Deciding Conditional Termination

Marius Bozga, Radu Iosif, Filip Konecny

Deciding Conditional Termination*

Marius Bozga[1], Radu Iosif[1], and Filip Konečný[1,2]

[1] VERIMAG, CNRS, 2 av. de Vignate, 38610 Gières, France
{bozga,iosif}@imag.fr
[2] FIT BUT, Božetěchova 2, 61266 Brno, Czech Republic
ikonecny@fit.vutbr.cz

Abstract. This paper addresses the problem of conditional termination, which is that of defining the set of initial configurations from which a given program terminates. First we define the dual set, of initial configurations, from which a non-terminating execution exists, as the greatest fixpoint of the pre-image of the transition relation. This definition enables the representation of this set, whenever the closed form of the relation of the loop is definable in a logic that has quantifier elimination. This entails the decidability of the termination problem for such loops. Second, we present effective ways to compute the weakest precondition for non-termination for difference bounds and octagonal (non-deterministic) relations, by avoiding complex quantifier eliminations. We also investigate the existence of linear ranking functions for such loops. Finally, we study the class of linear affine relations and give a method of under-approximating the termination precondition for a non-trivial subclass of affine relations. We have performed preliminary experiments on transition systems modeling real-life systems, and have obtained encouraging results.

1 Introduction

The termination problem asks whether every computation of a given program ends in a halting state. The universal termination asks whether a given program stops for every possible input configuration. Both problems are among the first ever to be shown undecidable, by A. Turing [24]. In many cases however, programs will terminate when started in certain configurations, and may[3] run forever, when started in other configurations. The problem of determining the set of configurations from which a program terminates on all paths is called *conditional termination*.

In program analysis, the presence of non-terminating runs has been traditionally considered faulty. However, more recently, with the advent of *reactive systems*, accidental termination can be an equally serious error. For instance, when designing a web server, a developer would like to make sure that the main program loop will not exit

* This work was supported by the French national project ANR-09-SEGI-016 VERIDYC, by the Czech Science Foundation (projects P103/10/0306 and 102/09/H042), the Czech Ministry of Education (projects COST OC10009 and MSM 0021630528), the Barrande project MEB021023, and the EU/Czech IT4Innovations Centre of Excellence CZ.1.05/1.1.00/02.0070.

[3] If the program is non-deterministic, the existence of a single infinite run, among other finite runs, suffices to consider an initial configuration non-terminating.

unless a stopping request has been issued. These facts lead us to considering the *conditional non-termination* problem, which is determining the set of initial configurations which guarantee that the program will not exit.

In this paper we focus on programs that handle integer variables, performing linear arithmetic tests and (possibly non-deterministic) updates. A first observation is that the set of configurations guaranteeing non-termination is the greatest fixpoint of the pre-image of the program's transition relation[4] R. This set, called the *weakest recurrent set*, and denoted $wrs(R)$ in our paper, can be defined in first-order arithmetic, provided that the closed form of the infinite sequence of relations $\{R^i\}_{i\geq 0}$, obtained by composing the transition relation with itself $0, 1, 2, \ldots$ times, can also be defined using first-order arithmetic. Moreover, if the fragment of arithmetic we use has quantifier elimination, the weakest recurrent set can be expressed in a quantifier-free decidable fragment of arithmetic. This also means that the problem $wrs(R) \stackrel{?}{=} \emptyset$ is decidable, yielding universal termination decidability proofs for free.

Contributions of this paper The main novelty in this paper is of rather theoretical nature: we show that the non-termination preconditions for integer transition relations defined as either *octagons* or *linear affine loops with finite monoid property* are definable in quantifier-free Presburger arithmetic. Thus, the universal termination problem for such program loops is decidable. However, since quantifier elimination in Presburger arithmetic is a complex procedure, we have developed alternative ways of deriving the preconditions for non-termination, and in particular:

- for *difference bounds*, we reduce the problem of finding the weakest recurrent set to finding the maximal solution of a system of inequalities in the complete lattice of integers extended with $\pm\infty$, where the right-hand sides use addition and $\min$ operators. Efficient algorithms for finding such maximal solutions are based on policy iteration [14]. This encoding gives us a worst-case time complexity of $\mathcal{O}(n^2 \cdot 2^n)$ in the number of variables n, for the computation of the weakest recurrent set for difference bounds relations.
- for *octagonal relations* (and implicitly for difference bounds relations, which are a subclass), we use a result from [5], namely that the sequence $\{R^i\}_{i\geq 0}$ is, in some sense, periodic. We give here a simple quantifier elimination method, targeted for the class of formulae defining weakest recursive sets. The algorithm suggested here runs in worst-case time complexity of $\mathcal{O}(n^3 \cdot 5^n)$ in the number of variables n. Moreover, we investigate the existence of linear ranking functions, and prove that, for each well-founded octagonal relations, there exists an effectively computable witness relation i.e., a well-founded relation that has a linear ranking function.
- for *linear affine relations*, weakest recurrent sets can be defined in Presburger arithmetic if we consider several restrictions concerning the transformation matrix. If the matrix A defining R has eigenvalues which are either zeros or roots of unity, all non-zero eigenvalues being of multiplicity one (these conditions are equivalent to the finite monoid property of [2, 12]), then $wrs(R)$ is Presburger definable. Otherwise, if all non-zero eigenvalues of A are roots of unity, of multiplicities greater or equal to one, $wrs(R)$ can be expressed using polynomial terms. In this case, we

[4] This definition is the dual of the *reachability set*, needed for checking safety properties: the reachability set is the least fixpoint of the post-image of the transition relation.

can systematically issue termination preconditions, which are of significant practical importance, as noted in [10].

For space reasons, all proofs are deferred to the Appendix.

Practical applications Unfortunately, in practice, the cases in which the closed form of the sequence $\{R^i\}_{i\geq 0}$ is definable in a logic that has quantifier elimination, are fairly rare. All relations considered so far are conjunctive, meaning that they can represent only simple program loops of the form `while(condition){body}`, where the loop body contains no further conditional constructs. In order to deal with more complicated program loops, one can use the results from this paper in several ways:

- use the decision procedures as a back-end of a termination analyzer, in order to detect spurious non-termination counterexamples consisting of a finite prefix (stem) and a conjunctive loop body (lasso). The spurious counterexamples can be discarded by intersecting the program model with the complement of the weak deterministic Büchi automaton representing the counterexample, as in [16].
- abstract a disjunctive loop body $R_1 \vee \ldots \vee R_n$ by a non-deterministic difference bounds or octagonal[5] relation $R^{\#} \supseteq R_{1,\ldots,n}$ and compute the weakest recurrent set of the latter. The complement of this set is a set of configurations from which the original loop terminates.
- attempt to compute a *transition invariant* i.e., an overapproximations of the transitive closure of the disjunctive loop body $(R_1 \vee \ldots \vee R_n)^+$ (using e.g., the semi-algorithmic unfolding technique described in [6]) and overapproximate it by a disjunction $R_1^{\#} \vee \ldots \vee R_m^{\#}$ of difference bounds or octagonal relations. Then compute the weakest recurrent set of each relation in the latter disjunction. If $wrs(R_1^{\#}) = \ldots = wrs(R_m^{\#}) = \emptyset$, the original loop terminates on any input, following the principle of transition invariants [19].

1.1 Related Work

The literature on program termination is vast. Most work focuses however on universal termination, such as the techniques for synthesizing linear ranking functions of Sohn and Van Gelder [22] or Podelski and Rybalchenko [18], and the more sophisticated method of Bradley, Manna and Sipma [8], which synthesizes lexicographic polynomial ranking functions, suitable when dealing with disjunctive loops. However, not every terminating program (loop) has a linear (polynomial) ranking function. In this paper we show that, for an entire class of non-deterministic linear relations, defined using octagons, termination is always witnessed by a computable octagonal relation that has a linear ranking function.

Another line of work considers the decidability of termination for simple (conjunctive) linear loops. Initially Tiwari [23] shows decidability of termination for affine linear loops interpreted over *reals*, while Braverman [9] refines this result by showing decidability over *rationals* and over *integers*, for homogeneous relations of the form $C_1\mathbf{x} > 0 \ \wedge \ C_2\mathbf{x} \geq 0 \ \wedge \ \mathbf{x}' = A\mathbf{x}$. The non-homogeneous integer case seems to be

[5] The linear affine relations considered in this paper are deterministic, which makes them unsuitable for abstraction.

much more difficult as it is closely related to the open *Skolem's Problem* [15]: given a linear recurrence $\{u_i\}_{i\geq 0}$, determine whether $u_i = 0$ for some $i \geq 0$.

Our work is concerned mostly with proofs of decidability: we show that the termination problem for a program loop becomes decidable if the closed form of the sequence of iterations of the loop can be defined in a decidable logic. As shown in [5], for octagonal and linear affine relations with the finite monoid property, this closed form is Presburger definable, which now entails the decidability of the termination problem for these classes of relations.

The work which is closest to ours is probably that of Cook et al. [10]. In this paper, the authors develop an algorithm for deriving termination preconditions, by first guessing a ranking function candidate (typically the linear term from the loop condition) and then inferring a supporting assertion, which guarantees that the candidate function decreases with each iteration. The step of finding a supporting assertion requires a fixpoint iteration, in order to find an invariant condition. Unlike our work, the authors of [10] do not address issues related to completeness: the method is not guaranteed to find the weakest precondition for termination, even in cases when this set can be computed. On the other hand, it is applicable to a large range of programs, extracted from real-life software. To compare our method with theirs, we tried all examples available in [10]. Since most of them are linear affine relations, we used our under-approximation method and have computed termination preconditions, which turn out to be slightly more general than the ones reported in [10].

2 Preconditions for Non-termination

In the rest of this paper we denote by $\mathbf{x} = \{x_1, \ldots, x_n\}$ the set of working variables, ranging over a domain of values denoted as $\mathcal{D}$. A *state* is a valuation $s : \mathbf{x} \rightarrow \mathcal{D}$, or equivalently, an n-tuple of values from $\mathcal{D}$. An *execution step* is a relation $R \subseteq \mathcal{D}^n \times \mathcal{D}^n$ defined by an *arithmetic formula* $\mathcal{R}(\mathbf{x}, \mathbf{x}')$, where the set $\mathbf{x}' = \{x'_1, \ldots, x'_n\}$ denotes the values of the variables after executing R once. If s and s' are valuations of the sets $\mathbf{x}$ and $\mathbf{x}'$, we denote by $\mathcal{R}(s, s')$ the fact that $(s, s') \in R$. A relation R is said to be *consistent* if there exist states s, s' such that $\mathcal{R}(s, s')$.

