

An Algebraic Algorithm for the Identification of Glass Networks with Periodic Orbits Along Cyclic Attractors^{*}

Igor Zinovik, Daniel Kroening, and Yury Chebiryak

Computer Systems Institute, ETH Zurich, 8092 Zurich, Switzerland
{izinovik,daniel.kroening,yury.chebiryak}@inf.ethz.ch

Abstract. Glass piecewise linear ODE models are frequently used for simulation of neural and gene regulatory networks. Efficient computational tools for automatic synthesis of such models are highly desirable. However, the existing algorithms for the identification of desired models are limited to four-dimensional networks, and rely on numerical solutions of eigenvalue problems. We suggest a novel algebraic criterion to detect the type of the phase flow along network cyclic attractors that is based on a corollary of the Perron-Frobenius theorem. We show an application of the criterion to the analysis of bifurcations in the networks. We propose to encode the identification of models with periodic orbits along cyclic attractors as a propositional formula, and solving it using state-of-the-art SAT-based tools for real linear arithmetic. New lower bounds for the number of equivalence classes are calculated for cyclic attractors in six-dimensional networks. Experimental results indicate that the runtime of our algorithm increases slower than the size of the search space of the problem.

1 Introduction

Many biological models can be formulated as hybrid systems in which the switch-like behavior of genes is approximated by discontinuous step functions, while the other state variables still change continuously in time. Piecewise-linear differential equations (PLDE) were proposed by Glass and Kaufmann as an approximation for systems in the context of gene regulation [1,2]. These equations are applied to the analysis of gene regulatory networks [3,4,5,6] and neural networks [7,8]. The piecewise linear approach for describing complex nonlinear dynamics is actively studied and utilized in control theory, design of electric and electronic circuits, and embedded software.

The main distinction of biological phenomena is that the interactions are characterized by very localized coupling of the state variables, unlike complex couplings in the context of control and electronic circuit problems. In the resulting model, interactions between genes are present only in the piecewise constant

^{*} This research is supported in part by an award from IBM Research and by ETH Research Grant TH-19 06-3.

terms of the PLDE [9]. Let n denote the number of genes and x_i denote the concentration of the product of gene i . The vector of the x_i -s is denoted by \mathbf{x} . The equations can be written in the form

$$\dot{x}_i = -g_i(\mathbf{x}) - \gamma_i x_i \quad \text{for } 1 \leq i \leq n,$$

where $\gamma_i > 0$ is the degradation rate of x_i . The function $g_i : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$ describes the coupling of the variables and is defined as

$$g_i(\mathbf{x}) = \sum_{l \in L} k_{il} b_{il}(\mathbf{x})$$

where $k_{il} \geq 0$ is a rate parameter, L is a set of indexes, and $b_{il} : \mathbb{R}_{\geq 0}^n \rightarrow \{0, 1\}$ is a composition of step functions with the steps located at the prescribed threshold concentrations $x_i = \theta_{il}$. The function b_{il} expresses the conditions under which the gene causes production of the protein at a rate k_{il} . The constant θ_{il} denotes the l -th threshold concentration of the protein encoded by gene i . The thresholds induce a partitioning of the phase space into a set of n -dimensional boxes. In each box, the protein concentrations are described by ODEs with a constant production term μ_i and a rate parameter γ_i :

$$\dot{x}_i = \mu_i - \gamma_i x_i \quad \text{for } 1 \leq i \leq n$$

The global behavior of PLDE with several thresholds for every continuous variable are actively studied in the context of modeling of gene regulatory networks [10] and the qualitative theory of differential equations [11]. If the model of the gene activity is restricted to on/off expressions and the decay rates are identical for all reactions, the PLDE system is reduced to a Glass model [12]. The general form of a Glass network is

$$\dot{x}_i = G_i(\tilde{x}_1, \dots, \tilde{x}_n) - \alpha x_i \quad \text{for } 1 \leq i \leq n \text{ and } \alpha > 0.$$

The protein production rates are defined via the interaction functions G_i , where $\tilde{x}_i = a$ if $x_i < \theta_i$, and $x_i = b$ if $x_i > \theta_i$ with real constants $a < b$. Using appropriate scaling of the variables, the PLDE can be transformed into the system

$$\dot{y}_i = F_i(\tilde{y}_1, \dots, \tilde{y}_n) - y_i \quad \text{for } 1 \leq i \leq n,$$

where $\tilde{y}_i = 0$ if $y_i < 0$, and $\tilde{y}_i = 1$ if $y_i > 0$ [12]. The equations describe a network with all thresholds equal 0 and unit decay rate. The equations can be easily integrated, and the trajectories are straight lines in every orthant¹ \mathcal{O}_k , $k \in \{1, 2, 3, \dots, 2^n\}$, of the phase space. The phase flow in each orthant \mathcal{O}_k is defined by its focal point $\mathbf{f}^k = (f_1^k, f_2^k, \dots, f_n^k) \in \mathbb{R}^n$ where $f_i^k = F_i(\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_n)|_{\mathcal{O}_k}$. Thus, the Glass network can be specified by a choice of a set of focal points $\{\mathbf{f}^{(k)}\}$, $k \in \{1, 2, 3, \dots, 2^n\}$.

