

# Multistability equivalence between gene regulatory networks of different dimensionality

Daniella Schittler, Frank Allgöwer, and Steffen Waldherr

**Abstract**—This paper formulates and addresses the problem of equivalence in terms of multistability properties between gene regulatory systems of different dimensionality. Given a dynamical system of a gene regulatory network, and the structure of another, higher-dimensional gene regulatory network, the aim is to find a dynamical system for the latter which has the same stability properties as the former. We propose construction rules for the dynamics of a high-dimensional system, given the low-dimensional system and the high-dimensional network structure. These construction rules yield a multistability-equivalent system, as we prove in this work. The usefulness of our method is illustrated with an exemplary gene regulatory network, for which we construct a dynamical system which has the same multistability properties as a given, low-dimensional model of gene regulation.

## I. INTRODUCTION

Gene regulatory networks (GRNs) in biological cells are high-dimensional systems. However, their dynamics are commonly studied via ordinary differential equation (ODE)-based models that represent abstracted, low-dimensional systems (e.g., [2], [3], [5], [6], [7], [8]). This approach is justified by the identification of core motifs that are responsible for generating certain properties such as bistability, switch-like behavior, oscillations, or adaptivity [11], [4].

A central property of GRNs is that many of them exhibit multistability, thereby realizing the coexistence of several stable states of the system, of which a suitable one can be adopted under the respective biological circumstances ([2], [5], [7], [8]). The analysis of multistability properties of such systems is of high interest and has attracted the development of new specially suited methods [3], [1], [9], [10].

It is desired that the analysis of a conceptual, low-dimensional model should allow for conclusions about a higher-dimensional model. On the one hand, low-dimensional systems are amenable to systems theoretic tools such as multistability and bifurcation analysis. On the other hand, a high-dimensional dynamic model might be required to allow for fitting to data from gene expression measurements, and thus for quantitative comparison with and predictions about the real biological system. From this the task arises how a dynamic model for a high-dimensional system can be obtained if only a core motif model of the gene

regulatory dynamics, implemented as a low-dimensional ODE system, is available. The high-dimensional ODE system to be obtained should have the same number of steady states, and each steady state should have the same stability properties. This is somewhat complementary to the field of model reduction, where one aims to reduce a high-dimensional model to a low-dimensional one.

If this question of multistability equivalence could be answered, methods and results from systems analysis of low-dimensional systems and modeling via high-dimensional systems could be combined and thus greatly enhance the understanding of a particular GRN system at hand. The aim of this paper is to propose construction rules for the expansion of a low-dimensional to a higher-dimensional ODE system, while ensuring multistability equivalence.

The outline of this paper is as follows: In Sec. II, the problem statement of multistability-equivalence in systems of different dimensionality is formulated. The next Sec. III, addresses this problem by proposing construction rules. In Sec. IV, we prove that a thereby constructed system is indeed multistability-equivalent. We apply our method to an example in Sec. V, and conclude with a discussion and outlook in Sec. VI.

## Notation

In the following, we denote by  $|\mathcal{A}|$  the cardinality of the set  $\mathcal{A}$ ; and by  $\|v\|$  the 1-norm of the vector  $v$ . By  $e_i^T$ , we denote the unit row vector with the  $i$ -th entry 1, all other entries 0. We define  $\mathbb{R}_+ := [0, \infty)$  and  $\mathbb{R}_{++} := (0, \infty)$ .  $\Re(\lambda)$  denotes the real part of the complex number  $\lambda$ . A complex number is said to be positive (non-negative), if its real part is positive (non-negative). Furthermore, we denote  $\mathbf{j} := \sqrt{-1}$ .

## II. PROBLEM STATEMENT

In this section we formulate the problem to be addressed. Given an ODE system

$$\Sigma_f : \dot{z} = f(z) = a(z) - d(z), \quad z \in \mathbb{R}_+^n, \quad (1)$$

with the activation rate  $a(z) \in \mathbb{R}_+^n$ , and the degradation rate  $d(z) = kz$ ,  $k = \text{diag}(k_1, \dots, k_n)$ , with  $k_i > 0$ .

We assume the system to have  $R$  steady states  $z^{*(r)}$ ,  $r = 1 \dots R$ :  $f(z^{*(r)}) = 0$ . The system's Jacobian at steady state  $z^{*(r)}$  is denoted by  $J_f^{(r)} := \frac{\partial f}{\partial z}(z^{*(r)})$ . Furthermore, we assume that the interaction sign matrix  $S_a := \text{sgn}\left(\frac{\partial a}{\partial z}\right) \in \{-1, 0, +1\}^{n \times n}$  is constant over  $z \in \mathbb{R}_{++}^n$ , and that the Jacobians  $J_f^{(r)}$ ,  $r = 1, \dots, R$  do not have eigenvalues on the imaginary axis.

D. Schittler and F. Allgöwer are with the Institute for Systems Theory and Automatic Control, Universität Stuttgart, 70550 Stuttgart, Germany daniella.schittler@ist.uni-stuttgart.de, frank.allgower@ist.uni-stuttgart.de

S. Waldherr is with the Institute for Automation Engineering, Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany steffen.waldherr@ovgu.de

The number of unstable modes (positive eigenvalues) for each steady state is given by  $|\{\lambda_{f,v}^{(r)} | \Re(\lambda_{f,v}^{(r)}) > 0\}|$ , with  $\lambda_{f,v}^{(r)}, v = 1 \dots n$  the eigenvalues of the Jacobian  $J_f^{(r)}$ .

An interaction sign matrix of a higher-dimensional system is given as

$$S_A \in \{-1, 0, +1\}^{N \times N}, N \geq n. \quad (2)$$

*Remark 1:* If, e.g., qualitative knowledge about gene interactions is available, which can be represented by an interaction graph, this can be directly translated into a sign matrix (2).

