UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mathematical Modeling of Biological Networks

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Vahid Mardanlou

March 2017

Dissertation Committee:

Dr. Elisa Franco, Chairperson
Dr. Ertem Tuncel
Dr. Fabio Pasqualetti

The Dissertation of Vahid Mardanlou is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

Before starting my graduate school, I used to search the web for some recommendations to be more productive during doctorate program. It was interesting to me first I saw the phrase “PhD life’. Now, after these years I understand why they call it “life rather than just a “degree. I think it is a snapshot of life, full of ups and downs. A path to the goal where the main requirements are staying positive and never giving up on the dreams.

I am grateful to my advisor Dr. Elisa Franco, without whose help, I would not have been here. Many thanks to Dr. Ertem Tuncel, Dr. Fabio Pasqualetti and Dr. Tao Jiang for their great feedbacks and being in my thesis committee. I am also thankful of all my lab mates in Franco Group and especially Dr. Leopold Green and Dr. Hari Subramanian whom I had a close collaboration in my research. I learned a lot from our collaborators and especially Dr. Franco Blanchini, Dr. Rizal Hariadi, Dr. Jongmin Kim and Dr. Giulia Giordano.

I owe a lot to my family, my lovely mom and my wonderful sister. Thousands of miles away from home, their supports and encouragements make me motivated to work harder. I also thank my uncles (Hassan and Karim) and their wonderful families who helped me to settle down in US. During my PhD, I met a wonderful girl who brought joy and happiness to my life. Thank you Sara, for all your love and support.

And finally, I am really lucky that have awesome friends. They are not only a friend but like a family member to me. They stood beside me in any situations. Thanks to all of you as you make the life easier for me.

To my mom who taught me to follow my dreams

ABSTRACT OF THE DISSERTATION

Mathematical Modeling of Biological Networks

by

Vahid Mardanlou

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2017
Dr. Elisa Franco, Chairperson

Many biological scaffolds, such as the cytoskeleton, are built with filamentous polymers that are constantly assembling and disassembling in response to environmental inputs and cellular instructions. DNA nanotechnology has produced a variety of artificial, rationally designed tubular structures whose dimensions and mechanical properties are comparable to those of natural filaments such as actin and microtubules. DNA nanotubes self-assemble from tiles. We derive a coarse-grained model that captures the temporal evolution of DNA nanotube length distribution during growth experiments. This research is motivated by the rapid expansion of dynamic DNA nanotechnology, which offers exciting opportunities to use DNA strand displacement circuits to control tile self-assembly.

We also studied biomolecular circuits with the potential for bistable behavior. Bistable networks govern many important biological processes, including the cell cycle, molecular signal transduction and metabolism. We find sufficient conditions on key parameters that guarantee bistability in toggle switches. Having the goal of designing a circuit with desirable level of robustness and toggling speed, we consider the problem of parameter optimization in biomolecular bistable circuits.

We show that, under some assumptions often satisfied by bistable biological networks, it is possible to derive algebraic conditions on the parameters that determine when bistability occurs. These algebraic conditions can be included as nonlinear constraints in a parameter optimization problem.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Many biological scaffolds, such as the cytoskeleton, are built with filamentous polymers that are constantly assembling and disassembling in response to environmental inputs and cellular instructions. DNA nanotechnology has produced a variety of artificial, rationally designed tubular structures whose dimensions and mechanical properties are comparable to those of natural filaments such as actin and microtubules [44, 32, 37, 47, 13]. DNA nanotubes self-assemble from tiles that can be single-stranded or multi-stranded; inter-tile binding patterns are determined by programmable single stranded interaction domains. Here, we focus on DNA nanotubes assembling from multi-stranded tiles. We propose a continuous time, coarse-grained approach to describe how the nanotube length distribution varies over time in a population of nanotubes. This research is motivated by the rapid expansion of dynamic DNA nanotechnology, which offers exciting opportunities to use DNA strand displacement circuits to control tile self-assembly [63]. Existing methods to model tile or nanotube assembly are not suited to track length distributions, and cannot handle dynamically varying free tile concentration.

DNA tile assembly has been primarily modeled with algorithmic approaches, which explore the capacity of various tile interaction rules to produce desired patterns, and investigate their computational power [61, 43, 52]. The famous Winfree kTAM model models the kinetics of individual tiles binding and unbinding to a growing structure [61]; single molecule [16] and single-filament [27] measurements have been used to validate this model. Nanotube length distributions can be computed using kTAM stochastic simulations [46, 38]; nucleation and end-joining events can be taken into account, however this approach is only viable at low (nanomolar) tile concentrations. The most notable limitation of this model is the fact that it cannot handle time-varying concentration of free tiles.

Bistable networks govern many important biological processes, including the cell cycle [12], molecular signal transduction [2], and metabolism [45]. Although simplified models often describe their behavior in a satisfactory manner, *in vivo* bistable systems are remarkably complex due to the large number of species directly or indirectly participating in the network dynamics. In contrast, a controlled *in vitro* environment is ideal to build minimal, rationally designed mechanisms for bistability, which may be more scalable and tractable than natural systems [29, 42]. Bistable chemical networks with minimal size have been demonstrated [60, 57], however their dynamics are not immediately suited for a biomolecular implementation as they include single-reactant second and third order rates.

Here we consider an experimentally plausible implementation of a bistable network built solely with first and second order reaction rates. A simplified version of this network was analyzed in our previous work [34]. The dynamics are based *in vitro* transcription reactions [29, 54] and nucleic acid strand displacement [64]. Like other well known networks [29, 21], the circuit is characterized by

an overall positive feedback loop created by two branches of transcriptional repression. However, regulation here is not obtained by modulating the activity of the genes' promoter region, but by modulating the activity of the transcribing enzymes with RNA aptamer monomeric repressors [40, 39] (*i.e.* a single copy of the RNA aptamer is required to reduce activity of its target). Recovery of enzyme activity is achieved via strand displacement. Potentially, this circuit may function *in vivo*, in contrast with other recently proposed *in vitro* designs [29, 42, 54].

We show that our network has the appropriate structure to exhibit bistability [8] for arbitrary parameter choices. In particular, we show that if the circuit admits simple transitions to instability, they can only be driven by a real eigenvalue. We also derive sufficient conditions on the constant term of the characteristic polynomial to destabilize the equilibria. Our results are relevant because networks using exclusively monomeric repressors do not generally yield bistable behaviors [49, 31]. We conclude with numerical simulations which identify the region in parameter space where the system is bistable.

In vitro synthetic molecular circuits are attractive because they allow us to build new biological functions in a simplified context [29]. Circuits that rely on the rational design of nucleic acids are particularly powerful because a variety of reaction primitives are available [51]. Using RNA and DNA as regulatory molecules [9] we design two canonical dynamic networks in biology: a bistable switch (toggle switch) and an oscillator. We use synthetic genes whose RNA outputs create regulatory loops by modulating the activity of the enzymes producing them.

In this thesis, we design and analyze the model of a toggle switch where two synthetic genes produce two different RNA transcripts relying on different enzyme species. Each RNA species is designed to inhibit the transcribing enzyme of the other RNA species, overall creating a positive feedback

loop (Fig. 3.1). This type of regulation is experimentally achievable using recently published RNA sequences [40, 39], and our laboratory is actively pursuing the construction of this circuit.

One of the goals of synthetic biology is the design and construction of *de novo* biomolecular systems with behaviors that satisfy user specifications. An important class of functional behaviors is bistability, which is essential to build signaling cascades, developmental networks, and memory elements [3]. Once the general structure of the desired bistable system has been identified, it may be desirable to optimize concentrations and reaction rates of its components to satisfy given performance specifications, while preserving bistability. Because bistability is a complex, global dynamic behavior, the most straightforward option to optimize parameters is by trial and error. To our knowledge, no systematic approaches to this problem are available. (We remark that this challenge is distinct from the problem of parameter optimization in the context of data fitting [5], or network reconstruction [56].)

Through out the thesis, we address the problem of defining a systematic procedure to optimize parameters in biomolecular bistable systems: we want this procedure to identify parameters that maximize or minimize a given objective function chosen by the user, while preserving the bistable behavior of the network. For instance, it may be desirable to maximize or minimize the total concentration of some reagents, or to speed up or slow down certain reaction rates based on the specific components used. We formulate our optimization problem so that desired performance is expressed in the objective function, while bistability is guaranteed via the introduction of nonlinear constraints. Nonlinear constraints are obtained from algebraic conditions on the equilibria and on their stability.

Chapter 2

A coarse-grained model captures the temporal evolution of DNA nanotube length distributions

2.1 Introduction

Many biological scaffolds, such as the cytoskeleton, are built with filamentous polymers that are constantly assembling and disassembling in response to environmental inputs and cellular instructions. DNA nanotechnology has produced a variety of artificial, rationally designed tubular structures whose dimensions and mechanical properties are comparable to those of natural filaments such as actin and microtubules [44, 32, 37, 47, 13]. DNA nanotubes self-assemble from tiles that can be single-stranded or multi-stranded; inter-tile binding patterns are determined by programmable single stranded interaction domains. Here, we focus on DNA nanotubes assembling from multi-

stranded tiles. We propose a continuous time, coarse-grained approach to describe how the nanotube length distribution varies over time in a population of nanotubes. This research is motivated by the rapid expansion of dynamic DNA nanotechnology, which offers exciting opportunities to use DNA strand displacement circuits to control tile self-assembly [63]. Existing methods to model tile or nanotube assembly are not suited to track length distributions, and cannot handle dynamically varying free tile concentration.

DNA tile assembly has been primarily modeled with algorithmic approaches, which explore the capacity of various tile interaction rules to produce desired patterns, and investigate their computational power [61, 43, 52]. The famous Winfree kTAM model models the kinetics of individual tiles binding and unbinding to a growing structure [61]; single molecule [16] and single-filament [27] measurements have been used to validate this model. Nanotube length distributions can be computed using kTAM stochastic simulations [46, 38]; nucleation and end-joining events can be taken into account, however this approach is only viable at low (nanomolar) tile concentrations. The most notable limitation of this model is the fact that it cannot handle time-varying concentration of free tiles.

We formulate and validate a phenomenological model for self-assembling nanotubes that describes how the length varies over time in a population of nanotubes. This model is coarse-grained in the sense that the nanotube population is segmented by length in a number of bins, and we use ordinary differential equations to describe how the population of each bin varies over time, an approach akin to concentration flux models [23]. We model several processes that are known to affect nanotube length: nucleation, polymerization and depolymerization, end-joining, and fragmentation.

The model is fitted and validated using growth experiments of DNA nanotubes assembled from a DAE-E tile variant originally proposed by Rothemund et al. [44], where the total tile concentration is constant. However, our model can easily handle scenarios where the total tile concentration varies over time (for example, by activation or deactivation [63]), because it is naturally compatible with ordinary differential equation (ODE) chemical reaction networks. These scenarios will be considered in future work. We expect that this model will be useful to describe other self-assembling polymer systems operating at different time and length scales.

2.2 Results

2.2.1 Model derivation

We consider a solution including assembled nanotubes and unpolymerized tiles. The real distribution of nanotubes is continuous, because our sample includes tubes having any length $l \in [0, l_{max}]$, where l_{max} is the maximum observed length (or a physically meaningful upper bound for length). To build a model that is computationally tractable, we segment the population of molecular species present in the system. We assume the sample includes tiles, whose concentration is indicated as T; nucleated assemblies of tiles, or nuclei, whose concentration is L_0; nanotubes, which are binned by length, so that variable L_n indicates the concentration of nanotubes in bin n. The bin width, which we indicate as l_b, can be chosen depending on the acceptable level of coarseness (and complexity) of the model, because it determines the number of species. For example, if l_b is 300 nm, variable L_1 is the concentration of tubes of length 300 nm. If $l_{max} = 30\,\mu m$, the number of variables in the model is $n_{max} = [l_{max}/l_b] = 100$. Segmentation introduces implicitly the assumption that a tube can switch from bin n to bin $n \pm 1$ only if it acquires or loses a number n_b of tiles, which

are the tiles forming a tube segment of length l_b. As an example, let us take again $l_b = 300$ nm; let us assume that the nanotube circumference is on average 7 tiles, each ≈ 14 nm wide; then we find $n_b = 147$ (these figures are based on previous measurements on DAE-E tile nanotubes [44], and were confirmed in our experiments).

We finally assume that tiles, nuclei, and tubes form and interact via several processes, which cause changes in the segmented distribution of nanotube length. An overview of these processes is provided below, together with phenomenological expressions for the rates at which these processes occur. For each description, we identify an equivalent, phenomenological reaction that describes how tiles, nuclei, and nanotubes interact in our model.

Nucleation It is generally accepted that nucleation of DNA tiles is a cooperative process [38]: there is a minimum number of tiles that need to bind simultaneously to form a nucleus, from which polymerization of a nanotube can be initiated. We assume that nucleation depends on the concentration of tiles, and proceeds with rate $k_{nucl}T^{n_{nucl}}$, where n_{nucl} is the critical nucleation size. The equivalent phenomenological reaction describing nucleation is:

$$n_{nucl}T \xrightarrow{k_{nucl}} L_0. \tag{2.1}$$

Polymerization and depolymerization Nuclei and nanotubes grow as tiles bind to accessible sites. The polymerization rate depends on the concentration of tiles as well as the availability of binding sites: for tubes of length n, polymerization occurs at rate k_pTL_n. For nuclei, which are smaller patches of tiles, we hypothesize a different polymerization rate $k_{p_0}TL_0$. Tiles can also dissociate from tubes (and nuclei) at a rate that depends exclusively on the concentration of tubes: for tubes of length n, the depolymerization rate is k_dL_n; for nuclei, we consider a different de-

polymerization rate $k_{d_0}L_0$. Equivalent phenomenological reactions describing polymerization and depolymerization are:

$$n_bT + L_n \xrightarrow{k_p} L_{n+1}, \quad n \geq 1 \tag{2.2}$$

$$(n_b - n_{nucl})T + L_0 \xrightarrow{k_{p_0}} L_1 \tag{2.3}$$

$$L_1 \xrightarrow{k_d} L_0 + (n_b - n_{nucl})T \tag{2.4}$$

$$L_{n+1} \xrightarrow{k_d} L_n + n_bT \text{ , for } n \geq 1 \tag{2.5}$$

$$L_0 \xrightarrow{k_{d_0}} n_{nucl}T \tag{2.6}$$

Here, the stoichiometric coefficients indicate how many tiles need to be added or removed from a nanotube in a certain bin length so that it moves to an adjacent bin. These coefficients are however not related to the order of the reaction rate; for example, in reaction (2.2) it is not required that n_b tiles bind simultaneously to the tube, therefore this is a second order reaction. We also note that these polymerization and depolymerization rates are meant to capture addition or dissociation of tiles at both ends of the growing segment (we lump the factor 2 in the corresponding reaction rates).

End-joining Nucleated nanotubes diffusing in solution grow not only by polymerization, but also by end-joining, as demonstrated experimentally in [14]. We assume that the joining rate depends on the length of the nanotubes, on their diameter d, and on the concentration of nanotubes in the corresponding length bins. For example, if we consider length bins n and m, we postulate that the joining rate of nanotubes in these bins is $k_{join}(n,m)L_nL_m$. An estimate for $k_{join}(n,m)$ is given in expression (2.7), which is derived in detail in reference [26]. This expression assumes that DNA

nanotubes are rigid rods, and that their end-joining is a diffusion controlled reaction:

$$k_{join}(m,n) = \frac{\alpha}{l_b}\left[\frac{1}{m}\ln\left(\frac{ml_b}{d}\right) + \frac{1}{n}\ln\left(\frac{nl_b}{d}\right)\right] \tag{2.7}$$

where $\alpha = (\kappa 12 k_B \mathscr{T} d)/\eta$. Here η is the dynamic viscosity of the liquid, k_B is the Boltzmann constant, $\mathscr{T}$ is the absolute temperature, d is the nanotube diameter, and κ is a factor accounting for the fraction of productive nanotube collisions. Note that each joining reaction can occur by joining of either end of each nanotube, so every reaction should be accounted for twice.

The concentration of nanotubes in a given bin n increases when shorter nanotubes end-join; an example equivalent reaction is:

$$L_{n-m} + L_m \xrightarrow{k_{join}(n-m,m)} L_n,$$

$$\text{for } \ 1 \le m \le \min\{n_{max} - n, n-1\}.$$

We observe that order of the reactants in the above reaction does not matter. For example, consider the bin of nanotubes having length $5l_b$. The end-joining reactions that contribute to an increase in the concentration L_5 are:

$$L_1 + L_4 \xrightarrow{k_{join}(1,4)} L_5$$

$$L_2 + L_3 \xrightarrow{k_{join}(2,3)} L_5$$

Reactions $L_4 + L_1 \xrightarrow{k_{join}(1,4)} L_5$ and $L_3 + L_2 \xrightarrow{k_{join}(2,3)} L_5$ are redundant and should not be included in the mass balance. The concentration of tubes in bin of length nl_b is therefore incremented by only $\lfloor \frac{n}{2} \rfloor$ end-joining reactions, where "$\lfloor\ \rfloor$" is the largest integer less than or equal to $\frac{n}{2}$.

The concentration of nanotubes in bin n also decreases due to end-joining events, as exemplified in this reaction:

$$L_n + L_m \xrightarrow{k_{join}(m,n)} L_{n+m} \ \text{ for } \ 1 \le m \le n_{max} - n$$

Fragmentation Formed nanotubes can spontaneously break into shorter fragments. We assume that breakage rate of nanotubes of length $n\,l_b$ depends on the concentration of nanotubes in bin length n via a constant rate k_{break}. For simplicity, we also assume fragments can only form to fall into the given length bins, and that a nanotube can break into at most two fragments per unit time. For example, a nanotube of length $n\,l_b$ can break into two fragments of length ml_b and $(n-m)l_b$. The equivalent phenomenological reaction for fragmentation is:

$$L_n \xrightarrow{k_{break}} L_m + L_{n-m}.$$

The rate at which the concentration of nanotubes in bin n decreases due to fragmentation is $(n-1)k_{break}L_n$, because there are $n-1$ sites at which breakage may occur. For example, consider the concentration of tubes having length $5l_b$. In this bin, fragmentation occurs according to the following equivalent reactions:

$$L_5 \xrightarrow{k_{break}} L_1 + L_4, \quad L_5 \xrightarrow{k_{break}} L_2 + L_3,$$
$$L_5 \xrightarrow{k_{break}} L_3 + L_2, \quad L_5 \xrightarrow{k_{break}} L_4 + L_1.$$

We note that the concentration of nanotubes in bin n also increases due to breakage of longer nanotubes (bins $m \geq n+1$).

Tile activation and deactivation Tiles may be injected or removed from the system, for example by activation or deactivation reactions that could be performed via strand displacement. Expression to model these phenomena depend on the mechanism chosen for activation and deactivation; at this stage we assume they are some continuous functions $a(t)$ (activation) and $d(t)$ (deactivation).

