Analysis of Discrete Bioregulatory Networks Using Symbolic Steady States

Heike Siebert

DFG Research Center Matheon,
Freie Universität Berlin, Arnimallee 6, D-14195 Berlin, Germany
siebert@mi.fu-berlin.de

Abstract. A discrete model of a biological regulatory network can be represented as a discrete function f that contains all available information on interactions between network components and the rules governing the evolution of the network in the discrete state space. Both the information on the structure as well as the dynamics of the system can be represented as directed graphs. Since the state space size grows exponentially with the number of network components, analysis of large networks is a complex problem. In this paper, we introduce the notion of symbolic steady state that allows us to identify subnetworks that govern the dynamics of the original network in at least a subset of state space. We then state rules to explicitly construct attractors of the system from subnetwork attractors. A further application of the underlying concept allows us to formulate sufficient conditions for the existence of multiple attractors resp. a cyclic attractor based on the existence of positive resp. negative feedback circuits in the structure graph. All results are discussed for dynamics derived from f via the synchronous as well as the asynchronous update rule.

1 Introduction

Discrete methods of modeling biological regulatory networks are often used if the available data is rather qualitative in nature. Each component of the network is associated with a finite number of activity levels representing e. g. a concentration interval of a substance, activity of a gene or presence or absence of a signal. The state space of the system then consists of vectors of the component activity levels, and the network dynamics is derived from a discrete function f capturing the rules of component interactions in the system. Here, two fundamentally different methods of calculating trajectories of the system are in use. The so-called synchronous update renders a deterministic representation by defining the successor of a given state as its image under f. In contrast, we obtain a non-deterministic version if we require, motivated by the assumption of distinct time delays associated with component value changes, that a state and its successor differ in one component only, but consider all successor possibilities in agreement with f. Both approaches have been used successfully, the synchronous method having advantages in particular regarding the complexity of the analysis, the asynchronous update often allowing for a more realistic representation of trajectories of the system (see e. g. [5], [16], [6] and references therein).

When analyzing large and complex networks, one is often interested in identifying subnetworks that can be associated with definite biological functions and play crucial roles in the dynamics of the system. Analysis of such networks in isolation may then yield information on the dynamical behavior of the original network. Clearly, the difficulty is that further components and interactions influence such a network building block once it is again embedded in the network. Conditions to identify suitable subnetworks that retain their behavior in the context of the complex network are needed to derive useful information on the network dynamics. In this paper we generalize and extend corresponding ideas developed for Boolean functions and asynchronous dynamics in [14] to multi-valued discrete functions, considering synchronous as well as asynchronous dynamics. We identify subsets of state space where a number of network components remain fixed in the dynamics independent of the values of the remaining components by exploiting the properties of *symbolic steady states*. These are fixed points of a adapted function f^θ that coincides with f for the most part, but also allows the consideration of a symbolic value θ for the network components. The value θ can be identified with the whole activity level range of a given component, representing uncertainty of the actual value of that component in a network state. The *regular* components, i. e. those with a specified activity level, of a symbolic steady state act as a boundary between dynamically active subnetworks similar to the notion of *frozen core* introduced in the context of random Boolean networks (see [5]). We obtain a fine structural representation of the active subnetworks by considering the local interaction graph associated with the symbolic steady state. Assembly of attractors of the isolated subnetworks with respect to the symbolic steady state then yields attractors of the original network. With this fundamental property in mind, we are able to proof more general statements concerning the relation between structure and dynamics of the network. Here, we proof that the existence of a positive resp. a negative circuit under certain conditions implies the existence of multiple attractors resp. a cyclic attractor in the synchronous as well as the asynchronous dynamics.

The paper is organized as follows. In the next section we introduce the modeling framework used in this paper. In Sect. 3 we establish the notion of symbolic steady state as well as some important properties, followed by the main results on compositional attractors of subnetworks derived from symbolic steady states in Sect. 4. The results are then used to obtain statements linking the existence of feedback circuits in the network to number and size of attractors in Sect.5. We end the paper with concluding remarks and perspectives for future work.

2 Regulatory Networks

Throughout the text let us consider a network with $n \in \mathbb{N}$ components $\alpha_1, \ldots, \alpha_n$. To simplify notation, we identify each component α_i with its index i. We view each component as a discrete variable the values of which signify the different activity levels of that component in the network. Activity levels may represent different biological characteristics, e. g. substance concentration, gene activity, absence or presence of a signal and so on. The number of activity levels of different components may differ, depending on function of components and available data. Thus, every component α_i is associated with a *range* $X_i := \{0, 1, \ldots, p_i\}$ of activity levels, where $p_i \in \mathbb{N}$ denotes the maximal activity level of α_i. The set $X := X_1 \times \cdots \times X_n$ comprises all possible vectors of activity levels of the network and thus represents the state space of the system. Interaction of network components and rules governing the network's dynamics are then captured by a discrete function $f = (f_1, \ldots, f_n) : X \to X$. Note that if for all maximal activity levels $p_i = 1$ holds, then f is a Boolean function.

2.1 Structure

In a next step, we want to derive the network structure from the function f. As commonly done, we represent the structure as a signed directed (multi-)graph with vertex set $V := \{\alpha_1, \ldots, \alpha_n\}$ and edges representing interactions between components. The sign of an edge describes the character of the interaction, negative sign signifying an inhibiting, positive an activating effect. However, it is often observed in biological systems that the influence of one component on another depends on the current state of the network. For example, if two substances form a complex that activates some target gene, then in general the presence of only one of those substances is not sufficient to induce gene expression. So, one of the substances can only effectively influence the gene when the other substance is present. Another possibility is that the character of an interaction changes depending on the state of the network. A well-known example is the DNA-binding protein TCF which can repress as well as activate the same target genes. TCF acts as activator in the presence of β-catenin, induced by WNT signaling, while the co-expression of the protein TLE converts TCF into a repressor. Such refined structural information is of great interest when linking structural and dynamical aspects and thus we want to include it in the structural representation of the network. This is done by considering *local interaction graphs*. This notion was introduced for Boolean functions in [8] and is used for multi-value functions in the form considered here in [10].