The following definition is made in order to formulate the concept of a higher-dimensional ODE system

$$\Sigma_F : \dot{x} = F(x) = A(x) - D(x), x \in \mathbb{R}_+^N, \quad (3)$$

with equivalent multistability properties.

*Definition 1:* A system (3) is called an  $N$ -dimensional multistability-equivalent system to (1) and consistent with the interaction structure given by (2), if the following hold.

- (i) The derivative of the activation function  $A(x)$  has signs as given by the interaction sign matrix (2):

$$\text{sgn} \left( \frac{\partial A}{\partial x} \right) = S_A \quad \forall x \in \mathbb{R}_{++}^N. \quad (4)$$

- (ii) There exists an injective map  $h : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^N, z^* \mapsto h(z^*)$ , with:
- $$f(z^*) = 0 \Leftrightarrow F(h(z^*)) = 0. \quad (5)$$

- (iii) The number of unstable modes (positive eigenvalues) in both systems (1) and (3), for each pair of steady states  $z^{*(r)}, x^{*(r)} = h(z^{*(r)})$ , is equal:

$$|\{\lambda_{F,u}^{(r)} | \Re(\lambda_{F,u}^{(r)}) > 0\}| = |\{\lambda_{f,v}^{(r)} | \Re(\lambda_{f,v}^{(r)}) > 0\}|. \quad (6)$$

In the remainder, a system that meets the properties of this Def. 1 will be called *multistability-equivalent* for short.

The class of considered systems is restricted for technical reasons, as follows:

*Assumption 1: (Modular structure)* Assume that, possibly by reordering the state space variables of system (3), the interaction sign matrix (2) fulfills the following structural property. There exist numbers  $m_i \in \{1, \dots, n\}$ ,  $i = n + 1, \dots, N$ , such that

- in rows  $i = n + 1, \dots, N$ , columns  $j = 1, \dots, n$ :  $S_A|_{i,m_i} = +1$  for no more than one  $m_i$ .
- in rows  $i = n + 1, \dots, N$ , columns  $j = n + 1, \dots, N$ :  $S_A|_{i,j} \in \{0, +1\}$  for all  $j$  where  $m_i = m_j$ , and  $S_A|_{i,j} = 0$  otherwise.

The general structure of such a matrix is given in Fig. 1. An example of a matrix fulfilling this property with  $n = 4$ ,  $N = 5$ ,  $m_3 = m_4 = 1$ , and  $m_5 = 2$  is given later in the example 25.

In the remainder, the first  $n$  state variables  $x_i$ ,  $i \in \{1, \dots, n\}$  will be referred to as “master genes”, whereas the remaining  $N - n$  state variables  $\{x_{n+1}, \dots, x_N\}$  will be referred to as “module genes”. In this way, each index  $i \in \{n + 1, \dots, N\}$  is uniquely assigned to one interaction module:  $i \in \mathcal{M}_{m_i}$ , such that

$i =$	1	-1,0,+1																							
	$\vdots$																								
	$n$																								
	$n + 1$	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px 10px;"><math>P_1</math></td> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;">0</td> </tr> <tr> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;"><math>P_{\dots}</math></td> <td style="padding: 2px 10px;">0</td> </tr> <tr> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;"><math>P_n</math></td> </tr> </table>			$P_1$	0	0	0	$P_{\dots}$	0	0	0	$P_n$	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px 10px;">0,+1</td> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;">0</td> </tr> <tr> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;">0,+1</td> <td style="padding: 2px 10px;">0</td> </tr> <tr> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;">0</td> <td style="padding: 2px 10px;">0,+1</td> </tr> </table>			0,+1	0	0	0	0,+1	0	0	0	0,+1
$P_1$	0				0																				
0	$P_{\dots}$				0																				
0	0	$P_n$																							
0,+1	0	0																							
0	0,+1	0																							
0	0	0,+1																							
$S_A =$	$\vdots$																								
	$\vdots$																								
	$\vdots$																								
	$\vdots$																								
	$\vdots$																								
	$N$																								

Fig. 1. General structure of a matrix fulfilling the structural requirements. Each  $P_m, m = 1 \dots n$  is a column vector of 0 and +1.

- interactions affecting  $x_i$  come either from  $x_{m_i}$ , or from other  $x_j$  belonging to the same interaction module,
- interactions from  $x_i$  go to genes that are either master genes or belong to the same module,
- interactions between genes belonging to the same module are nonnegative.

The classification into master and module genes with an according network structure might be predetermined from biological knowledge. If not, the selection of master genes could be addressed as a separate problem which is beyond the scope of this paper.

*Assumption 2: (Consistency of sign matrices  $S_a$  and  $S_A$ )* For each  $(i, j) \in \{1, \dots, n\} \times \{1, \dots, n\}$  with  $S_a|_{i,j} \neq 0$ , there must exist some simple path from  $j \in \{n + 1, \dots, N\}$  to  $i \in \{1, \dots, n\}$

$$p_{ij} := ((i, \iota_1), (\iota_1, \iota_2), \dots, (\iota_{\omega_{ij}-1}, \iota_{\omega_{ij}}), (\iota_{\omega_{ij}}, j))$$

$$\subseteq (\{1, \dots, N\} \times \{1, \dots, N\})^{(\omega_{ij}+1)}, \text{ for which}$$

$$\prod_{(\iota', \iota'') \in p_{ij}} S_A|_{\iota', \iota''} = S_a|_{i,j}.$$

Moreover, if there exists  $\iota_1 \in \{n + 1, \dots, N\}$  on some path from  $j$  to  $i$ :  $(i, \iota_1) \in p_{ij}$ , with  $S_A|_{i, \iota_1} \neq 0$ , there has to exist an interaction in the low-dimensional system

$$S_a|_{i,j} \neq 0. \quad (7)$$

If one of these assumptions is not fulfilled, then the proposed construction of an  $N$ -dimensional multistability-equivalent system can not be done in the same way, and therefore the existence of an  $N$ -dimensional multistability-equivalent system cannot be guaranteed.