We now combine the phenomena described above and their rates to write a set of differential equations describing the evolution of our segmented nanotube length distribution. Equations (2.8)-(2.11) describe the time derivative of the concentration of tiles (T), nuclei (L_0) and tubes in bin n (L_n).

$$\frac{dT}{dt} = a(t) - d(t) - \underbrace{n_{nucl}k_{nucl}T^{n_{nucl}}}_{\text{nucleation}} - \underbrace{n_b k_p T \sum_{i=1}^{n_{max}-1} L_i \quad - (n_b - n_{nucl})k_{p_0}TL_0}_{\text{polymerization}} \tag{2.8}$$

$$\underbrace{+ n_b k_d \sum_{i=2}^{n_{max}} L_i + (n_b - n_{nucl})k_d L_1 + n_{nucl}k_{d_0}L_0,}_{\text{depolymerization}}$$

$$\frac{dL_0}{dt} = \underbrace{k_{nucl}T^{n_{nucl}}}_{\text{nucleation}} - \underbrace{k_{p_0}TL_0}_{\text{polymerization}} + \underbrace{k_d L_1 - k_{d_0}L_0,}_{\text{depolymerization}} \tag{2.9}$$

$$\frac{dL_1}{dt} = \underbrace{k_{p_0}TL_0 - k_p TL_1}_{\text{polymerization}} + \underbrace{k_d L_2 - k_d L_1}_{\text{depolymerization}} \underbrace{- 2\sum_{m=1}^{n_{max}-1} k_{join}(1,m)L_1 L_m - 2k_{join}(1,1)L_1^2}_{\text{end-joining}} \tag{2.10}$$

$$\underbrace{+ 2\sum_{m=2}^{n_{max}} k_{break}(L_m \to L_1 + L_{m-1})L_m,}_{\text{fragmentation}}$$

$$\frac{dL_n}{dt} = \underbrace{k_p T(L_{n-1} - L_{n|_{n<n_{max}}})}_{\text{polymerization}} + \underbrace{k_d(L_{n+1|_{n<n_{max}}} - L_n)}_{\text{depolymerization}} \tag{2.11}$$

$$\underbrace{- 2\sum_{m=1}^{n_{max}-n} k_{join}(n,m)L_n L_m - 2k_{join}(n,n)L_{n|_{n\leq \frac{n_{max}}{2}}}^2 + 2\sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} k_{join}(n-m,m)L_{n-m}L_m}_{\text{end-joining}}$$

$$\underbrace{+ 2\sum_{m\geq n+1}^{n_{max}} k_{break}(L_m \to L_n + L_{m-n})L_m - (n-1)k_{break}(L_n \to L_m + L_{n-m})L_n,}_{\text{fragmentation}} \qquad \text{for} \quad n = 2, \ldots, n_{max}.$$

We clarify the presence of end-joining term:

$$-2k_{join}(n,n)L_n^2$$

in equation (2.11). This term is needed for two reasons: first, end-joining can occur at either end of a nanotube; second, when two nanotubes of length n end-join, two of them are lost from bin n. Thus, overall equation (2.11) should include a term $-4k_{join}(n,n)L_n^2$: we chose to split it between sum $-2\sum_{m=1}^{n_{max}-n} k_{join}(n,m)L_nL_m$ and the isolated term $-2k_{join}(n,n)L_n^2$.

In this model, the mass transfer (species concentration conversion) is carefully balanced. If $a(t) = d(t) = 0$, *i.e.* there is no net injection or removal of tiles, the total concentration of tiles T^{tot} remains constant. Equation (2.12) below describes tile mass conservation:

$$T^{tot} = T + n_{nucl}L_0 + n_b \sum_{n=1}^{n_{max}} nL_n, \tag{2.12}$$

when $a(t) = d(t) = 0$. Our model naturally guarantees tile mass conservation, because differential equations (2.8)-(2.11) satisfy the following equality:

$$\frac{dT}{dt} + n_{nucl}\frac{dL_0}{dt} + n_b \sum_{n=1}^{n_{max}} n\frac{dL_n}{dt} = 0,$$

where we take $a(t) = d(t) = 0$.

2.2.2 Modeling nanotube growth

We test the capacity of model (2.8)–(2.11) to fit nanotube length distributions in a sample of growing nanotubes where there is no injection or removal of tiles, *i.e.* $a(t) = d(t) = 0$.

Experiments

We consider a two-tile variant of DAE-E nanotubes described in [44, 14] (Fig. 2.1 A). Each tile assembles from five distinct DNA strands, and is about 14.3 nm wide. The two types

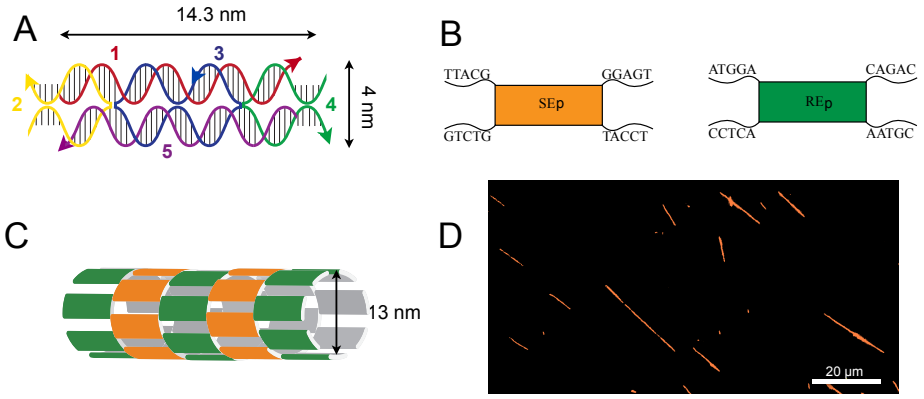

Figure 2.1: **We used a two-tile system assembling into DNA nanotubes [44] to train and validate our model.** A: Schematic of the DAE-E tile used in our experiments. B: Tile abstractions: two distinct types of tiles, SEp and REp, interact via their complementary sticky end domains. Strand 3 (blue) of the SEp tile is labeled with a Cy3 fluorophore. C: Tiles interact to form tubes via their complementary sticky ends [44] forming stripes perpendicular to the nanotube long axis; D: Example fluorescence microscopy image of nanotubes after 10 hour incubation.

of tiles (REp and SEp) differ in their sticky end sequences, and are designed so that each type of tile cannot self-assemble in isolation (Fig. 2.1 B); however, when mixed together, they assemble forming a "striped" pattern perpendicular to the long axis of the lattice (Fig. 2.1 C). These tiles form nanotubes that generally have a 7-tile circumference, and a diameter $d = 13 \times 10^{-3}(\mu m)$. Strand 3 of tile SEp is labeled with Cy3, so that nanotube length can be measured in fluorescence microscopy experiments [44, 14]. Sequences for both types of tiles and experimental protocols are reported in the Methods section. The two types of tiles are annealed and stored separately. Nanotube nucleation and growth is triggered only when the two tile types are mixed, thereby allowing the experimenter to control when the assembly process starts. We note that this is not possible in single tile-type nanotube variants that were considered in earlier growth studies [14, 33]: in these cases, to monitor tile assembly it was necessary to rapidly quench the anneal process upon reaching the melting temperature of sticky ends (estimated to be around 45-47°C). The rapid temperature quench may, however, increase assembly error rates in individual tiles and lattices, and variability in experimental

results; these issues can be eliminated by using a two-tile system.

Samples were incubated at room temperature, and imaged using a fluorescence microscope during the following 30 hours to assess changes in the nanotube length distribution over time. This experiment was run in triplicate samples; Fig. 2.2 D shows mean length and standard deviation of the mean measured across experiments. We observe rapid growth in the first 10 hours, when the mean nanotube length reaches 4–7 μm and plateaus in the following 20 hours. The measured growth curve in our experiments is qualitatively similar to previous experiments by Ekani-Nkodo et al. [14].

Data fitting

We adapted and fitted Model (2.8)-(2.11) to the nanotube distributions that were experimentally measured over time when the concentration of each tile type (SEp and REp) is 750 nM. The model equations were slightly adapted to take into account the fact that two distinct tile types are present. First, we made the simplifying assumption that both tile populations evolve with the same kinetic rates; then, we assumed that half the tiles in a given nanotube segment are of type SEp, and half are of type REp. These assumptions require that equation (2.8) be modified as follows (assuming there is no net tile generation $a(t)$ or depletion $d(t)$):

$$\frac{dT}{dt} = \underbrace{-\frac{n_{nucl}}{2}k_{nucl}T^{n_{nucl}}}_{\text{nucleation}} \tag{2.13}$$

$$\underbrace{-\frac{n_b}{2}k_pT\sum_{i=1}^{n_{max}-1}L_i \quad -\frac{(n_b-n_{nucl})}{2}k_{p_0}TL_0}_{\text{polymerization}} \tag{2.14}$$

$$\underbrace{+\frac{n_b}{2}k_d\sum_{i=2}^{n_{max}}L_i+\frac{(n_b-n_{nucl})}{2}k_dL_1+\frac{n_{nucl}}{2}k_{d_0}L_0}_{\text{depolymerization}}.$$

This equation now models the evolution over time of the concentration of each tile type, either SEp or REp. We note that the nucleation term results from the product: $\frac{n_{nucl}}{2} k_{nucl} T_{SEp}^{\frac{n_{nucl}}{2}} T_{REp}^{\frac{n_{nucl}}{2}}$, where for simplicity we indicated both tile species with the same symbol. Equation (2.12), describing tile mass conservation, must also be updated as follows:

$$T^{tot} = T + \frac{n_{nucl}}{2} L_0 + \frac{n_b}{2} \sum_{n=1}^{n_{max}} n L_n. \tag{2.15}$$

We do not need to modify the remaining equations of the model: because SEp and REp tiles assemble forming "rings", it is reasonable to assume that polymerization at a growing nanotube edge promotes formation of bonds with exclusively one tile type at a time. As for the dynamics of nuclei L_0, captured by equation (2.9), the nucleation term can be again interpreted as the product of two terms, as explained earlier.

Then, we established parameters of our model that depend on the type of nanotubes considered, and on the desired level of coarse graining. We choose the length bin: $l_b = 300$ nm; the maximum nanotube length is set as $l_{max} = 30\ \mu m$. These choices imply that we have a number of bins $n_{max} = 100$; thus the model includes 102 differential equations (including the ODEs for tiles and nuclei). The number of tiles in a tube chunk of length l_b is $n_b = 147$. (This follows from our assumption that the tubes have a 7-tile circumference, with ≈ 14 nm tile width). ODEs were integrated with in-house MATLAB scripts. We chose an integration step of 10 seconds.

Predicted nanotube distributions at time t are generated deterministically by integrating ODEs (2.13), (2.9)-(2.11) (up to time t), and using the updated mass conservation equation (2.15). Rather than comparing length histograms generated by the model to experimental histograms, we compare their cumulative distributions. This choice is motivated by the following observations: 1) Cumulative distributions are by definition scaled by the sample size (in our case, nanotube sample

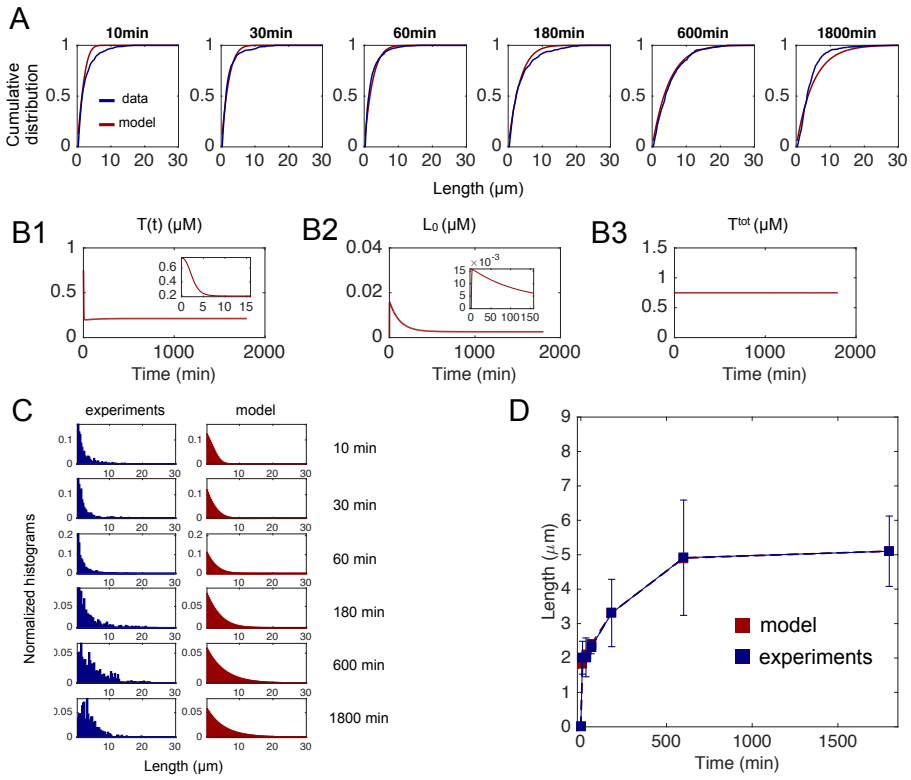

Figure 2.2: **Performance of our coarse-grained model when fitted to distributions of growing nanotubes at** 750 **nM tile concentration.** A: Comparison between a representative experimental (blue) cumulative length distributions gathered during nanotube growth and the cumulative distribution generated by our model (red). B1: Time course of free tile concentration; inset shows a more detailed view of the transient. B2: Nuclei concentration. B3: Total tile concentration remains constant. C: Histograms of representative growth experiment (blue) and histograms generated by our model (red). D: Mean and standard deviation of the mean lengths in our nanotube growth experiments, compared with the mean length predicted by our fitted model.

number). This implies that cumulative distributions of different samples can be immediately compared without requiring *ad hoc* normalization (which would be necessary to compare histograms). 2) Cumulative distributions do not require binning of data like a histogram, thus the fitting procedure is not biased by the choice of bin width.

To compare experimental and simulated cumulative distributions at the same points (nanotube length), we interpolated experimental cumulative length distributions. After interpolation, the model and experimental distributions at any given time are described by comparable vectors:

$$V_{sim}(t) = \frac{1}{\mathscr{L}_{sim}} \cdot \begin{bmatrix} L_{1,sim}(t) \\ L_{1,sim}(t) + L_{2,sim}(t) \\ \vdots \\ \sum_1^{n_{max}} L_{i,sim}(t) \end{bmatrix},$$

$$V_{exp}(t) = \frac{1}{\mathscr{L}_{exp}} \cdot \begin{bmatrix} L_{1,exp}(t) \\ L_{1,exp}(t) + L_{2,exp}(t) \\ \vdots \\ \sum_1^{n_{max}} L_{i,exp}(t) \end{bmatrix},$$

$$\mathscr{L}_{sim} = \sum_{i=1}^{n_{max}} L_{i,sim}(t), \quad \mathscr{L}_{exp} = \sum_{i=1}^{n_{max}} L_{i,exp}(t).$$

These distributions are to be compared at times t =10, 30, 60, 180, 600, 1800 minutes; we assume there are no nanotubes at time t=0 (*i.e.* the time at which the two types of tiles are mixed).

Fitting is done to identify several parameters in our model. Specifically, we fit the nucleation rate (k_{nucl}) and critical nucleation size (n_{nucl}), the polymerization and depolymerization rates for tubes and nuclei (k_p, k_{p_0}, k_d, and k_{d_0}), the breakage rate (k_{break}), and parameter α in the joining rate expression (2.7). These parameters can be stacked in a vector p, and we set up our fitting prob-

lem as the minimization of the objective function, which simultaneously compares the simulated distribution to the distributions of three separate experiments:

$$\min_{p} J = \sum_{j=1}^{3} \sum_{t} (V_{sim}(t) - V_{exp}^{j}(t))^{\top} (V_{sim}(t) - V_{exp}^{j}(t)). \tag{2.16}$$

Minimization was done using the `fmincon` routine in MATLAB. Initial conditions for the parameters were sampled uniformly in physically plausible intervals delimited by lower and upper bounds listed in Table 2.1; parameters were also constrained to fall within these bounds. We also imposed a lower bound of 10 nM on the admissible free tile concentration. Initial conditions for the model variables were chosen as $T(0) = 750$ nM (each tile type), $L_i(0) = 0$ μM for i = 0, ..., n_{max}, *i.e.* we assumed that initially only free tiles are present; this assumption is consistent with the fact that no nanotubes can form prior to mixing the two tile types.

Several fitting campaigns were launched and evaluated. Minimizing the objective function (2.16) is a non-convex problem, thus the fitting routine is likely to converge to local minima that depend on the randomly chosen initial conditions for the parameter vector. We report one of the best fitting results in Table 2.1.

Fig. 2.2 A shows the performance of the model in fitting cumulative distributions in one of our experiments. The concentration of free tiles and nuclei as a function of time are shown in Fig. 2.2 B1 and B2; during the time course, the total concentration of tiles remains constant as shown in Fig. 2.2 B3 (this is a sanity check that mass conservation equation (2.15) holds).

We also tested the capacity of our model to generate histograms that reflect the measurements; experimental data were binned consistently to our model bin width $l_b = 300$ nm; results are shown in Fig. 2.2 C. We also computed the mean nanotube length from simulated distributions, and we compared it with the measured mean length in Fig. 2.2 D (the standard deviation is computed

over triplicate experiments).

Table 2.1: **Fitted parameters.** This table lists the parameters of Model (2.13), (2.9)-(2.11) that were fitted to measured length distributions of DNA nanotubes where each tile concentration is at 750 nM. The table includes lower bounds (L.B.) and upper bounds (U.B.) used in the fitting procedure.

Parameter	Units	L.B.	U.B.	Fitted value	Definition
k_p	$M^{-1}s^{-1}$	10^1	10^7	1.319×10^5	Polymerization rate
k_d	s^{-1}	10^{-8}	10^4	2.959×10^{-2}	Depolymerization rate
α	$\mu m\, M^{-1}s^{-1}$	10^{-4}	10^{11}	25.9	End-joining parameter
k_{break}	s^{-1}	10^{-10}	10^2	2.242×10^{-9}	Fragmentation rate
k_{nucl}	$M^{1-n_{nucl}}s^{-1}$	10^{-10}	10^4	8.272×10^2	Nucleation rate
k_{p_0}	$M^{-1}s^{-1}$	10^1	10^7	1.338×10^3	Nuclei polymerization rate
k_{d_0}	s^{-1}	10^{-8}	10^1	1.369×10^{-4}	Nuclei depolymerization rate
n_{nucl}	N/A	2	5	5	Critical nucleation size

Validation

We tested the capacity of our fitted model to predict the distribution of nanotube length when the tile concentration differs from the fitting conditions. We integrated the differential equations using parameters listed in Table 2.1, but total tile (each type) concentration at 500 nM or 1000 nM. The simulations were compared to experimental data as shown in Figs. 2.3 and 2.4. Predictions generally match well the measured cumulative distributions and histograms as they evolve over time, as well as the mean length.

