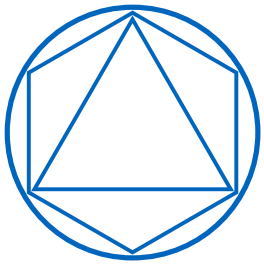

Technische Universität München
Fakultät für Mathematik

Modelling the Influence of the Anti-Activator System on Quorum Sensing in Pseudomonas aeruginosa

Masterarbeit von Philipp Städter

Aufgabensteller:	Prof. Dr. Christina Kuttler
Betreuer:	Prof. Dr. Martin Schuster
Abgabetermin:	

Ich erkläre hiermit, dass ich diese Arbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

Ort Datum Unterschrift

Zusammenfassung auf Deutsch

Summary

Contents

1 Biological Background **2**

2 Mathematical Theory **5**
2.1 Law of mass-action . . . 5
2.2 Michaelis-Menten Kinetics . . . 6
2.3 Time-Scale-Analysis . . . 9
2.4 Relaxation Oscillations . . . 11
2.5 Parameter estimation . . . 13
2.6 Optimisation . . . 16
2.7 Numerical Integration of ODEs . . . 20

3 Mathematical approach to analyse the biological circuit of *Pseudomonas aeruginosa* **23**
3.1 Mathematical model . . . 23
3.2 Model analysis . . . 27
3.2.1 Basic model . . . 27
3.2.2 Extended QslA model . . . 34
3.2.3 Extended QteE model . . . 45
3.2.4 Full model . . . 51
3.3 Stability analysis . . . 51
3.3.1 Basic model . . . 52
3.3.2 Extended QslA model . . . 55
3.4 Analytical limiting processes . . . 59

4 Estimation of all unknown parameters **61**
4.1 The complexity of parameter estimation . . . 66
4.2 Expanding the full model . . . 68
4.2.1 Double LasRd protein binding to the lasI gene . . . 68
4.2.2 RsaL pathway extension . . . 70
4.3 Comparison to Goryachev et al. . . . 74
4.4 Fitting all parameters . . . 81
4.5 The reliability of all estimated parameter values . . . 85

5 Discussion and Outlook **89**

6 Implementation **90**

1 Biological Background

Prokaryotic cells are one of the first living organisms on our planet [20]. Broadly speaking, they consist of circular-shaped DNA and ribosoms, embedded in cytoplasm, a cell membrane and cell wall, as well as multiple flagella. Thus, procaryotic cells can adapt to various situations by synthesising proteins. One integral sub-class of procaryotic cells are bacteria. They are considered to be precursors of special components in eukaryotic cells, laying the foundation of life as we know it today by enabling photosynthesis via chloroplasts or providing energy in form of ATP via mitochondria [20]. Through their adaptability, billions of subspecies can be found in almost any habitat making them one of the largest population class on earth.
But bacteria are not only responsible for the creation of today's eucaryotic cells. Many species interact with animals such as humans in a symbiotic way. Famous examples are bacteria living in our gut flora to help us digest food or bacteria living on our skin to protect us from other microorganisms [12]. Additionally, many bacteria are quite harmless due to our advanced immune system. However, some species of bacteria have pathogenic effects leading to infections of vital organs such as the heart or the lung. The latter being the case for the common bacteria *Pseudomonas aeruginosa* which induces in particular pneumonia. One of the reasons why an infection by these types of bacteria is that severe is because of their tendency to attack the immune system only after they exist in a density high enough to cause a super-infection [4].

Recently, it has been discovered that bacteria can regulate their own cell density through cell-cell communication which we call Quorum Sensing [11]. Hereby, through a complex network of genes, proteins and metabolites, cells produce autoinducer proteins that are then being transported through the cell membrane to the extracellular space. Nearby cells can take up those proteins which increase their own gene expression to form more autoinducer molecules themselves.
Through extensive experimental research a core biological circuit network in *Pseudomonas aeruginosa* has been worked out which represents its Quorum Sensing mechanism (Figure 1.1). Biologically speaking, the network can be divided into two main components, a combination of pathways which increase autoinducer production and two pathways that inhibit it. On the one hand, the transcription and translation of the lasI gene leads to LasI protein that enables enzymatic transformation of substrate into intracellular 3-oxo-C12-HSL autoinducer molecules. Through a transmembrane process those autoinducer molecules are transported into the extracellular space which can than be taken up by other cells. Once inside the cell, autoinducer molecules form a complex with LasR protein into the autoinducer-LasR heterodimer LasR*. This heterodimer then dimerizes into the homodimer LasRd which increases the gene expression of the

LasI gene. Meanwhile, both heterodimer and homodimer can dissociate into their singular proteins. In total, this results in a so-called positive feedback loop. On the other hand, the protein QteE and homodimer QslAd function as gene expression inhibitors by binding the LasR protein and autoinducer-LasR complex and inactivating them. Moreover, it has been established that the LasR dimer also regulates the gene expression of the qteE gene leading to a so-called negative feedback loop [1]. Therefore, both underlying genes are called anti-activator genes. Here, the question arises why bacteria have genes that inhibit their own cell-cell communication? Theories suggest that bacteria can avoid intracellular self-activation while being able to benefit from delaying Quorum Sensing to slowly increase their population density [1].

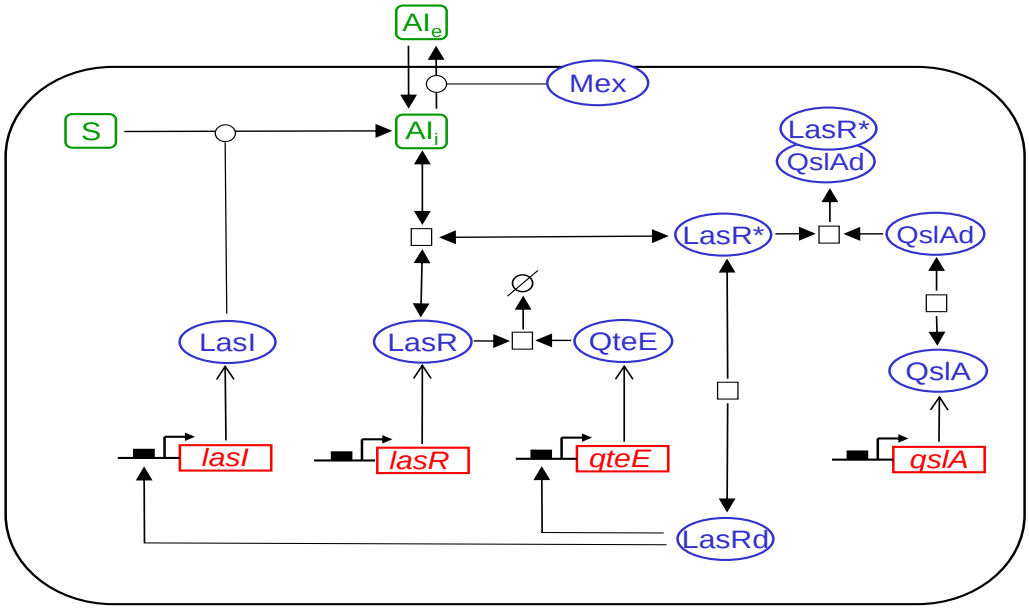

Figure 1.1: Inner biological circuit showing the connection of all important genes and their gene expression. Red rectangles represent mRNA, blue ellipses depict proteins and green squares portray metabolites. Furthermore, arrows represent the flow of information while small squares mark the biomolecular reactions and small circles stand for enzymatic activation. Lastly, the greek letter ϕ depicts inactivation and degradation of the protein complex.
This figure is directly taken from (citation?).

After a super-infection where pathogenic bacteria stress or even overwhelm the immune-system, often a medical solution such as an antibiotics helps to overcome the infection and aid our natural defense mechanism. These medications inhibit the bacteria's proliferation or kill them while aiming to preserve uninfected tissue [12]. However, besides negatively affecting surrounding healthy tissue, a severe side effect of antibiotics is the

possibility for bacteria to adapt and develop a resistance against them. This is also the case for *Pseudomonas aeruginosa* [4]. Roughly speaking, this happens due to transposable DNA elements which amplify enzyme production against antibiotics by copying and randomly moving genetic material [12]. Furthermore, is has been suggested that Quorum Sensing is even advantageous for their survival under antibiotic treatment by introducing a stochastic (maybe it doesn't have to be stochastic!?) effect on the interaction between bacteria [11]. Hence, an alternative and very effective way of treating a bacterial infection could be to make use of anti-activators to inhibit or at least control Quorum Sensing [10]. This could allow a beneficial decrease in gene expression of autoinducer-production-genes to avoid an accumulation of too many bacteria while preserving healthy tissue. This could be done by decreasing the formation of autoinducer molecules or by increasing the production of anti-activator proteins.

Hence, a mathematical analysis of the temporal progression of autoinducer concentration inside and outside of the cell such as the effect of specific anti-activator genes on autoinducer production can be insightful. In this thesis, a mathematical model for the portrayed circuit network of *Pseudomonas aeruginosa* is formulated and analytically as well as numerically analysed using various mathematical tools.

2 Mathematical Theory

2.1 Law of mass-action

Given the biochemical reaction network in Figure 1.1, the question arises how to translate it into mathematical language with all its components such as reactants, products and reaction rates. Here and later on reactants and products are referred to as species.
First of all, the most general reaction of a complex biochemical network of species X_i, $\forall i \in \{1, ..., n_X\}$, can be visualized as follows

$$\sum_{i=1}^{n_X} \mu_{ij} X_i \quad \underset{k_{+j}}{\overset{k_{-j}}{\rightleftharpoons}} \quad \sum_{i=1}^{n_X} \nu_{ij} X_i \tag{2.1}$$

where $1 \leq j \leq k_R$, $k_R, n_X \in \mathbb{N}$.
Furthermore, $\mu_{ij}, \nu_{ij} \in \mathbb{N}_0$ are the stoichiometric coefficients of the species X_i in the reaction R_j with reaction rates $k_{+j}, k_{-j} \in \mathbb{R}_+$.
It is important to mention that two different species can react to a species which was one of the original reactants. Therefore, a species X_i for some $i \in n_X$ can be reactant and product at the same time. Only the stoichiometric coefficients represent mathematically whether such a reaction exists - e.g. $\mu_{i,j} \neq 0$ for some i, j - or does not exist - e.g. $\mu_{i,j} = 0$ for some i, j - in the network.

The integral part of transforming reaction (2.1) into mathematical formalism is called the **Law of Mass Action**.

Assumption 2.1. The rate of molecular collisions of two chemical species in a dilute gas or solution is proportional to the product of the two concentrations [6].

Let furthermore $x_i = [X_i]$ be the concentration of the species on which the mathematical model will be build upon.
Then, by applying Definition 2.1 to reaction (2.1) one gets the frequency of the reaction

$$f_j(x) := k_{+j} \prod_{i=1}^{n_X} x_i^{\mu_{ij}} - k_{-j} \prod_{i=1}^{n_X} x_i^{\nu_{ij}} \tag{2.2}$$

Still, one needs to account for the sign of a reaction, i.e. if products are produced more than they are depleted or the other way around. This leads to the following expression

$$S_{ij} f(x) := (\nu_{ij} - \mu_{ij}) f_j(x) \tag{2.3}$$

where $S_{ij} \in \mathbb{R}^{n_X \times k_R}$ is the stoichiometric matrix.

For a complete characterization of the network, one does not only need the frequency of one isolated reaction R_j but the combination of all of them given for each concentration x_i. Thus, the discrete translation for the network in two states - at time t and after time Δt in state $t + \Delta t$ - is given by

$$x_i(t + \Delta t) = x_i(t) + \sum_{j=1}^{k_R} (\nu_{ij} - \mu_{ij}) f_j(x) \Delta t \tag{2.4}$$

Here, we have the interval $[t, t + \Delta t] \subset \mathbb{R}$ in which $\Delta t > 0$ arbitrarily small marks a change of the network from the first state to the second one by letting the species react to products at their reaction frequency.

Although the system of equations (2.4) is a nice representation of all states of the biochemical reaction network after equidistant time steps, it is not a sufficient biological description since most chemical reactions in nature change their states continuously over time.
One can account for that by rearranging the last equation

$$(2.4) \qquad \Longleftrightarrow \qquad \frac{x_i(t + \Delta t) - x_i(t)}{\Delta t} = \sum_{j=1}^{k_R} (\nu_{ij} - \mu_{ij}) f_j(x)$$

and then taking the limit $\Delta t \to 0$

$$\dot{x}_i(t) = \sum_{j=1}^{k_R} (\nu_{ij} - \mu_{ij}) f_j(x) \tag{2.5}$$

Visually speaking, the step-size Δt between different states will be infinitely small such that infinitely many discrete states appear as one continuous representation of the species' course over time. This results into a system of $n_X \in \mathbb{N}_0$ distinct **Ordinary Differential Equations (ODEs)**, one for each species that is described in the network.
The mathematical theory described in this section can be found in [13].

2.2 Michaelis-Menten Kinetics

This master thesis aims at describing a specific biochemical network where DNA of certain genes is transcribed to mRNA which itself is translated to proteins. Under dimerization and dissociation these proteins form different complexes which regulate among under genes their own gene expression. This network is thus called a regulatory network. To translate this network into mathematical language the theory of mass

action is not enough since it only covers protein reactions but not the gene-mRNA-protein pathway. Here, the theory of **Michaelis-Menten** is used to describe enzyme kinetics and can be adapted to model the dynamics of our pathways.

Definition 2.2. At the most basic level, in a reaction using substrate and enzyme, the substrate binds reversibly to the enzyme building a substrate-enzyme complex which is then converted irreversibly into a product and the original enzyme. Hence, the total concentration of unbound and complex-bound enzyme is constant [6].

A proper reaction network describing this condition is given by

$$S + E \quad \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} \quad SE \quad \xrightarrow{k_{+2}} \quad P + E \tag{2.6}$$

where S, E, SE and P represent the substrate, enzyme, substrate-enzyme complex and product respectively. By using the law of mass action as described in section 2.1 one gets a system of ODEs.
Therefore, let $s = [S], e = [E], c = [C], p = [P]$ be the concentrations of the species, it is

$$\begin{aligned} \dot{s} &= -k_{+1}se + k_{-1}c \\ \dot{e} &= -k_{+1}se + k_{-1}c + k_{+2}c \\ \dot{c} &= k_{+1}se - k_{-1}c - k_{+2}c \\ \dot{p} &= k_{+2}c \end{aligned} \tag{2.7}$$

By Definition 2.2, all enzyme is preserved, i.e. $e + c = e_0 \equiv$ constant. For simplicity, one can omit the equation for $\dot{e}$ since it is implicitly given by $\dot{e} = -\dot{c}$. Thus, one gets

$$\begin{aligned} \dot{s} &= -k_{+1}se_0 + (k_{+1}s + k_{-1})c \\ \dot{c} &= k_{+1}se_0 - (k_{+1} - k_{-1} - k_{+2})c \\ \dot{p} &= k_{+2}c \end{aligned} \tag{2.8}$$

At this point we can use a heuristic argument to get the solution in a faster and more understandable way, namely that after a sufficient period of time the concentration of the substrate-enzyme complex is in equilibrium. Hence we can assume $\dot{c} = 0$. This is of course not mathematically rigorous but since for this thesis the result is far more important than the lengthy proof, it can be found in [19].
Therefore, one gets an equation for the change of the substrate and product over time

$$\begin{aligned} \dot{c} = 0 \quad &\Longleftrightarrow \quad c = \frac{k_{+1}se_0}{k_{+1} + k_{-1} + k_{+2}} \\ \dot{s} &= -k_{+2}e_0 \frac{s}{s + \frac{k_{-1}+k_{+2}}{k_{+1}}} \\ \dot{p} &= k_{+2}e_0 \frac{s}{s + \frac{k_{-1}+k_{+2}}{k_{+1}}} \end{aligned} \tag{2.9}$$

The above mentioned heuristic argument further implies that the change of substrate and product over time is equal, thus $\dot{s}+\dot{p}=0$ as seen above. By re-defining the constant $\frac{k_{-1}+k_{+2}}{k_{+1}} := K$ which can be interpreted as the inverse of the affinity of the substrate and the enzyme one gets the main result

$$V := \dot{p} = -\dot{s} = k_{+2}e_0\frac{s}{s+K} \quad \sim \quad \beta\frac{s}{s+K}, \quad \beta \in \mathbb{R}_+ \tag{2.10}$$

We call V the rate at which the product is produced over time. By looking at $\dot{p} = -\dot{s}$ it is of course also the rate at which the substrate is depleted over time.

In the context of gene transcription, if one looks at DNA, genes are transcribed to mRNA if certain regulatory proteins bind to the promoter of those genes. Hence, the regulatory protein can be compared to the substrate while the promoter takes the role of the enzyme. Only if a regulatory protein binds to the promoter and forms a complex the gene is being transcribed to mRNA which functions as the product. With regards to modelling of this regulatory pathway, one can use the equation for product formation (2.10).
Let x = [X] and z = [Z] be the concentrations of the regulatory protein and the mRNA respectively, then

$$\dot{z} = \beta\frac{x}{x+K} \tag{2.11}$$

where $\beta \in \mathbb{R}_+$ is the transcription rate of the mRNA and $K \in \mathbb{R}_+$ is known as the inverse binding affinity of the regulatory protein to the promoter. However, due to the short life span of mRNA, a mRNA degradation term is missing in the above equation. This leads to

$$\dot{z} = \beta\frac{x}{x+K} - \gamma_z z \tag{2.12}$$

where $\gamma_z \in \mathbb{R}_+$ describes the degradation rate of the mRNA.

Finally, the translation of the mRNA into protein is mathematically defined by the law of mass action. Due to the disassembling of not correctly functioning or excessive protein once again a degradation term is additionally needed.
Let y = [Y] be the concentration of the protein created by translation of mRNA, then

$$\dot{y} = \alpha z - \gamma_y y \tag{2.13}$$

where $\alpha \in \mathbb{R}_+$ describes the formation rate of the protein and $\gamma_y \in \mathbb{R}_+$ visualizes the degradation rate of the protein.
The mathematical theory described in this section can be found in [18].

2.3 Time-Scale-Analysis

Once one has translated a biochemical reaction network into mathematical language, i.e. ODEs via the law of mass-action as well as via Michaelis-Menten kinetics, the question arises if any qualitative behaviour of the complex system can be shown analytically. Besides determining stationary states and the stability of the system to get information on how the system behaves especially at points of no temporal change, the main goal would be to solve the system of ODEs by hand to get the trajectories of all variables over the whole time domain. Unfortunately, these types of systems are mostly non-linear and interconnected which makes a complete analytical solution almost impossible to work out.

One idea would be to look at the system at the steady state where the variable of an ODE does not change over time anymore, i.e. $\dot{x} = 0$ for some variable x. This allows to get an analytical solution of the remaining system $f(x) = 0$ more easily where f is a function depending on variable x since one does not have to solve the ODE anymore. One crucial method that makes use of this concept is called **Time-Scale-Analysis**. Here, the system of ODEs - where each ODE models the behaviour of a biological species over time - is being split into at least two sub-systems by introducing different time scales. This approach is motivated by in vivo biological systems where connected chemical reactions have various distinctive reaction times.
By introducing a small parameter $\epsilon > 0$ and two arbitrary time-scales t, τ, one can use a linear transformation of the time component to switch between the different time-scales

Definition 2.3. Let t,τ be two different time scales, let $\epsilon > 0$
Then the transformation is given by

$$\tau = \epsilon t$$

which leads to the transformation of an arbitrary variable $(\cdot)$

$$\frac{d}{dt}(\cdot) = \frac{d}{d\tau}(\cdot)\frac{d\tau}{dt} = \epsilon\frac{d}{d\tau}(\cdot)$$

By switching between the time scales one has created two equivalent formulations of the original system. This can be written in the following way

Definition 2.4. Let there exist a system of the form

$$\dot{x} = f(x, y)$$
$$\dot{y} = g(x, y)$$

where f,g are some functions that depend on variables $x \in \mathbb{R}^n, y \in \mathbb{R}^m$ where $\dot{x} = \frac{d}{dt}x(t)$ and $\dot{y} = \frac{d}{dt}y(t)$.

Then by using the time-scale parameter $\epsilon > 0$ and the above transformation one gets

$$\begin{aligned} \frac{d}{dt}x(t) &= f(x(t), y(t)) \\ \frac{d}{dt}y(t) &= \epsilon g(x(t), y(t)) \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \epsilon\frac{d}{d\tau}x(\tau) &= f(x(\tau), y(\tau)) \\ \frac{d}{d\tau}y(\tau) &= g(x(\tau), y(\tau)) \end{aligned}$$

Since ϵ is very small, the right-hand side of $\frac{d}{dt}y(t)$ in the left sub-system is very small and $\frac{d}{dt}y(t)$ is almost 0. This indicates that compared to $y(t)$ being almost constant, $x(t)$ changes quickly. Thus the left-system is called the **fast system**. On the other hand, the left hand-side of $\frac{d}{d\tau}x(\tau)$ in the right system is very small. This indicates that compared to $x(\tau)$ being almost instantly in the steady state, $y(\tau)$ changes slowly. Thus the right-system is called the **slow system**.

The main idea is that we can now formally let the time-scale parameter go to 0 in the limit by $\epsilon \to 0$. This allows a holistic understanding of the behaviour of the original system. However, it is important to mention that in the limit the two sub-systems are not equivalent anymore since they are then describing two separate sub-processes!

$$\begin{aligned} \frac{d}{dt}x(t) &= f(x(t), y(t)) \\ \frac{d}{dt}y(t) &= 0 \end{aligned} \quad \Longleftarrow\!\!\!\!\!\!\Longrightarrow \quad \begin{aligned} 0 &= f(x(\tau), y(\tau)) \\ \frac{d}{d\tau}y(\tau) &= g(x(\tau), y(\tau)) \end{aligned}$$

If we look at the sub-systems separately we can make some statements.
Under the **steady state assumptions**, which states that a process converges to a stationary point and not a periodic orbit or a more complex behaviour, the limit case of the fast system looks as follows

$$\begin{aligned} 0 &= f(x(t), y(t)) \\ y(t) &\equiv \text{constant} \end{aligned}$$

whereas the limit case of the slow system is given by

$$\begin{aligned} 0 &= f(x(\tau), y(\tau)) \\ \frac{d}{d\tau}y(\tau) &= g(x(\tau), y(\tau)) \end{aligned}$$

In the latter case the equation $0 = f(x(\tau), y(\tau))$ represents the **slow manifold** since it depicts the stationary points of the fast process in the slow system. From the fast system, we get by the steady state assumption that the fast process converges to the slow manifold (Figure 2.1a). If one assumes that the slow manifold can be at least locally

solved for x, i.e.

$$\begin{aligned} 0 &= f(x(\tau), y(\tau)) \\ \frac{d}{d\tau}y(\tau) &= g(x(\tau), y(\tau)) \end{aligned} \iff \begin{aligned} x(\tau) &= \Phi(y(\tau)) \\ \frac{d}{d\tau}y(\tau) &= g(x(\tau), y(\tau)) \end{aligned} \iff \frac{d}{d\tau}y(\tau) = g(\Phi(y(\tau)), y(\tau))$$

we get that all points of the slow system can be found on the slow manifold (Figure 2.1b).

In total, the behaviour of the system can be represented by analysing limit cases of sub-systems with different time scales. If one takes an arbitrary point (x^*, y^*) in the phase plane, it will jump through the fast system until it reaches the slow manifold and remains on it (Figure 2.1c). A sufficient condition for this behaviour is given if the slow manifold is **transversally stable**. This is given by the Theorem of Fenichel whose proof due to its complexity and length is not stated here but can be found in [14].
The mathematical theory described in this section can be found in [18].

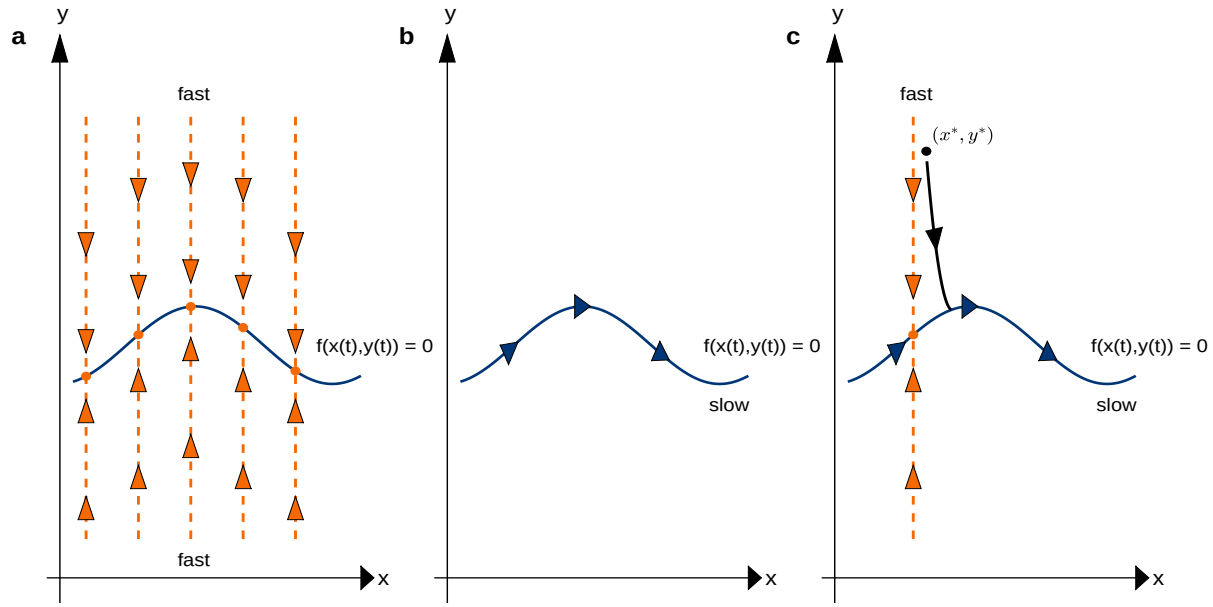

Figure 2.1: The theory of Time Scale analysis displaying the direction of points moving through the phase plane. (a) The fast system describes the flow of the system towards a point on the slow manifold. (b) The slow system describes the continuous flow of the system on the slow manifold. (c) An arbitrary point in the phase plane moves through the fast system and influenced by the slow system until it reaches the slow manifold and stays on it.

2.4 Relaxation Oscillations

However, a different situation arises if we deal with a non-hyperbolic slow manifold. This means that the slow manifold has eigenvalues whose real parts are 0. At these points, the

stability of the slow manifold can change from stable to unstable and back again. Then, we call the slow manifold neutrally stable. Therefore, a point (x^*, y^*) leaves the stable branch of the slow manifold at the non-hyperbolic point by jumping through the fast system to an once again stable branch of the same slow manifold (Figure 2.2). Hence, it is possible that the slow manifold can be concatenated by stable and unstable branches and thus the solution trajectory is made of slow and fast pieces. Under certain conditions, also the direction of the flow can change at two non-hyperbolic points, such that all initial points (x^*, y^*) of the system oscillate between two stable branches of the slow manifold. This behaviour is called **Relaxation Oscillations** which creates a periodic orbit in the system. In the context of Michaelis-Menten kinetics this phenomena is known as a **genetic switch** [19]. It models the ON and OFF switch of genes as hysteresis where the switch directly takes place at two non-hyperbolic points connecting two stable branches of the slow manifold.
The mathematical theory described in this section can be found in [18].

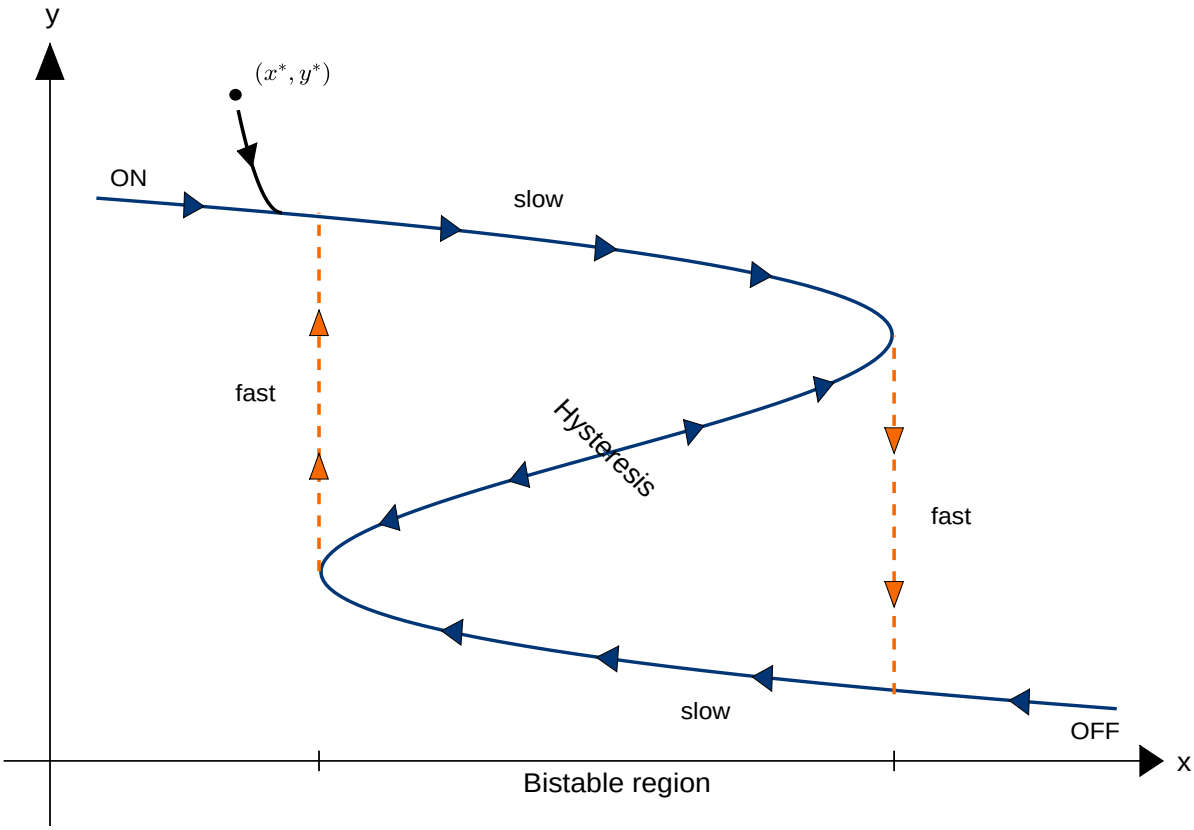

Figure 2.2: The theory of Relaxation Oscillations showing the direction of an arbitrary point moving through the phase plane. It moves on the slow manifold until it reaches non-hyperbolic points where it jumps through the fast system. In an oscillatory manner a bistable region as well as Hysteresis appear.