### III. CONSTRUCTION RULES

In this section we propose a construction procedure for an  $N$ -dimensional multistability-equivalent system  $\Sigma_F$ , given  $\Sigma_f$  (1) and  $S_A$  (2). The idea of the construction is as follows: Additional interactions are introduced via linear activation functions. The remaining interactions are defined

in terms of the interactions from the low-dimensional system, with specific mappings between the state spaces of different dimensionality. These interactions are constructed such that the steady state gains of the additional interactions are exactly compensated when the system is at steady state.

**Step 1:** Construct functions  $F_i(x)$  for the module genes indices  $i = n + 1, \dots, N$ , as follows:

$$\begin{aligned} F_i(x) &= A_i(x) - D_i(x) \quad \text{with} \\ A_i(x) &= \sum_{j=1 \dots N} S_A|_{i,j} x_j, \\ D_i(x) &= K_i x_i, \text{ with } K_i \in \mathbb{R}_{++} \text{ to be chosen.} \end{aligned} \quad (8)$$

**Step 2:** Next, the influence of master genes on the module genes is captured by defining the following transfer gains. For all  $(i, k)$ ,  $k \in \{1, \dots, n\}$ ,  $i \in \mathcal{M}_k \subseteq \{n + 1, \dots, N\}$ , given a system (3), denote the transfer gain with input  $x_k$  and output  $x_i$ , via the system matrix  $(\partial F / \partial x)_{l \in \mathcal{M}_k, m \in \mathcal{M}_k}$ , i.e. restricted to the state variables in the module  $\mathcal{M}_k$ :

$$\begin{aligned} G_{k \rightarrow i}(\lambda) &= (e_i^T |_m)_{m \in \mathcal{M}_k} \\ &\cdot \left( \lambda I_{|\mathcal{M}_k|} - \left( \frac{\partial F}{\partial x} \Big|_{l, m} \right)_{l \in \mathcal{M}_k, m \in \mathcal{M}_k} \right)^{-1} (S_A|_{l, k})_{l \in \mathcal{M}_k}. \end{aligned} \quad (9)$$

Then define parameters  $\gamma_{ik}$  which give the corresponding steady state gain for input  $x_k$  and output  $x_i$ :

$$\gamma_{ik} := G_{k \rightarrow i}(0). \quad (10)$$

**Step 3:** In this step, functions  $F_i(x)$  for the master genes indices  $i = 1, \dots, n$  are constructed. Therefore, construct the  $2n^2$   $N$ -dimensional index vectors  $J_{i, \nu}^-$  and  $J_{i, \nu}^\neq$ ,  $i, \nu = 1, \dots, n$  as follows: The  $j$ -th entry of  $J_{i, \nu}^-$  is set to

$$J_{i, \nu}^-|_j := \delta(S_A|_{i, j}, S_A|_{i, \nu}), \quad (11)$$

and the  $j$ -th entry of  $J_{i, \nu}^\neq$  to

$$J_{i, \nu}^\neq|_j := \delta(-S_A|_{i, j}, S_A|_{i, \nu}), \quad (12)$$

with  $\delta(a, b)$  the Kronecker delta of  $a$  and  $b$ . These vectors  $J_{i, \nu}^-$  ( $J_{i, \nu}^\neq$ ) denote which sign matrix entries of the high-dimensional system have the same sign as in the low-dimensional sign matrix, and which have the opposite sign respectively.

Then let the functions  $F_i(x)$ ,  $i = 1, \dots, n$  be constructed as follows:

$$\begin{aligned} F_i(x) &= A_i(x) - D_i(x) \quad \text{with} \\ A_i(x) &= a_i(\mu_i(x)), \text{ and} \\ D_i(x) &= k_i x_i, \end{aligned} \quad (13)$$

wherein the auxiliary map  $\mu_i : \mathbb{R}^N \rightarrow \mathbb{R}^n$  is used to map from the higher-dimensional state space to the lower-dimensional state space, as follows:

$$\begin{aligned} x \mapsto \mu_i(x) &:= \left( \left( \frac{1 + \epsilon_i \|J_{i, \nu}^\neq\|}{\|J_{i, \nu}^-\|} \right) (J_{i, \nu}^-)^T (\gamma_{j\nu}^{-1} x_j)_{j=1 \dots N} \right. \\ &\quad \left. - \frac{\epsilon_i}{\|J_{i, \nu}^-\|} (J_{i, \nu}^\neq)^T (\gamma_{j\nu}^{-1} x_j)_{j=1 \dots N} \right)_{\nu=1 \dots n} \end{aligned} \quad (14)$$

with  $\gamma_{j\nu} \in \mathbb{R}_{++}$  as determined in (10).

**Step 4:** As a last step, the remaining free parameters are chosen. The parameters  $K_j$ , for  $j \in \{n + 1, \dots, N\}$ , must be chosen sufficiently large,  $K_j > K_j^{min}$ , and the parameters  $\epsilon_i$ , for  $i \in \{1, \dots, n\}$ , sufficiently small,  $0 < \epsilon_i < \epsilon_i^{max}$ , such that it holds that  $\mu_i(x(t)) \in \mathbb{R}_{++}^n$ , for all  $x \in \mathbb{R}_{++}^N$ . It will be proven later in Lemma 2 that indeed such  $K_j^{min}$  and  $\epsilon_i^{max}$  exist.

