# Modelling Biochemical Components as Communicating Processes


**Riccardo Fiorista**

Supervisor: Dr. Ian Stark

School of Informatics
University of Edinburgh

This dissertation is submitted for the degree of
*BSc. Hons. Artificial Intelligence and Computer Science*

April 2020

# Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text.

Riccardo Fiorista
April 2020

# Acknowledgements

I would like to thank everyone who contributed to this thesis and allowed me to learn so much these past months. Very special thanks to Ian Stark and Thomas Wright, for the constant guidance and support and for always answering all of my "quick questions". A great thanks also goes to John Tyson for his availability and support on questions regarding his publications as well as sharing his current research with me.

Finally, I thank my family for their support as well as to Nielja for her positivity and all the suggestions and proof-reading.

# Abstract

We present a collection of reactive biochemical system components of increasing complexity modelled with the *Bond-Calculus*, a compositional process algebra continuous in space and time. This high-level language allows us to express biological systems in terms of agents and their interactions within certain processes. We show that all core components found in biochemical systems can be modeled accurately and composed to more complex systems such as oscillators. The models generated are expressed as systems of non-linear ordinary differential equations, describing system behaviour over time. Furthermore, we characterise system behaviour through model-checking with the *Logic of Behaviour in Uncertain Contexts (*$\mathcal{LBUC}$*)*. We show that the properties we devised for a set of biochemical systems can be quantitatively evaluated over uncertain initial states and mixture perturbations. Finally, this thesis should provide a comprehensive framework for future work on biochemical system modelling in the bond-calculus and serve as a reference for applied use of the $\mathcal{LBUC}$.

# Contents

# List of Figures

# Nomenclature

**Roman Symbols**

| | |
|---|---|
| $I$ | real-valued interval $I \in \mathbb{R}_{\geq 0}$ |

**Greek Symbols**

| | |
|---|---|
| $\Pi$ | bond-calculus or continuous $\pi$-calculus *process* |

**Subscripts**

| | |
|---|---|
| 0 | initial |
| P | phosphorylated |
| SS | steady-state |
| T | total |

**Other Symbols**

| | |
|---|---|
| $[A]$ | concentration of chemical species $A$ |

**Acronyms / Abbreviations**

| | |
|---|---|
| ADP | Adenosine Diphosphate |
| ATP | Adenosine Triphosphate |
| bondwb | The Bond-Calculus Workbench |
| c$\pi$ | Continuous $\pi$-Calculus |
| cAMP | cyclic Adenosine Monophosphate |
| DNA | Deoxyribonucleic acid |
| Flow* | Flowstar Framework |
| GK | Goldbeter-Koshland Function |
| IVP | Initial Value Problem |

LTL Linear Temporal Logic

$\mathcal{LBUC}$ Logic of Behaviour in Uncertain Contexts

MA Mass Action Kinetics

MPF Maturation Promotion Factor

mRNA messenger Ribonucleic Acid

MITL Metric Interval Temporal Logic

MM Michaelis-Menten Kinetics

ODE Ordinary Differential Equation

PDE Partial Differential Equation

SBML The Systems Biology Markup Language

TL Temporal Logic

tQSSA total Quasi Steady State Assumption

# Chapter 1

# Introduction

Formally describing biochemical system behaviour is at the core of system biology and vital in applications such as biomedical, pharmaceutical, or metabolic research. Expressing systems of ordinary differential equations (ODEs) to model the behaviour of complex systems such as entire cells or gene regulation networks poses a challenge which for complex systems cannot be solved merely by intuition or knowledge about the system. Several approaches over the past decades such as quantitative process algebras or rule-based approaches emerged. These attempt to provide a formalization process, allowing researchers to frame biochemical component interaction in a high-level syntax. Throughout this work we will focus on two process algebras, namely the *continuous $\pi$-calculus* by Kwiatkowski and Stark [49] and the *bond-calculus* by Wright and Stark [88]. The latter builds upon the former and attempts to address shortcomings such as limitations on the use of general kinetic laws and the ability to model symmetric multi-way interactions.

The bond-calculus attempts to unify concepts from both agent- and rule-based modeling approaches. It provides flexibility through multi-way interaction of agents, dynamic bonding, and the support for general kinetic laws. Thus, modeling biochemical systems in the bond-calculus allows us to succinctly encode the governing components and their interactions in a high-level language.

In this work we show that the bond-calculus is capable of modeling system components at the core of biochemical networks as well as more complex combinations of those. Our process is guided by the work of Tyson et al. [82] as well as related literature extending the work [19, 60, 80]. In Tyson et al. [82], 7 basic components and their biochemical signal-response interactions, also referred to as their *functional motifs*, and 3 oscillators composed from these components are discussed. Thus, we present our implementations for all 10 systems. We show that our models match the low-level ODE models in [82] and compare the resulting signal-response graphs visually. While this thesis aims to summarize the work done, it is also thought as reference and guiding literature for future applications of the bond-calculus.

As the bond-calculus originates from the continuous $\pi$-Calculus (c$\pi$) by Kwiatkowski and Stark [49], we build on the work of Wang [84] in which a subset of the systems presented here

are modelled in c$\pi$. Hence, we will show analogies and differences between the two process calculi, as well as the improvements of the bond-calculus over c$\pi$.

Finally, we present a novel approach of evaluating some of the systems presented through the use of *Logic of Behaviour in Uncertain Contexts* ($\mathcal{LBUC}$). It builds on the work by Banks and Stark [8] on *Logic of Behaviour in Context* ($\mathcal{LBC}$) and allows us to quantify uncertainty within a system. By capturing the characteristics of a system's behaviour with a *temporal logic* syntax, we show how the system behaves under perturbation and can show under which circumstances the functional motifs retain their behaviour. To this end we characterize the properties of the implemented biochemical systems and define under which circumstances we believe these are satisfied. As $\mathcal{LBUC}$ allows us to quantify uncertainty on both the initial values of the systems and the perturbations occurring during a simulation, we then verify the bounds of our assumptions. Finally, we show how we can use $\mathcal{LBUC}$ to pin-point values of constants such that systems reflect the expected behaviour.

## 1.1 Context and Motivation

System biology focuses on providing a formal way of modeling biochemical systems as a composition of their components. With this, the system's internal interactions can be simulated and their contribution to the system behaviour analysed. It presents an economic and fast way to test assumptions on systems which otherwise would be costly to reproduce in living cells, if at all possible. Recent developments such as by Karr et al. [44], indicate that ever more complex systems, such as entire cells, can be modelled. In such systems, the study of emergent behaviour is of particular interest. Emergent indicates that, although the different behaviours of the components within a system are known separately, the behaviour of the entire system emerges from their interaction and cannot be formulated without it.

Tyson et al. [82] call for closer collaboration between computational/system biologists and experimental biologists. It can be argued that the interaction between these distinct research groups can be bridged by a common language of interaction which makes steps towards closing the gap between natural language and mathematics.

The bond-calculus attempts to narrow the gap between the two. Since it provides a high-level language which seeks to allow computational and experimental biologists alike to express the components and their interactions of biochemical systems. It implements compositional semantics based on vector fields and linear operators. Through these, it generates sets of non-linear ODEs describing the time evolution of the modelled systems [88]. Hence, the first research question (Section 1.2) we try to answer is whether the bond-calculus is capable of modeling the systems presented in [82].

However, many limitations are still faced when modeling biochemical systems. Often, the exact components interacting in *in-vivo* systems cannot be determined. Furthermore, there are cases where stochastic perturbations cause random system behaviour [72]. Therefore, over the past century a multitude of simplifications and assumptions were expressed to reduce the

complexity of modelled systems and to make their computation tractable. This comes to a specific price in terms of accuracy [45, 16, 46]. Hence, our second research question (Section 1.2) focuses on the validity of the simplifications and assumptions made by Tyson et al. [82] and how our models are affected when less assumptions are made.

Ultimately, to motivate our novel approach to evaluate our modelled systems quantitatively, we argue that verifying system behaviour is notoriously difficult, especially if emergent behaviour plays a substantial role [91]. Once a model is devised, it is not trivial to identify under which circumstances its behaviour is guaranteed to reflect what is expected. Some approach this problem like Periwal [62], "We figuratively threw a bunch of models all at once at the experimental data and picked the model that came closest for further evaluation.". Even in our work presented, we first resort to qualitative analysis through the visual comparison of the obtained functional motifs. Hence, formalizing this step is of crucial importance. $\mathcal{LBUC}$ allows us to encode the system behaviour, such as "the response rate will increase linearly with the increase of signal concentration" and verify on which time scale and under which external perturbations this will hold eventually or globally (through the whole time of simulation). It furthermore allows us to explore under which initial conditions the expected behaviour is obtained. This motivates our last research question (Section 1.2) on whether it is possible to use $\mathcal{LBUC}$ to verify the properties of the emergent behaviour excerpted by the systems modelled.

## 1.2 Objectives and Aims

Throughout this thesis, our aim is to cover three main research questions,

1. Is the bond-calculus capable of modeling non-linear behaviour of complex systems such as oscillators with custom kinetic laws providing an abstraction of the underlying processes?

2. How do simplifications and abstractions affect our models and is it readily possible to employ more precise and verbose approximations to these in the bond-calculus?

3. Provided with a bond-calculus model, is it possible to qualitatively, formally verify and guarantee its behaviour through the use of $\mathcal{LBUC}$?

In order to address these, we identified the following objectives,

1. Conduct a literature review on the biological background, system biology, and biochemical process modelling and build a thorough understanding of what is being modelled.

2. Implement the components presented by Tyson et al. [82] in the bond-calculus and provide a cookbook of biochemical components for the calculus.

3. Retrieve additional information needed for the modeling of the systems from the Biomodels Database [51].

4. Devise appropriate qualitative evaluation criteria to assess the performance of the systems.

5. Compare the results obtained through our implementation in the bond-calculus with the ones presented by Tyson et al. [82] using the evaluation criteria.
6. Compare the results obtained through our implementation in the bond-calculus with the ones presented by Wang [84] using the evaluation criteria.
7. Devise appropriate quantitative evaluation criteria through the use of $\mathcal{LBUC}$.
8. Capture behaviour properties of some systems modelled with the $\mathcal{LBUC}$ syntax and evaluate these.
9. Provide a cookbook of how evaluation with $\mathcal{LBUC}$ can be achieved for biochemical models in the bond-calculus.

## 1.3 Contributions

Throughout this thesis we present the following contributions,

1. Elaboration on the biochemical implications on our models of the assumptions made by Tyson et al..
2. The quantitative characterization in natural language of all bond-calculus models presented.
3. The full implementation of all models presented by Tyson et al. [82] in the bond-calculus.
4. The formal characterization in $\mathcal{LBUC}$ of the first 4 bond-calculus models namely, linear response, hyperbolic response, sigmoidal response, and perfect adaptation.
5. A qualitative evaluation of our generated bond-calculus models
6. A qualitative evaluation of the four formally characterized bond-calculus models using $\mathcal{LBUC}$.

## 1.4 Thesis Structure

- **Chapter 1: Introduction** Introducing our work, its context and motivations as well as its aims and objectives.
- **Chapter 2: Background and Literature Review** Introducing the biological and biochemical background for the work, presenting c$\pi$, the bond-calculus and $\mathcal{LBUC}$, as well as existing tools.
- **Chapter 3: Presentation of Models** Providing a short overview of the work by Tyson et al. [82] and the biochemical assumptions and simplifications assumed by the authors. Presenting the models we chose to implement.

- **Chapter 4: Implementation** Presenting our implementation and the related code. Giving context on the complications found during implementation and illustrate the adaptions undertaken to mitigate these. Ultimately, characterizing system behaviour in $\mathcal{LBUC}$ for a subset of the systems modelled.
- **Chapter 5: Evaluation** Devising our approach to qualitative and quantitative evaluation and providing the results generated for each model discussed in Chapter 4. Relating our implementation to the results obtained by Tyson et al. and Wang in c$\pi$. Presenting our results on the quantitative evaluation in $\mathcal{LBUC}$. Ultimately, assessing the suitability of this latter form of evaluation.
- **Chapter 6: Conclusion** Evaluating the overall project. Assessing whether the aims and objectives have been reached and devising directives for future work.

# Chapter 2

# Background and Literature Review

This chapter presents the necessary background in the three fields discussed throughout our work, namely *system biology and biochemistry*, *process calculi*, and *formal verification*. While system biology is the field our work resides in, biochemistry is the basis for the biochemical systems discussed in this thesis and originally presented by Tyson et al. [82]. Process calculi such as the c$\pi$ and the bond-calculus provide a framework to model such systems in a high-level syntax. By extracting from the models the (non-linear) ordinary differential equations (ODEs) describing the temporal behaviour of the systems, we can reason about the system's behaviour over time. We can use paradigms from the field of formal verification to quantitatively evaluate whether certain properties are satisfied by the generated systems. Finally, we present the available tools which allow us to implement the models and conduct formal verification on these.

## 2.1 Systems Biology

By definition, systems biology is the exploration of a biological system by analyzing the interactions among its parts [85]. It emerged over the past century from the desire to add a quantitative dimension to experimental biology [76]. Every living organism is a biological system and while the field is widely spread, we will specifically consider *molecular systems biology*, focusing on biochemical systems. Through the aid of mathematical and computational methods, we seek to study and formalize the dynamic behaviour of complex molecular processes and regulatory systems comprising genes, proteins, and metabolites.

While research on providing mathematical formalization of observed biochemical processes started early with the discoveries by Guldberg and Waage [34] and Michaelis and Menten [56], the first successful mathematical model of an observed biochemical system was presented by Hodgkin and Huxley describing the simulation of neuron reactivity [38]. Progress over the decades encompassed the modelling of processes in yeast cells [13] and more recently lead to successes such as models of complete cellular behaviour of a human pathogen [44]. As a result, systems biology increasingly impacts fields such as pharmacology, agriculture, and metabolic

research, to name a few. However, fundamental difficulties in our understanding of biochemical interactions and limitations in our representation of such are still faced [45].

Kell and Knowles [45] argue that models of *in-vivo* systems require stark simplifications and abstractions to allow (partial) replication of the observed behaviour. Due to the stochastic nature of biological processes, there are quantitative challenges such as, multiple levels of signal processing, wide ranges of sensitivities to mixture perturbations, as well as multi-functionality of components depending on their current state and context. The attempt to incorporate all of these, if at all possible, would lead to computationally intractable problems.

In order to correctly incorporate adaptions to these challenges, we require in-depth knowledge of a system's components. This knowledge, poses challenges itself which are quantified as *knowledge challenges*. Kell and Knowles [45] present the following:

- finding the appropriate level of abstraction
- finding common basis to relate knowledge gained using different experimental techniques on the same system
- the ability to incorporate knowledge incrementally as new data is analyzed

It becomes clear that systems biology is reliant on experimental biology and aims to aid it, not to substitute it. Models cannot be formed merely by intuition and cannot be treated as oracles as it is put by Laursen [50].

### 2.1.1 Reductionism and Holism

Within the field of biology, there is a fundamental divide between the notion of *reductionism* and *holism* [32].

- **Reductionism** postulates that every system can be decomposed into its constituents. Hence, system behaviour can be deduced summatively from the behaviour of the parts.
- **Holism** on the contrary assumes that system behaviour can only be fully understood when seen as a whole. The whole is bigger than the sum of its parts.

Systems biology emerged as an attempt to provide a holistic approach to the modelling and the analysis of biochemical systems. However, some approaches ultimately use (reductionistic) mathematical models as building blocks. This allows us to model emergent behaviour which could not be replicated by analyze the parts of a system separately.

While more of a philosophical question, we introduce the differentiation here in order to acknowledge both and put them in context with our research. Throughout this thesis we show models of increasing complexity, building upon one another. We also re-use certain components and interaction patterns, re-arranging them to create new behaviour. However, ultimately we reason about the response behaviour of the ensemble, the system as a whole, exploring behaviour emergent from the interaction between the parts.

## 2.2 Biochemical Systems

We now elaborate the biochemical context necessary to reason about the systems presented throughout our work by Tyson et al. [82]. Hence, we hierarchically present first the basic molecular components as well as their possible interactions in networks composed of these. Here, general ideas and concepts are taken from textbooks such as [11] of [58]. Finally, we discuss the complex behaviours that can be observed from a holistic view-point of the component interaction as well as how the time component is vital for biochemical systems to stabilize.

### 2.2.1 Components

Cells are separated from their surrounding by a semi-permeable membrane. Within, cytoplasma represents an inert liquid which houses the internal cellular components, the so-called organelles. The cell's nucleus is a membrane-bound organelle containing the DNA (deoxyribonucleic acid), the biochemical encoding of the cell's genome. It consists of nucleotides, chemical building blocks that characterize the genetic code. Four different bases can be part of a nucleotide and their sequence encodes the genetic information. Individual genes are specific sections of the genome and contain information for the production of biologically active RNA (ribonucleic acid) [11].

In the process of protein-biosynthesis, this genetic information is transformed into proteins in a two-step process.

1. **Transcription:** The base sequence of the DNA is transcribed into a complementary transport form, the messenger RNA (mRNA) within the nucleus. The mRNA then leaves the nucleus to be processed further.

2. **Translation:** The mRNA is translated into a polypeptide, a chain of amino acids, by the ribosomes which can be thought of as molecular assembly unit. The constructed polypeptide folds up into a specific 3D-structure and is termed a protein.

Enzymes are a special biochemically active class of proteins and play an important role in cellular biochemical networks by catalyzing reactions of the substrate they bind to. Enzymatic catalysis reduces the activation energy of a reaction of one or multiple substrates. Substrates are molecules on which the enzyme acts. Often reactions are inhibited by an activation energy threshold and hence rely on enzyme catalysis. Only certain enzymes can interact with certain substrates and how well they can react is referred to as their affinity [58].

Two important classes of enzymes, kinases and phosphatases, are key to the biochemical networks discussed in this thesis. They drive a phosphorylation/dephosphorylation process which can also result in a metaphoric *activation* and *deactivation* of proteins respectively. These state changes allow then interactions with other molecules within the system.

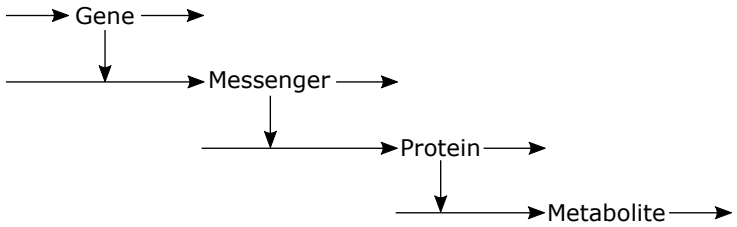



Figure 2.1 Signalling cascade from DNA (deoxyribonucleic acid) to mRNA (messenger ribonucleic acid) to enzymes to metabolites. Each stage can consist of multiple output signals. Modified from Savageau in [67]

- **Kinases** belong to the group of transferases and catalyze the transfer of phosphoryl groups from ATP (adenosine triphosphate) to a substrate. This transfer can lead to an activation of the substrate and is crucial to signal transduction pathways as described in Section 2.2.2.
- **Phosphatases** are part of the hydrolase enzyme class that can split chemical bonds via a reaction with water. Their task is to remove a phosphoryl group from a substrate and transfer it to the water molecule. Hence, phosphatases catalyze the reverse reaction of the process catalyzed by kinases.

### 2.2.2 Biochemical Networks

Interlacing the components discussed, we obtain signalling cascades resulting in Protein Interaction Networks (PINs). Upon external or internal stimuli, these cascades are triggered. External triggers are perceived through membrane receptors interacting with hormones and other molecules. Once triggered, they initialize a *transduction pathway*, a cascade of protein state changes within the cytoplasma. At the final stage, the process of protein-biosynthesis is initiated and a specific enzyme is produced as response to the external stimulus. This enzyme can then initiate another cellular pathway, leading to the production of metabolites, which are intermediate products of metabolic pathways. We will focus here on the activations and enzymatic responses as consequence of protein-biosynthesis and hence treat mRNA as the initial signal [11, 58]. This is shown in Figure 2.1. Such PINs enable behaviour such as movement, reaction to external stimuli, and the processing of internal reactions necessary for survival, to name a few [11].

### 2.2.3 Reaction and Enzyme Kinetics

Biochemical reactions and interactions occur at specific rates which are described, or governed by, *reaction kinetics*. Here, the law of *mass action* (MA) is at the basis of most molecular reactions. However, repetitive reaction patterns can be found and approximations such as the *Michaelis-Menten* kinetics (MM), the *Goldbeter-Koshland* function (GK), or other quasi-steady state assumptions are used. This reduces the number of explicit reactions that have to be modelled, hence the complexity of the model. Throughout this report we will be referencing to a species's concentration by square brackets, e.g. $[A]$. The unit for concentrations is Mol per liter or M.

### Mass Action

The law of Mass Action states that the rate of a chemical reaction is directly proportional to the concentration of the reactants [34]. It is universal and theoretically every chemical reaction can be expressed in terms of MA governed interactions [81]. To illustrate this dynamic, we consider an exemplary chemical Reaction 2.1 with reactants $A$ and $A$ binding in an irreversible (arrow to the right) reaction to $A$. Equations 2.2 and 2.3 describe the rate of change of each component. Observe that $[A]$ and $[B]$ reduce at the same rate as $[C]$ increases. The *rate constant* $k_f$ describes the speed of the reaction and is found *experimentally*.

$$A + B \rightarrow C \tag{2.1}$$

$$\frac{d[A]}{dt} = \frac{dB}{dt} = -k_f AB \tag{2.2}$$

$$\frac{dC}{dt} = k_f AB \tag{2.3}$$

### Michaelis-Menten Enzyme Kinetics

Enzymatic reactions follow a pattern as in Reaction 2.4. We categorize the forward, reverse, and catalyzation reactions by indices $f, r, cat$ respectively. The enzyme $E$ binds with substrate $S$ at rate $k_f$ to form an intermediate product $C$, also referred to as enzyme-substrate complex. This reaction is reversible and at rate $k_r$, the complex dissociates again into its initial reactants. However, with rate $k_{cat}$, the complex engages in an irreversible reaction to form a mixture containing the unmodified enzyme and a product $P$. The enzyme itself, being the catalyzer, is not affected by the reaction in terms of total concentration, $[E_T] = [E] + [E_C]$ (with $E_C$ being the amount of enzyme bound in complex $C$).

$$E + S \underset{k_r}{\overset{k_f}{\rightleftharpoons}} C \overset{k_{cat}}{\rightarrow} E + P \tag{2.4}$$

Michaelis and Menten [56] and later Briggs and Haldane [10] approximated the reaction above and expressed the *reaction velocity* $v$, hence the synthesis velocity of $P$, as a function over the substrate concentration as illustrated in Figure 2.2. The concentration of the substrate-enzyme-complex $C$ approaches a steady-state quickly resulting in $\frac{dC}{dt} \approx 0$. This leads to Equation 2.5 which we will refer to when assuming MM kinetics. Here, $K_m = \frac{k_r + k_{cat}}{k_f}$ is denoted as the *Michaelis-Menten (MM) constant*. It describes at what value of $[S]$ the reaction runs at half of its *maximal velocity* and is hence a measure of the affinity of the enzyme for the substrate.

At $[S] \ll K_m$, not all enzymes are saturated with substrate and increasing $[S]$ increases the overall reaction speed. Here, $[E_T] \gg [S]$ does not hold and we could also use MA kinetics to describe the system. However, at $[S] \gg K_m$, all enzymes are saturated with substrate. That is, $[E_T] \ll [S]$, hence an increase in $[S]$ does not affect the reaction speed which is at maximum velocity $V_{max}$. This depends on the specific properties of the enzyme in question. Thus, at

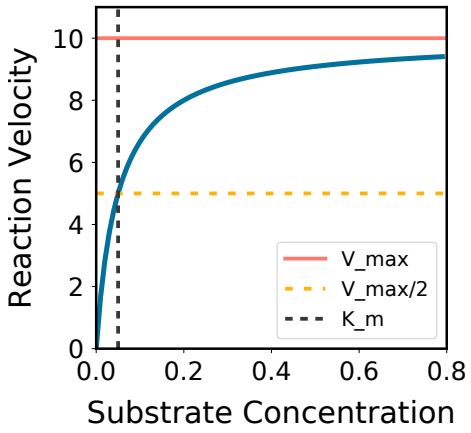



Figure 2.2 Reaction velocity of a MM governed enzymatic reaction. $K_m = 0.05$, $V_{max} = 10$

high substrate concentrations, using MM kinetics is more accurate. As discussed by Ciliberto et al. [16], the assumption that only the substrate concentration is limiting for the reaction speed, might not be representative for PINs with multiple substrates and enzymes.

$$\frac{dP}{dt} = \frac{k_{cat} E_T S}{K_m + S} \tag{2.5}$$

**Goldbeter-Koshland**

If two MM governed reactions are coupled in a forward/backward setting (e.g. $X \rightleftharpoons X_P$; species $X$ and its phosphorylated form $X_P$), the GK equation is used as a steady-state solution for the concentration of the product species.

Often, enzymatic reactions are found in coupled systems such as phsophorylation and dephosphorylation. Here a substrate-product pair (e.g. $X$ and $X_P$) interconvert through the acting of a kinase and phosphatase respectively. At constant levels of kinase and phosphatase, this system eventually reaches a steady-state of concentrations $[X]$,$[X_P]$. Hence, Goldbeter and Koshland devised Equation 2.6 in [33] by setting $\frac{dX}{dt} = \frac{dX_P}{dt} = 0$. While the derivation can be found in their work, we note that variables $u$ and $v$ expresses the influence on the coupled system by the kinase and phosphatase respectively. Moreover, $J_1 = \frac{K_{m1}}{X_T}$ and $J_2 = \frac{K_{m2}}{X_T}$ encode the MM constants for each reaction.

$$G(u,v,J_1,J_2) = \frac{2uK}{v-u+vJ+uK+\sqrt{(v-u+vJ+uK)^2-4(v-u)uK}} \tag{2.6}$$

Goldbeter and Koshland refer to this as ultra-sensitive switch due to the resulting sigmoidal functional shape of both concentrations $[X]$ and $[X_P]$. This ultra-sensitivity is however only guaranteed for $J_1, J_2 \ll 1$ [33, 16].

### 2.2.4 Biochemical Feedback and Oscillations

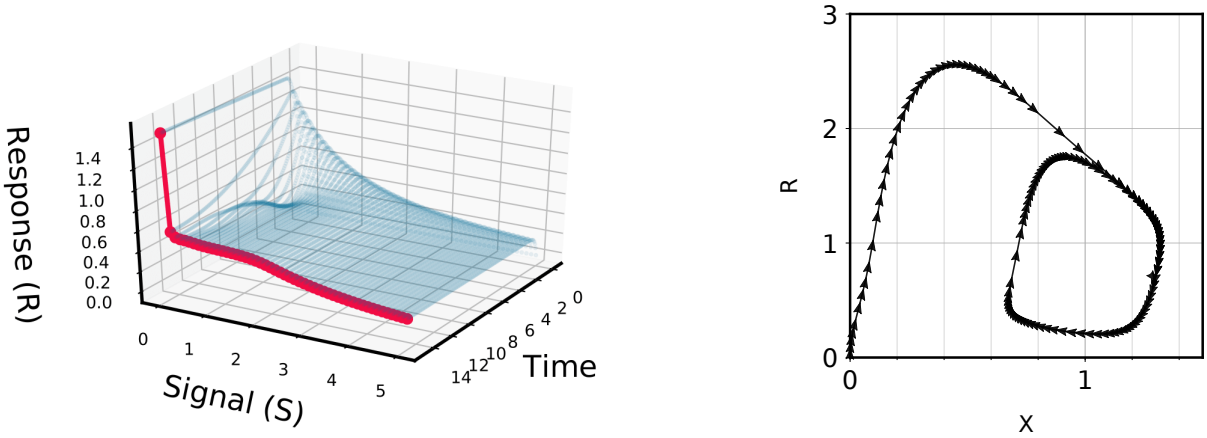
When in biochemical networks, the output of one reaction affects the progress of an upstream reaction, a biochemical *feedback* loops arise [19]. They can cause bi-stabilities in the concentration of species within a mixture and are fundamental to oscillating behaviour. There are three types of feedback, namely *positive*, *negative*, or *mutual inhibition* (also referred to as antagonism) and we shall discuss these in Chapter 3 [19]. While the behaviours of the single components constituting these loops are known, the behaviours of the whole is of an *emergent* nature. Hence it is no triviality to identify these without further mathematical simulation [19]. In fact, systems interwinding these feedback loops show exceedingly complex dynamic properties [19, 45, 43]. As Novák and Tyson [60] reason, this interwinding of feedback loops results in oscillatory behaviour within certain bounds of the bi-stabilities of the underlying feedback motifs. Three basic oscillators were identified, namely the *negative-feedback*, *activator-inhibitor*, and *substrate-depletion* oscillator. Such oscillators regulate internal cycles of organisms such as day-night production cycles in the case of the circadian clock [70], or entire cell life cycles as in the case of the MPF (maturation promoting factor) cycle of eukaryotes [9].

### 2.2.5 Time Dependent Stabilization

Generally, a model's behaviour over time is expressed through a set of ODEs. However, in certain cases, such as for the components presented in [82], we are interested in the dynamic steady-state concentrations of the system reached as it stabilizes. Here, stabilization refers to a dynamic equilibrium of synthesis (introduction) of a species and its degradation (decay or degradation). While components modelled with MM kinetics require a certain time to stabilize, approximations such as the GK function provide immediate steady-state results. However, systems such as oscillators do not reach such a steady-state. Instead the concentrations periodically oscillate around an imaginary steady-state point [19]. In Figure 2.3a we illustrate how the concentrations of a system driven by MM eventually reaches a dynamic steady-state. Figure 2.3b shows a phase-plane portrait of a substrate-depletion oscillator. Here, the phase-plane is a visual illustration of the dynamic behaviour of state variables $X$ and $R$. The portrait of this phase-plane refers to the illustration of trajectories (marked by the arrows) of the state variables in question [42].

## 2.3 Modeling in the Biochemical Context

As we have now covered the biochemical background, we here first present the general approaches to modeling. Subsequently, we focus on the c$\pi$ as well as the bond-calculus as our work is based on the latter and builds on the former. We introduce both by implementing the same MM kinetics reaction from Section 2.2.3 and reason about their differences.




(a) (b)

Figure 2.3 Visualization of time dependent stabilization. **(a)** Stabilization over time of homeostatic system discussed in Section 3.3.7. **(b)** Phase-plane diagram of the activator-inhibitor oscillator presented in Section 3.4.2. Concentrations $[R]$ and $[X]$ have initial value 0 and after sufficient time circulate around an imaginary steady-state. Time axis is perpendicular to the plane.

### 2.3.1 Approaches

Many approaches to biochemical modeling exist in both quantitative and qualitative forms. For the former, modeling through systems of partial differential equations (PDEs) to capture spatial locality of molecules or using stochastic simulations to model concentration fluctuations exist [45]. For the latter, it has been experimented to model system behaviour through Boolean networks and discrete variables [73]. While depending on the context, the currently most researched modeling approach uses (non-linear) ODEs, allowing to describe continuous state changes over continuous time [75, 67, 79, 66].

However, there are severe shortcomings of directly devising ODE systems from observed behaviour. For one, as mentioned in Section 2.1 certain system behaviour is emergent and it is not possible to express by mere intuition the underlying mathematical relations between species. For another, ODEs are non-compositional. Hence, although a system might incorporate components from another system, it is not possible to simply add new components to it and obtain a new sensible set of ODEs [49, 8]. Therefore, a multitude of high level languages arose building on different concepts to provide a formalization process for biochemical systems modeling.

The $\pi$-calculus was first presented by Milner et al. [57] as a calculus for mobile processes in the setting of concurrent computing. However, Regev et al. [65] identified that the $\pi$-calculus can also be used for the modeling of biochemical systems. Many agent-based approaches emerged over the past decades and languages such as BioPEPA, BlendX, the continuous $\pi$-calculus (c$\pi$), and most recently the bond-calculus, to name a few, all build on the concepts of processes, channels, and message passing [63, 17, 49, 88, 65].

The idea here is that a system is a combination of processes or *agents* which can communicate over channels. These can allow one-way or, as in the case of the bond-calculus for example,

multi-way communications between the agents. Actions of a process depend on the concept of message passing or mobility. We note however that, for the specific application in the biochemical context, we require agents to exist in continuous amounts (concentrations).

Rule based modeling represents the counter-part to the above. It takes its roots in the fundamentals of chemistry where interactions are represented by graphical connections. Frameworks such as the Kappa language [39] and BioNetGen [36] provide accessible solutions for this approach. Recentls, graph transformation models gained particular importance as they provide a convenient framework for rule based modeling [6]. We note here that the systems biology markup language (*SBML*) is an important rule-based language to express biochemical models. However, it is software agnostic and does not provide any modeling capabilities [40].

Despite the differentiation between agent- and rule-based modeling, there is an ever looser boundary between the two approaches. An example is the notion of agents in the Kappa language, or the rule-based inspired notion of affinity networks in the bond-calculus.

### 2.3.2 The Continuous $\pi$-Calculus

As the bond-calculus finds its roots in the work by Kwiatkowski and Stark [49], we first introduce its syntax through a working example of the MM kinetics as presented by Banks and Stark [8]. c$\pi$ itself builds upon the work by Regev et al. [65] and the *process-as-molecule* approach, where individual molecules are represented as processes and their interactions as binary message-passing [64, 89]. We refer to the thesis of Kwiatkowski [48] as canonical reference for the c$\pi$. A succinct summary of the syntax used here can be found in the work by Banks and Stark [8]. In Appendix D.1 we provide the code for the implementation of the following example in the CPi-IDE [55].

$$E + S \underset{k_r}{\overset{k_f}{\rightleftharpoons}} C \overset{k_{cat}}{\rightarrow} E + P \tag{2.7}$$

Recall Reaction 2.7 as the chemical reaction which the MM kinetic law functionally captures (as in Section 2.2.3). The system is composed of a substrate $S$, enzyme $E$, and product $P$. We first define the system's chemical species as:

$$S \triangleq s(u,r).(u.S + r.P) \tag{2.8}$$

$$E \triangleq (\nu \{t <_r^u\})\, e\langle u,r\rangle.t.E \tag{2.9}$$

$$P \triangleq \tau_{r_{deg}}.0 \tag{2.10}$$

$s,u,r,v$ here are channels and in c$\pi$ channels represent reaction sites of a species. Anything followed by a . (dot) is an interaction in c$\pi$ which can either be a *communication prefix*, as in species definition $S$ and $E$, or a *silent prefix* (or *autonomous reaction*) $\tau_k$ (a MA reaction at rate $k$), such as in species definition $P$.

In the case of substrate $S$, $s(u,r)$ indicates that after it reacts on channel $s$, it can either interact on $u$ or on $r$. The choice is denoted by the $+$ (plus) symbol in $(u.S + r.P)$. If $S$ reacts on $u$, the

species turns back into itself ($S$). Similarly, if it reacts on $r$ into the product $P$. Enzyme $E$, after reacting on $e$ can only react on $t$, leading to turn into itself. The $\nu$, also referred to as *binder*, represents a *name constriction* and indicates that names $u, r, t$ are local to the enzyme. We can also refer to this restriction as the local *affinity graph*, expressing that a reaction between $t$ and $u$ at rate $r_u$ or between $t$ and $r$ at rate $r_r$ is allowed. Note the **or** here as this indicates a limitation by the c$\pi$, allowing only binary message passing. When the enzymes reacts on $e$ it sends out two channels $u$ and $r$ [8].

The species definition of the product $P$ is the simplest and provides an example for the use of an autonomous reaction $\tau_{r_{deg}}$. It indicates that $P$ turns into the *null species*, an always present species incapable of any further action, at rate $r_{deg}$. This is also referred to as *degradation* or decay of the species [8].

