Analysing Control-theoretic Properties of Nonlinear Synthetic Biology Circuits ⋆

Antón Pardo [**] **Sandra Díaz Seoane** [**] **Dorin A. Ionescu** [**]
Antonis Papachristodoulou [***] **Alejandro F. Villaverde** [*,**]

[*] *CITMAga, 15782 Santiago de Compostela, Galicia, Spain*
[**] *Universidade de Vigo, Department of Systems Engineering and Control, 36310 Vigo, Galicia, Spain*
[***] *Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK*

Abstract: Synthetic biology is a recent area of biological engineering, whose aim is to provide cells with novel functionalities. A number of important results regarding the development of control circuits in synthetic biology have been achieved during the last decade. A differential geometry approach can be used for the analysis of said systems, which are often nonlinear. Here we demonstrate the application of such tools to analyse the structural identifiability, observability, accessibility, and controllability of several biomolecular systems. We focus on a set of synthetic circuits of current interest, which can perform several tasks, both in open loop and closed loop settings. We analyse their properties with our own methods and tools; further, we describe a new open-source implementation of the techniques.

Keywords: Accessibility, controllability, structural identifiability, observability, geometric control, nonlinear systems, system identification, synthetic biology.

1. INTRODUCTION

Synthetic biology is an interdisciplinary research field concerned with "re-programming" cells, providing them with new or modified functionalities (Qian et al., 2018). In this context, key goals are to achieve some desired dynamics and to reduce the effect of uncertainty. Since these systems are usually nonlinear, tools from nonlinear control theory are required for this aim. The theoretical ability to drive a system to a final state is given by its *accessibility* and *controllability* (Lewis, 2001). Biological systems often have components with a high degree of uncertainty. To obtain a good characterization, it is important to be able to determine correctly the values of any unknown parameters and unmeasured variables. The possibility of performing these tasks successfully is given by a model's *identifiability* and *observability* (Chatzis et al., 2015). In this work we are interested in analysing such control theoretic properties from a *structural* point of view, i.e., focusing on the constraints imposed by the equations that define system dynamics, rather than on *practical* limitations introduced by measurement uncertainty (Wieland et al., 2021). To model the systems under study we use nonlinear ordinary differential equations (ODEs), which are adequate for biomolecular systems as long as the number of molecules is sufficiently large. They are also helpful as approximate models of systems that are known to exhibit stochastic behaviour. Under these assumptions, the aforementioned structural properties – accessibility, controllability, structural identifiability and observability – can be studied using a differential geometry approach (Hermann and Krener, 1977).

According to Qian et al. (2018), many of the remaining challenges in synthetic biology can be addressed by a control-theoretic approach, but theory is currently lagging behind the development of biological tools. Likewise, Baetica et al. (2019) claim that control theory has yet to be fully applied to the understanding and engineering of biological systems. Indeed, there are few works reporting these types of analyses of synthetic biology systems, despite a few recent examples (Díaz-Seoane et al., 2023a; Haus et al., 2023). In this paper we contribute to fill this gap by applying techniques from nonlinear control theory to a set of recently presented models of synthetic biology circuits. We show how the analyses can be performed with different tools, and present a new Python implementation of our own accessibility and controllability code, which was previously available only in MATLAB.

2. THEORY AND METHODS

2.1 Notation and systems

We consider nonlinear systems of ordinary differential equations of the form

$$M : \begin{cases} \dot{x}(t) = & f\left(u(t), x(t), \theta\right), \\ y(t) = & h\left(u(t), x(t), \theta\right), \end{cases} \tag{1}$$

where f and h are analytic functions; $x(t) \in \mathbb{R}^{n_x}$ is the vector of state variables at time t; $u(t) \in \mathbb{R}^{n_u}$ is the

⋆ This work has been supported by grant PID2020-113992RA-I00 funded by MCIN/AEI/ 10.13039/501100011033, grant PID2023-146275NB-C21 funded by MICIU/AEI/10.13039/501100011033 and ERDF/EU, and grant CNS2023-144886 funded by MICIU/AEI /10.13039/501100011033 and the European Union NextGenerationEU/PRTR.

input vector, which is assumed to consist of infinitely differentiable functions; $y(t) \in \mathbb{R}^{n_y}$, the output vector; $\theta \in \mathbb{R}^{n_\theta}$, the parameter vector. In the following, we may omit the dependence on time for ease of notation, i.e., we may simply write x, y, u. A particular case of (1) are input-affine systems, the dynamics of which can be written as follows:

$$\dot{x} = f(x, \theta) \; + \; \Sigma_{i=1}^{n_u} \; u_i \; g_i(x, \theta) \tag{2}$$

2.2 Differential geometric concepts

Differential geometric concepts for the analysis of nonlinear systems are described e.g. by Sontag (2013). Below we provide the definitions required in this work.

