Quorum Sensing Modelling for *Pseudomonas aeruginosa*

Ehrlich Manuel

July 2, 2022

Abstract

We aim in this study to simulate the cooperative behaviour of bacterium *Pseudomonas aeruginosa*, a particularly dangerous pathogen in hospital settings. This bacterium employs several cooperative behavioral strategies to regulate virulence. Among them is Quorum Sensing, where the bacteria act differently depending on the cell concentration, measured through some signalling molecules. We aim to explain, predict and reproduce Quorum Sensing through our models. In addition, we attempt to replicate the bi-stable behaviour observed in experimental data. The model we have developed explains several dynamics at different scales using Ordinary Differential Equations (ODEs) and stochastic processes. The global dynamics at the colony level are explained through a deterministic model, while a stochastic model predicts the local cell-level dynamics. With experimental data and bifurcation analysis, we chose a parameter set that could show bi-stable behaviour. We analysed the stochastic model by simulating the distribution of the Master equations resulting from the stochastic process. All the simulations were implemented in Python using standard numerical integration techniques. The model has been successful in explaining quorum sensing behaviour in agreement with experimental data.

Contents

1 Introduction

1.1 Bacterium *Pseudomonas aeruginosa*

Pseudomonas aeruginosa is a pathogen bacterium commonly found in soil and water. It is rod-shaped and is 0.5 to 0.8 µm wide and 1.5 to 3.0 µm long [1]; it's usually able to move on its own thanks to a long beating tail called *flagellum*. It is a gram-negative bacterium, which means that its cell boundary is made of a thick polymer with no lipids (fats) coating it; this has experimental implications when marking the bacteria. The bacterium employs several complex group behaviour strategies called "Quorum Sensing", where bacteria release signalling molecules in the environment and measure their concentration to determine their cell density. Once the molecule concentration surpasses a threshold value, it activates a specific behaviour in the colony.

The bacterium is an opportunistic pathogen in humans because it attacks tissues belonging to weak (already sick) individuals. Therefore, *P. aeruginosa* is particularly dangerous in hospital settings, where it can contaminate food or medical instruments. Pneumonia and infections of the bloodstream, the urinary tract, and surgical wound are among the most common illness from *P. aeruginosa*. According to 2019 report by the United States' Centre for Disease Control (CDC) [2] there are annually 32,600 *P. aeruginosa* infections from hospitalised patients and 2,700 estimated deaths in the United States. Usually, antibiotics cure pathogen infections, but many strains of *P. aeruginosa* have become drug-resistant. Therefore current research is searching for alternative therapies to target *P. aeruginosa*, as well as other pathogens.

A possible treatment is to target *P. aeruginosa's* iron absorption on which they depend [3]. Another possibility is to target their capacity to cooperate as a group. In this context, it is essential to understand their communication capacity. The thesis attempts to add understanding by constructing mathematical models that explain, predict and reproduce some of this behaviour.

1.2 Quorum Sensing

Needs for Cooperation. Bacterial cooperation can be found throughout various species and in various forms. *Vibrio Fischeri* is a bacterium that often lives in symbiosis with sea animals. For example, it lives in colonies inside an organ of *Euprymna scolopes*, a small squid from Hawaii [4]. While the squid provides shelter and nutrients, the bacteria produce bio-luminescent molecules at night so that the squid does not cast a shadow against the backdrop of the moonshine. This way, the squid can avert detection by predators. The behaviour is effective because it operates switch-like: the colony goes from (almost) no production of bio-luminescent molecules to the activation in (almost) all bacteria.

P. aeruginosa cooperate differently. The bacteria stay quiescent in the first phase of the infection while its invading population is still in low concentrations. When the population reaches a threshold concentration, it attacks the tissue in a coordinated effort. This way, it postpones a response by the immune system from its most vulnerable phase to one when more significant numbers result in a higher success rate. *P. aeruginosa* also has a different density-dependent cooperative behaviour which regulates the formation of so-called biofilms. Bio-films are a protective membrane created on a pre-existing surface; the bio-film encapsulates the bacteria and defends them from attacks. In addition to protection from predators, the bio-film helps interchange substances such as nutrients within the bacterial community.

All these behaviours also depend on environmental conditions, but most notably on the local bacterial concentration, and they manifest themselves in a switch-like manner. The bacteria are inactive at low concentrations; once they surpass a threshold concentration, they activate the behaviour. For this reason, in 1994, Dr Winans labelled the behaviour: "Quorum Sensing". In addition to these behaviours, Quorum Sensing regulates other behaviours such as swarming. Swarming is a group behaviour also found in animals, especially insects, where individuals that can direct their movement migrate in a specific direction without any central coordination.

