Guaranteed multistability in a microRNA-based genetic network by formal methods

Nicholas Nolan, Emma L. Peterman, Kate E. Galloway,
Inigo Incer, Eduardo D. Sontag, and Domitilla Del Vecchio

***Abstract*— The development of genetic memory devices in synthetic biology is a challenging process that requires extensive analysis and characterization. In mammalian systems, this complexity is compounded by the need for a small DNA payload for efficient delivery into the cell. Previous genetic memory devices have relied exclusively on protein-based regulation, which are limited by their large size; in this paper, we propose a microRNA-based multistable network, which effectively halves the payload size for more efficient delivery. We demonstrate that the system can be multistable, and use formal methods to characterize constraints on design parameters that guarantee multistability. Our results provide a new genetic network topology that can achieve multistability and demonstrate the use of formal methods in the design of sophisticated genetic network architectures against non-convex top-level specifications.**

I. Introduction

Synthetic biology relies on fundamental biomolecular processes to develop biological networks that achieve an end goal. One notable example of a desirable end goal is the development of cellular memory, which can be used in applications such as biomolecular sensing and medical diagnostics [1], [2], [3], [4]. In the natural context, cellular memory is also critical in the healthy development of mammalian cells, allowing for differentiation and other complex endogenous processes [5], [6]. For engineering applications, cellular memory is most commonly achieved through the development of bistable circuits, for which the design and development process remains lengthy, *ad hoc*, and time-consuming [7].

One key bottleneck in translating bistable circuits to application-relevant contexts is the large cargo size of protein-based gene circuits. As the number of proteins required for the circuit to function increases, the difficulty of reliably delivering the circuit to mammalian cells increases due to the cargo size limitations of commonly used genomic integration methods such as lentivirus and CRISPR / Cas9 [8]. Here, we propose a multistable circuit that substitutes protein-based transcriptional repression for microRNA-based repression. MicroRNAs are a class of short RNA molecules that can bind to target mRNA molecules and induce their degradation via the RNA-induced silencing complex (RISC) [9]. These microRNA sequences can be compactly encoded in a synthetic intron of a gene and only add about 200 base pairs to the circuit cargo size [10]. As such, utilizing microRNA-based repression in place of transcriptional repression could reduce the cargo size of previously designed bistable circuits approximately by half [7].

Our proposed design utilizes two genes, each encoding a transcriptional activator with a microRNA-containing intron and microRNA target sites for the opposite microRNA. Multistability is achieved through the combination of mutual inhibition via microRNA-based repression and positive feedback via transcriptional activation. The microRNA-based repression can be tuned through varying the microRNA binding affinity with different microRNA sequences or numbers of target sites [10]. Similarly, the transcriptional activation can be tuned by varying the binding domain of the transcriptional activator as well as the number and spacing of binding sites [11]. With these design parameters, we seek to compute the parameter region that guarantees multistability.

Previous attempts to perform systematic genetic circuit design against a top-level specification include Cello, which provides for bacterial cells a circuit design that implements a specified logic map [12], [13]. However, these methods do not allow to specify multistability and cannot account for context-dependence, such as due to resource competition [14], [15]. Other approaches have allowed for a wider range of flexibility in including context-dependence such as resource sharing and offer some limited testing of parameters [16], [17]. These, however, do not provide formal guarantees on design parameters against a top-level specification. To address this question, we use an analytical method that uses linear programming and leverages the monotone structure of the genetic network architecture [18], [19]. This approach provides a computationally tractable method for guaranteeing a top-level specification while accounting for context effects.

Contributions. In this paper, we consider a novel genetic network that can achieve multistability. We formulate multistability as a formal specification on the system's output and harness the system's monotonicity to explicit determine algebraic conditions on the parameters guaranteeing that the specification is met.

Paper outline. Section II introduces the chemical reactions that implement the multistable microRNA system. Section III proves global convergence properties of this system and shows that, under some parameter regimes, the system can exhibit multistability. Section IV applies formal methods to determine a set of tunable parameters that guarantees multistability. Section V concludes the paper.

This work was supported by AFOSR MURI Grant #FA9550-22-1-0316.

N. Nolan, E. L. Peterman, K. E. Galloway, and D. Del Vecchio are with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

I. Incer is with the University of Michigan, Ann Arbor, MI 48109 USA.

E. D. Sontag is with Northeastern University, Boston, MA 02115 USA.

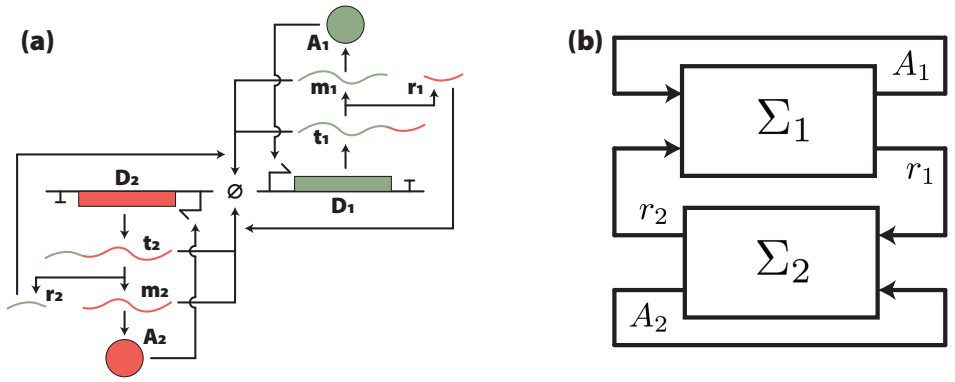

Fig. 1: The proposed multistable gene network. (a) Genetic diagram of the microRNA-based network. Each module's DNA molecule D_i encodes a transcript RNA t_i, which splices into an active mRNA molecule m_i and a microRNA r_i. The active mRNA is translated into an activator protein A_i, which activates the DNA D_i. The other module's microRNA r_j induces the degradation of the original module's transcript and active messenger RNA molecules. (b) Block diagram representation of the genetic network in (a), comprising self activation (via A_1 and A_2) and mutual inhibition (via r_1 and r_2).