Thereby interaction functions for all  $i = 1 \dots N$  are now determined in terms of function classes. The choice of the free parameters within the constraints as given in the last step, provide degrees of freedom that can be exploited to, for example, fit the dynamics of a model to data.

#### IV. THEOREM AND PROOF

Before stating the central theorem, we will need the following two lemmata.

*Lemma 1:* If  $K_j = K$  for all  $j \in \{n + 1, \dots, N\}$ , then  $\lim_{K \rightarrow \infty} \gamma_{jk}^{-1} G_{k \rightarrow j}(\lambda) = 1$  for all  $k \in \{1, \dots, n\}$ ,  $j \in \mathcal{M}_k$ .

*Proof:* In this case, omitting the subindices and writing  $\bar{F} := \frac{\partial F}{\partial x}$ ,  $\bar{A} := \frac{\partial A}{\partial x}$ , it is  $\gamma_{jk}^{-1} G_{k \rightarrow j}(\lambda) = \frac{K}{\lambda + K} \left( e_j^T (-\frac{\bar{A}}{K} + I)^{-1} S_A|_{l, k} \right)^{-1} e_j^T \left( \frac{-\bar{A}}{\lambda + K} + I \right)^{-1} S_A|_{l, k}$ , which is equal to 1 in the limit  $K \rightarrow \infty$ . ■

*Lemma 2:* There exist  $K_j^{min} > 0 \forall j \in \{n + 1, \dots, N\}$ ,  $\epsilon_i^{max} > 0 \forall i \in \{1, \dots, n\}$ , such that for all  $K_j > K_j^{min}$  and  $0 < \epsilon_i < \epsilon_i^{max}$  it is  $\mu_i(x(t)) \in \mathbb{R}_{++}^n \forall x \in \mathbb{R}_{++}^N$ .

*Proof:* For appropriate choices of  $K_j^{min}$  with  $K_j > K_j^{min}$ , if  $x_{k_j}$  is bounded, then also every  $x_j$ ,  $j \in \mathcal{M}_{k_j}$ , is bounded. Thus, for every  $t > 0$  there exists some  $M_{j, k_j}(t) > 0$  for which:  $M_{j, k_j}(t) x_{k_j}(t) \geq x_j(t)$ . With finite  $M \geq \max_{t > 0, j \in \mathcal{M}_{k_j}} \{M_{j, k_j}(t)\}$  such that for all  $t > 0$ ,  $j_1 : J_{i, k}^-|_{j_1} = 1$  which lie on a path  $p_{ik}$  as in Assumption 2, and  $j_2 : J_{i, k}^\neq|_{j_2} = 1$ , it holds  $M x_{j_1}(t) \geq x_{j_2}(t)$ . Choosing  $\epsilon_i \leq \frac{1}{M} \frac{\|J_{i, k}^-\| \gamma_{j_1 k}}{\|J_{i, k}^\neq\| \gamma_{j_2 k}}$ , it is:

$$\begin{aligned} \mu_i(x)|_k &\geq \frac{\|J_{i, k}^-\|}{\|J_{i, k}^\neq\|} \gamma_{j_1 k} x_{j_1} + \frac{1}{M} \frac{\|J_{i, k}^-\| \|J_{i, k}^\neq\|}{\|J_{i, k}^\neq\| \|J_{i, k}^-\|} \|J_{i, k}^-\| \gamma_{j_1 k} x_{j_1} \\ &\quad - \frac{1}{M} \frac{\|J_{i, k}^-\| \|J_{i, k}^\neq\|}{\|J_{i, k}^\neq\| \|J_{i, k}^-\|} \frac{\gamma_{j_1 k}}{\gamma_{j_2 k}} \gamma_{j_2 k} x_{j_2} > 0 \\ &\forall x_{j_1} = x_{j_1}(t), t > 0. \end{aligned} \quad \blacksquare$$

We now propose and prove that this construction indeed yields multistability-equivalent systems, by showing that a thereby constructed system has properties (i)-(iii) in Def. 1.

*Theorem 1:* If (2) fulfills the Assumptions 1,2, then for any system (3) as defined in Section II, there exist  $K_j^{min}$ ,  $j = n + 1, \dots, N$  such that every system constructed by the procedure steps 1-4 with  $K_j > K_j^{min}$ ,  $j = n + 1, \dots, N$  is multistability-equivalent to system (1).

*Proof:*

We subsequently prove the properties (i)-(iii) from Def. 1. (i) For  $i \in \{n + 1, \dots, N\}$ , this property can be seen directly from the construction in (8), thus  $\frac{\partial A_i}{\partial x_j} = S_A|_{i, j}$ .

For the remaining indices  $i \in \{1, \dots, n\}$  by construction (13), it is  $A_i(x) = (a_i \circ \mu_i)(x)$ . From Lemma 2, it follows that  $\frac{\partial a_i(z)}{\partial z_i}|_{z=\mu_i(x)} = (S_a)_{i,1\dots n}$ . Furthermore, from (14) it follows that for  $x \in \mathbb{R}_{++}$

$$\frac{\partial \mu_i(x)}{\partial x_j}|_{\nu} = \begin{cases} \frac{1+\epsilon_i \|J_{i,\nu}^{\neq}\|}{\|J_{i,\nu}^{\neq}\|} (\gamma_{j\nu}^{-1}) > 0 & , S_A|_{i,j} = S_a|_{i,\nu} \\ -\frac{\epsilon_i}{\|J_{i,\nu}^{\neq}\|} (\gamma_{j\nu}^{-1}) < 0 & , S_A|_{i,j} = -S_a|_{i,\nu} \\ 0 & , \text{otherwise} \end{cases}$$

since for all  $j \in \mathcal{M}_\nu$  the steady state gain is always nonnegative,  $\gamma_{j\nu} \geq 0$ . With this, one can write

$$\begin{aligned} \text{sgn} \left( \frac{\partial A_i}{\partial x_j} \right) &= \text{sgn} \left( \frac{\partial a_i(\mu_i(x))}{\partial \mu_i(x)} \right) \cdot \text{sgn} \left( \frac{\partial \mu_i(x)}{\partial x_j} \right) \\ &= S_A|_{i,j} \quad \forall i \in \{1, \dots, n\}, j \in \{1, \dots, N\}. \end{aligned}$$

In summary, then (4) holds.