The phase flow in Glass networks is studied using a *state transition diagram*, which is represented by an n -cube with directed edges. Each orthant of the phase

¹ Generalization of a quadrant to the n -dimensional Euclidean space.

space is associated with a vertex of the n -cube, and each common boundary of the orthants corresponds to an edge of the cube. The edge is directed according to the direction of the phase flow across the boundary [13]. Figure 1 illustrates a phase flow with two trajectories of a two-dimensional Glass network. The state transition diagram for a 3-dimensional Glass network is shown in Fig. 2. The vertices of the n -cube are labeled by tuples of n binary variables $(\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_n)$, which define a valuation of the network interaction functions F_i . Periodic trajectories of the networks correspond to closed cycles in the transition graphs (e.g., see thick line in Fig. 2).

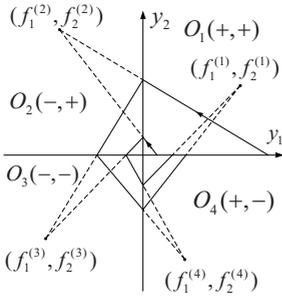


Fig. 1. 2-d phase flow

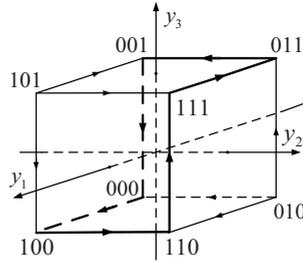


Fig. 2. 3-d transition diagram

The global phase flow in Glass networks can be quite complex. Oscillations towards equilibrium states, cycles and limit cycles may occur when linear parts of the trajectories are connected continuously over sequences of orthants [13,14,12,15]. Numerical simulations [13,16] indicate that for dimensions greater than 4, Glass networks may exhibit aperiodic and chaotic behavior. Studies of the periodic solutions for Glass models show that there are networks that possess a special type of stable limit cycles: the flow between the orthants along these cycles is unambiguous, i.e., for each orthant along the cycle, all trajectories must go to the same successor. In other words, the basin of attraction of the periodic trajectory is composed of all orthants spanned by the trajectory. Networks with such stable cycles are called networks with *cyclic attractors* [13].

Definition 1 (Cyclic Attractor). *A cycle in the state diagram is called a cyclic attractor if a) it is a chord-free simple cycle in the n -cube², and b) all edges adjacent to the cycle have to be directed towards the cycle nodes.*

As example, the cycle shown in Fig. 2 is a cyclic attractor.

Models for gene regulatory networks with equilibrium states and stable limit cycles are of special interest in Systems Biology because the models serve for simulation of cell differentiation processes and variability of cell types [17,18]. The classification of the cyclic attractors with respect to symmetry transformation of the n -cube up to dimension 5 has been completed [13]. Numerical studies of the

² Every edge in the graph that joins two vertices of the cycle is an edge of this cycle.

3-dimensional cyclic attractor identified unique stable oscillations for the value of the bifurcation parameter greater than the Hopf bifurcation point [19]. The network with cyclic attractors was integrated numerically for the 4-dimensional state space to simulate a neural network [20]. Three stable periodic trajectories were found by the parametric study of PLDE models, and the period of each cycle was computed for a set of thresholds θ_i .

To summarize, the classification of the transition diagrams has been obtained for Glass networks up to dimension five. Analytical results on phase flow are presented for three- and four-dimensional networks. The analysis relies on the integration of PLDE and numerical solutions of eigenvalue problems for the matrix associated with the Poincaré return map. Models for the simulation of gene regulatory and neural networks utilize Glass networks with cyclic attractors. Phase flow along cyclic attractors was proven to admit either a stable periodic orbit or to converge to the origin. If the focal coordinates $\{f_i^{(k)}\}$ for the system with cyclic attractor equal ± 1 , the flow always is attracted by the unique 1-period trajectory.

The determination of the parameters of gene regulation models based on experimental observations is known to be highly desirable [17], and is a computationally difficult problem [21]. A solution of the inverse problem of 4-dimensional Glass network reconstruction based on a partial information about the transition diagram and the signs of focal coordinates is shown in [3]. The objective of this paper is to suggest an efficient method for the identification of networks with cyclic attractors that exhibit phase flow of a prescribed type for a given set of focal point coordinates.

The problem is stated as follows: based on a given sequence of absolute values of focal point coordinates $\{|f_i^{(k)}|\}$ and the desired flow type, identify a Glass network with an attractor that admits the prescribed flow.

As there are straight-forward upper bounds for the length of cyclic attractors for a given dimension, we propose to use an encoding into propositional satisfiability (SAT) for the search. There are two contributions we present:

1. We propose an algebraic method for analysis of structural stability of phase flow for Glass networks with cyclic attractors. Our method utilizes a corollary of the Perron-Frobenius theorem and gives a criterion for the identification of the flow type along the cyclic attractors.
2. We propose a scalable SAT-based algorithm for identification of the networks with cyclic attractors.

Outline. The paper is organized as follows. In Section 2, we extend a sufficient condition for the identification of networks with cyclic attractors [13] to an algebraic criterion (i.e., necessary and sufficient condition), which is derived from properties of the state transition diagram of the networks. We also present an application of the criterion to the analysis of structural stability of the phase flow as an example, which is useful later on for the construction of the algorithm for network identification. In Section 3, we introduce an algorithm based on SAT for the search for cyclic attractors in the state transition diagram. In Section 4, we

integrate the proposed criterion as a part of the algorithm for the identification of Glass networks with periodic orbits along cyclic attractors. We also present experimental results that indicate that the algorithm scales well in the network dimension.