2.5 Parameter estimation

Still the problem exists that the time-scale analysis of a system only gives insight into the trajectories of the sub-systems at the steady state. To get the time trajectories of the full system over the whole time domain one has to solve the system numerically. This system can be written in the form of an **Initial Value Problem** (IVP).

$$\frac{dx(t,\theta)}{dt} = f(x(t,\theta),\theta,t), \quad x(t_0) = x_0(\theta) \tag{2.14}$$

where $D \subset \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \times \mathbb{R}$, $f : D \to \mathbb{R}^{n_x}$, $(x_0(\theta), \theta, t_0) \in D$, $x \in \mathbb{R}^{n_x}$.

Here, θ is called a **model parameter** and its concrete value is crucial for numerically solving the IVP (2.14).
Another problem is that most parameter values are unknown due to a lack of experimental data since measuring those parameters precisely is extremely difficult or even impossible. Thus, estimating those parameter values is mandatory which can be done in the following way:

Although the parameters themselves can not be measured directly, one can measure the state variables over a period of time which is called an **observable**

$$y(t,\theta) = h(x(t,\theta),\theta,t) \tag{2.15}$$

where $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \times \mathbb{R}_+ \to \mathbb{R}^{n_y}$.

In combination, the IVP and the observables build the **model**

$$\mathcal{M}(\theta) : \begin{cases} \frac{dx(t,\theta)}{dt} = f(x(t,\theta),\theta,t), & x(t_0) = x_0(\theta) \\ y(t,\theta) = h(x(t,\theta),\theta,t) \end{cases} \tag{2.16}$$

The experimental assessment of that observable is called a **measurement**

$$\bar{y} \sim p(\bar{y}|y,\sigma) \tag{2.17}$$

where σ is the **noise parameter** with which the measurement is corrupted. In the following, the probability function from (2.17) is represented by a p.
The complete **experimental data** then consists of tuples of time points and noise-corrupted measurements

$$D = \{(t_k, \bar{y}_k)\}_{k=1}^{n_t} \tag{2.18}$$

where $n_t \in \mathbb{N}$.
The main idea of parameter estimation is now to compare and align the observable $y(t,\theta)$ with the measurement $\bar{y}$ for all in the measurement data existing time points $t = t_k$ by

a method of choice. The easiest idea would be taking the least square method.
For that we have to introduce the conditional probability of data given the model and model parameters, the so called **likelihood**

$$p(D|\theta) := p(\bar{y}|y(t,\theta),\sigma(\theta)) \tag{2.19}$$

where $L : \Theta \to [0,1], \theta \mapsto p(D|\theta)$ is called the **Likelihood function** which depends on the parameters θ and not on D, with the parameter space $\Theta := \mathbb{R}^{n_\theta}$.
If we assume independent additive normally distributed noise

$$\bar{y}_{jk} = y_j(t_k,\theta) + \epsilon_{jk}, \qquad \epsilon_{jk} \sim \mathcal{N}(0,\sigma_{jk}^2) \tag{2.20}$$

and independent and uniformly distributed measurements $\bar{y}_{jk}$, the Likelihood function becomes

$$L(\theta) = p(D|\theta) = \prod_{k=1}^{n_t}\prod_{j=1}^{n_y} p(\bar{y}_{jk}|y_j(t_k,\theta),\sigma_{jk}(\theta)) = \prod_{k=1}^{n_t}\prod_{j=1}^{n_y} \frac{1}{\sqrt{2\pi}\sigma_{jk}(\theta)} \exp\left[-\frac{1}{2}\left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2\right] \tag{2.21}$$

In the last equation the so-called **residual** has been used which is commonly defined by

$$r(t_k,\theta) := \bar{y}_k - y(t_k,\theta) \tag{2.22}$$

Generally speaking, the Likelihood function is nothing but the probability to get the experimental data D for a certain parameter value θ. Thus, to get as close as possible to the experimental data D one has to maximize the Likelihood function. This method is called the **Maximum Likelihood Estimator** (MLE)

$$\theta^{ml} := \arg\max_{\theta\in\Omega} L(\theta) = \arg\max_{\theta\in\Omega} p(D|\theta), \qquad \text{subject to } \mathcal{M}(\theta) \tag{2.23}$$

where $\theta^{ml} \in \Omega \subseteq {\mathbb{R}_+}^{n_\theta}$ is the **maximum likelihood estimation** of θ.
From the construction of the MLE (2.23) it becomes clear that parameter estimation is an optimisation problem. In the multi-dimensional parameter space Θ, a (local) optimum is generally found by calculating the Jacobian of (2.21) with respect to θ. Since calculating the products and derivatives of exponential functions with squared exponents can be computationally expensive, the **negative log-likelihood function** is often preferred

$$J(\theta) := -\log(L(\theta)) \tag{2.24}$$

Because the log function is monotonically increasing the location of the optimum is conserved whereby due to the negative sign a maximum becomes a minimum and vice versa. Therefore, one has to take the $\arg\min_{\theta\in\Omega} J(\theta)$ instead of the $\arg\max_{\theta\in\Omega} L(\theta)$ later on.

By using the same assumptions as in (2.21) one gets

$$\begin{aligned}
J(\theta) :&= -\log p(D|\theta) \\
&= -\log\left(\prod_{k=1}^{n_t}\prod_{j=1}^{n_y}\frac{1}{\sqrt{2\pi}\sigma_{jk}(\theta)}\exp\left[-\frac{1}{2}\left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2\right]\right) \\
&= \sum_{k=1}^{n_t}\sum_{j=1}^{n_y} -\log\left(\frac{1}{\sqrt{2\pi}\sigma_{jk}(\theta)}\exp\left[-\frac{1}{2}\left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2\right]\right) \\
&= \sum_{k=1}^{n_t}\sum_{j=1}^{n_y} -\log\left(\frac{1}{\sqrt{2\pi}\sigma_{jk}(\theta)}\right) - \log\left(\exp\left[-\frac{1}{2}\left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2\right]\right) \\
&= \sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\log\left(\sqrt{2\pi}\sigma_{jk}(\theta)\right) + \left[\frac{1}{2}\left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2\right] \\
&= \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\log\left(2\pi\sigma_{jk}^2(\theta)\right) + \left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2
\end{aligned} \tag{2.25}$$

If we assume that the noise variance is known beforehand, it does not depend on θ and thus $\sigma_{jk}(\theta) = \sigma_{jk}$. Therefore, the term $\log(2\pi\sigma_{jk}^2(\theta))$ is constant and can be omitted since for the optimisation only the derivatives of (2.25) are needed.
Hence, the negative log-likelihood function is just a **weighted least squares function**

$$J(\theta) = \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}}\right)^2 \tag{2.26}$$

Once again, we have to optimize our negative log-likelihood function where now due to the negative log transformation of $L(\theta)$ the MLE is the minimiser of $J(\theta)$

$$\theta^{ml} = \arg\min_{\theta\in\Omega} J(\theta) = \arg\min_{\theta\in\Omega}\Big(-\log p(D|\theta)\Big), \qquad \text{subject to } \mathcal{M}(\theta) \tag{2.27}$$

where again $\theta^{ml} \in \Omega \subseteq {\mathbb{R}_+}^{n_\theta}$ is the **maximum likelihood estimation** of θ.
Furthermore, we get the log-transformation of (positive) parameters

$$\begin{aligned}
\theta^{ml} &= \exp(\zeta^{ml}) \\
\zeta^{ml} &= \arg\min_{\zeta\in\log(\Omega)} J(\exp(\zeta)), \qquad \text{subject to } \mathcal{M}(\exp(\zeta))
\end{aligned} \tag{2.28}$$

The mathematical theory described in this section can be found in [13].
A broader and deeper understanding of parameter estimation is stated in [8].

2.6 Optimisation

The optimisation problem (2.27) can be solved by using a **local optimisation** method. The underlying idea is to construct a sequence of parameter values $\{\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, ...\}$ starting from a guess value $\theta^{(0)}$ such that the objective function values decrease monotonically $J(\theta^{(0)}) > J(\theta^{(1)}) > J(\theta^{(2)}) > J(\theta^{(3)}) > ...$ until a certain stopping criterion is fulfilled.
The so-called **optimisation path** is constructed by using an iterative approach

$$\theta^{(k+1)} := \theta^{(k)} + \Delta^{(k)} t^{(k)} \tag{2.29}$$

where $\Delta^{(k)} \in \mathbb{R}^{n_t}$ is the search direction, $t^{(k)} \in \mathbb{R}_+$ is the step size and $t \in \mathbb{N}$ is the current iterative step of the optimisation.

The **search direction** is usually acquired by calculating the gradient of the objective function given the current parameter $\theta^{(k)}$ and a vector v along the function. Since the gradient is the multi-dimensional equivalent of the steepest curve along the function, by minimising over the vector v the "fastest way" to the minimum is found

$$\Delta^{(k)} := \arg \min_{v \in \mathbb{R}^{n_x}, ||v||=1} \nabla_\theta J(\theta^{(k)})^T v \tag{2.30}$$

This definition is possible since $(\mathbb{R}^{n_x}, ||\cdot||)$ is a normed vector space.
The **step size** is generally defined by taking a minimal step in the search direction such that the current parameter $\theta^{(k)}$ gets updated and leads to a decreased objective function value

$$t^{(k)} := \arg \min_{t \in \mathbb{R}_+} J(\theta^{(k)} + \Delta^{(k)} t) \tag{2.31}$$

As mentioned earlier, this iterative local optimisation method terminates only when a **stopping criterion** is fulfilled. Viable and often implemented stopping criteria are one or multiple of the below

$$\begin{aligned} ||J(\theta^{(k+1)}) - J(\theta^{(k)})|| &< \epsilon_J \\ ||\theta^{(k+1)} - \theta^{(k)}|| &< \epsilon_\theta \\ k &> N \end{aligned} \tag{2.32}$$

where $\epsilon_J, \epsilon_\theta \in \mathbb{R}_+$ are sufficiently small and N is the beforehand determined maximal number of steps.

Still, the problem remains how to calculate the gradient of the Jacobian $\nabla_\theta J(\theta)$ which is needed to calculate the search direction. In the literature, there exist two main distinct methods, gradient-based optimizer and non-gradient based optimizer. Since gradient-based optimizer are the common choice by being very reliable - although having high computation times - for calculating Jacobians and Hessians [5] they are discussed in the

following.
One can directly calculate the directional derivative of $J(\theta)$ with respect to θ in direction i from equation (2.25) by

$$\begin{aligned}
\frac{\partial J(\theta)}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i}\left(\frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\log\left(2\pi\sigma_{jk}^2(\theta)\right) + \left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2\right) \\
&= \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\frac{\partial}{\partial \theta_i}\left(\log\left(2\pi\sigma_{jk}^2(\theta)\right)\right) + \frac{\partial}{\partial \theta_i}\left(\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\right)^2 \\
&= \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\frac{4\pi\sigma_{jk}(\theta)}{2\pi\sigma_{jk}(\theta)^2}\frac{\partial}{\partial \theta_i}\sigma_{jk}(\theta) + 2\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)}\frac{-\sigma_{jk}(\theta)\frac{\partial}{\partial \theta_i}y_j(t_k,\theta) - \left(\bar{y}_{jk}-y_j(t_k,\theta)\right)\frac{\partial}{\partial \theta_i}\sigma_{jk}(\theta)}{\sigma_{jk}(\theta)^2} \\
&= \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\frac{2}{\sigma_{jk}(\theta)}\frac{\partial}{\partial \theta_i}\sigma_{jk}(\theta) + 2\frac{\left(2\bar{y}_{jk}y_j(t_k,\theta)-\bar{y}_{jk}^2-y_j(t_k,\theta)^2\right)\frac{\partial}{\partial \theta_i}\sigma_{jk}(\theta) - \left(\bar{y}_{jk}-y_j(t_k,\theta)\right)\sigma_{jk}(\theta)\frac{\partial}{\partial \theta_i}y_j(t_k,\theta))}{\sigma_{jk}(\theta)^3} \\
&= \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}2\frac{1}{\sigma_{jk}(\theta)}\frac{\partial}{\partial \theta_i}\sigma_{jk}(\theta) + 2\frac{\left(\bar{y}_{jk}-y_j(t_k,\theta)\right)^2\frac{\partial}{\partial \theta_i}\sigma_{jk}(\theta)}{\sigma_{jk}(\theta)^3} - 2\frac{\left(\bar{y}_{jk}-y_j(t_k,\theta)\right)\sigma_{jk}(\theta)\frac{\partial}{\partial \theta_i}y_j(t_k,\theta)}{\sigma_{jk}(\theta)^3} \\
&= \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}2\frac{1}{\sigma_{jk}(\theta)}\left(1-\frac{\left(\bar{y}_{jk}-y_j(t_k,\theta)\right)^2}{\sigma_{jk}(\theta)^2}\right)\frac{\partial}{\partial \theta_i}\sigma_{jk}(\theta) - 2\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)^2}\frac{\partial}{\partial \theta_i}y_j(t_k,\theta) \\
&= \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\frac{1}{\sigma_{jk}(\theta)^2}\left(1-\frac{\left(\bar{y}_{jk}-y_j(t_k,\theta)\right)^2}{\sigma_{jk}(\theta)^2}\right)\frac{\partial}{\partial \theta_i}\left(\sigma_{jk}(\theta)^2\right) - 2\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)^2}\frac{\partial}{\partial \theta_i}y_j(t_k,\theta)
\end{aligned} \tag{2.33}$$

We can do the same calculation in vectorial form by taking θ in all n_θ directions since $\theta \in \mathbb{R}^{n_\theta}$ and get

$$\nabla_\theta J(\theta) = \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y}\frac{1}{\sigma_{jk}(\theta)^2}\left(1-\frac{\left(\bar{y}_{jk}-y_j(t_k,\theta)\right)^2}{\sigma_{jk}(\theta)^2}\right)\nabla_\theta\left(\sigma_{jk}(\theta)^2\right) - 2\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)^2}\nabla_\theta y_j(t_k,\theta) \tag{2.34}$$

If we assume that the noise variance is known beforehand as we did in (2.26) we have $\sigma_{jk}(\theta) = \sigma_{jk} \equiv$ constant and thus $\nabla_\theta\sigma_{jk}(\theta) = \nabla_\theta\left(\sigma_{jk}(\theta)^2\right) = 0$. In total, we get

$$\nabla_\theta J(\theta) = \frac{1}{2}\sum_{k=1}^{n_t}\sum_{j=1}^{n_y} -2\frac{\bar{y}_{jk}-y_j(t_k,\theta)}{\sigma_{jk}(\theta)^2}\nabla_\theta y_j(t_k,\theta) \tag{2.35}$$

Generally, it is straight forward to calculate $\nabla_\theta\sigma_{jk}(\theta)$ but how does one evaluate $\nabla_\theta y_j(t_k,\theta)$, the gradient of the observable?

Therefore, one needs the mathematical theory of **Forward Sensitivity Analysis** which is only covered very briefly due to its complexity.
We can define the sensitivity matrix of the state variable vector

$$S^x(t) := \begin{pmatrix} (\nabla_\theta x_1(t))^T \\ \vdots \\ (\nabla_\theta x_{n_x}(t))^T \end{pmatrix} = (s_1^x(t), \dots, s_{n_x}^x(t)) \in \mathbb{R}^{n_x \times n_\theta}, \quad \text{with} \quad s_i^x(t) = \begin{pmatrix} \frac{\partial x_1(t)}{\partial \theta_i} \\ \vdots \\ \frac{\partial x_{n_x}(t)}{\partial \theta_i} \end{pmatrix} \tag{2.36}$$

as well as the sensitivity matrix of the observable vector

$$S^y(t) := \begin{pmatrix} (\nabla_\theta y_1(t))^T \\ \vdots \\ (\nabla_\theta y_{n_y}(t))^T \end{pmatrix} = (s_1^y(t), \dots, s_{n_y}^y(t)) \in \mathbb{R}^{n_y \times n_\theta}, \quad \text{with} \quad s_i^y(t) = \begin{pmatrix} \frac{\partial y_1(t)}{\partial \theta_i} \\ \vdots \\ \frac{\partial y_{n_y}(t)}{\partial \theta_i} \end{pmatrix} \tag{2.37}$$

By taking a closer look, the sensitivity matrix is nothing but $\nabla_\theta y(t,\theta)$ or $\nabla_\theta x(t,\theta)$ in its component form and thus only an equivalent representation.
The idea is now to calculate the gradient of the observable vector by using its relation to the gradient of the state variables (2.15) and the original IVP (2.14). For each component one gets

$$\begin{aligned} s_i^y(t) &= \frac{\partial y(t,\theta)}{\partial \theta_i} = \frac{\partial h(x(t,\theta),\theta,t)}{\partial \theta_i} = \frac{\partial h(x(t,\theta),\theta,t)}{\partial x}\frac{\partial x(t,\theta)}{\partial \theta_i} + \frac{\partial h(x(t,\theta),\theta,t)}{\partial \theta_i} \\ &= \frac{\partial h(x(t,\theta),\theta,t)}{\partial x} s_i^x(t) + \frac{\partial h(x(t,\theta),\theta,t)}{\partial \theta_i} \end{aligned} \tag{2.38}$$

While the function h is known and thus also its derivatives one gets the gradient of the state vector through the following ODE with initial condition

$$\begin{aligned} \frac{\partial s_i^x(t)}{\partial t} &= \frac{\partial}{\partial t}\frac{\partial x(t,\theta)}{\partial \theta_i} = \frac{\partial}{\partial \theta_i}\frac{\partial x(t,\theta)}{\partial t} = \frac{\partial f(x(t,\theta),\theta,t)}{\partial \theta_i} = \frac{\partial f(x(t,\theta),\theta,t)}{\partial x}\frac{\partial x(t,\theta)}{\partial \theta_i} + \frac{\partial f(x(t,\theta),\theta,t)}{\partial \theta_i} \\ &= \frac{\partial f(x(t,\theta),\theta,t)}{\partial x} s_i^x(t) + \frac{\partial f(x(t,\theta),\theta,t)}{\partial \theta_i}, \\ s_i^x(t_0) &= \frac{\partial x(t_0)}{\partial \theta_i} = \frac{\partial x_0(\theta)}{\partial \theta_i} \end{aligned} \tag{2.39}$$

where in the second equality the theorem of Schwarz has been used which states that for at least twice continuously differentiable functions on a subset of $\mathbb{R}^n$ their derivatives can be freely interchanged.
In total, one can calculate $\nabla_\theta y(t,\theta)$ by solving another IVP for the forward sensitivities (2.39) and calculate the expression in (2.38).

Another problem one has while using the local optimisation method is that given one parameter guess value $\theta^{(0)}$ the optimisation routine can find a local optimum. However, it is not clear if this is also the global optimum on the whole parameter search space ${\mathbb{R}_+}^{n_\theta}$.
To successfully find the global optimum, a method called **Multi-start Local Optimisation** is used. Here, the start parameter $\theta^{(0)}$ is chosen from a sufficiently large collection of random uniformly distributed values $p(\theta^{(0)})$ on the search space by

$$\theta_j^{i,(0)} := \theta_j^{lb} + (\theta_j^{ub} - \theta_j^{lb}) r_{i,j}, \qquad r_{i,j} \sim \mathcal{U}(0,1) \tag{2.40}$$

where θ^{lb} and θ^{ub} are the lower and upper boundary values of $\theta^{(0)}$. With these initial parameter values a full local optimisation is performed. By repeating this process for all remaining starting values the probability increases to find all local optima on the search space and thus also the global optimum. The global optimum $\theta^{ml}_{\text{global}}$ is the parameter with the smallest final objective function value, e.g. $J({\theta^{(N)}}_{\text{run }1}) > J({\theta^{(N)}}_{\text{run }4}) > ... > J({\theta^{(N)}}_{\text{run }2}) > J({\theta^{(N)}}_{\text{run }7}) =: J(\theta^{ml}_{\text{global}})$.

Sometimes, even constraint optimisation has to be used if one can restrict the domain based on biological insight or, more often, if the problem is too complex to be solved on the entire search domain and one has to divide it into separate subspaces.
Here, the question arises how to calculate $J(\theta)$ for $\theta \in [\theta_{lb}, \theta_{ub}]^{n_\theta} \subset {\mathbb{R}_+}^{n_\theta}$?
This bound now applies to every iteration of the optimisation path and not just for the starting value θ_0 as seen previously in the Multi-start Local Optimisation.
Therefore, one has to use ODE approximating polynomials

$$x_i(t,\theta) \approx \tilde{x}_i(t,c) := \sum_{j=1}^{m} c_j t^j \tag{2.41}$$

to fit the coefficients $c_j \in \mathbb{R}$ together with the parameters θ. By applying (2.41) to the Likelihood function (2.21), one gets the new MLE for constrained optimisation

$$\begin{aligned}(\theta^{ml}, c^{ml}) = \arg\max_{\theta,c \in \Omega \times \mathbb{R}} L(\theta) = \arg\max_{\theta,c \in \Omega \times \mathbb{R}} \prod_{k=1}^{n_t} \prod_{j=1}^{n_y} p(\bar{y}_{jk} | y_j(t_k,\theta), \sigma_{jk}(\theta)), \\ \text{subject to } \mathcal{M}(\theta, c)\end{aligned} \tag{2.42}$$

which includes the adapted model

$$\mathcal{M}(\theta,c): \begin{cases} \frac{d\tilde{x}(t,\theta)}{dt} = f(\tilde{x}(t,c),\theta,t), & \tilde{x}(t_0,c) = x_0(\theta) \\ y(t,\theta) = h(\tilde{x}(t,c),\theta,t) \end{cases} \tag{2.43}$$

The mathematical theory described in this section can be found in [13].
For a closer look at optimization in parameter estimation one can look at [8].

2.7 Numerical Integration of ODEs

To optimize one or several parameters the underlying system of ODEs has to be solved many times to align its solution step-by-step to the experimental data. Here, in the first optimisation step an initial guess $\theta = \theta_0$ has to be made to start the optimisation. After solving the system of ODEs with this θ_0 once, an optimized value θ_1 is being created. By recursively updating θ the optimal parameter values are found. Hence, in each step θ is not a variable but a known parameter.
Therefore, by considering the general theory of numerically solving ODEs, the initial IVP (2.14) can be reformulated slightly by dropping θ

$$\frac{dx(t)}{dt} = f(x(t), t), \quad x(t_0) = x_0 \tag{2.44}$$

where $D \subset \mathbb{R}^{n_x} \times \mathbb{R}$, $f : D \to \mathbb{R}^{n_x}$, $(x_0, t_0) \in D$, $x \in \mathbb{R}^{n_x}$.

Finding a solution to the system of ODEs is equivalent to numerically solving the new IVP (2.44). Thus, one must proof beforehand that such a solution exists and is even unique. By the following theorems the existence and uniqueness of such a solution is guaranteed

Definition 2.5. (Lipschitz-Continuity) Let $D \subset \mathbb{R}^{n_x} \times \mathbb{R}$.
A function $f : D \to \mathbb{R}^{n_x}$ is called Lipschitz-continuous for x in D, if there exists a $L > 0$ for which

$$\|f(x,t) - f(y,t)\|_\infty \leq L\|x - y\|_\infty \quad \forall (x,t),(y,t) \in D \tag{2.45}$$

f is called locally Lipschitz-continuous for x in D, if for all $(x_0, t_0) \in D$ there exists an open set $U \subset \mathbb{R}^{n_x} \times \mathbb{R}$ with $(x_0, t_0) \in U$ so that $f : D \cap U \to \mathbb{R}$ is Lipschitz-continuous.

Theorem 2.6. (Existence)
Let $D \subset \mathbb{R}^{n_x} \times \mathbb{R}$ open, let $f : D \to \mathbb{R}^{n_x}$ continuous and locally Lipschitz-continuous. Then for all $(x_0, t_0) \in D$ there exists an $\epsilon > 0$ and a solution $\phi : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}^n$ to the differential equation

$$\dot{x} = f(x,t), \ \phi(t_0) = x_0 \tag{2.46}$$

Theorem 2.7. (Uniqueness) Let $D \subset \mathbb{R}^{n_x} \times \mathbb{R}$, let $f : D \to \mathbb{R}^{n_x}$ continuous and locally Lipschitz-continuous, let $I \subset \mathbb{R}$.
Let $\phi, \psi : I \to \mathbb{R}^{n_x}$ be two solutions to the differential equation $\dot{x} = f(x,t)$ and there exists a $t_0 \in I$ so that $\phi(t_0) = \psi(t_0)$, then

$$\phi(t) = \psi(t) \ \forall t \in I \tag{2.47}$$

Proof: A complete proof to both theorems can be found in [7]. □

Hence, under conditions which can be fulfilled fairly easy a solution exists and it is also the only one. However, an analytical solution can not be derived by hand in most cases and numerical tools must be used. Numerically solving the IVP (2.44) is done by choosing hyperparameters such as an integration algorithm, a non-linear solver, a linear solver and error tolerances. In addition, the solver should use an adaptive time-stepping method to account for stiff dynamics foremost found in biological systems. Furthermore, it has recently been pointed out that in terms of reliability, accuracy and computation time a combination of the BDF integration algorithm, the Newton-type non-linear solver, a sparse direct linear solver and error tolerances in the range of 10^{-8} to 10^{-14} works best [26].

The **Backwards Differentiation Formula (BDF)** is a linear implicit multistep method of order $s \in \mathbb{N}$ that computes the solution x(t) of the IVP by starting at the initial point t_0 and $x(t_0) = x_0$ and approximates the k-th step $x(t_k)$ by x_k from all s previous steps $t_{k-1}, ..., t_{k-s}$. Given the IVP, its formula is given by

$$\begin{aligned} x_{k+1} &= \sum_{j=0}^{s} \alpha_j x_{k-j} + h_k \beta_{-1} f(t_{k+1}, x_{k+1}), \\ h_k &:= t_k - t_{k-1} \end{aligned} \tag{2.48}$$

where $1 \leq s \leq 5$, $\alpha_j \in \mathbb{R}$, $\beta_{-1} \neq 0$ and h_k being the step size.
It is important to mention that the BDF method has an upper limit for its order where above that boundary the method is not convergent anymore.
In most cases, the vector field f is non-linear hence equation (2.48) is non-linear and has to be solved by e.g. the **Newton-type non-linear solver** which in its most general multi-dimensional form is given by

$$x_{k+1} := x_k - \frac{F(x_k)}{\nabla_x F(x_k)} \tag{2.49}$$

where $F : \mathbb{R}^l \to \mathbb{R}^l$ has the invertible Jacobian $\nabla_x F \in \mathbb{R}^{l \times l}$ and $f(0, x) = F(x)$.
Therefore, the non-linear vector field f by applying (2.49) will be reduced in each k-th step to a linear term of the form

$$\begin{aligned} f(t_k, x_k) &= -\nabla_x f(t_k, x_k) \Delta x_k, \\ \Delta x_k &:= x_{k+1} - x_k \end{aligned} \tag{2.50}$$

The resulting linear problem can be solved by ideally applying some form of sparse **LU**

decomposition algorithm to the remaining matrix

$$\nabla_x f(t_k, x_k) = L \cdot U \tag{2.51}$$

where $\nabla_x f(t_k, x_k) \in \mathbb{R}^{n_x \times n_x}$ by construction and thus $L, U \in \mathbb{R}^{n_x \times n_x}$ being lower and upper triangle matrices connected by matrix multiplication. This linear solver is applied until certain **error tolerances** are met that control the precision to which the IVP has to be solved. The most common criteria are absolute and relative error tolerance

$$\begin{aligned} \frac{|x_{n+1} - x_n|}{|x_{n+1}|} &\leq \text{rtol} \\ |x_{n+1} - x_n| &\leq \text{atol} \end{aligned} \tag{2.52}$$

The mathematical theory described in this section can be found in [21].
In particular, the proof to use the formula of a non-linear solver (2.49) on the BDF formula (2.48) to get the linear problem (2.50) is explained in detail.

3 Mathematical approach to analyse the biological circuit of *Pseudomonas aeruginosa*

By looking at the mathematical theory of sections 2.1 and 2.2 we can translate the gene regulation circuit of *Pseudomonas aeruginosa* (Figure 1.1) into mathematical language. Afterwards, due to the theory of Time-Scale-Analysis from section 2.3 as well as Relaxation-Oscillations from section 2.4 the main goal consists of showing the existence of a genetic switch.

3.1 Mathematical model

Hereby, it is advantageous to split the full system into sub-systems which are then easier to analyse analytically due to their reduced complexity. In the following, the full system will be divided into three sub-systems, the **Basic model** which has no connection to any anti-activator gene and their expressed proteins, the **extended QslA model** and the **extended QteE model** which extend the Basic model by their respective anti-activator pathways. If one combines all sub-systems one gets the **full model**.

The Basic model for Quorum Sensing in *Pseudomonas aeruginosa* can be written as follows:

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{x} &= \beta_x + \beta_1 \frac{R_d}{R_d + K_x} - \gamma_x x \\
\dot{r} &= \beta_r - \gamma_r r \\
\dot{X} &= \alpha_X x - \gamma_X X \\
\dot{R} &= \alpha_R r - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - \gamma_{R_2} R_2 \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 - \gamma_{R_d} R_d
\end{aligned} \tag{3.1}$$

By taking a closer look, this sub-system does indeed represent a positive feedback loop. If X is increased, so is S_i, R_2 and R_d which itself increases the gene expression and pro-

duction of X. Furthermore, the transfer of the autoinducer through the cell membrane in both directions is modelled by the interconnected ODEs of S_e and S_i. By including the time dependent parameter ρ_V the system is able to incorporate the bacterial growth due to cell division which is limited by the total volume of a finite environment.

This Basic model (3.1) can be extended by including the anti-activator qteE and by accordingly increasing the ODE of the $LasR$ protein

$$\begin{aligned} \dot{q}_e &= \beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_x} - \gamma_{q_e} q_e \\ \dot{Q}_e &= \alpha_{Q_e} q_e - \gamma_{Q_e} Q_e - k^{(3)} R Q_e \\ \dot{R} &= \alpha_R r - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e \end{aligned} \tag{3.2}$$

Also here, the extension incorporates a negative feedback loop. If R is increased, so is R_2 and R_d which itself increases the gene expression and production of Q_e. However, Q_e decreases the concentration of R.