The cases described above about bio-luminescence production, virulence regulation, biofilm formation and swarming, regulated by Quorum Sensing, are found in several species of bacteria.

Research interests Understanding these behaviours is of great theoretical and practical interest. By understanding the communication of pathogens attacking body tissue, it would be possible to devise

strategies for disrupting communication, and the disruption would lead to a weaker attack and aid natural body defences in stopping the pathogen. Bacterial behaviour [5] is crucial as bacteria develop antibiotic resistance.

Humans have not yet been able to explain how complex multi-cellular microorganisms have evolved from single microorganisms. Increasing the understanding of group behaviour in prokaryotic bacteria, the simplest of living beings, is a step in that direction. In non-biological settings, it could be helpful to understand how complex collective behaviours emerge from simple units, and it could be applied to sociological contexts or artificial settings such as autonomous machines.

The "switch behaviour" is complicated to replicate, and being able to do so effectively would be of great use for bio-engineering and curing diseases where homeostasis is needed. For example, a bio-chip could measure glucose concentration in the blood and release insulin when it reaches a certain threshold.

1.2.1 Quorum Sensing Gene Regulatory Networks

Gene expression Every cell contains DNA, from highly specialised cells belonging to multi-cellular organisms to simple prokaryotic bacteria. The DNA, short for Deoxyribonucleic acid, is a molecule made by two complementary chains (strands) bound together through hydrogen bonds and wrapped around each other in a helix shape. Monomers called nucleotides constitute each strand. Each strand is a polymer of nucleotides*made of deoxyribose sugar, a phosphate group, and of four different bases: cytosine, guanine, adenine, or thymine. Cytosine pairs only with guanine, and adenine pairs only with thymine. The specific sequence of these chemical bases in the DNA gives the genetic information for cells' development, functioning, growth and reproduction. A gene is a segment of the DNA that codes (provides the information) for the cellular synthesis of a specific protein.

The protein synthesis occurs in two phases: transcription and translation. In the transcription phase, a DNA segment is split along its axis and used as a template for producing a copy called messenger RNA (mRNA). In the translation phase, the mRNA and other molecules in the cell are used to produce the related protein. Both transcription and translation require several sub-steps. Other DNA regions called Promoters regulate the start of the transcription phase and are located near the gene they regulate. Transcription starts when one or more molecules of the correct shape bind to the promoter region. The entire process, starting with the activation of transcription and ending with translation, is called Gene Expression.

Feedback Loops Cells are complex beings, and most gene expression processes coexist, sometimes interfering with each other. The activation of some gene might result in the production of some protein that enhances the activation of some other gene, or conversely, the protein might inhibit the activation of another gene. In some cases, the protein might even, directly or indirectly, enhance or inhibit the activation of the gene responsible for its production in a feedback loop.

QS in *P. aeruginosa* Quorum sensing works by cells producing signalling molecules called "autoinducers". *P. aeruginosa* has four different interconnected gene expression systems [6]: the *las*, *rhl*, *pqs*, *iqs*; these systems produce different signalling molecules that regulate several quorum sensing behaviours. The systems are hierarchical, as the *rhl* and other systems require the activation of the *las* system. The *las* and *rhl* system have a strong influence on virulence regulation [7].

as seen in Figure 2 for a scheme for the regulatory network within the las system. The autoinducer molecule is from the AHL family (acyl-homoserine lactone), and it is called N-[3-oxododecanoyl]-L-homoserine lactone (OdDHL in Figure 2). The autoinducer diffuses in and out of the cell and, within the cell, binds with the LasR protein to form a dimer. This dimer regulates many reactions. It activates the *rhl* system and, directly and indirectly, promotes cell virulence. Furthermore, the dimer is involved with several feedback loops. It activates transcription from the *lasR* gene and therefore synthesis of the LasR proteins, which then is present in higher concentrations to react with the autoinducer and form the production. The dimer activates gene *lasI* and therefore synthesis of the protein LasI. The dimer also activates gene *Rsal* and synthesis of the protein RsaL, which deactivates gene *lasI*. The LasI protein is an enzyme that facilitates the production of the autoinducer from molecules found within the cell.

