$ for some $\sigma \in R$, which follows from:

$$\begin{aligned} & q(\mathbf{u}) \xrightarrow{*} q(\mathbf{v}) \\ \implies & q(\mathbf{u}) \xrightarrow{*} q(\mathbf{v}) \\ \stackrel{\text{Proposition 3}}{\iff} & q(\mathbf{u}) \xrightarrow{\rho} q(\mathbf{v}) \text{ for some } \rho \in S \\ \stackrel{\text{Lemma 7 (i)}}{\implies} & q(\mathbf{u}) \xrightarrow{\sigma} q(\mathbf{v}) \text{ for some } \sigma \in R_\rho, \rho \in S \\ \stackrel{(2)}{\implies} & q(\mathbf{u}) \xrightarrow{\sigma} q(\mathbf{v}) \text{ for some } \sigma \in R \\ \implies & q(\mathbf{u}) \xrightarrow{*} q(\mathbf{v}) \end{aligned}$$

Proof of (b): Suppose that $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$. Then π can be factorized as $\pi = \alpha_0\beta_1\alpha_1\cdots\beta_k\alpha_k$ such that

$$\begin{aligned} p(\mathbf{u}) \xrightarrow{\alpha_0} q_1(\mathbf{u}_1) \xrightarrow{\beta_1} q_1(\mathbf{u}'_1) \xrightarrow{\alpha_1} q_2(\mathbf{u}_2) \\ \cdots q_k(\mathbf{u}_k) \xrightarrow{\beta_k} q_k(\mathbf{u}'_k) \xrightarrow{\alpha_k} q(\mathbf{v}) \end{aligned}$$

where $|\alpha_0|, |\alpha_1|, \dots, |\alpha_k| \leq |Q|$, each β_i is a cycle from q_i to q_i for some $q_i \in Q$, and $k \leq |Q|$. Since $\mathbf{u}_i, \mathbf{u}'_i \in \mathbb{O}$ for all $i \in [1, k]$, by Part (a) of Proposition 5 we have $q_i(\mathbf{u}_i) \xrightarrow{\rho_i} q_i(\mathbf{u}'_i)$ for some linear path scheme $\rho_i \in R$. Consequently, we define X as

$$\begin{aligned} X \stackrel{\text{def}}{=} \{ \text{linear path scheme } \alpha_0\rho_1\alpha_1\cdots\rho_k\alpha_k : \\ k \leq |Q|, \alpha_i \in T^*, |\alpha_i| \leq |Q|, \rho_i \in R \}. \end{aligned}$$

Let $\rho \in X$, then we have $|\rho| \leq |Q|^2 + |Q| \cdot (|Q| + \|T\|)^{O(1)} = (|Q| + \|T\|)^{O(1)}$, and $|\rho|_* \leq 2 \cdot |Q|$. \square

C. Reachability in 2-VASS with One Bounded Component

The purpose of this section is to establish the following result on reachability within L-shaped bands, as illustrated in the right-most picture of Fig. 4.

Proposition 8. Let $V = (Q, T)$ be a 2-VASS, $D \in \mathbb{N}$ and $\mathbb{L} = ([0, D] \times \mathbb{N}) \cup (\mathbb{N} \times [0, D])$. There exists a finite set $Y_{\mathbb{L}}$ of linear path schemes such that

- (i) $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$ implies $p(\mathbf{u}) \xrightarrow{Y_{\mathbb{L}}} q(\mathbf{v})$, and
- (ii) $|\rho| \leq (|Q| + \|T\| + D)^{O(1)}$ and $|\rho|_* \leq 2$ for every $\rho \in Y_{\mathbb{L}}$.

We briefly sketch the proof of Proposition 8 here. In its essence, restricting the set of admissible values of one of the two counters of a 2-VASS to $[0, D]$ gives rise to a 1-VASS. This observation enables us to resort to techniques and results developed for 1-VASS. In particular, to the following lemma established by Valiant and Paterson.

Lemma 9 (Lemma 2 in [26]). Let $V = (Q, T)$ be a 1-VASS such that $T \subseteq Q \times \{-1, 0, 1\} \times Q$ and let $p(\mathbf{u}) \xrightarrow{*} q(\mathbf{v})$

for some configurations $p(u)$ and $q(v)$ such that $|u - v| \geq |Q| + |Q|^2$. There exist $\alpha, \beta, \gamma \in T^*$ and $\pi \in T^*$ such that $p(u) \xrightarrow{\pi} p(v)$ and π has the the following properties,

- (i) $\pi = \alpha\beta^i\gamma$ for some $i > 0$,
- (ii) $\alpha\beta^*\gamma$ is a linear path scheme of $*$ -length one, and
- (iii) $|\alpha\gamma| < |Q|^2$ and β is a cycle with $|\beta| \leq |Q|$ and $|\delta(\beta)| \in [1, |Q|]$.

As a consequence of Lemma 9, one can show that the reachability relation of 1-VASS can be captured by linear path schemes with the following properties.

Lemma 10. Let $V = (Q, T)$ be a 1-VASS. There exists a finite set Y of linear path schemes such that

- (i) $p(u) \xrightarrow{*} q(v)$ if, and only if, $p(u) \xrightarrow{Y} q(v)$, and
- (ii) $|\rho| \leq (|Q| + \|T\|)^{O(1)}$ and $|\rho|_* \leq 1$ for each $\rho \in Y$.

This, in turn, allows us to prove Proposition 8.