Lie derivative: Given the model (1), the Lie derivative of the output function h along the vector field f is:

$$L_f h = \frac{\partial h}{\partial x} f + \frac{\partial h}{\partial u} \dot{u}.$$

Setting $L_f^0 h = h$, the i-order extended Lie derivative of h can be recursively computed as:

$$L_f^i h = \frac{\partial L_f^{i-1} h}{\partial x} f + \sum_{j=0}^{i-1} \frac{\partial L_f^{i-1} h}{\partial u^{(j)}} u^{(j+1)}.$$

where $u^{(j+1)}$ $(j \geq 0)$ stands for the $(j+1)$-th time derivative of u.

Lie bracket: The Lie bracket of two vector fields f, g is another vector field given by:

$$[f, g] = \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g,$$

Lie algebra: A set of vector fields $\mathcal{L}$ is a Lie algebra when it is a linear subspace of vector fields (that is, $\alpha f + \beta g \in \mathcal{L}$ when $f, g \in \mathcal{L}$) and when the Lie bracket is well defined (i.e. $[f, g] \in \mathcal{L}$ when $f, g \in \mathcal{L}$). The Lie algebra generated by a family of vector fields $\mathcal{P}$ is the smallest Lie algebra containing $\mathcal{P}$, and it is written as $\mathrm{LA}[\mathcal{P}]$.

Distribution: A distribution is a map $\mathcal{S}$ between each $x \in X$ and a subspace $\mathcal{S}(x) \subset \mathbb{R}^n$.

2.3 Observability and structural local identifiability

Observability: A state $x_i(\tau)$ (that is, the i^{th} element of the state variables vector) is observable if it can be determined from the output $y(t)$ and the input $u(t)$ in an interval $t_0 \leq \tau \leq t \leq t_f$, for a finite t_f. Otherwise, it is unobservable. A model is observable if all its states are observable. Observability is usually considered as a local property, i.e. an observable state can be distinguished from any other states in a neighbourhood, but possibly not from all distant states. Here we adopt this viewpoint and consider observability and identifiability in the local sense. This enables the use of differential geometric techniques.

Structural local identifiability: A parameter θ_i is structurally locally identifiable (SLI) if, for almost any parameter vector $\theta^* \in \mathbb{R}^{n_\theta}$, there is a neighbourhood $\mathcal{N}(\theta^*)$ where the following condition holds (DiStefano III, 2015):

$$\hat{\theta} \in \mathcal{N}(\theta^*) \quad \text{and} \quad y(t, \hat{\theta}) = y(t, \theta^*) \Rightarrow \hat{\theta}_i = \theta_i^* \tag{3}$$

If (3) is not true in any neighborhood of θ^*, the parameter θ_i is *structurally unidentifiable* (SU). If (3) is true for all model parameters, the model is said to be SLI as well, and SU otherwise. We will use the acronym SIO to refer to Structural Identifiability and Observability.

Structural local identifiability and observability (SIO): As noted by Tunali and Tarn (1987), structural local identifiability can be treated as a particular case of observability by considering the parameters as state variables that happen to be constant, i.e. their dynamics are given by $\dot{\theta}_i = f(u(t), x(t), \theta) = 0$. Thus, the SIO of a model can be evaluated using a version of the observability rank condition introduced by Hermann and Krener (1977). Here we follow this approach, which we have previously implemented in the MATLAB toolbox STRIKE-GOLDD (Díaz-Seoane et al., 2023b). The core idea is to build an observability-identifiability matrix, O_I, and calculate its rank. To this end, we augment the state vector as $\tilde{x} = [x, \theta]$, defining $n_{\tilde{x}} = n_x + n_\theta$. The observability-identifiability matrix O_I of a model (1) is:

$$\mathcal{O}_I(\tilde{x}) = \begin{pmatrix} \dfrac{\partial}{\partial \tilde{x}} h(\tilde{x}, u) \\ \dfrac{\partial}{\partial \tilde{x}} \left(\mathcal{L}_f h(\tilde{x}, u) \right) \\ \vdots \\ \dfrac{\partial}{\partial \tilde{x}} \left(\mathcal{L}_f^{n_{\tilde{x}}-1} h(\tilde{x}, u) \right) \end{pmatrix}. \tag{4}$$