II. System Model Description

The reaction network constituting our proposed multistable network comprises two identical modules i, where $i \in \{1,2\}$. In each module, a DNA molecule D_i is transcribed to produce an mRNA transcript molecule t_i. The transcript t_i is spliced to form an active mRNA m_i and a microRNA r_i. The microRNA of the other module $\mathrm{r}_j, j \neq i$, binds to both t_i and m_i, forming complexes T_i and M_i, respectively, which lead to the degradation of t_i and m_i. The active mRNA m_i is translated to produce a protein A_i, which binds to D_i to form a complex C_i, which produces t_i. All RNA species degrade, and the activators dilute. This system is shown in Figure 1. The following chemical reactions correspond to this model for $i, j \in \{1,2\}, j \neq i$:

$$
\begin{aligned}
&\mathrm{D}_i \xrightarrow{\alpha_{0,i}} \mathrm{D}_i + \mathrm{t}_i \qquad && \mathrm{A}_i + \mathrm{D}_i \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C}_i \xrightarrow{\alpha_i} \mathrm{C}_i + \mathrm{t}_i \\
&\mathrm{t}_i \xrightarrow{k_s} \mathrm{m}_i + \mathrm{r}_i \qquad && \mathrm{t}_i + \mathrm{r}_j \underset{d_i'}{\overset{a_i'}{\rightleftharpoons}} \mathrm{T}_i \xrightarrow{k_i'} \mathrm{r}_j \\
&\mathrm{m}_i \xrightarrow{\kappa} \mathrm{m}_i + \mathrm{A}_i \qquad && \mathrm{m}_i + \mathrm{r}_j \underset{d_i'}{\overset{a_i'}{\rightleftharpoons}} \mathrm{M}_i \xrightarrow{k_i'} \mathrm{r}_j \\
&\mathrm{t}_i \xrightarrow{\delta} \emptyset \qquad && \mathrm{m}_i \xrightarrow{\delta} \emptyset \\
&\mathrm{r}_i \xrightarrow{\delta} \emptyset \qquad && \mathrm{A}_i \xrightarrow{\gamma} \emptyset,
\end{aligned}
$$

from which the following reaction rate equations can be derived using mass-action kinetics [20]:

$$
\begin{aligned}
\dot{t}_i &= \alpha_{0,i} D_i + \alpha_i C_i + d_i' T_i - (k_s + \delta + a_i' r_j) t_i \\
\dot{r}_i &= k_s t_i + (d_i' + k_i')(T_j + M_j) - (\delta + a_i'(t_j + m_j)) r_i \\
\dot{m}_i &= k_s t_i + d_i' M_i - (\delta + a' r_j) m_i \\
\dot{A}_i &= \kappa m_i + d C_i - a A_i D_i - \gamma A_i \\
\dot{C}_i &= a A_i D_i - d C_i \\
\dot{T}_i &= a_i' t_i r_j - (d_i' + k_i') T_i \\
\dot{M}_i &= a_i' m_i r_j - (d_i' + k_i') M_i,
\end{aligned}
\tag{1}
$$

where $i \in \{1,2\}$ and $\alpha_i > \alpha_{i,0}$, since $\alpha_{i,0}$ represents basal expression while α_i represents the maximal transcription rate post-activation. We define D_{it} as the total concentration of the i-th module's DNA — which is conserved — so that we have the following conservation law:

$$
D_{it} = D_i + C_i, \tag{2}
$$

This system contains both fast and slow variables. In general, complexation reactions take place very quickly, and RNA dynamics are typically also fast compared to protein dynamics [20]. We will thus simplify our system sequentially, first by taking C_i, T_i, and M_i to their quasi-steady states, and then by taking m_i, t_i, and r_i to their quasi-steady states. To begin, we consider the quasi-steady state values of C_i, T_i, and M_i, which are obtained by setting the respective derivatives in (1) to 0:

$$
C_i = \frac{A_i D_i}{K_i} \qquad T_i = \frac{t_i r_j}{K_i'} \qquad M_i = \frac{m_i r_j}{K_i'}, \tag{3}
$$

where $K_i = (d + \alpha_i)/a$ and $K_i' = (d_i' + k_i')/a_i'$. Plugging (3) into (2), we find the concentrations of free DNA D_i as:

$$
D_i = \frac{D_{it}}{1 + \frac{A_i}{K_i}}. \tag{4}
$$

We substitute (2), (3), and (4) into (1), and take species t_i, then m_i, and finally r_i to the quasi-steady state to arrive at the system:

$$
\begin{aligned}
\dot{A}_1 &= \kappa \frac{F_1(A_1)}{\left(1 + \frac{\delta}{k_s} + \frac{k_1'}{k_s K_1'} r_2\right)\left(\delta + \frac{k_1'}{K_1'} r_2\right)} - \gamma A_1 \\
\dot{A}_2 &= \kappa \frac{F_2(A_2)}{\left(1 + \frac{\delta}{k_s} + \frac{k_2'}{k_s K_2'} r_1\right)\left(\delta + \frac{k_2'}{K_2'} r_1\right)} - \gamma A_2,
\end{aligned}
\tag{5}
$$

