University of California
Santa Barbara

A modular approach to analyzing biological networks

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

Doctor of Philosophy
in
Electrical and Computer Engineering

by

Hari Sivakumar

Committee in charge:

Professor João P. Hespanha, Chair
Professor Andrew R. Teel
Professor Bassam Bamieh
Professor Jeffery Moehlis
Professor Stephen R. Proulx

December 2016

The Dissertation of
Hari Sivakumar is approved:

Professor Andrew R. Teel

Professor Bassam Bamieh

Professor Jeffery Moehlis

Professor Stephen R. Proulx

Professor João P. Hespanha, Committee Chairperson

May 2016

A modular approach to analyzing biological networks

by

Hari Sivakumar

to the supreme director BSSSB

Acknowledgements

There is a saying in my culture which goes "Maathaa, Pithaa, Guru, Deivam". Literally translated, this means "Mother, Father, Teacher, God", and refers to the order in which reverence is offered.

I have indeed been blessed with a mother who never stops inculcating good values and important lessons into me, no matter how old I get. My mother's strong advice has allowed courage and self-confidence to arise from deep within me, which is the primary reason that I was able to overcome many challenges and complete my PhD. Despite her age and health, she braved many a challenge to stay with me and give my life meaning and perspective, enabling me to overcome my mental blocks and fears. There aren't enough words to express my gratitude towards my mother, who always has me in her thoughts and prayers.

I would also like to express my gratitude to my father, who through these few years, has been my rock of support. My father has a knack for seeing the positives in every situation and providing me the encouragement to keep moving forward. When times were challenging, my father would never stop reminding me to have self-belief and self-confidence. My father's life story, from a poor village boy to a Professor in a globally reputed university, has always been a source of inspiration that I look up to.

I am very grateful to have had the honor and opportunity to work for Professor Hespanha. A simple thank you cannot do justice to the immense amount of knowledge I have gained through this PhD. Having started knowing so little (as I know now!), Professor Hespanha has taken me on a journey through topics as diverse as Networks and Graph theory, Systems Biology, Dynamical Systems theory, The Biology of Cancer, Behavioral Systems theory, Evolutionary Science and Numerical Algorithms; not to mention that I am a student in Electrical Engineering in the area of Control Theory! His guidance and

vision have been instrumental in shaping my research problem, my thinking with respect to my research problem, and my ability to write and present my work to a wide audience ranging from Biologists and Physicists to Engineers and Mathematicians. I am indeed indebted to him for molding me into the researcher I am today.

God has entered my life in many different ways. From the bottom of my heart, I would like to thank my wife Sneha for being so patient with me over the last few years. In spite of her own struggles with the PhD, and having to move so far away, she has been so patient and loving in calming me in the face of "adversity" and showing me that there really was no adversity in the first place! Her positive spirit and deep faith have been crucial during my PhD, and I am indeed fortunate to have her by my side for the rest of our lives.

I would also like to sincerely thank my brother Sriram. Our constant banter was a source of many laughs, and his light-hearted attitude to life always gives me perspective and allows me to take life a bit easier. I wholeheartedly wish him all the best in his future and career.

I also seek the blessings of my respected grandfather, whose wisdom and spiritual insights have been a shining light for our whole family.

A special thank you also to my in-laws, who have been very supportive of my work and constantly infuse positivity and love into my life.

My sincere thanks to Professor Stephen Proulx for many interesting discussions about the evolution of $p53$. Conversations with him really helped me to understand how to shape a convincing argument to biologists, which is always an important skill.

I would also like to express my sincere thank you to lab-mates old and new, Farshad, Jason, Steven and Michael for their initial guidance to help me set-up and brave the screening exam and qualifiers, Henrique, Kunihisa, Masashi, Kyriakos, David and Justin for making the last few years especially special with our banters and discussions about

so many topics, and of course our visitors Rodolfo, Adolfo, Marcio, Michelle, Victor and Lucienne, discussions with whom provided me a window to the world. I am also thankful for the relationships I have developed with our collaborators from the Proulx lab, Amy, Bjorn, Kelly and Lilla.

I am very grateful to all the Professors in the CCDC with whom I have had the pleasure of taking (very interesting) classes with. Special thanks also to Prof Andy Teel, Prof Jeff Moehlis, Prof Bassam Bamieh and Prof Stephen Proulx for agreeing to serve on my committee. Thank you also to Val who always has the answers to questions about every other aspect of graduate school.

Finally, I would like to thank the NSF for providing me with the financial support to carry out my research.

Curriculum Vitæ

Hari Sivakumar

Education

2010 – 2014	MS in Electrical and Computer Engineering, University of California, Santa Barbara, CA
2009 – 2010	MS in Electrical Engineering: Systems, University of Michigan, Ann Arbor, MI
2006 – 2009	BS in Electrical Engineering, University of Michigan, Ann Arbor, MI

Experience

2011 – 2016	Graduate Research Assistant, University of California, Santa Barbara, CA
2010 – 2011	Graduate Teaching Assistant, University of California, Santa Barbara, CA
2008	Test and Verification Intern, Motorola Home and Networks Mobility, Arlington Heights, IL

Publications

Sivakumar, H., Hespanha, J. P., Roh, K. and Proulx, S. R.: "Evolution of p53 network behavior" *Submitted for journal publication*

Sivakumar, H., Hespanha, J. P. and Proulx, S. R.: "Modular analysis of two cyclic biological circuits" *Submitted for the 55th Conf. on Dec. and Control*, December, 2016.

Sivakumar, H. and Hespanha, J. P.: "Towards modularity in biological networks while avoiding retroactivity" In *Proc. of the 2013 Amer. Contr. Conf.*, June, 2013.

Abstract

A modular approach to analyzing biological networks

by

Hari Sivakumar

This monograph addresses the decomposition of biochemical networks into functional modules that preserve their dynamic properties upon interconnection with other modules, which permits the inference of network behavior from the properties of its constituent modules. The modular decomposition method developed here also has the property that any changes in the parameters of a chemical reaction only affect the dynamics of a single module. To illustrate our results, we define and analyze a few key biological modules that arise in gene regulation, enzymatic networks, and signaling pathways. We also provide a collection of examples that demonstrate how the behavior of a biological network can be deduced from the properties of its constituent modules, based on results from control systems theory. We then use this modular decomposition method to analyze the $p53$ core regulation network, which plays a key role in tumor suppression in many organisms. By decomposing the network into modules, we study the evolution of the $p53$ core regulation network and conduct a formal analysis of the different network configurations that emerge in the evolutionary path to complexity from putative primordial organisms to more evolved vertebrates. We develop an algorithm to solve the system of equations that describe the network behavior by interconnecting the network modules systematically, as these equations are typically difficult to solve using standard numerical solvers. In the process, we qualitatively compare the distinct types of switching behaviors that each network can exhibit. We demonstrate how our novel model for the core regulation network matches experimentally observed phenomena, and contrast this with the plausible

behaviors that primordial network configurations can admit. Specifically, we show that the complexity of the $p53$ network in humans and evolved vertebrates permits a wide range of behaviors that can bring about distinct cell fate decisions, but that this is not the case for primordial organisms.

Contents

Acknowledgements **v**

Curriculum Vitæ **viii**

Abstract **ix**

1 Introduction **1**
1.1 Dissertation Overview 4
1.2 Organization 7

Part I Modular decomposition and analysis **10**

2 The decomposition of biological networks into modules **11**
2.1 Notation and network representations 12
2.2 Motivational examples 15
2.3 Rules for modular decomposition 17
2.4 Conclusion 25

3 Biological modules **26**
3.1 Module properties 26
3.2 Functional modules in common biological networks 29
3.3 Interconnection between biological modules 40
3.4 Conclusion 45

4 A modular study of two cyclic biological circuits **47**
4.1 Generalized Repressilator Network 48
4.2 Generalized Covalent Modification Network 68
4.3 Conclusion 74

5 Future Work **75**

Part II Modular analysis of the $p53$ network **77**

6 The $p53$ tumor suppressor protein **78**
6.1 Literature Review 80

7 Modeling the evolution of the $p53$ core regulation network **84**
7.1 Taxonomic representation of regulatory genes and network configurations 85
7.2 Modular models of the $p53$ pathways 89
7.3 Conclusion 96

8 Computation of the $p53$ network equilibrium points **97**
8.1 Switching behavior and computation of equilibrium points 98
8.2 An algorithm to determine $p53$ network bifurcation diagrams 99
8.3 Conclusion 108

9 The evolution of $p53$ network behavior **109**
9.1 The role of each $p53$ network configuration in determining switching behavior 110
9.2 Simple models of DNA damage transduction and repair explain experimentally observed results 119
9.3 Discussion 121
9.4 Conclusion 128

10 Future Work **130**

Bibliography **132**

A Module Characteristics **143**
A.1 Transcriptional regulation (TR) module 143
A.2 Covalent modification (CM) module 143
A.3 PD-cycle module 144

B $p53$ network model parameters **146**
B.1 Core regulation network models 146
B.2 Damage sensing and repair model 149

Chapter 1

Introduction

Biological networks are inherently very complex, consisting of several entities reacting in a nonlinear fashion [22]. While there has been a lot of study into the behavior of these discrete biological entities, rarely can biological function be attributed to a single molecular species alone [53]. Therefore, it was argued that there was a need for the recognition of functional components in biological network organization. These components are discrete entities whose biological function is separable from that of other components.

Early efforts to uncover the basic modular structures from data in large biological networks were made by identifying patterns of interconnections in these networks that occurred with much higher frequency than in complex random networks [86, 2]. The biological networks studied included gene regulatory networks in *E. coli* and *S. cerevisiae* and neuronal networks in *C. elegans* [86]. The basic modular structures found, such as feedforward loops, bi-fan and bi-parallel structures, were termed *network motifs*. The conclusions from this study were that evolution had converged on these same motifs for various systems and organization. It was even suggested that the behavior of biological networks might be interpreted from the behaviors of individual motifs, although there were no further studies on how this might be done in a general setting.

Another early study argued that the underlying modular organization in biological networks is critical in reverse engineering the interaction parameters between individual biological entities given experimental data about the global network responses to steady-state perturbations [21, 67, 5]. This technique is known as "Modular Response Analysis", and permits the inference of the existence (or non-existence) of pathways in large networks, using mathematical models of these networks. That said, the definition of a biological component was simply "units of the local type that do not share mass flow", and there were no specific ideas on how a given biological network should be decomposed into components in this study.

When decomposing a biological module into components, it would be ideal if the characteristics of each component remained the same before and after its interconnection with other components [104]. This would permit properties like stability and robustness to be predicted just from properties of each individual component in the network and knowledge of the interconnection structure. From a computational perspective, computing network parameters such as its equilibrium point(s) can be greatly simplified, since the computations can be done over a set of components as opposed to over the entire network. From an evolutionary standpoint, grouping a network into components is useful to analyze the evolvability of each component.

One approach to demarcate the network into components was said to be based on the absence of *retroactive effects* in the junctions between the different components [103, 102], i.e., when the inputs and outputs in a given junction are unidirectional. These retroactive effects are akin to an electrical module whose output voltage changes upon the addition of a load. An algorithm was also developed to decompose any given biological network into components in such a way that the number of junctions between components which have retroactive effects is minimized [101]. Further studies on retroactive effects between

pre-determined gene regulatory components [34] and signal transduction components [120, 94] revealed the exact mathematical expressions and conditions to minimize these effects.

As it turns out, retroactivity is mostly a problem in synthetic biology, where bottom-up construction techniques are used to engineer new biological networks, and modify or optimize the behavior of existing networks [99, 42, 46, 43] using pre-determined biological components [66], such as the transcriptional regulation component [35]. Retroactivity causes the dynamics of a component to change upon its interconnection with other components. To obviate this problem, synthetic biologists have designed biological *insulators* to effectively isolate these components from each other [33, 34], and these have proven to be useful [45].

The body of work on retroactivity led some researchers to believe that biological networks cannot be delimited into modules whose characteristics remain the same before and after interconnection [104]. However, from an analytical perspective, this is not true. In Mathematical biology, a biological network is generally represented by a system of ordinary differential equations (ODEs). Every one of these equations can be assigned to a component, with appropriate input-output relationships to ensure that the composition of all components will reconstruct the original system of ODEs. This method of decomposing a biological network has proven useful in deriving results pertaining to the existence of multiple equilibrium points in a network of interconnected *monotone* components [9, 8, 10], and the stability of these equilibrium points [12, 11]. A potentially undesirable feature of simply assigning ODEs to different components is that, while the components considered can be dynamically isolated from each other, the parameters of a particular chemical reaction in the network (such as the stoichiometric coefficients or the rate constants) could appear in more than one component. This turns out to be a

problem in many algorithms that automatically decompose biological networks into components using this approach [4, 64]. Consequently, a change in a single chemical reaction could result in several distinct components changing their internal dynamics.

Evolutionary biologists rely heavily on biological network parameters being associated with a unique component. For example, parameters that are "internal" to a functional biological component, but that do not affect significantly the input-output behavior of the component, are considered as neutral traits [96], meaning that their values can change because of genetic drift. Components that are parametrically isolated from each other also facilitate the study of evolutionary change, since this allows the core function of a component to be robust to changes in parameters, while allowing for the alteration of its role in bringing about distinct behaviors by changing its connections with other components [53]. As such, each component is able to evolve independently from the rest of the network in response to shock or stress, hence enhancing future evolvability. [44, 29].

1.1 Dissertation Overview

We provide a brief summary of the contributions in this monograph.

Two notions of modularity We characterize the notion of modularity in the context of biological networks, using an analytical approach. We say that a biological component is a *module* if it admits both *dynamic modularity* and *parametric modularity*. The former implies that the properties of each module do not change upon interconnection with other modules, and the latter implies that the network parameters within a module appear in no other module. In this sense, a synthetic component that undergoes loading effects upon interconnection with other components does *not* exhibit dynamic modularity,

and a component whose internal dynamics depend on parameters that also affect other components does *not* exhibit parametric modularity.

Decomposing a biological network A key contribution in this monograph is the development of a systematic method to decompose an arbitrary network of elementary biochemical reactions into modules that exhibit both dynamic and parametric modularity. This method is based on three rules that specify how to partition species and reactions into modules and also how to define the signals that connect the modules. An important novelty of our approach towards a modular decomposition is the use of reaction rates as the communicating signals between modules (as opposed to species concentrations). We show that aside from permitting parametric modularity, this allows for the use of summation junctions to combine alternative pathways that are used to produce or degrade particular species. To illustrate the use of our approach, we analyze some key biological modules that arise in gene regulatory networks, enzymatic networks, and signaling pathways, such as Transcriptional Regulator modules and Signal Transduction modules. We then analyze these modules from a systems theory perspective. Specifically, we derive their Input-output static characteristic function (IOSCF) and their Linearized Transfer Function (LTF), and explain how these functions can help us characterize a module in terms of properties such as monotonicity and stability.

Modularity as a tool to analyze two cyclic biological circuits We study a generalized repressilator which consists of a cyclic interconnection of an arbitrary number of transcriptional repressors. For the symmetric case, where all parameters across all the modules are the same (which is typical in an experimental setting [42]), we provide results for when the network will converge to a stable steady-state, both in the local and global sense. Our results show how the stable parameter region for the repressilator becomes

smaller as more transcriptional regulators are added to the network. These results were inspired by studies by Arcak and Sontag [12, 11]. We further study a generalized covalent modification network, consisting of a cascade of an arbitrary number of enzymatic reactions connected in feedback. We show that regardless of the parameters chosen for each enzyme-substrate interaction, the substrates will necessarily degrade away, in spite of being connected in positive feedback. These results were inspired by the work on Monotone Systems Theory by Angeli and Sontag [9, 10, 8].

A new model for the $p53$ tumor suppression network in humans We develop a new model for the $p53$ core regulation network in humans. The study of this model combines ideas from the modeling and bifurcation analyses of the $p53$ network done previously by researchers [118, 131, 28, 27, 98, 126, 132], more recent experimental results [17, 121, 25] and our novel modular approach. This model is shown to explain many experimental results in a way that previous models did not. In addition, this model makes novel predictions about the role of different proteins in bringing about various observed behaviors.

Modularity as a tool to study the evolution of the $p53$ tumor suppression network We study the evolution of the $p53$ core regulation network by using *Genbank* and *Uniprot* to uncover the phylogeny of four key core regulation genes include the gene corresponding to $p53$. We argue that the alternative network structures that we uncover provide a natural way to decompose the network into modules. Using this modular decomposition, we analyze the role that each module plays in bringing about the $p53$ response in humans. This provides insight into how different organisms with alternative network configurations could respond to the threat of tumors.

Modularity as a tool to compute the steady-state bifurcation of the $p53$ network model We study the distinct dynamic behaviors admitted by the alternative network configurations obtained from evolution. To do this, we determine the $p53$ equilibrium points, which reveals the bifurcations of $p53$ with respect to some network parameters. The computation of the equilibrium points cannot be obtained through standard numerical solvers for a range of parameters. To overcome this problem, we develop an algorithm to compute the equilibrium points of the network by systematically eliminating the latent variables in the network. The modular decomposition of the $p53$ network is an important tool in the operation of this algorithm.

1.2 Organization

In Chapter 2, we introduce the two notions of modularity which a biological component needs to satisfy in order for the behavior of a network to be inferred from the characteristics of individual modules, and explain the problems associated with not delimiting biological components carefully. We then provide a set of rules which if followed, allows for the decomposition of a biological network consisting of elementary chemical reactions into components or *modules* which satisfy both notions of modularity.

In Chapter 3, we introduce some modules that are commonly seen in biological networks, which satisfy both notions of modularity. We then characterize each module by properties which have implications on the characteristics of a network of interconnected modules such as stability and the number of equilibrium points. We finally demonstrate how these network characteristics can be inferred from the interconnection of modules.

In Chapter 4, we study two cyclic biological networks that arise from the interconnection of the biological modules introduced in Chapter 3. We then derive conditions for the number of equilibrium points each network can admit, and the stability of these

points.

In Chapter 5, we conclude Part I of this monograph by discussing future work that can be done on the modular analysis of biological networks.

In Chapter 6, we first introduce the $p53$ protein, and its role in preventing the emergence of tumors in organisms. We then introduce the relevant literature on experimentally observed $p53$ behavior in response to DNA damage, experiments revealing the role of the $p53$ dynamics in bringing about various cell fates, and quantitative models that attempted to capture the $p53$ behavior.

In Chapter 7, we explore evolutionary data that provides insights into the $p53$ behavior in the earlies organisms that had the gene that encoded for $p53$. Using this data, we analyze how the $p53$ *core regulation network* could have evolved from early ancestors to humans, and analyze the resulting intermediate network configurations that admit qualitatively distinct characteristics. We introduce a new model for the $p53$ tumor suppression network in humans, and demonstrate how the evolutionary pathway to complexity provides a natural way to decompose the network model into modules.

In Chapter 8, we seek to compute the equilibrium point(s) of the different network configurations that emerge from our evolutionary study in Chapter 7, to study the possible steady-state bifurcations that these networks can admit. As it turns out, the equilibrium points of some of these network configurations typically cannot be computed by standard numerical software. We introduce an algorithm to compute the equilibrium point of the network, and demonstrate how this algorithm is shaped by the modular decomposition from Chapter 7.

In Chapter 9, we demonstrate how the different network configurations that emerge from our evolutionary study in Chapter 7 admit qualitatively distinct bifurcations in response to exogenous stress inputs. We then show how the bifurcations admitted by the $p53$ network in humans can explain the experimentally observed behavior in humans, and

explore how this behavior provides insight into how $p53$ chooses between different cell fates in response to different levels of exogenous stress inputs. We then study how the $p53$ response in the primordial organisms would have been different from that observed in humans, and explain how this would affect the ability of $p53$ to bring about the range of cell fates that is possible in humans.

In Chapter 10, we conclude Part II of this monograph by discussing future work that can be done on analyzing the evolution of the $p53$ network.

Part I

Modular decomposition and analysis

Chapter 2

The decomposition of biological networks into modules

We recall that a biological component is a *module* if it admits both *dynamic modularity* and *parametric modularity*. The former implies that the properties of each module do not change upon interconnection with other modules, and the latter implies that the network parameters within a module appear in no other module. Dynamic modularity is essential to infer the behavior of a biological network from the behavior of its constituent parts, which can enable the analyses of large networks and also the design of novel networks. Parametric modularity, which has not been explicitly mentioned as often in the literature, is a useful property when identifying parameters in a biological network, like in [21, 67, 5], and also for evolutionary analysis of network modules. In this chapter, we first introduce the notation we use to represent a biological network, followed by motivational examples to illustrate problems that could occur when these biological networks are not decomposed into modules that admit dynamic and parametric modularity. We then introduce some rules that permit the decomposition of any network of elementary biochemical reactions into modules that admit these properties.

2.1 Notation and network representations

A biological network can be represented in multiple ways, including as a system of ordinary differential equations (ODEs) associated with the law of mass action kinetics, a directed bipartite species-reaction graph (DBSR), or a dynamic DBSR graph that we introduce below.

Mass Action Kinetics (MAK) Ordinary Differential Equations (ODEs)

A set of species involved in chemical reactions can be expressed as a system of ODEs using the law of mass action kinetics (MAK) when the species are well-mixed and their copy numbers are sufficiently large. For a network involving the species $\mathcal{S}_j$, $\forall j \in \{1, 2, \ldots, N_{\text{species}}\}$ and the reactions $\mathcal{R}_i$, $\forall i \in \{1, 2, \ldots, N_{\text{reactions}}\}$, the MAK results in a system of ODEs whose states are the concentrations $[\mathcal{S}_j]$, $\forall j$ of the different chemical species; the ODE representing the dynamics of a specific species $\mathcal{S}_j$ is given by

$$[\dot{\mathcal{S}}_j] = \sum_{i=1}^{N_{\text{reactions}}} \psi_{ji}(k_1, k_2, \ldots, [\mathcal{S}_1], [\mathcal{S}_2], \ldots) \tag{2.1}$$

where $\psi_{ji}(k_1, k_2, \ldots, [\mathcal{S}_1], [\mathcal{S}_2], \ldots)$ denotes the rate of production/destruction of $\mathcal{S}_j$ due to the reaction $\mathcal{R}_i$, which typically depends on the parameters $k_1, k_2, \ldots$ that are intrinsic to $\mathcal{R}_i$ (reaction rate constants and stoichiometric coefficients) and also on the concentrations of the reactants $[\mathcal{S}_1], [\mathcal{S}_2] \cdots$ of $\mathcal{R}_i$. The value of ψ_{ji} is either positive or negative depending on whether $\mathcal{S}_j$ is produced or consumed (respectively) by $\mathcal{R}_i$, or zero if $\mathcal{R}_i$ is not involved in the production or consumption of $\mathcal{S}_j$.

To facilitate the discussion, we use as a running example a simple biological network consisting of species $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$, represented by the following set of chemical reactions:

$$\begin{aligned} &\mathcal{R}_1 : \mathcal{S}_1 \xrightarrow{k_1[\mathcal{S}_1]} \mathcal{S}_2 \qquad \mathcal{R}_3 : \mathcal{S}_2 \xrightarrow{k_2[\mathcal{S}_2]} \mathcal{S}_1 \\ &\mathcal{R}_2 : \mathcal{S}_1 \xrightarrow{\gamma_1[\mathcal{S}_1]} \varnothing \qquad \mathcal{R}_4 : \mathcal{S}_2 \xrightarrow{\gamma_2[\mathcal{S}_2]} \varnothing \end{aligned} \qquad \mathcal{R}_5 : \mathcal{S}_3 \xrightarrow{k_3[\mathcal{S}_3]} \mathcal{S}_2, \tag{2.2}$$

which correspond to the following set of ODEs derived from MAK:

$$[\dot{\mathcal{S}}_1] = k_2\,[\mathcal{S}_2] - (\gamma_1 + k_1)\,[\mathcal{S}_1] \tag{2.3a}$$

$$[\dot{\mathcal{S}}_2] = k_1\,[\mathcal{S}_1] + k_3\,[\mathcal{S}_3] - (\gamma_2 + k_2)\,[\mathcal{S}_2] \tag{2.3b}$$

$$[\dot{\mathcal{S}}_3] = -k_3\,[\mathcal{S}_3]\,. \tag{2.3c}$$

With respect to the general model (2.1), the ψ_{ji} for $j \in \{1,2,3\}$ and $i \in \{1,2,3,4,5\}$ are given by

$$\begin{aligned} \psi_{11}\big(k_1,[\mathcal{S}_1]\big) &= -k_1\,[\mathcal{S}_1] & \psi_{21}\big(k_1,[\mathcal{S}_1]\big) &= k_1\,[\mathcal{S}_1] \\ \psi_{12}\big(\gamma_1,[\mathcal{S}_1]\big) &= -\gamma_1\,[\mathcal{S}_1] & \psi_{23}\big(k_2,[\mathcal{S}_2]\big) &= -k_2\,[\mathcal{S}_2] \\ \psi_{13}\big(k_2,[\mathcal{S}_2]\big) &= k_2\,[\mathcal{S}_2] & \psi_{24}\big(\gamma_2,[\mathcal{S}_2]\big) &= -\gamma_2\,[\mathcal{S}_2] \\ & & \psi_{25}\big(k_3,[\mathcal{S}_3]\big) &= k_3\,[\mathcal{S}_3] \end{aligned}$$

$$\psi_{35}\big(k_3,[\mathcal{S}_3]\big) = -k_3\,[\mathcal{S}_3]\,,$$

and 0 otherwise.

Directed Bipartite Species-Reactions (DBSR) graph

When a biological network is large, writing down the system of MAK ODEs is cumbersome and therefore much work has been done on understanding the behavior of chemical reaction networks from a graph-theoretic perspective [37]. The Directed Bipartite Species-Reaction (DBSR) graph representation of chemical reaction networks was developed in [63], and is closely related to the Species-Reaction (SR) graph introduced in [32]. In the construction of the DBSR graph, every species in the network is assigned to an

elliptical node, and every chemical reaction is assigned to a rectangular node. For every reaction in the network $\mathcal{R}_i$, there exist directed edges from the nodes corresponding to the reactants of $\mathcal{R}_i$ to the node $\mathcal{R}_i$, and from the node $\mathcal{R}_i$ to the nodes corresponding to the products of $\mathcal{R}_i$. It is worth noting that this formulation is similar to that in [103], using storages and currents. The DBSR graph of the network (2.2) is shown in Figure 2.1a. From this graph, we can infer, for example, that $\mathcal{S}_2$ is produced from $\mathcal{S}_1$ due to the reaction $\mathcal{R}_1$, and that $\mathcal{S}_1$ is a reactant in $\mathcal{R}_1$. Therefore, the concentration of $\mathcal{S}_1$ is required in the computation of the rate of the reaction $\mathcal{R}_1$.

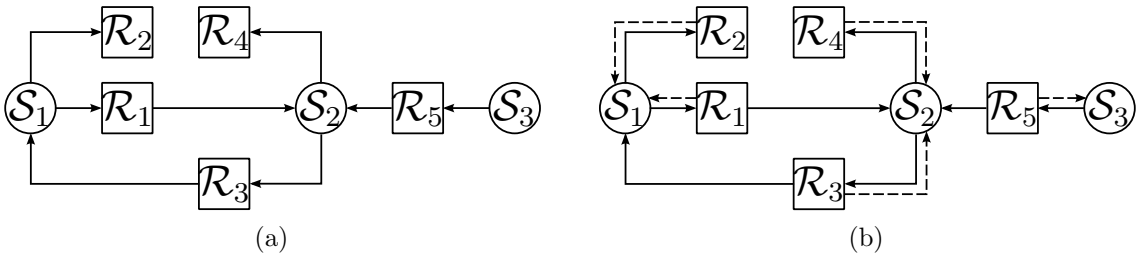


Figure 2.1: (a) DBSR graph and (b) Dynamic DBSR graph of the network represented by (2.2)

Dynamic DBSR graph

While the DBSR graph is useful in understanding the overall structure of a network of chemical reactions, it does not provide information about the flow of information in the network. For instance, the graph does not directly show whether the reaction $\mathcal{R}_1$ affects the dynamics of $\mathcal{S}_1$. To obviate this problem, we define the *dynamic* DBSR graph, which is a DBSR graph overlaid with arrows expressing the flow of information due to the dynamics of the network. In this graph, a dashed arrow from a reaction node $\mathcal{R}_i$ to a species node $\mathcal{S}_j$ indicates that $[\mathcal{S}_j]$ is affected by $\mathcal{R}_i$, usually by $\mathcal{S}_j$ being consumed in the reaction. Just like in the DBSR graph, a solid arrow from node $\mathcal{R}_i$ to node $\mathcal{S}_j$ indicates that $\mathcal{S}_j$ is produced by the reaction, while a solid arrow from node $\mathcal{S}_j$ to node

$\mathcal{R}_i$ indicates that $\mathcal{S}_j$ is a reactant of $\mathcal{R}_i$. The Dynamic DBSR graph of network (2.2) is shown in Figure 2.1b.

2.2 Motivational examples

Consider the transcriptional regulation component described as is typically done in synthetic biology [34, 68, 35]. A protein $\mathcal{S}_1$ acts as an activator for the production of another protein $\mathcal{S}_2$. Here, $[\mathcal{S}_1]$ is the input to the component. $\mathcal{S}_1$ binds to the promoter region of $\mathcal{G}_2$, the gene encoding for $\mathcal{S}_2$, to activate the transcription of $mRNA_2$. $mRNA_2$ in turn is translated into $\mathcal{S}_2$. This process can be captured by the set of chemical reactions

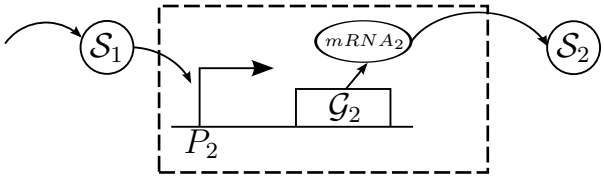
$$\varnothing \xrightarrow{f([\mathcal{S}_1])} mRNA_2 \qquad mRNA_2 \xrightarrow{\gamma_1[mRNA_2]} \varnothing$$
$$mRNA_2 \xrightarrow{\beta[mRNA_2]} mRNA_2 + \mathcal{S}_2 \qquad \mathcal{S}_2 \xrightarrow{\gamma_2[\mathcal{S}_2]} \varnothing.$$

This component is illustrated in Figure 2.2a. Suppose now, that this component is connected to another downstream transcriptional regulation component, whose input is $[\mathcal{S}_2]$. $\mathcal{S}_2$ binds to the promoter region of $\mathcal{G}_3$, which can be captured by the reaction

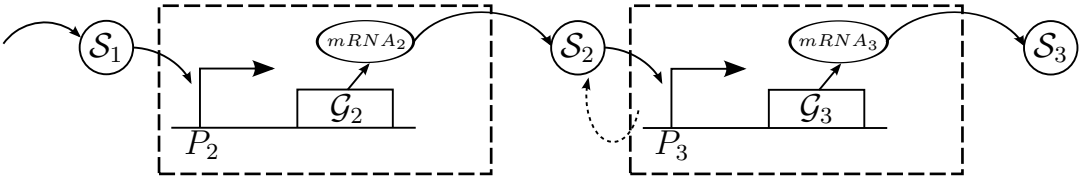
$$\mathcal{S}_2 + P_3 \underset{k^{\text{off}}[\mathcal{S}_2:P_3]}{\overset{k^{\text{on}}[\mathcal{S}_2][P_3]}{\rightleftharpoons}} \mathcal{S}_2{:}P_3. \tag{2.4}$$

As it turns out, (2.4) is a manifestation of the retroactivity phenomenon [102, 34], because the output dynamics of the upstream transcriptional regulator component changes upon the interconnection with the downstream component, as shown in Figure 2.2b. This implies that the two components defined in this way do not exhibit dynamic modularity, and therefore the properties of the interconnection of two or more components cannot be inferred from the properties of each individual component alone.

Consider also the biological network (2.2), corresponding to the MAK ODEs given



(a) Isolated transcriptional regulation component



(b) Transcriptional regulation component connected with another component downstream. The dashed line flowing in reverse signifies the change in the concentration of $\mathcal{S}_2$ upon the addition of the downstream component, due to the consumption of $\mathcal{S}_2$ when reacting with P_3

by (2.3), and suppose that we want to associate each of the three species $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ with a different component. One way this can be achieved is by partitioning the states into three components as shown in Figure 2.2, with the protein concentrations as the communicating signals between the components. Specifically, the outputs y_1, y_2, and y_3 are the concentrations of the species $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$, which in turn are the inputs u_2, u_1, and u_3 respectively. Each of these components satisfies the dynamic modularity property. That is, the concentration of each species appears in the state of one and only one component, and when the three components are combined, we obtain precisely the MAK ODEs in (2.3). However the components violate the parametric modularity property, because the rate parameters of the reactions $\mathcal{R}_1$, $\mathcal{R}_3$, and $\mathcal{R}_5$ (Equation (2.2)) appear in multiple components. Consequently, a change in the rate constant k_3 for reaction $\mathcal{R}_5$, for example, would change the internal dynamics of two components.

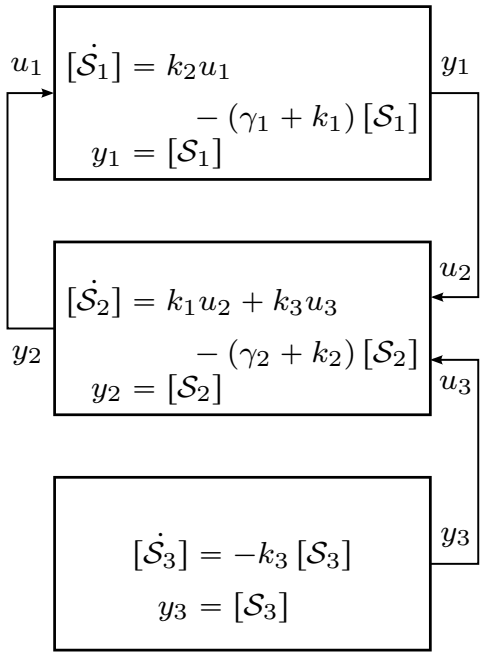


Figure 2.2: A decomposition of network (2.2) shows how parametric modularity is violated, since the parameters k_1, k_2 and k_3 appear in multiple blocks.

The examples above illustrate that the decomposition of biological networks into modules needs to be handled carefully. To ensure that every component admits both dynamic and parametric modularity, we come up with a set of rules to decompose a biological network.

2.3 Rules for modular decomposition

The modular decomposition of a biological network represented by the MAK ODE (2.1) entails the assignment of each chemical species and each chemical reaction in the network to modules. These modules then need to be interconnected appropriately such that the ODE (2.1) can be reconstructed from the module dynamics. In our framework, the assignment of a chemical species $\mathcal{S}_j$ to a module means that the species concentration $[\mathcal{S}_j]$ is part of the state of that module alone. The assignment of a chemical reaction to

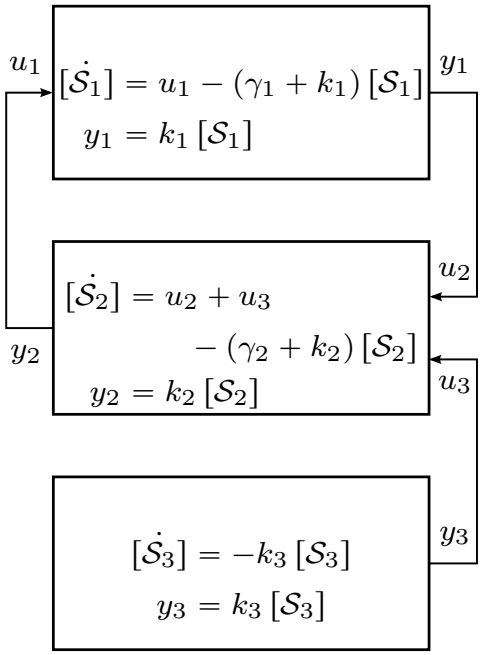
a module means that all the reaction parameters (such as stoichiometric coefficients and rate constants) appear only in that module. These observations lead to the formulation of two rules for a modular decomposition:

Rule 1 (Partition of species) *Each chemical species $\mathcal{S}_j$ must be associated with one and only one module $\mathcal{M}$, and the state of $\mathcal{M}$ is a vector containing the concentrations of all chemical species associated with $\mathcal{M}$.* □

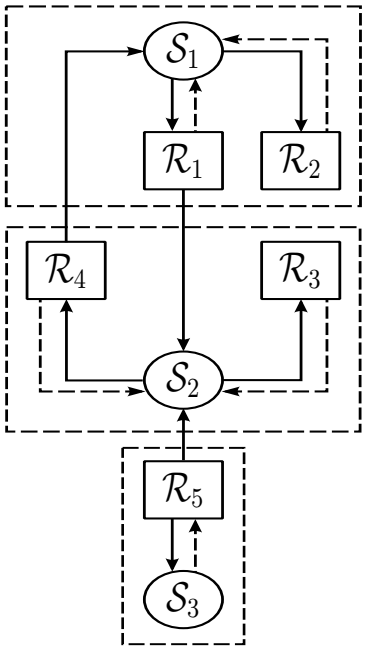
Rule 2 (Partition of reactions) *Each chemical reaction $\mathcal{R}$ must be associated with one and only one module $\mathcal{M}$, and the stoichiometric parameters and rate constants associated with $\mathcal{R}$ must only appear within the dynamics of module $\mathcal{M}$.* □

In terms of the DDBSR graphs, Rules 1 and 2 express that each node in the graph (corresponding to either a species or a reaction) must be associated with a single module. Therefore, our modular decomposition can be viewed as a *partition* of the nodes of the DDBSR graphs. We recall that a *partition* of a graph is an assignment of the nodes of the graph to disjoint sets.

The choice of signals used to communicate between modules has a direct impact on whether or not Rule 2 is violated, as is illustrated in Figure 2.2. In Figure 2.3, we illustrate a partition of the same network that satisfies Rules 1 and 2, by selecting a different set of communicating signals.



(a) Decomposition of network (2.2) using production rates as inputs and outputs, which satisfies Rules 1 and 2.



(b) Dynamic DBSR graph partition corresponding to the decomposition in Figure 2.3a

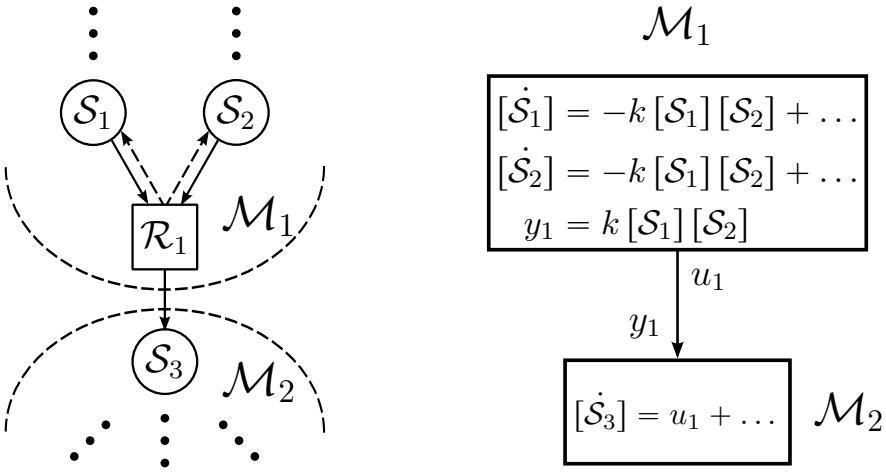
In this case, the communicating signals are the *rates of production* of the species. With this decomposition, we can now partition both the species and the reaction nodes among the different components so that the parameters of each reaction are confined to a single module, as illustrated in Figure 2.3b. We thus have a modular decomposition that simultaneously satisfies Rules 1 and 2.