Concentration	**Variable**
AI_e	S_e
AI_i	S_i
$lasI - mRNA$	x
$lasR - mRNA$	r
$LasI$	X
$LasR$	R
$LasR - AI_i$	R_2
$LasR - AI_i - LasR - AI_i$	R_d
$qteE - mRNA$	q_e
$QteE$	Q_e
$qslA - mRNA$	q_a
$QslA$	Q_a
$QslA - QslA$	Q_2
$LasR - AI_i - QslA - QslA$	$C_{R_2Q_2}$

Table 3.1: Variables used to identify concentrations

Instead of the last extension (3.2), one can include the anti-activator qlsA and accord-

ingly adapt the ODE for R_2

$$\begin{aligned}
\dot{q}_a &= \beta_{q_a} - \gamma_{q_a} q_a \\
\dot{Q}_a &= \alpha_{Q_a} q_a - \gamma_{Q_a} Q_a - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 \\
\dot{Q}_2 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - \gamma_{Q_2} Q_2 - k^{(5)} R_2 Q_2 \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - \gamma_{R_2} R_2 - k^{(5)} R_2 Q_2 \\
\dot{C}_{R_2Q_2} &= k^{(5)} R_2 Q_2 - \gamma_{R_2Q_2} C_{R_2Q_2}
\end{aligned} \tag{3.3}$$

Here, we don't have any feedback loops since the gene expression of qslA has by experimental evaluation no connection to any other part of the network.

If one includes the Basic model with all extensions in one system of ODEs one gets the full model

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{x} &= \beta_x + \beta_1 \frac{R_d}{R_d + K_x} - \gamma_x x \\
\dot{r} &= \beta_r - \gamma_r r \\
\dot{X} &= \alpha_X x - \gamma_X X \\
\dot{R} &= \alpha_R r - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - \gamma_{R_2} R_2 - k^{(5)} R_2 Q_2 \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 - \gamma_{R_d} R_d \\
\dot{q}_e &= \beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_x} - \gamma_{q_e} q_e \\
\dot{Q}_e &= \alpha_{Q_e} q_e - \gamma_{Q_e} Q_e - k^{(3)} R Q_e \\
\dot{q}_a &= \beta_{q_a} - \gamma_{q_a} q_a \\
\dot{Q}_a &= \alpha_{Q_a} q_a - \gamma_{Q_a} Q_a - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 \\
\dot{Q}_2 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - \gamma_{Q_2} Q_2 - k^{(5)} R_2 Q_2 \\
\dot{C}_{R_2Q_2} &= k^{(5)} R_2 Q_2 - \gamma_{R_2Q_2} C_{R_2Q_2}
\end{aligned} \tag{3.4}$$

In total, the full model consists of 14 ODEs with 14 variables and 37 parameters. The explanation for all variables and parameters can be found in Table 3.1 and Table 3.2 respectively. In all upper systems regarding an arbitrary variable Y we define $\dot{Y} := \frac{d}{dt} Y$ as the time derivative of Y.

Rate constant	Parameter
Volume fraction	ρ_V
Diffusion rate in / out	c_{sec}
Extracellular degradation rate	γ_{ex}
Intracellular degradation rate	γ_{in}
$LasI$ synthesis rate	A_{syn}
$AI_i - LasR$ binding rate	$k_+^{(1)}$
$AI_i - LasR$ dissociation rate	$k_-^{(1)}$
$(AI_i - LasR)_2$ binding rate	$k_+^{(2)}$
$(AI_i - LasR)_2$ dissociation rate	$k_-^{(2)}$
$AI_i - LasR$ degradation rate	γ_{R_2}
$(AI_i - LasR)_2$ degradation rate	γ_{R_d}
$lasI$ basal transcription rate	β_x
induced $lasI$ transcription rate	β_1
$lasI$ threshold	K_x
$lasI - mRNA$ degradation rate	γ_x
$LasI$ protein formation rate	α_X
$LasI$ degradation rate	γ_X
$lasR - mRNA$ basal transcription rate	β_r
$lasR - mRNA$ degradation rate	γ_r
$LasR$ protein formation rate	α_R
$LasR$ degradation rate	γ_R
$qteE - mRNA$ basal transcription rate	β_{q_e}
induced $qteE - mRNA$ transcription rate	β_2
$qteE$ threshold	K_{q_e}
$qteE - mRNA$ degradation rate	γ_{q_e}
$QteE$ protein formation rate	α_{Q_e}
$QteE$ degradation rate	γ_{Q_e}
$LasR - QteE$ binding rate	$k^{(3)}$
$qslA - mRNA$ basal transcription rate	β_{q_a}
$qslA - mRNA$ degradation rate	γ_{q_a}
$QslA$ protein formation rate	α_{Q_a}
$QslA$ degradation rate	γ_{Q_a}
$QslA - QslA$ binding rate	$k_+^{(4)}$
$QslA - QslA$ dissociation rate	$k_-^{(4)}$
$QslA - QslA$ degradation rate	γ_{Q_2}
$LasR - AI_i - QslA - QslA$ binding rate	$k^{(5)}$
$LasR - AI_i - QslA - QslA$ degradation rate	$\gamma_{C_{R_2Q_2}}$

Table 3.2: Parameters used to describe rate constants

3.2 Model analysis

3.2.1 Basic model

To analyse the Basic model (3.1) mathematically, we first use the biological fact that the mRNA has a shorter life-span than proteins [17]. Thus, the respective ODEs $\dot{x}$ and $\dot{r}$ can be assumed to be in a quasi-steady state:

$$\begin{aligned} 0 = \dot{x} &= \beta_x + \beta_1 \frac{R_d}{R_d + K_x} - \gamma_x x \\ 0 = \dot{r} &= \beta_r - \gamma_r r \end{aligned} \tag{3.5}$$

Solving for the quasi-stationary state (x^*, r^*) leads to:

$$\begin{aligned} x^* &= \frac{1}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) \\ r^* &= \frac{\beta_r}{\gamma_r} \end{aligned} \tag{3.6}$$

Re-substituting the quasi-stationary solution (3.6) into the Basic model (3.1) gives us a reformulated version:

$$\begin{aligned} \dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\ \dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\ \dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\ \dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\ \dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - \gamma_{R_2} R_2 \\ \dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 - \gamma_{R_d} R_d \end{aligned} \tag{3.7}$$

Furthermore, we make a final but crucial assumption about the degradation parameters of the autoinducer-LasR complex protein (LasR*) and the autoinducer-LasR homodimer protein (LasRd) which is biologically motivated and can be found in literature [15]:

$$\gamma_{R_2} = \gamma_{R_d} = 0 \tag{3.8}$$

This leads to a simpler version of system (3.7)

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2
\end{aligned} \tag{3.9}$$

The main idea how to completely analyse this reformulated model (3.9) is to introduce different time-scales on which the biological processes are taking place as it was stated in section 2.3.
Therefore, we introduce a parameter $\epsilon > 0$ but small. By multiplying this parameter to the right-hand-site of certain ODEs from system (3.9) we split the system into two sub-parts, a fast system and a slow system

$$\begin{aligned}
\dot{S}_e &= \epsilon[\rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e] \\
\dot{S}_i &= \epsilon[c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2] \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2
\end{aligned} \tag{3.10}$$

This is biologically equivalent to saying that the protein formation of LasI and LasR as well as the complex formation to LasR* and LasRd reaches a steady state much faster than the other processes. Thus, system (3.10) is called the **fast system**.

By using Definition (2.3) the fast system (3.10) can be equivalently transformed into

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\epsilon \dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\epsilon \dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\epsilon \dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d \\
\epsilon \dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2
\end{aligned} \tag{3.11}$$

We call system (3.11) the **slow system**. After transformation in the upper system regarding an arbitrary variable Y we define $\dot{Y} := \frac{d}{d\tau} Y$.

By using the limit $\epsilon \to 0$ and the quasi-steady state assumption (for those ODEs which have no ϵ) for both the fast and slow system one can mathematically evaluate how the whole biological system (3.9) will behave after a certain amount of time.

First, we use this approach on the fast system (3.10) which leads to the following system

$$\begin{aligned}
\dot{S}_e &= 0 && (6) \\
\dot{S}_i &= 0 && (5) \\
0 &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X && (4) \\
0 &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 && (3) \\
0 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d && (2) \\
0 &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 && (1)
\end{aligned} \tag{3.12a}$$

Now, we can solve each of the six equations to get a solution for every variable

$$(1) \iff R_d = \frac{k_+^{(2)}}{k_-^{(2)}} R_2^2 \tag{3.12b}$$

$$(3) \iff R = \frac{\alpha_R \beta_r + \gamma_r k_-^{(1)} R_2}{\gamma_r (\gamma_R + k_+^{(1)} S_i)} \tag{3.12c}$$

Substituting 3.12b and 3.12c into (2) gives us

$$(2) \Longleftrightarrow \boxed{R_2 = \frac{k_+^{(1)} \alpha_R \beta_r}{k_-^{(1)} \gamma_R \gamma_r} S_i} \tag{3.12d}$$

Re-substituting 3.12d into 3.12b and 3.12c leads to

$$3.12b \Longleftrightarrow \boxed{R_d = \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2}{k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2} S_i^2} \tag{3.12e}$$

$$3.12c \Longleftrightarrow \boxed{R = \frac{\alpha_R \beta_r}{\gamma_r \gamma_R}} \tag{3.12f}$$

Substituting 3.12e into (4) resolves in

$$(4) \Longleftrightarrow \boxed{X = \frac{\alpha_X}{\gamma_x \gamma_X} \left(\beta_x + \beta_1 \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2}{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2 + k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x} \right)} \tag{3.12g}$$

The other ODEs can be solved easily

$$(5) \Longleftrightarrow S_i \equiv const \tag{3.12h}$$

$$(6) \Longleftrightarrow S_e \equiv const \tag{3.12i}$$

In total, we get the stationary solutions for the fast system (3.12a)

$$\begin{aligned}
S_e &\equiv const \\
S_i &\equiv const \\
R &= \frac{\alpha_R \beta_r}{\gamma_r \gamma_R} \\
R_2 &= \frac{k_+^{(1)} \alpha_R \beta_r}{k_-^{(1)} \gamma_R \gamma_r} S_i \\
R_d &= \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2}{k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2} S_i^2 \\
X &= \frac{\alpha_X}{\gamma_x \gamma_X} \left(\beta_x + \beta_1 \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2}{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2 + k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x} \right)
\end{aligned} \tag{3.13}$$

To now get the non-constant solutions for the autoinducers S_i and S_e, we can use the limit case $\epsilon \to 0$ and the steady state for the slow system (3.11) which leads to the

following system

$$\begin{aligned}
0 &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e && (6') \\
0 &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 && (5') \\
0 &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X && (4') \\
0 &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 && (3') \\
0 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d && (2') \\
0 &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 && (1')
\end{aligned} \tag{3.14a}$$

If we now compare the last four equations (1'),(2'),(3'),(4') of this system (3.14a) with those (1),(2),(3),(4) of system (3.12a) we see that the equations are identical, so their solutions are identical too. Thus, we can look at the first two equations (5'), (6') and solve for their variables:

$$(6') \Longleftrightarrow \boxed{S_e = \frac{\rho_V c_{sec}}{\rho_V c_{sec} + \gamma_{ex}} S_i} \tag{3.14b}$$

By substituting 3.14b and 3.12d, 3.12f, 3.12g into (5') we get

$$0 = \left(\frac{\rho_V c_{sec}^2}{\rho_V c_{sec} + \gamma_{ex}} - c_{sec} - \gamma_{in}\right) S_i + \frac{\alpha_X A_{syn}}{\gamma_x \gamma_X}\left(\beta_x + \beta_1 \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2}{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2 + k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x}\right) \tag{3.14c}$$

We can bring this equation into a different form which will show the existence of a genetic switch

$$\frac{\gamma_x \gamma_X}{\alpha_X A_{syn} \beta_1}\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in}\right) S_i - \frac{\beta_x}{\beta_1} = \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2}{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2 + k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x} \tag{3.14d}$$

To proof the existence of a genetic switch we make the following definition where one can introduce new parameters just for shorter notation

Definition 3.1. Let equation (3.14d) hold where all parameters are positive, we define

$$\begin{aligned} m &:= \frac{\gamma_x \gamma_X}{\alpha_X A_{syn} \beta_1} \left(\frac{c_{sec} \gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in} \right) > 0 \\ u &:= \frac{\beta_x}{\beta_1} > 0 \\ c &:= k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 > 0 \\ d &:= k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x > 0 \end{aligned} \tag{3.15}$$

With this definition we can state and proof the following lemma

Lemma 3.2. Let $\exists\ m, u, c, d \in \mathbb{R}^+$, let $I = [0, \frac{u+1}{m}]$ be an interval, let $\exists$ a variable $Z \in I$, with

$$mZ - u = \frac{cZ^2}{cZ^2 + d} \tag{3.16}$$

Then: the equation has between one and three real stationary solutions in I. In particular, a necessary condition for having three real stationary solutions is given by

$$0 \geq c^2 u(u+1)^3 + m^2 cd(2u^2 - 5u - 0.25) + m^4 d^2, \ \ \forall m, c, d > 0, u \in \left(0, \frac{5 + 3\sqrt{3}}{4}\right)$$

Proof: Let us define $f(Z) := mZ - u$ and $g(Z) := \frac{cZ^2}{cZ^2+d}$. Then on I both functions are continuous and we have $f(0) < g(0)$ and $f(\frac{u+1}{m}) > g(\frac{u+1}{m})$. By defining $h(Z) := f(Z) - g(Z)$ we get $h(0) < 0$ and $h(\frac{u+1}{m}) > 0$. Thus, by applying the **Intermediate Value Theorem** $\exists\ Z^* \in I$: $h(Z^*) = 0 \quad \Longrightarrow \quad f(Z^*) = g(Z^*)$ Therefore, we can find at least one real stationary state.

One can also equivalently transform equation (3.16) into

$$mcZ^3 - c(u+1)Z^2 + mdZ - ud = 0$$

By using **Cardanos' formula** for solving cubic equations one could get all stationary states if all parameter values were known. If not, one can still find the number of real solutions by looking at the sign of its discriminant

$$\triangle := \left(\frac{3m^2cd - c^2(u+1)^2}{9m^2c^2} \right)^3 + \left(\frac{-2c^3(u+1)^3 + 9m^2c^2(u+1)d - 27m^2c^2ud}{54m^3c^3} \right)^2$$

By using Cardanos' Theorem we get

$$\begin{cases} \exists \text{ one real stationary solution} & , \triangle > 0 \\ \exists \text{ three real stationary solutions, two are the same} & , \triangle = 0 \\ \exists \text{ three different real stationary solutions} & , \triangle < 0 \end{cases}$$

Therefore, for certain parameter values, we can even find three real independent stationary solutions. An algebraic derivation of Cardano's theorem can be found in [2] . Although we don't know the parameter values we can propose necessary conditions on those to fulfill the different cases for the discriminate.

$\triangle = 0$: since the first term can be negative, the second term has to balance out the first one

$$-\left(\frac{3m^2cd - c^2(u+1)^2}{9m^2c^2}\right)^3 = \left(\frac{-2c^3(u+1)^3 + 9m^2c^2(u+1)d - 27m^2c^2ud}{54m^3c^3}\right)^2$$

$$\Longleftrightarrow$$

$$\left(c(u+1)^2 - 3m^2d\right)^3 = \frac{c}{4}\left(2c(u+1)^3 + 9m^2d(2u-1)\right)^2$$

$$\Longleftrightarrow$$

$$0 = c^2u(u+1)^3 + m^2cd(2u^2 - 5u - 0.25) + m^4d^2$$

Since the first and third term are always positive, it depends on the second term! Thus, we can formulate the necessary conditions

$$0 < u < \frac{5 + 3\sqrt{3}}{4}$$
$$0 = c^2u(u+1)^3 + m^2cd(2u^2 - 5u - 0.25) + m^4d^2, \ \ \forall m, c, d > 0$$

$\triangle > 0$: using the same calculations as before

$$0 < c^2u(u+1)^3 + m^2cd(2u^2 - 5u - 0.25) + m^4d^2$$

Here, we may use all parameter values such that the inequality is fulfilled!

$$0 < c^2u(u+1)^3 + m^2cd(2u^2 - 5u - 0.25) + m^4d^2, \ \ \forall m, u, c, d, > 0$$

$\triangle < 0$: using the same calculations as before

$$0 > c^2u(u+1)^3 + m^2cd(2u^2 - 5u - 0.25) + m^4d^2$$

Again, it depends on the negativity of the second term! Hence, the necessary conditions are again

$$\begin{aligned} &0 < u < \frac{5+3\sqrt{3}}{4} \\ &0 > c^2u(u+1)^3 + m^2cd(2u^2-5u-0.25) + m^4d^2, \ \ \forall m,c,d>0 \end{aligned}$$

In total, a necessary condition for having three real stationary solutions is directly given by the boundary of t and the combined condition of all parameter values

$$\boxed{\begin{aligned} &0 < u < \frac{5+3\sqrt{3}}{4} \\ &0 \geq c^2u(u+1)^3 + m^2cd(2u^2-5u-0.25) + m^4d^2, \ \ \forall m,c,d>0 \end{aligned}} \tag{3.17}$$

□

By using this result about the solutions of a cubic equation, we can look for our four different bifurcation parameters m, u, c, d at their resulting stationary states (Figure 3.1). By drawing the bifurcation parameter against the stationary states we get the slow manifold. To get the full slow manifold, one would have to visualize it for all bifurcation parameters as a hypersurface in five dimensions. The critical region in which the bifurcation parameter values lead to three stationary solutions ($\triangle \leq 0$) is called the **bistable region**. Below a certain threshold of autoinducer S_i the system remains in the OFF state. After the threshold is crossed the system jumps through the fast field and activates the system in the ON state. If the signal decreases over time the system will jump once more to the OFF state resulting in hysteresis in the bistable region (3.1). Thus, this process visualizes a so-called **genetic switch** [19].
It is important to notice that this process can only happen if the ON and OFF states are stable and the branch in between is unstable, resulting in two non-hyperbolic points where the activation and deactivation of the system happens. The respective calculations can be found in section 3.3.

3.2.2 Extended QslA model

Once again, we first use the biological fact that the mRNA has a shorter life-span than proteins [17]. Thus, the respective ODEs $\dot{x}$, $\dot{r}$ and now $\dot{q}_a$ can be assumed to be in a quasi-steady state:

$$\begin{aligned} 0 = \dot{x} &= \beta_x + \beta_1 \frac{R_d}{R_d + K_x} - \gamma_x x \\ 0 = \dot{r} &= \beta_r - \gamma_r r \\ 0 = \dot{q}_a &= \beta_{q_a} - \gamma_{q_a} q_a \end{aligned} \tag{3.18}$$

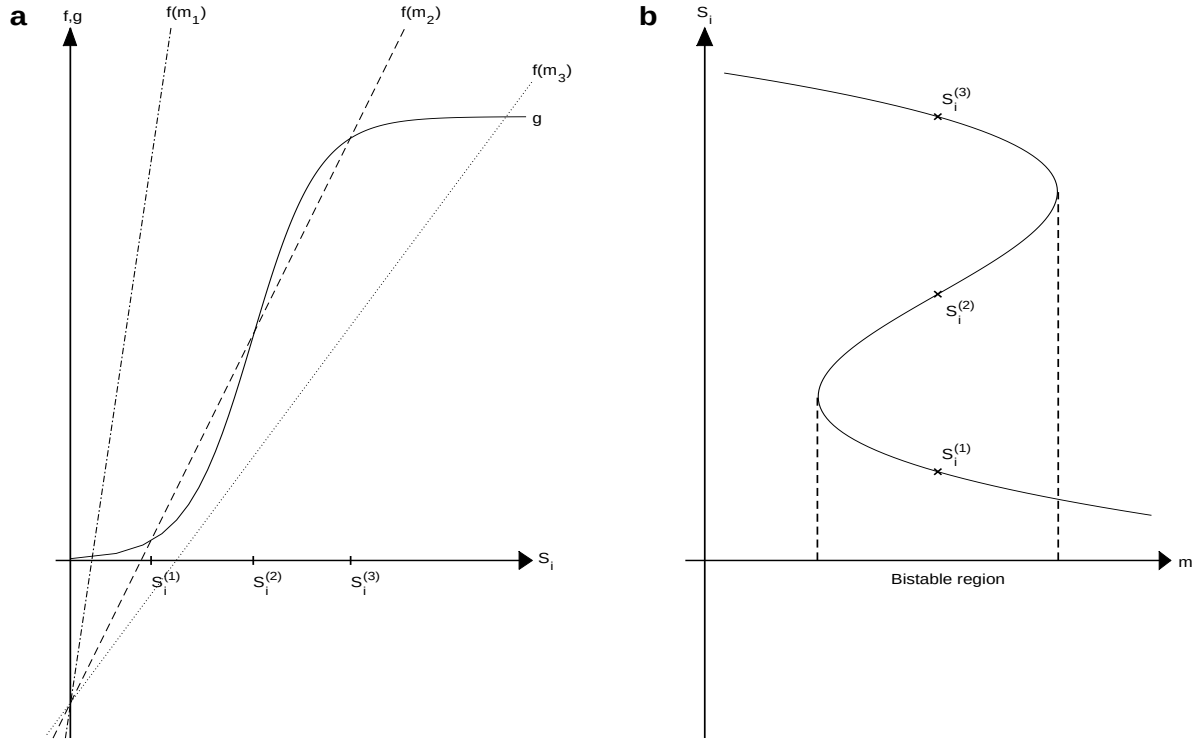

Figure 3.1: Bifurcation analysis showing the genetic switch of the Basic model. (a) For the bifurcation parameter m, functions f and g from equation (3.16) have different intersection points. If m is fixed up to three real solutions can be found. (b) Varying parameter m leads to a bifurcation diagram for all possible real solutions S_i laying on the slow manifold. It visualizes a genetic switch where the upper branch represents the ON state and the lower branch depicts the OFF state. Between certain thresholds a bistable region occurs.

Solving for the quasi-stationary state (x^*, r^*, q_a^*) leads to:

$$\begin{aligned} x^* &= \frac{1}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) \\ r^* &= \frac{\beta_r}{\gamma_r} \\ q_a^* &= \frac{\beta_{q_a}}{\gamma_{q_a}} \end{aligned} \tag{3.19}$$

Re-substituting the quasi-stationary solution (3.19) into the extended QslA containing model (3.3) gives us a reformulated version:

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - \gamma_{R_2} R_2 - k^{(5)} R_2 Q_2 \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 - \gamma_{R_d} R_d \\
\dot{Q}_a &= \frac{\alpha_{Q_a} \beta_{q_a}}{\gamma_{q_a}} - \gamma_{Q_a} Q_a - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 \\
\dot{Q}_2 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - \gamma_{Q_2} Q_2 - k^{(5)} R_2 Q_2 \\
\dot{C}_{R_2Q_2} &= k^{(5)} R_2 Q_2 - \gamma_{R_2Q_2} C_{R_2Q_2}
\end{aligned} \tag{3.20}$$

Furthermore, we make again crucial assumptions about the degradation parameters of the $LasR^*$ complex, $LasRd$ homodimer, $QslA$ protein and $QslA - QslA$ homodimer which is biologically motivated and can be found in literature [15]:

$$\gamma_{R_2} = \gamma_{R_d} = \gamma_{Q_a} = \gamma_{Q_2} = 0 \tag{3.21}$$

This leads to a simpler version of system (3.20)

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - k^{(5)} R_2 Q_2 \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 \\
\dot{Q}_a &= \frac{\alpha_{Q_a} \beta_{q_a}}{\gamma_{q_a}} - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 \\
\dot{Q}_2 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - k^{(5)} R_2 Q_2 \\
\dot{C}_{R_2Q_2} &= k^{(5)} R_2 Q_2 - \gamma_{R_2Q_2} C_{R_2Q_2}
\end{aligned} \tag{3.22}$$

The main idea how to completely analyse this reformulated model (3.22) is again using Time-Scale analysis from section 2.3.
Therefore, we introduce a parameter $\epsilon > 0$ but small. By multiplying this parameter to the right-hand-site of certain ODEs from system (3.22) we split the system into two sub-parts, a fast system and a slow system

$$\begin{aligned}
\dot{S}_e &= \epsilon[\rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e] \\
\dot{S}_i &= \epsilon[c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2] \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - k^{(5)} R_2 Q_2 \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 \\
\dot{Q}_a &= \frac{\alpha_{Q_a} \beta_{q_a}}{\gamma_{q_a}} - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 \\
\dot{Q}_2 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - k^{(5)} R_2 Q_2 \\
\dot{C}_{R_2Q_2} &= k^{(5)} R_2 Q_2 - \gamma_{R_2Q_2} C_{R_2Q_2}
\end{aligned} \tag{3.23}$$

This is biologically equivalent to saying that the protein formation of LasI, LasR and QslA as well as the complex formation to LasR*, LasRd, QslAd and LasR*-QslAd reaches a steady state much faster than the other processes. Thus, system (3.23) is called the **fast system**.

By using Definition (2.3) the fast system (3.23) can be equivalently transformed into

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\epsilon\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\epsilon\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\epsilon\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - k^{(5)} R_2 Q_2 \\
\epsilon\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 \\
\epsilon\dot{Q}_a &= \frac{\alpha_{Q_a} \beta_{q_a}}{\gamma_{q_a}} - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 \\
\epsilon\dot{Q}_2 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - k^{(5)} R_2 Q_2 \\
\epsilon\dot{C}_{R_2Q_2} &= k^{(5)} R_2 Q_2 - \gamma_{R_2Q_2} C_{R_2Q_2}
\end{aligned} \tag{3.24}$$

We call system (3.24) the **slow system**. After transformation in the upper system regarding an arbitrary variable Y we define $\dot{Y} := \frac{d}{d\tau}Y$.

By again using the limit $\epsilon \to 0$ and the quasi-steady state assumption on the fast system (3.23) it leads to

$$
\begin{aligned}
\dot{S}_e &= 0 && (9)\\
\dot{S}_i &= 0 && (8)\\
0 &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X && (7)\\
0 &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 && (6)\\
0 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - k^{(5)} R_2 Q_2 && (5)\\
0 &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 && (4)\\
0 &= \frac{\alpha_{Q_a} \beta_{q_a}}{\gamma_{q_a}} - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 && (3)\\
0 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - k^{(5)} R_2 Q_2 && (2)\\
0 &= k^{(5)} R_2 Q_2 - \gamma_{R_2 Q_2} C_{R_2 Q_2} && (1)
\end{aligned}
\tag{3.25a}
$$

Now, solving each of the nine equations to get a solution for every variable leads to

$$(1) \iff C_{R_2Q_2} = \frac{k^{(5)}}{\gamma_{R_2Q_2}} R_2 Q_2 \tag{3.25b}$$

$$(3) \iff Q_a = \sqrt{\frac{k_-^{(4)}}{k_+^{(4)}} Q_2 + \frac{\alpha_{Q_a}\beta_{q_a}}{2k_+^{(4)}\gamma_{q_a}}} \tag{3.25c}$$

Substituting 3.25c into (2) results in

$$(2) \iff Q_2 = \frac{\alpha_{Q_a}\beta_{q_a}}{2k^{(5)}\gamma_{q_a} R_2} \tag{3.25d}$$

Furthermore, we get

$$(4) \iff R_d = \frac{k_+^{(2)}}{k_-^{(2)}} R_2^2 \tag{3.25e}$$

$$(6) \iff R = \frac{\alpha_R \beta_r + \gamma_r k_-^{(1)} R_2}{\gamma_r(\gamma_R + k_+^{(1)} S_i)} \tag{3.25f}$$

Substituting 3.25e, 3.25f and 3.25d into (5) gives us

$$(5) \Longleftrightarrow \boxed{R_2 = \left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}} \tag{3.25g}$$

Re-substituting 3.25g into 3.25e, 3.25f and 3.25d leads to

$$3.25e \Longleftrightarrow \boxed{R_d = \frac{k_+^{(2)}}{k_-^{(2)}} \left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2} \tag{3.25h}$$

$$3.25f \Longleftrightarrow \boxed{R = \frac{\alpha_R \beta_r}{\gamma_r \gamma_R} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} \gamma_R}} \tag{3.25i}$$

$$3.25d \Longleftrightarrow \boxed{Q_2 = \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k^{(5)}} \frac{1}{\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}}} \tag{3.25j}$$

Substituting 3.25h into (7) resolves in

$$(7) \Longleftrightarrow \boxed{X = \frac{\alpha_X}{\gamma_x \gamma_X} \left(\beta_x + \beta_1 \frac{\left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2}{\left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2 + \frac{k_-^{(2)}}{k_+^{(2)}} K_x} \right)} \tag{3.25k}$$

Substituting 3.25j into 3.25c and 3.25j and 3.25g into 3.25b resolves in

$$3.25c \Longleftrightarrow \boxed{Q_a = \sqrt{\frac{k_-^{(4)}}{k_+^{(4)}} \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k^{(5)}} \frac{1}{\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}} + \frac{\alpha_{Q_a} \beta_{q_a}}{2k_+^{(4)} \gamma_{q_a}}}} \tag{3.25l}$$

$$3.25b \Longleftrightarrow \boxed{C_{R_2 Q_2} = \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} \gamma_{R_2 Q_2}}} \tag{3.25m}$$

The other ODEs can be solved easily

$$(8) \Longleftrightarrow S_i \equiv const \tag{3.25n}$$

$$(9) \Longleftrightarrow S_e \equiv const \tag{3.25o}$$

In total, we get the solution for the fast system (3.25a)

$$
\begin{aligned}
S_e &\equiv const \\
S_i &\equiv const \\
R &= \frac{\alpha_R \beta_r}{\gamma_r \gamma_R} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} \gamma_R} \\
R_2 &= \left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \\
R_d &= \frac{k_+^{(2)}}{k_-^{(2)}} \left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2 \\
X &= \frac{\alpha_X}{\gamma_x \gamma_X} \left(\beta_x + \beta_1 \frac{\left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2}{\left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2 + \frac{k_-^{(2)}}{k_+^{(2)}} K_x} \right) \\
Q_a &= \sqrt{\frac{k_-^{(4)}}{k_+^{(4)}} \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k^{(5)}} \frac{1}{\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}} + \frac{\alpha_{Q_a} \beta_{q_a}}{2k_+^{(4)} \gamma_{q_a}}} \\
Q_2 &= \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k^{(5)}} \frac{1}{\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}} \\
C_{R_2 Q_2} &= \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} \gamma_{R_2 Q_2}}
\end{aligned}
\tag{3.26}
$$

Again, we get the non-constant solutions for the autoinducers S_i and S_e by using the limit case $\epsilon \to 0$ and the quasi-steady state for the slow system

$$
\begin{aligned}
0 &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e && (9') \\
0 &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 && (8') \\
0 &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X && (7') \\
0 &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 && (6') \\
0 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - k^{(5)} R_2 Q_2 && (5') \\
0 &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 && (4') \\
0 &= \frac{\alpha_{Q_a} \beta_{q_a}}{\gamma_{q_a}} - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 && (3') \\
0 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - k^{(5)} R_2 Q_2 && (2') \\
0 &= k^{(5)} R_2 Q_2 - \gamma_{R_2 Q_2} C_{R_2 Q_2} && (1')
\end{aligned}
\tag{3.27a}
$$

If we now compare the last four equations (1'),(2'),(3'),(4'),(5'),(6'),(7') of this system (3.27a) with those (1),(2),(3),(4),(5),(6),(7) of system (3.25a) we see that the equations are identical, so their solutions are identical too. Thus, we can look at the first two equations (8'), (9') and solve for their variables:

$$
(9') \Longleftrightarrow \boxed{S_e = \frac{\rho_V c_{sec}}{\rho_V c_{sec} + \gamma_{ex}} S_i} \tag{3.27b}
$$

By substituting 3.27b and 3.25k, 3.25g, 3.25i into (8') we get

$$
\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in}\right) S_i + \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}} = \frac{\alpha_X A_{syn}}{\gamma_x \gamma_X}\left(\beta_x + \beta_1 \frac{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r} - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R} S_i - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}\right]^2}{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r} - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R} S_i - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}\right]^2 + \frac{k_-^{(2)}}{k_+^{(2)}} K_x}\right)
\tag{3.27c}
$$

For the equivalent form, under some conditions, we can show the existence of a genetic switch

$$\frac{\gamma_x\gamma_X}{\alpha_X A_{syn}\beta_1}\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right)S_i+\frac{\gamma_x\gamma_X}{\alpha_X A_{syn}\beta_1}\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}-\frac{\beta_x}{\beta_1}=\frac{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r}-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R}S_i-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k_-^{(1)}}\right]^2}{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r}-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R}S_i-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k_-^{(1)}}\right]^2+\frac{k_-^{(2)}}{k_+^{(2)}}K_x} \tag{3.27d}$$

Again, one can introduce new parameters just for shorter notation

Definition 3.3. Let equation (3.27d) hold, we define

$$\begin{aligned}
m &:= \frac{\gamma_x\gamma_X}{\alpha_X A_{syn}\beta_1}\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right)\\
u &:= \frac{\gamma_x\gamma_X}{\alpha_X A_{syn}\beta_1}\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}-\frac{\beta_x}{\beta_1}\\
c &:= \left(\frac{\alpha_R\beta_r}{\gamma_r}-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R}\\
d &:= \frac{k_-^{(2)}}{k_+^{(2)}}K_x\\
e &:= \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k_-^{(1)}}
\end{aligned} \tag{3.28}$$

Clearly, the newly defined parameters m, d and e are always positive, but what about the two remaining parameters?
If we assume $c \leq 0$ then this would imply $\frac{\alpha_R\beta_r}{\gamma_r} \leq \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}$ since every other parameter is positive.
Re-substituting this behaviour into the the solution of the fast (and thus also slow) system (3.26) yields $R_2 < 0$, $R < 0$, $Q_2 < 0$ and even $Q_a \in \mathbb{C}$ which is not biologically meaningful. Therefore, we have an important condition which has to be fulfilled

$$\frac{\alpha_R\beta_r}{\gamma_r} > \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}} \tag{3.29}$$

and so c is also always positive.
Unfortunately, for parameter u we can not tell its sign without any parameter values but we can derive a necessary condition for u such that a genetic switch exists.