Relational composition is defined as $R_1 \circ R_2 = \{(s, s') \in \mathcal{D}^n \times \mathcal{D}^n \mid \exists s'' \in \mathcal{D}^n \,.\, \mathcal{R}_1(s, s'') \wedge \mathcal{R}_2(s'', s')\}$. For any relation $R \in \mathcal{D}^n \times \mathcal{D}^n$, we consider R^0 to be the identity relation, and we define $R^{i+1} = R^i \circ R$, for all $i \geq 0$. The pre-image of a set $S \subseteq \mathcal{D}^n$ via R is the set $pre_R(S) = \{s \in \mathcal{D}^n \mid \exists s' \in S \,.\, \mathcal{R}(s, s')\}$. It is easy to check that $pre^i_R(S) = pre_{R^i}(S)$, for any $S \subseteq \mathcal{D}^n$ and for all $i \geq 0$. For any $i \geq 0$, we write $\mathcal{R}^i$ for the formula defining the relation R^i and $\mathcal{R}^{-i}(\top)$ for the formula defining the set $pre_{R^i}(\mathcal{D}^n)$.

Definition 1. *A relation R is said to be* $*$-consistent *if and only if, for any $k > 0$, there exists a sequence of states $s_1, \ldots, s_k$, such that $\mathcal{R}(s_i, s_{i+1})$, for all $i = 1, \ldots, k-1$. R is said to be* well-founded *if and only if there is no infinite sequence of states $\{s_i\}_{i>0}$, such that $\mathcal{R}(s_i, s_{i+1})$, for all $i > 0$.*

Notice that if a relation is not $*$-consistent, then it is also well-founded. However the dual is not true. For instance, the relation $R = \{(n, n-1) \mid n > 0\}$ is both $*$-consistent and well-founded.

Definition 2. *A set $S \subseteq \mathcal{D}^n$ is said to be a* non-termination precondition *for R if, for each state $s \in S$ there exists an infinite sequence of states $s_0, s_1, s_2, \ldots$ such that $s = s_0$ and $\mathcal{R}(s_i, s_{i+1})$, for all $i \geq 0$.*

If $S_0, S_1, \ldots$ are all preconditions for non-termination for R, then the (possibly infinite) union $\bigcup_{i=0,1,\ldots} S_i$ is a precondition for non-termination for R as well. The set $wnt(R) = \bigcup\{S \in \mathcal{D}^n \mid S \text{ is a precondition for non-termination for } R\}$ is called the *weakest non-termination precondition* for R. A relation R is well-founded if and only if $wnt(R) = \emptyset$. A set S such that $S \cap wnt(R) = \emptyset$ is called a *termination precondition*.

Definition 3. *A set $S \subseteq \mathcal{D}^n$ is said to be* recurrent *for a relation $R \in \mathcal{D}^n \times \mathcal{D}^n$ if and only if $S \subseteq pre_R(S)$.*

Proposition 1. *Let $S_0, S_1, \ldots \in \mathcal{D}^n$ be a (possibly infinite) sequence of sets, all of which are recurrent for a relation $R \in \mathcal{D}^n \times \mathcal{D}^n$. Then their union $\bigcup_{i=0,1,\ldots} S_i$ is recurrent for R as well.*

The set $wrs(R) = \bigcup\{S \in \mathcal{D}^n \mid S \text{ is a recurrent set for } R\}$ is called the *weakest recurrent set* for R. By Proposition 1, $wrs(R)$ is recurrent for R. Next we define the weakest recurrent set as the greatest fixpoint of the transition relation's pre-image.

Lemma 1. *Given a relation $R \in \mathcal{D}^n \times \mathcal{D}^n$, the weakest recurrent set for R is the greatest fixpoint of the function $X \mapsto pre_R(X)$.*

As a consequence, we obtain $wrs(R) = \bigcap_{i>0} pre_R^i(\mathcal{D}^n)$, by the Kleene Fixpoint Theorem. Since $pre_R^i = pre_{R^i}$, we have $wrs(R) = \bigcap_{i>0} pre_{R^i}(\mathcal{D}^n)$. In other words, from any state in the weakest recurrent set for a relation, an iteration of any finite length of the given relation is possible. The following lemma shows that in fact, this is exactly the set of states from which an infinite iteration is also possible.

Lemma 2. *Given a relation $R \in \mathcal{D}^n \times \mathcal{D}^n$, the weakest recurrent set for R equals its weakest precondition for non-termination.*

The characterization of weakest recurrent sets as greatest fixpoints of the pre-image function suggests a method for computing such sets. In this section we show that, for certain classes of relations, these sets are definable in Presburger arithmetic, which gives a decision procedure for the well-foundedness problem for certain classes of relations, and consequently, for the termination problem for several classes of program loops.

Definition 4. *Given a relation $R \in \mathcal{D}^n \times \mathcal{D}^n$ defined by an arithmetic formula $\mathcal{R}(\mathbf{x}, \mathbf{x}')$, the* closed form *of R is a formula $\mathcal{R}^{(k)}(\mathbf{x}, \mathbf{x}')$, with free variables $\mathbf{x} \cup \mathbf{x}' \cup \{k\}$, such that for every integer valuation $i > 0$ of k, $\mathcal{R}^{(i)}(\mathbf{x}, \mathbf{x}')$ defines the relation R^i.*

Example Consider for instance the relation $\mathcal{R}(x, x') \equiv x \geq 0 \wedge x' = x - 1$. Then we have $\mathcal{R}^{(k)}(x, x') \equiv x \geq k - 1 \wedge x' = x - k$. □

Since, by Lemma 1, we have $wrs(R) = \text{gfp}(pre_R) = \bigcap_{i>0} pre_{R^i}(\mathcal{D}^n)$, using the closed form of R, one can now define:

$$wrs(R) \equiv \forall k > 0\ \exists \mathbf{x}'\ .\ \mathcal{R}^{(k)}(\mathbf{x}, \mathbf{x}') \tag{1}$$

Because Presburger arithmetic has quantifier elimination, $wrs(R)$ can be defined in Presburger arithmetic[6] whenever $\mathcal{R}^{(k)}$ can. In [5] we show three classes of relations for which $\mathcal{R}^{(k)}$ is Presburger definable: difference bounds, octagonal and finite-monoid affine relations (the formal definitions of these classes are given in the next section). For each of these classes of loops termination is decidable, by the above argument.

Example Consider again the relation $\mathcal{R}(x, x') \equiv x \geq 0 \wedge x' = x - 1$ for which $\mathcal{R}^{(k)}(x, x') \equiv x \geq k - 1 \wedge x' = x - k$. Quantifier elimination yields $wrs(R) \equiv \forall k > 0\ \exists x' \,.\, x \geq k - 1 \wedge x' = x - k \equiv \forall k > 0 \,.\, x \geq k - 1 \equiv$ false. Hence the relation $\mathcal{R}$ is well-founded. □

3 Difference Bounds Relations

In this and the following sections, we assume that the variables $\mathbf{x} = \{x_1, \ldots, x_n\}$ range over integers i.e., that $\mathcal{D} = \mathbb{Z}$.

Definition 5. *A formula $\phi(\mathbf{x})$ is a* difference bounds constraint *if it is equivalent to a finite conjunction of atomic propositions of the form $x_i - x_j \leq a_{ij}$, for $1 \leq i, j \leq n, i \neq j$, where $a_{ij} \in \mathbb{Z}$.*

Given a difference bounds constraint ϕ, a *difference bounds matrix* (DBM) representing ϕ is a matrix $m_\phi \in \mathbb{Z}_\infty^{n \times n}$ such that $(m_\phi)_{ij} = a_{ij}$, if $x_i - x_j \leq a_{ij}$ is an atomic proposition in ϕ, and ∞, otherwise. If ϕ is inconsistent (logically equivalent to false) we also say that m_ϕ is inconsistent. The next definition gives a canonical form for consistent DBMs.

Definition 6. *A consistent DBM $m \in \mathbb{Z}_\infty^{n \times n}$ is said to be* closed *if and only if $m_{ii} = 0$ and $m_{ij} \leq m_{ik} + m_{kj}$, for all $1 \leq i, j, k \leq n$.*

Given a consistent DBM m, we denote by m^* the (unique) closed DBM equivalent with it. It is well-known that, if m is consistent, then m^* is unique, and can be computed from m in time $\mathcal{O}(n^3)$, by the classical Floyd-Warshall algorithm. The closure of DBM provides an efficient means to compare difference bounds constraints.

Proposition 2 ([17]). *Given two consistent difference bounds constraints $\varphi(\mathbf{x})$ and $\psi(\mathbf{x})$, the following conditions are equivalent:*

- $\forall \mathbf{x} \,.\, \varphi(\mathbf{x}) \rightarrow \psi(\mathbf{x})$
- $(m_\varphi^*)_{ij} \leq (m_\psi^*)_{ij}$, *for all* $1 \leq i, j \leq n$

In the following, let R be a relation defined by a difference bounds constraint. It is easy to show that, for any $i \geq 0$, the relation R^i is a difference bounds relation as well – in other words, difference bounds relations are closed under composition. Moreover, if S is a set defined by a difference bounds constraint, then the set $pre_{R^i}(S)$ is defined by a difference bounds constraint as well. But since $wrs(R) = \bigcap_{i>0} pre_{R^i}(\mathbb{Z}^n)$, it turns out that $wrs(R)$ can be defined by a difference bounds constraint, since the class of difference bounds constraints is closed under (possibly infinite) intersections.

We are now ready to describe the procedure computing the weakest recurrent set for a difference bounds relation R. Since $wrs(R)$ is a (possibly inconsistent) difference bounds constraint, we use the template $\mu(\mathbf{x}, \mathbf{p}) \equiv \bigwedge_{1 \leq i \neq j \leq n} x_i - x_j \leq p_{ij}$, where p_{ij}

[6] Or, for that matter, in any theory that has quantifier elimination.

are parameters ranging over $\mathbb{Z}_{\pm\infty}$ (we clearly do not need to track the constraints of the form $x_i - x_i \leq p_{ii}$). Moreover, we assume that the template is closed (Definition 6), which can be encoded as a system of inequalities of the form:

$$p_{ij} \leq \min\left\{p_{ik} + p_{kj} \mid k \neq i, k \neq j\right\} \tag{2}$$

Next, we compute the (symbolic) difference bounds constraint corresponding to the set $pre_R(\mu) \equiv \exists \mathbf{x}' \,.\, \mathcal{R}(\mathbf{x}, \mathbf{x}') \wedge \mu(\mathbf{x}', \mathbf{p})$. This step requires computing the closure of the DBM corresponding to $\mathcal{R} \wedge \mu$, and elimination of the $\mathbf{x}'$ variables. The result is a closed symbolic DBM π, whose entries are $\min$-terms consisting of sums of p_{ij} and integer constants. Further, we encode the recurrence condition $\mu \subseteq pre_R(\mu)$, again as a system of inequalities (Proposition 2) of the form:

$$p_{ij} \leq \pi_{ij}, i \neq j \tag{3}$$

By conjoining the inequalities (2) and (3), we obtain a system of inequalities with variables p_{ij}, whose right-hand sides are linear combinations of p_{ij} with addition and $\min$. We are interested in the maximal solution of this system, which can be obtained using an efficient policy iteration algorithm [14] in the complete lattice of $\mathbb{Z}_{\pm\infty}$ with addition, $\min$ and $\max$ operators. This solution defines the weakest recurrent set for R, and consequently, the weakest precondition for non-termination of the R loop. Since $wrs(R)$ is a difference bounds constraint, for any relation R definable by a difference bounds constraint, the maximal solution of the system is unique. It is to be noted that, if for some $1 \leq i \neq j \leq n$ we obtain $p_{ij} = -\infty$, then the weakest recurrent set is empty i.e., the relation R is well-founded, as shown by the following example.