As an additional sanity check, we estimated the nanotube density from the model pre-

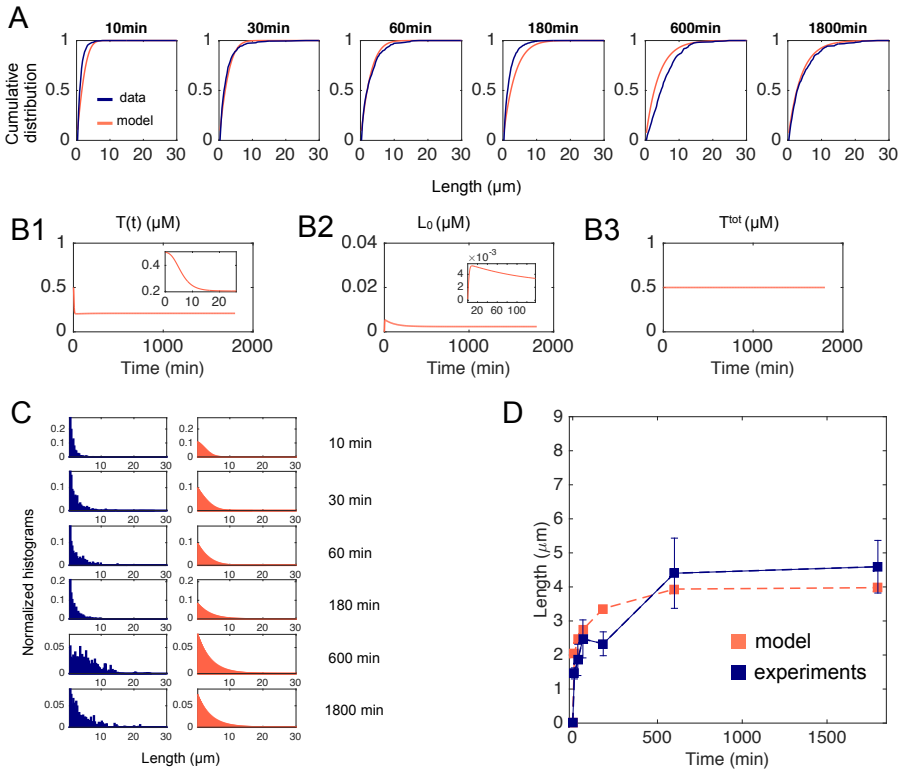

Figure 2.3: **We validate our model by predicting the evolution of the length distribution of nanotubes grown at** 500 **nM (each) tile concentration.** A: Representative experimental cumulative length distributions gathered during nanotube growth (blue), and predicted cumulative distribution generated by our model (orange). B1: Time course of free tile concentration; inset shows a more detailed view of the transient. B2: Nuclei concentration. B3: Total tile concentration remains constant. C: Histograms of representative growth experiment (blue) and histograms generated by our model (orange). D: Mean and standard deviation of the mean lengths in our nanotube growth experiments, compared with the mean length predicted by the model.

dictions; information on nanotube density was not included in the fitting procedure. We computed the expected sample volume of each experimental sample using the average measured thickness of an imaging sample (2.35 μm) and a standard imaging area of 100 μm $\times$ 100 μm. The predicted concentration was then converted to nanotube number in the sample volume, by taking into account dilution of the sample for imaging. We further made the simplifying assumption that all nanotubes land on either of the sample glass surface: because we only image one surface, we assumed that

only half the predicted number of nanotubes are measured. We compare our nanotube density prediction to the measured nanotube density (averaged over triplicate experiments) in Fig. 2.5. Initially the number of nanotube bursts (these are on average very short nanotubes, as can be deduced from Figs. 2.2, 2.3, and 2.4), and decreases over time due to end joining events. The model predicts a similar trend, and the correct order of magnitude for nanotube density. The simulated number is generally larger than the measured number; we hypothesize that this discrepancy is due in part to the fact that some nanotubes may not land on the imaging surface, and that overlapping and branching nanotubes are discarded by our image processing algorithm.

2.3 Methods

DNA sequences All DNA was purchased from Integrated DNA Technologies (Coralville, IA). Sequences were taken from Rothemund et al. [44] (our strand nomenclature slightly differs from Rothemund's notation). Following the notation introduced in Fig. 2.1, sequences are reported below.

SEp1: 5′-CTCAGTGGACAGCCGTTCTGGAGCGTTGGACGAAACT,

SEp2: 5′-GTCTGGTAGAGCACCACTGAGGCATT,

SEp3: 5′-Cy3/CCAGAACGGCTGTGGCTAAACAGTAACCGAAGC
ACCAACGCT,

SEp4: 5′-TGAGGAGTTTCGTGGTCATCGTACCT,

SEp5: 5′-CGATGACCTGCTTCGGTTACTGTTTAGCCTGCTCTAC,

REp1: 5′-CGTATTGGACATTTCCGTAGACCGACTGGACATCTTC,

REp2: 5′-CCTCACCTTCACACCAATACGAGGTA,

REp3: 5′−TCTACGGAAATGTGGCAGAATCAATCATAAGACACCA

GTCGG,

REp4: 5′−CAGACGAAGATGTGGTAGTGGAATGC,

REp5: 5′−CCACTACCTGTCTTATGATTGATTCTGCCTGTGAAGG.

Sample preparation Lyophilized DNA oligonucleotides were resuspended in water, quantitated by UV absorbance at 260 nm using Thermo Scientific Nanodrop 2000c Spectrophotometer, and stored at -20°C. Each tile solution was separately prepared to a target 2 μM tile concentration by adding 2 μM (final concentration) of the corresponding DNA strands 1, 2, 3, 4, and 5, 1x Tris-Acetate-EDTA (TAE) buffer and 12.5 mM $MgCl_2$. After vortexing for 60 seconds, each tile solution was placed in Eppendorf Mastercycler PCR machine and annealed by heating solution to 90°C, and cooled to room temperature over a 6 hour period. Annealed tile solutions were mixed to the desired target concentration using 1x TAE, 12.5 mM $MgCl_2$.

Image acquisition and processing Samples were imaged on an inverted microscope (Nikon Eclipse TI-E) with 60X/1.40 NA oil immersion objectives. Aliquots of nanotube solution (diluted to 50 nM tile concentration) were deposited on Fischerbrand microscope cover glass 12-545E No. 1 (thickness=0.13 to 0.17 mm; Size: 50 x 22 mm). VWR Micro Slides (Plain, Selected, Pre-cleaned, 25 x 75 mm, 1.0 mm thick) were placed gently on the cover glass. Nanotubes were imaged using a Cy3 filter cube (Semrock Brightline - Cy3-404C-NTE-ZERO). Upon image capturing, exposure time was set to 70 ms. ImageJ plugin AnalizeSkeleton was used to obtain nanotube contours, from which we estimated tube length using a MATLAB script. All branched nanotubes are automatically eliminated from our sample. Due to camera limitations, tubes below 0.3 μm in length were discarded from tube length distributions.

2.4 Conclusion

We have derived a deterministic, coarse-grained model to describe the growth kinetics of DNA nanotube populations. The model takes into account nucleation, polymerization and depolymerization, joining and fragmentation processes, for which we propose candidate reactions satisfying detailed balance. We obtained macroscopic reaction rates for each of these processes by fitting our model to experimental cumulative distributions of nanotube length at different time points. The model is trained on a set of experiments monitoring nanotube growth at a given tile concentration, and satisfactorily validated on two additional growth experiments conducted at different tile concentrations. We briefly discuss our fitting results in relation to the literature. Fitted rate constants are listed in Table 2.1; our fitted polymerization rate of tiles to growing tubes is on the same order of magnitude of polymerization rates estimated for individual tiles binding to a growing nanotube [16, 27]; this estimate is two orders of magnitude higher than the estimate we obtained in earlier preliminary fitting for single tile-type DNA nanotubes [33]. The fitted nanotube depolymerization rate is about two orders of magnitude slower than the estimates in [27] for DAO-O nanotubes measured at 33°C; this discrepancy may be due to the fact that our experiments were run at room temperature. Our critical nucleation size $n = 5$ is double the estimate provided by Zhang et al. [63], where single tile-type nanotubes were considered; our fitted nucleation rate $n_{nucl}\, k_{nucl} \approx 4.14 \times 10^3$ $\mathrm{M}^{1-n_{nucl}}$/s is two orders of magnitude slower than previous estimates in [63]; it is worth noting that tiles considered by Zhang et al. were annealed and gel extracted prior to triggering assembly, thereby minimizing the amount of misfolded tiles which could contribute to faster polymerization.

Polymerization and depolymerization rates on growing nuclei are slower than the corresponding rates of growing tubes; while this expectation is sensible for polymerization (due to a

reduced number of binding sites), depolymerization rates should be consistent with those of formed nanotubes. In our model we assume that the end-to-end joining rate depends on the nanotube length, according to (2.7) [26], and on the fitted parameter α. The rate k_{join} peaks when $n = m = 1$ and is a decreasing function of nanotube length, as shown in Fig. 2.6. We evaluated the expression for k_{join} over its domain n, m, and computed an average $k_{join} = 34.6$ /M/s, four orders of magnitude lower than earlier joining rate estimates for single-tile nanotubes [33], and three orders of magnitude slower than the joining rate estimate for DNA ribbons in [46], which is 3.5×10^4 /M/s. Finally, our fragmentation rate is on the order of 10^{-9}/M/s, suggesting that breakage events are extremely rare. Bayesian analysis will help gain further information on the most plausible range of fitted parameters [24].

Clearly, our fitted rates depend on the chosen level of granularity: further work is required to identify tradeoffs between model granularity and its predictive capacity. Additionally, the predictive capacity of the model should be tested when nanotubes are grown in a range of temperatures and ionic conditions, which we leave for future work. Further numerical sensitivity analysis will be useful to clarify the influence of individual processes (nucleation, polymerization/depolymerization, fragmentation and joining) on the kinetics and steady states of the length distributions, in particular considering different concentration regimes and alternative reaction mechanisms [36].

The complexity of DNA nanotube nucleation and growth justifies the formulation of a model based on length distribution segmentation. Unlike most polymers or copolymers that form linear chains where each link of the chain is a single chemical compound [19], DNA nanotubes are hollow cylinders where several monomers participate in each growth layer. This feature makes DNA nanotubes similar to microtubules [18], and makes kinetic modeling of assembly significantly

more complex.

Our modeling approach relies on ordinary differential equations that are typically used in polymerization process control models [19]: however, to track length distributions while maintaining our model numerically tractable, we define chemical species that represent nanotube populations in a certain length bin. Similar approaches have been widely used to model protein filament assembly (such as amyloid-like fibrils) and are known as concentration flux equations, where ODEs describe how the concentration of filaments of a given length vary over time as a function of nucleation, monomer addition, breakage, and end-to-end joining [23]; the complexity of these models can vary depending on the process being modeled. In some cases analytical solutions can be found for limiting cases or simplified filament models by using linearization and moment methods [30]. Conformational dynamics of DNA filaments are outside the scope of this research [1]. Future work will focus on modeling nanotube length distributions where tiles are activated or deactivated over time [63].

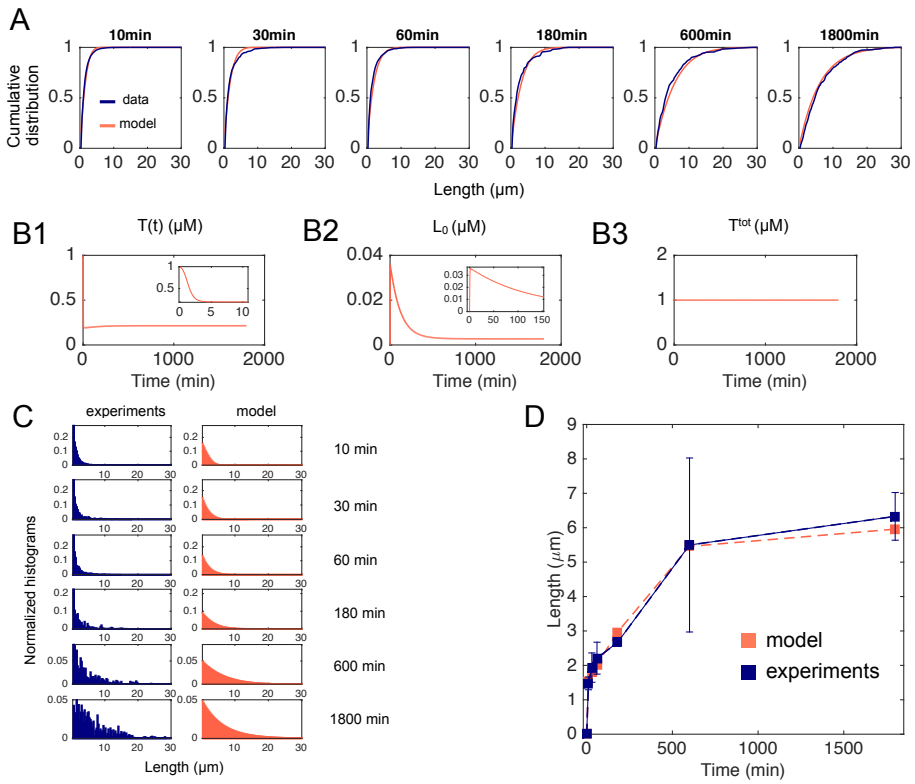

Figure 2.4: **Validation of our model is done by predicting the evolution of the length distribution of nanotubes grown at** 1000 **nM (each) tile concentration.** A: Example cumulative length distributions gathered during a representative nanotube growth experiment (blue), compared to the predicted cumulative distribution generated by our model (orange). B1: Time course of free tile concentration; inset shows a more detailed view of the transient. B2: Nuclei concentration. B3: Total tile concentration remains constant. C: Histograms of representative growth experiment (blue) and histograms generated by our model (orange). D: Mean and standard deviation of the mean lengths measured in our nanotube growth experiments, compared with the mean length predicted by the model.

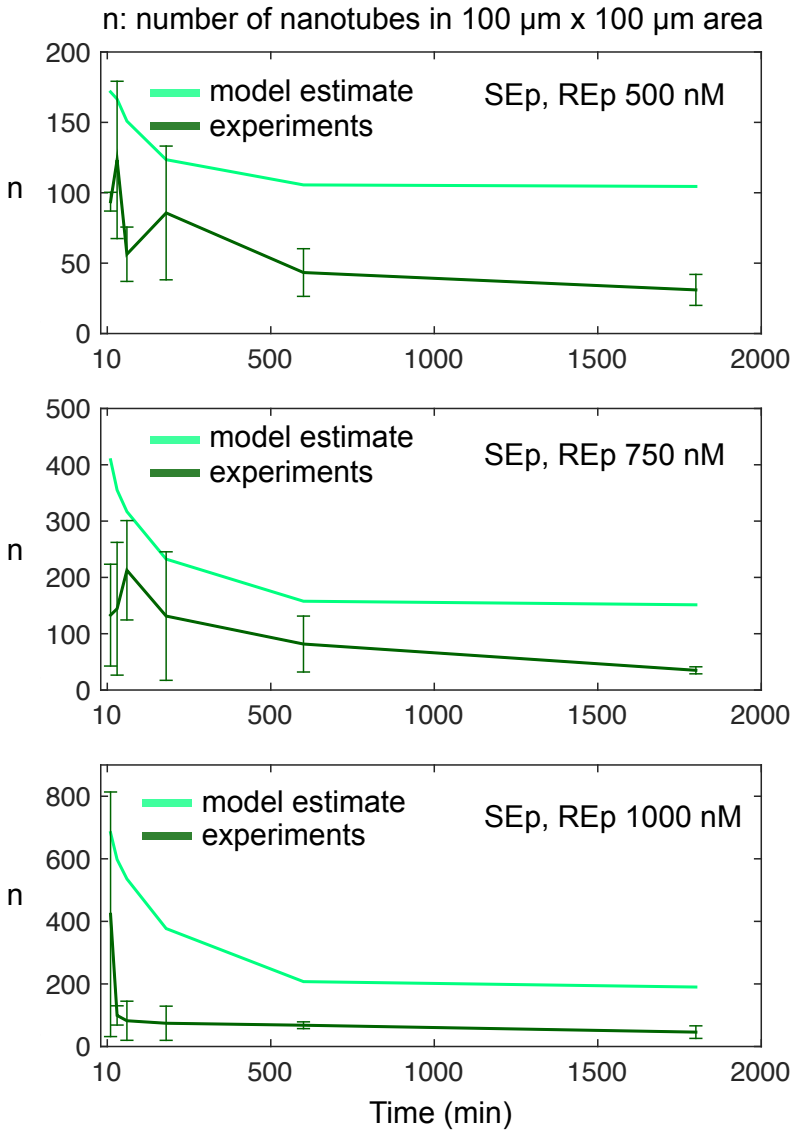

Figure 2.5: **Predicted imaging density of nanotubes using our coarse-grained model.** The predicted density follows the same trend as the imaged nanotube density.

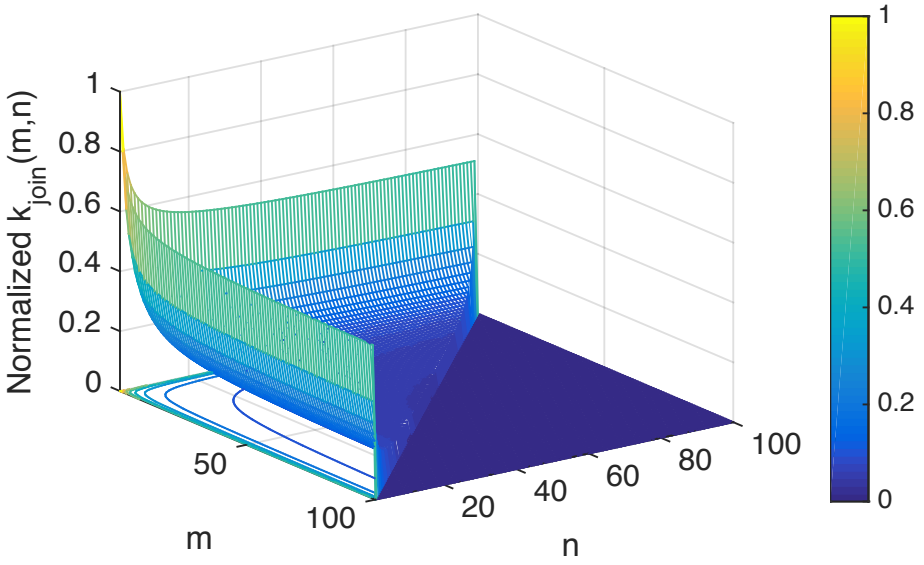

Figure 2.6: **Normalized plot of the joining rate** (2.7) [26].

Chapter 3

Design of a molecular bistable system with RNA-mediated regulation

3.1 Introduction

In vitro synthetic molecular circuits are attractive because they allow us to build new biological functions in a simplified context [29]. Circuits that rely on the rational design of nucleic acids are particularly powerful because a variety of reaction primitives are available [51]. Using RNA and DNA as regulatory molecules, in this and a companion paper [9] we design two canonical dynamic networks in biology: a bistable switch (toggle switch) and an oscillator. We use synthetic genes whose RNA outputs create regulatory loops by modulating the activity of the enzymes producing them.

In this chapter, we design and analyze the model of a toggle switch where two synthetic genes produce two different RNA transcripts relying on different enzyme species. Each RNA species

is designed to inhibit the transcribing enzyme of the other RNA species, overall creating a positive feedback loop (Fig. 3.1). This type of regulation is experimentally achievable using recently published RNA sequences [40, 39], and our laboratory is actively pursuing the construction of this circuit.

Here we consider a simple model capturing the essential interactions among molecular species. This model is not exhaustive, but is useful to provide intuition on what are the key reaction rates to achieve bistability. We focus on sufficient conditions for bistability that relate the RNA transcription rates and the enzyme activation rates in the system. We strive to derive analytical conditions that make a limited number of assumptions on the parameters [6, 7, 8]. Transcription rates are often not controllable in the system because they depend on the enzyme properties, which are difficult to tune. Mechanisms to achieve enzyme reactivation are currently under study in our laboratory, thus it is important for us to understand what is their role in the system, relative to parameters that are not under our control such as transcription rates.

In the past decade several bistable molecular circuits have been designed and built [21, 4], with recent *in vitro* implementations [29, 42, 54]. All these circuits are based on the creation of an overall positive feedback loop. Our circuit is based on transcription reactions as in [29, 54]. However regulation is not achieved by modulating the activity of the genes' promoter region, but by modulating the activity of the transcribing enzymes. Because it relies on reactions naturally occurring in the cellular environment, this circuit may function *in vivo*, in contrast with other recently proposed *in vitro* designs [29, 42, 54].