We can now define our *process* $\Pi$ with the initial mixture of the species as a *parallel composition* of its species (indicated as $P \parallel Q$ where $P, Q$ are species). Here, the concept of parallel composition indicates that a certain Molar concentration (e.g. $c_S, c_E, c_P$) of species (e.g. $S, E, P$) are within a mixture and that they can communicate on their declared channels. Note, that a process can also be defined as a single species or a composition of processes as well. Furthermore, we define the global *affinity graph* (or *affinity network*) $N$.

$$\Pi \triangleq c_S \cdot S \parallel \; c_E \cdot E \; \parallel \; c_P \cdot P \tag{2.11}$$

$$N = \{s \overset{r_b}{-} e\} \tag{2.12}$$

The global affinity graph is a weighted undirected graph establishing communication between channels. In our case, it states that channels $s$ and $e$ can react. As c$\pi$ supports dynamic bonding, the substrate-enzyme complex $C$ arises from the definitions in both, the local and global affinity graphs, hence doesn't have to be defined explicitly:

$$C \equiv (\nu \{t_r^u\})\,(t.E \mid (u.S + r.P)) \tag{2.13}$$

Note that it carries the local affinity graph defined in $E$ which binds all available channels to it. Hence, no external communication/reaction is possible. Ultimately, the local binding leads to a composition of the possible actions of $E$ and $S$ which in this case are internal (allosteric) reactions. Hence, $C$ turns into $E$ as it reacts on $t$ and it has the choice to, at rate $r_u$, react on $u$ and turn into $S$ or, at rate $r_r$, turn into the product species $P$. Thus, it encodes the forward, backward, and catalyzation reaction in Section 2.2.3.

From this definitions, the c$\pi$ is now able to derive the ODEs governing the behaviour of the whole process through its compositional vector-space semantics, continuous in time and space. It is beyond the scope of our work to define and illustrate the exact process involved. However, within the c$\pi$ we have a set $\mathcal{S}^{\#}$ called *prime species* of elementary species which cannot be broken down into a composition of two non-nil species. To guarantee this we compare species through *structural congruence*. The *process space* $\mathbb{P}$ is a vector space in $\mathbb{R}^{\mathcal{S}^{\#}}$. The gradient

vector $\frac{dP}{dt}$ describes a trajectory through this space and the dynamic behaviour of a system can be expressed through a trajectory from the initial state [8].

### 2.3.3 The Bond-Calculus

The bond-calculus addresses shortcomings of the c$\pi$ such as allowing for general kinetic laws (e.g. the MM kinetic law) as well as flexible, symmetric multi-way communications between species to more closely model chemical bonding. It introduces a novel way to group affinity patterns in one affinity network (instead of local and global affinity graphs). The canonical reference for the bond-calculus and its syntax is the work by Wright and Stark [88] and the corresponding technical annex. We refer to [89] as a further resource and elaborate a practical example of how the bond-calculus can be used to model a gene regulation network. We include the code for the following examples in Appendix D.2.

We illustrate the syntax of the bond-calculus through the implementation of a Michaelis-Menten system as above. The species definitions are:

$$S \triangleq s(\ell).\,(s^*@\ell.S + p^*@\ell.P) \quad E \triangleq e(\ell).e^*@\ell.E \quad P \triangleq p.\mathbf{0} \tag{2.14}$$

In the bond-calculus, possible interactions of a species are denoted by *sites* (such as $s, s^*, p^*, e^*, p$ in this case) which can be bound to specific *locations* ($\ell$) within the species. Both are defined in the *communication prefixes* of the form $s.S, s@\ell.S, s(\ell).S_\ell$ or more generally $s@\ell(m_1,\ldots,m_n).S_{m1,\ldots,m_n}$ which reads, species $S$ has a reaction site $s$ at location $\ell$ and upon reaction on site $s$ receives locations $m_1,\ldots,m_n$ to become species $S_{m1,\ldots,m_n}$ [88]. Two or more sites can interact with each other if they have compatible locations. Should this be the case of two or more sites within one species, they can interact allosterically (internally).

Species $S$ has a site $s$ which upon reaction receives location $\ell$. It then becomes specie $S^* \triangleq s^*@\ell.S + p^*@\ell.P$ where we observe that $S^*$ can either turn into $S$ or $P$ separated by the choice $+$ (plus). Both are possible as both sites have the same location $\ell$. Similarly, $E$ reacts on site $e$ and becomes $E^* \triangleq e^*@\ell.E$.

Ultimately, the species $P$ has a site $p$ at the *ambient location* $\top$, a location top level to the mixture which does not have to be explicitly stated. It degrades into the *null* species, following the same pattern as the c$\pi$.

The affinity network $\mathcal{A}_{\mathrm{MA}}$ consists of affinity patterns indicating the available interactions between sites (similar to biochemical site affinity) and the corresponding rates. Note, that all rate constants are encoded in the affinity network and not on the species definition level. The model can then be expressed as the process $\Pi_{\mathrm{MA}}$, providing the initial mixtures of the species.

$$\mathcal{A}_{\mathrm{MA}} \triangleq \left\{ s \parallel e@\mathrm{MA}_{k_f},\ s^* \mid e^*@\mathrm{MA}_{k_r},\ p^* \mid e^*@\mathrm{MA}_{k_{cat}},\ p@\mathrm{MA}_{k_{deg}} \right\} \tag{2.15}$$

$$\Pi_{\mathrm{MA}} \triangleq [S]S \parallel [E]E \parallel [P]P \tag{2.16}$$

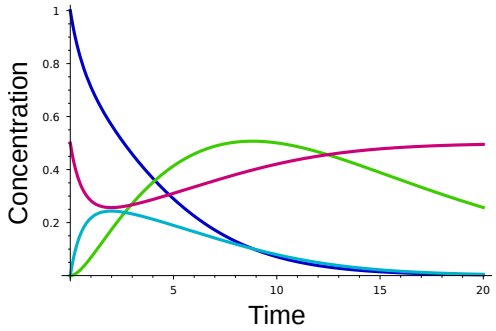



Figure 2.4 Test Enzyme. Temporal progression of concentrations of species $S$ in *blue*, $P$ in *green*, $E$ in *red*, and $C$ in *turquoise*.

Here, MA is indicative for mass action kinetics (Section 3.3.5). In the affinity network, $s \parallel e@\mathrm{MA}_{k_f}$ specifies that sites $s, e$ can interact at rate $k_f$ and that the reaction is governed by MA. While in $s^* \mid e^*@\mathrm{MA}_{k_r}$, we still refer to $\mid$ as parallel composition, it indicates that $s^* \mid e^*$ interact allosterically, hence within the molecule itself.

Similarly to the c$\pi$ example above, the dynamic bonding within the bond-calculus gives rise to the substrate-enzyme complex $C$ without explicit definition. The restriction $\nu\ell$ expresses the formation of collocated molecules and indicates that sites $s^*$ and $e^*$ or $p^*$ and $e^*$ react allosterically due to their same location [88]. Hence, $C$ reacts to $E$ as well as either $S$ or $P$. The resulting ODEs generate the plot illustrated in Figure 2.4.

$$C \triangleq (\nu\ell)\,(S^*_\ell \mid E^*_\ell) = (\nu\ell)\,((s^*@\ell.S + p^*@\ell.P) \mid e^*(@\ell.E) \tag{2.17}$$

However, we can simplify the species definitions and not rely on dynamic bonding by encoding the formation of the complex $C$ in the MM kinetic law. For this we can use species definitions:

$$S \triangleq s.P \quad E \triangleq e.E \quad P \triangleq p.\mathbf{0} \tag{2.18}$$

$$\mathrm{MM}_{V_{max},k}([S],[E]) \triangleq \frac{V_{max}[S][E]}{k+[E]} \tag{2.19}$$

By defining the rate law $\mathrm{MM}_{V_{max},k}$ as a function of the concentrations $[E]$ and $[S]$ (Equation 2.19), we can define the affinity network $\mathcal{A}_{\mathrm{MM}}$ and re-use the previous process definition as the species involved was not altered.

$$\mathcal{A}_{\mathrm{MM}} \triangleq \left\{ s \mid e@\mathrm{MM}_{V_{max},k}, p@\mathrm{MA}_{k_{deg}} \right\} \tag{2.20}$$

The bond-calculus is then capable of extracting the ODEs governing the specified system by the compositional semantics provided by Wright and Stark [88]. While an in-detail elaboration of the semantics would be outside the scope of this thesis, we want to note the most fundamental concepts here. Similar to the c$\pi$, the bond-calculus extracts a system's ODEs by observing the trace and trajectories of a mixture *evolution vector* $\frac{d\Pi}{dt}$ through the mixture space $\mathbb{M}$. The basis of $\mathbb{M}$ are the *prime species* vectors. These again are the elementary species and any other species

or mixture of species can be decomposed into a (linear) composition of these. However, the bond-calculus language has elements such as patterns, clusters (bags of sites), and transitions which have to be raised into the mixture space as well in order to account for these in the system dynamics. We differentiate between pattern space $\mathbb{P}$, cluster space $\mathbb{G}$, and transition space $\mathbb{T}$, which are all infinite dimensional real vector spaces. We therefore require to encode the definitions of a model respectively in transition matrix $\in \mathbb{T}$, interaction tensor (bi-linear map $\in \mathcal{M}(\mathbb{P}, \mathbb{T})$), site concentration vector $\in \mathbb{G}$, and rate vector $\in \mathbb{P}$. Ultimately, the evolution vector is composed by all these constructs and the governing set of ODEs can be extracted from the vector field generated. How this is achieved in detail can be found in the canonical reference and its supplementary material [88].

While not relevant for the work in this thesis, we note that it is possible to conduct numerical or stochastic simulation additionally to the ODE extraction using the bond-calculus. As indicated by Wright and Stark [88], frameworks SciPy [41] and StochPy [52] can respectively provide the tools necessary to implement this functionality.

### 2.3.4 Comparison Between the Bond-Calculus and the C$\pi$

Our work strongly focuses on the capabilities of biochemical modeling brought by the bond-calculus. However, as we build on top of the work by Wang [84] in the c$\pi$, we here present a comparison between the two process calculi.

**Syntax:** It is more intuitive to encode interactions, rate laws, species, and processes separately as done in the bond-calculus over the representation used in c$\pi$. Here, the novel concept of affinity networks borrowed from the rule-based modeling approach adds a central unit where molecular and inter-molecular reactions are defined together with their governing rates and rate constants. In comparison, the c$\pi$ requires rate laws and local affinity graphs to be included in the species definitions. This not only impacts the syntactic clarity but also restricts the re-usability of species definitions. However, we note that both process-calculi rely on compositional semantics, allowing systems to be of multiple smaller ones.

**Affinity:** We noted that c$\pi$ restricts us to using MA kinetics for species interaction. As discussed in Section 2.2.3, biologists often use approximations and simplifications such as MM and GK which allow for a smaller system representation. Each could be decomposed into simple MA kinetics. However, often reaction rates are encoded in function parameters (see GK in Section 2.2.3), hence are not stated explicitly in the system description or experimental results. Wang [84] raised this issue while implementing in c$\pi$ the systems presented in [82].

**Bonding** As c$\pi$ purely relies on directional communication channels, it is not possible to model symmetric bonding. Instead, we would have to model agents choosing non-deterministically in which bonds they engage [88]. Here, the bond-calculus offers a completely symmetric

communication operation as well as location binding on sites. Once internal sites have the same location, they can interact symmetrically and allosterically. A common example where such symmetric bonding is homooligomerization where multiple identical species (monomeres) bind into single units (polymeres) [88]. However, we note that both $c\pi$ and the bond-calculus incorporate the notion of *name passing* or *mobility* which allows to describe dynamically created processes [88, 8, 89].

**Overlapping sites:** Ultimately, we note that the bond-calculus is capable to model dependent sites within a single species. Through bonding, sites can change their affinity and not be receptive for certain other sites while they were so before. The $c\pi$, as do other process-algebra approaches, model reaction sites as channels and therefore assume these to be independent [88].

## 2.4 Model Evaluation

We now discussed what approaches exist and how to formally model a biochemical system in the bond-calculus. The sets of ODEs generated by this process calculus are also referred to as non-linear hybrid automata, describing our system's behaviour over time. This allows us to apply concepts such as model checking, formal verification, and to investigate behaviour equivalences [49]. Based on Metric Interval Temporal Logic (MITL) and Logic of Behaviour in Context ($\mathcal{LBC}$), Logic of Behaviour in Uncertain Contexts ($\mathcal{LBUC}$) provides a formal framework to evaluate temporal properties of such hybrid automata. We present a succinct syntax of $\mathcal{LBUC}$ as well as the underlying formalism of Flow*, a verification tool for cyber-physical systems.

### 2.4.1 Temporal Logic

For a basic overview of the general ideas of Linear Temporal Logic (LTL) and timed automata presented we refer the reader to [7]. Furthermore, we include a succinct summary of the LTL syntax as reference for the reader in Appendix D.3. LTL extends propositional or predicate logic into the temporal dimension. It introduces modalities to reason about the temporal behaviour of a reactive system and is a key component of formal software verification [7].

The non-linear ODE systems generated by bond-calculus models can be referred to as such as reactive system. Specifically, a non-linear *hybrid automata* on which formal verification can be conducted [47]. Furthermore, we can specify these systems as (finite) *timed automata*, a sub-class of hybrid automata. This allows us to include the concept of *clocks* which allow us to quantify the temporal dimension and keep track of time throughout simulations of the automaton [1].

However, the syntax of LTL itself does not accommodate for timed operators, hence cannot include time values from clocks. Thus, a multitude of temporal logics were devised, allowing to quantify the temporal aspect of simulations through timed operators [2]. One such logic is the

*Metric Temporal Logic* (MTL) which extends the LTL operators with time-bounded versions. It is defined over a point-based integer-time semantics [2].

Alur and Henzinger [3] discuss that due to the integer-time assumed by MTL, evaluating the satisfiability of real-timed system properties is *undecidable*. Informal relaxations such as abstracting timed automata to a finite discrete time domain were presented. However, formally, the MITL by Alur et al. [5] provides a framework where intervals rather than discrete time-steps are considered for the decision on satisfaction of temporal properties.

As both the syntax of $\mathcal{LBC}$ and $\mathcal{LBUC}$ combine LTL($\mathbb{R}$), LTL in continuous state-space, and MITL, we here present succinctly the MITL syntax:

$$\begin{aligned} \phi, \psi ::= \quad & \text{Atom} \mid \phi \wedge \psi \mid \phi \vee \psi \mid \phi \Longrightarrow \psi \mid \neg\phi \mid \\ & \mid \phi \mathbf{U}_I \psi \mid \mathbf{F}_I \phi \mid \mathbf{G}_I \phi \\ \text{Atom} ::= \quad & \top \mid \bot \end{aligned} \tag{2.21}$$

Here $\phi, \psi$ can either be atomic prepositions or MITM formulas which are recursively defined by the same syntax. While the first line re-iterates the basic grammar from propositional logic, the second line presents a sub-set of the operators defined by LTL. The *next* operator was excluded by Alur and Henzinger while the operators *until* ($\phi \mathbf{U}_I \psi$), *eventually* ($\mathbf{F}_I \phi$), and *globally* ($\mathbf{G}_I \phi$) are altered to include interval $I$. Here, $I$ is any non-singular interval in the domain of $\mathbb{R}_{\geq 0}$ such that $[a,b] \mid a,b \in \mathbb{R}_{\geq 0}$ and $a \neq b$ (although the interval can also be non-inclusive on the lower and/or upper bound). As an example, $\mathbf{F}_{[0,4]}\phi$ would express, '$\phi$ is satisfied eventually within 0 and 4 time-steps'.

### 2.4.2 Logic of Behaviour in Uncertain Contexts - $\mathcal{LBUC}$

Banks and Stark [8] present a novel temporal logic, $\mathcal{LBC}$, to allow expressing properties of *behaviour in context*. $\mathcal{LBC}$ builds on top of the MITL syntax presented above and introduces the *context operator* $C \rhd \phi$. Here, $C$ is a process referred to as *context* and $\phi$ a $\mathcal{LBC}$ formula specifying a *behaviour*. A context can be thought of as a newly introduced process $Q$ to the process $P$ on which satisfaction of formula $\psi$ is currently being verified. Formally, we express this in Equation 2.22. With respects to biochemical modeling, this is analogous to the introduction of mixture perturbations or new species within the mixture. Note that the concentrations introduced by $Q$ are added ($+$) to the current concentrations of process $P$.

$$P \models Q \rhd \psi \quad \Longleftrightarrow \quad Q \parallel P \models \psi \tag{2.22}$$

From this, $\mathcal{LBUC}$ arises and introduces uncertainty to the syntax of $\mathcal{LBC}$. Through the quantification of uncertainty, we can reason about sets of possible system states rather than being bound to exact quantities. Following we present, the $\mathcal{LBUC}$ syntax extending the $\mathcal{LBC}$ syntax defined

in [8]. We note that there is no full reference for $\mathcal{LBUC}$ yet available.

$$\begin{aligned}
\phi, \psi ::= \;& \text{Atom} \mid \phi \wedge \psi \mid \phi \vee \psi \mid \phi \Longrightarrow \psi \mid \neg\phi \mid \\
& \mid \phi \mathbf{U}_I \psi \mid \mathbf{F}_I \phi \mid \mathbf{G}_I \phi \mid Q \rhd \phi \\
Q ::= \;& [S_1]S_1 \parallel \cdots \parallel [S_n]S_n \\
\text{Atom} ::= \;& \text{Val ROp Val} \\
\text{Val} ::= \;& v \in \mathbb{R} \mid [S_n] \mid \text{ Val AOp Val} \\
\text{ROp} ::= \;& > \mid < \mid \geqslant \mid \leqslant \\
\text{AOp} ::= \;& + \mid - \mid \times \mid \div
\end{aligned} \tag{2.23}$$

While $\phi, \psi$ follow the same MITL syntax defined above in Section 2.4.1, $\mathcal{LBUC}$ introduces $Q$, a process introduced by the context operator to formula $\phi$. $Q$ itself is formulated as the parallel composition of an arbitrary number of species $S_n$ and their concentrations $[S_n]$ (as described in Sections 2.3.2 and 2.3.3). $[S_n]$ in turn can be $[A] \in \mathbb{R}_{\geq 0}$ or $[A] \in [a,b] = \{x \in \mathbb{R}_{\geq 0} \mid a \leq x \leq b\}$. Hence, *quantifying uncertainty* through the definition of a real-valued interval. ROp and AOp represent the relational and arithmetic operators respectively.

### 2.4.3 Flowstar

In order to computationally show the satisfaction of $\mathcal{LBUC}$ formulae, we rely on the work by Chen [14]. The approach presented allows us to verify a formula on non-linear, non-polynomial hybrid CPSs by conducting *over-approximate reachability analysis* over the sytem's state-space constructs called flowpipes [14].

Hybrid automata often appear in context of safety-critical applications (there as cyber physical systems; CPSs). Thus, it is often required to verify whether they satisfy certain required safety properties. However, this is notoriously difficult and different approaches have been proposed using paradigms from the field of control theory and formal verification [14]. The work of Chen [14] presents a hybrid approach to reachability set over-approximation through the use of Taylor model flowpipes and support function calculus. The work is made accessible through the Flowstar (Flow*) framework.

To prove that no behaviour of the model violates a given property we use *reachability analysis*. Here, usually the set of all reachable states from a given start state of a model are computed. If no unsafe state is included, then the system is said to be safe. However, as shown by Alur et al. [4], the reachability of a state for a hybrid automata (the underlying mathematical system of CSPs) is not decidable. Therefore, *over-approximate reachability analysis* for hybrid automata has been studied intensively over the past decades using a scheme called *flowpipe construction* [14]. Here we build on the fact that, if the over-approximated set of reachable states satisfy a safety property, inherently the exact reachability set will do so as well. Flowpipes aid to formalize this approach over finite time and trace the progression of a group of state sets over a specific time interval.

## 2.5 Existing Tools

Throughout this thesis we rely on the bond-calculus work bench *bondwb* by Wright for our implementation [87]. It allows us to specify bond-calculus models with their species, rate laws, affinity networks, and processes in a syntax *closely resembling* the one outlined in Section 2.3.3. These are stored in *.bond* files which are passed to the bondwb. In turn, following the semantics of the bond-calculus it generates a system of ODEs expressing the rate of change of each species.

The bondwb also provides us with the ability to evaluate $\mathcal{LBUC}$ formulae on the generated systems. It hence supports real field intervals for the concentration values of the initial mixtures as well as the introduction of contexts into the system. Underlying the bondwb, frameworks such as Sage Math [74] and Flow* (Section 2.4.3) are required for the operation. The former provides a framework with a wide range of mathematical tools such as symbolic expressions and ODE solvers, to name a few. The latter is necessary for our work with $\mathcal{LBUC}$.

While other frameworks than Flow* were considered (e.g. SpaceEx [31] for the verification of $\mathcal{LBUC}$ formulae, its performance is comparatively equal to these and it provided unique features such as symbolic restriction of flowpipes without the need for re-computation.

Furthermore, Flow* guarantees that model checking is exact and no numerical approximation is needed due to the use of *validated integration* [14]. This implies that once a $\mathcal{LBUC}$ formula is satisfied, it is definite that the system has the property described. However, there is the possibility that systems do not converge, inhibiting Flow* to provide firm results.

While the bondwb itself is written in the functional programming language Haskell, we accessed it through its Python bindings, allowing us to use jupyter notebooks for the implementation of the various systems and the subsequent analysis. Our contribution in terms of software consists in a *plotting class* which allows to analyze the generated ODE systems visually. We accounted for the different settings and can plot continuous or discrete graphics over time, showing the system's progression and, if possible, it's ultimate, sometimes quasi, steady state. However, it is also possible to plot phase plane diagrams as in [82] allowing to inspect oscillatory behaviour.

# Chapter 3

# Presentation of Models

Following the functional motifs presented by Tyson et al. [82], this chapter presents the components and systems studied and ultimately modelled in Chapter 4. We first discuss the necessary simplifications and assumptions taken by both the authors and us to quantify the validity of the discussed biochemical components. Following, we present what we refer to as the *elementary building blocks* of biochemical systems in order of increasing complexity. We give a clear interpretation of each system and their constituents. Moreover, we provide biochemical context and illustrate which currently known biological systems make use of these.

Subsequently, we derive and discuss the relevant set of ODEs describing each system's behaviour. We do so by applying methods from system analysis [19] to infer differential equations from schematic representations of the occurring processes in so-called *wiring diagrams* as shown in Figure 3.3.1.

Furthermore, our contribution in this section is the quantitative descriptions of the characteristic features of the system's behaviour. We will refer back to these when encoding these formally in $\mathcal{LBUC}$ in Chapter 4. In order to provide concise statements, we refer to Table D.1 in Appendix D.5 which contains all numerical values of rate constants and initial concentrations for the systems discussed as presented in [82]. Note that our characterizations of the system behaviours presented here are non-exhaustive.

## 3.1 Simplifications and Assumptions

As discussed in Section 2.1, certain simplifications and assumptions are necessary in the context of biochemical modeling.

One is the assumption that each system is *well-mixed*, such that there are no spatial concentration gradients of the species. This is reasonable when considering extracts residing outside the context of a living cell or within, at least yeast-sized cells as devised by Conrad and Tyson [19] which is assumed in [82].

The second assumption is that the concentrations of species are continuous functions of time. This assumption is valid for sufficiently large number of molecules of each species within the

reaction volume. We here assume that each species is represented by at least a few thousand molecules in the respective mixture [19].

We refer to a species *S* as the *signal* or *stimulus* emitted by the cell nucleus to enact a specific cellular responses (see Section 2.2.2). Hence, the concentration of a signal is the concentration of a certain enzyme which was generated by the mRNA translation. The translation mechanism itself is abstracted away in order to simplify the models discussed. However, as Tyson elaborates via private communication, this simplification is not reducing the validity of the model [78]. We include further assumptions for the sake of completeness:

1. The concentration of a certain signal species is constant. Hence, no mRNA degradation, mRNA splicing variability, or translation inhibitions assumed [11].

2. The influence of affinity compatibility between signal and response species is neglected.

3. The presence of cooperative kinetics within the mixture is neglected [16].

## 3.2 Biomodels Database

As an additional resource to the models presented in [82, 19, 80], some of our models are informed by data available on the Biomodels database [53]. It contains a collection of curated and non-curated biochemical models expressed in SBML [40]. While the models are not by Tyson et al., we verified the consistency between the implementation and literature by comparing rate constant values and differential equations. We will be referring to the models by their official identifier throughout the text.

## 3.3 Elementary Building Blocks

### 3.3.1 Linear Response

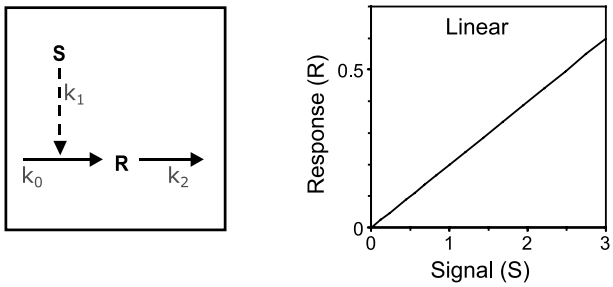



Figure 3.1 Linear Response. Left: wiring diagram of synthesis and degeneration of species *R* with respective rate constants. Species *S* is a signal enzyme resulting from mRNA translation and promotes the synthesis of species *R*. Right: the resulting steady-state signal-response diagram. Figure modified from [82] Figure 1 (a) with permission.

The first motif is the simultaneous synthesis and degeneration of molecules, fundamental to protein dynamics. It results in a linear signal-response curve at steady-state. Finding the motif's ODE system is easily achieved by inspecting the wiring diagram in Figure 3.1. Species $S$ represents an enzyme produced by the mRNA translation process as discussed in Section 2.2 and 3.1. For all systems the concentration of the signal, $[S]$, will not change over time. Hence, we set $\frac{dS}{dt} = 0$ and will not explicitly state this in the future.

In the case of this motif, $S$ promotes the synthesis of response species $R$ (e.g. a protein). $R$ furthermore is subject to an ambient synthesis as well as a constant decay. All these actions occur at certain, unknown rates $k_0, k_1, k_2$. These are usually found experimentally. However, in the case of [82], rates are chosen such that the systems discussed follow a desired signal-response curve. For this functional motif, $k_0 = 0.01$,$k_1 = 1$, $k_2 = 5$, and the initial concentration $[R_0] = 0$. However, for subsequent systems we refer to Table D.1 in Appendix D for the full presentation of the numerical values. Equation 3.1 formalizes this description. By setting $\frac{dR}{dt} = 0$ we can devise the steady-state concentration $R_{SS}$ as in Equation 3.1. Note that, as $[S] = 0$, $R_{SS} = \frac{k_0}{k_2}$ which evaluates to 0.002 with the rate constants provided by the authors. Hence, the signal-response curve in Figure 3.1 does not exhibit a significant y-intercept.

$$\frac{dR}{dt} = k_0 + k_1 S - k_2 R \tag{3.1}$$

$$R_{SS} = \frac{k_0}{k_2} + \frac{k_1}{k_2} S \tag{3.2}$$

Finally, we note that Tyson et al. only includes the ambient generation of a response species in this system as well as the systems exhibiting mutual inhibition and negative-feedback oscillation. For all other systems, although specified in the wiring diagram, the term quantifying ambient generation was omitted.

**Characterization:** For our quantitative evaluation with $\mathcal{LBUC}$ we here characterize the features of the system behaviour. In Chapter 4 we will devise the $\mathcal{LBUC}$ formulae which we subsequently evaluate in Chapter 5. We note about the system that:

1. As the concentration of species $S$ increases, we observe an increase in $[R_{SS}]$ proportional to $\frac{k_0}{k_2} + \frac{k_1}{k_2}[S] = 0.002 + 0.2 \cdot [S]$.
2. $[R_{SS}]$ is independent of $R$'s initial concentration.
3. This independence holds for any perturbation of $[R]$ introduced during the simulation of the system.
4. The linear relation holds for any perturbation of $[S]$ during the simulation.

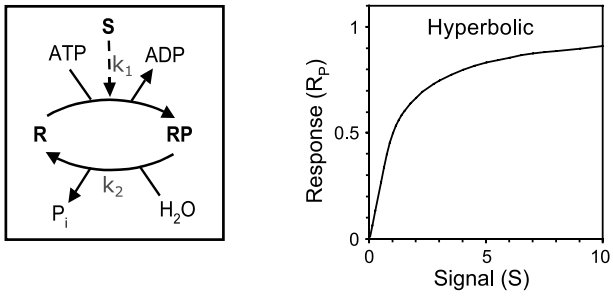



Figure 3.2 Hyperbolic Response. Left: wiring diagram with respective rate constants. Right: steady-state signal-response diagram. Modified from [82] Figure 1 (b) with permission.

### 3.3.2 Hyperbolic Response

The underlying process of this signal-response motif is phosphorylation and dephosphorylation of a response species $R$ through kinase $S$ governed by MA kinetics.

We described in Section 2.2.3 how this process changes the affinity of a protein by setting it to a state referred to as *active*. ATP and water play an important role in this reaction, however are abstracted for simplicity in this case. Examples are the functional regulation of enzymes in yeast cells [77], or the activation of necessary energy resources upon muscle contraction in mammals, to name a few [18].

The wiring diagram in Figure 3.2 illustrates how $S$, as kinase, promotes the phosphorylation of $R$ into $R_P$ at rate $k_1$. $R_P$ then dephosphorylates into $R$ at a constant rate $k_2$ without promotion through a phosphatase. We note, that this is an assumption by the authors violating this enzymatic reaction couple. The correct representation is illustrated in Figure 3.3 where a phosphatase $S_P$ drives the dephosphorylation reaction. However, we proceed with the interpretation of Tyson et al..

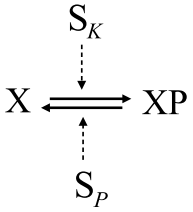



Figure 3.3 Correct phosphorylation/dephosphorylation wiring diagram including a phosphatase $S_P$. Modified from [19] Figure 6.6 (a).

$$\frac{dR}{dt} = -k_1 SR + k_2 R_P \tag{3.3}$$

$$\frac{dR_P}{dt} = k_1 SR - k_2 R_P = k_1 S(R_T - R_P) - k_2 R_P \tag{3.4}$$

$$dR_{P,SS} = \frac{R_T S}{\frac{k_2}{k_1} + S} \tag{3.5}$$

The governing ODEs in Equation 3.4 are rewritten in terms of the total concentration of the response element $R_T = R + R_P$. We note, that the rates of $R$ and $R_P$ are complimentary and while $R_T = const.$ the concentration of $R$ behaves inversely to the concentration of $R_P$.

As discussed in Section 2.2.3, at $[S] \ll K_m$, the reaction velocity is limited by the amount of available $S$, hence increasing $[S]$ increases the reaction speed. As Tyson et al. [82] did not provide enough information to determine $K_{m1}$ and $K_{m2}$ for the phosphorylation and dephosphorylation respectively, we assume that this assumption holds. Thus, we can use pure MA kinetics to describe the system reactions.

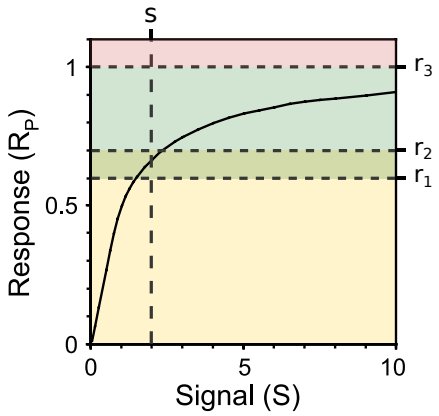



Figure 3.4 Hyperbolic signal-response curve with subdivisions for quantitative characterization. $r_1 = 0.6$, $r = 0.7$, $r = 3 = 1$, $s = 2$. **Colors:** Low values of $R_P$ in orange, high values of $R_P$ in *green*, not reachable values in *red*. Modified from [82] Figure 1 (b) with permission.

**Characterization:** Here we use a graphical interpretation of the signal response diagram to reason about the bounds characterizing the system behaviour (see Figure 3.4).

1. As we increase $[S]$, the steady-state concentration of the phosphorylated response will globally stay below an asymptotic maximum $\lim_{[S]\to\infty} dR_{P_{SS}}([S]) = [R_T]$. In our case, $[R_T] = 1$ and in Figure 3.4 denoted as $r_3$.

2. The asymptotic behaviour holds for any perturbation of $[S]$ during the simulation.

3. As we increase $[S]$ from low to high values, the increase in $[R_{P_{SS}}]$ is first rapid and then gradually slows down. In Figure 4.1, $[S] = s = 2$ acts as a threshold between low and high values of $[S]$. Bounds $r_1 = 0.6$ and $r_2 = 0.7$ reflect our bounds for high and low values in $[R_{P_{SS}}]$. Hence, as $[S] < s$, $[R_{P_{SS}}] < r_2$ and as $[S] \geq s$, $r_1 < [R_{P_{SS}}] < r_3$. The bounds for $[R_{P_{SS}}]$ have a margin of 0.1 to relax the boundaries for this property and therewith allow $\mathcal{LBUC}$ to compute results.

### 3.3.3 Sigmoidal Response

For this signal-response motif we reconsider the phosphorylation and dephosphorylation process presented above in Section 3.3.2. However, in this case the MM kinetics are governing the

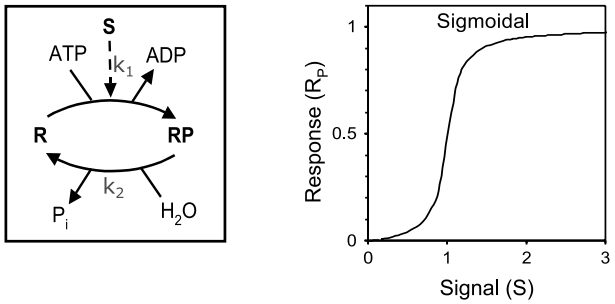



Figure 3.5 Sigmoidal Response. Left: wiring diagram with respective rate constants. Right: steady-state signal-response diagram. Modified from [82] Figure 1 (c) with permission.

reactions. This would be the case if the total response concentration $[R_T]$ is not significantly bigger than the MM constants of kinase and phosphatase. The steady-state signal-response curve is of sigmoidal shape and is referred to as zero-order ultras-sensitive switch [33]. However, as discussed in Section 2.2.3, the ultra-sensitivity can only be observed while the MM constants $K_{m1}, K_{m2} \ll 1$. Tyson et al. relate this motif to a *buzzer* component in electronics. This analogy refers to the very abrupt but reversible predominance of $R_P$ within the mixture once $[S]$ is big enough.

$$\frac{dR_P}{dt} = \frac{k_2 S_{K,total}\,(R_T - R_P)}{K_{m1} + R_T - R_P} - \frac{k_4 S_{P,total} R_P}{K_{m2} + R_P} \tag{3.6}$$

In order to devise the governing system of ODEs for this functional motif, we refer to the work by Wang [84] and present our own derivation of $\frac{dR_P}{dt}$ in Appendix D.4. Here, $S_{K,total}$ refers to the total concentration of kinase and $S_{P,total}$ to the total concentration of phosphatase (as in Figure 3.3).

From our work, we note that Tyson et al. neglected the presence of a phosphatase guiding the reaction $R_P \to R$ and assumed that this occurs at a constant rate $k_2$ with respective MM constant $K_{m2}$. Hence, $S_{P,total} = 1$ in our subsequent implementation and the equations provided in [82]. In Equation 3.6, the positive term describes the phosphorylation $R \to R_P$ while the negative term represented the dephosphorylation $R_P \to R$.

**Characterization:** We devised bounds for this system's behaviour illustrated in Figure 3.6. We deduct from these the following requirements:

1. Similar to the hyperbolic motif, $[R_{P_{SS}}]$ will globally stay below the asymptotic limit of $r_3 = [R_T] = 1$ as $[S] \to \infty$.

2. The asymptotic behaviour holds for any perturbation of $[S]$ during the simulation.

3. As $[S] \leq s_1$, $[R_{P_{SS}}] < r_1$. Once $s_1 < [S] < s_2$, $[R_{P_{SS}}]$ increases rapidly, reaching $[R_{P_{SS}}] = r_2$ quickly. Thus, within a small change of $[S]$, $[R_{P_{SS}}]$ rises quickly. We encode this by having a small difference of $[S]$ between a low value and a high value of $[R_{P_{SS}}]$, we choose $s_2 - s_1 = 0.5$. As $s_2 < [S]$, $r_2 < [R_{P_{SS}}] < r_3$.