2 Algebraic Criterion for Flow Identification

The flow along cyclic attractors is known either to converge to the origin or to admit a unique stable 1-period orbit. The type of the flow is identified by analyzing a Poincaré plane: the flow with periodic orbit has a unique fixed point, while the Poincaré map for flows converging to the origin does not have fixed points. Thus, a criterion for flow identification should distinguish between Poincaré maps with and without fixed points.

Poincaré return maps of Glass networks can be represented by the composition of *fractional linear maps* $M^{(k)} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ [13]. Following the notation in [12], the mapping can be presented as:

$$\mathbf{y}^{(k+1)} = M^{(k)}\mathbf{y}^{(k)} = B^{(k)}\mathbf{y}^{(k)} / (1 + \langle \psi^{(k)}, \mathbf{y}^{(k)} \rangle), \quad (1)$$

where $\mathbf{y}^{(k)}$ is the coordinate vector on the k -th orthant boundary crossed by the trajectory, and the matrix $B^{(k)}$ is calculated as

$$B^{(k)} = I - (\mathbf{f}^{(k)} \mathbf{e}_j^T) / f_j^{(k)} \quad (2)$$

where I is the identity matrix. The focal point $\mathbf{f}^{(k)}$ associated with the orthant being entered is assumed not to lie in that orthant, \mathbf{e}_j denotes the standard basis vector in \mathbb{R}^n , and the vector $\psi^{(k)}$ is defined to equal $-\mathbf{e}_j / f_j^{(k)}$. The angular brackets denote the Euclidean inner product. Thus, the return map for a cycle restricted to the orthant boundary $y_i = 0$ can be written as

$$M\mathbf{y} = A\mathbf{y} / (1 + \langle \phi, \mathbf{y} \rangle), \quad (3)$$

where $A = \|a_{mp}\|$ is the $(n-1) \times (n-1)$ matrix obtained by deleting the i -th column and row of the composition of $B^{(k)}$, and ϕ is the same reduction of the composition of $\psi^{(k)}$ [12].

The values of the matrix elements a_{mp} depend on the choice of the initial orthant boundary as well as on the order of enumeration of the variables, and the prescribed orientation of the basis vectors along the axes. The same N -node cycle in the state diagram may be represented by $N \cdot n! \cdot 2^n$ different matrices. In case of a cyclic attractor, the matrix can be obtained in such a way that all its elements are positive [13]. Subsequently, the Perron-Frobenius theorem guarantees that the flow admits a stable periodic orbit if the dominant eigenvalue r of the positive matrix A is greater than one, and converges to the origin otherwise. Therefore, if a cyclic attractor is represented by a positive matrix, the identification of the flow type does not require the calculation of eigenvalues, but only reasoning about satisfiability of the inequality $r > 1$ for positive matrices. For this purpose, we suggest to utilize a corollary to the Perron-Frobenius theorem. The corollary asserts [22]:

Corollary 1. *A real number λ is greater than the maximal characteristic value r of the (non-negative) matrix A if and only if for this value λ all the successive principal minors of the characteristic matrix $\lambda I - A$ are positive.*

If we are only interested in testing $r > 1$, we need to check the signs of the determinants of the k by k upper left matrices of $A - I$ being $(-1)^k$. Thus, the following algebraic criterion for the identification of flows in cyclic attractors can be used:

Criterion 1. *The flow of an n -dimensional cyclic attractor converges to the origin if and only if the signs of the determinants of the k by k upper left matrices of $A - I$ are $(-1)^k$ for $k = 1, 2, \dots, n - 2$ and the sign of $\det(A - I)$ is $(-1)^{n-1}$ or $\det(A - I) = 0$, where A is the positive matrix that defines the return map of the attractor by means of Eq. 3. Otherwise, the phase flow along the cyclic attractor admits a unique stable 1-period orbit.*

The analysis of the generic ways in which stable attractors undergo bifurcations in Glass networks is an open question listed in [15]. As a simple example of an application of Criterion 1 to bifurcation analysis, we can consider the structural stability of phase flow along a cyclic attractor for the 3-dimensional Boolean Glass network shown in Fig. 2.

First, we have to define the focal point coordinates of the network. Two conditions are assumed throughout the paper: focal points lie inside orthants and none of them on the orthant boundaries, and the i -th state variable does not change in sign when crossing an orthant boundary in direction i . The conditions ensure that the flow is unambiguous [12,11]³. In this case, the focal point for every orthant of the cycle lies inside the next cycle orthant.

Example 1. The attractor in Fig. 2 is represented by the orthant sequence

$$(111) \rightarrow (011) \rightarrow (001) \rightarrow (000) \rightarrow (100) \rightarrow (110)$$

Thus, the sequence of focal points is obtained by applying a one-step cyclic shift to the sequence of orthants, and replacing all 0-s by -1 , and is written as:

$$(-1, 1, 1) \rightarrow (-1, -1, 1) \rightarrow (-1, -1, -1) \rightarrow (1, -1, -1) \rightarrow (1, 1, -1) \rightarrow (1, 1, 1)$$