(ii) Choosing the map  $h : \mathbb{R}^n \rightarrow \mathbb{R}^N$  as

$$h : z^* \mapsto h(z^*) := \left( (z_i^*)_{i=1\dots n}, \sum_{k:i \in \{\mathcal{M}_k\}} (\gamma_{ik}^{-1} z_k^*)_{i=(n+1)\dots N} \right)^T \quad (15)$$

it is injective, since it is injective in the first  $n$  components.

Now, we prove  $f(z^*) = 0 \Rightarrow F(h(z^*)) = 0$  separately for  $i \in \{1, \dots, n\}$  and  $i \in \{n+1, \dots, N\}$ .

For  $i \in \{1, \dots, n\}$ , by construction (13) and with (14), (15) and the structure of the matrix  $S_A$  (Fig. 1), it is

$$\begin{aligned} \mu_i(h(z^*)) &= \left( \left( \frac{1 + \epsilon_i \|J_{i,\nu}^{\neq}\|}{\|J_{i,\nu}^{\neq}\|} \right) (J_{i,\nu}^{\neq})^T (\gamma_{j\nu}^{-1} h(z^*)|_j)_{j=1\dots N} \right. \\ &\quad \left. - \frac{\epsilon_i}{\|J_{i,\nu}^{\neq}\|} (J_{i,\nu}^{\neq})^T (\gamma_{j\nu}^{-1} h(z^*)|_j)_{j=1\dots N} \right)_{\nu=1\dots n} = z^*. \end{aligned}$$

With  $h(z^*)|_i = z_i^*$ , one arrives at  $F_i(h(z^*)) = f(z^*) = 0$ .

For the remaining indices  $i \in \{n+1, \dots, N\}$ , by construction (8) and since it has to hold  $z_{m_i}^* = z_{m_j}^* = z_m^*$  for all  $(i, j) : S_A|_{i,j} \neq 0$ , it is

$$F_i(h(z^*)) = \left( \sum_{\substack{j \in \mathcal{M}_m \\ j \neq i}} S_A|_{i,j} \gamma_{jm} + S_A|_{i,i} \gamma_{im} - K_i \gamma_{im} \right) z_m^*,$$

which, if divided by the sum of steady state gains of all influencing state variables,  $\sum_{\{j: S_A|_{i,j} \neq 0, j \neq i\}} \gamma_{jm}$ , is

$$\begin{aligned} &\left( \frac{\sum_{\{j: S_A|_{i,j} \neq 0, j \neq i\}} \gamma_{jm}}{\sum_{\{j: S_A|_{i,j} \neq 0, j \neq i\}} \gamma_{jm}} - K_i \frac{\gamma_{im}}{\sum_{\{j: S_A|_{i,j} \neq 0, j \neq i\}} \gamma_{jm}} \right. \\ &\left. + \frac{S_A|_{i,i} \gamma_{im}}{\sum_{\{j: S_A|_{i,j} \neq 0, j \neq i\}} \gamma_{jm}} \right) z_m^* = \left( 1 + \frac{S_A|_{i,i} - K_i}{K_i - S_A|_{i,i}} \right) z_m^* = 0. \end{aligned}$$

With this, it follows that  $f(z^*) = 0 \Rightarrow F(h(z^*)) = 0$ .

Finally, the reverse direction  $f(z^*) = 0 \Leftarrow F(h(z^*)) = 0$  for  $i \in \{1, \dots, n\}$ : If  $F_i(h(z^*)) = 0$  follows already from the previous part: If  $F_i(h(z^*)) = 0$ , then  $0 = a_i(\mu_i(h(z^*))) - k_i h(z^*)|_i = a_i(z^*) - k_i z_i^* = f_i(z_i^*) = 0$ . With this, (5) holds.

(iii) Finally, it remains to be shown that the number of unstable modes of system (3), if constructed by the steps 1-4, is equal to the number of unstable modes of the low-dimensional system (1).

Therefore, the eigenvalues of the low-dimensional system (1) and of the high-dimensional system (3) are analyzed by a loop breaking approach [10]: Breaking up the loops in the original system yields an open-loop system, and there is a direct relation between the eigenvalues of the original system and the transfer function of the open loop system. (For details the reader is referred to [10].) The original system is obtained by closing the loop via setting the input equal to the output. We restrict the proof to the case where all loops of the system can be captured by a single-input single-output (SISO) loop breaking, whereas a generalization to the multi-input multi-output (MIMO) case is discussed in Sec. VI.

Let us first restrict to the case where the loop to be broken depends on one state variable  $z_k$ , i.e. the output is given by a single state variable:  $y = e_k^T z$ . A generalization to an output being a linear combination of arbitrarily many state variables is straightforward.