4. The ultra-sensitivity of the switch is only retained when $K_{m1}, K_{m2} \ll 1$.

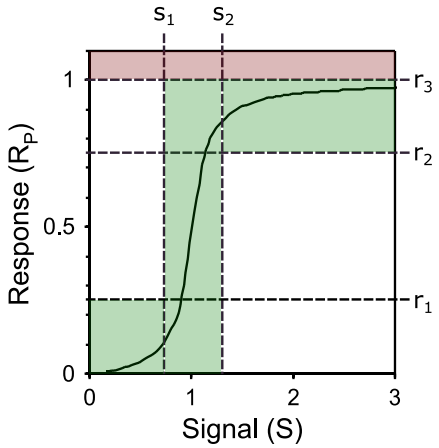



Figure 3.6 Sigmoidal signal-response curve subdivided into regions for quantitative evaluation. Sigmoidal signal-response curve with subdivisions for quantitative characterization. $r_1 = 0.25$, $r = 0.75$, $r = 3 = 1$, $s_1 = 0.7$, and $s_1 = 1.3$. **Colors:** Values of $[R_P]$ in *green* describe ultrasensitive behaviour, not reachable values are in *red*. Modified from [82] Figure 1 (c) with permission.

### 3.3.4 Perfect Adaptation

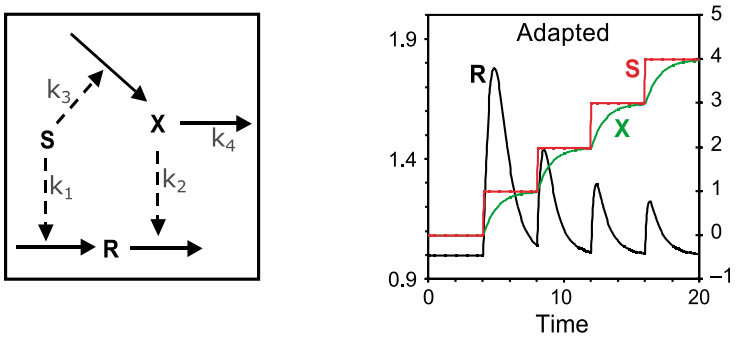



Figure 3.7 Perfect Adaptation. Left: wiring diagram with respective rate constants. Right: Behaviour of concentrations $[X], [R]$ over time as $[S]$ exhibits discrete step-wise increase. Modified from [82] Figure 1 (d) with permission.

As mentioned in the overview of this section, the components are in order of increasing complexity and building upon each other. In the case of perfect adaptation, we combine two linear response signaling pathways (Section 3.3.1) as illustrated in Figure 3.7. The response mechanism of perfect adaptation entails that the response species $R$ exhibits a *transient response* upon changes of $[S]$. However, the steady-state response of $R$ is independent of $[S]$. In [82], the authors build an analogy to the sense of smelling. It similarly reacts upon new odors, however adapts to persistent input. They refer to this functional motif as "sniffer".

As one of the best studied biological processes, *chemotaxis* represents one of the fundamental mechanisms relying on system adaptation. It is the biasing of movement of an organism towards environments that contain higher concentrations of beneficial, or lower concentrations of toxic, chemicals [83]. Mathematical models of such system behaviour were devised early in the context

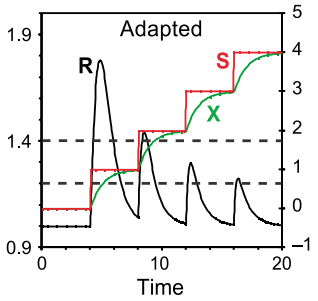



Figure 3.8 Perfect adaption temporal progression diagram illustrating reduction of the intensity of spiking of $[R]$ as $[S]$ is higher prior step-wise increase. Modified from [82] Figure 1 (c) with permission.

of system biology [71].

$$\frac{dR}{dt} = k_1 S - k_X R \tag{3.7}$$

$$\frac{dX}{dt} = k_3 S - k_4 X \tag{3.8}$$

To deduce the ODEs governing this system, we observe that $S$ promotes the synthesis of both response species $R$ and auxiliary species $X$ at some rates $k_1$ and $k_3$ respectively. However, while species $X$ degrades at a constant rate $k_4$, the degradation of $R$ is proportional to $X$ by factor $k_2$. When a sudden increase of $[S]$ occurs, the synthesis of $R$ is bigger than its degeneration. However, as $X$ adapts to the new concentration of $S$, synthesis and degradation of $R$ become gradually equal, resulting in $[R]$ to return to its steady-state value. In Equations 3.9 and 3.10 we identify the steady-states of $X, R$ and note that $[X_{SS}] = [S]$ and that $[R_{SS}]$ is equal to $R$'s initial value $R_0 = 1$ in our case. The following equations formally describe this:

$$R_{SS} = \frac{k_1 S}{k_2 X_{SS}} = \frac{k_1 k_4}{k_2 k_3} = 1 = [R_0] \tag{3.9}$$

$$X_{SS} = \frac{k_3 S}{k_4} = S \tag{3.10}$$

**Characterization:**

1. As we introduce a step-wise increase of $[S]$, we observe:
   (a) $[R]$ will rise as a direct response to the increased synthesis. It will exceed a threshold and eventually return to a steady-state value ($R_{SS} = \frac{k_1 k_4}{k_2 k_3} = 1$ in our case) as soon as $[X]$ adapted to the new $[S]$.
   (b) $[X]$ will adapt to $[S]$ and their concentrations will eventually be equal. This is due to $k_3 = k_4 = 1$ in our case. Hence, $[X] = [S]$ eventually.

2. The higher $[S]$ is *prior* the step-wise increase, the lower will be the spike in $[R]$ when an increase in $[S]$ is induced (observe reducing peaks in Figure 4.3).

3. The steady-state response $R_{SS}$ is independent of its initial value $R_0$. Hence, any perturbation of $[R]$ during a simulation, results in the same value for $R_{SS}$.

### 3.3.5 Mutual Activation

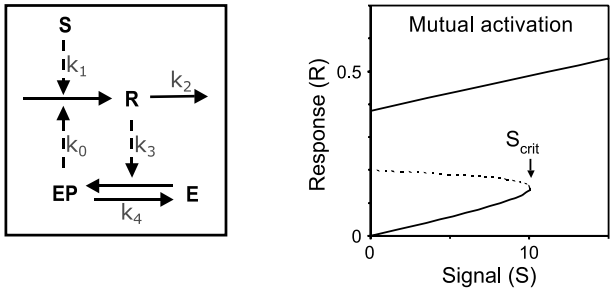



Figure 3.9 Mutual Activation. Left: wiring diagram with respective rate constants. Right: steady-state signal-response diagram. Modified from [82] Figure 1 (e) with permission.

Just like fuses or one-way switches in electronics, biochemical systems too can include components which lead to irreversible state changes within. Positive feedback, as discussed in Section 2.2.1, provides a framework for such behaviour. It allows for reactions, as a certain threshold of signal is reached, to auto-sustain their activity leading to a terminal state for the system unless counteracted. Well known examples of where such behaviour can be observed are apoptosis (controlled cell death) or stage changes of cellular growth [37, 11].

Observing the wiring diagram in Figure 3.9, we can recognize that the system consists of a combination of synthesis and degeneration (species $R$ and $S$) as well as phosphorylation and dephosphorylation ($E \rightleftharpoons E_P$). The synthesis of $R$ is steered by both, $[S]$ and the activated enzyme $E_P$ at rates $k_1$ and $k_0$ respectively. $R$ functions as the kinase of $E \rightleftharpoons E_P$ at rate $k_3$. We note again that Tyson et al. neglected the existence of a phosphatase regulating the dephosphorylation of species $EP$ (as discussed in Section 3.3.3). Hence, $E_P$ dephosphorylates at a constant rate $k_4$.

$$\frac{dR}{dt} = k_0 E_P + k_1 S - k_2 R \tag{3.11}$$

$$\frac{dE_P}{dt} = \frac{k_3 R\,(E_T - E_P)}{K_{m1} + E_T - E_P} - \frac{k_4 S_{P,total} E_P}{K_{m2} + E_P} \tag{3.12}$$

Equations 3.13 and 3.14 illustrate the ODEs governing the rate of the response element and the phosphorylated species $E_P$. We applied a similar procedure as in Section 3.3.3 (elaborated in Appendix D.4) and refer to $K_{m1}$ and $K_{m2}$ as the MM constants, and to $S_{K,total}$ as the total concentration of the kinase. Following the discussion in Section 2.2.3, the steady-state concentration of the phosphorylated component $E_P$ can be expressed in terms of $R$ with its biophysically

acceptable steady-state solution provided by the GK function.

$$\frac{dR}{dt} = k_0 E_P(R) + k_1 S - k_2 R \tag{3.13}$$

$$E_P(R) = G(k_3 R, k_4, J_3, J_4) \tag{3.14}$$

The behaviour can be illustrated in terms of a bifurcation diagram as in Figure 3.9 (right). Here, as we increase $[S]$ to a critical value $S_{crit}$, $[R]$ will be big enough to *switch on* the $E \rightleftharpoons E_P$ buzzer. $R$ then engages in an *autocatalytic* process with $E_P$ as a direct result from the positive feedback and synthesises enough of itself to be independent of $[S]$. As we decrease $[S]$, $[R]$ and $[E_P]$ will remain high, giving the system no possibility to reduce the response element anymore to a low concentration. In system's analysis, when a system's state is dependent on its history, it is referred to as *hysteresis* [19]. Mathematically, this is referred to as a one-parameter bifurcation where the system is bi-stable for values $[S] \in [0, S_{crit})$. It has two stable steady-state response values separated by an unstable steady-state [82].

**Characterization:**

1. While $[S] \in [0, S_{crit})$ we observe the steady-state concentration of $R$ to be $[R_{SS}] \lessapprox 0.14$.
2. As $[S]$ is increased to $S_{crit} \approx 10.2$, we observe a discontinuous jump (bifurcation) at from $[R_{SS}] \approx 0.15$ to $[R_{SS}] \approx 0.49$. The values refer to the results obtained given the rate constants in [82].
3. As $[S] \geq S_{crit}$, $[R_{SS}] \geq 0.49$.
4. As we decrease $[S] \to 0$, $[R_{SS}] \geq 0.38$. Hence, the steady-state response remains in a state of *high concentration* due to the autocatalytic process with $E_P$.
5. If a perturbations of $[R]$ is introduced during simulation such that $[R] > 0.2$, the steady-state concentration $[R_{SS}]$ will jump to a high concentration $[R_{SS}] \geq 0.38$.

### 3.3.6 Mutual Inhibition

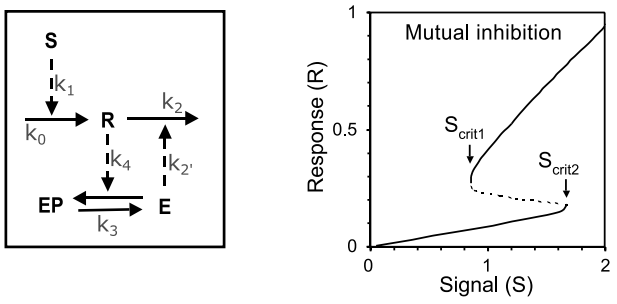



Figure 3.10 Mutual Inhibition. Left: wiring diagram with respective rate constants. Right: steady-state signal-response diagram. Modified from [82] Figure 1 (f) with permission.

For this motif we re-use the components of the previous system, however in this case $E$ promotes the degradation of $R$. However, as $[R]$ increases, more of $E$ will be present in its phosphorylated form $E_P$. Thus, a positive feedback on the degradation of $R$ is in-place, promoting the effect of the signal. The resulting system also exhibits hysteresis. Tyson et al. refer to it as a discontinuous two-way switch, or "toggle", and is the reversible pendant to the configuration above. Well known examples such behaviour can be observed in the cell cycle of budding yeast cells [13, 20]. Moreover, more research is recently devoted to the positive feedback mechanism as a whole and the bistability presented here was found to be fundamental for the motility (free motion) of eukaryotic cells [54].

$$\frac{dR}{dt} = k_0 + k_1 S - k_2 R - k_2' E(R) \cdot R \tag{3.15}$$

$$E(R) = G(k_3, k_4 R, J_3, J_4) \tag{3.16}$$

As in synthesis and degeneration, $S$ synthesises $R$ at rate $k_1$ which is also subject to an ambient synthesis at a constant rate $k_0$. Furthermore, $R$ also degrades at a constant rate $k_2$. $R$ promotes the phosphorylation of $E \rightleftharpoons EP$ at rate $k_4$ while $E$ promotes the degeneration of $R$ at rate $k_2'$. $R$ increases as $S$ increases. However, as more of the total concentration $E_T$ is in form of $E_P$, the degradation of $R$ decreases as the concentration of $E$ reduces. Hence, once the concentration of $S$ reaches a critical level $S_{crit1}$, the sigmoidal switch from Section 3.3.3 switches from high $E$ to low. Degradation of $R$ is quickly decreased and a sudden increase of its concentration can be observed. As we now reduce the concentration of $S$ again, the concentration of $R$ remains high until it does not suffice anymore to promote $E \rightarrow E_P$ to keep the switch low. Once $S$ reaches a value as low as $S_{crit2}$, the high concentration of $E$ promotes the degradation of $R$, resulting in a sudden decrease of $[R]$. Equations 3.11 and 3.16 are the governing ODEs. We again apply the GK function for the steady-state concentration of $E$.

As noted in Section 3.3.1, here $R$ has an explicit ambient synthesis as well as a degeneration dependent on its concentration, similar to the linear response system. These are additionally to the synthesis by $S$ and degeneration through $E$ respectively.

**Characterization:**

1. Similar to the system before, as $[S] \in [0, S_{crit2})$, the steady-state response $R$ will remain below a threshold of $[R_{SS}] \lessapprox 1.7$. In our case, $S_{crit2} \approx 1.7$.
2. However, as $S = S_{crit2}$, we observe a discontinuous jump in $[R_{SS}]$ such that $[R_{SS}] \approx 0.79$.
3. As $[S]$ is increased, $[R_{SS}] \gtrapprox 0.76$.
4. As $[S]$ is decreased and $[S] > S_{crit1}$, $[R_{SS}] \gtrapprox 0.28$.
5. As $[S] = S_{crit2} \approx 0.85$ a discontinuous jump in $[R_{SS}]$ is observed to the lower values previously observed.

6. Perturbations of $[R]$ introduced during simulation can result in a state change of $[R_{SS}]$ depending on $[S]$.

### 3.3.7 Homeostasis

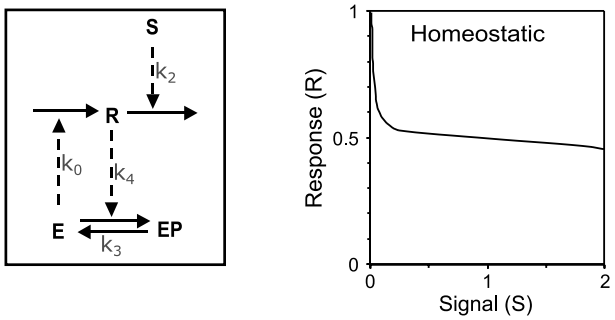



Figure 3.11 Homeostasis. Left: wiring diagram with respective rate constants. Right: steady-state signal-response diagram. Modified from [82] Figure 1 (g) with permission.

In the case of homeostasis, the concentration of the response species is confined by a narrow band for a wide range of signal concentration as in Figure 3.11. Underlying is a configuration referred to as negative feedback (Section 2.2.1). Here, the response element counteracts the effect produced by the signal. It has been shown early on, that components enacting homeostasis allow systems to gain resilience over varying input and the ability to reach plateau levels of certain species concentrations even with inaccurate stimulus concentrations [76]. An excellent example is the blood sugar regulation through insulin within our body, also referred to as *glucose homeostasis* [21]. Here, the pancreas secretes insulin as a response to elevated blood sugar until it reaches a plateau concentration upon which secretion is halted. Moreover, homeostasis is at the core of many other cellular and physiological activities such that dysfunctions of this negative feedback loop is core to much research in the field of pathogens [15].

$$\frac{dR}{dt} = k_0 E(R) - k_2 S \cdot R \tag{3.17}$$

$$E(R) = G(k_3, k_4 R, J_3, J_4) \tag{3.18}$$

Figure 3.11 illustrates how species $R$ is synthesized at rate $k_0$ by species $E$. Similar to both mutual activation and mutual inhibition, $R$ acts as kinase for the phosphorylation $E \rightarrow E_P$ at rate $k_4$. Hence, at high concentrations of $R$, $E_P$ is predominant in the system. The signal $S$ degrades $R$ at rate $k_2$. As we increase the concentration of $S$, the concentration of $R$ within the system decreases. As a result, the concentration of $E$ will rise due to $E_P$ dephosphorylating. This in turn promotes the synthesis of $R$ which results in a negative feedback to the stimulus. Hence, for a wide range of $S$, the synthesis of $R$ will be equivalent to its degeneration. The system of ODEs in Equations 3.17 and 3.18 describe the system over time.

**Characterization:**

1. With approximately $[S] \in [0.6, 2]$, the concentration of the steady-state response will be approximately $[R_{SS}] \in [0.6, 0.4]$ in our case. Hence, we characterize that for a wide band of $[S]$, $[R_{SS}]$ keeps within a tight bound.

2. $[R_{SS}]$ is independent of the initial concentration of $R$ as well as perturbation of its concentration throughout the simulation.

## 3.4 Oscillators

As discussed in Section 2.2, complex system behaviour can be traced back to positive and negative feedback, promoting and reducing effects between component interactions. Oscillators are crucial conveyors of this statement. Hence, we discuss here three elementary oscillatory motifs which are believed to be at the basis of every oscillator that can be observed in biochemical systems such as PINs and gene regulation networks [60, 12, 90]. The systems presented are build upon the previously presented elementary building blocks.

### 3.4.1 Negative-Feedback

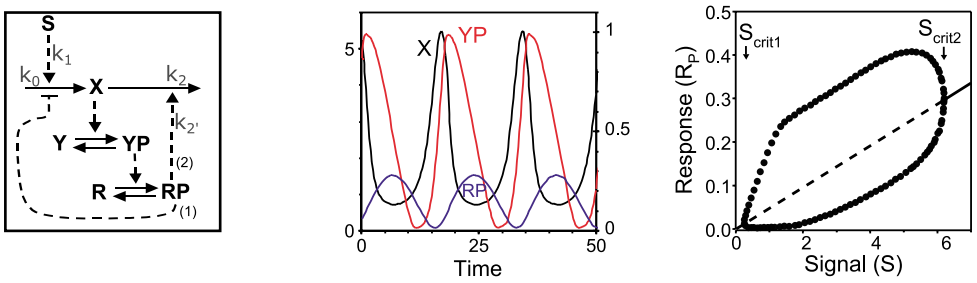



Figure 3.12 Negative-feedback oscillator. Left: wiring diagram with respective rate constants. Center: diagram illustrating oscillation over time of species concentrations $[X], [Y_P], [R_P]$ at $[S] = 2$. Right: signal-response diagram showing the effect of the Hopf-bifurcation; points illustrate maximum and minimum values of oscillating $[R_P]$ as $S_{crit1} < S < S_{crit2}$. Modified from [82] Figure 2 (a) with permission.

Building upon the notion of negative feedback, this oscillator consists of two interlaced phosphorylation/dephosphorylation, or buzzer, circuits which ultimately provide negative feedback to the stimulus. The oscillation is only observed as $S_{crit1} < S < S_{crit2}$ resulting in what is referred to as *Hopf-bifurcation* in the signal-response diagram (Figure 3.12 right). $S_{crit1}, S_{crit2}$ are referred to as *subcritical Hopf bifurcation points* [82]. If $[S]$ is not within these boundaries, the system ceases to oscillate and the species concentrations reach a dynamic steady-state. This kind of oscillators are also referred to as *repressilators* [12] and can be found in MAPK (mitogen-activated protein kinases) signaling pathways as well as circadian rhythms [82]. Following an experiment by Elowitz and Leibler [22] in which they reproduced an artificial oscillator of this kind with a

fluorescent response component, Tyson et al. refer to this oscillator as a *blinker*.

$$\frac{dX}{dt} = k_0 + k_1 S - k_2 X + k_2' R_P \cdot X \tag{3.19}$$

$$\frac{dY_P}{dt} = \frac{k_3 X (Y_T - Y_P)}{K_{m3} + Y_T - Y_P} - \frac{k_4 Y_P}{K_{m4} + Y_P} \tag{3.20}$$

$$\frac{dR_P}{dt} = \frac{k_5 Y_P (R_T - R_P)}{K_{m5} + R_T - R_P} - \frac{k_6 R_P}{K_{m6} + R_P} \tag{3.21}$$

To specify the system, we first observe synthesis and degeneration governing the rate of change of the concentration of $X$ and MM governed phosphorylation and dephosphorylation of $Y \rightleftharpoons Y_P$ and $R \rightleftharpoons R_P$. Again, any phosphatase species are omitted by Tyson et al.. This oscillator can have two configurations. In Figure 3.12, dashed line (1) refers to an inhibition of the synthesis while (2) refers to promoted degeneration of $X$, which we chose for this model. Both are equivalent in how they affect system behaviour. $Y \rightleftharpoons Y_P$ and $R \rightleftharpoons R_P$ can be seen as buffers, delaying the signal propagation.

With $S_{crit1} < S < S_{crit2}$, $X$ is synthesized to a certain concentration at which it phosphorylates enough $Y \rightarrow Y_P$ to *turn on* $Y \rightleftharpoons Y_P$ buzzer (set $Y_P$ to a high concentration). $Y_P$ in turn phosphorylates $R \rightarrow R_P$, turning the $R \rightleftharpoons R_P$ buzzer on but delayed in time from the original stimulus by $X$. $R_P$ subsequently degenerates $X$ which causes the two buzzers to be *turned off* sequentially. Delays are dependent on the reaction constants $k_0, \ldots, k_6$ chosen. This ultimately leads to an oscillating behaviour of $X$, $Y_P$, and $R_P$. As soon as $S > S_{crit2}$, the generation of $X$ is high enough to keep the phoshporylated component $Y_P$ predominant (hence the buzzer doesn't turn off anymore). Therefore, $R_P$ remains on a constant high as well and its promotion of the degeneration of $X$ does not result in a significant concentration change of $Y_P$. Oscillation ceases.

**Characterization:**

1. As long as $S_{crit1} < S < S_{crit2}$, we can observe an oscillation. We assume for the following statements that $S$ is within these bounds.

2. Species $X$, $Y_P$, and $R_P$ (hence also $Y$ and $R$) will oscillate infinitely often between their high and low values while oscillation is possible.

3. Once $[X]$ is high, eventually $[Y_P]$ will be high, and eventually $[R_P]$ will be high, repeating infinitely often.

4. As $[R_P]$ is at a high, $[X]$ is at a low.

5. The initial values of the system do not influence the oscillatory behaviour which will eventually be reached as long as $[R] + [R_P] = [Y] + [Y_P] = 1$. The same reasoning holds for perturbations of the mixture during the simulation.

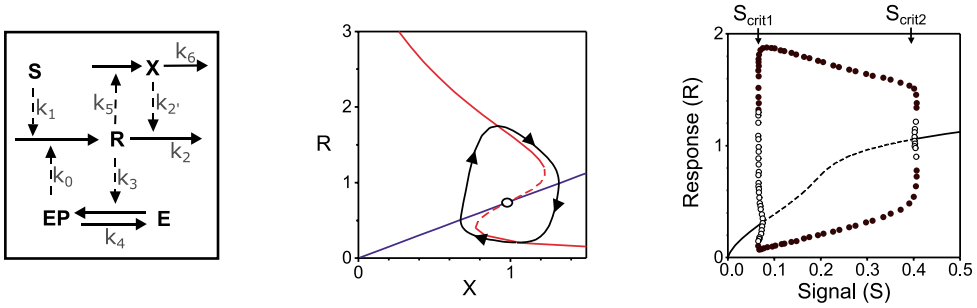



Figure 3.13 Activator-Inhibitor Oscillator. Left: wiring diagram with respective rate constants. Center: phase-plane portrait illustrating the oscillating behaviour between species $R$ and $X$ at $[S] = 2$ in black; $(X,R)$ pairs satisfying $dR/dt = 0$ in red; $(X,R)$ pairs satisfying $dX/dt = 0$ in blue. Right: signal-response diagram illustrating the Hopf-bifurcation at $S_{crit1} < S < S_{crit2}$. Modified from [82] Figure 2 (b) with permission.

### 3.4.2 Activator-Inhibitor

As presented by Tyson et al. [82], oscillators can consist of a combination of positive and negative feedback. The activator-inhibitor and the substrate-depletion oscillators (in Section 3.4.3 below) are the two varieties of oscillators that can arise from such configurations. Tyson et al. [82] presents the cyclic adenosine monophosphate (cAMP) production in slime mold, *Dictyostelium discoideum*, as a classic example of such an activator-inhibitor oscillation. Here, external cAMP binds to surface receptors promoting the synthesis of more cAMP. Concurrently, cAMP pushes the receptor into an inactive form. As the synthesis ceases, cAMP degrades, allowing the receptor to return into its active, cAMP secreting state. More recent literature shows that the oscillatory interplay of activators and inhibitors also play a significant role in cell membrane behavior and cell surface properties [90].

$$\begin{aligned} \frac{dR}{dt} &= k_o E_P(R) + k_1 S - k_2 R - k_2' X \cdot R \\ \frac{dX}{dt} &= k_5 R - k_6 X \\ E_P(R) &= G\left(k_3 R, k_4, J_3, J_4\right) \end{aligned} \tag{3.22}$$

From the wiring diagram in Figure 3.13 we deduce that $R$ is synthesized by $S$ and $E_P$ while being subject to constant degradation at rate $k_2$ which is also promoted by species $X$. As $[R]$ increases, the positive feedback loop with $E_P$ causes an autocatalytic process as found in the mutual activation system (Section 3.3.5). However, more of the inhibitor species $X$ is generated concurrently which results in a degradation of $R$ until near complete depletion. As $X$ degrades, $R$ can newly accumulate and, in the fashion of the receptor in the cAMP cycle above, return to promote its own synthesis. The oscillation described is again only supported for concentrations of the signal $S_{crit1} < S < S_{crit2}$ while the instability allowing for the oscillation is introduced by the positive feedback loop between $R$ and $E_P$. However, from the the signal response curve we see that the transition between the stable and oscillatory behaviour is rather abrupt compared to

the previous system. This is mostly due to the ultra-sensitive buzzer component $E \rightleftharpoons E_P$ which steers the positive feedback loop which in turn causes the instability of the system. The same reasoning applies to the following oscillator.

**Characterization:**

1. As long as $S_{crit1} < S < S_{crit2}$, we can observe an oscillation. We assume for the following statements that $S$ is within these bounds.

2. Once the oscillation stabilizes for a given value of $[S]$, we can imagine that a particle at approximately $([X],[R]) \approx (1,0.3)$ will eventually return to the same point after one full period of oscillation.

3. The oscillatory behaviour which will *eventually* be reached is independent of the initial values of $R$ and $X$ as well as perturbations of their concentrations during the simulation.

### 3.4.3 Substrate-Depletion

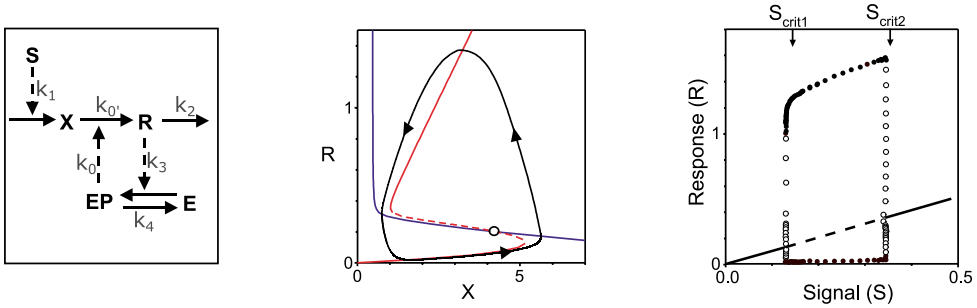



Figure 3.14 Substrate-Depletion Oscillator. Left: wiring diagram with respective rate constants. Center: phase-plane portrait illustrating the oscillating behaviour between species $R$ and $X$ at $[S] = 2$ in black; $(X,R)$ pairs satisfying $dR/dt = 0$ in red; $(X,R)$ pairs satisfying $dX/dt = 0$ in blue. Right: signal-response diagram illustrating the Hopf-bifurcation at $S_{crit1} < S < S_{crit2}$. Modified from [82] Figure 2 (c) with permission.

The last fundamental oscillatory motif relies upon a positive feedback loop between $R, E_P$, depleting the substrate $X$ from which $R$ is synthesized. This type of oscillation was the first one to be discovered, governing the periodic sugar-alcohol conversion in yeast cells [69, 68]. However, it was also found that this motif describes the underlying mechanism of MPF (maturation promoting factor) oscillations in frog egg extracts, steering embryo growth [59, 82].

$$\begin{aligned} \frac{dX}{dt} &= k_1 S - \left[k_0' + k_0 E_P(R)\right] \cdot X \\ \frac{dR}{dt} &= \left[k_0' + k_0 E_P(R)\right] \cdot X - k_2 R \\ E_P(R) &= G\left(k_3 R, k_4, J_3, J_4\right) \end{aligned} \tag{3.23}$$

By inspecting the wiring diagram, $X$ is the substrate from which $R$ is synthesized. The signal $S$ steers the synthesis of this substrate which in turn synthesizes $R$. Given $S_{crit1} < S < S_{crit2}$, $R$ promotes its own synthesis through the positive feedback loop by phosphorylating $E$ (as in mutual activation Section 3.3.5). However, as the synthesis of $R$ increases, it depletes its substrate $X$. As the concentration of $X$ does no longer sustain the synthesis of $R$ which in turn cannot phosphorylate E, the positive feedback collapses and the concentration of $X$ can recover to newly initiate this oscillation. Equations 3.23 represent the resulting ODE system with the respective reaction rates.

**Characterization:** The characterization of this system closely resembles the reasoning presented in the previous activator-inhibitor oscillator (Section 3.4.2). We here add:

1. The particle to observe is at $([X],[R]) \approx (1.5, 0)$ which too will eventually return to the same point after one full period of oscillation.

## 3.5 Summary

In this chapter we have illustrated the elementary functional motifs required to build models capable of explaining and predicting complex emergent behaviour of larger biochemical systems. This is also the justification of why we chose the systems presented by Tyson et al. [82]. We discussed seven different signal-response elements which build upon each other, extending the complexity of the resulting system behaviour. Subsequently, we built the three fundamental oscillators which can be found at the base of every observed biochemical oscillatory or periodic behaviour. For all systems we were capable of devising the governing systems of ODEs by mathematical reasoning using system analysis, following the simplifying assumptions made in the beginning of this chapter. Furthermore, we contributed formal quantitative descriptions to the systems presented, characterizing the features of the resulting system behaviours. What follows next is the automatization of the reasoning presented here through the use of the high level language provided by the bond-calculus.

# Chapter 4

# Implementation

This chapter displays the implementation of the previously discussed biochemical systems in the bond-calculus, following the syntax outlined in Section 2.3.3. By using the bond-calculus workbench (bondwb), we follow the order of increasing complexity presented before and introduce new code components as needed. Box 1 and Figure 2 in [82] provide the set of parameters used for each model. Furthermore, we include the code for all models in Appendix A. The relevant reaction rates and initial values can be found in Table D.1 in Appendix D.5. This chapter is split into elementary building blocks, oscillators, as well as the implementation in $\mathcal{LBUC}$ of the characterized systems: linear response, hyperbolic response, sigmoidal response, and perfect adaptation.

## 4.1 Elementary Building Blocks

Following the structure from the previous chapter, we first introduce the implementation of the elementary building blocks. Each system is presented with code excerpts and description to justify the design choices made. We note that for the systems mutual activation, mutual inhibition, and homeostasis we provide two implementations. One model implements the system with the GK function (see Section 2.2.3) while the other uses explicit MM kinetics instead. This choice was taken to overcome a temporary bug in the bondwb (further elaborated in Section 4.1.5). However, we only present the additional explicit implementation for the mutual activation system for the sake of brevity. The code for the explicit models as well as all others can be found in Section A.

### 4.1.1 Linear Response

Synthesis and degeneration, being the basic component at the core of any signal-response network, consist of a species for the signal $S$ as well as response $R$. All interactions are based on simple MA kinetics (as in Section 2.2.3). Hence, we first define the *species* and their respective *sites*.

```
species S = ssynR -> (S | R);
species R = degradeR -> 0;
species A = asynR -> (R | A);
```

Following the syntax, species $S$ has site `ssynR` which upon reaction results in a parallel composition of $S$ and $R$. Thus, it remains within the system and is not lost or degraded while synthesizing $R$. Species $R$ can only react to become the *null species*, hence degrade (or decay). Note that variable names for sites and locations are required to begin with a lower-case letter when using the bondwb.

An important aspect of this motif is the ambient synthesis of $R$. As the bond-calculus does not provide a generic ambient species, we chose to model this aspect through the ambient species $A$. Here, $\frac{dA}{dt} = 0$ and initial concentration $A_0 = 1$, we ensure that it synthesises $R$ by any rate constant provided in the MA kinetics. We again express the synthesis of $R$ as a parallel composition between $A$ and $R$ such that the concentration of $A$ is not affected by interaction on this site.

Given that we are using MA, which already is implemented in the system by default as `MA(.)`, we can proceed to define our *affinity network* N with *rate constants* `a, s, r`.

```
affinity network N(a, s, r) = {
    asynR at rate MA(a);
    ssynR at rate MA(s);
    degradeR at rate MA(r);
}
process P1 = [1.0] A || [1.0] S || [0] R
                with network N(0.01,1,5);
```

Finally, we define the *process* ( `P1`) with its initial mixture, the respective affinity network, and the real-valued reaction constants. We used $[k_0, k_1, k_2] = [a, s, r] = [0.01, 1, 5]$ as provided in [82] Box 1. The entire code for this model is included in Appendix A.1 as a coherent file.

### 4.1.2 Hyperbolic Response

We can make use of the previous system in order to model the MA governed phosphorylation and dephosphorylation underlying this signal-response pathway referred to as *buzzer* in [82]. We do not require ambient synthesis of any species and neither $R$ nor $R_P$ are directly synthesized by signal $S$. Hence, $S$ has a site which upon reaction does not lead to any change in concentration of itself. However, $R$ and $R_P$ can phosphorylate and dephosphorylate respectively.

```
species S = kinaseR -> S;
species R = phosphorylate -> RP;
species RP = dephosphorylate -> R;

affinity network N(k1, k2) = {
```

```
6       kinaseR || phosphorylate at rate MA(k1);
7       dephosphorylate at rate MA(k2);
8   }
```

Within the affinity network we define `signal || phosphorylate at rate MA(k1)` as the affinity pattern that site `kinaseR` can interact with site `phosphorylate` at a specific rate `MA(k1)`. The parallel composition indicates that, once the two sites are available, they can react as discussed in Section 2.3.3.

As a last step, we define a process with initial values `[1.0] S || [1.0] R || [0.0] RP` and rate constants $k_1 = k_2 = 1$. While Tyson et al. do not provide the specific initial values for species $R$ and $R_P$, they indicate that the total Molar concentration of $R$ is $R_T = 1$. Given that only an initial amount of $R = 1$ is available in our model and given that no additional $R$ and $R_P$ is introduced, it is guaranteed by the law of mass conservation, that $R_T = 1$ is satisfied.

### 4.1.3 Sigmoidal Response

The sigmoidal functional motif is very similar to the previous system, including the same species definitions. With the phosphorylation and dephosphorylation being governed by MM kinetics, we need to implement a custom kinetic law reflecting the rate law devised in Section 3.3.3.

```
1   kinetic law MM1(k,km; S,R) =
2                       (k*S*R) / (km + R);
```

We do so by using the key words `kinetic law` in the bondwb, followed by a function name and its arguments `MM1(k,km; S,R)`. In this example the first two arguments `k,km` are rate constants, passed into the the function by the corresponding affinity pattern in the affinity network. The `;` separates rate constants from the species passed as arguments. Following, species `S,R` are abstracted names of the species onto which the law is applied to in the affinity pattern. All arguments are positional arguments.