Proof sketch of Proposition 8. Let $V = (Q, T)$ be a 2-VASS and $D \in \mathbb{N}$. Let $\mathbb{B}_1 = \mathbb{N} \times [0, D]$, $\mathbb{B}_2 = [0, D] \times \mathbb{N}$ and $\mathbb{L} = \mathbb{B}_1 \cup \mathbb{B}_2$. Consider a run $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ of minimal length. By making \mathbb{L} slightly larger, we can decompose π as a sequence of runs that remain within either the vertical band, or the horizontal band. Formally, let $\mathbb{B}'_1 = \mathbb{N} \times [0, D + \|T\|]$, $\mathbb{B}'_2 = [0, D + \|T\|] \times \mathbb{N}$, $\mathbb{L}' = \mathbb{B}'_1 \cup \mathbb{B}'_2$ and $H = \mathbb{B}'_1 \cap \mathbb{B}'_2$. Due to minimality of $|\pi|$ we can factorize π as $\pi = \pi_1 \cdots \pi_k$, where $p_0(\mathbf{u}_0) \xrightarrow{\pi_1} p_1(\mathbf{u}_1) \cdots \xrightarrow{\pi_k} p_k(\mathbf{u}_k)$ and

- $p_0 = p, p_k = q, \mathbf{u}_0 = \mathbf{u}, \mathbf{u}_k = \mathbf{v}$,
- $\mathbb{C}_i \in \{\mathbb{B}'_1, \mathbb{B}'_2\}$ for each $i \in [1, k]$
- $\mathbf{u}_i \in H$ for each $i \in [1, k-1]$, and
- $k \leq |H| = (D + \|T\| + 1)^2$.

Note that a 2-VASS in which the value of at least one counter is bounded by some $E \in \mathbb{N}$ can be simulated by a 1-VASS with $|Q| \cdot (E+1)$ states. This allows us, by applying Lemma 10, to replace $\pi_0, \pi_1, \dots, \pi_k$ by some linear path schemes $\rho_0, \rho_1, \dots, \rho_k$. However, this would yield a linear path scheme of $*$ -length k . In fact we can restrict only ρ_0 and ρ_k to have $*$ -length at most one, where ρ_1, \dots, ρ_k can be chosen to have $*$ -length zero. Indeed, since $\pi_1, \pi_2, \dots, \pi_{k-1}$ are runs from H to H , which is finite, and $\rho_1, \rho_2, \dots, \rho_{k-1}$ only have one cycle, then $\pi_1, \pi_2, \dots, \pi_{k-1}$ can each be replaced with runs of length at most $(|Q| + \|T\| + D)^{O(1)}$. Therefore, $p(\mathbf{u}) \xrightarrow{\rho_0 \sigma \rho_k} q(\mathbf{v})$ where $\sigma \in T^*$ and $|\sigma| \leq (|Q| + \|T\| + D)^{O(1)}$. \square

D. Factorizing arbitrary runs: Proof of Theorem 1

By application of the results established in Sections IV-B and IV-C, we will now prove Theorem 1. More precisely, we will show that any run can be factorized into few runs of types (1), (2) or (3). To this end, let us fix a 2-VASS $V = (Q, T)$. Let $D \leq (|Q| + \|T\|)^{O(1)}$ be the constant from Proposition 5. Informally speaking, we have hereby defined that “sufficiently large” means to be greater or equal to D . Moreover we set $\mathbb{L} \stackrel{\text{def}}{=} ([0, D + \|T\|] \times \mathbb{N}) \cup (\mathbb{N} \times [0, D + \|T\|])$, $\mathbb{O} \stackrel{\text{def}}{=} [D, \infty)^2$, and $\mathbb{B} \stackrel{\text{def}}{=} \mathbb{L} \cap \mathbb{O} = ([D, D + \|T\|] \times \mathbb{N}) \cup (\mathbb{N} \times [D, D + \|T\|])$. Again, informally speaking, we have hereby defined that “small” means to be less or equal to $D + \|T\|$.

Let us summarize what we have proven in Sections IV-B and IV-C:

- Runs of type (1) can be captured by a set of linear path schemes R , where each $\rho \in R$ has $*$ -length at most two and length at most $(|Q| + \|T\|)^{O(1)}$ by Proposition 5 (a).
- Runs of type (2) can be captured by a set of linear path schemes X , where each $\rho \in X$ has $*$ -length at most $2 \cdot |Q|$ and length at most $(|Q| + \|T\|)^{O(1)}$ by Proposition 5 (b).
- Runs of type (3) can be captured by a set of linear path schemes $Y_{\mathbb{L}}$, where each $\rho \in Y_{\mathbb{L}}$ has $*$ -length at most two and length at most $(|Q| + \|T\| + D)^{O(1)} = (|Q| + \|T\|)^{O(1)}$ by Proposition 8.

Given $p(\mathbf{u})$ and $q(\mathbf{v})$, let us fix an arbitrary run $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$, where $\pi = t_1 \cdots t_k \in T^k$ and

$$p(\mathbf{u}) = q_0(\mathbf{u}_0) \xrightarrow{t_1} q_1(\mathbf{u}_1) \cdots \xrightarrow{t_k} q_k(\mathbf{u}_k) = q(\mathbf{v}) .$$

We will be interested in the indices of configurations whose counter values lie in \mathbb{B} and define $I \stackrel{\text{def}}{=} \{i \in [0, k] : \mathbf{u}_i \in \mathbb{B}\}$. Let us define the function $x : I \rightarrow I$ that maps each index $i \in I$ to the smallest element in I larger than i (and i if $i = \max I$), i.e.

$$x(i) \stackrel{\text{def}}{=} \begin{cases} \min\{j \in I : j > i\} & \text{if } i < \max I, \\ i & \text{otherwise .} \end{cases}$$

We also define the function $\ell : \{q_i \in Q : i \in I\} \rightarrow I$ that maps each state q that appears in a configuration in $Q \times \mathbb{B}$ to the largest index in I where it appears, i.e. $\ell(q) \stackrel{\text{def}}{=} \max\{i \in I : q = q_i\}$. We are now interested in factorizing the run $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ into runs between configurations that start and end in $\mathbb{B} = \mathbb{L} \cap \mathbb{O}$. More precisely, by the choice of \mathbb{O} , \mathbb{L} and \mathbb{B} and by the pigeonhole principle there exist indices $i_1, \dots, i_h \in I$ such that the run $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ can be factorized as (cf. Fig. 5):