(1) is SLI and observable if $\mathrm{rank}(\mathcal{O}_I(\tilde{x})) = n_{\tilde{x}}$. If $\mathrm{rank}(\mathcal{O}_I(\tilde{x})) < n_{\tilde{x}}$, there is at least one unobservable variable and/or one unidentifiable parameter. Since the i^{th} column of $\mathcal{O}_I(\tilde{x})$ represents the partial derivative with respect to the i^{th} element of $\tilde{x}$, the SIO of an individual variable, $\tilde{x}_i$, can be determined by removing the i^{th} column and recalculating the rank. If the rank decreases, $\tilde{x}_i$ is observable (or SLI, if it is a parameter); if the rank remains unchanged, $\tilde{x}_i$ is observable (or SU).

2.4 Accessibility and controllability

Reachable set: The set of all points $x_f = x(t)$ with $t \leq T$ that a system can reach from an initial point x_0 in time at most T is called the reachable set:

$$\mathrm{Reach}\,(M, \leq T, x_0) = \bigcup_{0 \leq t \leq T} \mathrm{Reach}\,(\Sigma, t, x_0)$$

Accessibility: The system M (1) has the accessibility property from $x_0 \in X$ if for every $T > 0$ the set $\mathrm{Reach}\,(M, \leq T, x_0)$ has a nonempty (full dimensional) interior (Sussmann, 1987).

Controllability: The system M (1) is small-time locally controllable (STLC) from $x_0 \in X$ if for every $T > 0$ the set $\mathrm{Reach}\,(M, \leq T, x_0)$ contains x_0 in its non-empty interior.

To analyse accessibility and controllability we adopt the methodology described by Díaz-Seoane et al. (2023a), which was originally implemented in MATLAB. To widen its adoption we have developed a new version in Python. We provide both implementations as open source software (`https://github.com/afvillaverde/NLcontrollability`). The test for accessibility is based on determining whether certain

distributions defined by the Lie algebras generated by the vector fields of a system are full-dimensional. We use the Lie Algebraic Rank Condition (LARC) described by Díaz-Seoane et al. (2023a), which provides a sufficient and necessary condition for accessibility. For controllability we consider the General Sufficient Condition (GSC). These tests are applicable to systems of the form (2), i.e. which are affine in the inputs. This analytical approach informs about *structural* properties, which are locally valid for a generic point, and can also inform about specific points by replacing the symbolic variables with numerical values. If we were interested in a specific operating region, we could use numerical approaches such as empirical Gramians, as e.g. (Himpe, 2018).

3. CONTROL CIRCUITS IN SYNTHETIC BIOLOGY

Importantly, biological systems are typically positive, a property that has implications for their control (Briat, 2020). However, this feature is not an obstacle for obtaining signals that can in principle be negative, such as derivatives, since this goal can be achieved by adding a bias (Alexis et al., 2021); likewise, it is not an obstacle for performing the analyses, as noted by Díaz-Seoane et al. (2023a). Haus et al. (2023) analysed the structural identifiability of several models of biomolecular controller motifs, classified either as 'basic' or 'antithetic'. In this paper we focus on a set of circuits of which we analyse their accessibility, controllability (when possible), structural identifiability and observability. We provide their mathematical description in the remainder of this section, and their topologies in Fig. 1. These ODE models are approximations of the real systems, which are inherently stochastic. The extent to which this modelling assumption is adequate is worthy of investigation. In this regard, Kelly et al. (2018) found that the effect of extrinsic noise (the one resulting from cell-wide variations) was stronger than that of the intrinsic noise (the one resulting from stochasticity). It should also be noted that the dichotomous feedback architecture described in Section 3.2 decreases intrinsic noise (Sootla et al., 2022).

3.1 Molecular topologies for signal differentiation

Alexis et al. (2021) presented three topologies that perform signal differentiation. They can serve several purposes, such as acting as speed biosensors or implementing derivative control actions. We refer to them as BioSD-I, BioSD-II and BioSD-III, and give their equations below. We use a generic notation for their description, where the parameters are written as p_i, and the derivative signal is $x_1(t) \approx \dot{u}(t)$.