Tyson et al. did not include the phosphatase component in the dephosphorylation step (as noted in Section 3.3.3). Hence, we define an additional rate law, reflecting the formula.

```
3   kinetic law MM2(k,km; RP) =
4                       (k*RP) / (km + RP);
```

While the affinity matrix is equivalent to the one in the previous model, we substitute MA with our custom rate laws. The parallel composition `signal || phosphorylate` respects the positional order of the signal species $S$ and response species $R$ defined in the kinetic law `MM1(.)` as `kinase` and `phosphorylate` are their respective sites.

```
5   affinity network N(k1, k2, km1, km2) = {
6       kinaseR || phosphorylate at rate MM1(k1,km1);
7       dephosphorylate at rate MM2(k2,km2);
8   }
```

Finally, the process with initial values `[1.0] S || [1.0] R || [0.0] RP`, rate constants $k_1 = k_2 = 1$ and MM constants $k_{m1} = k_{m2} = 0.05$ is defined.

### 4.1.4 Perfect Adaptation

The *sniffer* functional motif exhibited by systems with perfect adaptation can be built from the components provided in Section 4.1.1. All interactions are governed by MA kinetics. Following the wiring diagram (Section 3.3.4) for this component we obtain the following species definitions.

```
1  species S = ssynR -> (S | R)
2            + ssynX -> (S | X);
3  species R = degradeR -> 0;
4  species X = degradeX -> 0
5            + stayX -> X;
```

For species $S$ and $X$ we define multiple sites separated by a choice using a mathematical *OR* operator (denoted by +). Hence, $S$ can either synthesise $R$ or synthesise $X$ without changing its own concentration. $X$ and $R$ degrade to the null species. However, species $X$ promotes the degradation of $R$ while not being affected itself. By introducing the site `stayX` we access the concentration of $X$ without changing it.

```
6  affinity network N(k1, k2, k3, k4) = {
7      ssynR at rate MA(k1);
8      ssynX at rate MA(k3);
9      degradeR || stayX at rate MA(k2);
10     degradeX at rate MA(k4);
11 }
```

The underlying mixture to the process is `[0] S || [0] X || [1.0] R`, with rate constants $k_1 = k_2 = 2$ and $k_3 = k_4 = 1$.

### 4.1.5 Mutual Activation

This system is the first of the discussed to exhibit positive feedback by combining the concept of synthesis and degeneration with the one of the phosphorylation/dephosphorylation buzzer. We refer to it as the *irreversible switch*. Species $R$ acts as the stimulus, the kinase, for the $E \rightarrow E_P$ phosphorylation. Once $E_P$ surpasses a certain threshold value, it synthesises enough $R$ to create the one-way switch intended by Tyson et al.. The authors assume that the $E \rightleftharpoons E_P$ reactions fulfill the MM requirements (see Section 2.2.3) such that the GK function can be used to express the steady-state concentration of $E_P$. Following their description, we build a model consisting only of species $S$ and $R$, while the concentration of $E_P$ is computed as presented in Section 3.3.5 through the GK function.

```
species S = ssynS -> (S | R);
species R = degradeR -> 0
          + epsynR -> (R | R);
```

The species definition for $S$ is as before. However, species $R$ now includes the site `epsynR` which allows it to synthesise itself. While biochemically this does not seem to be a valid statement, this is required to accommodate for the missing explicit species definition of $E_P$.

```
kinetic law G(f,u,v,J,K; R) =
      (2*u*K*f*R)/(v-R*u+v*J+R*u*K+((v-R*u+v*J+R*u*K)**2
                    -4*(v-R*u)*R*u*K))**0.5;
```

Here, our implementation of the GK function according to its definition (Equation 2.6 in Section 2.2.3). It includes the following extensions due to some particularities of the bondwb and the bond-calculus itself:

- Given the function parameters for this system in [82] (replicated in Equation 3.14, Section 3.3.5) we are required to set parameter $u = u \cdot R$. Thus, we included species $R$ in the kinetic law to access its concentration and multiply each occurrence of $u$ by it.
- In Equation 3.13 (Section 3.3.5), we are also required to multiply the computed concentration of $E_P$ by rate constant $k_0$. As the bond-calculus does not support arbitrary multiplication of affinity patterns, we encoded this multiplication as an additional factor $f$ in the kinetic law definition.

The affinity network consists of the synthesis of $R$ through $S$, $R$'s decay, as well as the self-synthesis of $R$ by $k_0 E(R)$ with $E(R) = G(k_3 R, k_4, J_3, J_4)$. The resulting process follows the reaction rates provided.

```
affinity network N(k0, k1, k2, k3, k4, j3, j4) = {
    ssynR at rate MA(k1);
    degradeR at rate MA(k2);
    epsynR at rate G(k0,k3,k4,j3,j4);
}
process Pi = [1.0] S || [0.0] R
                with network N(0.4,0.01,1,1,0.2,0.05,0.05);
```

A bug in the symbolic simplification algorithm of the bondwb initially caused the resulting sets of ODEs to not conform with the ones devised in [82]. Hence, we concluded that by explicitly modeling the MM kinetics governing the $E \rightleftharpoons E_P$ reaction, we can circumvent that temporary limitation.

The resulting model uses the same species definition for $S$ as before. For species $R$ we define a site for its degradation as well as a site to access its concentration. This is similar to species $X$

in the model for perfect adaptation (Section 4.1.4). Species $E$ and $E_P$ are similar to $R$ and $R_P$ in Section 4.1.3 as they follow the same phosphorylation/dephosphorylation pattern. However, species $E_P$ in this case also has a site to synthesize $R$.

```
species R = degradeR -> 0
          + stayR -> R;
species E = phosphorylateE -> EP;
species EP = dephosphorylateEP -> E
           + epsynR -> (EP | R);

affinity network N(k0, k1, k2, k3, k4, j3, j4) = {
    epsynR at rate MA(k0);
    ssynR at rate MA(k1);
    degradeR at rate MA(k2);
    stayR || phosphorylateE at rate MM1(k3,j3);
    dephosphorylateEP at rate MM2(k4,j4);
}
```

The affinity network reflects the combination of the concepts from the linear and sigmoidal response motif by first defining the synthesis and the decay of $R$, then the $E \rightleftharpoons E_P$ reaction where $R$ acts as the kinase as mentioned before. Rate laws `MM1(.)` and `MM2(.)` are re-used from Section 4.1.3. The process is equivalent to the one of the previous model, however we now also include species $E$ and $E_P$ in the mixture with initial values `[1.0] E || [0] EP`. This fulfills the requirement $E_T = E + E_P$ (as in Section 4.1.3). As Tyson et al. do not specify the initial values, we used a curated SBML entry of this component on the Biomodels database as reference [25].

The aforementioned bug had been fixed throughout the project by Wright. Hence, the phosphorylation/dephosphorylation reactions of following models are modelled using the GK function as specified by Tyson et al. in order to replicate the sets of ODEs presented. However, in Appendix A we also include for each model an explicit implementation where the phosphorylation/dephosphorylation reactions are modelled using MM kinetics and the relevant species.

### 4.1.6 Mutual Inhibition

The wiring diagram in Section 3.3.6 shows that the *reversible switch* component only differs slightly from the previous model. In this case, species $E$ is responsible for promoting the degradation of species $R$. Thus, we compute the concentration of $E$ using the GK function and degrade $R$ following simple MA kinetics with a rate constant $k_{2'}$. Furthermore, Tyson et al. re-introduced an ambient synthesising of $R$ at rate $k_0$.

```
species R = degradeR -> 0
          + stayR -> R;
```

We re-use the species definitions for $S$ and $A$ from the linear response motif in Section 4.1.1. Similar to the mutual activation model above, we set $v = v \cdot R$ and include the additional parameter $f$ in our GK function. Furthermore, to encode that $R \cdot E(R)$, where $E(R) = G(k_3, k_4 R, J_3, J_4)$ and $G$ the GK function, we multiply the numerator of our function with the concentration of $R$. This way we comply with the equations devised by Tyson et al. (replicated in Equations 3.16 and 3.15, Section 3.3.5).

```
kinetic law G(f,u,v,J,K; R) =
    2*u*K*f*R / (v*R-u+v*J*R+u*K+((v*R-u+v*J*R+u*K)**2-
        4*(v*R-u)*u*K)**0.5);

affinity network N(k0, k1, k2, k21, k3, k4, j3, j4) = {
    asynR at rate MA(k0);
    ssynR at rate MA(k1);
    degradeR at rate MA(k2);
    degradeR at rate G(k21,k3,k4,j3,j4);
}
process Pi = [0.0] S || [0.0] R || [1.0] A
    with network N(0,0.05,0.1,0.5,0.2,1,0.05,0.05);
```

### 4.1.7 Homeostasis

Finally, the last elementary building block component exerts homeostasis resulting in a semi-stationary concentration of $R$ for a broad range of signal. While the setup is similar to the one in mutual activation and inhibition, here we observe negative feedback. As $E$ counteracts the effect of $S$ by synthesizing more $R$ as $S$ degrades it. $R$ again acts as kinase for the reaction $E \rightarrow E_P$.

We re-use species definition for $S$ and $R$ from the mutual activation GK implementation (in Section 4.1.5). The GK function itself is re-used from the definition in the mutual activation system (in Section 4.1.6). However, the additional $R$ term in the numerator is omitted to comply with the function definition in [82].

The affinity network reflects the interactions from the wiring diagram. Note that $S$ now is part of the decay of $R$. Hence, the first pattern would evaluate to $[S] \cdot [R] \cdot k_2$. Moreover, we again use the fact that $R$ has a site to synthesize itself to model the synthesis by $E$ given the value resulting from the GK function.

```
affinity network N(k0, k2, k3, k4, j3, j4) = {
    sdegradeR || degradeR at rate MA(k2);
    synR at rate G(k0,k3,k4,j3,j4);
}
```

```
17
18 process Pi = [1.0] S || [0.0] R with network N(1,1,0.5,1,0.01,0.01);
```

## 4.2 Oscillators

As oscillators are at the very core of fundamental periodic processes within cells as well as organs and whole metabolisms, showing whether the bond-calculus is capable of modeling their emergent behaviour accurately is of special interest. This section is devoted to the implementation of the three core oscillators found in biochemical systems using the previously presented elementary building blocks. We will follow the description of the systems provided in Section 3.4. Note that for the activation-inhibitor and substrate-depletion oscillator we also implemented two explicit models using MM kinetics and the respective species instead of the GK function. These can be found in Appendix A.

### 4.2.1 Negative-Feedback

For this oscillator we re-use the concepts of the sigmoidal response (Section 4.1.3) and the reversible-switch motifs (Section 4.1.6). Hence, we devise a species $A$ for the ambient synthesis of $X$. Species $S$, the signal, synthesises $X$ as well. Species $X$ has a site to degrade. Both $Y$ and $R$ have sites to phosphorylate while their activated form $Y_P$ and $R_P$ have sites to turn back into their initial forms.

```
1  species A = asynX -> (A | X);
2  species S = ssynX -> (S | X);
3  species X = stayX -> X
4            + degradeX -> 0;
5  species Y = phosphorylateY -> YP;
6  species YP = stayYP -> YP
7             + dephosphorylateYP -> Y;
8  species R = phosphorylateR -> RP;
9  species RP = stayRP -> RP
10            + dephosphorylateRP -> R;
```

In the affinity network, we incorporate the signal and ambient synthesis of $X$ as well as its synthesis by using MA. The $Y \rightleftharpoons Y_P$ reaction is steered by the concentration of $X$ using the MM kinetics from the sigmoidal response motif (Section 4.1.3). We do so equally for the $R \rightleftharpoons R_P$ reaction which is steered by the concentration of $Y_P$. Finally, the decay of $X$ is additionally driven by the concentration of $R_P$ with MA kinetics.

```
11 affinity network N(k0,  k1,  k2, k21, k3, k4, k5,
12                    k6, km3, km4, km5, km6) = {
```

```
13      asynX at rate MA(k0);
14      ssynX at rate MA(k1);
15      degradeX at rate MA(k2);
16      stayX || phosphorylateY at rate MM1(k3, km3);
17      dephosphorylateYP at rate MM2(k4, km4);
18      stayYP || phosphorylateR at rate MM1(k5, km5);
19      dephosphorylateRP at rate MM2(k6, km6);
20      stayRP || degradeX at rate MA(k21);
21  }
```

Following the initial concentrations documented in the corresponding Biomodel [27] as well as the values in [82], we create a process with mixture `[1.0] A || [2.0] S || [0] X || [1.0] Y || [0.0] YP || [1.0] R || [0.0] RP`. Here, we have $S \in (S_{crit1}, S_{crit2})$ such that oscillation occurs. Furthermore, we note that initially species $Y$ and $R$ are predominant in their inactive (non-phosphorylated) form.

### 4.2.2 Activator-Inhibitor

We refer to our description of this oscillator in Section 3.4.2 and build the system by re-using components from the mutual activation (Section 4.1.5) and the perfect adaptation motifs (Section C.4).

The signal species definition as well as the GK function follow from the model for mutual activation. The species definition for $X$ follows from perfect adaption. However, for species $R$ we need to define sites for degradation, the synthesis of $X$, and the self-synthesis used with the GK function in the mutual activation motif.

```
1  species R = rsynR -> (R|R)
2            + degradeR -> 0
3            + rsynX -> (R | X);
```

The affinity network includes affinity patterns for the synthesis of $R$ by $S$ and $E_P$. The latter is modelled using the GK function. Furthermore, the degradation of $X$ and $R$. The latter is both by a constant value as well as steered by $[X]$.

```
4  affinity network N(k0, k1, k2,
5                     k21, k3, k4,
6                      k5, k6, j3, j4) = {
7      ssynR at rate MA(k1);
8      degradeR at rate MA(k2);
9      degradeX at rate MA(k6);
10     rsynX at rate MA(k5);
11     stayX || degradeR at rate MA(k21);
```

```
12      rsynR at rate G(k0,k3,k4,j3,j4);
13  }
14  process Pi = [0.2] S || [0.0] X || [0.0] R
15      with network N(4,1,1,1,1,1,0.1,0.075,0.3,0.3);
```

### 4.2.3 Substrate-Depletion

For the last system modelled, we observe two sequential synthesis and degeneration components coupled with an auto-catalytic reaction $R \rightleftharpoons E_P$ promoting the synthesis of $R$ and depletion of $X$. Following the species definitions where $S$ synthesises $X$, $X$ turns into $R$, and $R$ degenerates to the null species.

```
1  species S = ssynX -> (S | X);
2  species R = stayR -> R
3            + degradeR -> 0;
4  species X = xturnR -> R;
5
6  kinetic law G(f,u,v,J,K; R,X) =
7          (f * 2*u*R*K*X) / (v-u*R+v*J+u*R*K+((v-u*R+v*J+u*R*K)**2-
8              4*(v - u*R) * u*R *K)**0.5);
```

Although we re-use the GK function from the activator-inhibitor system discussed before, we include the multiplication of $[X]$ in the nominator to replicate Equation 3.23 (Section 3.4.3). The affinity network reflects the MA synthesis of $X$ as well as its MA governed conversion $X \to R$ and the MA degeneration of $R$. Ultimately, we include the conversion governed by the concentration of $E_P$ modelled by the GK function.

```
9  affinity network N(k01, k0, k1, k2, k3, k4, j3, j4) = {
10     ssynX at rate MA(k1);
11     degradeR at rate MA(k2);
12     xturnR at rate MA(k01);
13     stayR || xturnR at rate G(k0,k3,k4,j3,j4);
14 }
15 process Pi = [0.2] S || [0.0] X
16             || [0.0] R with network
17             N(0.01,0.4,1,1,1,0.3,0.05,0.05);
```

## 4.3 $\mathcal{LBUC}$

In this section we present formal $\mathcal{LBUC}$ formulae encoding the characterizations provided for the systems of linear response, hyperbolic response, sigmoidal response, and perfect adaptation. These characterizations were elaborated in Sections 3.3.1, 3.3.2, 3.3.3 and 3.4.1 respectively.

### 4.3.1 Linear Response

For this motif, the steady-state response follows a simple linear curve for different values of the signal. The following properties were identified:

1. As the concentration of species $S$ increases, we observe an increase in $[R_{SS}]$ proportional to $\frac{k_0}{k_2} + \frac{k_1}{k_2}[S] = 0.002 + 0.2 \cdot [S]$.
2. $[R_{SS}]$ is independent of $R$'s initial concentration.

For property 1, consider the steady-state response derived in Section 3.3.1. Rearranging its analytic solution we obtain $R_{SS} - \frac{k_0}{k_2} - \frac{k_1}{k_2} S = 0$. Thus, it describes the proportionality between the steady-state of $R$ and $S$.

$$P_1 = F_{[0,t_1]} \left( G_{[0,t_2]} \left( \varepsilon > \left( R - \frac{k_0}{k_2} - \frac{k_1}{k_2} S \right)^2 \right) \right) \tag{4.1}$$

Proposition $P_1$ in Equation 4.1 formally encodes the first property. It reads, *eventually* $[R]$ will *globally* (unchanged within the defined time interval) be *close to* the expected analytic value $[R_{SS}]$. As we work in a setting of uncertainty, we encode *"... close to ..."* as a difference between the value of $[R]$ during simulation and its analytic steady-state $[R_{SS}]$ (blue in Equation 4.1). Thus, we define a small value $\varepsilon$ to quantify an admissible deviation between the two values. Note that in $\mathcal{LBUC}$ it is not possible to *increase* a specific value. Hence, we assume that, if $P_1$ is satisfied on the whole range of $[S]$, we consider it to hold also for increases of $S$.

We choose values $t_1 = 4$, $t_2 = 10$, and $\varepsilon = 0.1$. $t_1$ provides a good margin to the time for stabilization we devised in Table D.1 in Appendix D. In this case, $t_2$ describes the time of observation for which we want the system to satisfy the property (in case of the global quantifier). However, $t_2$ and $\varepsilon$ were chosen arbitrarily. These values assure that our property reflects the behaviour we intend to verify.

Note that we use the squared value of the difference for two reasons. One, we can verify whether the difference, negative or positive, is within the bound set by $\varepsilon$. Two, as opposed to taking the *absolute value*, we remain in the polynomial domain. This prevents Flow* to engage in computationally expensive re-computations of a flowpipe as symbolic subdivision is available.

For property 2, note that the analytic result for $R_{SS}$ shows no dependency on its own initial concentration $R_0$. In order to verify this, we set $[R_0]$ of the process to a real-valued interval $[0, 1000]$. We also set $[S_0] = [0, 3]$ to reflect the interval of signal values observed by Tyson et al.. We then re-evaluate whether the system still satisfies $P_1$.

Finally, we want to verify that:

3. This independence holds for any perturbation of $[R]$ introduced during the simulation of the system.

4. The linear relation holds for any perturbation of $[S]$ during the simulation.

To verify that property 3 and 4 above holds for any perturbation of $[R]$, we make use of $\mathcal{LBUC}$'s context operator.

$$C_{LR} \models \Pi_{R,pert.} \rhd P_1 \iff C_{LR} \parallel \Pi_{R,pert.} \models P_1 \tag{4.2}$$

$$\Pi_{R,pert.} \triangleq [0,1000]R \parallel [0,3]S \tag{4.3}$$

Let $C_{LR}$ be our current process and $P_1$ the proposition we want $C_{LR}$ to satisfy. In Equation 4.2 we introduce a new process $\Pi_{R,pert.}$ (Equation 4.3) as a context to the satisfaction relation. It quantifies perturbations in $[R]$ and $[S]$ during the simulation. If this composition evaluates to true, we showed that also this statement holds.

### 4.3.2 Hyperbolic Response

Recall the first properties of this system's behaviour:

1. As we increase $[S]$, the steady-state concentration of the phosphorylated response will globally stay below an asymptotic maximum $\lim_{[S]\to\infty} dR_{P_{SS}}([S]) = [R_T]$. In our case, $[R_T] = 1$ and in Figure 3.4 denoted as $r_3$.

2. The asymptotic behaviour holds for any perturbation of $[S]$ during the simulation.

To encode the first property, we make use of the steady-state solution for $R_P$ and note that the asymptote is $R_T = R + R_P$ ($R_T = 1$ in our case). By using $R_T$ instead of a fixed value, we can verify that this behaviour holds even with other initial values for $R$ and $R_P$ or whether perturbations of these are introduced.

$$P_2 = G_{[0,t_1]}\,(R + R_P > R_P) \tag{4.4}$$

$$C_{HR} \models \Pi_{S,pert.} \rhd P_2 \tag{4.5}$$

$$\Pi_{S,pert.} \triangleq [0,10]S \tag{4.6}$$

$P_2$ in Equation 4.4 reflects that, *at all times (globally)* (throughout $[0,t_1]$) $[R_P]$ is below $[R_T]$. We evaluate the formula over an interval of $[S_0] = [0,10]$, such that we cover the range in $[S]$ presented by Tyson et al.. The initial concentrations for $R$ and $R_P$ are $[R_0] = 1$ and $[R_{P,0}] = 0$.

For the second statement, we define a process $C_{HR}$ (Equation 4.5) used for verification. Moreover, we define context process $\Pi_{S,pert.}$ (Equation 4.6) introducing a perturbation of $[S] =$

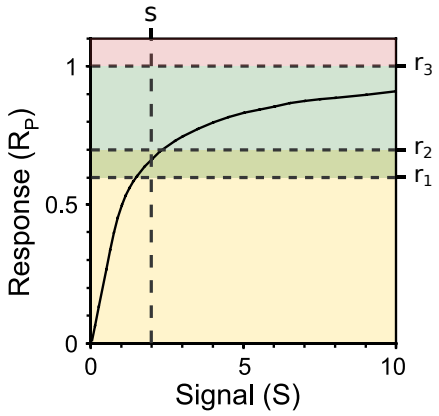



Figure 4.1 Hyperbolic signal-response curve with subdivisions for quantitative characterization. $r_1 = 0.6$, $r = 0.7$, $r = 3 = 1$, $s = 2$. **Colors:** Low values of $R_P$ in orange, high values of $R_P$ in *green*, not reachable values in *red*. Modified from [82] Figure 1 (b) with permission.

$[0,10]$ to the process. To verify that $P_2$ holds with the applied perturbations, we use $t_1 = 10$ (arbitrarily chosen).

Recall the subdivision in Figure 4.1 (originally in Section 3.3.2) of the signal-response diagram for this motif. Based on it, we devised the following steady-state properties between signal and response:

3. As we increase $[S]$ from low to high values, the increase in $[R_{P_{SS}}]$ is first rapid and then gradually slows down. In Figure 4.1, $[S] = s = 2$ acts as a threshold between low and high values of $[S]$. Bounds $r_1 = 0.6$ and $r_2 = 0.7$ reflect our bounds for high and low values in $[R_{P_{SS}}]$. Hence, as $[S] < s$, $[R_{P_{SS}}] < r_2$ and as $[S] \geq s$, $r_1 < [R_{P_{SS}}] < r_3$. The bounds for $[R_{P_{SS}}]$ have a margin of 0.1 to relax the boundaries for this property and therewith allow $\mathcal{LBUC}$ to compute results.

We hence identified the following formula:

$$P_3 = G_{[0,t_2]}\left(P_2 \wedge F_{[0,t_1]}\left(G_{[0,t_2]}\left((S < s \implies R_P < r_2) \wedge (S > s \implies R_P > r_1)\right)\right)\right) \tag{4.7}$$

In $P_3$, we first evaluate whether $[R_P]$ is *globally* lower than the asymptote $[R_T]$. Note, that the global operator is *distributive* over the *and* ($\wedge$) and we make use of the identity $G_I(G_{I'}(\psi)) \equiv G_{max[I,I']}(\psi)$. Subsequently, we say that *globally* (for $[0,t_2]$), *eventually* (within $[0,t_1]$), *globally* (for $[0,t_2]$), the enclosed proposition holds, allowing the system to reach the steady-state within $[0,t_1]$. Note that the clock for the second global statement starts at *relative* time 0 while the *global clock* is already at some point in $[0,t_1]$ from the *eventually* quantifier. The enclosed proposition reads, when $[S] < s$, then $[R_P] < r_2$ and respectively if $[S] > s$, then $[R_P] > r_1$. Hence, describing the bounds in the description. We evaluate whether this system satisfies $P_3$ using process $C_{HR}$ with initial values $[S_0] = [0,10]$,$[R] = 1$, and $[R_P] = 0$.

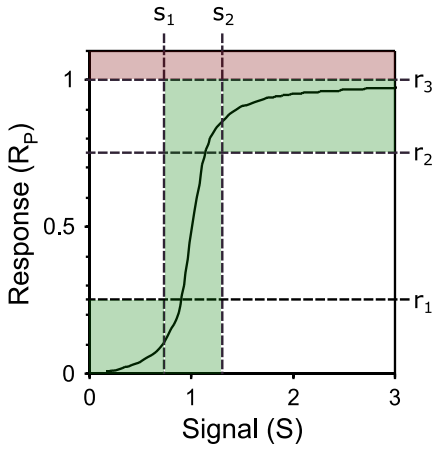



Figure 4.2 Sigmoidal signal-response curve subdivided into regions for quantitative evaluation. Sigmoidal signal-response curve with subdivisions for quantitative characterization. $r_1 = 0.25$, $r = 0.75$, $r = 3 = 1$, $s_1 = 0.7$, and $s_1 = 1.3$. **Colors:** Values of $[R_P]$ in *green* describe ultra-sensitive behaviour, not reachable values are in *red*. Modified from [82] Figure 1 (c) with permission.

We use $t_1 = 2$, allowing enough time for the system to stabilize, and $t_2 = 10$ (arbitrarily chosen). Furthermore, we note that for our implementation, as the bondwb does not provide with an implication operator, we make use of the identity $\phi \implies \psi \iff \neg\phi \vee \psi$.

### 4.3.3 Sigmoidal Response

For the sigmoidal response motif, also referred to as ultra-sensitive switch by Goldbeter and Koshland [33] or *buzzer* by Tyson et al., we initially characterized the following properties:

1. Similar to the hyperbolic motif, $[R_{P_{SS}}]$ will globally stay below the asymptotic limit of $r_3 = [R_T] = 1$ as $[S] \to \infty$.

2. The asymptotic behaviour holds for any perturbation of $[S]$ during the simulation.

Here we can re-use the equal properties presented in the previous Section 4.3.2 as the statements are equivalent. Hence, we use $P_2$ (Equation 4.5) and the context process $\Pi_{S,pert.}$ (Equation 4.6) to verify that this holds. However, we set the initial concentration interval as well as the perturbation interval of $\Pi_{S,pert.}$ to $[S] = [0,3]$. This reflects the interval observed by Tyson et al. [82]. Finally, we require $t_1 = 20$ as this system requires a longer stabilization time (see Table D.1 in Appendix D.5).

3. As $[S] \leq s_1$, $[R_{P_{SS}}] < r_1$. Once $s_1 < [S] < s_2$, $[R_{P_{SS}}]$ increases rapidly, reaching $[R_{P_{SS}}] = r_2$ quickly. Thus, within a small change of $[S]$, $[R_{P_{SS}}]$ rises quickly. We encode this by having a small difference of $[S]$ between a low value and a high value of $[R_{P_{SS}}]$, we choose $s_2 - s_1 = 0.5$. As $s_2 < [S]$, $r_2 < [R_{P_{SS}}] < r_3$.

For this property, we use the subdivision of the signal-response diagram presented in Figure 4.2 (originally in Section 3.3.3). Here, we have signal thresholds $s_1 = 0.7$ and $s_2 = 1.3$ as well as $[R_{P_{SS}}]$ threshold $r_1 = 0.25$,$r_2 = 0.75$, and $r_3 = [R_T] = 1$. These value were chosen arbitrarily to determine the bounds within which we consider the switch to be ultra-sensitive. We note that there is no formal mathematical definition of ultra-sensitivity in literature.

$$P_4 = G_{[0,t_2]}\left(P_2 \wedge F_{[0,t_1]}\left(G_{[0,t_2]}\left((S \leq s_1 \implies R_P < r_1) \wedge (S \geq s_2 \implies R_P > r_2)\right)\right)\right) \tag{4.8}$$

Similarly to the approach in the previous section, $P_4$ encodes the property we want to verify. Here, $P_2$ from Section 4.3.2, guarantees that $[R_P]$ is *globally* below the asymptote. Verifying this over $[S_0] = [0,3]$, guarantees that the ultra-sensitivity is satisfied. We use $t_1 = 20$ (also for $P_2$) to allow the system to stabilize. Furthermore, we use $t_2 = 10$ (arbitrarily chosen).

4. The ultra-sensitivity of the switch is only retained when $K_{m1}, K_{m2} \ll 1$.

For this interesting characteristic, we used a slightly altered implementation of the sigmoidal response system which can be found in Appendix A.3.1. The model encodes the rate constants $K_{m1}$ and $K_{m2}$ as species *KM1* and *KM2* with sites `stayKM1 -> KM1` and `stayKM2 -> KM2`. This guarantees that, once added to the mixture, $\frac{dK_{m1}}{dt} = \frac{dK_{m2}}{dt} = 0$. The kinetic laws `MM1(.)` and `MM2(.)` are altered such that they access the concentration of *KM1* and *KM2* and use these as the MM constants. Using this model, we can verify for which values the ultra-sensitive behaviour holds. We used an interval on both *KM1* and *KM2* of $[KM1_0] = [KM2_0] = [0.05, 0.3]$ to explore the range of values under which $P_4$ is satisfied.

### 4.3.4 Perfect Adaptation

Recall the first two properties of this system's behaviour defined in Section 3.3.4 as well as the illustration in Figure 4.3:

1. As we introduce a step-wise increase of $[S]$, we observe:
   (a) $[R]$ will rise as a direct response to the increased synthesis. It will exceed a threshold and eventually return to a steady-state value ($R_{SS} = \frac{k_1 k_4}{k_2 k_3} = 1$ in our case) as soon as $[X]$ adapted to the new $[S]$.
   (b) $[X]$ will adapt to $[S]$ and their concentrations will eventually be equal. This is due to $k_3 = k_4 = 1$ in our case. Hence, $[X] = [S]$ eventually.

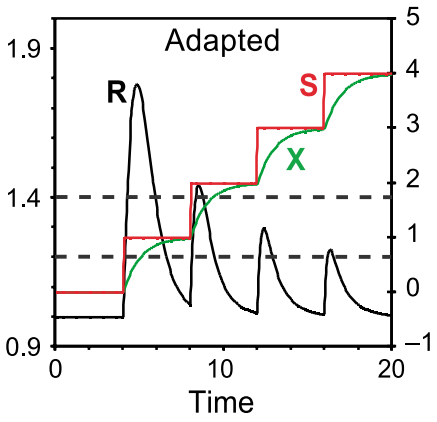



Figure 4.3 Perfect adaption temporal progression diagram illustrating reduction of the intensity of spiking of $[R]$ as $[S]$ is higher prior step-wise increase. Modified from [82] Figure 1 (c) with permission.

We first encode property 1.a. Thus, we devise $P_5$ in Equation 4.9.

$$P_5 = G_{[0,t_2]}\left(\Pi_{S,pert.} \rhd F_{[0,t_1]}\left((R > r) \wedge F_{[0,t_2]}\left(G_{[0,t_2]}\left(\left(\varepsilon_2 > R - \frac{k_1 k_4}{k_2 k_3}\right)\right)\right)\right)\right) \tag{4.9}$$

$$C_{PA} \models P_5 \tag{4.10}$$

$$\Pi_{S,pert.} \triangleq [1,10]S \tag{4.11}$$

Here, we apply a context $\Pi_{S,pert.}$ (Equation 4.11) to a proposition within $P_5$ (Equation 4.9) to quantify that, no matter what perturbation is added to $[S]$, we observe *globally* the expected behaviour. Furthermore, if process $C_{PA}$ (Equation 4.10) has a real-valued interval over $[R_0]$, we are able to verify that $P_5$ holds for any initial value of $[R_0]$ as well.

$P_5$ in Equation 4.9 reads, *globally* (throughout $[0,t_2]$), as we introduce a context $\Pi_{S,pert.}$, we will *eventually* (within $[0,t_1]$) observe that $[R]$ is higher than a certain threshold $r$ *and* that *eventually* (within $[0,t_2]$ of that having occurred), we will reach a state where $[R]$ *globally* (throughout $[0,t_2]$ after we first observed $\varepsilon_2 > R - \frac{k_1 k_4}{k_2 k_3}$) returns to its steady-state value $R_{SS}$ (blue in Equation 4.9). In essence, this formally summarizes the property stated above. We chose $t_1 = 1$ to guarantee the *shortness* of the initial spike and $t_2 = 10$ to allow the system to stabilize. Furthermore, $\varepsilon_1 = \varepsilon_2 = 0.1$, arbitrarily chosen.

$$P_6 = (X < S) \wedge F_{[0,t_2]}\left(\varepsilon_1 > (X - S)^2\right) \tag{4.12}$$

$$C_{PA} \models \Pi_{S,pert.} \rhd P_6 \tag{4.13}$$

For property 1.b we devise $P_6$ (Equation 4.11). The first part reads, $[X]$ is smaller than $[S]$. In Metric Interval Temporal Logic (similarly in LTL), an atomic proposition is satisfied if the first state in the sequence satisfies the property. As we increase $[S]$ step-wise, the first state of the hybrid-automaton will be that $[X] < [S]$. Thus, the first part is satisfied. The second part of $P_6$

reads, *eventually* within $[t_1, t_2]$, the squared difference will be smaller than a small value $\varepsilon_1$. Here we again make use of the concept presented in Section 4.3.1 and use the squared difference. By evaluating this proposition as in Equation 4.13, we guarantee that this hold for any perturbation $[S] \in [1, 10]$.
The next property we characterized is:

2. The higher $[S]$ is *prior* the step-wise increase, the lower will be the spike in $[R]$ when an increase in $[S]$ is induced (observe reducing peaks in Figure 4.3).

$$P_5' = \Pi_{S,step} \rhd F_{[0,t_1]}\left((R > r) \wedge F_{[0,t_2]}\left(G_{[0,t_2]}\left(\varepsilon_2 > R - \frac{k_1 k_4}{k_2 k_3}\right)\right)\right) \tag{4.14}$$

$$C_{PA} \models P_5' \tag{4.15}$$

$$\Pi_{S,step} \triangleq [1]S \tag{4.16}$$

As $\mathcal{LBUC}$ does not provide a functionality to recursively add contexts, we are limited to manual checking of property 2. Hence, we use a relaxed version of $P_5$, $P_5'$ (Equation 4.14), as well as a real-valued constant for the concentration of $[S]$ in $\Pi_{S,pert.}$ (Equation 4.16).

We verify that the current process $C_{PA} \models P_5'$ by first setting $[S_0] = 0, [X_0] = 0$, and $[R_0] = 1$. As the concentration of $S$ is perturbed, we note the result of the evaluation. In the next step, we update the initial values $[S_0] = 1$ and $[X_0] = 1$. Thus, we re-run the verification and observe whether the proposition is satisfied. This is done recursively for a finite amount of increments. Recall threshold value $r$ which describes the value of $[R]$ which has to be reached at least by $R$ such that we consider it a spike. In Figure 4.3, the dotted lines show hoe this value has to be lowered as $[S]$ increases gradually. Thus, we devise multiple $r$ values in order to quantify the reductinon in spike magnitude of $[R]$.
The last property we characterised is that:

3. The steady-state response $R_{SS}$ is independent of its initial value $R_0$. Hence, any perturbation of $[R]$ during a simulation, results in the same value for $R_{SS}$.

Here we devise $P_7$ (Equation 4.17) and process $\Pi_{R,pert.}$ (Equation 4.18) to quantify the perturbation of $R$. Process $C_{PA}$ in Equation 4.19 represents the current process on which we verify the satisfaction of $P_7$.

$$P_7 = F_{[0,t_2]}\left(G_{[0,t_2]}\left(\varepsilon_2 > R - \frac{k_1 k_4}{k_2 k_3}\right)\right) \tag{4.17}$$

$$C_{PA} \models \Pi_{R,pert.} \rhd P_7 \tag{4.18}$$

$$\Pi_{R,pert.} \triangleq [0, 100]R \tag{4.19}$$

# Chapter 5

# Evaluation

Provided the illustrations in Figure 5.1 and 5.2 for the elementary building blocks and the oscillators respectively, this chapter focuses on the evaluation of the generated behaviour. First, we conduct a qualitative analysis of the results by comparing our results in shape and values with the results presented by Tyson et al. [82]. As Wang [84] modelled a subset of these components in the c$\pi$, we also relate our results to the ones provided in his work. This allows us to subsequently discuss our first research question.