$$\begin{aligned} q_0(\mathbf{u}_0) &\xrightarrow{\pi_{0,1}}_{\mathbb{D}_0} q_{i_1}(\mathbf{u}_{i_1}) \xrightarrow{\pi_1}_{\mathbb{N}^2} q_{\ell(q_{i_1})}(\mathbf{u}_{\ell(q_{i_1})}) \\ &\xrightarrow{\pi_{1,2}}_{\mathbb{D}_1} q_{i_2}(\mathbf{u}_{i_2}) \xrightarrow{\pi_2}_{\mathbb{N}^2} q_{\ell(q_{i_2})}(\mathbf{u}_{\ell(q_{i_2})}) \\ &\cdots \xrightarrow{\pi_{h-1,h}}_{\mathbb{D}_{h-1}} q_{i_h}(\mathbf{u}_{i_h}) \xrightarrow{\pi_h}_{\mathbb{N}^2} q_{\ell(q_{i_h})}(\mathbf{u}_{\ell(q_{i_h})}) \\ &\xrightarrow{\pi_{h,h+1}}_{\mathbb{D}_h} q_k(\mathbf{u}_k) , \end{aligned}$$

where

- (a) $h \leq |Q|$,
- (b) $i_t \in I$ and thus we have $\mathbf{u}_{i_t} \in \mathbb{B}$ and $q_{i_t} = q_{\ell(q_{i_t})}$ for each $t \in [1, h]$,
- (c) $\mathbb{D}_t \in \{\mathbb{O}, \mathbb{L}\}$ for each $t \in [0, h]$, and
- (d) $i_{t+1} = x(\ell(q_{i_t}))$ for each $t \in [1, h-1]$.

By (b) each run of the form $q_{i_t}(\mathbf{u}_{i_t}) \xrightarrow{\pi_t}_{\mathbb{N}^2} q_{\ell(q_{i_t})}(\mathbf{u}_{\ell(q_{i_t})})$ is a run of type (1) and can hence be replaced by some linear path scheme from R (recall that $\mathbb{B} \subseteq \mathbb{O}$). By (c) and (d), each run of the form $\xrightarrow{\pi_{t,t+1}}_{\mathbb{D}_t}$ is a run of type (2) or of type (3) and can hence be replaced by some linear path scheme from $X \cup Y_{\mathbb{L}}$. In summary, the run $p(\mathbf{u}) \xrightarrow{\pi}_{\mathbb{N}^2} q(\mathbf{v})$ can be replaced by a linear path scheme of $*$ -length at most $(h+1) \cdot 2 \cdot |Q| \leq O(|Q|^2)$ and length at most $(h+1) \cdot (|Q| + \|T\|)^{O(1)} = (|Q| + \|T\|)^{O(1)}$. This concludes the proof of Theorem 1.

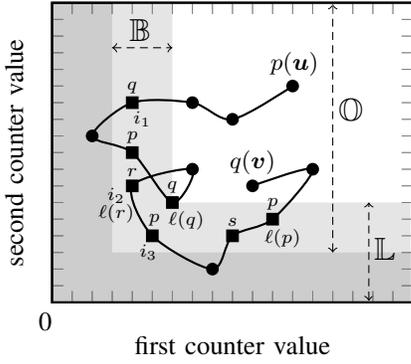


Fig. 5. Example of the decomposition of a path in the proof of Theorem 1. The region depicted is the positive quadrant in the Cartesian plane. Here, $I = \{3, 5, 6, 8, 9, 11, 12\}$ is marked with squares, and $i_1 = 3$, $\ell(q) = 6$, $i_2 = \ell(r) = 8$, $i_3 = 9$ and $\ell(p) = 12$.

V. COMPLEXITY RESULTS

Having established Theorem 1, it is now not difficult to show that reachability in 2-VASS is in PSPACE by application of bounds from integer linear programming. A complementary lower bound follows via a reduction from reachability in bounded one-counter automata, which is known to be PSPACE-complete [3]. This is the subject of Section V-A below which proves Theorem 2. The PSPACE lower bound does, however, crucially depend on binary encoding of numbers. In fact, we show in Section V-B that reachability in unary 2-VASS is in NP and NL-hard. The precise complexity of this problem is left as an open problem by this paper. Finally, for the sake of completeness, in Section V-C we briefly state some corollaries of our results on the complexity of reachability in \mathbb{Z} -VASS, and on coverability and boundedness in 2-VASS.

Before we begin, let us recall some definitions and results from integer linear programming. Let A be a $d \times k$ integer matrix and $c \in \mathbb{Z}^d$. A *system of linear Diophantine inequalities* (resp. a *system of linear Diophantine equations*) is given as $\mathcal{I} : A \cdot x \geq c$ (resp. as $\mathcal{E} : A \cdot x = c$) and we say that \mathcal{I} (resp. \mathcal{E}) is *feasible* if there exists some $e \in \mathbb{N}^k$ such that $A \cdot e \geq c$ (resp. $A \cdot e = c$), i.e., every inequality (resp. equality) holds in every row of \mathcal{I} (resp. \mathcal{E}). Subsequently, we refer to e as a *solution* of \mathcal{I} or \mathcal{E} , respectively. By $[\mathcal{I}] \subseteq \mathbb{N}^k$ we denote the set of all solutions of \mathcal{I} , the set of solutions $[\mathcal{E}] \subseteq \mathbb{N}^k$ is defined analogously.

Let us now recall two bounds on solutions of systems of linear Diophantine inequalities and equations that we subsequently rely upon. The first bound we use in this paper concerns systems of linear Diophantine inequalities.

Proposition 11 ([24], p. 239). Let $\mathcal{I} : A \cdot x \geq c$ be a feasible system of linear Diophantine inequalities, where A is a $d \times k$ matrix. Then there exists a solution $e \in \mathbb{N}^k$ of \mathcal{I} such that $\|e\| \leq 2^{k^{O(1)}} \cdot O(\|A\| + \|c\|)$.

Next, we consider a bound for feasible homogeneous systems of linear Diophantine equations.

Proposition 12 ([20], Theorem 1). Let $\mathcal{E} : A \cdot x = 0$ be a system of linear Diophantine equations, where A is a $d \times k$ integer matrix. Then there exists $P \subseteq \mathbb{N}^k$ such that $\|P\| \leq (\|A\| + 1)^d$ and $[\mathcal{E}] = \text{cone}_{\mathbb{N}}(P)$.

The previous proposition is easily generalized to the non-homogeneous case.

Corollary 13. Let $\mathcal{E} : A \cdot x = c$ be a feasible system of linear Diophantine equations such that A is a $d \times k$ matrix. Then there exists a solution $e \in \mathbb{N}^k$ of \mathcal{E} such that $\|e\| \leq (\|A\| + \|c\|)^{O(d)}$.