Definition 1. *Let $x \in X$. By $G(x) := G(f)(x)$ we denote the directed signed (multi-)graph with vertex set $\{\alpha_1, \ldots, \alpha_n\}$ and edge set $E(x) \subseteq V \times V \times \{+, -\}$. An edge (i, j, ε) belongs to $E(x)$ iff there exists $c_i \in \{-1, +1\}$ such that $x_i + c_i \in X_i$ and*

$$\operatorname{sgn} \frac{f_j((x_1, \ldots, x_{i-1}, x_i + c_i, x_{i+1}, \ldots, x_n)) - f_j(x)}{c_i} = \varepsilon \, .$$

We call $G(x)$ the local interaction graph of f in x.

To obtain the local interaction graph we consider changes in the values of the coordinate functions depending on small changes, i. e. changes by absolute value

1, in one component. The local interaction graph in x is thus closely related to the discrete Jacobian matrix, which was introduced in [13] in the Boolean case. Note that in the multi-value other than in the Boolean case it is possible that $G(x)$ contains parallel edges. There are at most two parallel edges from one vertex to another which then have opposite sign.

If we combine the structural information of the local interaction graphs for a set of states M we obtain a graph that contains all interactions influencing the network's dynamics in M.

Definition 2. *Let $M \subseteq X$. We denote by $G(M) := G(f)(M)$ the union of the graphs $G(x)$, $x \in M$. For $M = X$ we set $G(f) := G(X)$ and call $G(f)$ the* global interaction graph of f.

The global interaction graph contains all interactions influencing the network's dynamics in at least some part of state space.

When analyzing interaction graphs we are interested in certain structural motives. We focus on so-called *(feedback) circuits*. Here, a circuit is a tuple $(e_1, \dots, e_r)$ of edges $e_i = (k^i, l^i, \varepsilon) \in E$ such that all k^i, $i \in \{1, \dots, r\}$, are pairwise distinct, and $l^i = k^{i+1}$ for all $i \in \{1, \dots, r\}$ modulo r. The *sign of a circuit* is the product of the signs of its edges. Note that in a multigraph a circuit is not uniquely determined by its vertices.

In Fig. 1 we see on the left interaction graphs of the function $f = (f_1, f_2, f_3) : X \to X$, $X := \{0,1\}^2 \times \{0,1,2\}$, introduced in the caption of the figure. Here, f_2 models the situation that α_2 is influenced by α_1 and α_3 via an OR-gate as long as the activity level of α_3 is below 2. However, if $\alpha_3 = 2$, then α_2 is repressed. The way α_3 influences α_2 thus depends on the current state of the system. If the system is in state $(1,0,0)$, then a small change in the α_3 value is not enough to reach the value 2. Thus, f_2 corresponds to a logical OR-function. Since $\alpha_1 = 1$, we have $f_2((1,0,x_3)) = 1$ for $x_3 \in \{0,1\}$. It follows that in the state $(1,0,0)$ small changes in α_3 do not influence the component value α_2. As a result, we see in Fig. 1 (a) that there is no edge from α_3 to α_2 in $G((1,0,0))$. In comparison, if we look at the state $(0,0,0)$, we have $f_2((0,0,1)) - f_2((0,0,0)) = 1$ and therefore we get a positive edge from α_3 to α_2 in $G((0,0,0))$. Lastly, when looking at state (0,0,1), α_3 influences α_2 via a positive edge, since the argument we just made is still valid. However, if we increase the activity level of α_3 to 2, then by definition of f_2 we have a negative influence of α_3 on α_2. Thus in the local interaction graph $G((0,0,1))$ there is a negative as well as a positive edge from α_3 to α_1. By definition all the local interaction graphs are subgraphs of the global interaction graph of f, which is shown in Fig. 1(d).

Here, edges in an interaction graph are not labeled with additional information pertaining the activity level values of the tail vertex which allow that edge to have an effect on the dynamics. Thus, an edge (i, j, ε) may represent several influences of sign ε from i on j, which may differ in strength and depend on the current value of α_i. Based on this observation we introduce the following notion.

Definition 3. *Let $M \subseteq X$ and let $e := (i, j, \varepsilon)$ be an edge in $G(M)$. We call e* unique in M *if there exists $t^{ij} \in \{0, \dots, p_{i-1}\}$ such that $f_j(x) = f_j(x')$ for all*

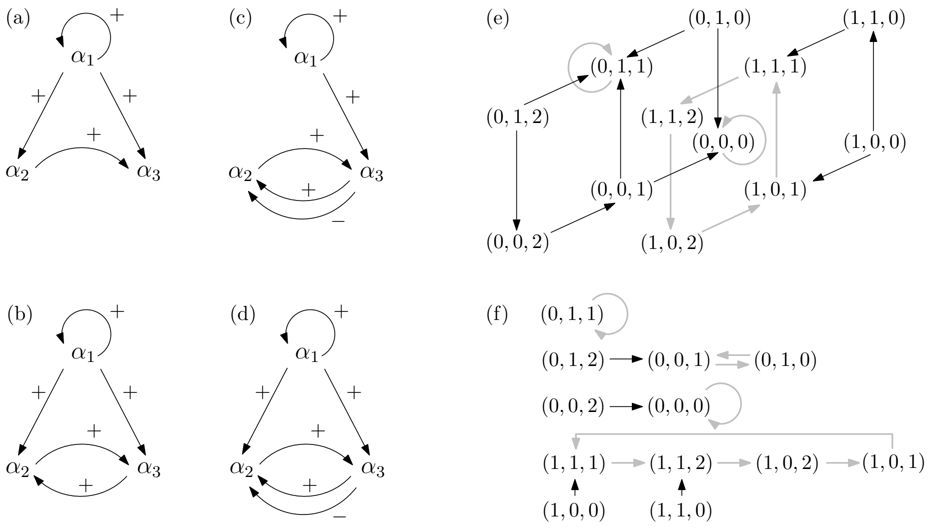


Fig. 1. In (a) to (c) local interaction graphs in states (1,0,0), (0,0,0) and (0,0,1), respectively, of the function $f = (f_1, f_2, f_3) : \{0,1\}^2 \times \{0,1,2\} \to \{0,1\}^2 \times \{0,1,2\}$ with $f_1(x) = x_1$, $f_2(x) = 0$ if $x_3 = 2$ and $f_2(x) = x_1 + x_3 - x_1 \cdot x_3$ otherwise, and $f_3(x) = x_1 + x_2$ for $x = (x_1, x_2, x_3) \in \{0,1\}^2 \times \{0,1,2\}$. In (d) the global interaction graph of f. In (e) and (f) the asynchronous and synchronous state transition graph, respectively. Heavier gray edges indicate attractors.