Due to an initial bug in the bondwb, we modelled certain systems using two different assumptions, namely the MM kinetic law and the GK function. While the ultimate behaviour should be equivalent, we show the differences between these in our implementation and hence also address our second research question.

Finally, we employ a novel approach of formal verification using $\mathcal{LBUC}$ and provide a quantitative analysis of a sub-set of the models. We present our results on the characterizations of the system behaviours devised in Chapter 3 and encoded as $\mathcal{LBUC}$ formulae in Chapter 4. Thus, elaborating our third and last research question.

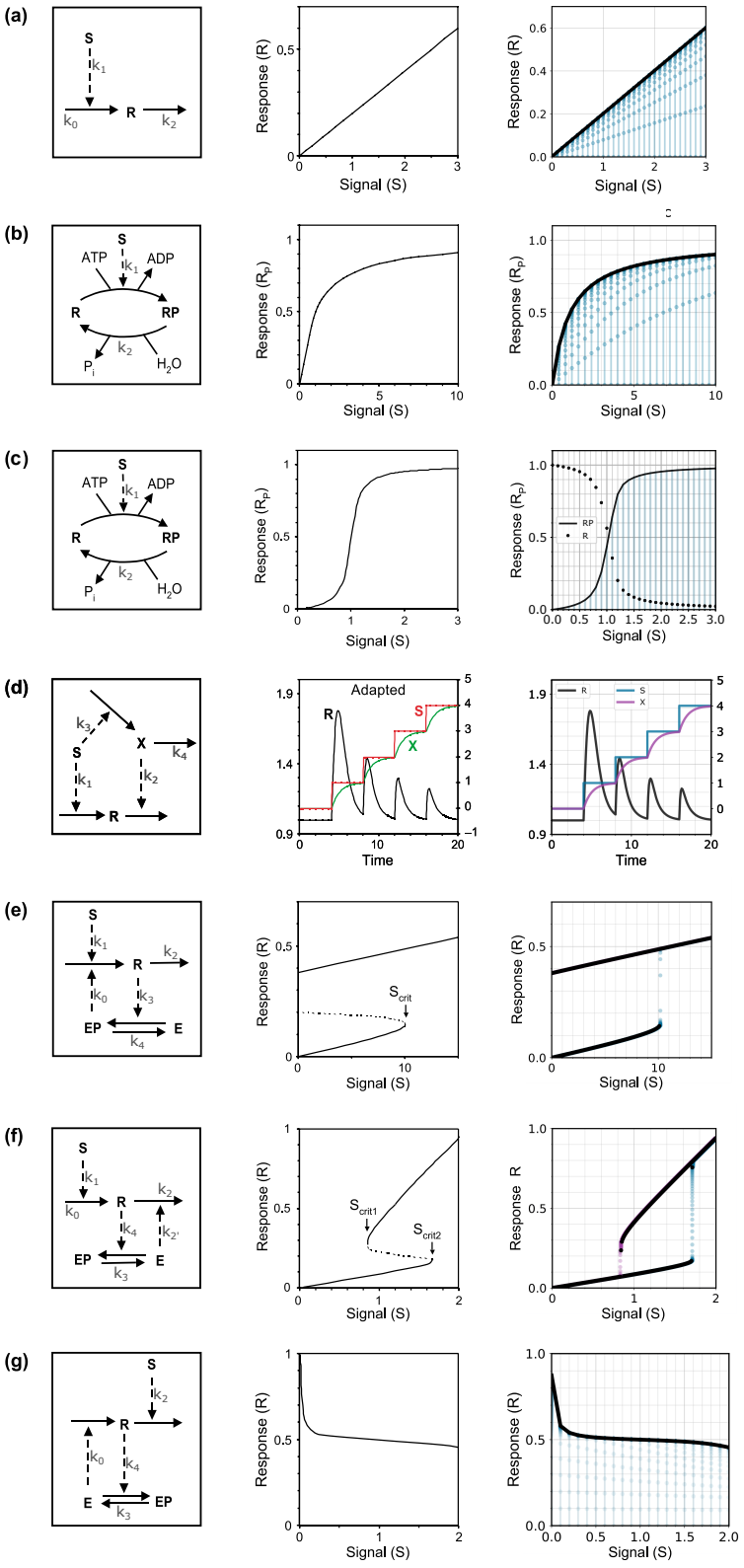



Figure 5.1 Elementary Building Blocks.
**Systems:**
(a) linear response,
(b) hyperbolic response,
(c) sigmoidal response,
(d) perfect adaptation,
(e) mutual activation,
(f) mutual inhibition,
(g) homeostasis.
**Columns: Left:** wiring diagrams with respective rate laws; **Center:** signal-response or (in case of system d) time progression diagrams of systems provided in [82]; **Right:** diagrams obtained from our implementation in the bond-calculus. Figures in left and center column modified from [82] Figure 1 with permission.

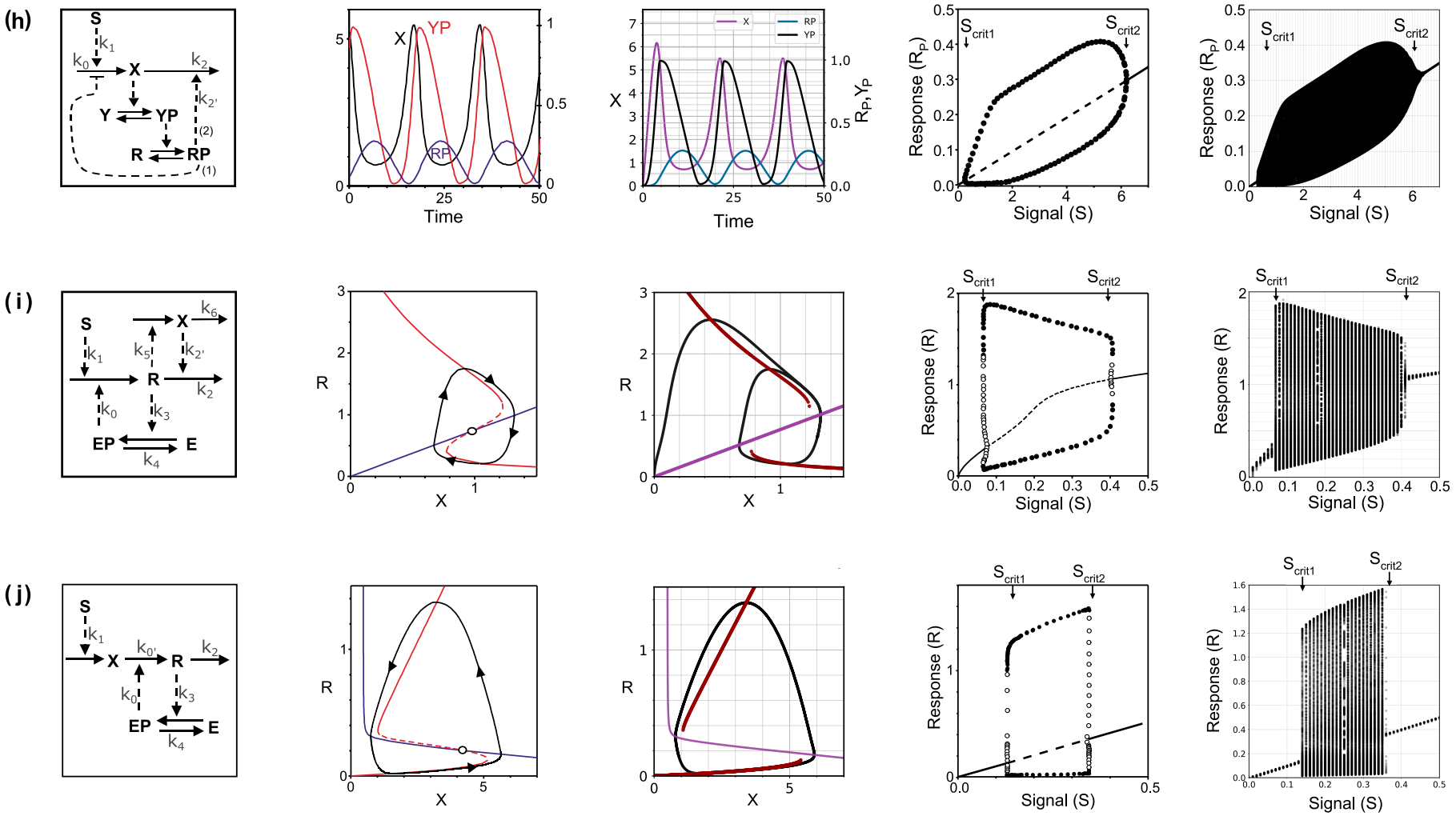



Figure 5.2 Oscillators. Systems: (h) negative-feedback oscillator, (i) activator-inhibitor oscillator, (j) substrate-depletion oscillator. **Columns from left to right:** wiring diagrams with respective rate laws; time progression diagrams for system (h) and phase-plane portraits for systems (i,j) from [82]; respective diagrams obtained from our implementation in the bond-calculus; signal-response Hopf-bifurcation diagrams from [82]; respective signal-response diagrams obtained from our implementation. Relevant figures modified from [82] with permission.

## 5.1 Evaluation Methods

First, we justify our choice of qualitative and quantitative evaluation metrics.

System behaviour is fully determined by the set of governing ODEs and the initial values. Hence, we present the generated sets of ODEs for all implemented systems in Appendix B. However, Tyson et al. strongly focus on the visual aspect of the steady-state signal-response diagrams of the systems. Thus, we opt to first conduct a visual qualitative analysis. We compare the steady-state curves generated by our models with the ones presented by Tyson et al. and Wang. Furthermore, we compare the results from our explicit models, using MM kinetics for the underlying reactions, and abstracted models where the concentration of certain species is calculated through the GK function (as in Section 4.1.5). Note that more detailed illustrations of Figures 5.1 and 5.2 can be found in Appendix C.

While the qualitative analysis already provides insight, we cannot argue about system behaviour under uncertainty. To verify that the systems modelled indeed behave as expected, we chose to conduct a further quantitative evaluation using $\mathcal{LBUC}$. Thus, we present the evaluation for the systems with linear, hyperbolic, and sigmoidal response, as well as perfect adaptation.

## 5.2 Qualitative Analysis

The comparisons in Figure 5.1 and 5.2 show that we were able to accurately reproduce all system behaviours presented in [82]. As Wang [84] noted in his evaluation of the elementary building block components in c$\pi$, Tyson et al. chose different illustrations for the different systems.

For the linear (a), hyperbolic (b), sigmoidal (c) motifs as well as mutual activation (e), mutual inhibition (f), and homeostasis (g), Tyson et al. use steady-state signal-response diagrams. However, for perfect adaptation (d) and the negative-feedback oscillator (h), the temporal progression of species's concentrations are depicted. For the remaining oscillators, phase-plane portraits at a specific concentration of the signal are shown to indicate the oscillating concentrations. We therefore employed different approaches to produce equivalent illustrations with our generated ODE systems.

For systems (a,b,c,g), we generated the signal-response diagram by conducting a simulation of the system with a specific initial concentration of the signal species. For each time step, with step-size 0.1, we record the current state of the response concentration with a blue dot. The simulation is run until the (quasi) steady-state, or dynamic-equilibrium (as discussed in Section 2.2.5) of the system is reached which we then highlighted in black. For each system's signal concentration range, we repeat this operation with a 0.1 increment in signal concentration at each simulation. Finally, the black dots representing the steady-state concentrations are then connected to form the steady-state curves relevant for comparison with the curves in [82].

However, systems (e,f) exhibit hysteresis, hence the current state depends on the system's history. For these systems we repeat the process as before, however instead of returning to the

initial values and $t = 0$, the simulation is sequential. Thus, once the steady-state is reached after time $t$ for a certain concentration of $[S]$, we retrieve all species's concentrations. Subsequently, we increase the concentration of $S$ in this set of concentrations by 0.01 and proceed with the simulation. Once the highest value for $[S]$ is reached, we repeat the process but after each stabilization deduct 0.01 from $[S]$. The black lines represent each steady-state value of $[R]$ for a given $[S]$.

For system (d) we conduct each simulation for $t = 4$, retrieve the values for the different species, increment $[S]$ by 1 and use this new set of concentrations as the input for a new run of the same duration. This is repeated 4 times, resulting in the temporal progression in Figure 5.1 (d). For the temporal progression plot in system (h) we simply simulated the system for $t = 50$ without further altering of the concentrations during the run.

The phase-plane diagrams of systems (i,j), as presented by Tyson et al., involve three components,$(X,R)$ pairs satisfying $\frac{dR}{dt} = 0$ in *red*, those satisfying $\frac{dX}{dt} = 0$ in *purple*, and the temporal oscillation of$(X,R)$ in *black*. For the first two curves we chose the same approach as for systems (e,f) above. For the third curve, we ran our simulation for $t = 200$ and projected the obtained values into the $(R,X)$ plane, removing the temporal information of the progression. An example of this for system (i) is in Figure 5.3.

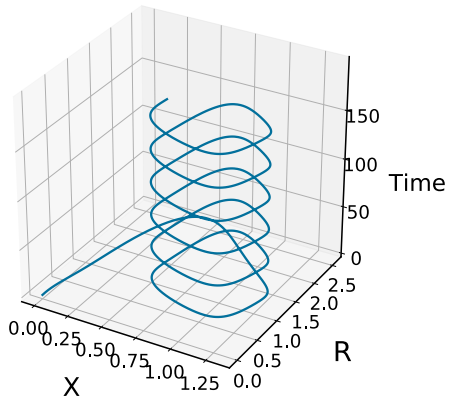



Figure 5.3 Temporal progression of oscillatory behaviour in activator-inhibitor oscillator in 3D.

Note that for all three oscillators (h,i,j) we were able to replicate the signal-response diagrams, illustrating the critical bounds on $[S]$, $S_{crit,1}$ and $S_{crit,1}$, which allow oscillation to occur. For these diagrams we simulated the system at a certain value for $[S]$ and record the values of response species concentration over time. By then gradually increasing $[S]$, we can project the obtained values for the response species onto the 2-dimensional $(R,S)$ plane.

### 5.2.1 Elementary Building Blocks

**Linear response:** This is the simplest model and uses pure MA kinetics. For this and the following models, the same set of ODEs as presented by Tyson et al. are extracted from our bond-calculus implementations. Hence, visually our results coincide with the ones presented by the authors. The same result was achieved by Wang, showing that the system was successfully implemented in the c$\pi$. We note that we used the same approach as Wang and modelled the ambient generation through an ambient, constant species $A$.

**Hyperbolic response:** Here, we were able to employ the MA kinetics directly and obtain the same steady-state signal-response curve as the authors. Wang was also able to achieve the same

results. However, due to a bug in the CPi-IDE he was required to explicitly bind the response and signal to a complex at a high rate which then turned into the phosphorylated response at a high rate.

**Sigmoidal response:** Here, Wang correctly identified that the assumptions made by Tyson et al. about the composition of $R_T$ are incorrect. Tyson et al. argue that $R_T = R + R_P$ while the MM kinetics derived by Briggs and Haldane [10] clearly state, that $R_T = R + R_P + RS$ where $RS$ is the intermediate substrate-enzyme complex. Furthermore, we note in Section 3.3.3, that the authors incorrectly formulated the MM kinetics for dephosphorylation ($R_P \rightarrow R$) and omitted the existence of a phosphatase steering the reaction.

Wang had to account for this incorrect assumption on $R_T$ when deriving the MA interactions underlying the MM kinetics as $c\pi$ only allows for MA kinetics. Hence, he had to scale the initial concentrations for $S$ as well as the rate constants to explicitly model the association and dissociation of the intermediary complex. Finally, he was able to achieve the same results as Tyson et al.. In our work, thanks to the flexibility provided by the bond-calculus, we were able to formulate the rate laws to reflect the ones used by Tyson et al.. Thus, we obtain the same results as the authors without further changes. Not that in Figure 5.1 (c) we also include the concentration of species $R$ (the unphosphorylated response) to show the adherence to $[R_T] = [R] + [R_P]$.

However, as the original publication is from 2003, Tyson acknowledges that their implementation of MM was erroneous as well as some of their assumptions on the systems.

**Perfect Adaptation:** As this system represents a combination of two, inter-dependent synthesis and degeneration components, pure MA kinetics are used. Both our implementation and the one presented by Wang reflect the results obtained by Tyson et al..

**Mutual Activation:** During the implementation of this system, we identified an issue with the symbolic simplification algorithm of the bondwb. This bug caused the GK kinetic law definition to not be correctly simplified, resulting in an erroneous set of ODEs. We therefore opted to model the system explicitly using the MM kinetics underlying the GK function while the issue persisted. Once this bug had been solved, we re-implemented the models using the GK function as intended by Tyson et al..

This additional work gave us the opportunity to raise our second research question. We ask whether there are observable differences between modeling explicit interactions based on MM or using strong assumptions and encoding the same interactions with GK. Recall that GK computes the steady-state concentration of a species in a phosphorylation/dephosphorylation couple given the concentration of kinase (enzyme promoting $E \rightarrow E_P$).

Both implementation proved to provide the same bifurcation, hence irreversible switch motif presented in [82]. However, the time for stabilization is $t = 35$ for the GK implementation and $t = 125$ for the MM implementation.

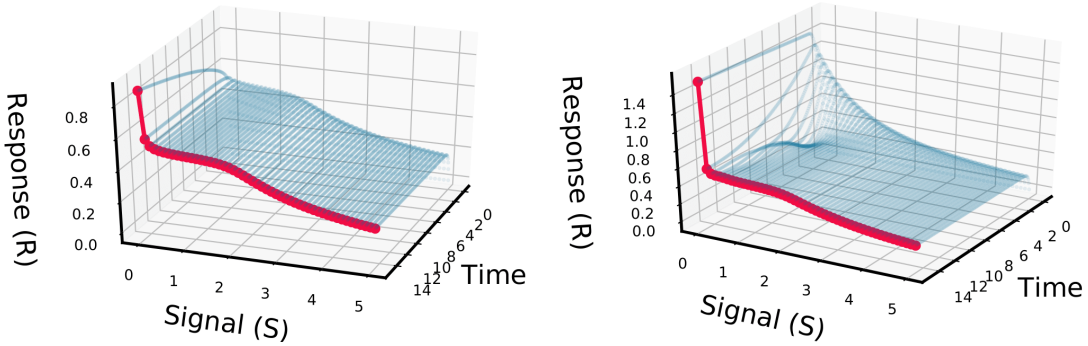



Figure 5.4 Stabilization behaviour of homeostasis in 3D. Left: system implemented with the GK function. Right: system implemented with explicit MM kinetics

For this motif, Wang again encountered difficulties to convert Tyson et al.'s assumptions to the underlying MA kinetics. Through repeated scaling of concentrations and rate constants, he was able to obtain a qualitative replication of the behaviour in [82]. However, his results did not match numerically to the ones presented in the paper. Low concentrations of $R$ were significantly smaller than the ones obtained by our implementation or Tyson et al.. We argue, that through the use of more accurate assumptions such as the tQSSA, Wang would have been able to replicate the intended behaviour through the use of MA kinetics.

**Mutual Inhibition:** The same rationale as for mutual activation applies here. We provide two systems, one modelled explicitly with the MM kinetics governing $E \leftrightharpoons E_P$ and one using the GK function to express the steady-state concentration of $E$. Both systems lead to the same steady-state behaviour as presented by Tyson et al.. However, the time required for the two systems to stabilize is different ($t \approx 80$ for GK, $t \approx 100$ for MM). We again note that Wang's results again required scaling of initial concentrations and rate constants, leading to a conformity in shape but not in value to our results or the ones in [82].

**Homeostasis:** Similar to the systems above, we provide two implementations of this system, one using MM kinetics and the other the GK function abstraction to compute the steady-state concentration of $E$. We obtain the same result as Tyson et al.. However, Wang again achieves an implementation in the c$\pi$ reflecting the behaviour but not close to the numerical values presented by Tyson et al..

We again observed different stabilization times of the MM and the GK system. In Figure 5.4 above, we depict the 3-dimensional stabilization behaviour for each run of the simulation using a different value for $[S]$. While in the MM system on the right we observe local fluctuations and a stabilized response after 14 seconds, the system implementing the GK function on the left stabilizes after 1 second. While this behaviour can be favorable when considering the steady-state results, it can lead to skewed results as we note in the evaluation of the activator-inhibitor and the substrate-depletion oscillator.

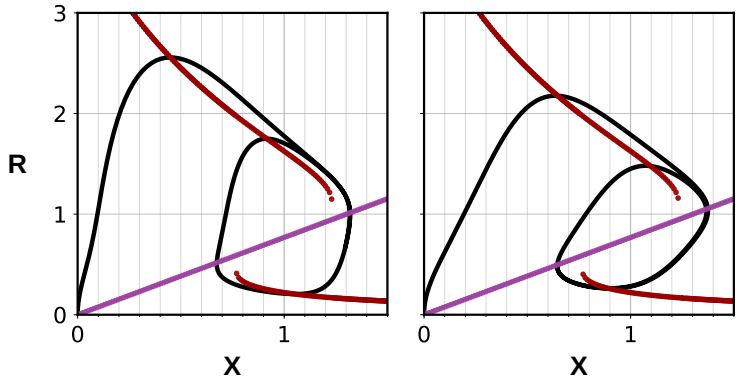



Figure 5.5 Phase-plane portrait for activator-inhibitor oscillator. **Left:** system implemented using GK function. **Right:** system implemented using explicit MM kinetics. **Colors:**$(X,R)$ pairs satisfying $\frac{dR}{dt} = 0$ in *red*; $(X,R)$ pairs satisfying $\frac{dX}{dt} = 0$ in *purple*; temporal oscillation of $(X,R)$ in *black*.

### 5.2.2 Oscillators

Wang was not able to reproduce the following systems in [84] due to the limitation presented by the c$\pi$. Hence, we do not present a comparison between his work and ours.

**Negative-feedback:** As can be noted in Figure 5.2 (h), we were not able to correctly reproduce the exact oscillation phases of the different species as in [82]. We argue that this is due to the fact that the authors did not disclose the initial concentrations for the species. Furthermore, we suspect that although the authors declare the start of the plot to be at $t = 0$, they used a systems which already stabilized its oscillation. Our system's oscillation starts with initial concentrations for $[X],[Y],[Y_P] = 0$ and require one period to align to the correct phase. Nevertheless, we are able to replicate the same behaviour and sets of ODEs. Furthermore, our signal-response diagram in the rightmost column of Figure 5.2 shows close resemblance with the one provided by Tyson et al..

**Activator-inhibitor:** Here, we again provide two models, one explicitly modeling the MM kinetics governing the $E \leftrightharpoons E_P$ reaction and one using the GK function for modeling the steady-state concentration of $E_P$. As for this and the next oscillator, Tyson et al. chose to use a phase-plane portrait as well as the signal-response graph to illustrate the system's behaviour. We have presented the same illustrations using our implementation. Both the explicit and the model using GK replicate the same behaviour as intended by the authors. Note that our circular motion of $(X,R)$ (center column in Figure 5.2 (i)) first starts at $[X] = [R] = 0$ and then, progressively reaches the oscillatory behaviour.

However, we observe a clear divide between the MM and the GK approximation in Figure 5.5. While the model on the left, using the GK function, exhibits the same behaviour as presented by Tyson et al., the MM implementation shows a decisive skew. This is due to the time required to

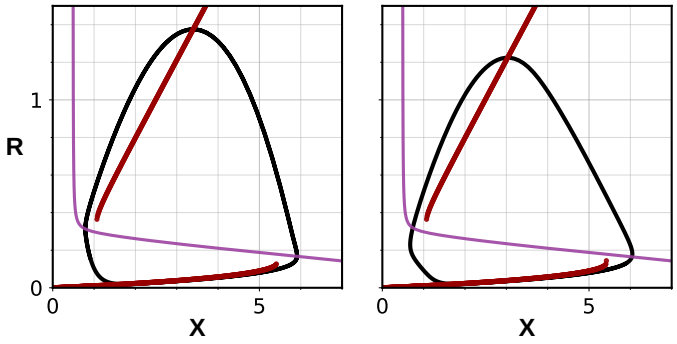



Figure 5.6 Phase-plane portrait for substrate-depletion oscillator. **Left:** system implemented using GK function. **Right:** system implemented using explicit MM kinetics. **Colors:**$(X,R)$ pairs satisfying $\frac{dR}{dt} = 0$ in *red*; $(X,R)$ pairs satisfying $\frac{dX}{dt} = 0$ in *purple*; temporal oscillation of $(X,R)$ in *black*.

stabilize the concentrations of the different species when using MM kinetics. Thus, oscillatory feedback components already act upon other components before reaching their steady-state.

Finally, note that the signal-response diagram in Figure 5.2 (rightmost column) replicates the one presented by Tyson et al.. Both $S_{crit,1}$ and $S_{crit,2}$ agree with the ones devised by the authors. However, in Appendix C.16 we provide the same signal-response diagram obtained from our MM governed implementation. There, the values are not equivalent and the minima and maxima values are as well skewed as expected.

**Substrate-depletion:** The last system modelled shows a similar behaviour as the activator-inhibitor oscillator above. In Figure 5.6 we relate the GK governed model (left) to the MM governed model (right). Again, the values of $[X]$ and $[R]$ reached are damped due to the lack of stabilization of the species concentrations.
However, we note that in Figure 5.2 (j) our signal-response diagram shows the same values for $S_{crit,1}$ and $S_{crit,2}$ as presented by Tyson et al. but the maxima values reached exceed the ones illustrated by the authors. This is in both, the signal-response diagram obtained from the GK as well as the MM governed systems (for MM governed system see Figure C.20 in Appendix C). While the cause is unknown, we assume that the error lies within the value reported by Tyson et al. for a specific rate law, allowing our system to reach higher values of $R$.

## 5.3 Quantitative Analysis

In this section we lay the basis for discussion of our last research question on whether it is possible to verify and guarantee bond-calculus model behaviour through the use of $\mathcal{LBUC}$. Using the bondwb and the underlying Flow* framework (Section 2.4.3), we attempt to evaluate the different system properties devised in Section 4.3. Our implementation showed the limits of the current implementation and we conclude by discussing alternative approaches taken.

Note that Tyson et al. provide both the set of governing ODEs as well as a description of the behavioural characteristics of the systems presented in [82]. We use this information to verify that our results obtained by our quantitative analysis coincide with the ones expected.

### 5.3.1 Procedures and Nomenclature in the bondwb

We here define the constructs and concepts used within the bondwb which we use to verify system properties.

$\mathcal{LBUC}$ **uncertainty:** Can be introduced either for the initial values of a mixture (1) or using a context process in the form of a perturbation of the mixture during simulation (2):

1. A vector of real-valued intervals expressing the uncertain range of each species' initial value in the mixture is defined. Thus, when conducting integration of the ODEs, Flow* symbolically quantifies an uncertainty of the initial state (at $t = 0$) of the system and propagates it through a flowpipe.
2. Using a context embedded within a logical expression, we quantify uncertain perturbations of a system, again described via real-valued intervals, occurring at some instant $t$ instant $t$ during the simulation of the system.

**Reach object:** Computes the flowpipe representing the set of initial values and its progression through the space of (prime) species concentrations and time. This is equivalent to integrating the ODE system given a set of initial states, either defined by real constants or intervals, over time.

**Signal:** Represent the evaluation of a proposition propagated through a flowpipe at every time-step. Signals can evaluate at a given timepoint to three different states, *true*, *false*, and *uncertain*. True and false here are definite. When a flowpipe fits into a region of species concentration space for which the proposition holds, it evaluates to true and if not, to false. When a flowpipe covers the region but isn't fully contained in it, the signal evaluates to *uncertain*.

**Signal for System:** Uses a reach object to compute the flowpipe for a specific proposition we want to verify. It then uses the flowpipe to compute a signal, representing the truth of a proposition at each point in time, according to the flowpipe.

**Context Signal for System:** Or *context signal* works similarly to signal for system. However, it allows us to improve the quality (precision) of the signal by spatially subdividing the initial set, computing a signal for each sub-domain, and combining them appropriately.

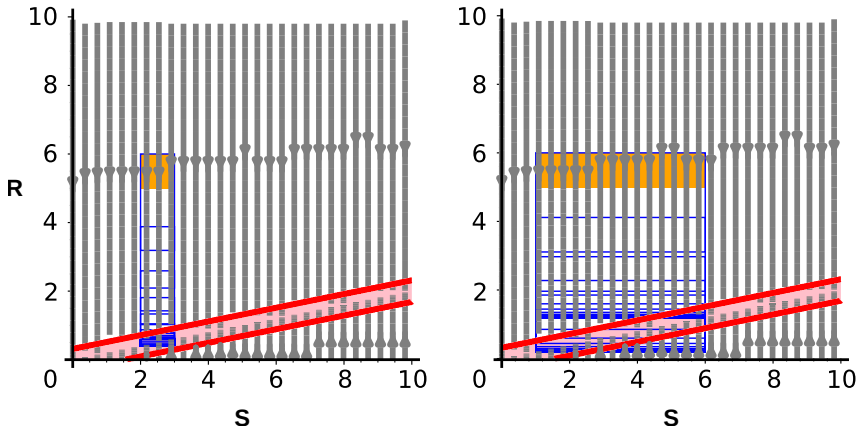



Figure 5.7 Linear response. 2-dimensional vector field for species $S$ and $R$. In *red* the region described by $\mathcal{LBUC}$ proposition $P_1'$. **Left:** *orange* square quantifying initial values $[5,6]R \parallel [2,3]S$. *Blue* squares visualizing flowpipe. **Right:** *orange* square quantifying initial values $[5,6]R \parallel [1,6]S$. *Blue* squares visualizing flowpipe.

When changing the spatial signal refinement, we need to re-compute our signal. This can be done in two ways. Either, by using *symbolic subdivision* of the initial set to consider different initial conditions. Or, by *re-computing* the entire flowpipe. The former allows us to make use of the symbolic flowpipe representation provided by Flow* and reduce computation significantly.

### 5.3.2 Linear Response

Recall from Section 4.1.1 that we used species $S$ for the signal, $R$ for the response, and $A$ as the ambient species to constantly generate a minimal quantity of $R$. However, in Appendix A.1.1 we include a model where only two species are stated explicitly, encoding the ambient generation in a custom MA rate law used by the signal species. This model results in the same ODEs and allows us to visualize the vector field in 2-dimensions as shown in Figure 5.7. We immediately note that all vectors point towards a straight line. Recall that the properties we want to verify are:

1. As the concentration of species $S$ increases, we observe an increase in $[R_{SS}]$ proportional to $\frac{k_0}{k_2} + \frac{k_1}{k_2}[S] = 0.002 + 0.2 \cdot [S]$.
2. $[R_{SS}]$ is independent of $R$'s initial concentration.
3. This independence holds for any perturbation of $[R]$ introduced during the simulation of the system.
4. The linear relation holds for any perturbation of $[S]$ during the simulation.

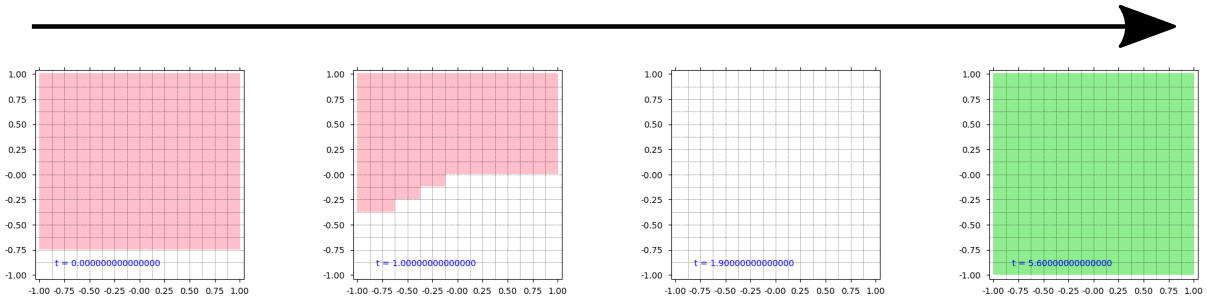


Figure 5.8 Temporal progression of refined signal (refinement 4) for $P_1'$ (Equation 5.2) evaluated on the linear response system. X-axis describes the subdivision of the interval on $[S_0]$. Y-axis describes the subdivision of the interval on $[R_0]$. **Colors:** green for *true*, red for *false*, white for *uncertain*. Time-steps from left to right: $t = 0, t = 1, t = 1.9, t = 5.6$

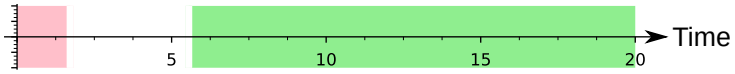



Figure 5.9 1-dimensional temporal progression of refined signal (refinement 4) for $P_1'$ (Equation 5.2) evaluated on the linear response system. X-axis describes the subdivision of the interval on $[S_0]$.**Colors:** green for *true*, red for *false*, white for *uncertain*.

For properties 1 and 2 we devised proposition $P_1$ in Equation 4.1 (Section 4.3.1) which is replicated in Equation 5.1 below.

$$P_1 = F_{[0,t_1]}\left(G_{[0,t_2]}\left(\varepsilon > \left(R - \frac{k_0}{k_2} - \frac{k_1}{k_2}S\right)^2\right)\right) \tag{5.1}$$

$$P_1' = G_{[0,t_2]}\left(\varepsilon > \left(R - \frac{k_0}{k_2} - \frac{k_1}{k_2}S\right)^2\right) \tag{5.2}$$

By setting $[R_0] = [0, 1000]$, we can verify that property 1 indeed holds for any initial concentration of $R$ within that interval. By setting $[S_0] = [0, 10]$, we encode the *increase* of $[S]$ in property 1. Note that throughout this example we will use $P_1'$ (Equation 5.2) such that we can visualize the temporal progression. $P_1$ would simply evaluate to a single value due to the encapsulating *eventually*.

Note that in Figure 5.7 we visualize two different sets of initial values and their flowpipe progression in the vector field. On the left of the figure, interval $[S_0] = [2, 3]$, allows the flowpipe to fit within the bounds set by the term $\varepsilon > \left(R - \frac{k_0}{k_2} - \frac{k_1}{k_2}S\right)^2$ which encodes the linear behaviour. On the right however, the resulting flowpipe doesn't fit into the region. Here we use *signal for system* for the signal computation. Hence, if we use this approach, we would be required to manually refine the initial interval values for the flowpipe to fit into the bounds of the linear behaviour. Otherwise, the signal for the larger region $[S_0] = [1, 6]$ would evaluate to *uncertain*.

In order to overcome this limitation and avoid manual subdivision, we use a refined signal using the *context signal for system* object. Thus, we spatially subdivide the concentration interval of both $R_0$ and $S_0$ by 4. Figure 5.8 illustrates this progression over time from left to right. We observe that more and more subdivided signals reach a state of uncertainty until all signals are

uncertain up until time-step $t = 5.5$. The signals,as they further progress in the flowpipe, then abruptly reach the state of satisfaction of $P_1'$.

Figure 5.9 (bottom) represents the same progression on a 1-dimensional plot which we refer to as *italo-vexillar diagram*. Both plots illustrate, that property 1 and 2, when expressed with $P_1'$, are satisfied within a certain time. Thus, $P_1$ too evaluates to *true*.

For properties 3 and 4, we are not able to use *context signals* as applying a context on a proposition is not yet supported by the bondwb. However, we can manually subdivide the regions as previously mentioned and use *signal for system* for each manually subdivided $[S_0]$. By choosing $[S_0] = [n \cdot 0.01, (n \cdot 0.01) + 0.01] \mid n \in [0, 1, .., 300]$ we can define a small interval of width 0.1 in concentration $S$ and evaluate $P_1$ with an introduced context process $\Pi_{R,pert.}$ (Equations 4.3 and 4.2 in Section 4.3.1). This context process then has an initial mixture of $[0, 1000]R \parallel [n \cdot 0.01, (n \cdot 0.01) + 0.01]S \mid n \in [0, 1, .., 300]$. Hence, we quantify uncertainty on both, the initial value of $S$ and an uncertainty on a perturbation of it during the simulation. As both intervals are "back-to-back", we can safely reason about the property evaluation for the full interval $[S_0] = [0, 3]$ given the evaluations of our manual sub-intervals. We found, that each segment evaluates to *true*, thus propositions 2 and 3 are satisfied.

