NTNU

Norwegian University of
Science and Technology

Modelling and Analysis of a Synthetic Bistable Genetic Switch

Sigurd Hagen Johansen

Biotechnology
Submission date: May 2011
Supervisor: Eivind Almaas, IBT

Norwegian University of Science and Technology
Department of Biotechnology

Acknowledgements

This work was conducted at the Department of Biotechnology at the Norwegian University of Science and Technology (NTNU) in the period August 2010 - May 2011.

First of all, I would like to thank my supervisor Professor Eivind Almaas for inspiring guidance and for convincing me to choose this very interesting field as the topic of my Master thesis.

Following, I would like to thank Rahmi Lale, PhD, for insightful comments regarding my thesis and planning of experimental conduction of the system described.

A great thanks to my housemates Kaan Yabar and Kristian Jenssen for a good introduction to programming and helpful comments while revising my thesis. Also thanks to Ida Maria Evensen for constructive comments.

Additionally, I would like to thank all my friends and family for giving me moral support and believing in me. Finally, I would like to thank Eli for being there for me whenever needed.

Trondheim, May 8, 2011

Sigurd Hagen Joansen

Abstract

In the field of systems and synthetic biology there has been an increasing interest for the use of genetic circuits during the last decade. Several circuits have been successfully put together, many of which were based on models. During this thesis a model for a toggle switch was analysed both deterministically and stochastically.

The HOM2-circuit approximation for a bistable tuneable switch from Ghim and Almaas (2009) was re-derived in order to make it asymmetrical. Deterministic analysis was conducted yielding stability diagrams, describing the phase plane showing bistability for the genetic switch. Furthermore, stochastic simulations of the approximation were conducted. This gave a somewhat narrower bistable area than the deterministic analysis, possibly due to the nature of saddle-node bifurcations. Parameter values for a switch based on experiments were estimated for the approximation, and these were used in a stochastic simulation. The result from this simulation was in correspondence with the deterministic analysis. A stochastic simulation of the complete circuit was conducted based on parameter values found in literature. For this simulation bistability was not shown.

In order to further explore the circuit, and validity of the approximation, experimental investigation is needed. This has been planned together with Rahmi Lale, PhD, and Professor Eivind Almaas at the Department of Biotechnology NTNU.

Abbreviations

aa Amino acid

bp Base pairs

DNA Deoxyribonucleic acid

FPP Farnesyl pyrophosphate

iGEM International Genetically Engineered Machine

MIT Massachusetts Institute of Technology

mRNA messenger RNA

ODE Ordinary differential equation

PySCeS Python Simulator for Cellular Systems

RBS Ribosomal binding site

RNA Ribonucleic acid

RNAp RNA polymerase

SBML Systems Biology Markup Language

TetR Tet repressor

Contents

1 Introduction

1.1 Thesis objectives

During the 90s the high throughput technologies in genetics laid the foundation for the fields of systems and synthetic biology, and in year 2000 there were published two genetic circuits. Among these was the genetic toggle switch made by Gardner *et al.* (2000) consisting of two repressors able to control each others expression [1].

In Ghim and Almaas (2009) there was introduced a model for a symmetric genetic switch, named HOM2, having very similar design as the one produced by Gardner *et al.* (2000). In the model they included dimerisation of the repressor before cooperative binding at two operator sites within the promoter of the opposite gene. The model was made by starting out with a total of 40 reactions and 20 coupled ordinary differential equations (ODEs) and then making an approximation reducing the system to a set of only two ODEs. Simulations of this approximation showed good agreement compared to simulations for the entire system. In the approximation a few important parameters were introduced which can be readily tuned by genetic modification, such as base gene expression and promoter leakage [2]. However, the model assumed symmetrical values for the parameters for both promoter and repressor pairs, somewhat restricting the ease in which the model can be related to an acutal circuit. During this thesis the circuit will be made asymmetrical, allowing the different promoters to have different values to their parameters. The parameter space allowing for bistability, where the circuit will function as a switch, will be predicted using deterministic analysis, followed by stochastic simulations of both the approximation and the complete circuit, having estimated parameters for both.

The use of genetic switches in genetic circuits have been proposed in different systems like for instance biosensors [3]. In order to put these together and achieve the desired functionality they need to have predictable dynamics. This is often done through modelling, although a lack of kinetic parameters highly restricts the use of such models in addition to a high need for computational power. If the approximation is shown to be predictive it could become a base for further development of genetic circuits consisting of genetic switches, needing only to estimate a few parameters instead of about 40 and also reducing the amount of computational power. Additionally, the approach used for making the approximation could also be used for other components in a genetic circuit.

The following parts will introduce the fields of systems and synthetic biology, give a general description of basic modelling approaches, some more

details regarding how genetic circuits are put together, description of robustness and stochasticity in biological systems and additionally some of the basics to mathematics behind the analysis will be described.

1.2 Systems and synthetic biology

The scientific discipline of systems biology is described as a merging of the growing array of biological data with sophisticated engineering applications and analysis methods [4]. The systems approach provide descriptions of organisms as integrated and interacting network of genes, as opposed to describing the isolated genetic elements and their products alone [5]. From a biochemical perspective, it is described as a study of the many biochemical changes that occur in a cell as a function of genetic or environmental stress [6]. For the purpose of this thesis the term *systems biology* will be used to describe *biological research focusing on the interactions between different biological components and how these interactions make a complete system.* This research aims to generate predictive knowledge about the system to be applied in both research and industry.

Attempts of understanding the nature of biology in terms of mathematics was made already in 1948 by Norbert Wiener, while he at the same time laid the foundations for the field known as cybernetics [7]. A little more than a decade later, Jacques Monod and François Jacob suggested that the total level of enzymes in a cell was regulated by feedback mechanisms on the transcription of elements resident at the level of genes [8]. This was one of the reasons for that they, together with André Lwoff, in 1965 was awarded the Nobel Prize in Physiology or Medicine, "for their discoveries concerning genetic control of enzyme and virus synthesis" [9]. Further computational studies of biological systems was performed throughout the 1970s, for instance by making biologically descriptive models by the use of logical (Boolean) algebra [10] and the nonlinearity of the chemical reactions within a cell [11]. Although descriptive, the amounts of data needed to verify these models somewhat restricted their use.

During the 90s high-throughput technologies emerged, allowing the genetic interactions of many thousands of genes to be investigated at the same time [4]. These technologies enabled the researchers to experimentally check the agreement between computational models and the real biological systems [12].The high-throughput technologies have caused enormous amounts of data to be produced, as illustrated by the number of sequenced genomes and base pairs of DNA stored in GenBank as shown in Fig. 1 [13]. These data needs to be analysed in order to have any significance, and this can not be accomplished by the reductionist approach that was very popular before

the emerging of the high-throughput methods. In order to get useful information from these huge amounts of data there became a need to make use of an approach focusing on the system as a whole. In this approach one gives less attention to the individual parts themselves and rather focus on how their interactions formed a complete functional biological unit (this biological unit may be anything from a single metabolic pathway to a multicellular organism) [14]. As the gene copy numbers are normally small, the cells are vulnerable to stochastic effects affecting their genes, potentially this could completely alter the phenotypic behaviour of a cell. This has been theoretically proposed previously, although it has been difficult to prove experimentally. Due to recent technological advancements in developing techniques for single-cell experiments the stochastic effects have been validated and have put down a platform giving insight to what processes that give rise to these effects [15, 16]. The topic of stochasticity in gene expression is discussed more carefully in Sec. 1.6.

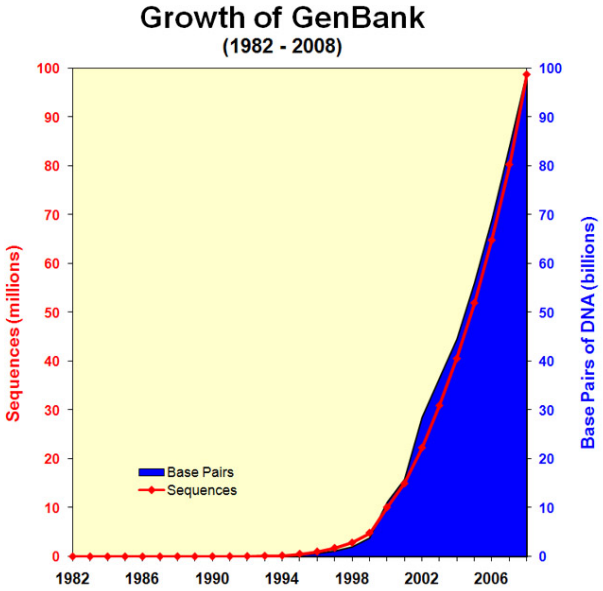


Figure 1: **The growth of GenBank from 1982–2008.** The number of sequences and base pairs obtained in GenBank during the period 1982–2008 are shown in the graph by a red dotted line and a blue bar [13]

The genetic toggle switch and the repressilator, see Sec. 1.4, were published in the same volume of Nature in 2000, establishing functional genetic circuits with properties reflecting computational models that were made in advance. These computational models predicted the circuit dynamics before

the actual circuits were built [1, 17]. The modelling that was performed in the making of the genetic toggle switch and the repressilator was carried out using traditional methods from engineering and their tools, like Matlab (Mathworks). During the rise of systems biology there have been developed a myriad of programs that are specifically designed to model and analyze biological systems. Examples include Dizzy for making stochastic simulations of genetic regulatory networks [18], Cytoscape for visualization and analysis of complex networks [19] and the cellular modelling software PySCeS [20]. To enable the researchers to share their models and cooperate, the Systems Biology Markup Language (SBML) was established in 2003 as a common language all platforms could be translated to [21].

The modelling of systems can be used to generate synthetic circuits from novel genes and transcription factors as in the repressilator, consisting of three repressors acting on each others expression in a cyclic fashion, as described in Sec. 1.4 [17]. These genetic circuits can be called synthetic in the sense that they are not occurring in nature, and the creation of such genetic circuits is a part of the field called synthetic biology. *Synthetic biology is the creation of novel biological systems based on principles from engineering and components from biology.* A defining goal for this new field appears to be the generation of new genetic systems based upon computational modelling [22]. The biological system that is being built can be in any scale, from a simple genetic circuit to a whole multicellular organism. This is usually accomplished by making use of rational modifications to the genetic material. Compared to classical biotechnology the advance lies in the use of the engineering methodology allowing for rational modifications to complex systems with effects elucidated before the system is put together [23, 24, 25].

An illustration of the power and accessibility of synthetic biology is the creation of *Escherichia coli* cells that are able to process images. This was performed by the introduction of three genes from different origins put together with the existing signalling cascades of the host *E. coli*. Bacteria expressing these genes were then enabled to produce pigments if not exposed to light [26]. It has been said that an ultimate goal for synthetic biology would be to create a living cell from only synthetic components [23]. A major leap towards that goal was made by J. Craig Venter and his team in 2010, with the synthetic production of an entire functional genome that was transplanted into restriction-minus *Mycoplasma capricolum* recipient cell [27].

In order to complete the project Venter and his co-workers had to engineer an entire genome. This has, in close resemblance to the sequencing of genomes, become significantly more feasible through techniques developed in the later years. The development of synthesis of synthetic genes and genomes

has gone through several steps from producing only short oligo-sequences by means of organic synthesis to synthesizing entire genes by biochemical assistance. Recent improvements have led to the use of *in vivo* synthesis of DNA, enabling synthesis of complete genomes. These advances have made the production cost of synthetic DNA drop significantly. This reduction in sequence cost may prove crucial in verifying different models of genetic circuits and help understanding the basics of life [28, 29, 30].

One of the most famous contributions from synthetic biology to medicine so far is the production of the antimalarial drug precursor artemisinic acid in *Saccharomyces cerevisiae*. The summarized modifications that were made to the metabolic pathways of *S. cerevisiae* to produce the artemisinin drug precursor artemisinic acid are depicted in Fig. 2 [31]. Although this undoubtedly was a great achievement for the field, the amount of work it took to get there provides a glimpse of the complexity of such tasks. Jay Keasling, the leader of the group who made this strain, has made a rough estimate that 150 person-years have been spent in the making of the final pathway.

There are great expectations related to synthetic biology, although some major limitations still remain. One of them is concerning that many of the parts or modules in use are not defined sufficiently to perform as they are supposed to. It seems like many of the elements that are used are not characterized well enough to the parameters needed for accurate modelling of for instance a genetic circuit [32, 33]. Host incompatibility is also named as a challenge where for instance codon-bias may affect the success rate for having a functional genetic element in a different species than its origin [32]. Additionally, the design of genetic circuits based on classical engineering principles may not be compatible with the fact that the cells are self replicating, making evolution of implemented sequences to a potential challenge [33]. Another hindrance is that parts put together may not function as one would expect. This problem might be reduced by modelling a system before putting it together [32]. Further problems are caused by the enormous complexity of biological systems. Labour-intensive work is required for elucidating how genes, proteins and metabolites affect each other in order to make precise and effective changes to the systems. Finally, the process of making the cells behave as expected is impeded by the cells vulnerability to stochastic fluctuations [32, 33]. All of these challenges need to be addressed, and using computational tools of systems biology could be at least part of the solution for some of them.

Important aspects that could help the development of synthetic biology include expanding the available toolkit for synthetic biology and standardizing it, modelling and fine-tuning the properties of the synthetic gene networks, development of probes/tags to quantify the behaviour of the synthetic

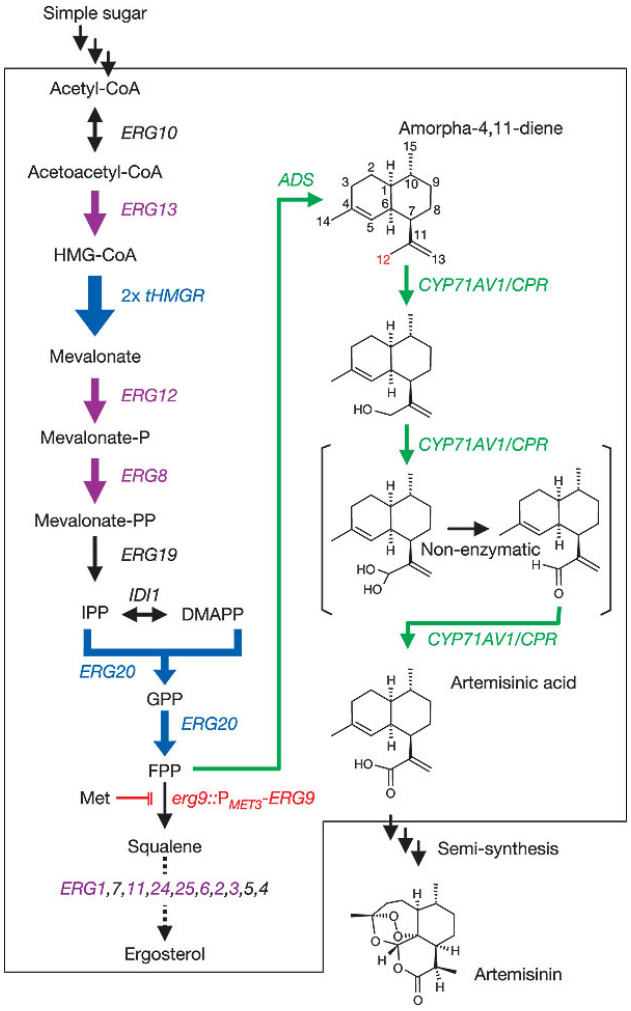


Figure 2: **The engineered pathway in *Saccharomyces cerevisiae* to produce artemisinic acid, a precursor to the antimalarial drug artemisinin.** The genes from the mevalonate pathway that are directly upregulated are shown in blue, the ones that are indirectly upregulated are shown in purple and the downregulated ones are shown in red. Genes that are shown in green were introduced to *S. cerevisiae* from *Artemisia annua* L. and constitute the biochemical pathway from farnesyl pyrophosphate (FPP) to artemisinic acid. The artemisinic acid is efficiently transported out of the cell, and can be purified from the culture. The conversion of arteminisic acid to arteminisin is done by known high yielding chemical reactions [31].

network and creating test platforms for characterizing the interactions within the network. Furthermore, decoupling of the field's processes could show useful and proving to standardize the parts even further. In this sense decoupling involves the separation of different steps in the process of creating synthetic systems, having different groups or companies specializing on one field, for instance the production of DNA sequences could be one such field [33, 3, 34].