With this intention in mind we can state and proof the following lemma

Lemma 3.4. Let $\exists\ m, c, d, e \in \mathbb{R}^+$, let $\exists\ u \in \mathbb{R}$, let $I = [0, \infty)$ be an interval, let $\exists$ a variable $Z \in I$, with

$$mZ + sgn(u)|u| = \frac{(cZ - e)^2}{(cZ - e)^2 + d} \tag{3.30}$$

where the signum function is defined by $sgn(t) := \begin{cases} -1, & \text{if } u < 0 \\ 0, & \text{if } u = 0 \\ +1, & \text{if } u > 0 \end{cases}$

Then: the equation has between zero and three real stationary solutions in I. In particular, a necessary condition for having three real stationary solutions and a genetic switch is given by

$$u < 0$$

Thus, for simplicity, we can re-define $u := -\tilde{u}$, $\tilde{u} > 0$ from 3.3 by pulling out a negative sign to formally drop the absolute value of t, resulting in

$$mZ - \tilde{u} = \frac{(cZ - e)^2}{(cZ - e)^2 + d}, \qquad \tilde{u} > 0$$

Proof: We can start by showing the amount of reals stationary solutions. Therefore, let us define $f(Z) := mZ + sgn(u)|u|$ and $g(Z) := \frac{(cZ-e)^2}{(cZ-e)^2+d}$. We look at three distinct cases.
In the first case, wlog. let $sgn(u)|u| > \frac{e^2}{e^2+d}$. Since $f(0) = sgn(u)|u|$ and $g(0) = \frac{e^2}{e^2+d}$, $\exists m > 0$, large enough : $mZ + sgn(u)|u| \neq \frac{(cZ-e)^2}{(cZ-e)^2+d}$ in I $\forall c > 0$. Hence, under certain conditions, there can be no real stationary solutions. However, if $m > 0$ but small enough, there can exist up to two stationary states (Figure 3.2).
In the second case, let $sgn(u)|u| = \frac{e^2}{e^2+d}$. Then $f(0) = g(0)$ and $Z = 0$ is at least one stationary solution.
In the third case, let $sgn(u)|u| < \frac{e^2}{e^2+d}$. Then on I both functions are continuous and we have $f(0) < g(0)$. Furthermore, g is bounded from above by 1 since $g(Z) < 1$ $\forall Z \in I$ and $\lim_{Z\to\infty} g(Z) = 1$. By taking the arbitrary point $\tilde{Z} = \frac{2-sgn(u)|u|}{m}$ we get $f(\tilde{Z}) = 2 > 1 > g(\tilde{Z})$. By construction we have indeed $\tilde{Z} \in I$ since $\frac{2-sgn(u)|u|}{m} > \frac{2-\frac{e^2}{e^2+d}}{m} > \frac{2-1}{m} > 0$ $\forall m > 0$. Thus, by defining $h(Z) := f(Z) - g(Z)$ we get $h(0) < 0$ and $h(\tilde{Z}) > 0$. Thus, by applying the **Intermediate Value Theorem** $\exists\ Z^* \in I$: $h(Z^*) = 0 \implies f(Z^*) = g(Z^*)$
Therefore, we can find at least one real stationary state (Figure 3.2).

As seen in the different cases above, three real stationary states can only exist if $sgn(u)|u| \leq \frac{e^2}{e^2+d}$ and only for a precise choice of parameters $m, c, d, e > 0$. However, the more interesting question is for which parameter values equation (3.27d) shows bistable behaviour and a proper genetic switch. By performing a bifurcation analysis with the

bifurcation parameter m, one gets the first necessary condition from Figure 3.2e that

$$\boxed{u < 0} \tag{3.31}$$

Otherwise, visually speaking, an arbitrary point on the slow manifold would eventually hit the disconnected second branch where the genetic switch would remain in the OFF state and could not be switched on again (Figure 3.2c,d). In two other cases, no genetic switch appears because once only two real stationary solutions can be found (Figure 3.2b) and once no there are no stationary solutions. This can only be seen in (Figure 3.2a) since no bifurcation diagram exists for that case. □

Similar to the Basic model, one can also equivalently transform equation (3.30) into

$$mc^2Z^3 - [2mce + c^2(u+1)]Z^2 + [2(u+1)ce + m(e^2+d)]Z + (u+1)e^2 - ud = 0$$

to find the number of real solutions. By looking at the sign of the discriminant, one gets

$$\begin{aligned}\triangle := &\left(\frac{2mc^3e(u+1) - m^2c^2e^2 + 3m^2c^2d - c^4(u+1)^2}{9m^2c^24}\right)^3 + \\ &+\left(\frac{-2c^6(u+1)^3 + 9m^2c^4d - 18m^2c^4ud + 18m^3c^3ed + 2m^3c^3e^3 + 48m^2c^4e^2)(u+1) + 6mc^5e(u+1)^2}{54m^3c^6}\right)^2\end{aligned}$$

For the bistable region of a genetic switch one must have three solutions. For the non-hyperbolic points where the system jumps through the fast field one needs two solutions. Therefore, one is interested in $\triangle \geq 0$ and after some algebraic calculations one eventually gets

$$\begin{aligned}0 \geq &150mc^3e^3(\tilde{u}+1)^3 + 576m^2c^2e^4(\tilde{u}+1)^2 + 54m^3ce^5(\tilde{u}+1) + \\ &27m^4d^3 + 27m^2c^2d^2(2\tilde{u}^2 - 5\tilde{u} - 0.25) + 27c^4d(\tilde{u}+1)^3\tilde{u} - \\ &54c^4e^2(\tilde{u}+1)^4 + 18m^2c^2de^2(5\tilde{u}^2 - 14t + 17) + \\ &27m^3cde^3(14\tilde{u}+15) + 12mc^3e^3(\tilde{u}+1)^4 - \\ &27m^3cd^2e(4\tilde{u}-5) + 54m^4d^2e^2 - 27mc^3de(\tilde{u}+1)^2(4\tilde{u}+1) + \\ &18m^2c^2de^2(\tilde{u}+1)^2 + 72m^3de^4\end{aligned} \tag{3.32}$$

Here, at least two and up to four terms can be negative to balance out all the other positive terms. Thus, the result underlines the existence of a genetic switch.
However, this genetic switch only exists if the ON and OFF states are stable and the branch in between is unstable, resulting again in two non-hyperbolic points where the switch of the system happens. The respective calculations can be found in section 3.3.

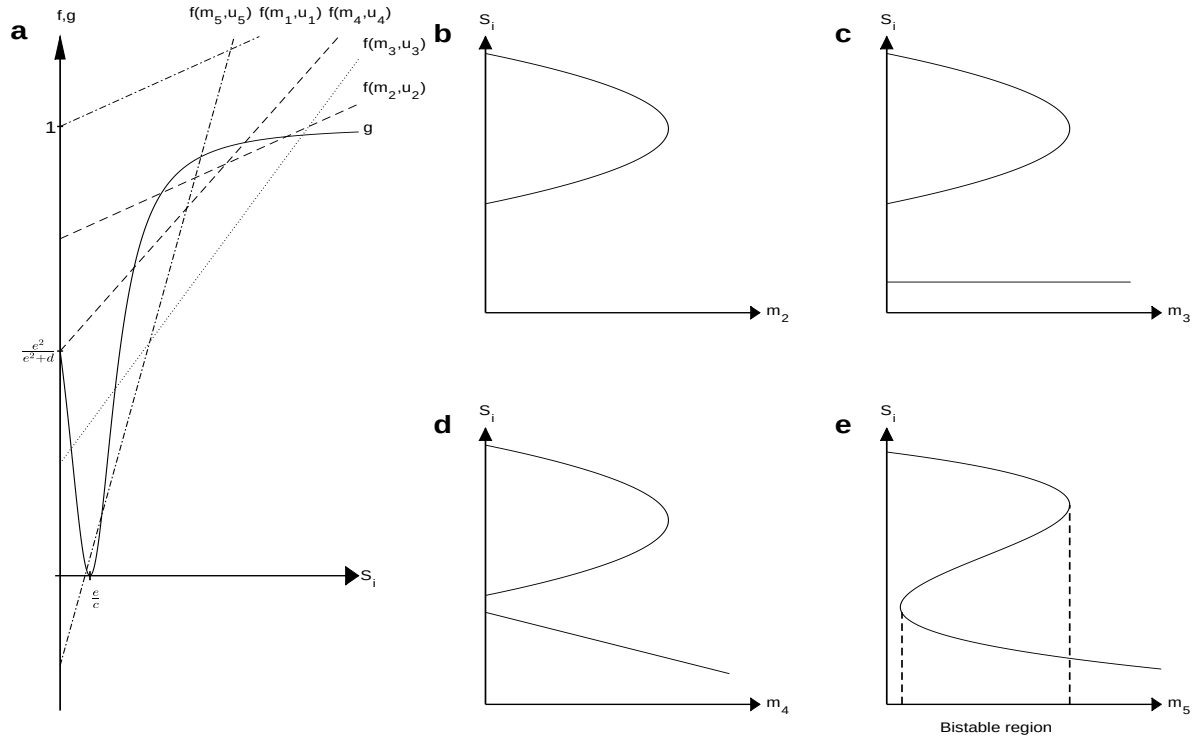

Figure 3.2: Bifurcation analysis showing the genetic switch of the extended QslA model. (a) By varying the bifurcation parameter m and by fixing u, functions f and g from equation (3.30) have up to three real intersection points for five different cases of t. (b) Parameter m_2 shows the bifurcation diagram for $\frac{e^2}{e^2+d} < u < 1$. (c) Parameter m_3 shows the bifurcation diagram for $u = \frac{e^2}{e^2+d}$. (d) Parameter m_4 shows the bifurcation diagram for $0 \leq u < \frac{e^2}{e^2+d}$. (e) Parameter m_5 shows the bifurcation diagram for $u < 0$. Only here a genetic switch appears where the upper branch represents the ON state and the lower branch depicts the OFF state. Between certain thresholds a bistable region occurs.

3.2.3 Extended QteE model

For the third time, we first use the biological fact that the mRNA has a shorter life-span than proteins [17]. Thus, the respective ODEs $\dot{x}$, $\dot{r}$ and this time $\dot{q}_e$ can be assumed to be in a quasi-steady state:

$$\begin{aligned} 0 = \dot{x} &= \beta_x + \beta_1 \frac{R_d}{R_d + K_x} - \gamma_x x \\ 0 = \dot{r} &= \beta_r - \gamma_r r \\ 0 = \dot{q}_e &= \beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}} - \gamma_{q_e} q_e \end{aligned} \tag{3.33}$$

Solving for the quasi-stationary state (x^*, r^*, q_e^*) leads to:

$$\begin{aligned} x^* &= \frac{1}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) \\ r^* &= \frac{\beta_r}{\gamma_r} \\ q_e^* &= \frac{1}{\gamma_{q_e}}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) \end{aligned} \tag{3.34}$$

Re-substituting the quasi-stationary solution (3.34) into the extended QteE containing model (3.2) gives us a reformulated version:

$$\begin{aligned} \dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\ \dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\ \dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\ \dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e \\ \dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - \gamma_{R_2} R_2 \\ \dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 - \gamma_{R_d} R_d \\ \dot{Q}_e &= \frac{\alpha_{Q_e}}{\gamma_{q_e}}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) - \gamma_{Q_e} Q_e - k^{(3)} R Q_e \end{aligned} \tag{3.35}$$

Furthermore, we make again crucial assumptions about the degradation parameters of the $LasR^*$ complex, $LasRd$ homodimer and $QteE$ protein which is biologically motivated and can be found in literature [15]:

$$\gamma_{R_2} = \gamma_{R_d} = \gamma_{Q_e} = 0 \tag{3.36}$$

This leads to a simpler version of system (3.35)

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 \\
\dot{Q}_e &= \frac{\alpha_{Q_e}}{\gamma_{q_e}}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) - k^{(3)} R Q_e
\end{aligned} \tag{3.37}$$

Once again, the main idea how to completely analyse this reformulated model (3.37) is using Time-Scale analysis from section 2.3.
Therefore, we introduce a parameter $\epsilon > 0$ but small. By multiplying this parameter to the right-hand-site of certain ODEs from system (3.37) we split the system into two sub-parts, a fast system and a slow system

$$\begin{aligned}
\dot{S}_e &= \epsilon[\rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e] \\
\dot{S}_i &= \epsilon[c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2] \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 \\
\dot{Q}_e &= \frac{\alpha_{Q_e}}{\gamma_{q_e}}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) - k^{(3)} R Q_e
\end{aligned} \tag{3.38}$$

This is biologically equivalent to saying that the protein formation of LasI, LasR and QteE as well as the complex formation to LasR* and LasRd reaches a steady state much faster than the other processes. Thus, system (3.38) is called the **fast system**.

By using Definition (2.3) the fast system (3.38) can be equivalently transformed into

$$\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\epsilon \dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\epsilon \dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e \\
\epsilon \dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d \\
\epsilon \dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 \\
\epsilon \dot{Q}_e &= \frac{\alpha_{Q_e}}{\gamma_{q_e}}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) - k^{(3)} R Q_e
\end{aligned} \tag{3.39}$$

We call system (3.39) the **slow system**. After transformation in the upper system regarding an arbitrary variable Y we define $\dot{Y} := \frac{d}{d\tau} Y$.

By again using the limit $\epsilon \to 0$ and the quasi-steady state assumption on the fast system (3.38) it leads to

$$\begin{aligned}
\dot{S}_e &= 0 && (7) \\
\dot{S}_i &= 0 && (6) \\
0 &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X && (5) \\
0 &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e && (4) \\
0 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d && (3) \\
0 &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 && (2) \\
0 &= \frac{\alpha_{Q_e}}{\gamma_{q_e}}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) - k^{(3)} R Q_e && (1)
\end{aligned} \tag{3.40a}$$

Now, solving each of the seven equations to get a solution for every variable leads to

$$(1) \iff Q_e = \frac{\alpha_{Q_e}}{k^{(3)} \gamma_{q_e} R}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) \tag{3.40b}$$

$$(2) \iff R_d = \frac{k_+^{(2)}}{k_-^{(2)}} R_2^2 \tag{3.40c}$$

Substituting 3.40b and 3.40*c* into (4) gives us

$$(4) \Longleftrightarrow R = \frac{\frac{\alpha_R \beta_r}{\gamma_r} + k_-^{(1)} R_2 - \frac{\alpha_{Q_e}}{\gamma_{q_e}} \left(\beta_{q_e} + \beta_2 \frac{R_2^2}{R_2^2 + \frac{k_-^{(2)}}{k_+^{(2)}} K_{q_e}} \right)}{\gamma_R + k_+^{(1)} S_i} \tag{3.40d}$$

Substituting 3.40*c* and 3.40*d* into (3) gives us

$$(3) \Longleftrightarrow 0 = -\frac{k_-^{(1)} k_+^{(2)}}{k_+^{(1)} \beta_2} R_2^3 - k_+^{(2)} \left[\frac{S_i}{\beta_2} \left(\frac{\alpha_{Q_e} \beta_{q_e}}{\gamma_{q_e}} - \frac{\alpha_R \beta_r}{\gamma_r} \right) + 1 \right] R_2^2 - \frac{k_-^{(1)} k_-^{(2)} K_{q_e}}{k_+^{(1)} \beta_2} R_2 - \frac{k_-^{(2)} K_{q_e}}{\beta_2} \left(\frac{\alpha_{Q_e} \beta_{q_e}}{\gamma_{q_e}} - \frac{\alpha_R \beta_r}{\gamma_r} \right) S_i \tag{3.40e}$$

This is a problem since we don't have a unique real solution for which R_2 is positive but maybe even three different real solutions. To determine the amount of real solutions we can equivalently transform the last equation into

$$-\frac{k_-^{(1)} \gamma_R}{k_+^{(1)} \beta_2} R_2 + \frac{S_i}{\beta_2} \left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_e} \beta_{q_e}}{\gamma_{q_e}} \right) = \frac{k_+^{(2)} R_2^2}{k_+^{(2)} R_2^2 + k_-^{(2)} K_{q_e}} \tag{3.40f}$$

Once again, one can introduce new parameters just for shorter notation

Definition 3.5. Let equation (3.40*f*) hold, we define

$$\begin{aligned} m &:= \frac{k_-^{(1)} \gamma_R}{k_+^{(1)} \beta_2} \\ u &:= \frac{S_i}{\beta_2} \left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_e} \beta_{q_e}}{\gamma_{q_e}} \right) \\ c &:= k_+^{(2)} \\ d &:= k_-^{(2)} K_{q_e} \end{aligned} \tag{3.41}$$

Here, the defined parameters m, c and d are always positive, but what about u?
If we assume $u < 0$ then this would imply that the function $f(R_2) := -mR_2 + u$ defined as the left-hand-side of (3.40*f*) has its only stationary state for $R_2^* = \frac{u}{m} < 0$, since $u < 0$. Thus, the only existing real solution of (3.40*f*) would be negative which is not biologically meaningful. Therefore, u must be non-negative which gives us an important condition that has to be fulfilled

$$\frac{\alpha_R \beta_r}{\gamma_r} \geq \frac{\alpha_{Q_e} \beta_{q_e}}{\gamma_{q_e}} \tag{3.42}$$

This leads to the following important observation

Lemma 3.6. Let $\exists\ m, c, d \in \mathbb{R}^+$, let $\exists\ u \in \mathbb{R} \geq 0$, let $I = [0, \frac{u+1}{m}]$ be an interval, let $\exists$ a variable $Z \in I$, with

$$-mZ + u = \frac{cZ^2}{cZ^2 + d} \tag{3.43}$$

Then: the equation has exactly one real stationary solution in I.

Proof: Let us define $f(Z) := -mZ + u$ and $g(Z) := \frac{cZ^2}{cZ^2+d}$. Then on I both functions are continuous and we have $f(0) < g(0)$ as well as $f(\frac{u+1}{m}) > g(\frac{u+1}{m})$. By defining $h(Z) := f(Z) - g(Z)$ we get $h(0) > 0$ and $h(\frac{u+1}{m}) < 0$. Thus, by applying the **Intermediate Value Theorem** $\exists\ Z^* \in I\colon\ h(Z^*) = 0 \quad \Longrightarrow \quad f(Z^*) = g(Z^*)$.
This shows the existence of at least one stationary state. However, it is also the only one $\forall m, c, d, u > 0$. A simple argument would be to look at the course of $f(Z)$ and $g(Z)$ on I where $f(Z)$ is strictly monotonically decreasing while $g(Z)$ is strictly monotonically increasing. Let us assume there would be two distinctive intersection points a and b between $f(Z)$ and $g(Z)$ where wlog $a < b$. Then we have $f(a) > f(b) = g(b) > g(a)$ which implies $f(a) > g(a)$. So a is not an intersection point and hence there can not be more than one intersection point. Since we know one exists it is the only one. □

In the last lemma we showed the existence of exactly one positive solution for the LasR* protein R_2. Thus, we know that system (3.40a) has at least one biologically meaningful solution. However, the problem is that depending on the exact parameter values up to two more real solutions for R_2 can exist. Due to the structure of equation (3.43) all other solutions must be negative leaving in all cases exactly one real positive solution for R_2. Still, only by knowing the sign of Δ the correct solution can be stated by using Cardano's Theorem.
Calculating the discriminant for equation (3.43) leads to

$$\begin{aligned}
\Delta :&= \left(\frac{3m^2cd - c^2(u-1)^2}{9m^2c^2}\right)^3 + \left(\frac{-2c^3(u-1)^3 + 9m^2c^2(u-1)d - 27m^2c^2ud}{54m^3c^3}\right)^2 \\
&= \frac{1}{729m^6c^3}\left(27m^6d^3 - \frac{27}{4}m^4cd^2\Big((u-1)^2 + 18u(u-1) - 27u^2\Big) + \frac{9}{2}m^2c^2d(u-1)^3(u-2)\right) \\
&= \frac{d}{81m^4c^3}\left(3m^4d^2 - 6m^2cd\Big(u - \frac{-5+\sqrt{537}}{8}\Big)\Big(u - \frac{-5-\sqrt{537}}{8}\Big) + \frac{1}{2}c^2(u-1)^3(u+2)\right)
\end{aligned} \tag{3.44}$$

It becomes clear that only by knowing $m, c, d > 0$ and $u \geq 0$ without their actual values determining the sign is not possible. Hence, also stating the exact positive real solution for R_2 is impossible. Thus, a proper analytical analysis and a proof of the existence of a genetic switch for this subsystem is not possible.

Nonetheless, an interesting case with an important biological interpretation arises if we consider $u = 0$, which by Definition 3.5 is equivalent to $\frac{\alpha_R \beta_r}{\gamma_r} = \frac{\alpha_{Q_e} \beta_{q_e}}{\gamma_{q_e}}$. This leads for equation 3.42 in the context of the model (3.40f) with $Z := R_2$ in I to the following situation

$$\begin{aligned} mR_2 = \frac{cR_2^2}{cR_2^2 + d} \qquad &\Longleftrightarrow \qquad R_2 = 0 \quad \forall m, c, d > 0 \\ &\overset{3.40c}{\Longleftrightarrow} \qquad R_d = 0 \\ &\overset{3.40a}{\Longleftrightarrow} \qquad X = \frac{\alpha_X \beta_x}{\gamma_x \gamma_X} \\ &\overset{3.40d}{\Longleftrightarrow} \qquad R = 0 \\ &\overset{3.40b}{\Longleftrightarrow} \qquad Q_e = \infty \end{aligned} \tag{3.45}$$

We can see that the mathematical model breaks down since we get an infinite amount of Q_e. This is not surprising since in (3.36) we did set $\gamma_{Q_e} := 0$ so besides inactivating the LasR protein R it just grows unlimited. However, if one took the more complicated model before this simplification (3.35) but assumes that γ_{Q_e} is small, the values for R, R_2 and R_d would be very small but non-zero while the value for Q_e would be very large but finite. We can interpret this result in the following way. If the transcription, translation and mRNA-degradation of LasR protein and QteE protein are identical, the Quorum Sensing system is inactivated completely. Only a small amount of autoinducer can be produced.

3.2.4 Full model

As we know from the full model (3.4) both anti-activators QslA and QteE are independent in the sense that they only influence each other over the amount of LasR or LasR* protein still present in the system. Therefore, trying to analyse the full system with the mathematical theory used for all submodels will eventually run into the same problem one has for the extended QteE subsystem. Hence, also for the full model it is not possible to analytically prove the existence of a genetic switch.

In total, the estimation and usage of actual parameter values is crucial to proceed the determination of the qualitative behaviour for this sub-system as well as the full system. This can only be done numerically and will be shown in chapter 4.

3.3 Stability analysis

To determine the stability of the genetic switches 3.14d and 3.27d and thus verify the existence of relaxation oscillations as seen in section 2.3 one has to calculate the stability

of the turning point as well as the change of stability at the non-hyperbolic points, i.e. those points where the system looses track of the slow manifold. For a genetic switch to have relaxation oscillations the turning point must be unstable and the non-hyperbolic points must indicate a switch from stable to unstable or vice versa. Moreover, the whole branch between the non-hyperbolic points must be unstable whereas the other branches must be stable. This can be verified by determining the flow direction on the slow manifold. Only then an arbitrary initial point starting somewhere on the slow manifold will jump through the fast system from one stable branch to the other and avoid the unstable branch around the turning point. As seen in section 2.4, the trajectory will converge to a periodic orbit.

3.3.1 Basic model

The main goal is to determine the slow manifold of the Basic model (3.9). Herefore, one has to look at the limit case $\epsilon \to 0$ for the fast system (3.10) since all points in the fast system will then converge to a set of new points, the so-called slow manifold of the Basic system.

Since the analysis of a system containing all six ODEs will become difficult very quickly, one can make use of a nice simplification. The idea is to reduce the system to two ODEs containing only the variables (S_i, X). Thus, the different time scales in the fast system are preserved. Let $\dot{R}$, $\dot{R}_2$, $\dot{R}_d$ and $\dot{S}_e$ be in quasi-steady state, i.e. $\dot{R} = \dot{R}_2 = \dot{R}_d = \dot{S}_e = 0$. As previously seen, by solving these equations, we get

$$3.12f \iff R = \frac{\alpha_R \beta_r}{\gamma_r \gamma_R} \tag{3.46}$$

$$3.12d \iff R_2 = \frac{k_+^{(1)} \alpha_R \beta_r}{k_-^{(1)} \gamma_R \gamma_r} S_i \tag{3.47}$$

$$3.12c \iff R_d = \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2}{k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2} S_i^2 \tag{3.48}$$

$$3.14b \iff S_e = \frac{\rho_V c_{sec}}{\rho_V c_{sec} + \gamma_{ex}} S_i \tag{3.49}$$

Re-substituting all latter equations into (3.10) leads to the reduced 2D-case of the fast system

$$\begin{aligned} \dot{X} &= \frac{\alpha_X}{\gamma_x \gamma_X} \left(\beta_x + \beta_1 \frac{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2}{k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2 + k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x} \right) - \gamma_X X \\ \dot{S}_i &= \epsilon \left[- \left(\frac{c_{sec} \gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in} \right) S_i + A_{syn} X \right] \end{aligned} \tag{3.50}$$

Now, we can take the limit case $\epsilon \to 0$ to make the transition to the slow manifold.

$$\begin{aligned} \dot{X} &= \frac{\alpha_X}{\gamma_x \gamma_X}\left(\beta_x + \beta_1 \frac{k_+^{(2)}(k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2}{k_+^{(2)}(k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2 + k_-^{(2)}(k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x}\right) - \gamma_X X \\ \dot{S}_i &= 0 \end{aligned} \tag{3.51}$$

Afterwards, we need to see what happens at the steady state where $t \to \infty$. Since $\dot{S}_i = 0$ implies that $S_i \equiv \text{const}$ we only take $\dot{X} = 0$. Then, one gets

$$\begin{aligned} 0 &= \frac{\alpha_X}{\gamma_x \gamma_X}\left(\beta_x + \beta_1 \frac{k_+^{(2)}(k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2}{k_+^{(2)}(k_+^{(1)})^2 \alpha_R^2 \beta_r^2 S_i^2 + k_-^{(2)}(k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x}\right) - \gamma_X X \\ S_i &\equiv \text{const} \end{aligned} \tag{3.52}$$

Finally, we arrive at the slow manifold given by a combined two-dimensional expression of the upper system. For simplicity, we introduce the same parameter re-definitions as seen in Definition 3.1.

$$\boxed{\{(X, S_i) | \gamma_x \gamma_X c S_i^2 X - \alpha_X c(\beta_X + \beta_1) S_i^2 + \gamma_x \gamma_X d X - \alpha_X \beta_x d = 0\}} \tag{3.53}$$

Now, one has to determine the non-hyperbolic points and the turning point as well as their stability according to the slow manifold, newly defined as
$\Phi(X, S_i) := \gamma_x \gamma_X c S_i^2 X - \alpha_X c(\beta_X + \beta_1) S_i^2 + \gamma_x \gamma_X d X - \alpha_X \beta_x d$

By starting with the non-hyperbolic points, we get

$$\nabla \Phi(X, S_i) = \begin{pmatrix} \gamma_x \gamma_X c S_i^2 + \gamma_x \gamma_X d \\ 2\Big(\gamma_x \gamma_X c X - \alpha_X c(\beta_X + \beta_1)\Big) S_i \end{pmatrix} \tag{3.54}$$

and

$$\nabla \Phi(X, S_i) = 0 \qquad \Longleftrightarrow \qquad \begin{pmatrix} X_{\text{nh}}^* \\ S_{i,\text{nh}}^* \end{pmatrix} = \begin{pmatrix} \frac{\alpha_X}{\gamma_x \gamma_X}(\beta_X + \beta_1) \\ \pm i \sqrt{\frac{d}{c}} \end{pmatrix} \tag{3.55}$$

By continuing with the turning point, one gets

$$\Delta \Phi(X, S_i) = 2\Big(\gamma_x \gamma_X c X - \alpha_X c(\beta_X + \beta_1)\Big) \tag{3.56}$$

and

$$\Delta \Phi(X, S_i) = 0 \qquad \Longleftrightarrow \qquad X_{\text{tp}}^* = \frac{\alpha_X}{\gamma_x \gamma_X}(\beta_X + \beta_1) \;\; \forall S_{i,\text{tp}}^* \tag{3.57}$$

Both calculations are possible since all variables are continuously differentiable functions in time thus also the slow manifold is continuously differentiable.