BioSD-I:

$$\begin{aligned}\dot{X} &= k_{in} \cdot U + b - k_1 \cdot X \cdot Z - \delta \cdot X \\ \dot{Z} &= k_2 \cdot X - k_3\end{aligned}$$

BioSD-II:

$$\begin{aligned}\dot{X} &= k_{in} \cdot U + b - k_1 \cdot X \cdot Z_1 - \delta \cdot X \\ \dot{Z}_1 &= k_2 \cdot X - \eta \cdot Z_1 \cdot Z_2 \\ \dot{Z}_2 &= k_3 - \eta \cdot Z_1 \cdot Z_2\end{aligned}$$

BioSD-III:

$$\begin{aligned}\dot{X} &= k_{in} \cdot U + b - k_1 \cdot X \cdot Z_1 + k_1 \cdot X \cdot Z_2 - \delta \cdot X \\ \dot{Z}_1 &= k_2 \cdot X - \eta \cdot Z_1 \cdot Z_2 \\ \dot{Z}_2 &= k_3 - \eta \cdot Z_1 \cdot Z_2\end{aligned}$$

Additionally, Alexis et al. (2021) introduced a more realistic version of BioSD-II, in which the activation of x_2 by x_1 takes place with Michaelis-Menten kinetics:

BioSD-II-MM-simple:

$$\begin{aligned}\dot{X} &= k_{in} \cdot U + b - k_1 \cdot X \cdot Z_1 - \delta \cdot X \\ \dot{Z}_1 &= \frac{V_{max} \cdot X}{X + K_m} - \eta \cdot Z_1 \cdot Z_2 \\ \dot{Z}_2 &= k_3 - \eta \cdot Z_1 \cdot Z_2\end{aligned}$$

BioSD-II-MM-complex:

$$\begin{aligned}\dot{X} &= k_{in} \cdot U + b - k_1 \cdot X \cdot Z_1 - (\delta + \gamma) \cdot X \\ \dot{Z}_1 &= \frac{V_{max} \cdot X}{X + K_m} - \eta \cdot Z_1 \cdot Z_2 - \gamma \cdot Z_1 \\ \dot{Z}_2 &= k_3 - \eta \cdot Z_1 \cdot Z_2 - \gamma \cdot Z_2\end{aligned}$$

3.2 Dichotomous Feedback

Natural biological systems may exhibit dichotomous feedback, which works through sequestration of a molecule or a signal. Sootla et al. (2022) proposed several ways of implementing this functionality. Here we study the following model, which is described in equations (2.7) of their article:

$$\begin{aligned}\dot{HK} &= \beta_{HK} - \delta \cdot HK - k_{ap}(I) \cdot HK \\ &\quad + k_t \cdot \left(\frac{\beta_{HK}}{\delta} - HK\right) \cdot RR + k_{tc} \cdot \left(\frac{\beta_{HK}}{\delta} - HK\right) \cdot SR \\ \dot{RR} &= \beta_{RR} - \delta \cdot RR - k_t \cdot \left(\frac{\beta_{HK}}{\delta} - HK\right) \cdot RR \\ &\quad + k_p \cdot HK \cdot \left(\frac{\beta_{RR}}{\delta} - RR\right) \\ \dot{SR} &= \beta_{SR} - \delta \cdot SR - k_{tc} \cdot \left(\frac{\beta_{HK}}{\delta} - HK\right) \cdot SR \\ &\quad + k_{pc} \cdot HK \cdot \left(\frac{\beta_{SR}}{\delta} - SR\right)\end{aligned}$$

where $k_{ap}(I) = k_{ap-max} \frac{I}{I + K_{da}}$, and I is the input signal. Note that taking k_{ap} as the input signal yields a model that is affine in the inputs. We will also analyse this model considering the production rates β_* as inputs that can be modified.

3.3 Negative Feedback

We consider two synthetic circuits based on engineered small RNAs (sRNAs) presented by Kelly et al. (2018).