A. Reachability in 2-VASS is PSPACE-complete

In this section, we prove Theorem 2 and show that reachability in 2-VASS is PSPACE-complete. Given an instance $p(u) \xrightarrow{*}_{\mathbb{N}^2} q(v)$ of reachability, by Theorem 1 we have that $p(u) \xrightarrow{\rho}_{\mathbb{N}^2} q(v)$ for some linear path scheme ρ such that $|\rho| \leq (|Q| + \|T\|)^{O(1)}$ and $|\rho|_* \leq O(|Q|^2)$. Let $\rho = \alpha_0 \beta_1^* \alpha_1 \cdots \beta_k^* \alpha_k$. Then we have $p(u) \xrightarrow{\rho}_{\mathbb{N}^2} q(v)$ if, and only if,

$$\exists e_1, \dots, e_k \in \mathbb{N} \text{ s.t. } p(u) \xrightarrow{\alpha_0 \beta_1^{e_1} \alpha_1 \cdots \beta_k^{e_k} \alpha_k}_{\mathbb{N}^2} q(v). \quad (3)$$

Consequently, obtaining a PSPACE upper bound for reachability reduces to bounding the binary representation of the e_i polynomially in the sizes of V , u and v .

Our approach is straightforward: we rephrase the existential question from (3) in terms of finding solutions to a system of linear Diophantine inequalities and then apply the aforementioned bounds from integer linear programming in order to bound the e_i . For technical convenience we distinguish for each linear path scheme and every cycle of the linear path scheme whether the cycle is taken at least once or not at all. To this end, let us define the function $\text{sign} : \mathbb{N} \rightarrow \{0, 1\}$ as $\text{sign}(n) = 1$ if $n \geq 1$ and $\text{sign}(n) = 0$ if $n = 0$. Our approach is formalized by the following lemma.

Lemma 14. Let $V = (Q, T)$ be a d -VASS, $u \in \mathbb{N}^d$ and $\rho = \alpha_0 \beta_1^* \alpha_1 \cdots \beta_k^* \alpha_k$ be a linear path scheme from p to q and let $\chi : [1, k] \rightarrow \{0, 1\}$. Then there exists a system of linear Diophantine inequalities $\mathcal{I} = \mathcal{I}(u, \rho, \chi)$ of the form $\mathcal{I} : A \cdot x \geq c$ such that

- $e \in [\mathcal{I}]$ if, and only if, $\pi = \alpha_0 \beta_1^{e_1} \alpha_1 \cdots \beta_k^{e_k} \alpha_k$, $p(u) \xrightarrow{\pi}_{\mathbb{N}^d} q(u + \delta(\pi))$ and $\chi(i) = \text{sign}(e_i)$ for every $e = (e_1, \dots, e_k) \in \mathbb{N}^k$,
- A is a $((d+1) \cdot k) \times k$ -matrix, and
- $\|A\| \leq k \cdot |\rho| \cdot \|T\|$ and $\|c\| \leq O(\|u\| + |\rho| \cdot \|T\|)$.

Proof. We only prove the lemma for the concrete function $\chi : [1, k] \rightarrow \{0, 1\}$, where $\chi(i) = 1$ for all $i \in [1, k]$. In the following, we write $x = (x_1, \dots, x_k)$. First, we assert that the solutions e_i are greater or equal to 1, i.e.,

$$I_k \cdot x \geq \mathbf{1}, \quad (4)$$

where I_k is the $k \times k$ unit matrix and $\mathbf{1} = (1, \dots, 1)$. Next, informally speaking, we have to construct \mathcal{I} in a way such that we assert that the counter value does not drop below zero on

any infix of ρ in any dimension. For segments of ρ between cycles, this can be ensured by the following constraints for every $j \in [0, k]$ and $\ell \in [1, |\alpha_j|]$, which simply enforce the accumulated counter value to be non-negative:

$$\mathbf{u} + \sum_{0 \leq i < j} (\delta(\alpha_i) + \delta(\beta_{i+1}) \cdot x_{i+1}) + \delta(\alpha_j[1, \ell]) \geq \mathbf{0}$$

if, and only if,

$$\sum_{1 \leq i \leq j} \delta(\beta_i) \cdot x_i \geq -\mathbf{u} - \sum_{0 \leq i < j} \delta(\alpha_i) - \delta(\alpha_j[1, \ell]) \quad (5)$$

For counter values which, informally speaking, occur along cycles β_j of ρ , it is sufficient to only check whether their initial and final segments lead to counter values greater or equal to zero. Formally, we assert the following constraints for every $j \in [1, k]$ and $\ell \in [1, |\beta_j|]$:

$$\begin{aligned} \mathbf{u} + \delta(\alpha_0) + \sum_{1 \leq i < j} (\delta(\beta_i) \cdot x_i + \delta(\alpha_i)) + \delta(\beta_j[1, \ell]) &\geq \mathbf{0} \\ \mathbf{u} + \delta(\alpha_0) + \sum_{1 \leq i < j} (\delta(\beta_i) \cdot x_i + \delta(\alpha_i)) + \\ &\delta(\beta_j) \cdot (x_j - 1) + \delta(\beta_j[1, \ell]) \geq \mathbf{0} \end{aligned}$$

if, and only if,

$$\begin{aligned} \sum_{1 \leq i \leq j-1} \delta(\beta_i) \cdot x_i &\geq -\mathbf{u} - \sum_{0 \leq i < j} \delta(\alpha_i) - \delta(\beta_j[1, \ell]) \quad (6) \\ \sum_{1 \leq i \leq j} \delta(\beta_i) \cdot x_i &\geq -\mathbf{u} - \sum_{0 \leq i < j} \delta(\alpha_i) + \delta(\beta_j) - \\ &\delta(\beta_j[1, \ell]) \quad (7) \end{aligned}$$

By our construction, it is easily verified that for every $\mathbf{e} = (e_1, \dots, e_k) \in \mathbb{N}^k$ we have $\chi(i) = 1$ for all $i \in [1, k]$ and $p(\mathbf{u}) \xrightarrow{\alpha_0 \beta_1^{e_1} \alpha_1 \dots \beta_k^{e_k} \alpha_k} q(\mathbf{u} + \delta(\pi))$ if, and only if, \mathbf{e} fulfills all constraints defined in (4), (5), (6) and (7). It thus remains to, informally speaking, extract the required system \mathcal{I} of linear Diophantine inequalities from those constraints.