To determine the stability of the non-hyperbolic points as well as the turning point one has to calculate the Jacobian of the system (3.50), re-substitute the respective point (X^*, S_i^*) and look at the eigenvalues or the trace-determinant-criterion.

$$J(X, S_i) = \begin{pmatrix} -\gamma_X & \frac{\alpha_X \beta_1}{\gamma_x} \frac{2cdS_i}{(cS_i^2+d)^2} \\ \epsilon A_{syn} & -\epsilon\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}} + \gamma_{in}\right) \end{pmatrix} \tag{3.58}$$

The Jacobian (3.58) has two interesting properties:

1. $J(X, S_i) = J(S_i) \quad \forall (X, S_i)$
2. $\mathrm{tr}(J(X, S_i)) < 0 \quad \forall (X, S_i)$

Thus, by the trace-determinant-criterion based on the sign of the determinant one can only encounter unstable saddle points or stable nodes such as sinks, stars or spirals [16].

In the case of our non-hyperbolic points, they can be seen as the transition points from stable to unstable branches without actually having a clear stability. However, one can make the argument that the fast system looses track at the non-hyperbolic points giving them an unstable behaviour. Therefore, one gets the following condition for the non-hyperbolic points to be unstable

$$\begin{aligned} \det(J(X, S_i)|_{(X^*_{\mathrm{nh}}, S^*_{i,\mathrm{nh}})}) < 0 &\iff \epsilon \left[\gamma_X \left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in}\right) - A_{syn}\frac{\alpha_X \beta_1}{\gamma_x}\frac{2cd\sqrt{\frac{d}{c}}}{4d^2}\right] < 0 \\ &\iff \frac{\gamma_x \gamma_X}{A_{syn}\alpha_X \beta_1}\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in}\right) < \frac{1}{2}\sqrt{\frac{c}{d}} \\ &\iff \boxed{\sqrt{\frac{d}{c}} < \frac{1}{2m}} \end{aligned} \tag{3.59}$$

where in the last transformation again for simplicity the parameter were re-defined according to (3.1). not true!!!, due to complex solution one divides by 0 in (1,2) matrix entry. maybe just take real part?

In the case of the turning point, it has to be unstable. Hence, due to the Jacobian being independent of X as well as the turning point being independent of S_i one gets a

condition on $S^*_{i,\mathrm{tp}}$

$$\begin{aligned}\det(J(X,S_i)|_{(X^*_{\mathrm{tp}},S^*_{i,\mathrm{tp}})})<0 &\iff \epsilon\left[\gamma_X\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right)-A_{syn}\frac{\alpha_X\beta_1}{\gamma_x}\frac{2cdS^*_{i,\mathrm{tp}}}{(c{S^*_{i,\mathrm{tp}}}^2+d)^2}\right]<0\\ &\iff \frac{\gamma_x\gamma_X}{A_{syn}\alpha_X\beta_1}\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right)<\frac{2cdS^*_{i,\mathrm{tp}}}{(c{S^*_{i,\mathrm{tp}}}^2+d)^2}\\ &\iff \boxed{mc^2{S^*_{i,\mathrm{tp}}}^4+2cdm{S^*_{i,\mathrm{tp}}}^2-2cd{S^*_{i,\mathrm{tp}}}^2+md^2<0}\end{aligned} \tag{3.60}$$

Finally, the flow direction on the slow manifold can be determined by looking at the sign of $\dot{X}$ in the limit case of the fast system (3.50) and $\dot{S}_i$ in the limit case of the slow system (3.61) by substituting points left, right or in between of the non-hyperbolic points in their respective right-hand sides. Herefore, the required slow system in the limit case $\epsilon\to 0$ is briefly stated

$$\begin{aligned}0&=\frac{\alpha_X}{\gamma_x\gamma_X}\left(\beta_x+\beta_1\frac{k^{(2)}_+(k^{(1)}_+)^2\alpha_R^2\beta_r^2S_i^2}{k^{(2)}_+(k^{(1)}_+)^2\alpha_R^2\beta_r^2S_i^2+k^{(2)}_-(k^{(1)}_-)^2\gamma_R^2\gamma_r^2K_x}\right)-\gamma_X X\\ \dot{S}_i&=-\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right)S_i+A_{syn}X\end{aligned} \tag{3.61}$$

However, due to the time constraints of this thesis and the necessary lengthy calculations the last part of the analysis is omitted at this point.

3.3.2 Extended QslA model

One can use the same arguments given in the last section to determine the slow manifold of the extended QslA model (3.22). Again, one has to look at the limit case $\epsilon\to 0$ for the fast system (3.23) since all points in the fast system will then converge to a set of new points, the so-called slow manifold of the Basic system.

Since the analysis of a system containing all nine ODEs will become difficult very quickly, one can make use of the same simplification as before. The idea is to reduce the system to two ODEs containing only the variables (S_i, X). Still, the different time scales in the fast system are preserved. Let $\dot{R}$, $\dot{R}_2$, $\dot{R}_d$, $\dot{Q}_a$, $\dot{Q}_2$, $\dot{C}_{R_2Q_2}$ and $\dot{S}_e$ be in quasi-steady state, i.e. $\dot{R}=\dot{R}_2=\dot{R}_d=\dot{Q}_a=\dot{Q}_2=\dot{C}_{R_2Q_2}=\dot{S}_e=0$. As previously seen, by solving these

equations, we get

$$3.25i \Longleftrightarrow R = \frac{\alpha_R \beta_r}{\gamma_r \gamma_R} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} \gamma_R} \tag{3.62}$$

$$3.25g \Longleftrightarrow R_2 = \left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \tag{3.63}$$

$$3.25h \Longleftrightarrow R_d = \frac{k_+^{(2)}}{k_-^{(2)}} \left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2 \tag{3.64}$$

$$3.25l \Longleftrightarrow Q_a = \sqrt{\frac{k_-^{(4)}}{k_+^{(4)}} \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k^{(5)}} \frac{1}{\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}} + \frac{\alpha_{Q_a} \beta_{q_a}}{2k_+^{(4)} \gamma_{q_a}}} \tag{3.65}$$

$$3.25j \Longleftrightarrow Q_2 = \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k^{(5)}} \frac{1}{\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}}} \tag{3.66}$$

$$3.25m \Longleftrightarrow C_{R_2 Q_2} = \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} \gamma_{R_2 Q_2}} \tag{3.67}$$

$$3.27b \Longleftrightarrow S_e = \frac{\rho_V c_{sec}}{\rho_V c_{sec} + \gamma_{ex}} S_i \tag{3.68}$$

Re-substituting all latter equations into (3.23) leads to the reduced 2D-case of the fast system

$$\begin{aligned} \dot{X} &= \frac{\alpha_X A_{syn}}{\gamma_x \gamma_X} \left(\beta_x + \beta_1 \frac{\left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2}{\left[\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} S_i - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} \right]^2 + \frac{k_-^{(2)}}{k_+^{(2)}} K_x} \right) - \gamma_X X \\ \dot{S}_i &= \epsilon \left[- \left(\frac{c_{sec} \gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in} \right) S_i + A_{syn} X - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right] \end{aligned} \tag{3.69}$$

Now, we can take the limit case $\epsilon \to 0$ to make the transition to the slow manifold.

$$\begin{aligned}\dot{X} &= \frac{\alpha_X A_{syn}}{\gamma_x \gamma_X}\left(\beta_x + \beta_1 \frac{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r} - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R}S_i - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k_-^{(1)}}\right]^2}{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r} - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R}S_i - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k_-^{(1)}}\right]^2 + \frac{k_-^{(2)}}{k_+^{(2)}}K_x}\right) - \gamma_X X \\ \dot{S}_i &= 0\end{aligned} \tag{3.70}$$

Afterwards, we need to see what happens at the steady state after $t \to \infty$. Since $\dot{S}_i = 0$ implies that $S_i \equiv \text{const}$ we only take $\dot{X} = 0$. Then, one gets

$$\begin{aligned}0 &= \frac{\alpha_X A_{syn}}{\gamma_x \gamma_X}\left(\beta_x + \beta_1 \frac{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r} - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R}S_i - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k_-^{(1)}}\right]^2}{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r} - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k_+^{(1)}}{k_-^{(1)}\gamma_R}S_i - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k_-^{(1)}}\right]^2 + \frac{k_-^{(2)}}{k_+^{(2)}}K_x}\right) - \gamma_X X \\ S_i &\equiv \text{const}\end{aligned} \tag{3.71}$$

Finally, we arrive at the slow manifold given by a combined two-dimensional expression of the upper system. For simplicity, we introduce the same parameter re-definitions as seen in Definition 3.3.

$$\boxed{\{(X, S_i) | \gamma_x\gamma_X(cS_i - e)^2 X - \alpha_X(\beta_X + \beta_1)(cS_i - e)^2 + \gamma_x\gamma_X dX - \alpha_X\beta_x d = 0\}} \tag{3.72}$$

Now, one has to determine the non-hyperbolic points and the turning point as well as their stability according to the slow manifold, newly defined as
$\Phi(X, S_i) := \gamma_x\gamma_X(cS_i - e)^2 X - \alpha_X(\beta_X + \beta_1)(cS_i - e)^2 + \gamma_x\gamma_X dX - \alpha_X\beta_x d$

By starting with the non-hyperbolic points, we get

$$\nabla\Phi(X, S_i) = \begin{pmatrix} \gamma_x\gamma_X(cS_i - e)^2 + \gamma_x\gamma_X d \\ 2c\Big(\gamma_x\gamma_X cX - \alpha_X c(\beta_X + \beta_1)\Big)(cS_i - e)\end{pmatrix} \tag{3.73}$$

and

$$\nabla\Phi(X, S_i) = 0 \qquad \Longleftrightarrow \qquad \begin{pmatrix} X^*_{\text{nh}} \\ S^*_{i,\text{nh}}\end{pmatrix} = \begin{pmatrix} \frac{\alpha_X}{\gamma_x\gamma_X}(\beta_X + \beta_1) \\ \frac{\pm i\sqrt{d} + e}{d}\end{pmatrix} \tag{3.74}$$

By continuing with the turning point, one gets

$$\Delta\Phi(X, S_i) = 2c^2\Big(\gamma_x\gamma_X cX - \alpha_X c(\beta_X + \beta_1)\Big) \tag{3.75}$$

and

$$\Delta\Phi(X, S_i) = 0 \qquad \Longleftrightarrow \qquad X^*_{\text{tp}} = \frac{\alpha_X}{\gamma_x\gamma_X}(\beta_X + \beta_1) \quad \forall S^*_{i,\text{tp}} \tag{3.76}$$

Both calculations are possible since all variables are continuously differentiable functions in time thus also the slow manifold is continuously differentiable.

To determine the stability of the non-hyperbolic points as well as the turning point one has to calculate the Jacobian of the system (3.69), re-substitute the respective point (X^*, S_i^*) and look at the eigenvalues or the trace-determinant-criterion.

$$J(X, S_i) = \begin{pmatrix} -\gamma_X & \frac{\alpha_X\beta_1}{\gamma_x}\frac{2cd(cS_i - e)}{\left((cS_i-e)^2+d\right)^2} \\ \epsilon A_{syn} & -\epsilon\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}} + \gamma_{in}\right) \end{pmatrix} \tag{3.77}$$

Similar to (3.58), the Jacobian (3.77) has two interesting properties:

1. $J(X, S_i) = J(S_i) \quad \forall (X, S_i)$
2. $\text{tr}(J(X, S_i)) < 0 \quad \forall (X, S_i)$

Thus, by the trace-determinant-criterion based on the sign of the determinant one can once again only encounter unstable saddle points or stable nodes such as sinks, stars or spirals [16].

In the case of our non-hyperbolic points, they can be seen as the transition points from stable to unstable branches without actually having a clear stability. However, one can make the argument that the fast system looses track at the non-hyperbolic points giving them an unstable behaviour. Therefore, one gets the following condition for the non-hyperbolic points to be unstable

$$\begin{aligned} \det(J(X, S_i)|_{(X^*_{\text{nh}}, S^*_{i,\text{nh}})}) < 0 &\Longleftrightarrow \epsilon\left[\gamma_X\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in}\right) - A_{syn}\frac{\alpha_X\beta_1}{\gamma_x}\frac{2cd\sqrt{\frac{d}{c}}}{4d^2}\right] < 0 \\ &\Longleftrightarrow \frac{\gamma_x\gamma_X}{A_{syn}\alpha_X\beta_1}\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in}\right) < \frac{1}{2}\sqrt{\frac{c}{d}} \\ &\Longleftrightarrow \boxed{\sqrt{\frac{d}{c}} < \frac{1}{2m}} \end{aligned} \tag{3.78}$$

where in the last transformation again for simplicity the parameter were re-defined according to (3.3). not true!!!, due to complex solution one divides by 0 in (1,2) matrix entry. maybe just take real part?

In the case of the turning point, it has to be unstable. Hence, due to the Jacobian being independent of X as well as the turning point being independent of S_i one gets a condition on $S^*_{i,\mathrm{tp}}$

$$\begin{aligned}
\det(J(X,S_i)|_{(X^*_{\mathrm{tp}},S^*_{i,\mathrm{tp}})}) < 0 &\iff \epsilon\left[\gamma_X\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right) - A_{syn}\frac{\alpha_X\beta_1}{\gamma_x}\frac{2cd(cS^*_{i,\mathrm{tp}}-e)}{\left((cS^*_{i,\mathrm{tp}}-e)^2+d\right)^2}\right] < 0 \\
&\iff \frac{\gamma_x\gamma_X}{A_{syn}\alpha_X\beta_1}\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right) < \frac{2cd(cS^*_{i,\mathrm{tp}}-e)}{\left((cS^*_{i,\mathrm{tp}}-e)^2+d\right)^2} \\
&\iff \boxed{m(cS^*_{i,\mathrm{tp}}-e)^4+2dm(cS^*_{i,\mathrm{tp}}-e)^2-2cd(cS^*_{i,\mathrm{tp}}-e)^2+md^2<0}
\end{aligned} \tag{3.79}$$

Finally, the flow direction on the slow manifold can once more be determined by looking at the sign of $\dot{X}$ in the limit case of the fast system (3.69) and $\dot{S}_i$ in the limit case of the slow system (3.80) by substituting points left, right or in between of the non-hyperbolic points in their respective right-hand sides. Herefore, the required slow system in the limit case $\epsilon \to 0$ is briefly stated

$$\begin{aligned}
0 &= \frac{\alpha_X A_{syn}}{\gamma_x\gamma_X}\left(\beta_x+\beta_1\frac{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r}-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k^{(1)}_+}{k^{(1)}_-\gamma_R}S_i-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k^{(1)}_-}\right]^2}{\left[\left(\frac{\alpha_R\beta_r}{\gamma_r}-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}\right)\frac{k^{(1)}_+}{k^{(1)}_-\gamma_R}S_i-\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}k^{(1)}_-}\right]^2+\frac{k^{(2)}_-}{k^{(2)}_+}K_x}\right)-\gamma_X X \\
\dot{S}_i &= -\left(\frac{c_{sec}\gamma_{ex}}{\rho_V c_{sec}+\gamma_{ex}}+\gamma_{in}\right)S_i + A_{syn}X - \frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a}}
\end{aligned} \tag{3.80}$$

However, due to the time constraints of this thesis and the necessary lengthy calculations the last part of the analysis is omitted at this point.

3.4 Analytical limiting processes

Last but not least, we can look at different limiting cases. The aim is to see what happens analytically in the submodels if certain parameters obtain special values.

QslA submodel → Basic model

By taking $\beta_{q_a} \to 0$ or $\alpha_{Q_a} \to 0$ or $\gamma_{q_a} \to \infty$ in the extended QslA model (3.22) its solution trajectories after time scale analysis (3.26) resolve to be those of the Basic model (3.13). Furthermore, all calculations regarding the genetic switch of the QslA submodel (3.27d) are simplified to those of the Basic model (3.14d). Thus, also the bifurcation analysis in Figure 3.2e is equal to Figure 3.1.

QteE submodel → Basic model

Since the qteE gene expression is regulated by LasRd protein, by taking $\beta_{q_e} \to 0$ and $\beta_2 \to 0$ or $\alpha_{Q_e} \to 0$ or $\gamma_{q_e} \to \infty$ in the extended QteE model (3.37) its solution trajectories at least to the point of determining R_2 after time scale analysis (3.40f) resolve to be those of the Basic model (3.13). Although for the QteE submodel it was not possible to determine the existence of a genetic switch analytically, by taking the above limits the genetic switch and the bifurcation diagram are the same as in (3.14d) and Figure 3.1.

Basal lasI transcription only

By taking $\beta_1 \to 0$ in the Basic model (3.9) only basal gene transcription is considered. Interestingly, although the autoinducer can still be produced, its time course is constant and a genetic switch can not be formed.

$$
\begin{aligned}
(3.14c) \qquad & \xLeftrightarrow{\beta_1 \to 0} \qquad 0 = \left(\frac{\rho_V c_{sec}^2}{\rho_V c_{sec} + \gamma_{ex}} - c_{sec} - \gamma_{in} \right) S_i + \frac{\alpha_X A_{syn}}{\gamma_x \gamma_X} \beta_x \\
& \Longleftrightarrow \qquad S_i = \frac{\alpha_X A_{syn} \beta_x}{\gamma_x \gamma_X \left(\frac{c_{sec} \gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in} \right)}
\end{aligned}
\tag{3.81}
$$

High QslA concentration

By taking $\beta_{q_a} \to \infty$ or $\alpha_{Q_a} \to \infty$ or $\gamma_{q_a} \to 0$ in the extended QslA model (3.22) its solution trajectories after time scale analysis (3.26) show the behaviour described in (3.29). $R_2 < 0$, $R < 0$, $Q_2 < 0$ and for certain parameter values $Q_a \in \mathbb{C}$ lead to a non biologically meaningful solution. Furthermore, the genetic switch of the QslA submodel (3.27d) does not exist since we have $t > 0$ instead of the required condition $t < 0$ from (3.30).

High QteE concentration

By taking $\beta_{q_e} \to \infty$ and $\beta_2 \to \infty$ or $\alpha_{Q_e} \to \infty$ or $\gamma_{q_e} \to 0$ in the extended QteE model (3.37) its solution trajectories at least to the point of determining R_2 after time scale analysis (3.40f) depict the effects seen in (3.42). One gets $R_2 < 0$ and $R < 0$ which is not biologically meaningful. Similar to the calculations for $\frac{\alpha_R \beta_r}{\gamma_r} = \frac{\alpha_{Q_e} \beta_{q_e}}{\gamma_{q_e}}$ in (3.45) in our case we would still get a formation of autoinducer, however no genetic switch can be formed.

4 Estimation of all unknown parameters

To exactly estimate all unknown parameters listed in Table 4.1 where in most cases only certain upper and lower bounds are already known, one has to use experimental data and apply it to the theory of parameter estimation as stated in section 2.5.

The experimental data used in this thesis has been acquired by (citation?). In several independent and reproducible experiments the DNA of the green fluorescent protein (GFP) has been added in-vivo to the gene lasI, which is the main indicator for measuring the intracellular autoinducer concentration. The produced GFP is an important protein used in experimental biology to measure the amount of gene expression of a gene over time. By exposing the resulting transcriptome or proteome to ultraviolet light it will exhibit green fluorescence. Hence, the experimental data contains a median GFP value which was measured after certain time intervals. In addition, the optical density (OD) after the same time steps was measured. An OD is a magnitude of cell density where 1 OD is equivalent to $1 \cdot 10^9 [\frac{\text{Zellen}}{\text{ml}}]$. Overall, both GFP and OD were measured for wild-types, gene knockouts of only one anti-activator as well as gene knockouts of both anti-activators. For each type the experiment was conducted three times to ensure reliability of the produced data. While a wild-type is a variant where all genes are active, a gene-knockout is a technique where at least one target gene has been manually inactivated. Thus, the experimental data represents the impact that one or more missing genes have on the gene expression of the whole system.

Since the experimental data (citation?) contains only measurements for GFP and not the actual gene expression of lasI, we have to extend the full model (3.4) by an ODE that translates the time development of lasI gene expression to measured GFP. Furthermore, the temporal bacterial growth due to cell division has not been addressed properly. Until now, the time dependent parameter $\rho_V := \rho_V(t)$ in the ODE for S_e states this bacteria growth without actually defining a necessary ODE. Both GFP development and bacterial growth have not yet been included in the analytical analysis of the model in section 3.2 since they have not been able to offer any further qualitative insight e.g. in the existence of a genetic switch.

Rate constant	Parameter	Value Range	Unit	Reference
Volume fraction	ρ_V	$\rho_V(t)$	cell	
volume of a single cell	V_{cell}	$6 \cdot 10^{-16}$	l	
total culture volume	V_{tot}	$1 \cdot 10^{-3}$	l	
cell culture growth rate	μ	$9.58 \cdot 10^{-1}$	s^{-1}	[25]
initial number of cells in the culture	P_0	$5.1 \cdot 10^3$	cell	
initial concentration of GFP	G_0	$1 \cdot 10^{-5}$	$\mu mol/l$	
Diffusion rate in / out	c_{sec}	$2.1 \cdot 10^{-3}$	s^{-1}	[9]
Extracellular degradation rate	γ_{ex}	$2 \cdot 10^{-4}$	s^{-1}	[9]
$LasI$ synthesis rate	A_{syn}	$5.2 \cdot 10^{-4}$	s^{-1}	[23]
$lasI$ threshold	K_x	$1.7 \cdot 10^{-1}$	$\mu mol/l$	[25]
$QteE$ degradation rate	γ_{Q_e}	0	s^{-1}	
$QslA$ degradation rate	γ_{Q_a}	0	s^{-1}	
$QslA - QslA$ degradation rate	γ_{Q_2}	0	s^{-1}	
$AI_i - LasR$ degradation rate	γ_{R_2}	0	s^{-1}	
$(AI_i - LasR)_2$ degradation rate	γ_{R_d}	0	s^{-1}	
$AI_i - LasR$ binding rate	$k_+^{(1)}$	$0.1 \cdot 10^{-3} - 5 \cdot 10^{-3}$	$(\mu mol/l)^{-1} s^{-1}$	[11]
$(AI_i - LasR)_2$ binding rate	$k_+^{(2)}$	$0.5 \cdot 10^{-1} - 5 \cdot 10^{-1}$	$(\mu mol/l)^{-1} s^{-1}$	[11]
$(AI_i - LasR)_2$ dissociation rate	$k_-^{(2)}$	$0.5 \cdot 10^{-4} - 5 \cdot 10^{-4}$	s^{-1}	[11]
$lasI$ basal transcription rate	β_x	$0.1 \cdot 10^{-6} - 4.5 \cdot 10^{-6}$	$(\mu mol/l) s^{-1}$	[11]
induced $lasI$ transcription rate	β_1	$1 \cdot 10^{-4} - 4 \cdot 10^{-4}$	$(\mu mol/l) s^{-1}$	[11]
$lasI - mRNA$ degradation rate	γ_x	$1.4 \cdot 10^{-3} - 7 \cdot 10^{-3}$	s^{-1}	[11]
$LasI$ protein formation rate	α_X	$0.8 \cdot 10^{-2} - 3.2 \cdot 10^{-2}$	s^{-1}	[11]
$LasI$ degradation rate	γ_X	$0.2 \cdot 10^{-4} - 4 \cdot 10^{-4}$	s^{-1}	[11]
$lasR - mRNA$ degradation rate	γ_r	$1.4 \cdot 10^{-3} - 7 \cdot 10^{-3}$	s^{-1}	[11]
$LasR$ protein formation rate	α_R	$0.8 \cdot 10^{-2} - 3.2 \cdot 10^{-2}$	s^{-1}	[11]
maximal capacity of cells in the culture	$K_{P-basic}$	$4.33 \cdot 10^9$	$\mu mol/l$	
$lasR - mRNA$ basal transcription rate	β_r	$0.9 \cdot 10^{-7} - 4.5 \cdot 10^{-5}$	$(\mu mol/l) s^{-1}$	
Intracellular degradation rate	γ_{in}	$2 \cdot 10^{-5} - 2 \cdot 10^{-3}$	s^{-1}	
$AI_i - LasR$ dissociation rate	$k_-^{(1)}$	$0.1 \cdot 10^{-7} - 5 \cdot 10^{-5}$	s^{-1}	
$LasR$ degradation rate	γ_R	$0.2 \cdot 10^{-5} - 4 \cdot 10^{-3}$	s^{-1}	
maximal capacity of cells in the culture	K_{P-qteE}	$3.85 \cdot 10^9$	$\mu mol/l$	
$qteE - mRNA$ basal transcription rate	β_{q_e}	$0.9 \cdot 10^{-7} - 4.5 \cdot 10^{-5}$	$(\mu mol/l) s^{-1}$	
induced $qteE - mRNA$ transcription rate	β_2	$1 \cdot 10^{-5} - 4 \cdot 10^{-3}$	$(\mu mol/l) s^{-1}$	
$qteE$ threshold	K_{q_e}	$1.7 \cdot 10^{-2} - 1.7 \cdot 10^0$	$\mu mol/l$	
$qteE - mRNA$ degradation rate	γ_{q_e}	$1.4 \cdot 10^{-4} - 7 \cdot 10^{-2}$	s^{-1}	
$QteE$ protein formation rate	α_{Q_e}	$0.8 \cdot 10^{-3} - 3.2 \cdot 10^{-1}$	s^{-1}	
$LasR - QteE$ binding rate	$k^{(3)}$	$7 \cdot 10^{-3} - 7 \cdot 10^{-1}$	$(\mu mol/l)^{-1} s^{-1}$	
maximal capacity of cells in the culture	K_{P-qslA}	$4.19 \cdot 10^9$	$\mu mol/l$	
$qslA - mRNA$ basal transcription rate	β_{q_a}	$0.9 \cdot 10^{-7} - 4.5 \cdot 10^{-5}$	$(\mu mol/l) s^{-1}$	
$qslA - mRNA$ degradation rate	γ_{q_a}	$1.4 \cdot 10^{-4} - 7 \cdot 10^{-2}$	s^{-1}	
$QslA$ protein formation rate	α_{Q_a}	$0.8 \cdot 10^{-3} - 3.2 \cdot 10^{-1}$	s^{-1}	
$QslA - QslA$ binding rate	$k_+^{(4)}$	$0.5 \cdot 10^{-2} - 5 \cdot 10^0$	$(\mu mol/l)^{-1} s^{-1}$	
$QslA - QslA$ dissociation rate	$k_-^{(4)}$	$0.5 \cdot 10^{-5} - 5 \cdot 10^{-3}$	s^{-1}	
$LasR - AI_i - QslA - QslA$ binding rate	$k^{(5)}$	$7 \cdot 10^{-3} - 7 \cdot 10^{-1}$	$(\mu mol/l)^{-1} s^{-1}$	
maximal capacity of cells in the culture	K_{P-full}	$3.51 \cdot 10^9$	$\mu mol/l$	
$LasR - AI_i - QslA - QslA$ degradation rate	$\gamma_{C_{R_2Q_2}}$	$0 - \infty$	s^{-1}	
GFP formation	η	$0 - \infty$	dimensionless	
GFP dilution	ζ	$0 - \infty$	$(\mu mol/l)^{-1}$	

Table 4.1: Unknown parameters used to describe rate constants with their value ranges, units and references. All parameters listed above the first horizontal line are exact values as well as biologically assumed values from the analytical section. All parameters listed between the first and second horizontal line have exact value ranges given in literature. All parameters listed between the second and third horizontal line have similar value ranges to biologically similar components from the first two parts since there are no bounds stated in literature. All parameters listed below the third horizontal line have no references to any biologically similar component or they are introduced solely for this model and are thus unbounded.

The bacterial population growth can be modelled as a simple logistic ODE with saturation effect to represent a finite environment

$$\dot{P} = \mu P\left(1 - \frac{P}{K_P}\right) \tag{4.1}$$

With this ODE we can finally give an exact definition of what ρ_V actually is

$$\rho_V = \rho_V(t) := \frac{V_{cell}}{V_{tot}} P(t) \tag{4.2}$$

where $P(t)$ is the solution to the IVP consisting of ODE (4.1) with initial value $P(0) = P_0$.