where:

$$
F_i(A_i) = \frac{\alpha_{0,i} + \alpha_i \frac{A_i}{K_i}}{1 + \frac{A_i}{K_i}} D_{it} \tag{6}
$$

is the traditional Hill function for transcriptional activation [20], and each of r_1 and r_2 is the unique positive solution to the quadratic equations:

$$\frac{k_s F_1(A_1)}{\delta k_2'/K_2'} = r_1^2 + \left(\frac{k_s+\delta}{k_2'/K_2'} + \frac{k_s k_1'/K_1'}{\delta k_2'/K_2'} F_2(A_2) - F_1(A_1)\right) r_1$$
$$\frac{k_s F_2(A_2)}{\delta k_1'/K_1'} = r_2^2 + \left(\frac{k_s+\delta}{k_1'/K_1'} + \frac{k_s k_2'/K_2'}{\delta k_1'/K_1'} F_1(A_1) - F_2(A_2)\right) r_1.$$

The parameters α_i, D_{it}, k_i', and K_i' can be tuned to design the system against desired specifications. The other parameters we regard as fixed; the values that we use for these fixed parameters in all subsequent simulations are reported in Table I.

Given this system model, we next investigate the asymptotic behavior of the system's trajectories and determine whether it can be multistable.

III. Convergence of system to equilibria

A. Global convergence of system's trajectories

We now show that the system's trajectories converge globally to a set of equilibria. To this end, we first prove the following Lemma, which is an intermediate result to showing that the system is strongly monotone [21].

Lemma 1. *Suppose that the functions r and c are differentiable functions $J \to \mathbb{R}$, where $J \subseteq \mathbb{R}$, such that:*

- *$r(x)$, $c(x)$, and $c'(x)$ are positive for all x,*
- *$b'(x)$ is negative for all x, and*
- *the following equality holds for all x:*

$$r(x)^2 + b(x)r(x) - c(x) = 0. \tag{7}$$

Then $r'(x) > 0$ for all x.

Proof. Taking derivatives in (7), we have:

$$r'(x) = \frac{c'(x) - b'(x)r(x)}{2r(x) + b(x)}.$$

Since $c'(x)$ and $r'(x)$ are positive and $b'(x)$ is negative, the numerator is positive. Thus all we need to prove is that $2r(x) + b(x) > 0$ for all x. Since $r(x) > 0$, it is enough to show that $r(x) + b(x) > 0$. Now, from (7), we have that $r(x) + b(x) = c(x)/r(x)$, so this is true by positivity of $c(x)$ and $r(x)$. □

We apply this Lemma to system (5), for which we define $\dot{A}_1 := \Phi_1(A_1, A_2)$ and $\dot{A}_2 := \Phi_2(A_1, A_2)$. We want to show that $\partial r_i/\partial A_i > 0$. It then follows that $\partial\Phi_1/\partial A_2 < 0$ and $\partial\Phi_2/\partial A_1 < 0$, showing that this planar system is competitive [21], and hence monotone. In fact, the strict inequalities establish strong monotonicity in the sense of [21].

We show this for $i = 1$, the proof is analogous for $i = 2$. Recall that we are assuming that $\alpha_{i,0} < \alpha_i$ for $i \in \{1, 2\}$, which ensures that $dF_i(A_i)/dA_i > 0$. Let $x = A_1$, $b(x) = -F_1(A_1) + \frac{k_s+\delta}{k_2'/K_2'} + \frac{k_s k_1'/K_1'}{\delta k_2'/K_2'} F_2(A_2)$, and $c(x) = \frac{k_s}{\delta k_2' K_2'} F_1(A_1)$. The result then follows from applying Lemma 1.

Parameter	Value	Source
α_i	**Design Parameter**	[11]
D_{it}	**Design Parameter**	[10]
k_i'/K_i'	**Design Parameter**	[10]
$\alpha_{0,i}$	$4.62 \cdot 10^{-3}\text{s}^{-1}$	[11]
K_i	$[10, 1000]$ pM·s^{-1}	[11]
k_s	$2.00 \cdot 10^{-3}\text{s}^{-1}$	[10]
δ	$2.88 \cdot 10^{-4}\text{s}^{-1}$	[10]
κ	$3.33 \cdot 10^{-4}\text{s}^{-1}$	[10]
γ	$9.67 \cdot 10^{-5}\text{s}^{-1}$	[10]

TABLE I: Numerical values of the parameters. Design parameters are explicitly listed as such. The parameter K_i is, unless otherwise noted, taken as 100.

According to Theorem 2.2 in [21], any planar competitive system with precompact trajectories has the property that all solutions converge to equilibria. Precompactess of trajectories follows from the fact that, for each i, $\dot{A}_i$ is negative for large enough A_i, since the production term saturates and the degradation term is linear.

A subtle issue is to show that the system obtained when substituting $r_i(A_1, A_2)$ into the A_1 and A_2 equations results in a well-posed system. This follows from the fact that these are smooth functions of A_1, A_2, which in turn follows from the following observation regarding Lemma 1. For any fixed x, Equation (7) has a unique positive solution $r(x)$, because the parabola $P_x(r) = r^2 + b(x)r - c(x)$ is negative at $r = 0$. Moreover, if we take the partial derivative of $P_x(r)$ with respect to r, then $rP_x'(r) = 2r^2 + b(x)r = r^2 + (r^2 + b(x)r) = r^2 + c(x) > 0$ whenever $P_x(r) = 0$, and hence also $P_x(r)$ is nonzero. By the Implicit Function Theorem we know that r is a smooth function of x. In other words, differentiability of r is automatically satisfied because b and c are differentiable.