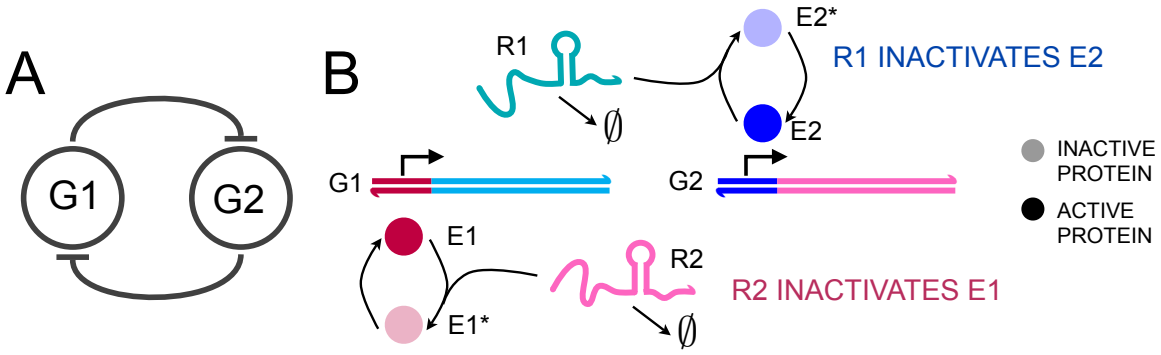

Figure 3.1: A: Structure of our two node transcriptional toggle switch; G_1 and G_2 are synthetic genes. B: Proposed experimental implementation of our RNA-based switch, where the transcription enzymes are inhibited by RNA aptamers.

3.2 Circuit description and modeling

Biomolecular bistable switches are generally built based on the principle that a positive feedback loop is required; the presence of a positive feedback loop is a well known necessary condition set forth by Thomas [55, 50]. We propose to design a positive loop using *in vitro* transcription reactions, designing RNA inputs to regulate the ability of synthetic genes to generate RNA outputs. We will then discuss sufficient conditions on the reaction parameters to achieve bistability. Regulation of transcription is achieved by modulating the activity of RNA polymerases (RNAP), rather than the activity of promoters. Using specific RNA aptamers (short, noncoding RNA sequences that bind to a desired target [59, 11]) we can inhibit the activity of two well characterized and commercially available bacteriophage RNAP species [40, 39].

Fig. 3.1 shows our circuit structure: two synthetic genes produce RNA outputs R_1 and R_2. The RNA outputs are transcribed by two different enzymes E_1 and E_2; for example, one could use bacteriophage SP6 and T7 RNAPs, and the corresponding promoter sequences (purple and blue domains on the genes in Fig. 3.1). We assume that enzymes can either be in an active or inactive state, and their

mass is conserved: $[E_i^{tot}] = [E_i] + [E_i^*]$, $i = 1,2$. Each enzyme E_i is inhibited by RNA species R_j (through a direct aptamer binding pathway), and we assume it reverts to its active state at a given rate which could be driven by a strand displacement reaction [64] (*i.e.* another RNA or DNA specie binds to R_2 and releases the enzyme).

The designed reactions are:

$$E_i + g_i \xrightarrow{k_i} E_i + R_i + g_i \tag{3.1}$$

$$R_i \xrightarrow{\delta_i} 0 \tag{3.2}$$

$$E_i + R_j \xrightarrow{\gamma_i} E_i^* \tag{3.3}$$

$$E_i^* \xrightarrow{\beta_i} E_i \tag{3.4}$$

for i = 1,2 and j = 2,1, where E_i are active enzymes, E_i^* are inactive enzymes, R_i are RNA species, g_i are genes (constant). From now on, for simplicity we assume that both subcircuits have the same reaction rates, i.e. $k_1 = k_2 = k$, $\delta_1 = \delta_2 = \delta$ etc. (This simplifying assumption may be violated in experimental implementations of the circuit.) Using mass action laws, we can derive the ODEs describing the dynamics of the system:

$$[\dot{R}_i] = k[E_i][g_i] - \delta[R_i] - \gamma[E_j][R_i] \tag{3.5}$$

$$[\dot{E}_i] = \beta([E_i^{tot}] - [E_i]) - \gamma[E_i][R_j] \tag{3.6}$$

We now switch to a more compact notation, and define $x_1 = [R_1]$, $x_2 = [E_1]$, $x_3 = [R_2]$, and $x_4 = [E_2]$. Because the concentration of genes is constant, we define $\kappa_1 = kg_1$ and $\kappa_2 = kg_2$. The system of

ODEs modeling our system can be rewritten as:

$$\begin{cases} \dot{x_1} = \kappa_1 x_2 - \delta x_1 - \gamma x_4 x_1 \\ \dot{x_2} = \beta x_2^{tot} - \beta x_2 - \gamma x_2 x_3 \\ \dot{x_3} = \kappa_2 x_4 - \delta x_3 - \gamma x_2 x_3 \\ \dot{x_4} = \beta x_4^{tot} - \beta x_4 - \gamma x_4 x_1 \end{cases} \tag{3.7}$$

In the next sections, we investigate analytically the potential for bistability of model (3.7), supporting our findings with simple numerical simulations. Under some simplifying assumptions on the parameters, we will find conditions on RNA production rates and enzyme recovery rates that can guarantee bistability.

3.3 Stability analysis

We begin by deriving equilibrium conditions. We set each ODE in model (3.7) equal to zero for $i = 1, ..., 4$, and derive the equations below which describe the equilibrium relationship between the enzyme concentrations x_2 and x_4:

$$\beta\gamma x_4^2 + (\gamma\kappa_1 x_2 - \beta\gamma x_4^{tot} + \beta\delta)x_4 - \beta\delta x_4^{tot} = 0, \tag{3.8}$$

$$\beta\gamma x_2^2 + (\gamma\kappa_2 x_4 - \beta\gamma x_2^{tot} + \beta\delta)x_2 - \beta\delta x_2^{tot} = 0. \tag{3.9}$$

It is easy to verify that each of these second order equations has a unique positive solution; thus from the first equation we can find an expression of x_4 as a monotonic strictly decreasing function of x_2, and from the second equation we can find x_2 as a monotonic strictly decreasing function of x_4. Expressions are reported in Appendix A. We numerically plot these equilibrium conditions in Fig. 3.2; for illustrative purposes, we choose the following parameters: $\beta = 4 \cdot 10^{-3}$ /s, $\gamma = 5 \cdot 10^5$

/M/s, $\delta = 3 \cdot 10^{-3}$ /s, $\kappa_1 = \kappa_2 = 1 \cdot 10^{-2}$ /s, $x_2^{tot} = x_4^{tot} = 100$ nM. These parameters are close to realistic values for *in vitro* circuits [20], however our assumed inhibition rate γ may be relatively high for aptamer systems. The equilibrium values $\bar{x}_2$ and $\bar{x}_4$ satisfying the equations above can be used to find the equilibria of x_1 and x_3, using again our model (3.7).

$$\bar{x}_1 = \beta \frac{x_4^{tot} - \bar{x}_4}{\gamma \bar{x}_4}, \quad \bar{x}_3 = \beta \frac{x_2^{tot} - \bar{x}_2}{\gamma \bar{x}_2}. \tag{3.10}$$

To investigate the local stable or unstable nature of each equilibrium, we look at the linearized

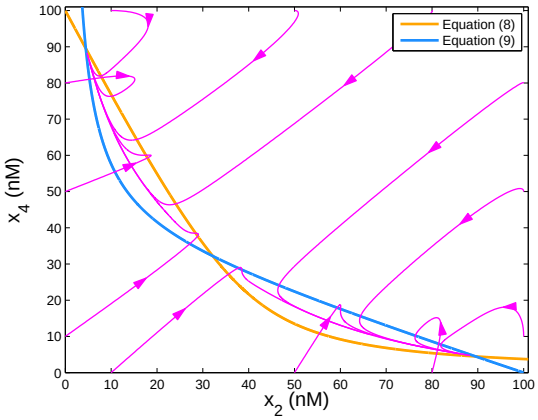

Figure 3.2: Equilibrium conditions (3.8) (orange) and (3.9) (cyan), and overlapped selection of trajectories (magenta).

dynamics. The Jacobian is:

$$J = \begin{bmatrix} -\gamma\bar{x}_4 - \delta & \kappa_1 & 0 & -\gamma\bar{x}_1 \\ 0 & -\beta - \gamma\bar{x}_3 & -\gamma\bar{x}_2 & 0 \\ 0 & -\gamma\bar{x}_3 & -\gamma\bar{x}_2 - \delta & \kappa_2 \\ -\gamma\bar{x}_4 & 0 & 0 & -\beta - \gamma\bar{x}_1 \end{bmatrix} \tag{3.11}$$

By examining the Jacobian and its characteristic polynomial, we find that instability of the linearized system can only occur through transition of the real dominant eigenvalue to the positive half plane.

Lemma 1 *Consider system* (3.7)*, whose characteristic polynomial has all positive coefficients except the constant term (which could be positive or negative). Any equilibrium of system* (3.7) *is unstable if and only if the constant term of the characteristic polynomial is negative. Instability can only be driven by a simple, real (positive) eigenvalue.*

Proof The Jacobian is similar to a Metzler matrix through a transformation $T = \text{diag}(1,1,-1,-1)$:

$$T^{-1}JT = \begin{bmatrix} -\gamma\bar{x}_4 - \delta & \kappa_1 & 0 & \gamma\bar{x}_1 \\ 0 & -\beta - \gamma\bar{x}_3 & \gamma\bar{x}_2 & 0 \\ 0 & \gamma\bar{x}_3 & -\gamma\bar{x}_2 - \delta & \kappa_2 \\ \gamma\bar{x}_4 & 0 & 0 & -\beta - \gamma\bar{x}_1 \end{bmatrix}$$

Therefore, the Jacobian always has a real dominant eigenvalue, i.e. $\lambda_{max} > \mathscr{R}e(\lambda_i), \forall \lambda_i \in J$ [53]. The characteristic polynomial of the Jacobian (3.11) is computed in Appendix B, expression (A.4). All its coefficients are real and are positive with the exception of the constant term, which can be positive or negative.

Assume the constant term of the characteristic polynomial $p_\lambda(s)$ is negative. Then, we know that $p_\lambda(0) < 0$ and it is real. Take $s \to \infty$, $p_\lambda(s) > 0$, because all other coefficients are positive. Thus, there must be at least one point in the right half plane that is a root of p_λ. Thus, the system is unstable. Because the matrix is Metzler, the largest root must be real.

If the system is unstable, then the characteristic polynomial must have at least one root with positive real part. *Ab absurdo*, suppose the constant term is positive. Then instability can only occur with a pair of complex conjugate eigenvalues with positive real part. This is impossible because the Jacobian is a Metzler matrix and the dominant eigenvalue must be real. Then, the constant term of the characteristic polynomial must be negative. ■

3.4 Sufficient conditions for bistability

We now seek any conditions on the transcription and recovery reaction rates that can guarantee bistability of the system. We begin by making two additional simplifying assumptions:

1) The total amount of enzyme is identical for the two subcircuits: $E_1^{tot} = E_2^{tot} = E^{tot}$, *i.e.* $x_2^{tot} = x_4^{tot} = E^{tot}$.

2) The rates of transcription of the RNA species are identical (see reaction (3.1)): $\kappa_1 = \kappa_2 = \kappa$.

Bistability requires the presence of two stable and one unstable equilibrium. From our equilibrium conditions (3.8) and (3.9), we establish that the system presents three equilibria when $\kappa > 2\beta$. Calculations are shown in Appendix.

We then derive some additional conditions on the equilibria. First, by subtracting the equations (3.8) and (3.9) (equilibrium conditions), we get:

$$\beta\gamma(x_4^2 - x_2^2) + (\beta\delta - \beta\gamma E^{tot})(x_4 - x_2) = 0. \tag{3.12}$$

This equation is satisfied when either of these two equations holds:

$$x_4 = x_2, \tag{3.13}$$

$$x_4 = -x_2 + \left(E^{tot} - \frac{\delta}{\gamma}\right). \tag{3.14}$$

Note that $x_4 + x_2$ must be a positive quantity, thus $\left(E^{tot} - \frac{\delta}{\gamma}\right)$ must be positive for our problem to be physically meaningful. We plot equations (3.13) and (3.14) overlapped to the equilibrium conditions (3.8) and (3.9) in Fig. 3.3. If we pick $\kappa > 2\beta$, we find three intersection points A, B, and C which define the equilibria of the system. Based on Lemma 1, we know that stability of any equilibrium depends on the constant term a_0 of the characteristic polynomial (A.5):

$$a_0 = \left(\beta\gamma\bar{x}_4 + \beta\delta + \delta\gamma\bar{x}_1\right)\left(\beta\gamma\bar{x}_2 + \beta\delta + \delta\gamma\bar{x}_3\right) - \gamma^2\kappa^2\bar{x}_2\bar{x}_4. \tag{3.15}$$

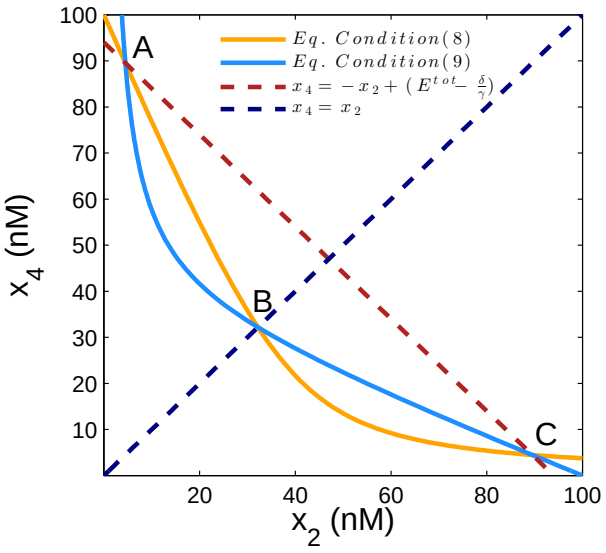

Figure 3.3: Equilibrium conditions (3.13), (3.14) and their intersections with equilibrium conditions (3.8) and (3.9).

If $a_0 > 0$, then the equilibrium is stable; if $a_0 < 0$, then the equilibrium is unstable. We further manipulate expression (3.15), after multiplying it by the positive factor $\bar{x}_2\bar{x}_4$. The modified constant term is now $P_0 = a_0\bar{x}_2\bar{x}_4$.

$$P_0 = \left(\beta\gamma\bar{x}_4^2 + \beta\delta\bar{x}_4 + \delta\gamma\bar{x}_1\bar{x}_4\right)\left(\beta\gamma\bar{x}_2^2 + \beta\delta\bar{x}_2 + \delta\gamma\bar{x}_3\bar{x}_2\right) - \gamma^2\kappa^2\bar{x}_2^2\bar{x}_4^2 \tag{3.16}$$

We finally substitute the equilibrium expressions for $\bar{x}_1$ and $\bar{x}_3$ found in equation (3.10), thus equation (3.16) becomes:

$$P_0 = \left(\beta\gamma\bar{x}_4^2 + \beta\delta E^{tot}\right)\left(\beta\gamma\bar{x}_2^2 + \beta\delta E^{tot}\right) - \gamma^2\kappa^2\bar{x}_2^2\bar{x}_4^2 \tag{3.17}$$

Unstable Equilibrium(Point B) For bistability, we require that point B is unstable. This occurs when $P_0 < 0$. Because (3.8) and (3.9) are both monotonically decreasing (Proof in Appendix A), point B is the only intersection of equation (3.13) and the equilibrium conditions (3.8) and (3.9).

We substitute $\bar{x}_2 = \bar{x}_4 = z$ in equation (3.18), and we find a fourth order polynomial in $\bar{x}_4$.

$$P_0^B(z) = \left(\beta\gamma z^2 + \beta\delta E^{tot}\right)^2 - \gamma^2\kappa^2\bar{z}^4. \tag{3.18}$$

Keeping in mind that $z > 0$, we can rewrite $P_0^B(z)$ as $P_0^B(t) = -at^2 + bt + c$, where $t = z^2$ and $a, b, c > 0$. Due to the negative sign in front of a, this polynomial is negative for t larger than its biggest root. Thus, $P_0^B(t) < 0$ when:

$$t = z^2 = \bar{x}_4 > \frac{\beta\delta E^{tot}}{\gamma}\frac{1}{\kappa - \beta}. \tag{3.19}$$

We also know, from (3.14), that

$$\bar{x_4} < \left(E^{tot} - \frac{\delta}{\gamma}\right). \tag{3.20}$$

Combining these bounds (3.19) and (3.20), we find the following sufficient condition that guarantees $P_0^B < 0$ and thus instability of point B:

$$\kappa > \beta(1+\Theta), \qquad \Theta = \frac{\gamma\delta E^{tot}}{(\gamma E^{tot} - \delta)^2}. \tag{3.21}$$

Stable Equilibriua (Points A and C) Stable equilibria result from the intersection of equilibrium conditions (3.8) and (3.9) with equation (3.14). As done for the unstable node, we use these conditions on the equilibria within the expression for term P_0, in order to find the relationship between the transcription and recovery parameters that guarantee stability of points A and C.

We begin by summing equations (3.8) and (3.9). It is convenient to rewrite their sum as:

$$\begin{aligned}\beta\gamma(x_2+x_4)^2 + (\beta\delta - \beta\gamma E^{tot})(x_2+x_4) \\ + 2\gamma(\kappa-\beta)x_2x_4 - 2\beta\delta E^{tot} = 0\end{aligned} \tag{3.22}$$

Now, we can substitute term $x_2 + x_4$ with the equivalent parametric value that can be found from equation (3.14). Thus, we find a parametric expression for the product x_2x_4:

$$x_2x_4 = \frac{\beta\delta E^{tot}}{\gamma(\kappa-\beta)} \tag{3.23}$$

We then rewrite (3.17) as:

$$P_0 = (\beta\delta E^{tot})(\beta\gamma)(x_2^2 + x_4^2) + \gamma^2(\beta^2 - \kappa^2)x_2^2x_4^2 + (\beta\delta E^{tot})^2, \tag{3.24}$$

and as done before we substitute the parametric expressions for $x_2 + x_4$ and x_2x_4 found above, and we impose that the resulting new expression for P_0 be positive. We find:

$$(\beta\gamma)(\beta\delta E^{tot})\left(E^{tot} - \frac{\delta}{\gamma}\right)^2 - 4\beta\frac{(\beta\delta E^{tot})^2}{\kappa - \beta} > 0 \tag{3.25}$$

We find the following sufficient condition relating κ and β,

$$\kappa > \beta(1 + \Psi), \quad \Psi = 4\frac{\gamma\delta E^{tot}}{(\gamma E^{tot} - \delta)^2} = 4\Theta, \tag{3.26}$$

where Θ was defined at condition (3.21). By imposing condition (3.26) we guarantee that the system is bistable.

3.5 Numerical Analysis

We ran numerical simulations using MATLAB, checking the bistability region for the system in a range of values for β and κ. First, we compute equilibria using our equilibrium conditions described in the previous sections, discarding negative solutions which are not physically acceptable. For each equilibrium point, we find the eigenvalues of the corresponding Jacobian. A combination of parameters is classified as yielding bistability if we find two stable and one unstable equilibrium between them. Our results are shown in Fig. 3.4, and they are compared with the analytical bounds for κ and β derived at the previous sections. While we coarsely vary κ and β between 0.001 and 0.9, all other parameters are held constant as: $\gamma = 5 \cdot 10^5$ /M/s, $\delta = 3 \cdot 10^{-3}$ /s, $x_2^{tot} = x_4^{tot} = 100$ nM (these are the same parameters used to plot Figs. 3.2 and 3.3).