Example Let us compute $wrs(R)$ for $\mathcal{R}(x, x') \equiv z - x \leq 0 \wedge x' = x - 1 \wedge z' = z$. The template used is $\mu(x, z) \equiv x - z \leq p_1 \wedge z - x \leq p_2$. We compute $pre_R(\mu) \equiv x - z \leq \min(p_1 + 1, 1) \wedge z - x \leq p_2 - 1$. The recurrence condition $\mu \subseteq pre_R(\mu)$ is given by the following system: $\{p_1 \leq \min(p_1 + 1, 1), p_2 \leq p_2 - 1\}$, whose unique maximal solution is $p_1 = 1, p_2 = -\infty$. This proves the well-foundedness of $\mathcal{R}$. □

Lemma 3. *Computing the weakest recurrent set of a difference bounds relation can be done in time $\mathcal{O}(n^2 \cdot 2^n)$, where n is the number of variables.*

4 Octagonal Relations

Definition 7. *A formula $\phi(\mathbf{x})$ is an* octagonal constraint *if it is equivalent to a finite conjunction of terms of the form $\pm x_i \pm x_j \leq a_{ij}$, where $a_{ij} \in \mathbb{Z}$ and $1 \leq i, j \leq n$.*

We represent octagons as difference bounds constraints over the *dual* set of variables $\mathbf{y} = \{y_1, y_2, \ldots, y_{2n}\}$, with the convention that y_{2i-1} stands for x_i and y_{2i} for $-x_i$, respectively. For example, the octagonal constraint $x_1 + x_2 = 3$ is represented as $y_1 - y_4 \leq 3 \wedge y_2 - y_3 \leq -3$. To handle the dual variables in the following, we define $\bar{i} = i - 1$, if i is even, and $\bar{i} = i + 1$ if i is odd. We say that a DBM $m \in \mathbb{Z}_{\infty}^{2n \times 2n}$ is *coherent* iff $m_{ij} = m_{\bar{j}\bar{i}}$ for all $1 \leq i, j \leq 2n$. The coherence property is needed because any atomic proposition $x_i - x_j \leq a$, in ϕ can be represented as both $y_{2i-1} - y_{2j-1} \leq a$ and $y_{2j} - y_{2i} \leq a$, $1 \leq i, j \leq n$. We denote by $\overline{\phi}$ the difference bounds formula

$\phi[y_1/x_1, y_2/-x_1, \ldots, y_{2n-1}/x_n, y_{2n}/-x_n]$ with free variables $\mathbf{y}$. The following equivalence relates ϕ and $\overline{\phi}$:

$$\phi(\mathbf{x}) \Leftrightarrow (\exists y_2, y_4, \ldots, y_{2n} \,.\, \overline{\phi} \wedge \bigwedge_{i=1}^{n} y_{2i-1} + y_{2i} = 0)[x_1/y_1, \ldots, x_n/y_{2n-1}] \quad (4)$$

Given a coherent DBM m representing $\overline{\phi}$, we say that m is *octagonal-consistent* if and only if ϕ is consistent. The following definition gives the canonical form of a DBM representing an octagonal-consistent constraint.

Definition 8. *An octagonal-consistent coherent DBM $m \in \mathbb{Z}^{2n \times 2n}$ is said to be* tightly closed *if and only if the following hold:*

1. $m_{ii} = 0,\ \forall 1 \leq i \leq 2n$
2. $m_{i\bar{i}}$ *is even*, $\forall 1 \leq i \leq 2n$
3. $m_{ij} \leq m_{ik} + m_{kj},\ \forall 1 \leq i, j, k \leq 2n$
4. $m_{ij} \leq \lfloor \frac{m_{i\bar{i}}}{2} \rfloor + \lfloor \frac{m_{\bar{j}j}}{2} \rfloor,\ \forall 1 \leq i, j \leq 2n$

Given an octagonal-consistent DBM m, we denote by m^t the equivalent tightly closed DBM. The tight closure of an octagonal-consistent DBM m is unique and can be computed in time $\mathcal{O}(n^3)$ as $m^t_{i,j} = min\left\{m^*_{i,j}, \left\lfloor \frac{m^*_{i,\bar{i}}}{2} \right\rfloor + \left\lfloor \frac{m^*_{\bar{j},j}}{2} \right\rfloor\right\}$ [1]. This generalizes to unbounded finite compositions of octagonal relations [4]:

$$\forall k \geq 0 \,.\, (m^t_{R^k})_{i,j} = min\left\{(m^*_{\overline{R}^k})_{i,j}, \left\lfloor \frac{(m^*_{\overline{R}^k})_{i,\bar{i}}}{2} \right\rfloor + \left\lfloor \frac{(m^*_{\overline{R}^k})_{\bar{j},j}}{2} \right\rfloor\right\} \quad (5)$$

Notice that the above relates the entries of the tightly closed DBM representation of R^k with the entries of the closed DBM representation of the relation defined by $\overline{\mathcal{R}}^k$.

We are now ready to introduce a result [5] that defines the "shape" of the closed form $\mathcal{R}^{(k)}$ for an octagonal relation R. Intuitively, for each $i \geq 0$, R^i is an octagon, whose bounds evolve in a periodic way. The following definition gives the precise meaning of periodicity for relations that have a matrix representation.

Definition 9. *An infinite sequence of matrices $\{M_k\}_{k=1}^{\infty} \in \mathbb{Z}_{\infty}^{m \times m}$ is said to be* ultimately periodic *if and only if:*

$$\exists b > 0\ \exists c > 0\ \exists \Lambda_0, \Lambda_1, \ldots, \Lambda_{c-1} \in \mathbb{Z}_{\infty}^{m \times m} \,.\, M_{b+(k+1)c+i} = \Lambda_i + M_{b+kc+i}$$

for all $k \geq 0$ and $i = 0, 1, \ldots, c-1$. The smallest b, c for which the above holds are called prefix *and* period *of the $\{M_k\}_{k=1}^{\infty}$ sequence, respectively.*

A result reported in [5] is that the sequence $\{m^t_{R^i}\}_{i \geq 0}$ (5) of tightly closed matrices representing the sequence $\{R^i\}_{i \geq 0}$ of powers of an octagonal relation R is ultimately periodic, in the sense of the above definition. The constants b and c from Definition 9 will also be called the *prefix and period of the octagonal relation R*, throughout this section.

For a set $\mathbf{v}$ of variables, let $U(\mathbf{v}) = \{\pm v_1 \pm v_2 \mid v_1, v_2 \in \mathbf{v}\}$ denote the set of octagonal terms over $\mathbf{v}$. As a first remark, by the periodicity of the sequence $\{m^t_{R^i}\}_{i \geq 0}$, the closed form of the subsequence $\{R^{b+c\ell}\}_{\ell \geq 0}$ (of $\{R^i\}_{i \geq 0}$) can be defined as:

$$\mathcal{R}^{(\ell)}_{b,c} \equiv \bigwedge_{u \in U(\mathbf{x} \cup \mathbf{x}')} u \leq a_u \ell + d_u \quad (6)$$

for all $\ell \geq 0$, where a_u and d_u are entries (i, j) corresponding to the term $u = y_i - y_j$ in the octagonal DBMs Λ_0 and $m^t_{R^b}$, respectively. This is the case, since the matrix sequence $\{m^t_{R^{b+c\ell}}\}_{\ell \geq 0}$ is ultimately periodic i.e., $m^t_{R^{b+c\ell}} = m^t_{R^b} + \ell\Lambda_0$, for all $\ell \geq 0$.

Example Given an octagonal relation $\mathcal{R} \equiv x{+}y{\leq}5 \wedge x'{=}x{+}2 \wedge y'{=}y{-}1$, we compute

$$\begin{aligned}\mathcal{R}^{(\ell)}_{b,c} \equiv\ & x{+}y{\leq}{-}\ell{+}5\ \wedge\ x'{-}x{=}2\ell{+}2\ \wedge\ y{-}y'{=}\ell+1\\ & \wedge\ x{+}y'{\leq}{-}2\ell{+}4\ \wedge\ x'{+}y{\leq}\ell{+}7\ \wedge\ x'{+}y'{\leq}6\end{aligned}$$

We have $b{=}c{=}1$, and $\mathcal{R}^{(k)} \equiv \mathcal{R}^{(k)}_{b,c}$. □

Second, we notice that the greatest fixpoint of a monotonic[7] function can be computed by an infinite subsequence of the classical decreasing Kleene iteration. Concretely, we have that $wrs(R) = \bigcap_{k>0} pre^k_R(\mathbb{Z}^n) = \bigcap_{\ell \geq 0} pre^{b+c\ell}_R(\mathbb{Z}^n)$. The latter set can now be defined using the closed form of the subsequence (6) i.e., $wrs(R) \equiv \forall \ell \geq 0\ \exists \mathbf{x}'\ .\ \mathcal{R}^{(\ell)}_{b,c}$.

The proof of periodicity from [5] relies on the fact that the DBM encoding of the closed form of R is tightly closed for any unfolding length k, see (5). Hence, the existential quantifier $\exists \mathbf{x}'$ can be eliminated by simply deleting all atomic propositions involving primed variables from (6). Further, we obtain:

$$wrs(R) \equiv \forall \ell \geq 0\ \textstyle\bigwedge_{u \in U(\mathbf{x})} u \leq a_u\ell + d_u \equiv \bigwedge_{u \in U(\mathbf{x})} u \leq \inf\{a_u\ell + d_u \mid \ell \geq 0\}$$

where, for a set $S \subseteq \mathbb{Z}$, $\inf S$ denotes the minimal element of S, if one exists, or $-\infty$, otherwise. We have

$$\inf\{a_u\ell + d_u \mid \ell \geq 0\} = \begin{cases} -\infty \text{ if } a_u < 0 \\ d_u \quad \text{otherwise} \end{cases}$$

Hence $wrs(R)$ is the empty set, if $a_u < 0$ for some $u \in U(\mathbf{x})$. Otherwise, we obtain $wrs(R) \equiv \bigwedge_{u \in U(\mathbf{x})} u \leq d_u$. However, this is exactly the set defined by $\mathcal{R}^{-b}(\top) \equiv \exists \mathbf{x}'\ .\ \mathcal{R}^b(\mathbf{x}, \mathbf{x}') \equiv \exists \mathbf{x}'\ .\ \mathcal{R}^{(0)}_{b,c}$, by (6). The following complexity upper bound is a consequence of this fact.