All molecules can be degraded by chemical decomposition. In addition, the autoinducer molecule can diffuse in and out of the cell membrane; hence, they can be used as a proxy to know the bacterial

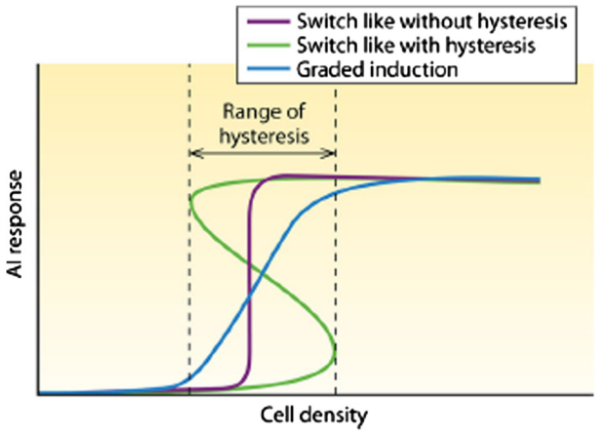

Figure 1: Curve depicting different induction behaviours, taken from Hense and Schuster (2013) [8].

concentration. When diffusivity is low, the autoinducer molecules may stay close or diffuse away from the cell, and they might diffuse into another cell in its vicinity, affecting the production of that cell's synthetase LasI and, consequently, its autoinducer.

At low bacterial concentrations, cells produce a small number of autoinducer molecules. As the bacterial concentration increases, they release more autoinducer molecules into the environment. Therefore the autoinducer concentration within cells increases, and gene expression is activated more often. Hence, the synthetase, and as a consequence, autoinducer molecules too, are produced in greater quantity in a positive feedback loop. At some point, the autoinducer (and bacterial) concentration raises over some threshold value, which sets the autoinducer into a positive feedback loop. The loop pushes the autoinducer concentration to spike up, and, in turn, the bacteria start enacting some specific behaviour, such as activating virulence behaviour.

1.2.2 Hysteresis

We mentioned already that the behaviour activation happens switch-like after reaching a certain threshold of local bacterial concentration. However, there are several ways this can happens. As schematised in Figure 1, some biological systems react gradually (blue curve) to changes in parameters such as cell concentration. Quorum Sensing has two possible switch behaviours (purple and green lines). For the purple-like biological systems, oscillations around the threshold parameter (concentration in this context) result in a constant change ("flickering") between the inactive (OFF) and active (ON) states. This flickering is advantageous for some systems, while for some others, it is undesirable. In order to counter this, some biological systems have evolved to switch between states only after sufficient change in the system has happened (green line). Such systems have an intermediate range of parameters where both the "ON" state and the "OFF" state are acceptable. This behaviour is a trait seen in biological systems other than quorum sensing, such as neuron firing (as discussed, for example, in [9]).

The behaviour results in a system where most bacteria are inactive at low cell density. At high cell density, most bacteria are active; for cell density values in an intermediate range, both the ON state and the OFF state are possible. In that intermediate region, whether the bacteria are active or inactive depends on the bacteria's previous state (or initial state). If they start in the OFF state at low cell density and the cell density increases, the bacteria stay inactive throughout the intermediate region until the concentration reaches the second threshold. At that point, they "jump" to the ON state. Once in the ON state, the bacteria "jump" to the OFF state only if the cell density decreases until the first threshold. The behaviour is called hysteresis and it is mathematically explained.

1.2.3 Bi-stability

Bacterial colonies are not spread homogeneously in a medium; instead, there are lower and higher local cell concentrations. Therefore, given some average cell concentration, some cells will lie in a cluster with a higher cell density, while others will lie in a region with a lower cell density. If quorum sensing happens switch-like with hysteresis, a portion of the population will be entirely in the ON state and a portion of the population in the OFF state. Since this situation is stable for a fixed cell density, this phenomenon is called bi-stability and can be observed experimentally.

1.3 Self-Sensing

Bacteria do not know whether the sensed autoinducer molecules are produced by other close-by bacteria or by themselves. In an environment with low diffusivity, a bacterium primarily reacts to signalling molecules produced by itself and not by other bacteria. In this context, the bacteria, instead of measuring bacterial concentration, measures the speed with which molecules diffuse away. This behaviour is called in many different ways, such as "self-sensing" or "diffusion sensing". In this thesis, we will refer to it as self-sensing. There is no consensus on what came first [10], whether Self-Sensing or Quorum Sensing. It is possible that, at first, a mechanism to measure environment diffusivity evolved, and then the feedback mechanism became a vehicle that allowed for complex behaviours such as quorum sensing to exist. In any case, both behaviours can be observed, and both behaviours can coexist [11].