Referring to Fig. 3.4, light green dots correspond to the bistability region. Dark blue dots correspond to a region where the system is not bistable. The orange line plots condition (3.26) for our specific choice of parameters, while the red line marks the $\kappa > 2\beta$ conservative condition that guarantees the presence of three equilibria. We remark that three equilibria may occur also for $\beta < \kappa < 2\beta$, as highlighted by our numerical results; unfortunately, we were not able to analytically find a condition less conservative than $\kappa > 2\beta$.

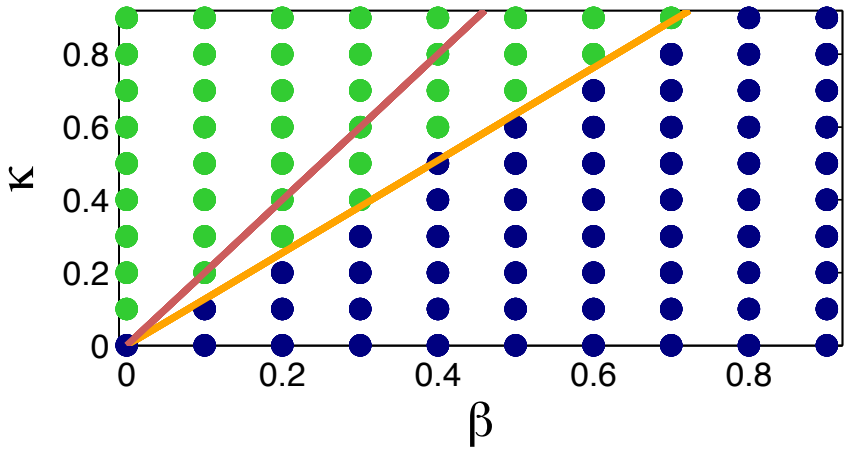

Figure 3.4: Simulation results: we numerically computed the bistability region (green) of model (3.7) as a function of β and κ; the orange line corresponds to condition (3.26); the red line corresponds to the conservative condition $\kappa > 2\beta$ which guarantees three equilibria.

3.6 Conclusions

We have described a design idea for a new molecular bistable system, which relies on the use of RNA aptamers for regulation. A simple model for this switch is thoroughly analyzed and its potential for bistability is demonstrated. In particular, we derive sufficient conditions for bistability that focus on RNA transcription and enzyme reactivation rates, which we consider key experimental parameters. Our laboratory is actively pursuing the construction of this circuit. We expect that a

more exhaustive model will be required for the purpose of data fitting; however analysis of this simple model provided significant intuition on the potential for bistability of this circuit.

Chapter 4

An algebraic approach to parameter optimization in biomolecular bistable systems

4.1 Introduction

One of the goals of synthetic biology is the design and construction of *de novo* biomolecular systems with behaviors that satisfy user specifications. An important class of functional behaviors is bistability, which is essential to build signaling cascades, developmental networks, and memory elements [3]. Once the general structure of the desired bistable system has been identified, it may be desirable to optimize concentrations and reaction rates of its components to satisfy given performance specifications, while preserving bistability. Because bistability is a complex, global dynamic behavior, the most straightforward option to optimize parameters is by trial and error. To

our knowledge, no systematic approaches to this problem are available. (We remark that this challenge is distinct from the problem of parameter optimization in the context of data fitting [5], or network reconstruction [56].)

In this paper we address the problem of defining a systematic procedure to optimize parameters in biomolecular bistable systems: we want this procedure to identify parameters that maximize or minimize a given objective function chosen by the user, while preserving the bistable behavior of the network. For instance, it may be desirable to maximize or minimize the total concentration of some reagents, or to speed up or slow down certain reaction rates based on the specific components used. We formulate our optimization problem so that desired performance is expressed in the objective function, while bistability is guaranteed via the introduction of nonlinear constraints. Nonlinear constraints are obtained from algebraic conditions on the equilibria and on their stability.

Algebraic approaches to optimization have been previously proposed in the context of biochemical networks. In [62], the metabolic flux of a biochemical network is optimized defining an objective function where a single flux, multiplication of fluxes or concentrations of reactants are either maximized or minimized. The network is described using an S-system differential equation model. Algebraic equilibrium conditions are derived and used as constraints within the optimization problem, so that fluxes are optimized while the network is guaranteed to stay in the neighborhood of the desired steady state. In [10], the stationary behavior of a metabolic network is optimized by imposing stability of the equilibrium. The decision variables in this problem can be fluxes, reaction rates or the concentration of some of the reactants. Algebraic equilibrium and stability conditions (derived from a Routh-Hurwitz table) are included in the problem as part of the optimization constraints. The solution of the problem is a stable equilibrium which is optimal with respect to a

chosen objective function.

The main limitation of existing approaches is that they consider optimization of *local* (near equilibrium) behaviors. Here, we show that optimization of a network ensuring a *global* behavior like bistability can still be done using an algebraic approach when the system satisfies some structural conditions. Specifically we focus on a class of bistable systems for which the bistability property can be related solely to the number of equilibria of the system and to the sign of the coefficients of the characteristic polynomial. Equilibrium polynomials can be analyzed using Sturm's theorem [48] or a Routh table [35], obtaining a set of inequalities that guarantee a desired number of positive roots. Bistability conditions (global) can thus be expressed as nonlinear constraints to be enforced in the optimization procedure.

This chapter is organized as follows. Several notions from algebraic geometry and degree theory are reviewed in Section 4.2; at the end of this section we formulate our optimization problem. Section 4.3 describes two case studies, for which we derive algebraic conditions for bistability. Results for parameter optimization using our algebraic approach on the described case studies are presented in Section 4.4. We summarize and briefly discuss our results in Section 4.5.

4.2 Background and problem formulation

We review definitions and theorems, and formulate assumptions we need to establish algebraic conditions for bistability in biomolecular networks. These conditions are exploited as constraints in the parameter optimization problem described later.

4.2.1 Definitions and assumptions

We consider nonlinear biological dynamical systems that are successfully modeled with ODEs, which may be derived from chemical reactions by applying the law of mass-action or from phenomenological observations. These systems can be modeled as:

$$\dot{x} = f(x),\ \ x(0) = x_0,\ x \in \mathbb{R}^n_{\geq 0} \tag{4.1}$$

In biomolecular systems, the concentration of the reactants and products are the state variables of the system. Because states are positive concentrations, the state space is the non-negative orthant of $\mathbb{R}$.

Equilibria (steady states) are states $\bar{x}$ of system (4.1) satisfying $\dot{\bar{x}} = f(\bar{x}) = 0$. For a system of order n, equilibrium conditions are n equations in n unknowns [41]. If we can merge all equilibrium conditions into a single equation in the form of polynomial function w.r.t. either of the state variables, we refer to it as *master equilibrium condition*.

Definition 1 (Regular equilibrium) *An equilibrium point $\bar{x}$ of system* (4.1) *is regular if $det(J(\bar{x})) \neq 0$ (in other words, $J(\bar{x})$ must be invertible; alternatively, $J(\bar{x})$ must not have eigenvalues at the origin).*

For every equilibrium point $\bar{x}$, $J(\bar{x})$ is the Jacobian of system (4.1) evaluated at the equilibrium. Polynomial $P_{\bar{x}}(\lambda) = det(\lambda I - J(\bar{x})) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ is the characteristic polynomial of the Jacobian evaluated at $\bar{x}$. In many practical cases, the characteristic polynomials of bistable networks have a unique coefficient sign pattern that we summarize in the definition below [48].

Definition 2 (F-polynomial) *A polynomial whose coefficients are all non-negative (non-positive)*

except the constant term a_0, which can be either positive or negative, is an F-polynomial. Because the constant term is $a_0 = det(-J(\bar{x}))$, its sign depends on the determinant of $J(\bar{x})$.

This paper focuses on nonlinear systems (4.1) that are *bistable*.

Definition 3 (Bistability) *System* (4.1) *is bistable if it presents two asymptotically stable and one unstable equilibrium points.*

We consider biomolecular networks whose linearization is a positive system. This property will be used to derive algebraic conditions for bistability.

Definition 4 (Positive system) *A linear system $\dot{x} = Ax$ is positive if for every nonnegative initial state the solution $x(t)$ is nonnegative.*

Theorem 1 *[17] A linear system $\dot{x} = Ax$ is positive if and only if matrix A is a Metzler matrix, i.e., its elements satisfy: $a_{ij} \geq 0, \forall (i, j)$ such that $i \neq j$.*

Finally, we consider systems that are dissipative. In general, *dissipativity* (see, e.g., [25]) is defined based on the existence of compact, forward invariant subsets of $\mathbb{R}^n_+$ that absorb the system trajectories. The following definition ([28]) is equivalent and easier to verify:

Definition 5 (Dissipative system) *A system $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ is dissipative if its solutions are eventually uniformly bounded, i.e., there exists a constant $k > 0$ such that:*

$$\lim_{t \to +\infty} \sup x_j(t) \leq k, \quad j = 1, ..., n.$$

We conclude this section with a list of assumptions we make in the rest of this manuscript.

Assumption 1 *We assume the linearization of system* (4.1) *is a positive system at each equilibrium point.*

Assumption 2 *We assume the characteristic polynomial of system* (4.1) *is an F-polynomial at each equilibrium point.*

Assumption 3 *We assume that system* (4.1) *is dissipative.*

4.2.2 Bistability conditions

Algebraic equilibrium conditions. The number of equilibria of system (4.1) can be in many cases established by finding the roots of a master (polynomial) equilibrium condition [48]. The number of positive roots of a polynomial can be established using methods such as Routh table [35] or Sturm's theorem [48] on the master equilibrium condition of the system. These methods can provide parametric conditions for the presence of a desired number of positive roots, *i.e.* of equilibria.

Routh table

The Routh table is typically built to check stability of linear feedback systems. The number of sign changes in the first column of the table provides the number of roots of a polynomial that have positive real part. We can apply this method on the master equilibrium condition of system (4.1) to find parametric conditions to obtain a desired number of roots with positive real part (equilibrium points of the system). Complex positive roots can often be excluded by imposing additional conditions derived from knowledge of the system [35].

Sturm's Theorem

The Sturm sequence associated to a polynomial $p(x)$ is a set of polynomials $\mathscr{P} = \{p_0, p_1, \ldots, p_m\}$ defined as:

$$
\begin{aligned}
p_0 &= p(x), \\
p_1 &= dp(x)/dx, \\
p_2 &= -\text{rem}(p_0, p_1), \\
p_3 &= -\text{rem}(p_1, p_2), \\
&\vdots \\
p_m &= -\text{rem}(p_{m-2}, p_{m-1}), \\
0 &= -\text{rem}(p_{m-1}, p_m),
\end{aligned}
$$

where the symbol $rem(p_{i-2}, p_{i-1})$ indicates the remainder of the polynomial long division of p_{i-2} by p_{i-1} for $i \neq 0, 1$. It is assumed that p_0 and p_1 are univariate polynomials. The sequence ends when p_{m-1} divided by p_m gives a remainder of zero. For a polynomial $p(x)$ of degree n, the sequence has $m \leq n+1$ Sturm polynomials.

Theorem 2 (Sturm's theorem) *Let $p(x)$ be a real-valued univariate polynomial and $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$, with $a < b$ and $p(a), p(b) \neq 0$. Then the number of zeros of $p(x)$ in the interval (a, b) is the difference*

$$var(\mathscr{P}, a) - var(\mathscr{P}, b)$$

where $\mathscr{P}$ is the Sturm sequence of $p(x)$ and the variations $var(\mathscr{P}, a)$ and $var(\mathscr{P}, b)$ are the number of times that consecutive nonzero elements of the Sturm sequence — evaluated at a and b, respectively—have opposite signs.

Sturm's theorem can thus be used to find conditions for the master equilibrium polynomial to present a desired number of non-negative, real roots by considering the interval $[0,\infty)$.

Algebraic equilibrium conditions are also bistability conditions in a class of bistable networks.

A bistable system is characterized by the presence of three equilibria; in the previous paragraphs we listed two methods to obtain algebraic parametric conditions to ensure the presence of three equilibria. Of course, the correct number of equilibria does not in general guarantee bistability.

However, biomolecular bistable networks often satisfy Assumptions 1 – 3, *i.e.* their linearization is a positive system, their characteristic polynomial is an F-polynomial at each equilibrium point, and they are dissipative; several examples can be found in [48], and include two variants of a two-node negative feedback genetic toggle switch, and a positive feedback circuit based on the bacteriophage λ promoter P_{RM}. For these systems, bistability is guaranteed when three equilibria are present. For the reader's convenience, we provide additional background on this topic.

Theorem 3 *Consider a linear positive system $\dot{x} = Ax$, whose characteristic polynomial is an F-polynomial. This system has an unstable equilibrium if and only if the constant term of the characteristic polynomial is negative [35].*

Definition 6 (Index of an equilibrium point) *The index of a regular equilibrium point $\bar{x}$ is the sign of the determinant of $-J(\bar{x})$:*

$$ind(\bar{x}) = sign\,[det(-J(\bar{x}))]$$

Note that the above definition can be rewritten as $sign\,(det(\lambda I - J(\bar{x}_j)))\,|_{\lambda=0}$, which highlights that the index is the sign of the constant term of the characteristic polynomial at the equilibrium considered.

Definition 7 (Degree of a system) *The degree of a dynamical system $\dot{x} = f(x)$, over a set $U \in \mathbb{R}^n$, having equilibria $\bar{x}_i$, $i = 1,...,m$, is defined as:*

$$\deg(f) = \sum_{i=1}^{m} \{\text{ind}(\bar{x}_i), \bar{x}_i \in U, f(\bar{x}_i) = 0\} ,$$

where $\bar{x}_i$ are regular equilibria.

The following theorem is the last result we need to quote:

Theorem 4 *[28] A dissipative dynamical system $\dot{x} = f(x)$ defined on $\mathbb{R}^n$ has degree $+1$ with respect to any bounded open set containing all its equilibria [48].*

Theorem 5 *Consider a dynamical system* (4.1)*, satisfying Assumptions 1–3. If the system has three equilibria, then it is bistable.*

Proof. Due to Theorem 3, since the index is equal to the sign of the constant term in the characteristic polynomial, a positive index is associated with a stable equilibrium and a negative index is associated with an unstable equilibrium. Due to Theorem 4, we can conclude that when three regular equilibria are present, our systems are bistable. ■ We conclude remarking that as a consequence of Theorem 5, any algebraic equilibrium conditions that enforce the presence of three equilibria become global algebraic bistability conditions.

4.2.3 Optimization problem

We illustrate the general formulation of our optimization problem. We consider dynamical systems $\dot{x} = f(x, p_d)$ where p_d are model parameters (for example, reaction rates), and $p_d \in \Omega_d$. The state space of all admissible decision variables of the optimization problem is defined as Ω, such that $\Omega_d \subseteq \Omega$. We define a vector $p_{aux} \in \Omega_{aux}$ of auxiliary decision variables, which are combinations of

model parameters ($p_{aux}(p_d)$) or additional variables of the objective function which may be relevant to specify the system's behavior. In general, we define decision variables as $p = \{p_d, p_{aux}\} \in \Omega$, where $\Omega = \Omega_d \cup \Omega_{aux}$.

Our concern is to optimize an objective function $\mathscr{F}(p)$ which could be a function of a single or a combination of decision variables. We might need to optimize a single (or multi) objective function(s) regarding the goal(s) of the problem.

$$\begin{aligned} \min_{p\in\Omega}/\max_{p\in\Omega} \quad & \mathscr{F}(p) \\ \text{subject to} \quad & \mathbb{I}_i(p) < 0,\ i = 1,\ldots,q. \\ & \mathbb{E}_j(p) = 0,\ j = 1,\ldots,m. \\ & l_k \leq p_{d,k} \leq u_k,\ k = 1,\ldots,N. \\ & l_{k'} \leq p_{aux,k'} \leq u_{k'},\ k' = 1,\ldots,M. \end{aligned} \tag{4.2}$$

The constraints in the problem above include bistability algebraic conditions ($\mathbb{I}_i$ are inequality constraints and $\mathbb{E}_j$ are equality constraints), upper and lower bounds for parameters (l_k and u_k) and for decision variables ($l_{k'}, u_{k'}$). Because of Assumptions 1–3, our constraints are independent from the value of the equilibria.

4.3 Case studies

In this section we introduce two examples of bistable biological networks and describe the derivation of global algebraic conditions for bistability. The first example is the well known toggle switch by Gardner and Collins [22]; the second example is a simple model for an RNA-based bistable system proposed in our earlier work [35]. These examples are considered for the purpose of illustrating our approach step by step.

4.3.1 Gardner and Collins toggle switch

The circuit is composed by two mutually repressing genes, and its nondimensionalized model is:

$$\begin{cases} \dot{u} = \frac{\alpha_1}{1+v^\beta} - u \\ \dot{v} = \frac{\alpha_2}{1+u^\gamma} - v, \end{cases} \tag{4.3}$$

where α_i are the maximal synthesis rates of the i-th repressor, β and γ are Hill coefficients. The equilibrium conditions of the toggle switch are $\bar{u} = \frac{\alpha_1}{1+\bar{v}^\beta}$ and $\bar{v} = \frac{\alpha_2}{1+\bar{u}^\gamma}$ where $\bar{u}$ and $\bar{v}$ are equilibrium points. To simplify our notation, we will denote $\bar{u}$ and $\bar{v}$ as u and v, dropping the bar. Again for simplicity, we take the value of $\beta = \gamma = 2$, *i.e.* the repressors are dimers.

Application of Sturm's theorem to find algebraic equilibrium conditions

The master equilibrium condition, expressed with respect to variable v, is:

$$\alpha_1^2 v = (\alpha_2 - v)(1+v^2)^2.$$

We rewrite it as a fifth order polynomial w.r.t. v,

$$P_0 = v^5 - \alpha_2 v^4 + 2v^3 - 2\alpha_2 v^2 + (1+\alpha_1^2)v - \alpha_2 \tag{4.4}$$

We use Sturm's theorem to find algebraic conditions guaranteeing that the circuit has exactly three positive roots. First, we find the Sturm sequence $\{P_0, \cdots, P_5\}$, where $P_1 = dP_0/dv$ and $P_i = -rem(P_{i-2}, P_{i-1})$ for $i \in \{2, \cdots, n\}$. Applying Sturm's theorem, we find that 6 possible Sturm sequences (3 cases for $v = 0$ and 2 cases for $v \to \infty$) are admissible and yield exactly three positive solutions. These admissible sequences were found with the support of *Wolfram Mathematica*. Furthermore, for

$v \to \infty$ the sign of P_2 doesn't have any effect on the total number of sign variations of the sequence, so we are left with 3 distinct cases $S_I = \{-,+,-,-,+,-\}$, $S_{II} = \{-,+,+,-,+,-\}$, $S_{III} = \{-,+,-,+,+,-\}$, which are reported in Table 4.1. In Table 4.1, P_0 and P_1 are sign definite regardless of the value of α_1 and α_2. These elements are indicated by circles. For the other polynomials in the sequence, we find conditions on the parameters so that the appropriate sign changes are obtained. With some tedious derivations (Appendix A) we identify a single sufficient condition we can use to guarantee that the master equilibrium polynomial (4.4) has exactly three positive roots.