Lemma 4. *Computing the weakest recurrent set of an octagonal relation can be done in time $\mathcal{O}(b \cdot n^3)$, where b is its prefix and n is the number of variables. Alternatively, this problem has $\mathcal{O}(n^3 \cdot 5^n)$ worst-case time complexity.*

Example (continued) Following the decision procedure above, we obtain

$$wrs(R) \equiv \forall \ell \geq 0\ \exists \mathbf{x}'\ .\ \mathcal{R}^{(\ell)}_{b,c} \equiv \forall \ell \geq 0\ .\ x{+}y{\leq}{-}\ell{+}5$$

Hence $wrs(R) = \emptyset$ i.e., $\mathcal{R}$ is well-founded. □

4.1 On the Existence of Linear Ranking Functions

A ranking function for a given relation R constitutes a proof of the fact that R is well-founded. We distinguish here two cases. If R is not $*$-consistent, then the well-foundedness of R is witnessed simply by an integer constant $i > 0$ such that $R^i = \emptyset$.

[7] In our case, $pre^{k_1}_R(\mathbb{Z}^n) \supseteq pre^{k_2}_R(\mathbb{Z}^n)$, for $k_1 \leq k_2$.

Otherwise, if R is $*$-consistent, we need a better argument for well-foundedness. In this section we show that, for any $*$-consistent well-founded octagonal relation R, the (strenghtened) relation defined by $\mathcal{R}^{-b}(\top) \wedge \mathcal{R}$ is well-founded and has a linear ranking function, even when R alone does not have one. For space reasons, we do not give here all the details of the construction of such a function. However, the existence proof suffices, as one can use *complete* ranking function extraction tools (such as e.g. RankFinder [18]) in order to find them.

Definition 10. *Given a relation $R \subseteq \mathbb{Z}^n \times \mathbb{Z}^n$, a* linear ranking function *for R is a term $f(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i$ such that, for all states $s, s' : \mathbf{x} \to \mathbb{Z}$:*

1. *f is* decreasing*: $\mathcal{R}(s, s') \to f(s) > f(s')$*
2. *f is* bounded*: $\mathcal{R}(s, s') \to (f(s) > h \ \wedge \ f(s') > h)$, for some $h \in \mathbb{Z}$.*

The main result of this section is the following:

Theorem 1. *Let $R \subseteq \mathbb{Z}^n \times \mathbb{Z}^n$ be a $*$-consistent and well-founded octagonal relation, with prefix $b \geq 0$. Then, the relation defined by $\mathcal{R}^{-b}(\top) \wedge \mathcal{R}$ is well founded and has a linear ranking function.*

The first part of the theorem is proved by the following lemma:

Lemma 5. *Let $R \subseteq \mathbb{Z}^n \times \mathbb{Z}^n$ be a relation, and $m > 0$ be an integer. Then $wrs(R) = \emptyset$ if and only if $wrs(R_m) = \emptyset$, where R_m is the relation defined by $\mathcal{R}^{-m}(\top) \wedge \mathcal{R}$.*

It remains to prove that the witness relation defined by $\mathcal{R}^{-b}(\top) \wedge \mathcal{R}$ has a linear ranking function, provided that it is well-founded. The proof is organized as follows. First we show that well-foundedness of an octagonal relation R is equivalent to the well-foundedness of its difference bounds representation $\overline{\mathcal{R}}$ (Lemma 6). Second, we use a result from [7], that the constraints in the sequence of iterated difference bounds relations $\{\overline{\mathcal{R}}^i\}_{i \geq 0}$ can be represented by a finite-state weighted automaton, called the *zigzag automaton* in the sequel. If the relation defined by $\overline{\mathcal{R}}$ is well-founded, then this weighted automaton must have a cycle of negative weight. The structure of this cycle, representing several of the constraints in $\overline{\mathcal{R}}$, is used to show the existence of the linear ranking function for the witness relation $\mathcal{R}^{-b}(\top) \wedge \mathcal{R}$.

Lemma 6. *Let $R \subseteq \mathbb{Z}^n \times \mathbb{Z}^n$ be an octagonal relation and R_{db} be the difference bounds relation defined by $\overline{\mathcal{R}}$. Then R is well-founded if and only if R_{db} is well-founded.*

The above lemma reduces the problem of showing existence of a ranking function for an octagonal relation $\mathcal{R}(\mathbf{x}, \mathbf{x}')$ to showing existence of a ranking function for its difference bounds encoding $\overline{\mathcal{R}}(\mathbf{y}, \mathbf{y}')$. Assume that $f(\mathbf{y})$ is a ranking function for $\overline{\mathcal{R}}$. Then $f[x_i/y_{2i-1}, -x_i/y_{2i}]_{i=1}^{n}$ is a linear ranking function for R. Hence, in the rest of this section, we consider without loss of generality that R is a difference bounds relation.

Zigzag Automata For the later developments, we need to introduce the *zigzag automaton* corresponding to a difference bounds relation R. Intuitivelly, for any $i > 0$, the relation R^i can be represented by a constraint graph which is the i-times repetition of the constraint graph of R. The constraints induced by R^i can be represented as shortest paths in this graph, and can be recognized (in the classical automata-theoretic sense)

by a weighted automaton $\mathcal{A}_R$. The structure of this automaton is needed to show the existence of a linear ranking function.

For a difference bounds relation R, we define the directed graph $\mathcal{G}_R$, whose set of vertices is the set $\mathbf{x} \cup \mathbf{x}'$, and in which there is an edge from x_i to x_j labeled a_{ij} if and only if the atomic proposition $x_i - x_j \leq a_{ij}$ occurs in R. Clearly, m_R is the incidence matrix of $\mathcal{G}_R$. We define the concatenation of $\mathcal{G}_R$ with itself as the disjoint union of two copies of $\mathcal{G}_R$, in which the $\mathbf{x}$ vertices of the second copy overlap with the $\mathbf{x}'$ vertices of the first copy. Then R^m corresponds to the graph $\mathcal{G}_R^m$, obtained by concatenating the graph of R to itself $m > 0$ times.

Example Let $R \equiv x_3 - x_2' \leq 0 \wedge x_1 - x_3' \leq 0 \wedge x_4' - x_4 \leq 0 \wedge x_3' - x_2 \leq 0$. Figure 1 (b) shows $\mathcal{G}_{R^8}$, the 8-times unfolding of the graph $\mathcal{G}_R$ representing R. □

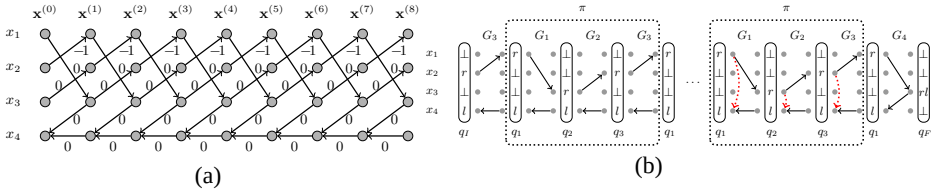

Fig. 1. (a) Unfolding of $\mathcal{G}_R$. Here $\mathbf{x}^{(i)} = \{x^{(i)} \mid x \in \mathbf{x}\}$. (b) A run of the zigzag automaton over a path in $\mathcal{G}_{R^8}$.

Given a difference bounds relation R, the *zigzag automaton* $\mathcal{A}_R$ recognizes all paths from x_i to x_j in $\mathcal{G}_R^k$. Intuitively, a path π between x_i and x_j in $\mathcal{G}_R^k$, is represented by a word w of length k, as follows: the w_l symbol represents *simultaneously* all edges of π that involve only nodes from $\mathbf{x}^{(l)} \cup \mathbf{x}^{(l+1)}$, for all $0 \leq l < k$. The alphabet of the zigzag automaton consists of subgraphs of $\mathcal{G}_R$, where the weight of a subgraph is the sum of the weights on its edges. The set of control states of the zigzag automaton is[8] $\{l, r, lr, rl, \perp\}^n$. Clearly, the size of the zigzag automaton is at most 5^n. For a complete definition, the interested reader may consult [7].

Example Consider the relation $\mathcal{R} \equiv x_2 - x_1' \leq -1 \wedge x_3 - x_2' \leq 0 \wedge x_1 - x_3' \leq 0 \wedge x_4' - x_4 \leq 0 \wedge x_3' - x_2 \leq 0$. An example of a run of $\mathcal{A}_R$ recognizing a path of constraints in $\mathcal{G}_{\mathcal{R}}^8$ is given in Figure 1 (b). The word accepted by π is a subgraph of $\mathcal{G}_{\mathcal{R}}^8$ shown in Figure 1 (a). The cycle $\pi : q_1 \xrightarrow{G_1} q_2 \xrightarrow{G_2} q_3 \xrightarrow{G_3} q_1$ is taken several times in this run. The weights of the symbols on the run are $w(G_1) = w(G_2) = w(G_4) = 0$ and $w(G_3) = -1$. □

The following lemma proves the existence of a negative weight cycle in the zigzag automata corresponding to well-founded difference bounds relation. The intuition behind this fact is that the rates of the DBM sequence $\{m_{R^i}\}_{i>0}$ are weights of optimal

[8] The intuition behind the names $\{l, r, lr, rl, \perp\}$ of components of control states is that they capture the direction of incoming and outgoing edges (l for left, r for right).

ratio (weight per length) cycles in the zigzag automaton. According to the previous section, if R is well-founded, there exists a negative rate for $\{m_{R^i}\}_{i>0}$, which implies the existence of a negative cycle in the zigzag automaton.

Lemma 7. *If R is a $*$-consistent well-founded difference bounds relation of prefix $b \geq 0$, and $\mathcal{A}_R$ is its corresponding zigzag automaton, then there exists a cycle π from a state q to itself, such that $w(\pi) < 0$ and there exists paths π_i from an initial state to q, and π_f from q to a final state, such that $|\pi_i| + |\pi_f| = b$.*

Next we prove the existence of a linear decreasing function, based on the existence of a negative weight cycle in the zigzag automaton.

Lemma 8. *If R is a $*$-consistent well-founded difference bounds relation of prefix $b \geq 0$, then there exists a linear function $f(\mathbf{x})$ such that, for all states $s, s' : \mathbf{x} \to \mathbb{Z}$ we have $\mathcal{R}^{-b}(\top)(s) \wedge \mathcal{R}(s, s') \to f(s) > f(s')$.*

Example We illustrate the construction of linear decreasing function on the relation $\mathcal{R} \equiv x_2 - x'_1 \leq -1 \ \wedge \ x_3 - x'_2 \leq 0 \ \wedge \ x_1 - x'_3 \leq 0 \ \wedge \ x'_4 - x_4 \leq 0 \ \wedge \ x'_3 - x_2 \leq 0$ from the previous example. Summing the edges in π, we obtain $x_2 - x'_1 + x_1 - x'_3 + x_3 - x'_2 + x'_4 - x_4 + x'_4 - x_4 + x'_4 - x_4 \leq -1$, which simplifies to $x_1 + x_2 + x_3 - 3x_4 - (x'_1 + x'_2 + x'_3 - 3x_4) \leq -1$. Letting $f(\mathbf{x}) = -(x_1 + x_2 + x_3 - 3x_4)$, we have that $f(\mathbf{x}) > f(\mathbf{x}')$. □

Last, we prove that the function f of Lemma 8 is bounded from below, concluding that it is indeed a ranking function. Since each run in the zigzag automaton recognizes a path from some x_i to some x_j, a run that repeats a cycle can be decomposed into a prefix, the cycle itself and a suffix. The path recognized may traverse the cycle several times, however each exit point from the cycle must match a subsequent entry point. These paths from the exit to the corresponding entries gives us the necessary lower bound. In fact, these paths appear already on graphs $\mathcal{G}_{R^i}$ for $i \geq b$, where b is the prefix of R (Lemma 9). Hence the need for a strenghtened witness $\mathcal{R}^{-b}(\top) \wedge \mathcal{R}$, as $\mathcal{R}$ alone is not enough for proving boundedness of f.