B. Numerical demonstration of multistability

We now perform a numerical study to demonstrate multistability in this microRNA-based system. Specifically, we plot the number of distinct positive solutions as the design parameters k_i'/K_i' vary logarithmically from 10^{-5} to 1 in Fig. 2(a). As it can be seen from this diagram, increasing k_i'/K_i' allows for up to 5 possible steady states. In Figures 2(b-d), a selection of nullclines that confer either 1, 3, or 5 steady states is shown. In regimes of 3 or more steady states, multiple concurrent stable steady states are observed, thus confirming that this microRNA-based system can exhibit multistability.

IV. Parameter space that guarantees multistability

In the previous section, we have demonstrated that the system can be multistable and that two of the stable steady states appear in regions where either one or the other subsystem's outputs A_i are high. In this section, we formally introduce the system's specification and find the set of parameters that guarantee the specification. In this system, we have access to three pairs of design parameters, namely D_{it}, α_i, and k_i'/K_i'. The total DNA of a module can be tuned

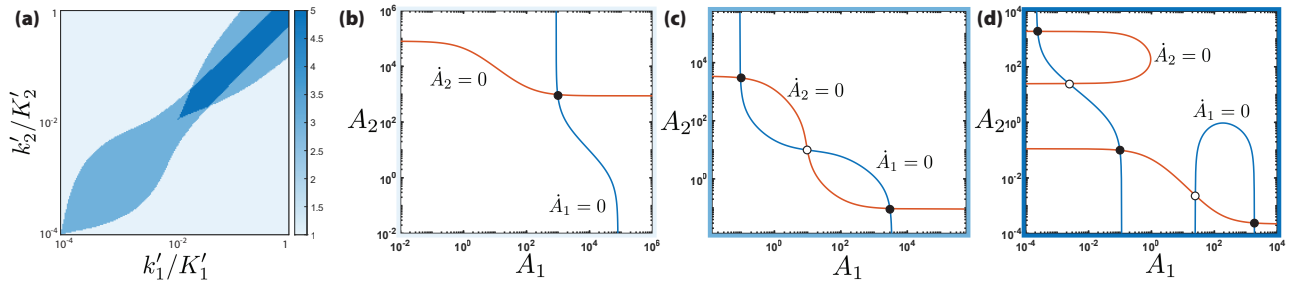

Fig. 2: Demonstration that the system can exhibit multistability. (a) Bifurcation plot as the two design parameters $\frac{k_i'}{K_i'}$ are logarithmically increased from 10^{-5} to 1. The number of equilibria in the system increases with these design parameters, up to a total of 5. (b-d) For each possible number of steady states, a sample $\left(\frac{k_1'}{K_1'}, \frac{k_2'}{K_2'}\right)$ is taken and its nullcline is plotted. The stability of each equilibrium was computed and is shown as either a black fill at the nullcline intersection (indicating the equilibrium is stable) or a white fill (indicating the equilibrium is unstable). Design parameters used: $\alpha_1 = \alpha_2 = 4 \cdot 10^{-1}\text{s}^{-1}$ and $D_{1t} = D_{2t} = 10\text{pM}$. $\left(\frac{k_1'}{K_1'}, \frac{k_2'}{K_2'}\right) =$ (b) $\left(10^{-5}, 10^{-5}\right)$, (c) $\left(10^{-3}, 10^{-3}\right)$, (d) $\left(10^{-1}, 10^{-1}\right)$.

via the DNA copy number, the maximal transcription rate can be tuned by the choice of promoter, and the catalytic rate/dissociation constant can be tuned by variations in the microRNA sequence itself [10]. We wish to characterize a parameter regime over these design parameters such that multistability is guaranteed. To this end, we represent the system by the block diagram in Figure 1(b), so that we have two subsystems Σ_i with the following dynamics:

$$\begin{aligned} \dot{r}_i &= \frac{F_i(A_i)}{1 + \frac{\delta}{k_s} + \frac{k_i'}{k_s K_i'} r_i} - \delta r_i \\ \dot{A}_i &= \kappa \frac{F_i(A_i)}{\left(1 + \frac{\delta}{k_s} + \frac{k_i'}{k_s K_i'} r_i\right)\left(\delta + \frac{k_i'}{K_i'} r_i\right)} - \gamma A_i, \end{aligned} \tag{8}$$

with outputs:

$$y_i = (r_i, A_i),$$

and inputs:

$$u_i = (y_{i2}, y_{j1}).$$

We wish to select our design parameters so that the system contains steady states in each of two regions, where either y_{11}, y_{12} are high and y_{21}, y_{22} are low, or y_{11}, y_{12} are low and y_{21}, y_{22} are high. Therefore, letting $(y_{11}, y_{12}, y_{21}, y_{22})$ be a steady state of (8), we seek to determine the set of parameters such that the specification:

$$\phi^A(y)\colon \left(y_{11} \geq y_{11}^H \wedge y_{21} \leq y_{21}^L \wedge y_{12} \geq y_{12}^H \wedge y_{22} \leq y_{22}^L\right), \tag{9}$$

is satisfied, where the superscripts indicate high (H) or low (L) limits. These can be arbitrarily set, and for our example we use the values listed in Table II. This ensures the existence of a steady state in Region A of Fig. 3(a). Then, we find the set of parameters for which the specification:

$$\phi^B(y)\colon \left(y_{11} \leq y_{11}^L \wedge y_{21} \geq y_{21}^H \wedge y_{12} \leq y_{12}^L \wedge y_{22} \geq y_{22}^H\right), \tag{10}$$

is satisfied, so that there is a steady state in Region B of Fig. 3(a). Then, we will take the intersection of the two sets