Let us consider the perturbations of the first focal point when it remains inside the same orthant (011). The focal point sequence undergoing the perturbations has the form:

$$(-\epsilon_1, \epsilon_2, \epsilon_3) \rightarrow (-1, -1, 1) \rightarrow (-1, -1, -1) \rightarrow (1, -1, -1) \rightarrow (1, 1, -1) \rightarrow (1, 1, 1),$$

where $\epsilon_1 > 0$, $\epsilon_2 > 0$, and $\epsilon_3 > 0$ are free parameters of the network. The matrix $A = ||a_{mp}||$ for the return map is calculated using Equations (1-3):

$$\begin{pmatrix} \frac{8\epsilon_2}{\epsilon_1} + \frac{5\epsilon_3}{\epsilon_1} & 8 \\ \frac{5\epsilon_2}{\epsilon_1} + \frac{3\epsilon_3}{\epsilon_1} & 4 \end{pmatrix}.$$

³ The conditions can be relaxed using set-valued Filippov solutions. The application of differential inclusions to PLDE is still a current research topic [10], and is not considered in this paper.

All elements of the matrix are positive due to the definition of the perturbation via positive ϵ -s. Thus, Criterion 1 is applicable to the matrix above, and it asserts that the cyclic attractor admits the flow converging to the origin if and only if

$$a_{11} < 0 \wedge (a_{11}a_{22} - a_{21}a_{12} > 0 \vee a_{11}a_{22} - a_{21}a_{12} = 0). \tag{4}$$

The corresponding systems of inequalities are written as

$$\left\{ \begin{array}{l} \frac{-\epsilon_1 + 8\epsilon_2 + 5\epsilon_3}{\epsilon_1} < 0 \\ \frac{4(\epsilon_1 + 2\epsilon_2 + \epsilon_3)}{\epsilon_1} > 0 \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{l} \frac{-\epsilon_1 + 8\epsilon_2 + 5\epsilon_3}{\epsilon_1} < 0 \\ \frac{4(\epsilon_1 + 2\epsilon_2 + \epsilon_3)}{\epsilon_1} = 0 \end{array} \right.$$

Both systems are inconsistent, and therefore, the flow admits a stable periodic orbit, i.e., the network is stable under any perturbations of the first focal point that leave the point inside the orthant $y_1 < 0, y_2 > 0, y_3 > 0$.

The perturbations of any single focal point within the orthants have been found to preserve the flow type along 3- and 4-dimensional cyclic attractors for Boolean Glass networks. In contrast, simultaneous perturbation of two coordinates of different focal points may change the flow from “periodic” to “converging to the origin”. As an example, consider perturbations of the second coordinate of the fifth focal point and the third coordinate of the sixth focal point.

Example 2. Let us consider the sequence of the focal points which is written as $(-1, 1, 1) \rightarrow (-1, -1, 1) \rightarrow (-1, -1, -1) \rightarrow (1, -1, -1) \rightarrow (1, \epsilon_1, -1) \rightarrow (1, 1, \epsilon_2)$

The corresponding positive matrix A is

$$\begin{pmatrix} \frac{5}{\epsilon_1} + \frac{3}{\epsilon_2} + \frac{5}{\epsilon_1\epsilon_2} & \frac{3}{\epsilon_1} + \frac{2}{\epsilon_2} + \frac{3}{\epsilon_1\epsilon_2} \\ \frac{3}{\epsilon_2} + \frac{5}{\epsilon_1\epsilon_2} & \frac{2}{\epsilon_2} + \frac{3}{\epsilon_1\epsilon_2} \end{pmatrix}$$

The system (4) that represents the criterion is simplified by cylindrical decomposition implemented in *Mathematica*. The sufficient condition for converging flow along the cyclic attractor and the bifurcation condition are written as:

$$\left\{ \begin{array}{l} \epsilon_1 > 5 \\ \epsilon_2 > \frac{7+5\epsilon_1}{-5+\epsilon_1} \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{l} \epsilon_1 > 5 \\ \epsilon_2 = \frac{7+5\epsilon_1}{-5+\epsilon_1} \end{array} \right.$$

Any solution of this system defines a network with the flow converging to the origin. As an example of one parametric bifurcation diagram we consider a solution of the second system with $\epsilon_1 = 6$ and $\epsilon_2 = 37$. In this case, the dominant eigenvalue r is 1 and the phase flow converges to the origin. The sequence of focal points with the bifurcation parameter μ is written as:

$$\begin{aligned} &(-1, 1, 1) \rightarrow (-1, -1, 1) \rightarrow (-1, -1, -1) \\ &\rightarrow (1, -1, -1) \rightarrow (1, 6, -1) \rightarrow (1, 1, 37 - \mu) \end{aligned}$$

If $\mu \leq 0$, the cyclic attractor admits the flow converging to the origin, and if $\mu > 0$, the location of the fixed point $\mathbf{y}^*(\mu)$ on the Poincaré plane ($y_1 > 0, y_2 > 0, y_3 = 0$) is computed as [12]:

$$\mathbf{y}^*(\mu) = \frac{(r - 1)\mathbf{v}}{\langle \phi, \mathbf{v} \rangle},$$

where \mathbf{v} is the eigenvector corresponding to the dominant eigenvalue r . The characteristic polynomial for the matrix A is quadratic, and thus, a bifurcation diagram that represents the fixed point coordinate $\mathbf{y}^*(\mu)$ can be obtained in a closed analytical form. The bifurcation diagram was found to be similar to a Hopf supercritical bifurcation for non-linear ODE (see Fig. 5 in [23]).