$x, x' \in M$ satisfying $x_i, x'_i \in \{0, \dots, t^{ij}\}$ or $x_i, x'_i \in \{t^{ij}+1, \dots, p_i\}$, and $x_k = x'_k$ for all $k \neq i$.

Whether or not the edge e has an impact on the dynamical behavior may still be dictated by the values of components other than α_i. However, if all component values x_l, $l \neq i$, are fixed, the value of $f_j(x)$ solely depends on whether x_i is above or below the threshold t^{ij}.

2.2 Dynamics

The function f determines the behavior of the network. However, there are different possibilities to derive the actual trajectories of the system. The entirety of the dynamical behavior, in any case, is captured in a *state transition graph* the paths of which represent all possible behaviors. The most straight forward approach leads to the following definition.

Definition 4. *With $S^s := S^s(f)$ we denote the directed graph with vertex set X and edge set $\{(x, f(x)) \mid x \in X\}$. We call S^s the* synchronous state transition graph *of f.*

Here, each state x has a unique successor, $f(x)$, and trajectories are obtained by applying f iteratively on an initial state. The underlying assumption concerning the evolution of the system is that all activity level changes indicated by f are executed concurrently. This is a highly simplifying assumption. Changes in activity level may represent very different biological processes and it is not realistic to assume that these processes have the exact same duration. If we want to incorporate this observation, we can make the assumption that a state differs from its successor in at most one component. Since data on such time delays is often lacking, we have no way to decide which of the components in a state where multiple component activity level changes are indicated by f should be executed. Therefore, we consider all possibilities and derive a non-deterministic representation of the dynamical behavior. Furthermore, we take into account that although f may indicate an activity level change of absolute value greater than one, the system nevertheless will behave in some sense continuously. That is, the activity levels of a state and its successor should differ by at most 1.

Definition 5. *Let $S^a := S^a(f)$ be a directed graph with vertex set X. For states $x = (x_1, \dots, x_n), x' = (x'_1, \dots, x'_n) \in X$ there is an edge $x \to x'$ if and only if $x' = f(x) = x$ or $x'_i = x_i + \mathrm{sgn}(f_i(x) - x_i)$ for some $i \in \{1, \dots, n\}$ satisfying $x_i \neq f_i(x)$, and $x'_j = x_j$ for all $j \neq i$. We call S^a the* asynchronous state transition graph *of f.*

To analyze state transition graphs we use, in addition to standard terminology from graph theory such as paths and cycles, the following concepts.

Definition 6. *Let $S \in \{S^s, S^a\}$. An infinite path $(x_0, x_1, \dots)$ in S is called* trajectory. *A nonempty set of states D is called* trap set *if every trajectory starting in D never leaves D. A trap set A is called* attractor *if for all $x^1, x^2 \in A$ there is a path from x^1 to x^2 in S. Attractors of cardinality greater than one are called* cyclic *attractors. A cycle $C := (x^1, \dots, x^r, x^1)$, $r \geq 2$, is called a* trap cycle *if every x^j, $j \in \{1, \dots, r\}$, has only one outgoing edge in S, i. e., the trajectory starting in x^1 is unique. A state x is called* steady state, *if there exists an edge $x \to x$, i. e. if $f(x) = x$.*

In other words, the attractors correspond to the terminal strongly connected components of the graph. In a synchronous state transition graph the trajectory starting from some initial state is unique. In consequence, every attractor is either a fixed point of f or a trap cycle, i. e. a periodic point of f. Since the state space is finite, every trajectory leads to an attractor. This is not true for asynchronous state transition graphs. However, it is easy to see that for every state x exists a trajectory starting in x leading to an attractor. Steady states and trap cycles are attractors, but there may also be attractors of cardinality greater than one which are not trap cycles. Since steady states are fixed points of f, the steady states, other than attractors of cardinality greater than one, in the synchronous and the asynchronous state transition graph coincide.

In Fig. 1 (e) and (f) we see the asynchronous and synchronous state transition graph of the function f defined in the caption. The system has two steady states, (0,1,1) and (0,0,0). The asynchronous state transition graph contains one cyclic attractor, namely $\{(1,0,1), (1,1,1), (1,1,2), (1,0,2)\}$, which is also an attractor in S^s. In the synchronous state transition graph we find a further cyclic attractor, $\{(0,0,1),(0,1,0)\}$.

We close this section with the following observation. If some coordinate function f_i is constant with value c, then $x_i = c$ for every state x in an attractor Similarly, we know the values x_j of every component j such that f_j depends only on values of components whose dynamics are described by constant coordinate functions. That is, we can easily determine the dynamical behavior of such components, which leads to the same fixed values of those components in every attractor of the system. Throughout the remainder of the paper we assume that *no coordinate function of f is constant.* We still allow the system to have input values in the sense of components maintaining their current activity level independent of the values of the other components. They can be modeled with the coordinate function $f_i(x) = x_i$.

3 Symbolic Steady States

Analysis of complex network dynamics, in particular of asynchronous state transition graphs, is costly. However, complex networks are often composed of smaller building blocks – modules and motifs (see e. g. [2, 1]). When analyzed in isolation such building blocks reveal specific biological functions. The question of interest is whether or not the behavior observed in isolation can be rediscovered in the complex network. Here, the building blocks themselves interact and influence each others behavior. One goal of this paper is to find conditions that allow to infer behavioral properties of the complex system from the dynamics of suitable subnetworks. The central notions in this endeavor are introduced in the following two definitions. It has already been used for Boolean functions in [15, 14] and uses notation first introduced in [12].