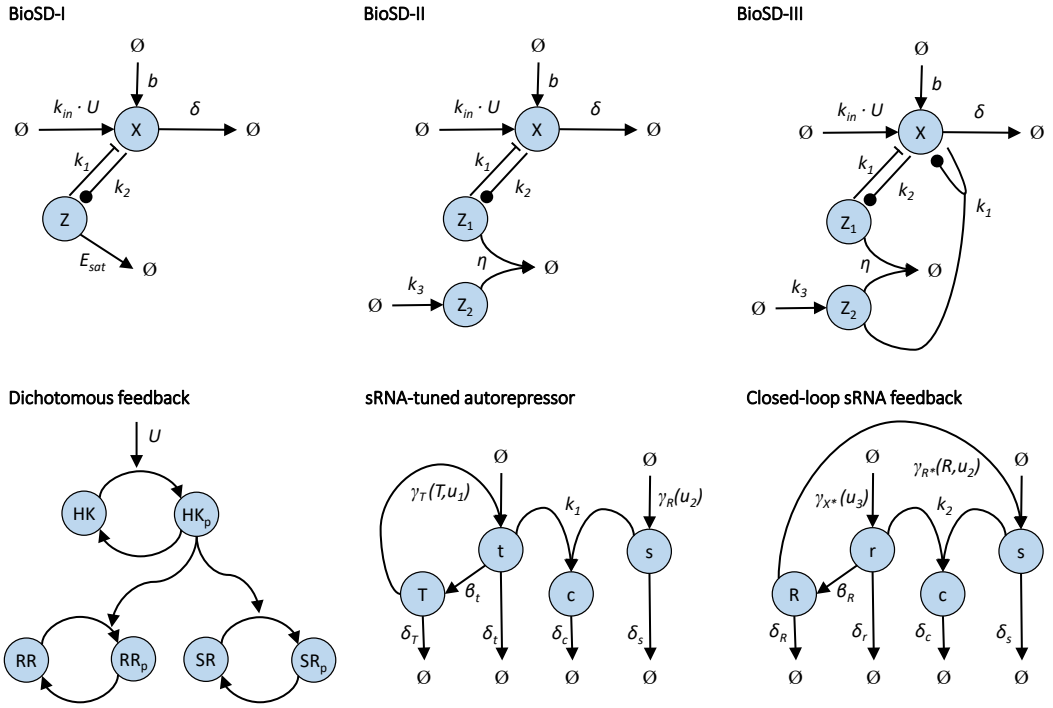

Fig. 1. Diagrams of the synthetic biology circuits analysed in this work.

sRNA-tuned autorepressor:

$$\begin{aligned}\dot{t} &= \gamma_T - \delta_t \cdot t - K_1 \cdot t \cdot s\\ \dot{s} &= \gamma_R - \delta_s \cdot s - K_1 \cdot t \cdot s\\ \dot{c} &= K_1 \cdot t \cdot s - \delta_c \cdot c\\ \dot{T} &= \beta_T \cdot t - \delta_T \cdot T\end{aligned}$$

Where γ_T and γ_R are defined as follows:

$$\gamma_T = \frac{\alpha_t}{1+\left(\frac{T}{K_T\left(1+\left(u_1/K_{u_1}\right)^{n_{u_1}}\right)}\right)^{n_T}} + \frac{\alpha_L\left(\frac{T}{K_T\left(1+\left(u_1/K_{u_1}\right)^{n_{u_1}}\right)}\right)^{n_T}}{1+\left(\frac{T}{K_T\left(1+\left(u_1/K_{u_1}\right)^{n_{u_1}}\right)}\right)^{n_T}}$$

$$\gamma_R = \alpha_r \frac{u_2}{K_{u_2}+u_2} \tag{5}$$

The states are: t, mRNA concentration; s, sRNA concentration; c, sRNA-mRNA complex concentration; and T, TetR-GFP complex concentration. Similarly to the models in Section 3.2, this model is only affine in the inputs if these are taken to be γ_T and γ_R instead of u_1 and u_2. The same applies to the following model.

Closed-loop sRNA Feedback Circuit:

$$\begin{aligned}\dot{r} &= \gamma_X^* - \delta_r \cdot r - K_2 \cdot r \cdot s\\ \dot{s} &= \gamma_R^* - \delta_s \cdot s - K_2 \cdot r \cdot s\\ \dot{c} &= K_2 \cdot r \cdot s - \delta_c \cdot c\\ \dot{R} &= \beta_R \cdot r - \delta_R \cdot R\end{aligned}$$

Where γ_X^* and γ_R^* are defined as follows:

$$\gamma_X^* = \alpha_X^* \frac{u_3}{K_{u_3}+(u_3)} \tag{6}$$

$$\gamma_R^* = \alpha_r^* \frac{\left(\frac{R\cdot u_2}{K_{u_2}^*+u_2}\right)}{K_R + \frac{R\cdot u_2}{K_{u_2}^*+u_2}} \tag{7}$$

The new states are r, mRNA concentration; and R, RhaS-GFP complex concentration.

4. RESULTS

4.1 Accessibility and controllability

Table 1. Results: accessibility, controllability.