Criterion 1 relies on the condition that matrix A is positive, and thus, the first step of any application of the criterion is to find the sequence of nodes in the n -cube that determines the cyclic attractors with a positive matrix. Such sequences have to satisfy condition (a) of Def. 1, and are called *induced cycles*. The problem of finding longest induced paths in graphs is known to be *NP*-complete [24], and the problem of detecting longest induced cycles in n -cubes is open for dimensions greater than 7 [25]. We propose to encode the search for induced cycles into a satisfiability (SAT) problem for propositional logic. Thus, the computationally demanding calculations can be handled by the state-of-the-art SAT solvers, which are known to be very efficient for problems with large, tightly constrained search spaces.

3 Computing Induced Cycles

The search for an induced cycle in the network state transition diagram relies on the identification of a cycle with desired properties on n -cubes. The length of the cycle N and the dimension n serve as input parameters. We propose to apply propositional SAT to the search for the attractors.

A state corresponds to a coordinate vector labeling the nodes on the n -cube, i.e., an n -tuple of Boolean variables. Let $s_{i,j}$ with $i \in \{1, \dots, N\}$, $j \in \{1, \dots, n\}$ denote the value of bit j in step i . The transitions on the n -cube correspond to sequences of states that satisfy a Gray code condition: the Hamming distance between two neighboring states equals one. We write $H_{k,l}^\alpha$ if the Hamming distance between the states s_k and s_l is α . The Gray code condition is then written as the following conjunction:

$$\Psi^{gray} = \bigwedge_{i=1}^{N-1} H_{i,i+1}^1 \wedge H_{1,N}^1$$

The constraints that eliminate the chords from the paths are represented by the following formula:

$$\Psi^{cycle} = \bigwedge_i \bigwedge_j [H_{i,j}^1 \Leftrightarrow (H_{i-1,j}^0 \vee H_{i+1,j}^0)]$$

The constraints guarantee that the Hamming distance for two of the cycle nodes equals one if and only if one of the nodes is either the previous or the next in the cycle with respect to the other one. A satisfying assignment to

$$\Psi^{ind} = \Psi^{gray} \wedge \Psi^{cycle} \tag{5}$$

identifies an induced cycle of an n -cube. The set of all attractors is represented by the set of all satisfying assignments of formula (5).

Due to the symmetries in the n -cube, the set of cycles that corresponds to the solutions of (5) is highly redundant. Glass proposes equivalence classes that are defined as sets of induced cycles such that all cycles in every set can be obtained via n -cube symmetry transformations of any cycle in the set [2]. The classification for 5-dimensional networks was obtained by Glass [2] using an enumeration approach. We utilize (5) to extend the classification to 6-dimensional networks.

The computation of the equivalence classes utilizes *coordinate* and *interval sequences* for Gray codes and paths on n -cubes. The *coordinate sequence* is a listing of the coordinates that change as the cycle is traversed. The *interval sequence* of a coordinate is a tuple giving the number of coordinates intervening between each successive appearance of the coordinate in the coordinate sequence. A necessary but not sufficient condition that two induced cycles are equivalent is that the set of interval sequences for one cycle are in a one-to-one correspondence with the set of interval sequences of the second cycle, where the interval sequence for any one coordinate can be cyclically permuted [2]. We apply the condition to compute lower bounds for the number of equivalence classes as follows:

1. We obtain the set of induced cycles of a given length N in the n -cube by computing all satisfying assignments of (5). This is an *all-SAT* problem.
2. We construct the set of equivalence classes as follows: every satisfying assignment is decoded back to the coordinates of the induced cycle on n -cube that it represents; if this induced cycle does not belong to any of the computed classes, it is added to the set as the representative.

The pseudocode for the computation is shown in [23]. The all-SAT problem is solved using the *blocking clause* algorithm [26] for the MiniSAT SAT-solver [27]. The algorithm computes a satisfying assignment of the given formula, saves it, and constructs a clause that eliminates the assignment. The clause is added to the formula as an additional constraint, and the previous step is repeated until no satisfying assignment can be found. There are more efficient algorithms available for the all-SAT problem, but these techniques are beyond the scope of this paper.

We applied the algorithm to 5- and 6-dimensional cubes (see the results in Table 1 in [23]). The lower bound for the total number of equivalence classes for six dimensions has been found to increase from 17 to 3007. The computed bound for 5-dimensional networks differs from the exact number of classes [2] by just one class. To the best of our knowledge, these lower bounds for the number of equivalence classes for dimension 6 are presented for the first time.

The network identification may require the evaluation of all induced cycles, even if they belong to the same equivalence class (see the example in [23]). The results of the all-SAT computation indicate that the number of induced cycles increases rapidly with the network dimension: the total number of cycles is 238 and 706336 for 5- and 6-dimensional networks, respectively. Thus, the search over the set of cycles becomes computationally demanding with increasing network dimension.

The size of the search space grows in the order of $2^{n \cdot N}$, i.e., exponentially in the network dimension and the length of the induced cycle. On the other hand, the number of induced cycles *decreases* when the cycle length approaches its maximum value. These opposite trends compromise the efficiency of any algorithm if it identifies Glass networks by enumerating cyclic attractors and applying Criterion 1 subsequently. In the next section, we propose combining the criterion for the flow detection and the search of induced cycles into an identification algorithm that efficiently scales with the network dimension.