Table 4.1: Admissible Sturm sequences for $v = \{0, +\infty\}$ that yield 3 positive roots of the master equilibrium condition (4.4)

	$v \to 0$			$v \to +\infty$
	I	II	III	
P_0	⊖	⊖	⊖	⊕
P_1	⊕	⊕	⊕	⊕
P_2	$-$	$+$	$-$	$-/+$
P_3	$-$	$-$	$+$	$-$
P_4	$+$	$+$	$+$	$-$
P_5	$-$	$-$	$-$	$-$
$var(\mathscr{G}, v)$	4			1

$$\{\alpha_1 > 2 \ \wedge \ P_5(\alpha_1, \alpha_2) < 0\}, \tag{4.5}$$

Condition (4.5) was derived with the support of *Wolfram Mathematica*; details can be found in Appendix A. We will use condition (4.5) as an inequality constraint of type $\mathbb{I}_1(\alpha_1, \alpha_2)$ in the optimization problem (4.2) for the system.

In Fig. 4.1 the mono-stability and bistability regions are plotted using *MATLAB* in a range of $\alpha_i \in (0, 200]$ for $i = 1, 2$. In this plot, different colors are associated to the different Sturm sequences

in Table 4.1.

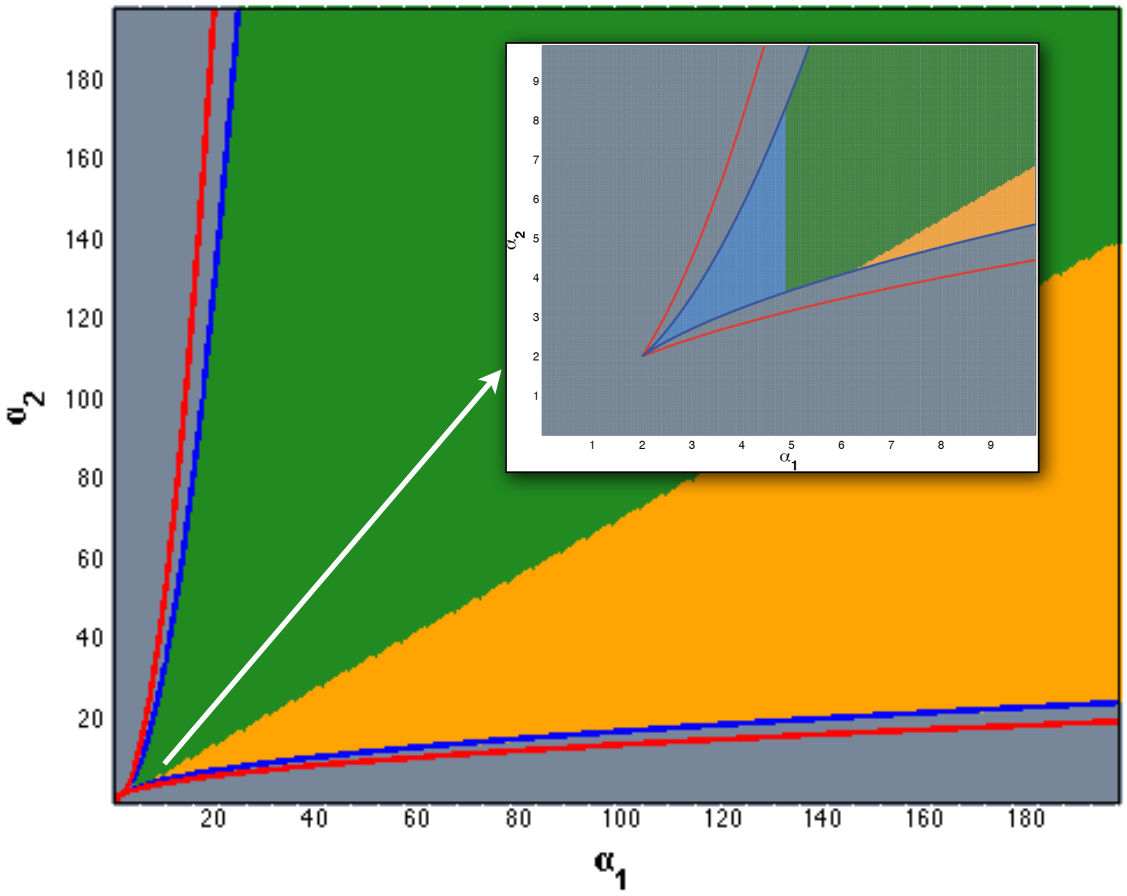

Figure 4.1: Stability domains of Gardner's toggle switch. Gray area: mono-stable domain. Green, light blue, and orange regions are bistability areas. Inset shows details for the bistability region for $\alpha_i \in (0, 10]$. The green region is associated to Sturm sequence (I) in Table 4.1, light blue is associated to sequence (II) and orange is associated to sequence (III). The blue curves are the boundaries of condition (4.5). The red curves represent the boundaries of conditions $\beta_u, \beta_v > 2$ derived in Section III A. 3) examining the constant term of the characteristic polynomial.

Verification of Assumptions 1 – 3

• Assumption 1: The Jacobian of system (4.3) is,

$$J = \begin{bmatrix} -1 & -a \\ -b & -1 \end{bmatrix},$$

where $a = \frac{2\alpha_1 v}{(1+v^2)^2}$ and $b = \frac{2\alpha_2 u}{(1+u^2)^2}$. This Jacobian can be transformed into a Metzler matrix by changing sign to its first row and first column. This Assumption is verified regardless of the system's parameter values.

• Assumption 2: The characteristic polynomial has positive coefficients except the constant term (a_0) which is not sign definite. This Assumption is verified for arbitrary values of parameters.

$$a_0 = 1 - ab = 1 - \frac{4\alpha_1^2 v^3}{\alpha_2 (1+v^2)^3} \tag{4.6}$$

Theorem 3 tells us that the stability of each equilibrium depends on the sign of a_0 : the equilibrium is stable if $a_0 > 0$, and it is unstable if $a_0 < 0$.

• Assumption 3: It is easy to show that both states of system (4.3) are eventually uniformly bounded, for arbitrary parameter values. For instance, $\dot{v} \leq \alpha - v$, thus $v(t) \leq \max\{v(0), \alpha e^{-t}\}, \forall t \geq 0$.

Algebraic conditions to determine admissible locations of equilibria

Theorem 3 indicates that the stability of each equilibrium point is determined by the constant term a_0 of the characteristic polynomial. We exploit this fact to derive algebraic conditions to determine the location of stable and unstable equilibria. We can rewrite a_0 as:

$$a_0 = \frac{n_0}{d_0} = \frac{(1+v^2)^3 - \beta_v^3 v^3}{(1+v^2)^3}, \quad \beta_v \triangleq \sqrt[3]{\frac{4\alpha_1^2}{\alpha_2}}$$

Because the denominator d_0 is positive, the sign of a_0 is only determined by its numerator n_0. The sign of n_0 can be studied by rewriting it as:

$$n_0 = (v^2 - \beta_v v + 1)\left[(1+v^2)^2 + \beta_v v(1+v^2) + \beta_v^2 v^2\right] \tag{4.7}$$

The factor in square brackets is positive, thus the sign of n_0 depends on the sign of $q(v) = (v^2 - \beta_v v + 1)$, a second order polynomial. Its discriminant $\Delta_q = \beta_v^2 - 4$ is positive for $\beta_v > 2$ and negative for $\beta_v < 2$. Thus, when $\beta_v < 2$, $\Delta_q < 0$, $q(v) > 0 \forall v$ and the sign of a_0 is always positive in all the domain; we conclude that in this case the system is never bistable, because any existing equilibrium is stable. If $\beta_v > 2$ and $\Delta_q > 0$, then the sign of $q(v)$ changes as follows:

Stable eq. $\quad q(v) > 0$	$v < \dfrac{\beta_v - \sqrt{\Delta_q}}{2} \vee \dfrac{\beta_v + \sqrt{\Delta_q}}{2} < v$
Unstable eq. $\quad q(v) < 0$	$\dfrac{\beta_v - \sqrt{\Delta_q}}{2} < v < \dfrac{\beta_v + \sqrt{\Delta_q}}{2}$.

We can derive similar inequalities for variable u, by defining the parameter $\beta_u = \sqrt[3]{\frac{4\alpha_2^2}{\alpha_1}}$. These results indicate that stable equilibria always "sandwich" the only unstable equilibrium, as depicted in Fig. 4.2.

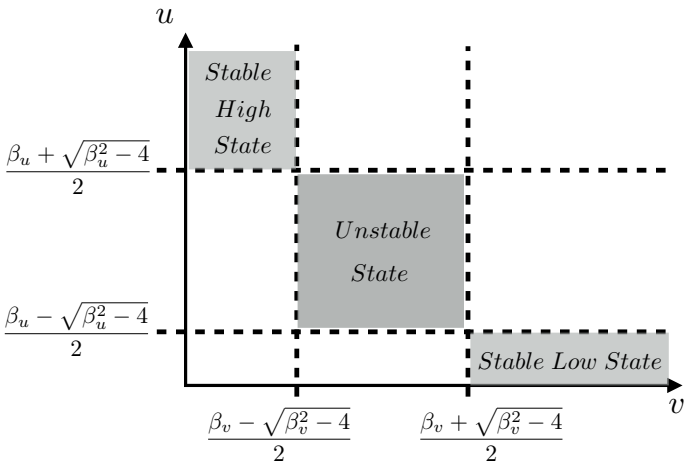

Figure 4.2: Admissible regions for stable and unstable equilibria determined by analyzing the constant term of the characteristic polynomial.

The admissible location of equilibrium points may be a feature we want to automatically optimize. A large distance between the stable equilibria, for instance, means that there is a more significant distinction between the two admissible states; it also means that switching due to perturbations may be more difficult. In contrast, a small distance between stable equilibria may favor rapid toggling between the stable states. We can influence the size of the admissible regions for equilibria by maximizing or minimizing $\beta_v > 2$ and $\beta_u > 2$.

We note that values of v and u for which $n_0 = 0$ (dashed lines in Fig. 4.2), and thus $a_0 = 0$, are not roots of the master equilibrium condition (4.4) (*Wolfram Mathematica* worksheets are available at [15]). Thus, all admissible equilibria are regular equilibria.

4.3.2 RNA-based toggle switch

This biomolecular circuit that is proposed by the authors in [35] models the two synthetic genes produce RNA outputs x_1 and x_3. The RNA outputs are transcribed by two different enzymes x_2 and x_4. The enzymes are either active or inactive and their total concentration is constant (E^{tot}). Each enzyme x_2 (x_4) is inhibited by RNA species x_3 (x_1) and it reverts to its active state at a given rate. The mathematical model is:

$$\begin{cases} \dot{x}_1 = \kappa x_2 - \delta x_1 - \gamma x_4 x_1 \\ \dot{x}_2 = \beta E^{tot} - \beta x_2 - \gamma x_2 x_3 \\ \dot{x}_3 = \kappa x_4 - \delta x_3 - \gamma x_2 x_3 \\ \dot{x}_4 = \beta E^{tot} - \beta x_4 - \gamma x_4 x_1 \end{cases}$$

In this model we assume for simplicity that the subcircuits have the same reaction rates, where κ is the transcription rate, δ is the RNA degradation rate, γ is the enzyme inhibition rate and β is the

enzymes activation rate.

Using the Routh table combined with additional observations on the system, we proved in [35] that the circuit is bistable if the following parametric inequalities hold:

$$E^{tot} - \frac{\delta}{\gamma} > 0 \ , \ \kappa > \beta \left(1 + 4 \frac{\gamma \delta E^{tot}}{(\gamma E^{tot} - \delta)^2} \right) \tag{4.8}$$

These results were confirmed in [48] using Sturm's theorem. In [35] we verified that Assumptions 1–3 hold for arbitrary parameter values. It can also be verified that if inequalities (4.8) hold, then the constant term of the characteristic polynomial is never zero, and therefore all equilibria are regular.

4.4 Parameter optimization examples

In this section we formulate two optimization problems on our case studies. In these problems, global algebraic conditions for bistability are enforced as part of the optimization constraints as explained in Section 4.2.3. Scripts were written in MATLAB; optimizations were solved using the genetic algorithm (GA) toolbox and the nonlinear constrained optimization toolbox. Simulations were done using a 2.13 GHz Intel Core 2 Duo CPU on Linux interface. The length of each single optimization run when using the GA optimization is 300 seconds on average and using the nonlinear constrained optimization toolbox it is 5 seconds on average.

4.4.1 Optimizing the Gardner and Collins toggle switch for fast toggling

We want to optimize the Gardner and Collins toggle switch so that admissible regions for its stable steady states are at a minimal distance. If the purpose of the switch is to toggle rapidly, neighboring stable steady states are likely to favor fast switching. As discussed in Section 4.3.1, this

can be done by minimizing parameters $\beta_u = \sqrt[3]{\frac{4\alpha_2^2}{\alpha_1}}$ and $\beta_v = \sqrt[3]{\frac{4\alpha_1^2}{\alpha_2}}$. In addition, we want to identify nominal values α_i^n that are sufficiently far from the boundary of the bistability regions so the system is less sensitive to the perturbations of the parameters: thus we define a "distance" parameter d_i which we want to maximize so that deviations of α_i from the nominal values will still fall in the bistable regions, i.e. $|\alpha_i - \alpha_i^n| < d_i$. By maximizing d_i we improve the robustness of the bistable behavior. This goal can be described as:

$$\mathscr{F} = \{\min_{\beta_u \in \Omega} \beta_u,\ \min_{\beta_v \in \Omega} \beta_v,\ \max_{d_i \in \Omega} d_i\}$$

This problem can be converted into a single-objective optimization problem:

$$\min_{\beta_u, \beta_v, d_1, d_2 \in \Omega} \quad \lambda_1 \beta_u^3 + \lambda_2 \beta_v^3 - \lambda_3 d_1 - \lambda_4 d_2 \tag{4.9}$$

$$\begin{aligned} \text{subject to} \quad & P_5(\alpha_1, \alpha_2) < 0, \\ & P_5(\alpha_1 - d_1, \alpha_2 + d_2) < 0, \\ & P_5(\alpha_1 + d_1, \alpha_2 - d_2) < 0, \\ & l_{\alpha_i} \leq \alpha_i \leq u_{\alpha_i},\ i = 1,2 \\ & l_{d_i} \leq d_i \leq u_{d_i},\ i = 1,2 \\ & l_\beta \leq \beta \leq u_\beta,\ \beta = \{\beta_u, \beta_v\} \end{aligned} \tag{4.10}$$

Coefficients λ_is in equation (4.9) are weights used to scale the importance of each parameter in the optimization problem. For instance, in Fig. 4.3 we compare changes for the pair λ_3 and λ_4. The i-th rectangle on the right side shows the bistability region w.r.t. the nominal rates (α_1^n, α_2^n) in the center (c_i) where the $d_1 = \frac{h_i}{2}$ and $d_2 = \frac{v_i}{2}$. On left side plot in Fig. 4.3 we have three different cases. In the 1^{st} case, the region is a square which means the value for λ_3 equals to λ_4. In this case we consider the robustness issue in the same level for both α_i rates. But in 2^{nd} case the region is a

vertical rectangle that shows $v_2 > h_2$. This means less sensitivity of α_2 for larger perturbations of its nominal value is the goal of the optimization problem and it is obvious $\lambda_3 < \lambda_4$. We can conclude similarly for 3^{rd} case where $\lambda_3 > \lambda_4$. Here, β_u, β_v are combinations of parameters, thus can be considered auxiliary variables (see Section 4.2.3).

Variables are scaled to be comparable; while both α_is have same dimensions and operational ranges as well as d_is, β_u and β_v are scaled to their third power to be comparable with the other variables.

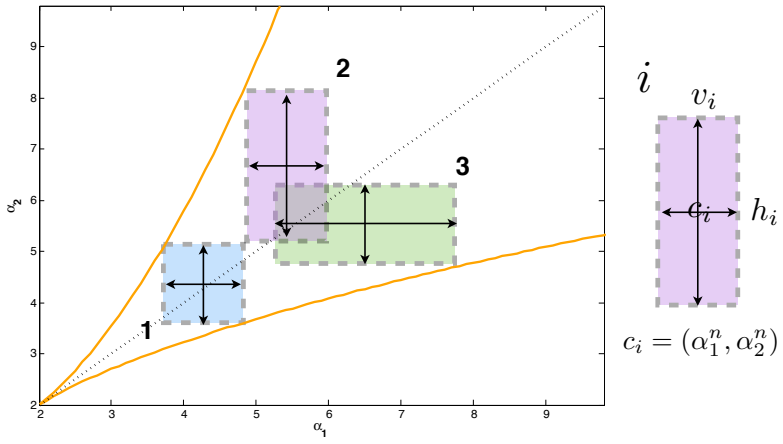

Figure 4.3: Effects of the weighting coefficients λ_3 and λ_4 in objective function (4.9).

Regarding the constraints (4.10), they include the bistability condition (4.5) which must be valid for the new coordinates $(\alpha_1 - d_1, \alpha_2 + d_2)$ and $(\alpha_1 + d_1, \alpha_2 - d_2)$. These coordinates are the diagonal vertices of the rectangles surrounding the nominal rates in Fig. 4.3. It is obvious from figure that if the inequality holds for the mentioned coordinates then for any single point inside the rectangular region the circuit will behave as a bistable switch.

The last three inequalities in (4.10) are the boundary constraints on the parameters of the problem.

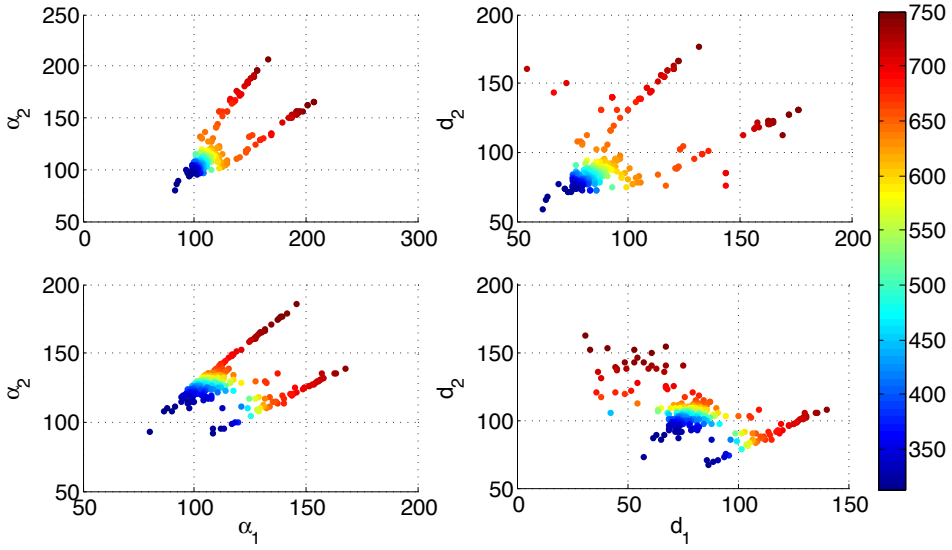

Figure 4.4: The Genetic Algorithm optimization result for 350 uniformly random selected initial points after 150 generations. (Left): $\alpha_2 - \alpha_1$ plots. (Right): $d_2 - d_1$ plots. (Upper): $\Lambda = (\lambda_1, \lambda_2, ...) = (0.5, 0.5, 0.1, 0.1)$ and (Bottom): $\Lambda = (0.5, 0.5, 0.1, 0.4)$.