For every fixed $j \in [1, k]$, by combining the constraints from (5), (6) and (7), we obtain systems of linear Diophantine inequalities $\mathcal{I}'_j : B_j \cdot \mathbf{x} \geq \mathbf{d}_j$ such that B_j consists of at most d different rows, since every x_i is multiplied by the same $\delta(\beta_i)$. Let A_j be the following $(d \times k)$ -matrix: $A_j \stackrel{\text{def}}{=} [\delta(\beta_1) \dots \delta(\beta_j) \mathbf{0} \dots \mathbf{0}]$. For the i -th row of A_j , let $c_{j,i} \in \mathbb{Z}$ be the maximum value in \mathbf{d}_j of the rows with the same coefficients in \mathcal{I}'_j . We define $\mathbf{c}_j \stackrel{\text{def}}{=} (c_{j,1}, \dots, c_{j,d})$ and set $\mathcal{I}_j : A_j \cdot \mathbf{x} \geq \mathbf{c}_j$. By construction, we now have that $\mathbf{e} \in \mathbb{N}^k$ is a solution of \mathcal{I}_j if, and only if, \mathbf{e} is a solution to \mathcal{I}'_j and in particular fulfills all relevant constraints in (5), (6) and (7).

In order to obtain the matrix A and \mathbf{c} required in the lemma, we define

$$A \stackrel{\text{def}}{=} \begin{bmatrix} I_k \\ A_1 \\ \vdots \\ A_k \end{bmatrix} \quad \text{and} \quad \mathbf{c} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{1} \\ \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_k \end{pmatrix}.$$

The dimension of A and \mathbf{c} is as required. It thus remains to estimate the norm of A and \mathbf{c} . We have $\|A\| \leq \sum_{1 \leq i \leq k} \|\delta(\beta_i)\| \leq k \cdot |\rho| \cdot \|T\|$. For \mathbf{c} , the following inequality bounds the norm of the right-hand sides of (5), (6) and (7): $\|\mathbf{c}\| \leq \|\mathbf{u}\| + 2 \cdot |\rho| \cdot \|T\|$. \square

By application of Proposition 11, this lemma now enables us to give bounds on the length of a run witnessing reachability for two given configurations.

Lemma 15. Let $V = (Q, T)$ be a d -VASS, let $p(\mathbf{u})$ and $q(\mathbf{v})$ be configurations of V , and let $\rho = \alpha_0 \beta_1^* \alpha_1 \dots \beta_k^* \alpha_k$ be a linear path scheme from p to q . Then $p(\mathbf{u}) \xrightarrow{\rho} q(\mathbf{v})$ if, and only if, $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ for some $\pi = \alpha_0 \beta_1^{e_1} \alpha_1 \dots \beta_k^{e_k} \alpha_k$ such that $e_i \leq 2^{k^{O(1)}} \cdot O(\|\mathbf{u}\| + \|\mathbf{v}\| + |\rho| \cdot \|T\|)$ for each $i \in [1, k]$.

Proof. The “if”-direction is trivial. For the “only-if”-direction assume $p(\mathbf{u}) \xrightarrow{\rho} q(\mathbf{v})$. Then $p(\mathbf{u}) \xrightarrow{\psi} q(\mathbf{v})$, where $\psi = \alpha_0 \beta_1^{f_1} \alpha_1 \dots \beta_k^{f_k} \alpha_k$ for some $f_1, \dots, f_k \in \mathbb{N}$. Let $\chi : [1, k] \rightarrow \{0, 1\}$ be defined as $\chi(i) \stackrel{\text{def}}{=} \text{sign}(f_i)$ for each $i \in [1, k]$.

The set of those $e_1, \dots, e_k \in \mathbb{N}$ that achieve $\mathbf{u} + \delta(\pi) = \mathbf{v}$ can be obtained from the set of solutions of the system $\mathcal{E} : B \cdot \mathbf{x} = \mathbf{d}$ of linear Diophantine inequalities with unknowns $\mathbf{x} = (x_1, \dots, x_k)$, where $B \stackrel{\text{def}}{=} [\delta(\beta_1) \dots \delta(\beta_k)]$ and $\mathbf{d} \stackrel{\text{def}}{=} \mathbf{v} - \mathbf{u} - \sum_{0 \leq i \leq k} \delta(\alpha_i)$. The constraint matrix B is of dimension $d \times k$ and has norm bounded by $|\rho| \cdot \|T\|$. The norm of \mathbf{d} is bounded by $\|\mathbf{u}\| + \|\mathbf{v}\| + |\rho| \cdot \|T\|$.

Lemma 14 yields a system of linear Diophantine inequalities $\mathcal{I} = \mathcal{I}(\mathbf{u}, \rho, \chi)$ of the form $\mathcal{I} : A \cdot \mathbf{x} \geq \mathbf{c}$ whose set of solutions $\mathbf{e} = (e_1, \dots, e_k) \in \mathbb{N}$ corresponds to all runs $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{u} + \delta(\pi))$, where $\pi = \alpha_0 \beta_1^{e_1} \alpha_1 \dots \beta_k^{e_k} \alpha_k$ and $\chi(i) = \text{sign}(e_i)$ for all $i \in [1, k]$. The constraint matrix A is of dimension $((d+1) \cdot k) \times k$ and has norm at most $k \cdot |\rho| \cdot \|T\|$. The norm of \mathbf{c} is bounded by $O(\|\mathbf{u}\| + \|\mathbf{v}\| + |\rho| \cdot \|T\|)$.