4 Algorithm for Network Identification

4.1 Implementation Using SMT

An algorithm that simultaneously detects the flow type and identifies the cyclic attractors is required to conduct a search over both the continuous and discrete parts of the problem. We propose to utilize solvers for *Satisfiability Modulo Theories* for this problem. State-of-the-art solvers for Satisfiability Modulo Theories (SMT) decide logical satisfiability (or dually, validity) with respect to a background theory expressed in classical first-order logic with equality. These theories include real or integer arithmetic, which makes SMT solvers a successful tool for the analysis of problems that include linear inequalities over reals [28]. We propose to encode the identification of networks by adding the inequalities that represent the criterion for flow detection to the propositional formula (5). The Boolean structure of the inequality system for the three-dimensional network is defined by formula (4). In case of an arbitrary dimension n , the formula is written as

$$\Psi^{con} = \Psi^{suf} \vee \Psi^{bif},$$

where a sufficient condition for the converging flow is defined by

$$\Psi^{con} = (\det(A - I)^{(1)} < 0) \wedge (\det(A - I)^{(2)} > 0) \wedge (\det(A - I)^{(3)} < 0) \wedge \dots \wedge (\det(A - I)^{(n-1)} \geq 0),$$

and the condition for the bifurcation point is

$$\Psi^{bif} = (\det(A - I)^{(1)} < 0) \wedge (\det(A - I)^{(2)} > 0) \wedge (\det(A - I)^{(3)} < 0) \wedge \dots \wedge (\det(A - I)^{(n-2)} \geq 0) \wedge (\det(A - I)^{(n-1)} = 0).$$

Here, $\det(A - I)^{(k)}$ denotes the determinant of the upper left $k \times k$ matrix of $A - I$ and \geq changes accordingly with the sign of $(-1)^k$.

The criterion is applicable if $A = ||a_{mp}||$ is a positive matrix. The following condition guarantees that the cyclic attractor induces a matrix with positive entries:

$$\Psi^{pos} = \bigwedge_m \bigwedge_p (a_{mp} > 0).$$

The matrix elements a_{mp} are calculated using (1-3) based on the prescribed sequence of absolute values $\{|f_i^k|\}$ for the focal point coordinates and a satisfying

assignment $s_{k,i}^*$ of the propositional formula (5). The assignment defines the signs in the sequence of focal coordinates $\{|f_i^{(k)}|\}$ such that the focal point for every orthant of the induced cycle is located inside the next orthant along the cycle:

$$f_i^{(k)} = \begin{cases} -|f_i^{(k)}| & : \text{if } \neg s_{k-1,i}^* \\ |f_i^{(k)}| & : \text{otherwise} \end{cases}$$

A Glass network with converging flow along the cyclic attractor is identified from a satisfying assignment for the propositional formula

$$\Psi^{ind} \wedge \Psi^{pos} \wedge \Psi^{con}, \quad (6)$$

and a network with stable periodic orbit is specified by a satisfying assignment for the formula

$$\Psi^{ind} \wedge \Psi^{pos} \wedge \neg \Psi^{con}. \quad (7)$$

An assignment for (6) or (7) solves the corresponding identification problem if all coordinates $\{|f_i^{(k)}|\}$ are given as a sequence of positive real numbers. On the other hand, the formulae allow for an analysis of the structural stability of the network if one coordinate of the sequence is a positive parameter ϵ that undergoes the perturbation. The periodic flow along the cyclic attractor is structurally unstable if there is a satisfying assignment for the formula

$$\Psi^{ind} \wedge \Psi^{pos} \wedge \Psi^{con} \wedge (\epsilon > 0). \quad (8)$$

The perturbation of two or more focal coordinates causes the polynomial inequalities in the criterion to appear (see the examples in Section 2). Non-linear inequalities are not supported by any of the existing SMT solvers, and thus, the calculations are restricted to the case of one parameter. Additional constraints may be added to limit the analysis of structural stability to a particular equivalence class of the n -cube.

4.2 Experiments

We evaluated the Yices and CVCL SMT solvers [29,30]. Yices won the Satisfiability Modulo Theories competition⁴ in 2006 in the relevant category. As first step of the experimental evaluation, we compare the run-time of the search for a single induced cycle, i.e., checking satisfiability of the purely propositional formula (5) for various instances. The test cases include the search for the induced cycles of different length in the networks of dimension 4, 5, and 6 (see Table 2 in [23]). A PC with a 1.4 GHz processor and 2 GB RAM was used for the evaluation. We also recorded the run-time of MiniSAT on the same instance as a reference point.

The difference between MiniSAT and the SMT solvers is that MiniSAT accepts conjunctive normal form (CNF) as input directly, while the SMT solvers use rich

⁴ Computer-Aided Verification Conference, SMT-COMP, <http://www.csl.sri.com/users/demoura/smt-comp/>

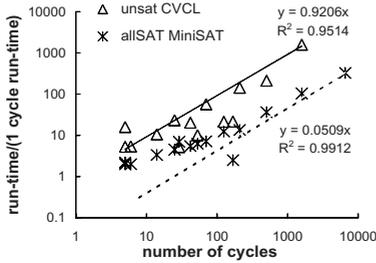


Fig. 3.