Concluding, we showed that we could implement all three identified properties of this system's behaviour in $\mathcal{LBUC}$ and that these are satisfied by the system. We identified technical limitations of

### 5.3.3 Hyperbolic Response

Recall the properties for this system:

1. As we increase $[S]$, the steady-state concentration of the phosphorylated response will globally stay below an asymptotic maximum $\lim_{[S]\to\infty} dR_{P_{SS}}([S]) = [R_T]$. In our case, $[R_T] = 1$ and in Figure 3.4 denoted as $r_3$.

2. The asymptotic behaviour holds for any perturbation of $[S]$ during the simulation.

3. As we increase $[S]$ from low to high values, the increase in $[R_{P_{SS}}]$ is first rapid and then gradually slows down. In Figure 4.1, $[S] = s = 2$ acts as a threshold between low and high values of $[S]$. Bounds $r_1 = 0.6$ and $r_2 = 0.7$ reflect our bounds for high and low values in $[R_{P_{SS}}]$. Hence, as $[S] < s$, $[R_{P_{SS}}] < r_2$ and as $[S] \geq s$, $r_1 < [R_{P_{SS}}] < r_3$. The bounds for $[R_{P_{SS}}]$ have a margin of 0.1 to relax the boundaries for this property and therewith allow $\mathcal{LBUC}$ to compute results.

As this system is composed of 3 species, we cannot visually represent the vector field with the current implementation of the bondwb. While properties 1 and 2 concern the asymptotic behaviour of the response $[R_{P_{SS}}]$, property 3 encodes the hyperbolic shape of the signal-response illustrated in Figure 5.10.

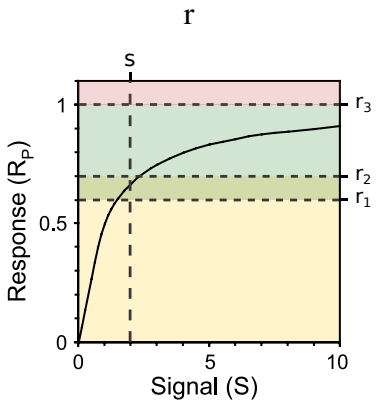



Figure 5.10 Hyperbolic signal-response curve with subdivisions for quantitative characterization. $r_1 = 0.6$, $r = 0.7$, $r = 3 = 1$, $s = 2$. **Colors:** Low values of $R_P$ in orange, high values of $R_P$ in *green*, not reachable values in *red*. Modified from [82] Figure 1 (b) with permission.

$$P_2 = G_{[0,t_1]}\,(R + R_P > R_P) \tag{5.3}$$

$$P_3 = G_{[0,t_2]}\,\big(P_2 \wedge F_{[0,t_1]}\,\big(G_{[0,t_2]}\,((S < s \implies R_P < r_2) \wedge (S > s \implies R_P > r_1))\big)\big) \tag{5.4}$$

For property 1 and 3, we verify $P_2$ and $P_3$ respectively replicated in Equation 5.3 and 5.4 (originally in Section 4.3.2). We evaluate these over $[S_0] = [0, 10]$ by using a manually subdivided *signal for system* as above using $[S_0] = [n \cdot 0.1, (n \cdot 0.1) + 0.1] \mid n \in [0, 1, .., 100]$. To better illustrate the underlying evaluations of $P_2$ and $P_3$, we additionally evaluated the propositions $P_{i..iv}$ in Equations 5.5.

$$P_i = R_P < r_2 \qquad P_{ii} = R_P > r_1 \qquad P_{iii} = S < s \qquad P_{iv} = S > s \tag{5.5}$$

In Figure 5.11, we illustrate the values that each sub-interval takes as we progress in signal concentration. Note that at $[S] \approx s$, $\mathcal{LBUC}$ is *uncertain* whether $[S] < s$ or $[S] > s$ which is as expected. Here, one interval ends at $[S] = s$ while the next begins with it. Thus, we cannot reason about $P_{iii}$ and $P_{iv}$ in sub-ranges including $S = s$. Note how the validation of each proposition $P_{i..iv}$ reflects what is encoded in $P_2$ and $P_3$. Ultimately, both $P_2$ and $P_3$ are satisfied over the whole range of $S$. Hence, properties 1 and 3 are satisfied.

Property 2 expresses that the asymptotic behaviour of the signal-response curve holds for any perturbation of $[S]$ during the simulation. Hence, we use the same subdivided *signal for system* as above to evaluate $P_2$ with introduced context $\Pi_{S,pert.}$ (Equation 5.6). Again, we face the limitation that $[S_0] = [0, 10]$ is too large for Flow* to compute. Thus, we use $[S_0] = [0, 0.1]$ as shown in Equation 5.7. While this does not fully represent what we initially encoded, we reason that this comes sufficiently close. For every interval of $[S]$ of size $[0.1]$ we observe whether an introduced context of size $[S] = [0, 0.1]$ added to the current signal still satisfies the proposition. This ultimately also evaluates to *true*.

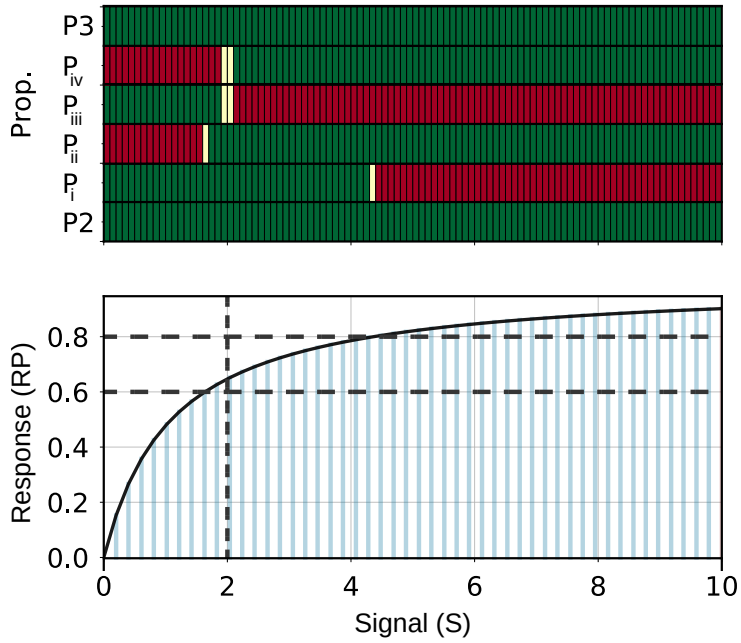



Figure 5.11 Evaluation of propositions $P_2$, $P_3$, $P_i..P_{iv}$ over the whole range of observed signal $[S] = [0, 10]$ and evaluated with a resolution of 0.1 concentration step-size. Red for *false*, green for *true*, white for *uncertain*

Finally, properties 1 and 3 were shown to hold by using *signal for system* on properties $P_2$ and $P_3$. However, we were required to manually subdivide the concentration space of $S_0$ as otherwise the flowpipes did not converge. For property 2 we had to notably reduce the perturbation interval of $[S]$ from $[0, 10]$ to $[0, 0.1]$. Thus, while we can question the validity of our approach, also this property seems to hold for the conducted experiment.

$$C_{HR} \models \Pi_{S,pert.} \rhd P_2 \tag{5.6}$$

$$\Pi_{S,pert.} \triangleq [0, 0.1]S \tag{5.7}$$

### 5.3.4 Sigmoidal Response

Recall the properties for this system:

1. Similar to the hyperbolic motif, $[R_{P_{SS}}]$ will globally stay below the asymptotic limit of $r_3 = [R_T] = 1$ as $[S] \to \infty$.

2. The asymptotic behaviour holds for any perturbation of $[S]$ during the simulation.

3. As $[S] \leq s_1$, $[R_{P_{SS}}] < r_1$. Once $s_1 < [S] < s_2$, $[R_{P_{SS}}]$ increases rapidly, reaching $[R_{P_{SS}}] = r_2$ quickly. Thus, within a small change of $[S]$, $[R_{P_{SS}}]$ rises quickly. We encode this by having a small difference of $[S]$ between a low value and a high value of $[R_{P_{SS}}]$, we choose $s_2 - s_1 = 0.5$. As $s_2 < [S]$, $r_2 < [R_{P_{SS}}] < r_3$.

4. The ultra-sensitivity of the switch is only retained when $K_{m1}, K_{m2} \ll 1$.

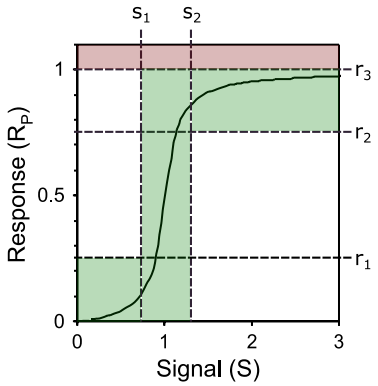



Figure 5.12 Sigmoidal signal-response curve subdivided into regions for quantitative evaluation. Sigmoidal signal-response curve with subdivisions for quantitative characterization. $r_1 = 0.25$, $r = 0.75$, $r = 3 = 1$, $s_1 = 0.7$, and $s_1 = 1.3$. **Colors:** Values of $[R_P]$ in *green* describe ultra-sensitive behaviour, not reachable values are in *red*. Modified from [82] Figure 1 (c) with permission.

Properties 1 and 2 are equal to the above in Section 5.3.3 and by using the same approach we were able to verify that these evaluate to *true* as well for this system. Property 3 describes the sigmoidal shape of the signal-response curve as depicted in Figure 5.12 and encoded in formula $P_4$ replicated in Equation 5.8. By using the same approach as in Section 5.3.3 and $[S_0] = [n \cdot 0.1, (n \cdot 0.1) + 0.1] \mid n \in [0, 1, .., 30]$, we verify that $P_4$ evaluates to true over the range of signal.

$$P_4 = G_{[0,t_2]} \left( P_2 \wedge F_{[0,t_1]} \left( G_{[0,t_2]} \left( (S \leq s_1 \implies R_P < r_1) \wedge (S \geq s_2 \implies R_P > r_2) \right) \right) \right) \tag{5.8}$$

The possibly most interesting is property 4, stating that the ultra-sensitivity of the switch is only retained when $K_{m1}, K_{m2} \ll 1$. To verify this property, we use formula $P_4$ as before but encode $K_{m1}$ and $K_{m2}$ as separate species. We described the implementation itself in Section 4.3.2. As the two critical points are when $[S] = s_1$ and $[S] = s_2$, we chose to evaluate $P_4$ at these concentrations of $S$ rather than separate intervals. This approach was chosen as bondwb exhibits an underlying memory leakage causing our random access memory to be completely filled after a certain amount of evaluations using *signal for system* objects.

Furthermore, as the current implementation of $\mathcal{LBUC}$ fails when real-valued intervals are provided as initial values for $K_{m1}$ and $K_{m2}$, we observed discrete values $[K_{m1}] = [K_{m1}] = [0.05, 0.1, 0.15, 0.2, 0.25, 0.3]$ and evaluated $P_4$ at the two distinct signal concentrations $S = s_1$ and $S = s_2$. The results of this parameter exploration are visualized in Figure 5.13.

We clearly see that for value pairs $(K_{m1}, K_{m2}) = [(0.05, 0.05), (0.05, 0.10), (0.10, 0.10), (0.10, 0.15)]$ the property holds. We illustrate as an example $(K_{m1}, K_{m2}) = (0.05, 0.30)$ on the left of the graphic, showing that values $(K_{m1}, K_{m2}) = (0.05, 0.3)$ result in a system which does not fulfill the property of ultra-sensitivity.

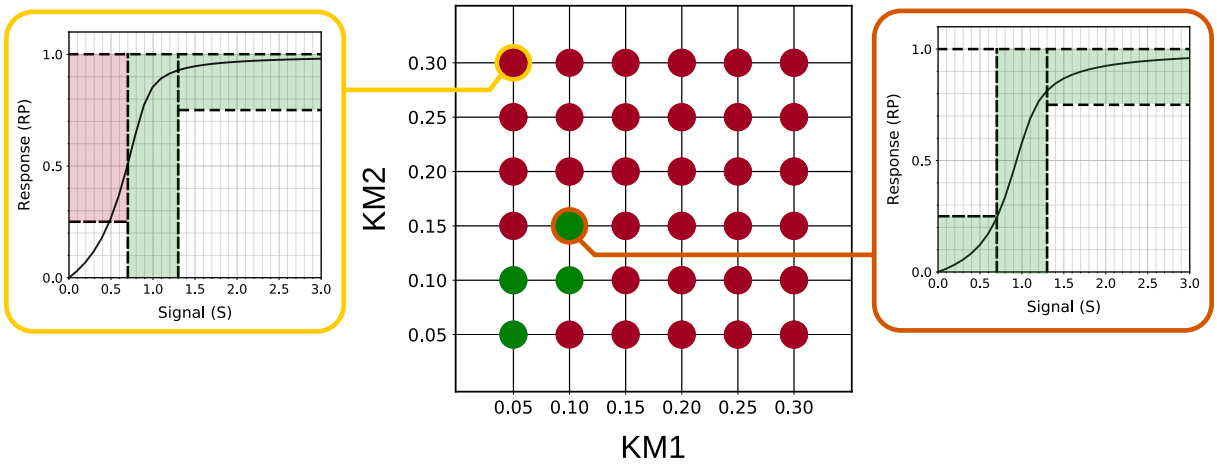



Figure 5.13 **Center:** Discrete value exploration for MM rate constants $K_{m1}$ and $K_{m2}$. Colors: dots in green indicates that ultra-sensitivity is maintained; dots in red indicate that ultra-sensitivity is not maintained. **Left:** signal-response curve at $(K_{m1}, K_{m2}) = (0.05, 0.3)$; red area indicating the violation of the ultra-sensitivity. **Right:** signal-response curve at $(K_{m1}, K_{m2}) = (0.1, 0.15)$; all green areas indicating the satisfaction of the ultra-sensitivity

Concluding, we show that properties 1,2, and 3 are satisfied by the same means of evaluation shown before in Section 5.3.3 and all three evaluate to *true*. For property 4 we showed that indeed, ultra-sensitivity is only given within the bounds of $K_{m1}, K_{m2} \ll 1$ when using the GK function.

### 5.3.5 Perfect Adaptation

Recall the properties for this system:

1. As we introduce a step-wise increase of $[S]$, we observe:
   (a) $[R]$ will rise as a direct response to the increased synthesis. It will exceed a threshold and eventually return to a steady-state value ($R_{SS} = \frac{k_1 k_4}{k_2 k_3} = 1$ in our case) as soon as $[X]$ adapted to the new $[S]$.
   (b) $[X]$ will adapt to $[S]$ and their concentrations will eventually be equal. This is due to $k_3 = k_4 = 1$ in our case. Hence, $[X] = [S]$ eventually.
2. The higher $[S]$ is *prior* the step-wise increase, the lower will be the spike in $[R]$ when an increase in $[S]$ is induced (observe reducing peaks in Figure 4.3).
3. The steady-state response $R_{SS}$ is independent of its initial value $R_0$. Hence, any perturbation of $[R]$ during a simulation, results in the same value for $R_{SS}$.

For properties 1.a and 1.b, we evaluate $P_5$ and $P_6$ from Equations 4.9 and 4.12 using context $\Pi_{S,pert.}$ from Equation 4.10 (see Section 4.3.4). Thus, we use a *signal for system* to validate both and finally obtain *true* in both cases. Here we use $\Pi_{S,pert.} \triangleq [1, 10]S$.

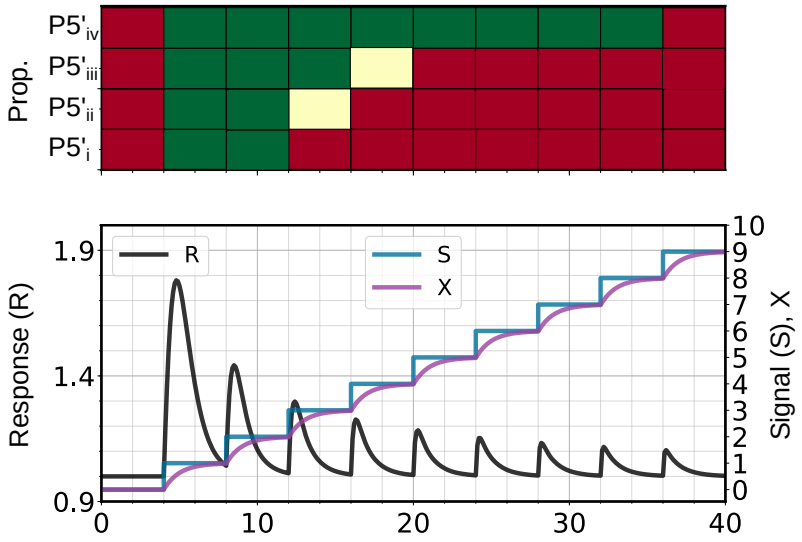



Figure 5.14 Evaluation of propositions $P'_{5,i..iv}$ over 9 step-wise increases of $[S]$ through context $\Pi_{S,pert.}$. Red for *false*, green for *true*, white for *uncertain*

However, we found that $\mathcal{LBUC}$ is not able to verify $P_5$ with context $\Pi_{S,pert.}$ if we set it the context mixture to $[0,10]S$. We assume that this occurs due to the inclusion of $[S_0] = 0$. Here, both $P_5$ and $P_6$ evaluate to *false* as no state change of $R$ and $X$ takes place. Thus, Flow* is probably not able to compute the flowpipe because of this behaviour. As we cannot refine our signal through the use of a *context signal* object, we resort to our previous result with $\Pi_{S,pert.}$.

$$P'_5 = \Pi_{S,step} \rhd F_{[0,t_1]}\left((R > r) \wedge F_{[0,t_2]}\left(G_{[0,t_2]}\left(\varepsilon_2 > R - \frac{k_1 k_4}{k_2 k_3}\right)\right)\right) \tag{5.9}$$

$$C_{PA} \models P'_5 \tag{5.10}$$

$$\Pi_{S,step} \triangleq [1]S \tag{5.11}$$

For property 2 we devise property $P'_5$ in Equation 5.9 (reproducing Equation 4.14 in Section 4.3.4). Here, when introducing $\Pi_{S,step} \triangleq [1]S$ (Equation 5.11) to our current process $C_{PA}$ (Equation 5.10), we add 1 to the current value of $[S] = [S_0]$. This reflects a discrete step increase in the concentration of $S$. However, we still face the issue to not capture the second feature of property 2. Therefore, we conduct the evaluation on 4 different versions of $P'_5$ namely, $P_{5,i..iv}$, each one with respective threshold values $r \in [1.4, 1.3, 1.2, 1.1]$.

In Figure 5.14, we introduce the context to 9 different systems $C_{PA1..9}$ where for each $C_{PAi} \mid i \in [1,9]$ we use as initial values for each specie concentration the concentrations gathered from the previous run $C_{PA(i-1)}$. This results in a step-wise increase in $[S]$ up to value $[S] = 9$ and allows us to evaluate each formula. Note that the first evaluation correctly reflects that, as no perturbation of the signal is introduced, $P'_5$ (as well as $P_5$ and $P_6$) evaluate to *false*. Now, as we increase $[S_0]$, we clearly see that the spike in $[R]$ reduces in value. Thus, the first to evaluate to

*false* is $P'_{5i}$ with threshold value $r = 1.4$. Respectively, the last to do so is $P'_{5iv}$. This shows that the property holds as we expect, although some manual work was required.

$$P_7 = F_{[0,t_2]}\left(G_{[0,t_2]}\left(\varepsilon_2 > R - \frac{k_1 k_4}{k_2 k_3}\right)\right) \tag{5.12}$$

$$C_{PA} \models \Pi_{R,pert.} \rhd P_7 \tag{5.13}$$

$$\Pi_{R,pert.} \triangleq [0,100]R \tag{5.14}$$

For the last property, we simply used a *signal for system* and a context process $\Pi_{R,pert.}$ (Equation 5.14) with $[0,100]R$ introduced to $P_7$ in Equation 5.13 (replicated from Section 4.3.4). $P_7$ (Equation 5.12) verifies that $[R]$ will *eventually* return to a steady-state value. Together with the context, we quantify that no matter what perturbation of $[R] \in [0,100]$ we experience, this is the case. This evaluates to *true* in the bondwb. However we note, that both concentration $[S],[X] \neq 0$ as otherwise no degradation of $[R]$ occurs, thus $P_7$ would not satisfy.

Concluding, we were able to show that properties 1.a and 1.b hold for the system although only for $[S] > 0$. For property 2 we devise 4 different propositions, capturing the reducing of the spike in $[R]$ as $[S]$ gradually increases. This approach serves our purposes and evaluates our propositions to *true*, however limits us in the band of validity. Ultimately, we show that $[R_S S]$ indeed is independent of any perturbations in $[R]$.

## 5.4 Summary

This chapter presents our qualitative and quantitative evaluation on the models presented. The qualitative evaluation compares our work, mostly visually, to the presentation by Tyson et al. [82]. We also compare some results with the work of Wang [84] who implemented a subset of the models presented here in c$\pi$.

However, we also illustrate the contrasts between different implementations of the same systems using the GK function and more explicit MM kinetics. Through the use of our novel approach of formal verification through $\mathcal{LBUC}$, we show that our characterizations of the first four systems (linear, hyperbolic, sigmoidal response, and perfect adaptation) hold within specific bounds.

# Chapter 6

# Discussion and Conclusion

We first discuss the three research questions initially posed and discuss to what degree we were able to answer these through our work. Subsequently, we present possible future work as well as an outlook on the bigger picture for both the bond-calculus and $\mathcal{LBUC}$.

## 6.1 Discussion

**1. Is the bond-calculus capable of modeling non-linear behaviour of complex systems such as oscillators with custom kinetic laws providing an abstraction of the underlying processes?** We were able to fully and accurately replicate all models presented by Tyson et al. [82] in the bond-calculus and to generate the exact same sets of ODEs as provided by the authors. Given our high-level bond-calculus models, we obtain the same system behaviours as identified by Tyson et al. using a low-level modeling approach.

By comparing our work with the work done by Wang [84], we furthermore show that the bond-calculus indeed improves on the c$\pi$. What we noticed most is the flexibility provided through custom rate laws. We found that the separation of species definition, rate laws, affinity networks, and processes brings bond-calculus models and the natural language description closer together. This is especially explicit when comparing our implementations with the system descriptions.

However, we also faced one major technical challenges during the implementation. The symbolic simplification algorithm of the bondwb caused it to return erroneous ODEs when using the GK function. To circumvent this, we extracted the MM kinetics underlying the function and modelled the systems explicitly. Ultimately, the issue was found and fixed and allowed us to implement the systems using the GK function. Nevertheless, being able to compare the explicit model using MM kinetics and the model using the GK function gave rise to our second research question.

**2. How do simplifications and abstractions affect our models and is it readily possible to employ more precise and verbose approximations to these in the bond-calculus?** We attempted to understand the biochemical validity of the systems presented by Tyson et al. [82] and compare the results from both, the MM and GK governed models of the same systems.

Abstractions and simplifications are key to modeling in system biology. However, the assumptions by the authors such as excluding phosphatase species in MM kinetics (derived in Section D.4) or presenting models which do not agree with the necessary assumptions of MM or GK kinetics (see Section 5.2), hinder reproducibility. In fact, Wang [84] who implemented the elementary building blocks in c$\pi$, supporting only MA kinetics, was not able to exactly replicate the behaviour. We were only able to do so as the bond-calculus allows us to encode arbitrary rate-laws.

For our systems exhibiting mutual activation, mutual inhibition, and homeostasis, we were able to replicate the GK function behaviour with MM kinetics. However, this is at the price of much longer stabilization times when using explicit MM than in the version with the GK function. Thus, presented with experimental data of in vivo systems, models with the GK function could reveal to work on a completely different time-scale. We have seen this especially for our implementations of the activator-inhibitor and substrate-depletion oscillators. Here, the different time scales of stabilization lead to skewed results in concentration values. Thus, we argue that certain assumptions and simplifications undertaken by the authors inhibit us from easily replicating the same behaviour with more precise and verbose methods. New steady state approximations might provide more accurate results in this regard as discussed further in Section 6.2.

**3. Provided with a bond-calculus model, is it possible to qualitatively, formally verify and guarantee its behaviour through the use of $\mathcal{LBUC}$?** Shifting our focus to the $\mathcal{LBUC}$ characterization of our systems, it should be stated that formulating these expressions can prove difficult. Although the syntax of $\mathcal{LBUC}$ is already narrowing the gap between natural language and formal statements with the context operator, one still requires considerable knowledge of the system and capabilities in encoding properties in temporal logic. In fact, the knowledge limitation also limited our characterizations in such that we do not believe them to be exhaustive.

Apart from this difficulty, we encountered limitations posed by the current state of the $\mathcal{LBUC}$ implementation in the bondwb. It proved easy to verify properties with big ranges in uncertainty on simple systems such as the linear response and perfect adaption. Here we showed that for big ranges in initial values of a certain species, its steady-state is unaffected. However, as the ODEs became more complex, we had to strongly limit the size of uncertainty we sought to verify our properties for. For the value exploration conducted on the ultra-sensitivity of the sigmoidal response, we had to resort to single-point values and even here, limitations due to memory leakage were encountered. While this might be due to the bondwb itself, we believe that for the more complex systems such as the negative-feedback oscillator, we reached Flow*'s limits of flowpipe computation, inhibiting us to further analyze these.

Ultimately, there was also a limitation from the point of few of available resources for both, the bondwb and $\mathcal{LBUC}$ itself. Being a novel logic specifically designed for biochemical systems, considerable trial and error was required to present the results in Chapter 5. However, we were able to identify bugs in both, the *context for signal* and the *reach object* computation, leading us to support the development.

Concluding, we note that it is possible to evaluate properties of simple system behaviour with $\mathcal{LBUC}$. Given a bond-calculus model, we are able to verify all properties devised to a satisfactory level of generalization. However, for more complicated system, the limitations of the $\mathcal{LBUC}$ implementation still prohibits extensive experimentation.

## 6.2 Future Work

Throughout our work, we had to comply with a variety of assumptions and approximations to reliably model the systems presented by Tyson et al. [82]. For the future, it would be interesting to implement the new findings by Kim and Tyson in their yet to be published paper on misuses of the MM kinetics [46]. Research is ongoing on alternative steady-state approximations [86], such as the total Quasi Steady-State Approximations (tQSSA) [16] which could help provide more biochemically accurate systems.

In terms of the bond-calculus, we think to merely have scratched the surface. It would be intriguing to model more complex, real-world systems and make use of the multi-way dynamic bonding capabilities of the calculus. Furthermore, we see great opportunities in further exploring the compositions feature of the bond-calculus. Processes can be composed from many smaller processes and could in the future lead to a database of bond-calculus models such as the ones presented here, to be used as plug-in components for complex systems. Moreover, leveraging the intuitive semantics of the bond-calculus, a user interface allowing system biologists to visually compose bond-calculus models could further lower the threshold on using it to formalize biochemical systems.

We want to emphasize the opportunities arising through the introduction of $\mathcal{LBUC}$ and its specific capability to verify biochemical systems. While our work merely introduced this framework, it would be interesting to further develop a work similar to ours but focusing solely on the formal characterization and verification of the models presented. Furthermore, we believe that this form of formal verification could prove essential for fields such as bio-medicine or pharmacy.

Finally, we terminate this work with the prospect that the bond-calculus and $\mathcal{LBUC}$ are openly available through the bondwb [87]. We want to increase the audience aware of this as well as increase its accessibility. For this reason in [30], we provide an informal and accessible blog post on the website medium.com to encourage further engagement by the public. Especially in contexts of crises such as the current *COVID-19* pandemic [61], system biology research on antibodies and vaccines becomes essential for quick action as shown by Guo et al. [35]. Although

we are aware that the bond-calculus will not provide game changing results immediately, we believe it has prospects to provide a useful framework in the future.

## 6.3 Conclusion

The main aim of this project was to obtain a working understanding of the bond-calculus and demonstrate that it is capable of modeling biochemical systems incorporating multiple species interacting through general kinetic laws. We base our work on multiple sources presenting core components of biochemical systems [82, 19, 60, 80], incorporating processes such as synthesis and degeneration, positive and negative feedback, as well as oscillatory networks tying multiple concepts together. The systems presented by Tyson et al. (2003) in [82] here present our starting point. We then formally characterize system behaviours of a subset of the systems modelled in Logic of Behaviour in Uncertain Contexts ($\mathcal{LBUC}$) and verified their satisfaction through the bondwb.

We accomplished to model and simulate all 7 basic systems as well as all 3 oscillators presented in [82] in the bond-calculus and verify their correctness qualitatively. Due to initial limitations we found alternative implementations for some of these which however in the course of our work were lifted and we were capable to model all systems as intended by Tyson et al.. Furthermore, we were able to characterize 3 systems out of the 10 formally and verify our formalized believes through the framework of $\mathcal{LBUC}$. Finally, we seek with this thesis to provide a comprehensive framework for future studies on biochemical system modeling in the bond-calculus as well as formal characterization of system behaviour with $\mathcal{LBUC}$.

# Bibliography


[1] Rajeev Alur and David L Dill. A theory of timed automata. *Theoretical computer science*, 126(2):183–235, 1994.

[2] Rajeev Alur and Thomas A Henzinger. Logics and models of real time: A survey. In *Workshop/School/Symposium of the REX Project (Research and Education in Concurrent Systems)*, pages 74–106. Springer, 1991.

[3] Rajeev Alur and Thomas A Henzinger. A really temporal logic. *Journal of the ACM (JACM)*, 41(1):181–203, 1994.

[4] Rajeev Alur, Costas Courcoubetis, Nicolas Halbwachs, Thomas A Henzinger, Pei-Hsin Ho, Xavier Nicollin, Alfredo Olivero, Joseph Sifakis, and Sergio Yovine. The algorithmic analysis of hybrid systems. *Theoretical computer science*, 138(1):3–34, 1995.

[5] Rajeev Alur, Tomás Feder, and Thomas A Henzinger. The benefits of relaxing punctuality. *Journal of the ACM (JACM)*, 43(1):116–146, 1996.

[6] Jakob L Andersen, Christoph Flamm, Daniel Merkle, and Peter F Stadler. Rule composition in graph transformation models of chemical reactions. *Match*, 80(3):661–704, 2018.

[7] Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT press, 2008.

[8] Chris J Banks and Ian Stark. A logic of behaviour in context. *Information and Computation*, 236:3–18, 2014.

[9] Attila Becskei, Bertrand Séraphin, and Luis Serrano. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *The EMBO journal*, 20(10):2528–2535, 2001.

[10] GE Briggs and JBS Haldane. A criticism of michaelis and menten's equation for enzyme action. *J. Biochem.*, 19:338, 1925.

[11] Neil A Campbell, Lisa A Urry, Michael L Cain, Steven A Wasserman, and Peter V Minorsky. *Biology: a global approach*. Pearson Higher Ed, 2017.

[12] Yuansheng Cao, Hongli Wang, Qi Ouyang, and Yuhai Tu. The free-energy cost of accurate biochemical oscillations. *Nature Physics*, 11(9):772–778, 2015.

[13] Katherine C Chen, Attila Csikasz-Nagy, Bela Gyorffy, John Val, Bela Novak, and John J Tyson. Kinetic analysis of a molecular model of the budding yeast cell cycle. *Molecular biology of the cell*, 11(1):369–391, 2000.

[14] Xin Chen. *Reachability Analysis of Non-Linear Hybrid Systems Using Taylor Models*. PhD thesis, Fachgruppe Informatik, RWTH Aachen University, 2015.

[15] Yoomi Chun and Joungmok Kim. Autophagy: an essential degradation program for cellular homeostasis and life. *Cells*, 7(12):278, 2018.

[16] Andrea Ciliberto, Fabrizio Capuani, and John J Tyson. Modeling networks of coupled enzymatic reactions using the total quasi-steady state approximation. *PLoS computational biology*, 3(3), 2007.

[17] Federica Ciocchetta and Jane Hillston. Bio-pepa: A framework for the modelling and analysis of biological systems. *Theoretical Computer Science*, 410(33-34):3065–3084, 2009.

[18] Philip Cohen. The origins of protein phosphorylation. *Nature cell biology*, 4(5):E127–E130, 2002.

[19] Emery D. Conrad and John J. Tyson. Modeling molecular interaction networks with nonlinear ordinary differential equations. In Zoltan Szallasi, Jörg Stelling, and Vipul Periwal, editors, *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*, chapter 6, pages 97–123. The MIT Press, 2006.

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

[21] G. T. Dodd and T. Tiganis. Insulin action in the brain: Roles in energy and glucose homeostasis. *Journal of Neuroendocrinology*, 29(10):e12513, 2017. doi: 10.1111/jne.12513. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jne.12513.

[22] Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[23] Lukas Endler. Biomd0000000306 - tyson2003 activator inhibitor, 2014.

[24] Lukas Endler. Biomd0000000309 - tyson2003 negfb homeostasis, 2014.

[25] Lukas Endler. Biomd0000000311 - tyson2003 mutual activation, 2014.

[26] Lukas Endler. Biomd0000000306 - tyson2003 mutual inhibition, 2014.

[27] Lukas Endler. Biomd0000000308 - tyson2003 negfb oscillator, 2014.

[28] Lukas Endler. Biomd0000000312 - tyson2003 perfect adaption, 2014.

[29] Lukas Endler. Biomd0000000307 - tyson2003 substrate depletion osc, 2014.

[30] Riccardo Fiorista. Modelling biochemical systems as communicating processes, April 2020. URL https://medium.com/@fiorista.riccardo/mbsacp-2020. [Online; posted 22-April-2020].

[31] Goran Frehse, Colas Le Guernic, Alexandre Donzé, Scott Cotton, Rajarshi Ray, Olivier Lebeltel, Rodolfo Ripado, Antoine Girard, Thao Dang, and Oded Maler. Spaceex: Scalable verification of hybrid systems. In *International Conference on Computer Aided Verification*, pages 379–395. Springer, 2011.

[32] Derek Gatherer. So what do we really mean when we say that systems biology is holistic? *BMC Systems Biology*, 4(1):22, 2010. ISSN 1752-0509. doi: 10.1186/1752-0509-4-22. URL https://doi.org/10.1186/1752-0509-4-22.

[33] Albert Goldbeter and Daniel E Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, 78 (11):6840–6844, 1981.

[34] Cato M Guldberg and Peter Waage. Studies concerning affinity. *CM Forhandlinger: Videnskabs-Selskabet i Christiana*, 35(1864):1864, 1864.

[35] Li Guo, Lili Ren, Siyuan Yang, Meng Xiao, De Chang, Fan Yang, Charles S Dela Cruz, Yingying Wang, Chao Wu, Yan Xiao, Lulu Zhang, Lianlian Han, Shengyuan Dang, Yan Xu, Qiwen Yang, Shengyong Xu, Huadong Zhu, Yingchun Xu, Qi Jin, Lokesh Sharma, Linghang Wang, and Jianwei Wang. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). *Clinical Infectious Diseases*, 03 2020. ISSN 1058-4838. doi: 10.1093/cid/ciaa310. URL https://doi.org/10.1093/cid/ciaa310. ciaa310.

[36] Leonard A Harris, Justin S Hogg, José-Juan Tapia, John AP Sekar, Sanjana Gupta, Ilya Korsunsky, Arshi Arora, Dipak Barua, Robert P Sheehan, and James R Faeder. Bionetgen 2.2: advances in rule-based modeling. *Bioinformatics*, 32(21):3366–3368, 2016.

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

[38] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117 (4):500–544, 1952.

[39] Alexander Hoffmann, Andre Levchenko, Martin L Scott, and David Baltimore. The i$\kappa$b-nf-$\kappa$b signaling module: temporal control and selective gene activation. *Science*, 298(5596): 1241–1245, 2002.

[40] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J.-H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novère, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, J. Wang, and the rest of the SBML Forum:. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 03 2003. ISSN 1367-4803. doi: 10.1093/bioinformatics/btg015. URL https://doi.org/10.1093/bioinformatics/btg015.