For further development of the field it seems necessary to produce a strong educational system for synthetic biology. In 2003 the first International Genetically Engineered Machine (iGEM) competition was arranged at Massachusetts Institute of Technology (MIT). This competition has grown a lot from having only participants from MIT, to expanding having 155 teams participating in 2011. During the iGEM undergraduate students are given standardized parts from the Registry of Standard Biological Parts to develop novel biological devices or parts. These are supposed to be stored in the registry and be freely accessible for anyone afterwards. The 2010 competition winners, Slovenia, were able to create a repressilator using zinc-finger repressors. Other examples from 2010 include the Kyoto team's E.coli pen, which make use of fluorescent proteins produced from H_2O_2 inducible proteins to make colourful drawings. The competition demonstrates the accessibility of synthetic biology by having undergraduate students able to produce completely new devices from standardized parts, and further establishes the standard which is produced by The Registry of Standard Biological Parts [34, 35, 36, 37]

The field of synthetic biology has many potential applications, ranging from environmental and medical purposes, food and pharmaceutical industry, to the production of biomaterials and biosensing using bioreporters [34, 25, 38]. The possible environmental benefits from synthetic biology are vast. For instance, chemical synthesis may be performed with a much lower energy requirement, leading to a more sustainable synthesis industry. Other examples include biodegradation of waste products and creation of biodegradable plastics and biofuels [34, 25]. Bioreporters have been made to detect both heavy metals and organic compounds that are potentially harmful or toxic for the environment and humans. For improving the generation of new and more efficient bioreporters there is a need for more streamlined production, mechanisms to exploit noise, enhancement of the properties of the regulatory cascades that enhance the signal and tune the strength of the reporter signal. Even though these bioreporter assays prove to be both efficient and cheap there are legislations in many countries restricting their use because of their synthetic nature [39]. Medical purposes include both virus detection and vaccine development. The food industry could benefit from synthetic biology through the production of important metabolites

and nutrients, development of better preservatives and decrease the waste production from the food industry. Biosensors could also be used in food industry for detection of molecules giving rise to smell or taste in food products, may it be to control that the good taste is still preserved or that the degradation of the food compounds have not started [34, 25]. Analysis tools have been utilized to explore expression patterns of tomatoes when comparing metabolic patterns of tomatoes overexpressing the *Psy*-1 gene compared to the wild type, and its effect on carotenoid levels in the tomatoes. Results of such system analysis together with modelling of the metabolic pathways may help set future directions for where in the pathway the most effective changes can be done to improve productivity of important nutrients, laying the foundations for synthetic biology [40].

1.3 Modelling approaches in systems biology

When modelling genetic circuits there are three common approaches; logical (Boolean), continuous and stochastic modelling [41].

Boolean descriptions of genetic regulatory networks are constructed of discrete entities which are either on (1) or off (0). Predictions can be made to such networks by the use of well developed analysis techniques for discrete mathematics [41]. This type of logical analysis has been employed for decades already, as mentioned previously. Although seemingly simple, it can still provide useful information and is highly applied by systems biologists around the world. For instance Réka Albert and her group at Pennsylvania State University have recently published Boolean descriptions of G-protein action in plants based on transcriptome data [42] and how segment polarity genes affect the development of *Drosophila* segmentation [43]. These Boolean models are well suited for making large-scale descriptions as they are less computational costly than more complex modelling. Although, the Boolean models can only give qualitative information about a circuit, and for smaller systems continous models are often preferred as they can give more accurate predictions [41].

There are several different continuous modelling approaches including linear models and models of transcription factor activity, although the most common apporoach for continuous modelling is by the use of ordinary diffrential equations (ODEs). By describing the systems in a continous manner the rate constants and concentration of all the species involved in each reaction describing the circuit are accounted for. The high usage of detail enables the models to be very accurate in their predictions giving good quantitative predictions as an addition compared to Boolean models. ODEs can be used for calculating steady state solutions for the system in hand and their

corresponding stabilities, as described in Sec. 1.7. Examples for different genetic circuits and how to model them can be found in [44]. Although the predictions will seem accurate and can give good results compared to experimental results, there are still areas that need improvement. As the genetic elements involved in genetic circuits may be scarce in numbers (low copy numbers), the circuits may be highly affected by stochastic mechanisms and the determinism produced by the continuous models may not be sufficiently descriptive [41].

When the stochastic effects become significant a stochastic modelling approach may be able to capture a good picture of the behaviour of the circuit. Good examples of descriptions of genetic regultatory circuits described by stochastic models are the development of the sea urchin [45] and of the lambda phage [46]. Stochastic models are also called single-molecule level models as they take the fluctuating concentrations of single molecules into account when describing a circuit. The stochastic models are built up much like the ODEs but instead of a reaction rate they make use of a reaction probability. The system can then be run with a stochastic simulator, like Dizzy [18], using algorithms made for stochastic simulation of coupled chemical reactions like the Gillespie Direct [47] or Gibson-Bruck algorithms [48].

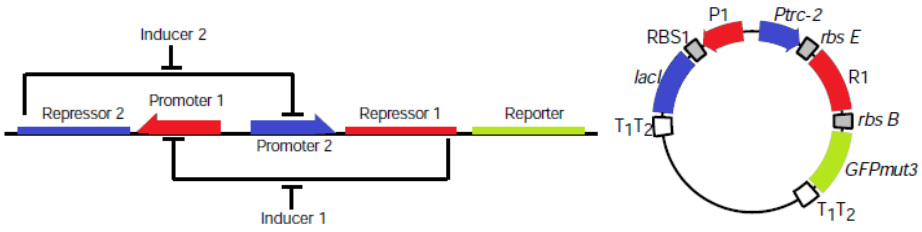
In summary the choise of what modelling appproach one wants to use depends on what is being modelled. If the system is big and complex and mostly qualitative information is required the ideal modelling approach would probably be Boolean. If the system is of intermediate size and the amount of quantitative information is higher one would prefer a continous representation of the system, possibly using ODEs to describe the system. If the system is relatively small it could be possible to perform single-molecule level simulations of the system taking stochasticity into account, giving highly accurate predictions of the system [41].

1.4 Genetic circuits

As synthetic biology is a merging between biology and engineering, so is genetic circuits a merging between genetics and electrical engineering. Electrical circuits are based upon mathematical models and so are genetic circuits, and many of the techniques in predicting outcomes of genetic circuits are directly derived from electric circuits. Electric circuits often contain modular parts such as switches and oscillators, which have strong resemblance to the two first published genetic circuits, a genetic toggle switch and the repressilator respectively.

The genetic toggle switch consists of two mutually repressible promoters, as schematically illustrated in Fig. 3a. In the experiment the LacI-repressor

was used as Repressor 2, repressing the promoter *Ptrc-2* being inducible by IPTG working as Inducer 2. Promoter 2 encodes either a heat inducible cI repressor or anhydortetracycline (aTc) inducible tetR repressor that will repress Promoter 1. If there is expression from Promoter 2 (the *Ptrc-2* promoter) there will be expression of a reporter protein, in this case in the form of the *GFPmut3* gene. The vector design used by Gardner *et al.* is shown in Fig. 3b. By using the construct design in *E. coli* strain JM2.300, they created a genetic circuit with two separate stable expression states (bistable) in which could be switched between by adding an inducer (chemical or physical) to the medium [1].



(a) (b)

Figure 3: **A genetic toggle switch.** (a) A schematic presentation of the genetic toggle switch made by Gardner *et al.* Promoter 1 is repressed by the Inducer 1 inducible Repressor 1. Promoter 2 is repressed by the Inducer 2 inducible Repressor 2. (b) The vector design applied for demonstration of a genetic toggle switch. Both figures are adapted from [1].

The repressilator was constructed by Elowitz and Leibler as a synthetic oscillatory network of trancriptional regulators, a biological oscillator. Here it was used three repressors acting in sequence on each other, as depiced in Fig. 4a. If it is being transcribed from the $P_L tet01$ promoter in the beginning there will be produced λ cI repressor and reporter protein. The cI repressor will repress the λ P_R promoter, and therefore there will be no lacI repressor produced and the RNA polymerase will transcribe from the $P_L lac01$ promoter, and the tetR repressor will be produced. This repressor will repress both $P_L tet01$ promoters stopping the production of the reporter protein and the λ cI repressor. This will leave open the λ P_R promoter and there will be produced lacI repressors. This will stop the production of the tetR repressor, and thereby rendering the transcription from the $P_L tet01$ promoters again producing the reporter protein. This cyclic fashion of repression gives the circuit an oscillating behaviour, as the fluorescence density graph in Fig. 4b illustrates. The two plasmids illustrated in Fig. 4c was contained by a culture

of *E. coli* strain MC4100 to produce the oscillating behaviour.

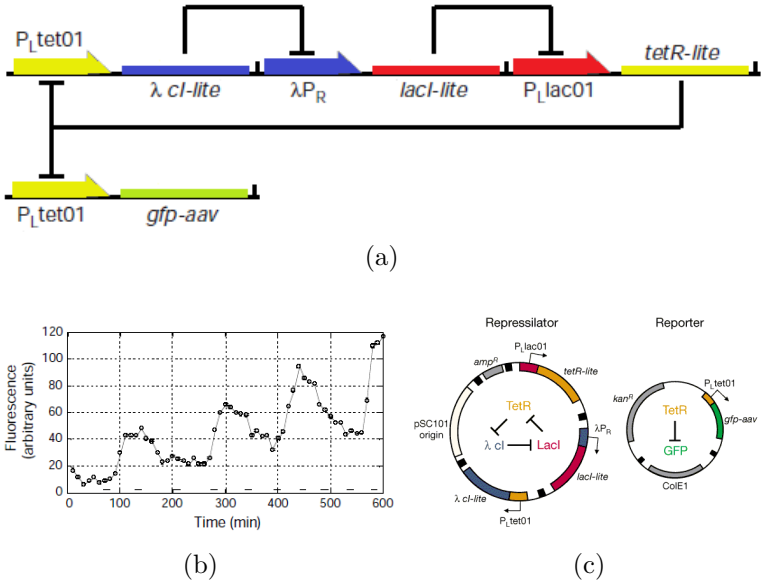


Figure 4: **The repressilator.** (a) A schematic representation of the repressilator. One of the two $P_L tet01$ promoters encodes the λ *cI-lite* gene. The λ cI-lite repressor can repress the λ P_R promoter which encodes the *lacI-lite* gene. The lacI-lite repressor can repress the $P_L lac01$ promoter which encodes the *tetR-lite* gene. The tetR-lite repressor can repress both the $P_L tet01$ promoters and thereby the expression of the reporter gene *gfp-aav*. This repression will occur in a cyclic fashion creating oscillating patterns of reporter gene readouts. The -lite notation on the repressors refer to the attachment of LVA-tails to the repressors to give rapid degradation. Adapted from [17]. (b) The repressilator produced the fluorescence plotted in the graph. Fluorescence is produced by the GFP-aav protein. From [17]. (c) The vetor design applied for the creation of the oscillating repressilator and reporter construct. From [17].

The inspiration from electronic circuits has continued in the last decade and some of the circuits can be described as digital logic evaluators [49, 50]. There has also been produced a genetic bandpass filter, by the combination of lowpass and highpass filters in series [51], more oscillators [22] and several other electronic circuit inspired genetic circuits. In addition one can imagine getting these working as modular parts generating bigger systems, as described by Lu *et al.* (2009) [3].

1.5 Robustness in biological systems

The term robustness is used differently in litterature, but for the purpose of systems biology it was defined by Hiroaki Kitano as; "Robustness is a property that allows a system to mantain its functions despite external and internal pertubations" [52].

Robustness is a feature often observed when studying biological systems. Observable phenomena that characterize such robustness is adaptability, insensitivity to parameter changes and resistance to structural damages. Adaptive biological systems have the ability to change mode in a changing environment but still maintain the same phenotype. It is often observed that organisms have a wide range of rate parameters for the catalysation of the same biologcal reactions, although still able to produce very similar phenotypes. Damage to the network structure of a biological organism is not necessarily fatal, and usually consecutive random mutations would lead to a gradual decrease in functional phenotype, called graceful degradation [12, 52, 53].

These robust phenomena emerges from the underlying properties of system control, alternative mechanisms, decoupling and modularity resident in the organism. System control is the primary feature for having a robust system, and consist of mechanisms such as feedback and feed-forward loops. This control makes it possible to regulate the flow of metabolites through a system with a relatively constant flux, regardless of changes in internal parameters or fluctuations in the metabolite availability. Alternative mechanisms in biological systems can be divided in two different categories; one is the existence of isozymes, enzymes able to catalyse the same reaction, meaning the system will be able to survive despite a nonsense mutation in one of the isozymes; the other is the existence of alternative pathways, where several pathways lead to the same end product. Robust biological systems have often developed ways to decouple the phenotype from the genetic material to a certain extent. This allows mutations to occur in the DNA, although the mutated DNA is not necessarily affecting the phenotype of the organism, as proteins such as Hsp90 identifies and destroys misfolded proteins. Such decoupling mechanisms therefore allow genetic diversity while maintaining the phenotype. Modularity of biological systems is a property derived from the spacial distribution of components separating different chemical species from interaction. This can be illustrated by both membranebound proteins, protein complexes and the compartementalization observed in eucaryotes. When designing and studying biological systems these properties will be able to ensure or explain the robustness of the system [12, 52, 53].

Robustness against external pertubations can be highly advantageous in certian conditions, but it could also increase the fragility of an organism [52].

This can be illustrated for for extremophiles, organisms adapted to extreme conditions, as they are extremely difficult to cultivate. This is because they are very robust to the pertubations that may happen in the extreme environment, but when exposed to an unexpected pertubation (more normal growth conditions) they are unable to survive [23]. In additon, robustness can cause proliferation of unwanted organisms and cells. For instance cancer cells are highly robust, and the structure of the cell's metabolism and defence mechanisms of the body itself help them prevail [52].

1.6 Stochasticity in genetic circuits

Gene expression is exposed to stochasticity, caused by fluctuations in transcription and translation, despite constant environmental conditions giving rise to diversity and diffrentiation of cell types. As the cells only have a few copies of every gene, the gene expression is vulnerable to fluctuations and these can significantly alter the cells phenotypical behaviour [15]. The total noise in a cellular environment can be divided in the noise arising from the gene expression itself, intrinsic noise, and that of the fluctuations in all the other components of a cell, extrinsic noise, like transcription factors and RNA polymerase abundancy. These have been experimentally validated and measured [16]. The sources for intrinsic noise in gene expression have been shown to mainly be caused by translation, as each copy of mRNA can give rise to many proteins [54]. The cellular processes have to be robust in order to cope with the noise, and the general mechanisms described in Sec. 1.5 ensures that the cells are finetuned despite all the stochastic processes. Additionally, cells are able to exploit noise by using it to give a phenotypic diversity in cell cultures that seems to give the species an edge for surviving changes in the environmental conditions. As an example many pathogenic organisms show stochastically driven phase variations that makes it more difficult for the body to create antibodies against them. For instance *Neisseria gonorrheae* have two different pili genes and which one being the active at the time seems to be stochastically driven [55].

There are different ways to measure stochasticity, but some common measures include the coefficient of variation (CoV) and Fano Factor,

$$\text{CoV} = \sigma/\overline{x} \tag{1}$$

$$\text{Fano Factor} = \sigma^2/\overline{x} \tag{2}$$

where σ is the standard deviation and $\overline{x}$ is the mean [15]. Additionally a presentation of the effective potential lanscape is sometimes used [2].

In order to model the effects of stochasticity in gene expression there have been developed stochasitc models as described in Sec. 1.3 and there has also been developed software to handle simulations of these models as described in Sec. 2.2.