2.3.1 Rates as communicating signals

In the context of a simple example, we have seen that using *rates* as the communicating signals between modules (as opposed to protein concentrations) enables a modular decomposition that simultaneously satisfies Rules 1 and 2. We now generalize these ideas to arbitrary biological networks.

When partitioning a dynamic DBSR graph into modules, each arrow of the graph "severed" by the partition corresponds to an interconnecting signal flowing between the resulting modules, in the direction of the arrow. For example, one can see two arrows being severed in Figure 2.3b by the partition between modules $\mathcal{M}_1$ and $\mathcal{M}_2$ and then two signals (y_2 and y_1) connecting the corresponding modules in Figure 2.3a. In the remainder of this section, we present two basic scenarios that can arise in partitioning a network into two modules and discuss the signals that must flow between these modules. These two cases can be applied iteratively to partition a general network into an arbitrary number of modules.

Partition at the output of a reaction node.



(c) Dynamic DBSR graph partition

(d) Block diagram decomposition.

Figure 2.3: Modular decomposition corresponding to the partition of a biochemical network at the output of a reaction node of the dynamic DBSR graph.

Suppose first that we partition a biological network into two modules $\mathcal{M}_1$ and $\mathcal{M}_2$ at the *output of a reaction node* of the dynamic DBSR graph, corresponding to a generic

elementary reaction of the form

$$\mathcal{R}_1 : \mathcal{S}_1 + \mathcal{S}_2 \xrightarrow{k[\mathcal{S}_1][\mathcal{S}_2]} \mathcal{S}_3, \tag{2.5}$$

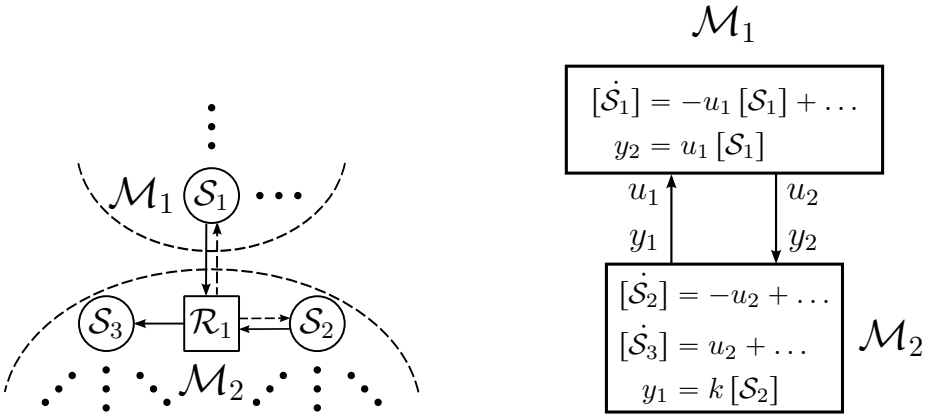
as shown in Figure 2.3c. This can be accomplished by connecting an output y_1 from module $\mathcal{M}_1$ to an input u_2 of module $\mathcal{M}_2$ that is equal to the rate of production of $\mathcal{S}_3$ due to $\mathcal{R}_1$, which is given by $k\,[\mathcal{S}_1]\,[\mathcal{S}_2]$. The block diagram representation of this partition is shown in Figure 2.3d. In this configuration, the reaction rate parameter k only appears inside the module $\mathcal{M}_1$ and we thus have parametric modularity. As far as $\mathcal{M}_2$ is concerned, the rate of production of $\mathcal{S}_3$ in *molecules per unit time* is given by the abstract chemical reaction

$$\varnothing \xrightarrow{u_1} \mathcal{S}_3,$$

where the rate u_1 is an input to the module. In this decomposition, we also have dynamic modularity, since when we combine the dynamics of the two modules in Figure 2.3d, we recover the MAK ODEs. This partition therefore ensures that both Rules 1 and 2 are satisfied. We emphasize that the *single arrow* from node $\mathcal{R}_1$ to node $\mathcal{S}_3$ that is "severed" by the partition in Figure 2.3c, gives rise to *one signal* flowing from $\mathcal{M}_1$ to $\mathcal{M}_2$ in Figure 2.3d.

Partition at the input of a reaction node.

Now consider the case where we partition the network at an *input of a reaction node* of the dynamic DBSR graph, corresponding to a generic elementary reaction of the form in (2.5), as shown in Figure 2.4a. This can be accomplished by a *bidirectional* connection between the two modules: The output y_1 from $\mathcal{M}_2$ is connected to the input u_1 of $\mathcal{M}_1$ and is equal to $k\,[\mathcal{S}_2]$, which is the *degradation* rate of a single molecule of $\mathcal{S}_1$ due to the reaction $\mathcal{R}_1$, in *molecules per molecule of* $\mathcal{S}_1$ *per unit time*. The output y_2 from $\mathcal{M}_1$ is



(a) Dynamic DBSR graph partition

(b) Block diagram decomposition.

Figure 2.4: Modular decomposition corresponding to the partition of a biochemical network at the output of a species node of the dynamic DBSR graph.

connected to the input u_2 of $\mathcal{M}_2$ and is equal to $u_1\,[\mathcal{S}_1]$, which is the rate of production of $\mathcal{S}_3$ and the net consumption rate of $\mathcal{S}_2$ due to $\mathcal{R}_1$, both in *molecules per unit time*. The block diagram representation of this modular decomposition is shown in Figure 2.4b. In isolation, the block $\mathcal{M}_1$ corresponds to a chemical reaction of the form

$$\mathcal{S}_1 \xrightarrow{u_1[\mathcal{S}_1]} \varnothing,$$

where the input u_1 determines the degradation rate of the species $\mathcal{S}_1$, and the block $\mathcal{M}_2$ corresponds to an abstract chemical reaction of the form

$$\mathcal{S}_2 \xrightarrow{u_2} \mathcal{S}_3,$$

where the input u_2 determines the production rate of the species $\mathcal{S}_3$, which is also the net consumption rate of $\mathcal{S}_2$. This decomposition is parametrically modular since the reaction rate parameter k is only part of the module $\mathcal{S}_2$, which contains the reaction $\mathcal{R}_1$. We emphasize that the *two arrows* between $\mathcal{R}_1$ to $\mathcal{S}_1$ that are "severed" by the partition in

the dynamic DBSR in Figure 2.4a give rise to the *two signals* flowing between $\mathcal{M}_1$ and $\mathcal{M}_2$ in Figure 2.4b.

The discussion above gives rise to the following general rule that governs the communicating signals between modules.

Rule 3 (Signals between modules) *Each arrow of the dynamic DBSR graph that is "severed" by the partition that defines the modular decomposition gives rise to one signal that must flow between the corresponding modules. Specifically,*

1. *When the modular decomposition cuts the dynamic DBSR graph between a reaction node $\mathcal{R}_i$ and a product species node $\mathcal{S}_j$ at the output of node $\mathcal{R}_i$, one signal must flow between the modules: the module with the reaction must have an output equal to the rate [in molecules per unit time] at which the product $\mathcal{S}_j$ is produced by the reaction.*

2. *When the modular decomposition cuts the dynamic DBSR graph between a reaction node $\mathcal{R}_i$ and a reactant species node $\mathcal{S}_j$ at the output of node $\mathcal{S}_j$, two signals must flow between the corresponding modules: the module with the reaction must have an output equal to the rate at which each molecule of $\mathcal{S}_j$ is degraded [in molecules per molecule of $\mathcal{S}_j$ per unit time], and the module with $\mathcal{S}_j$ must have an output equal to the total rate at which the molecules of $\mathcal{S}_j$ are consumed [in molecules per unit time].* □

When Rules 1–3 are followed, the decomposition of the network into modules will exhibit both parametric and dynamic modularity. Furthermore, the network can be decomposed into any number of modules less than or equal to the total number of reacting species in the network.

Remark 1 *In exploring the different possible cases, we restricted our discussion to elementary reactions (at most two reactants) and assumed that at least one of the reactants remains in the same module as the reaction. We made these assumptions mostly for simplicity, as otherwise one would have to consider a large number of cases.* □

2.3.2 Summation Junctions

In biological networks, it is not uncommon for a particular species to be produced or degraded by two or more distinct pathways. In fact, we have already encountered this in the biochemical network (2.2) corresponding to the dynamic DBSR graph depicted in Figure 2.1b, where the species $\mathcal{S}_2$ is produced both by reactions $\mathcal{R}_1$ and $\mathcal{R}_5$. The use of rates as communicating signals between modules allows for the use of summation junctions outside modules to combine different mechanisms to produce/degrade a chemical species.

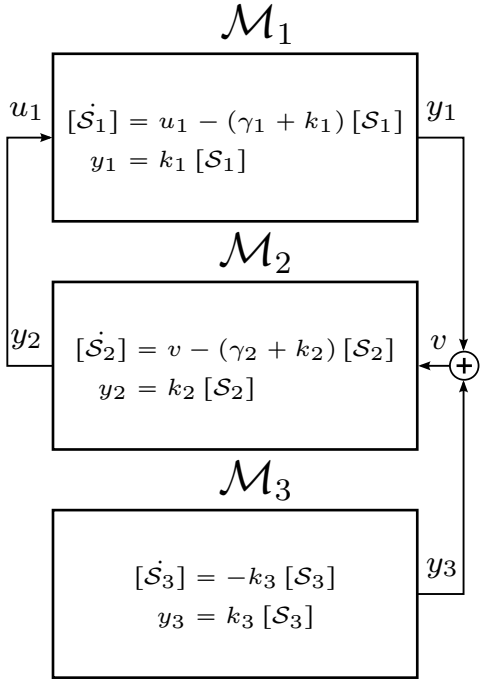


Figure 2.5: Modular decomposition from Figure 2.3a, simplified using summation junctions.

This is illustrated in Figure 2.5, where we provide a modular decomposition alternative to that shown in Figure 2.3a. This decomposition still preserves the properties of dynamic and parametric modularity, but permits simpler blocks than those in Figure 2.3a, since each module now only has a single input and a single output (SISO).

It is worth noting that when species concentrations are used as the communicating signals between modules [as in Figure 2.2], it is generally not possible to use summation junctions to combine two or more distinct mechanisms to produce or degrade a chemical species. Even if parametric modularity were not an issue, this limitation would typically lead to more complicated modules with a larger number of inputs and outputs.

2.4 Conclusion

We used examples to show that decomposition of a biological network into components would need to be done carefully to ensure that the resulting components don't violate dynamic and parametric modularity, as this violation would not permit the inference of properties of the network from properties of each individual component. We discovered a set of rules that when followed, guarantees that any modular decomposition would satisfy both dynamic and parametric modularity. The fact that communicating signals between modules are rates as opposed to protein concentrations permits the use of summation junctions between modules, which could reduce the complexity of each module in a network.

Chapter 3

Biological modules

In Chapter 2, we derived a set of rules to decompose any biological network consisting of elementary biochemical reactions into modules that satisfy dynamic and parametric modularity. In this chapter, we first recall some system theoretic properties that can be used to establish properties of complex interconnections involving these modules. We then introduce modules that correspond to important biological functions and characterize them using these properties.

3.1 Module properties

Consider a generic input-output module, expressed by an ODE of the form

$$\dot{x} = A(x,u), \qquad y = B(x,u), \qquad x \in \mathbb{R}^n, u \in \mathbb{R}^k, y \in \mathbb{R}^m, \tag{3.1}$$

where $x(t)$ denotes the n-vector state of the module, $u(t)$ the k-vector input to the module, and $y(t)$ the m-vector output from the module.

We say that the module described by (3.1) is *positive* if the entries of its state vector

$x(t)$ and output vector $y(t)$ never take negative values, as long as all the entries of the initial condition $x(0)$ and of the input vector $u(t)$, $\forall t \geqslant 0$ never take negative values. All the modules described in this section are positive.

We say that the module described by (3.1) is *cooperative* (also known as *monotone with respect to the positive orthant*) if for all initial conditions $x_0, \bar{x}_0 \in \mathbb{R}^n$ and inputs $u(t), \bar{u}(t) \in \mathbb{R}^k$, $\forall t \geqslant 0$, we have that

$$x_0 \gg \bar{x}_0 \quad \& \quad u(t) \succeq \bar{u}(t),\ \forall t \geqslant 0 \quad \Rightarrow \quad x(t; x_0, u) \gg x(t; \bar{x}_0, \bar{u}),\ \forall t > 0$$

where $x(t; x_0, u)$ denotes the solution to (3.1) at time t, starting from the initial condition $x(0) = x_0$ and with the input u. Given two vectors $v, \bar{v}$, we write $v \gg \bar{v}$ if every entry of v is strictly larger than the corresponding entry of $\bar{v}$ and we write $v \succeq \bar{v}$ if every entry of v is larger than or equal to the corresponding entry of $\bar{v}$. The reader is referred to [9, 8, 10] for a more comprehensive treatment of monotone dynamical systems, including simple conditions to test for monotonicity and results that allow one to infer monotonicity of a complex network from the monotonicity of its constituent parts. Several modules described in this section are cooperative.

The *Input-to-State Static Characteristic Function (ISSCF) $g(u^*)$ of* (3.1) specifies how a constant input $u(t) = u^*$, $\forall t \geqslant 0$ to the module maps to the corresponding equilibrium value of the state $x(t) = x^*$, $\forall t \geqslant 0$. In terms of (3.1), the value of $g(u^*)$ is the (unique) solution x^* to the steady-state equation $A(x^*, u^*) = 0$. When this equation has multiple solutions x^*, the ISSCF is not well defined.

For modules with a well-defined ISSCF, the *Input-to-Output Static Characteristic Function (IOSCF) $f(u^*)$ of* (3.1) specifies how a constant input $u(t) = u^*$, $\forall t \geqslant 0$ to the module maps to the corresponding equilibrium value of the output $y(t) = y^*$, $\forall t \geqslant 0$. In terms of (3.1), the value of $f(u^*)$ is given by $B(g(u^*), u^*)$. We shall see in

Section 3.3 that one can determine the equilibrium point of a network obtained from the interconnection of several input-output modules like (3.1), from the IOSCFs and the ISSCFs of the constituent modules.

For systems with a well-defined ISSCF, the *Linearized Transfer Function (LTF)* $H(s)$ *of* (3.1) *around an equilibrium defined by the input* u^* determines how a small perturbation $\delta u(t) \coloneqq u(t) - u^*$ of the input $u(t)$ around the constant input $u(t) = u^*$, $\forall t \geqslant 0$ leads to a perturbation $\delta y(t) \coloneqq y(t) - y^*$ of the output $y(t)$ around the constant equilibrium output $y(t) = y^* \coloneqq f(u^*)$, $\forall t \geqslant 0$. In particular, $\delta y(t) = \mathcal{L}^{-1}[H(s)] \star \delta u(t)$, where $\mathcal{L}^{-1}[H(s)]$ denotes the inverse Laplace transform of $H(s)$ and $\star$ the convolution operator [55]. We shall also see in Section 3.3 that one can determine the LTF of a network obtained from the interconnection of several input-output modules like (3.1), from the LTFs of the constituent modules.

The LTF of a module like (3.1) is given by a rational function and generally, the (local) stability of the equilibrium defined by the input u^* can be inferred from the roots of the denominator of the LTF. Specifically, if all the roots have strictly negative real parts, in which case the LTF is *bounded-input/bounded-output (BIBO) stable*, then the equilibrium point is locally asymptotically stable, which means solutions starting close to the equilibrium will converge to it as $t \to \infty$; this assumes that the McMillan degree of the LTF equals the size n of the state x of (3.1) [55], which is generically true, but should be tested.

3.2 Functional modules in common biological networks

In this section, we consider a few key biological units that arise in gene regulatory networks, enzymatic networks, and signaling pathways. We then derive the corresponding modules that admit dynamic and parametric modularity, while preserving the functionality of the unit.

3.2.1 Transcriptional regulation module

A gene regulatory network consists of a collection of transcription factor proteins, each involved in the regulation of other proteins in the network. Such a network can be decomposed into *transcriptional regulation (TR) modules*, each containing a transcription factor $\mathcal{S}_0$, the promoter regions of a set of genes $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_F$ that $\mathcal{S}_0$ up-regulates or down-regulates, and the corresponding mRNA molecules $mRNA_1, mRNA_2, \ldots, mRNA_F$ transcribed. The case $F > 1$ is referred to in the literature as *fan-out* [69].

The input u to a TR module is the rate of production of $\mathcal{S}_0$ due to exogenous processes such as regulation from other TR modules, and can be associated with a generic reaction of the form

$$\varnothing \xrightarrow{u} \mathcal{S}_0.$$

A number $q_j \geqslant 1$ of molecules of the transcription factor $\mathcal{S}_0$ can bind to the promoter region P_j of the gene $\mathcal{G}_j$, which is represented by the reaction

$$q_j\mathcal{S}_0 + P_j \underset{k_j^{\text{off}}[\mathcal{S}_0{:}P_j]}{\overset{k_j^{\text{on}}[\mathcal{S}_0]^{q_j}[P_j]}{\rightleftharpoons}} \mathcal{S}_0{:}P_j, \qquad \forall j \in \{1, 2, \ldots, m\}.$$

The total concentration of promoter regions $P_j^{\text{tot}} = [P_j] + [\mathcal{S}_0{:}P_j]$ for the gene $\mathcal{G}_j$ (bound and unbound to the transcription factor) is assumed to remain constant. When $\mathcal{S}_0$ *activates* the gene $\mathcal{G}_j$, the bound complex $\mathcal{S}_0{:}P_j$ gives rise to transcription, which is expressed by a reaction of the form

$$\mathcal{S}_0{:}P_j \xrightarrow{\alpha_j[\mathcal{S}_0:P_j]} \mathcal{S}_0{:}P_j + mRNA_j.$$

Alternatively, when $\mathcal{S}_0$ *represses* the gene $\mathcal{G}_j$, it is the unbounded promoter P_j that gives rise to transcription, which is expressed by a reaction of the form

$$P_j \xrightarrow{\alpha_j[P_j]} P_j + mRNA_j.$$

Additional reactions in the module include the translation of $mRNA_j$ to $\mathcal{S}_j$

$$mRNA_j \xrightarrow{\beta_j[mRNA_j]} mRNA_j + \mathcal{S}_j,$$

and the protein and mRNA degradation reactions

$$\mathcal{S}_0 \xrightarrow{\bar{\beta}[\mathcal{S}_0]} \varnothing,$$
$$mRNA_j \xrightarrow{\gamma_j[mRNA_j]} \varnothing$$

The TR module has F outputs $y_1, y_2, \ldots, y_F$ that are equal to the rates of translations of the proteins $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_F$, respectively. In particular,

$$y_j = \beta_j \, [mRNA_j]$$

When $F = 1$, we refer to each module simply as a TR activator or TR repressor module,

and the subscript j's can be omitted.

We derive the system of ODEs that correspond to the TR module, as well as its IOSCF and LTF, under the following assumptions:

Assumption 1 (Homogeneity in TR module) *For simplicity of presentation, it is assumed that the association and dissociation constants, the total promoter concentration, and the stoichiometric coefficients are the same for every gene, i.e., that* $k_j^{\text{off}} = k^{\text{off}}$, $k_j^{\text{on}} = k^{\text{on}}$, $P_j^{\text{tot}} = P^{\text{tot}}$ *and* $q_j = q$, $\forall j \in \{1, 2, \ldots, F\}$.

Assumption 2 (Parameters in TR module) *The following assumptions on the parameter values are considered:*

1. *The binding-unbinding reactions are on timescales much faster than those of the transcription, translation, and decay reactions, i.e.,* $k^{\text{off}}, k^{\text{on}} \gg \gamma$, β_j, $\bar{\beta}$ *and* $u(t)$ $\forall t \geqslant 0$ *[33]. This assumption simplifies the LTF of the module.*

2. *The dissociation constant* $K = \frac{k^{\text{off}}}{k^{\text{on}}}$ *is much higher than the total promoter concentration i.e* $K \gg P^{\text{tot}}$*, implying that the affinity of each binding site is small.*

□

An interesting consequence of Assumption 2 is that the LTFs from a perturbation in the input to a perturbation in each of the outputs do not depend on the fan-out, which is not true for the original dynamics without this assumption. For completeness, we include in Appendix A.1 the LTF of the TR module computed without Assumption 2.

The dynamics of this module under Assumption 1 are given by

$$[\dot{\mathcal{S}_0}] = u - \bar{\beta}\,[\mathcal{S}_0] + \sum_{j=1}^{m} q\big(k^{\text{off}}\,[\mathcal{S}_0{:}P_j] - k^{\text{on}}\,[\mathcal{S}_0]^q\,(P^{\text{tot}} - [\mathcal{S}_0{:}P_j])\big)$$

$$[\dot{\mathcal{S}_0{:}P_j}] = -k^{\text{off}}\,[\mathcal{S}_0{:}P_j] + k^{\text{on}}\,[\mathcal{S}_0]^q\,(P^{\text{tot}} - [\mathcal{S}_0{:}P_j]),$$

where

$$[\dot{mRNA_j}] = \alpha_j \left[\mathcal{S}_0{:}P_j\right] - \gamma_j \left[mRNA_j\right]$$

when $\mathcal{S}_0$ activates $\mathcal{G}_j$ and

$$[\dot{mRNA_j}] = \alpha_j (P^{\text{tot}} - \left[\mathcal{S}_0{:}P_j\right]) - \gamma_j \left[mRNA_j\right]$$

when $\mathcal{S}_0$ represses $\mathcal{G}_j$.

The dynamics of this module under Assumptions 1–2 are given by

$$[\dot{\mathcal{S}}_0] = u - \bar{\beta} \left[\mathcal{S}_0\right],$$

with

$$[\dot{mRNA_j}] = \frac{\alpha_j P_j^{\text{tot}}}{1 + \frac{K}{[\mathcal{S}_0]^q}} - \gamma_j \left[mRNA_j\right]$$

when $\mathcal{S}_0$ activates $\mathcal{G}_j$ and

$$[\dot{mRNA_j}] = \frac{\alpha_j P_j^{\text{tot}}}{1 + \frac{[\mathcal{S}_0]^q}{K}} - \gamma_j \left[mRNA_j\right]$$

when $\mathcal{S}_0$ represses $\mathcal{G}_j$.

The IOSCF of the module under Assumption 1 is given by

$$y_j^* = \frac{\alpha_j \beta_j P^{\text{tot}}}{\gamma_j \left(1 + \frac{K}{(\theta u^*)^q}\right)}$$

when $\mathcal{S}_0$ activates $\mathcal{G}_j$ and

$$y_j^* = \frac{\alpha_j \beta_j P^{\text{tot}}}{\gamma_j \left(1 + \frac{(\theta u^*)^q}{K}\right)}$$

when $\mathcal{S}_0$ represses $\mathcal{G}_j$, where

$$K \coloneqq \frac{k^{\text{off}}}{k^{\text{on}}}$$
$$\theta \coloneqq \frac{1}{\bar{\beta}}.$$

The LTF of the module under Assumptions 1–2 is given by

$$H_j(s) = \frac{qKP^{\text{tot}}\alpha_j\beta_j(\theta u^*)^{q-1}}{(K+(\theta u^*)^q)^2(s+\gamma_j)(s+\bar{\beta})}$$

when $\mathcal{S}_0$ activates $\mathcal{G}_j$ and

$$H_j(s) = -\frac{qKP^{\text{tot}}\alpha_j\beta_j(\theta u^*)^{q-1}}{(K+(\theta u^*)^q)^2(s+\gamma_j)(s+\bar{\beta})}$$

when $\mathcal{S}_0$ represses $\mathcal{G}_j$.

Figure 3.1 shows a biological representation and the corresponding DBSR graph of a TR module. The reader may verify that the module satisfies Rules 1–3 in that (i) its state contains the concentrations of all the chemical species associated with the module, (ii) the parameters of all the chemical reactions associated with the module are not needed outside this module, and (iii) the inputs and outputs of the module are rates of protein production/degradation. Furthermore, this module is positive and the equilibrium defined by any constant input u^* is locally asymptotically stable, under Assumptions 1–2. The module is also cooperative under these assumptions when all outputs are activating.

3.2.2 Covalent modification module

The *covalent modification (CM) module* represents the process by which a substrate protein $\mathcal{S}_0$ is covalently modified by an enzyme E into an alternative form $\mathcal{S}_1$. The

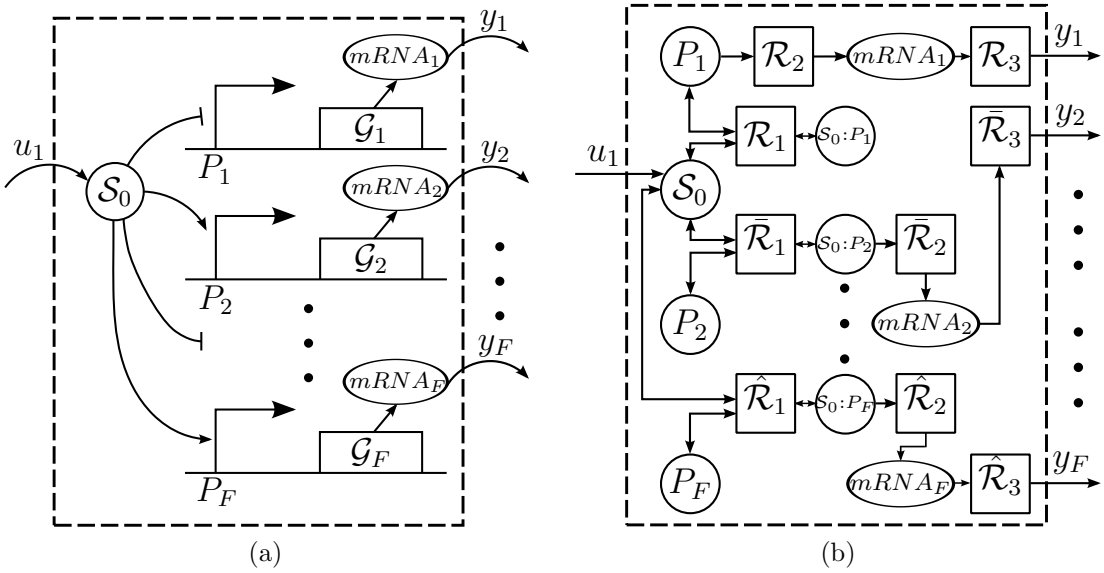


Figure 3.1: (a) Biological representation and (b) DBSR graph of a TR module. For simplicity, protein degradation reactions are shown neither in Figure 3.1b nor in all subsequent DBSR graphs presented in this monograph.

chemical species associated with this module are the substrate protein $\mathcal{S}_0$, the enzyme E, and the complex $\mathcal{S}_0{:}E$ formed by the enzyme-substrate binding. The input u to the CM module is the rate of production of the substrate $\mathcal{S}_0$ due to an exogenous process (e.g., a TR or another CM module) and can be associated with the generic reaction

$$\varnothing \xrightarrow{u} \mathcal{S}_0.$$

The additional reactions associated with the CM module include the $\mathcal{S}_0$ degradation reaction

$$\mathcal{S}_0 \xrightarrow{\gamma[\mathcal{S}_0]} \varnothing,$$

and the reactions involved in the Michaeles-Menten model for the enzyme-substrate interaction:

$$\mathcal{S}_0 + E \underset{k^{\mathrm{r}}[\mathcal{S}_0{:}E]}{\overset{k^{\mathrm{f}}[\mathcal{S}_0][E]}{\rightleftharpoons}} \mathcal{S}_0{:}E \xrightarrow{k^{\mathrm{cat}}[\mathcal{S}_0{:}E]} \mathcal{S}_1 + E,$$

where the total concentration of enzyme $E^{\mathrm{tot}} = [E] + [\mathcal{S}_0{:}E]$ is assumed to remain constant. The output y of the CM module is the rate of production of the modified substrate $\mathcal{S}_1$, given by

$$y = k^{\mathrm{cat}} [\mathcal{S}_0{:}E] .$$

For simplicity, instead of presenting the exact dynamics of the module (which are straightforward to derive using MAK), we present two common approximations to the Michaeles-Menten model:

Assumption 3 (Equilibrium Approximation [112]) *The reversible reaction is in thermodynamic equilibrium (i.e., $k^{\mathrm{f}} [\mathcal{S}_0] [E] = k^{\mathrm{r}} [\mathcal{S}_0{:}E]$), which is valid when $k^{\mathrm{r}} \gg k^{\mathrm{cat}}$.* □

Assumption 4 (Quasi Steady-State Approximation [20]) *The concentration of the complex $\mathcal{S}_0{:}E$ does not change on the timescale of product formation (i.e., $k^{\mathrm{f}} [\mathcal{S}_0] [E] = k^{\mathrm{r}} [\mathcal{S}_0{:}E] + k^{\mathrm{cat}} [\mathcal{S}_0{:}E]$), which is valid either when $k^{\mathrm{r}} + k^{\mathrm{cat}} \gg k^{\mathrm{f}}$ or when $[\mathcal{S}_0] \gg E^{\mathrm{tot}}$.* □

Under Assumption 3, the dynamics of the CM module are given by

$$[\dot{\mathcal{S}}_0] = \frac{u - \gamma [\mathcal{S}_0] - \frac{k^{\mathrm{cat}} E^{\mathrm{tot}} [\mathcal{S}_0]}{K^{\mathrm{d}} + [\mathcal{S}_0]}}{1 + \frac{K^{\mathrm{d}} E^{\mathrm{tot}}}{(K^{\mathrm{d}} + [\mathcal{S}_0]^2)}} y \qquad = \frac{k^{\mathrm{cat}} E^{\mathrm{tot}} [\mathcal{S}_0]}{K^{\mathrm{d}} + [\mathcal{S}_0]},$$

where

$$K^{\mathrm{d}} \coloneqq \frac{k^{\mathrm{r}}}{k^{\mathrm{f}}}.$$

Under Assumption 4, the dynamics of the CM module are given by

$$[\dot{\mathcal{S}}_0] = u - \gamma [\mathcal{S}_0] - \frac{k^{\mathrm{cat}} E^{\mathrm{tot}} [\mathcal{S}_0]}{K^{\mathrm{m}} + [\mathcal{S}_0]} y \qquad = \frac{k^{\mathrm{cat}} E^{\mathrm{tot}} [\mathcal{S}_0]}{K^{\mathrm{m}} + [\mathcal{S}_0]},$$

where

$$K^{\mathrm{m}} \coloneqq \frac{k^{\mathrm{r}} + k^{\mathrm{cat}}}{k^{\mathrm{f}}}.$$

The IOSCF under Assumption 3 and Assumption 4 is given by

$$y^* = \frac{1}{2}(2u^* + K^{\mathrm{e}} - \sqrt{(K^{\mathrm{e}})^2 + 4K^{\mathrm{d}}u^*\gamma})$$

and

$$y^* = \frac{1}{2}(2u^* + K^{\mathrm{f}} - \sqrt{(K^{\mathrm{f}})^2 + 4K^{\mathrm{m}}u^*\gamma})$$

respectively, where

$$K^{\mathrm{e}} \coloneqq -u^* + k^{\mathrm{cat}}E^{\mathrm{tot}} + K^{\mathrm{d}}\gamma$$

and

$$K^{\mathrm{f}} \coloneqq -u^* + k^{\mathrm{cat}}E^{\mathrm{tot}} + K^{\mathrm{m}}\gamma.$$

The LTF under Assumption 3 and Assumption 4 is given by

$$H(s) = \frac{K^{\mathrm{p}}(u^*)}{s + K^{\mathrm{q}}(u^*)}$$

and

$$H(s) = \frac{K^{\mathrm{h}}}{s + K^{\mathrm{h}} + \gamma}$$

respectively, where

$$K^{\mathrm{h}} \coloneqq \frac{k^{\mathrm{cat}}E^{\mathrm{tot}}K^{\mathrm{m}}}{(K^{\mathrm{m}} + \frac{1}{2\gamma}(K^{\mathrm{f}} + \sqrt{(K^{\mathrm{f}})^2 + 4K^{\mathrm{m}}u^*\gamma}))^2}.$$

The constants $K^{\mathrm{p}}(u^*)$ and $K^{\mathrm{q}}(u^*)$ that appear in the module LTF under Assumption 3 are included in Appendix A.2.

Figure 3.2 shows a biological representation and the corresponding DBSR graph of an CM module. This module is positive, cooperative, and the equilibrium defined by any constant input u^* is globally asymptotically stable under both assumptions.

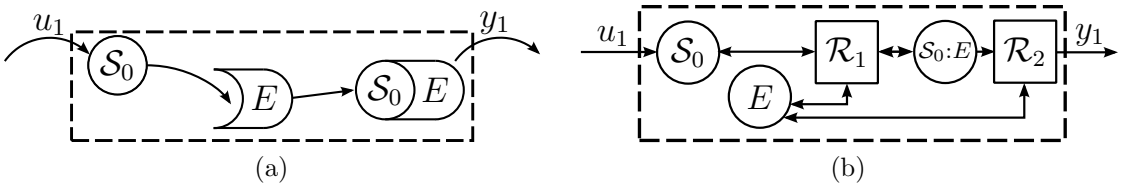


Figure 3.2: (a) Biological representation and (b) DBSR graph of an CM module.

3.2.3 Phosphorylation-dephosphorylation (PD) cascade module

Signaling pathways are common networks used by cells to transmit and receive information. A well-known signal transduction pathway, known as a signaling cascade, consists of the series of phosphorylation and dephosphorylation (PD) cycles shown in Figure 3.3.

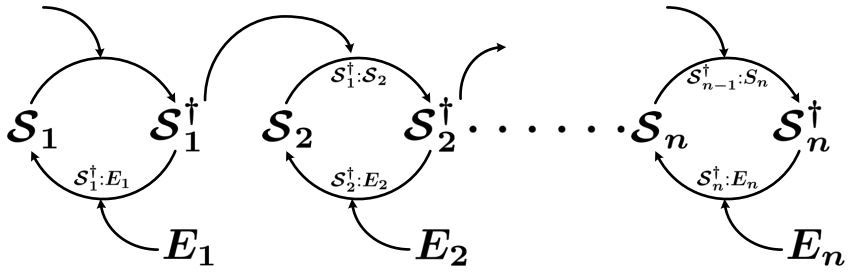


Figure 3.3: Representation of an n-stage signaling cascade, adapted from [94].

Each of these cycles, typically referred to as a *stage* in the cascade, consists of a signaling protein that can exist in either an inactive ($\mathcal{S}_i$) or an active form ($\mathcal{S}_i^\dagger$). The protein is activated by the addition of a phosphoryl group and is inactivated by its removal

[120]. The activated protein $\mathcal{S}_i^\dagger$ then goes on to act as a kinase for the phosphorylation or activation of the protein $\mathcal{S}_{i+1}$ at the next stage in the cascade. At each stage, there is also a phosphatase that removes the phosphoryl group to deactivate the activated protein.

The signaling cascade depicted in Figure 3.3 can be decomposed into n *PD-cycle modules*, each including the active protein $\mathcal{S}_i^\dagger$, the inactive protein $\mathcal{S}_{i+1}$, the complex $\mathcal{S}_i^\dagger{:}\mathcal{S}_{i+1}$ involved in the activation of the protein in the next module, and the associated chemical reactions:

$$\mathcal{S}_i^\dagger + \mathcal{S}_{i+1} \underset{k_i^{\mathrm{r}}[\mathcal{S}_i^\dagger{:}\mathcal{S}_{i+1}]}{\overset{k_i^{\mathrm{f}}[\mathcal{S}_i^\dagger][\mathcal{S}_{i+1}]}{\rightleftarrows}} \mathcal{S}_i^\dagger{:}\mathcal{S}_{i+1} \xrightarrow{\alpha_i[\mathcal{S}_i^\dagger{:}\mathcal{S}_{i+1}]} \mathcal{S}_i^\dagger + \mathcal{S}_{i+1}^\dagger.$$

In addition, the module also includes the phosphatase enzyme E_i, the complex $\mathcal{S}_i^\dagger{:}E_i$ involved in the deactivation of $\mathcal{S}_i^\dagger$, and the associated chemical reactions:

$$\mathcal{S}_i^\dagger + E_i \underset{\bar{k}_i^{\mathrm{r}}[\mathcal{S}_i^\dagger{:}E_i]}{\overset{\bar{k}_i^{\mathrm{f}}[\mathcal{S}_i^\dagger][E_i]}{\rightleftarrows}} \mathcal{S}_i^\dagger{:}E_i \xrightarrow{\bar{\alpha}_i[\mathcal{S}_i^\dagger{:}E_i]} \mathcal{S}_i + E_i,$$

where the total concentration of the enzyme $E_i^{\mathrm{tot}} = [\mathcal{S}_i^\dagger{:}E_i] + [E_i]$ is assumed to remain constant. The two inputs to this PD-cycle module are the rate of production u_i of the active protein $\mathcal{S}_i^\dagger$ due to the activation of $\mathcal{S}_i$ in the preceding module of the cascade, and the rate of production v_i of the inactive protein $\mathcal{S}_{i+1}$ due to dephosphorylation of $\mathcal{S}_{i+1}^\dagger$ in the subsequent module of the cascade. Consistently, the two outputs of this PD-cycle are the rate y_i of production of $\mathcal{S}_{i+1}^\dagger$ due to the activation of $\mathcal{S}_{i+1}$, to be used as an input to a subsequent module; and the rate z_i of production of $\mathcal{S}_i$ due to the deactivation of $\mathcal{S}_i^\dagger$ to be used as an input to a preceding module.