Case study	Eq. point	Accessible	Controllable
BioSD-I	param.	yes	yes
BioSD-II	param.	yes	yes
BioSD-II-MM-simple	param.	yes	yes
BioSD-II-MM-complex	not found	yes $^{(\text{neq})}$	NA
BioSD-III	param.	yes	yes
Dichotomous Feedback	param.	yes	yes
sRNA-tuned auto-repressor (input: γ_R)	param.	yes	yes
Closed-loop sRNA (inputs: γ_R^*, γ_X^*)	param.	yes	yes

Table 1 shows the list of biosystems along with the results produced by the LARC and GSC tests, which assess accessibility and small-time local controllability, respectively. The GSC is not applicable (NA) in non-equilibrium points, and it cannot test if the model is inaccessible. In principle, tests are performed at the equilibrium (in the table, 'param.' means that the equilibrium point depends on the value of the parameters and inputs). When an equilibrium is not found, the tests are performed around a point $x_i = 1$, i= 1, ... , n_x; this is denoted as $^{(\text{neq})}$. We analysed the negative feedback models of Section 3.3 by treating the γ functions as inputs. Alternatively, one could consider

as inputs the u_* variables; however, doing so would make these models non-affine in the inputs, which would prevent the application of the accessibility and controllability tests.

4.2 Structural local identifiability and observability

Since identifiability and observability (SIO) depend on the model outputs, we have considered several possible output configurations for each system. First, we considered the output defined in the original publications. Additionally, we explored combinations of several states, always keeping the original output among the measured states. Thus we are able to assess how measurement availability influences the SIO of the unmeasured states and parameters. Table 2 summarizes the results of these analyses. We analyse two versions of the Dichotomous Feedback circuit: one that replaces $k_{ap}(I)$ with $k_{ap-max}\frac{I}{I+K_{da}}$ in the equations, which we denote with (I), and another one which uses directly k_{ap}. In each of these versions we perform the analyses when said variable is the input, and also when each of the production rates β_* is the input (when a β_* is not an input, it is treated as a parameter). Likewise, we consider similar variants of the sRNA models, i.e., taking either the γ or the u variables as inputs. We note that the sRNA-tuned autorepressor model could only be analysed by assuming that at least two states can be measured; analyses with only one output required too much memory.

The SIO results inform about the possibility of identifying the parameters for every possible output configuration. Let us consider for example the closed-loop sRNA circuit. A typical choice could be to take measurements on the RhaS-GFP complex, i.e. R. As shown in Table 2, this would make it impossible to identify the sRNA-mRNA binding strength (K_2), the translation rate of RhaS-GFP mRNA (β_R), the degradation rate of the mRNA-sRNA complex (δ_c), and the transcription rates α_r^* and α_X^*. Furthermore, it would be impossible to infer any of the unmeasured state variables. In contrast, if one measures not only R but also c, which is the sRNA-mRNA complex concentration, all the parameters become identifiable, and all state variables become observable. Similar insights can be extracted for each of the output configurations shown in the aforementioned table.

5. DISCUSSION

In this paper we have demonstrated the use of symbolic computation to analyse structural properties of synthetic biological circuits. Our analyses have shed light on how the availability of output measurements affects parameter identifiability. While certain outputs yield identifiable and observable models, others achieve only partial identifiability. In contrast, all systems were found to be accessible, as one could expect; our analyses did not find any unforeseen deficiencies. While the structural identifiability of a number of synthetic biology models had already been analysed by Haus et al. (2023), here we have considered a different set of models. In regard to accessibility and controllability, to the best of our knowledge the results reported in this paper represent the first systematic study of these properties in synthetic biology circuits.

An additional contribution of this work is the implementation of the methods in open source software toolboxes. The main novelty is the development of a Python version of NLcontrollability, which we had previously made available in MATLAB. Together with STRIKE-GOLDD and STRIKEpy, these tools provide implementations of the tests reported here in MATLAB and Python.

It should be noted that some models could only be analysed under assumptions that render them affine in the inputs, which is a requirement for the application of these accessibility and controllability tests. This transformation is, in principle, practically achievable, but it introduces additional challenges for the design of control laws. Furthermore, the results should be taken as an initial approximation to the properties being studied: our analyses adopt a *structural* viewpoint; they do not consider *practical* limitations that can affect their numerical versions. In particular, the analyses assume that the inputs are continuous, time-varying, and sufficiently exciting; in contrast, in real applications inputs may be restricted to e.g. constant or piecewise constant functions. Taking those limitations into account requires a different set of methods.

DATA AND CODE AVAILABILITY STATEMENT

All the code developed and used for the analyses reported in this paper is available at: `https://github.com/afvillaverde/NLcontrollability` (for accessibility and controllability) and `https://github.com/afvillaverde/strike-goldd` (for structural identifiability and observability). The models analysed in this paper can be found in the corresponding `models` folders.