Threshold	Value
y_{i1}^H	1×10^2
y_{j1}^L	1×10^1
y_{i2}^H	1×10^2
y_{j2}^L	1×10^{-1}
y_{i1}^H	1×10^2
y_{j1}^L	1×10^1
y_{i2}^H	1×10^2
y_{j2}^L	1×10^{-1}

TABLE II: Numerical values for threshold constants.

of parameters identified to guarantee the co-existence of at least two steady states, one in each of Regions A and B (Fig. 3(b)).

To this end, we use a linear programming tool known as Pacti to follow a similar strategy as used to find parameter spaces for a variety of genetic circuits in our previous work [19]. Specifically, we consider a circuit with m inputs u_i and n outputs y_j, in which we want our system to satisfy the specification $\phi(y)\colon \bigwedge_j \phi_j(y_j)$ as long as the inputs u satisfy $\sigma(u)\colon \bigwedge_i \sigma_i(u_i)$, where $\sigma_i\colon u_i \in U_i$ for given sets U_i. We use the following theorem:

Theorem 1. [19] *Assume that $\phi_k(y_k)$ is of the form $F_k(G_1^k(u_1), \ldots, G_m^k(u_m), \theta) \geq 0$, where θ is a vector of parameters and the functions $F_k(x_1, \ldots, x_m, \theta)$ are monotone in their first m arguments, then the most relaxed specification $\psi(\theta)$ that ensures that the system satisfies $\phi(y)$ when the inputs satisfy $\sigma(u)$ is given by*

$$\psi(\theta)\colon \bigwedge_k F_k(g_1^k, \ldots, g_m^k, \theta) \geq 0,$$

where:

$$g_i^j = \begin{cases} \min\limits_{u \models \sigma} G_i^k(u_i) & \text{if } \frac{\partial F_k}{\partial x_i}(G_1^k(u_1), \ldots, G_m^k(u_m), \theta) \geq 0 \\ & \text{for all } u_p \models \sigma_p \\ \max\limits_{u \models \sigma} G_i^k(u_i) & \text{otherwise.} \end{cases} \tag{11}$$

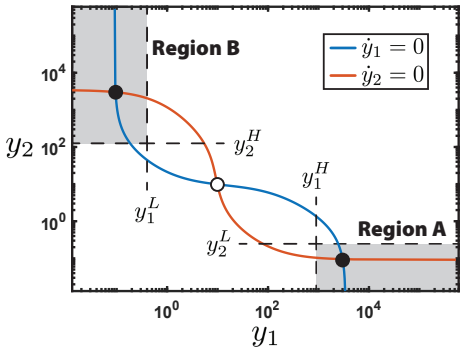

Fig. 3: Pacti-based approach to characterize parameters that ensure multistability. (a) Pacti allows the user to specify regions of parameters over which a user-specificied condition is met via quantifier elimination. We will use this to first identify the parameters for which there exists a steady state in region I; after that, we will perform the same process to identify a parameter set for which there exists a steady state in region II. (b) We will then take the intersection of these two design parameter sets to find parameter regimes for which a steady state exists in both regions, and will manually verify that this yields at least two stable steady states.

The utility of this theorem lies in the fact thar $\psi(\theta)$ becomes an explicit set of constraints on θ, thereby removing the need to solve a quantifier elimination problem [19]. In order to apply this theorem, we consider each of the specifications ϕ^A and ϕ^B separately. For the first one in expression (9), we note that it is in the form $\phi(y)\colon \bigwedge_k \phi_k(y)$, specifically:

$$\phi^A(y)\colon \left(\underbrace{y_{i1} \geq y_{i1}^H}_{\phi_1^A} \wedge \underbrace{y_{j1} \leq y_{j1}^L}_{\phi_2^A} \wedge \underbrace{y_{i2} \geq y_{i2}^H}_{\phi_3^A} \wedge \underbrace{y_{j2} \leq y_{j2}^L}_{\phi_4^A} \right). \tag{12}$$

We then write the steady-state conditions of the system as:

$$\delta y_{i1} = \frac{\alpha_{0,i} D_{it} + \alpha_i \frac{y_{i2}}{K_i} D_{it}}{\left(1 + \frac{y_{i2}}{K_i}\right)\left(1 + \frac{\delta}{k_s} + \frac{k_i'}{k_s K_i'} y_{j1}\right)} \tag{13}$$

$$\gamma y_{i2} = \kappa \frac{\alpha_{0,i} D_{it} + \alpha_i \frac{y_{i2}}{K_i} D_{it}}{\left(1 + \frac{y_{i2}}{K_i}\right)\left(1 + \frac{\delta}{k_s} + \frac{k_i'}{k_s K_i'} y_{j1}\right)\left(\delta + \frac{k_i'}{K_i'} y_{j1}\right)}, \tag{14}$$

and define new parameters $p_i := \alpha_{0,i} D_{it}$, $q_i := (\alpha_i - \alpha_{0,i}) D_{it}$, and $s_i := \frac{k_i'}{K_i'}$, which encapsulate all tunable parameters in our system. Using these definitions, and substituting (13) and (14) into (12), we have our top-level specification as:

$$\phi^A(y)\colon \left(\frac{\left(p_i + q_i \frac{y_{i2}}{K_i}\right) / \left(1 + \frac{y_{i2}}{K_j}\right)}{\left(1 + \frac{\delta}{k_s} + s_i \frac{y_{j1}}{k_s}\right)} - \delta y_{i1}^H \geq 0 \right) \wedge \tag{15}$$

$$\left(\delta y_{j1}^L - \frac{\left(p_j + q_j \frac{y_{j2}}{K_j}\right) / \left(1 + \frac{y_{j2}}{K_j}\right)}{\left(1 + \frac{\delta}{k_s} + s_j \frac{y_{i1}}{k_s}\right)} \geq 0 \right) \wedge \tag{16}$$

$$\left(\frac{\left(p_i + q_i \frac{y_{i2}}{K_i}\right) / \left(1 + \frac{y_{i2}}{K_j}\right)}{\left((\delta + k_s)\frac{\delta}{k_s} + s_i y_{j1}\left(1 + \frac{2\delta}{k_s}\right) + s_i^2 y_{j1}^2\right)} - \frac{\gamma}{\kappa} y_{i2}^H \geq 0 \right) \wedge \tag{17}$$

$$\left(\frac{\gamma}{\kappa} y_{j2}^L - \frac{\left(p_j + q_j \frac{y_{j2}}{K_j}\right) / \left(1 + \frac{y_{j2}}{K_j}\right)}{\left((\delta + k_s)\frac{\delta}{k_s} + s_j y_{i1}\left(1 + \frac{2\delta}{k_s}\right) + s_j^2 y_{i1}^2\right)} \geq 0 \right), \tag{18}$$

for which each $\phi_k^A(y)$ is of the form $F_k(y_{11}, \ldots, y_{22}, \theta) \geq 0$, where $\theta = (p_i, q_i, s_i), i \in \{1, 2\}$, is our vector of system parameters. In order to apply Theorem 1, we must have that each F_k is monotone in $(y_{11}, y_{12}, y_{21}, y_{22})$, which we prove in the following claim.

Claim 1. Let $F_k(y_{11}, \ldots, y_{22}, \theta)$ for $k \in \{1, 2, 3, 4\}$ be given in expressions (15-18). Then, they are monotonic in their arguments.

Proof. The right-hand sides of both (13) and (14) are monotonically decreasing with y_{j1}. Regarding monotonicity with respect to y_{i2}, we note that $p_i = \alpha_{0,i} D_{it}$ and $q_i = \alpha D_{it}$ are such that $\alpha_{0,i} < \alpha$ (assumption in Section II). Therefore, we further have that $p_i < q_i$ and thus each F_k is monotonically increasing with y_{i2}. □

There are no exogenous inputs to the system, since $u = y$, so we additionally have the constraints:

$$\sigma(y)\colon \left(y_{i1} \geq y_{i1}^H\right) \wedge \left(0 \leq y_{j1} \leq y_{j1}^L\right) \wedge \left(y_{i2} \geq y_{i2}^H\right) \wedge \left(0 \leq y_{j2} \leq y_{j2}^L\right).$$

By applying Theorem 1, we arrive at the most relaxed specification that ensures satisfaction of $\phi^A(y)$, that is $\psi^A(p, q, s)$ given by

$$\psi^A\colon \left(\frac{\left(p_i + q_i \frac{y_{i2}^H}{K_i}\right) / \left(1 + \frac{y_{i2}^H}{K_j}\right)}{\left(1 + \frac{\delta}{k_s} + s_i \frac{y_{j1}^L}{k_s}\right)} - \delta y_{i1}^H \geq 0 \right) \wedge$$

$$\left(\delta y_{j1}^L - \frac{\left(p_j + q_j \frac{y_{j2}^L}{K_j}\right) / \left(1 + \frac{y_{j2}^L}{K_j}\right)}{\left(1 + \frac{\delta}{k_s} + s_j \frac{y_{i1}^H}{k_s}\right)} \geq 0 \right) \wedge$$

$$\left(\frac{\left(p_i + q_i \frac{y_{i2}^H}{K_i}\right) / \left(1 + \frac{y_{i2}^H}{K_j}\right)}{\left((\delta + k_s)\frac{\delta}{k_s} + s_i y_{j1}^L\left(1 + \frac{2\delta}{k_s}\right) + s_i^2 (y_{j1}^L)^2\right)} - \frac{\gamma}{\kappa} y_{i2}^H \geq 0 \right) \wedge$$

$$\left(\frac{\gamma}{\kappa} y_{j2}^L - \frac{\left(p_j + q_j \frac{y_{j2}^L}{K_j}\right) / \left(1 + \frac{y_{j2}^L}{K_j}\right)}{\left((\delta + k_s)\frac{\delta}{k_s} + s_j y_{i1}^H\left(1 + \frac{2\delta}{k_s}\right) + s_j^2 (y_{i1}^H)^2\right)} \geq 0 \right). \tag{19}$$

Using this result, we now substitute in the values for each of the fixed parameters and thresholds as defined in Tables I

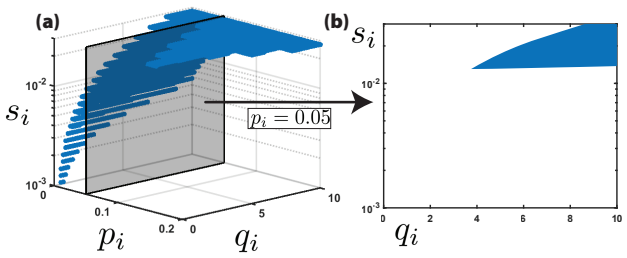

Fig. 4: Feasible region of parameters for which the microRNA-based toggle switch system exhibits multistability. (a) Plot of the 3-dimensional design parameter space for one module of the multistable system. (b) By taking $p_i = 0.05$, we can reduce the system to a 2-dimensional plane for easier visualization.