Consequently, for any $(e_1, \dots, e_k) \in \llbracket \mathcal{I} \rrbracket \cap \llbracket \mathcal{E} \rrbracket$ and $\pi = \alpha_0 \beta_1^{e_1} \alpha_1 \dots \beta_k^{e_k} \alpha_k$, we have $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ and $\chi(i) = \text{sign}(e_i)$ for all $i \in [1, k]$. Now we obtain $\mathcal{I} \cap \mathcal{E}$ as

$$\mathcal{I} \cap \mathcal{E} : \begin{bmatrix} A \\ B \\ -B \end{bmatrix} \cdot \mathbf{x} \geq \begin{pmatrix} \mathbf{c} \\ \mathbf{d} \\ -\mathbf{d} \end{pmatrix}.$$

We conclude that the constraint matrix of $\mathcal{I} \cap \mathcal{E}$ is of dimension $((d+1) \cdot k + 2d) \times k$ and has a norm that is bounded by $k \cdot |\rho| \cdot \|T\|$. Moreover, the norm on right-hand side of $\mathcal{I} \cap \mathcal{E}$ is bounded by $O(\|\mathbf{u}\| + \|\mathbf{v}\| + |\rho| \cdot \|T\|)$. By application of Proposition 11, the bounds on the solutions of $\mathcal{I} \cap \mathcal{E}$ follow. \square

Corollary 16. Reachability in 2-VASS is in PSPACE.

Proof. Let $V = (Q, T)$ be a 2-VASS and $p(\mathbf{u}), q(\mathbf{v})$ be configurations of V . Let S be the set of linear path schemes obtained from Theorem 1. By Lemma 15, if $p(\mathbf{u}) \xrightarrow{\rho} q(\mathbf{v})$ for some $\rho = \alpha_0 \beta_1^* \alpha_1 \dots \beta_k^* \alpha_k \in S$ then $p(\mathbf{u}) \xrightarrow{\pi} q(\mathbf{v})$ for

some $\pi = \alpha_0 \beta_1^{e_1} \alpha_1 \cdots \beta_k^{e_k} \alpha_k$ where $e_1, \dots, e_k \in [0, e]$, for some $e \in \mathbb{N}$ that is bounded by

$$\begin{aligned} e &\leq 2^{|Q|^{O(1)}} \cdot O\left(\|\mathbf{u}\| + \|\mathbf{v}\| + (|Q| + \|T\|)^{O(1)} \cdot \|T\|\right) \\ &\leq 2^{(|V| + \log\|\mathbf{u}\| + \log\|\mathbf{v}\|)^{O(1)}}. \end{aligned}$$

Since $|\pi| \leq |\rho| \cdot e$, the run $p(\mathbf{u}) \xrightarrow{\pi}_{\mathbb{N}^2} q(\mathbf{v})$ can be guessed nondeterministically in polynomial space by storing only the intermediate configurations in an on-the-fly manner. Consequently, reachability in 2-VASS in PSPACE. \square

In order to complete the proof of Theorem 2, it remains to show hardness for PSPACE. We reduce from reachability in bounded one-counter automata, which is known to be PSPACE-complete [3]. A bounded one-counter automaton is given by a tuple $V = (Q, T, b)$, where (Q, T) is a 1-VASS and $b \in \mathbb{N}$ is a bound encoded in binary. Let $\mathbb{B} = [0, b]$, given configurations $p(u), q(v)$ of V such that $u, v \in \mathbb{B}$, *reachability* is to decide whether $p(u) \xrightarrow{*}_{\mathbb{B}} q(v)$.

Lemma 17. Reachability in 2-VASS is PSPACE-hard.

Proof. Let $V = (Q, T, b)$ be a bounded one-counter automaton, and let $V' \stackrel{\text{def}}{=} (Q, T')$ be the 2-VASS obtained from V by setting $T' \stackrel{\text{def}}{=} \{h(t) : t \in T\}$, where $h(p, z, q) \stackrel{\text{def}}{=} (p, (z, -z), q)$. We define an injection φ from configurations of V to configurations of V' as $\varphi(q(z)) \stackrel{\text{def}}{=} q(z, b - z)$. It is now easily verified that $p(u) \xrightarrow{*}_{\mathbb{B}} q(v)$ if, and only if, $\varphi(p(u)) \xrightarrow{*}_{\mathbb{N}^2} \varphi(q(v))$. \square

B. Reachability in 2-VASS with Unary Updates

For unary 2-VASS we can show that reachability is in NP and NL-hard.

Given a unary 2-VASS V , whenever $p(\mathbf{u}) \xrightarrow{*}_{\mathbb{N}^2} q(\mathbf{v})$ then by Theorem 1 there exists a linear path scheme $\rho = \alpha_0 \beta_1^* \alpha_1 \cdots \beta_k^* \alpha_k$ whose length is *polynomial* in $|V|$ such that $p(\mathbf{u}) \xrightarrow{\rho}_{\mathbb{N}^2} q(\mathbf{v})$. Moreover, the proof of Corollary 16 shows that there exist $e_1, \dots, e_k \leq 2^{(|V| + \log\|\mathbf{u}\| + \log\|\mathbf{v}\|)^{O(1)}}$ such that for $\pi = \alpha_0 \beta_1^{e_1} \alpha_1 \cdots \beta_k^{e_k} \alpha_k$, we have $p(\mathbf{u}) \xrightarrow{\pi}_{\mathbb{N}^2} q(\mathbf{v})$. In particular, every e_i can be represented using a polynomial number of bits. Hence, (ρ, e_1, \dots, e_k) may serve as a certificate that can be guessed in polynomial time. It remains to show that this certificate can be verified in polynomial time. Checking that ρ is a linear path scheme is easily verified in polynomial time. In order to check if $p(\mathbf{u}) \xrightarrow{\pi}_{\mathbb{N}^2} q(\mathbf{v})$ in polynomial time we can construct the system of linear Diophantine equations from Lemma 14 and verify that $e = (e_1, \dots, e_k)$ is a solution to this system. This shows that reachability in unary 2-VASS is in NP.