Lemma 9. *Let R be a $*$-consistent octagonal relation with prefix b and period c. Then, for any $1 \leq i, j \leq 2n$ and $k \geq 1$, we have $(m_{\mathcal{R}^{-k}(\top)})_{i,j} < \infty \to (m_{\mathcal{R}^{-b}(\top)})_{i,j} < \infty$.*

Lemma 10. *If R is a $*$-consistent well-founded difference bounds relation of prefix b, and $f(\mathbf{x})$ is the linear decreasing function from Lemma 8, there exists an integer h such that, for all states $s, s' : \mathbf{x} \to \mathbb{Z}$, $(\mathcal{R}^b(\top)(s) \ \wedge \ \mathcal{R}(s, s')) \to (f(s) \geq h \ \wedge \ f(s') \geq h)$.*

Example (continued) We will continue the previous example and illustrate the boundedness of $f = -(x_1 + x_2 + x_3 - 3x_4)$ (see Figure 1b). Since there is a path from $\mathbf{x}_2^{(6)}$ to $\mathbf{x}_4^{(6)}$ in G_3G_4 (and hence in $\mathcal{G}_R^2$), then $\mathcal{R}^2 \to (x_2 - x_4 \leq -1)$, and by Lemma 9, we obtain $\mathcal{R}^b \to (x_2 - x_4 \leq -1)$. Similarly, since there is a path $\mathbf{x}_3^{(5)} \rightsquigarrow \mathbf{x}_4^{(5)}$ in $G_2G_3G_4$ (and hence in $\mathcal{G}_R^3$), we obtain $\mathcal{R}^b \to (x_3 - x_4 \leq -1)$. Similarly, since there is a path $\mathbf{x}_1^{(4)} \rightsquigarrow \mathbf{x}_4^{(4)}$ in $G_1G_2G_3G_4$ (and hence in $\mathcal{G}_R^4$), we obtain $\mathcal{R}^b \to (x_3 - x_4 \leq -1)$. Summing up these inequalities, we obtain that $f(\mathbf{x}) = -(x_1 + x_2 + x_3 - 3x_4) \geq 3$ and, thus $\mathcal{R}^{-b}(\top) \wedge \mathcal{R} \to (f \geq 3)$.

As an experiment, we have tried the RANKFINDER [18] tool (complete for linear ranking functions), which failed to discover a ranking function on this example. This comes with no surprise, since no linear decreasing function that is bounded after the first iteration exists. □

5 Linear Affine Relations

Let $\mathbf{x} = \langle x_1, \ldots, x_n \rangle^\top$ be a column vector of variables ranging over integers. A linear affine relation is a relation of the form $\mathcal{R}(\mathbf{x}, \mathbf{x}') \equiv C\mathbf{x} \geq \mathbf{d} \ \wedge \ \mathbf{x}' = A\mathbf{x} + \mathbf{b}$, where $A \in \mathbb{Z}^{n \times n}$, $C \in \mathbb{Z}^{p \times n}$ are matrices, and $\mathbf{b} \in \mathbb{Z}^n$, $\mathbf{d} \in \mathbb{Z}^p$ are column vectors of integer constants. Notice that we consider linear affine relations to be deterministic, unlike the difference bounds and octagonal relations considered in the previous. In the following, it is convenient to work with the equivalent homogeneous form:

$$\begin{gathered} \mathcal{R}(\mathbf{x}, \mathbf{x}') \equiv C_h \mathbf{x}_h \geq \mathbf{0} \ \wedge \ \mathbf{x}'_h = A_h \mathbf{x}_h \\ A_h = \begin{pmatrix} A & \mathbf{b} \\ 0 & 1 \end{pmatrix} \ C_h = \begin{pmatrix} C & -\mathbf{d} \end{pmatrix} \ \mathbf{x}_h = \begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix} \end{gathered} \tag{7}$$

The closed form of a linear affine relation is defined by the following formula:

$$\mathcal{R}^{(k)}(\mathbf{x}, \mathbf{x}') \ \equiv \ \exists x_{n+1}, x'_{n+1}. \mathbf{x}'_h = A_h^k \mathbf{x}_h \wedge \forall 0 \leq \ell < k. C A_h^\ell \mathbf{x} \geq \mathbf{0} \wedge x_{n+1} = 1 \tag{8}$$

Intuitively, the first conjunct defines the (unique) outcome of iterating the relation $\mathbf{x}' = A\mathbf{x} + \mathbf{b}$ for k steps, while the second (universally quantified) conjunct ensures that the condition ($C\mathbf{x} \geq \mathbf{d}$) has been always satisfied all along the way. The definition of the weakest recursive set of a linear affine relation is (after the elimination of the trailing existential quantifier):

$$wrs(R)(\mathbf{x}) \ \equiv \ \exists x_{n+1} \forall k > 0 \ . \ C_h A_h^k \mathbf{x} \geq \mathbf{0} \ \wedge \ x_{n+1} = 1 \tag{9}$$

The main difficulty with the form (9) comes from the fact that the powers of a matrix A cannot usually be defined in a known decidable theory of arithmetic. In the following, we discuss the case of A having the finite monoid property [2, 25], which leads to $wrs(R)$ being Presburger definable. Further, we relax the finite monoid condition and describe a method for generating sufficient termination conditions, i.e. sets $S \in \mathbb{Z}^n$ such that $S \cap wrs(R) = \emptyset$.

Some basic notions of linear algebra are needed in the following. If $A \in \mathbb{Z}^{n \times n}$ is a square matrix, and $\mathbf{v} \in \mathbb{Z}^n$ is a column vector of integer constants, then any complex number $\lambda \in \mathbb{C}$ such that $A\mathbf{v} = \lambda\mathbf{v}$, for some complex vector $\mathbf{v}\mathbb{C}^n$, is called an *eigenvalue* of A. The vector $\mathbf{v}$ in this case is called an *eigenvector* of A. It is known that the eigenvalues of A are the roots of the *characteristic polynomial* $\det(A - \lambda I_n) = 0$, which is an effectively computable univariate polynomial in λ. A complex number r is said to be a *root of the unity* if $r^d = 1$ for some integer $d > 0$.

In the previous work of Weber and Seidl [25], Boigelot [2], and Finkel and Leroux [12], a restriction of linear affine relations has been introduced, with the goal of defining the closed form of relations in Presburger arithmetic. A matrix $A \in \mathbb{Z}^{n \times n}$ is said to have the *finite monoid property* if and only if its set of powers $\{A^i \mid i \geq 0\}$ is finite. A linear affine relation has the finite monoid property if and only if the matrix A defining the update has the finite monoid property.

Lemma 11 ([12, 2]). *A matrix $A \in \mathbb{Z}^{n \times n}$ has the finite monoid property iff:*

1. *all eigenvalues of A are either zero or roots of the unity, and*
2. *all non-zero eigenvalues are of multiplicity one.*

Both conditions are decidable.

In the following, we drop the second requirement of Lemma 11, and consider only linear relations, such that all non-zero eigenvalues of A are roots of the unity. In this case, $\mathcal{R}^{(k)}$ cannot be defined in Presburger arithmetic any longer, thus we renounce defining $wrs(R)$ precisely, and content ourselves with the discovery of *sufficient conditions for termination*. Basically given a linear affine relation R, we aim at finding a disjunction $\phi(\mathbf{x})$ of linear constraints on $\mathbf{x}$, such that $\phi \wedge wrs(R)$ is inconsistent, without explicitly computing $wrs(R)$.

Lemma 12. *Given a square matrix $A \in \mathbb{Z}^{n\times n}$, whose non-zero eigenvalues are all roots of the unity. Then $(A^m)_{i,j} \in \mathbb{Q}[m]$, for all $1 \leq i, j \leq n$, are effectively computable polynomials with rational coefficients.*

We turn now back to the problem of defining $wrs(R)$ for linear affine relations R of the form (9). First notice that, if all non-zero eigenvalues of A are roots of the unity, then the same holds for A_h (7). By Lemma 12, one can find rational polynomials $p_{i,j}(k)$ defining $(A_h^k)_{i,j}$, for all $1 \leq i, j \leq n$. The condition (9) resumes to a conjunction of the form:

$$wrs(R)(\mathbf{x}) \equiv \bigwedge_{i=1}^{n} \forall k > 0 \,.\, P_i(k, \mathbf{x}) \geq 0 \tag{10}$$

where each $P_i = a_{i,d}(\mathbf{x}) \cdot k^d + \ldots + a_{i,1}(\mathbf{x}) \cdot k + a_{i,0}(\mathbf{x})$ is a polynomial in k whose coefficients are the linear combinations $a_{i,d} \in \mathbb{Q}[\mathbf{x}]$. We are looking after a sufficient condition for termination, which is, in this case, any set of valuations of $\mathbf{x}$ that would invalidate (10). The following proposition gives sufficient invalidating clauses for each conjunct above. By taking the disjunction of all these clauses we obtain a sufficient termination condition for R.

Lemma 13. *Given a polynomial $P(k, \mathbf{x}) = a_d(\mathbf{x}) \cdot k^d + \ldots + a_1(\mathbf{x}) \cdot k + a_0(\mathbf{x})$, there exists $n > 0$ such that $P(n, \mathbf{x}) < 0$ if, for some $i = 0, 1, \ldots, d$, we have $a_{d-i}(\mathbf{x}) < 0$ and $a_d(\mathbf{x}) = a_{d-1}(\mathbf{x}) = \ldots = a_{d-i+1}(\mathbf{x}) = 0$.*

Example Consider the following program [10], and its linear transformation matrix A.

$$\begin{array}{l}\text{while } (x \geq 0)\\ \quad x' = x + y\\ \quad y' = y + z\end{array} \qquad A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad A^k = \begin{pmatrix} 1 & k & \frac{k(k-1)}{2} \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}$$

The characteristic polynomial of A is $\det(A - \lambda I_3) = (1 - \lambda)^3$, hence the only eigenvalue is 1, with multiplicity 3. Then we compute A^k (see above), and $x' = x + k \cdot y + \frac{k(k-1)}{2} z$ gives the value of x after k iterations of the loop. Hence the (precise) non-termination condition is: $\forall k > 0 \,.\, \frac{z}{2} \cdot k^2 + (y - \frac{z}{2}) \cdot k + x \geq 0$ The sufficient condition for termination is: $(z < 0) \vee (z = 0 \wedge y < 0) \vee (z = 0 \wedge y = 0 \wedge x < 0)$ □

We can generalize this method further to the case where all eigenvalues of A are of the form $q \cdot r$, with $q \in \mathbb{R}$ and $r \in \mathbb{C}$ being a root of the unity. The main reason for not using this condition from the beginning is that we are, to this point, unaware of its decidability status. With this condition instead, it is sufficient to consider only the eigenvalues with the maximal absolute value, and the polynomials obtained as sums of the polynomial coefficients of these eigenvalues (see Corollary 2 in Appendix 5). The result of Lemma 12 and the sufficient condition of Lemma 13 carry over when using these polynomials instead.