1.7 Mathematical approaches

1.7.1 Characetization of points in a in linear systems

A definition of *two-dimensional linear systems* is stated as

$$\begin{aligned} \dot{x}_1 &= ax_1 + bx_2 \\ \dot{x}_2 &= cx_1 + dx_2 \end{aligned} \tag{3}$$

where the dot notation represents the operation $\frac{d}{dt}$. By introducing boldface notation for vectors the system above becomes

$$\dot{\mathbf{x}} = A\mathbf{x} \tag{4}$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{5}$$

Solutions of the two-dimensional linear system can be illustrated as trajectories in a phase plane, making a phase portait, with vectors from the trajectories having a direction in relation to fixed points, $\mathbf{x}^*$. The fixed points, $\mathbf{x}^*$, for a system are the values of $\mathbf{x}$ that satisfies $A\mathbf{x} = 0$ meaning both time derivatives are zero. For a linear system the point $\mathbf{x}^* = 0$ will always be a fixed point. These fixed points can have different behaviours as; nodes, spirals, centers, stars, non-isolated fixed points, saddle points or degenerate nodes. Their individual stability can be either stable or unstable, except saddle points which are always unstable. To classify a fixed point it can be evaluated as a two-dimensional linear system, $\dot{\mathbf{x}} = A\mathbf{x}$, and find the trace, τ, and the determinant, Δ, to the matrix A. For a matrix descibed as A the trace and the determinant values are found by

$$\tau = \text{trace}(A) = a + d \tag{6}$$

$$\Delta = \det(A) = ad - bc \tag{7}$$

and the vaules can be evaluated in Fig. 5 to determine what type of point the fixed point is in the phase plane. From these predictions a phase portrait of the system could be created [56].

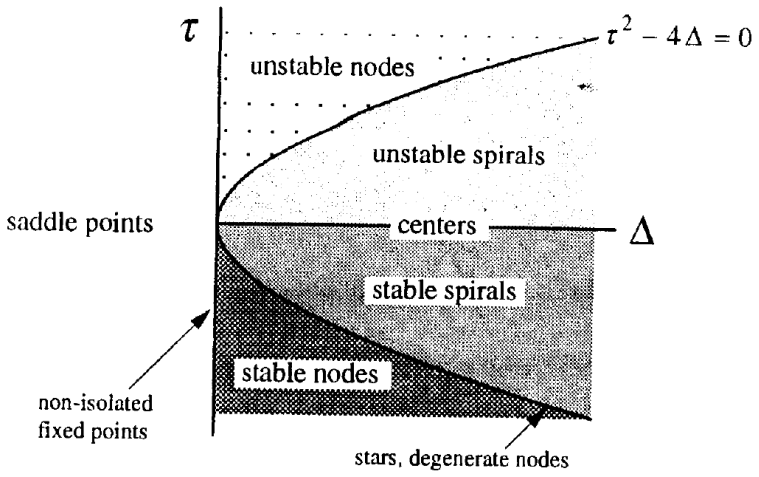


Figure 5: **Classification of fixed points from the Δ and τ values.** By deciding the value of the trace, τ, and the determinant, Δ, of the matrix A for a fixed point, the diagram plotted can be used to decide what type of fixed point $\mathbf{x}^*$ is. A fixed point with $\Delta < 0$ is a saddle point, and if $\Delta = 0$ it is a non-isolated fixed point. For a fixed point with a $\Delta > 0$ the classification depends upon the values of τ and $\tau^2 - 4\Delta$. A fixed point with $\tau = 0$ is a center. If $\tau^2 - 4\Delta = 0$ the point is either a star or a degenerate node, if $\tau^2 - 4\Delta > 0$ it is a node and if $\tau^2 - 4\Delta < 0$ it is a spiral. The stability of a spiral, a node, a star or a degenerate node is dependent upon the value of τ. If $\tau < 0$ it is a stable point but if $\tau > 0$ it is an unstable point. From [56].

1.7.2 Non-linear systems

Non-linear systems can be expressed as a vector field on a phase plane as

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_1, x_2) \end{aligned} \tag{8}$$

where f_1 and f_2 are given functions. This can be written as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{9}$$

where $\mathbf{x} = (x_1, x_2)$, and $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}))$. For this system $\mathbf{x}$ represents a point in the phase plane and $\dot{\mathbf{x}}$ represents the velocity vector at that point. The fixed points, $\mathbf{x}^*$, for this system are the points that satisfy $\mathbf{f}(\mathbf{x}) = 0$, thereby representin the steady state of a system.

The system described in Eq. (8) can be linearized to

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{10}$$

Where

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix}_{(x_1^*, x_2^*)} \tag{11}$$

is called the Jacobian matrix. This matrix can be used to classify fixed points for the system using the same methods as described for fixed point classification for linear systems above. Although, the fixed points described in Fig. 5 as non-isolated fixed points, centers, star nodes and degenerate nodes are borderline cases, in which for a non linear system based solely on the Jacobian are not necessarily correct. Methods for correct characterization of such borderline cases can be found in the litterature, like "Nonlinear Dynamics and Chaos" by Steven H. Strogatz (1994) [56].

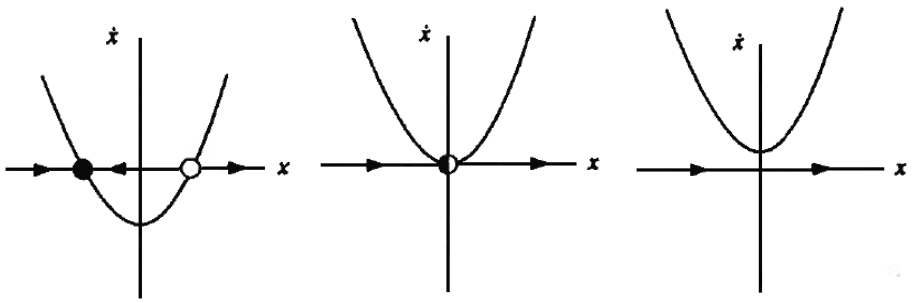
1.7.3 Bifurcations

The number and stability of the steady states may change as the value of some control paramter changes value. The critical value at which the qualitative change of the steady states occur is called a *bifurcation point*. There are different types of bifurcations, and they are described below using the simple base functions of one-dimensional bifurcations to establish the terms. All the bifurcations described are in essence the same for any dimensionality of the system.

The *saddle-node* bifurcations are the bifurcations where by changing the control parameter two steady states, one stable and one unstable, will coalesce and disappear. This can also happen the other way around where two steady states suddenly appears as the control parameter is changed through some critical value. An example of a saddle-node bifurcation is served by the one-dimensional system described in Eq. (12).

$$\dot{x} = r + x^2 \tag{12}$$

where r is the control parameter. This system will have two steady states at $r < 0$, one "half-stable" fixed point for $r = 0$ and when $r > 0$ there are no steady states, as shown in Fig. 6.



(a) r < 0 (b) r = 0 (c) r > 0

Figure 6: **Saddle-node bifurcation.** Stable steady states are marked as filled circles and unstable ones as open circles. (a) When $r < 0$ there are two steady states, one unstable and one stable. (b) When $r = 0$ there is only one fixed point, which is "half stable". (c) When $r > 0$ there are no steady states, shown by the curve for $\dot{x}$ that never cross the x-axis From [56]

A different way of visualising the bifurcation is by plotting the variable values of the steady states as a function of the control parameter as shown in Fig. 7.

Transcritical bifurcations will have qualititative change of the stability of a steady state. This can be exemplified by the one-dimensional system

$$\dot{x} = rx - x^2 \tag{13}$$

where r again is the control parameter. This system will always have one steady state $x^* = 0$. In addition, there will be one fixed point at $r = x^*$. The respective stabilities of the steady states will change as the control parameter changes. While $r < 0$ the steady state $x^* = 0$ will be stable and the other steady state have a negative value and will be unstable. When $r > 0$ the

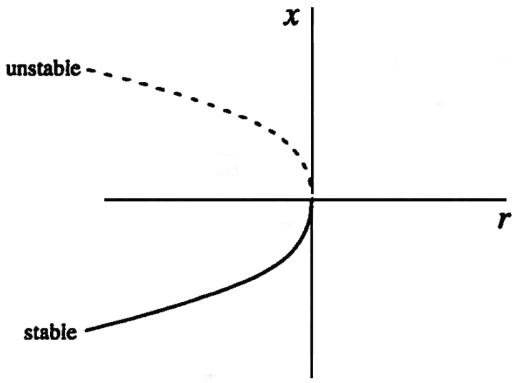
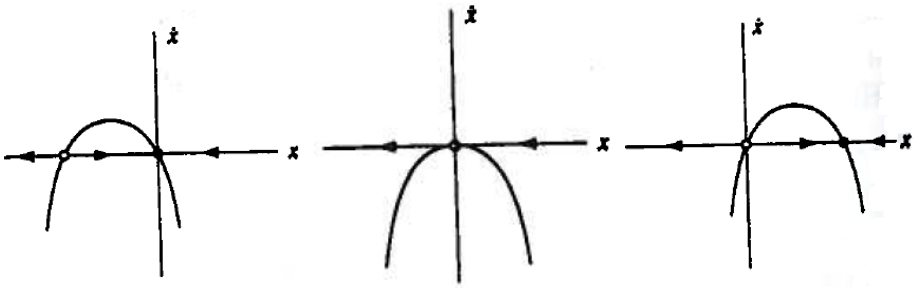


Figure 7: **Saddle-node bifurcation diagram.** Shows how the steady state values of x varies as a function of the control parameter r. The stippled curve represents unstable steady states, while the unbroken curve is stable steady states. For $r > 0$ the two steady states that existed for $r < 0$ have coalesced and disappeared, leaving no steady states. From [56]

steady state $x^* = 0$ will become unstable, while the other steady state will have a positive value and will be stable. The stability have been transferred and hence the bifurcation point is transcritical. Fig. 8 shows the vector field around the steady states as r changes, while the bifurcation diagram for the system is shown in Fig. 9.



(a) r < 0 (b) r = 0 (c) r > 0

Figure 8: **Transcritical bifurcation.** (a) When $r < 0$ there are two steady states, with one stable residing in $x = 0$ and one unstable at $x < 0$. (b) When $r = 0$ there is only one steady state in $x = 0$ (c) When $r > 0$ there are two steady states, whith one unstable residing in $x = 0$ and one stable at $x > 0$. From [56]

There also exists bifurcations where one steady state gives rise to three new steady states. These are called a *pitchfork* bifurcations, as the bifurcation diagram will resemble a pitchfork. There exists two different pitchfork

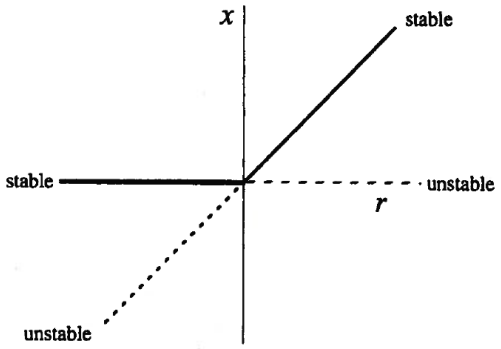


Figure 9: **Transcritical bifurcation diagram.** Shows how the steady state values of x varies as a function of the control parameter r. The steady state in $x = 0$ changes stability in $r = 0$, and goes from stable to unstable for increasing r. There is another steady state for $r \neq 0$ which is unstable and have $x < 0$ for $r < 0$, but becomes stable and $x > 0$ for $r > 0$. From [56]

bifurcations; *supercritical* and *subcritical*. In the supercritcal bifurcation one stable steady state will give rise to two new stable steady states and one unstable steady state in the middle of them. One supercritical pitchfork is the system

$$\dot{x} = rx - x^3 \tag{14}$$

where r is the control parameter. The vector plot of the system in Fig. 10 describes the system for different values of r. When $r < 0$ there is only one steady state which is stable, at the bifurcation point $r = 0$ the system changes, and for $r > 0$ there is 3 steady states where the steady state at $x = 0$ is unstable and there are two stable steady states symmetrically placed around the unstable steady state.

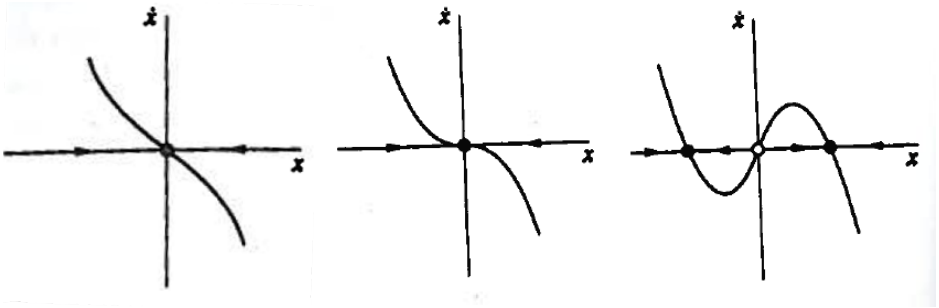
The bifurcation diagram for the supercritical pitchfork is shown in Fig. 11, illustrating why the bifurcation is called a pitchfork bifurcation.

A subcritical pitchfork bifurcation will have one unstable steady state giving rise to one stable and two unstable steady states. The steady states at different parameter values for the system described by

$$\dot{x} = rx + x^3 \tag{15}$$

where r is the control parameter, will give rise to the bifurcation diagram in Fig. 12.

Stability diagrams are often used to illustrate the behaviour of a system with more than one control parameter. For instance the system



(a) r < 0 (b) r $= 0$ (c) r > 0

Figure 10: **Supercritical pitchfork bifurcation.** (a) When $r < 0$ there is only one steady state in $x = 0$ which is stable. (b) When $r = 0$ there is still only one steady state. (c) When $r > 0$ there are three steady states, where the steady state in $x = 0$ is unstable, and two stable steady states are at equal distance from $x = 0$.From [56]

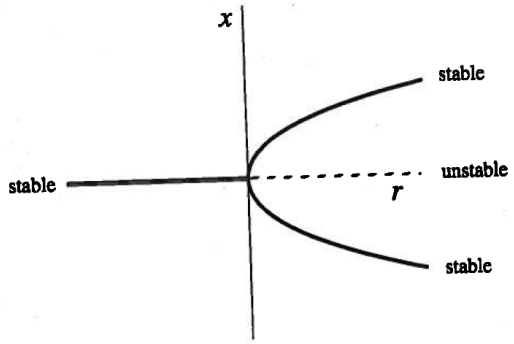


Figure 11: **Supercritical pitchfork bifurcation diagram.** The stable steady state in $x = 0$ splits up into three steady states for $r > 0$. The shape of the curve resembles a pitchfork. From [56]

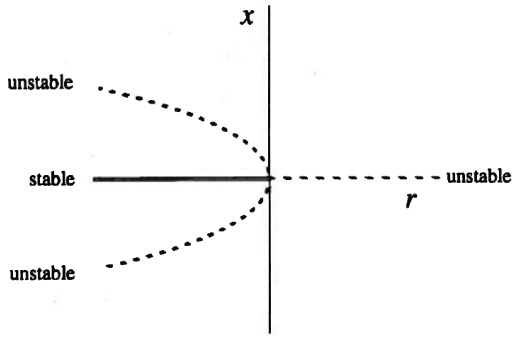


Figure 12: **Subcritical pitchfork bifurcation diagram.** Three steady states, one stable at $x = 0$ and two unstable symmetrically aligned around $x = 0$ merges when $r = 0$, and for $r > 0$ there is only one steady state which is unstable. From [56]

$$\dot{x} = h + rx - x^3 \tag{16}$$

have the two control parameters h and r. As one may vary either of the parameters it could be of use visualising where the bifurcations occur at certian values of each parameter. This can be illustrated in a stability diagram as in Fig. 13 for Eq. (16) [56].