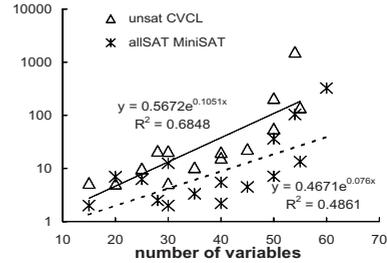


Fig. 4.

input languages that include all Boolean logic operators. The results indicate that a hand-tuned translation of the propositional formula (5) into CNF may increase the speed of the SMT solver up to the efficiency of MiniSAT, which is currently one of the fastest tools for satisfiability analysis of propositional formulae (see Fig. 6 in [23]).

We chose CVCL over Yices for the evaluation of the identification algorithm because we found the CVCL language more convenient than that of Yices for arithmetic expressions that define the elements of matrix A . The benchmark problem is to identify a Glass model with converging phase flow along cyclic attractors for the same networks as presented above. The satisfiability of formula (6) was evaluated for the input sequence of focal points that were specified to contain only unit coordinates. Such an input restricts the search to Boolean Glass networks that are known not to have flow converging to the origin. Thus, the instance is unsatisfiable, i.e., there is no Glass model satisfying the problem specifications. In this case, the solution requires the evaluation of all induced cycles of the instance and the run-time provides a conservative estimate of the efficiency of the algorithm.

The run-time of CVCL for the benchmark increases linearly with the number of induced cycles in the network (see the solid line in Fig. 3). A linear increase was also observed when MiniSAT was used to solve the corresponding *all-SAT* problems (dashed line in Fig. 3). The same trend indicates that the proposed network identification algorithm scales in the number of network cycles, just as MiniSAT scales well for the problem of computing all induced cycles of this network.

The scalability of the algorithm in the size of search space is estimated using the least square interpolation of the run-time as a function of the number L of Boolean variables in the instance. The interpolation using exponential trend lines is depicted in Fig. 4 by a solid and a dashed line for CVCL and MiniSAT, respectively. The run-time increases approximately as $e^{0.1L}$, while the growth of the size of the search space of a set of L Boolean variables is proportional to $2^L = e^{\ln(2)L} \approx e^{0.7L}$. Thus, the experimental results indicate that the run-time of our algorithm increases about 7 times more slowly than the volume of the problem search space.

5 Conclusion

The proposed algorithm belongs to the methods that utilize propositional logic for reasoning about properties of ordinary differential equations. Such methods are widely applied to the analysis of biological networks and hybrid systems. The existing computational tools, developed for the propositional analysis of biological networks, approximate the ODE trajectories using the numerical Runge-Kutta procedure [31], Taylor series [32], or an approximate partitioning of phase space of continuous variables [4]. The computation of the reachable states for hybrid systems also relies on approximations of the PLDE solution [33,34]. We show that an exact algebraic algorithm can be applied for reasoning about the phase flow in a subclass of PLDE that is utilized in the Glass model.

The algorithm is applicable in the case when the PLDE system is near a bifurcation point, where the approximate methods may be inconclusive. We conducted an analysis of the structural stability of the phase flow for Glass networks with cyclic attractors. Cylindrical decomposition has been used for the evaluation of the criterion for identification of the phase flow. The flow for Boolean Glass models has been shown to be stable under the perturbations of any single focal point along the cyclic attractor. The cylindrical decomposition is known to be a powerful tool for evaluating the structural stability of partial and ordinary differential equations [35,36,37]. To the best of our knowledge, the presented stability analysis is a first attempt to apply cylindrical decomposition for the identification of bifurcations in Glass networks.

The proposed algorithm has been found to benefit from the scalability of Bounded Model Checking: new lower bounds for the number of equivalence classes are calculated for cyclic attractors in 6-dimensional networks. Our experimental results also indicate that the run-time of our algorithm increases slower than the size of the search space of the problem.

Acknowledgements. We would like to thank Felix Friedrich for the help with *Mathematica* calculations, and Clark Barrett for kind and detailed answers about CVCL.

References

1. Glass, L., Kaufmann, S.: The logical analysis of continuous non-linear biochemical control networks. *J. Theor. Biol.* 39, 103–129 (1973)
2. Glass, L.: Combinatorial aspects of dynamics in biological systems. In: *Stat. Mech. Stat. Methods in Theory and Application*, pp. 585–611. Plenum Press, New York (1976)
3. Edwards, R.: Symbolic dynamics and computation in model gene networks. *Chaos* 11, 160–169 (2001)
4. Ghosh, R., Tiwari, A., Tomlin, C.: Automated symbolic reachability analysis; with application to delta-notch signalic automata. In: Maler, O., Pnueli, A. (eds.) *HSCC 2003*. LNCS, vol. 2623, pp. 233–248. Springer, Heidelberg (2003)
5. Mason, J., Linsay, P., Collins, J., Glass, L.: Evolving complex dynamics in electronic models of genetic networks. *Chaos* 14, 707–715 (2004)