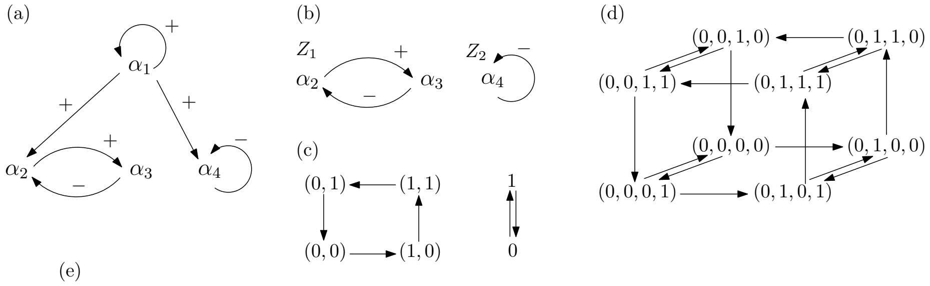
Definition 7. *For all $i \in \{1,\dots,n\}$ we set $X_i^\theta := \{0,\dots,p_i,\theta\}$ and $X^\theta := X_1^\theta \times \dots \times X_n^\theta$, where θ is a symbolic value. We call the elements of X^θ* states. *If no component of a state has value θ, the state is called* regular state, *otherwise it is called* symbolic state. *We denote $J(x) := \{i \in \{1,\dots,n\} \mid x_i = \theta\}$ for all $x \in X^\theta$.*

The value θ is used to describe uncertainty of a component value. Following this idea, we define a so-called *qualitative value* $|a,b|$ for $a,b \in \{0,\dots,\max_i p_i\}$, $a \le b$ by setting $|a,a| := a$ and $|a,b| := \theta$ if $a < b$. Furthermore, we denote $[x] := \{x' \in X \mid x'_j = x_j \text{ for all } j \notin J(x)\}$ for all $x \in X^\theta$.

Definition 8. *For all $i \in \{1,\dots,n\}$ we define $f_i^{\min} : X^\theta \to X_i^\theta$, $f_i^{\min}(x) := \min\{f_i(x') \mid x' \in [x]\}$ and $f_i^{\max} : X^\theta \to X_i^\theta$, $f_i^{\max}(x) := \max\{f_i(x') \mid x' \in [x]\}$. Then we define*

$$f^\theta : X^\theta \to X^\theta \quad \text{by} \quad f_i^\theta(x) = |f_i^{\min}(x), f_i^{\max}(x)|\,,\ i \in \{1,\dots,n\}.$$

We call $x \in X^\theta$ a steady state *if $f^\theta(x) = x$.*



(e)

$f(x) = x$	attractors (async)	attractors (sync)
$(0, \theta, \theta, \theta)$	$\{x \in X \mid x_1 = 0\}$	$\{(0,0,0,0), (0,1,0,1), (0,1,1,0), (0,0,1,1)\}$ $\{(0,0,0,1), (0,1,0,0), (0,1,1,1), (0,0,1,0)\}$
$(1, 1, 1, \theta)$	$\{(1,1,1,0), (1,1,1,1)\}$	$\{(1,1,1,0), (1,1,1,1)\}$
$(2, 1, 1, 1)$	$\{(2,1,1,1)\}$	$\{(2,1,1,1)\}$

Fig. 2. In (a) the global interaction graph of $f = (f_1, f_2, f_3, f_4) : X \to X$, $X := \{0, 1, 2\} \times \{0, 1\}^3$ with $f_1(x) = x_1$, $f_2(x) = 1$ if $x_1 \geq 1$ or $x_3 = 0$, and $f_2(x) = 0$ otherwise, $f_3(x) = x_2$, and $f_4 = 1$ if $x_1 = 2$ and $f_4(x) = 0$ otherwise. The graph $G^\theta([(0, \theta, \theta, \theta)])$ is shown in (b). The corresponding state transition graphs, and the derived compositional attractor in S^a of f in (c) and (d), respectively. In (e) a table of attractors derived from subgraphs induced by steady states given in the left column.

By definition we have $f^\theta|_X = f$, since regular states are mapped to regular states. If $f_i(x) \in X_i$ for a symbolic state x, we can deduce that $f_i(x) = f_i(y)$ for all $y \in [x]$. That is, the information inherent in the regular components of x is sufficient to determine the evolution of the i-th component. This can be visualized by looking at the local interaction graph $G([x])$, which then does not contain edges from any $\alpha_j \in J(x)$ to α_i.

Since we made the assumption that no coordinate function of f is constant, we know that the state $(\theta, \ldots, \theta)$ is a steady state. In the following we are interested in symbolic steady states x with $J(x) \neq \{1, \ldots, n\}$. The set of components not belonging to $J(s)$ remains dynamically stable regardless of value changes in the symbolic components. This allows us to consider a reduced network for analyzing the system's behavior in the subset $[x]$ of state space without loosing information. The regular components of the symbolic steady state act as a stable or *frozen core*, as was described first by S. Kauffman for random Boolean networks (see [5] for an overview).

To formalize the impact of a symbolic steady state on the dynamics and relate certain structural characteristics, we introduce the following notation for a symbolic steady state x. By $G^\theta([x])$ we denote the (multi-)graph with vertex set $V^\theta[x] := J(x)$ and edge set $E^\theta[x] := \{(i, j, \varepsilon) \in G([x]) \mid i, j \in J(x)\}$. We call a graph $Z = (V_Z, E_Z)$ component of $G^\theta([x])$, if the undirected graph derived from Z is a maximal connected subgraph of the undirected graph derived from $G^\theta([x])$. Figure 2 (b) shows the graph $G^\theta([x])$ for the symbolic steady state $x := (0, \theta, \theta, \theta)$ of the function f defined in the corresponding caption.

Proposition 1. *Let $x \in X^\theta$ be a symbolic steady state and let Z be a component of $G^\theta([x])$. Then every vertex of Z has at least one predecessor in Z. In particular, Z contains a circuit.*

Proof. Let α_i be a vertex in Z. Since $f_i(x) = x_i = \theta$, we have $f_i^{\min}(x) \neq f_i^{\max}(x)$. Thus, according to the definition, f_i depends on some $\alpha_j \in J(x)$ and we find an edge (j, i, ε) in Z for some $\varepsilon \in \{+, -\}$. Since the vertex set of Z is finite, there has to be a circuit in Z. $\square$

In the next lemma we show that the regular components of a symbolic steady state x stay fixed regardless of value changes in $J(x)$, and that the components of $G^\theta([x])$ act dynamically independent from each other in the state set $[x]$. This has already been shown for Boolean functions in [14] and the proof can be adapted easily.