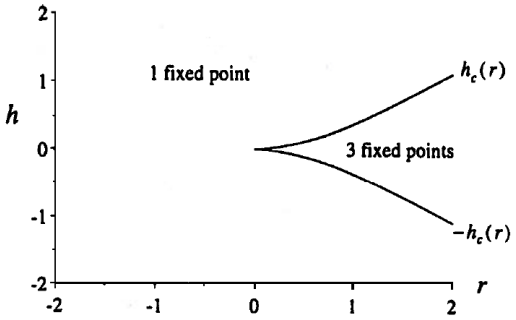


Figure 13: **Stability diagram.** Two separarate areas of the parameter space, where one space contains 1 fixed point, while the other contains three fixed points. Bifurcations giving rise to such patterns may be a saddle-node bifurcation or a pitchfork bifurcation. From [56]

1.7.4 Nondimensionalisation of an equation

When describing a large system using ODEs there are sometimes numerous parameters which are hard to handle and will complicate the analysis of the

system. Then a possible solution would be to nondimensionalise the system, using dimensionless parameters. This would involve using the existing parameters to formulate new dimensionless parameters. A famous example is the spruce budworm outbreak described by Ludwig *et al.* (1979) [57] and also described in Strogatz (1994) [56]. Here a one-dimensional description is formulated as,

$$\dot{N} = RN\left(1 - \frac{N}{K}\right) - \frac{BN^2}{A^2 + N^2} \tag{17}$$

where the change in the variable N ([# budworms]) per unit of time (the dot operation $\frac{d}{dt}$, [months^{-1}]) is described by the parameters R ([# budworms/month]) for growth rate, K ([# budworms]) population carrying capacity, A ([# budworms]) critical level for preditation and B ([# budworms/month]). This equation can be nondimensionalised by first substituting the variable to a dimensionless variable

$$x = \frac{N}{A}$$

giving

$$\frac{A}{B}\frac{dx}{dt} = \frac{R}{B}Ax\left(1 - \frac{Ax}{K}\right) - \frac{x^2}{1 + x^2} \tag{18}$$

Furthermore, one can introudce dimensionless time τ and two dimensionless parameters r and k

$$\tau = \frac{Bt}{A}, \quad r = \frac{RA}{B}, \quad k = \frac{K}{A}$$

and inserting them into Eq. (18) yielding,

$$\frac{dx}{d\tau} = rx\left(1 - \frac{x}{k}\right) - \frac{x^2}{1 + x^2}. \tag{19}$$

And the expression from Eq. (17) have been nondimensionalised and having only two parameters instead of four, and the expression has become much more feasible for further analysis like the ones of bifurcation described above.

2 Materials and Methods

2.1 Deterministic analysis - MATLAB

The deterministic analysis was used to find bifurcation points for different values of certain parameter values, and thereby giving stability diagrams (also called phase planes) as described in Sec. 1.7.3.

Deterministic analysis of the genetic circuit was performed using MATLAB. Especially, bifurcation analysis was performed using the built in function *fsolve*, solving a given non-linear set of equations for zero. The 'trust-region-dogleg' algorithm was chosen to be utilised with *fsolve*, as this is the only algorithm specially designed to solve non-linear problems. This works by giving *fsolve* an equation, *fun*, and initial values, *x0*, and tries to solve the equations described by *fun*. The function proceeds in an iterative fashion until the equation is solved within some predefinable limits. These limits can be described as how close to zero the solution must be in order to be accepted. *fsolve* gives the output solution as x [58]. The *fsolve*-function is a numerical solver, as opposed to an analytical solver, such as Maple or Mathematica.

A general approach to how the stability diagrams were made can be described as follows. First, the function was defined along with several limits for *fsolve* and some other constants used by user defined functions. The function *stestasea* can be described as a brute force way of finding steady states of the system at the given parameter values. Solving the system at some initial parameter values gave a solution that was employed by later functions.

Following, a log-log linear variation of the parameter values was performed, and steady states were found using the *rwrthom2nextss* function. This function takes in *fun*, the last computed solutions for steady states and some values describing how far from the previous solution the new *x0* values will be. The funtion will return the steady states at the next set of parameter values. The *rwrthom2nextss* function uses the *simplhom2scan* function with different resolution, depending upon if *simplhom2scan* gives satisfying results at the first set of resolution. If there is observed a change in the number of steady states the resolution is increased to verify this change. The stability of the individual steady states was determined by using the *stabilityss* function.

The log-log linear scan was then used as a base to find bifurcation points. Starting from points that were bistable, the parameter values were varied one at a time until monostability was found, again using the *rwrthom2nextss* function. When monostability occurred, the last bistable point was again used as input, and then taking smaller steps towards the monostable point, thereby increasing the resolution to what point the bifurcation had occurred.

The mean between the next identified monostable state and bistable state was determined as a bifurcation point. The resulting set of bifurcation points were used to plot stability diagrams as depicted in Sec. 3.5.

All the above mentioned functions, except the built-in function *fsolve*, were defined by the candidate.

2.2 Stochastic analysis - Dizzy

Dizzy is a stochastic simulator of chemical reactions, and was used for stochastic simulations of the genetic circuit described in Sec. 3.1. Both the complete circuit and the approximation were simulated. All simulations were performed using the Gibson-Bruck stochastic algorithm for stochastic simulation of chemical systems [18, 48].

3 Results

3.1 General circuit description

The analysed system is a genetic circuit composing of two genes each encoding a repressor controlling the other gene as homodimers, as depicted in Fig. 14. Each promoter has two operator domains for repressor binding, specific for the repressor encoded by the other gene. The homodimers will bind cooperatively at the two binding sites.

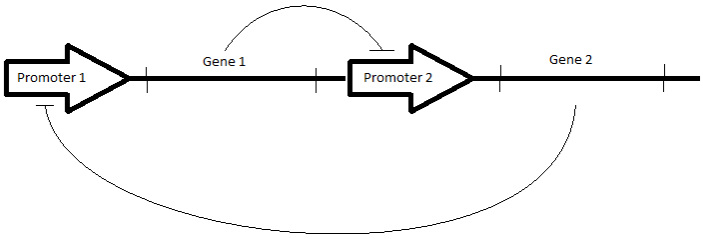


Figure 14: **Genetic Toggle Switch.** Promoter 1, D^1_{ab}, is the promoter for Gene 1, which encodes a repressor with specific binding for Promoter 2. Promoter 2, D^2_{ab}, is the promoter for Gene 2, which encodes a repressor with specific binding for Promoter 1.

Reactions used for describing the system are listed in Tab. 1. Here, the promoter encoding the gene i is described as D^i_{ab}, where a = number of repressors bound at the promoter (0,1 or 2), and b is representing if the RNA polymerase (RNAp) is bound (1), or unbound (0) at the promoter. All these reactions can be described by using ordinary diffrential equations (ODEs) as can be seen in the Eqs. (20) – (40). In the ODEs the dot notation will be used, where $\dot{x} = \frac{dx}{dt}$, is describing the change in species x per unit of time. The reversible reactions are listed in the table with arrows pointing in both directions. For these reactions the parameters are k_i for the forward rate and q_i for the reverse rate. These can be combined in one dissociation parameter $K_i = q_i/k_i$ which can be seen later in the derivation. Many of the species are in squrare brackets denoting that it describes the concentration of that species. Although, note that the promoter elements are not in brackets, but describe the exact number of that species.

Table 1: **Reactions for Toggle Switch with homodimerizarion.** In the first coulumn the reaction is described in words. The reactions associated with *Gene 1* are listed in the second column. The reactions associatied with *Gene 2* are listed in the third coulumn. The reaction rates are written above or below the arrows describing the forward and reverse rate respectively.

Type of reaction	***Gene 1***	***Gene 2***
Repressor binding	$D^1_{00} + P^2_2 \underset{q_{1.2}}{\overset{k_{1.2}}{\rightleftharpoons}} D^1_{10}$ $D^1_{10} + P^2_2 \underset{q_{1.4}}{\overset{k_{1.4}}{\rightleftharpoons}} D^1_{20}$	$D^2_{00} + P^1_2 \underset{q_{2.2}}{\overset{k_{2.2}}{\rightleftharpoons}} D^2_{10}$ $D^2_{10} + P^1_2 \underset{q_{2.4}}{\overset{k_{2.4}}{\rightleftharpoons}} D^2_{20}$
RNAp binding	$D^1_{00} + RNAp \underset{q_{1.3}}{\overset{k_{1.3}}{\rightleftharpoons}} D^1_{01}$ $D^1_{10} + RNAp \underset{q_{1.5}}{\overset{k_{1.5}}{\rightleftharpoons}} D^1_{11}$ $D^1_{20} + RNAp \underset{q_{1.7}}{\overset{k_{1.7}}{\rightleftharpoons}} D^1_{21}$	$D^2_{00} + RNAp \underset{q_{2.3}}{\overset{k_{2.3}}{\rightleftharpoons}} D^2_{01}$ $D^2_{10} + RNAp \underset{q_{2.5}}{\overset{k_{2.5}}{\rightleftharpoons}} D^2_{11}$ $D^2_{20} + RNAp \underset{q_{2.7}}{\overset{k_{2.7}}{\rightleftharpoons}} D^2_{21}$
Transcription initiation	$D^1_{01} \overset{\alpha_{1.m}}{\rightarrow} E^1 + D^1_{00}$ $D^1_{11} \overset{\alpha_{1.m}}{\rightarrow} E^1 + D^1_{10}$ $D^1_{21} \overset{\alpha_{1.m}}{\rightarrow} E^1 + D^1_{20}$	$D^2_{01} \overset{\alpha_{2.m}}{\rightarrow} E^2 + D^2_{00}$ $D^2_{11} \overset{\alpha_{2.m}}{\rightarrow} E^2 + D^2_{10}$ $D^2_{21} \overset{\alpha_{2.m}}{\rightarrow} E^2 + D^2_{20}$
Elongation	$E^1 \overset{\alpha'_{1.m}}{\rightarrow} M^1 + RNAp$	$E^2 \overset{\alpha'_{2.m}}{\rightarrow} M^2 + RNAp$
Translation	$M^1 \overset{\alpha_{1.p}}{\rightarrow} P^1 + M^1$	$M^2 \overset{\alpha_{2.p}}{\rightarrow} P^2 + M^2$
Dimerization	$P^1 + P^1 \underset{q_{1.1}}{\overset{k_{1.1}}{\rightleftharpoons}} P^1_2$	$P^2 + P^2 \underset{q_{2.1}}{\overset{k_{2.1}}{\rightleftharpoons}} P^2_2$
Degradation	$M^1 \overset{\gamma_{1.m}}{\rightarrow} \oslash$ $P^1 \overset{\gamma_{1.p}}{\rightarrow} \oslash$ $P^1_2 \overset{\gamma_{1.p}/\sigma_1}{\rightarrow} \oslash$	$M^2 \overset{\gamma_{2.m}}{\rightarrow} \oslash$ $P^2 \overset{\gamma_{2.p}}{\rightarrow} \oslash$ $P^2_2 \overset{\gamma_{2.p}/\sigma_2}{\rightarrow} \oslash$

$$
\begin{aligned}
[\dot{P_2^1}] =& k_{1.1}[P^1]^2 + q_{2.2}D_{10}^2 + q_{2.4}D_{20}^2 - (\gamma_{1.p}/\sigma_1)[P_2^1] \\
& - q_{1.1}[P_2^1] - k_{2.2}D_{00}^2[P_2^2] - k_{2.4}D_{10}^2[P_2^1] && (20)\\
[\dot{P^1}] =& \alpha_{1.p}[M^1] + 2q_{1.1}[P_2^1] - 2k_{1.1}[P^1]^2 - \gamma_{1.p}[P^1] && (21)\\
\dot{D_{00}^1} =& q_{1.2}D_{10}^1 + q_{1.3}D_{01}^1 + \alpha_{1.m}D_{01}^1 \\
& - k_{1.2}D_{00}^1[P_2^2] - k_{1.3}D_{00}^1[\mathrm{RNAp}] && (22)\\
\dot{D_{10}^1} =& k_{1.2}D_{00}^1[P_2^2] + q_{1.4}D_{20}^1 + q_{1.5}D_{11}^1 + \alpha_{1.m}D_{11}^1 \\
& - q_{1.2}D_{10}^1 - k_{1.4}D_{10}^1[P_2^2] - k_{1.5}D_{10}^1[\mathrm{RNAp}] && (23)\\
\dot{D_{20}^1} =& k_{1.4}D_{10}^1[P_2^2] + q_{1.7}D_{21}^1 + \alpha_{1.m}D_{21}^1 \\
& - q_{1.4}D_{20}^1 - k_{1.7}D_{20}^1[\mathrm{RNAp}] && (24)\\
\dot{D_{01}^1} =& k_{1.3}D_{00}^1[\mathrm{RNAp}] - q_{1.3}D_{01}^1 - \alpha_{1.m}D_{01}^1 && (25)\\
\dot{D_{11}^1} =& k_{1.5}D_{10}^1[\mathrm{RNAp}] - q_{1.5}D_{11}^1 - \alpha_{1.m}D_{11}^1 && (26)\\
\dot{D_{21}^1} =& k_{1.7}D_{20}^1[\mathrm{RNAp}] - q_{1.7}D_{21}^1 - \alpha_{1.m}D_{21}^1 && (27)\\
[\dot{E^1}] =& \alpha_{1.m}(D_{01}^1 + D_{11}^1 + D_{21}^1) - \alpha_{1.m}^{'}[E^1] && (28)\\
[\dot{M^1}] =& \alpha_{1.m}^{'}[E^1] - \gamma_{1.m}[M^1] && (29)\\
[\dot{P_2^2}] =& k_{2.1}[P^2]^2 + q_{1.2}D_{10}^1 + q_{1.4}D_{20}^1 - (\gamma_{2.p}/\sigma_2)[P_2^2] \\
& - q_{2.1}[P_2^2] - k_{1.2}D_{00}^1[P_2^2] - k_{1.4}D_{10}^1[P_2^2] && (30)\\
[\dot{P^2}] =& \alpha_{2.p}[M^2] + 2q_{2.1}[P_2^2] - 2k_{2.1}[P^2]^2 - \gamma_{2.p}[P^2] && (31)\\
\dot{D_{00}^2} =& q_{2.2}D_{10}^2 + q_{2.3}D_{01}^2 + \alpha_{2.m}D_{01}^2 \\
& - k_{2.2}D_{00}^2[P_2^1] - k_{2.3}D_{00}^2[\mathrm{RNAp}] && (32)\\
\dot{D_{10}^2} =& k_{2.2}D_{00}^2[P_2^1] + q_{2.4}D_{20}^2 + q_{2.5}D_{11}^2 + \alpha_{2.m}D_{11}^2 \\
& - q_{2.2}D_{10}^2 - k_{2.4}D_{10}^2[P_2^1] - k_{2.5}D_{10}^2[\mathrm{RNAp}] && (33)\\
\dot{D_{20}^2} =& k_{2.4}D_{10}^2[P_2^1] + q_{2.7}D_{21}^2 + \alpha_{2.m}D_{21}^2 \\
& - q_{2.4}D_{20}^2 - k_{2.7}D_{20}^2[\mathrm{RNAp}] && (34)\\
\dot{D_{01}^2} =& k_{2.3}D_{00}^2[\mathrm{RNAp}] - q_{2.3}D_{01}^2 - \alpha_{2.m}D_{01}^2 && (35)\\
\dot{D_{11}^2} =& k_{2.5}D_{10}^2[\mathrm{RNAp}] - q_{2.5}D_{11}^2 - \alpha_{2.m}D_{11}^2 && (36)\\
\dot{D_{21}^2} =& k_{2.7}D_{20}^2[\mathrm{RNAp}] - q_{2.7}D_{21}^2 - \alpha_{2.m}D_{21}^2 && (37)\\
[\dot{E^2}] =& \alpha_{2.m}(D_{01}^2 + D_{11}^2 + D_{21}^2] - \alpha_{2.m}^{'}[E^2] && (38)\\
[\dot{M^2}] =& \alpha_{2.m}^{'}[E^2] - \gamma_{2.m}[M^2] && (39)
\end{aligned}
$$

$$\begin{aligned}[\dot{\text{RNAp}}] =& q_{1.3}D^1_{01} + q_{1.5}D^1_{11} + q_{1.7}D^1_{21} + \alpha'_{1.m}[E^1] \\ &+ q_{2.3}D^2_{01} + q_{2.5}D^2_{11} + q_{2.7}D^2_{21} + \alpha'_{2.m}[E^2] \\ &- k_{1.3}D^1_{00}[\text{RNAp}] - k_{1.5}D^1_{10}[\text{RNAp}] - k_{1.7}D^1_{20}[\text{RNAp}] \\ &- k_{2.3}D^2_{00}[\text{RNAp}] - k_{2.5}D^2_{10}[\text{RNAp}] - k_{2.7}D^2_{20}[\text{RNAp}] \end{aligned} \tag{40}$$

The total reaction set consist of 40 reactions described by 21 coupled differential equations. Calculation the steady states for this system requires a lot of computational power and a high degree of insight to the values of the rate constants. Although, this set of ODEs can be re-written quite extensively by making the following assumptions.