6 Experimental Evaluation

We have validated the methods described in this paper by automatically verifying termination of all the octagonal running examples, and of several integer programs synthesized from (i) programs with lists [3] and (ii) VHDL models [21]. We have first computed automatically their precise transition invariant $\mathcal{T}$ by adapting the method for reachability analysis for counter automata, described in [6], and implemented in the

FLATA tool [13]. Then we automatically proved that $\mathcal{T}$ is contained in a disjunction of octagonal relations, which are found to be well-founded by the procedure described in Section 4.

We first verified termination of the LISTCOUNTER and LISTREVERSAL programs, which were obtained using a translation scheme from [3], which generates an integer program from a program manipulating dynamically allocated single-selector linked lists. Using the same technique, we also verified the COUNTER and SYNLIFO programs, obtained by translating VHDL designs of hardware counter and synchronous LIFO [21]. These models have infinite runs for any input values, which is to be expected, as they encode the behavior of synchronous reactive circuits.

Second, we compared (Table 1) our method for termination of linear affine loops with the examples given in [10], and found the same termination preconditions as they do, with one exception, in which we can prove universal termination in integer input values (row 3 of Table 1). The last example from [10] is the Euclidean Greatest Common Divisor algorithm, for which we infer automatically the correct termination preconditions using a disjunctively well-founded octagonal abstraction of the transition invariant.

PROGRAM	COOK ET. AL [10]	LINEAR AFFINE LOOPS

if (lvar $\geq$ 0)
while (lvar $< 2^{30}$)
lvar = lvar $\ll$ 1;
	$lvar > 0 \vee lvar < 0 \vee lvar \geq 2^{30}$	$\frac{lvar^{(k)}=2^k \cdot lvar^{(0)}}{\neg(lvar=0)\vee lvar\geq 2^{30}}$

while (x $\leq$ N)
if (*) {
x=2*x+y;
y=y+1;
} else
x ++;
	$x > N \vee x + y \geq 0$	$\frac{\begin{array}{c}x^{(k)}\geq 2^k\cdot(x^{(0)}+y^{(0)}+1)-k-y^{(0)}-1 \\ y^{(k)}\leq y^{(0)}+k\end{array}}{x>N\vee x+y\geq 0}$

while (x $\geq$ N) x = -2*x + 10;	$x > 5 \vee x + y \geq 0$	$\frac{x^{(k)}=(-2)^k\cdot(x^{(0)}-\frac{10}{3})+\frac{10}{3}}{x \neq \frac{10}{3} \iff \text{true}}$

//@ requires $n > 200$
x = 0;
while (1)
if (x < n) {
x=x+y;
if (x $\geq$ 200) break; }
	$y > 0$	$\frac{x^{(k)}=x^{(0)}+k\cdot y^{(0)}}{y>0}$

Table 1. Termination preconditions for several program fragments from [10]. The even rows of column 3 represent the closed form of the transition relation, while the rows give the termination preconditions.

7 Conclusions

We have presented several methods for deciding conditional termination of several classes of program loops manipulating integer variables. The universal termination problem has been found to be decidable for octagonal relations and linear affine loops with the finite monoid property. In other cases of linear affine loops, we give sufficient termination conditions. We have implemented our method in the FLATA tool [13] and performed a number of preliminary experiments.

References

1. R. Bagnara, P. M. Hill, and E. Zaffanella. An improved tight closure algorithm for integer octagonal constraints. In *VMCAI'08*, pages 8–21, 2008.
2. B. Boigelot. *Symbolic Methods for Exploring Infinite State Spaces*, volume PhD Thesis, Vol. 189. Collection des Publications de l'Université de Liège, 1999.
3. A. Bouajjani, M. Bozga, P. Habermehl, R. Iosif, P. Moro, and T. Vojnar. Programs with lists are counter automata. In *CAV'06*, pages 517–531, 2006.
4. M. Bozga, C. Gîrlea, and R. Iosif. Iterating octagons. In *TACAS'09*, pages 337–351, 2009.

5. M. Bozga, R. Iosif, and F. Konečný. Fast acceleration of ultimately periodic relations. In *CAV'10*, pages 227–242, 2010.
6. M. Bozga, R. Iosif, and F. Konečný. Transitive Closures Ultimately Periodic Relations. Technical Report TR-2011-14, Verimag, Grenoble, France, 2011.
7. M. Bozga, R. Iosif, and Y. Lakhnech. Flat parametric counter automata. *Fundamenta Informaticae*, 91:275–303, 2009.
8. A. R. Bradley, Z. Manna, and H. B. Sipma. Linear ranking with reachability. In *CAV'05*, pages 491–504, 2005.
9. M. Braverman. Termination of integer linear programs. In *CAV'06*, pages 372–385, 2006.
10. B. Cook, S. Gulwani, T. Lev-Ami, A. Rybalchenko, and M. Sagiv. Proving conditional termination. In *CAV'08*, pages 328–340, 2008.
11. G. Everest. *Recurrence sequences*. American Mathematical Soc., 2003.
12. A. Finkel and J. Leroux. How to compose presburger-accelerations: Applications to broadcast protocols. In *FST TCS'02*, pages 145–156, 2002.
13. http://www-verimag.imag.fr/FLATA.html.
14. T. Gawlitza and H. Seidl. Precise fixpoint computation through strategy iteration. In *ESOP'07*, pages 300–315, 2007.
15. V. Halava, T. Harju, M. Hirvensalo, and J. Karhumaki. Skolem's problem – on the border between decidability and undecidability, 2005.
16. R. Iosif and A. Rogalewicz. Automata-based termination proofs. In *CIAA'09*, pages 165–177, 2009.
17. A. Miné. The octagon abstract domain. *Higher-Order and Symbolic Computation*, 19(1):31–100, 2006.
18. A. Podelski and A. Rybalchenko. A complete method for the synthesis of linear ranking functions. In *VMCAI'04*, 2004.
19. A. Podelski and A. Rybalchenko. Transition invariants. In *LICS'04*, pages 32–41, 2004.
20. B. De Schutter. On the ultimate behavior of the sequence of consecutive powers of a matrix in the max-plus algebra. *Linear Algebra and its Applications*, 307:103–117, 2000.
21. A. Smrcka and T. Vojnar. Verifying parametrised hardware designs via counter automata. In *HVC'07*, pages 51–68, 2007.
22. K. Sohn and A. Van Gelder. Termination detection in logic programs using argument sizes. In *PODS'91*, 1991.
23. A. Tiwari. Termination of linear programs. In *CAV'04*, pages 70–82, 2004.
24. A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42:230–265, 1936.
25. A. Weber and H. Seidl. On finitely generated monoids of matrices with entries in n. *ITA'91*, pages 19–38, 1991.

A Proofs from Section 2

Proof of Proposition 1 For each i we have $S_i \subseteq pre_R(S_i) \subseteq pre_R(\bigcup_{j=0,1,\ldots} S_j)$. The last inclusion is by the monotonicity of pre_R. Hence $\bigcup_{j=0,1,\ldots} S_j \subseteq pre_R(\bigcup_{j=0,1,\ldots} S_j)$. □

Proof of Lemma 1 By the Knaster-Tarski Fixpoint Theorem, $\text{gfp}(pre_R) = \bigcup\{S \mid S \subseteq pre_R(S)\} = wrs(R)$. □

Proof of Lemma 2 "$wrs(R) \subseteq wnt(R)$" Let $s_0 \in wrs(R)$ be a state. Then there exists $s_1 \in wrs(R)$ such that $\mathcal{R}(s_0, s_1)$. Applying this argument infinitely many times, one can construct an infinite sequence $s_0, s_1, s_2, \ldots$ such that $\mathcal{R}(s_i, s_{i+1})$, for all $i \geq 0$. Hence $s_0 \in wnt(R)$.

"$wnt(R) \subseteq wrs(R)$" Let $s_0 \in wnt(R)$. Then there exists an infinite sequence $s_0, s_1, s_2, \ldots$ such that $\mathcal{R}(s_i, s_{i+1})$ for all $i \geq 0$. Then, for all $i \geq 0$, $s_0 \in pre^i_R(s_i) \subseteq pre^i_R(\mathcal{D}^n)$, by monotonicity of pre_R. Hence $s_0 \in \bigcap_{i \geq 0} pre^i_R(\mathcal{D}^n) = \text{gfp}(pre_R)$. By Lemma 1, $s_0 \in wrs(R)$. □

B Proofs from Section 3

Proof of Lemma 3 The algorithm of Gawlitza and Seidl [14] runs in time at most $\mathcal{O}(n \cdot |\mathcal{S}|)$, where n is the number of variables in the system, and $|\mathcal{S}|$ denotes the size of the system i.e., the sum of the expression sizes of the right-hand sides. In our case, the min-terms in the first system (2) are of size at most $\mathcal{O}(n)$, while in the second system (3), we have min-terms of size at most $\mathcal{O}(2^n)$. Since the number of variables in both systems is n^2, the result follows. □

C Proofs from Section 4

Proof of Lemma 4 Let b and c be the prefix and period of R, respectively. If Λ_0 has at least one negative entry, then $wrs(R) = \emptyset$. Otherwise, $wrs(R) \equiv \mathcal{R}^{-b}(\top)$, and we can compute it by composing R with itself b times. Each composition requires a quantifier elimination which takes at most $\mathcal{O}(n^3)$. Since there are exactly b quantifier eliminations, we have the result. For the second part, $b + c$ is the length of a path in the zigzag automaton $\mathcal{A}_R$, with no repeated states. Hence $b + c$ is bounded by the number of states in the zigzag automaton $\mathcal{A}_R$, which is at most 5^n. Also, computing Λ_0 of R requires iterating R up to $b + c$. □

Proof of Lemma 5 "$\Rightarrow$" By the fact that $\mathcal{R} \leftarrow \mathcal{R}^{-m}(\top) \wedge \mathcal{R}$ and the monotonicity of wrs. "$\Leftarrow$" We prove the dual. Assume that $wrs(R) \neq \emptyset$ i.e., there exists an infinite sequence $s_1, s_2, \ldots$ such that $\mathcal{R}(s_i, s_{i+1})$, for all $i > 0$. Then all s_i belong to the set defined by $\mathcal{R}^{-m}(\top)$, hence $s_1, s_2, \ldots$ is an infinite sequence for the relation defined by $\mathcal{R}^{-m}(\top) \wedge \mathcal{R}$ as well. □