The dynamics of the PD-cycle module are given by

$$
\begin{aligned}
[\dot{\mathcal{S}_{i+1}}] &= v_i - k_i^{\mathrm{f}}\,[\mathcal{S}_{i+1}]\,[\mathcal{S}_i^{\dagger}] + k_i^{\mathrm{r}}[\mathcal{S}_i^{\dagger}{:}\mathcal{S}_{i+1}] \\
[\mathcal{S}_i^{\dagger}{:}\dot{\mathcal{S}}_{i+1}] &= k_i^{\mathrm{f}}\,[\mathcal{S}_{i+1}]\,[\mathcal{S}_i^{\dagger}] - (k_i^{\mathrm{r}} + \alpha_i)[\mathcal{S}_i^{\dagger}{:}\mathcal{S}_{i+1}] \\
[\dot{\mathcal{S}_i^{\dagger}}] &= u_i - k_i^{\mathrm{f}}\,[\mathcal{S}_{i+1}]\,[\mathcal{S}_i^{\dagger}] + (k_i^{\mathrm{r}} + \alpha_i)[\mathcal{S}_i^{\dagger}{:}\mathcal{S}_{i+1}] - \bar{k}_i^{\mathrm{f}}[\mathcal{S}_i^{\dagger}]\,[E_i] \\
&\qquad + \bar{k}_i^{\mathrm{r}}(E_i^{\mathrm{tot}} - [E_i]) \\
[\dot{E_i}] &= -\bar{k}_i^{\mathrm{f}}[\mathcal{S}_i^{\dagger}]\,[E_i] + (\bar{k}_i^{\mathrm{r}} + \bar{\alpha}_i)(E_i^{\mathrm{tot}} - [E_i]),
\end{aligned}
$$

with output equations

$$
\begin{aligned}
y_i &= \alpha_i[\mathcal{S}_i^{\dagger}{:}\mathcal{S}_{i+1}] \\
z_i &= \bar{\alpha}_i(E_i^{\mathrm{tot}} - [E_i]),
\end{aligned}
$$

and IOSCF

$$
y_i^* = v_i^*, \qquad\qquad z_i^* = u_i^*
$$

The LTFs of this module are straightforward to compute and can be found in Appendix A.3. While our decomposition of the signaling cascade network satisfies Rules 1–3 and hence enjoys both types of modularity, this is not the case for the decompositions of the same biochemical system found in [120, 93, 94]. Figure 3.4 shows a biological representation and the corresponding DBSR graph of a PD cycle module.

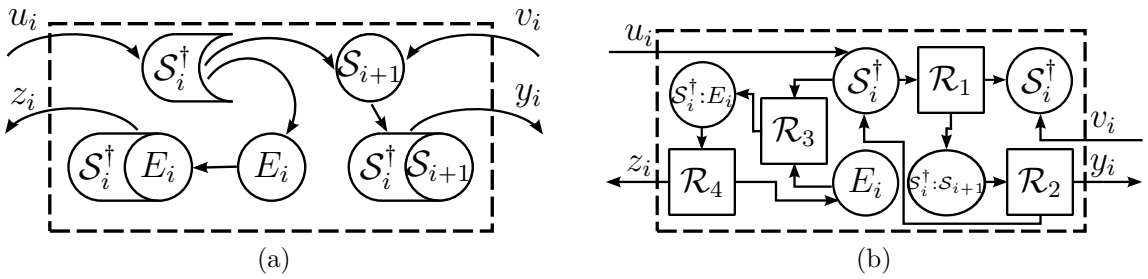


Figure 3.4: (a) Biological representation and (b) DBSR graph of a PD cycle module.

3.3 Interconnection between biological modules

A key motivation to decompose a biological network into dynamically and parametrically isolated modules is to predict properties of the overall network from the properties its constituent modules. In this section, we review three basic mechanisms for the interconnection of modules: cascade, parallel, and feedback interconnections. While simple, these three mechanisms can be combined to obtain arbitrarily complex networks. As we introduce the three interconnection mechanisms, we present a few theoretical tools that can be used to establish properties of the resulting interconnected networks.

3.3.1 Cascade interconnections

In a *cascade* interconnection between two modules $\mathcal{M}_1$ and $\mathcal{M}_2$, the output of the of the up-stream module $\mathcal{M}_1$ is connected to the input of the downstream module $\mathcal{M}_2$, as shown in Figure 3.5.

For the cascade to be well defined, the number of outputs of $\mathcal{M}_1$ must match the number of inputs of $\mathcal{M}_2$. Standard results in systems theory allow us to compute the IOSCF and LTF of the cascade from the IOSCFs and LTFs of the constituent modules:

Lemma 1 (Cascade) *Consider the cascade interconnection between an up-stream module $\mathcal{M}_1$ and a down-stream module $\mathcal{M}_2$ depicted in Figure 3.5a.*

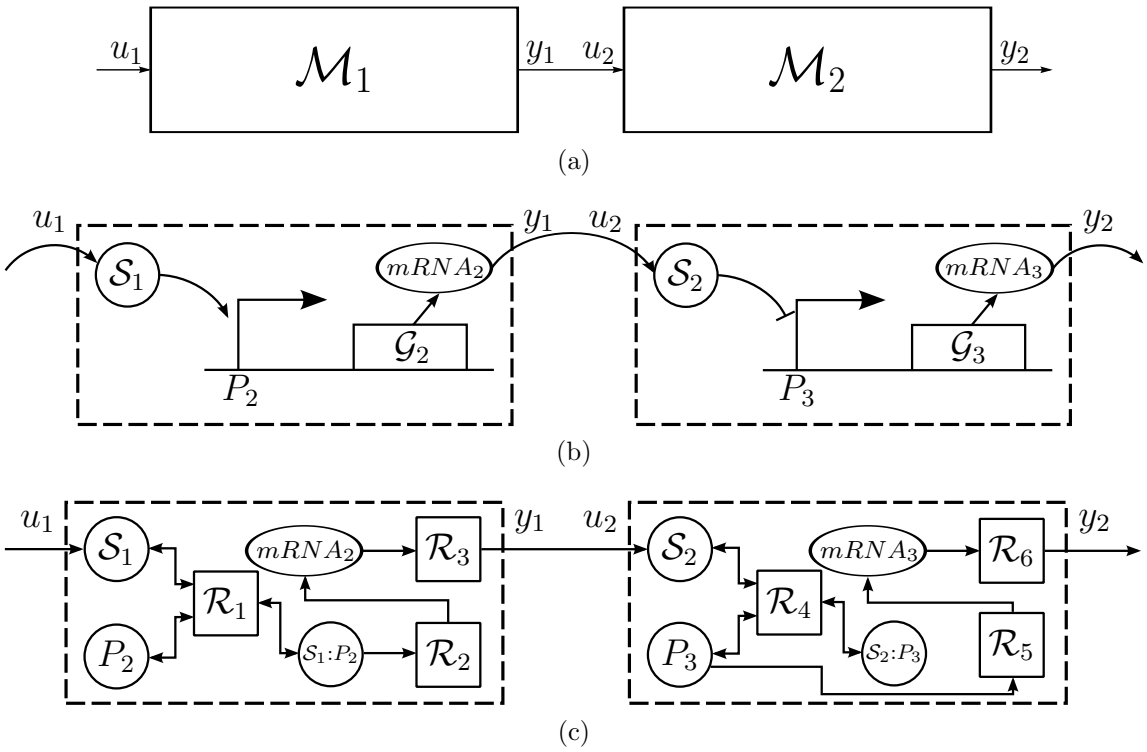


Figure 3.5: (a) Cascade interconnection between a module $\mathcal{M}_1$ and another module $\mathcal{M}_2$. (b) Example of a cascade interconnection between a TR activator module and a TR repressor module and (c) corresponding DBSR graph partition.

1. *Denoting by $f_1(u_1^*)$ and $f_2(u_2^*)$ the IOSCFs of the modules $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, the IOSCF of the cascade is well defined and given by $f(u_1^*) = f_2\big(f_1(u_1^*)\big)$, $\forall u_1^*$.*

2. *Denoting by $H_1(s)$ and $H_2(s)$ the LTFs of the modules $\mathcal{M}_1$ and $\mathcal{M}_2$ around the equilibria defined by the inputs u_1^* and $f_1(u_1^*)$, respectively, the LTF of the cascade around the equilibrium defined by the input u_1^* is given by $H(s) = H_2(s)H_1(s)$. Consequently, $H(s)$ will be BIBO stable if both $H_1(s)$ and $H_2(s)$ are BIBO stable.*

□

3.3.2 Parallel interconnections

In a *parallel* interconnection between the modules $\mathcal{M}_1$ and $\mathcal{M}_2$, the outputs of the modules are added, as shown in Figure 3.6.

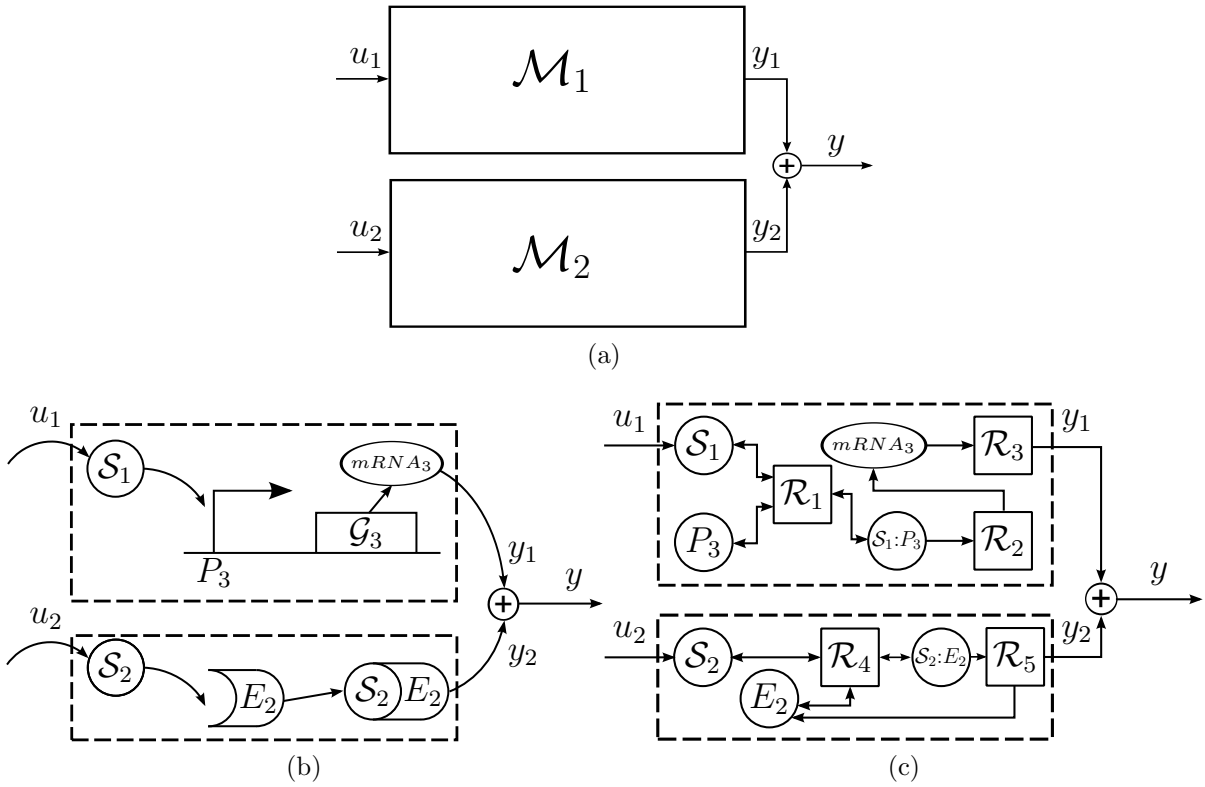


Figure 3.6: (a) Parallel interconnection of two modules $\mathcal{M}_1$ and $\mathcal{M}_2$. (b) Example of a parallel interconnection between an activator TR module and an CM module and (c) corresponding DBSR graph partition.

For the parallel interconnection to be well defined, both blocks must have the same number of outputs. Standard results in systems theory allow us to compute the IOSCF and LTF of the cascade from the IOSCFs and LTFs of the constituent modules:

Lemma 2 (Parallel) *Consider the parallel interconnection between two modules $\mathcal{M}_1$ and a $\mathcal{M}_2$ depicted in Figure 3.6a.*

1. *Denoting by $f_1(u_1^*)$ and $f_2(u_2^*)$ the IOSCFs of modules $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively,*

the (multi-input) IOSCF of the parallel interconnection is well defined and given by $f(u_1^*, u_2^*) = f_1(u_1^*) + f_2(u_2^*)$, $\forall u_1^*, u_2^*$.

2. *Denoting by* $H_1(s)$ *and* $H_2(s)$ *the LTFs of the modules* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *around the equilibria defined by the inputs* u_1^* *and* u_2^**, respectively, the (multi-input) LTF of the parallel interconnection around the equilibrium defined by the input pair* (u_1^*, u_2^*) *is obtained by stacking* $H_1(s)$, $H_2(s)$ *side by side, as in* $H(s) = \begin{bmatrix} H_1(s) & H_2(s) \end{bmatrix}$. *Consequently,* $H(s)$ *will be BIBO stable if both* $H_1(s)$ *and* $H_2(s)$ *are BIBO stable.* □

3.3.3 Feedback interconnection structures

In a *feedback* interconnection, the output of a module $\mathcal{M}_1$ is connected back to its input through a summation block, as shown in Figure 3.7. In the control theory literature, feedback interconnections are typically defined by *subtracting* the output from the input (instead of adding, as in Figure 3.7). However, since a summation is more natural in the context of biological networks, here we deviate from that standard approach. For the feedback interconnection to be well defined, the module $\mathcal{M}_1$ must have the same number of inputs and outputs. Standard results in systems theory allow us to compute the IOSCF and LTF of the feedback from the IOSCF and LTF of the constituent module:

Lemma 3 (Feedback) *Consider the feedback interconnection of a module* $\mathcal{M}_1$ *depicted in Figure 3.7a.*

1. *Denoting by* $f_1(u_1^*)$ *the IOSCF of the module* $\mathcal{M}_1$*, the IOSCF of the feedback interconnection is given by* $f(u^*) = y^*$*, where* y^* *is the solution to the equation* $y^* = f_1(u^* + y^*)$. *When this equation has multiple solutions, the IOSCF of the feedback is not well defined.*

2. *Denoting by* $H_1(s)$ *the LTFs of the module* $\mathcal{M}_1$ *around the equilibria defined by*

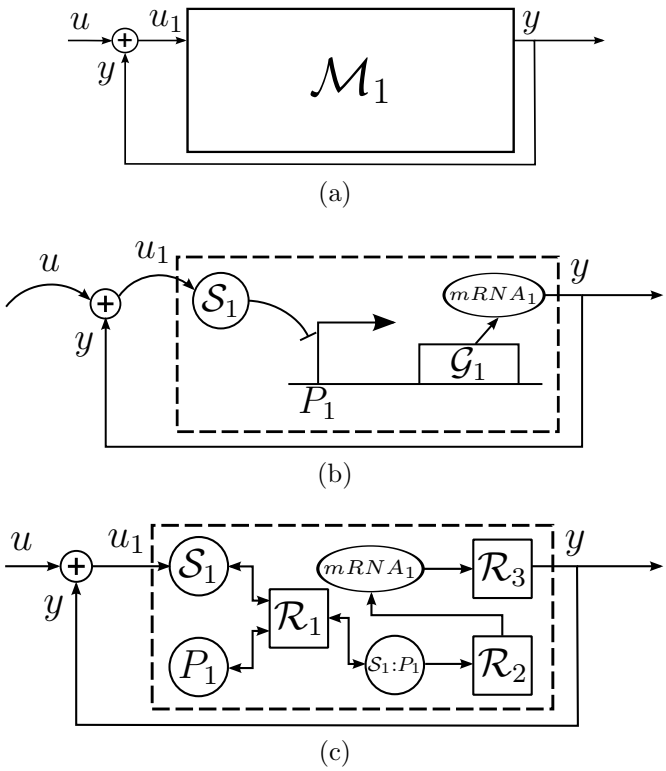


Figure 3.7: (a) Feedback interconnection of module $\mathcal{M}_1$. (b) Example of a feedback interconnection of a TR repressor module and (c) corresponding DBSR graph partition.

the input $u_1^* = u^* + f(u^*)$*, the LTF of the feedback interconnection around the equilibrium defined by the input* u^* *is given by* $H(s) = \big(I - H_1(s)\big)^{-1} H_1(s) = H_1(s)\big(I - H_1(s)\big)^{-1}$. □

For the cascade and parallel interconnections, when the IOSCFs of the constituent modules were well defined, the IOSCFs of the interconnections were also well defined. Moreover, if the constituent modules had BIBO stable LTFs, the LTF of the interconnection was also BIBO stable. This is no longer the true for feedback connections, where the block $\mathcal{M}_1$ may have a well-defined IOSCF and a BIBO stable LTF, but the interconnection may have multiple equilibria and the LTFs around these equilibria may or may not be BIBO stable. This will become more evident in the examples we study in Chapter 4.

3.3.4 Nested interconnection structures

In general, arbitrarily complex biological networks can be decomposed into combinations of cascade, parallel and feedback interconnection structures, also known as *nested* interconnection structures. One such example, involving TR and CM modules, is depicted in Figure 3.8. In this case, a protein $\mathcal{S}_1$, which is produced by some exogenous process at a rate u, activates a gene $\mathcal{G}_2$ while simultaneously repressing $\mathcal{G}_3$. The protein $\mathcal{S}_2$ is then covalently modified to $\mathcal{S}_4$, while $\mathcal{S}_3$ activates $\mathcal{G}_4$. $\mathcal{S}_4$ then goes on to repress $\mathcal{G}_1$.

3.4 Conclusion

We introduced biological units that commonly appear in biological networks, namely transcriptional regulation, covalent modification and phosphorylation-dephosphorylation. We also introduced modules that represent these functions while preserving the dynamic and parametric modularity properties. We characterized these modules by their IOSCFs and LTFs, and demonstrated how these properties are useful in determining the equilibrium points and their stability for a range of interconnection structures. We further characterized these modules by their cooperativity and positivity, which are important in deriving the results that will be seen in Chapter 4.

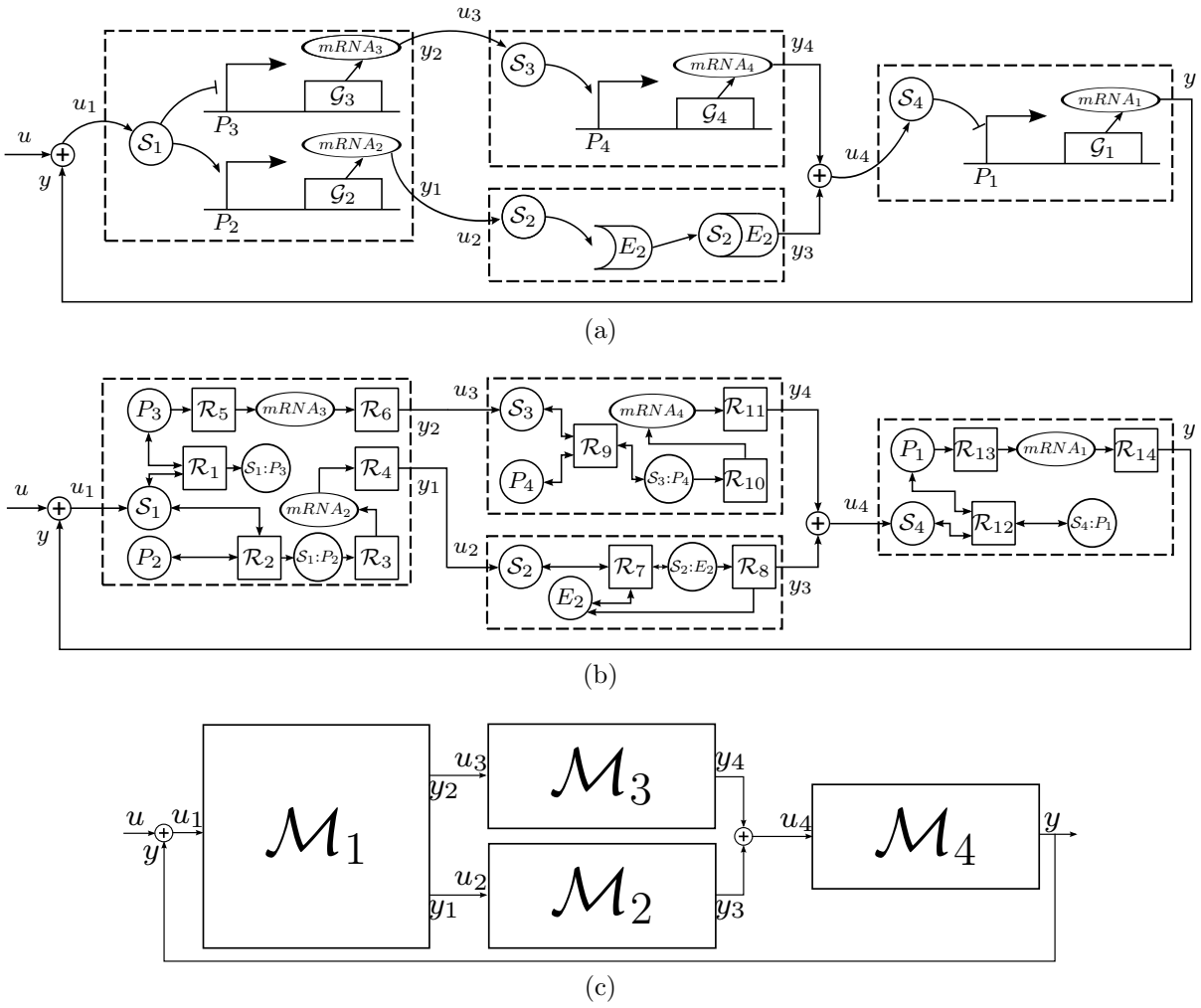


Figure 3.8: (a) Biological representation of the nested interconnection structure and (b) corresponding DBSR graph partition. (c) Block diagram representation of network decomposition, which shows $\mathcal{M}_1$ in cascade with the parallel interconnection of $\mathcal{M}_2$ and $\mathcal{M}_3$, connected in cascade with $\mathcal{M}_4$ and in turn, connected in a feedback loop.

Chapter 4

A modular study of two cyclic biological circuits

Some interesting classes of biological networks that are commonly studied are gene regulatory networks and enzymatic networks, both of which play important roles in many vital life processes [90] and are also well-characterized mathematically [65, 112]. A class of gene regulatory networks that has been the subject of many theoretical studies is known as the repressilator [42], which was the first known synthetic oscillator to be built [40]. Repressilators are topologically equivalent to a class of cyclic negative gain networks [40], and numerical analysis of the model revealed the parameter regions for which the network would oscillate and when it would converge to a stable steady-state [42]. Since then, there has been a plethora of work on theoretically establishing oscillatory regions of the network, for varying degrees of generalizations of the network model [40, 57, 88, 115]. While the results of these studies provide deep insights about when the repressilator oscillates, these results are often complex and require the satisfaction of multiple assumptions.

The Goldbeter-Koshland model for covalent modification of proteins [49] is another

enzymatic network that has been the subject of many studies [129, 78]. In this network, there exists a protein X_1 that is modified to another form X_2 through an interaction with an enzyme E_1. X_2 is also modified back to X_1 through an interaction with an enzyme E_2. The dynamical model describing this behavior is simple, and yet reveals interesting results about sensitivity of the protein concentrations to changes in enzyme concentration. Typically, these covalent modification reactions happen in signaling cascades [94], and the enzymes and substrates are subject to degradation.

In this chapter, we study a generalized repressilator which consists of a cyclic interconnection of an arbitrary number of transcriptional repressor modules. For the symmetric case, where all parameters across all the modules have the same value (which is typical in an experimental setting [42]), we provide results for when the network will converge to a stable steady-state, both in the local and global sense. Our results show how the stable parameter region for the repressilator becomes smaller as more transcriptional repressor modules are added to the network. These results were inspired by studies by Arcak and Sontag [12, 11].

We further study a generalized covalent modification network, consisting of a cascade of an arbitrary number of covalent modification modules connected in feedback. We show that regardless of the parameters chosen for each enzyme-substrate interaction, the substrates will necessarily degrade away, in spite of being connected in positive feedback. These results were inspired by the work on Monotone Systems Theory by Angeli and Sontag [9, 10, 8].

4.1 Generalized Repressilator Network

The Repressilator is a synthetic network designed to gain insight into the behavior of biological oscillators [42]. In its generalized form, this network consists of an odd number

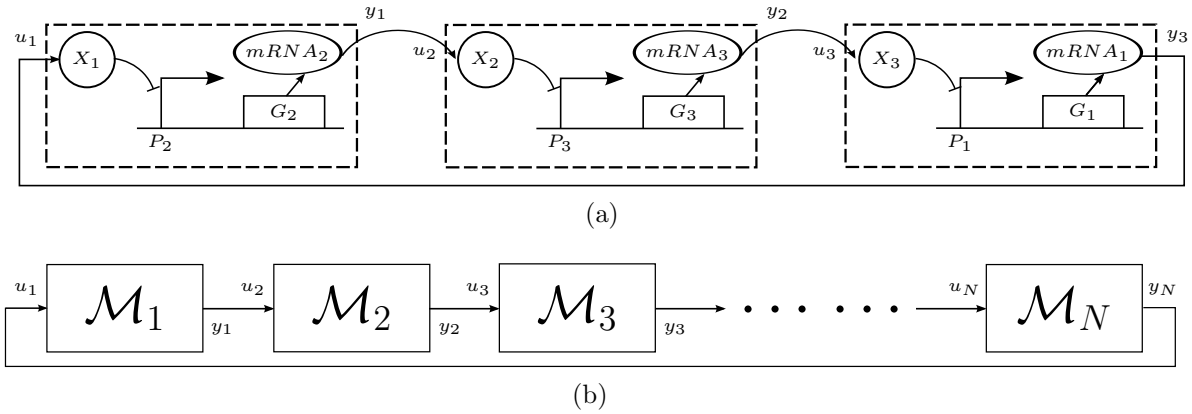


Figure 4.1: (a) Biological realization of a three gene Repressilator network. In experiment, this was done with three naturally-occurring repressor proteins $LacI$, $tetR$ and cI, each corresponding to a TR repressor module. (b) Cascade of N TR repressor modules connected in feedback, where each module is denoted by $\mathcal{M}_i$, $i \in \{1, \cdots, N\}$.

N of repressor proteins $\mathcal{S}_1, \cdots, \mathcal{S}_N$, where $\mathcal{S}_i$ represses the gene $\mathcal{G}_{i+1}$, for $i \in \{1, \cdots, N-1\}$ and $\mathcal{S}_N$ represses the gene $\mathcal{G}_1$. These networks can be decomposed into a cascade of N single-gene TR repressor modules connected in a (negative) feedback loop with no exogenous input. The biological realization of this decomposition with $N = 3$ (as in [42]) is shown in Figure 4.1a, with the corresponding block diagram representation depicted in Figure 4.1b.

The equilibrium point of the network must satisfy the equation

$$f_N\big(\cdots f_2\big(f_1(u_1^*)\big)\cdots\big) = u_1^*, \tag{4.1}$$

where $f_i(\cdot)$ denotes the IOSCF of the ith TR module (see Section 3.2.1). Since each $f_i(\cdot)$ is monotone decreasing, the composition of the N (odd) functions is also monotone decreasing and we have a feedback interconnection with a unique solution u_1^* to (4.1) *for all values of the parameters.* For simplicity, in the remainder of this section we assume that the parameters of the chemical reactions within each TR repressor module are exactly the same, implying that each module is identical and the network is *symmetric.*

However, it is straightforward to modify the results below for networks consisting of non-identical modules.

For this example, we follow two alternative approaches to determine whether or not the concentrations of the species converge to the unique equilibrium. The first approach is based on the *Nyquist Stability Criterion* [38] and will allow us to determine whether trajectories that start close to the equilibrium eventually converge to it.

4.1.1 Nyquist-criterion based approach to determine necessary and sufficient conditions for LAS

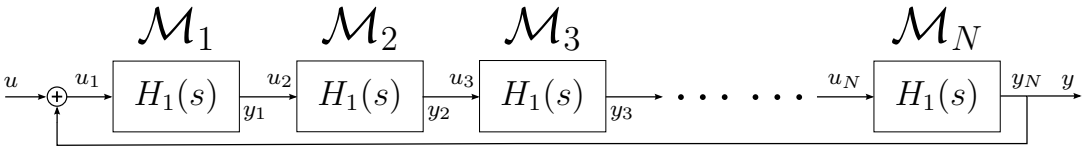


Figure 4.2: Cascade of N modules, each with an equal LTF $H_1(s)$, connected in feedback. Lemma 4 provides conditions for the BIBO stability of the linearized network, from a small perturbation in u to a small perturbation in y.

Lemma 4 *Consider the feedback interconnection of a cascade of N modules depicted in Figure 4.2, all with the same LTF $H_1(s)$ around a given equilibrium point of the feedback interconnection. Then the LTF of the feedback connection is BIBO stable if and only if*

$$\#OUP = -\frac{1}{N}\sum_{\ell=1}^{N} \#END[e^{j\frac{2\pi\ell}{N}}],$$

where $\#OUP$ represents the number of (open-loop) unstable poles of $H_1(s)$ and $\#END[e^{j\frac{2\pi\ell}{N}}]$ denotes the number of clockwise encirclements of the Nyquist contour of $H_1(j\omega)$, $\omega \in \mathbb{R}$ around the point $e^{j\frac{2\pi\ell}{N}}$ on the complex plane.[1] □

[1]We assume here that $H_1(s)$ has no poles on the imaginary axis. If this were the case, the standard "trick" of considering an infinitesimally perturbed system with the poles moved off the axis can be applied [38].

Proof of Lemma 4. To investigate the BIBO stability of the LTF of the network, we consider the characteristic equation of the feedback loop: $1 - H_1(s)^N = 0$. The number of unstable poles is thus given by the unstable solutions to the equation:

$$1 - H_1(s)^N = 0 \quad \Leftrightarrow \quad \exists i \in \{1, 2, \cdots, n\},\ H_1(s) = z_i,$$

where $z_\ell \coloneqq e^{j\frac{2\pi\ell}{N}}$ are the N roots to the equation $z^N = 1$.
To count the number of unstable poles of the network, we must then add the number of unstable poles of each of the N equations

$$H_1(s) = z_\ell, \quad \ell \in \{1, 2, \cdots, N\},$$

which can be done using Cauchy's argument principle by counting the number of clockwise encirclements of the point $z_\ell \in \mathbb{C}$ for the Nyquist contour of $H_1(j\omega)$, $\omega \in \mathbb{R}$. ■

Theorem 1 *Consider a Repressilator network, that consists of an odd number N of equal single-gene TR repressor modules ($F = 1$) connected in feedback as in Figure 4.1b. The network has a unique equilibrium point that is locally asymptotically stable if and only if*

$$\sum_{\ell=1}^{N} \#END[e^{j\frac{2\pi\ell}{N}}] = 0, \tag{4.2}$$

where $\#END[e^{j\frac{2\pi\ell}{N}}]$ denotes the number of clockwise encirclements of the Nyquist plot of the LTF of a single TR repressor module around the point $e^{j\frac{2\pi\ell}{N}}$. □

Proof of Theorem 1. We assume that Assumption 2 is satisfied for simplicity, although it is straightforward to extend the proof for the case when it is not. Theorem 1 follows from Lemma 4 by recognizing that a single-gene repressor TR module can be represented

by the LTF

$$H_1(s) = -\frac{qKP^{\text{tot}}\alpha_1\beta_1(\frac{u^*}{\bar{\beta}})^{q-1}}{(K+(\frac{u^*}{\bar{\beta}})^q)^2(s+\gamma_1)(s+\bar{\beta})} \tag{4.3}$$

under Assumption 2, as was shown in Section 3.2.1. Moreover, since the network input $u(t) = 0\ \forall t \geqslant 0$ and all modules have identical parameters, the values of the inputs and outputs of each module at equilibrium will be the same, each given by the unique solution u^* to (4.5). Therefore all modules have equal LTFs, and this enables us to use Lemma 4 to analyze when the LTF of the Repressilator is BIBO stable.

To complete the proof, we need to show that if the LTF of the Repressilator network is BIBO stable, then the equilibrium point is locally asymptotically stable. To do this, we need to show that the realization of the linearized network is minimal [55]. Defining the state of the network to be

$$x = \begin{bmatrix} [\mathcal{S}_1] \\ [mRNA_1] \\ [\mathcal{S}_2] \\ [mRNA_2] \\ \vdots \\ \vdots \\ [\mathcal{S}_N] \\ [mRNA_N] \end{bmatrix},$$

the state-space realization of the linearized closed-loop Repressilator network is given by

$$\dot{\delta x} = A\delta x + B\delta u \qquad \delta y = C\delta x$$

where

$$A_{2N\times 2N} = \begin{bmatrix} -\bar{\beta} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \beta_1 \\ P & -\gamma_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta_1 & -\bar{\beta} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & P & -\gamma_1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \beta_1 & -\bar{\beta} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & P & -\gamma_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \beta_1 & -\bar{\beta} & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & P & -\gamma \end{bmatrix}$$

$$B_{2N\times 1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad C_{1\times 2N} = \begin{bmatrix} 0 & 0 & \cdots & 0 & \bar{\beta} \end{bmatrix} \qquad P = -\frac{q\alpha_1 P^{\text{tot}} K([\mathcal{S}_i]^*)^{q-1}}{(K + ([\mathcal{S}_i]^*)^q)^2} \quad \forall i.$$

Both the controllability and observability matrices of this system have full rank and therefore the realization is minimal [55]. In this case, BIBO stability of the LTF implies that the realization is exponentially stable and therefore the equilibrium is locally asymptotically stable. □

The symmetric repressilator we analyze has a convenient property that the equilibrium point of the network, and hence the Nyquist plot of the LTF of a TR repressor module, remains the same regardless of the number of modules added to the network. Because of this, the greater the number of TR repressor modules in the network, the larger the parameter range over which the repressilator will be unstable. This is demonstrated in Figure 4.3, which shows the Nyquist plot for $\beta = 10^{-1.5}$ to be stable when there are only three TR repressor modules, but becomes unstable when there are five TR repressor

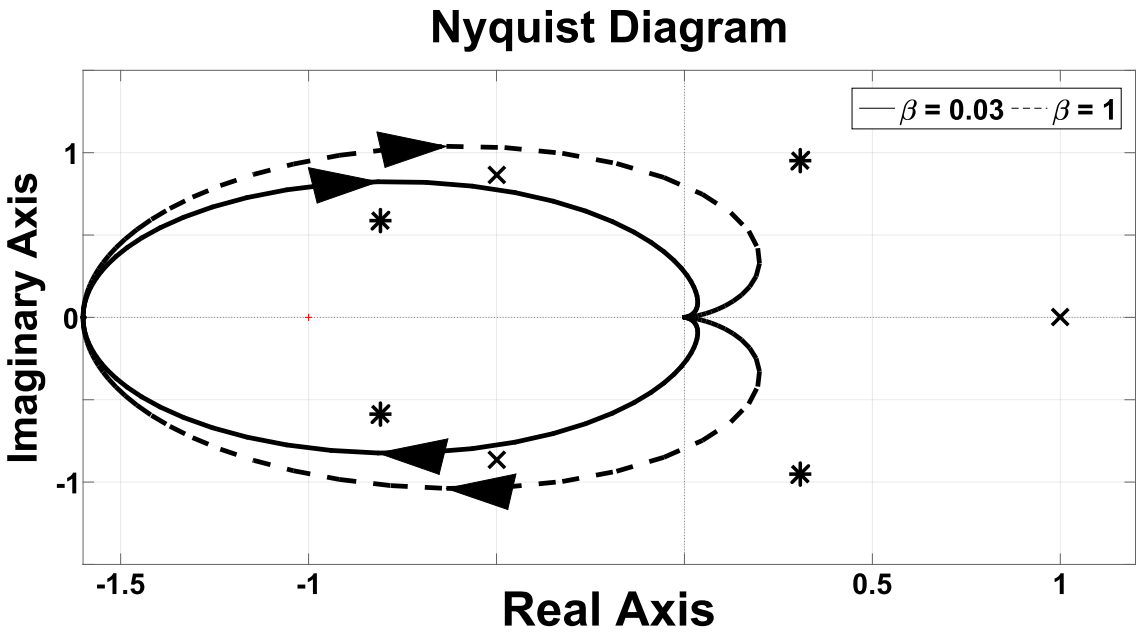


Figure 4.3: Nyquist plots of the LTF of a TR repressor module which satisfies Assumption 2 for two different sets of parameters: $\beta = \bar{\beta} = 10^{-1.5}$ (solid) and $\beta = \bar{\beta} = 1$ (dashed). The remaining parameters are the same for both plots: $Ptot = 1$, $K = 100$, $\gamma = 1$, $q = 2$, $\alpha = 100$. The points $e^{j\frac{2\pi\ell}{3}}$, $j \in \{1, 2, 3\}$ that appear in the criterion (4.2) are marked with "X", and the points $e^{j\frac{2\pi\ell}{5}}$, $j \in \{1, \cdots, 5\}$ are marked with "*". For the network with $N = 3$, the solid plot does not encircle any of the three points so we have BIBO stability for the feedback LTF, whereas for the dashed Nyquist plot this is not the case. When N becomes 5, the repressilator represented by the solid plot becomes unstable

modules in the network. As more TR repressor modules are added to a network, there will be a wider parameter region for which the Nyquist plot which encircles the points $e^{j\frac{2\pi l}{N}}$ $\forall l \in \{1, 2, \cdots, N\}$.

4.1.2 Secant-criterion based approach to determine sufficient conditions for LAS and GAS

An alternative approach that can be used to determine global convergence to the equilibrium of a negative feedback interconnection is based on the *Secant Criterion* [119, 117] and some of its more recent variations [12].

Theorem 2 *Consider a Repressilator network, that consists of an odd number N of equal single-gene TR repressor modules ($F = 1$) connected in feedback as in Figure 4.1b, with $q = 2$. Under Assumption 2 for the TR repressor modules, the network has a unique equilibrium point that is LAS if*

$$\frac{2P^{\text{tot}}\alpha\beta\frac{u^*}{\bar{\beta}}}{\bar{\beta}K\gamma\left(1+\frac{1}{K}(\frac{u^*}{\bar{\beta}})^2\right)^2} < \sec\left(\frac{\pi}{2N}\right)^2, \tag{4.4}$$

where u^* is the unique solution to

$$\frac{\alpha\beta P^{\text{tot}}}{\gamma\left(1+\frac{(\frac{u^*}{\bar{\beta}})^2}{K}\right)} = u^*, \tag{4.5}$$

and is GAS if

$$\frac{3P^{\text{tot}}\alpha\beta}{8\bar{\beta}\gamma}\sqrt{\frac{3}{K}} < \sec\left(\frac{\pi}{2N}\right)^2. \tag{4.6}$$

Both of these conditions are sufficient but not necessary. (4.6) in particular provides a condition which guarantees global convergence of the repressilator to its steady-state. This result again shows for a symmetric repressilator that the stability regions of the network shrink as N grows. Figure 4.4 shows the stability regions for a 3-gene Repressilator as we vary two of the TR module parameters.

To prove Theorem 2, we first prove the following result based on the Secant Criterion [119, 117] and some of its more recent variations [12], by choosing an appropriate coordinate transformation.