3.2 Assumptions related to the circuit

Firstly the concentration of free RNAp, [RNAp], was assumed to be constant.

$$[\dot{\text{RNAp}}] = 0 \tag{41}$$

As the circuit is symmetrical all the following assumptions and the derivation later on will relate to both genetic elements and their transcripts, but for simplicity only the regulation, transcription, translation etc. at genetic element one will be described and the species index will be emitted for simplicity. The active dimeric repressor controlling the *gene 1*-expression, P_2^2, will be named TF. Further, it is assumed that there is only one copy of each genetic element.

$$\sum_{ab} D_{ab} = D_{00} + D_{10} + D_{20} + D_{01} + D_{11} + D_{21} = 1 \tag{42}$$

With one copy of each gene in each cell, there will be only two binding sites for each repressor. It can be condsidered a reasonable assumption that with so few binding sites there will be almost no binding if the concentration of the repressor is low. If the repressor concentration is high there might be binding at the operator sequences, although at a high concentration the binding of one repressor will cause only a small change in the total amount of repressor. This lead to the assumption that the terms describing repressor binding and unbinding to the promoter was left out from the expression in Eq. (20) giving the following expression,

$$[\dot{P}_2] = k_1[P]^2 - q_1[P_2] - (\gamma_p/\sigma)[P_2] \tag{43}$$

In order to simplify algebraic aspects of non-dimensionalising the system (see Sec. 3.3) two more assumptions were made; the repressor dimerisation dissociation constants are assumed to be equal *i.e.* $K_{1.2} = K_{2.2} = K_2$ and the degradation rates of each monomer is assumed to be equal *i.e.* $\gamma_{1.p} = \gamma_{2.p} = \gamma_p$. If both protein repressors contain the LVA degradation tail [59] the assumption of equal degradation rate seems more reasonable, although not unreasonable by itself.

Additionally, in order to decrease the amount of different non-dimensionalised parameters K_5 is assumed to be equal to K_7.

3.3 Derivation of the approximative expression

By using the above mentioned assumptions it were possible to derive a system consisting of two coupled differential equations instead of the 21 mentioned in Sec. 3.1. This was done by assuming steady state for all the equations, apart from the two final ones, and also a non-dimensionalisation of the system.

As a starter, it was assumed that all the binding reactions at the DNA are in steady state, giving the following expression for $[\dot{E}]$

$$\begin{aligned} [\dot{E}] &= \alpha_m(D_{01} + D_{11} + D_{21}) - \alpha_m^{'}[E] = 0 \\ \alpha_m^{'}[E] &= f([TF]) = \alpha_m(D_{01} + D_{11} + D_{21}) \end{aligned} \tag{44}$$

By then setting Eqs. (22)-(27) = 0 (at steady state) and solving for ${D_{10}}^*$, ${D_{20}}^*$, ${D_{01}}^*$, ${D_{11}}^*$ and ${D_{21}}^*$, where * denotes that it represents the steady state condition of the related species, one gets

$${D_{10}}^* = \frac{k_2 D_{00}[TF]}{q_2} = \frac{D_{00}[TF]}{K_2} \tag{45}$$

$${D_{20}}^* = \frac{k_2 k_4 D_{00}[TF]^2}{q_2 q_4} = \frac{D_{00}[TF]^2}{K_2 K_4} \tag{46}$$

$${D_{01}}^* = \frac{k_3 D_{00}[\mathrm{RNAp}]}{q_3 + \alpha_m} = \frac{D_{00}[\mathrm{RNAp}]}{K_3} \tag{47}$$

$${D_{11}}^* = \frac{k_2 k_5 D_{00}^1[\mathrm{RNAp}][TF]}{q_2(q_5 + \alpha_m)} = \frac{D_{00}[\mathrm{RNAp}][TF]}{K_2 K_5} \tag{48}$$

$${D_{21}}^* = \frac{k_2 k_4 k_7 D_{00}[\mathrm{RNAp}][TF]^2}{q_2 q_4 (q_7 + \alpha_m)} = \frac{D_{00}[\mathrm{RNAp}][TF]^2}{K_2 K_4 K_7} \tag{49}$$

where the dissociation constant have been defined as $K_i = q_i/k_i$, and as $K_i >> \alpha_m$ (see Sec. 3.8), $(q_i + \alpha_m)/k_i \approx q_i/k_i = K_i$. The expressions in Eqs. (47), (48) and (49) can be inserted into Eq. (44)

$$f([TF]) = \alpha_m D_{00}[\text{RNAp}](\frac{1}{K_3} + \frac{[TF]}{K_2K_5} + \frac{[TF]^2}{K_2K_4K_7}) \tag{50}$$

The steady state concentrations of the different states of the promoter, Eqs. (45)–(49), can be inserted into Eq. (42), before solving for the number of free promoter

$$\begin{aligned} 1 &= D_{00}\left(1 + \frac{[TF]}{K_2} + \frac{[TF]^2}{K_2K_4} + [\text{RNAp}](\frac{1}{K_3} + \frac{[TF]}{K_2K_5} + \frac{[TF]^2}{K_2K_4K_7})\right) \\ {D_{00}}^{-1} &= 1 + \frac{[\text{RNAp}]}{K_3} + (1 + \frac{[\text{RNAp}]}{K_5})\frac{[TF]}{K_2} + \frac{K_2}{K_4}(1 + \frac{[\text{RNAp}]}{K_7})(\frac{[TF]}{K_2})^2 \end{aligned} \tag{51}$$

By introducing the following dimensionless parameters

$$s = \frac{K_3}{K_5}, \quad u = \frac{K_3}{[\text{RNAp}]}, \quad T = \frac{[TF]}{K_2}, \quad r = \frac{K_2}{K_4}, \quad \mu = \frac{u+s}{1+u}$$

where s is a measure for promoter leakage, u is the RNAp-promoter dissociation constant scaled by the concentration of free RNA, T is the dimensionless concentration of the repressor and r is a measure for cooperativity in repressor-DNA binding. By using the assumption that $K_5 = K_7$, these parameters can be substituted into Eq. (51) and give the following expression

$$\begin{aligned} {D_{00}}^{-1} &= 1 + u^{-1} + (1 + su^{-1})(T + rT^2) \\ &= (1 + u^{-1})(1 + \mu(T + rT^2)) \end{aligned} \tag{52}$$

Plugging this expression back into Eq. (50) and making use of the same parameters again gives

$$\begin{aligned} f([TF]) &= \alpha_m D_{00}(\frac{[\text{RNAp}]}{K_3} + \frac{[\text{RNAp}][TF]}{K_2K_5} + \frac{[\text{RNAp}][TF]^2}{K_2K_4K_7}) \\ \frac{f([TF])}{\alpha_m} &= D_{00}(u^{-1} + u^{-1}sT + u^{-1}srT^2) \\ &= \frac{u^{-1} + su^{-1}(T + rT^2)}{(1 + u^{-1})(1 + \mu(T + rT^2))} \\ &= \frac{1}{(1 + u/s)}\left(1 + \frac{\nu}{1 + \mu(T + rT^2)}\right) \end{aligned} \tag{53}$$

where $\nu = \frac{u(1-s)}{s(1+u)}$. The steady state concentration of mRNA, $[M]^*$, can be expressed as

$$\begin{aligned}
&\alpha_m^{'}[E] - \gamma_m[M] = 0 \\
&[M]^* = (\gamma_m)^{-1} f([TF]) \\
&[M]^* = \frac{\alpha_m}{\gamma_m(1+u/s)} \left(1 + \frac{\nu}{1+\mu(T+rT^2)}\right)
\end{aligned} \tag{54}$$

The system can be non-dimensionalised by scaling all concentrations with K_2 and time with γ_p^{-1}. By further restoring the species indices, Eq. (21) can be rewritten as

$$\begin{aligned}
\dot{p_1} =& \frac{\alpha_{1.p}[M^1]^* + 2q_{1.1}K_2T_1 - 2k_{1.1}(K_2p_1)^2 - \gamma_p K_2 p_1}{\gamma_p K_2} \\
=& \frac{\alpha_{1.p}[M^1]^*}{K_2\gamma_p} - p_1 - \frac{2}{K_2\gamma_p}(k_{1.1}(K_2p_1)^2 - q_{1.1}K_2T_1) \\
=& \frac{\alpha_{1.p}\alpha_{1.m}}{K_2\gamma_p\gamma_{1.m}(1+u_1/s_1)} \left(1 + \frac{\nu_1}{1+\mu_1(T_2+r_1T_2^2)}\right) - p_1 - 2\psi(p_1,T_1)
\end{aligned} \tag{55}$$

Now the variable p_1 is dimensionless and the dot over the variable denotes the operation $\gamma_p^{-1}\frac{d}{dt}$ and ψ is a function of p_1 and T_1. By further introducing the parameters

$$\lambda_1 = \frac{\beta_1}{1+(u_1/s_1)}, \quad \beta_1 = \frac{\alpha_{1.p}\alpha_{1.m}}{K_2\gamma_{1.m}\gamma_p}$$

where β_1 is the gene expression efficiency of *gene 1*, the expression becomes

$$\dot{p_1} = \lambda_1 \left(1 + \frac{\nu_1}{1+\mu_1(p_1^2+r_1T_2^2)}\right) - p_1 - 2\psi(p_1,T_1) \tag{56}$$

By then calculating the steady state for the dimensionless form of Eq. (43)

$$\begin{aligned}
\dot{T_1} &= \frac{k_{1.1}(K_2p_1)^2 - q_{1.1}T_1K_2 - \gamma_p/\sigma_1 T_1 K_2}{\gamma_p K_2} \\
&= -(1/\sigma_1)T_1 + \frac{1}{K_2\gamma_p}(k_{1.1}(K_2p_1)^2 - 2q_{1.1}K_2T_1) \\
&= -(1/\sigma_1)T_1 + \psi(p_1,T_1) = 0 \\
\psi(p_1,T_1) &= (1/\sigma_1)T_1
\end{aligned} \tag{57}$$

Inserting Eq. (57) into Eq. (56) gives

$$\dot{p_1} = \lambda_1 \left(1 + \frac{\nu_1}{1 + \mu_1(T_2 + r_1 T_2^2)}\right) - (p_1 + (2/\sigma_1)T_1) \tag{58}$$

The steady state for $\dot{T_1}$ also gives

$$\begin{aligned}
\dot{T_1} &= \frac{k_{1.1}(K_2 p_1)^2 - q_{1.1} T_1 K_2 - (\gamma_p/\sigma_1) T_1 K_2}{\gamma_p K_2} \\
&= \frac{k_{1.1} K_2 p_1^2 - q_{1.1} T_1 - (\gamma_p/\sigma_1) T_1}{\gamma_p} = 0 \\
p_1^2 &= \frac{T_1(q_{1.1} - \gamma_p/\sigma_1)}{K_2 k_{1.1}} \approx \frac{T_1 K_{1.1}}{K_2} = T_1/\theta_1 \\
p_1 &= \sqrt{T_1/\theta_1}
\end{aligned} \tag{59}$$

where $\theta_1 = K_2/K_{1.1}$. The following conditions must be satisfied

$$\begin{aligned}
&p_1 = \sqrt{T_1/\theta_1}, \\
&\dot{T_1} = \frac{dT_1}{dp_1}\frac{dp_1}{dt} \quad \Rightarrow \quad \dot{p_1} = \frac{dp_1}{dT_1}\dot{T_1} = \frac{1}{2\sqrt{\theta_1 T_1}}\dot{T_1}
\end{aligned}$$

in order to perform a variable change on Eq. (58);

$$\begin{aligned}
\dot{T_1} &= 2\sqrt{(\theta_1 T_1)}\lambda_1 \left(1 + \frac{\nu_1}{1 + \mu_1(T_2 + r_1 T_1^2)}\right) - 2\sqrt{\theta_1 T_1}\left(\sqrt{\frac{T_1}{\theta_1}} + \frac{2T_1}{\sigma_1}\right) \\
&= 2\sqrt{(\theta_1 T_1)}\lambda_1 \left(1 + \frac{\nu_1}{1 + \mu_1(T_2 + r_1 T_2^2)}\right) - 2\left(T_1 + \frac{2\sqrt{\theta_1} T_1^{(3/2)}}{\sigma_1}\right) \\
&= 2\sqrt{(\theta_1 T_1)}\lambda_1 \left(1 + \frac{\nu_1}{1 + \mu_1(T_2 + r_1 T_2^2)}\right) - 2\left(T_1 + \frac{2T_1^{(3/2)}}{\sqrt{\theta_1}\kappa_1}\right)
\end{aligned} \tag{60}$$

where $\kappa_1 = \sigma_1/\theta_1$. All the same operations can be performed on the equations describing reactions at *Gene 2*, giving the following dimensionless expression for the repressor 2

$$\dot{T_2} = 2\sqrt{(\theta_2 T_2)}\lambda_2 \left(1 + \frac{\nu_2}{1 + \mu_2(T_1 + r_2 T_1^2)}\right) - 2\left(T_2 + \frac{2T_2^{(3/2)}}{\sqrt{\theta_2}\kappa_2}\right) \tag{61}$$

By setting $T_1 = x$ and $T_2 = y$ one gets the expressions

$$\begin{aligned}\dot{x} &= 2\sqrt{(\theta_1 x)}\lambda_1\left(1+\frac{\nu_1}{1+\mu_1(y+r_1y^2)}\right)-2(x+\frac{2x^{(3/2)}}{\sqrt{\theta_1}\kappa_1})\\ \dot{y} &= 2\sqrt{(\theta_2 y)}\lambda_2\left(1+\frac{\nu_2}{1+\mu_2(x+r_2x^2)}\right)-2(y+\frac{2y^{(3/2)}}{\sqrt{\theta_2}\kappa_2})\end{aligned} \tag{62}$$

being close to the expression describing the HOM2 circuit in Ghim and Almaas (2009) [2]. In this article all parameters were assumed having the exact same values as it is investigated as a completely symmetrical circuit. This assumption is probably not too realistic, although the expressions describing the circuit could still be valid, even by making the circuit asymmetrical.

3.4 Numerical instability

The equation set in Eq. (62) was solved for steady state setting $\dot{x} = 0$ at different sets of parameter values. Although for some parameter values the solutions became very small ($<10^{-5}$) and these steady state solutions were difficult to find using the numerical solver in Matlab. This numerical problem was caused by having the variables in the denominators of the expression. In order to find the correct steady states the equation set had to be rewritten to

$$\begin{aligned}\dot{x} &= 2\sqrt{(\theta_1 x)}\lambda_1(1+\mu_1 y+\mu_1 r_1 y^2+\nu_1)-2x\left(1+\frac{2}{\kappa_1}\sqrt{\frac{x}{\theta_1}}\right)(1+\mu_1 y+\mu_1 r_1 y^2)\\ \dot{y} &= 2\sqrt{(\theta_2 y)}\lambda_2(1+\mu_2 x+\mu_2 r_2 x^2+\nu_2)-2y\left(1+\frac{2}{\kappa_2}\sqrt{\frac{y}{\theta_2}}\right)(1+\mu_2 x+\mu_2 r_2 x^2)\end{aligned} \tag{63}$$

3.5 Deterministic analysis

By solving a set of ODEs for the steady state at different parameter values the number of steady states was deduced. By determining the stability of each steady state the number of stable steady states at those particular parameter values was found. The parameter values at which the number of steady states change is called the critical value (a bifurcation point). These critical can be used for analysing system behaviour at different parameters, and these values correspond to the curves in stability diagrams, like the one in Fig. 13 in Sec. 1.7.3. The deterministic analysis of the gentic circuit, described by

Eq. (63), was composed of such analysis giving stability diagrams for some important parameter values. Unless otherwise noted the parameter values used are the same as in [2] (based on the cI repressor from bacteriophage λ) and as listed in Tab. 2. These parameter values will also serve as a bistable reference point in the phase plane (stability diagrams).

Table 2: **Parameter Values.** The parameter values used for deterministic analysis of the system, in order to compute stability diagrams. All the values are based on the values used in [2]. The bistable state with all these parameter values serve as a reference point in the following graphs.

Parameter	**Value**
K_2	20 nM
$K_{1.1}$	10 nM
r_1	25
s_1	0.01
β_1	17.5
u_1	3
σ_1	10
$K_{2.1}$	10 nM
r_2	25
s_2	0.01
β_2	17.5
u_2	3
σ_2	10

The leakage from each of the promoters, described by the parameters s_1 and s_2, can have values independent of the other and can to a quite high extent be modified by genetic manipulation. Therefore, it could be very interesting to explore the stability diagram composed of these two parameters. The stability diagram for s_1 vs s_2 was computed using Matlab running the file s1vss2rwrthom2.m from the folder DeterminitsticAnalysis/s1s2 in the attached zip-file, and is shown in Fig. 15.