NL-hardness of reachability trivially follows from NL-hardness of reachability in directed graphs. Here, we wish to slightly strengthen this result and remark that reachability is NL-hard already for unary 2-VASS, whose underlying graph corresponds structurally to a linear path scheme (formally, every vertex lies on at most one cycle and the deletion of all cycles yields a union of isolated vertices and a cycle-free path, cf. Fig. 1 at the beginning of this paper). Let $G = (U, E)$

be a directed graph such that $U = \{u_0, \dots, u_{m-1}\}$ and $E = \{e_0, \dots, e_{m-1}\} \subseteq U \times U$. We define an injection $h : U \rightarrow [0, m-1]^2$ as $h(u_i) = (i, m-1-i)$ that relates vertices of G with vectors from bounded intervals. Let $\ell \stackrel{\text{def}}{=} m \cdot n - 1$. The flat unary 2-VASS $V = (Q, T)$ can now be defined as $Q \stackrel{\text{def}}{=} \{q_0, q'_0, \dots, q_\ell, q'_\ell\}$ and

$$\begin{aligned} T &\stackrel{\text{def}}{=} \{(q_j, \mathbf{0}, q_{j+1}) : j \in [0, \ell-1]\} \\ &\cup \{(q_j, -h(u_a), q'_j), \\ &\quad (q'_j, h(u_b), q_j) : e_{j \bmod n} = (u_a, u_b), j \in [0, \ell]\}. \end{aligned}$$

Suppose we wish to decide whether u_{m-1} is reachable from u_0 , we claim that this is the case if, and only if, $q_0(h(u_0)) \xrightarrow{*}_{\mathbb{N}^2} q_\ell(h(u_{m-1}))$. Informally speaking, the vertex currently visited along a path is encoded in the counter values of V . Every loop between q_j and q'_j allows for simulating the transition $e_{j \bmod n} = (u_a, u_b)$ of G . The transition from q_j to q'_j can only be traversed if the vertex encoded into the current counter values corresponds to u_a . If we are able to reach q'_j , the transition back to q_j then updates the currently visited vertex to u_b . Since a path from u_0 to u_{m-1} of minimal length in G traverses at most m vertices, $\ell + 1 = m \cdot n$ states q_j suffice.

Theorem 18. Reachability in unary 2-VASS is in NP and NL-hard.

C. Derived Results

Here, we explicitly remark some results that can additionally be derived from the technical results of this paper.

1) *\mathbb{Z} -Reachability in Unary d -VASS is NL-complete:* The complexity of \mathbb{Z} -reachability in d -VASS depends on the encoding of numbers as well as the dimension d . When numbers are encoded in binary, reachability is NP-complete even when $d = 1$ [5], [7], and reachability is also NP-complete when numbers are encoded in unary and d is part of the input to the problem [7]. We solve the case of reachability under unary encoding of numbers for each fixed dimension d .

Theorem 19. \mathbb{Z} -reachability in unary d -VASS is NL-complete for any fixed $d \geq 1$.

Proof. NL-hardness trivially follows from NL-hardness of reachability in directed graphs. Let $d \geq 1$ be fixed and $V = (Q, T)$ be a unary d -VASS. Suppose $p(\mathbf{u}) \xrightarrow{*}_{\mathbb{Z}^d} q(\mathbf{v})$, then by Proposition 3, there exists a linear path scheme $\rho = \alpha_0 \beta_1^* \alpha_1 \cdots \beta_k^* \alpha_k \in S$ with $k \leq |T|$ and $|\rho| \leq 2 \cdot |Q| \cdot |T|$ such that $p(\mathbf{u}) \xrightarrow{\rho}_{\mathbb{Z}^d} q(\mathbf{v})$.

Let $\mathcal{E} : A \cdot \mathbf{x} = \mathbf{c}$ be the system of linear Diophantine equations such that $A \stackrel{\text{def}}{=} [\delta(\beta_1) \cdots \delta(\beta_k)]$ and $\mathbf{c} \stackrel{\text{def}}{=} \mathbf{v} - (\mathbf{u} + \delta(\alpha_0 \alpha_1 \cdots \alpha_k))$. Then, $p(\mathbf{u}) \xrightarrow{\rho}_{\mathbb{Z}^d} q(\mathbf{v})$ if, and only if, $(e_1, e_2, \dots, e_k) \in \llbracket \mathcal{E} \rrbracket$. By Corollary 13, if $\llbracket \mathcal{E} \rrbracket \neq \emptyset$ then \mathcal{E} has a solution e such that $\|e\| \leq ((|T| + \|T\|)^{O(1)} + \|\mathbf{u}\| + \|\mathbf{v}\|)^{O(d)}$. Since $\|T\|, \|\mathbf{u}\|$ and $\|\mathbf{v}\|$ are encoded in unary and d is fixed, minimal runs are bounded by some $b \leq |V|^{O(1)}$. Thus, reachability can be

decided by guessing a path of polynomial length on-the-fly in logarithmic space. \square

2) *Boundedness and Coverability in d -VASS*: For the sake of completeness, here we wish to discuss some consequences of PSPACE-hardness of reachability in 2-VASS to the complexity of coverability and boundedness in d -VASS that were left open in the literature. The boundedness problem is to determine, given $p(\mathbf{u})$, whether $\{q(\mathbf{v}) : p(\mathbf{u}) \xrightarrow{*}_{\mathbb{N}^d} q(\mathbf{v})\}$ is infinite. The coverability problem is to determine, given $p(\mathbf{u})$ and $q(\mathbf{v})$, whether there exists $\mathbf{w} \geq \mathbf{v}$ such that $p(\mathbf{u}) \xrightarrow{*}_{\mathbb{N}^d} q(\mathbf{w})$.

The complexity of boundedness and coverability for d -VASS in a fixed dimension d has been studied by Rosier and Yen in [22]. They show that both problems are PSPACE-complete for any fixed $d \geq 4$. Chan [2] later noted that boundedness is already PSPACE-complete for $d = 3$, leaving the case $d = 2$ as an open problem. It is moreover known that for $d = 1$ those problems are NP-complete [6]. From the results in [3] and Lemma 17, it is easy to show that both problems are PSPACE-complete for every fixed $d \geq 2$. An instance of reachability between $p(\mathbf{u})$ and $q(\mathbf{v})$ in a bounded one-counter automaton with bound b can be reduced to boundedness and coverability in 2-VASS by using the construction of Lemma 17 as a gadget and adding an extra transition $(-v, v - b)$ from q to a fresh control state r which has a self-loop $(1, 1)$.

Corollary 20. Boundedness and coverability in d -VASS are PSPACE-complete for any fixed $d \geq 2$.