Lemma 5 *Consider the feedback interconnection depicted in Figure 4.1b, with each of the N modules $\mathcal{M}_i$ of the form*

$$[\dot{\mathcal{S}}_i] = u_i - c_i([\mathcal{S}_i]), \qquad y_i = d_i([\mathcal{S}_i]), \qquad i \in \{1, 2, \ldots, N\}, \tag{4.7}$$

where

1. *the $c_i(\cdot)$ are continuous and monotone strictly increasing;*

2. *an odd number $M \leqslant N$ of the $d_i(\cdot)$ are continuous and monotone strictly decreasing, while the remaining $N - M$ of the $d_i(\cdot)$ are continuous and monotone strictly increasing.*

Then the IOSCF of the cascade is monotone strictly decreasing and the feedback interconnection has a unique equilibrium. This equilibrium is locally asymptotically stable provided that

$$\prod_{i=1}^{N} \left| \frac{\frac{\partial d_i(s_i)}{\partial s_i}\Big|_{s_i=[\mathcal{S}_i]^*}}{\frac{\partial c_i(s_i)}{\partial s_i}\Big|_{s_i=[\mathcal{S}_i]^*}} \right| < \sec\left(\frac{\pi}{N}\right)^N, \tag{4.8}$$

where $[\mathcal{S}_i]^*$ denotes the value of $[\mathcal{S}_i]$ at the equilibrium; and it is globally asymptotically stable if there exist constants $\phi_i > 0$, $i \in \{1, 2, \ldots, N\}$ for which

$$\prod_{i=1}^{N} \phi_i < \sec\left(\frac{\pi}{N}\right)^N,$$

where ϕ_i satisfies either one of the following two conditions:

1. $$\left| \frac{d_i(z_i) - d_i([\mathcal{S}_i]^*)}{c_i(z_i) - c_i([\mathcal{S}_i]^*)} \right| \leqslant \phi_i \quad \forall z_i \neq [\mathcal{S}_i]^* \tag{4.9}$$

2. $$\left| \frac{\partial d_i(z_i)}{\partial z_i} \right| \leqslant \phi_i \frac{\partial c_i(z_i)}{\partial z_i}, \quad \forall z_i \neq [\mathcal{S}_i]^*. \tag{4.10}$$

□

Proof of Lemma 5. The proof of this result relies on making a coordinate transformation to our original system to take it into a form that allows us to use the secant criterion [119, 117, 12] and results in [12].

The dynamics of the feedback interconnection under consideration can be written as

$$[\dot{\mathcal{S}_1}] = -c_1([\mathcal{S}_1]) + d_N([\mathcal{S}_N]), \tag{4.11a}$$

$$[\dot{\mathcal{S}_2}] = -c_2([\mathcal{S}_2]) + d_1([\mathcal{S}_1]) \tag{4.11b}$$

$$\vdots \tag{4.11c}$$

$$[\dot{\mathcal{S}_N}] = -c_N([\mathcal{S}_N]) + d_{N-1}([\mathcal{S}_{N-1}]). \tag{4.11d}$$

with an equilibrium state defined by concentrations $[\mathcal{S}_i]^*$ for which

$$c_1([\mathcal{S}_1]^*) = d_N([\mathcal{S}_N]^*), \tag{4.12a}$$

$$c_2([\mathcal{S}_2]^*) = d_1([\mathcal{S}_1]^*), \tag{4.12b}$$

$$\vdots \tag{4.12c}$$

$$c_N([\mathcal{S}_N]^*) = d_{N-1}([\mathcal{S}_{N-1}]^*). \tag{4.12d}$$

To verify that such an equilibrium exists and is unique, note that we have a feedback interconnection of a cascade of N systems, each with an IOSCF given by

$$f_i(u_i^*) = d_i\big(c_i^{-1}(u_i^*)\big), \quad \forall u_1^* \in \mathbb{R},$$

where c_i^{-1} denotes the inverse function of c_i, which is invertible and monotone strictly increasing since c_i is monotone strictly increasing. Therefore, $f_i(u_i^*)$ has the same (strict) monotonicity as d_i. Since an odd number M of the d_i are monotone strictly decreasing, an odd number of the f_i are also monotone strictly decreasing and therefore the composition

of all the f_i is monotone strictly decreasing. This shows that we have a negative feedback interconnection and thus a unique equilibrium.

Using (4.12), we can re-write (4.11) as

$$\begin{aligned}
[\dot{\mathcal{S}}_1] &= -c_1([\mathcal{S}_1]) + c_1([\mathcal{S}_1]^*) + d_N([\mathcal{S}_N]) - d_N([\mathcal{S}_N]^*),\\
[\dot{\mathcal{S}}_2] &= -c_2([\mathcal{S}_2]) + c_2([\mathcal{S}_2]^*) + d_1([\mathcal{S}_1]) - d_1([\mathcal{S}_1]^*),\\
&\vdots\\
[\dot{\mathcal{S}_N}] &= -c_N([\mathcal{S}_N]) + c_N([\mathcal{S}_N]^*) + d_{N-1}([\mathcal{S}_{N-1}]) - d_{N-1}([\mathcal{S}_{N-1}]^*).
\end{aligned}$$

Since we have an odd number $M \geqslant 1$ of functions d_i that are monotone strictly decreasing and there is perfect symmetry in the cycle (4.11), we shall assume without loss of generality that d_N is monotone strictly decreasing; if that were not the case we could simply shift the numbering of the modules appropriately.

We consider a coordinate transformation with

$$x_1 = [\mathcal{S}_1] - [\mathcal{S}_1]^*, \qquad x_N = [\mathcal{S}_N] - [\mathcal{S}_N]^*, \tag{4.13}$$

and the remaining x_i, $i \in \{2, 3, \ldots, N-1\}$ either given by

$$x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^* \tag{4.14}$$

or given by

$$x_i = -[\mathcal{S}_i] + [\mathcal{S}_i]^*; \tag{4.15}$$

each to be determined shortly. The coordinate transformation (4.13) leads to

$$\begin{aligned}\dot{x}_1 &= -c_1([\mathcal{S}_1]) + c_1([\mathcal{S}_1]^*) + d_N([\mathcal{S}_N]) - d_N([\mathcal{S}_N]^*) \\ &= -c_1([\mathcal{S}_1]^* + x_1) + c_1([\mathcal{S}_1]^*) + d_N([\mathcal{S}_N]^* + x_N) - d_N([\mathcal{S}_N]^*) \\ &= -a_1(x_1) - b_N(x_N)\end{aligned}$$

with

$$a_1(x_1) \coloneqq c_1([\mathcal{S}_1]^* + x_1) - c_1([\mathcal{S}_1]^*), \qquad b_N(x_N) \coloneqq -d_N([\mathcal{S}_N]^* + x_N) + d_N([\mathcal{S}_N]^*).$$

Note that because c_1 is monotone strictly increasing and d_N is monotone strictly decreasing, we have that

$$a_1(x_1) \quad \begin{cases} > 0 & x_1 > 0 \\ = 0 & x_1 = 0 \\ < 0 & x_1 < 0. \end{cases}, \qquad b_N(x_N) \quad \begin{cases} > 0 & x_N > 0 \\ = 0 & x_N = 0 \\ < 0 & x_N < 0. \end{cases} \tag{4.16}$$

For the remaining variables x_i, $i \in \{2, 3, \ldots, N\}$, the coordinate transformation leads to

$$
\begin{aligned}
\dot{x}_i = & \\
& \begin{cases}
c_i([\mathcal{S}_i]^*) - c_i([\mathcal{S}_i]) + d_{i-1}([\mathcal{S}_{i-1}]) - d_{i-1}([\mathcal{S}_{i-1}]^*) & \\
\qquad\qquad \text{if } x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^*\,,\ x_{i-1} = [\mathcal{S}_{i-1}] - [\mathcal{S}_{i-1}]^* & \\
\qquad\qquad \text{or } x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^*\,,\ x_{i-1} = [\mathcal{S}_{i-1}]^* - [\mathcal{S}_{i-1}] & \\
c_i([\mathcal{S}_i]) - c_i([\mathcal{S}_i]^*) - d_{i-1}([\mathcal{S}_{i-1}]) + d_{i-1}([\mathcal{S}_{i-1}]^*) & \\
\qquad\qquad \text{if } x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i]\,,\ x_{i-1} = [\mathcal{S}_{i-1}] - [\mathcal{S}_{i-1}]^* & \\
\qquad\qquad \text{or } x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i]\,,\ x_{i-1} = [\mathcal{S}_{i-1}]^* - [\mathcal{S}_{i-1}] &
\end{cases} \\
& = -a_i(x_i) + b_{i-1}(x_{i-1})
\end{aligned}
$$

where, for every $i \in \{2, 3, \dots, N\}$,

$$a_i(x_i) \coloneqq \begin{cases} c_i([\mathcal{S}_i]^* + x_i) - c_i([\mathcal{S}_i]^*) & x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^* \\ -c_i([\mathcal{S}_i]^* - x_i) + c_i([\mathcal{S}_i]^*) & x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i] \end{cases}$$

$$b_{i-1}(x_{i-1}) \coloneqq$$

$$\begin{cases} d_{i-1}([\mathcal{S}_{i-1}]^* + x_{i-1}) - d_{i-1}([\mathcal{S}_{i-1}]^*) \\ \qquad\qquad \text{if } x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^*,\ x_{i-1} = [\mathcal{S}_{i-1}] - [\mathcal{S}_{i-1}]^* \\ d_{i-1}([\mathcal{S}_{i-1}]^* - x_{i-1}) - d_{i-1}([\mathcal{S}_{i-1}]^*) \\ \qquad\qquad \text{if } x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^*,\ x_{i-1} = [\mathcal{S}_{i-1}]^* - [\mathcal{S}_{i-1}] \\ -d_{i-1}([\mathcal{S}_{i-1}]^* + x_{i-1}) + d_{i-1}([\mathcal{S}_{i-1}]^*) \\ \qquad\qquad \text{if } x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i],\ x_{i-1} = [\mathcal{S}_{i-1}] - [\mathcal{S}_{i-1}]^* \\ -d_{i-1}([\mathcal{S}_{i-1}]^* - x_{i-1}) + d_{i-1}([\mathcal{S}_{i-1}]^*) \\ \qquad\qquad \text{if } x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i],\ x_{i-1} = [\mathcal{S}_{i-1}]^* - [\mathcal{S}_{i-1}]. \end{cases}$$

Since all the c_i are monotone strictly increasing, we have that

$$a_i(x_i) \quad \begin{cases} > 0 & x_i > 0 \\ = 0 & x_i = 0 \\ < 0 & x_i < 0. \end{cases}, \qquad\qquad \forall i \in \{2, 3, \dots, N\}.$$

We have already selected x_1 and x_N according to (4.13) to obtain (4.16). Our goal is now to select the remaining x_i, $i \in \{2, 3, \dots, N-1\}$ according to (4.14) or (4.15) so that

we also have

$$b_{i-1}(x_{i-1}) \quad \begin{cases} > 0 & x_{i-1} > 0 \\ = 0 & x_{i-1} = 0 \\ < 0 & x_{i-1} < 0, \end{cases} \qquad \forall i \in \{2, 3, \ldots, N\}.$$

which would require us to have d_{i-1} monotone strictly increasing if either

$$x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^*\,,\ x_{i-1} = [\mathcal{S}_{i-1}] - [\mathcal{S}_{i-1}]^*$$

or

$$x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i]\,,\ x_{i-1} = [\mathcal{S}_{i-1}]^* - [\mathcal{S}_{i-1}]$$

and d_{i-1} monotone strictly decreasing if either

$$x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^*\,,\ x_{i-1} = [\mathcal{S}_{i-1}]^* - [\mathcal{S}_{i-1}]$$

or

$$x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i]\,,\ x_{i-1} = [\mathcal{S}_{i-1}] - [\mathcal{S}_{i-1}]^*\,,$$

$\forall i \in \{2, 3, \ldots, N\}$. It turns out that this is always possible because there is an even number of the d_{i-1} with $i \in \{2, 3, \ldots, N\}$ that are monotone strictly decreasing (recall that d_N is monotone strictly decreasing and there are in total an odd number of d_i that are monotone strictly decreasing). All we need to do is to start with x_1 as in (4.13) and alternate between (4.14) and (4.15) each time d_{i-1} is monotone strictly decreasing. Since there is an even number of the d_{i-1} with $i \in \{2, 3, \ldots, N\}$, we will end up with x_N as in (4.13).

The coordinate transformation constructed above, leads us to a system of the following

form

$$\dot{x}_1 = -a_1(x_1) - b_N(x_N) \tag{4.17a}$$

$$\dot{x}_2 = -a_2(x_2) + b_1(x_1) \tag{4.17b}$$

$$\vdots \tag{4.17c}$$

$$\dot{x}_N = -a_N(x_N) + b_{N-1}(x_{N-1}) \tag{4.17d}$$

To prove the local stability result, we apply the secant criterion [119, 117, 12] to the local linearization of this system around the equilibrium $x_i = 0$, $\forall i$, which has a Jacobian matrix of the form

$$\begin{bmatrix} -\frac{\partial a_1(x_1)}{\partial x_1}\Big|_{x_1=0} & 0 & \cdots & 0 & -\frac{\partial b_N(x_N)}{\partial x_N}\Big|_{x_N=0} \\ \frac{\partial b_1(x_1)}{\partial x_1}\Big|_{x_1=0} & -\frac{\partial a_2(x_2)}{\partial x_2}\Big|_{x_2=0} & \ddots & & 0 \\ 0 & \frac{\partial b_2(x_2)}{\partial x_2}\Big|_{x_2=0} & -\frac{\partial a_3(x_3)}{\partial x_3}\Big|_{x_3=0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\partial b_{N-1}(x_{N-1})}{\partial x_{N-1}}\Big|_{x_{N-1}=0} & -\frac{\partial a_N(x_N)}{\partial x_N}\Big|_{x_N=0} \end{bmatrix}, \tag{4.18}$$

where

$$\frac{\partial a_i(x_i)}{\partial x_i}\Big|_{x_i=0} = \frac{\partial c_i(s_i)}{\partial s_i}\Big|_{s_i=[\mathcal{S}_i]^*} > 0$$

$$\frac{\partial b_i(x_i)}{\partial x_i}\Big|_{x_i=0} = \begin{cases} \frac{\partial d_i(s_i)}{\partial s_i}\Big|_{s_i=[\mathcal{S}_i]^*} > 0 & \text{if } d_i \text{ monotone increasing} \\ -\frac{\partial d_i(s_i)}{\partial s_i}\Big|_{s_i=[\mathcal{S}_i]^*} > 0 & \text{if } d_i \text{ monotone decreasing,} \end{cases}$$

$\forall i \in \{1, 2, \ldots, N\}$. This matrix matches precisely the one considered in the secant criteria, which states that the Jacobian matrix (4.18) is Hurwitz if (4.8) holds.

For the global asymptotic stability result we use [12, Corollary 3], which applies precisely to systems of the form (4.17) with

$$x_i a_i(x_i) > 0, \qquad x_i b_i(x_i) > 0, \qquad \forall x_i \neq 0, i \in \{1, 2, \ldots, N\}.$$

Two additional conditions are needed by [12, Corollary 3]:

$$\lim_{|x_i| \to \infty} \int_0^{x_i} b_i(\sigma) d\sigma = \infty. \tag{4.19}$$

and there must exist $\phi_i > 0$, $\forall i \in \{1, 2, \ldots, N\}$ for which

$$\frac{b_i(x_i)}{a_i(x_i)} \leqslant \phi_i, \quad \forall i,\ x_i \neq 0, \tag{4.20a}$$

$$\prod_{i=1}^{N} \phi_i < \sec\left(\frac{\pi}{N}\right)^N. \tag{4.20b}$$

The first condition (4.19) holds because our functions b_i are all zero at zero and monotone strictly increasing.

We then prove two conditions to be sufficient for (4.20a) to be satisfied. Before proceeding, we make the following observations:

$$\frac{b_i(x_i)}{a_i(x_i)} \coloneqq \begin{cases} \frac{d_i([\mathcal{S}_i]^* + x_i) - d_i([\mathcal{S}_i]^*)}{c_i([\mathcal{S}_i]^* + x_i) - c_i([\mathcal{S}_i]^*)} \quad \text{or} \quad \frac{-d_i([\mathcal{S}_i]^* + x_i) + d_i([\mathcal{S}_i]^*)}{c_i([\mathcal{S}_i]^* + x_i) - c_i([\mathcal{S}_i]^*)} & \text{if } x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^* \\ \\ \frac{d_i([\mathcal{S}_i]^* - x_i) - d_i([\mathcal{S}_i]^*)}{-c_i([\mathcal{S}_i]^* - x_i) + c_i([\mathcal{S}_i]^*)} \quad \text{or} \quad \frac{-d_i([\mathcal{S}_i]^* - x_i) + d_i([\mathcal{S}_i]^*)}{-c_i([\mathcal{S}_i]^* - x_i) + c_i([\mathcal{S}_i]^*)} & \text{if } x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i]\,. \end{cases} \tag{4.21}$$

$$\frac{\partial a_i(x_i)}{\partial x_i} \coloneqq \begin{cases} \frac{\partial}{\partial x_i} c_i([\mathcal{S}_i]^* + x_i) & \text{if } x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^* \\ -\frac{\partial}{\partial x_i} c_i([\mathcal{S}_i]^* - x_i) & \text{if } x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i]\,. \end{cases} \tag{4.22}$$

$$\frac{\partial b_i(x_i)}{\partial x_i} \coloneqq \begin{cases} \frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* + x_i) \quad \text{or} \quad -\frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* + x_i) & \text{if } x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^* \\ \frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* - x_i) \quad \text{or} \quad -\frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* - x_i) & \text{if } x_i = [\mathcal{S}_i]^* - [\mathcal{S}_i]\,. \end{cases} \tag{4.23}$$

First, we prove that condition (4.9) implies (4.20a). We see that (4.9) implies that

$$\frac{d_i(z_i) - d_i([\mathcal{S}_i]^*)}{c_i(z_i) - c_i([\mathcal{S}_i]^*)} \leqslant \phi_i \quad \text{and} \quad -\frac{d_i(z_i) - d_i([\mathcal{S}_i]^*)}{c_i(z_i) - c_i([\mathcal{S}_i]^*)} \leqslant \phi_i \qquad \forall z_i \neq [\mathcal{S}_i]^*\,. \tag{4.24}$$

With the change of co-ordinates $x_i = -[\mathcal{S}_i]^* + z_i$ and $x_i = [\mathcal{S}_i]^* - z_i$, we see that (4.24) implies that

$$\begin{aligned} &\frac{d_i([\mathcal{S}_i]^* + x_i) - d_i([\mathcal{S}_i]^*)}{c_i([\mathcal{S}_i]^* + x_i) - c_i([\mathcal{S}_i]^*)} \leqslant \phi_i\,, \quad &&\frac{-d_i([\mathcal{S}_i]^* + x_i) + d_i([\mathcal{S}_i]^*)}{c_i([\mathcal{S}_i]^* + x_i) - c_i([\mathcal{S}_i]^*)} \leqslant \phi_i \\ &\frac{d_i([\mathcal{S}_i]^* - x_i) - d_i([\mathcal{S}_i]^*)}{-c_i([\mathcal{S}_i]^* - x_i) + c_i([\mathcal{S}_i]^*)} \leqslant \phi_i\,, \quad &&\frac{-d_i([\mathcal{S}_i]^* - x_i) + d_i([\mathcal{S}_i]^*)}{-c_i([\mathcal{S}_i]^* - x_i) + c_i([\mathcal{S}_i]^*)} \leqslant \phi_i \qquad \forall x_i \neq 0. \end{aligned} \tag{4.25}$$

From (4.21), we conclude that (4.25) implies (4.20a).

We then prove that the condition (4.10) implies (4.20a). We see that (4.10) implies that

$$\frac{\partial d_i(z_i)}{\partial z_i} \leqslant \phi_i \frac{\partial c_i(z_i)}{\partial z_i} \quad \text{and} \quad -\frac{\partial d_i(z_i)}{\partial z_i} \leqslant \phi_i \frac{\partial c_i(z_i)}{\partial z_i} \qquad \forall z_i \neq [\mathcal{S}_i]^* \tag{4.26}$$

With the change of co-ordinates $x_i = -[\mathcal{S}_i]^* + z_i$ and $x_i = [\mathcal{S}_i]^* - z_i$, (4.26) can be seen

to imply that

$$\begin{aligned}
\frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* + x_i) &\leqslant \phi_i \frac{\partial}{\partial x_i} c_i([\mathcal{S}_i]^* + x_i) \\
-\frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* + x_i) &\leqslant \phi_i \frac{\partial}{\partial x_i} c_i([\mathcal{S}_i]^* + x_i) \\
\frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* - x_i) &\leqslant -\phi_i \frac{\partial}{\partial x_i} c_i([\mathcal{S}_i]^* - x_i) \\
-\frac{\partial}{\partial x_i} d_i([\mathcal{S}_i]^* - x_i) &\leqslant -\phi_i \frac{\partial}{\partial x_i} c_i([\mathcal{S}_i]^* - x_i) \quad \forall x_i \neq 0
\end{aligned} \tag{4.27}$$

From (4.22)–(4.23), we can observe that (4.27) implies that

$$\frac{\partial b_i(x_i)}{\partial x_i} < \phi_i \frac{\partial a_i(x_i)}{\partial x_i} \quad \forall x_i \neq 0 \tag{4.28}$$

Let $h_i(x_i) \coloneqq b_i(x_i) - \phi_i a_i(x_i)$. Then, (4.28) implies that

$$\frac{\partial h_i(x_i)}{\partial x_i} \leqslant 0 \quad \forall x_i \neq 0. \tag{4.29}$$

Since $a_i(0) = 0$ and $b_i(0) = 0$, we know that $h_i(0) = 0$. Therefore (4.29) implies that

$$h_i(x_i) \begin{cases} \leqslant 0 & \forall x_i > 0 \\ \geqslant 0 & \forall x_i < 0, \end{cases}$$

which further implies that

$$b_i(x_i) \begin{cases} \leqslant \phi_i a_i(x_i) & \forall x_i > 0 \\ \geqslant \phi_i a_i(x_i) & \forall x_i < 0 \end{cases} \tag{4.30}$$

Since

$$a_i(x_i) \quad \begin{cases} > 0 & x_i > 0 \\ < 0 & x_i < 0 \end{cases}$$

(4.30) implies (4.20a), hence completing our proof. ■

We can then apply Lemma 5 to prove Theorem 2.

Proof of Theorem 2. Each SISO repressor module can be further decomposed into two modules that can be described by the equations

$$\begin{aligned} [\dot{\mathcal{S}_0}] &= u_1 - \bar{\beta}\,[\mathcal{S}_0] \\ y_1 \coloneqq h([\mathcal{S}_0]) &= \begin{cases} \frac{\alpha P^{\text{tot}}}{1+\frac{1}{K}[\mathcal{S}_0]^2} & \text{if } [\mathcal{S}_0] \geqslant 0 \\ \alpha P^{\text{tot}} \frac{1+\frac{2}{K}[\mathcal{S}_0]^2}{1+\frac{1}{K}[\mathcal{S}_0]^2} & \text{if } [\mathcal{S}_0] < 0. \end{cases} \end{aligned}$$

and

$$\begin{aligned} [\dot{mRNA_1}] &= u_2 - \gamma\,[mRNA_1] \\ y_2 &= \beta\,[mRNA_1] \end{aligned}$$

respectively. It can be seen that a small modification has been made to y_1, the repressing output from the TR repressor module. Since our network is positive, this modification has no effect on the network behavior. However, this change makes it more straightforward to apply Lemma 5 to analyze this network, since the theorem relied on each output function being monotone strictly increasing or monotone strictly decreasing $\forall\,[\mathcal{S}_0]$.

From the (4.8) in Lemma 5, (4.4)–(4.5) guarantee that the equilibrium point of the Repressilator network will be LAS.

For GAS, we first need to pick ϕ_1 and ϕ_2 to satisfy

$$\left|\frac{\partial h(z)}{\partial z}\right| \leqslant \phi_1 \bar{\beta} \quad z \neq [\mathcal{S}_0]^*$$
$$\beta \leqslant \phi_2 \gamma.$$

It is straightforward to show that

$$\max_{z \neq [\mathcal{S}_0]^*} \left|\frac{\partial h(z)}{\partial z}\right| = \frac{3\alpha P^{\text{tot}}}{8}\sqrt{\frac{3}{K}},$$

so we can pick

$$\phi_1 = \frac{3\alpha P^{\text{tot}}}{8\bar{\beta}}\sqrt{\frac{3}{K}}, \quad \phi_2 = \frac{\beta}{\gamma}.$$

From the proof of Lemma 5, we can infer that (4.6) guarantees that the equilibrium point of the Repressilator network will be GAS. □

4.2 Generalized Covalent Modification Network

We derive a result for the covalent modification network shown in Figure 4.5. From a biological perspective, this result implies that regardless of the parameters chosen, substrates that are covalently modified in this cyclic manner will completely degrade. Despite the fact that each of the substrates is essentially "activating" the next substrate in the cascade, the degradation reactions dominate in this case.

Theorem 3 *Consider the covalent modification network shown in Figure 4.5, which consists of a cascade of N CM modules connected in (positive) feedback. The substrate concentrations in each module converges to* 0 *as* $t \to \infty$. □

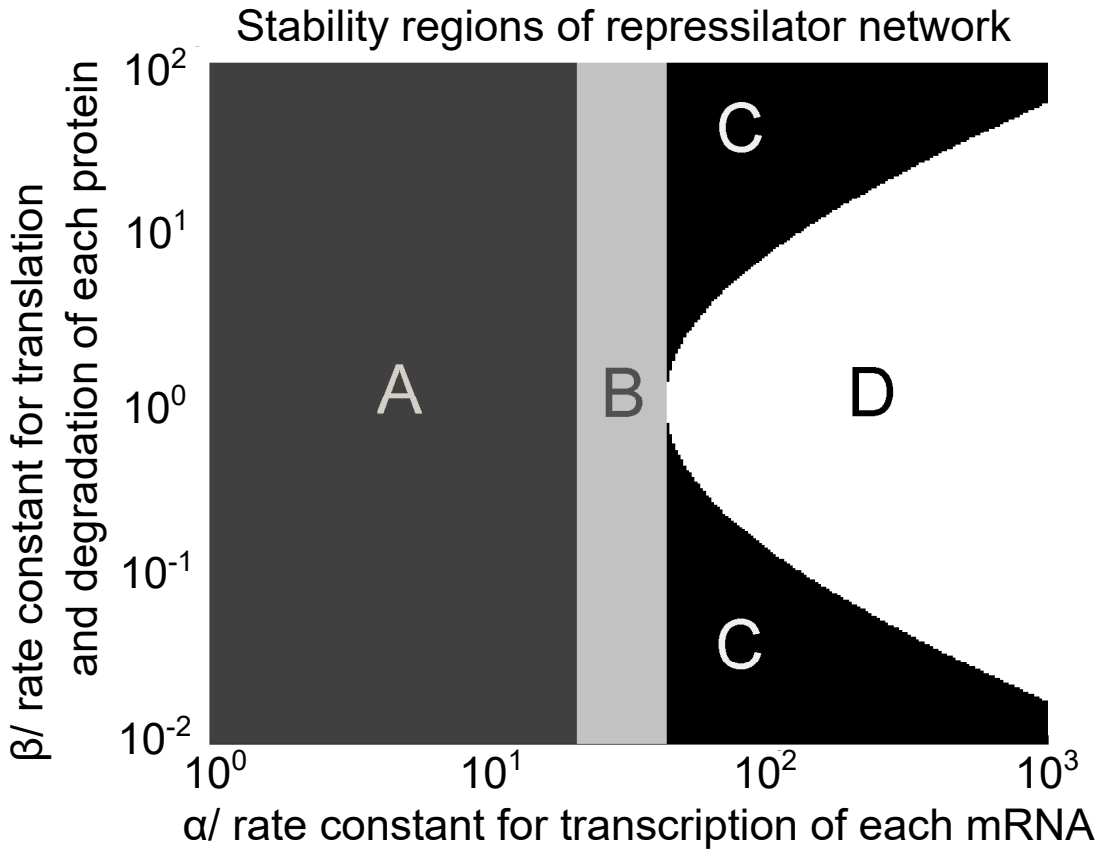


Figure 4.4: Stability regions for the 3-gene Repressilator, as a function of the parameters α and $\beta = \bar{\beta}$, under Assumption 2. The (sufficient) condition (4.6) allows us to conclude that the equilibrium is GAS in region "A;" the (sufficient) condition (4.4) allows us to conclude that the equilibrium is LAS in regions "A" and "B;" and the (necessary and sufficient) condition (4.2) allows us to conclude that the equilibrium is LAS in regions "A," "B" and "C;" and also that it is unstable in region "D.".

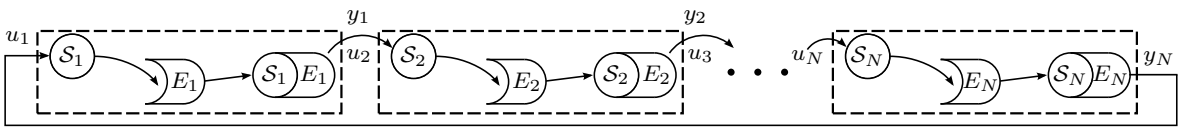


Figure 4.5: Biological representation of a Covalent Modification network decomposition.

To prove Theorem 3, we use the following result which is adapted from [10, Theorems 2-3], and provides conditions that can be used to establish the stability of the equilibrium points for positive feedback networks. To express the result, we need the following definitions which are closely related to that of cooperativity: We say that the system

$$\dot{x} = A(x,u), \qquad y = B(x), \qquad x \in \mathbb{R}^n, u \in \mathbb{R}^k, y \in \mathbb{R}^m, \tag{4.31}$$

is *excitable (with respect to the positive orthant)* if for every initial condition $x_0 \in \mathbb{R}^n$ and all inputs $u(t), \bar{u}(t) \in \mathbb{R}^k$, $\forall t \geqslant 0$, we have that

$$u(t) \succ \bar{u}(t),\ \forall t \geqslant 0 \quad \Rightarrow \quad x(t;x_0,u) \gg x(t;x_0,\bar{u}),\ \forall t > 0,$$

and it is *transparent (with respect to the positive orthant)* if for all initial conditions $x_0, \bar{x}_0 \in \mathbb{R}^n$ and every input $u(t) \in \mathbb{R}^k$, $\forall t \geqslant 0$, we have that

$$x_0 \succ \bar{x}_0 \quad \Rightarrow \quad x(t;x_0,u) \gg x(t;\bar{x}_0,u),\ \forall t > 0.$$

Given two vectors $v, \bar{v}$, we write $v \succ \bar{v}$ if every entry of v is larger than or equal to the corresponding entry of $\bar{v}$ and $v \neq \bar{v}$.

Lemma 6 *Consider the feedback interconnection of a SISO module $\mathcal{M}$ of the form* (4.31)*, whose output y is fed back to its input u. Assume that*

1. *$\mathcal{M}$ is excitable, transparent, and cooperative with a well-defined ISSCF and IOSCF;*
2. *for every constant input $u(t) = u^*$, $\forall t \geqslant 0$ to $\mathcal{M}$, the Jacobian matrix $\frac{\partial A(x,u)}{\partial x}$ is nonsingular at the corresponding equilibrium, which is globally asymptotically stable;*

3. *the IOSCF $f(u^*)$ of $\mathcal{M}$ has fixed points $\bar{u}^*$ for which $f_1(\bar{u}^*) = \bar{u}^*$, and the partial derivative $\frac{\partial f_1(u^*)}{\partial u^*}\Big|_{u^*=\bar{u}^*} \neq 1$;*

4. *all trajectories of the feedback interconnection are bounded.*

Then, for almost all initial conditions, the solutions converge to the set of points for which $f(u^*) = u^*$ and $\frac{\partial f(u^*)}{\partial u^*} < 1$. □

We now use Lemma 6 to prove Theorem 3.

Proof of Theorem 3. Every CM module $\mathcal{M}_i$ for $i \in \{1, \cdots, N\}$ is given by

$$\boldsymbol{\mathcal{M}_i} : [\dot{\mathcal{S}}_i] = A_i([\mathcal{S}_i], u_i), \qquad y_i = B_i([\mathcal{S}_i])$$

where

$$A_i([\mathcal{S}_i], u_i) = \frac{u_i - \gamma_i [\mathcal{S}_i] - \frac{k_i^{\mathrm{cat}} E_i^{\mathrm{tot}} [\mathcal{S}_i]}{K_i^{\mathrm{d}} + [\mathcal{S}_i]}}{1 + \frac{K_i^{\mathrm{d}} E_i^{\mathrm{tot}}}{(K_i^{\mathrm{d}} + [\mathcal{S}_i]^2)}} \qquad B_i([\mathcal{S}_i]) = \frac{k^{\mathrm{cat}} E_i^{\mathrm{tot}} [\mathcal{S}_i]}{K_i^{\mathrm{d}} + [\mathcal{S}_i]} \tag{4.32}$$

under Assumption 3, and

$$A_i([\mathcal{S}_i], u_i) = u_i - \gamma_i [\mathcal{S}_i] - \frac{k_i^{\mathrm{cat}} E_i^{\mathrm{tot}} [\mathcal{S}_i]}{K_i^{\mathrm{m}} + [\mathcal{S}_i]} \qquad B_i([\mathcal{S}_i]) = \frac{k_i^{\mathrm{cat}} E_i^{\mathrm{tot}} [\mathcal{S}_i]}{K_i^{\mathrm{m}} + [\mathcal{S}_i]} \tag{4.33}$$

under Assumption 4. The covalent modification system consists of the modules $\mathcal{M}_1$ and $\mathcal{M}_2$ interconnected with $y_1 = u_2$ and $y_2 = u_1$.

We first show that the cascade of two CM modules satisfies the properties in item 1 from Lemma 6. For the cascade network with $y_1 = u_2$, we have

$$\begin{aligned} &\frac{\partial A_1([\mathcal{S}_1], u_1)}{\partial u_1} = 1 \qquad \frac{\partial A_2([\mathcal{S}_2], u_2)}{\partial [\mathcal{S}_1]} > 0 \qquad \frac{\partial y_2}{\partial [\mathcal{S}_2]} > 0 \\ &\frac{\partial A_1([\mathcal{S}_1], u_1)}{\partial [\mathcal{S}_2]} = 0 \qquad \frac{\partial A_2([\mathcal{S}_2], u_2)}{\partial u_1} = 0 \qquad \frac{\partial y_2}{\partial [\mathcal{S}_1]} = 0, \end{aligned}$$

from which we can infer several properties: From [8, Propositon 1], we conclude that the cascade is cooperative, from [10, Theorem 4] that it is excitable, and from [10, Theorem 5] that it is transparent. The IOSCF of the cascade is well-defined and given by $f_2(f_1(u_1^*))$, where u_1^* denotes a constant input to the cascade and

$$f_i(u_i^*) \coloneqq y_i^* = \frac{1}{2}(u_i^* + k_i^{\mathrm{cat}} E_i^{\mathrm{tot}} + K_i^{\mathrm{d}}\gamma_i - \sqrt{(-u_i^* + k_i^{\mathrm{cat}} E_i^{\mathrm{tot}} + K_i^{\mathrm{d}}\gamma_i)^2 + 4K_i^{\mathrm{d}} u_i^* \gamma_i}),$$
$$i = \{1, 2\}.$$

The ISSCF of the cascade is also well-defined, and therefore, the cascade of two CM modules satisfies the properties in item 1 of Lemma 6.

We now show that the cascade of two CM modules satisfies the properties in item 2 from Lemma 6. For some constant input $u_1^* \geqslant 0$, the Jacobian matrix of the interconnection is given by the 2×2 lower triangular matrix

$$J = \begin{bmatrix} J_{11} & 0 \\ J_{21} & J_{22}, \end{bmatrix}$$

which is non-singular because $J_{11} > 0$ and $J_{22} > 0$, under the implicit assumption that all parameters within each module are positive.

Each of the modules $\mathcal{M}_1$ and $\mathcal{M}_2$ has an equilibrium point that is globally asymptotically stable for some constant input u_i^* into each module. This can be verified by doing the coordinate transformation $x_i = [\mathcal{S}_i] - [\mathcal{S}_i]^*$ and observing that the Lyapunov function $V(x_i) = x_i^2$ is zero-at-zero, locally positive definite and $\dot{V}(x_i) < 0$, $\forall x_i$, $i \in \{1, 2\}$. The cascade of both the modules also has a globally asymptotically stable equilibrium point, as can be seen from the argument in [116].

To verify that the cascade of the CM modules satisfies the property in item 3 from Lemma

6, we will show that $f_2(f_1(u_1^*)) = u_1^*$ has a unique solution at $u_1^* = 0$, and also that

$$\left.\frac{\partial f_2(f_1(u_1^*))}{\partial u_1^*}\right|_{u_1^*=0} < 1.$$

To show that $f_2(f_1(u_1^*)) = u_1^*$ has a unique solution at $u_1^* = 0$, we first observe that the IOSCFs of $\mathcal{M}_1$ and $\mathcal{M}_2$ are each monotone increasing and strictly concave, since $f_i'(u_i^*) > 0$ and $f_i''(u_i^*) < 0$, $\forall i \in \{1, 2\}$. Then from

$$\frac{\partial f_2\left(f_1\left(u_1^*\right)\right)}{\partial u_1^*} = f_1'(u_1^*) f_2'\left(f_1\left(u_1^*\right)\right) \tag{4.34}$$

$$\frac{\partial^2 f_2\left(f_1\left(u_1^*\right)\right)}{\partial (u_1^*)^2} = f_1''(u_1^*) f_2'\left(f_1\left(u_1^*\right)\right) + \left(f_1'\left(u_1^*\right)\right)^2 f_2''\left(f_1\left(u_1^*\right)\right), \tag{4.35}$$

it can be seen that the cascade of $\mathcal{M}_1$ and $\mathcal{M}_2$ is also monotone increasing and strictly concave because

1. $f_1'(u_1^*) > 0$ and $f_2'(u_2^*) > 0 \quad \forall u_1^*, u_2^* > 0 \implies \frac{\partial f_2\left(f_1\left(u_1^*\right)\right)}{\partial u_1^*} > 0$ from (4.34).
2. $f_1''(u_1^*) < 0$ and $f_2''(u_2^*) < 0 \quad \forall u_1^*, u_2^* > 0 \implies \frac{\partial^2 f_2\left(f_1\left(u_1^*\right)\right)}{\partial (u_1^*)^2} < 0$ from (4.35).

Therefore, there can be no other solution to $f_2(f_1(u_1^*)) = u_1^*$ other than $u_1^* = 0$. In addition, the derivative of the IOSCF of the cascade at this equilibrium is given by

$$\left.\frac{\partial f_2(f_1(u_1^*))}{\partial u_1^*}\right|_{u_1^*=0} = \frac{E_1^{\text{tot}} E_2^{\text{tot}} k_1^{\text{cat}} k_2^{\text{cat}}}{(E_1^{\text{tot}} k_1^{\text{cat}} + K_1^d \gamma_1)(E_2^{\text{tot}} k_2^{\text{cat}} + K_2^d \gamma_2)},$$

which is always less than 1 since all parameters are positive by definition.

The boundedness property in item 4 of Lemma 6 follows from techniques used to analyze MAK ODEs from [6, Main Technical Lemma], which can be applied to the covalent modification network.