First, we perform a loop breaking in the low-dimensional system  $\Sigma_f$ , linearized at a particular steady state  $z^*$ . The corresponding open-loop system reads

$$\begin{aligned} \dot{\tilde{z}} &= \tilde{f} \tilde{z} + B_f u \\ y &= C_f \tilde{z}, \end{aligned} \quad (16)$$

with a Jacobian matrix  $\tilde{f} := \frac{\partial a(\dots, u, \dots)}{\partial z}|_{z^*} - \frac{\partial d(z)}{\partial z}|_{z^*} \in \mathbb{R}^{n \times n}$ , and vectors  $B_f := \frac{\partial a}{\partial z_k}|_{z^*} \in \mathbb{R}^{n \times 1}$ ,  $C_f := e_k^T \in \mathbb{R}^{1 \times n}$ . The function  $a(\dots, u, \dots)$  reflects the activation functions of (1), but the argument  $z_k$  is replaced by  $u$  to realize the loop breaking.

The transfer function of this open-loop system (16) is

$$G_f(\lambda) = C_f (\lambda I - \tilde{f})^{-1} B_f \quad (17)$$

and for non-negative eigenvalues  $\lambda_f$  of the corresponding closed-loop system it holds that  $G_f(\lambda_f) = 1$  ([10], Lemma 2.3), since all loops are broken.

Next, we perform a loop breaking in the high-dimensional system  $\Sigma_F$ , linearized at a particular steady state  $x^* = h(z^*)$ , at the state variable  $x_k$ , with  $k \in \{1, \dots, n\}$  as chosen above (16). This results in an open-loop system

$$\begin{aligned} \dot{\tilde{x}} &= \tilde{F} \tilde{x} + B_F U \\ Y &= C_F \tilde{x}, \end{aligned} \quad (18)$$

again with a Jacobian matrix  $\tilde{F} := \frac{\partial A(\dots, U, \dots)}{\partial x}|_{x^*} - \frac{\partial D(x)}{\partial x}|_{x^*} \in \mathbb{R}^{N \times N}$ , and vectors  $B_F := \frac{\partial A}{\partial x_k}|_{x^*} \in \mathbb{R}^{N \times 1}$ ,  $C_F := e_k^T \in \mathbb{R}^{1 \times N}$ . Here, the function  $A(\dots, U, \dots)$  reflects the activation functions of (3), with  $x_k$  replaced by  $U$ .

The transfer function of this open-loop system (18) is

$$G_F(\lambda) = C_F (\lambda I - \tilde{F})^{-1} B_F \quad (19)$$

and  $G_F(\lambda_F) = 1$  for non-negative eigenvalues  $\lambda_F$  of the corresponding closed-loop system. Denote an index sequence of all indices that are not in  $\mathcal{M}_k$  by  $\overline{\mathcal{M}}_k := \{1, \dots, N\} \setminus \{\mathcal{M}_k\}$ . The set of state variables  $x_j$ ,  $j \in \mathcal{M}_k$  and the

remaining state variables  $x_j$ ,  $j \in \overline{\mathcal{M}}_k$  are in a series circuit, with a matrix  $P$  to pass the signal from  $\mathcal{M}_k$  to the rest of the system:

$$P := \left( \frac{\partial F_i}{\partial x_j} \right)_{\substack{i \in \overline{\mathcal{M}}_k \\ j \in \mathcal{M}_k}} \in \mathbb{R}^{(N-|\mathcal{M}_k|) \times |\mathcal{M}_k|}. \quad (20)$$

So this transfer function (19) can be rewritten:

$$\begin{aligned} G_F(\lambda) &= C_F(\lambda I - \tilde{F})^{-1} B_F \\ &= C_{\overline{\mathcal{M}}_k}(\lambda I - \tilde{F}_{\overline{\mathcal{M}}_k})^{-1} P (G_{k \rightarrow j}(\lambda))_{j \in \mathcal{M}_k} \end{aligned} \quad (21)$$

where the index  $\overline{\mathcal{M}}_k$  ( $\mathcal{M}_k$ , respectively) denotes that the subset of rows and/or columns of the matrix or vector are taken as indexed.

We first look at the case that all modules except one  $\mathcal{M}_k$  are empty, and afterwards iterate to the general case with arbitrarily many non-empty modules.

Then, the index set  $\overline{\mathcal{M}}_k$  equals the set of indices  $\{1 \dots n\}$ , and the transfer function becomes

$$\begin{aligned} G_F(\lambda) &= C_f(\lambda I - \tilde{f})^{-1} \tilde{B}_F(\lambda), \text{ with} \\ \tilde{B}_F(\lambda) &:= \left( \sum_{j \in \mathcal{M}_k} \frac{\partial A_i(x)}{\partial x_j} \Big|_{x^*} G_{k \rightarrow j}(\lambda) \right)_{i=1 \dots n} \\ &= \left( \frac{\partial a_i}{\partial z_k} \Big|_{z^*} \frac{1}{\|J_{i,k}^-\|} \sum_{\substack{j \in \mathcal{M}_k \\ S_A|_{i,j} \neq 0}} \gamma_{jk}^{-1} G_{k \rightarrow j}(\lambda) \right)_{i=1 \dots n}. \end{aligned} \quad (22)$$

We now argue that  $\Sigma_f^{(r)}$  has the same number of eigenvalue with positive real part (positive eigenvalue, for short) as  $\Sigma_F^{(s)}$ . Let  $\varepsilon = \min_{\omega} |G_f(j\omega) - 1|$ . Due to the assumption that  $\Sigma_f^{(r)}$  has no eigenvalues on the imaginary axis, we have  $\varepsilon > 0$ . By Lemma 1 and (22),  $G_F(\lambda) \rightarrow G_f(\lambda)$  as  $K_j \rightarrow \infty$ . This implies that there exist  $K_j^{min}$  such that  $G_F(j\omega) - 1$  lies within a tube of diameter  $\varepsilon$  around  $G_f(j\omega) - 1$  for sufficiently large  $K_j$ , and the winding numbers of these curves around the origin are equal. Then, by the argument principle,  $G_F(j\omega) - 1$  has the same number of zeros in the right half plane as  $G_f(j\omega) - 1$ , and the Jacobians  $\Sigma_f^{(r)}$  and  $\Sigma_F^{(r)}$  have the same number of positive eigenvalues.