Lemma 1. *Let $x \in X^\theta$ be a symbolic steady state, and let $Z_1, \ldots, Z_m$ be the components of $G^\theta([x])$. Consider a union Z of arbitrary components Z_j. Let $\tilde{x} \in X^\theta$ such that $\tilde{x}_i = x_i$ for all $i \notin Z$. Then $f_i^\theta(\tilde{x}) = f_i^\theta(x) = x_i = \tilde{x}_i$ for all $i \notin Z$.*

Proof. First, let us consider $i \notin J(x)$. Since $x_j = \theta$ for all $j \in Z$, we have $J(\tilde{x}) \subseteq J(x)$. Therefore, $[\tilde{x}] \subseteq [x]$. It follows that $f_i^{\min}(x) \leq f_i^{\min}(\tilde{x}) \leq f_i^{\max}(\tilde{x}) \leq f_i^{\max}(x)$. Since $f_i^\theta(x) = x_i$ is regular, we know $f_i^{\min}(x) = f_i^{\max}(x) = x_i$. Thus, $f_i^{\min}(\tilde{x}) = f_i^{\max}(\tilde{x}) = x_i$ and $f_i^\theta(\tilde{x}) = x_i = \tilde{x}_i$.

Now, let us consider $i \in J(x) \setminus Z$. We need to show that $f_i^\theta(\tilde{x}) = \theta$. Assume $f_i^\theta(\tilde{x}) = c \in X_i$, i.e. $f_i^{\min}(\tilde{x}) = f_i^{\max}(\tilde{x}) = c$. Since $f_i^\theta(x) = \theta$ there exist $y, y' \in [x]$ and $a, b \in X_i$ such that $f_i(y) = a < b = f_i(y')$. We may assume that $a \neq c$, since $a \neq c$ or $b \neq c$. It follows that $y \notin [\tilde{x}] \subseteq [x]$. Since x and $\tilde{x}$ only differ in Z-components, there exists $\tilde{y} \in [\tilde{x}]$ with $y_j = \tilde{y}_j$ for all $j \notin Z$. Then we find regular states $y =: y^1, y^2, \ldots, y^k := \tilde{y}$ in $[x]$ such that for each $l < k$ exists $j^l \in Z$ with $|y^l_{j^l} - y^{l+1}_{j^l}| = 1$ and $y^l_m = y^{l+1}_m$ for $m \neq j^l$, and $y^l \notin [\tilde{x}]$ for all $l < k$. Since $f_i(y^k) = c \neq a$, we find y^l such that $f_i(x^l) = a \neq f_i(y^{l+1})$. According to the definition of the local interaction graph, we then find an edge from j^l to i in $G(y^l)$, and thus also in $G([x])$. This is a contradition to Z being a union of components of $G^\theta([x])$, since $j^l \in J(x) \setminus Z$. $\square$

This lemma allows us to focus on the dynamics of the subnetworks represented by the components of $G^\theta([x])$ and to derive the dynamical behavior of the original network in $[x]$ by a simple reassembly, as we will see in the next section.

4 Compositional Attractors

Symbolic steady states allow us to identify subnetworks that govern the behavior of the original network at least in some part of state space. We focus in this section on finding attractors of f by looking at the dynamics restricted to the components of the local interaction graph of a symbolic steady state.

We need the following notation. Let $x \in X^\theta$ be a symbolic steady state, and let Z be a component of $G^\theta([x])$. Let $k := \operatorname{card} V_Z$ be the cardinality of V_Z. We may assume that $V_Z = \{\alpha_{l+1}, \dots, \alpha_{l+k}\}$ for some $l \in \{0, \dots, n-1\}$. Set $X^Z := X_{l+1} \times \cdots \times X_{l+k}$. We define $f^Z : X^Z \to X^Z$, $f^Z := \pi^Z \circ f^\theta \circ \rho^Z$, where $\rho^Z : X^Z \to X^\theta$ with $\rho^Z_i(z) = x_i$ for $i \notin Z$ and $\rho^Z_i(z) = z_{i-l}$ for $i \in Z$, and $\pi^Z : X^\theta \to X^Z$ is the projection on the components of Z. The function f^Z maps regular states to regular states, since Z is disjoint from all other components $\alpha_j \in J(x)$ in $G^\theta([x])$ and thus $f^\theta_i(\rho(z)) \in X_i$ for $i \in Z$. Proposition 1 ensures that no coordinate function of f^Z is constant. Furthermore, it is easy to see that the global interaction graph $G(f^z)$ is isomorphic to Z. We denote the synchronous and asynchronous state transition graph derived from f^Z by S^s_Z and S^a_Z, respectively.

In the remainder of the section let x be a symbolic steady state of f, and let $Z_1, \dots, Z_m$ be the components of $G^\theta([x])$. W. l. o. g. we may assume that Z_1 contains the vertices $\alpha_1, \dots, \alpha_{\operatorname{card} Z_1}$, $V_{Z_2} = \{\alpha_{\operatorname{card} Z_1 + 1}, \dots, \alpha_{\operatorname{card} Z_1 + \operatorname{card} Z_2}\}$, etc., and $k, \dots, n$ are the components of $\{1, \dots, n\} \setminus J(x)$ for some $k \in \{1, \dots, n\}$. To simplify notation we identify subsets of $X^{Z_1} \times \cdots \times X^{Z_m} \times \{(x_k, \dots, x_n)\}$ with subsets of X.

Theorem 1. *For all $i \in \{1, \dots, m\}$ let A_i be an attractor in $S^a_{Z_i}$. Then $A := A_1 \times \cdots \times A_m \times \{(x_k, \dots, x_n)\}$ is an attractor of the asynchronous state transition graph S^a of f. Moreover, every attractor in S^a with all its vertices in $[x]$ can be represented in this manner as Cartesian product of attractors in $S^a_{Z_i}$, $i \in \{1, \dots, m\}$, and $\{(x_k, \dots, x_n)\}$.*

Proof. We observe that $\pi^{Z_j}(f(x')) = \pi^{Z_j}(f^\theta(x')) = \pi^{Z_j}(f^\theta(\rho^{Z_j}(\pi^{Z_j}(x')))) = f^{Z_j}(\pi^{Z_j}(x'))$ for all $x' \in [x]$ and all $j \in \{1, \dots, m\}$ according to the definition of ρ^{Z_j}, π^{Z_j} and Lemma 1.