This argument can be extended over each of the N modules to complete the proof. □

4.3 Conclusion

In this chapter, we used the biological modules described in Chapter 3 as the building blocks to analyze larger biological networks inspired from the literature. We developed insights into how adding more repressors to the repressilator tends to shrink the region of parameters over which the repressilator is stable. Furthermore, we discovered that despite being connected in positive feedback, the substrates in a generalized covalent modification network will always degrade.

Chapter 5

Future Work

In Part I of this monograph, we introduced a method to decompose any biochemical network consisting of elementary chemical reactions into modules that admit dynamic and parametric modularity. Further work can be done to extend these results to biological networks for which the elementary reactions are unknown. Concretely, it is often the case that the existence of an interaction between two species is known through experimentation, but the exact form of the reaction is unknown. In this case, it would be interesting to partition biological networks into modules that admit dynamic and parametric modularity purely by analyzing experimental data. For this, it would be essential to design experiments smartly to obtain essential parameters of the network and analyze response dynamics of individual species to determine how the network partitions can be done.

Biological modules that admit dynamic and parametric modularity could be very useful in synthetic biology, where circuits are designed from bottom-up. The absence of retroactivity between modules, and the parametric independence between each module, would permit network behavior to be predicted from the characteristics of individual modules. This would aid synthetic circuit design to bring about novel behaviors.

In Chapter 3, we analyzed biological modules that occur frequently in biological

networks. One interesting but as yet unresolved problem is on the best way to partition a given biological network into modules. These networks could be modeled by ODEs, stochastic equations, boolean maps or even characterized by their interaction pathways determined by experimentation. Further work can be done on developing a metric that determines if a biological network partition would be useful in simplifying its analysis. Minimizing the number of junctions between the resulting modules that have bidirectional signals was one metric chosen in previous work [103], but it is unclear if this necessarily helps to reduce the complexity of analyzing a network. For example, we will see in Part II of this monograph biological modules that have bidirectional signals flowing between them that still provide useful insights into the network behavior. Similarly, ensuring monotonicity of the resulting modules was another metric chosen [64], but it is unclear if this is always possible to do in a meaningful way.

Finally, there is a large body of work that has also shown that biological processes are inherently noisy. In some cases, a stochastic approach may be more appropriate to model the dynamics of chemical reactions. Further work can be done to generalize the results presented in Part I of this monograph, taking into account stochasticity and noise.

Part II

Modular analysis of the $p53$ network

Chapter 6

The $p53$ tumor suppressor protein

A tumor is an abnormal growth of cells, which can become cancerous if the proliferation of these cells occurs in an uncontrolled fashion [30]. A major cause for the emergence of tumors is genetic mutations that occur when the cellular DNA is damaged and not repaired to its original state. When the mutations interfere with the normal regulation of cell division, then cell proliferation could cause the tumor to become cancerous [19]. While there is abundant evidence that exposure to chemical carcinogens and radiation has had an impact on cancer rates [50], there is also evidence that cancer is a general feature of multi-cellular organisms that has presented a long-standing evolutionary challenge [36].

A key component of the DNA damage response mechanism is the $TP53$ gene which expresses the $p53$ tumor suppressor protein, the so-called "guardian of the genome" [71]. $p53$ and its ancestors have been protecting metazoans from mutations arising from DNA damage for over one billion years [18]. Recent research has shown that $p53$ plays this role by first sensing DNA damage, and then mediating various downstream processes in response. $p53$ can promote cell survival by up-regulating genes that bring about cell-cycle arrest or DNA repair. However, $p53$ can also promote pro-elimination processes that bring about permanent cell-cycle arrest or cell death [122, 123, 39, 54].

One of the oldest known mechanisms cells employ in response to DNA damage is programmed cell death, also known as apoptosis [3]. Cells with DNA that is damaged above some "apoptotic" threshold level [91, 25, 70] could display sustained high levels of $p53$, inducing the expression of pro-elimination genes [91, 97, 70, 130] that would bring about apoptosis, thus limiting the risk that a cell lineage will become cancerous. This behavior is analogous with some basic concepts from Statistical Decision theory. One possible paradigm for capturing this behavior is by imagining that many cells employ the statistical hypothesis "When DNA damage crosses the apoptotic threshold, it is beyond repair and the cell must be killed to prevent the formation of a tumor".

It is clear that setting a low apoptotic threshold minimizes the risk of cancer since a cell that experiences potentially cancer-causing DNA damage would go to apoptosis. A feature of such networks is that they abet Type 1 errors. Such errors happen when the corresponding Statistical Null Hypothesis ("Cells can recover and operate normally even after DNA damage crosses the apoptotic threshold") is true, but the cell is killed anyway. While not ideal, Type 1 errors would still be preferable to Type 2 errors. In the latter, a cell carrying a cancer-causing mutation would be allowed to live. A single Type 2 error would lead to lethal cancer, and this is in fact known to be one of the "hallmarks" of cancerous clones [51, 50]. That said, overly responsive networks with a low apoptotic threshold may result in developmental defects, reduced tissue growth, or high metabolic costs to replace the dead cells. A major task for $p53$ and the DNA damage response mechanism is therefore to effectively suppress the onset of tumors by selecting different cell fates that would increase whole organism fitness. This is achieved by trading-off between minimizing the number of Type 1 errors that take place, while also trying to suppress all Type 2 errors.

6.1 Literature Review

Among the vast literature available on $p53$, we present results that are key to our work.

6.1.1 Experimental results on $p53$ response to DNA damage

The first time oscillations in $p53$ were reported in response to DNA damage was in a paper by Bar-Or et al in 2000 [14]. The authors claimed that $p53$ levels in cell populations displayed damped oscillations in response to damage beyond a threshold level. A few years later, a landmark experimental study in individual human breast cancer cells revealed that in fact, $p53$ pulses in response to DNA damage had the same mean period across different cells, with the number of oscillations increasing with increasing amounts of damage [72]. It was then revealed that the amplitude of these pulses varied between the different cells, and therefore in the earlier study, the oscillations appeared to be damped due to the averaging of $p53$ responses over the cell population [72]. Further studies revealed that $p53$ could oscillate for many days in response to large amounts of DNA damage [48]. At the time, it was well accepted that the primary reason for the oscillations was delayed negative feedback between $p53$ and another protein $MDM2$.

A subsequent study presented evidence that a simple delayed negative feedback between $p53$ and $MDM2$ would in fact not be sufficient to explain the experimentally observed behavior [17]. In this study, the upstream transduction kinases that relay DNA damage information to $p53$, such as ATM and $ChK2$, were also shown to undergo pulsing in response to damage, and these pulses were coupled with pulses of $p53$. It was also shown that a pulse of $p53$ could be initiated by inducing the kinases for one hour, with subsequent pulsing of $p53$ depending on continued kinase activity. The kinase activity is shut down by due to negative feedback with $p53$, for which the protein $Wip1$ was found

to be a crucial protein. However, kinase activity was found to be repeatedly activated in the presence of DNA damage. This is known as the *recurrent initiation* model for $p53$ pulsing, and suggests that the observed behavior is a result of excitable pulses in $p53$ due to kinase activity as opposed to autonomous oscillations.

A more recent study in human U2-OS cells reveals bimodal behavior in response to DNA damage [25]. In response to moderate amounts of damage, $p53$ admits pulsatile behavior. Large amounts of damage trigger a strong monotonic elevation of the $p53$ level, which is not seen to return to a low level subsequently.

6.1.2 The role of $p53$ dynamics in bringing about varied cell fates

The role of $p53$ in determining when to promote cell survival and when to induce programmed cell death (apoptosis) to DNA damage has been studied for many years. An early study revealed that moderate levels of $p53$ in cells were likely to bring about cell cycle arrest, while higher levels of $p53$ were more likely to bring about apoptosis [26]. In fact, it was even postulated that $p53$ was a necessary intermediate to bring about apoptosis in response to radiation [80]. While it was thought that $p53$ bringing about cell cycle arrest simply provided time for DNA repair to take place, studies revealed that $p53$ actually plays a more active role in DNA repair, and the loss of $p53$ function resulted in decreased DNA repair in response to damage [111]. A later study also suggested that DNA repair is activated much faster than $p53$-induced apoptosis [47, 41]. A study on $p53$'s role in inducing permanent cell cycle arrest (senescence) was also conducted around that time [62].

Many recent studies focused on the mechanisms through which $p53$ could be involved in the apparently antagonistic functions of promoting cell survival and cell death or

senescence. Studies such as [122, 60, 106, 58, 123, 39, 54] provide comprehensive reviews on the different mechanisms and intermediate proteins through which $p53$ can carry out these functions. While this knowledge is helpful in understanding the underlying molecular mechanisms, there are fewer studies relating the experimentally observed $p53$ response to DNA damage to the role of $p53$ in bringing about various cell fates. In one such study, cells that underwent $p53$ pulsing after damage were altered to produce a sustained high $p53$ response [97]. It was shown that cells in which $p53$ pulsed were able to recover from DNA damage, whereas cells exposed to sustained $p53$ signaling frequently underwent a terminal cell fate. Although the authors were only able to study senescence due to the limitations of their experimental method [97], numerous studies also reveal that sustained high $p53$ signaling can lead to apoptosis [56, 77, 26]. One of the accepted paradigms for the onset of apoptosis is sustained high levels of $p53$ signaling [91, 97, 70, 130], which in turn is triggered when the amount of DNA damage detected crosses some "apoptotic" threshold [91, 25, 70].

The recent study by [25] in which they show that moderate amounts of damage leads to $p53$ pulsing, and large amounts of damage leads to a strong monotonic elevation in $p53$ therefore provides a crucial link between the role of $p53$ in determining cell fate and experimental observations. The pulsing behavior in $p53$ allows cell cycle arrest to be initiated and promotes the induction of DNA repair genes. As long as damage is present, $p53$ will pulse, allowing for DNA to be repaired. However when the amount of damage crosses a threshold, $p53$ levels switch monotonically to a high value. This allows for the induction of irreversible cell fates such as senescence and apoptosis.

6.1.3 Earlier models of the $p53$ response to DNA damage

There have been many dynamical models proposed to explain $p53$ behavior in response to DNA damage. Most of the early models focused on explaining oscillatory behavior, since the plausible monotonic switching of $p53$ to a high level in response to high levels of damage had yet to be discovered.

Some of the early modeling work assumed that the $p53 - MDM2$ negative feedback loop was the basis behind the pulsatile behavior observed in $p53$ [81, 124]. It was noted that while negative feedback can generate sustained oscillations, these oscillations are typically not very robust [131]. Moreover as noted earlier, more recent experimental observations suggest that $p53$ pulses in an excitable fashion due to recurrent initiation of $p53$ by upstream kinases, and not due to the negative feedback between $p53$ and $MDM2$ [17].

More recent modeling work on the $p53$ core regulation network incorporates both the $p53 - MDM2$ negative feedback loop along with positive feedback around $p53$, although the species involved in the positive feedback loop are not explicitly modeled [28, 131, 27, 98]. These models assumed that the upstream kinase concentrations were proportional to damage and caused autonomous pulsing in $p53$ when the kinase levels were sufficiently high, again violating recent results from [17, 16].

In addition, all of the above models did not consider the possibility of $p53$ displaying both pulsing behavior and sustained high levels of signaling. There are other models of the $p53$ core regulation network in humans that display both the pulsing and sustained high levels of $p53$ [98, 126, 132]. However, these models imply that $p53$ would need to pulse before switching to a high value [98, 126, 132], contrasting more recent experimental results [25].

Chapter 7

Modeling the evolution of the $p53$ core regulation network

While $p53$ is a vital component of the damage response network, it acts as part of a network of interacting genes that determine cell fate. Extensive genetic studies have uncovered a complex network of hundreds of genes [75] that cooperate to carry out tumor suppression as soon as damage is registered in the genome [60]. There are a multitude of upstream sensors that detect DNA damage, the transducers that relay the information to $p53$ and the downstream actuators that are regulated by $p53$ to bring about different cell fates [58]. In addition, the dynamics of $p53$ is known to be governed by its interaction with a set of core regulation network proteins including $MDM2$, $PTEN$, AKT, ARF and $E2F1$ through multiple feedback pathways [52]. Mutations in some of the respective genes have even been implicated in human cancers, although not to the extent of $p53$ [106].

Im this chapter, we first uncover the phylogeny of organisms with varying evolutionary complexity with respect to $TP53$, $MDM2$, $PTEN$ and ARF, all of which play a crucial role in the core regulation of $p53$ [52, 75]. Based on these results, we map the paths by

which the $p53$ core regulation network could have evolved from the structures seen in early ancestors to those seen in more evolved vertebrates. This study unveils the intermediate core regulation network configurations that might have evolved, each of which could admit qualitatively distinct characteristics. We then introduce the reader to the dynamic models associated with these different network configurations. The evolutionary paths derived from the phylogenic data provide a natural way to decompose the different $p53$ networks into modules, which allows us to study the effect of removing different combinations of modules on the overall network behavior.

7.1 Taxonomic representation of regulatory genes and network configurations

Among clinical cancers, mutations in $p53$ are among the most commonly detected [71], and $p53$ is widely recognized as having a key role in tumor suppression in multicellular organisms [52, 123]. The $p53$ family of genes is at least one billion years old [18], appearing in organisms as diverse as choanoflagellates, sea anemones and mammals. The tumor supression role of the $p53$ ancestral gene appears to have been preserved over a large time span; its function in the sea anemone is to protect the germline gametes from DNA damage, and this function persists in insects, worms, clams, and vertebrates [18].

In humans, $p53$ is part of a sophisticated network of hundreds of genes [75] that cooperate to mediate cell fate decisions such as the initiation of cell-cycle arrest, DNA repair, senescence or apoptosis, as soon as damage is registered in the genome [60]. This network includes a core regulation network [52] consisting of the proteins $MDM2$ [15, 128, 114], $PTEN$, [59, 113], and ARF [95, 82, 121] among a host of other upstream, downstream and intermediate species involved in sensing, transduction and regulation

[52, 58, 60].

Recent studies have uncovered much about how these core regulation proteins interact. It is well known that $p53$ transcriptionally activates the $MDM2$ gene. $MDM2$ in turn antagonizes $p53$ forming a negative feedback loop around $p53$ [84, 15, 128]. $p53$ also activates the $PTEN$ gene, and $PTEN$ proceeds to down-regulate $MDM2$ through a series of interactions [113, 83, 126], which forms a positive feedback loop around $p53$. ARF is known to cause the translocation and eventual degradation of $MDM2$ [109, 82, 95], while it has been recently discovered that ARF can also mediate the degradation of $MDM2$ [121], leading to a positive feedback loop around $MDM2$. The network configuration of these four genes is shown in Figure 7.1, and will henceforth be referred to as the *full network* $\mathcal{N}_v$. Much of our understanding of the core regulation network interaction and dynamics in response to DNA damage comes from studies of human and mouse cell lines [76], and this network is most likely common to at least mammals. As we will see in this paper, this network configuration brings about various cell fate decisions in humans, and is crucial in explaining the experimental observations about $p53$ behavior in response to DNA damage.

A simple search for the presence of these four genes showed that homologs of $p53$, $PTEN$ and ARF are present in a large taxonomic group of animals, including non-vertebrates, while $MDM2$ does not appear in invertebrate species such as *C. Elegans*, *D. Melanogaster* and *E. Histolytica*. This is consistent with a recent study which found that $MDM2$ and its paralog $MDM4$ arose from a gene duplication event over 440 million years ago and became the primary negative regulators of the $p53$ family of genes [87]. A brief glance at Network $\mathcal{N}_v$ in Figure 7.1 reveals that the absence of this $MDM2$ would prevent $PTEN$ and ARF from affecting the dynamics of $p53$, because both $PTEN$ and ARF only regulate $p53$ through feedback connections involving $MDM2$. Thus, in the absence of $MDM2$, $p53$ regulates $PTEN$, but neither $PTEN$ nor ARF can directly

affect $p53$. This means that the response of $p53$ to DNA damage would only depend on the dynamics of $p53$. The fact that the loss of a single gene could potentially affect the functioning of the entire network motivated us to study the "paths to complexity" from the ancestral $p53$ network to the network inferred to exist mammals. In the remainder of this section, we explore the possible intermediate network configurations that might have emerged in the evolution from early species to humans, where each intermediate network configuration behaves in a qualitatively different manner from the other networks.

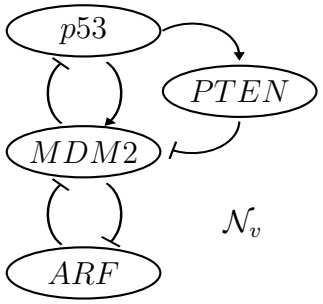


Figure 7.1: The $p53$ core regulation network configuration in humans, which we label $\mathcal{N}_v$.

The evolution of the p53 network

As was mentioned earlier, invertebrates were known not to possess the $MDM2$ gene, which would lead to the $p53$ response to DNA damage being governed by $p53$ alone. The network in Figure 7.2a corresponds to a primordial species without $MDM2$ and is labeled $\mathcal{N}_i$.

The study by Momand et al suggests that a gene duplication event 440 million years ago led to the emergence of $MDM2$, which became one of the primary negative regulators of $p53$ [89]. Figure 7.1 shows that in the Network $\mathcal{N}_v$ present in humans, $MDM2$ is involved in interactions with $PTEN$ (via AKT), ARF and $p53$.

Since it is statistically improbable that all the $MDM2$ interactions that appear in the human network $\mathcal{N}_v$ in Figure 7.1 appeared simultaneously, we consider here two possible

evolutionary paths from the ancestral network $\mathcal{N}_i$ to the mammalian network $\mathcal{N}_v$. In both evolutionary paths, the $p53 - MDM2$ negative feedback interaction evolves first (Figure 7.2b) since this interaction is central to the regulation of $p53$ by either $PTEN$ or ARF. In the first path, the interactions through which $p53$ inhibits $MDM2$ through $PTEN$ (Figure 7.2c) evolves first (leading to a positive feedback loop on $p53$), followed by the mutually inhibitive feedback loop between $MDM2$ and ARF (Figure 7.1). In the other path, the mutually inhibitive feedback loop between $MDM2$ and ARF evolves first (Figure 7.2d), followed by the interactions through which $p53$ inhibits $MDM2$ via $PTEN$.

This study will highlight the role of each of the regulatory interactions in bringing about the $p53$ response to DNA damage in human cells. It is worth noting that our emphasis is on providing insight into the behaviors permitted by these different network configurations from a dynamical systems standpoint. As such, we do not make any claims about the relative fitness of organisms that might express such structures. Moreover, we also do not make claims that the paths to complexity studied in this paper are the only possible paths that might have evolved. Rather, we choose to focus on exploring the different qualitative behaviors permitted by the intermediate network configurations.

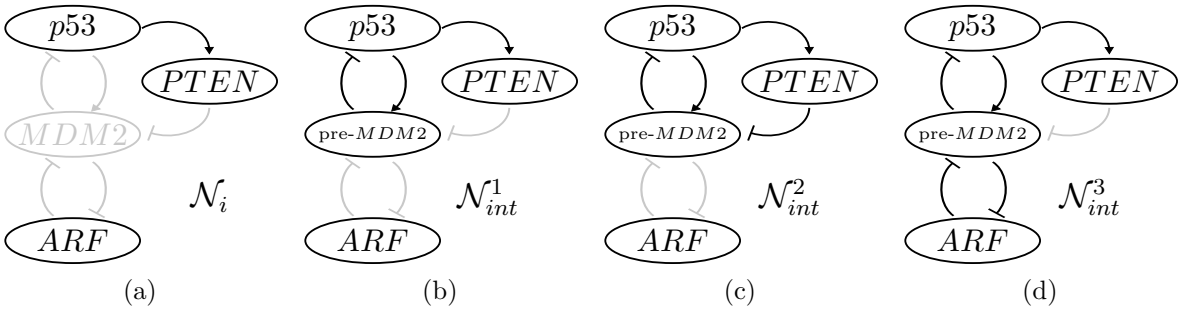


Figure 7.2: (a) $p53$ network configuration in invertebrates in the absence of $MDM2$. (b)–(d) Possible intermediate network configurations in the evolution to increased complexity.

7.2 Modular models of the $p53$ pathways

The diagrams shown in Figures 7.1–7.2 are conceptual in that while each block is labeled with the name of a single protein, it typically represents several interacting chemical species. For example, the block labeled $MDM2$ includes the dynamics of the $MDM2$ protein in its active an inactive forms, as well as its corresponding mRNA. We should thus think of each block as a "module" rather than an individual protein.

To predict how the network behavior changes as modules are added, the different modules in Figures 7.1–7.2 must be associated with precise mathematical models, so as to exhibit dynamic modularity and parametric modularity. Figures 7.3–7.7 show a mathematically precise "block-diagram" representation of the network configurations from Figures 7.1–7.2. Each module, labeled from $\mathcal{M}_1$ to $\mathcal{M}_4$, is characterized by a system of ordinary differential equations (ODEs) depicted in Figures 7.8–7.11 that describe the behavior of the species in that module. The solid arrows between the modules represent the signals that flow between them. For example, the arrow originating at $\mathcal{M}_2$ and ending at $\mathcal{M}_1$ represents the mechanism by which active $MDM2$ represses $p53$. The dashed arrows represent signals that originate upstream from the network or flow downstream, such as the transduction of DNA damage signals to activate $p53$ or the acceleration of DNA repair by $p53$. These signals are labeled $[D_1]$ and $[D_2]$, and their behavior is further discussed in Section 9.2.

The species associated with each module are as follows. Module $\mathcal{M}_1$ is associated only with the dynamics of the $p53$ protein. Module $\mathcal{M}_2$ consists of the dynamics of $MDM2$ in both its active ($MDM2a$) and inactive ($MDM2$) forms, along with its mRNA ($mRNA_m$). Module $\mathcal{M}_3$ consists of the dynamics of $PTEN$ and its mRNA ($mRNA_p$), along with the intermediate species $PIP2$ and its phosphorylated form $PIP3$, and AKT and its active phosphorylated form $AKTa$. Finally, Module $\mathcal{M}_4$ contains the dynamics

of the ARF protein.

The specific ODE dynamics of each module is illustrated in Figures 7.8–7.11, where the specific parameters and functions used in our work can be found in Appendix B. In $\mathcal{M}_1$, $p53$ is activated at some basal rate p_1 and also in an input-dependent manner as a function of $u_{11} = [D_1]$, which represents kinases that activate $p53$ in response to DNA damage. $p53$ has a basal degradation rate constant of d_1. The degradation of $p53$ is also mediated by the input u_{12}, which typically describes the antagonism of $p53$ by active $MDM2$ [15, 128]. This results in the interconnection between Modules $\mathcal{M}_1$ and $\mathcal{M}_2$ given by

$$y_{21} = u_{12}, \tag{7.1}$$

where y_{21} is an output from Module $\mathcal{M}_2$ given by the expression

$$y_{21} = b_1 \cdot [MDM2a]\,. \tag{7.2}$$

In Module $\mathcal{M}_2$, $mRNA_m$ has a basal transcription rate p_2 and a degradation rate constant d_2. In addition, $mRNA_m$ transcription is activated by $p53$ through the input u_{21}, which can be described through the interconnection

$$y_{11} = u_{21}, \tag{7.3}$$

where y_{11} is an output from Module $\mathcal{M}_1$ given by the expression

$$y_{11} = b_4 \cdot f_4([p53]). \tag{7.4}$$

$MDM2$ is translated at a rate of $p_3 \cdot [mRNA_m]$ and degraded with rate constant d_3. The phosphorylation of $MDM2$ into its active form $MDM2a$ is mediated by an input u_{22},

which is typically mediated by $AKTa$ that is associated with $\mathcal{M}_3$ [84, 83]. This results in the interconnection

$$y_{31} = u_{22} \tag{7.5}$$

where y_{31} is the only output from Module $\mathcal{M}_3$ given by the expression

$$y_{31} = b_2 \cdot [AKTa]\,. \tag{7.6}$$

The rate of dephosphorylation back to $MDM2$ is given by $b_3 \cdot f_3([MDM2a])$. $MDM2a$ has a basal degradation rate constant p_4. The degradation of $MDM2a$ is also mediated by two inputs u_{23} and u_{24}. $u_{23} = [D_1]$ represents the concentration of upstream transduction kinases like ATM, which mediate the down-regulation of active $MDM2$ [114, 126, 125]. u_{24} represents the ARF-mediated translocation and eventual degradation of active $MDM2$ by ARF [109, 82, 95], which results in the interconnection

$$y_{41} = u_{24}, \tag{7.7}$$

where y_{41} is the only output from Module $\mathcal{M}_4$ given by the expression

$$y_{41} = b_{11} \cdot [ARF]\,. \tag{7.8}$$

In Module $\mathcal{M}_3$, the mRNA of $PTEN$, represented by $mRNA_p$, is transcribed at some basal rate p_5, with a degradation rate constant d_4. The transcription of $mRNA_p$ is also governed by the input u_{31}, which describes the transcriptional activation of $PTEN$ by $p53$ [113, 83, 126] resulting in the interconnection

$$y_{12} = u_{31}, \tag{7.9}$$

where y_{12} is an output from Module $\mathcal{M}_1$ given by the expression

$$y_{12} = b_5 \cdot f_5([p53]). \tag{7.10}$$

$mRNA_p$ is translated to $PTEN$ with rate constant p_6, and degraded with rate constant d_5. $PTEN$ mediates the de-phosphorylation of $PIP3$ into $PIP2$. The total concentration of $PIP2$ and $PIP3$ is assumed to be the constant tot_1. The rate of phosphorylation of $PIP2$ is given by the term $b_6 \cdot f_6(tot_1 - [PIP3])$, while the rate of dephosphorylation of $PIP3$ is given by $b_7 \cdot [PTEN] \cdot f_7([PIP3])$. $PIP3$ then mediates the phosphorylation of AKT to $AKTa$ (active AKT). The total concentration of AKT and $AKTa$ is assumed to be the constant tot_2. The rate of phosphorylation of AKT is given by the term $b_9 \cdot [PIP3]\, f_9(tot_2 - [AKTa])$, while the rate of dephosphorylation of $AKTa$ is given by $b_8 \cdot f_8 7([AKTa])$ [98, 126]. The equations describing this module are shown in Figure 7.10.

In Module $\mathcal{M}_4$, p_7 is the basal rate of production of ARF, with $u_{41} = [D_2]$ the extrinsic rate of production that represents the activation of ARF in response to DNA damage through transducers like $PARP1$ [92] and $E2F1$ [107] . d_6 is the rate constant for the degradation of ARF. ARF is also degraded by an extrinsic process mediated by u_{42}, which represents a novel recently-discovered mechanism through which $MDM2$ can mediate the degradation of ARF [121]. This results in the interconnection

$$y_{22} = u_{42}, \tag{7.11}$$

where $y_2 2$ is an output of Module $\mathcal{M}_2$ given by the expression

$$y_{22} = b_{10} \cdot [MDM2a]\,. \tag{7.12}$$

To preserve simplicity of the model, we assume that both $MDM2$-ARF interactions from (7.7) and (7.11) involve active $MDM2$, since active $MDM2$ directly influences the temporal behavior of $p53$. This is a similar assumption to the one made in [126], where the upstream kinases were assumed to accelerate the degradation of active $MDM2$. Moreover, since there are multiple known $MDM2 - ARF$ binding sites [121], we assume that these two interactions occur at separate binding sites.

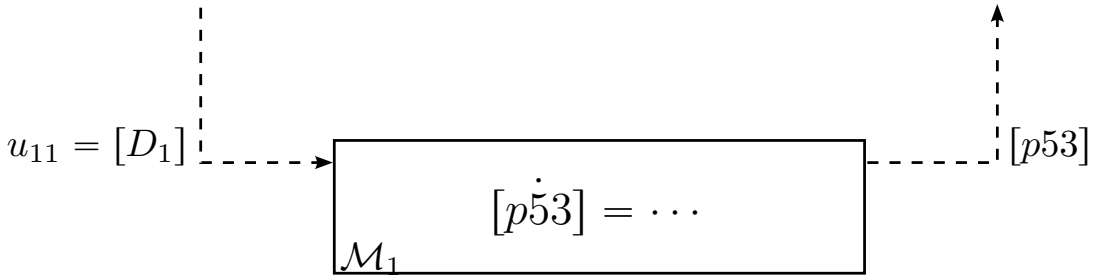


Figure 7.3: Modular block-diagram representations of Network $\mathcal{N}_i$. The specific ODEs describing each module are shown in Figures 7.8–7.11.

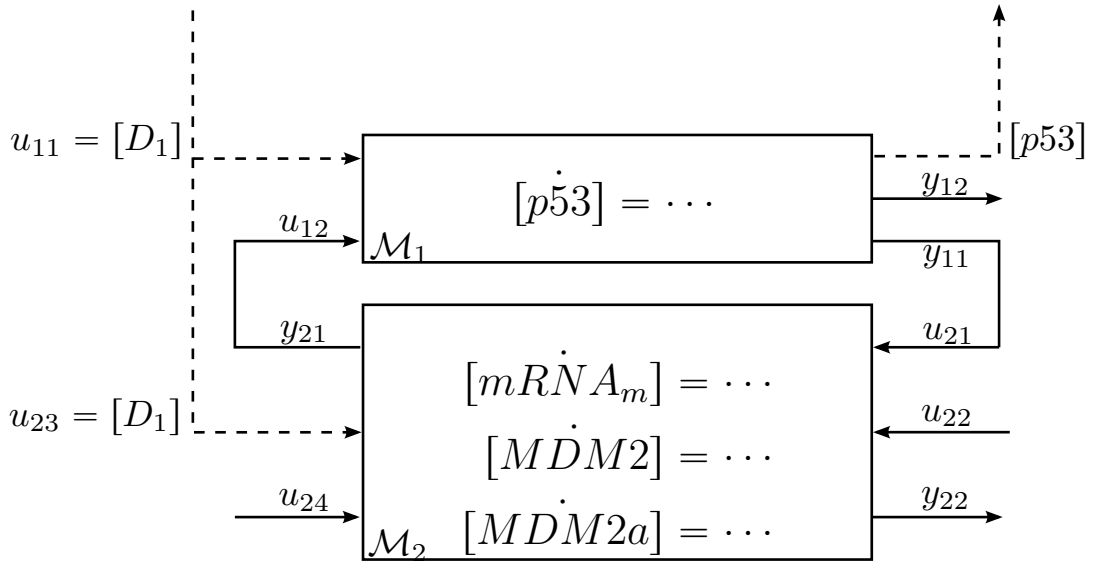


Figure 7.4: Modular block-diagram representations of Network $\mathcal{N}^1_{int}$. The specific ODEs describing each module are shown in Figures 7.8–7.11.

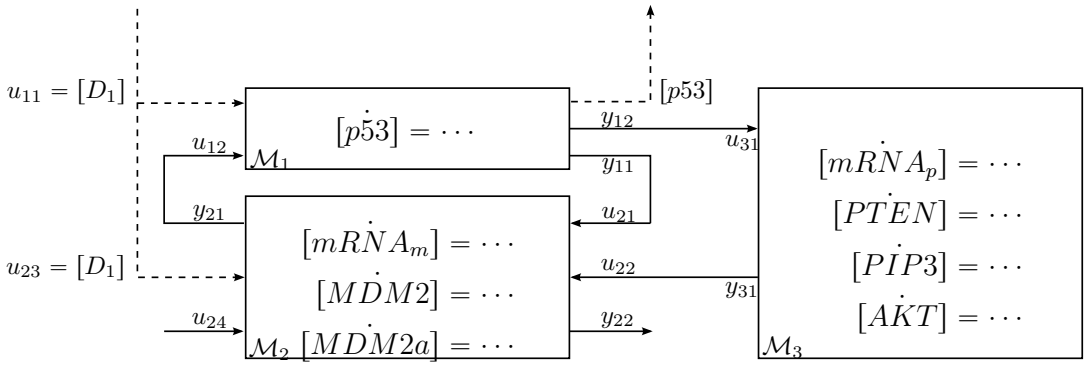


Figure 7.5: Modular block-diagram representations of Network $\mathcal{N}_{int}^2$. The specific ODEs describing each module are shown in Figures 7.8–7.11.

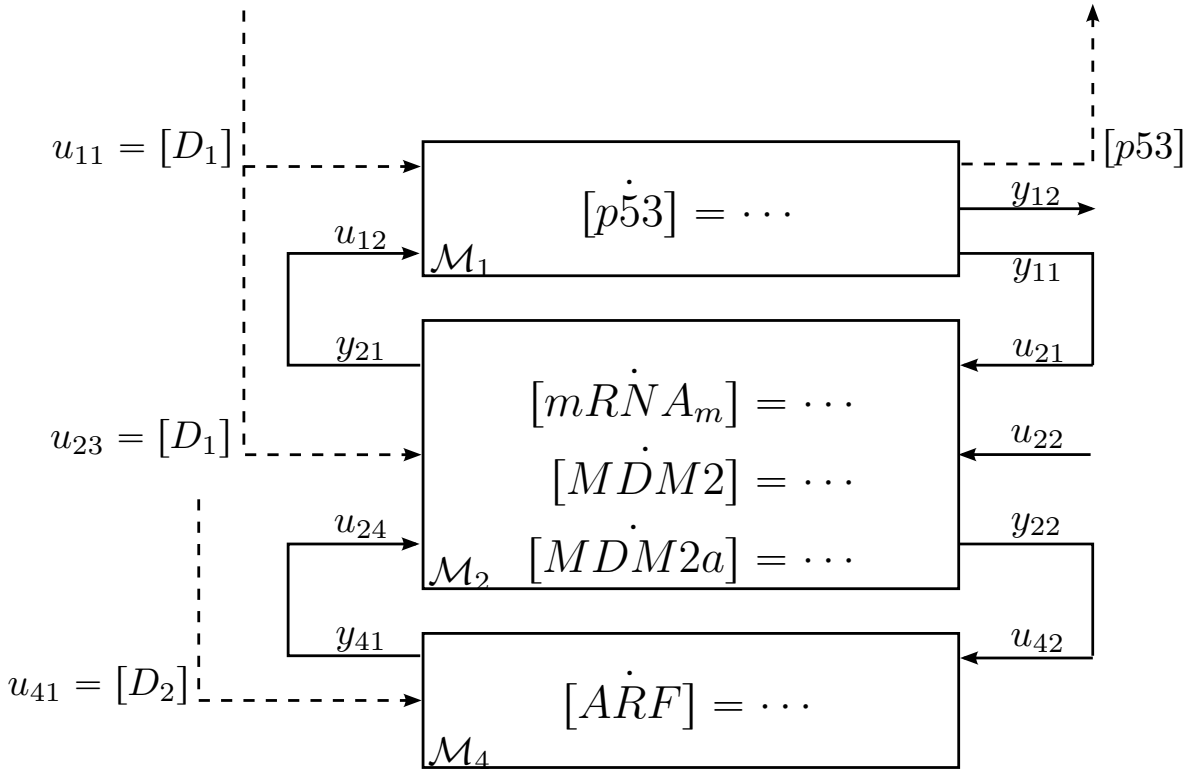


Figure 7.6: Modular block-diagram representations of Network $\mathcal{N}_{int}^3$. The specific ODEs describing each module are shown in Figures 7.8–7.11.

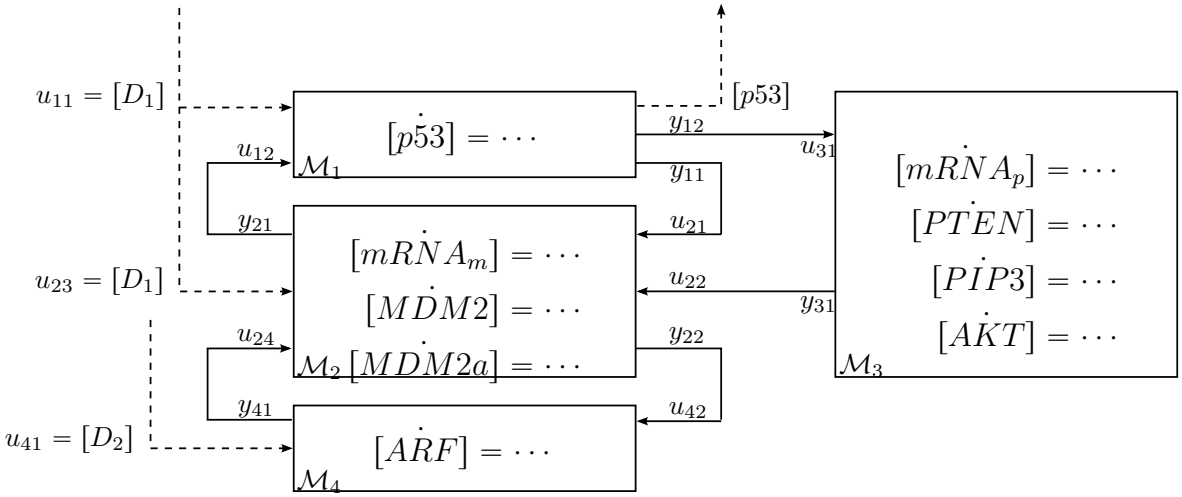


Figure 7.7: Modular block-diagram representations of Network $\mathcal{N}_v$. The specific ODEs describing each module are shown in Figures 7.8–7.11.