VI. CONCLUSION

This paper established the precise complexity, *i.e.*, PSPACE-completeness, of the reachability problem for 2-VASS. We also noted that the coverability and boundedness problems for 2-VASS are PSPACE-complete. When numbers are encoded in unary we showed that \mathbb{Z} -reachability in d -VASS is NL-complete for fixed d . Reachability for unary 2-VASS was shown to be NL-hard and in NP. Our approach does not immediately lead to a better upper bound than NP mainly due to the following reason. Our proof showed that the reachability relation can be captured by a set of linear path schemes whose $*$ -length is quadratic in the number of control states. The matrix of the resulting system of linear Diophantine inequalities thus has quadratically many columns and its smallest solution — which corresponds to the exponents of the cycles of the linear path scheme and hence of the length of the path — can thus become exponentially large.

REFERENCES

- [1] T. Ball, S. Chaki, and S. Rajamani, “Parameterized verification of multithreaded software libraries,” in *Tools and Algorithms for the Construction and Analysis of Systems*, ser. Lecture Notes in Computer Science, vol. 2031. Springer, 2001, pp. 158–173.
- [2] T. Chan, “The boundedness problem for three-dimensional vector addition systems with states,” *Information Processing Letters*, vol. 26, no. 6, pp. 287–289, 1988.
- [3] J. Fearnley and M. Jurdziński, “Reachability in two-clock timed automata is PSPACE-complete,” in *Automata, Languages, and Programming*, ser. Lecture Notes in Computer Science, vol. 7966. Springer, 2013, pp. 212–223.

- [4] S. M. German and A. P. Sistla, “Reasoning about systems with many processes,” *Journal of the ACM*, vol. 39, no. 3, pp. 675–735, Jul. 1992.
- [5] C. Haase, S. Kreutzer, J. Ouaknine, and J. Worrell, “Reachability in succinct and parametric one-counter automata,” in *CONCUR 2009 – Concurrency Theory*, ser. Lecture Notes in Computer Science, vol. 5710. Springer, 2009, pp. 369–383.
- [6] C. Haase, “On the complexity of model checking counter automata,” Ph.D. dissertation, University of Oxford, 2012.
- [7] C. Haase and S. Halfon, “Integer vector addition systems with states,” in *Reachability Problems - 8th International Workshop*, ser. Lecture Notes in Computer Science, vol. 8762. Springer, 2014, pp. 112–124.
- [8] M. Heiner, D. Gilbert, and R. Donaldson, “Petri nets for systems and synthetic biology,” in *Formal Methods for Computational Systems Biology*, 2008, pp. 215–264.
- [9] J. Hopcroft and J.-J. Pansiot, “On the reachability problem for 5-dimensional vector addition systems,” *Theoretical Computer Science*, vol. 8, no. 2, pp. 135–159, 1979.
- [10] R. R. Howell, L. E. Rosier, D. T. Huynh, and H.-C. Yen, “Some complexity bounds for problems concerning finite and 2-dimensional vector addition systems with states,” *Theoretical Computer Science*, vol. 46, no. 3, pp. 107–140, 1986.
- [11] S. R. Kosaraju, “Decidability of reachability in vector addition systems (preliminary version),” in *Proceedings of the 14th annual ACM symposium on Theory of computing*. ACM, 1982, pp. 267–281.
- [12] J.-L. Lambert, “A structure to decide reachability in Petri nets,” *Theoretical Computer Science*, vol. 99, no. 1, pp. 79–104, 1992.
- [13] J. Leroux and S. Schmitz, “Reachability in vector addition systems demystified,” in *Proceedings of the 30th Annual ACM/IEEE Symposium on Logic in Computer Science*. ACM, 2015, to appear.
- [14] J. Leroux and G. Sutre, “On flatness for 2-dimensional vector addition systems with states,” in *CONCUR 2004 – Concurrency Theory*, ser. Lecture Notes in Computer Science, vol. 3170. Springer, 2004, pp. 402–416.
- [15] J. Leroux, “Vector addition systems reachability problem (a simpler solution),” in *Turing-100*, ser. EPIc Series, vol. 10. EasyChair, 2012, pp. 214–228.
- [16] —, “The general vector addition system reachability problem by presburger inductive invariants,” in *Proceedings of the 24th Annual IEEE Symposium on Symposium on Logic in Computer Science*. IEEE, 2009, pp. 4–13.
- [17] R. J. Lipton, “The reachability problem is exponential-space-hard,” Department of Computer Science, Yale University, Tech. Rep. 62, 1976.
- [18] E. W. Mayr, “An algorithm for the general Petri net reachability problem,” in *Proceedings of the 13th Annual ACM Symposium on Theory of Computing*. ACM, 1981, pp. 238–246.
- [19] C. A. Petri, *Kommunikation mit Automaten*. Bonn: Institut für Instrumentelle Mathematik, Schriften des IIM Nr. 2, 1962.
- [20] L. Pottier, “Minimal solutions of linear Diophantine systems : bounds and algorithms,” in *Rewriting Techniques and Applications*, ser. Lecture Notes in Computer Science, vol. 488. Springer, 1991, pp. 162–173.
- [21] V. Reddy, M. Liebman, and M. Mavrouniotis, “Qualitative analysis of biochemical reaction systems,” *Computers in biology and medicine*, vol. 26, no. 1, pp. 9–24, 1996.
- [22] L. E. Rosier and H.-C. Yen, “A multiparameter analysis of the boundedness problem for vector addition systems,” *Journal of Computer and System Sciences*, vol. 32, no. 1, pp. 105–135, 1986.
- [23] G. S. Sacerdote and R. L. Tenney, “The decidability of the reachability problem for vector addition systems (preliminary version),” in *Proceedings of the 9th annual ACM symposium on Theory of computing*. ACM, 1977, pp. 61–76.
- [24] A. Schrijver, *Theory of linear and integer programming*. John Wiley & Sons, 1998.
- [25] A. W. To, “Model checking infinite-state systems: generic and specific approaches,” Ph.D. dissertation, University of Edinburgh, 2010.
- [26] L. G. Valiant and M. Paterson, “Deterministic one-counter automata,” *Journal of Computer and System Sciences*, vol. 10, no. 3, pp. 340–350, 1975.
- [27] W. van der Aalst, “The application of Petri nets to workflow management,” *Journal of circuits, systems, and computers*, vol. 8, no. 1, pp. 21–66, 1998.