Similarly, the individual gene expression from each of the promoters, described by the parameters β_1 and β_2, are able to vary independently of each other. This can also be highly modified by genetic manipulation, especially by modifiying the 5'UTR region, making the β_1 vs β_2 stability diagram highly interesting. This was computed using Matlab running the file beta1beta2rwrthom2.m from the folder DeterministicAnalysis/beta1beta2assym in the attached zip-file and is shown in Fig. 16.

Furthermore, one of the promoters and the corresponding gene can be

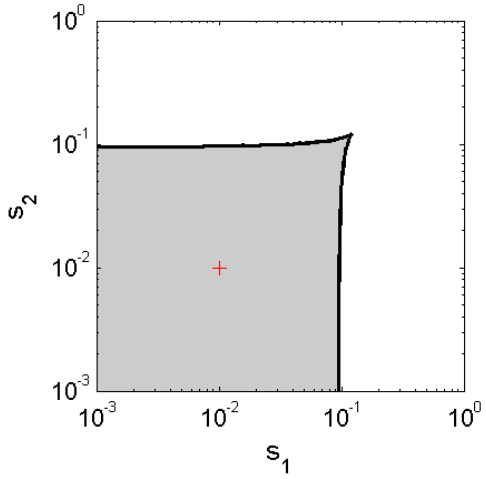


Figure 15: **Stability diagram for s_1 vs s_2.** The system is bistable in the shaded area. All other parameters are as noted in Tab. 2. The red cross indicates the reference point, where all the parameter values are as in Tab. 2.

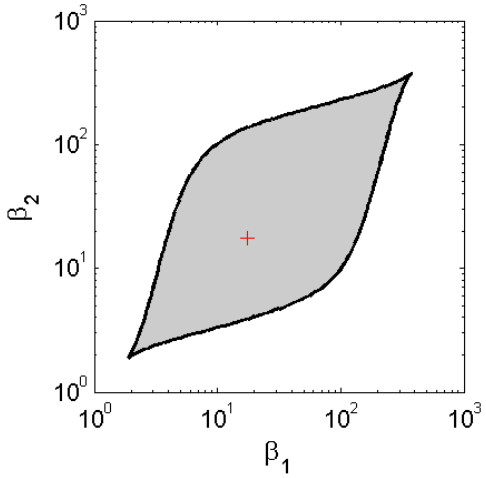


Figure 16: **Stability diagram for β_1 vs β_2.** The system is bistable in the shaded area. All other parameters are as noted in Tab. 2. The red cross indicates the reference point, where all the parameter values are as in Tab. 2.

kept untouched, while only modifying the properties of leakage and gene expression of the other gene, corresponding to the parameters s_2 and β_2 respectively. By predicting the gene expression and the leakage of another repressor and promoter pair (other than the cI-repressor, with two binding sites for the repressor) relative to the cI parameters, it could be possible to predict if the system would be bistable or not using a stability diagram mapping the s_2 vs β_2. This was computed using Matlab running the file s2vsbeta2.m from the folder DetetrministicAnalysis/s2beta2 in the attached zip-file and is shown in Fig. 17.

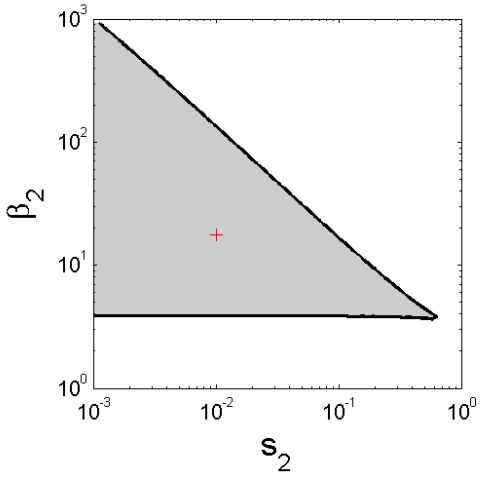


Figure 17: **Stability diagram for s_2 vs β_2.** The system is bistable in the shaded area. All other parameters are as noted in Tab. 2. The red cross indicates the reference point, where all the parameter values are as in Tab. 2.

Additionally, stability diagrams for s_1 vs s_2 with different values for both β-values and for β_1 vs β_2 with different values for s was generated. The s_1 vs s_2-diagrams were was computed by running the file runthemall.m from the folder DeterministicAnalysis/s1s2assym. The β_1 vs β_2-diagrams were computed running the file beta1beta2rwrthom2.m from each of the subfolders in DeterministicAnalysis/beta1beta2assym. The resulting plots are illustrated in Figs. 18 and 19 respectively.

3.6 Approximative parameter values

For the HOM2 circuit described by Ghim and Almaas (2009) all the parameters were derived from the phage lambda repressor, cI. This lead to the following values,

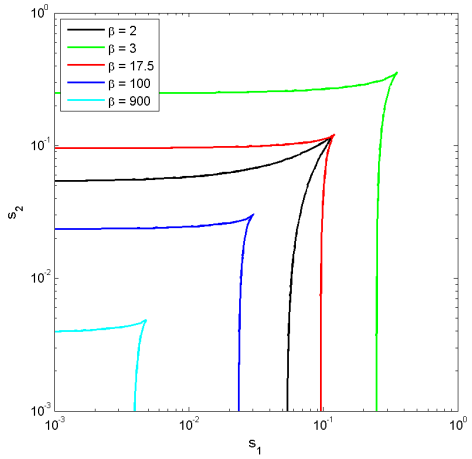


Figure 18: **Stability diagram for s_1 vs s_2 with different values for β.** The system is bistable within the borders of each curve, as in Fig. 15. The different curves corresponds to different values of β_1 and β_2 (both parameters are set to the same value). The values of β are 2, 3, 17.5, 100 and 900 represented by the black, green, red, blue and cyan curves respectively. All other parameters are as noted in Tab. 2.

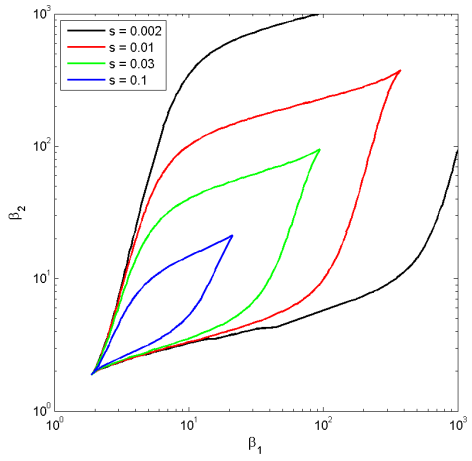


Figure 19: **Stability diagram for β_1 vs β_2 with different values for s.**The system is bistable in the area of the curves, as in Fig. 16. The different curves correspond to different values of s_1 and s_2 (both paramters are equal for each curve). The values of s are 0.002, 0.01, 0.03 and 0.1 represented by the black, red, green and blue curves respectively. All other parameters are as noted in Tab. 2.

$$s = 0.01, \quad u = 3, \quad \beta = 17.5, \quad r = 25.$$

In the supplementary material to the first genetic toggle switch there were made several measurements for promoter expression and leakage. Here they tested both λ cI- (cI for simplicity) and TetR-controlled promoters. This can be used as a base for designing a system composing of these two repressors operating on each others promoters, fitting well with the model. Some chosen measurements for the cI-controlled promoters were assumed to be proportional to the gene expression (β_1) and leakage (s_1) of one of the genes in the model. These reference values were estimated from the experiments involving plasmids pBRT123 and pTAK107. By comparison to the measurements made with the TetR-controlled promoters in the plasmids pBAG103 and pIKE108, the parameters β_2 and s_2 were estimated. The following relationship was used to estimate the β_2 parameter;

$$\begin{aligned}
\frac{\beta_1}{\beta_2} &= \frac{\text{expression from bare cI-controlled promoter}}{\text{expression from bare tetR-controlled promoter}} \\
\beta_2 &= \frac{\beta_1 \cdot \text{bare tetR-controlled}}{\text{bare cI-controlled}} \\
\beta_2 &= \frac{17.5 \cdot 660}{390} = 30
\end{aligned} \tag{64}$$

Further was be assumed that s is proportional to the leakage from the repressed promoters relative to the expression from the bare promoters, so that

$$\begin{aligned}
\frac{s_1}{s_2} &= \frac{(\text{repressed cI-controlled /bare cI-controlled})}{(\text{repressed tetR-controlled /bare tetR-controlled})} \\
s_2 &= \frac{s_1 \cdot (\text{repressed tetR-controlled /bare tetR-controlled})}{(\text{repressed cI-controlled /bare cI-controlled})} \\
s_2 &= \frac{0.01 \cdot (5.8/660)}{(2.0/387)} = 0.005
\end{aligned} \tag{65}$$

These values assume a linear relationship between the expression and leakage of cells containing the promoters in high and low copy numbers, which is usually not true [60].

3.7 Stochastic analysis of the approximation

The approximation can be explored further by exposing it to stochastic fluctuations, to verify the existence of bistable regions and the stability of the

stable steady states therein. In the stochastic simulations an expression using the monomers as variables were used and the expression was redimensionalised. For p_1 this is done by inserting the relation from Eq. (59) into Eq. (58)

$$\dot{p_1} = \lambda_1 \left(1 + \frac{\nu_1}{1 + \mu_1(\theta_2 p_2^2 + r_1 \theta_2^2 p_2^4)}\right) - (p_1 + (2/\sigma_1)\theta_1 p_1^2) \tag{66}$$

As the expression can be divided into one positive and one negative term, $\dot{p_1} = F(p_2) - G(p_1)$, this can be interpreted as one term for the synthesis and one for the degradation of the monomer. The syntesis and degradation terms were redimensionalised

$$\begin{aligned} F(P_2) &= K_2 \lambda_1 \left(1 + \frac{\nu_1}{1 + \mu_1(\theta_2 (P_2/K_2)^2 + r_1 \theta_2^2 (P_2/K_2)^4)}\right) \\ G(P_1) &= (K2(P_1/K_2)(1 + (2/\sigma_1)\theta_1 (P_1/K_2)) \\ &= P_1(1 + \frac{2(P_1/K_2)}{\kappa_1}) \end{aligned} \tag{67}$$

before incorporating them in the Dizzy model files. After each simulation the monomers numbers were converted into dimer numbers by using the equilibrium relation stated as

$$[P_2^1] = \frac{[P^1]^2}{K_{1.1}} \tag{68}$$

The stability diagram in Fig. 15 can be explored by keeping one of the leakage parameters at the reference value, $s_2 = 0.01$ while varying the other leakage parameter from a the lower to the higher limits of the stability diagram, $s_1 = [0.001...1]$. The variation in the parameter was done in discrete steps,$s_1 = 10^{(-3+0.1(floor(time*0.0001)))}$. This simulation have two possible different starting points, the protein repressor (the variable) can be in either high state with many copies (switch is turned on) or in low state (switch is turned off), as can be seen in the attached files 'cItetRmodelvars1cIhigh.cmdl' and 'cItetRmodelvars1cIlow.cmdl' respectively. The simulations were performed in the Dizzy package with the gibson-bruck algorithm, with start = 0.0, stop = 3000000, number of result points = 300000 and stochastic ensemble size = 1, giving the data for the plot in Fig. 20. The data for these two simulations can be combined in one dataset to calculate the potential effective landscape,

$$V_{eff} = -ln(\mathrm{P}(P^1 - P^2)) \tag{69}$$

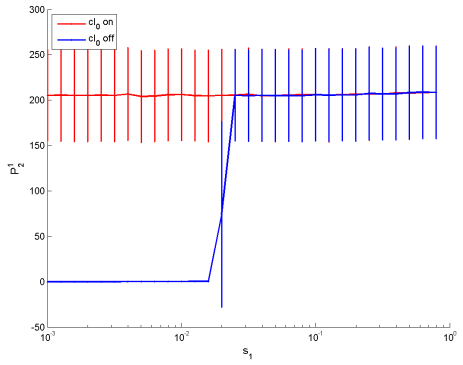


Figure 20: **Stochastic simulation result plot, potential landscape at different s_1 values.** At low s_1-values there are two possible states for the P_2^1-variable. As the s_1-value increase the lower stable steady state suddenly disappears, in strong resemblance to a saddle-node bifurcation. The errorbars represent the standard deviation around the mean value.

[2], at different values of s_1 giving rise to the plot shown in Fig. 21.

Furthermore, one can inspect the stability diagram in Fig. 16 in the same manner keeping one of the gene expression parameters at the reference value, $\beta_2 = 17.5$, while varying the other gene expression parameter, $\beta_1 = [1...100]$. In this situation, it could seem from the stability diagram, Fig. 16, that there should be two bifurcations, and bistability only in an intermediate area. This can be explored by performing two stochastic simulations. In one simulation $\beta_1 = 10^{2-0.1*floor(time*0.0001)}$, while in the other $\beta_1 = 10^{0+0.1*floor(time*0.0001)}$, as can be seen in the attached files 'cItetRmodelbeta1vardecreasing.cmdl' and 'cItetRmodelbeta1varincreasing.cmdl' respectively. This means one will make discrete steps in the β_1-value for every 10000th time point, starting at $\beta_1 = 100$ and $\beta_1 = 1$ for the two simulations respectively. By running the simulations with the gibson-bruck algorithm, start = 0.0, stop = 200000, number of result points = 200000 and stochastic ensemble size = 1 data for the plot in Fig. 22 was obtained. Here the simulation corresponding to starting with $\beta_1 = 100$ is represented by the red curve, and the one starting at $\beta_1 = 1$ by the blue curve. By combining the two datasets the effective potential landscape can be calculated for different values of β_1 as well, as illustrated in Fig.23.

Finally, the estimated values for β and s from Sec. 3.6 can be used in a stochastic simulation. This was done by running two simulations where the expression of cI was turned on and off in the files 'cItetRmodelassymGardnercIon.cmdl' and 'cItetRmodelassymGardnercIoff.cmdl' respectively. Both simultaions were performed with the gibson-bruck algorithm, start = 0, stop

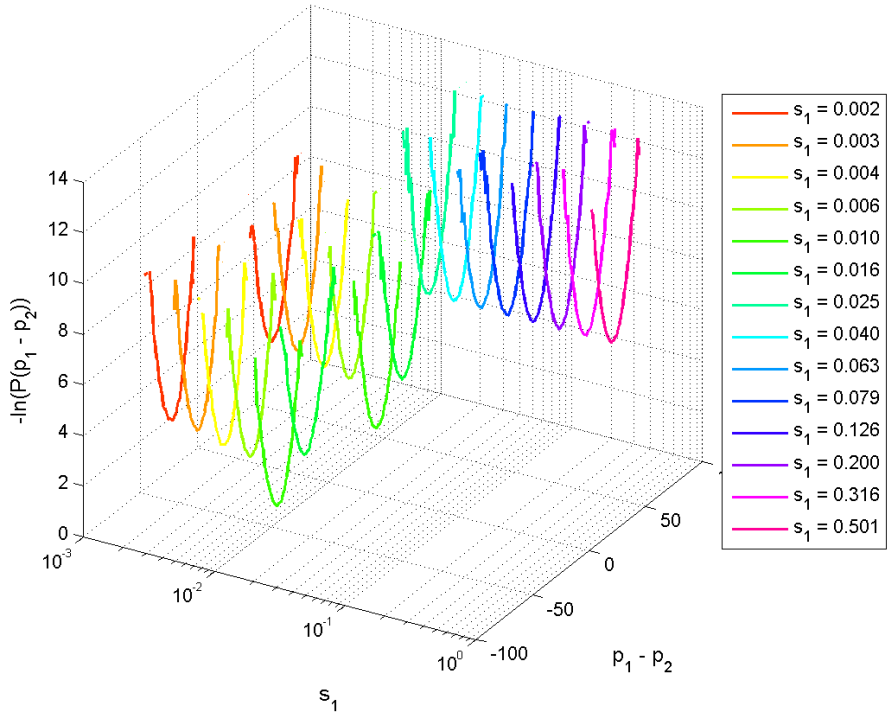


Figure 21: **Stochastic simulation result plot, effective potential landscape at different s_1-values.** The double welled potential at lower values describe the presence of two stable steady states. As one disappears with increasing s_1 the other still remains, now with twice as many result points taking the well to a higher potential, although the well remains at almost the same values for p_1 and p_2.