[41] Eric Jones, Travis Oliphant, and Pearu Peterson. Scipy: Open source scientific tools for python. 2001.

[42] Dominic Jordan, Peter Smith, Peter Smith, et al. *Nonlinear Ordinary Differential Equations: An Introduction for Scientists and Engineers*, volume 10. Oxford University Press on Demand, 2007.

[43] Francis J. Doyle III Jörg Stelling, Uwe Sauer and John Doyle. Complexity and robustness of cellular systems. In Zoltan Szallasi, Jörg Stelling, and Vipul Periwal, editors, *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*, chapter 2, pages 19–40. The MIT Press, 2006.

[44] Jonathan R. Karr, Jayodita C. Sanghvi, Derek N. Macklin, Miriam V. Gutschow, Jared M. Jacobs, Benjamin Bolival, Nacyra Assad-Garcia, John I. Glass, and Markus W. Covert. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389 – 401, 2012. ISSN 0092-8674. doi: https://doi.org/10.1016/j.cell.2012.05.044. URL http://www.sciencedirect.com/science/article/pii/S0092867412007763.

[45] Douglas B. Kell and Joshua D. Knowles. The role of modeling in systems biology. In Zoltan Szallasi, Jörg Stelling, and Vipul Periwal, editors, *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*, chapter 1, pages 3–18. The MIT Press, 2006.

[46] Jae Kyoung Kim and John J. Tyson. Misuse of the michaelis-menten rate law for protein-interaction networks and its remedy. Obtained through private communication with John. J. Tyson (tyson@vt.edu).

[47] Shankara Narayanan Krishna and Ashutosh Trivedi. Hybrid automata for formal modeling and verification of cyber-physical systems. *arXiv preprint arXiv:1503.04928*, 2015.

[48] Marek Kwiatkowski. Formal computational framework for the study of molecular evolution. 2010.

[49] Marek Kwiatkowski and Ian Stark. The continuous $\pi$-calculus: A process algebra for biochemical modelling. In Monika Heiner and Adelinde M. Uhrmacher, editors, *Computational Methods in Systems Biology*, pages 103–122, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. ISBN 978-3-540-88562-7.

[50] Lucas Laursen. Biological logic. *Nature*, 462(26):408–410, 2009.

[51] Nicolas Le Novere, Benjamin Bornstein, Alexander Broicher, Melanie Courtot, Marco Donizelli, Harish Dharuri, Lu Li, Herbert Sauro, Maria Schilstra, Bruce Shapiro, et al. Biomodels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic acids research*, 34(suppl 1):D689–D691, 2006.

[52] Timo R Maarleveld, Brett G Olivier, and Frank J Bruggeman. Stochpy: a comprehensive, user-friendly tool for simulating stochastic biological processes. *PloS one*, 8(11), 2013.

[53] Rahuman S Malik-Sheriff, Mihai Glont, Tung V N Nguyen, Krishna Tiwari, Matthew G Roberts, Ashley Xavier, Manh T Vu, Jinghao Men, Matthieu Maire, Sarubini Kananathan, Emma L Fairbanks, Johannes P Meyer, Chinmay Arankalle, Thawfeek M Varusai, Vincent Knight-Schrijver, Lu Li, Corina Dueñas-Roca, Gaurhari Dass, Sarah M Keating, Young M Park, Nicola Buso, Nicolas Rodriguez, Michael Hucka, and Henning Hermjakob. BioModels—15 years of sharing computational models in life science. *Nucleic Acids Research*, 11 2019. ISSN 0305-1048. doi: 10.1093/nar/gkz1055. URL https://doi.org/10.1093/nar/gkz1055. gkz1055.

[54] Satomi Matsuoka and Masahiro Ueda. Mutual inhibition between pten and pip3 generates bistability for polarity in motile cells. *Nature communications*, 9(1):1–15, 2018.

[55] Andrew McCrae. Cpi-ide: An integrated development environment for the continuous pi-calculus. https://github.com/continuouspi/cpi-ide/blob/master/models/testEnzyme.cpi, 2014.

[56] Leonor Michaelis and Maud Leonora Menten. *Die Kinetik der Invertinwirkung*. Universitätsbibliothek Johann Christian Senckenberg, 1913.

[57] Robin Milner, Joachim Parrow, and David Walker. A calculus of mobile processes, ii. *Information and computation*, 100(1):41–77, 1992.

[58] David L Nelson, Michael M Cox, and Albert L Lehninger. *Principles of biochemistry*. Freeman New York:, 2008.

[59] Bela Novak and John J Tyson. Numerical analysis of a comprehensive model of m-phase control in xenopus oocyte extracts and intact embryos. *Journal of cell science*, 106(4): 1153–1168, 1993.

[60] Béla Novák and John J Tyson. Design principles of biochemical oscillators. *Nature reviews Molecular cell biology*, 9(12):981–991, 2008.

[61] World Health Organization et al. Coronavirus disease 2019 (covid-19): situation report, 72. 2020.

[62] Vipul Periwal. Computational constraints on modeling in systems biology. In Zoltan Szallasi, Jörg Stelling, and Vipul Periwal, editors, *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*, chapter 15, pages 315–330. The MIT Press, 2006.

[63] Corrado Priami, Paolo Ballarini, and Paola Quaglia. Blenx4bio – blenx for biologists. In Pierpaolo Degano and Roberto Gorrieri, editors, *Computational Methods in Systems Biology*, pages 26–51, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. ISBN 978-3-642-03845-7.

[64] Aviv Regev and Ehud Shapiro. Cellular abstractions: Cells as computation. *Nature*, 419 (6905):343–343, 2002.

[65] Aviv Regev, William Silverman, and Ehud Shapiro. Representation and simulation of biochemical processes using the $\pi$-calculus process algebra. In *Biocomputing 2001*, pages 459–470. World Scientific, 2000.

[66] Xin Chen Rwth, Sriram Sankaranarayanan, and Erika Ábrahám. Under-approximate flowpipes for non-linear continuous systems. In *2014 Formal Methods in Computer-Aided Design (FMCAD)*, pages 59–66. IEEE, 2014.

[67] Michael A Savageau. Design principles for elementary gene circuits: Elements, methods, and examples. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(1):142–159, 2001.

[68] Prof. Dr. Ulrich Schwarz. Course manuscript: Non-linear dynamics. 2017. URL https://www.thphys.uni-heidelberg.de/~biophys/PDF/Skripte/NonlinearDynamics.pdf.

[69] EE Sel'Kov. Self-oscillations in glycolysis 1. a simple kinetic model. *European Journal of Biochemistry*, 4(1):79–86, 1968.

[70] Paul Smolen, Douglas A Baxter, and John H Byrne. Modeling circadian oscillations with interlocking positive and negative feedback loops. *Journal of Neuroscience*, 21(17): 6644–6656, 2001.

[71] Peter A Spiro, John S Parkinson, and Hans G Othmer. A model of excitation and adaptation in bacterial chemotaxis. *Proceedings of the National Academy of Sciences*, 94(14):7263–7268, 1997.

[72] Zoltan Szallasi, Jörg Stelling, and Vipul Periwal. System modeling in cellular biology. *From Concepts to*, 2006.

[73] LUCAS SÁNCHEZ and DENIS THIEFFRY. A logical analysis of the drosophila gap-gene system. *Journal of Theoretical Biology*, 211(2):115 – 141, 2001. ISSN 0022-5193. doi: https://doi.org/10.1006/jtbi.2001.2335. URL http://www.sciencedirect.com/science/article/pii/S0022519301923355.

[74] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version x.y.z)*, YYYY. `https://www.sagemath.org`.

[75] René Thomas and Morten Kaufman. Multistationarity, the basis of cell differentiation and memory. i. structural conditions of multistationarity and other nontrivial behavior. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(1):170–179, 2001.

[76] Anthony Trewavas. A brief history of systems biology. *The Plant Cell*, 18(10):2420–2430, 2006. ISSN 1040-4651. doi: 10.1105/tpc.106.042267. URL http://www.plantcell.org/content/18/10/2420.

[77] Farida Tripodi, Raffaele Nicastro, Veronica Reghellin, and Paola Coccetti. Post-translational modifications on yeast carbon metabolism: Regulatory mechanisms beyond transcriptional control. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1850(4):620 – 627, 2015. ISSN 0304-4165. doi: https://doi.org/10.1016/j.bbagen.2014.12.010. URL http://www.sciencedirect.com/science/article/pii/S0304416514004103.

[78] John Tyson. Re: Regarding sniffers, buzzers, toggles and blinkers (2003). private communication, 2020.

[79] John J Tyson. Classification of instabilities in chemical reaction systems. *The Journal of Chemical Physics*, 62(3):1010–1015, 1975.

[80] John J Tyson and Béla Novák. Functional motifs in biochemical reaction networks. *Annual review of physical chemistry*, 61:219–240, 2010.

[81] John J Tyson, Bela Novak, Garrett M Odell, Kathy Chen, and C Dennis Thron. Chemical kinetic theory: understanding cell-cycle regulation. *Trends in biochemical sciences*, 21(3):89–96, 1996.

[82] John J Tyson, Katherine C Chen, and Bela Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15(2):221 – 231, 2003. ISSN 0955-0674. doi: https://doi.org/10.1016/S0955-0674(03)00017-6. URL http://www.sciencedirect.com/science/article/pii/S0955067403000176.

[83] George H Wadhams and Judith P Armitage. Making sense of it all: bacterial chemotaxis. *Nature reviews Molecular cell biology*, 5(12):1024–1037, 2004.

[84] Stanley Wang. Modelling Biological Systems as Communicating Processes, 2016.

[85] C Wanjek. Systems biology as defined by nih: an intellectual resource for integrative biology. *The NIH Catalyst*, 19(6):1–12, 2011.

[86] Martin Kin Lok Wong, James Robert Krycer, James Geoffrey Burchfield, David Ernest James, and Zdenka Kuncic. A generalised enzyme kinetic model for predicting the behaviour of complex biochemical systems. *FEBS open bio*, 5:226–239, 2015.

[87] Thomas Wright. Bond calculus workbench and continuous pi workbench combined repo. https://github.com/twright/bondwb, 2020.

[88] Thomas Wright and Ian Stark. The bond-calculus: A process algebra for complex biological interaction dynamics. *arXiv preprint arXiv:1804.07603*, 2018.

[89] Thomas Wright and Ian Stark. Modelling patterns of gene regulation in the bond-calculus. *Elsevier*, 2018.

[90] Yang Yang and Min Wu. Rhythmicity and waves in the cortex of single cells. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1747):20170116, 2018.

[91] Vipul Periwal Zoltan Szallasi and Jörg Stelling. On modules and modularity. In Zoltan Szallasi, Jörg Stelling, and Vipul Periwal, editors, *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*, chapter 3, pages 41–50. The MIT Press, 2006.

# Appendix A

# Bond Calculus Models

## A.1 Linear Response

### A.1.1 Ambient Species

```
species A = asynR -> (R | A);
species S = ssynR -> (S | R);
species R = degradeR -> 0;

kinetic law CL(k; A) = k*A;

affinity network N(a, s, r) = {
    asynR at rate CL(a);
    ssynR at rate MA(s);
    degradeR at rate MA(r);
}
process Pi = [1.0] A || [1.0] S || [0] R with network N(0.01, 1, 5);
```

### A.1.2 Simplified

```
species S = synthR -> (S | R);
species R = decayR -> 0;

kinetic law CL(a, k; X) = a + k*X;

affinity network N(a, s, r) = {
    synthR at rate CL(a,s);
    decayR at rate MA(r);
}
process Pi = [1.0] S || [0] R with network N(0.01, 1, 5);
```

## A.2 Hyperbolic Response

```
species S = kinaseR -> S;
species R = phosphorylate -> RP;
species RP = dephosphorylate -> R;

affinity network N(k, m) = {
    phosphorylate || kinaseR at rate MA(k);
    dephosphorylate at rate MA(m);
}
process Pi = [1.0] S || [1.0] R || [0.0] RP with network N(1.1,1.2);
```

## A.3 Sigmoidal Response

### A.3.1 Regular

```
species S = kinaseR -> S;
species R = phosphorylate -> RP;
species RP = dephosphorylate -> R;

kinetic law MM1(k,km; S,R) = (k*S*R) / (km + R);
kinetic law MM2(k,km; RP)  = (k*RP) / (km + RP);

affinity network N(k1, k2, km1, km2) = {
    kinaseR || phosphorylate at rate MM1(k1,km1);
    dephosphorylate at rate MM2(k2,km2);
}
process Pi = [1.0] S || [1.0] R || [0.0] RP with network N(1,1,0.05,0.05);
```

### A.3.2 Rate Exploration

```
species S = kinaseR -> S;
species R = phosphorylate -> RP;
species RP = dephosphorylate -> R;
species KM1 = stayKM1 -> KM1;
species KM2 = stayKM2 -> KM2;

kinetic law MM1(k; S,R,km) = (k*S*R) / (km + R);
kinetic law MM2(k; RP,km)  = (k*RP) / (km + RP);

```

```
affinity network N(k1, k2) = {
    kinaseR || phosphorylate, stayKM1 at rate MM1(k1);
    dephosphorylate, stayKM2 at rate MM2(k2);
}

process Pi = [1.0] S || [1.0] R || [0.0] RP
            || [0.05] KM1 || [0.05] KM2 with network N(1,1);
```

## A.4 Perfect Adaptation

```
species S = ssynR -> (S | R)
          + ssynX -> (S | X);
species R = degradeR -> 0;
species X = degradeX -> 0
          + stayX -> X;

affinity network N(k1, k2, k3, k4) = {
    ssynR at rate MA(k1);
    ssynX at rate MA(k3);
    degradeR || stayX at rate MA(k2);
    degradeX at rate MA(k4);
}
process Pi = [1.0] S || [1.0] X || [1.0] R with network N(2,2,1,1);
```

## A.5 Mutual Activation

### A.5.1 GK Governed

```
species S = ssynR -> (S | R);
species R = degradeR -> 0
          + epsynR -> (R | R);

kinetic law G(f,u,v,J,K; R) =
        (f*2*u*R*K) / (v-u*R+v*J+u*R*K+((v-u*R+v*J+u*R*K)**2
            -4*(v-u*R)*u*R*K)**0.5);

affinity network N(k0, k1, k2, k3, k4, j3, j4) = {
    ssynR at rate MA(k1);
    degradeR at rate MA(k2);
    epsynR at rate G(k0,k3,k4,j3,j4);
```

```
}
process Pi = [0.0] S || [0.0] R
            with network N(0.4,0.01,1,1,0.2,0.05,0.05);
```

### A.5.2 MM Governed

```
species S = ssynR -> (S | R);
species R = degradeR -> 0
          + stayR -> R;
species E = phosphorylateE -> EP;
species EP = dephosphorylateEP -> E
           + epsynR -> (EP | R);

kinetic law MM1(k,km; R,E) = (k*R*E) / (km + E);
kinetic law MM2(k,km; EP) = (k*EP) / (km + EP);

affinity network N(k0, k1, k2, k3, k4, j3, j4) = {
    epsynR at rate MA(k0);
    ssynR at rate MA(k1);
    degradeR at rate MA(k2);
    stayR || phosphorylateE at rate MM1(k3,j3);
    dephosphorylateEP at rate MM2(k4,j4);
}
process Pi = [0.0] S || [0.0] R
            || [1.0] E || [0.0] EP
            with network N(0.4,0.01,1,1,0.2,0.05,0.05);
```

## A.6 Mutual Inhibition

### A.6.1 GK Governed

```
species S = ssynthR -> (S | R);
species R = decayR -> 0
          + stayR -> R;
species A = asynthR -> (A | R);

kinetic law G(f,u,v,J,K; R) =
        (2*u*K*f*R) / (v*R-u+v*J*R+u*K+((v*R-u+v*J*R+u*K)**2
            -4*(v*R-u)*u*K)**0.5);

```

```
affinity network N(k0, k1, k2, k21, k3, k4, j3, j4) = {
    asynthR at rate MA(k0);
    ssynthR at rate MA(k1);
    decayR at rate MA(k2);
    decayR at rate G(k21,k3,k4,j3,j4);
}
process Pi = [0.0] S || [0.0] R || [1.0] A
            with network N(0,0.05,0.1,0.5,0.2,1,0.05,0.05);
```

### A.6.2 MM Governed

```
species S = ssynthR -> (S | R);
species R = decayR -> 0
          + stayR -> R;
species E = phosphE -> EP
          + stayE -> E;
species EP = dephosphEP -> E;
species A = asynthR -> (A | R);

kinetic law MM1(k,km; E, R) = (k*R*E) / (km + E);
kinetic law MM2(k,km; EP) = (k*EP) / (km + EP);

affinity network N(k0, k1, k2, k21, k3, k4, j3, j4) = {
    asynthR at rate MA(k0);
    ssynthR at rate MA(k1);
    decayR at rate MA(k2);
    decayR || stayE at rate MA(k21);
    phosphE || stayR at rate MM1(k3,j3);
    dephosphEP at rate MM2(k4,j4);
}
process Pi = [0.0] S || [0.0] R || [0.0] E
            || [1.0] EP || [1.0] A
            with network N(0,0.05,0.1,0.5,1,0.2,0.05,0.05);
```

## A.7 Homeostasis

### A.7.1 GK Governed

```
species S = sdegradeR -> S;
species R = degradeR -> 0
```

```
3           + synR -> (R | R);
4
5  kinetic law G(f,u,v,J,K; R) =
6        2*u*K*f / (v*R-u+v*J*R+u*K+((v*R-u+v*J*R+u*K)**2
7              -4*(v*R-u)*u*K)**0.5);
8
9  affinity network N(k0, k2, k3, k4, j3, j4) = {
10     sdegradeR || degradeR at rate MA(k2);
11     synR at rate G(k0,k3,k4,j3,j4);
12 }
13
14 process Pi = [1.0] S || [0.0] R
15                 with network N(1,1,0.5,1,0.01,0.01);
```

### A.7.2 MM Governed

```
1  species S = sdecayR -> S;
2  species R = decayR -> 0
3            + stayR -> R;
4  species E = phosphE -> EP
5            + esynthR -> (E | R);
6  species EP = dephosphEP -> E;
7
8  kinetic law MM1(k,km; R,E) = (k*R*E) / (km + E);
9  kinetic law MM2(k,km; EP) = (k*EP) / (km + EP);
10
11 affinity network N(k0, k2, k3, k4, j3, j4) = {
12     esynthR at rate MA(k0);
13     sdecayR || decayR at rate MA(k2);
14     stayR || phosphE at rate MM1(k4,j4);
15     dephosphEP at rate MM2(k3,j3);
16 }
17 process Pi = [0.0] S || [0.0] R
18              || [1.0] E || [0.0] EP
19              with network N(1,1,0.5,1,0.01,0.01);
```

## A.8 Negative-Feedback Oscillator

```
1  species A = asynX -> (A | X);
2  species S = ssynX -> (S | X);
```

```
species X = stayX -> X
          + degradeX -> 0;
species Y = phosphorylateY -> YP;
species YP = stayYP -> YP
           + dephosphorylateYP -> Y;
species R = phosphorylateR -> RP;
species RP = stayRP -> RP
           + dephosphorylateRP -> R;

kinetic law MM1(k,km; S,E) = (k*S*E) / (km + E);
kinetic law MM2(k,km; EP) = (k*EP) / (km + EP);

affinity network N(k0, k1, k2, k21, k3, k4, k5, k6, km3, km4, km5, km6) = {
    asynX at rate MA(k0);
    ssynX at rate MA(k1);
    degradeX at rate MA(k2);
    stayX || phosphorylateY at rate MM1(k3, km3);
    dephosphorylateYP at rate MM2(k4, km4);
    stayYP || phosphorylateR at rate MM1(k5, km5);
    dephosphorylateRP at rate MM2(k6, km6);
    stayRP || degradeX at rate MA(k21);
}
process Pi = [1.0] A || [2.0] S || [0] X
            || [1.0] Y || [0.0] YP || [1.0] R
            || [0.0] RP with network
            N(0   , 1   , 0.01, 10.0,
              0.1 , 0.2 , 0.1 , 0.05,
              0.01, 0.01, 0.01, 0.01);
```

# A.9 Activator-Inhibitor Oscillator

## A.9.1 GK Governed

```
species S = ssynR -> (S | R);
species R = rsynR -> (R|R)
          + degradeR -> 0
          + rsynX -> (R | X);
species X = stayX -> X
          + degradeX -> 0;

```

```
kinetic law G(f,u,v,J,K; R) =
        (f*2*u*R*K) / (v-u*R+v*J+u*R*K+((v-u*R+v*J+u*R*K)**2
            -4*(v-u*R)*u*R*K)**0.5);

affinity network N(k0, k1, k2, k21, k3, k4, k5, k6, j3, j4) = {
    ssynR at rate MA(k1);
    degradeR at rate MA(k2);
    degradeX at rate MA(k6);
    rsynX at rate MA(k5);
    stayX || degradeR at rate MA(k21);
    rsynR at rate G(k0,k3,k4,j3,j4);
}

process Pi = [0.2] S || [0.0] X || [0.0] R
        with network N(4,1,1,1,1 ,1, 0.1,0.075,0.3,0.3);
```

### A.9.2 MM Governed

```
species S = ssynthR -> (S | R);
species R = stayR -> R
          + decayR -> 0
          + rsynthX -> (R | X);
species X = stayX -> X
          + decayX -> 0;
species E = phospE -> EP;
species EP = stayEP -> EP
           + epsynthR -> (EP | R)
           + dephosphEP -> E;

kinetic law MM1(k,km; R,E) = (k*R*E) / (km + E);
kinetic law MM2(k,km; EP) = (k*EP) / (km + EP);

affinity network N(k0, k1, k2, k21, k3, k4, k5, k6, km3, km4) = {
    ssynthR at rate MA(k1);
    decayR at rate MA(k2);
    decayX at rate MA(k6);
    rsynthX at rate MA(k5);
    stayX || decayR at rate MA(k21);
    stayR || phospE at rate MM1(k3,km3);
    dephosphEP at rate MM2(k4,km4);
```

```
    epsynthR at rate MA(k0);
}

process Pi = [0.2] S || [0.0] X || [1.0] E
            || [0.0] EP || [0.0] R with
            network N(4, 1   , 1    , 1   , 1 ,
                      1, 0.1 , 0.075, 0.3 , 0.3);
```

## A.10 Substrate-Depletion Oscillator

### A.10.1 GK Governed

```
species S = ssynX -> (S | X);
species R = stayR -> R
          + degradeR -> 0;
species X = xturnR -> R;

kinetic law G(f,u,v,J,K; R) =
        (f*2*u*R*K) / (v-u*R+v*J+u*R*K+((v-u*R+v*J+u*R*K)**2
            -4*(v-u*R)*u*R*K)**0.5);

affinity network N(k01, k0, k1, k2, k3, k4, j3, j4) = {
    ssynX at rate MA(k1);
    degradeR at rate MA(k2);
    xturnR at rate MA(k01);
    stayR || xturnR at rate G(k0,k3,k4,j3,j4);
}

process Pi = [0.2] S || [0.0] X || [0.0] R
        with network N(0.01,0.4,1,1,1,0.3,0.05,0.05);
```

### A.10.2 MM Governed

```
species S = ssynthX -> (S | X);
species R = stayR -> R
          + decayR -> 0;
species X = xturnR -> R;
species E = phosphE -> EP;
species EP = stayEP -> EP
           + dephosphEP -> E;
```

```
8
9  kinetic law MM1(k,km; R,E) = (k*R*E) / (km + E);
10 kinetic law MM2(k,km; EP) = (k*EP) / (km + EP);
11
12 affinity network N(k01, k0, k1, k2, k3, k4, km3, km4) = {
13     ssynthX at rate MA(k1);
14     decayR at rate MA(k2);
15     xturnR at rate MA(k01);
16     xturnR || stayEP at rate MA(k0);
17     stayR || phosphE at rate MM1(k3,km3);
18     dephosphEP at rate MM2(k4,km4);
19 }
20 process Pi = [0.2] S || [0.0] X || [1.0] E
21             || [0.0] EP || [0.0] R
22             with network N(0.01, 0.4, 1   , 1,
23                            1   , 0.3, 0.05, 0.05);
```

# Appendix B

# Generated Systems of ODEs

## B.1 Linear Response

### B.1.1 Ambient Species

$$\frac{dS}{dt} = 0 \tag{B.1}$$
$$\frac{dR}{dt} = 1 \cdot S - 5 \cdot R + 0.001 \cdot A \tag{B.2}$$
$$\frac{dA}{dt} = 0 \tag{B.3}$$

### B.1.2 Simplified

$$\frac{dS}{dt} = 0 \tag{B.4}$$
$$\frac{dR}{dt} = 1 \cdot S - 5 \cdot R + 0.001 \tag{B.5}$$

## B.2 Hyperbolic Response

$$\frac{dS}{dt} = 0 \tag{B.6}$$
$$\frac{dR}{dt} = -1.101 \cdot S \cdot R + 1.202 \cdot R_P \tag{B.7}$$
$$\frac{dR_P}{dt} = 1.101 \cdot S \cdot R - 1.202 \cdot R_P \tag{B.8}$$

## B.3 Sigmoidal Response

### B.3.1 Regular

$$\frac{dR_P}{dt} = \frac{(R_P + 0.05) \cdot S \cdot R - R_P \cdot (R + 0.05)}{(R_P + 0.05) \cdot (R + 0.05)} \tag{B.9}$$

$$\frac{dS}{dt} = 0 \tag{B.10}$$

$$\frac{dR}{dt} = -\frac{(R_P + 0.05) \cdot S \cdot R - R_P \cdot (R + 0.05)}{(R_P + 0.05) \cdot (R + 0.05)} \tag{B.11}$$

### B.3.2 Rate Exploration

$$\frac{dR}{dt} = -((K_{m2} + R_P) \cdot S \cdot R - (K_{m1} + R) \cdot R_P)/((K_{m2} + R_P) \cdot (K_{m1} + R)) \tag{B.12}$$

$$\frac{dR_P}{dt} = ((K_{m2} + R_P) \cdot S \cdot R - (K_{m1} + R) \cdot R_P)/((K_{m2} + R_P) \cdot (K_{m1} + R)) \tag{B.13}$$

$$\frac{dS}{dt} = 0 \tag{B.14}$$

$$\frac{dK_{m1}}{dt} = 0 \tag{B.15}$$

$$\frac{dK_{m2}}{dt} = 0 \tag{B.16}$$

## B.4 Perfect Adaptation

$$\frac{dS}{dt} = 0 \tag{B.17}$$

$$\frac{dR}{dt} = -2 \cdot R \cdot X + 2 \cdot S \tag{B.18}$$

$$\frac{dX}{dt} = 1 \cdot S - 1 \cdot X \tag{B.19}$$

## B.5 Mutual Activation

### B.5.1 GK Governed

$$\frac{dS}{dt} = 0 \tag{B.20}$$

$$\frac{dR}{dt} = 0.01 \cdot S - R + 0.04 \cdot \frac{R}{-0.95 \cdot R + \sqrt{0.20 \cdot R^2 + (-0.95 \cdot R + 0.21)^2.0 - 0.04 \cdot R} + 0.21} \tag{B.21}$$

### B.5.2 MM Governed

$$\frac{dS}{dt} = 0 \tag{B.22}$$

$$\frac{dR}{dt} = 0.01 \cdot S - R + 0.40 \cdot E_P \tag{B.23}$$

$$\frac{dE}{dt} = -\frac{R \cdot E \cdot (E_P + 0.05) - 0.20 \cdot (E + 0.05) \cdot E_P}{(E + 0.05) \cdot (E_P + 0.05)} \tag{B.24}$$

$$\frac{dE_P}{dt} = \frac{R \cdot E \cdot (E_P + 0.05) - 0.20 \cdot (E + 0.05) \cdot E_P}{(E + 0.05) \cdot (E_P + 0.05)} \tag{B.25}$$

## B.6 Mutual Inhibition

### B.6.1 GK Governed

$$\frac{dR}{dt} = 0.05 \cdot S - 0.10 \cdot R - 0.01 \cdot \frac{R}{(1.05 \cdot R + \sqrt{(1.05 \cdot R - 0.19)^2.0 - 0.04 \cdot R + 0.008} - 0.19} \tag{B.26}$$

$$\frac{dS}{dt} = 0 \tag{B.27}$$

$$\frac{dA}{dt} = 0 \tag{B.28}$$

### B.6.2 MM Governed

$$\frac{dS}{dt} = 0 \tag{B.29}$$
$$\frac{dR}{dt} = -0.50 \cdot R \cdot E + 0.05 \cdot S - 0.10 \cdot R \tag{B.30}$$
$$\frac{dA}{dt} = 0 \tag{B.31}$$
$$\frac{dE}{dt} = -\frac{R \cdot E \cdot (E_P + 0.05) - 0.20 \cdot (E + 0.05) \cdot E_P}{(E + 0.05) \cdot (E_P + 0.05)} \tag{B.32}$$
$$\frac{dE_P}{dt} = \frac{R \cdot E \cdot (E_P + 0.05) - 0.20 \cdot (E + 0.05) \cdot E_P}{(E + 0.05) \cdot (E_P + 0.05)} \tag{B.33}$$

## B.7 Homeostasis

### B.7.1 GK Governed

$$\frac{dS}{dt} = 0 \tag{B.34}$$
$$\frac{dR}{dt} = -S \cdot R + \frac{0.01}{1.01 \cdot R + \sqrt{(1.01 \cdot R - 0.495)^2.0 - 0.020 \cdot R + 0.01} - 0.495} \tag{B.35}$$

### B.7.2 MM Governed

$$\frac{dS}{dt} = 0 \tag{B.36}$$
$$\frac{dR}{dt} = -S \cdot R + E \tag{B.37}$$
$$\frac{dE}{dt} = -\frac{R \cdot E \cdot (E_P + 0.01) - 0.50 \cdot (E + 0.01) \cdot E_P}{(E + 0.01) \cdot (E_P + 0.01)} \tag{B.38}$$
$$\frac{dE_P}{dt} = \frac{R \cdot E \cdot (E_P + 0.01) - 0.50 \cdot (E + 0.01) \cdot E_P}{(E + 0.01) \cdot (E_P + 0.01)} \tag{B.39}$$

## B.8 Negative-Feedback Oscillator

$$\frac{dS}{dt} = 0 \tag{B.40}$$
$$\frac{dA}{dt} = 0 \tag{B.41}$$
$$\frac{dX}{dt} = -10. \cdot R_P \cdot X + S - 0.01 \cdot X \tag{B.42}$$
$$\frac{dR}{dt} = \frac{-0.10 \cdot (R_P + 0.01) \cdot R \cdot YP + 0.05 \cdot R_P \cdot (R + 0.01)}{(R_P + 0.01) \cdot (R + 0.01)} \tag{B.43}$$
$$\frac{dR_P}{dt} = \frac{0.10 \cdot (R_P + 0.01) \cdot R \cdot YP - 0.05 \cdot R_P \cdot (R + 0.01)}{(R_P + 0.01) \cdot (R + 0.01)} \tag{B.44}$$
$$\frac{dY}{dt} = \frac{-0.10 \cdot Y \cdot X \cdot (YP + 0.01) + 0.20 \cdot (Y + 0.01) \cdot YP}{(Y + 0.01) \cdot (YP + 0.01)} \tag{B.45}$$
$$\frac{dY_P}{dt} = \frac{0.10 \cdot Y \cdot X \cdot (YP + 0.01) - 0.20 \cdot (Y + 0.01) \cdot YP}{(Y + 0.01) \cdot (YP + 0.01)} \tag{B.46}$$

## B.9 Activator-Inhibitor Oscillator

### B.9.1 GK Governed

$$\frac{dS}{dt} = 0 \tag{B.47}$$
$$\frac{dR}{dt} = -R \cdot X + S - R + 2.4 \cdot \frac{R}{(-0.70 \cdot R + \sqrt{1.2 \cdot R^2 + (-0.70 \cdot R + 1.3)^2.0 - 1.2 \cdot R} + 1.3} \tag{B.48}$$
$$\frac{dX}{dt} = 0.10 \cdot R - 0.075 \cdot X \tag{B.49}$$

### B.9.2 MM Governed

$$\frac{dS}{dt} = 0 \tag{B.50}$$
$$\frac{dR}{dt} = -R \cdot X + S - R + 4.0 \cdot E_P \tag{B.51}$$
$$\frac{dX}{dt} = 0.10 \cdot R - 0.075 \cdot X \tag{B.52}$$
$$\frac{dE}{dt} = -\frac{R \cdot E \cdot (E_P + 0.30) - (E + 0.30) \cdot E_P}{(E + 0.30) \cdot (E_P + 0.30)} \tag{B.53}$$
$$\frac{dE_P}{dt} = \frac{R \cdot E \cdot (E_P + 0.30) - (E + 0.30) \cdot E_P}{(E + 0.30) \cdot (E_P + 0.30)} \tag{B.54}$$

# B.10 Substrate-Depletion Oscillator

## B.10.1 GK Governed

$$\frac{dS}{dt} = 0 \tag{B.55}$$

$$\frac{dR}{dt} = -R + 0.040 \cdot R \cdot \frac{X}{-0.950 \cdot R + \sqrt{0.200 \cdot R^2 + (-0.950 \cdot R + 0.315)^2.00 - 0.0600 \cdot R} + 0.315} + 0.010 \cdot X \tag{B.56}$$

$$\frac{dX}{dt} = S - 0.040 \cdot R \cdot \frac{X}{-0.950 \cdot R + \sqrt{0.200 \cdot R^2 + (-0.950 \cdot R + 0.315)^2.00 - 0.0600 \cdot R} + 0.315} - 0.010 \cdot X \tag{B.57}$$

## B.10.2 MM Governed

$$\frac{dS}{dt} = 0 \tag{B.58}$$

$$\frac{dE}{dt} = -\frac{R \cdot E \cdot (E_P + 0.05) - 0.30 \cdot (E + 0.05) \cdot E_P}{(E + 0.05) \cdot (E_P + 0.05)} \tag{B.59}$$

$$\frac{dE_P}{dt} = \frac{R \cdot E \cdot (E_P + 0.05) - 0.30 \cdot (E + 0.05) \cdot E_P}{(E + 0.05) \cdot (E_P + 0.05)} \tag{B.60}$$

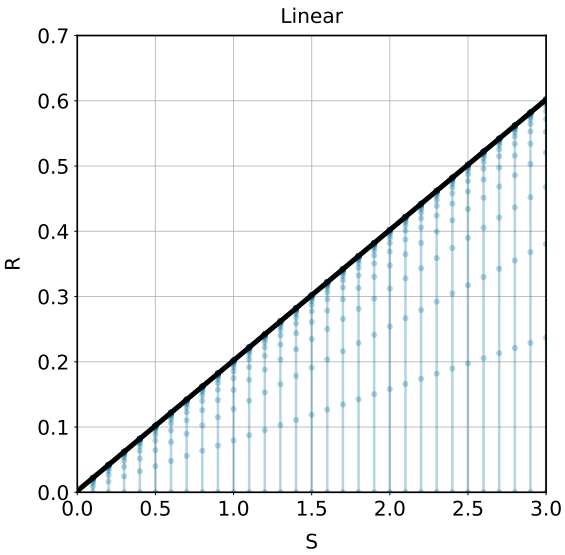
$$\frac{dR}{dt} = 0.400 \cdot X \cdot E_P - R + 0.010 \cdot X \tag{B.61}$$

$$\frac{dX}{dt} = -0.400 \cdot X \cdot E_P + S - 0.010 \cdot X \tag{B.62}$$

# Appendix C

# Figures and Plots

## C.1 (a) Linear Response




[H]

Figure C.1 Linear response. Signal-response curve in *black*, temporal progression in *blue*.