Modelling the time development of GFP however is a bit more complicated. By adding GFP-DNA to the lasI gene, its transcription and translation has to be interlinked with that of lasI. Thus, modelling the production of GFP has the same mathematical structure of protein formation as the LasI protein. Moreover, in contrast to LasI, GFP almost never degrades making it a viable option as a long-term fluorescent that can be detected wherever the protein is carried to even after a long period of time. However, incorporating the bacterial cell growth of a cell and its cell division, GFP will dilute into all daughter cells resulting in a decrease of GFP concentration in the mother cell. This can be modelled by taking the right-hand site of (4.1) coupled with GFP. This results in the ODE for temporal GFP development

$$\dot{G} = \eta\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_X}\right) - \zeta\mu P\left(1 - \frac{P}{K_P}\right) G \tag{4.3}$$

All newly introduced parameters from (4.1) and (4.3) can be also found in Table 4.1

The **full model for parameter estimation and simulation** can then be written by taking the full model (3.4), including all biologically motivated assumptions from

section 3.2 and adding bacterial growth (4.1) as well as GFP development (4.3)

$$
\begin{aligned}
\dot{S}_e &= \rho_V c_{sec}(S_i - S_e) - \gamma_{ex} S_e \\
\dot{S}_i &= c_{sec}(S_e - S_i) + A_{syn} X - \gamma_{in} S_i - k_+^{(1)} R S_i + k_-^{(1)} R_2 \\
\dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \gamma_X X \\
\dot{R} &= \frac{\alpha_R \beta_r}{\gamma_r} - \gamma_R R - k_+^{(1)} R S_i + k_-^{(1)} R_2 - k^{(3)} R Q_e \\
\dot{R}_2 &= -k_-^{(1)} R_2 + k_+^{(1)} R S_i - 2k_+^{(2)} R_2^2 + 2k_-^{(2)} R_d - \gamma_{R_2} R_2 - k^{(5)} R_2 Q_2 \\
\dot{R}_d &= -k_-^{(2)} R_d + k_+^{(2)} R_2^2 - \gamma_{R_d} R_d \\
\dot{Q}_e &= \frac{\alpha_{Q_e}}{\gamma_{q_e}}\left(\beta_{q_e} + \beta_2 \frac{R_d}{R_d + K_{q_e}}\right) - k^{(3)} R Q_e \\
\dot{Q}_a &= \frac{\alpha_{Q_a} \beta_{q_a}}{\gamma_{q_a}} - 2k_+^{(4)} Q_a^2 + 2k_-^{(4)} Q_2 \\
\dot{Q}_2 &= -k_-^{(4)} Q_2 + k_+^{(4)} Q_a^2 - \gamma_{Q_2} Q_2 - k^{(5)} R_2 Q_2 \\
\dot{C}_{R_2Q_2} &= k^{(5)} R_2 Q_2 - \gamma_{R_2Q_2} C_{R_2Q_2} \\
\dot{P} &= \mu P\left(1 - \frac{P}{K_P}\right) \\
\dot{G} &= \eta\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\right) - \zeta \mu P\left(1 - \frac{P}{K_P}\right) G
\end{aligned}
\tag{4.4}
$$

For numerically solving the above system of ODEs one needs **Initial Conditions** (IC). Herefore, we make the assumption that small amounts of LasI, LasR, QteE and QslA protein concentrations are present directly at the beginning due to basal DNA transcription and protein formation. Moreover, we assume GFP has already been added to the system in small amounts. Furthermore, the experimental data states an initial amount of bacteria cells in the culture. All other concentrations are set to zero. Thus,

the respective IC can be stated in the following way

$$
\begin{aligned}
S_e(0) &= 0 \\
S_i(0) &= 0 \\
X(0) &= \frac{\alpha_X \beta_x}{\gamma_x \gamma_X} \\
R(0) &= \frac{1}{\gamma_R}\left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_e}\beta_{q_e}}{\gamma_{q_e}}\right) \\
R_2(0) &= 0 \\
R_d(0) &= 0 \\
Q_e(0) &= \frac{\gamma_R}{k^{(3)}} \frac{1}{\frac{\alpha_R \beta_r \gamma_{q_e}}{\alpha_{Q_e}\beta_{q_e}\gamma_r} - 1} \\
Q_a(0) &= \sqrt{\frac{\alpha_{Q_a}\beta_{q_a}}{2\gamma_{q_a} k_+^{(4)}}} \\
Q_2(0) &= 0 \\
C_{R_2Q_2} &= 0 \\
P(0) &= P_0 \\
G(0) &= G_0
\end{aligned}
\tag{4.5}
$$

In particular, the IC for LasI, LasR, QteE and QslA were calculated by taking the quasi-steady state in (4.4) for their respective ODE, re-substituting every known IC for the respective variable and then solving the equation for the main variable of the ODE. The IC for GFP is structurally the same as the one for LasI since GFP-DNA is added to the lasI gene and not calculated over its ODE. However, the level of initial concentration is reduced by some orders of magnitude. Last but not least, the IC for the bacterial population is given by the experimental data.

Before estimating all parameters, one can make further useful assumptions to increase the probability that the optimization process converges in a reasonable time while decreasing the search domain. First, one should use all upper and lower parameter boundaries one can find in literature. This reduces the search space $\mathbb{R}_+^\theta$ to $I \subset \mathbb{R}_+^\theta$ where $I := \bigtimes_{i=1}^{\theta} I_i = [lb_1, ub_1] \times ... \times [lb_\theta, ub_\theta]$, as seen in Table 4.1. Second, parameters which describe similar processes are likely to have similar parameter values (citation?). Thus, also similar parameter ranges for the parameter estimation can be assumed. Third, it has been experimentally verified that newly synthesized GFP needs to undergo self-modifications to be fluorescent which takes only around ten minutes if one uses modified GFP (citation?). Still, this produces a lag in comparison to the LasI protein formation. Hence, during parameter estimation the experimental data has to be adapted accordingly. Fourth, one can directly use the additional information about measured OD given in the experimental data to estimate the maximal capacity

threshold of cells in the culture. The exact values are stated in Table 4.1 while the estimation is visualized in Figure 4.1. By realizing that all four estimated parameter values are slightly different it becomes clear that the ODE (4.1) is too simple and can not incorporate all effects shown in the data perfectly. However, the estimation is good enough since it depicts the growth phase sufficiently well. Fifth, to further reduce the complexity one starts by estimating parameter values of the smallest model and later on one increases the model size until the full model is reached. Hence, all estimated parameters from the Basic model can be used in the other submodels to reduce their search space immensely.

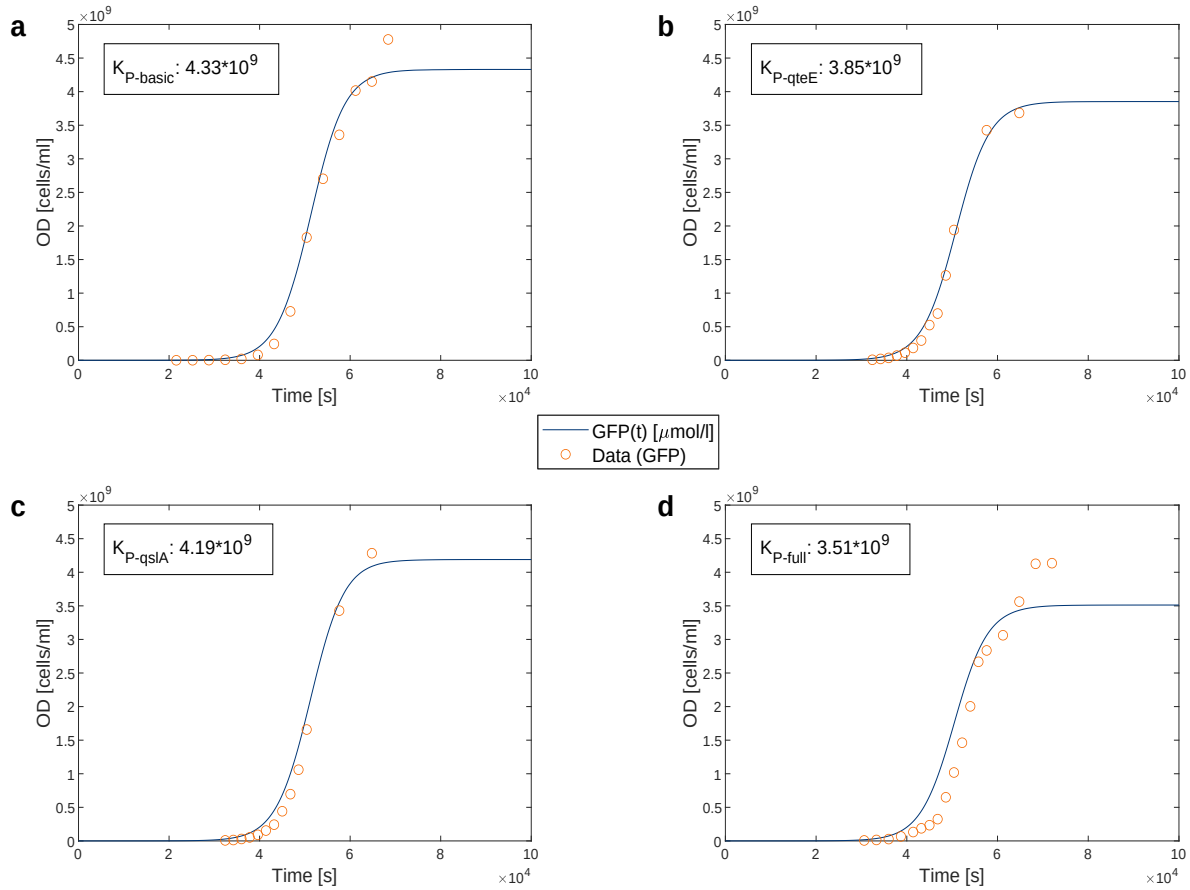

Figure 4.1: Estimation of the maximal capacity threshold of bacterial cells in the culture. (a) Estimation of $K_{P-basic}$ for the Basic model. (b) Estimation of K_{P-qteE} for the QteE submodel. (c) Estimation of K_{P-qslA} for the QslA submodel. (d) Estimation of K_{P-full} for the full model.

4.1 The complexity of parameter estimation

Using all previously stated assumptions as well as the mathematical theory of parameter estimation from section 2.5 one gets a not very convincing result for estimating all

parameters from the Basic model (Figure 4.2). One can see that all multistarts do not quite capture the course of the experimental data. Even the best estimates which show the highest convergence due to the lowest objective function value underestimate the highest peek and the drop in the middle of the time course. Moreover, the temporal delay at the beginning is underestimated.

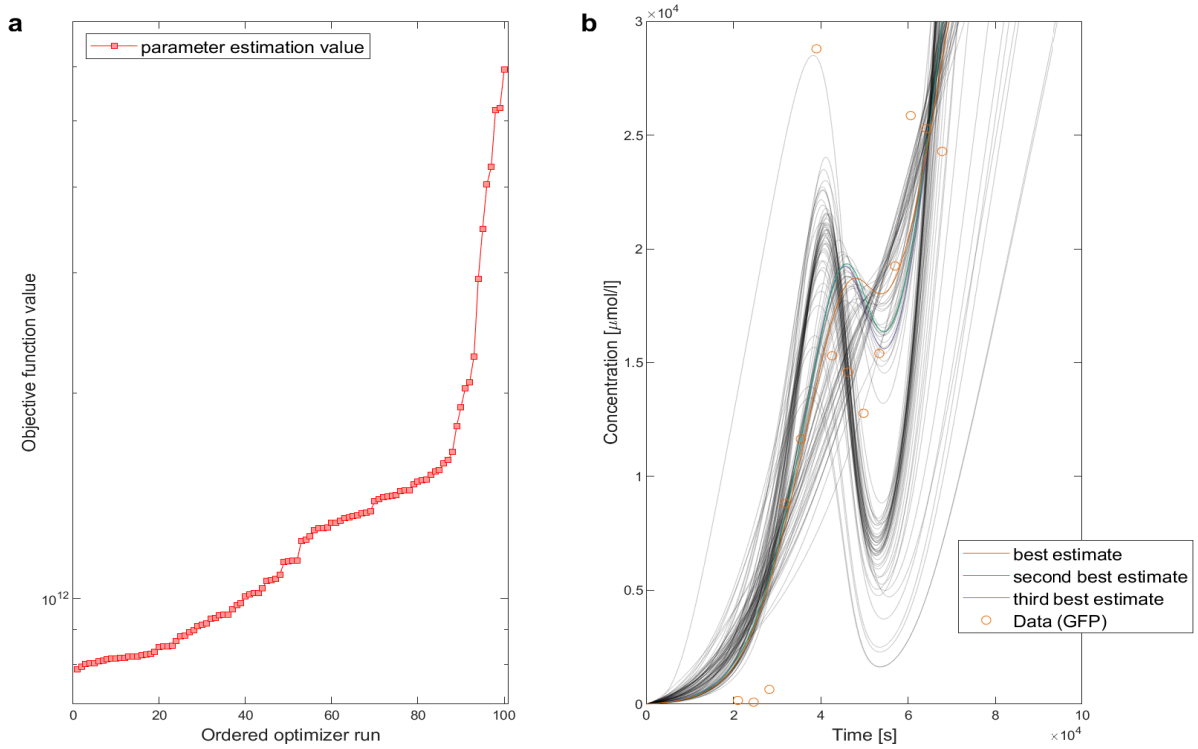

Figure 4.2: Estimation of all unknown parameters from the Basic model. (a) A total of 100 multistart optimizations sorted after their respective objective function value from low to high. (b) GFP trajectories as well as the experimental data points for the Basic model. The best, second best and third best trajectory ranked after their respective objective function value are colored.

Since 100 multistarts are frankly not enough such that the solution of the GFP ODE (4.3) converges to the experimental data it becomes necessary to rely on a different method. The problem seems to be that the search space $I \subset \mathbb{R}_+^{16}$ is still too large to find perfect solution values just from beginning at random uniformly distributed starting values. Thus, a solution could be to adapt the starting values manually until the solution trajectory resembles the course of the experimental data better. These manually determined parameter starting values then become the initial position from where the parameter estimation routine is started. The results for all submodels as well as the full model can be seen in Figure 4.3.

On the one hand, it becomes clear that manually setting all parameter values achieves a better fit visually speaking. Also, at least for the Basic model we see that the

objective function value of the manual set converges better than 64% of all multistarts we have seen in Figure 4.2. This means that by implication the parameter estimation routine starts off at a better position than by randomly guessing starting values. This also shows objectively that choosing parameter values manually can improve parameter estimation. On the other hand, if one compares the objective function value of the best fit from Figure 4.2 with both fits from Figure 4.3a for the Basic model, the latter fit is converging worse. The problem might be that setting the initial parameter values manually does not improve the objective function value enough. Another problem is depicted in both Figure 4.2 and Figure 4.3. Neither randomly nor manually chosen initial parameter values achieve a convergence of the GFP trajectory to the earliest experimental data points. However, accurately representing what the network does in the onset of Quorum Sensing is crucial for predicting its temporal progression. In total, we can conclude that only setting parameter values manually is still no solution for the problem because the parameter estimation routine is not able to find the perfect values. Although it could be possible to manually find better or even perfect parameter values, it is highly unlikely and based on a 13 to 22 dimensional hyperspace - depending on the submodel or full model - where an infinite amount of combinations can be taken this approach is far too time consuming.

4.2 Expanding the full model

The biggest issue that could correlate and maybe even cause the main problem, that the GFP trajectory does not properly converge to the experimental data, is the missing temporal delay which is visible in all three submodels as well as the full model (Figure 4.3). Unfortunately, it is not straight forward how to solve this problem. Several reasons such as a wrong biological connection between species or a missing extension to the full model might cause the problem and should therefore be addressed.

4.2.1 Double LasRd protein binding to the lasI gene

In chapter 1 it is stated that the protein LasRd binds to the promoter of the lasI gene and regulates its gene expression. Hence, a positive feedback loop inside the system is created which allows the existence of a genetic switch and activates Quorum Sensing. However, it is biologically possible that the promoter can even bind two LasRd homodimers at the same time (citation?). Only if both proteins are present the gene is activated and enhanced gene expression is possible. This concept slightly changes the mathematical approach, i.e. the Michaelis-Menten kinetics established in section 2.2 in the following way.

Let x = [X] and z = [Z] be the concentrations of the regulatory protein and the mRNA respectively. Since we have two binding sites let there exist two binding affinities $K_1, K_2 \in \mathbb{R}_+$. If we assume these binding sites to be independent e.g. the binding site

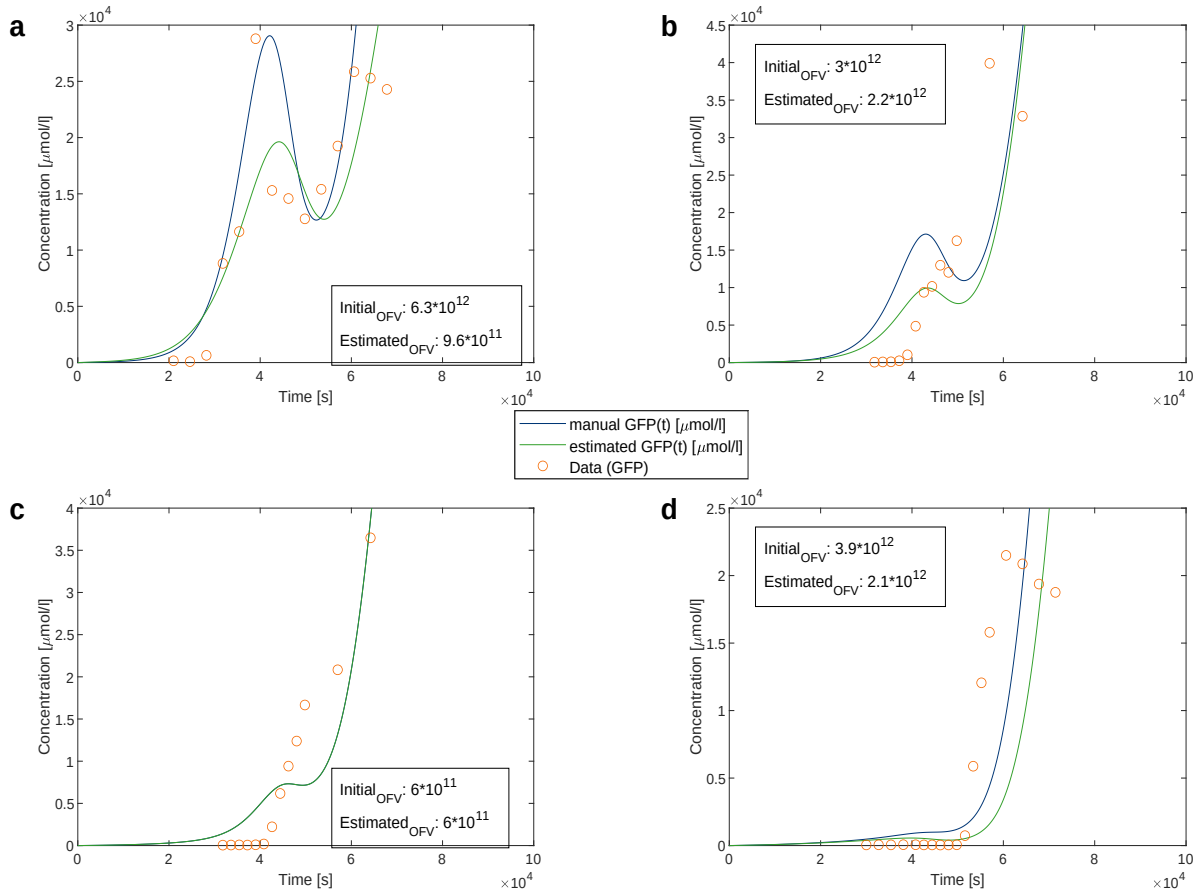

Figure 4.3: Manual setting and estimation of all unknown parameters for all submodels as well as the full model. (a) GFP trajectory and experimental data points for the Basic model (b) GFP trajectory and experimental data points for the QteE submodel (c) GFP trajectory and experimental data points for the QslA submodel. Both lines are plotted but overlap perfectly so only one line is visible. (d) GFP trajectory and experimental data points for the full model.
In all cases, the initial and estimated objective function value are stated.

as well as the binding order do not matter for the protein, then

$$\dot{z} = \beta \frac{x}{x + K_1} \frac{x}{x + K_2} - \gamma_z z \tag{4.6}$$

where again $\beta \in \mathbb{R}_+$ represents the transcription rate of the mRNA and $\gamma_z \in \mathbb{R}_+$ depicts the degradation rate of the protein.
Unfortunately, is is often impossible that both binding sites are independent. In most cases, the binding of one protein regulates the binding ability of another protein due to some sort of change in spatial conformation of the promoter [18]. Thus, equation (4.6)

is generalized into

$$\dot{z} = \beta \frac{x^2}{x^2 + \kappa^2} - \gamma_z z \tag{4.7}$$

using the same variables as before while $\kappa \in \mathbb{R}_+$ is the half-activating concentration of the regulatory protein. We can also express κ in terms of K_1, K_2 as $\kappa^2 = K_1 \cdot K_2$ [18].

The idea behind choosing this approach of double LasRd binding to the lasI promoter is that we can model the situation using (4.7) with a Hill coefficient of $n = 2$. Mathematically speaking, the higher the Hill coefficient the steeper the sigmoidal rise is at the beginning. Hence, the model is better suited to correctly represent the temporal delay that the experimental data requires (Figure 4.3).

Before running the parameter estimation one has to correctly adapt the full model (4.4). In this case, one has to rewrite the ODEs for the LasI protein $\dot{X}$ as well as for GFP $\dot{G}$ in the following way

$$\begin{aligned} \dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d^2}{R_d^2 + K_x^2}\right) - \gamma_X X \\ \dot{G} &= \eta\left(\beta_x + \beta_1 \frac{R_d^2}{R_d^2 + K_x^2}\right) - \zeta\mu P\left(1 - \frac{P}{K_P}\right) G \end{aligned} \tag{4.8}$$

All other ODEs from the full model (4.4) remain the same.

By looking at the results, one can see that the objective function value after the parameter estimation of the double LasRd system has hardly decreased (Figure 4.4c,d) and has even increased (Figure 4.4a,b) for some submodels. This means that relying on a double binding of LasRd protein does not increase the convergence of the model to the experimental results. Even visually speaking the trajectories after the parameter estimation for each submodel show a worse representation of the data time course. However, it is well visible that by having a double LasRd protein bind the lasI gene it leads to a temporal delay of the GFP trajectory in all submodels as well as the full model. Thus, the idea that an increase of the Hill coefficient may help to delay the rapid increase of GFP was correct but apparently not well suited for this model.

4.2.2 RsaL pathway extension

Another idea to help fix the problem would be that the full model (4.4) adequately represents the biological components but another pathway is missing which is crucial to describe the given experimental data. A well-researched pathway that inhibits lasI transcription and limits LasI protein formation is given by the transcriptional repressor gene rsaL. By also having the LasRd protein bind to the rsaL gene promotor, rsaL

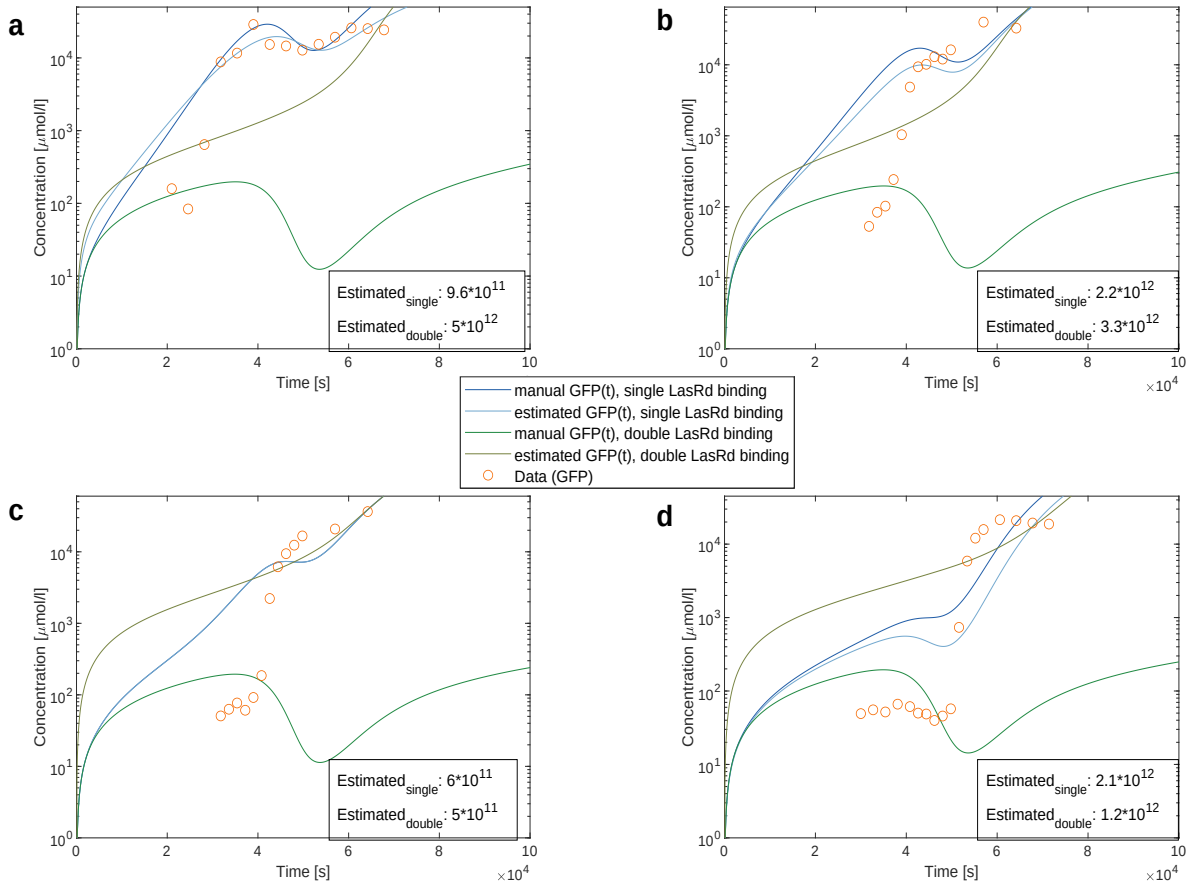

Figure 4.4: Manual setting and estimation of all unknown parameters for all submodels as well as the full model expanded by a single and double LasRd homodimer binding to the promoter of the lasI gene. (a) GFP trajectory and experimental data points for the Basic model (b) GFP trajectory and experimental data points for the QteE submodel (c) GFP trajectory and experimental data points for the Qsla submodel. Both lines for the single LasRd binding submodel are plotted but overlap perfectly so only one line is visible. (d) GFP trajectory and experimental data points for the full model.
In all cases, the estimated objective function values for both the single as well as the double protein binding are stated.

DNA is transcribed to rsal-mRNA and translated into RsaL protein. This protein then decreases LasI protein formation by binding to the lasI promotor, thus creating a second negative feedback loop. Meanwhile, the LasR protein formation remains unaffected. This pathway is described extensively in [3]. Again, this concept changes the mathematical approach, i.e. the Michaelis-Menten kinetics established in section 2.2 in the following way.

While one can model the transcription of lasR-mRNA and formation of LasR

protein directly as described in section 2.2

$$\begin{aligned} \dot{l} &= \beta_l + \beta_0 \frac{R_d}{R_d + K_l} - \gamma_l l \\ \dot{L} &= \alpha_L l - \gamma_L L \end{aligned} \tag{4.9}$$

where the explanation of all new variables and parameters can be found in Table 4.2, it is a novelty in this thesis to model an inhibitory effect of a protein binding to a promoter.

Let z = [Z] be the concentrations of the mRNA. Let x = [X] as well as y = [Y] be the concentration of the binding protein with an increasing as well as decreasing effect. Since we have two binding sites let there exist two binding affinities $K_+, K_- \in \mathbb{R}_+$. If we assume these binding sites to be independent e.g. the binding site as well as the binding order do not matter for the protein, then

$$\dot{z} = \beta \frac{x}{x + K_+} \frac{K_-}{y + K_-} - \gamma_z z \tag{4.10}$$

where again $\beta \in \mathbb{R}_+$ represents the transcription rate of the mRNA and $\gamma_z \in \mathbb{R}_+$ depicts the degradation rate of the protein [18].

The idea behind adding the RsaL pathway is that in the early stages the small amount of available LasRd protein increases the existence of RsaL protein which itself decreases the lasI gene expression. However, in the later stages of the time course, LasRd will increase LasI protein formation more than RsaL can decrease, resulting in a rapid growth of LasI protein concentration. By finding correct parameter values via parameter estimation, the extended model could correctly represent the temporal delay that the experimental data requires (Figure 4.3).

In the context of our model using the rsaL-mRNA, the LasRd protein and the RsaL protein we get the extended ODEs for lasI-mRNA and LasI protein formation

$$\begin{aligned} \dot{x} &= \beta_x + \beta_1 \frac{R_d}{R_d + K_x} \frac{K_H}{L + K_H} - \gamma_x x \\ \dot{X} &= \alpha_X x - \gamma_X X \end{aligned} \tag{4.11}$$

By using the steady state assumption similar to section 3.2 we get

$$\begin{aligned} 0 = \dot{l} &= \beta_l + \beta_0 \frac{R_d}{R_d + K_l} - \gamma_l l \\ 0 = \dot{x} &= \beta_x + \beta_1 \frac{R_d}{R_d + K_x} \frac{K_H}{L + K_H} - \gamma_x x \end{aligned} \tag{4.12}$$

and hence solving for (l^*, x^*)

$$\begin{aligned} l^* &= \frac{1}{\gamma_l}\left(\beta_l + \beta_0 \frac{R_d}{R_d + K_l}\right) \\ x^* &= \frac{1}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\frac{K_H}{L + K_H}\right) \end{aligned} \tag{4.13}$$

we get the new ODEs for Rsal and LasI protein formation over time by substituting the steady states (4.13) into the respective equations for protein formation in (4.9) and (4.11), resulting in

$$\begin{aligned} \dot{L} &= \frac{\alpha_L}{\gamma_l}\left(\beta_l + \beta_0 \frac{R_d}{R_d + K_l}\right) - \gamma_L L \\ \dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\frac{K_H}{L + K_H}\right) - \gamma_X X \end{aligned} \tag{4.14}$$

With equations (4.14) we can formulate the changes one needs to make before performing parameter estimation. In this case, one has to add an ODE for the RsaL protein $\dot{L}$ and rewrite the ODEs for the LasI protein $\dot{X}$ as well as for GFP $\dot{G}$ in the following way

$$\begin{aligned} \dot{L} &= \frac{\alpha_L}{\gamma_l}\left(\beta_l + \beta_0 \frac{R_d}{R_d + K_l}\right) - \gamma_L L \\ \dot{X} &= \frac{\alpha_X}{\gamma_x}\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\frac{K_H}{L + K_H}\right) - \gamma_X X \\ \dot{G} &= \eta\left(\beta_x + \beta_1 \frac{R_d}{R_d + K_x}\frac{K_H}{L + K_H}\right) - \zeta\mu P\left(1 - \frac{P}{K_P}\right)G \end{aligned} \tag{4.15}$$

All other ODEs from the full model (4.4) remain the same.

By looking at the results, one can see that the objective function value after the parameter estimation of the system with RsaL pathway inclusion has hardly decreased (Figure 4.4a,d) and has even increased (Figure 4.4b,c) for some submodels. This means that the addition of this particular pathway does not increase the convergence of the model to the experimental results. Visually speaking, for each submodel the trajectories with the RsaL pathway show an almost exact representation of the trajectories without it. In total, a temporal delay of the GFP trajectory by including the RsaL pathway is not visible at all in neither one of the submodels nor the full model. Thus, the original idea that a second negative feedback loop may help to delay the rapid increase of GFP at the beginning turned out incorrect, at least for this specific pathway.