By definition we have $A \subseteq [x]$. We show that A is a trap set in S^a which is strongly connected. Let $a \in A$. If $f(a) = a$ then A is an attractor. Otherwise choose $i \in \{1, \dots, n\}$ such that $f_i(a) \neq a_i$. Consider $a' \in X$ with $a'_i = a_i + \operatorname{sgn}(f_i(a) - a_i)$ and $a'_j = a_j$ for $j \neq i$. Since $a \in [x]$ we have $f_j(a) = a_j$ for all $j \in \{k, \dots, n\}$. Thus there is $l \in \{1, \dots, m\}$ with $i \in Z_l$. Then we choose k^l such that $i = j^l + k^l$ with $j^l := \sum_{j=1}^{l-1} \operatorname{card} Z_j$. We have $a_i \neq f_i(a) = f^{Z_l}_{k^l}(\pi^{Z^l}(a))$. It follows that $a_i + \operatorname{sgn}(f_i(a) - a_i) = (\pi^{Z_l}(a))_{k^l} + \operatorname{sgn}(f^{Z_l}_{k^l}(\pi^{Z_l}(a)) - (\pi^{Z_l}(a))_{k^l})$. Since A^l is a trap set in $S^a_{Z_l}$, it follows that $\pi^{Z_l}(a') \in A_l$, and thus $a' \in A$.

To obtain a path from $a, a' \in A$ in S^a we construct the path componentwise in the attractors A_i. More precisely, we exploit the fact that if γ is path in $S^a_{Z_i}$ from z to z' for $z, z' \in X^{Z_i}$, then we find a path γ' in S^a from y to y' with $y_j = y'_j$ for all $j \notin Z_i$ which is projected on γ by π^{Z_i}. This is possible since the dynamics in $Z_1, \dots, Z_m$ do not influence each other according to Lemma 1 and since in the asynchronous state transition graph state changes indicated by f are executed componentwise. For a detailed elaboration of this argument in the Boolean case see [14], proof of Theor. 5.6. It follows, that A is strongly connected.

The same reasoning ensures that the projection $\pi^{Z_j}(A)$, $j \in \{1, \dots, m\}$, of an arbitrary attractor A of S^a is an attractor of $S^a_{Z_j}$. Edges leaving $\pi^{Z_j}(A)$ in

$S^a_{Z_j}$ would generate edges leaving A in S^a, and paths in A are projected on sequences of states in X^{Z_j} that constitute a path in $S^a_{Z_j}$, if we eliminate all but one consecutive identical states (due to non-injectivity of π^{Z_l}) in the sequence. □

From the proof we get an idea that the compositional properties inherent in x are not restricted to the construction of attractors. We have shown in [14] in the Boolean case that in fact the subgraph of S^a with vertex set $[x]$ is the composition of the graphs $S^a_{Z_i}$, $i \in \{1, \ldots, m\}$, and $\{x_k, \ldots, x_n\}$. The reasoning can be adopted for multi-valued functions.

In Fig. 2 (e) we see in the second column of the table attractors derived from the corresponding symbolic steady state in the first column. For the state $(0, \theta, \theta, \theta)$ we obtain two components of $G^\theta([x])$ as shown in Fig. 2 (b). In (c) the corresponding graphs $S^a_{Z_1}$ and $S^a_{Z_2}$ are shown, and in (d) we see the compositional attractor in the state transition graph S^a of f.

In [4, 3] the authors consider the case that $G(f)$ is not connected and show how to derive attractors of the synchronous state transition graph from the subnetwork dynamics corresponding to the graph components. Lemma 1 allows us to apply their results to the components $Z_1, \ldots, Z_m$ and obtain the following statement concerning the synchronous state transition graph of f.

Theorem 2. *For all $i \in \{1, \ldots, m\}$ let A_i be an attractor in $S^s_{Z_i}$. Then the set $\mathcal{A} := A_1 \times \cdots \times A_m \times \{(x_k, \ldots, x_n)\}$ is a union of attractors in S^s such that the cardinality of each attractor $A \in \mathcal{A}$ is the least common multiple of the cardinalities of the attractors A_i, $i \in \{1, \ldots, m\}$, and, as a consequence, the number of attractors in $\mathcal{A}$ is $\prod_{j=2}^{m}((\operatorname{card} A_1 \bullet \operatorname{card} A_2) \bullet \cdots \bullet \operatorname{card} A_{j-1}) \star \operatorname{card} A_j$, where $\bullet$ denotes the least common multiple and $\star$ the greatest common divisor operation. Moreover, for every attractor $A \subseteq [x]$ in S^s and every $j \in \{1, \ldots, m\}$ is $\pi^{Z_j}(A)$ an attractor in $S^s_{Z_j}$.*

It is easy, if tedious, to describe the states of a compositional attractor in S^s. It is basically a concatenation of the steady states resp. cycles of the components. Since we update all components at once in every step, cycles in $S^s_{Z_i}$ can generate more than one cycle in S^s. This is illustrated by the example in Fig. 2. The last column of the table in (e) shows attractors in S^s derived from different symbolic steady states. The composition of the two attractors shown in (c) yield two cycles in the synchronous dynamics.

When comparing the synchronous and the asynchronous case the differences become apparent when looking at the attractors derived from component attractors of cardinality greater than one. In the synchronous case composition of component attractors may result in a greater number of attractors, while in the asynchronous case only the attractor size increases. A more detailed description of the relation between cyclic attractors in synchronous and asynchronous dynamics necessarily would include observations on the way edges are generated during composition of attractors in the asynchronous case as well as on effects of the gradual activity level change of the asynchronous update.

We close this section with a simple corollary from the two preceding theorems, which for the synchronous case was already formulated in [3].