REFERENCES

Alexis, E., Schulte, C.C., Cardelli, L., and Papachristodoulou, A. (2021). Biomolecular mechanisms for signal differentiation. *Iscience*, 24(12).

Baetica, A.A., Westbrook, A., and El-Samad, H. (2019). Control theoretical concepts for synthetic and systems biology. *Curr.t Opin. Syst. Biol.*, 14, 50–57.

Briat, C. (2020). A biology-inspired approach to the positive integral control of positive systems: The antithetic, exponential, and logistic integral controllers. *SIAM J. Appl. Dyn. Syst.*, 19(1), 619–664.

Chatzis, M.N., Chatzi, E.N., and Smyth, A.W. (2015). On the observability and identifiability of nonlinear structural and mechanical systems. *Struct. Contr. Health Monitor.*, 22(3), 574–593.

Díaz-Seoane, S., Blas, A.B., and Villaverde, A.F. (2023a). Controllability and accessibility analysis of nonlinear biosystems. *Computer Methods and Programs in Biomedicine*, 242, 107837.

Díaz-Seoane, S., Rey-Barreiro, X., and Villaverde, A.F. (2023b). STRIKE-GOLDD 4.0: user-friendly, efficient analysis of structural identifiability and observability. *Bioinformatics*, 39(1), btac748.

DiStefano III, J. (2015). *Dynamic systems biology modeling and simulation.* Academic Press: Amsterdam, Netherlands.

Haus, E.S., Drengstig, T., and Thorsen, K. (2023). Structural identifiability of biomolecular controller motifs with and without flow measurements as model output. *PLOS Comput. Biol.*, 19(8), e1011398.

Table 2. Results: structural identifiability, observability. A slash (e.g. A/B) indicates that the result holds when either variable (A or B) is taken as an input or output. Parameters are 'Identifiable' or 'Non-identifiable'; states are observable ('Obs') or non-observable ('Non-obs').