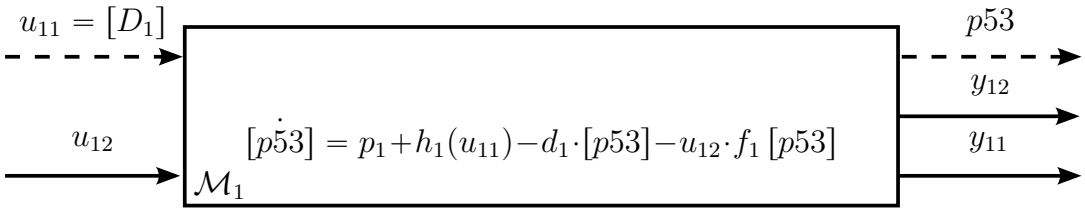


Figure 7.8: ODE associated with Module $\mathcal{M}_1$

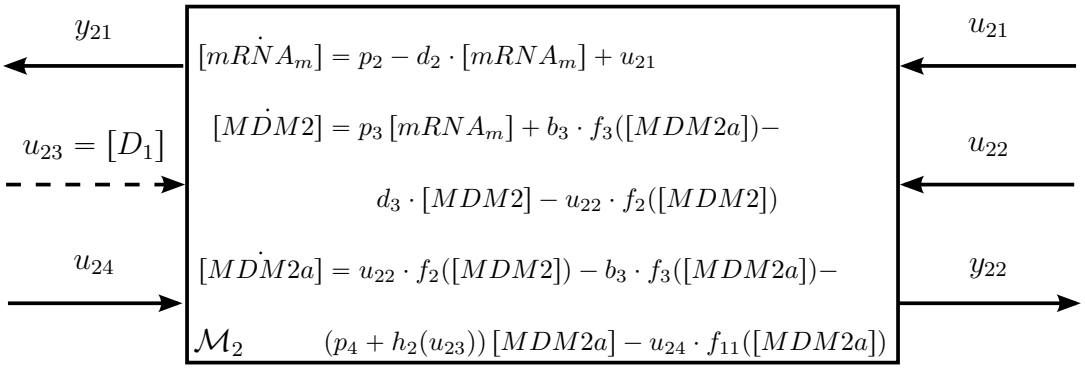


Figure 7.9: ODE associated with Module $\mathcal{M}_2$

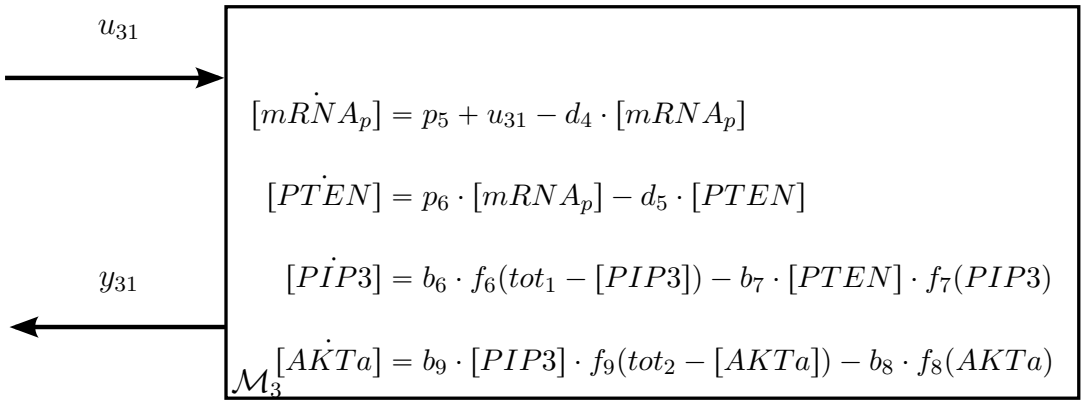


Figure 7.10: ODE associated with Module $\mathcal{M}_3$

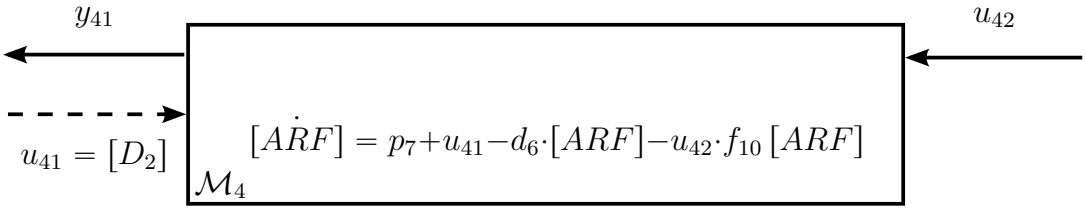


Figure 7.11: ODE associated with Module $\mathcal{M}_4$

7.3 Conclusion

While homologs of $TP53$, $PTEN$ and ARF are inferred to have been present in the earliest metazoan ancestors, the same cannot necessarily be said about $MDM2$. Using our evolutionary data, we mapped the pathways to complexity of the network from early ancestors to the network configuration in vertebrates, and introduced a dynamical model to capture the interactions between the corresponding proteins, that can be partitioned into modules that admit dynamic and parametric modularity. To our knowledge, this is the first evolutionary study of the $p53$ pathway that considers the $PTEN$ and ARF genes. A key novelty is that we incorporate in our model the recent discovery of a mechanism through which $MDM2$ mediates the degradation of ARF [121].

Chapter 8

Computation of the $p53$ network equilibrium points

In our study of $p53$ network evolution, our focus is on whether the alternative network structures introduced in Chapter 7, which we infer as ancestral to the mammalian network, would allow for qualitatively distinct forms of $p53$ regulation, which in turn would lead to qualitatively different strategies for dealing with the risk of cancer. Specifically, we seek to study the different forms of switching behavior each of these networks could admit, for which we would need to study the steady-state bifurcations of each network with respect to the DNA damage input parameters introduced in Chapter 7. As it turns out, standard numerical solvers are unable provide equilibrium solutions for the some of these network configurations for some parameters. In this chapter, we introduce an algorithm that allows us to compute the equilibrium points of the network.

8.1 Switching behavior and computation of equilibrium points

Biological networks commonly use switch-like behavior to turn a continuous input signal (usually an extrinsic stimuli) into an "all-or-nothing" output signal response (usually species or compound concentrations). A common mechanism employed by these networks to bring about this behavior is *bistability*. A bistable network has two sets of stable steady-states, in which the state of the network transitions from a neighborhood of one set of steady-states (which we call A) to the other (which we call B) when an input signal grows above a certain threshold (which we call T_1). A unique property of these networks is that even if the input later falls below T_1, the state may still remain in a neighborhood of B. Reversion to a neighborhood of A is generally observed only when the input signal falls below another threshold T_0, where $T_0 < T_1$. This property is known as hysteresis, and refers to the ability of a bistable system to "remember" the input stimulus that was above T_1 long after that stimulus is removed, until it falls below T_0 [74]. The lac operon network in E. Coli, the Wee1-Cdc2 network, the $NF\kappa\beta$ response in an apoptosis network and the cell cycle oscillator in *Xenopus laevis* are among the many naturally occurring biological networks whose responses have been modeled as bistable switches [74, 7, 24]. In some cases, the lower threshold T_0 does not exist, which makes the transition from A to B *irreversible*. It is also plausible to have more than two sets of stable steady-states, which is known as *multi-stability*.

Monostable networks on the other hand have states that slide along a continuum of steady states [7] in response to input signaling. Such networks can only exhibit switch-like behavior if the system is *ultrasenstive*; that is, if the input-output relationship in the network is described by a sharp sigmoidal curve [74]. When the input signal crosses the exponential phase of the sigmoid, the output response of the network transitions

from a low to a high state. In fact, the dynamics of transcription factor target genes as a function of transcription factor concentration is typically modeled as a sigmoidal "Hill function", which can be viewed as an ultrasensitive switch. However, such switches are *memoryless*; once the input signal is removed or is reduced, the system immediately returns to its original state [74]. For example, when the concentration of transcription factor crosses a certain threshold, the target gene turns on. As soon as the transcription factor concentration goes below that same threshold, the target gene immediately turns off.

As it turns out, the two damage sensing inputs $[D_1]$ and $[D_2]$ are *bifurcation parameters* for the $p53$ core regulation network. As was seen in Section 7.2, the rate of production of $p53$ is directly proportional to $[D_1]$, and in networks with Module $\mathcal{M}_2$, $[D_1]$ is also directly proportional to the rate of degradation of active (phosphorylated) $MDM2$. Moreover, in networks with Module $\mathcal{M}_4$, an increase in DNA damage causes an increase in $[D_2]$, which in turn leads to an increase in the rate of production of the ARF protein.

8.2 An algorithm to determine $p53$ network bifurcation diagrams

The fact that $p53$ can exhibit bifurcations with respect to $[D_1]$ and $[D_2]$ implies that the different $p53$ network configurations can exhibit qualitatively distinct switching behavior in response to damage. The computation of these bifurcation diagrams requires the computation of equilibrium points for the system of ODEs described in Chapter 7. In practice, this reduces to the problem of computing the solutions to multivariate polynomial equations, which is known to be an NP-complete problem [31]. While simple

equations can often be solved using commercially available numerical solvers, more complicated systems such as the ones that arise in our analysis, exhibit multiple equations with many possible solutions that are close to each other. In this case, solvers or continuation software often miss solutions that might exist, which was the case when attempting to evaluate the equilibrium points of multi-stable networks.

To overcome this problem, we developed an algorithm to solve for the equilibrium points of a network as a function of the inputs to the network. In our case, the algorithm would solve for the $p53$ equilibrium points as a function of the concentrations of the transducer inputs $[D_1]$ and $[D_2]$, to produce the plots in Figures 9.2 and 9.6, that will be seen in Chapter 9.

8.2.1 The Steady-State Behavior of a module

The procedure that the algorithm employs involves *gridding* the equilibrium values of the input and output signals for each module to form a table known as the *Steady-State Behavior* (SSB) of a module [127]. Typically, this table is obtained by gridding the equilibrium values of the outputs for a range of constant inputs, and this special case of the SSB when there is only one unique output for every input was called the IOSCF of the module in Chapter 3. It is worth noting that these tables could be multi-valued, as there could be multiple output equilibria for a given input as is evidenced in the bistable and tri-stable networks.

The number of degrees of freedom of each module is the total number of inputs to that module. Given a module with p inputs and q outputs, it is often more convenient to perform the grid the equilibrium values of an input and $q-1$ outputs for a range of constant values of the remaining output and $p-1$ inputs. We illustrate this using $\mathcal{M}_1$, where the equilibrium equation is given by

$$\begin{aligned} 0 &= p_1 + [D_1] - d_1\,[p53] - u_{12} f_1([p53]) \\ y_{11} &= b_4 \cdot f_4([p53]) \\ y_{12} &= b_5 \cdot f_5([p53]). \end{aligned} \tag{8.1}$$

Given that f_1 is a Hill function of $[p53]$ (see Appendix B), it is more straightforward to solve for u_{12} as a function of $[p53]$ and $[D_1]$. We can then obtain a table of equilibrium values of the signals, that can be formally represented by the SSB of the module as

$$\mathfrak{B}_1 = \big\{([D_1], u_{12}, y_{11}, y_{12}, [p53]) \in \mathbb{R}^5_{\geqslant 0} \mid (8.1) \text{ is satisfied}\big\}. \tag{8.2}$$

A tabular representation of (8.2) is illustrated in Table 8.1.

$[D_1]$	$[p53]$	u_{12}	y_{11}	y_{12}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$[D_1]^i$	$[p53]^i$	u_{12}^i	y_{11}^i	y_{12}^i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 8.1: Table of equilibrium values for $\mathcal{M}_1$. The superscript i denotes the i^{th} row of the table.

We can obtain the SSBs of the other modules as well, given by $\mathfrak{B}_2 - \mathfrak{B}_4$ for modules $\mathcal{M}_2 - \mathcal{M}_4$ respectively, with corresponding tabular representations shown in Tables 8.2–

8.4.

$$\mathfrak{B}_2 = \{([D_1], u_{22}, u_{24}, y_{21}, y_{22}, u_{21}) \in \mathbb{R}^6_{\geqslant 0} \mid \exists([mRNA_m], [MDM2], [MDM2a]) \in \mathbb{R}^3_{\geqslant 0}$$
$$\text{s.t. } [m\dot{RNA}_m] = 0, [\dot{MDM2}] = 0, [\dot{MDM2a}] = 0,$$
$$y_{21} = b1\,[MDM2a], y_{22} = b_{11}\,[MDM2a]\}$$

$$\mathfrak{B}_3 = \{(y_{31}, u_{31}) \in \mathbb{R}^2_{\geqslant 0} \mid \exists([mRNA_p], [PTEN], [PIP3], [AKTa]) \in \mathbb{R}^4_{\geqslant 0} \text{ s.t.}$$
$$[m\dot{RNA}_p] = 0, [\dot{PTEN}] = 0, [\dot{PIP3}] = 0, [\dot{AKTa}] = 0, y_{31} = b_2\,[AKTa]\}$$

$$\mathfrak{B}_4 = \{([D_2], y_{41}, u_{42}) \in \mathbb{R}^3_{\geqslant 0} \mid \exists\,[ARF] \in \mathbb{R}_{\geqslant 0} \text{ s.t. } [\dot{ARF}] = 0, y_{41} = b_{11}\,[ARF]\}$$

$[D_1]$	u_{22}	u_{24}	y_{21}	y_{22}	u_{21}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$[D_1]^j$	u_{22}^j	u_{24}^j	y_{21}^j	y_{22}^j	u_{21}^j
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Table 8.2: Table of equilibrium values for $\mathcal{M}_2$. The superscript j denotes the j^{th} row of the table.

y_{31}	u_{31}
$\vdots$	$\vdots$
y_{31}^k	u_{31}^k
$\vdots$	$\vdots$

Table 8.3: Table of equilibrium values for $\mathcal{M}_3$. The superscript k denotes the k^{th} row of the table.

$[D_2]$	y_{41}	u_{42}
$\vdots$	$\vdots$	$\vdots$
$[D_2]^l$	y_{41}^l	u_{42}^l
$\vdots$	$\vdots$	$\vdots$

Table 8.4: Table of equilibrium values for $\mathcal{M}_4$. The superscript l denotes the k^{th} row of the table.

8.2.2 A behavioral approach to computing the equilibrium point of a network

To solve for the equilibrium value of $[p53]$ of the network $\mathcal{N}_v$ as a function of $[D_1]$ and $[D_2]$, we need to systematically eliminate the input and output signals flowing between the modules, also known as the *latent variables* in the network [127]. We demonstrate this by first focusing on the interconnection between Modules $\mathcal{M}_1$ and $\mathcal{M}_2$ as is shown in Figure 7.7, where we see the associations $y_{11} = u_{21}$ and $y_{21} = u_{12}$. To eliminate these latent variables, we identify the rows i and j in Tables 8.1 and 8.2 respectively such that

$$\begin{aligned}[D_1]^i &= [D_1]^j \\ y_{11}^i &= u_{21}^j \\ y_{21}^j &= u_{12}^i,\end{aligned}$$

leading to the Table 8.5. Mathematically, the construction of Table 8.5 corresponds to computing the SSB set $\mathfrak{B}_{12}$, where

$$\begin{aligned}\mathfrak{B}_{12} = \big\{([D_1], y_{12}, [p53], u_{22}, u_{24}, y_{22}) \in \mathbb{R}^6_{\geqslant 0} \mid \exists (u_{21}, u_{12}) \in \mathbb{R}^2_{\geqslant 0} \text{ s.t.} \\ ([D_1], u_{12}, u_{21}, y_{12}, [p53]) \in \mathfrak{B}_1, ([D_1], u_{22}, u_{24}, u_{12}, y_{22}, u_{21}) \in \mathfrak{B}_2\big\}.\end{aligned}$$

Computationally, the construction of Table 8.5 is achieved by finding the intersection points between each of the lines mapping from u_{12} to y_{11} from Table 8.1, and each of the lines mapping from u_{21} to y_{21} from Table 8.2, for all combination of other variables, as outlined in Figure 8.1.

$[D_1]$	$[p53]$	u_{22}	u_{24}	y_{22}	y_{12}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$[D_1]^i$	$[p53]^i$	u_{22}^j	u_{24}^j	y_{22}^j	y_{12}^i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 8.5: Table of equilibrium values when modules $\mathcal{M}_1$ and $\mathcal{M}_2$ are combined.

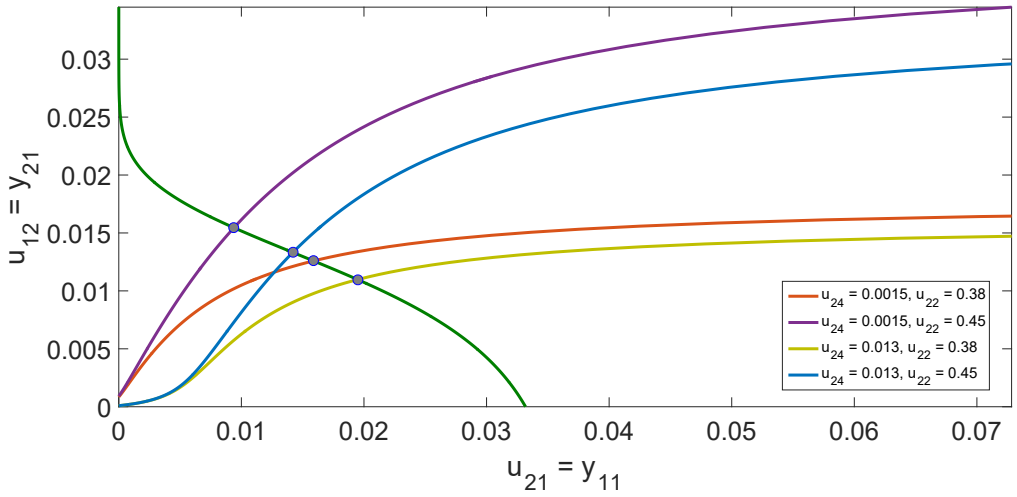


Figure 8.1: Projections of Table 8.1 and Table 8.2 into the u_{12} vs y_{11} (Green plot) and y_{21} vs u_{21} (All other plots) planes respectively, for different values of u_{24} and u_{22} with $[D_1] = 0.4$. The gray intersection points correspond to the equilibrium values of u_{12} and u_{21} for the given u_{24}, u_{22} and $[D_1]$.

The next step is to identify the rows m and k from Tables 8.5 and 8.3 respectively such that

$$y_{12}^m = u_{31}^k$$

$$y_{31}^k = u_{22}^m.$$

The elimination of these latent variables results in Table 8.6, whose construction can be mathematically represented by the SSB

$$\mathfrak{B}_{123} = \big\{([D_1], [p53], u_{24}, y_{22}) \in \mathbb{R}^4_{\geqslant 0} \mid \exists (u_{22}, u_{31}) \in \mathbb{R}^2_{\geqslant 0} \text{ s.t.}$$
$$([D_1], u_{31}, [p53], u_{22}, u_{24}, y_{22}) \in \mathfrak{B}_{12}, (u_{22}, u_{31}) \in \mathfrak{B}_3\big\}.$$

Computationally, the construction of Table 8.6 is achieved by finding the intersection points between each of the lines mapping from y_{12} to u_{22} from Table 8.5, and each of the lines mapping from u_{31} to y_{31} from Table 8.3, for all combination of other variables, as outlined in Figure 8.2.

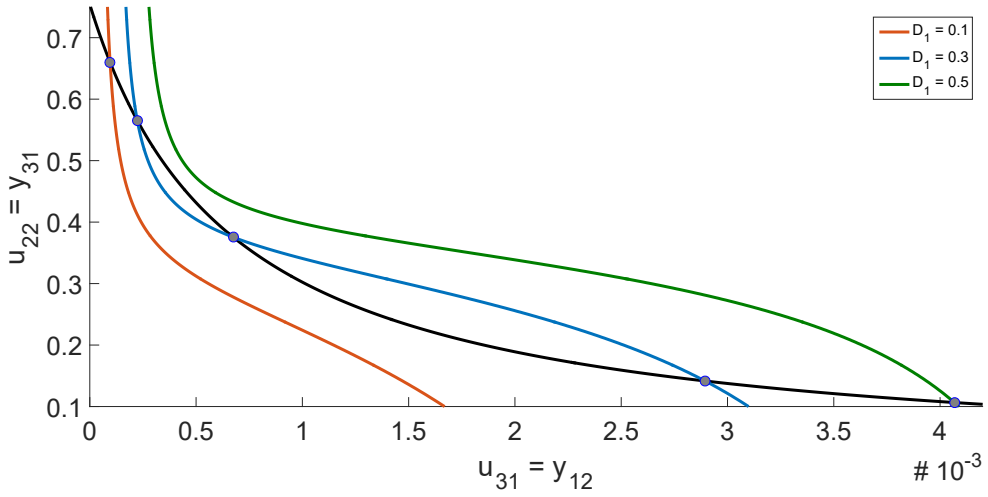


Figure 8.2: Projections of Table 8.5 and Table 8.3 into the u_{22} vs y_{12} (Black plot) and y_{31} vs u_{31} (All other plots) planes respectively, for different values of $[D_1]$ with $u_{24} = 0.0015$. The gray intersection points correspond to the equilibrium values of u_{31} and u_{22} for the given u_{24} and $[D_1]$. The bifurcation can clearly be seen as $[D_1]$ increases, with $[D_1] = 0.1$ and $[D_1] = 0.5$ corresponding to a single equilibrium point, while $[D_1] = 0.3$ corresponds to three equilibrium points.

$[D_1]$	$[p53]$	u_{24}	y_{22}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$[D_1]^m$	$[p53]^m$	u_{24}^m	y_{12}^m
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 8.6: Table of equilibrium values when modules $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ are combined.

Finally, we identify the rows n and l from Tables 8.6 and 8.4 respectively such that

$$y_{22}^n = u_{42}^l$$
$$y_{41}^l = u_{24}^n,$$

The elimination of these latent variables results in Table 8.7, which is the combined steady-state behavior of the entire network $\mathcal{N}_v$. Mathematically, Table 8.7 corresponds to the set

$$\mathfrak{B}_{1234} = \big\{([D_1], [D_2], [p53]) \in \mathbb{R}^3_{\geqslant 0} \mid \exists (u_{24}, u_{42}) \in \mathbb{R}^2_{\geqslant 0} \text{ s.t. } ([D_1], [p53], u_{24}, u_{42}) \in \mathfrak{B}_{123}, ([D_2], u_{24}, u_{42}) \in \mathfrak{B}_4\big\},$$

which was precisely what we wanted to compute. Computationally, the construction of Table 8.7 is achieved by finding the intersection points between each of the lines mapping from u_{24} to y_{22} from Table 8.6, and each of the lines mapping from y_{41} to u_{41} from Table 8.3, for all combination of $[D_1]$ and $[D_2]$, as outlined in Figure 8.3.

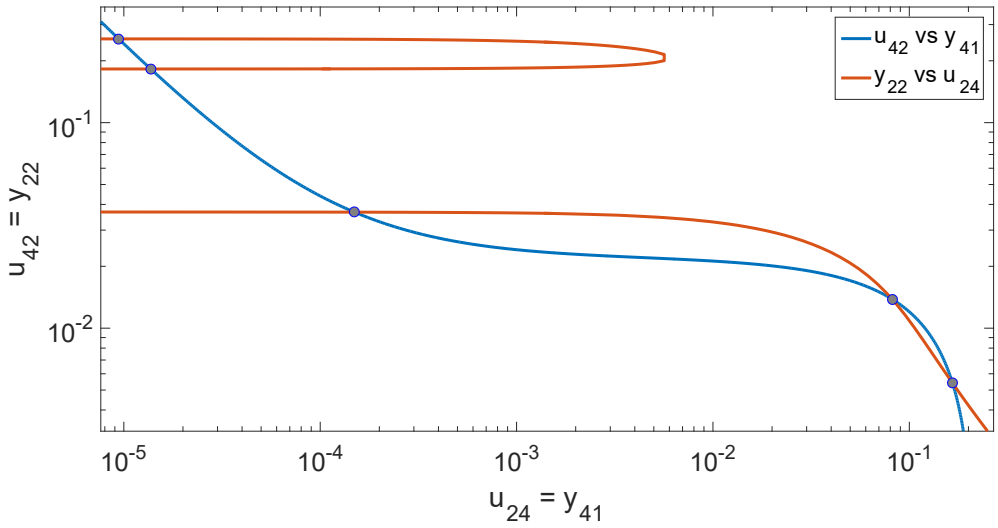


Figure 8.3: Projections of Table 8.6 and Table 8.4 into the y_{22} vs u_{24} (Red plot) and u_{42} vs y_{41} (Blue plot) planes respectively, for $[D_1] = 0.3$ with $[D_2] = 0.002$. The gray intersection points correspond to the equilibrium values of u_{42} and u_{24} for the given $[D_1]$ and $[D_2]$. The equilibrium values can be found for a range of $[D_1]$ and $[D_2]$, and then used to compute the corresponding $p53$ equilibrium values.

$[D_1]$	$[D_2]$	$[p53]$
$\vdots$	$\vdots$	$\vdots$
$[D_1]^n$	$[D_2]^l$	$[p53]^n$
$\vdots$	$\vdots$	$\vdots$

Table 8.7: Table of equilibrium values for the entire network $\mathcal{N}_v$.

A key part of this algorithm was automating the process of computing intersection points between lines, which was done using a MATLAB program available on-line [105]. The stability of each equilibrium point can then be determined through an eigenvalue analysis of the network linearized about each equilibrium point.

8.3 Conclusion

In this chapter, we provided an introduction to biological switching behavior. To analyze the qualitatively distinct switching behaviors admitted by each of the $p53$ core regulation network configurations required the computation of the equilibrium points of the networks, as this would reveal the bifurcations in $p53$ with respect to the DNA damage sensing parameters. Since standard numerical solvers were unable to compute the equilibrium points for some of the network configurations, we developed an algorithm to perform this task. A key aspect of the working of this algorithm was that it relied on the network being partitioned into modules that admitted dynamic and parametric modularity, so that the equilibrium point of the network could be computed from the SSBs of each individual module alone. Further work is required to find an optimal partition of an arbitrary network that minimizes the computational complexity of this algorithm.

Chapter 9

The evolution of $p53$ network behavior

In this chapter, we study how the topology of alternative $p53$ core regulation networks integrate DNA damage information and determine the range of dynamic responses to DNA damage, and explore how this could affect an individual organism's fitness when dealing with DNA damage. We then demonstrate how the $p53$ responses exhibited by different networks relate to the potential costs and fitness benefits of the different core regulation network structures. Clearly, an individual's fitness will be low if it is at a high risk of developing fatal cancer. However, a tumor suppression policy whereby $p53$ and its core regulation proteins bring about apoptosis of too many cells at the slightest sign of DNA damage could also incur high metabolic costs. A core question in the evolution of tumor suppression mechanisms is how the relevant genetic networks can optimally control the response in order to balance the risk of cancer with the risk of false responses to signals of DNA damage.

9.1 The role of each $p53$ network configuration in determining switching behavior

In this section, we discuss the different switching behaviors that are possible for each $p53$ network configuration we identified in Chapter 7, as determined by our algorithm.

9.1.1 The network configurations in the early phases of evolution behave as monostable switches

Our modular study revealed that as $[D_1]$ is varied, the network configurations $\mathcal{N}_i$ and $\mathcal{N}_{int}^1$ in Figures 7.3 and 7.4 operate as monostable switches, in the sense that when an increase in $[D_1]$ makes the $p53$ level rise, a subsequent decrease in $[D_1]$ to its original value causes the concentration of $p53$ to revert to its initial level. The steady-state concentration of $p53$ as a function of $[D_1]$ is shown in Figure 9.1. The main difference between the two networks $\mathcal{N}_i$ and $\mathcal{N}_{int}^1$ is that negative feedback of $p53$ with $MDM2$ typically makes the $p53$ steady-state in $\mathcal{N}_{int}^1$ to be lower than that of $\mathcal{N}_i$.

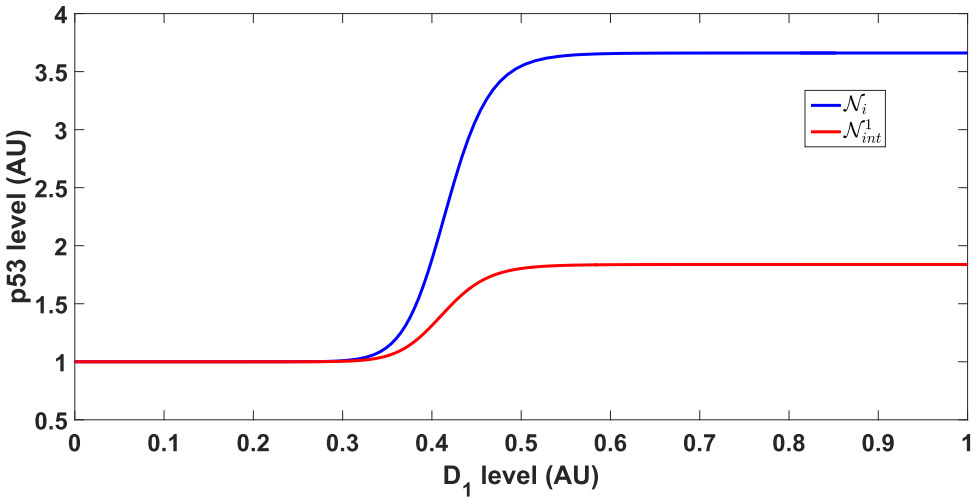


Figure 9.1: Steady-state concentration of $p53$ as a function of $[D_1]$ shows a continuum of steady-states in both $\mathcal{N}_i$ and $\mathcal{N}_{int}^1$. Due to the negative feedback on $p53$, $\mathcal{N}_{int}^1$ admits lower $p53$ steady-state values than $\mathcal{N}_i$. The $p53$ levels are normalized by their value when $[D_1] = 0$.

9.1.2 Networks $\mathcal{N}_{int}^2$ and $\mathcal{N}_{int}^3$ can operate as bistable switches

With model parameters similar to those from [126], the network $\mathcal{N}_{int}^2$ operates as a bistable switch with respect to the input $[D_1]$. From this study, we can plot the *bifurcation diagram* of $\mathcal{N}_{int}^2$ with respect to the parameter $[D_1]$ in Figure 9.2. The solid blue line signifies a continuum of stable steady-state concentrations of $p53$ as a function of $[D_1]$, while the red dashed lines represent unstable equilibrium points. It can be seen that when $[D_1]$ is between the two thresholds T_0 and T_1, there are two stable steady-states the $p53$ level can converge to. To illustrate the behavior of the bistable network, we run a simulation as shown in the red curve of Figure 9.3b. We observe the manifestation of the "memory" property of bistable networks, where after rising to a point in set "B", the $p53$ level returns to a point in set "A" after $[D_1]$ crosses the threshold T_0 from above. Figure 9.3b further contrasts this to the behavior of the monostable networks $\mathcal{N}_i$ (blue curve), where the switch has no memory.

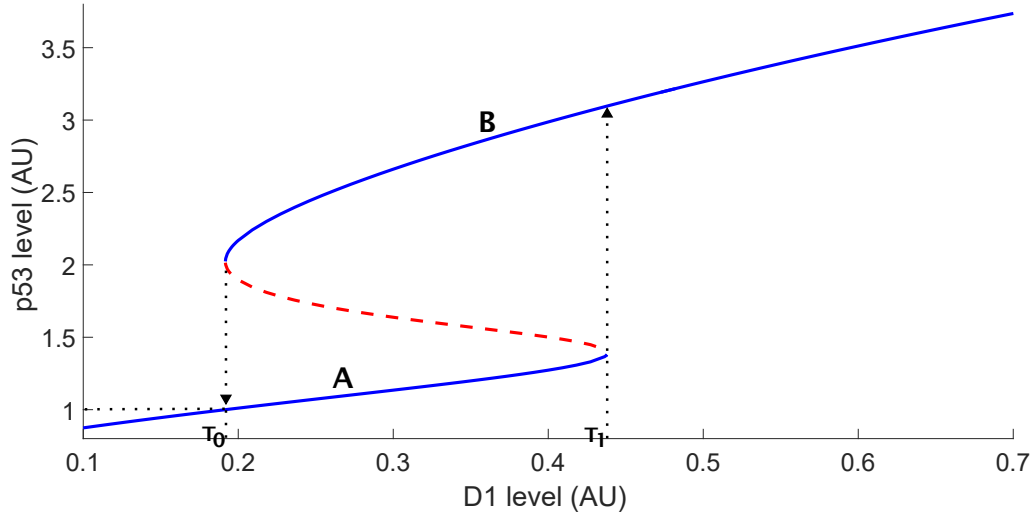
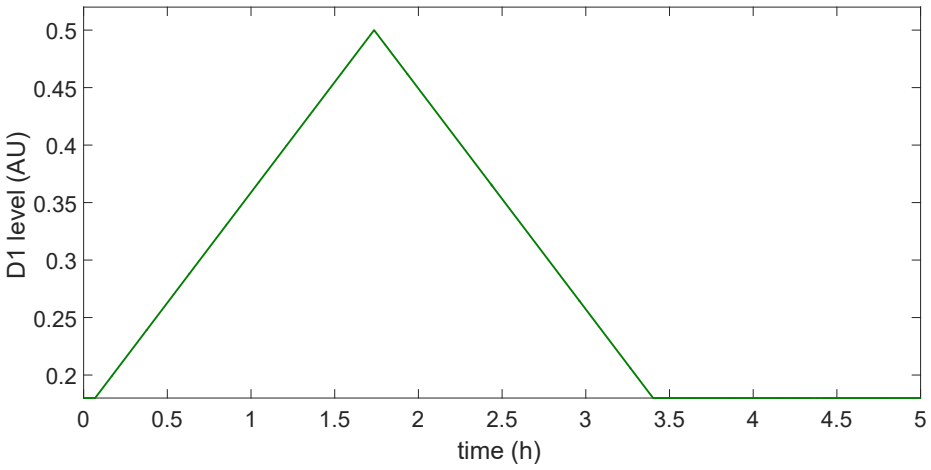
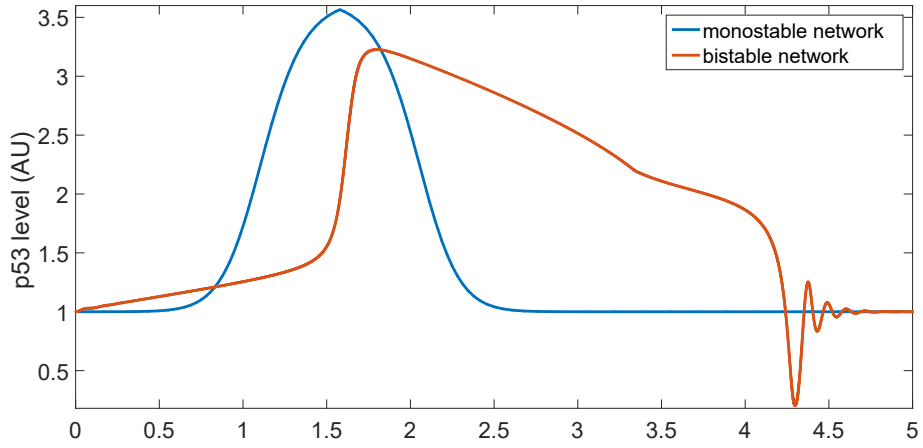


Figure 9.2: Bifurcation diagram of $\mathcal{N}_{int}^2$.



(a)



(b)

Figure 9.3: (a) The level of D_1 in our simulation. (b) The corresponding $p53$ response contrasting the behaviors of the monostable network $\mathcal{N}_i$ to the bistable network $\mathcal{N}_{int}^2$ as $[D_1]$ is varied. While the bistable switch has memory, the monostable switch does not.

A change in the parameters corresponding to the rate of initiation of $p53$ and ubiquitination of $MDM2$ causes T_0 to become less than 0. In this case, the network exhibits an irreversible switch from the set of points in Region "A" to those in Region "B". This is illustrated in Figures 9.4–9.5.

Network $\mathcal{N}_{int}^3$ can also behave as a bistable switch with respect to $[D_1]$ provided that $[D_2]$

also increases along with $[D_1]$. This is a reasonable assumption, since both $[D_1]$ and $[D_2]$ will increase with DNA damage. Qualitatively, this network admits the same behavior as $\mathcal{N}_{int}^2$. The bifurcation diagram for this network is provided in the Appendix B.1.2.

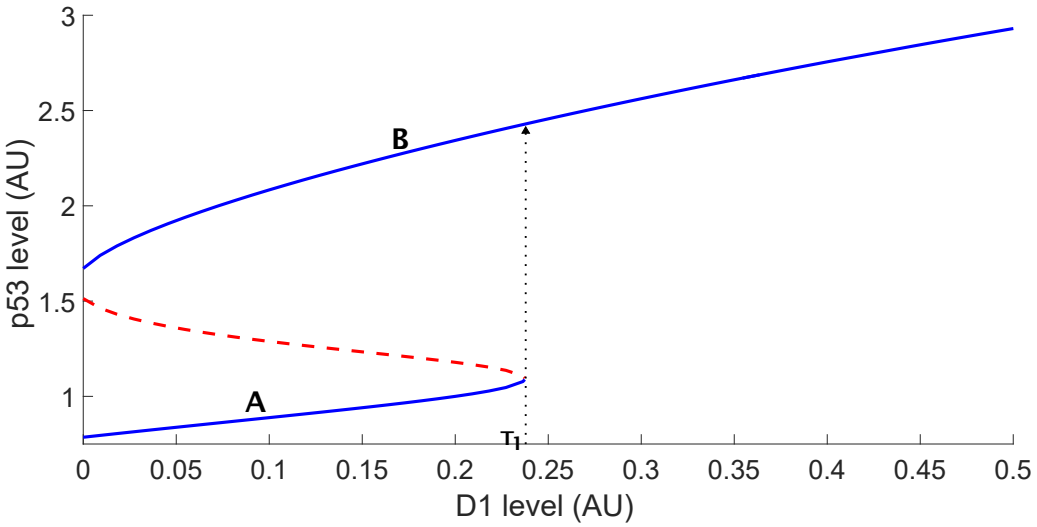
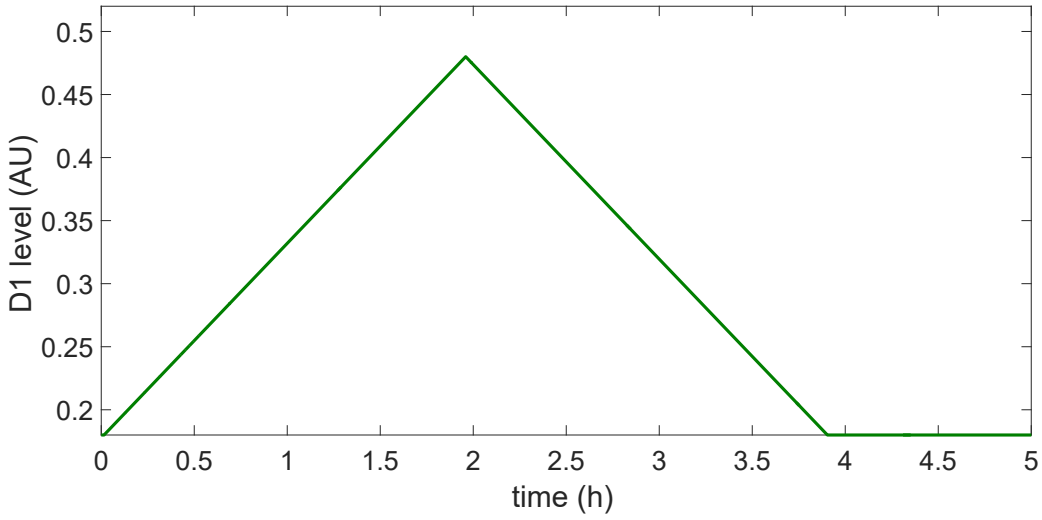
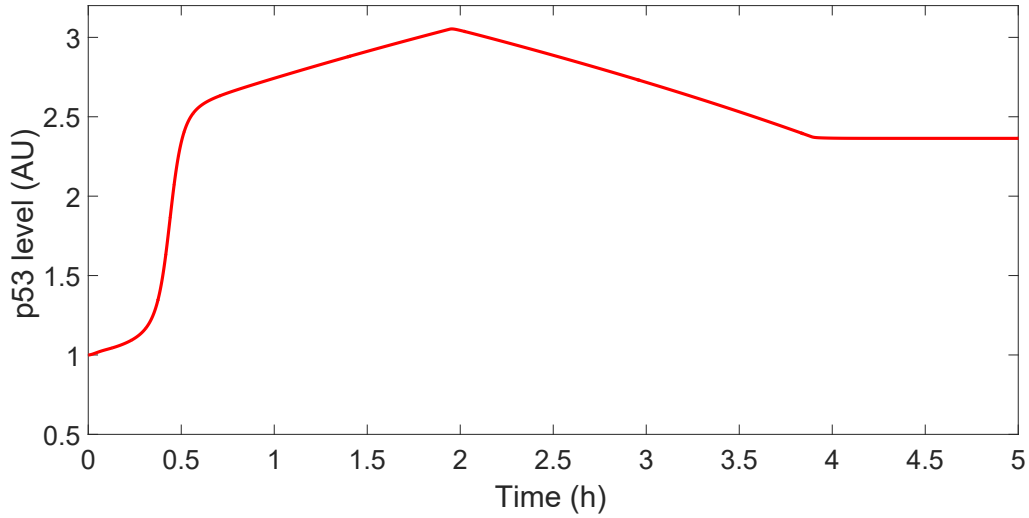


Figure 9.4: Bifurcation diagram of $\mathcal{N}_{int}^2$, illustrating its ability to behave as an irreversible bistable switch.