Corollary 1. *For $i \in \{1, \dots, m\}$ let $S^{\delta}_{Z_i}$, $\delta \in \{a, s\}$, contain N_i attractors with cardinalities L_{ij}, $j \in \{1, \dots, N_i\}$. Set $I := I_1 \times \dots \times I_m$ with $I_l := \{1, \dots, N_l\}$.*

- *In the asynchronous state transition graph of f, the number of attractors with vertices in $[x]$ is $\prod_{j=1}^{m} N_j$ and the maximal attractor cardinality in $[x]$ is $\max_{(k_1,\dots,k_m)\in I} \prod_{j=1}^{m} L_{jk_j}$.*
- *In the synchronous state transition graph of f, the number of attractors with vertices in $[x]$ is $\sum_{(k_1,\dots,k_m)\in I} \prod_{j=2}^{m}((L_{1k_1} \bullet L_{2k_2}) \bullet \dots \bullet L_{j-1k_{j-1}}) \star L_{jk_j}$, and $\max_{(k_1,\dots,k_m)\in I}((L_{1k_1} \bullet L_{2k_2}) \bullet \dots \bullet L_{mk_m})$ is the maximal attractor cardinality in $[x]$.*

5 Circuits and Attractors

In the preceding section we constructed attractors in the dynamics of f from attractors of functions f^{Z_i} associated with certain subnetworks. The same reasoning allows us to formulate more general relations between structural characteristics of the subnetworks and the dynamics of f. We focus in this section on the impact of feedback circuits in the structure on number and size of attractors of f.

For multi-valued discrete functions it was shown in [11] that the existence of a positive circuit in the global interaction graph (and even in certain local interaction graphs) is a necessary condition for the existence of two attractors in the asynchronous state transition graph. Complementary, the existence of a negative circuit in the global interaction graph is necessary for f to have a cyclic attractor (see [9]). Obviously, the result on positive circuits also holds in the synchronous case if we specify the attractors to be steady states. However, simple examples show that the second result is false in the synchronous case.

We now focus on functions the global interaction graph of which basically consist of a circuit.

Lemma 2. *Assume the global interaction graph $G(f)$ contains only one circuit C, and let V_C be the set of vertices visited by C. Then $f_i(x) = f_i(x')$ for all $i \in V_C$ and $x, x' \in X$ with $x_j = x'_j$ for all $j \in V_C$.*

Proof. Assume we find $i \in V_C$ and $x, x' \in X$ such that $x_j = x'_j$ for all $j \in V_C$ and $f_i(x) \neq f_i(x')$. Then there exist $y, y' \in X$ and $k \notin V_C$ such that $y_j = y'_j$ for all $j \neq k$, $|y_k - y'_k| = 1$ and $f_i(y) \neq f_i(y'_k)$ (see proof of Lemma 1). It follows that $G(f)$ contains an edge from k to i. Since we always assume that no coordinate function is constant, every vertex in $G(f)$ has a predecessor. Since the set of vertices is finite, we then find a circuit other than C in $G(f)$ which is a contradiction. □

In [7] it is shown that Boolean functions associated with isolated circuits always display a characteristic behavior depending on their sign, both in the synchronous and the asynchronous case. We use this result to prove the following lemma.

Lemma 3. *Assume the global interaction graph $G(f)$ contains only one circuit C and that all edges of C are unique in X. Then there are at least two attractors in S^s as well as S^a, if C is a positive circuit. If C is negative, there exists a cyclic attractor in S^s and in S^a.*

Proof. Let us assume C corresponds to the vertex sequence $(\alpha_1, \ldots, \alpha_r)$. Set $X_C := X_1 \times \cdots \times X_r$ and let $z_i \in X_i$ for $i \in \{r+1, \ldots, n\}$. We set $f^C : X^C \to X^C$ by $f^C := \pi^C \circ f \circ \rho^C$, where $\rho^C : X^C \to X, y \mapsto (y_1, \ldots, y_r, z_{r+1}, \ldots, z_n)$ and π^C is the projection on the components of C. According to Lemma 2, we have $f_i^C(\pi^C(x)) = \pi^C(f(x))$. It follows that $G(f^C) = C$, i.e., values of coordinate function f_i^C depend only on the value of the predecessor in C. Since furthermore all edges of C are unique (see Def. 3), we find for all $i \in \{1, \ldots, r\}$ values $t^{i,i+1} \in \{0, \ldots, p_i - 1\}$ such that $f_{i+1}(y) = f_{i+1}(y')$ for all $y, y' \in X^C$ with either $y_i, y'_i \leq t^{i,i+1}$ or $y_i, y'_i > t^{i,i+1}$, indices taken modulo r. This allows us to define a Boolean function $f^{\mathcal{B}} : \{0,1\}^r \to \{0,1\}^r$ by $f^{\mathcal{B}} := \pi^t \circ f^C \circ \rho^t$ with $\rho^t : \{0,1\}^r \to X^C$, $\rho_j^t(0) = c_j$, $\rho_j^t(1) = c'_j$ for arbitrary but fixed $c_j \in \{0, \ldots, t^{j,j+1}\}$, $c'_j \in \{t^{j,j+1} + 1, \ldots, p_j\}$, and $\pi^t : X^C \to \{0,1\}^r$, $\pi_j^t(y) = 0$ if $y_j \leq t^{j,j+1}$ and $\pi_j^t(y) = 1$ if $y_j > t^{j,j+1}$. Note that $f^{\mathcal{B}}$ does not depend on the choice of c_j, c'_j. It is easy to see that $G(f^{\mathcal{B}})$ coincides with C.

In [7] was shown that $f^{\mathcal{B}}$ has at least two fixed points, if C is positive. Let b be a fixed point of $f^{\mathcal{B}}$. Choose $y' \in (\pi^t)^{-1}(b)$ and set $y_i := f_i^C(y')$ for all $i \in \{1, \ldots, r\}$. Since b is a fixed point, we can deduce that $y = (y_1, \ldots, y_r) = f^C(y') \in (\pi^t)^{-1}(b)$. It follows from the definition of π^t that for all $i \in \{1, \ldots, r\}$ either $y_i, y'_i \leq t^{i,i+1}$ or $y_i, y'_i > t^{i,i+1}$. The uniqueness condition then yields $f_i^C(y) = f_i^C(y') = y_i$. Thus y is a fixed point of f^C.

We then find a trajectory in S^δ, $\delta \in \{a, s\}$, starting in y and leading to an attractor A. According to Lemma 2 we then have $f_i(a) = y_i$ for all $i \in \{1, \ldots, r\}$ and every $a \in A$. Since we have two different fixed points of f^C, if C is positive, we also find two different attractors in S^δ.