2 Current Research

In the past twenty years, many authors have modelled quorum sensing behaviours in mathematical terms. In the following section, we present some of these papers that focused on *P. aeruginosa* or bacteria with similar quorum-sensing gene regulatory networks. We present a few papers that model different biological aspects of quorum sensing and employ different mathematical tools with that scope in mind.

Fagerlind et al. (2003) [12] constructed an 8D ODE system modelling the *las* and *rhl* regulatory network in *P. aeruginosa*. They include diffusion of autoinducer molecules and chemical reactions internal to the regulatory network. The authors employ theoretical parameters to reproduce the switch behaviour and investigate for which parameters the model predicts hysteresis.

Dilanji et al. (2012) [13] explicitly include spatial processes via PDEs by having a spatially heterogeneous bacterial and autoinducer concentration.

Fujimoto and Sawai (2013) [14] also model autoinducer diffusion and chemical reactions, but they can do so through a smaller ODE system, and they are also able to reproduce the switch behaviour. They apply a bifurcation analysis to find a range of parameters that show bistability.

Weber & Buceta (2013) [15] study more directly the effect of noise by constructing a stochastic model as well as a deterministic one. The authors also include cell division in the model and simulate the stochastic model via the Gillespie algorithm.

2.1 Fagerlind et al. (2003)

In the 2003 paper by Fagerlind et al. [12] a deterministic model is constructed for the *las* and *rhl* gene regulatory networks in *P. aeruginosa*, both of which help regulate the virulence and biofilm formation of the pathogen.

The model investigates the dynamics of the concentrations of eight molecules: the two autoinducers OdDHL (N-(3-oxododecanoyl)-L-homoserine lactone) and BHL (N-butyryl-L-homoserine lactone), the transcriptional activators for the *las* (LasR) and *rhl* (RhlR) systems, the three autoinducer-transcriptional activator dimers (LasR-OdDHL, RhlR-BHL, RhlR-OdDHL) and, lastly, the Rsal protein which has a repressive effect on the autoinducer synthesiser LasI. The authors ignore intermediate mRNA steps such that it is possible to write the production of some protein as a function of the molecule activating (or inactivating) the gene that produces it. The building blocks for the production of the autoinducer molecules are assumed to be infinite. The model focuses on one bacterium and assumes that the bacterial population does not change in the model time period. The authors set up eight coupled Ordinary Differential Equations (ODEs). Some terms of the equations follow the Law of Mass Action, which states that the rate of a chemical equation is directly proportional to the product of the concentration of the reactants. Some production terms are motivated by Michaelis-Menten kinetics that explain the reactant-product concentration when a reaction's equilibrium depends on some enzyme concentration. In that case, the reaction rate depends on the reactant concentrations and does not increase beyond some saturation value. The autoinducers are assumed to diffuse without impediment through the cell membrane against the concentration gradient. This assumption is a modelling simplification of the process whereby biological pumps move autoinducers across the cellular membrane. Lastly, all molecules are assumed to degrade according to some molecule-specific rate. The authors treat the concentrations of extracellular autoinducers, which replicates a population increase, as a parameter.

The model can reproduce the switch behaviour, where the system switches between an uninduced and induced state after varying the extracellular autoinducer concentration slightly. Furthermore, the system displays hysteresis for some parameter sets. The authors take theoretical parameters rather than from literature. However, as explained by the authors, the analysis is valid nonetheless because the relationship between the parameters determines the system behaviour rather than the values of the parameters themselves.

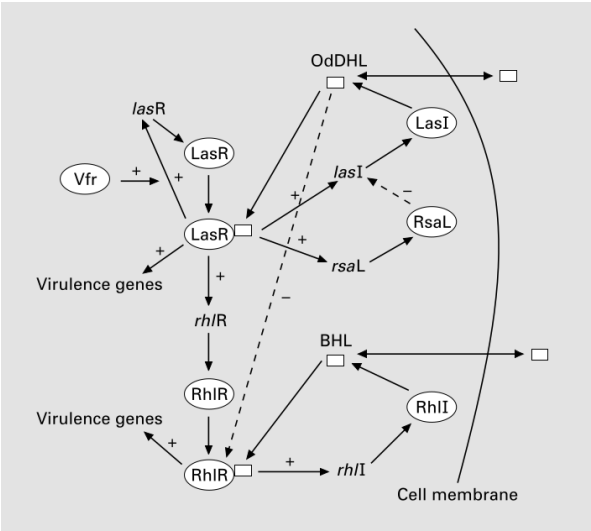

Figure 2: Schematic diagram representing the functioning of the *las* and *rhl* systems involved in the virulence regulation for *P. aeruginosa*, taken from Fagerlind et al. (2003) [12]. The publisher granted reproduction permission.