Model	Inputs (u)	Outputs (y)	Identifiable	Non-identifiable	Obs.	Non-obs.
BioSD I	U	X	k_{in} b	$k_1\ k_2\ k_3\ \delta$	X	Z
		Z	$k_1\ k_3\ \delta$	$k_2\ k_{in}$ b	Z	X
		all	all	-	all	-
BioSD II	U	X	k_{in} b δ	$k_1\ k_2\ k_3\ \eta$	X	$Z_1\ Z_2$
		$Z_1/Z_2/(Z_1\ Z_2)$	$k_1\ k_3\ \delta\ \eta$	$k_2\ k_{in}$ b	$Z_1\ Z_2$	X
		X Z_1	all	-	all	-
		$X\ Z_2$				
BioSD II MM simple	U	X	k_{in} b $\delta\ K_m$	$k_1\ k_3\ \eta\ V_{max}$	X	$Z_1\ Z_2$
		$Z_1/Z_2/(Z_1\ Z_2)$	$k_1\ k_3\ \delta\ \eta\ V_{max}$	k_{in} b K_m	$Z_1\ Z_2$	X
		$X\ Z_1$	all	-	all	-
		$X\ Z_2$				
BioSD II MM complex	U	X	k_{in} b $\delta\ K_m\ \gamma$	$k_1\ k_3\ \eta\ V_{max}$	X	$Z_1\ Z_2$
		$Z_1/Z_2/(Z_1\ Z_2)$	$k_1\ k_3\ \delta\ \eta\ V_{max}\ \gamma$	k_{in} b K_m	$Z_1\ Z_2$	X
		$X\ Z_1$	all	-	all	-
		$X\ Z_2$				
BioSD III	U	X	k_{in} b	$k_1\ k_2\ k_3\ \delta\ \eta$	X	$Z_1\ Z_2$
		$Z_1/Z_2/(Z_1\ Z_2)$	$k_1\ k_3\ \delta\ \eta$	$k_2\ k_{in}$ b	$Z_1\ Z_2$	X
		$X\ Z_1$	all	-	all	-
		$X\ Z_2$				
Dichotomous Feedback (I)	I	HK/RR/SR	all	-	all	-
	$\beta_{HK}/\beta_{RR}/\beta_{SR}$	HK/RR/SR	$\delta\ k_t\ k_{tc}\ k_p\ k_{pc}$ β_*	$I\ k_{ap-max}\ K_{da}$		
Dichotomous Feedback (k_{ap})	$k_{ap}/\beta_{HK}/\beta_{RR}/\beta_{SR}$	HK/RR/SR	all	-	all	-
sRNA-tuned autorepressor	$u_1\ u_2$	all	all	-	all	-
	γ_R	$s, c, T/c, T$	$\alpha_t\ \alpha_L\ n_T\ K_1$ $\beta_T\ \delta_t\ \delta_s\ \delta_c\ \delta_T$	$K_T\ K_{u_1}\ n_{u_1}\ u_1$		
		t, T	$\alpha_t\ \alpha_L\ n_T\ K_1$ $\beta_T\ \delta_t\ \delta_s\ \delta_c\ \delta_T$	$K_T\ K_{u_1}\ n_{u_1}\ u_1$	t,s,T	c
		s, T	$\alpha_t\ \alpha_L\ n_T\ K_1$ $\beta_T\ \delta_t\ \delta_s\ \delta_T$	$K_T\ K_{u_1}\ n_{u_1}\ u_1\ \delta_c$		
Closed-loop sRNA	$u_2\ u_3$	R	$K_{u_2}\ K_{u_3}\ K_R$ $\delta_R\ \delta_r\ \delta_s$	$\alpha_r^*\ \alpha_X^*\ K_2\ \delta_c\ \beta_R$	R	r s c
		$r/s/r\ s$	$\alpha_r^*\ \alpha_X^*\ K_{u_2}\ K_{u_3}$ $K_2\ \delta_R\ \delta_r\ \delta_s$	$K_R\ \delta_c\ \beta_R$	r s	c R
		$c/r\ c/s\ c/$ $r\ s\ c$	$\alpha_r^*\ \alpha_X^*\ K_{u_2}\ K_{u_3}$ $K_2\ \delta_R\ \delta_r\ \delta_s\ \delta_c$	$K_R\ \beta_R$	r s c	R
		$R\ r/R\ s/R\ r\ s$	$\alpha_r^*\ \alpha_X^*\ K_{u_2}\ K_{u_3}$ $K_2\ K_R\ \delta_R\ \delta_r\ \delta_s\ \beta_R$	δ_c	R r s	c
		$R\ c$	all	-	all	-
	$\gamma_X^*\ \gamma_R^*$	$R/R\ r/R\ s/R\ r\ s$	$K_2\ \delta_R\ \delta_r\ \delta_s\ \beta_R$	δ_c	R r s	c
		$r/s/r\ s$	$K_2\ \delta_r\ \delta_s$	$\delta_R\ \delta_c\ \beta_R$	r s	c R
		$c/r\ c/s\ c/r\ s\ c$	$K_2\ \delta_r\ \delta_s\ \delta_c$	$\delta_R\ \beta_R$	r s c	R
		$R\ c$	all	-	all	-

Hermann, R. and Krener, A.J. (1977). Nonlinear controllability and observability. *IEEE Trans. Autom. Contr.*, 22(5), 728–740.

Himpe, C. (2018). emgr—the empirical gramian framework. *Algorithms*, 11(7), 91.

Kelly, C.L., Harris, A.W.K., Steel, H., Hancock, E.J., Heap, J.T., and Papachristodoulou, A. (2018). Synthetic negative feedback circuits using engineered small rnas. *Nucleic acids research*, 46(18), 9875–9889.

Lewis, A.D. (2001). A brief on controllability of nonlinear systems. *Preprint, https://mast.queensu.ca/ andrew/notes/abstracts/2001a.html*.

Qian, Y., McBride, C., and Del Vecchio, D. (2018). Programming cells to work for us. *Annu. Rev. Control Robot. Auton. Syst.*, 1, 411–440.

Sontag, E.D. (2013). *Mathematical control theory: deterministic finite dimensional systems*. Springer Science & Business Media.

Sootla, A., Delalez, N., Alexis, E., Norman, A., Steel, H., Wadhams, G.H., and Papachristodoulou, A. (2022). Dichotomous feedback: a signal sequestration-based feedback mechanism for biocontroller design. *J. R. Soc. Interface*, 19(189), 20210737.

Sussmann, H.J. (1987). A general theorem on local controllability. *SIAM J. Contr. Optim.*, 25(1), 158–194.

Tunali, E.T. and Tarn, T.J. (1987). New results for identifiability of nonlinear systems. *IEEE Trans. Autom. Contr.*, 32(2), 146–154.

Wieland, F.G., Hauber, A.L., Rosenblatt, M., Tönsing, C., and Timmer, J. (2021). On structural and practical identifiability. *Curr. Opin. Syst. Biol.*, 25, 60–69.