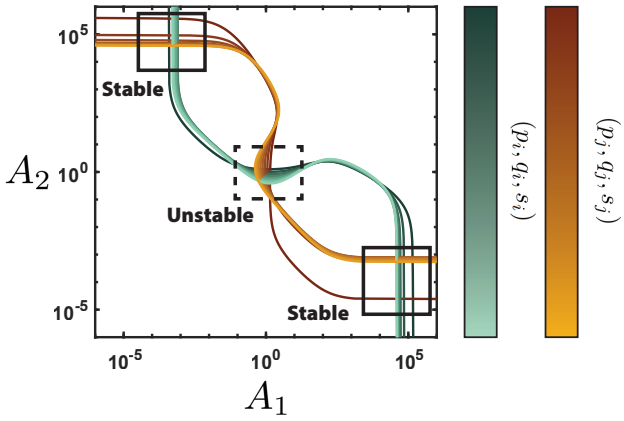

Fig. 5: System behavior predictions based on values from within the feasible region seen in Fig. 4. The intersecting nullclines show equilibria; for all nullclines sampled here, we observe three intersections, of which two are stable and one is unstable. The parameters $(p_i, q_i, s_i, p_j, q_j, s_j)$ are varied linearly from $(10^{-3}, 4, 7 \times 10^{-3}, 2 \times 10^{-3}, 4.4, 7.9 \times 10^{-3})$ (darker) to $(2.8 \times 10^{-2}, 9.6, 1.91 \times 10^{-2}, 3 \times 10^{-2}, 10, 2 \times 10^{-2})$ (lighter). All other parameters are set to the nominal values noted in Table I.

and II. We subsequently use the solver from [19] to rearrange (19) and identify the following explicit constraints on design parameters:

$$\begin{aligned} 0.5p_i + 0.5q_i - 144s_i &\geq 3.30 \times 10^{-2} \\ 0.999p_j + 9.99 \times 10^{-4}q_j - 144s_j &\leq 3.30 \times 10^{-3} \\ 0.5p_i + 0.5q_i - 37.4s_i - 2.90s_i^2 &\geq 9.57 \times 10^{-4} \\ 0.999p_j + 9.99 \times 10^{-4}q_j - 0.374s_j - 290s_j^2 &\leq 9.57 \times 10^{-7}. \end{aligned}$$

We perform a similar analysis considering the specification $\phi^B(y)$ (with y_{i1}, y_{i2} low and y_{j1}, y_{j2} high) to obtain a set of feasible parameters that ensure the system contains a steady state in Region B of Fig. 3. By taking the intersections of these two regions, we can identify a final set of constraints for which the system contains a steady state in both regions. We then check the Jacobian of the system to confirm that each of these regional steady states is stable, thus ensuring that we have identified a multistable system. This region is visualized in Fig. 4, and a sampling of nullclines that demonstrate this multistability is shown in Fig. 5.

To demonstrate how one can then use this formulation to assess the impact of other parameters on the design space, we investigate the role of selecting an activator by varying the dissociation constant K_i. To this end, we consider the change of the feasible region in Fig. 6. Here, we see that activators with low dissociation constants require a higher value of s_i to lead to satisfaction of the specification; increasing the dissociation constant offers a tradeoff between the minimal s_i to be feasible, and the maximal p_i before feasibility is lost.

V. Conclusions and Future Work

Here, we considered a novel microRNA-based multistable genetic network to achieve cellular memory in mammalian cells while minimizing the genetic cargo size. We mathematically demonstrated that the system can be multistable and then used formal methods along the monotinicity of the system to explicitly comute the parameter space guaranteeing multistability. We additionally characterized the sensitivity of this parameter region to the user's choice of activator protein, finding a tradeoff between the microRNA's affinity to its target and the maximal leaky expression tolerable before feasibility is lost. In the future, we will consider experimental implementation of this microRNA-based circuit, following the parameter design guidelines provided here, and subsequent validation of the model's predictions of mono-, bi-, and tristability under different parameter regimes.

References

[1] G. Rong, S. R. Corrie, and H. A. Clark, "In vivo biosensing: progress and perspectives," *ACS sensors*, vol. 2, no. 3, pp. 327–338, 2017.

[2] K. Pardee, A. A. Green, M. K. Takahashi, D. Braff, G. Lambert, J. W. Lee, T. Ferrante, D. Ma, N. Donghia, M. Fan *et al.*, "Rapid, low-cost detection of zika virus using programmable biomolecular components," *Cell*, vol. 165, no. 5, pp. 1255–1266, 2016.

[3] K. Pardee, A. A. Green, T. Ferrante, D. E. Cameron, A. DaleyKeyser, P. Yin, and J. J. Collins, "based synthetic gene networks," *Cell*, vol. 159, no. 4, pp. 940–954, 2014.

[4] D. T. Riglar, T. W. Giessen, M. Baym, S. J. Kerns, M. J. Niederhuber, R. T. Bronson, J. W. Kotula, G. K. Gerber, J. C. Way, and P. A. Silver, "Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation," *Nature biotechnology*, vol. 35, no. 7, pp. 653–658, 2017.