If the system  $\Sigma_F^{(s)}$  contains several modules  $\mathcal{M}_i$ ,  $i = 1 \dots n$ , then the index set  $\overline{\mathcal{M}}_k$  contains also the indices of all other modules except for  $k$ :  $\overline{\mathcal{M}}_k = \{1, \dots, n\} \cup \bigcup_{i \neq k} \{\mathcal{M}_i\}$ . Therefore, additional blocks of  $G_{i \rightarrow j}(\lambda)$  appear in  $C_{\overline{\mathcal{M}}_k}$  and can be factored out iteratively for each  $j \notin \{k\} \cup \{1, \dots, n\}$ , until one arrives again at  $C_f$ . Each of these factors  $G_{i \rightarrow j}(\lambda)$  also fulfills the properties (a) and (b), and the same conclusion can be drawn.

This concludes the proof.  $\blacksquare$

## V. EXAMPLE

To illustrate the usefulness of our method proposed in the previous section, we show with an exemplary system how a multistability-equivalent, higher-dimensional system can be constructed.

Let the following low-dimensional system  $\Sigma_f$  of dimension  $n = 2$  be given:

$$\begin{aligned} \dot{z}_1 &= \frac{z_1^2}{1 + 0.1z_1^2 + 0.1z_2^2} - z_1 \\ \dot{z}_2 &= \frac{2}{1 + 0.1z_1^2} - z_2. \end{aligned} \quad (23)$$

The resulting interaction sign matrix is

$$S_a = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}. \quad (24)$$

This system represents the well-known motif of two mutually inhibiting master genes, where in addition one master gene shows also positive auto-feedback. The system has three steady states and is bistable (Fig. 2, upper table).

Now let the gene regulatory network structure for a higher-dimensional system  $\Sigma_F$  of dimension  $N = 5$  be given, as by the interaction sign matrix

$$S_A = \begin{bmatrix} 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \quad (25)$$

The task is to construct a higher-dimensional system consistent with the interaction structure (25) and multistability-equivalent to the low-dimensional system (23), for which we apply the construction rules as presented in Sec. III.

The matrix is given in a structure as by Assumption 1, with modules  $\mathcal{M}_1 = \{3, 4\}$  and  $\mathcal{M}_2 = \{5\}$ . It also meets Assumption 2 on the sign matrices.

Step 1 of the construction yields the functions

$$\begin{aligned} \dot{x}_3 &= x_1 - K_3 x_3 \\ \dot{x}_4 &= x_1 - K_4 x_4 \\ \dot{x}_5 &= x_2 + x_5 - K_5 x_5. \end{aligned} \quad (26)$$

From step 2, the steady state gain values are derived

$$\gamma_{31} = \frac{1}{K_3}, \quad \gamma_{41} = \frac{1}{K_4}, \quad \gamma_{52} = \frac{1}{K_5 - 1}. \quad (27)$$

In step 3, we construct the remaining functions  $f_1(x), f_2(x)$ : The index vectors as defined by (11) are

$$\begin{aligned} J_{1,1}^- &= e_3^T, \quad J_{1,2}^- = e_5^T, \quad J_{2,1}^- = e_3^T, \quad J_{2,1}^\# = e_4^T \\ J_{1,1}^\# &= J_{1,2}^\# = J_{2,2}^- = J_{2,2}^\# = 0. \end{aligned} \quad (28)$$

With this the functions  $\mu_i(x)$  to map to the lower-dimensional space can be determined:

$$\begin{aligned} \mu_1(x) &= [K_3 x_3, (K_5 - 1)x_5], \\ \mu_2(x) &= [(1 + \epsilon_2)K_3 x_3 - \epsilon_2 K_4 x_4, 0]. \end{aligned} \quad (29)$$

In the last step 4, the free parameters are assigned to specific values. For initial values, e.g.,  $x_3(0), x_4(0), x_5(0) \in \mathbb{R}_+$  and  $x_3(0) \geq 0.5x_4(0)$ , choosing  $K_3 = K_4 = 1, K_5 = 5$  and  $\epsilon_2 = 0.5 \leq \min_{t \geq 0} \left\{ \frac{x_4(t)}{x_3(t)} \right\} \cdot \frac{K_4}{K_3}$  ensures that the functions  $\mu_i(x(t))$  are always nonnegative for any  $x(t)$ ,  $t \geq 0$ . (Any initial values  $x(0) \in \mathbb{R}_+^5$  can be covered, by appropriate choice of  $K_j$  and/or  $\epsilon_j$  that ensure  $\mu_i(x(t)) \geq 0 \forall t \geq 0$ .)

With this, the functions become  $\mu_1(x) = [x_3, 4x_5]$ ,  $\mu_2(x) = [1.5x_3 - 0.5x_4, 0]$ .

Thus, we obtain the activation functions

$$A_1(x) = a_1(\mu_1(x)) = \frac{x_3^2}{1 + 0.1x_3^2 + 1.6x_5^2} \quad (30)$$

$$A_2(x) = a_2(\mu_2(x)) = \frac{2}{1 + 0.1(1.5x_3 - 0.5x_4)^2}.$$

and we can now write a system  $\Sigma_F : \dot{x} = A(x) - D(x)$ .

The steady states and corresponding eigenvalues, both computed numerically, are given in a table in Fig. 2.