Fig. 4.4 shows solutions of the optimization problem found using *MATLAB GA Optimization Toolbox*. We ran 350 simulations with randomly selected initial points (uniformly sampled). Each dot corresponds to a local minimum of a single run of the GA after 150 generations; colors from blue to red indicate an increasing value of the objective function. Dots with the same color on the different panels of Fig. 4.4 correspond to the same optimization round. We note that the optimization results are clustered in parameter space; choosing nominal parameters at the center of these clusters would guarantee a robust performance in the presence of uncertainty.

4.4.2 Selection of robust nominal parameters for the RNA based bistable switch

We want to find robust nominal values for some uncertain parameters, such that variations around the nominal parameters do not result in loss of bistability. We consider κ and E^{total} as

parameters for which we seek nominal κ_n and E_n^{tot}, and large r_1 and r_2 such that:

$$|\kappa - \kappa_n| < r_1 \ , \ |E^{tot} - E_n^{tot}| < r_2. \tag{4.11}$$

We are allowed to vary parameters (δ, β, γ) as long as we stay in feasible sets. Expressions (4.11) prompt us to modify the bistability conditions (4.8) as follows:

$$\mathbb{I}_1 : \ (E_n^{tot} - r_2) - \frac{\delta}{\gamma} > 0 \tag{4.12}$$

$$\mathbb{I}_2 : (\kappa_n - r_1) - \beta \left(1 + \frac{4\gamma\delta(E_n^{tot} - r_2)}{(\gamma(E_n^{tot} - r_2) - \delta)^2} \right) > 0 \tag{4.13}$$

In Appendix B we show that these inequalities hold for any point in the interval of $E^{tot} \in [E_n^{tot} - r_2, E_n^{tot} + r_2]$ and $\kappa \in [\kappa_n - r_1, \kappa_n + r_1]$.

The proposed objective function of the optimization problem is as equation (4.14). The other goal of the optimization problem is to reduce the concentration of E^{tot} that is the total amount of enzymes and are expensive reagents; we also seek to increase the transcription rate to speed up the system behavior.

$$\min_{E_n^{tot}, \kappa_n, \beta, \gamma, \delta, r_1, r_2 \in \Omega} \quad \lambda_1 E_n^{tot} - \lambda_2 \kappa_n - \lambda_3 r_1 - \lambda_4 r_2 \tag{4.14}$$

$$\begin{aligned} \text{subject to} \quad & \mathbb{I}_i(p) > 0, \ i = 1, \ldots, 2. \\ & Ax \leq b, \ x = [E_n^{tot}, \kappa_n, \beta, \gamma, \delta, r_1, r_2]^T \\ & l_{r_i} \leq r_i \leq u_{r_i}, \ i = 1, 2 \\ & l_{p_d} \leq p_d \leq u_{p_d}, p_d = \{E_n^{tot}, \kappa_n, \beta, \gamma, \delta\} \end{aligned} \tag{4.15}$$

The constraints $\mathbb{I}_i$ for $i = 1, 2$ come from equations (4.12)-(4.13). Linear constraints $Ax \leq b$ include constraints for $E_n^{tot} \in (l_{E^{tot}}, u_{E^{tot}})$ and $\kappa_n \in (l_\kappa, u_\kappa)$:

$$\begin{aligned} & \kappa_n - r_1 > l_\kappa \ , \ \kappa_n + r_1 < u_\kappa \\ & E_n^{tot} - r_2 > l_{E^{tot}} \ , \ E_n^{tot} + r_2 < u_{E^{tot}} \end{aligned} \tag{4.16}$$

We include additional constraints to obtain experimentally plausible parameters for strand displacement reactions and transcription (see, for instance, [29]): $\kappa \in [10^{-4}, 10^{-1}]$, $\gamma \in [10^{4}, 10^{5}]$, $\delta \in [10^{-5}, 10^{-3}]$, $\beta \in [10^{-5}, 10^{-2}]$, $E^{tot} \in [10^{-7}, 10^{-6}]$. Then we need to select $r_1 \in [0, 0.05]$ and $r_2 \in [0, 5 \cdot 10^{-7}]$.

Because decision variables have very different magnitudes, we scale them as suggested in [58] so that they fall in the interval $(0,1)$. The mapping function is $x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}}$ where x is the decision variable of the system and $x \in \Omega$.

Results of our optimization problem are shown in Fig. 4.5. We ran 1000 simulations with randomly selected initial points (uniformly sampled). Colors from blue to red indicate an increasing value of the objective function, and the same color is used to represent an individual run in the different panels of Fig. 4.5. As in the Gardner and Collins toggle switch, the optimization results are clustered in parameter space and indicate that they would provide robust behavior in the presence of parametric uncertainty.

4.5 Conclusions

We have described an approach to parameter optimization for bistable biological networks. We showed that global bistability conditions derived from standard algebraic methods can be included as nonlinear constraints in the optimization procedure. These conditions can be derived for systems that satisfy a series of assumptions: dissipativity and positivity are relatively mild assumptions in biomolecular bistable networks, and are often structural properties [35] that do not depend on the parameters. Assumption 2, which prescribes the positivity of all coefficients of the characteristic polynomial, may limit the applicability of our approach, but several biomolecular net-

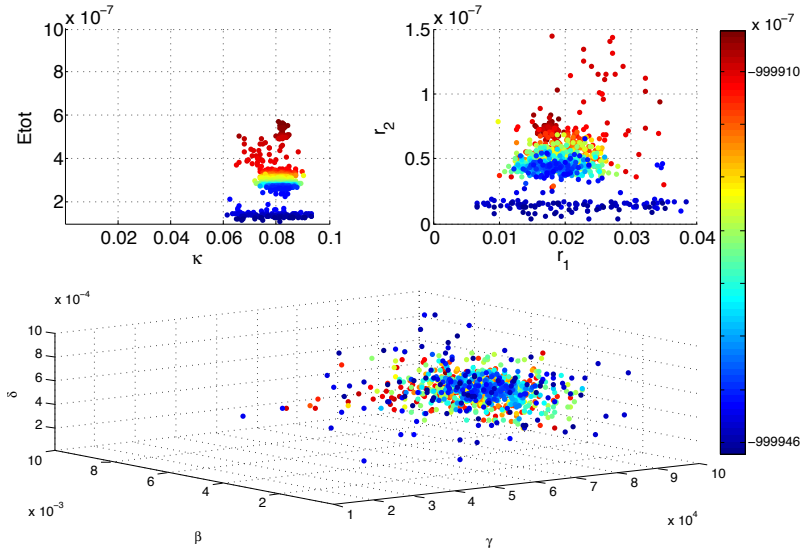

Figure 4.5: Optimization by MATLAB Nonlinear constrained optimization toolbox. (Upper left): $E^{tot} - \kappa$ plots. (Upper right): $r_2 - r_1$ plots. (Bottom): The 3D plot regarding the parameters $\delta - \beta - \gamma$. The color bar is measuring the objective function value. This simulation is for $\Lambda = (\lambda_1, \lambda_2 ...) = (8,1,1,1)$.

works in the literature are found to satisfy this requirement [48]. Clearly, algebraic tools such as Sturm's theorem can become cumbersome to apply to systems of large order. However, this method is applicable to bistable modules within larger networks, in which inputs or interconnections with the large networks may be regarded as parameters of the circuit.

Bibliography

[1] Steven S Andrews. Methods for modeling cytoskeletal and DNA filaments. *Physical biology*, 11(1):011001, 2014.

[2] David Angeli, James E Ferrell, and Eduardo D Sontag. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences of the United States of America*, 101(7):1822–1827, 2004.

[3] David Angeli, James E Ferrell, and Eduardo D Sontag. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences of the United States of America*, 101(7):1822–1827, 2004.

[4] Mariette R Atkinson, Michael A Savageau, Jesse T Myers, and Alexander J Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, 113(5):597–607, 2003.

[5] Julio R Banga. Optimization in computational systems biology. *BMC systems biology*, 2(1):47, 2008.

[6] Franco Blanchini and Elisa Franco. Structurally robust biological networks. *BMC systems biology*, 5(1):74, 2011.

[7] Franco Blanchini, Elisa Franco, and Giulia Giordano. Determining the structural properties of a class of biological models. In *Proceedings of the IEEE Conference on Decision and Control*, pages 5505–5510, 2012.

[8] Franco Blanchini, Elisa Franco, and Giulia Giordano. A structural classification of candidate oscillators and multistationary systems. *bioRxiv doi:10.1101/000562*, 2013.

[9] Franco Blanchini, Christian Cuba Samaniego, Elisa Franco, and Giulia Giordano. Design of a molecular clock with RNA-mediated regulation.

[10] YoungJung Chang and Nikolaos V Sahinidis. Optimization of metabolic pathways under stability considerations. *Computers & chemical engineering*, 29(3):467–479, 2005.

[11] Eun Jeong Cho, Joo-Woon Lee, and Andrew D. Ellington. Applications of Aptamers as Sensors. *ANNUAL REVIEW OF ANALYTICAL CHEMISTRY*, 2:241–264, 2009.

[12] Frederick R Cross, Vincent Archambault, Mary Miller, and Martha Klovstad. Testing a mathematical model of the yeast cell cycle. *Molecular biology of the cell*, 13(1):52–70, 2002.

[13] Shawn M Douglas, James J Chou, and William M Shih. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proceedings of the National Academy of Sciences*, 104(16):6644–6648, 2007.

[14] Axel Ekani-Nkodo, Ashish Kumar, and Deborah Kuchnir Fygenson. Joining and scission in the self-assembly of nanotubes from DNA tiles. *Physical review letters*, 93(26):268301, 2004.

[15] *Wolfram Mathematica* worksheets:
http://www.engr.ucr.edu/∼efranco/ACC16MathematicaFiles.html.

[16] Constantine G Evans, Rizal F Hariadi, and Erik Winfree. Direct atomic force microscopy observation of DNA tile crystal growth at the single-molecule level. *Journal of the American Chemical Society*, 134(25):10485–10492, 2012.

[17] Lorenzo Farina and Sergio Rinaldi. *Positive linear systems: theory and applications*, volume 50. John Wiley & Sons, 2011.

[18] Henrik Flyvbjerg, Elmar Jobs, and Stanislas Leibler. Kinetics of self-assembling microtubules: an "inverse problem" in biochemistry. *Proceedings of the National Academy of Sciences*, 93(12):5975–5979, 1996.

[19] H Scott Fogler. *Elements of chemical reaction engineering*. Prentice-Hall International London; 4^{th} edition., 2005.

[20] Elisa Franco, Giulia Giordano, Per-Ola Forsberg, and Richard M Murray. Negative autoregulation matches production and demand in synthetic transcriptional networks. *ACS Synthetic Biology*, 2014.

[21] Timothy S Gardner, Charles R Cantor, and James J Collins. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403(6767):339–342, 2000.

[22] Timothy S Gardner, Charles R Cantor, and James J Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403(6767):339–342, 2000.

[23] JE Gillam and CE MacPhee. Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. *Journal of Physics: Condensed Matter*, 25(37):373101, 2013.

[24] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol*, 3(10):e189, 2007.

[25] J.K. Hale. *Theory of functional differential equations*. Number v. 3 in Applied Mathematical Sciences Series. Springer Verlag GmbH, 1977.

[26] Rizal F Hariadi, Erik Winfree, and Bernard Yurke. Determining hydrodynamic forces in bursting bubbles using DNA nanotube mechanics. *Proceedings of the National Academy of Sciences*, 112(45):E6086–E6095, 2015.

[27] Rizal F Hariadi, Bernard Yurke, and Erik Winfree. Thermodynamics and kinetics of DNA nanotube polymerization from single-filament measurements. *Chemical Science*, 6(4):2252–2267, 2015.

[28] Josef Hofbauer. An index theorem for dissipative semiflows. *Journal of Mathematics*, 20(4), 1990.

[29] Jongmin Kim, Kristin S White, and Erik Winfree. Construction of an *in vitro* bistable circuit from synthetic transcriptional switches. *Molecular systems biology*, 2(1), 2006.

[30] Tuomas PJ Knowles, Christopher A Waudby, Glyn L Devlin, Samuel IA Cohen, Adriano Aguzzi, Michele Vendruscolo, Eugene M Terentjev, Mark E Welland, and Christopher M Dobson. An analytical solution to the kinetics of breakable filament assembly. *Science*, 326(5959):1533–1537, 2009.

[31] Tina Lebar, Urban Bezeljak, Anja Golob, Miha Jerala, Lucija Kadunc, Boštjan Pirš, Martin Stražar, Dušan Vučko, Uroš Zupančič, Mojca Benčina, Vida Forstnerič, Rok Gaber, Jan Lonzarić, Andreja Majerle, Alja Oblak, Anže Smole, and Roman Jerala. A bistable genetic switch based on designable DNA-binding domains. *Nat Commun*, 5:5007, 2014.

[32] Dage Liu, Sung Ha Park, John H Reif, and Thomas H LaBean. DNA nanotubes self-assembled from triple-crossover tiles as templates for conductive nanowires. *Proceedings of the National Academy of Sciences of the United States of America*, 101(3):717–722, 2004.

[33] Vahid Mardanlou, Leopold N Green, Hari KK Subramanian, Rizal F Hariadi, Jongmin Kim, and Elisa Franco. A coarse-grained model of DNA nanotube population growth. In *International Conference on DNA-Based Computers*, pages 135–147. Springer, 2016.

[34] Vahid Mardanlou, Claire H. Tran, and Elisa Franco. Design of a molecular bistable system with rna-mediated regulation.

[35] Vahid Mardanlou, Claire H Tran, and Elisa Franco. Design of a molecular bistable system with rna-mediated regulation. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 4605–4610. IEEE, 2014.

[36] Albert J Markvoort, Huub MM ten Eikelder, Peter AJ Hilbers, and Tom FA de Greef. Fragmentation and coagulation in supramolecular (Co) polymerization kinetics. *ACS central science*, 2(4):232–241, 2016.

[37] James C Mitchell, J Robin Harris, Jonathan Malo, Jonathan Bath, and Andrew J Turberfield. Self-assembly of chiral DNA nanotubes. *Journal of the American Chemical Society*, 126(50):16342–16343, 2004.

[38] Abdul M Mohammed and Rebecca Schulman. Directing self-assembly of DNA nanotubes using programmable seeds. *Nano letters*, 13(9):4006–4013, 2013.

[39] Yusuke Mori, Yoshikazu Nakamura, and Shoji Ohuchi. Inhibitory RNA aptamer against SP6 RNA polymerase. *Biochemical and Biophysical Research Communications*, 420(2):440–443, 2012.

[40] Shoji Ohuchi, Yusuke Mori, and Yoshikazu Nakamura. Evolution of an inhibitory RNA aptamer against T7 RNA polymerase. *FEBS open bio*, 2012.

[41] Sarah P Otto and Troy Day. *A biologist's guide to mathematical modeling in ecology and evolution*, volume 13. Princeton University Press, 2007.

[42] Adrien Padirac, Teruo Fujii, and Yannick Rondelez. Bottom-up construction of *in vitro* switchable memories. *Proceedings of the National Academy of Sciences*, 109(47):E3212–E3220, 2012.

[43] John H Reif, Sudheer Sahu, and Peng Yin. Compact error-resilient computational DNA tilings. In *Nanotechnology: science and computation*, pages 79–103. Springer, 2006.

[44] Paul W. K. Rothemund, Axel Ekani-Nkodo, Nick Papadakis, Ashish Kumar, Deborah Kuchnir Fygenson, and Erik Winfree. Design and characterization of programmable DNA nanotubes. *Journal of the American Chemical Society*, 126(50):16344–16352, 2004.

[45] Moisés Santillán and Michael C Mackey. Influence of catabolite repression and inducer exclusion on the bistable behavior of the lac operon. *Biophysical journal*, 86(3):1282–1292, 2004.

[46] Rebecca Schulman and Erik Winfree. Synthesis of crystals with a programmable kinetic barrier to nucleation. *Proceedings of the National Academy of Sciences*, 104(39):15236–15241, 2007.

[47] Jaswinder Sharma, Rahul Chhabra, Anchi Cheng, Jonathan Brownell, Yan Liu, and Hao Yan. Control of self-assembly of DNA tubules through integration of gold nanoparticles. *Science*, 323(5910):112–116, 2009.

[48] Dan Siegal-Gaskins, Elisa Franco, Tiffany Zhou, and Richard M. Murray. An analytical approach to bistable biological circuit discrimination using real algebraic geometry. *Journal of The Royal Society Interface*, 12(108), 2015.

[49] Dan Siegal-Gaskins, Erich Grotewold, and Gregory D Smith. The capacity for multistability in small gene regulatory networks. *BMC systems biology*, 3(1):96, 2009.

[50] E.H. Snoussi. Necessary conditions for multistationarity and stable periodicity. *Journal of Biological Systems*, 6:3–9, 1998.

[51] David Soloveichik, Georg Seelig, and Erik Winfree. DNA as a universal substrate for chemical kinetics. *Proceedings of the National Academy of Sciences*, 107(12):5393–5398, 2010.

[52] David Soloveichik and Erik Winfree. Complexity of self-assembled shapes. *SIAM Journal on Computing*, 36(6):1544–1569, 2007.

[53] Nguyen Khoa Son and Diederich Hinrichsen. Robust stability of positive continuous time systems. *Numerical functional analysis and optimization*, 17(5-6):649–659, 1996.

[54] Pakpoom Subsoontorn, Jongmin Kim, and Erik Winfree. Ensemble bayesian analysis of bistability in a synthetic transcriptional switch. *ACS synthetic biology*, 1(8):299–316, 2012.

[55] R. Thomas. On the relation between the logical structure of systems and their ability to generate multiple steady states or sustained oscillations. In Jean Dora, Jacques Demongeot, and Bernard Lacolle, editors, *Numerical Methods in the Study of Critical Phenomena*, volume 9 of *Springer Series in Synergetics*, pages 180–193. Springer Berlin Heidelberg, 1981.

[56] Spencer Angus Thomas and Yaochu Jin. Reconstructing biological gene regulatory networks: where optimization meets big data. *Evolutionary Intelligence*, 7(1):29–47, 2014.

[57] Melissa Vellela and Hong Qian. Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the schlögl model revisited. *Journal of The Royal Society Interface*, 6(39):925–940, 2009.

[58] Panchapakesan Venkataraman. *Applied optimization with MATLAB programming*. John Wiley & Sons, 2009.

[59] Jan L Vinkenborg, Nora Karnowski, and Michael Famulok. Aptamers for allosteric regulation. *Nature Chemical Biology*, 7(8):519–527, 2011.

[60] Thomas Wilhelm. The smallest chemical reaction system with bistability. *BMC Systems Biology*, 3(1):90, 2009.

[61] Erik Winfree. Simulations of computing by self-assembly. Technical report, California Institute of Technology, 1998.

[62] Gongxian Xu. Steady-state optimization of biochemical systems through geometric programming. *European Journal of Operational Research*, 225(1):12–20, 2013.

[63] David Yu Zhang, Rizal F. Hariadi, Harry M. T. Choi, and Erik Winfree. Integrating DNA strand-displacement circuitry with DNA tile self-assembly. *Nature Communications*, 4, 2013.

[64] David Yu Zhang and Georg Seelig. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chemistry*, 3(2):103–113, 2011.