Proposition 3. *Let $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$ be sequences. Then the following hold:*

- *if* $\inf\{\lfloor \frac{a_n}{2} \rfloor\}_{n=0}^{\infty} = -\infty$ *then* $\inf\{a_n\}_{n=0}^{\infty} = -\infty$

– *if* $\inf\{a_n + n_b\}_{n=0}^{\infty} = -\infty$ *then* $\inf\{a_n\}_{n=0}^{\infty} = -\infty$ *or* $\inf\{b_n\}_{n=0}^{\infty} = -\infty$
– *if* $\inf\{\min(a_n, b_n)\}_{n=0}^{\infty} = -\infty$ *then* $\inf\{a_n\}_{n=0}^{\infty} = -\infty$ *or* $\inf\{b_n\}_{n=0}^{\infty} = -\infty$

Proof. By contraposition. Suppose that $\inf \mathcal{S}_1 \neq -\infty$ and $\inf \mathcal{S}_2 \neq -\infty$. Then, $\exists k_1 \geq 1 \,.\, \forall l_1 \geq k_1 \,.\, a_{k_1} = a_{l_1}$ and $\exists k_2 \geq 1 \,.\, \forall l_2 \geq k_2 \,.\, b_{k_2} = b_{l_2}$. Let $k = max\{k_1, k_2\}$. Then clearly, $\forall l \geq k \,.\, \left\lfloor \frac{a_k}{2} \right\rfloor = \left\lfloor \frac{a_l}{2} \right\rfloor \wedge a_k + b_k = a_l + b_l \wedge min\{a_k, b_k\} = min\{a_l, b_l\}$. Hence $\inf \mathcal{S}_t \neq -\infty$, $\inf \mathcal{S}_p \neq -\infty$, and $\inf \mathcal{S}_m \neq -\infty$. □

Proof of Lemma 6 Let b and c be the prefix and period of R. In Section 4 we proved that a $*$-consistent octagon R is well-founded if and only if the closed form of the sequence $\{\mathcal{R}^{b+c\ell}\}_{\ell \geq 0}$ contains an atomic proposition of the form $u \leq a_u \ell + d_u$, where $u \in U(\mathbf{x})$ is an octagonal term, and $a_u < 0$. We will show that the same holds if we use the closed form of the sequence $\{\overline{\mathcal{R}}^{b+c\ell}\}_{\ell \geq 0}$ instead. Notice that for any $k \geq 0$, the difference bounds encoding of $\mathcal{R}^{-k}(\top)$ is the projection of $m^t_{\mathcal{R}^k}$ on the entries corresponding to unprimed variables i.e., $(m^t_{\mathcal{R}^k})_{\downarrow \mathbf{y}}$. By the monotonicity of pre_R, the sequence $\{(m^t_{\mathcal{R}^k})_{\downarrow \mathbf{y}}\}_{k \geq 0}$ is decreasing. Since the elements of the sequence are defined by (5), we can apply Proposition 3 and observe that if for some $1 \leq i, j \leq 2n$, $\inf\{(m^t_{\mathcal{R}^k})_{i,j}\}_{k \geq 0} = -\infty$, then also $\inf\{(m^*_{\overline{\mathcal{R}}^k})_{i_2,j_2}\}_{k \geq 0} = -\infty$ for some $(i_2, j_2) \in \{(i, j), (i, \bar{i}), (\bar{j}, j)\}$. Hence R is well-founded iff there exists a negative coefficient a_u in the closed form of $\{\overline{\mathcal{R}}^{b+c\ell}\}_{\ell \geq 0}$ iff R_{db} is well founded. □

Definition 11. *The sequence $\{g_k\}_{k=0}^{\infty}$ is ultimately geometric if there exist $b \in \mathbb{N}_0$, $c \in \mathbb{N}$, and $\lambda \in \mathbb{Q}_\infty$ such that*

$$\forall k \geq 0, \forall s \in \{0, \ldots, c-1\} \,.\, g_{b+s+(k+1)c} = \lambda + g_{b+s+kc}$$

Definition 12. *The sequence $\{g_k\}_{k=0}^{\infty}$ is ultimately periodic if there exist $b \in \mathbb{N}_0$, $c \in \mathbb{N}$, and $\lambda_0, \ldots, \lambda_{c-1} \in \mathbb{Q}_\infty$ such that*

$$\forall k \geq 0, \forall s \in \{0, \ldots, c-1\} \,.\, g_{b+s+(k+1)c} = \lambda_s + g_{b+s+kc}$$

For the sake of completeness, we present key results of [20]. Let $G = (V, E, \nu : E \to \mathbb{Z})$ be a weighted digraph and M_G the associated incidence matrix. Let $\mathcal{G}(V'), V' \subseteq V$ be a subgraph induced by V'. We say that $\mathcal{G}(V')$ is strongly connected if for any two *different* vertices $v_1, v_2 \in V', v_1 \neq v_2$ there exists a path from v_1 to v_2. $\mathcal{G}(V')$ is a strongly connected component of G if there is no $V' \subset V'' \subseteq V$ such that $\mathcal{G}(V'')$ is strongly connected.

Given a path $\pi : v_0 \xrightarrow{c_1} v_1 \xrightarrow{c_2} v_2 \ldots v_{p-1} \xrightarrow{c_p} v_p$, the length of π is $|\pi| = p$, the weight of π is $w(\pi) = \sum_{i=1}^{p} c_i$, average weight of π is $\frac{w(\pi)}{|\pi|}$. A cycle is a path where $v_0 = v_p$. A cycle of a strongly connected graph G is critical if it has maximum average weight among all cycles of G.

A cycle of a strongly connected graph G is critical if it has minimum average weight and cyclicity of G is the greatest common divisor of lengths of critical cycles in G.

Theorem 2. *(Theorem 2.4 in [20]) Let G be a strongly connected digraph, c its cyclicity and λ the minimum average weight of critical cycles in G. Then, M_G is ultimately geometric with period c and rate Λ, where $\Lambda_{i,j} = c\lambda$ for all i, j.*

Theorem 3. *(Corollary of Theorem 3.3 in [20]) Let G be a digraph and let $\{c_1, \ldots, c_m\}$ and $\{\lambda_1, \ldots, \lambda_m\}$ be cyclicities and minimum average weights of critical cycles of strongly connected components of G. Then, M_G is ultimately periodic with period $c = lcm(c_1, \ldots, c_m)$ and rates $\{\Lambda'_0, \ldots, \Lambda'_{c-1}\}$, where $(\Lambda'_k)_{i,j} \in \{c\lambda_1, \ldots, c\lambda_m\}$ for all $0 \leq k < c$ and for all i, j.*

Corollary 1. *For all $v_1, v_2 \in V$ s.t. $(\Lambda_0)_{v_1,v_2} \neq \infty$ there exists a critical cycle $\pi : v \rightsquigarrow v$ of length p and paths $\pi_i : v_1 \rightsquigarrow v$ and $\pi_f : v \rightsquigarrow v_2$ s.t. $|\pi_i| + |\pi_f| = b$ s.t.*

$$\forall k \geq 0 \,.\, (M_G^{b+pk})_{v_1,v_2} = w(\pi_i) + kw(\pi) + w(\pi_f)$$

Notice that zigzag automaton can be viewed as a digraph and hence Theorems 2 and 3 apply to them. This means that difference bounds relations are ultimately periodic. Let $b, c, \Lambda_0, \ldots, \Lambda_{c-1}$ be the prefix, period, and rates of R. Then, $b+2, c, \Lambda_0, \ldots, \Lambda_{c-1}$ are the prefix, period, and rates of $\mathcal{M}_R$, the incidence matrix of a zigzag automaton $\mathcal{A}_\mathcal{R}$. Moreover, the closed form of $\{\mathcal{R}^{b+c\ell}\}_{\ell \geq 0}$ is

$$\mathcal{R}^{(l)}_{b,c} \equiv \bigwedge_{i \neq j} x_i - x_j \leq (\Lambda_0)_{I_{i,j},F_{i,j}}\ell + (M_{\mathcal{R}^b})_{i,j} \tag{11}$$

or, equivalently

$$\mathcal{R}^{(l)}_{b,c} \equiv \bigwedge_{i \neq j} x_i - x_j \leq (\Lambda_0)_{I_{i,j},F_{i,j}}\ell + (\mathcal{M}^{b+2}_\mathcal{R})_{I_{i,j},F_{i,j}} \tag{12}$$

Proof of Lemma 7 By the decision procedure in Section 4, if $\mathcal{R}$ is $*$-consistent and well-founded, then the closed form $\varphi^{(l)}_r$ of $\{\mathcal{R}^{b+c\ell}\}_{\ell \geq 0}$ contains an atomic formula $x_i - x_j \leq a\ell + d$ where $a, d \in \mathbb{Z}$, $a < 0$. By the Equation (11), $a = (\Lambda_0)_{I_{i,j},F_{i,j}}$. By Theorems 2 and 3, $(\Lambda_0)_{I_{i,j},F_{i,j}}$ is a c-multiple of the average weight of some critical cycle in some SCC of $\mathcal{A}_\mathcal{R}$. One of these cycles is π of length p from Corollary 1 (since $b + 2$ in this lemma corresponds to b in Corollary 1, due to special initial and final transitions in $\mathcal{A}_\mathcal{R}$). $a < 0$ implies $(\Lambda_0)_{I_{i,j},F_{i,j}} < 0$, which in turn implies $w(\pi) = \frac{p}{c}(\Lambda_0)_{I_{i,j},F_{i,j}} < 0$. Other properties stated in this lemma follow from Corollary 1 for special case $k = 1$. □

Proof of Lemma 8 By Lemma 7, there is a negative critical cycle π of length p in the zigzag automaton: $q_1 \xrightarrow{G_1} q_2 \ldots q_p \xrightarrow{G_p} q_1$. Let $G_j = (\mathbf{x} \cup \mathbf{x}', E_j)$ for all $1 \leq j \leq p$.