(a)



(b)

Figure 9.5: (a) The level of D_1 in our simulation.(b) The corresponding response of $p53$ demonstrating the irreversible switch.

Unlike the monostable networks, the bistable networks allow for the $p53$ levels to remain high for a sustained period of time once the damage sensor level has crossed threshold T_1, even if the damage sensor level subsequently falls below this level. This provides enough time for $p53$ to induce pathways that bring about cell-cycle arrest and DNA repair. $p53$ can also

induce pathways that bring about apoptosis or senescence, which are typically slower than repair pathways [47]. As such, it is more likely that these pro-elimination pathways will be induced either if the damage sensor level remains high for sufficiently long, or if the network operates as an irreversible bistable switch, since a sustained high levels of $p53$ expression is known to bring about apoptosis [91, 97, 130].

9.1.3 $\mathcal{N}_v$ admits tri-stable behavior

The addition of $\mathcal{M}_4$ to $\mathcal{N}_{int}^2$, or $\mathcal{M}_3$ to $\mathcal{N}_{int}^3$, can bring about *tri-stability* in the $p53$ levels with respect to $[D_1]$. That is, there are three sets of stable steady-states between which the $p53$ level can toggle, which we label "A", "B" and "C".

The bifurcation diagram illustrating the steady-states of $p53$ in this network $\mathcal{N}_v$ is shown in Figure 9.6. From this diagram, we observe that the original switching behavior which was a property of the bistable networks is retained; when $[D_1]$ crosses T_1 the $p53$ level switches to a high value, and this switch is "turned-off" only when $[D_1]$ falls below T_0. In addition to this behavior, the tri-stable network admits a third threshold T_2, that can be seen in Figure 9.6. Once $[D_1]$ crosses T_1, $p53$ transitions *irreversibly* to a high value; no matter how $[D_1]$ changes following this transition, $p53$ will not return to its nominal level. This behavior is illustrated in Figure 9.7.

It is known that when the $p53$ concentration remains high for a sufficiently long time, the cell goes into senescence or apoptosis [97, 77, 111]. From the perspective of our network, the level of damage-sensor $[D_1]$ crossing T_2 is likely associated with a scenario in which too much damage has been sustained, causing an irreversible switch in the $p53$ level to a high value and eventually bringing about either of these two cell fates. From this perspective, the role of $[D_2]$ is important. As $[D_2]$ increases, the middle branch of stable equilibrium points (labeled 'B') "shrinks", as illustrated in Figure 9.6. For high damage when $[D_2]$ is high, the rate of production of ARF is high. In turn, the threshold T_2 decreases, meaning that the cell has a higher chance of going to senescence or apoptosis as compared to when damage is lower. This

is consistent with the known role of ARF in triggering apoptosis [107].

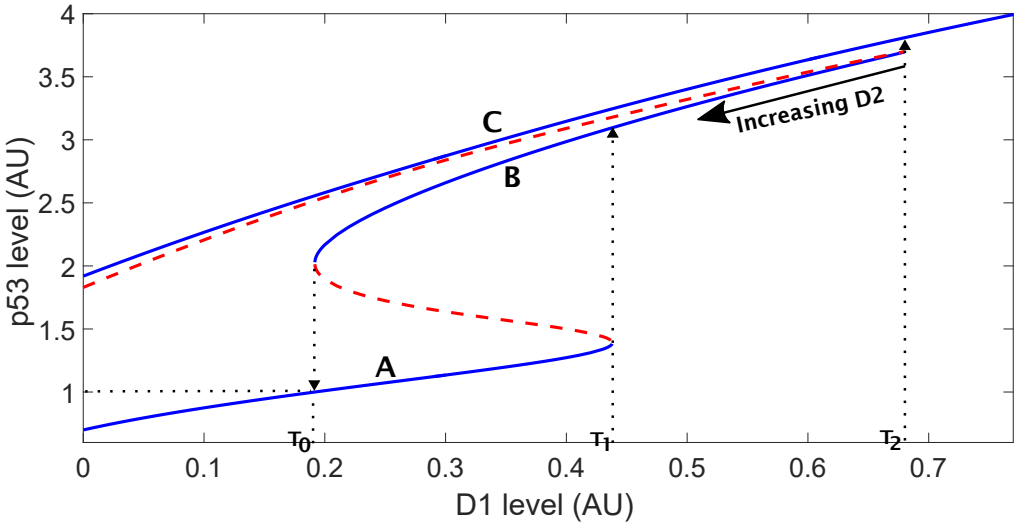
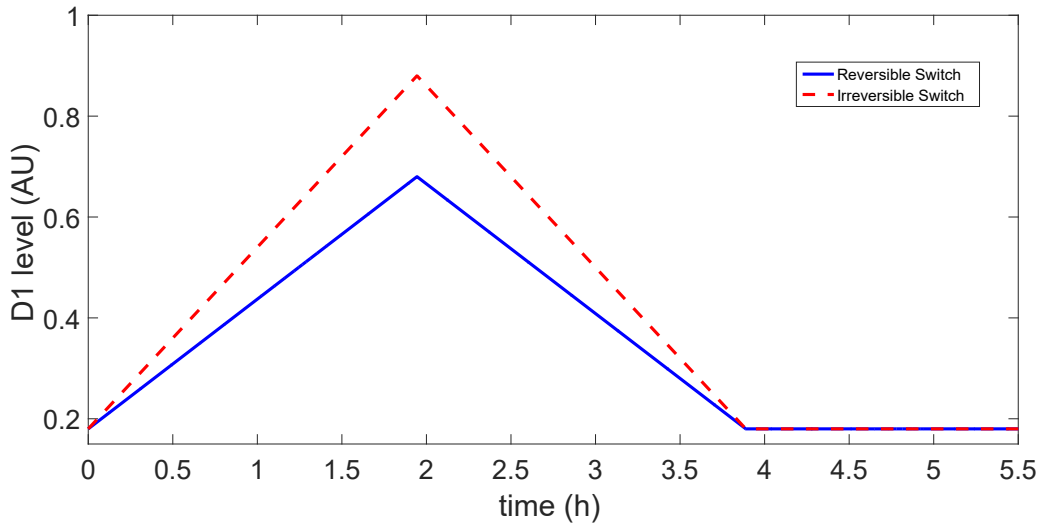
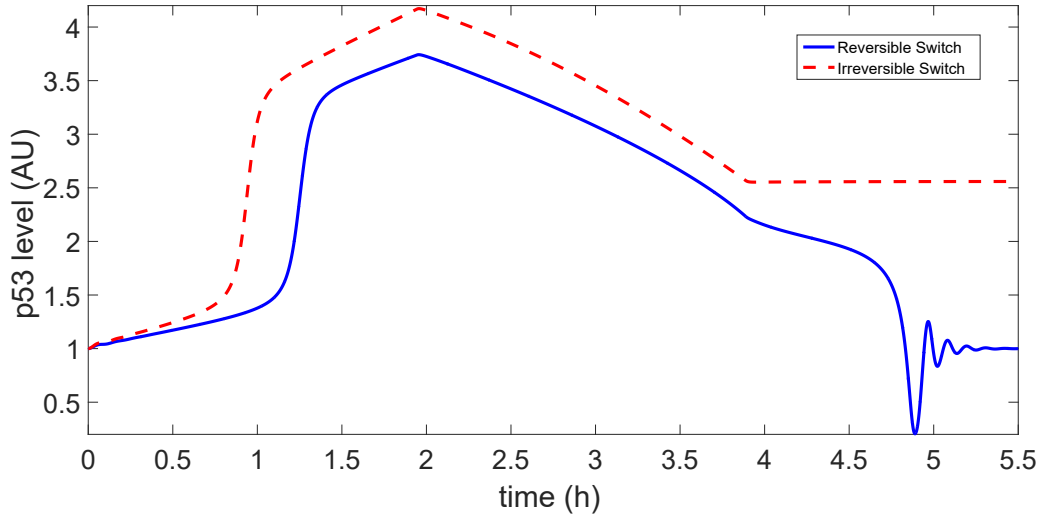


Figure 9.6: Bifurcation diagram of the network $\mathcal{N}_v$, illustrating the tristable switch. The switching behavior of the bistable network is retained. However, there is an additional threshold T_2 which when $[D_1]$ crosses, results in an irreversible switch of the $p53$ level to the set of points "C".



(a)



(b)

Figure 9.7: (a) Simulated D_1 levels, with $[D_2]$ kept fixed. (b) The resulting $p53$ behavior shows that the tri-stable network can behave both as a reversible and an irreversible switch.

9.2 Simple models of DNA damage transduction and repair explain experimentally observed results

In this section, we show how our tri-stable model for the $p53$ core regulation network in vertebrates $\mathcal{N}_v$, coupled with simple models of DNA damage transduction and repair, explains many experimentally observed results. We model the activity of the transduction kinases such as ATM and $CHK1$ using a single variable D_1, similar to what was done in [17]. We model the dynamics of D_1 as

$$[\dot{D_1}] = g_1(\#DSB) - g_2([p53])\,[D_1] - \gamma\,[D_1]\,. \tag{9.1}$$

In this equation, the variable $\#DSB$ represents the number of DNA double-strand breaks (DSBs), which activate D_1 through the function $g_1(\#DSB)$. D_1 then goes on to activate $p53$ and cause the accelerated degradation of $MDM2$ [114, 126], per the model in Figure 7.7. The feedback term $g_2([p53])$ represents negative feedback between $p53$ and the transduction kinases [17, 16], and the rate constant γ represents the $p53$-independent degradation of the kinases. We assume that the dynamics of (9.1) is slow compared to the dynamics of the $p53$ core regulation network, which is a key element in explaining the experimentally observed behavior.

We model the dynamics of $\#DSB$ as a simple deterministic death process,

$$\#\dot{D}SB = -\mu\#DSB\,[p53]\,, \tag{9.2}$$

where $\mu\#DSB\,[p53]$ is the rate of repair, that is known to be regulated by $p53$ [111, 85, 23]. Since damage is induced in the cell on the order of minutes [28], while the response can last for many hours or even days [48, 72], the number of DSBs induced by damaging agents is assumed to be the initial condition of (9.2). The specific parameters and functions from Equations (9.1)–(9.2) are given in Appendix B.

The parameter D_2 represents the activity of proteins such as $PARP1$ [92] and $E2F1$ [107],

which are known to play a role in activating ARF in response to DNA damage. We model D_2 as a simple monotonically increasing function of the number of DNA DSBs

$$[D_2] = g_3(\#DSB), \tag{9.3}$$

with the rate of production of ARF governed by D_2. The specific parameters and functions from Equations (9.1)–(9.3) are given in the Supplemental text. We then couple these models with $\mathcal{N}_v$ and run simulations.

In Figure 9.8, we show that the $p53$ levels remain low for damage resulting 5 DNA DSBs, which is a low level of DNA damage. In this case, the DNA damage is repaired before a $p53$ pulse is initiated. For damage resulting in 10 DSBs, a single $p53$ pulse is observed as shown in Figure 9.9. This corresponds to experimental results from [72], which show that 0.3Gγ of damage is sufficient to initiate a pulse, where 1Gγ of damage results in between 30-35 DSBs [100, 13, 48]. In Figure 9.10, our simulation shows that DNA damage that brings about 150 DSBs causes the initiation of multiple $p53$ pulses for over 60 hours until the damage is repaired, matching experimental results that suggest that the number of pulses increases as the number of DSBs increases [72, 48, 71], and that these pulses could last for days after damage is induced [48]. All pulses in our simulation have a period of between 4-7 hours, matching these experimental results in human breast cancer cells.

Early modeling work suggested that these pulses were primarily a result of the negative feedback interaction between $p53$ and $MDM2$, along with some positive feedback or delays [81, 124, 28, 131, 27]. However, more recent experimental results, in which pulses in $p53$ are observed to be highly coupled with pulses in the upstream transduction kinases, suggest that $p53$ pulses are brought about by the interaction of $p53$ with these kinases [17, 16]. Another important observation from [17] is that transient activation of the kinases results in a pulse of p53, suggesting that there is an excitable mechanism controlling p53 pulses [17, 16]. It was noted that while a single negative feedback loop between $p53$ and these kinases (with delays) would be sufficient to generate sustained oscillations, other $p53$ interactions would be necessary

to explain the observed excitability.

In our model, the interaction between D_1 and $\mathcal{N}_v$ can be described as an excitable mechanism known as a relaxation oscillator [61]. The slow negative feedback between D_1 and $p53$ described in Equation 9.1 operates over the faster dynamics of the $p53$ core regulation network. The pulses in $p53$ are observed because of the resulting cyclic transfers between the $p53$ equilibrium points in Regions 'A' and 'B' (Figure 9.6). The pulses in D_1 levels are also seen to be highly coupled with the pulses in $p53$ levels. In this way, our simulation results strongly matches both experimental observations, and hypotheses about the underlying mechanism that brings about the observed behavior without recourse to additional mechanisms or interactions.

In Figure 9.11, our simulation shows the $p53$ levels monotonically elevating to a high level for a large number of DNA DSBs, and remain high without pulsing. Even with the progression of the DNA repair process, the $p53$ levels remain high, due to the irreversible switch permitted by the tri-stable model. This behavior matches recent experimental observations in human U2-OS cells, in which the $p53$ pulsing is observed for moderate levels of DNA damage, but very high levels of DNA damage trigger a strong monotonic elevation of p53 level which doesn't return to a low level after rising [25].

It is worth noting that the experimental results are from different cells. As such, our goal is to show that the experimentally observed behavior is possible, and not to identify specific parameter values for the different cell types.

9.3 Discussion

In this section, we first discuss some of the features of our model of the full network that are novel as compared to previous models, and how these features allow us to predict the various experimentally observed behaviors. We then discuss the features of this network that could not have been present in putative primordial organisms, and discuss how the tumor suppression strategy might have been different in these organisms. Finally, we conclude the chapter with

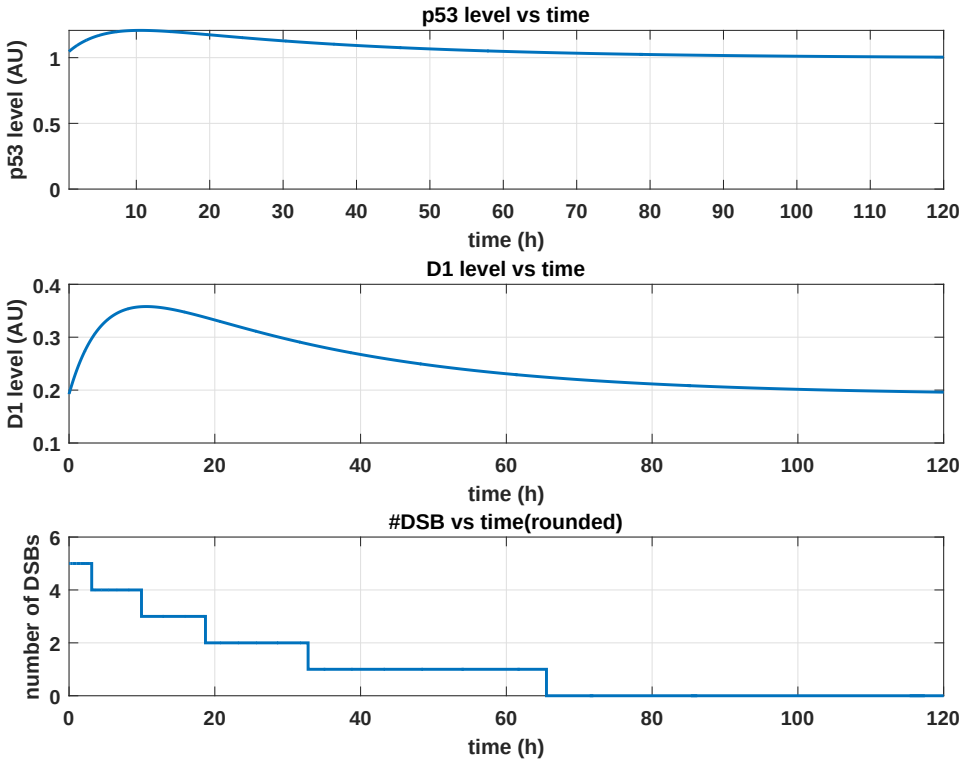


Figure 9.8: $p53$ and D_1 responses, along with DSB repair process for initial $\#DSB = 5$ in $\mathcal{N}_v$.

some possible future directions.

9.3.1 A novel model to explain experimentally observed behavior

Some of the earlier $p53$ modeling work was carried out under the assumption that the transduction kinase levels were proportional to the amount of damage, and that pulses were primarily observed due to the negative feedback between $p53$ and $MDM2$ [81, 124]. Other modeling work, while considering the positive feedback loop, assumed that the upstream kinase concentrations were proportional to damage and caused autonomous pulsing in $p53$ when the kinase levels were sufficiently high [28, 131, 27, 98]. In contrast, our model takes into account recent experimental results which suggest that upstream transduction kinases are responsible for initiating $p53$ pulses, and that the pulses of these kinases are highly coupled with pulses of

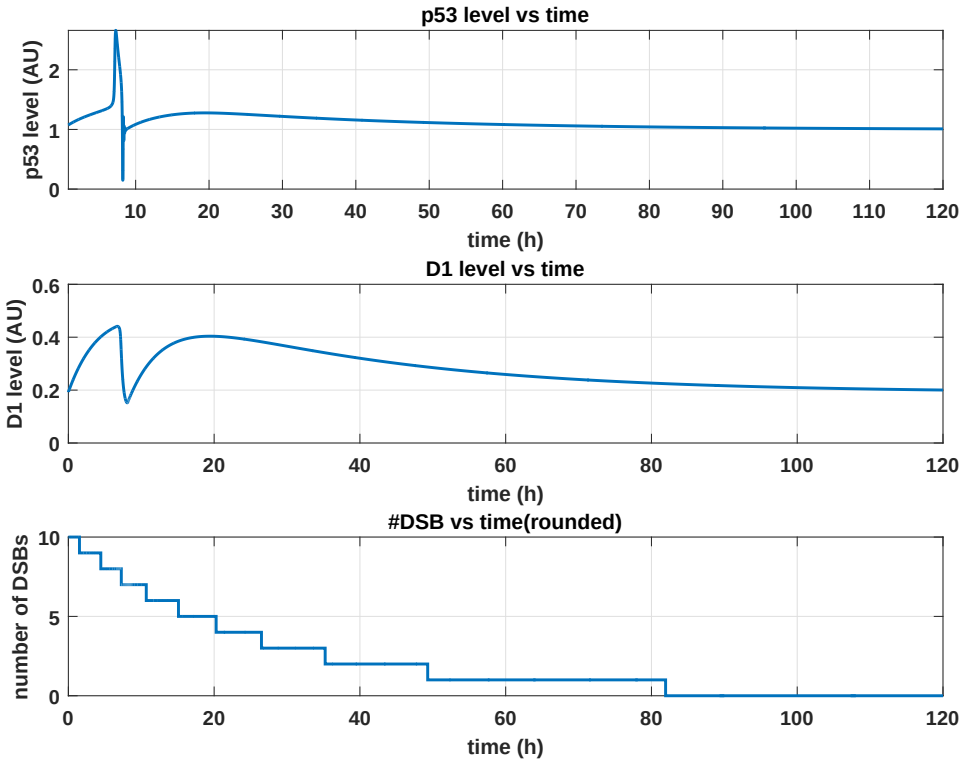


Figure 9.9: $p53$ and D_1 responses, along with DSB repair process for initial $\#DSB = 10$ in $\mathcal{N}_v$.

$p53$ [17, 16]. Moreover, these previous models primarily explain plausible mechanisms for the observation of oscillations in $p53$, and do not explore the irreversible switch to a high level.

There are other models of the $p53$ core regulation network in humans that display both the pulsing and sustained high levels of $p53$ [98, 126, 132] as described in Section 9.2. However, our model matches recent experimental observations that demonstrate pulsing in $p53$ for a moderate amount of damage, and a strong monotonic elevation in $p53$ for high levels of damage [25], contrasting previous models which imply that $p53$ would need to pulse before switching to a high value [98, 126, 132].

To our knowledge, our model of $p53$ core regulation in humans is the first where the irreversible switch of $p53$ to a high level is an inherent property of the tri-stable network structure, as opposed to a consequence of temporal behavior. It is also worth noting that all the observed behavior can be explained precisely using the dynamical properties of the ODEs, and do not

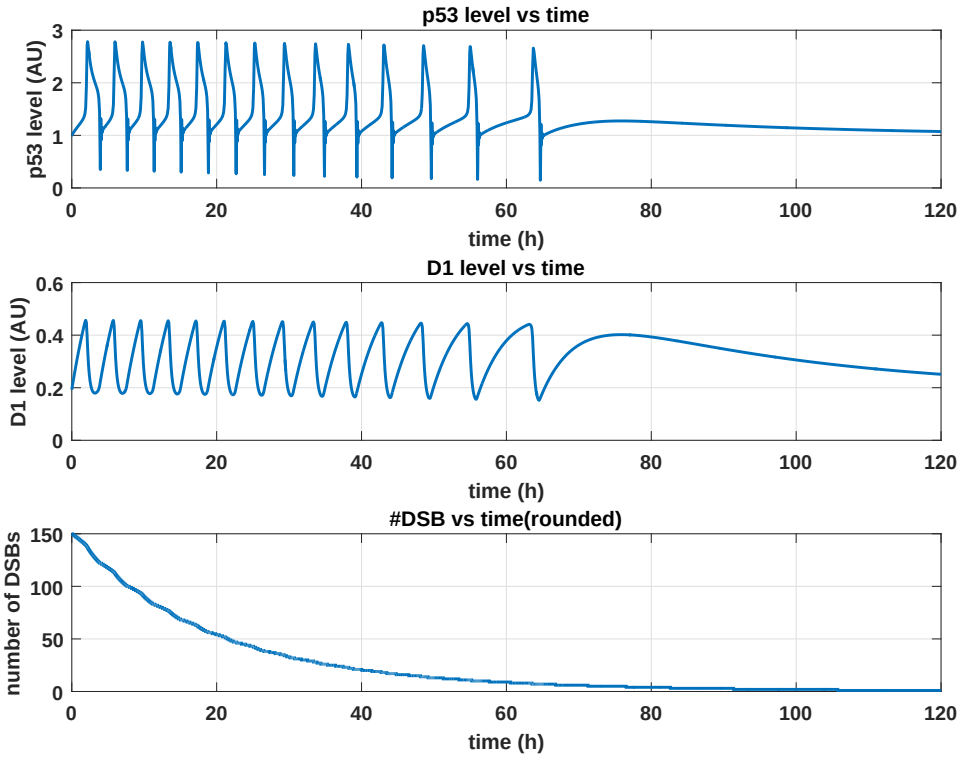


Figure 9.10: $p53$ and D_1 responses, along with DSB repair process for initial $\#DSB = 150$ in $\mathcal{N}_v$.

necessitate the artificial introduction of external signals, switches or clamps.

A novel behavior that our model predicts is the role played by ARF in bringing about an irreversible cell fate. It has long been postulated that ARF enhances apoptosis by sequestering away $MDM2$, thereby causing an increase in $p53$ levels. Although models to capture this behavior have been proposed [95, 68], the details of the underlying dynamics are still largely unknown [108]. A recent study has further pointed to a novel mechanism through which $MDM2$ can inhibit ARF [121]. In our model, this mutually inhibitory feedback between active $MDM2$ and ARF is crucial in bringing about tri-stability. Our model predicts that an increased rate of production of ARF is an important factor in bringing about the irreversible switch of $p53$ to a high level. In this way, our model proposes a novel but mathematically sound reasoning for how ARF might behave as an apoptosis enhancer.

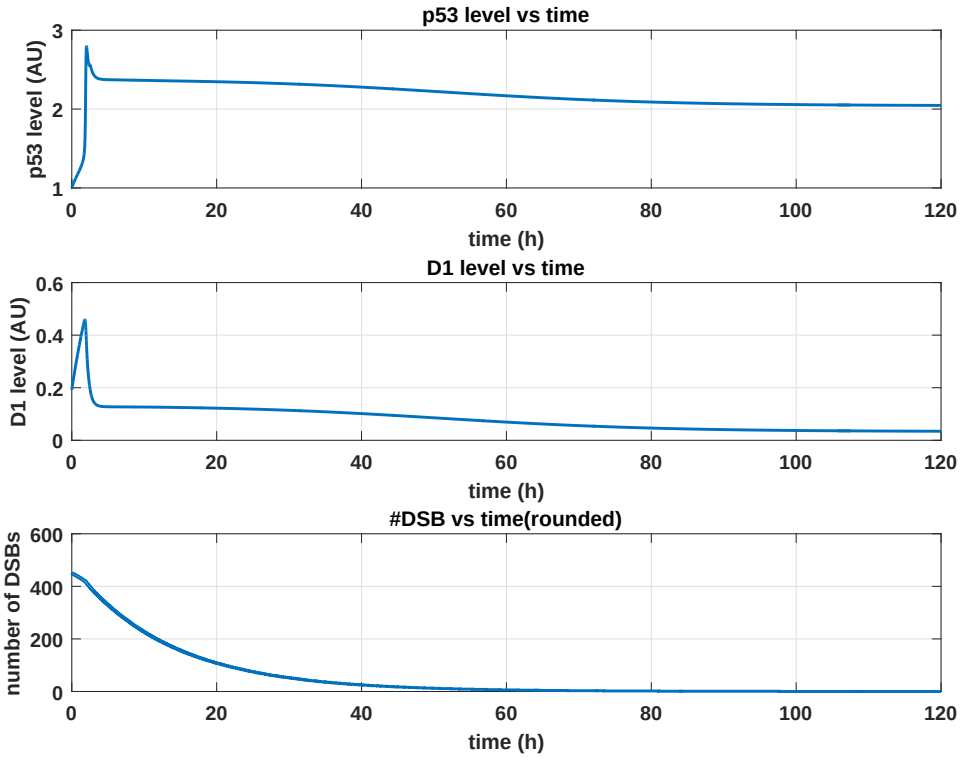


Figure 9.11: $p53$ and D_1 responses, along with DSB repair process for initial $\#DSB = 450$ in $\mathcal{N}_v$.

9.3.2 Tumor suppression strategy in humans

We recall that from an evolutionary standpoint, a cell has to achieve a balance between minimizing the number of Type 1 errors, while eradicating Type 2 errors. Type 1 errors occur when the Statistical Null Hypothesis, "Cells can recover and operate normally even after DNA damage crosses the apoptotic threshold" is true but the cells are killed anyway, while Type 2 errors occur when the Null Hypothesis is false, but the cells are not killed. While Type 1 errors could have metabolic costs in that too many cells are killed, Type 2 errors are potentially fatal.

From this perspective, the functional role of the $p53$ core regulation network in humans is clear. For low amounts of damage, the $p53$ level remains low. This is associated with the fact that the damage is deemed too low for a tumor to develop, and hence a significant $p53$ response is not initiated. For moderate amounts of damage, $p53$ exhibits pulsatile behavior, with the number of pulses proportional to the amount of damage. At this point, the damage

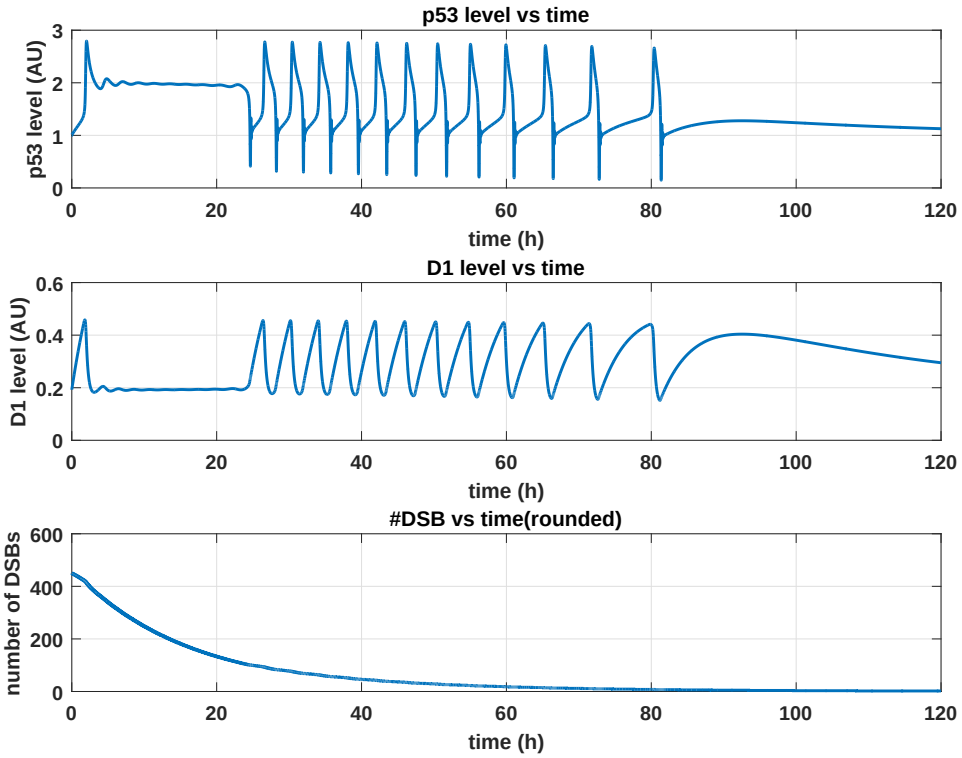


Figure 9.12: $p53$ and D_1 responses, along with DSB repair process for initial $\#DSB = 450$ in $\mathcal{N}_{int}^2$.

is deemed large enough to be repaired, triggering pulses in $p53$ until all the damage has been repaired. This is consistent with experimental results which show that majority of cells that exhibit pulsed p53 are able to grow and divide after recovery from DNA damage [97]. For a sufficiently large amount of damage, the $p53$ level switches monotonically to a high level and remains high. In cells that exhibit sustained high levels of p53 signaling, most cells go into senescence or apoptosis [97, 25, 130]. In this way, the cell is able to minimize the number of Type 1 errors by repairing DNA damage for a range of moderate levels of damage, but also has a clear threshold to bring about an irreversible switch to apoptosis when the amount of damage sustained is too high.

9.3.3 Tumor suppression strategy in other organisms

The networks $\mathcal{N}_{int}^2$ and $\mathcal{N}_{int}^3$ were found to admit both reversible and irreversible bistable switching behavior. It is worth noting that these networks are not capable of exhibiting tristability regardless of the values of the model parameters.

For the reversible switch, the $p53$ dynamics of these networks would be qualitatively similar to that of $\mathcal{N}_v$ in the presence of low and moderate amounts of DNA damage when coupled with the transducer and repair dynamics from Section 9.2. When the DNA damage is very high, the $p53$ level monotonically elevates to a high level. However, this switch is reversible if the DNA damage is repaired before the cell goes into senescence or apoptosis. This is illustrated for $\mathcal{N}_{int}^2$ in Figure 9.12. At this point, however, the integrity of the repair can be brought into question since the cell has likely experienced a large amount of damage. Hence, while the reversible bistable switches can minimize Type 1 errors by repairing moderate levels of damage, the elimination of Type 2 errors in the bistable networks requires the fast initiation of apoptotic pathways once the $p53$ level becomes high.

In the case of the irreversible switch, $p53$ switches irreversibly to a high level once $[D_1]$ crosses threshold T_1. However, in this case $p53$ is not able to display a phase of pulsatile behavior, increasing the chance of the cell going to senescence or apoptosis. Hence, this policy would be very effective in eliminating Type 2 errors, but could also increase the frequency of Type 1 errors. In essence, the bistable network is either able to admit $p53$ pulses or an irreversible switch of $p53$ to a high level triggering apoptosis, but not both phases of behavior as the tristable network is able to. These networks would probably have been more likely to behave as irreversible switches, since the elimination of Type 2 errors is crucial in the cell's ability to avoid tumors.

The network configurations $\mathcal{N}_i$ and $\mathcal{N}_{int}^1$ will only admit a single continuum of stable equilibrium points [79], and hence can only exhibit ultrasensitive switching behavior. In the absence of more complex mechanisms, the $p53$ levels are incapable of pulsing. Moreover, when the $p53$ level is high, it is more sensitive to changes in D_1 levels than the bistable switch. In these

organisms, it is therefore likely that the exponential phase of switching happens for reasonably low amounts of damage, and that the pro-apoptotic pathways were very sensitive to rising levels of $p53$, to ensure that Type 2 errors do not occur. This hypothesis is consistent with an earlier proposal on the operation of the $p53$ family of genes in early organisms [18].

9.4 Conclusion

In this chapter, we discovered how the different network configurations play a role in determining the $p53$ switching behaviors admitted, and therefore the $p53$ response to DNA damage. This study further revealed a novel method by which ARF could play the role of an apoptosis enhancer [107]. By incorporating our $p53$ network models into a simple model for DNA damage induction, transduction and repair, our new model for $p53$ response to DNA damage predicted many recent experimental results in a way that previous models were unable to do. Specifically, our model predicts that for a moderate amount of damage, $p53$ will pulse, with the number of pulses proportional to the amount of damage. When the damage becomes too high, the $p53$ level elevates monotonically to a high level and remains there. This matches early experiments on $p53$ in human breast cancer cells [72, 48, 17], and also more recent experiments on human U2-OS cells [25]. Our model also predicts that the upstream transduction kinases are responsible for activating $p53$ pulses when damage is present and that pulses in kinase levels are highly coupled with $p53$, matching recent experimental observations [17, 16].

We then discussed the relationship between the dynamical behaviors admitted by the different network configurations, and the potential implications of our findings with respect to the fitness and behavior of different organisms. In conclusion, we believe that one adaptive value of the $p53$ core regulation network in vertebrates is its ability to achieve a balance between eradicating Type 2 errors and minimizing Type 1 errors, and this behavior is clearly observed in experiments on human cells. On the other hand, the alternative network configurations in putative primordial organisms would not be able to exhibit this range of behaviors that are

possible in vertebrates and hence, would not be able to achieve this trade-off.

Chapter 10

Future Work

In Part II, we studied the evolution of the $p53$ core regulation network by postulating a path to complexity from primordial organisms to evolved vertebrates using evolutionary data from some invertebrate organisms. Recent studies have shown that some invertebrates might have carried a homolog MDM gene, prior to its duplication to form $MDM2$ 440 million years ago. Studies on invertebrates such as the *M. Trossulus* [89], *Enteropneusta* [87], *Strongylocentrotus purpuratus* and *Nematostella vectensis* [73] suggest that these invertebrates, this MDM homolog might have developed the ability to negatively regulate $p53$. While the interactions between this MDM homolog and $p53$ were analyzed deeply, further studies will be required to see if MDM would be up-regulated by $p53$, and if it could interact with $PTEN$ and ARF. This would then provide a higher resolution and further insight into the evolution of the core regulation network.

Further studies can also be carried out on the evolution of the binding sites of all the core-regulation network proteins. This would involve using bioinformatics approaches such as the BLAST algorithm to identify the presence or absence of different binding sites between the core regulation species in primordial organisms, and their rates of evolution. A key challenge in doing this would be that many of the reaction mechanisms are only partially known. For example, it is known that there are many binding sites between $MDM2$ and ARF, and further studies are needed to precisely model how important each of these binding sites are for their

interaction.

It was recently discovered that cells in elephants tend to go to apoptosis at a much higher rate than human cells, for the same amounts of DNA damage [1]. The fact that the cells in elephants have 20 copies of the $TP53$ gene, as compared to a single copy in humans, means that the $p53$ levels in elephant cells would probably be more sensitive to DNA damage; this in turn means that the irreversible switch of $p53$ to a high level (Figure 9.7) happens for much lower levels of damage as observed. Further study is required to understand how this adds adaptive value to the fitness of elephants. This can also pave the ways for studying how the same vertebrate core regulation network can operate in distinct ways in different vertebrates.

Finally, further work is required to generalize the algorithm introduced in Chapter 8. In particular, work is ongoing to determine the computational complexity of the algorithm when networks are partitioned in different ways. In our case, the modules were chosen by determining the evolutionary paths to complexity. Further work is needed on how to choose the modules over which the algorithm computes the network equilibrium points. From this study, we seek to find the optimal partition that minimizes the computational complexity of the algorithm.

Bibliography

[1] Lisa M Abegglen, Aleah F Caulin, Ashley Chan, Kristy Lee, Rosann Robinson, Michael S Campbell, Wendy K Kiso, Dennis L Schmitt, Peter J Waddell, Srividya Bhaskara, et al. Potential mechanisms for cancer resistance in elephants and comparative cellular response to dna damage in humans. *JAMA*, 314(17):1850–1860, 2015.

[2] Uri Alon. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.*, 8(6):450–461, 2007.

[3] J C Ameisen. On the origin, evolution, and nature of programmed cell death: a timeline of four billion years. *Cell death and differentiation*, 9(4):367–93, 2002.

[4] James Anderson, Yo-Cheng Chang, and Antonis Papachristodoulou. Model decomposition and reduction tools for large-scale networks in systems biology. *Automatica*, 47(6):1165–1174, June 2011.

[5] M. Andrec, B. N. Kholodenko, R. M. Levy, and E. D. Sontag. Inference of signaling and gene regulatory networks by steady-state perturbation experiments: structure and accuracy. *J. Theoret. Biol.*, 232(3):427–441, 2005.

[6] David Angeli. Boundedness analysis for open chemical reaction networks with mass-action kinetics. *Natural Computing*, 10(2):751–774, 2011.

[7] 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.

[8] David Angeli and Eduardo Sontag. Interconnections of monotone systems with steady-state characteristics. In Michael Malisoff Marcio de Queiroz and Peter Wolenski, editors, *Lect. Notes Contr. Inf.*, pages 135–154. Springer, 2004.

[9] David Angeli and Eduardo D Sontag. Monotone control systems. *IEEE T. Automat. Contr.*, 48(10):1684–1698, 2003.

[10] David Angeli and Eduardo D Sontag. Multi-stability in monotone input/output systems. *Syst. Control Lett.*, 51(3):185–202, 2004.

[11] Murat Arcak and Eduardo Sontag. A passivity-based stability criterion for a class of biochemical reaction networks. *Math Biosci. Eng.*, 5(1):1, 2008.

[12] Murat Arcak and Eduardo D Sontag. Diagonal stability of a class of cyclic systems and its connection with the secant criterion. *Automatica*, 42(9):1531–1537, 2006.