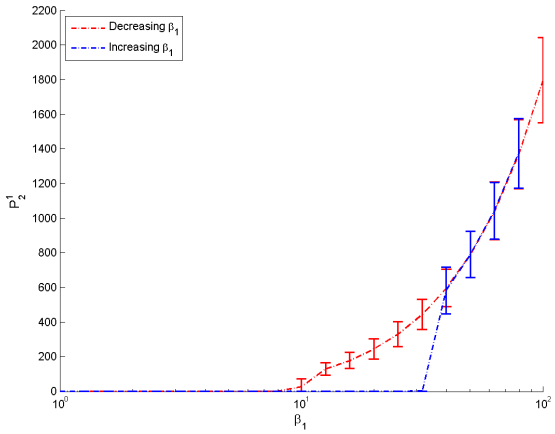


Figure 22: **Stochastic simulation result plot of P_2^1 as a function of β_1.** For increasing values of β_1 the steady state suddenly shifts to a higher state at $\beta \approx 40$, as if the lower state disappeared like a saddle-node bifurcation. For decreasing values of β_1 the high state is reduced until it the P_2^1-number is reduced to 0. The errorbars represent the standard deviation around the mean value.

= 1000000, number of result points = 100000, stochastic ensemble size = 1. It was not observed a single switch even with those log run times, and the effective potential landscape created by using the data from the two simulations is shown in Fig. 24.

3.8 Complete circuit parameters

In order to do a full system simulation (simulating the entire system as described in Tab. 1) there was a need for intensive litterature studies for correct rate parameters. The repressors chosen were the cI and TetR repressors, as both of them have two operator sites for their respective homodimers and are frequently used in genetic circuits [1, 17, 22]. For most of the reversible reaction there was no information about the single rates, so these were chosen abritarily from the dissociation constants, trying to keep all single rate constants > 1.

3.8.1 Dimerisation of the cI-repressor — $K_{1.1}$

In Burz *et al.* (1994) the dimerization equilibrium constant of the λ cI repressor was measured to $K_{eq} = 1.8 \cdot 10^8$ M. Since the $K_{1.1}$ parameter is the dissociation constant of this reaction following the relationship,

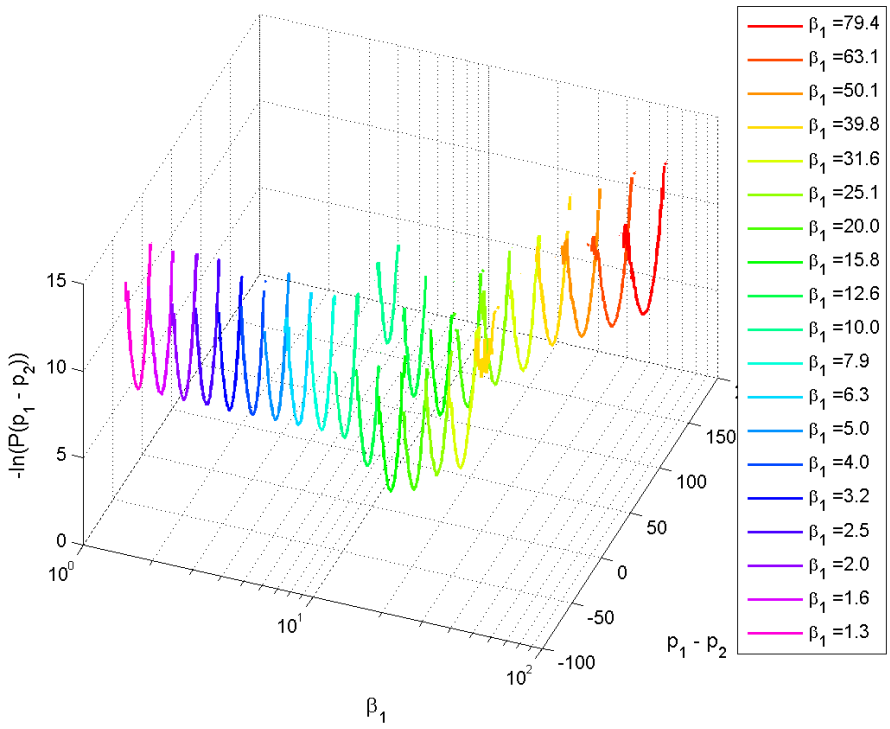


Figure 23: **Stochastic simulation result plot, effective potential landscape at different β_1-values.** At low β_1-values there is only one well in the potential landscape. As the β_1-value increases another arises, giving the system the double-welled landscape characterising a bistability. At even higher values for β_1 the well where p_2 is highest disappears, and the p_1 is now far higher in numbers as the system is now monostable having only one well.

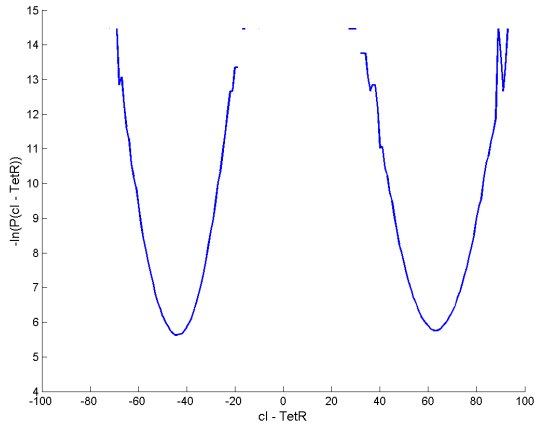


Figure 24: **Stochastic simulation result plot, effective potential lanscape at $\beta_1 = 30$ and $s_1 = 0.01$.** The simulations showed a distinct double-well pattern, as there was no switching between the two different states. The asymmetry of the system is also evident, as the well with cI dominating TetR is located further from the center than the well where TetR is higher than cI.

$$K_d = {K_{eq}}^{-1} \tag{70}$$

the parameter $K_{1.1} = 5.6 \cdot 10^{-9}$ M = 5.6 nM. Although in the same article they also showed that single-site mutations could cause a change in the dimerization dissociation constant up to a value of $K_{1.1} = 2.7$ µM [61]. As this large interval has been shown experimentally it was assumed resonable to keep using the $K_{1.1}$-value used in Ghim and Almaas (2009) [2] where $K_{1.1}$ = 10 nM. As no dimerisation constant for TetR was found it was assumed to be equal to the one for cI, yielding $K_{2.1} = 10$ nM.

3.8.2 Binding of the cI repressor to DNA — K_2

Sauer reported in 1979 that $K_d = 20$ nM for cI_2 dissociation from the O_{R1} operator site. However, this reference was not avilable, but several others have cited the same value [46, 62]. The value was assumed to be the same for O_{L1} the operator site in the P_L-promoter, which gives a $K_2 = 20$ nM.

Although, the assumption that $K_{1.2} = K_{2.2}$ is not really valid when comparing cI to the Tet repressor (TetR). The TetR have a dissociation constant of $K_d = 0.005$ nM [63], being a 4000-fold difference to the one reported for cI.

As cI was chosen as repressor 1, it will bind to promoter 2, thereby the rates for the model became $K_{2.2} = 20$ nM and $K_{1.1} = 0.005$ nM.

3.8.3 Co-operative binding to the second operator site — r and K_4

According to Johnson *et al.* (1979) the co-operativity between two cI repressor dimers when binding to O_{R1} and O_{R2} will decrease the two dissociation constants with a factor of two and 12.5 respectively [64]. As the model is made for sequential binding to the two binding sites these two factors are combined to one co-operative factor $r_2 = 25$. As $r = \frac{K_2}{K_4}$ this gives $K_{2.4} = 0.8$ nM.

In order to explain the operation of the Tn*10* regulon, regulated by TetR, it is not necessary to include cooperativity to the binding of the repressor to the two repressor sites. This gives $r_1 = 1$ giving $K_{1.4} = K1.2 = 0.005$ nM.

3.8.4 Binding of the RNAp to the P_L-promoter and transcription initation — K_3 and α_m

Giladi *et al.* (1990) reported both the equilibrium constant for RNAp binding to the P_L-promoter and the forward rate constant for the isomerization of closed to open RNAp-promoter complexes. The reported values were $8.94 \cdot 10^7$ M and $4.38 \cdot 10^{-3}$ s^{-1} respectively. These values can be used to estimate the parameters $K_{1.3} = 11.2$ nM and $\alpha_{1.m} = 6.6 \cdot 10^{-3}$ nM s^{-1} [65]. The $K_{1.3}$-parameter was estimated using the relationship from Eq. (70), while the $\alpha_{1.m}$-parameter was estimated based on the existence of one promoter in one *E. coli*-cell with average cell volume on 1.1 µm^3 [66].

As the planned second promoter is P_L*tet-O* both values can be assumed to be equal for the second promoter as well, which gives $K_{2.3} = 11.2$ nM and $\alpha_{2.m} = 6.6 \cdot 10^{-3}$ nM s^{-1}.

3.8.5 Rate of transcription — α'_m

The rate of transcription in *E. coli* is usually 50–90 bp/s. Assuming an average transcript length of 1000 bp, and an average transcription rate of 70 bp/s, gives $\alpha'_m = 0.07$ s^{-1} [6].

3.8.6 Rate of translation — α_p

The rate of translation have been determined to about 17 amino acids/second (aa/s) [67]. Although, as the step of ribosome binding to the mRNA is not included in this model the translation rate is assumed to be 15 aa/s. As the coding regions of TetR and λ cI is 660 and 750 base pairs (bp) respectively [68], the translation rates becomes $\alpha_{1.p} = 0.068$ and $\alpha_{2.p} = 0.060$ s^{-1}.

3.8.7 Rate of mRNA degradation — γ_m

An average half-life for mRNA in *E. coli* have been determined to 3.69 min [69]. In lack of more accurate data this was assumed to be an average value for both mRNA transcripts, giving $\gamma_m = 0.0031\ \mathrm{s}^{-1}$.

3.8.8 Rate of protein degradation — γ_p

The rate of protein degradation was assumed to be the same for both transcripts and in correspondence with previous modelling of the λ-phage switch [46], giving $\gamma_p = 0.0007\ \mathrm{s}^{-1}$.

3.8.9 Concentration of free RNAp — [RNAp]

The concentration of free RNAp was assumed to be the same as for previous modelling of the λ-phage switch [46], [RNAp] = 30 nM.

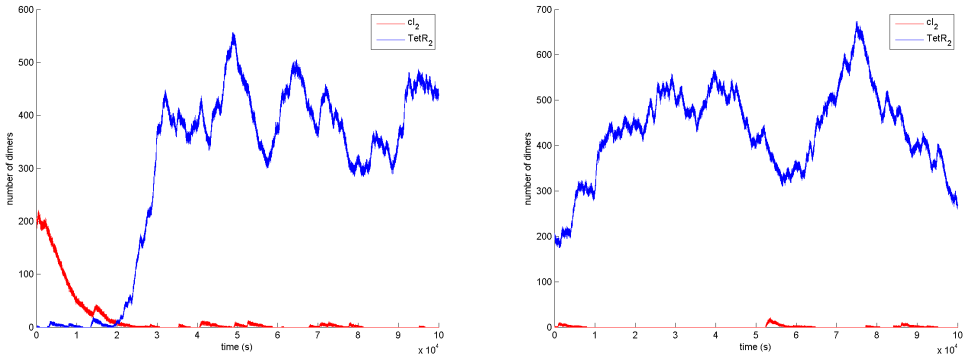
3.8.10 Remaining parameters — s, K_5, K_7, σ

Due to limited time the remaining parameters were chosen as in Ghim and Almaas (2009), meaning a monomer to dimer lifetime rato $\sigma = 10$ and the leakage parameter $s = 0.01$. The leakage parameter is defined as $s = K_3/K_5$, thereby giving a $K_5 = 1100$ nM [2].

3.9 Stocahstic analysis of the complete circuit

A Dizzy-model file was created with parameter-values according to Sec. 3.8. Two simulations were performed having different initial conditions, having the cI-repressor being turned on or off as shown if Figs. 25a and 25b respectively. The simulations of model files 'cItetRfullcIon_notetrcoop.cmdl' and 'cItetRfullcIoff_notetrcoop.cmdl' were performed with the gibson-bruck algoritm, start = 0, stop = 100000, number of result points = 100000 and stochastic ensemble size = 1. Both the model files and the files containing the data results are in the attached zip-file folder StochasticSimulations/Full.

It seems from Fig. 25 as the circuit is dominated by the presence of TetR_2 which will after a while control the switch, regardless of the circuits initial state.



(a) (b)

Figure 25: **Stochastic simulation result plot of complete circuit.** (a) The simulation started with cI_2 initially being in high state, although as time in the simulation pass the cI_2-concentration decrease, and the $TetR_2$ takes over as active repressor. (b) The simulation started with cI_2 initially being in low state, and staying that way for the entire simulation.

4 Discussion

4.1 General description of the circuit

What level of detail used when making a model depends on the level of accuracy sought after and what parameters that are reasonable to include. For instance, in the model described in this thesis it was assumed average translation rates, instead of incorporating all the different amino acids addition to the growing polypeptide chain. Although this model incorporates most of the processes at the protein level, the folding of the peptide chains to active proteins is not included. This is a step that has previously been reported to have a good impact on improving the accuracy of a model [22]. Additionally, the model could be improved by adding a step for ribosome binding to transcripts before the translation process.

In order to make the model better suited for switching, an inducer could be included to each of the two repressors. This would enable simulations with inducers to investigate whether some of the less stable steady states (regions where switching occurred) could be more stable if inducer was added to the system, restricting the binding of the induced repressor at its operator sites.

4.2 Validity of assumptions

Most of the assuptions made in Sec. 3.2 were made so that the nondimensionalisation of the system would become easier. Here it was suggested using LVA-tail to both the repressors in order to make sure the γ_p is very similar in both cases. Although, LVA-tails are not ideal for systems with low copy number, *i.e.* $\sum D = 1$, giving a 30-fold increase in the degradation rate [70]. As the degradation rate of the proteins, γ_p, is inversely proportional to the β-parameters, this gives an unfortunate effect to the range of the leakage parameters, s_1 and s_2, giving bistability, as can be observed in Fig. 18. The discrepancies between the native degradation rates are not necessarily too different, so this could be a reasonable assumption even without the LVA-tail.

The assumption that $K_{1.2} = K_{2.2} = K_2$ was not necessarily a good one, as binding affinities of repressors to DNA vary widely. For instance the TetR repressor have reportedly a $K_D = 0.005\,\text{nM}$ [63], about 4000 times lower than the one reported for cI [61]. However, this property can be modified by manipulation of both the operator sequences and the coding sequences of the repressors in order to match the model more tightly.

The assumption that $K_5 = K_7$ can be considered somewhat reasonable as both constants probably have quite high values. Assuming an average for

both, somewhat higher for K_5 and lower for K_7, would give a decent picture of the system.

The concentration can seemingly be considered to be in constant supply, as have been done previously [46, 2]. This also seems reasonable as it is indispensable for the cell, being responsible for the extremely important process of transcription.

4.3 Stochastic analysis compared to deterministic

From Figs. 20 and 22 it was observed that when the P_1-expression is repressed (switched off) there are no dimers present, in correspondence with the bifurcation plots giving rise to the stability diagrams, in Sec. 3.5. Although these bifurcation plots were deterministic and gave rise to a steady state concentration for the dimers, these were all <1, meaning they were virtually not present.

The stochastic analysis gave a narrower bistable area than what was predicted using the deterministic approach. This could be explained by the nature of the bifurcation giving rise to the bistability. Most of the bifurcations for the asymmetrical model are saddle-node bifurcations, as described in Sec. 1.7.3. As the P_2^1-concentration is lowered with a decreasing β-value the stable steady state is approaching an unstable steady state (a saddle point), while the "unoccupied" steady state for P_2^1 is getting further away from its unstable steady state. A stochastic fluctuation pushing either one of the two variables beyond the saddle point would help pushing the system towards the other stable steady state. Hence, a stable steady state lying close to an unstable one is less stable than a stable steady state far away from the unstable steady state.

4.4 Full circuit simulations

The modelling of the full circuit with all steps included did not show a bistable system. This could be explained by the fact that many of the parameters were not properly accounted for. Furthermore, the repressor-promoter dissociation constant for TetR, $K_{1.2}$, was estimated for values for its operator sequences in its native system, the Tn*10* regulon [63]. In the planned system the operator sequences are incorporated with the P_L-promoter from the bacteriophage λ, the $P_L tetO-1$ [1, 60]. It can be assumed that the dissociation constant to the operator sequences in this promoter is not as low as to the native operator sequences.