Consider the following sum of all constraints represented by edges appearing in the zigzag cycle (note that the sum of weights of these edges equals $w(\pi)$):

$$\sum_{1 \leq j \leq p} \left(\sum_{(x_i \rightarrow x'_j) \in E_j} (x_i - x'_j) + \sum_{(x'_i \rightarrow x_j) \in E_j} (x'_i - x_j) \right) \leq w(\pi) \tag{13}$$

The left-hand side of (13) can be written equivalently as

$$\sum_{1 \leq j \leq p} \left(\sum_{\substack{1 \leq i \leq n, \\ (q_j)_i = r}} (x_i - x'_i) + \sum_{\substack{1 \leq i \leq n, \\ (q_j)_i = l}} (-x_i + x'_i) + \sum_{\substack{1 \leq i \leq n, \\ (q_j)_i = lr}} (-x_i + x_i) + \sum_{\substack{1 \leq i \leq n, \\ (q_j)_i = rl}} (-x'_i + x'_i) \right) \tag{14}$$

and thus, after simplifications ($-x_i+x_i=0, -x'_i+x'_i=0$), (13) can be written equivalently as

$$\sum_{1\leq j\leq p}\left(\sum_{\substack{1\leq i\leq n,\\ (q_j)_i=r}}(x_i-x'_i)+\sum_{\substack{1\leq i\leq n,\\ (q_j)_i=l}}(-x_i+x'_i)\right)\leq w(\pi) \tag{15}$$

Let f denote the negated sum of all unprimed terms in (14) and f' denote the negated sum of all primed terms in (14). Then, clearly $f'=-f[\mathbf{x}'/\mathbf{x}]$. Thus, (15) can be written as

$$f'-f\leq w(\pi) \tag{16}$$

Notice that since $w(\pi)<0$, we establish that $f'-f<0$ hence f is strictly decreasing. Since, for all states s,s' we have $\mathcal{R}(s,s')\rightarrow f(s)>f'(s)$, we have that $\mathcal{R}^{-b}(\top)(s)\wedge\mathcal{R}(s,s')\rightarrow f(s)>f'(s)$. □

Proof of Lemma 9 (Case $1\leq k\leq b$) By monotonicity of pre_R, $(m_{\mathcal{R}^{-k}(\top)})_{i,j}\geq(m_{\mathcal{R}^{-b}(\top)})_{i,j}$. Thus if $(m_{\mathcal{R}^{-k}(\top)})_{i,j}\neq\infty$, then clearly $(m_{\mathcal{R}^{-b}(\top)})_{i,j}<\infty$.
(Case $k>b$) Let $p=\lceil\frac{k-b}{c}\rceil$, and $k'=b+pc$. Note that $\mathcal{R}^{k'}=\mathcal{R}^{(\ell)}_{b,c}[p/\ell]$. Since $k'\geq k$, by the same argument as for case ($1\leq k\leq b$), $(m_{\mathcal{R}^{-k'}(\top)})_{i,j}<\infty$. Since $\mathcal{R}^{(\ell)}_{b,c}$ is closed, $y_i-y_j\leq a\ell+d$, where $a,d\in\mathbb{Z}$, is one of its conjuncts. Since $\mathcal{R}^{(\ell)}_{b,c}$ is closed, $(m_{\mathcal{R}^{-b}(\top)})_{i,j}=d\neq\infty$. □

Proof of Lemma 10 Let f be a linear decreasing function from Lemma 8. Let $\pi: q_1\xrightarrow{G_1}q_2\ldots q_p\xrightarrow{G_p}q_1$ be the negative cycle used to construct f, and π_f be the suffix from Lemma 8. By construction of the zigzag automaton, for any $1\leq j\leq p$,

$$|\{i\mid(q_j)_i=r\}|=|\{i\mid(q_j)_i=l\}|$$

It follows from (15) that each $(q_j)_i=r$ contributes to f with a term $-x_i$ and that each $(q_j)_i=l$ contributes to f with a term $+x_i$ and that each $(q_j)_i\notin\{r,l\}$ doesn't contribute at all. We now demonstrate that for each $1\leq j\leq p$, there exists a bijective matching $\beta_j:\{i\mid(q_j)_i=r\}\rightarrow\{i\mid(q_j)_i=l\}$ such that for any $1\leq i_1\leq n$ s.t. $\beta_j(i_1)=i_2$, the difference $x_{i_2}-x_{i_1}$ is bounded in $\mathcal{R}^{-b}(\top)\wedge\mathcal{R}$, formally $\mathcal{R}^{-b}(\top)\wedge\mathcal{R}\rightarrow(x_{i_2}-x_{i_1}\geq h)$ for some $h\in\mathbb{Z}$.

Let $j\in\{1,\ldots,p\}$. By construction of the zigzag automaton, the concatenated graph $G_jG_{j+1}\ldots G_p\pi_f$ connects each $(q_j)_{i_1}$ s.t. $(q_j)_{i_1}=r$ with a unique $(q_j)_{i_2}$ s.t. $(q_j)_{i_2}=l$. This induces the required bijection β_j. Since $G_jG_{j+1}\ldots G_p\pi_f$ is a subgraph of $\mathcal{G}_R^{p+|\pi_f|}$, it follows that there is a path $\mathbf{x}_{i_1}^{(0)}\rightsquigarrow\mathbf{x}_{i_2}^{(0)}$ in $\mathcal{G}_R^{p+|\pi_f|}$, in other words, $\mathcal{R}^{p+|\pi_f|}\rightarrow x_{i_1}-x_{i_2}\leq h$ for some $h\in\mathbb{Z}$. By Lemma 9, $\mathcal{R}^b\rightarrow x_{i_1}-x_{i_2}\leq h'$ for some $h'\in\mathbb{Z}$ too.

Clearly, $\mathcal{R}^{-b}(\top)\wedge\mathcal{R}\rightarrow x_{i_1}-x_{i_2}\leq h'$ too. Since $x_{i_1}-x_{i_2}\leq h'$ if and only if $x_{i_2}-x_{i_1}\geq -h'$, we obtain the required property.

Now since $f=\sum_{1\leq j\leq p}\sum_{\substack{1\leq i_1,i_2\leq n\\ \beta_j(i_1)=i_2}}(x_{i_2}-x_{i_1})$ and since we proved that each of the differences $x_{i_2}-x_{i_1}$ in the sum is bounded in $\mathcal{R}^{-b}(\top)\wedge\mathcal{R}$, it follows that f is bounded in $\mathcal{R}^{-b}(\top)\wedge\mathcal{R}$ too, formally $\mathcal{R}^{-b}(\top)\wedge\mathcal{R}\rightarrow(f\geq h)$ for some $h\in\mathbb{Z}$. □

D Proofs from Section 5

Definition 13. *A function $f : \mathbb{N} \to \mathbb{C}$ is said to be a* C-finite recurrence *if and only if:*

$$f(m+d) = a_{d-1}f(m+d-1) + \ldots + a_1 f(m+1) + a_0 f(n),\ \forall m \geq 0$$

for some $d \in \mathbb{N}$ and $a_0, a_1, \ldots, a_{d-1} \in \mathbb{C}$, with $a_{d-1} \neq 0$. The polynomial $x^d - a_{d-1}x^{d-1} - \ldots a_1 x - a_0$ is called the characteristic polynomial *of f.*

A C-finite recurrence always admits a closed form.

Theorem 4 ([11]). *The closed form of a C-finite recurrence is:*

$$f(m) = p_1(m)\lambda_1^m + \ldots + p_s(m)\lambda_s^m$$

where $\lambda_1, \ldots, \lambda_s \in \mathbb{C}$ are non-zero distinct roots of the characteristic polynomial of f, and $p_1, \ldots, p_s \in \mathbb{C}[m]$ are polynomials of degree less than the multiplicities of $\lambda_1, \ldots, \lambda_s$, respectively.

Next, we define the closed form for the sequence of powers of A.

Corollary 2. *Given a square matrix $A \in \mathbb{Z}^{n\times n}$, we have:*

$$(A^m)_{i,j} = p_{1,i,j}(m)\lambda_1^m + \ldots + p_{s,i,j}(m)\lambda_s^m$$

where $\lambda_1, \ldots, \lambda_s \in \mathbb{C}$ are non-zero distinct eigenvalues of A, and $p_{1,i,j}, \ldots, p_{s,i,j} \in \mathbb{C}[m]$ are polynomials of degree less than the multiplicities of $\lambda_1, \ldots, \lambda_s$, respectively.

Proof. If $\det(A - xI_n) = x^d - a_{d-1}x^{d-1} - \ldots - a_1 x - a_0$ is the characteristic polynomial of A, then we have

$$A^d - a_{d-1}A^{d-1} - \ldots - a_1 A - a_0 = 0$$

by the Cayley-Hamilton Theorem. If we define $f_{i,j}(m) = (A^m)_{i,j}$, then we have

$$\begin{aligned} xA^{m+d} &= a_{d-1}A^{m+d-1} + \ldots + a_1 A^{m+1} + a_0 A^m \\ f_{i,j}(m+d) &= a_{d-1}f_{i,j}(m+d-1) + \ldots + a_1 f_{i,j}(m+1) + a_0 f_{i,j}(m) \end{aligned}$$

By Theorem 4, we have that

$$(A^m)_{i,j} = p_{1,i,j}(m)\lambda_1^m + \ldots + p_{s,i,j}(m)\lambda_s^m$$

for some polynomials $p_{1,i,j}, \ldots, p_{s,i,j} \in \mathbb{C}[m]$ of degrees less than the multiplicities of $\lambda_1, \ldots, \lambda_s$, respectively. □

Proof of Lemma 12 Assume from now on that all non-zero eigenvalues $\lambda_1, \ldots, \lambda_s$ of A are such that $\lambda_1^{d_1} = \ldots = \lambda_s^{d_s} = 1$, for some integers $d_1, \ldots, d_s > 0$. The method given in [2] for testing the finite monoid condition for A gives also bounds for $d_1, \ldots, d_s$. Then we have $\lambda_1^L = \ldots \lambda_s^L = 1$, where $L = \text{lcm}(d_1, \ldots, d_s)$. As $d_1, \ldots, d_s$ are effectively bounded, so is L. By Corollary 2, we have that, if m is a multiple of L, then $(A^m)_{i,j} = p_{i,j}(m)$ for some effectively computable polynomial $p_{i,j} \in \mathbb{C}[m]$ i.e., for m multiple of L, A^m is polynomially definable. But since $p_{i,j}(m)$ assumes real

values in an infinity of points $m = kL,\ k > 0$, the it must be that its coefficients are all real numbers, i.e. $p_{i,j} \in \mathbb{R}[m]$. Moreover, these coefficients are the solutions of the integer system:

$$\begin{cases} p_{i,j}(L) = (A^L)_{i,j} \\ \dots \\ p_{i,j}((d+1)L) = (A^{(d+1)L})_{i,j} \end{cases}$$

Clearly, since $A \in \mathbb{Z}^{n\times n}$, $A^p \in \mathbb{Z}^{n\times n}$, for any $p \geq 0$. Hence $p_{i,j} \in \mathbb{Q}[m]$. □

Proof of Lemma 13 Assuming the condition $a_{d-i}(\mathbf{x}) < 0$ and $a_d(\mathbf{x}) = a_{d-1}(\mathbf{x}) = \ldots = a_{d-i+1}(\mathbf{x}) = 0$, for some $0 \leq i \leq d$, we have $P(k, \mathbf{x}) = a_{d-i}(\mathbf{x}) \cdot k^d + \ldots + a_1(\mathbf{x}) \cdot k + a_0(\mathbf{x})$. Since the dominant coefficient $a_{d-i}(\mathbf{x})$ is negative, the polynomial will assume only negative values, from some point on. □