Appendix A

Appendix

A.1 Monotonicity of equilibrium conditions (3.8) and (3.9)

From equilibrium condition (3.9) we derive an expression for x_4 as a function of x_2:

$$x_4 = \beta \frac{\delta E^{tot} - \delta x_2 + \gamma E^{tot} x_2 - \gamma x_2^2}{\gamma \kappa x_2}. \tag{A.1}$$

By defining $p_1 = \beta\gamma$, $p_2 = \left(E^{tot} - \frac{\delta}{\gamma}\right)$, $p_3 = \beta\delta E^{tot}$ and $p_4 = \gamma\kappa$, we can rewrite (A.1) as:

$$x_4 = \frac{p_3 + p_1 p_2 x_2 - p_1 x_2^2}{p_4 x_2} = \frac{p_3}{p_4 x_2} + \frac{p_1 p_2}{p_4} - \frac{p_1 x_2}{p_4}, \tag{A.2}$$

where $x_2 \neq 0$ and $p_i > 0$, $i = 1, ..., 4$. We now take the derivative of x_4 with respect to x_2:

$$\frac{\partial x_4}{\partial x_2} = -\frac{1}{p_4}\left(\frac{p_3}{x_2^2} + p_1\right) \tag{A.3}$$

The above derivative is strictly negative if all parameters are positive, thus equilibrium condition (A.1) is a monotonic strictly decreasing function. The same procedure can be repeated for the equilibrium condition of x_2 as a function of x_4.

A.2 Analysis of the characteristic polynomial

We compute the characteristic polynomial of the Jacobian matrix (3.11), reported in equation (A.4). We find that all the coefficients of the polynomial are positive, with the exception of the constant term. To explore the existence of any parameter-independent, we first rewrite the polynomial with a set of aggregate coefficients, then use the Routh-Hurwitz criterion to derive stability/instability conditions. We define the positive coefficients:

$$\begin{aligned}\det[sI-J] = s^4 &+ \Big(2\beta+2\delta+\gamma(\bar{x}_1+\bar{x}_2+\bar{x}_3+\bar{x}_4)\Big)s^3 \\ &+ \Big(\beta\gamma(\bar{x}_2+\bar{x}_4)+2\beta\delta+\delta\gamma(\bar{x}_1+\bar{x}_3)+\Big(\beta+\gamma(\bar{x}_1+\bar{x}_4)+\delta\Big)\Big(\beta+\gamma(\bar{x}_2+\bar{x}_3)+\delta\Big)\Big)s^2 \\ &+ \Big(\Big(\beta+\gamma(\bar{x}_1+\bar{x}_4)+\delta\Big)\Big(\beta\gamma\bar{x}_2+\delta\gamma\bar{x}_3+\beta\delta\Big)+\Big(\beta\gamma\bar{x}_4+\beta\delta+\delta\gamma\bar{x}_1\Big)\Big(\beta+\gamma(\bar{x}_2+\bar{x}_3)+\delta\Big)\Big)s \\ &+ \Big(\Big(\beta\gamma\bar{x}_4+\beta\delta+\delta\gamma\bar{x}_1\Big)\Big(\beta\gamma\bar{x}_2+\beta\delta+\delta\gamma\bar{x}_3\Big)-\gamma^2\kappa_1\kappa_2\bar{x}_2\bar{x}_4\Big)=0.\end{aligned} \tag{A.4}$$

$$P := \beta+\delta+\gamma(\bar{x}_2+\bar{x}_3), \quad Q := \beta+\delta+\gamma(\bar{x}_1+\bar{x}_4),$$

$$W := \beta\delta+\beta\gamma\bar{x}_2+\delta\gamma\bar{x}_3, \quad Z := \beta\delta+\delta\gamma\bar{x}_1+\beta\gamma\bar{x}_4.$$

Then, the characteristic polynomial (A.4) can be rewritten as:

$$\begin{aligned}p(s) =& s^4+s^3(P+Q)+s^2(W+Z+PQ) \\ &+s(QW+PZ)+(WZ-\gamma^2\kappa_1\kappa_2\bar{x}_2\bar{x}_4) \\ =& s^4+a_3s^3+a_2s^2+a_1s+a_0\end{aligned} \tag{A.5}$$

All coefficients of the polynomial are real; all coefficients are positive, with the exception of the constant term a_0 which may be positive or negative.

A.3 Admissible number of equilibria

We rewrite the equilibrium condition (3.8) as:

$$a_1 x_4^2 + b_1 x_4 + c_1 = 0, \tag{A.6}$$

where $a_1 = \beta\gamma$, $b_1 = \gamma\kappa x_2 - \beta\gamma E^{tot} + \beta\delta$, $c_1 = -\beta\delta E^{tot}$. This equation has a single positive root $x_4 = \frac{-b_1+\sqrt{\Delta_1}}{2a_1}$.

We can equate the expression above with expression (A.1) derived at Appendix A. We thus find a polynomial of variable x_2. After several computations we find:

$$A_2 x_2^4 + B_2 x_2^3 + C_2 x_2^2 + D_2 x_2 + E_2 = 0, \tag{A.7}$$

where:

$$A_2 := {A_1}^2 - (\gamma\kappa)^2 f_{12}^2, \quad B_2 := 2A_1B_1 - 2(\gamma\kappa)^2 f_{11} f_{12},$$

$$C_2 := B_1^2 + 2A_1C_1 - (\gamma\kappa)^2 (f_{11}^2 - 4a_1c_1),$$

$$D_2 := 2B_1C_1, \quad E_2 := {C_1}^2,$$

$$A_1 := -2a_1\beta\gamma + (\gamma\kappa) f_{12},$$

$$B_1 := 2a_1\beta\gamma E^{tot} - 2a_1\beta\delta + (\gamma\kappa) f_{11},$$

$$C_1 := 2a_1\beta\delta E^{tot}, \quad f_{11} := \beta\delta - \beta\gamma E^{tot}, \quad f_{12} := \gamma\kappa.$$

The Routh table for polynomial (A.7) is:

row 4:	A_2	C_2	E_2
row 3:	B_2	D_2	
row 2:	$R_1 = \frac{B_2C_2 - A_2D_2}{B_2}$	E_2	
row 1:	$R_2 = \frac{R_1D_2 - B_2E_2}{R1}$		
row 0:	E_2		

As previously noted, $E^{tot} - \frac{\delta}{\gamma}$ must be positive; then we can show that the possible sign changes in the first row of the table are:

row 4:	$\pm$
row 3:	$+$
row 2:	$R_1 : \pm$
row 1:	$R_2 : \pm$
row 0:	$+$

The first row of the table has at most 3 sign changes. We conclude that the polynomial which describes our equilibrium conditions has *at most* three positive roots, thus system (3.7) has at most three positive equilibria. To have at least one positive root, then A_2 must be negative. We examine A_2 and rewrite it as:

$$\begin{aligned} A_2 = A_1^2 - (\kappa\gamma)^2 f_{12}^2 &= \left((\kappa\gamma)^2 - 2(\beta\gamma)^2\right)^2 - (\kappa\gamma)^2 \\ &= 4a_1^2\left(a_1^2 - (\kappa\gamma)^2\right) \end{aligned} \tag{A.8}$$

If $\kappa > \beta$, then $A_2 < 0$ and we have at least one positive equilibrium. If R_1 or R_2, or both R_1 and R_2 are negative, then we have three positive roots. Analytical computations are long and tedious for this case. However, we are able to show that if $k > 2\beta$, then $R_1 < 0$. This is a conservative condition. We may find three equilibria also for $\beta < \kappa < 2\beta$, as highlighted in the numerical simulation in Fig. (3.4).

First, take $\alpha = \left(E^{tot} - \frac{\delta}{\gamma}\right) > 0$. If we substitute this term in B_2, we find that B_2 is always positive:

$$B_2 = 4\alpha a_1^2(a_1(\gamma\kappa) - 2a_1^2 + (\gamma\kappa)^2) = 4\alpha a_1^2\gamma^2(\kappa^2 + \beta\kappa - 2\beta^2). \tag{A.9}$$

If $\kappa > \beta$, then $B_2 > 0$.

We can also simplify the term E_2 which is equal $4(\beta\gamma)^2(\beta\delta E^{tot})^2$ and it is always positive.

The sign of R_1 now depends on the sign of C_2 and D_2. Before discussing those elements, we derive simplified expressions for B_1,

$$\begin{aligned} B_1 &= 2a_1\beta\gamma E^{tot} - 2a_1\beta\delta + (\kappa\gamma)(\beta\delta - \beta\gamma E^{tot}) \\ &= a_1\left(2\beta\gamma E^{tot} - 2\beta\delta + \kappa(\delta - \gamma E^{tot})\right) \\ &= a_1\gamma\left(E^{tot} - \frac{\delta}{\gamma}\right)(2\beta - \kappa) \end{aligned} \tag{A.10}$$

We now simplify the expression of $f_{11}^2 - 4a_1c_1$ as we will need it for simplifying the C_2. By squaring the f_{11} and subtracting the term $4a_1c_1$ from that, we will have the equivalent term for the expression as

$f_{11}^2 - 4a_1c_1 = (\beta\delta + \beta\gamma E^{tot})^2$.

$$\begin{aligned} f_{11}^2 &= (\beta\delta - \beta\gamma E^{tot})^2 \\ 4a_1c_1 &= -4(\beta\gamma)(\beta\delta E^{tot}) \\ f_{11}^2 - 4a_1c_1 &= (\beta\delta + \beta\gamma E^{tot})^2 \end{aligned} \tag{A.11}$$

Then, we discuss the sign of C_2 (for brevity we omit a few steps):

$$\begin{aligned} C_2 &= B_1^2 + 2A_1C_1 - (\kappa\gamma)^2(f_{11}^2 - 4a_1c_1) \\ &= a_1^2\gamma^2\left(E^{tot} - \frac{\delta}{\gamma}\right)^2(2\beta - \kappa)^2 + \\ &\quad + 4a_1\beta\delta E^{tot}(-2a_1^2 + (\gamma\kappa)^2) - (\kappa\gamma)^2(\beta\delta + \beta\gamma E^{tot})^2 \\ &= ... \\ &= 4a_1^2\beta\left(\gamma^2\left(E^{tot} - \frac{\delta}{\gamma}\right)^2(\beta - \kappa) - 2\delta a_1 E^{tot}\right) \end{aligned} \tag{A.12}$$

If $\kappa > \beta$, then $C_2 < 0$.

Finally, we inspect D_2:

$$D_2 = 2B_1C_1 = 4a_1^2\delta\gamma\beta E^{tot}\left(E^{tot} - \frac{\delta}{\gamma}\right)(2\beta - \kappa) \tag{A.13}$$

We finally find that:

$$D_2 > 0 \text{ if } \kappa < 2\beta, \qquad D_2 < 0 \text{ if } \kappa > 2\beta$$

Thus, to guarantee that $R_1 < 0$, and that there are three sign positive roots, we ask $\kappa > 2\beta$.

We consider this bound to be conservative. A thorough analysis of the sign of R_2 may provide a less conservative condition; we were not able to find analytically any such condition.

A.4 Sturm sequences for the Gardner and Collins toggle switch

In the following sections the symbol "$\vee$" means "*or*" and symbol "$\wedge$" means "*and*". From the Sturm sequence polynomials, we isolated terms that determine the sign of the sequence element (as a function

of α_1 and α_2). Term P_i^0 (P_i^∞), for instance, is the term that determines the sign of $P_i(v \to 0)$ ($P_i(v \to \infty)$). Due to space limitations, we do not report the full expressions for the Sturm sequence polynomials.

- $P_0(v \to 0) = -\alpha_2 < 0, \quad \forall \alpha_2 > 0.$
- $P_0(v \to \infty) \to \infty > 0.$
- $P_1(v \to 0) = dP_0/dv = (1 + \alpha_1^2) > 0.$
- $P_1(v \to \infty) \to \infty > 0.$
- $P_2(v \to 0)$: its sign is determined by $P_2^0 = -(\alpha_1^2 - 24)$. $P_2^0 > 0$ when $\alpha_1 < 2\sqrt{6}$, $P_2^0 < 0$ when $\alpha_1 > 2\sqrt{6}$. The sign of $P_2(v \to 0)$ does not depend on α_2.
- $P_2(v \to \infty)$: Its sign is determined by $P_2^\infty = -5 + \alpha_2^2$. $P_2^\infty > 0$ if $\alpha_2 > \sqrt{5}$ and $P_2^\infty < 0$ if $\alpha_2 < \sqrt{5}$. As shown in Table (4.1), the sign of $P_2(v \to \infty)$ doesn't have any effect on the number of positive real roots of the master equilibrium polynomial.
- $P_3(v \to 0)$. The sign of P_3 is determined by:

$$P_3^0(x) = a_{P_3^0} x^2 + b_{P_3^0} x + c_{P_3^0},$$

$$x = \alpha_2^2, a_{P_3^0} = -2, b_{P_3^0} = (\alpha_1^2 - 4), c_{P_3^0} = -2(\alpha_1^2 + 1)$$

The discriminant is $\Delta_{P_3^0} = b_{P_3^0}^2 - 4 a_{P_3^0} c_{P_3^0}$ and we define $L_{P_3^0}(\alpha_1) = \frac{-b_{P_3^0} - \sqrt{\Delta_{P_3^0}}}{2a_{P_3^0}} = \frac{(\alpha_1^2-4)+\sqrt{\alpha_1^4 - 24\alpha_1^2}}{4}$ and $U_{P_3^0}(\alpha_1) = \frac{-b_{P_3^0} + \sqrt{\Delta_{P_3^0}}}{2a_{P_3^0}} = \frac{(\alpha_1^2-4)-\sqrt{\alpha_1^4 - 24\alpha_1^2}}{4}$.

$$\begin{cases} P_3^0 > 0 & \alpha_1 > 2\sqrt{6} \wedge U_{P_3^0}(\alpha_1) < \alpha_2^2 < L_{P_3^0}(\alpha_1) \\ P_3^0 < 0 & \alpha_1 > 2\sqrt{6} \wedge \{\alpha_2^2 < U_{P_3^0}(\alpha_1) \vee \alpha_2^2 > L_{P_3^0}(\alpha_1)\} \end{cases}$$

For $\Delta_{P_3^0} < 0$ we find that:

$$P_3^0 < 0 \; if \; \alpha_1 < 2\sqrt{6}$$

- $P_3(v \to \infty)$: its sign is determined by $P_3^\infty = -(\alpha_1^2(-5+\alpha_2^2)+(1+\alpha_2^2)^2)$

$$\begin{cases} P_3^\infty < 0 & \alpha_2 > 0 \wedge \alpha_1 \leq \frac{1}{\sqrt{5}} \\ P_3^\infty < 0 & \alpha_1 > \frac{1}{\sqrt{5}} \wedge \alpha_2^2 > U_{P_3^\infty}(a_1) \end{cases}$$

$U_{P_3^\infty}(a_1) = \frac{-(\alpha_1^2+2)+\sqrt{(\alpha_1^2+2)^2+4(5\alpha_1^2-1)}}{2}$

- $P_4(v \to 0)$. The sign of P_4 is determined by:

$$P_4^0(x) = a_{P_4^0}x^2 + b_{P_4^0}x + c_{P_4^0}$$

$$x = \alpha_1^2, a_{P_4^0} = 2, b_{P_4^0} = (\alpha_2^2 - 62), c_{P_4^0} = 8(\alpha_2^2+1)^2$$

If $L_{P_4^0}(\alpha_2) = \frac{(62-\alpha_2^2)-\sqrt{63(60-4\alpha_2^2-\alpha_2^4)}}{4}$ and $U_{P_4^0}(\alpha_2) = \frac{(62-\alpha_2^2)+\sqrt{63(60-4\alpha_2^2-\alpha_2^4)}}{4}$ then,

$$\begin{cases} P_4^0 > 0 & \alpha_1 > 2 \wedge \alpha_2 > \sqrt{6} \\ P_4^0 > 0 & \alpha_2 < \sqrt{6} \wedge \{\alpha_1^2 < L_{P_4^0}(\alpha_2) \vee \alpha_1^2 > U_{P_4^0}(\alpha_2)\} \end{cases}$$

- $P_4(v \to \infty)$. The sign of this term is determined by

$$P_4^\infty(x) = a_{P_4^\infty}x^2 + b_{P_4^\infty}x + c_{P_4^\infty}, x = \alpha_2^2,$$

$$a_{P_4^\infty} = 16 + 9\alpha_1^2, b_{P_4^\infty} = 32 + 35\alpha_1^2,$$

$$c_{P_4^\infty} = 16 - 64\alpha_1^2 - 80\alpha_1^4$$

Then we define $L_{P_4^\infty}(\alpha_1) = \frac{-b_{P_4^\infty}+\sqrt{\Delta_{P_4^\infty}}}{2a_{P_4^\infty}}$ where $\Delta_{P_4^\infty} = b_{P_4^\infty}^2 - 4a_{P_4^\infty}c_{P_4^\infty}$. The conditions on the sign of P_4^∞ are:

$$\begin{cases} P_4^\infty < 0 & \alpha_2 > 0 \wedge \alpha_1 < \frac{1}{\sqrt{5}} \\ P_4^\infty < 0 & \alpha_2^2 > L_{P_4^\infty}(\alpha_1) \wedge \alpha_1 > \frac{1}{\sqrt{5}} \end{cases}$$

- $P_5(v \to 0, \infty)$: their sign is determined by:

$$P_5^0 = P_5^\infty = 256\alpha_1^6 - 3\alpha_1^4(9\alpha_2^4 + 32\alpha_2^2 - 256)$$

$$- 96\alpha_1^2(\alpha_2^4 + 29\alpha_2^2 - 8) + 256(1+\alpha_2^2)^3$$

The last polynomial is $P_5(v \to \infty) = P_5(v \to 0) < 0$ if $\alpha_1 > 2$ and $\min(R(\alpha_1)) < \alpha_2^2 < \max(R(\alpha_1))$. The $R(\alpha_1)$ is the set of positive real roots of the polynomial $P_5^{\infty}(x)$ where $x = \alpha_2^2$.

$$R(\alpha_1) = \{y \in \mathbb{R} | \forall y > 0 \ \wedge \ P_5^{\infty}(y) = 0\}$$

It is easy to prove that the aforementioned polynomial has a single negative real root or 2 positive and 1 negative real roots depending on the value of α_1.

A.5 Deriving the robust constraints for the circuit with RNA-mediated regulations

Inequality $E^{tot} - \frac{\delta}{\gamma} > 0$ (first inequality in (4.8)) is positive for all values of $E^{tot} \in [E_n^{tot} - r_2, E_n^{tot} + r_2]$, if it holds for the lower bound$E^{tot} = E_n^{tot} - r_2$. This yields inequality (4.12).

Next, we need to verify that for all $E^{tot} \in [E_n^{tot} - r_2, E_n^{tot} + r_2]$ and $\kappa \in [\kappa_n - r_1, \kappa_n + r_1]$, the second inequality in (4.8) holds if it is verified for:

$$\min[\kappa] = \kappa_n - r_1 > \max\left[\beta\left(1 + 4\frac{\gamma\delta E^{tot}}{(\gamma E^{tot} - \delta)^2}\right)\right], \tag{A.14}$$

$$\kappa \in [\kappa_n - r_1, \kappa_n + r_1], E^{tot} \in [E_n^{tot} - r_2, E_n^{tot} + r_2]$$

The right side of the inequality is a function of E^{tot}, thus evaluate its derivative with respect to E^{tot}.

$$\frac{\partial}{\partial E^{tot}}\left[\beta\left(1 + 4\frac{\gamma\delta E^{tot}}{(\gamma E^{tot} - \delta)^2}\right)\right] = \frac{-4\beta\gamma\delta(\gamma E^{tot} + \delta)}{(\gamma E^{tot} - \delta)^3} \tag{A.15}$$

The derivative is negative for any value of admissible parameters (its denominator is always positive using (4.12)). Thus, to maximize the right term of (A.14) we should substitute E^{tot} with its lowest value in the admissible interval $E_n^{tot} - r_2$. Thus we get inequality (4.13).