The lack of bistability to the full system is in correspondence with the approximative model, as the K_2 parameter is inversely proportional to the

β parameter. The β-value is extremely high, $\beta_1 = 36000$, and far outside of the bistable region depicted in Fig. 17.

4.5 Future prospects

As to this model itself there are many interesting features that could be tested. For instance one could try adding a folding step to the model, as suggested in Sec. 4.1, to increase accuracy to models. Another interesting feature would be to change the copy number to a somewhat higher value, as many expression vectors have high copy numbers.

Switches have been proposed to be used as submodules to many more complex circuits, such as bioswiches and digital counters [3], and investigations to check if the model is able to predict these kind of patterns putting several switches together could show very profitable.

Last but not least, experimental testing of the model is planned in correspondence with NTNU researcher Rahmi Lale, PhD, and Professor Eivind Almaas. As the effective potential landscape illustrated in Fig. 24 was clearly double-welled (bistable) one would expect a combination of cI and TetR as repressors having the right ribosomal binding sites (RBS) could function as a switch. Although, as mentioned in Sec. 3.6 there is probably not a linear relationship between the expression and leakage from high copy number plasmids (as applied by Gardner *et al.* [1]) and low copy numbers. Therefore, prior to actually putting the repressor genes together to a switching module, the expression and leakage should be tested using either low copy number plasmids or chromosomal integration of the genes and promoters. These results could give further directions to if the switch would function properly or not.

5 Conclusion

The HOM2-circuit approximation for a bistable tuneable switch from Ghim and Almaas (2009) was re-derived in order to make it asymmetrical. Deterministic analysis was conducted yielding stability diagrams, describing the phase plane showing bistability for the genetic switch. Furthermore, stochastic simulations of the approximation were conducted. This gave a somewhat narrower bistable area than the deterministic analysis, possibly due to the nature of saddle-node bifurcations. Parameter values for a switch based on experiments were estimated for the approximation, and these were used in a stochastic simulation. The result from this simulation was in correspondence with the deterministic analysis. A stochastic simulation of the complete circuit was conducted based on parameter values found in literature. For this simulation bistability was not shown.

In order to further explore the circuit, and validity of the approximation, experimental investigation is needed. This has been planned together with Rahmi Lale, PhD, and Professor Eivind Almaas at the Department of Biotechnology NTNU.

References

[1] Gardner T.S., Cantor C.R., and Collins J.J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, **403**:339–42, 2000.

[2] Ghim C.-M. and Almaas E. Two-component genetic switch as a synthetic module with tunable stability. *Pys. Rev. Letters*, **103**:028101, 2009.

[3] Lu T.K., Khalil A.S., and Collins J.J. Next-generation synthetic gene networks. *Nature biotechnolgy*, **27**:1139–50, 2009.

[4] Klug W.S., Cummings M.R., Spencer C.A., and Palladino M.A. *Concepts of Genetics*. Pearson Education,Inc., publishing as Pearson Benjamin Cummings, San Francisco, 9th edition, 2009. 779 pp.

[5] Institute for Systems Biology (2010). *Institute for Systems Biology: Systems Biology – the 21st Century Science* [online]. source: Institute for Systems Biology. Available from: `http://systemsbiology.org/Intro_to_ISB_and_Systems_Biology/Systems_Biology_--_the_21st_Century_Science` [Downloaded 20. November 2010].

[6] Nelson D.L. and Cox M.M. *Principles of Biochemistry*. W.H. Freeman and Company, New York, 5th edition, 2008.

[7] Wiener N. *Cybernetics : or Control and communication in the animal and the machine*. The technology press, Cambridge, Mass, 1st edition, 1948. 194 pp.

[8] Monod J. and Jacob F. General conclusions: Telenomic mechanisms in cellular metabolism, growth and diffrentiation. *Cold Spring Harbor symposia on quantitative biology*, **26**:389–401, 1961.

[9] nobelprize.org (2010). *"The Nobel Prize in Physiology or Medicine 1965"* [online]. source: nobelprize.org. Available from: `http://nobelprize.org/nobel_prizes/medicine/laureates/1965/` [Downloaded 22. November 2010].

[10] Kauffman S. The large scale structure and dynamics of gene control circuits: An ensemble approach. *Journal of theoretical Biology*, **44**:167–90, 1974.

[11] Glass L. Combinatorial and topological methods in nonlinear chemical kinetics. *The Journal of chemical physics*, **63**:1325–35, 1975.

[12] Kitano H. Computational systems biology. *Nature*, **420**:206–10, 2002.

[13] NCBI (2009). *GenBank Growth* [online]. source: NCBI. Available from: `http://www.ncbi.nlm.nih.gov/genbank/genbankstats.html` [Downloaded 17. November 2010].

[14] Palsson B.Ø. *Systems Biology - Properties of Reconstructed Networks*. Cambridge University Press, New York, 1 st edition, 2005. pp 322.

[15] Kærn M., Elston T.C., William J.B., and Collins J.J. Stochastisity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, **6**:451–64, 2005.

[16] Elowitz M.B., Levine A.J., Siggia E.D., and Swain P.S. Stochastic gene expression in a single cell. *Science*, **297**:1183–86, 2002.

[17] Elowitz M.B. and Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**:335–38, 2000.

[18] Ramsey S., Orrell D, and Bolouri H. Dizzy: stochastic simulation of large-scale genetic regulatory networks. *Journal of Bioinformatics and Computational Biology*, **3**:415–36, 2005.

[19] Shannon P., Markiel A., Ozier O., Baliga N.S., Wang J.T., Ramage D., Amin N., Schwikowski B., and Ideker T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, **13**(11):2498–2504, 2003.

[20] Hofmeyr J.-H.S. Olivier B.G., Rohwer J.M. Modelling cellular systems with pysces. *Bioinformatics*, **21**:560–61, 2005.

[21] 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, and J. Wang. The systems biology markup language (sbml): A medium for representation and exchange of biochemical network models. *Bioinformatics*, **19**(4):524–531, 2003.

[22] Stricker J., Cookson S., Bennett M.R., Mather W.H., Tsimring L.S., and Hasty J. A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**:516–19, 2008.

[23] Madigan M.T., Martinko J.M., Dunlap P.V., and Clark D.P. *Brock - Biology of Microorganisms*. Pearson Benjamin Cummings, San Francisco, 12th edition, 2009. 1061 pp.

[24] Alberts B., Johnson A., Lewis J., Raff M., Roberts K., and Walter P. *Molecular Biology of THE CELL*. Garland Science, New York, 5th edition, 2008. 1268 pp.

[25] Serrano L., Vancompernolle K., Ball P., Benenson Y., Holliger P., Panke S., Rettberg R. Stelling J., Weiss R., and Krassnig C. (2005). *Synthetic Biology - Applying Engineering to Biology* [online]. source: European Commission. Available from: `ftp://ftp.cordis.europa.eu/pub/nest/docs/syntheticbiology_b5_eur21796_en.pdf` [Downloaded 16. February 2011].

[26] Levskaya A., Chevailer A.A., Tabor J.J., Simpson Z.B., Lavery L.A., Levy M., Davidson E.A., Scouras A., Ellington A.D., Marcotte E.M., and Voigt C.A. Engineering *Esherichia coli* to see light. *Nature*, **438**:441–42, 2005.

[27] Gibson D.G., Glass J.I., Lartigue C., Noskov V., Chuang R.-Y., Algire M.A., Benders G.A., Montague M.G., Ma L., Moodie M.M., Merryman C., Vashee S., Krishnakumar R., Assad-Garcia N., Adrews-Pfannkoch C., Denisova E.A., Young L., Qi Z.-Q., Segall-Shapiro T.H., Calvey C.H., Parmar P.P., Hutchinson III C.A., Smith H. O., and Venter J.C. Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science*, **329**:52–56, 2010.

[28] Carr P.A. and Church G.M. Genome engineering. *Nature biotechnology*, **27**:1151–62, 2009.

[29] Wang H.H., Isaacs F.J., Carr P.A., Sun Z.Z., Xu G., Forest C.R., and Chruch G.M. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, **460**:894–98, 2009.

[30] Norvig P., Relman D.A., Goldstein D.B., Kammen D.M., Weinberger D.R., Aiello L.C., Church G., Hennessy J.L., Sachs J., Burrows A., Pisano G.P., Goldstein J.R., Anastas P., Klausner R., Baltimore D., Montgomery D.R., Baer T.M., Bigelow N.P., Holt R.D., and Nicholson J.K. 2020 visions. *Nature*, **463**:26–32, 2010.

[31] Ro D.-K., Paradise E.M., Ouellet M., Fisher K.J., Newman K.L., Ndungu J.M., Ho K.A., Eachus R.A., Ham T.S., Kirby J., Chang M.C.Y., Withers S.T., Shiba Y., Sarpong R., and Keasling J.D. Production of the antimalarial drug precursor atremisnic acid in engineered yeast. *Nature*, **440**:940–43, 2006.

[32] Kwok R. Five hard truths for synthetic biology. *Nature*, **463**:288–90, 2010.

[33] Endy D. Foundations for engineering biology. *Nature*, **438**:449–453, 2005.

[34] OECD (2010). *Symposium on Opportunities and Challenges in the Emerging Field of Synthetic Biology* [online]. source: OECD Publishing. Available from: `http://dx.doi.org/10.1787/9789264086265-en` [Downloaded 14. April 2011].

[35] ung.igem.org (2010). *"About - ung.igem.org"* [online]. source: ung.igem.org. Available from: `http://ung.igem.org/About` [Downloaded 08. March 2011].

[36] 2010.igem.org (2010). *"Sloveina"* [online]. source: 2010.igem.org. Available from: `http://2010.igem.org/files/poster/Slovenia.pdf` [Downloaded 08. March 2011].

[37] 2010.igem.org (2010). *"KIT - Kyoto"* [online]. source: 2010.igem.org. Available from: `http://2010.igem.org/files/poster/KIT-Kyoto.pdf` [Downloaded 08. March 2011].

[38] Khalil A.S. and Collins J.J. Synthetic biology: applications come of age. *Nature Reviews Genetics*, **11**:367–79, 2010.

[39] van der Meer J.R. and Belkin S. Where microbiology meets microengineering: design and applications of reporter bacteria. *Nature Reviews Microbiology*, **8**:511–521, 2010.

[40] Fraser P.D, Enfissi E.M.A., and Bramley P.M. Genetic engineering of carotenoid formation in tomato fruit and the potential application of systems and synthetic biology approaches. *Archives of Biochemistry and Biophysics*, **483**(2):196–204, 2009.

[41] Karlebach G. and Shamir R. Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology*, **9**:770–80, 2008.

[42] Pandey S., Wang R.-S., Wilson L., Li S., Zhao Z., Gookin T.E., Assmann S.M., and Albert R. Boolean modeling of transcriptome data reveals novel modes of heterotrimeric G-protein action. *Molecular Systems Biology*, **6**:372, 2010.

[43] Chaves M. and Albert R. Studying the effect of cell division on expression patterns of the segment polarity genes. *Journal of the Royal Society Interface*, **6**:5, 2008.

[44] Conrad E.D and Tyson J.J. Modeling molecular interaction networks with nonlinar ordinary diffrential equaitons. In Szallasi Z., Stelling J., and Periwal V., editors, *System Modelling in Cellular Biology*, chapter 6, pages 97–123. The MIT Press, 2006.

[45] Davidson E.H., Rast J.P., Oliveri P., Ransick A., Calestani C., Yuh C.-H., Minokawa T., Amore G., Hinman V., Arenas-Mena C., Otim O., Brown C.T., Livi C.B., Lee P.Y., Revilla R., Rust A.G., Pan Z., Schilstra M.J., Clarke P.J.C., Arnone M.I., Rowen L., Cameron R.A., McClay D.R., Hood L., and Bolouri H. A genomic regulatory network for development. *Science*, **295**:1669–78, 2002.

[46] Arkin A., Ross J., and McAdams H.H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics*, **149**:1633–48, 1998.

[47] Gillespie D.T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal Computational Physics*, **22**:403–34, 1976.

[48] Gibson M.A. and Bruck J. Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry*, **104**:1876–1889, 2000.

[49] Rinaudo K., Bleris L., Maddamsetti R., Subramanian S., Weiss R., and Benenson Y. A universal RNAi-based logic evaluator that operates in mammalian cells. *Nature Biotechnology*, **25**(7):795–801, 2007.

[50] Win M.N. and Smolke C.D. Higher-order cellular information processing with synthetic RNA devices. *Science*, **322**:456–460, 2008.

[51] Sohka T., Heins R.A., Phelan R.M., Greisler J.M., Townsend C.A., and Ostermeier M. An externally tunable bacterial band-pass filter. *Proceedings of the National Academy of Sciences of the USA*, **106**(25):10135–10140, 2009.

[52] Kitano H. Biolgical robustness. *Nature Reviews Genetics*, **5**:826–37, 2004.

[53] Kitano H. Systems biology: A brief overview. *Science*, **295**:1662–64, 2002.

[54] Ozbudak E.M., Thattai M., Kurtser I., Grossman A.D., and Van Oudenaarden A. Regulation of noise in the expression of a single gene. *Nature Genetics*, **31**(1):69–73, 2002.

[55] Rao C.V., Wolf D.M., and Arkin A.P. Control, exploration and tolerance of intracellular noise. *Nature*, **420**:231–10, 2002.

[56] Strogatz S.H. *Nonlinear Dynamics and Chaos*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1 st edition, 1994. pp 498.

[57] Ludwig D., Aronson D.G., and Weinberger H.F. Spatial patterning of the spruce budworm. *Journal of Mathematical Biology*, 8(3):217–258, 1979.

[58] Mathworks (2011). *Solve system of non-linear equations* [online]. source: mathworks.com. Available from: `http://www.mathworks.com/help/toolbox/optim/ug/fsolve.html` [Downloaded 08. May 2011].

[59] Karzai A.W., Roche E.D., and Sauer R.T. The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. *Nature Structural Biology*, **7**(6):449–55, 2000.

[60] Lutz R. and Bujard H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucleic Acids Research*, **25**(6):1203–10, 1997.

[61] Burz D.S., Beckett D, Benson N., and Ackers G.K. Self-assembly of bacteriophage λ cI repressor: Effects of single-site mutations on the monomer-dimer equilibrium. *Nature Structural Biology*, **7**(6):449–55, 1994.

[62] Hawley D.K., Johnson A.D., and McClure W.R. Functional and physical characterization of transcription initiation complexes in the bacteriophage λ o_r region. *The Journal of Biological Chemistry*, **260**(14):8618–26, 1985.

[63] Hillen W. and Berens C. Mechanisms underlying expression of tn*10* encoded tetracycline resistance. *Annual Review of Microbiology*, **48**:345–69, 1994.

[64] Johnson A.D., Meyer B.J., and Ptashne M. Interactions between dna-bound repressors govern regulation by the λ phage repressor. *Proceedings of the National Academy of Sciences of the USA*, **76**(10):5061–65, 1979.

[65] Giladi H., Gottesman M., and Oppenheim A.B. Integration host factor stimulates the phage lambda pl promoter. *Journal of Molecular Biology*, **213**:109–21, 1990.

[66] Kubitschek H.E. and Friske J.A. Determination of bacterial cell volume with the coulter counter. *Journal of Bacteriology*, **168**(3):1466–7, 1986.

[67] Young R.Y. and Bremer H. Polypeptide-chain-elongation rate in *Escherichia coli* b/r as a function of growht rate. *Biochemical Journal*, **160**:185–94, 1976.

[68] The Registry of Standard Biological Parts (2011). *Protein coding sequences/Transcriptional regulators* [online]. source: partsregistry.org. Available from: `http://partsregistry.org/Protein_coding_sequences/Transcriptional_regulators` [Downloaded 07. May 2011].

[69] Bernstein J.A., Lin P.-H., Cohen S.N., and Lin-Chao S. Global analysis of *Escherichia coli* rna degradosome function using dna microarrays. *Proceedings of the National Academy of Sciences of the United States of America*, **101**:2758–2763, 2004.

[70] Lies M. and Maurizi M.R. Turnover of endogenous ssra-tagged proteins mediated by atp-dependent proteases in *Escherichia coli*. *The Journal of Biological Chemistry*, **283**(34):22918–29, 2008.