$z^*(r)$	$\lambda_{f,1}^{(r)}$	$\lambda_{f,2}^{(r)}$
(0,2)	-1	-1
(1.49,1.64)	+0.78	-1.07
(8.87,0.23)	-0.76	-1.01

  

$x^*(r)$	$\lambda_{F,1}^{(r)}$	$\lambda_{F,2}^{(r)}$	$\lambda_{F,3}^{(r)}$	$\lambda_{F,4}^{(r)}$	$\lambda_{F,5}^{(r)}$
(0,2,0,0,0.5)	-1	-1	-1	-4	-1
(1.49,1.64,1.49,1.49,0.41)	+0.37	-1.23	-2.10	-4.05	-1
(8.87,0.23,8.87,8.87,0.06)	-4.00	-0.51	-1.46	-1.02	-1

Fig. 2. Steady states and associated eigenvalues of systems  $\Sigma_f$  and  $\Sigma_F$ .

As can be seen, system  $\Sigma_f$  and system  $\Sigma_F$  have in each pair of corresponding steady states the same number of positive eigenvalues. With this, we have successfully established a five-dimensional multistability-equivalent system with an interaction network structure as given by the sign matrix.

## VI. CONCLUSION

The proposed method allows to transfer systems-theoretic results on low-dimensional core motif models to more detailed networks, as usually encountered in reconstructed gene networks. These in turn can, for example, be fitted to experimental data.

While equal results might hold for other choices of construction rules, we briefly discuss the specific choices of the proposed construction procedure. For example, the activation functions for module genes are chosen as linear functions, whereas for master genes they are assembled via the auxiliary map which entries are linear combinations in the module genes. The construction leads to a positive steady state gain independent of the steady state value, such that the amplification of the signal induced by the module genes can be “factored out”. In turn, the activation functions for the master genes are unbounded in the values of the module genes, which allows for a full signal transmission between master genes. The impact of an interaction from module genes onto master genes with a sign opposed to the corresponding interaction in the low-dimensional system is counterbalanced by enhancing interactions with equal sign.

The module structure of the interaction sign matrix is required to exclude contradictions for the activation functions. However, this structural requirement may not be restrictive when modeling real gene regulatory systems: Least significant interactions might be chosen such that the interaction sign matrix can be adjusted by setting these entries to zero.

Several generalizations of our method are possible, which could not be covered in this work, but should be mentioned

shortly. More general interaction functions for the module genes can be implemented by weighted sums in the linear activation rates (8). Similarly, the degradation rates for master genes in (13) might be chosen more flexible, by adjusting the maps  $\mu_i$  and  $h$  accordingly. For a generalization to cases where a MIMO loop breaking is required, one has to exploit the MIMO Nyquist criterion.

With the example given in Sec. V, we have demonstrated how the construction of a multistability-equivalent system can be conducted, and how the results from stability analysis of the low-dimensional system translate into properties of the high-dimensional system. While the number of positive eigenvalues in each steady state is equal for the low- and high-dimensional systems, the specific values of these eigenvalues may differ. Hence, the transient dynamics may differ substantially. It will be the subject of future work to investigate how bifurcations and dynamical properties in such multistability-equivalent systems are related.

## ACKNOWLEDGMENT

The authors would like to thank for financial support from the German Research Foundation (DFG) within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart and the Matheon research center at FU Berlin, from the German Federal Ministry of Education and Research (BMBF) within the SysTec program (grant nr. 0315506A), and from The MathWorks Foundation of Science and Engineering. The authors thank Taouba Jouini for helpful comments on a preliminary version of the paper.

## REFERENCES

- [1] C. Breindl and F. Allgöwer. Verification of multistability in gene regulation networks: A combinatorial approach. Proc. 48th IEEE Conf. Decision and Control (CDC), 2009.
- [2] S. Huang, Y.-P. Guo, G. May and T. Enver. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental Biology*, vol. 305, pp. 695–713, 2007.
- [3] C. Li, L. Chen and K. Aihara. Stability of genetic networks with SUM regulatory logic: Lur’e system and LMI approach. *IEEE Transactions on Circuits and Systems*, vol 53, pp.2451–2458, 2006.
- [4] B.D. MacArthur, A. Ma’ayan and I.R. Lemischka. Systems biology of stem cell fate and cellular reprogramming. *Nature Reviews Molecular Cell Biology*, vol. 10, pp. 672–681, 2009.
- [5] J. Narula, A.M. Smith, B. Gottgens and O.A. Igoshin. Modeling reveals bistability and low-pass filtering in the network module determining blood stem cell fate. *PLoS Computational Biology*, vol. 6, e1000771, 2010.
- [6] J. Peltier and D.V. Schaffer. Systems biology approaches to understand stem cell fate choice. *IET Systems Biology*, vol. 4, pp. 1–11, 2010.
- [7] I. Roeder and I. Glauche. Towards an understanding of lineage specification in hematopoietic stem cells: A mathematical model for the interaction of transcription factors GATA-1 and PU.1. *Journal of Theoretical Biology*, vol. 241, pp. 852–865, 2006.
- [8] D. Schittler, J. Hasenauer, F. Allgöwer, and S. Waldherr. Cell differentiation modeled via a coupled two-switch regulatory network. *Chaos*, vol. 20(4):045121, 2010.
- [9] D. Schittler, C. Breindl, S. Waldherr, and F. Allgöwer. Structural requirements and discrimination of cell differentiation networks. Proc. 18th IFAC World Congress, 2011, arXiv:1107.0900.
- [10] S. Waldherr and F. Allgöwer. Searching bifurcations in high-dimensional parameter space via a feedback loop breaking approach. *International Journal of Systems Science*, vol. 40, pp. 769–782, 2009.
- [11] W. Xiong and J.E. Ferrell, Jr.. A positive-feedback-based bistable memory module that governs a cell fate decision. *Nature*, vol. 426, pp. 460–465, 2003.