If C is negative, then $f^{\mathcal{B}}$ does not have a fixed point (see again [7]). It follows from the definition of $f^{\mathcal{B}}$, in particular from its independence of the choice of c_j and c'_j, that f^C does not have a fixed point either. Thus, the synchronous as well as the asynchronous state transition graph of f^C contain a cyclic attractor. Lemma 2 then again yields the existence of a cyclic attractor in the asynchronous and synchronous state transition graph of f. □

The uniqueness of edges is exploited in the proof to obtain a suitable projection on the Boolean case. However, it seems likely that the statement remains true when dropping the condition. A finer partition of the range of a component corresponding to multiple thresholds could clarify the situation.

By applying the above lemma and Theorems 1 and 2 or Cor. 1 we immediately obtain the following theorem.

Theorem 3. *Let x be a symbolic steady state. Assume a component Z of $G^{\theta}([x])$ contains only one circuit C and that all edges of C are unique. If C is positive, then there exist at least two attractors in S^a as well as in S^s. If C is negative, then there is a cyclic attractor in both S^a and S^s.*

The theorem basically states that circuits embedded in complex networks imprint some of the characteristics they show in isolation on the whole network, if we can in some sense recover isolation in at least a part of state space. The statement does not hold, if the circuit C is not the only circuit in Z (see Boolean examples in [15, 14]) Furthermore, the hypothesis given in the last theorem is a sufficient but not a necessary condition for the existence of multiple resp. cyclic attractors. This is illustrated by an example given in [15], Fig. 4.

6 Conclusion

Analyzing complex networks is a difficult task. Even if the number of components of a discrete regulatory network is in some sense manageable, we have to deal with the problem of analyzing the dynamics in an exponentially large state space. A well known idea to approach this difficulty is to identify smaller building blocks of the system the study of which in isolation still renders information on the dynamics of the whole network. In this paper, we introduce the notion of symbolic steady state which allows us to identify such building blocks, systematically extending ideas developed for Boolean functions and asynchronous dynamics in [15] and [14]. We state explicit rules how to derive attractors of the network from subnetwork attractors valid for synchronous as well as asynchronous dynamics. Illustrating those rules, we derive general conditions for circuits embedded in the network to transfer their behavioral characteristics pertaining number and size of attractors observed in isolation to the complex network.

Several aspects can be addressed in future work. Stronger results are certainly possible if we refine the representation of component values via the symbolic value. Instead of merging the whole range of a component to one symbolic value we could partition it into several symbolic values, which would allow for a more precise localization in phase space. On the structural side, an even more accurate understanding of the interactions governing asymptotic behavior in the area of state space associated with a symbolic steady state x would be possible when considering the local interaction graph $G(M)$ for a set $M \subseteq [x]$ derived from $[x]$ by eliminating in some sense dynamically irrelevant states. Here a comparison of $[x]$ with its forward orbit under f may be helpful. Besides refining the results derived from the existence of symbolic steady states, we would also like to provide criteria that guarantee the existence and allow explicit calculation of symbolic steady states. In [14] we have provided some results in this direction for a certain class of Boolean networks. Further results seem possible when focussing on (nested) canalyzing functions, which are known for generating frozen cores in Kauffman networks. A related question of interest is whether we can link existence and properties of symbolic steady states to general characteristics of the function f. Lastly, a very important step is to apply the methods to established biological network models. This would allow not only for testing the suitability of the approach to the dynamical analysis, but also for a comparison of the subnetworks derived from symbolic steady states with network modules of known biological importance.

References

1. U. Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8:450–461, 2007.
2. C. Christensen and R. Albert. Using graph concepts to understand the organization of complex systems. *IJBC*, 17(7):2201–2214, 2007.
3. E. Dubrova. Random multiple-valued networks: Theory and applications. In *Proceedings of 36th IEEE International Symposium on Multiple-Valued Logic (ISMVL'2006), Singapore*, pages 27–33, 2006.
4. E. Dubrova and M. Teslenko. Compositional properties of random boolean networks. *Physical Review E*, 71:056116, 2005.
5. S. Kauffman. *The Origins of Order.* Oxford University Press, 1993.
6. R. Laubenbacher and P. Mendes. A discrete approach to top-down modeling of biochemical networks. In *Computational Systems Biology*, pages 229–247. Elsevier Academic Press, 2005.
7. É. Remy, B. Mossé, C. Chaouiya, and D. Thieffry. A description of dynamical graphs associated to elementary regulatory circuits. *Bioinform.*, 19:172–178, 2003.
8. É. Remy, P. Ruet, and D. Thieffry. Graphic requirements for multistability and attractive cycles in a boolean dynamical framework. *Adv. Appl. Math.*, 41(3):335–350, 2008.
9. A. Richard. On the link between oscillations and negative circuits in discrete genetic regulatory networks. JOBIM, 10-12 juillet, Marseille, France, 2007.
10. A. Richard. Positive circuits and maximal number of fixed points in discrete dynamical systems. Preprint arXiv:0807.4229v3 [cs.DM], 2008.
11. A. Richard and J.-P. Comet. Necessary conditions for multistationarity in discrete dynamical systems. *Discr. Appl. Math.*, 155(18):2403–2413, 2007.
12. A. Richard, J.-P. Comet, and G.Bernot. R. Thomas' modeling of biological regulatory networks: introduction of singular states in the qualitative dynamics. *Fundamenta Informaticae*, 65:373–392, 2005.
13. F. Robert. *Discrete Iterations: A Metric Study*, volume 6 of *Springer Series in Computational Mathematics.* Springer, 1986.
14. H. Siebert. Deriving behavior of boolean bioregulatory networks from subnetwork dynamics. *Math. Comp. Sci.* To appear.
15. H. Siebert and A. Bockmayr. Relating attractors and singular steady states in the logical analysis of bioregulatory networks. In *Algebraic Biology, AB 2007, Castle of Hagenberg, Austria*, volume 4545 of *LNCS*, pages 36–50. Springer, 2007.
16. R. Thomas and M. Kaufman. Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits. *Chaos*, 11:180–195, 2001.