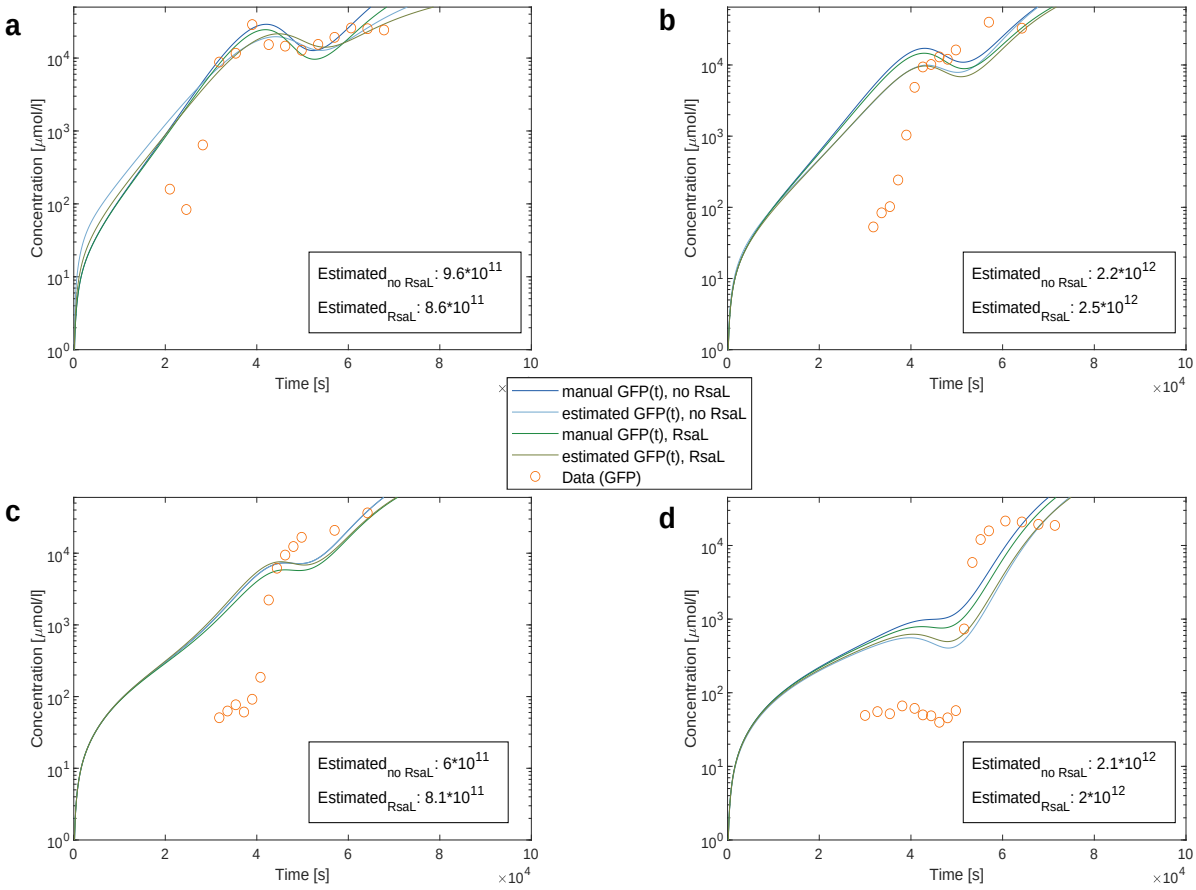

Figure 4.5: Manual setting and estimation of all unknown parameters for all submodels as well as the full model expanded by the RsaL pathway which inhibits lasI gene expression. (a) GFP trajectory and experimental data points for the Basic model (b) GFP trajectory and experimental data points for the QteE submodel (c) GFP trajectory and experimental data points for the Qsla submodel. Both lines for the submodel without RsaL pathway are plotted but overlap perfectly so only one line is visible. (d) GFP trajectory and experimental data points for the full model. In all cases, the estimated objective function value for the model without as well as with additional RsaL pathway are stated.

4.3 Comparison to Goryachev et al.

Another possibility to achieve a sufficient temporal delay in the early stages of the time trajectories could be to compare the model design of the full model to another model which already shows such a delay. A prime example for such a model has been introduced in the paper of Goryachev et al. [11] and models the Tra-system for Quorum Sensing in the bacteria *Agrobacterium tumefaciens*.
Similar to *Pseudomonas aeruginosa* it is a Gram-negative bacteria that is controlled by a small gene expression network which shows an almost identical construction of interacting mRNA, proteins and metabolites resulting in one important positive and

Concentration	**Variable**
$rsaL - mRNA$	l
$RsaL$	L
Rate Constant	**Parameter**
$RsaL$ protein formation rate	α_L
$rsaL - mRNA$ degradation rate	γ_l
$rsaL$ basal transcription rate	β_l
induced $rsaL$ transcription rate	β_0
$RsaL$ threshold	K_l
$RsaL$ degradation rate	γ_L
$lasI$ threshold	K_H

Table 4.2: Variables and Parameters used to identify concentrations and rate constants for the RsaL pathway extension.

one negative feedback loop (Figure 4.6). Hence, it is possible to compare both circuits and their modelling processes based on the law of mass action (section 2.1) and the Michaelis-Menten kinetics (section 2.2). Herefore, we can identify TraI with LasI, TraR with LasR, TraRd with LasRd, AAI with S_i and S_e and finally we can compare the anti-activator TraM with QslA or QteE. In case of the anti-activator, it can not be identified directly to either one of our anti-activators since its binding to LasRd can not be found in the circuit of *Pseudomonas aeruginosa*. However, its pathway is similar and thus at least its parameter values and its behaviour over time can be compared.
Furthermore, the paper depicts the trajectories of LasRd and TraM as seen in Figure 4.7 after a stochastic as well as deterministic numerical simulation. Although it shows the trajectories as amount of molecules over cell density, the time trajectories are implicitly given. Due to the cell division of bacteria the growth in cell density goes hand in hand with the temporal progression. Thus, every point on the displayed trajectories show a certain point in time. Therefore, the temporal delay of the LasRd trajectory can be seen. One can further assume that this would also hold for the GFP trajectory since its ODE representation is modelled similarly. Still, this has to be shown.

The main goal is to compare both modelling processes and results to find their similarities and differences. In particular, by looking at their deviations it can become clear why our modelling approach results in such an early ascent of the GFP time trajectories. Therefore, based on the model of Goryachev one still has to model and analyse the ODE for GFP. In the end, by then applying the precise concept which is responsible for the delay from Goryachev's model to our model this problem might be fixed and the parameter estimation routine could have a higher possibility to map the experimental data better.
Therefore, it is mandatory to look at the ODE system established from the circuit

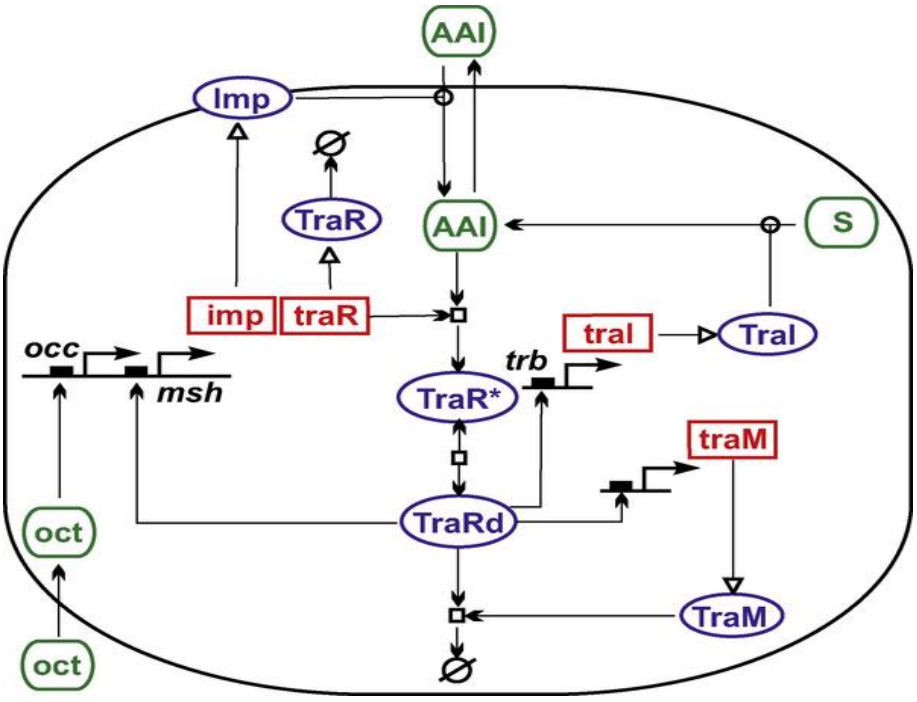

Figure 4.6: Quorum-Sensing Network of an Octopine-Type Agrobacterium Ti Plasmid
Blue ellipses represent proteins, red rectangles indicate mRNA species, and green flattened circles denote metabolites. Open circle arrowheads represent enzymatic activation of a reaction or transport, and open triangle arrowheads denote translation. Essential bimolecular reactions are shown explicitly as open squares. S represents two substrates of the autoinducer synthetase TraI, and Ø denotes protein degradation. DOI: 10.1371/journal.pcbi.0010037.g001.
This figure as well as the caption are directly taken from [11].

(Figure 4.6), add an equation for GFP and do a numerical analysis which can be compared to (Figure 4.7). Only then, by alternating different parameters or even pathways it is possible to determine the differences in our model.

The paper states a core system containing two ODEs which the authors derived from their full model of 16 ODEs with 30 parameters. However, on the one hand the documentation on how the reduced model of two ODEs has been calculated is very sparse. On the other hand the Figure 4.7 shows the behaviour of TraM whose ODE is not stated anywhere. In addition, the ODE for the extracellular autoinducer is stated after the simplification process but not in its simplified form. Because of these difficulties one has to re-derive the ODEs for the extracellular autoinducer and TraM by hand and adapt the ODEs for TraRd accordingly. After lenghty calculations one get

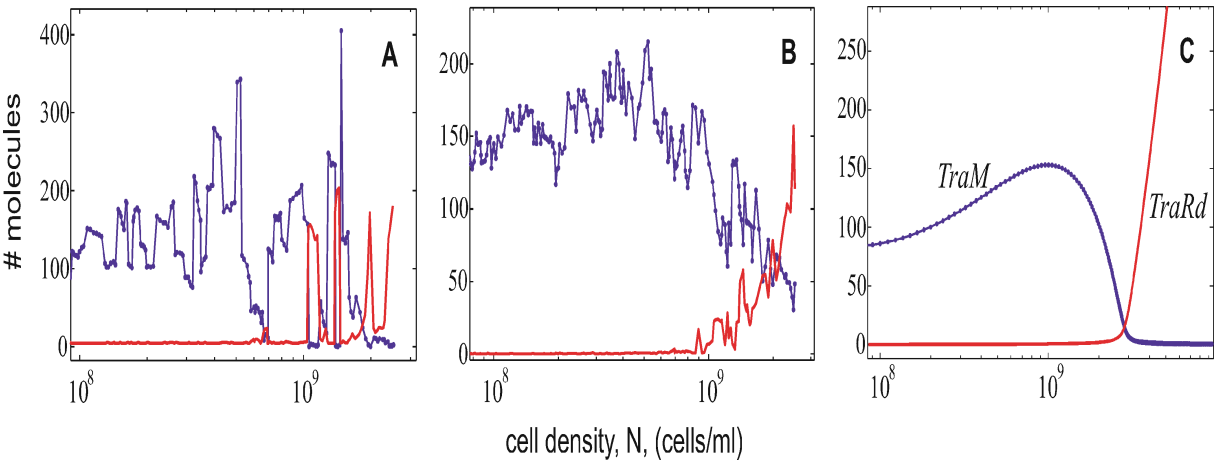

Figure 4.7: Transition of a Model Bacterial Population to Quorum Sensing in a Homogeneous Liquid Medium
Intracellular concentrations of TraRd (thick red line) and TraM (thin blue line, filled circles) are plotted against the population density that exponentially grows with time. (A) Dynamics of an individual cell in the stochastic population model. (B) Dynamics of the stochastic population model averaged over ten bacteria. (C) Behavior of the deterministic population model. DOI: 10.1371/journal.pcbi.0010037.g005.
This figure as well as the caption are directly taken from [11].

the following reduced system consisting of five ODEs

$$\begin{aligned}
\dot{A}_e &= V'N\left[k_2(A_i - A_e) + \left(\left(k_7 + k_{28}\frac{D}{\frac{k_{30}}{k_{29}} + D}\right) + k_4 A_e\right)\left(\frac{k_4 k_5}{k_3 + k_5}A_e - \frac{A_e}{\frac{k_{14}}{k_4} + A_e}\right)\right] \\
\dot{A}_i &= \frac{k_{21}k_{25}}{k_{26}k_{24}}\left(k_{27} + k_{21}\frac{D}{\frac{k_{23}}{k_{22}} + D}\right) + k_2(A_e - A_i) - A_i\frac{k_6}{k_{12}}\left(k_7 + k_{28}\frac{D}{\frac{k_{30}}{k_{29}} + D}\right) \\
\dot{D} &= A_i\frac{k_6}{2k_{12}}\left(k_7 + k_{28}\frac{D}{\frac{k_{30}}{k_{29}} + D}\right) - k_{11}DM - k_{10}D \\
\dot{M} &= \frac{k_{19}k_{15}}{k_{18}}\frac{D}{\frac{k_{17}}{k_{16}} + D} - k_{11}DM - k_{20}M \\
\dot{N} &= \alpha N
\end{aligned} \tag{4.16}$$

In addition, one can derive the ODE for GFP based on the considerations from (4.3)

$$\dot{G} = \eta\left(k_{27} + k_{21}\frac{D}{D + \frac{k_{23}}{k_{22}}}\right) - \zeta\alpha NG \tag{4.17}$$

All calculations as well as the explanation of all used parameters as well as all their

parameter values can be found in [11]. The explanation of all variables is also briefly stated here for the sake of clarity (Table 4.3).

Concentration	**Variable**
extracellular autoinducer	A_e
intracellular autoinducer	A_i
$TraRd$	D
$TraM$	M
number of cells	N
green fluorescence protein	G

Table 4.3: Variables used to identify concentrations for Goryachev's model [11].

With the ODE system containing equations (4.16) and (4.17) one can now look at the results of the numerical simulation (Figure 4.8). It is interesting to see that for the first simulation (Figure 4.8a,b,c) which uses the ODE for unlimited bacterial growth a numerical error appears. While the autoinducer concentration goes to infinity almost instantly, the GFP concentration simultaneously goes to zero. Since this representation of a continuous biological system is not valid, the first approach and easiest way to fix this problem could be to bound the bacterial growth by a saturation effect. Although this is not done in the paper of Goryachev it makes sense from a biological point of view that the bacterial growth is always limited by environmental constraints. Therefore, we can introduce the adapted ODE for bacterial growth using the logistic ODE

$$\dot{N} = \alpha N \left(1 - \frac{N}{K_N}\right) \tag{4.18}$$

where K_N is the maximal possible capacity of cells.

Using this ODE in the otherwise identical system we get the adapted numerical simulation result (Figure 4.8d,e,f) where we see that all trajectories are continuous without any jumps or sudden changes in behaviour. Moreover, one can identify two main results. First of all, the Genetic Switch of Quorum Sensing is clearly visible as stated in the paper. While the concentration of internal and external autoinducer as well as TraRd is small at the beginning, it does increase in a switch-like behaviour once the bacterial density reaches a certain threshold. At this point the anti-activator TraM decreases and the autoinducer production is not largely inhibited anymore. Furthermore, if one compares the plot with the deterministic simulation from the paper (Figure 4.7) with the plot using its system of ODEs (Figure 4.8e) it matches nicely. Second of all, by introducing GFP into the system the temporal delay of its concentration at the beginning of the simulation is clearly visible (Figure 4.8f). Therefore, this underlines the capability of this type of modelling approach depicted in

the paper to achieve such a delay.

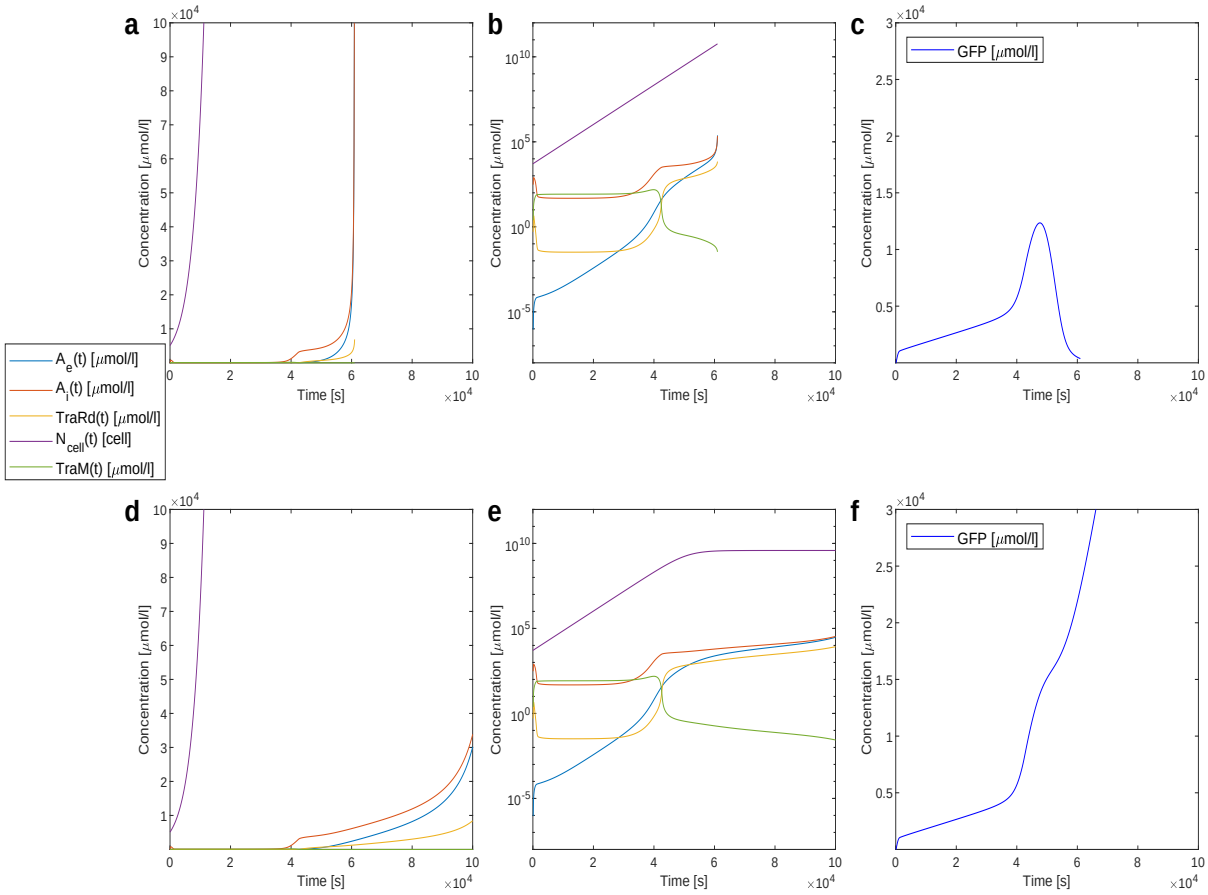

Figure 4.8: Numerical simulation of the ODE system stated in Goryachev's paper [11] as well as the extension (4.18) due to a numerical integration error. (a) The simulation of Goryachev's ODE system without GFP in linear y-axis scale. (b) The simulation of Goryachev's ODE system without GFP in log-transformed y-axis scale. (c) The simulation of just the GFP ODE for Goryachev's ODE system. (d) The simulation of the extended ODE system without GFP in linear y-axis scale. (e) The simulation of the extended ODE system without GFP in log-transformed y-axis scale. (f) The simulation of just the GFP ODE for the extended ODE system.

To find the precise component in the system of Goryachev that is responsible for the delay one has to look at the differences in creating this system. Thus, by comparing the network structure of the circuit for *Agrobacterium tumefacien* (Figure 4.6) with that of *Pseuomonas aeruginosa* (Figure 1.1) the biggest differences are

- no degradation of internal autoinducer A_i
- no dissociation of TraR* complex
- active transport of internal autoinducer A_i out of the cell

- differentiating vacant and occupied transcription binding sites
- no involvement of TraR in the system
- the only anti-activator TraM binds directly to TraRd

In the following, the possible impact of these differences on the temporal delay of the GFP trajectory is discussed. In both cases, the respective exclusion of degradation as well as dissociation does not have a big impact on the trajectories since in our model these rates $\gamma_{in} \in [2 \cdot 10^{-5}, 2 \cdot 10^{-3}]$ as well as $k_{-}^{(2)} \in [0.5 \cdot 10^{-4}, 5 \cdot 10^{-4}]$ are very small. The active transport of internal autoinducer A_i outside of the cell is modelled by an enzymatic reaction. In the system (4.16) in the ODE for A_i this transport is modelled by $-A_i \frac{k_6}{k_{12}} \left(k_7 + k_{28} \frac{D}{\frac{k_{30}}{k_{29}} + D} \right)$. Thus, in addition to the linear outflow of A_i we also have in our system (3.1), there is a sigmoidal decline based on LasRd production. Since a sigmoidal curve is small at the start, the outflow of A_i is not increased much at the beginning which makes a significant temporal delay based on this change not possible. Similarly, using ODEs to model the differences in vacant and occupied transcription binding factors as connection between proteins and gene promoters leads to a comparable result to our model. While our system directly models the transcription of DNA to mRNA by the binding process of a protein i.e. LasRd to the promotor of the lasI gene after the Michaelis-Menten kinetics (2.2) as

$$\dot{x} = \beta_x + \beta_1 \frac{R_d}{R_d + K_x} - \gamma_x x \tag{4.19}$$

the paper of Goryachev states a different possible way by introducing vacant and occupied binding sited. Since the actual calculations can not be found in the paper [11] they are stated here

$$\begin{aligned} \dot{x} &= k_{27} + k_{21} trb_o - k_{24} x \\ \dot{trb}_o &= k_{22} R_d trb_v - k_{23} trb_o \\ \dot{trb}_v &= -k_{22} R_d trb_v + k_{23} trb_o \end{aligned} \tag{4.20}$$

This modelling approach makes sense since actual mRNA transcription can only happen if the binding site is occupied by a protein. Hence, $\dot{x}$ depends on trb_o. By using the quasi-steady state assumption on $\dot{trb}_o$ and $\dot{trb}_v$ one gets

$$\begin{aligned} \dot{x} &= k_{27} + k_{21} trb_o - k_{24} x \\ 0 &= k_{22} R_d trb_v - k_{23} trb_o \\ 0 &= -k_{22} R_d trb_v + k_{23} trb_o \end{aligned} \tag{4.21}$$

Finally, one can use the argument that either a transcription side is vacant or closed, so their states always sum up to the available transcription side, i.e. $trb_v + trb_o = 1$. Thus,

we get in particular for the equation of trb_o

$$0 = k_{22}R_d(1 - trb_o) - k_{23}trb_o \qquad \Longleftrightarrow \qquad trb_o = \frac{R_d}{R_d + \frac{k_{23}}{k_{22}}} \tag{4.22}$$

and hence

$$\dot{x} = k_{27} + k_{21}\frac{R_d}{R_d + \frac{k_{23}}{k_{22}}} - k_{24}x \tag{4.23}$$

depicts the same result as in our case. Therefore, introducing vacant and occupied transcription binding sites will have no impact on the temporal delay of our trajectories. The biggest differences between both networks are definitely the involvement of the TraR or LasR protein as well as the exact protein where the anti-activator binds. This does not only lead to a change in one ODE but in the whole system. Without actually changing our system in several key components and running a numerical simulation for every change the resulting consequences can not simply be predicted. Due to time constraints on this master thesis these excessive modifications are not investigated further. Although by changing only one of these components a temporal delay in the trajectory of GFP might by visible, the resulting system of ODEs would probably not represent the original experimentally verified network of *Pseudomonas Aeruginosa* anymore.

4.4 Fitting all parameters

In the last three sections, we have seen that the main problem of having the numerical simulation of the GFP ODE fit better to the experimental data depends on the ability of the system to achieve a temporal delay at the beginning. However, all attempts to fix the problem have either not really been applicable to our model or have had no huge impact on the behaviour of the GFP trajectory.
Although this problem still persists the question arises whether an imperfect parameter estimation can still generate trajectories which incorporate valid results and statements about Quorum sensing and the genetic switch. Our best attempt so far has been to manually set the parameter values and then use the parameter estimation routine to improve the results (Figure 4.3). Here, one can see that the GFP trajectory roughly imitates the course of the experimental data over time. Therefore, one could get valuable insights into the behaviour of the system in general by taking a look at all other trajectories.

Still, one can make one further useful adjustment. By taking a look at the experimental data for the Basic model at (citation?) or directly in Figure 4.3a there exists one outlier which neither aligns with the remaining time course of this submodel nor with that of any other submodel. The problem is that this outlier could also be the reason for the bad fit since the parameter estimation tries to fit all data points although the model does not have the capability. This then leads to a bad objective

function value and a visually poor fit. Moreover, by falsely determining the parameter values for the Basic model all further submodels become inaccurate since they all use the estimated parameter values from the Basic model. Thus, this specific outlier can be omitted to see how the parameter estimation performs. The results can be seen in the following plot (Figure 4.9). Again, all parameter values for the submodels were set manually to achieve a visually nice fit of the GFP trajectory and the experimental data. Afterwards, the parameter estimation routine tried to improve all parameter values. In comparison to the first try in Figure 4.3 this fit has clearly reduced objective function values in both the manual fit as well as the actual parameter estimation. This indicates a far better convergence of the trajectory to the experimental data. Also visually the improvement can be seen in all submodels. Moreover, in contrast to all earlier parameter estimations, this parameter estimation noticeably reduced the objective function value of the estimation compared to the manual fit. This implies that dropping the outlier led to a parameter search space where local minima can be found more easily. Only now it becomes clear that the outlier has a major impact on the performance of parameter estimation. The estimated parameter values resulting in the so far best fit of the GFP trajectory and the experimental data are stated in Table 4.4.

With the results from Figure 4.3 one can finally look at the rest of the time trajectories for all other concentrations inside and outside the cell. Therefore, all estimated parameter values from Table 4.4 are substituted into the full system of *Pseudomonas aeruginosa* (4.4) as fixed values. Hence, all submodels as well as the full model can be simulated for the given IC (4.5). The results can be seen in Figure 4.10 which are very interesting for several reasons.
Looking at the Basic model where during the experiment both anti-activator genes were knocked-out, the trajectories depict a sigmoidal growth over time. At some point, the density of bacteria is high enough and a switch-like behaviour drastically increases the production of internal autoinducer. This explains the huge growth in external autoinducer because at that point more internal autoinducer can be transported through the cell membrane. Interestingly, the concentration of LasR is the only concentration that declines from its initial value. This makes sense since LasR protein formation happens only via basal DNA transcription and after the onset of the genetic switch more LasR is needed to form the LasR* complex than it is produced.
Looking at the QteE submodel where during the experiment the qslA anti-activator gene was knocked-out, we can see the same sigmoidal growth in all concentrations except for LasR and QteE. Especially the trajectory for the anti-activator QteE depicts the genetic switch really well. Until the critical threshold of bacteria cells in the culture it stays at the level but after the onset of the genetic switch and the growth of the autoinducer it declines. However, due to the descent of LasR over time, QteE protein can not inactivate LasR that quickly and in addition it is produced more than just basally. This explains the rise of the QteE trajectory near the steady state.