## C.2 (b) Hyperbolic Response

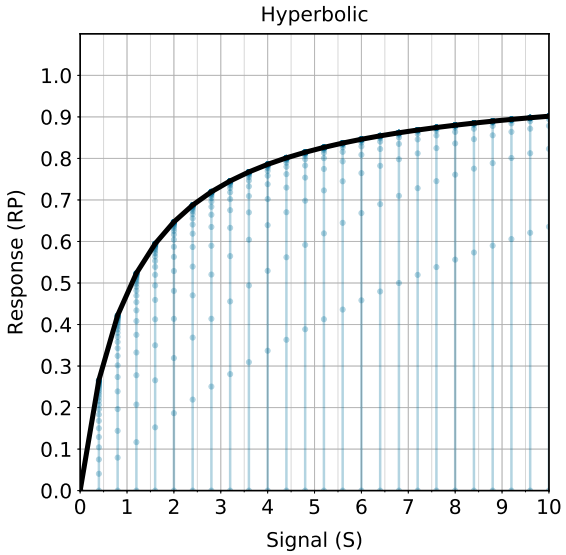



Figure C.2 Hyperbolic Response. Signal-response curve in *black*, temporal progression in *blue*.

## C.3 (c) Sigmoidal Response

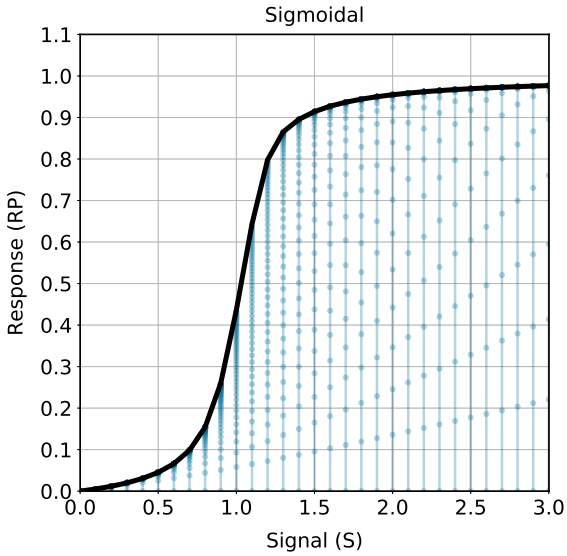



Figure C.3 Sigmoidal Response. Signal-response curve in *black*, temporal progression in *blue*.

## C.4 (d) Perfect Adaptation

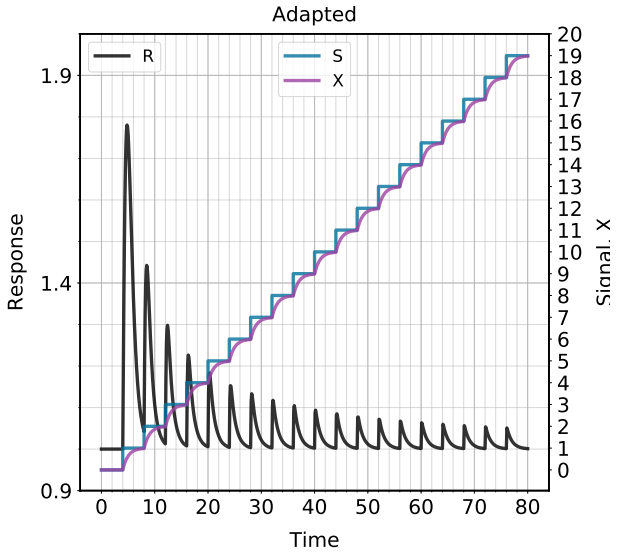



Figure C.4 Perfect Adaptation. Signal $S$ in *blue*, species $X$ in *purple*, species $R$ in *black*.

## C.5 (e) Mutual Activation

### C.5.1 GK Governed

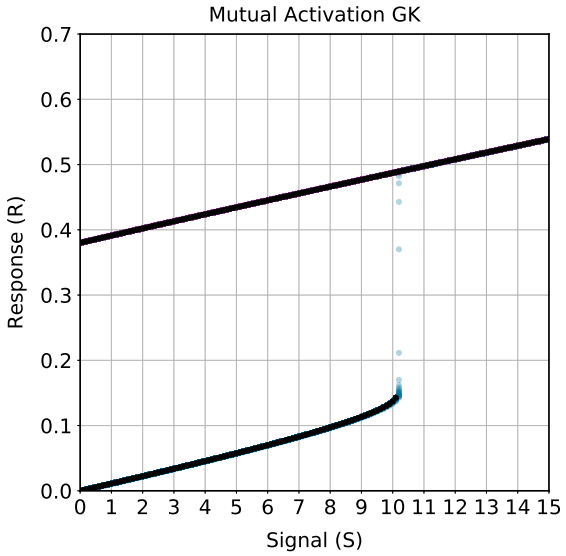



Figure C.5 Mutual Activation governed by GK function. Bi-stable signal-response curve in *black*, temporal progression in *blue*.

### C.5.2 MM Governed

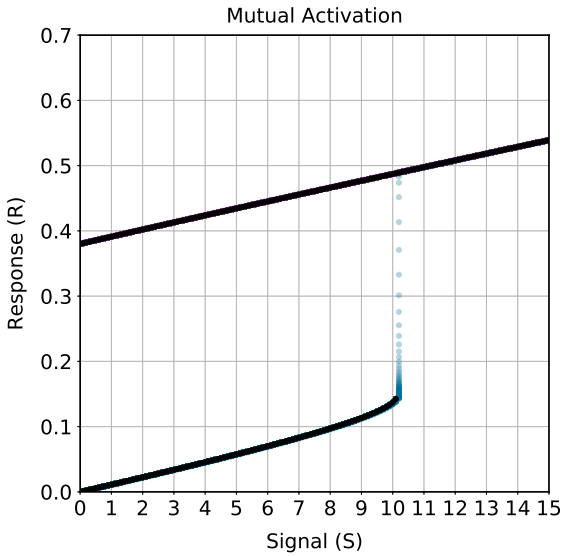



Figure C.6 Mutual Activation governed by MM kinetics. Bi-stable signal-response curve in *black*, temporal progression in *blue*.

## C.6 (f) Mutual Inhibition

### C.6.1 GK Governed

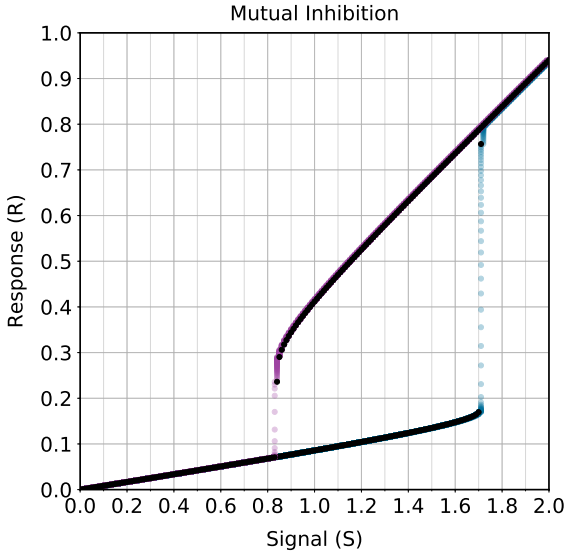



Figure C.7 Mutual Inhibition governed by GK function. Bi-stable signal-response curve in *black*, temporal progression in *blue* and *purple*.

### C.6.2 MM Governed

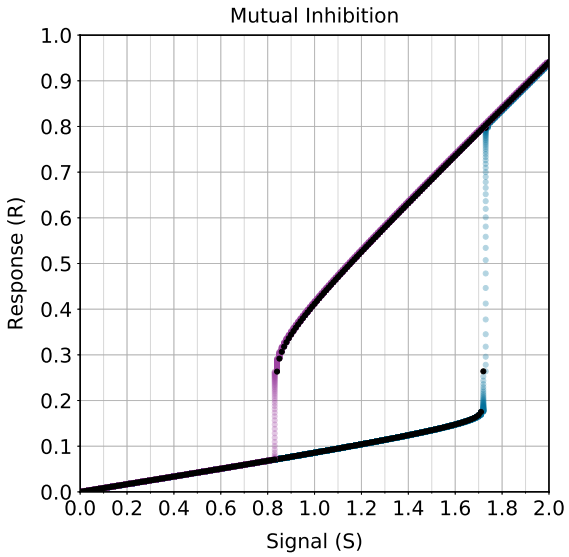



Figure C.8 Mutual Inhibition governed by GK function. Bi-stable signal-response curve in *black*, temporal progression in *blue* and *purple*.

## C.7 (g) Homeostasis

### C.7.1 GK Governed

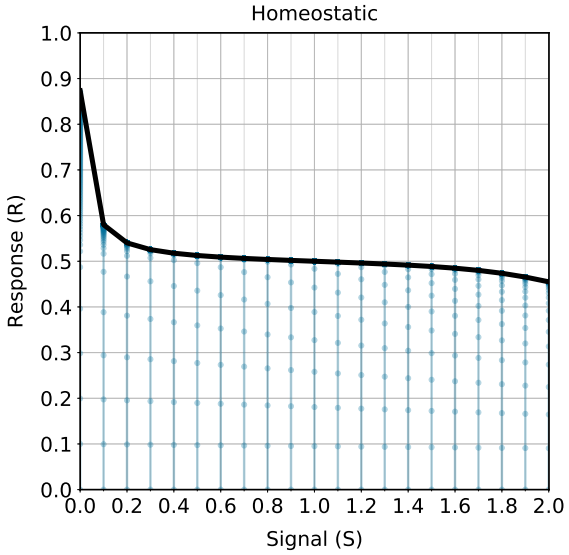



Figure C.9 Homeostasis governed by GK function.Signal-response curve in *black*, temporal progression in *blue*.

### C.7.2 MM Governed

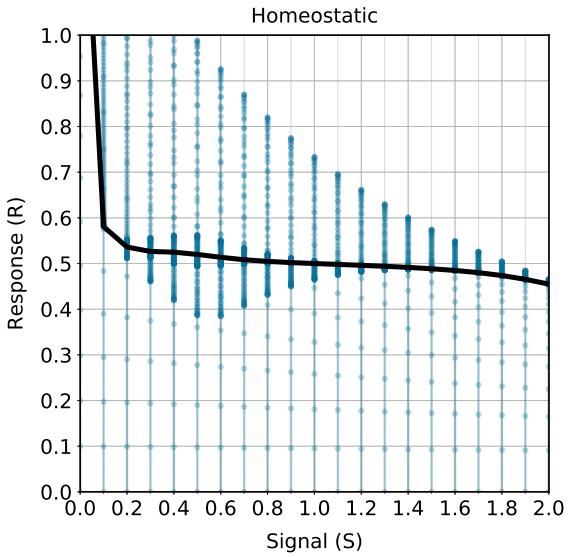



Figure C.10 Homeostasis governed by GK function.Signal-response curve in *black*, temporal progression in *blue*.

## C.8 (h) Negative-Feedback Oscillators

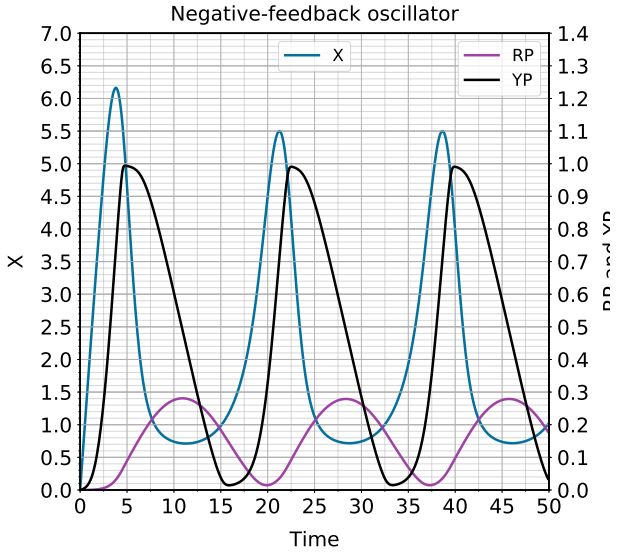



Figure C.11 Negative-Feedback Oscillator. Temporal progression of oscillating species. Species $X$ in *blue*, species $R_P$ in *pink*, species $Y_P$ in *black*.

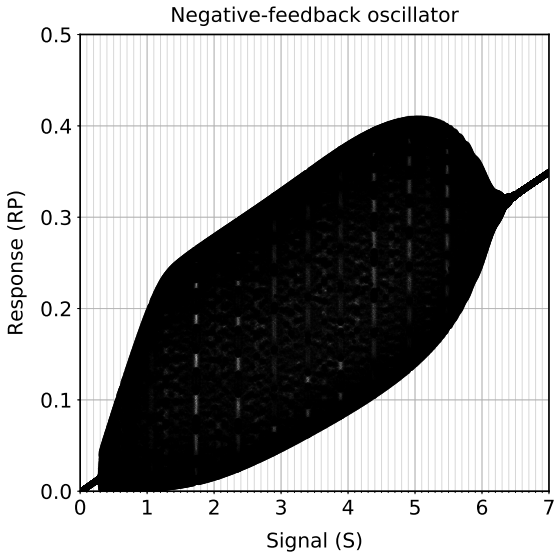



Figure C.12 Negative-Feedback Oscillator. Signal-response diagram showing minima and maxima values achieved by oscillations at specific $[S]$.

# C.9 (i) Activator-Inhibitor Oscillator

## C.9.1 GK Governed

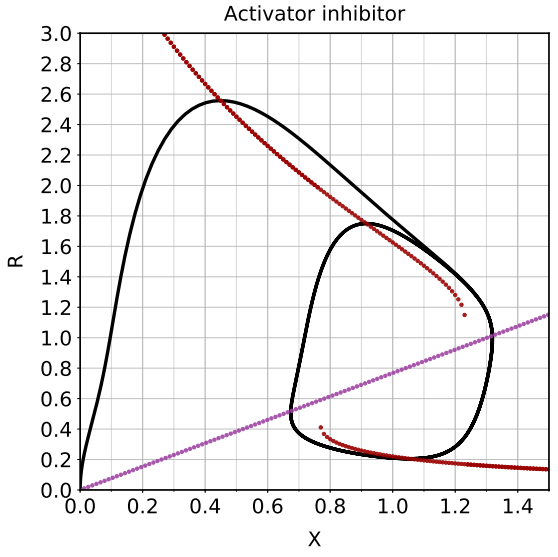



Figure C.13 Activator-Inhibitor Oscillator governed by GK function. Phase-plane portrait. $(X,R)$ pairs satisfying $\frac{dR}{dt} = 0$ in *red*; $(X,R)$ pairs satisfying $\frac{dX}{dt} = 0$ in *purple*; temporal oscillation of $(X,R)$ in *black*.

### C.9.2 MM Governed

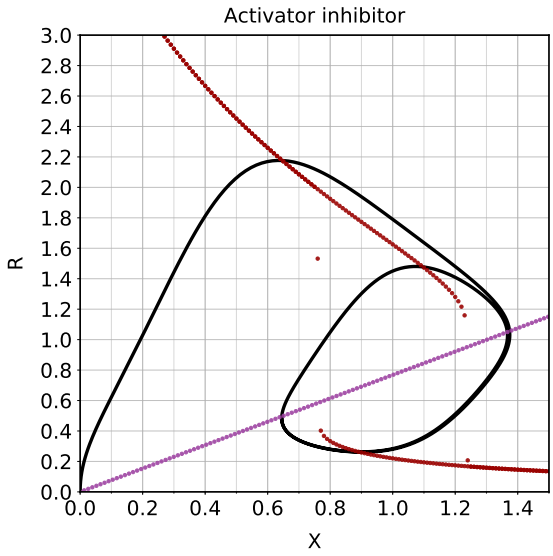



Figure C.14 Activator-Inhibitor Oscillator governed by MM kinetics. $(X,R)$ pairs satisfying $\frac{dR}{dt} = 0$ in *red*; $(X,R)$ pairs satisfying $\frac{dX}{dt} = 0$ in *purple*; temporal oscillation of $(X,R)$ in *black*.

### C.9.3 GK Governed Signal-Response

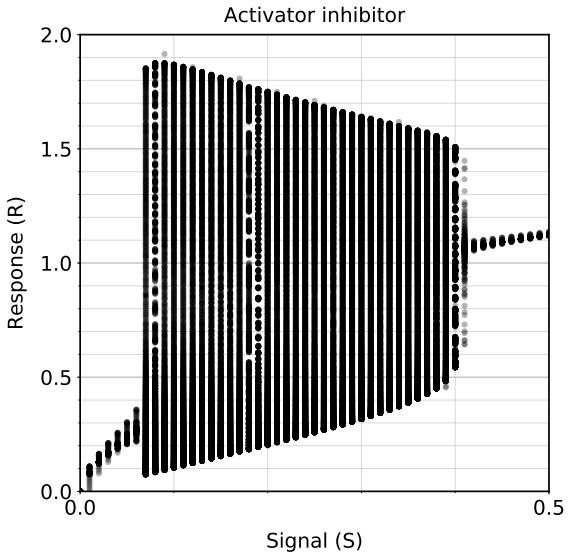



Figure C.15 Activator-Inhibitor oscillator governed by GK function. Signal-response diagram showing minima and maxima values achieved by oscillations at specific $[S]$.

### C.9.4 MM Governed Signal-Response

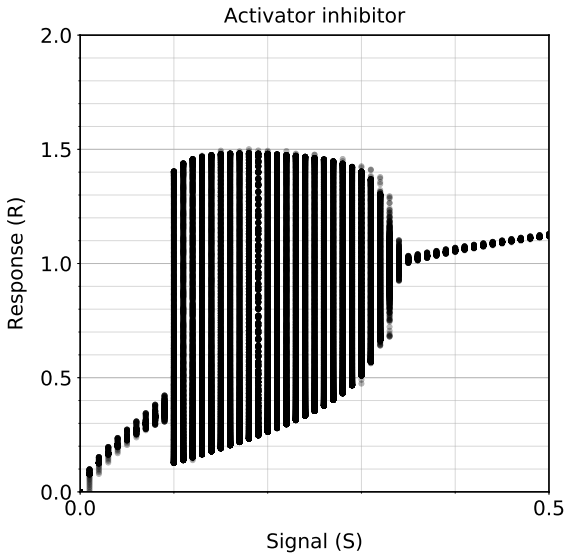



Figure C.16 Activator-Inhibitor oscillator governed by MM kinetics. Signal-response diagram showing minima and maxima values achieved by oscillations at specific $[S]$.

## C.10 (j) Substrate-Depletion Oscillator

### C.10.1 GK Governed

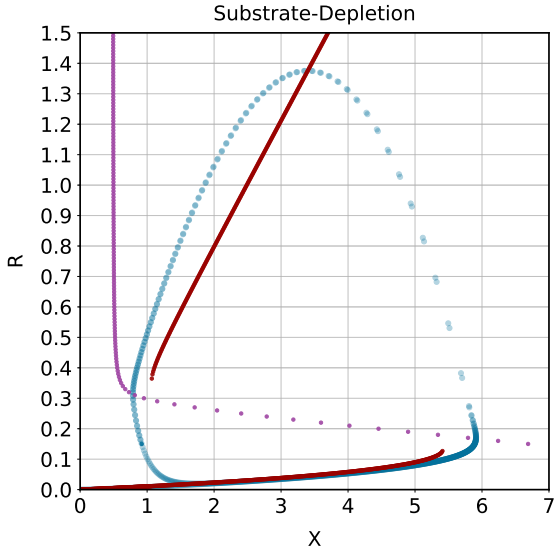



Figure C.17 Substrate-depletion oscillator governed by GK function. $(X,R)$ pairs satisfying $\frac{dR}{dt} = 0$ in *red*; $(X,R)$ pairs satisfying $\frac{dX}{dt} = 0$ in *purple*; temporal oscillation of $(X,R)$ in *blue*.

### C.10.2 MM Governed

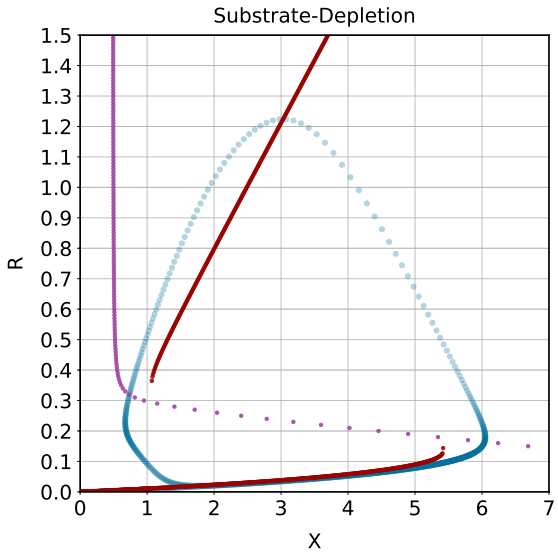



Figure C.18 Substrate-depletion oscillator governed by MM kinetics. $(X,R)$ pairs satisfying $\frac{dR}{dt} = 0$ in *red*; $(X,R)$ pairs satisfying $\frac{dX}{dt} = 0$ in *purple*; temporal oscillation of $(X,R)$ in *blue*.

### C.10.3 GK Governed Signal-Response

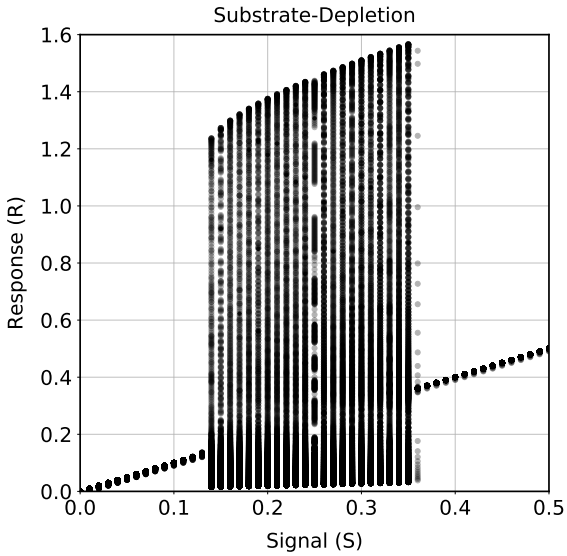



Figure C.19 Substrate-depletion oscillator governed by GK function. Signal-response diagram showing minima and maxima values achieved by oscillations at specific $[S]$.

### C.10.4 MM Governed Signal-Response

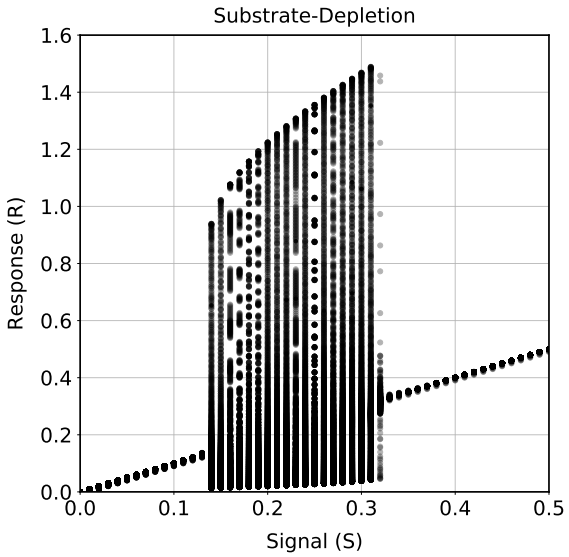



Figure C.20 Substrate-depletion oscillator governed by MM kinetics. Signal-response diagram showing minima and maxima values achieved by oscillations at specific $[S]$.

# Appendix D

# Additional Material

## D.1 cπ Example Code - *testEnzyme.cpi*

Code reproduced from the CPi-IDE GitHub repository by McCrae [55]:

```
-- Species (S) + Enzyme (E) <-> Complex -> Product (P)

species S(s) = s(x,y).(x.S(s) + y.P());
species E(e) = {a-u@0.5,a-t@1.0}
               e<u,t>.a.E(e);
species P() = tau<0.5>.0;

process Pi = [1.0] S(s) || [0.5] E(e) || [0] P()
           : {e-s@1.0};

-- NB: Complex = {a-u@0.5,a-t@1.0} a.E(e)|u.S(s)+t.P()
```

## D.2 Bond-Calculus Code

### D.2.1 Modeling with Dynamic Bonding

```
species E = e(l) -> ep@l -> E;
species S = s(l) -> (pp@l -> P + sp@l -> S);
species P = p -> 0;

affinity network M1(k1,m1,k2, k3) = {
  e, s at rate MA(k1);
  ep + sp at rate MA(m1);
  ep + pp at rate MA(k2);
  p at rate MA(k3);
```

```
}
process Pi = [0.5] E || [1.0] S || [0.0] P
             with network M1(1.0, 0.1, 0.5, 0.1);
```

### D.2.2 Modeling with Michaelis-Menten Kinetic Law

```
species E = e -> E;
species S = s -> P;
species P = p -> 0;

kinetic law MM(k,km; S,E) = (k*S*E) / (km + E);

affinity network M1(k1,km1,k2) = {
  s || e at rate MM(k1,km1);
  p at rate MA(k2);
}
process Pi = [0.5] E || [1.0] S || [0.0] P
             with network M1(1.0, 0.1, 0.1);
```

## D.3 LTL Syntax

$$\phi, \psi ::= \quad \text{Atom} \quad | \ \phi \wedge \psi \mid \phi \vee \psi \mid \phi \Longrightarrow \psi \mid \neg\phi \mid \tag{D.1}$$

$$\mathbf{X}\phi \quad | \ \phi \mathbf{U}_l \psi \mid \mathbf{F}_l \phi \mid \mathbf{G}_l \phi \tag{D.2}$$

$$\text{Atom} ::= \quad \top \quad | \ \bot \tag{D.3}$$

The first row represents the basic operations in propositional logic. LTL introduces 4 temporal operators, namely *next* ($\mathbf{X}\phi$), *until* ($\phi \mathbf{U}_l \psi$), *globally* ($\mathbf{G}_l \phi$) and *eventually* ($\mathbf{F}_l \phi$). They can be combined to form expressions such as "infinitely often proposition $\phi$ holds" ($\mathbf{G}_l \mathbf{F}_l \phi$) or "eventually proposition $\phi$ holds infinitely often" ($\mathbf{F}_l \mathbf{G}_l \phi$). The *next* operator allows to access a subsequent state. With the *until* operator we can express that a certain proposition ($\phi$) holds until another proposition ($\psi$) evaluates to true. True and false are denoted by $\top$ and $\bot$ respectively. The subscript $l$ denotes a time interval $[0, l)$ in which the respective operator performs the checking. Any expression which is not satisfied within that time interval evaluates to false.
Given a transition system $P$ and a linear time (LT) property $\gamma$, we formulate a *satisfaction relation* as $P \models \gamma$. Here, $\gamma$ could be any formula satisfying the syntax in Definition D.3 [7].

## D.4 Sigmoidal Response - ODE System Derivation

To devise the system of ODEs governing the sigmoidal signal-response motif we rely on the work by Wang [84]. The phosphorylation, referred to as forward reaction ($R \rightarrow R_P$) and dephosphorylation, or backward reaction ($R_P \rightarrow R$) can be expressed by the chemical Reactions D.4 and D.5 respectively. The indices $K$ indicate all components involved with the forward reaction, the kinase, and indices $P$ all components in the backward reaction, the phosphatase. However, $R_P$ still refers to the phosphorylated response of protein $R$.

$$S_K + R \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C_K \xrightarrow{k_2} S_K + R_P \tag{D.4}$$

$$S_P + R_P \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} C_P \xrightarrow{k_4} S_P + R \tag{D.5}$$

The forward reaction is expressed by the ODEs in Equations D.16, D.17, D.18, and D.19 while the ODEs for the backward reaction are Equations D.20, D.21, D.22, and D.23. The reaction constant $k_1$ and $k_{-1}$ describe the formation and decomposition of the temporary enzyme-substrate complex C, while $k_2$ indicates the rate at which C reacts into enzyme and the phosphorylated response. Analog to this, reaction constants $k_3$, $k_{-3}$, and $k_4$ follow the same purpose in the backward reaction. The equations are formed in a similar fashion to the ones in Sections 3.3.1 and 3.3.2.

If the concentration of the enzyme is much smaller than the concentration of the substrate ($S_T \ll R_T$), the quasi-steady state assumption by Briggs & Haldane can be applied [10]. It states that the concentration of the enzyme-substrate complex C remains constant throughout the reaction, hence $\frac{dC}{dt} = 0$. Applying this to both equations in Equation D.21, we obtain Equations D.11 and D.14 respectively for the forward and backward reactions.

$$\frac{dC_K}{dt} = -k_2 C_K - k_{-1} C_K + k_1 SR = 0 \tag{D.6}$$

$$\Longrightarrow C_K (k_2 + k_{-1}) = k_1 SR \tag{D.7}$$

$$\Longrightarrow C_K (k_2 + k_{-1}) = k_1 (S_{\text{total}} - C_K) R \tag{D.8}$$

$$\Longrightarrow C_K (k_2 + k_{-1} + k_1 R) = k_1 S_{\text{total}} R \tag{D.9}$$

$$\Longrightarrow C_K \left( \frac{k_2 + k_{-1}}{k_1} + R \right) = S_{\text{total}} R \tag{D.10}$$

$$\Longrightarrow C_K = \frac{S_{K,\text{total}} R}{\frac{k_2 + k_{-1}}{k_1} + R} \tag{D.11}$$

$$\Longrightarrow C_K = \frac{S_{K,\text{total}} R}{K_{m1} + R} \tag{D.12}$$

$$\frac{dC_P}{dt} = k_3 R_P S_P - k_{-3} C_P - k_4 C_P = 0 \tag{D.13}$$

$$\Longrightarrow C_P = \frac{S_{P,\text{total}} R_P}{\frac{k_4 + k_{-3}}{k_3} + R_P} \tag{D.14}$$

$$\Longrightarrow C_P = \frac{S_{P,\text{total}} R_P}{K_{m2} + R_P} \tag{D.15}$$

$$\frac{dR_P}{dt} = k_2 C_K \tag{D.16}$$

$$\frac{dC_K}{dt} = -k_2 C_K - k_{-1} C_K + k_2 S_K R \tag{D.17}$$

$$\frac{dS_K}{dt} = -k_1 S_K R + k_{-1} C_K + k_2 C_K \tag{D.18}$$

$$\frac{dR}{dt} = -k_1 S R + k_{-1} C_K \tag{D.19}$$

$$\frac{dR_P}{dt} = -k_3 R_P S_P + k_{-3} C_P \tag{D.20}$$

$$\frac{dC_P}{dt} = k_3 R_P S_P - k_{-3} C_P - k_4 C_P \tag{D.21}$$

$$\frac{dS_P}{dt} = k_4 C_P + k_{-3} C_P - k_3 R_P S_K \tag{D.22}$$

$$\frac{dR}{dt} = k_4 C_P \tag{D.23}$$

The terms in the denominator of Equations D.11 and D.14 can be rewritten as the MM constants $K_{m1} = \frac{k_2 + k_{-1}}{k_1}$ and $K_{m2} = \frac{k_4 + k_{-3}}{3}$ respectively.

**For the forward reaction** we now can insert Equation D.12 into Equation D.16. With the expression $R = R_T + R_P$ we can express the rate of $R_P$ in terms of the total concentration of the response and $R_P$ itself:

$$\frac{dR_P}{dt} = \frac{k_c S_{P,\text{total}} \left(R_T - R_P\right)}{K_{m1} + R_T - R_P} \tag{D.24}$$

**For the backward reaction** we can similarly express the rate of change of the concentration of species $R_P$. As pointed out by Wang, we note that Equation D.21 is the reverse of Equation D.20 with an added $k_4 C_P$ term. Hence we rewrite:

$$\frac{dR_P}{dt} = -\frac{dC_P}{dt} - k_4 C_P \tag{D.25}$$

$$\longrightarrow \frac{dR_P}{dt} = -k_4 C_P \tag{D.26}$$

$$\longrightarrow \frac{dR_P}{dt} = -\frac{k_4 S_{P,\text{total}} R_P}{K_{m2} + R_P} \tag{D.27}$$

Combining Equations D.24 and D.27 results in the governing ODE (Equation 3.6 for the concentration of $R_P$. It has to be noted, that this equation does not correspond with the one in [82]. Similar to the discussion in Section 3.3.2, the authors neglect the phosphatase enzyme $S_P$ (hence also $S_{P,\text{total}}$) in their derivation and assume that the MM kinetics drives the dephosphorylation

allosterically without any catalizing enzyme. In order to ease the qualitative evaluation in Chapter 5, we will follow the results presented in [82].

## D.5 Complete Table of Rate Constants

| System | Rate Constants | | | | | | | | | | | | Initial Concentrations | Time for Stabilization |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $k_0$ | $k_{0'}$ | $k_1$ | $k_2$ | $k_{2'}$ | $k_3$ | $k_4$ | $k_5$ | $k_6$ | $J_3 = J_4$ | $K_{m1} = K_{m2}$ | $K_{m3} = K_{m4} = K_{m5} = K_{m6}$ | | |
| Linear (synthesis and degeneration) | 0.01 | - | 1 | 5 | - | - | - | - | - | - | - | - | $[1]A \parallel [1]S \parallel [0]R$ | 1.2 |
| Hyperbolic (phsophorylation & dephsophorylation) | - | - | 1 | 1 | - | - | - | - | - | - | - | - | $[1]S \parallel [1]R \parallel [0]RP$ | 3 |
| Sigmoidal (phsophorylation & dephsophorylation) | - | - | 1 | 1 | - | - | - | - | - | - | 0.05 | - | $[1]S \parallel [1]R \parallel [0]RP$ | 20 |
| Perfect adaptation | - | - | 2 | 2 | - | 1 | 1 | - | - | - | - | - | $[1]S \parallel [1]R \parallel [1]X$ | - |
| Mutual activation GK | 0.4 | - | 0.01 | 1 | - | 1 | 0.2 | - | - | 0.01 | - | - | $[1]S \parallel [0]R$ | 35 |
| Mutual activation MM | 0.4 | - | 0.01 | 1 | - | 1 | 0.2 | - | - | 0.01 | - | - | $[1]S \parallel [0]R \parallel [1]E \parallel [0]EP$ | 125 |
| Mutual Inhibition GK | 0 | - | 0.05 | 0.1 | 0.5 | 0.2 | 1 | - | - | 0.05 | - | - | $[1]A \parallel [1]S \parallel [0]R$ | 80 |
| Mutual Inhibition MM | 0 | - | 0.05 | 0.1 | 0.5 | 0.2 | 1 | - | - | 0.05 | - | - | $[1]A \parallel [1]S \parallel [0]R \parallel [0]E \parallel [1]EP$ | 100 |
| Homeostatic GK | 1 | - | - | 1 | - | 0.5 | 1 | - | - | 0.01 | - | - | $[1]S \parallel [0]R$ | 1 |
| Homeostatic MM | 1 | - | - | 1 | - | 0.5 | 1 | - | - | 0.01 | - | - | $[1]S \parallel [0]R \parallel [1]E \parallel [0]EP$ | 14 |
| Negative-feedback osc. | 0 | - | 1 | 0.01 | 10 | 0.1 | 0.2 | 0.1 | 0.05 | - | - | 0.01 | $[1]A \parallel [2]S \parallel [1]R \parallel [0]RP \parallel [0]X \parallel [1]Y \parallel [0]YP$ | - |
| Activator-inhibitor osc. GK | 4 | - | 1 | 1 | 1 | 1 | 1 | 0.1 | 0.075 | 0.3 | - | - | $[0.2]S \parallel [0]R \parallel [0]X$ | - |
| Activator-inhibitor osc. MM | 4 | - | 1 | 1 | 1 | 1 | 1 | 0.1 | 0.075 | 0.3 | - | - | $[0.2]S \parallel [0]R \parallel []RP \parallel [0]X \parallel [1]E \parallel [0]EP$ | - |
| Substrate-depletion osc. GK | 0.4 | 0.01 | 1 | 1 | - | 1 | 0.3 | - | - | 0.05 | - | - | $[0.2]S \parallel [0]R \parallel [0]X$ | - |
| Substrate-depletion osc. MM | 0.4 | 0.01 | 1 | 1 | - | 1 | 0.3 | - | - | 0.05 | - | - | $[]A \parallel [0.2]S \parallel [0]R \parallel [0]X \parallel [1]E \parallel [0]EP$ | - |

Table D.1 Aggregation of rate constants, initial concentrations, and time needed for the systems to stabilize. Rate constants modified from [82] with permission - errata included. Initial concentrations obtained from [82] Figure 2 with permission and [29, 28, 27, 26, 25, 24, 23]. Time for stabilization are extracted from our experiments and represent an overestimate. Not applicable to oscillators.