6. Batt, G., Ropers, D., de Jong, H., Geiselmann, J., Mateescu, R., Page, M., Schneider, D.: Analysis and verification of qualitative models of genetic regulatory networks: A model-checking approach. In: 19th Int Joint Conference on Artificial Intelligence, pp. 370–375 (2005)
7. Gedeon, T.: Global dynamics of neural nets with infinite gain. *Physica D: Nonlinear Phenomena* 146, 200–212 (2000)
8. Gedeon, T.: Attractors in continuous time switching networks. *Communications on Pure and Applied Analysis* 2, 187–209 (2003)
9. de Jong, H.: Modeling and simulation of genetic regulatory systems: a literature review. *J. Comp. Biol.* 9, 67–103 (2002)
10. Casey, R., de Jong, H., Gouze, J.L.: Piecewise-linear models of genetic regulatory networks: equilibria and their stability. *J. Math. Biol.* 52, 27–56 (2006)
11. Farcot, E.: Geometric properties of a class of piecewise affine biological network models. *J. Math. Biol.* 52, 373–418 (2006)
12. Edwards, R.: Analysis of continuous-time switching networks. *Physica D* 146, 165–199 (2000)
13. Glass, L., Pasternack, J.: Stable oscillations in mathematical models of biological control systems. *J. Math. Biol.* 6, 207–223 (1978)
14. Mestl, T., Plahte, E., Omholt, S.: Periodic solutions in systems of piecewise-linear differential equations. *Dynam. Stabil. Syst.* 10, 179–193 (1995)
15. Edwards, R., Glass, L.: Combinatorial explosion in model gene networks. *Chaos* 10, 691–704 (2000)
16. Mestl, T., Lemay, C., Glass, L.: Chaos in high-dimensional neural and gene networks. *Physica D* 98, 33–52 (1996)
17. Thomas, R., Kaufman, M.: Multistationarity, the basis of cell differentiation and memory. *Chaos* 11, 170–195 (2001)
18. Kauffman, S.: A proposal for using the ensemble approach to understand genetic regulatory networks. *Theor. Biol.* 230, 581–590 (2004)
19. Glass, L.: Global analysis of nonlinear chemical kinetics. *Statistical mechanics, part B: time dependent processes*, 311–349 (1977)
20. Glass, L., Pasternack, J.: Prediction of limit cycles in mathematical models of biological oscillations. *Bull. Math. Biol.* 40, 27–44 (1978)
21. Laubenbacher, R., Stigler, B.: A computational algebra approach to the reverse engineering of gene regulatory networks. *J. Theor. Biol.* 229, 523–537 (2004)
22. Gantmacher, F.: *The Theory of Matrices*. vol. 2. Chelsea (1974)
23. Zinovik, I., Kroening, D., Chebiryak, Y.: An algebraic algorithm for the identification of Glass networks with periodic orbits along cyclic attractors. Technical Report 557, ETH Zurich, Computer Science Department (2007)
24. Rajan, D., Shende, A.: Maximal and reversible snakes in hypercubes. In: 24th Annual Australasian Conference on Combinatorial Mathematics and Combinatorial Computing (1999)
25. Casella, W.P.D.: Using evolutionary techniques to hunt for the snakes and coils. In: *The 2005 IEEE Congress on Evolutionary Computation*, vol. 3, pp. 2499–2505. IEEE Press, NJ (2005)
26. Plaisted, D., Biere, A., Zhu, Y.: A satisfiability tester for quantified boolean formulae. *J. Discrete Appl. Math.* 130, 291–328 (2003)
27. Eén, N., Sörensson, N.: An extendable SAT-solver. *Theory and Applications of Satisfiability Testing* 2919, 502–518 (2004)
28. Sheini, H., Sakallah, K.: From propositional satisfiability to satisfiability modulo theories. In: Biere, A., Gomes, C.P. (eds.) *SAT 2006*. LNCS, vol. 4121, pp. 1–9. Springer, Heidelberg (2006)

29. Stump, A., Barrett, C., Dill, D.: CVC: a cooperating validity checker. In: 14th Int. Conf. on Computer-Aided Verification (CAV), pp. 87–105. Springer, Heidelberg (2002)
30. Dutertre, B., de Moura, L.: A fast linear-arithmetic solver for DPLL(T). In: Ball, T., Jones, R.B. (eds.) CAV 2006. LNCS, vol. 4144, pp. 81–94. Springer, Heidelberg (2006)
31. Calzone, L., Chabrier-Rivier, N., Fages, F., Soliman, S.: Machine learning biochemical networks from temporal logic properties. In: Priami, C., Plotkin, G. (eds.) Transactions on Computational Systems Biology VI. LNCS (LNBI), vol. 4220, pp. 68–94. Springer, Heidelberg (2006)
32. Piazza, C., Antonioti, M., Mysore, V., Policriti, A., Winkler, F., Mishra, B.: Algorithmic algebraic model checking I: Challenges from systems biology. In: Etessami, K., Rajamani, S.K. (eds.) CAV 2005. LNCS, vol. 3576, pp. 5–19. Springer, Heidelberg (2005)
33. Henzinger, T., Preussig, J., Wong-Toi, H.: Some lessons from the HyTech experience. In: Proc of the 40th Annual Conference on Decision and Control (CDC), pp. 2887–2892. IEEE Press, NJ (2001)
34. Frehse, G.: Phaver: Algorithmic verification of hybrid systems past HyTech. In: Morari, M., Thiele, L. (eds.) HSCC 2005. LNCS, vol. 3414, pp. 258–273. Springer, Heidelberg (2005)
35. Hong, H., Liska, R., Steinberg, S.: Testing stability by quantifier elimination. *J. Symb. Comp.* 11, 1–26 (1996)
36. Wang, D.: Elimination theory, methods, and practice. In: Mathematics, and Mathematics-Mech, pp. 91–137. Shandong Education Publishing House, Jinan (2001)
37. Wang, D., Xia, B.: Stability analysis of biological systems with real solution classification. In: ISSAC'05, vol. 3414, pp. 354–361. ACM, New York (2005)