Rate constant	Parameter	Value Range	Unit	Best Fit
Volume fraction	ρ_V	$\rho_V(t)$	cell	$\rho_V(t)$
volume of a single cell	V_{cell}	$6 \cdot 10^{-16}$	l	$6 \cdot 10^{-16}$
total culture volume	V_{tot}	$1 \cdot 10^{-3}$	l	$1 \cdot 10^{-3}$
cell culture growth rate	μ	$9.58 \cdot 10^{-1}$	s^{-1}	$9.58 \cdot 10^{-1}$
initial number of cells in the culture	P_0	$5.1 \cdot 10^{3}$	cell	$5.1 \cdot 10^{3}$
initial concentration of GFP	G_0	$1 \cdot 10^{-5}$	$\mu mol/l$	$1 \cdot 10^{-5}$
Diffusion rate in / out	c_{sec}	$2.1 \cdot 10^{-3}$	s^{-1}	$2.1 \cdot 10^{-3}$
Extracellular degradation rate	γ_{ex}	$2 \cdot 10^{-4}$	s^{-1}	$2 \cdot 10^{-4}$
$LasI$ synthesis rate	A_{syn}	$5.2 \cdot 10^{-4}$	s^{-1}	$5.2 \cdot 10^{-4}$
$lasI$ threshold	K_x	$1.7 \cdot 10^{-1}$	$\mu mol/l$	$1.7 \cdot 10^{-1}$
$QteE$ degradation rate	γ_{Q_e}	0	s^{-1}	0
$QslA$ degradation rate	γ_{Q_a}	0	s^{-1}	0
$QslA - QslA$ degradation rate	γ_{Q_2}	0	s^{-1}	0
$AI_i - LasR$ degradation rate	γ_{R_2}	0	s^{-1}	0
$(AI_i - LasR)_2$ degradation rate	γ_{R_d}	0	s^{-1}	0
$AI_i - LasR$ binding rate	$k_+^{(1)}$	$0.1 \cdot 10^{-3} - 5 \cdot 10^{-3}$	$(\mu mol/l)^{-1} s^{-1}$	$5 \cdot 10^{-3}$
$(AI_i - LasR)_2$ binding rate	$k_+^{(2)}$	$0.5 \cdot 10^{-1} - 5 \cdot 10^{-1}$	$(\mu mol/l)^{-1} s^{-1}$	$5 \cdot 10^{-1}$
$(AI_i - LasR)_2$ dissociation rate	$k_-^{(2)}$	$0.5 \cdot 10^{-4} - 5 \cdot 10^{-4}$	s^{-1}	$1.29 \cdot 10^{-4}$
$lasI$ basal transcription rate	β_x	$0.1 \cdot 10^{-6} - 4.5 \cdot 10^{-6}$	$(\mu mol/l) s^{-1}$	$0.26 \cdot 10^{-6}$
induced $lasI$ transcription rate	β_1	$1 \cdot 10^{-4} - 4 \cdot 10^{-4}$	$(\mu mol/l) s^{-1}$	$3.54 \cdot 10^{-4}$
$lasI - mRNA$ degradation rate	γ_x	$1.4 \cdot 10^{-3} - 7 \cdot 10^{-3}$	s^{-1}	$3.64 \cdot 10^{-3}$
$LasI$ protein formation rate	α_X	$0.8 \cdot 10^{-2} - 3.2 \cdot 10^{-2}$	s^{-1}	$3.2 \cdot 10^{-2}$
$LasI$ degradation rate	γ_X	$0.2 \cdot 10^{-4} - 4 \cdot 10^{-4}$	s^{-1}	$3.62 \cdot 10^{-4}$
$lasR - mRNA$ degradation rate	γ_r	$1.4 \cdot 10^{-3} - 7 \cdot 10^{-3}$	s^{-1}	$7 \cdot 10^{-3}$
$LasR$ protein formation rate	α_R	$0.8 \cdot 10^{-2} - 3.2 \cdot 10^{-2}$	s^{-1}	$2.19 \cdot 10^{-2}$
maximal capacity of cells in the culture	$K_{P-basic}$	$4.33 \cdot 10^{9}$	$\mu mol/l$	$4.33 \cdot 10^{9}$
$lasR - mRNA$ basal transcription rate	β_r	$0.9 \cdot 10^{-7} - 4.5 \cdot 10^{-5}$	$(\mu mol/l) s^{-1}$	$1.34 \cdot 10^{-5}$
Intracellular degradation rate	γ_{in}	$2 \cdot 10^{-5} - 2 \cdot 10^{-3}$	s^{-1}	$7.8 \cdot 10^{-5}$
$AI_i - LasR$ dissociation rate	$k_-^{(1)}$	$0.1 \cdot 10^{-7} - 5 \cdot 10^{-5}$	s^{-1}	$2.34 \cdot 10^{-5}$
$LasR$ degradation rate	γ_R	$0.2 \cdot 10^{-5} - 4 \cdot 10^{-3}$	s^{-1}	$4 \cdot 10^{-3}$
maximal capacity of cells in the culture	K_{P-qteE}	$3.85 \cdot 10^{9}$	$\mu mol/l$	$3.85 \cdot 10^{9}$
$qteE - mRNA$ basal transcription rate	β_{q_e}	$0.9 \cdot 10^{-7} - 4.5 \cdot 10^{-5}$	$(\mu mol/l) s^{-1}$	$4.32 \cdot 10^{-6}$
induced $qteE - mRNA$ transcription rate	β_2	$1 \cdot 10^{-5} - 4 \cdot 10^{-3}$	$(\mu mol/l) s^{-1}$	$1 \cdot 10^{-5}$
$qteE$ threshold	K_{q_e}	$1.7 \cdot 10^{-2} - 1.7 \cdot 10^{0}$	$\mu mol/l$	$1.7 \cdot 10^{0}$
$qteE - mRNA$ degradation rate	γ_{q_e}	$1.4 \cdot 10^{-4} - 7 \cdot 10^{-2}$	s^{-1}	$6.61 \cdot 10^{-3}$
$QteE$ protein formation rate	α_{Q_e}	$0.8 \cdot 10^{-3} - 3.2 \cdot 10^{-1}$	s^{-1}	$8.13 \cdot 10^{-7}$
$LasR - QteE$ binding rate	$k^{(3)}$	$7 \cdot 10^{-3} - 7 \cdot 10^{-1}$	$(\mu mol/l)^{-1} s^{-1}$	$7 \cdot 10^{-1}$
maximal capacity of cells in the culture	K_{P-qslA}	$4.19 \cdot 10^{9}$	$\mu mol/l$	$4.19 \cdot 10^{9}$
$qslA - mRNA$ basal transcription rate	β_{q_a}	$0.9 \cdot 10^{-7} - 4.5 \cdot 10^{-5}$	$(\mu mol/l) s^{-1}$	$0.85 \cdot 10^{-7}$
$qslA - mRNA$ degradation rate	γ_{q_a}	$1.4 \cdot 10^{-4} - 7 \cdot 10^{-2}$	s^{-1}	$7.2 \cdot 10^{-4}$
$QslA$ protein formation rate	α_{Q_a}	$0.8 \cdot 10^{-3} - 3.2 \cdot 10^{-1}$	s^{-1}	$3.93 \cdot 10^{-3}$
$QslA - QslA$ binding rate	$k_+^{(4)}$	$0.5 \cdot 10^{-2} - 5 \cdot 10^{0}$	$(\mu mol/l)^{-1} s^{-1}$	$0.52 \cdot 10^{-2}$
$QslA - QslA$ dissociation rate	$k_-^{(4)}$	$0.5 \cdot 10^{-5} - 5 \cdot 10^{-3}$	s^{-1}	$1.82 \cdot 10^{-4}$
$LasR - AI_i - QslA - QslA$ binding rate	$k^{(5)}$	$7 \cdot 10^{-3} - 7 \cdot 10^{-1}$	$(\mu mol/l)^{-1} s^{-1}$	$7 \cdot 10^{-1}$
maximal capacity of cells in the culture	K_{P-full}	$3.51 \cdot 10^{9}$	$\mu mol/l$	$3.51 \cdot 10^{9}$
$LasR - AI_i - QslA - QslA$ degradation rate	$\gamma_{C_{R_2Q_2}}$	$0 - \infty$	s^{-1}	$10 \cdot 10^{0}$
GFP formation	η	$0 - \infty$	dimensionless	$7 \cdot 10^{3}$
GFP dilution	ζ	$0 - \infty$	$(\mu mol/l)^{-1}$	$1 \cdot 10^{-10}$

Table 4.4: Unknown parameters used to describe rate constants with their value ranges, units and best fits according to parameter estimation. Every parameter range was taken from literature stated in Table 4.1.
It has to be mentioned that while the estimated parameter values using MATLAB have a precision of 6 decimal places they are rounded here to two decimals. To get the precise values or values with an ever higher precision the scripts stated in (URL_{Github}) can be used and modified.

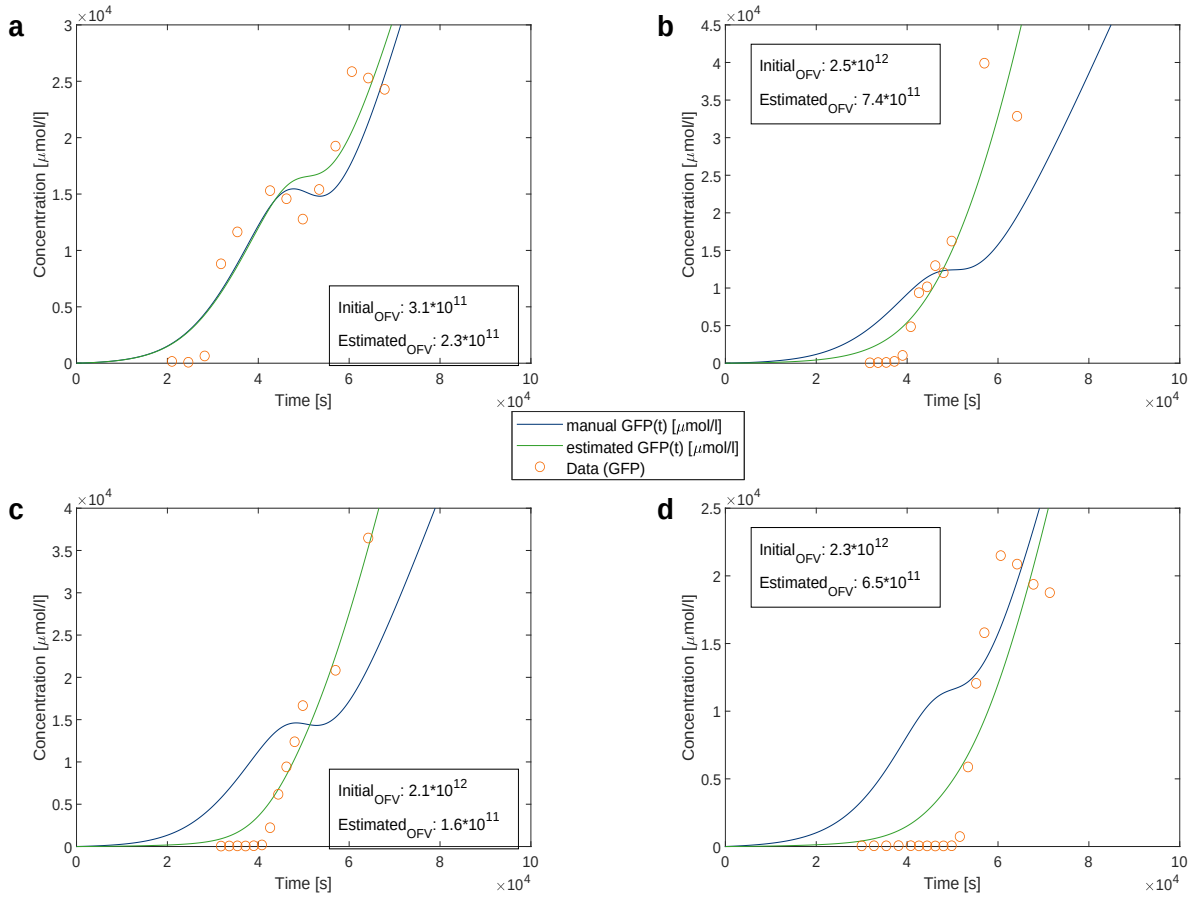

Figure 4.9: Manual setting and estimation of all unknown parameters for all submodels as well as the full model with the exclusion of the outlier. (a) GFP trajectory and experimental data points for the Basic model (b) GFP trajectory and experimental data points for the QteE submodel (c) GFP trajectory and experimental data points for the QslA submodel. Both lines are plotted but overlap perfectly so only one line is visible. (d) GFP trajectory and experimental data points for the full model.
In all cases, the initial and estimated objective function value are stated.

Looking at the QslA submodel where during the experiment the qteE anti-activator gene was knocked-out, one can see a similar behaviour as for the previous submodel.

However, the impact of this anti-activator is even better visible. Directly at the critical threshold the concentration of QslA declines rapidly while the concentrations of all other proteins rise. Similarly to the reason behind the temporal decline of LasR protein, also the QslA protein homodimer decreases because the qslA gene transcription happens only basally in our model. Thus, the inactivation of LasR* is near the steady state higher than its own production.
Finally looking at the full model where during the experiment no gene knockouts occurred, all previously discussed results can be found. In addition, it is very interesting to see that the anti-activators have no effect on themselves since their trajectories are identical to the ones displayed in the other submodels but they have a synergistic effect on the concentrations of the other trajectories. Both anti-activators combined have the highest temporal delay on the growth of the other protein concentrations. This means that a higher bacterial threshold has to be acquired to turn on the genetic switch and produce increased gene expression of our target genes. This is exactly what one would predict from all experimental insight.

In total, one can make some remarks about all submodels as well as the full model. We can identify a similar behaviour of one or both anti-activators and their impact on other protein concentrations compared to the plots of Goryachev's system (Figure 4.8). Moreover, if one looks closely at the plots or directly at the exact numerical result of the trajectories the QteE anti-activator has a bigger impact on the delay of the onset of the genetic switch than the QslA anti-activator. This effect could be caused by the amount of reactions that are inhibited by the anti-activator. Since QteE inactivates LasR, the formation of LasR* and LasRd are reduced while for QslA only the formation of LasRd is diminished. Furthermore, in all submodels as well as the full model all protein concentrations eventually run into the steady state due to the saturation effect of the bacterial growth in the finite environment. If one substitutes the estimated parameter values from Table 4.4 into the steady state solutions of the fast Basic system (3.13) as well as steady state solutions of the extended QslA system (3.26) their final values precisely coincide with the values seen in the previous plots. Also interesting is that the external autoinducer concentration is always lower than the internal autoinducer concentration. This is due to the fast production rate of the internal autoinducer compared to its transportation rate through the cell membrane to the extracellular space.
Since all previously described results seen in the submodels as well as the full model (Figure 4.10) match the expectations and experimental predictions of that system pretty well it shows that an imperfect parameter estimation still can offer a lot of insight into a highly complex system.

4.5 The reliability of all estimated parameter values

Last but not least, one can check how reliable all estimated parameter values from Table 4.4 are. In general, one has to calculate the model and prediction uncertainties

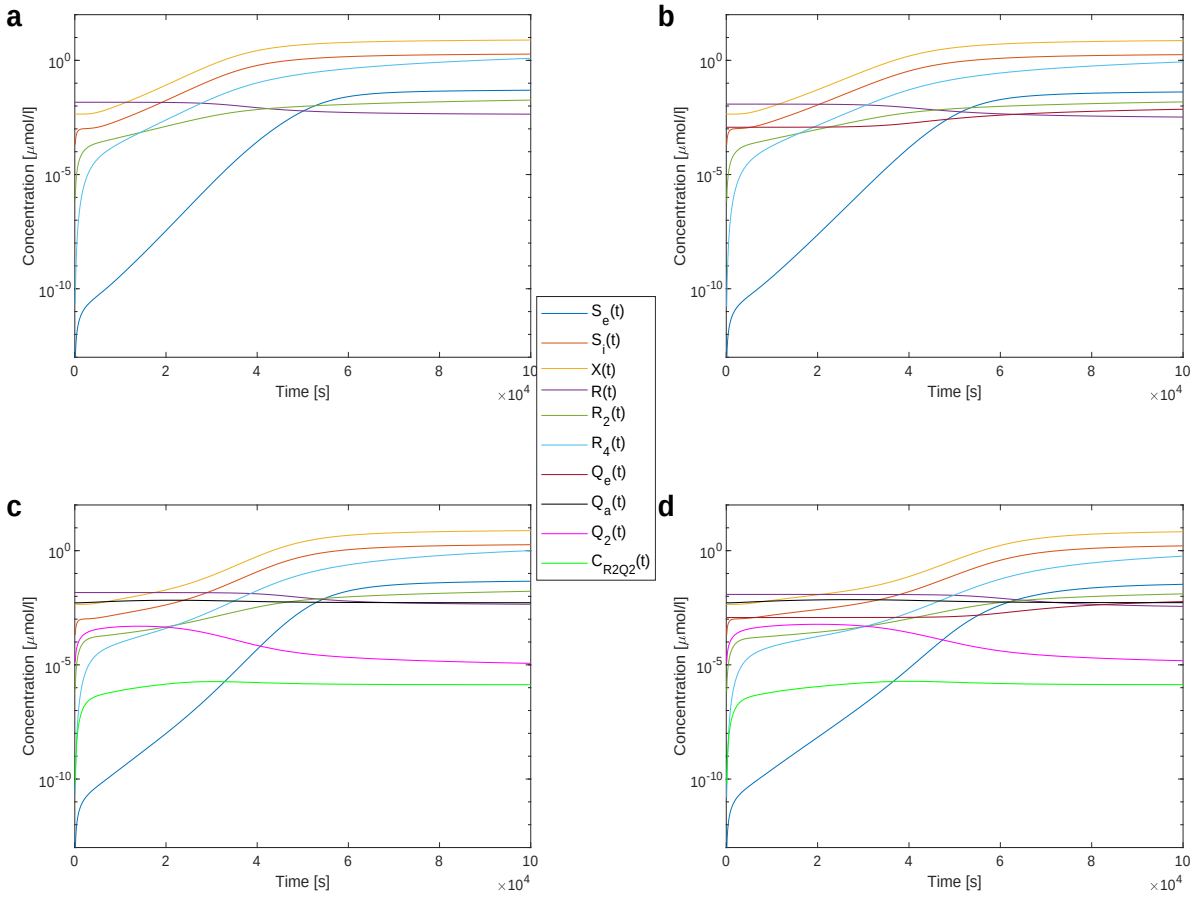

Figure 4.10: Simulation of all submodels as well as the full model using the estimated parameter values. (a) Trajectories for the Basic model (b) Trajectories for the QteE submodel (c) Trajectories for the Qsla submodel. (d) Trajectories for the full model.

but due to time constraints on this thesis this part is omitted at this point. However, we can check the necessary criteria for having a genetic switch stated in the model analysis chapter (3.2).

Basic model

By substituting all estimated parameter values into Definition 3.1 one gets

$$\begin{aligned} m &:= \frac{\gamma_x \gamma_X}{\alpha_X A_{syn} \beta_1} \left(\frac{c_{sec} \gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in} \right) \in [m_l, m_h] = [0.477, 0.487] \\ u &:= \frac{\beta_x}{\beta_1} = 7.34 \cdot 10^{-4} \\ c &:= k_+^{(2)} (k_+^{(1)})^2 \alpha_R^2 \beta_r^2 = 1.08 \cdot 10^{-18} \\ d &:= k_-^{(2)} (k_-^{(1)})^2 \gamma_R^2 \gamma_r^2 K_x = 9.41 \cdot 10^{-24} \end{aligned} \tag{4.24}$$

where $m = m(t)$ is time dependent due to $\rho_V = \rho_V(t)$ and therefore has values in an interval going from its lowest value m_l where $\rho_V(t) = \frac{V_{cell}}{V_{tot}} K_{Pfull} = 2.11 \cdot 10^{-3}$ is at its highest to its highest value m_h where $\rho_V(t) = \frac{V_{cell}}{V_{tot}} P_0 = 3.06 \cdot 10^{-9}$ is at its lowest.

If one substitutes these values into the condition for the existence of a genetic switch (3.17) one gets

$$\begin{aligned} &0 < 7.34 \cdot 10^{-4} < \frac{5 + 3\sqrt{3}}{4} \\ &0 \geq 8.5744 \cdot 10^{-40}, \quad \text{for } m = m_l \\ &0 \geq 8.5741 \cdot 10^{-40}, \quad \text{for } m = m_h \end{aligned} \tag{4.25}$$

While the first part of the condition is indeed fulfilled, the second part is not correct for any value of m in its interval. The problem is that the value for d is too small. Increasing d to a value around 10^{-19} would already be enough to fulfill the second part of the condition.

Extended QslA model

By substituting all estimated parameter values into Definition 3.3 one gets

$$\begin{aligned} m &:= \frac{\gamma_x \gamma_X}{\alpha_X A_{syn} \beta_1} \left(\frac{c_{sec} \gamma_{ex}}{\rho_V c_{sec} + \gamma_{ex}} + \gamma_{in} \right) \in [m_l, m_h] = [0.477, 0.487] \\ u &:= \frac{\gamma_x \gamma_X}{\alpha_X A_{syn} \beta_1} \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} - \frac{\beta_x}{\beta_1} = -6.81 \cdot 10^{-4} \\ c &:= \left(\frac{\alpha_R \beta_r}{\gamma_r} - \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a}} \right) \frac{k_+^{(1)}}{k_-^{(1)} \gamma_R} = 2.23 \\ d &:= \frac{k_-^{(2)}}{k_+^{(2)}} K_x = 4.39 \cdot 10^{-5} \\ e &:= \frac{\alpha_{Q_a} \beta_{q_a}}{2\gamma_{q_a} k_-^{(1)}} = 9.91 \cdot 10^{-3} \end{aligned} \tag{4.26}$$

where $m = m(t)$ is again time dependent due to $\rho_V = \rho_V(t)$ and therefore has values in an interval going from its lowest value m_l where $\rho_V(t) = \frac{V_{cell}}{V_{tot}} K_{Pfull} = 2.11 \cdot 10^{-3}$ is at its highest to its highest value m_h where $\rho_V(t) = \frac{V_{cell}}{V_{tot}} P_0 = 3.06 \cdot 10^{-9}$ is at its lowest.

If one substitutes these values into the conditions (3.31) and (3.32) for the existence of a genetic switch one gets

$$\begin{aligned} &0 > -6.81 \cdot 10^{-4} \\ &0 \geq -0.1306, \quad \text{for } m = m_l \\ &0 \geq -0.1306, \quad \text{for } m = m_h \end{aligned} \tag{4.27}$$

We can see that both conditions are fulfilled for any value of m in its interval.

Extended QteE model

By substituting all estimated parameter values into Definition 3.5 one gets

$$\begin{aligned} m &:= \frac{k_-^{(1)}\gamma_R}{k_+^{(1)}\beta_2} = 1.87 \\ u &:= \frac{S_i}{\beta_2}\left(\frac{\alpha_R\beta_r}{\gamma_r} - \frac{\alpha_{Q_e}\beta_{q_e}}{\gamma_{q_e}}\right) = 4.19 \cdot S_i \\ c &:= k_+^{(2)} = 5 \cdot 10^{-1} \\ d &:= k_-^{(2)} K_{q_e} = 2.19 \cdot 10^{-4} \end{aligned} \tag{4.28}$$

where in this case m is not time dependent since it does not depend on $\rho_V = \rho_V(t)$. The reason for this is that we do not calculate the actual genetic switch for the internal autoinducer S_i but rather the stationary solutions for R_2. Therefore, in the upper equation for t we still have S_i as a variable. As a reminder, calculating the genetic switch for the extended QteE submodel has not been possible yet since already solving the ODE for R_2 has led to a cubic equations where the sign of the discriminant was unknown.

If one substitutes these values into the conditions (3.42) and (3.44) one can finally determine the sign of the discriminant and thus the amount of real solutions R_2 has

$$\begin{aligned} &4.19 \cdot 10^{-5} \geq 9.6 \cdot 10^{-11} \\ &\Delta = 3.11 \cdot 10^{-12} - 4.06 \cdot 10^{-9}\left(4.19S_i - \frac{-5+\sqrt{537}}{8}\right)\left(4.19S_i - \frac{-5-\sqrt{537}}{8}\right) + \\ &\qquad 2.21 \cdot 10^{-7}(4.19S_i - 1)^3(4.19S_i + 2) \end{aligned} \tag{4.29}$$

We can see that the first condition is fulfilled. Because the sign of Δ depends on the value of S_i one can only calculate the conditions for the existence of a genetic switch in this submodel by determining the possible exact real solutions of R_2 in dependence of the cases for S_i. Since this is still too much effort for almost no benefit it is omitted.

In conclusion, the estimated parameter values fulfill most analytically derived conditions for the existence of a genetic switch. Still, there is some room for improvement since the one wrong condition indicates an inaccuracy of the estimated parameter values.

5 Discussion and Outlook

The goal of this master thesis was to mathematically model an experimentally verified gene network for the pathogenic bacteria *Pseudomonas aeruginosa* and to determine whether it expresses Quorum Sensing as well as a genetic switch. In particular, the influence of several anti-activator genes on the regulation of gene expression was investigated. Herefore, a system of coupled non-linear ODEs was created and analysed to get a holistic understanding of the underlined cellular processes. Furthermore, the full model was divided into three submodels to also gain qualitative insight in all integral components of the full system. While the analytical analysis of the system focused on proving the existence of a genetic switch and a region of bistable behaviour, the numerical analysis aimed at estimating previously unknown parameter values and solving the system to get trajectories for all variables over a full time course. During this process, there appeared some interesting results. The main observations are discussed in the following.

Firstly, in the Basic model as well as the QslA submodel we were able to derive conditions on the unknown parameter values under which the existence of a genetic switch and thus Quorum Sensing is guaranteed. For the QteE submodel as well as the full model these calculations were still possible however due to the complexity of the ODE system not practicable in the time frame of this thesis.
Secondly, in the Basic model as well as the QslA submodel it was possible to calculate further conditions on the unknown parameter values with regards to the stability of the genetic switch. Here, one was able to show that the genetic switch does not only exist but it is also a true genetic switch which exhibits relaxation oscillations.
Thirdly, a full parameter estimation for all 34 unknown parameter values was performed. Using these estimated values, the temporal behaviour of all important variables confirmed the existence of Quorum Sensing since at a certain bacteria threshold the gene expression was drastically increased. Due to the limited environmental resources of the bacteria all protein concentrations ran eventually into a stationary state.
Fourthly, both anti-activators had a delaying temporal effect on the gene expression and thus Quorum Sensing. While the effect of the qteE anti-activator was stronger than the one for the qslA anti-activator, both anti-activators combined showed a synergistic impact. By excluding both anti-activators the onset of Quorum Sensing appeared almost instantly at the beginning of gene expression.
Fifthly, by comparing the numerical simulation results to the analytical calculations we saw a great match between both results. After substituting the estimated parameter values into the derived calculations all conditions but one were fulfilled. Moreover, the one condition which did not coincide with the estimated parameter values showed that

the parameter estimation still has to be improved.

Although this master thesis showed that mathematical modelling of a biological system is vital to gain insight into complex cellular processes and to predict the interaction of several components over a vast time period it can be extended to show a deeper as well as broader picture.
Foremost, it is important to enhance the parameter estimation routine to acquire even better results based on the objective function value. One idea could be to start at the manually set parameter values from Figure 4.9 and use random uniformly distributed parameter values in a small environment around that initial set. Since that set has been the most promising one it is likely that by small perturbations an even better one can be found. Another idea could be to use a specifically designed toolbox for parameter estimations such as D2D [22] for MATLAB or pyPESTO [24] for Python which offers a variety of possibilities to find the global optimum.
Another important aspect would be to extend the admittedly complex but still arguably small model. By only modelling the interaction of 10 proteins including two anti-activators the overlap to the extensive gene machinery of *Pseudomonas aeruginosa* is small. Hence, all predictions of the behaviour of Quorum Sensing as well as the interaction between bacteria are only small approximations of reality. Therefore, including more experimentally verified components such as the RsaL pathway seen in section 4.2.2 or the LasB pathway [1] could contribute to a more holistic understanding of the biological processes. Although an analytical solution might then not exist, numerically analysing the system is still an option.
Finally, a very interesting question with regards to application in medicine arises whether it is possible to completely turn off Quorum Sensing in vivo. Then, bacteria would not be able to increase their own gene expression and accumulate a high enough density to overwhelm the immune system. Since we already see the synergistic effects of just two anti-activators on Quorum Sensing, it could be interesting to extend the model by an external source that can regulate the gene expression of anti-activator genes even further and find certain conditions under which Quorum Sensing is never possible.

6 Implementation

The numerical study was implemented using MATLAB version R2021b. The objective function used in the parameter estimation routine was a weighted least squares function. The optimizer used for the optimization was the built-in *fmincon* function. The solver used for the numerical simulation was the built-in *ode15s* solver. The experimental data file as well as the code of the implementation can be found in a repository on GitHub at URL_{GitHub}.

Bibliography

[1] Kyle L Asfahl and Martin Schuster. Additive effects of quorum sensing anti-activators on Pseudomonas aeruginosa virulence traits and transcriptome. Frontiers in Microbiology, 8:2654, 2018.

[2] Jörg Bewersdorff. Algebra für Einsteiger. Springer, 2002.

[3] Teresa De Kievit, Patrick C Seed, Jonathon Nezezon, Luciano Passador, and Barbara H Iglewski. Rsal, a novel repressor of virulence gene expression in Pseudomonas aeruginosa. Journal of Bacteriology, 181(7):2175–2184, 1999.

[4] Jack D Dockery and James P Keener. A mathematical model for quorum sensing in Pseudomonas aeruginosa. Bulletin of Mathematical Biology, 63(1):95–116, 2001.

[5] George S Dulikravich, Thomas J Martin, Brian H Dennis, and Norman F Foster. Multidisciplinary hybrid constrained GA optimization. Chapter, 12:231–260, 1999.

[6] Leah Edelstein-Keshet. Mathematical Models in Biology. SIAM, 2005.

[7] O. Forster. Analysis 2: Differentialrechnung im IRn, gewöhnliche Differentialgleichungen. Grundkurs Mathematik. Springer Fachmedien Wiesbaden, 2013.

[8] Fabian Fröhlich, Carolin Loos, and Jan Hasenauer. Scalable inference of ordinary differential equation models of biochemical processes. Gene Regulatory Networks, 1:385–422, 2019.

[9] Koichi Fujimoto and Satoshi Sawai. A design principle of group-level decision making in cell populations. PLoS Computational Biology, 9(6):e1003110, 2013.

[10] Andrew B Goryachev. Understanding bacterial cell- cell communication with computational modeling. Chemical reviews, 111(1):238–250, 2011.

[11] Andrew B Goryachev, Da-Jun Toh, Keng Boon Wee, Travis Lee, Hai-Bao Zhang, and Lian-Hui Zhang. Transition to quorum sensing in an Agrobacterium population: A stochastic model. PLoS Computational Biology, 1(4):e37, 2005.

[12] John F Griffiths, Anthony JF Griffiths, Susan R Wessler, Richard C Lewontin, William M Gelbart, David T Suzuki, Jeffrey H Miller, et al. An introduction to genetic analysis. Macmillan, 2005.

[13] Jan Hasenauer and Fabian Theis. Parameter estimation for dynamic biological systems. Lecture Script, TUM, Summer Semester 2016.

[14] Douglas Samuel Jones, Michael Plank, and Brian D Sleeman. Differential Equations and Mathematical Biology. Chapman and Hall/CRC, 2009.

[15] Peter Kumberger, Christina Kuttler, Peter Czuppon, and Burkhard A Hense. Multiple regulation mechanisms of bacterial quorum sensing. Biomath, 5(22):1607291, 2016.

[16] Christina Kuttler. Mathematical Models in Biology. Lecture Script, TUM, Winter Semester 2021.

[17] Christina Kuttler and Burkhard A Hense. Interplay of two quorum sensing regulation systems of Vibrio fischeri. Journal of Theoretical Biology, 251(1):167–180, 2008.

[18] Johannes Müller. Advanced Mathematical Biology. Lecture Script, TUM, Summer Semester 2021.

[19] Johannes Müller and Christina Kuttler. Methods and Models in Mathematical Biology. Lecture Notes on Mathematical Modelling in Life Sciences, Springer, Berlin, 2015.

[20] Helmut Plattner and Joachim Hentschel. Zellbiologie. Georg Thieme Verlag, 2006.

[21] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. Numerische Mathematik 2. Springer-Verlag, 2013.

[22] Andreas Raue, Bernhard Steiert, Max Schelker, Clemens Kreutz, Tim Maiwald, Helge Hass, Joep Vanlier, Christian Tönsing, Lorenz Adlung, Raphael Engesser, et al. Data2dynamics: a modeling environment tailored to parameter estimation in dynamical systems. Bioinformatics, 31(21):3558–3560, 2015.

[23] Amy L Schaefer, Dale L Val, Brian L Hanzelka, John E Cronan, and E Peter Greenberg. Generation of cell-to-cell signals in quorum sensing: acyl homoserine lactone synthase activity of a purified Vibrio fischeri luxi protein. Proceedings of the National Academy of Sciences, 93(18):9505–9509, 1996.

[24] Leonard Schmiester, Daniel Weindl, and Jan Hasenauer. Parameterization of mechanistic models from qualitative data using an efficient optimal scaling approach. Journal of Mathematical Biology, 81(2):603–623, 2020.

[25] Rebecca L Scholz and E Peter Greenberg. Positive autoregulation of an acyl-homoserine lactone quorum-sensing circuit synchronizes the population response. MBio, 8(4):e01079–17, 2017.

[26] Philipp Städter, Yannik Schälte, Leonard Schmiester, Jan Hasenauer, and Paul L Stapor. Benchmarking of numerical integration methods for ODE models of biological systems. Scientific reports, 11(1):1–11, 2021.