[5] S. Bruno, T. M. Schlaeger, and D. Del Vecchio, "Epigenetic oct4 regulatory network: stochastic analysis of cellular reprogramming," *npj Systems Biology and Applications*, vol. 10, no. 1, p. 3, 2024.

[6] K. Ilia, N. Shakiba, T. Bingham, R. D. Jones, M. M. Kaminski, E. Aravera, S. Bruno, S. Palacios, R. Weiss, J. J. Collins *et al.*, "Synthetic genetic circuits to uncover the oct4 trajectories of successful reprogramming of human fibroblasts," *Science Advances*, vol. 9, no. 48, p. eadg8495, 2023.

[7] T. Lebar, U. Bezeljak, A. Golob, M. Jerala, L. Kadunc, B. Pirš, M. Stražar, D. Vučko, U. Zupančič, M. Benčina, V. Forstnerič, R. Gaber, J. Lonzarić, A. Majerle, A. Oblak, A. Smole, and R. Jerala, "A bistable genetic switch based on designable DNA-binding domains," *Nature Communications*, vol. 5, no. 1, p. 5007, Sep. 2014.

[8] E. L. Peterman, D. S. Ploessl, and K. E. Galloway, "Accelerating Diverse Cell-Based Therapies Through Scalable Design," *Annual Review of Chemical and Biomolecular Engineering*, vol. 15, pp. 6.1–6.26, 2024.

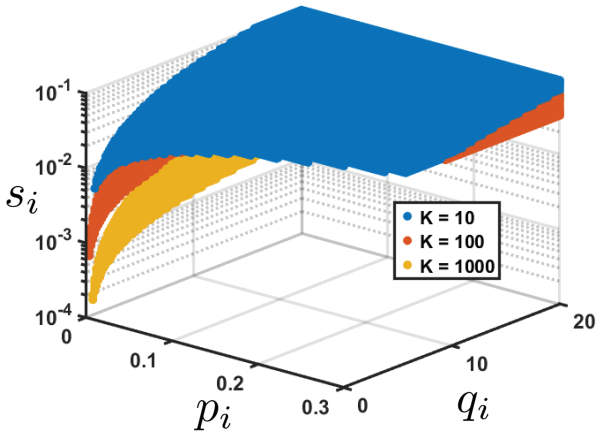

Fig. 6: Visualizing the impacts of varying the activator protein's dissociation constant K_i. The feasible region is plotted as K_i and K_j are each increased from 10 to 1000, encompassing the full range of values noted in Table I. As can be seen, increasing K_i increases the system's sensitivity to p_i.

[9] M. Leisner, L. Bleris, J. Lohmueller, Z. Xie, and Y. Benenson, "Rationally designed logic integration of regulatory signals in mammalian cells," *Nature Nanotechnology*, vol. 5, no. 9, pp. 666–670, Sep. 2010.

[10] K. S. Love, C. P. Johnstone, E. L. Peterman, S. Gaglione, and K. E. Galloway, "Model-guided design of microrna-based gene circuits supports precise dosage of transgenic cargoes into diverse primary cells," *bioRxiv*, 2024.

[11] P. S. Donahue, J. W. Draut, J. J. Muldoon, H. I. Edelstein, N. Bagheri, and J. N. Leonard, "The COMET toolkit for composing customizable genetic programs in mammalian cells," *Nature Communications*, vol. 11, no. 1, p. 779, Dec. 2020.

[12] A. A. Nielsen, B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. A. Strychalski, D. Ross, D. Densmore, and C. A. Voigt, "Genetic circuit design automation," *Science*, vol. 352, no. 6281, p. aac7341, 2016.

[13] T. S. Jones, S. M. Oliveira, C. J. Myers, C. A. Voigt, and D. Densmore, "Genetic circuit design automation with cello 2.0," *Nature protocols*, vol. 17, no. 4, pp. 1097–1113, 2022.

[14] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. Del Vecchio, "Resource competition shapes the response of genetic circuits," *ACS synthetic biology*, vol. 6, no. 7, pp. 1263–1272, 2017.

[15] R. D. Jones, Y. Qian, V. Siciliano, B. DiAndreth, J. Huh, R. Weiss, and D. Del Vecchio, "An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells," *Nature communications*, vol. 11, no. 1, p. 5690, 2020.

[16] W. Poole, A. Pandey, A. Shur, Z. A. Tuza, and R. M. Murray, "Biocrnpyler: Compiling chemical reaction networks from biomolecular parts in diverse contexts," *PLOS Computational Biology*, vol. 18, no. 4, p. e1009987, 2022.

[17] A. Swaminathan, W. Poole, A. Pandey, V. Hsiao, and R. M. Murray, "Fast and flexible simulation and parameter estimation for synthetic biology using bioscrape," *BioRxiv*, p. 121152, 2017.

[18] I. Incer, A. Badithela, J. Graebener, P. Mallozzi, A. Pandey, S.-J. Yu, A. Benveniste, B. Caillaud, R. M. Murray, A. Sangiovanni-Vincentelli *et al.*, "Pacti: Scaling assume-guarantee reasoning for system analysis and design," *arXiv preprint arXiv:2303.17751*, 2023.

[19] I. Incer, A. Pandey, N. Nolan, E. L. Peterman, K. E. Galloway, E. D. Sontag, and D. Del Vecchio, "Guaranteeing system-level properties in genetic circuits subject to context effects," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*. IEEE, 2024, pp. 5558–5565.

[20] D. Del Vecchio and R. M. Murray, *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[21] H. Smith, *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems, Mathematical Surveys and Monographs, vol. 41*. Providence, RI: AMS, 1995.