[13] Aroumougame Asaithamby and David J Chen. Cellular responses to dna double-strand breaks after low-dose γ-irradiation. *Nucleic acids research*, 37(12):3912–3923, 2009.

[14] Ruth Lev Bar-Or, Ruth Maya, Lee A Segel, Uri Alon, Arnold J Levine, and Moshe Oren. Generation of oscillations by the p53-mdm2 feedback loop: a theoretical and experimental study. *Proceedings of the National Academy of Sciences*, 97(21):11250–11255, 2000.

[15] Yaacov Barak, Tamar Juven, Rebecca Haffner, and Moshe Oren. mdm2 expression is induced by wild type p53 activity. *The EMBO journal*, 12(2):461, 1993.

[16] Eric Batchelor, Alexander Loewer, and Galit Lahav. The ups and downs of p53: understanding protein dynamics in single cells. *Nature Reviews Cancer*, 9(5):371–377, 2009.

[17] Eric Batchelor, Caroline S Mock, Irun Bhan, Alexander Loewer, and Galit Lahav. Recurrent initiation: a mechanism for triggering p53 pulses in response to dna damage. *Molecular cell*, 30(3):277–289, 2008.

[18] Vladimir A Belyi, Prashanth Ak, Elke Markert, Haijian Wang, Wenwei Hu, Anna Puzio-Kuter, and Arnold J Levine. The origins and evolution of the p53 family of genes. *Cold Spring Harbor perspectives in biology*, 2(6):a001198, 2010.

[19] Carol Bernstein, Anil R Prasad, Valentine Nfonsam, and Harris Bernstein. DNA Damage , DNA Repair and Cancer. *New Research Directions in DNA Repair*, pages 413–466, 2013.

[20] George Edward Briggs and John Burdon Sanderson Haldane. A note on the kinetics of enzyme action. *Biochem. J.*, 19(2):338, 1925.

[21] Frank J Bruggeman, Hans V Westerhoff, Jan B Hoek, and Boris N Kholodenko. Modular response analysis of cellular regulatory networks. *J. Theor. Biol.*, 218(4):507–520, 2002.

[22] N. Bellomo C. Bianca. *Towards a mathematical theory of complex biological systems*, volume 11 of *Series in Mathematical Biology and Medicine*. World Scientific Publishing Co. Pte. Ltd., 2011.

[23] Mauro AA Castro, Rodrigo JS Dalmolin, Jose CF Moreira, Jose CM Mombach, and Rita MC de Almeida. Evolutionary origins of human apoptosis and genome-stability gene networks. *Nucleic acids research*, 36(19):6269–6283, 2008.

[24] M. Chaves, T. Eissing, and F. Allgwer. Bistable Biological Systems: A Characterization Through Local Compact Input-to-State Stability. *Automatic Control, IEEE Transactions on*, 53:87–100, 2008.

[25] Xi Chen, Jia Chen, Siting Gan, Huaji Guan, Yuan Zhou, Qi Ouyang, and Jue Shi. Dna damage strength modulates a bimodal switch of p53 dynamics for cell-fate control. *BMC biology*, 11(1):73, 2013.

[26] Xinbin Chen, Linda J Ko, Lata Jayaraman, and Carol Prives. p53 levels, functional domains, and dna damage determine the extent of the apoptotic response of tumor cells. *Genes & development*, 10(19):2438–2451, 1996.

[27] Vijay Chickarmane, Animesh Ray, Herbert M Sauro, and Ali Nadim. A model for p53 dynamics triggered by dna damage. *SIAM Journal on Applied Dynamical Systems*, 6(1):61–78, 2007.

[28] Andrea Ciliberto, Béla Novak, and John J Tyson. Steady states and oscillations in the p53/mdm2 network. *Cell cycle*, 4(3):488–493, 2005.

[29] Jeff Clune, Jean-Baptiste Mouret, and Hod Lipson. The evolutionary origins of modularity. *P. Roy. Soc. B-Biol. Sci.*, 280(1755):20122863, 2013.

[30] Geoffrey M Cooper. *Elements of human cancer*. Jones & Bartlett Learning, 1992.

[31] Nicolas Courtois, Alexander Klimov, Jacques Patarin, and Adi Shamir. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In *Advances in Cryptology—EUROCRYPT 2000*, pages 392–407. Springer, 2000.

[32] Gheorghe Craciun and Martin Feinberg. Multiple equilibria in complex chemical reaction networks: Ii. the species-reaction graph. *SIAM J. Appl. Math*, 66(4):1321–1338, 2006.

[33] D. Del-Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.*, 4(161), 2008.

[34] Domitilla Del Vecchio and Shridhar Jayanthi. Retroactivity attenuation in transcriptional networks: Design and analysis of an insulation device. In *IEEE Decis. Contr. P.*, pages 774–780. IEEE, 2008.

[35] Domitilla Del Vecchio and Eduardo D Sontag. Synthetic biology: a systems engineering perspective. In Pablo A. Iglesias and Brian P. Ingalls, editors, *Control Theory and Systems Biology*, pages 101–124. The MIT Press, 2009.

[36] Tomislav Domazet-Lošo, Alexander Klimovich, Boris Anokhin, Friederike Anton-Erxleben, Mailin J. Hamm, Christina Lange, and Thomas C.G. Bosch. Naturally occurring tumours in the basal metazoan Hydra. *Nature Communications*, 5(May):1–8, 2014.

[37] Mirela Domijan and Markus Kirkilionis. Graph theory and qualitative analysis of reaction networks. *Netw Heterog Media*, 3(2):295–322, 2008.

[38] R. C. Dorf and R. H. Bishop. *Modern Control Systems*. Pearson Prentice-Hall, 2004.

[39] Alejo Efeyan and Manuel Serrano. p53: guardian of the genome and policeman of the oncogenes. *Cell cycle*, 6(9):1006–1010, 2007.

[40] H El Samad, D Del Vecchio, and M Khammash. Repressilators and promotilators: Loop dynamics in synthetic gene networks. In *P. Amer. Contr. Conf.*, pages 4405–4410. IEEE, 2005.

[41] Susan Elmore. Apoptosis: a review of programmed cell death. *Toxicologic pathology*, 35(4):495–516, 2007.

[42] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–8, January 2000.

[43] Drew Endy. Foundations for engineering biology. *Nature*, 438(7067):449–453, 2005.

[44] Carlos Espinosa-Soto and Andreas Wagner. Specialization can drive the evolution of modularity. *PLOS Comput. Biol.*, 6(3):e1000719, 2010.

[45] Elisa Franco, Eike Friedrichs, Jongmin Kim, Ralf Jungmann, Richard Murray, Erik Winfree, and Friedrich C Simmel. Timing molecular motion and production with a synthetic transcriptional clock. *P. Natl. Adad. Sci. USA*, 108(40):E784–E793, 2011.

[46] 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.

[47] FJ Geske, AC Nelson, R Lieberman, R Strange, T Sun, and LE Gerschenson. Dna repair is activated in early stages of p53-induced apoptosis. *Cell death and differentiation*, 7(4):393, 2000.

[48] Naama Geva-Zatorsky, Nitzan Rosenfeld, Shalev Itzkovitz, Ron Milo, Alex Sigal, Erez Dekel, Talia Yarnitzky, Yuvalal Liron, Paz Polak, Galit Lahav, et al. Oscillations and variability in the p53 system. *Molecular systems biology*, 2(1), 2006.

[49] Albert Goldbeter and Daniel E Koshland. An amplified sensitivity arising from covalent modification in biological systems. *P. Natl. Acad. Sci. USA*, 78(11):6840–6844, 1981.

[50] Mel Greaves. *Cancer: the evolutionary legacy*. Oxford University Press, 2001.

[51] Douglas Hanahan and Robert A Weinberg. The hallmarks of cancer. *cell*, 100(1):57–70, 2000.

[52] Sandra L Harris and Arnold J Levine. The p53 pathway: positive and negative feedback loops. *Oncogene*, 24(17):2899–2908, 2005.

[53] L. H. Hartwell, J. J. Hopfield, S. Liebler, and A. W. Murray. From molecular to modular cell biology. *Nature*, 402(6761):C47–52, December 1999.

[54] E Scott Helton and Xinbin Chen. p53 modulation of the dna damage response. *Journal of cellular biochemistry*, 100(4):883–896, 2007.

[55] J. P. Hespanha. *Linear Systems Theory*. Princeton Press, 2009.

[56] Khoo Kian Hoe, Chandra S Verma, and David P Lane. Drugging the p53 pathway: understanding the route to clinical efficacy. *Nature reviews Drug discovery*, 13(3):217–236, 2014.

[57] Yutaka Hori, Tae-Hyoung Kim, and Shinji Hara. Existence criteria of periodic oscillations in cyclic gene regulatory networks. *Automatica*, 47(6):1203–1209, 2011.

[58] Jaco H Houtgraaf, Jorie Versmissen, and Wim J van der Giessen. A concise review of dna damage checkpoints and repair in mammalian cells. *Cardiovascular Revascularization Medicine*, 7(3):165–172, 2006.

[59] He Huang, Christopher J Potter, Wufan Tao, Da-Ming Li, Walter Brogiolo, Ernst Hafen, Hong Sun, and Tian Xu. Pten affects cell size, cell proliferation and apoptosis during drosophila eye development. *Development*, 126(23):5365–5372, 1999.

[60] George Iliakis, YA Wang, Jun Guan, and Huichen Wang. Dna damage checkpoint control in cells exposed to ionizing radiation. *Oncogene*, 22(37):5834–5847, 2003.

[61] Brian P Ingalls. *Mathematical modeling in systems biology: an introduction*. MIT Press, 2013.

[62] K Itahana, G Dimri, and J Campisi. Regulation of cellular senescence by p53. *European journal of biochemistry/FEBS*, 268(10):2784, 2001.

[63] A. I. Ivanova. The condition for the uniqueness of the steady state of kinetic systems related to the structure of reaction scheme, part 1. *Kinet. Catal.*, 20:1019–1023, 1979.

[64] Hans-Michael Kaltenbach, Simona Constantinescu, Justin Feigelman, and Jörg Stelling. Graph-based decomposition of biochemical reaction networks into monotone subsystems. In *Algorithms in Bioinformatics*, pages 139–150. Springer, 2011.

[65] Guy Karlebach and Ron Shamir. Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology*, 9(10):770–780, 2008.

[66] Jason R. Kelly, Adam J. Rubin, Joseph H. Davis, Caroline M. Ajo-Franklin, John Cumbers, Michael J. Czar, Kim de Mora, Aaron L. Glieberman, Dileep D. Monie, and Drew Endy. Measuring the activity of BioBrick promoters using an in vivo reference standard. *Journal of Biological Engineering*, 3(1):4, 2009.

[67] Boris N Kholodenko, Anatoly Kiyatkin, Frank J Bruggeman, Eduardo Sontag, Hans V Westerhoff, and Jan B Hoek. Untangling the wires: a strategy to trace functional interactions in signaling and gene networks. *P. Natl. Acad. Sci. USA*, 99(20):12841–6, 2002.

[68] Do-Hyun Kim, Kyoohyoung Rho, and Sunghoon Kim. A theoretical model for p53 dynamics: identifying optimal therapeutic strategy for its activation and stabilization. *Cell Cycle*, 8(22):3707–3716, 2009.

[69] Kyung H Kim and Herbert M Sauro. Fan-out in gene regulatory networks. *Journal of Biological Engineering*, 4, 2010.

[70] M Kracikova, G Akiri, A George, R Sachidanandam, and SA Aaronson. A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis. *Cell Death & Differentiation*, 20(4):576–588, 2013.

[71] Galit Lahav. Oscillations by the p53-mdm2 feedback loop. In *Cellular Oscillatory Mechanisms*, pages 28–38. Springer, 2009.

[72] Galit Lahav, Nitzan Rosenfeld, Alex Sigal, Naama Geva-Zatorsky, Arnold J Levine, Michael B Elowitz, and Uri Alon. Dynamics of the p53-mdm2 feedback loop in individual cells. *Nature genetics*, 36(2):147–150, 2004.

[73] David P Lane and Chandra Verma. Mdm2 in evolution. *Genes & cancer*, 3(3-4):320–324, 2012.

[74] Michel Laurent and Nicolas Kellershohn. Multistability: a major means of differentiation and evolution in biological systems. *Trends in biochemical sciences*, 24(11):418–422, 1999.

[75] AJ Levine, W Hu, and Z Feng. The p53 pathway: what questions remain to be explored? *Cell Death & Differentiation*, 13(6):1027–1036, 2006.

[76] Arnold J Levine and Moshe Oren. The first 30 years of p53: growing ever more complex. *Nature Reviews Cancer*, 9(10):749–758, 2009.

[77] G Li and VC Ho. p53-dependent dna repair and apoptosis respond differently to high-and low-dose ultraviolet radiation. *British Journal of Dermatology*, 139(1):3–10, 1998.

[78] Yongfeng Li and Jeyaraman Srividhya. Goldbeter–koshland model for open signaling cascades: a mathematical study. *Journal of mathematical biology*, 61(6):781–803, 2010.

[79] Wendell Lim, Bruce Mayer, and Tony Pawson. *Cell Signaling: principles and mechanisms*. Taylor & Francis, 2014.

[80] Scott W Lowe, Earlene M Schmitt, Sallie W Smith, Barbara A Osborne, and Tyler Jacks. p53 is required for radiation-induced apoptosis in mouse thymocytes. *Nature*, 362(6423):847–849, 1993.

[81] Lan Ma, John Wagner, John Jeremy Rice, Wenwei Hu, Arnold J Levine, and Gustavo A Stolovitzky. A plausible model for the digital response of p53 to dna damage. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40):14266–14271, 2005.

[82] Ander Matheu, Antonio Maraver, and Manuel Serrano. The arf/p53 pathway in cancer and aging. *Cancer research*, 68(15):6031–6034, 2008.

[83] Lindsey D Mayo, Jack E Dixon, Donald L Durden, Nickolas K Tonks, and David B Donner. Pten protects p53 from mdm2 and sensitizes cancer cells to chemotherapy. *Journal of Biological Chemistry*, 277(7):5484–5489, 2002.

[84] Lindsey D Mayo and David B Donner. A phosphatidylinositol 3-kinase/akt pathway promotes translocation of mdm2 from the cytoplasm to the nucleus. *Proceedings of the National Academy of Sciences*, 98(20):11598–11603, 2001.

[85] Vijay Menon and Lawrence Povirk. Involvement of p53 in the repair of dna double strand breaks: Multifaceted roles of p53 in homologous recombination repair (hrr) and non-homologous end joining (nhej). In *Mutant p53 and MDM2 in Cancer*, pages 321–336. Springer, 2014.

[86] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[87] Jamil Momand, Alberto Villegas, and Vladimir A Belyi. The evolution of mdm2 family genes. *Gene*, 486(1):23–30, 2011.

[88] Stefan Müller, Josef Hofbauer, Lukas Endler, Christoph Flamm, Stefanie Widder, and Peter Schuster. A generalized model of the repressilator. *Journal of mathematical biology*, 53(6):905–937, 2006.

[89] Annette F Muttray, Teagan F O'Toole, Wendy Morrill, Rebecca J Van Beneden, and Susan A Baldwin. An invertebrate mdm homolog interacts with p53 and is differentially expressed together with p53 and ras in neoplastic mytilus trossulus haemocytes. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 156(4):298–308, 2010.

[90] David L Nelson, Albert L Lehninger, and Michael M Cox. *Lehninger principles of biochemistry*. Macmillan, 2008.

[91] Chris J Norbury and Boris Zhivotovsky. Dna damage-induced apoptosis. *Oncogene*, 23(16):2797–2808, 2004.

[92] Giulia Orlando, Svetlana V Khoronenkova, Irina I Dianova, Jason L Parsons, and Grigory L Dianov. Arf induction in response to dna strand breaks is regulated by parp1. *Nucleic acids research*, page gkt1185, 2013.

[93] Hamid R. Ossareh and Domitilla Del Vecchio. Retroactivity attenuation in signaling cascades. In *IEEE Decis. Contr. P.*, pages 2220–2226, 2011.

[94] Hamid R Ossareh, Alejandra C Ventura, Sofia D Merajver, and Domitilla Del Vecchio. Long signaling cascades tend to attenuate retroactivity. *Biophys. J.*, 100(7):1617–1626, 2011.

[95] Carole J Proctor and Douglas A Gray. Explaining oscillations and variability in the p53-mdm2 system. *BMC systems biology*, 2(1):75, 2008.

[96] SR Proulx and FR Adler. The standard of neutrality: still flapping in the breeze? *J. Evolution Biol.*, 23(7):1339–1350, 2010.

[97] Jeremy E Purvis, Kyle W Karhohs, Caroline Mock, Eric Batchelor, Alexander Loewer, and Galit Lahav. p53 dynamics control cell fate. *Science*, 336(6087):1440–1444, 2012.

[98] Krzysztof Puszyński, Beata Hat, and Tomasz Lipniacki. Oscillations and bistability in the stochastic model of p53 regulation. *Journal of Theoretical Biology*, 254(2):452–465, 2008.

[99] Paul WK Rothemund. Folding dna to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, 2006.

[100] Kai Rothkamm and Markus Löbrich. Evidence for a lack of dna double-strand break repair in human cells exposed to very low x-ray doses. *Proceedings of the National Academy of Sciences*, 100(9):5057–5062, 2003.

[101] Julio Saez-Rodriguez, Stefan Gayer, Martin Ginkel, and Ernst Dieter Gilles. Automatic decomposition of kinetic models of signaling networks minimizing the retroactivity among modules. *Bioinformatics*, 24(16):i213–i219, 2008.

[102] Julio Saez-Rodriguez, Andreas Kremling, Holger Conzelmann, Katja Bettenbrock, and Ernst Dieter Gilles. Modular analysis of signal transduction networks. *IEEE Contr. Syst. Mag.*, 24(4):35–52, 2004.

[103] Julio Saez-Rodriguez, Andreas Kremling, and Ernst Dieter Gilles. Dissecting the puzzle of life: modularization of signal transduction networks. *Comput. Chem. Eng.*, 29(3):619–629, 2005.

[104] H. M. Sauro. Modularity defined. *Mol. Syst. Biol.*, 4(166), March 2008.

[105] Douglas M. Schwarz. Fast and robust curve intersections. `http://www.mathworks.com/matlabcentral/fileexchange/11837-fast-and-robust-curve-intersections/content/intersections.m/`, 2010.

[106] Charles J Sherr. Principles of tumor suppression. *Cell*, 116(2):235–246, 2004.

[107] Charles J Sherr. Autophagy by arf: a short story. *Molecular cell*, 22(4):436–437, 2006.

[108] Charles J Sherr. Divorcing arf and p53: an unsettled case. *Nature Reviews Cancer*, 6(9):663–673, 2006.

[109] Charles J Sherr and Jason D Weber. The arf/p53 pathway. *Current opinion in genetics & development*, 10(1):94–99, 2000.

[110] Hari Sivakumar and Joao P Hespanha. Towards modularity in biological networks while avoiding retroactivity. In *P. Amer. Contr. Conf.*, pages 4550–4556. IEEE, 2013.

[111] Martin L Smith and Albert J Fornace Jr. The two faces of tumor suppressor p53. *The American journal of pathology*, 148(4):1019, 1996.

[112] James Sneyd and James Keener. *Mathematical physiology*. Springer-Verlag New York, 2008.

[113] V Stambolic, D MacPherson, D Sas, Y Lin, B Snow, Y Jang, S Benchimol, and TW Mak. Regulation of pten transcription by p53. *Molecular cell*, 8(2):317–325, 2001.

[114] Jayne M Stommel and Geoffrey M Wahl. Accelerated mdm2 auto-degradation induced by dna-damage kinases is required for p53 activation. *The EMBO journal*, 23(7):1547–1556, 2004.

[115] Natalja Strelkowa and Mauricio Barahona. Transient dynamics around unstable periodic orbits in the generalized repressilator model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(2):023104, 2011.

[116] V Sundarapandian. Global asymptotic stability of nonlinear cascade systems. *Appl. Math. Lett.*, 15(3):275–277, 2002.

[117] CD Thron. The secant condition for instability in biochemical feedback control—i. the role of cooperativity and saturability. *B. Math. Biol.*, 53(3):383–401, 1991.

[118] John J Tyson. Monitoring p53's pulse. *Nature genetics*, 36(2):113–114, 2004.

[119] John J Tyson and Hans G Othmer. The dynamics of feedback control circuits in biochemical pathways. *Progress in Theoretical Biology*, 5(1):62, 1978.

[120] Alejandra C Ventura, Jacques-A Sepulchre, and Sofía D Merajver. A hidden feedback in signaling cascades is revealed. *PLOS Comput. Biol.*, 4(3):e1000041, 2008.

[121] Maria Vivo, Maria Matarese, Maria Sepe, Rosaria Di Martino, Luisa Festa, Viola Calabrò, Girolama La Mantia, and Alessandra Pollice. Mdm2-mediated degradation of p14arf: A novel mechanism to control arf levels in cancer cells. *PloS one*, 10(2):e0117252, 2015.

[122] Bert Vogelstein, David Lane, and Arnold J Levine. Surfing the p53 network. *Nature*, 408(6810):307–310, 2000.

[123] Karen H Vousden and David P Lane. p53 in health and disease. *Nature Reviews Molecular Cell Biology*, 8(4):275–283, 2007.

[124] J Wagner, L Ma, JJ Rice, W Hu, AJ Levine, and GA Stolovitzky. p53–mdm2 loop controlled by a balance of its feedback strength and effective dampening using atm and delayed feedback. *IEE Proceedings-Systems Biology*, 152(3):109–118, 2005.

[125] Zhiwei Wang, Hiroyuki Inuzuka, Jiateng Zhong, Hidefumi Fukushima, Lixin Wan, Pengda Liu, and Wenyi Wei. Dna damage-induced activation of atm promotes β-trcp-mediated mdm2 ubiquitination and destruction. *Oncotarget*, 3(9):1026, 2012.

[126] Keng Boon Wee, Uttam Surana, and Baltazar D Aguda. Oscillations of the p53-akt network: implications on cell survival and death. *PLoS One*, 4(2):e4407, 2009.

[127] Jan C Willems. The behavioral approach to open and interconnected systems. *Control Systems, IEEE*, 27(6):46–99, 2007.

[128] Xiangwei Wu, J Henri Bayle, David Olson, and Arnold J Levine. The p53-mdm-2 autoregulatory feedback loop. *Genes & development*, 7(7a):1126–1132, 1993.

[129] Jianhua Xing and Jing Chen. The goldbeter-koshland switch in the first-order region and its response to dynamic disorder. *PLoS One*, 3(5):e2140, 2008.

[130] Kyoung Wan Yoon, Sanguine Byun, Eunjeong Kwon, So-Young Hwang, Kiki Chu, Masatsugu Hiraki, Seung-Hee Jo, Astrid Weins, Samy Hakroush, Angelika Cebulla, et al. Control of signaling-mediated clearance of apoptotic cells by the tumor suppressor p53. *Science*, 349(6247):1261669, 2015.

[131] Tongli Zhang, Paul Brazhnik, and John J Tyson. Exploring mechanisms of the dna-damage response: p53 pulses and their possible relevance to apoptosis. *Cell cycle*, 6(1):85–94, 2007.

[132] Xiao-Peng Zhang, Feng Liu, and Wei Wang. Two-phase dynamics of p53 in the dna damage response. *Proceedings of the National Academy of Sciences*, 108(22):8990–8995, 2011.

Appendix A

Module Characteristics

In this section, we present some properties of the modules introduced in Chapter 3.

A.1 Transcriptional regulation (TR) module

In Section 3.2.1, we presented the LTF of a TR module when Assumptions 1–2 were satisfied. A comprehensive explanation of how the dynamics of a TR module are simplified with Assumption 2 is provided in [33].

It is reasonably straightforward to compute the LTF of a TR module when only Assumption 1 is satisfied, and it is given by

$$H_j(s) = D_j \frac{q k^{\text{on}} \frac{K P^{\text{tot}} (\theta u^*)^q}{K + (\theta u^*)^q} \alpha_j \beta_j}{\Big(\theta u^* (s + k^{\text{off}})(s + \bar{\beta}) + k^{\text{on}} (\theta u^*)^q \big((q^2 F \frac{K P^{\text{tot}} (\theta u^*)^q}{K + (\theta u^*)^q} + \theta u^*) s + u^*\big)\Big)(s + \gamma_j)}$$

with

$$D_j \coloneqq \begin{cases} +1 & \text{if } \mathcal{S}_0 \text{ activates } \mathcal{G}_j \\ -1 & \text{if } \mathcal{S}_0 \text{ represses } \mathcal{G}_j, \end{cases} \qquad \theta \coloneqq \frac{1}{\bar{\beta}}.$$

This LTF now depends on the fan-out F. It is worth noting that the LTFs to each activating and to each repressing output are exactly the same because of Assumption 1. If this were not the case, the same method could be utilized to compute the LTF, but the expression would look different for each output. [110] contains more information on how to compute the LTF of the module given the ODEs.

A.2 Covalent modification (CM) module

The properties of CM modules are summarized in Section 3.2.2. The characteristics of these modules have been presented using two classical approximations to simplify the

reaction dynamics.

The form of the LTF obtained when using the quasi steady-state approximation (Assumption 4) was presented in Chapter 3, but not the LTF when using the equilibrium approximation (Assumption 3). Under Assumption 3, the LTF of a CM module is given by

$$\frac{K^{\mathrm{p}}(u^*)}{s + K^{\mathrm{q}}(u^*)}$$

where

$$K^{\mathrm{p}}(u^*) = \frac{E^{\mathrm{tot}} k^{\mathrm{cat}} K^{\mathrm{d}}}{(K^{\mathrm{d}} + [\mathcal{S}_0]^*)^2}$$
$$K^{\mathrm{q}}(u^*) = \frac{N}{\left(E^{\mathrm{tot}} K^{\mathrm{d}} + (K^{\mathrm{d}} + [\mathcal{S}_0]^*)^2\right)^2}$$

with

$$N = E^{\mathrm{tot}} K^{\mathrm{d}} \Big((K^{\mathrm{d}} + [\mathcal{S}_0]^*) \big(k^{\mathrm{cat}} (K^{\mathrm{d}} + [\mathcal{S}_0]^*) + \gamma (K^{\mathrm{d}} + 3\,[\mathcal{S}_0]^*) - 2u^* \big) + k^{\mathrm{cat}} E^{\mathrm{tot}} (K^{\mathrm{d}} + 2\,[\mathcal{S}_0]^*) \Big) + \gamma (K^{\mathrm{d}} + [\mathcal{S}_0]^*)^4$$

and

$$[\mathcal{S}_0]^* = \frac{-E^{\mathrm{tot}} k^{\mathrm{cat}} + u^* - \gamma K^{\mathrm{d}} + \sqrt{(-E^{\mathrm{tot}} k^{\mathrm{cat}} + u^* - \gamma K^{\mathrm{d}})^2 + 4u^* \gamma K^{\mathrm{d}}}}{2\gamma}.$$

A.3 PD-cycle module

The PD-cycle module was discussed in detail in Section 3.2.3. The equilibrium point of the module as a function of the constant inputs v_i^*, u_i^* are given by

$$[\mathcal{S}_{i+1}]^* = \frac{\bar{k}_i^{\mathrm{f}} (k_i^{\mathrm{r}} + \alpha_i) v_i^* (E_i^{\mathrm{tot}} \bar{\alpha}_i - u_i^*)}{k_i^{\mathrm{f}} \alpha_i (\bar{k}_i^{\mathrm{r}} + \bar{\alpha}_i) u_i^*}, \qquad [\mathcal{S}_i^{\dagger}{:}\mathcal{S}_{i+1}]^* = \frac{v_i^*}{\alpha_i}$$
$$[\mathcal{S}_i^{\dagger}]^* = \frac{u_i^* (\bar{k}_i^{\mathrm{r}} + \bar{\alpha}_i)}{\bar{k}_i^{\mathrm{f}} (E_i^{\mathrm{tot}} \bar{\alpha}_i - u_i^*)}, \qquad [E_i]^* = E_i^{\mathrm{tot}} - \frac{u_i^*}{\bar{\alpha}_i}$$

and the LTF around this equilibrium is given by

$$H(s) \coloneqq \frac{1}{D(s)} \begin{bmatrix} N_{11}(s) & N_{12}(s) \\ N_{21}(s) & N_{22}(s) \end{bmatrix}$$

where

$$
\begin{aligned}
N_{11}(s) &= k_i^{\mathrm{f}}[\mathcal{S}_i^{\dagger}]^*([E_i]^*\bar{k}_i^{\mathrm{f}}(s+\bar{\alpha}_i) + s(\bar{k}_i^{\mathrm{r}} + s + \bar{k}_i^{\mathrm{f}}[\mathcal{S}_i^{\dagger}]^* + \bar{\alpha}_i))\alpha_i \\
N_{12}(s) &= k_i^{\mathrm{f}} s[\mathcal{S}_{i+1}]^*(\bar{k}_i^{\mathrm{r}} + s + \bar{k}_i^{\mathrm{f}}[\mathcal{S}_i^{\dagger}]^* + \bar{\alpha}_i)\alpha_i \\
N_{21}(s) &= -[E_i]^*\bar{k}_i^{\mathrm{f}} k_i^{\mathrm{f}} s[\mathcal{S}_i^{\dagger}]^*\bar{\alpha}_i \\
N_{22}(s) &= [E_i]^*\bar{k}_i^{\mathrm{f}}\bar{\alpha}_i(k_i^{\mathrm{r}} s + (s + k_i^{\mathrm{f}}[\mathcal{S}_i^{\dagger}]^*)(s+\alpha_i)), \\
D(s) &= [E_i]^*\,\bar{k}_i^{\mathrm{f}}(s+\bar{\alpha}_i)\big(k_i^{\mathrm{r}} s + (s + k_i^{\mathrm{f}}\,[\mathcal{S}_i]^*)(s+\alpha_i)\big) + \\
&\qquad s\big(s + \bar{k}_i^{\mathrm{r}} + \bar{k}_i^{\mathrm{f}}\,[\mathcal{S}_i]^* + \bar{\alpha}_i\big)\big(s^2 + (k_i^{\mathrm{r}} + k_i^{\mathrm{f}}([\mathcal{S}_i^{\dagger}{:}\mathcal{S}_{i+1}]^* + [\mathcal{S}_i]^*) + \alpha_i)s + \alpha_i k_i^{\mathrm{f}}\,[\mathcal{S}_i]^*\,\big).
\end{aligned}
$$

Appendix B

$p53$ network model parameters

The parameters and functions for the $p53$ network are provided in this chapter.

B.1 Core regulation network models

The functions $f_i(x)$, $i \in \{1, 2, \cdots, 10, 11\}$ from the model, represented in the are given by

$$f_i(x) = \begin{cases} \frac{x^4}{x^4+j_i^4} & i \in \{4, 5\} \\ \frac{x}{x+j_i} & \text{otherwise.} \end{cases}$$

The functions $h_i(x)$, $i \in \{1, 2\}$, which represent the kinase activities, are given by

$$h_i(x) = c_i \frac{x^{n_i}}{x^{n_i} + k_i^{n_i}}.$$

B.1.1 Tristable network $\mathcal{N}_h$

The parameters for this network are given in Table B.1. Most parameters are consistent with the parameters and ranges from [126], except for the parameters relating to the ARF-$MDM2$ interactions and ARF reactions, that did not appear in [126]. Moreover the parameters relating to the entry of damage into the core regulation network through the kinases (p_1, p_4,c_1,c_2,k_1 and k_2) are chosen arbitrarily as was done in [126]. Moreover, while [126] chose to use μM as their unit of measurement, we used arbitrary unites (A.U.) in line with what is done in experimental studies [48, 17]. Since the amplitude of responses are known to vary significantly between different cells [72], the ratios between the parameters play a far more crucial role in bringing about the dynamical behavior as opposed to the parameters themselves. The steady-state solution to the network with these parameters results in the bifurcation diagram 9.6, which has been normalized with respect to the $p53$ level at T_0.

Parameter	Value	Units
b_1	0.0625	\h
b_2	22	/h
b_3	0.5	A.U./h
b_4	0.0375	A.U./h
b_5	0.006	A.U./h
b_6	0.15	A.U./h
b_7	73	/h
b_8	0.2	A.U./h
b_9	20	/h
b_{10}	1	/h
b_{11}	0.01	/h
j_1	0.01	A.U.
j_2	0.6	A.U.
j_3	0.1	A.U.
j_4	0.84	A.U.
j_5	1.19	A.U.
j_6	0.1	A.U.
j_7	0.5	A.U.
j_8	0.1	A.U.
j_9	0.1	A.U.
j_{10}	0.01	A.U.
j_{11}	0.01	A.U.

Parameter	Value	Units
d_1	0.02	/h
d_2	0.01	/h
d_3	0.005	/h
d_4	0.01	/h
d_5	0.0054	/h
d_6	0.001	/h
p_1	0.017	A.U./h
p_2	0.0009	A.U./h
p_3	0.02	/h
p_4	0.0039	/h
p_5	0.0009	A.U./h
p_6	0.02	/h
p_7	0.021	A.U./h
tot_1	1	A.U.
tot_2	1	A.U.
c_1	0.066	A.U./h
c_2	0.183	A.U./h
k_1	2	A.U.
k_2	3	A.U.
n_1	1	-
n_2	1	-

Table B.1: Parameters for core regulation network model in Network $\mathcal{N}_h$

B.1.2 Bistable networks $\mathcal{N}_{int}^2$ and $\mathcal{N}_{int}^3$

The study of $\mathcal{N}_{int}^2$ was performed with the same parameters as shown in Table B.1, with the external input u_{24} set to 0 (although the results could be generalized to any u_{24}). This results in the bifurcation diagram given by Figure 9.2, which has been normalized by the $p53$ level at T_0. To obtain the irreversible bistable switch, only five parameters corresponding to the entry of damage into the network were altered, and these are shown in Table B.2.

The study of $\mathcal{N}_{int}^3$ was performed by slightly altering some parameters with respect to Table B.1. These parameters are shown in Table B.3.

Most of the parameters altered are those which affect the entry of damage into the core regulation network through the kinases (p_1, p_4,c_1 and c_2), which were chosen arbitrarily. The strength of the negative feedback between $MDM2$ and $p53$ is strengthened by increasing the parameter b_1. With respect to this model, this does not qualitatively change the behavior other than to clearly demarcate the higher and lower branches of $p53$ equilibrium points by increasing the distance between these branches. The external

Parameter	Value	Units
p_1	0.0229	A.U./h
p_4	0.0147	/h
c_1	0.0559	A.U./h
c_2	0.1191	A.U./h
k_2	2	A.U.

Table B.2: Parameters altered to obtain irreversible bistable switch in Network $\mathcal{N}_{int}^2$.

Parameter	Value	Units
b_1	1.25	\h
p_1	0.017	A.U./h
p_4	0.0039	/h
c_1	0.066	A.U./h
c_2	0.183	A.U./h

Table B.3: Parameters altered for core regulation network model in Network $\mathcal{N}_{int}^3$

input to this network u_{22} is set to 0.1.

The bifurcation diagram of this network is shown in Figure B.1.1. It is also assumed that D_1 and D_2 are coupled, which is not unreasonable since they are both triggered by DNA double-strand breaks.

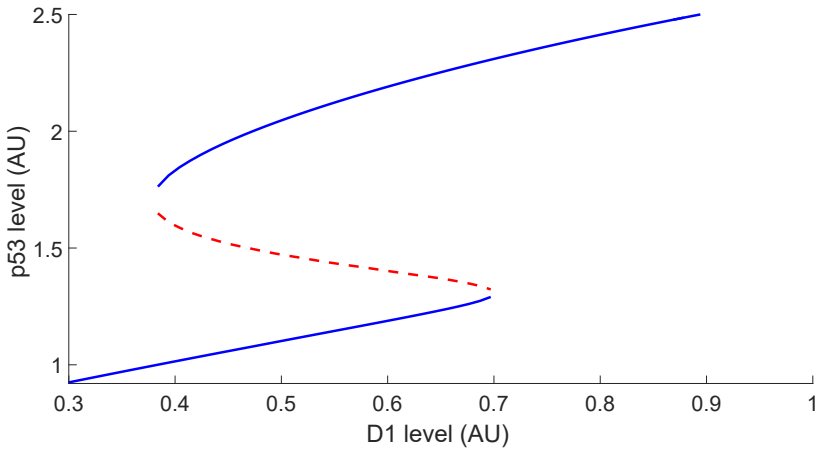


Figure B.1.1: Bifurcation diagram of $\mathcal{N}_{int}^3$.

B.1.3 Monostable networks $\mathcal{N}_{int}^1$ and $\mathcal{N}_i$

The study of $\mathcal{N}_{int}^1$ and $\mathcal{N}_i$ were performed with the same parameters as shown in Table B.1, except for those that affect the entry of damage into the core regulation network

through the kinases (p_1, p_4,c_1 and c_2). The parameters were chosen to demonstrate how these networks could operate as ultrasensitive switches, and are given in Table B.4. The resulting $p53$ responses to changes in $[D_1]$ are seen in Figure 9.1, where these figures have been normalized by the initial level of $p53$.

Parameter	Value	Units
p_1	0.0206	A.U./h
p_4	0.1	/h
c_1	0.0226	A.U./h
c_2	0.0531	A.U./h
k_1	$2.79 * 10^{-7}$	A.U.
k_2	$2.79 * 10^{-7}$	A.U.
n_1	17	-
n_2	17	-

Table B.4: Parameters altered for core regulation network models Network $\mathcal{N}_{int}^1$ and $\mathcal{N}_i$.

B.2 Damage sensing and repair model

The functions for (9.1)–(9.3) are given as follows:

$$g_i(x) = q_i + \alpha_i \frac{x^{m_i}}{\beta_i^{m_i} + x^{m_i}} \tag{B.2.1}$$

with parameters given in Table B.5.

As was noted in Section 9.2, the dynamics of D_1 are assumed to be very slow compared to that of the core regulation network, leading to the observed behavior. It is worth noting that we simplified our analysis by linearizing g_2 in the following way:

$$g_2(x) \approx max(0, -k_0 + k_1 * x),$$

where

$$k_0 = 6 * 10^{-4}$$
$$k_1 = 9.84 * 10^{-4}.$$

Parameter	Value	Units
q_1	$1.14 * 10^{-5}$	$\#DSB$/h
q_2	0	A.U./h
q_3	0.025	$\#DSB$/h
α_1	$5.25 * 10^{-5}$	$\#DSB$/h
α_2	$4.5 * 10^{-8}$	A.U./h
α_3	0.014	$\#DSB$/h
β_1	14.298	$\#DSB$
β_2	1.3167	A.U
β_3	300	$\#DSB$
m_1	1	-
m_2	4	-
m_3	6	-
μ	$2 * 10^{-5}$	/(h A.U.)
γ	$6 * 10^{-5}$	/h

Table B.5: Parameters altered for core regulation network models Network $\mathcal{N}_{int}^1$ and $\mathcal{N}_i$.