2.2 Fujimoto and Sawai (2013)

The 2013 paper by Fujimoto and Sawai [14] constructs several deterministic models for the *las* gene regulatory networks of several species of bacteria, including *P. aeruginosa*. The authors study how the OFF to ON state transition originates and what parameters affect the type of transition. They begin with a simple model, including solely one positive feedback loop. As such, they include autoinducer synthetase LasI and the autoinducer of the AHL family as variables only. To avoid constructing a partial differential equation (PDE), they explicitly include the autoinducer diffusion process by separately considering extracellular autoinducer and intracellular autoinducer. They construct an ODE system according to tested mathematical laws. Furthermore, they assume synthetase dynamics to be faster than autoinducer dynamics. The resulting deterministic model [14] after non-dimensionalisation is:

$$\frac{d}{dt}x_i = \left(1-\lambda^{-1}\right)\frac{s_i^2}{s_i^2+k_i^2} - \gamma_x x_i + \eta_i$$
$$s_i = \rho_v \bar{x} + \frac{\gamma_{ex}}{c_{sec}} x_i$$

where x and s are the synthetase and autoinducer concentrations, respectively. The parameters are: λ (minimal to maximal synthetase production ratio), k_i (determines how soon synthetase production reaches its plateau, specific to cell i), γ_x (synthetase degradation rate), η_i (cell i white noise), ρ_v (intracellular to extracellular volume ratio), γ_{ex} (extracellular degradation rate), c_{sec} (easyness of cell membrane crossing). The authors perform a bifurcation analysis and find a transition from no hysteresis to hysteresis for some bifurcation parameters. They also find that single-cells have an all-or-none transition between the OFF and ON state. However, by just including a positive feedback loop, bacteria have a graded transition on the colony level. This changes when including one additional feedback loop into the model.

2.3 Weber & Buceta (2013)

The 2013 paper by Weber and Buceta [15] models the LuxI/LuxR regulatory network found in *Vibrio fischeri*, *P. aeruginosa*, and others. The authors are interested in how noise affects bacterial behaviour. The modelling assumptions were similar to previous studies. In addition to previous studies, the authors include cell growth and division. Since dividing cells split the autoinducer molecules present in the mother cell between themselves, the authors also include a dilution term for the intracellular autoinducer concentration. They express the chemical reactions through a stochastic model, where three processes/events occur. For example, an autoinducer molecule can diffuse from the intracellular

to the extracellular space, and cells can split by reproduction at some cell-specific rate. The authors simulated the stochastic model through the Gillespie algorithm, which generates possible trajectories of a random walk. The Gillespie algorithm works by sampling a waiting time for the next event, choosing which event happens, and updating the variables according to the event and time change. The Gillespie algorithm is used for a sufficient number of simulations, and the results estimate the probability function of the model. The authors reproduced quorum sensing's switch-like behaviour through such a stochastic model. They found parameters that show hysteresis and analysed how different parameters affect system activation.

3 Methodology and Results

3.1 Mathematical Tools

Programming tools All the analysis and simulations within this thesis were performed in python 3.9, using standard python libraries such as NumPy, Sympy, Matplotlib, Pandas and Scipy. Python could not perform some computations due to very big or small values. In those situations, we performed arbitrary rescalings to bring the values to a range where Python would be able to perform calculations.

3.1.1 Ordinary Differential Equations (ODEs)

Ordinary Differential Equations are a class of equations that include ordinary derivatives of the dependent variable as part of the equation. The "ordinary" term distinguishes ODEs from Partial Differential Equations (PDEs), where there may be partial derivatives of one dependent variable with respect to more than one independent variable. The order of the ODE is the order of the highest derivative.

ODEs help understand natural and social phenomena where there is some dynamic change in behaviour. Sometimes, it is not easy to express the relationship between some variable (e.g. time) and the resulting behaviour of the quantity of interest. ODEs are then very useful in scenarios where it is easy to express the rate of change (e.g. in time) of some quantity, which depends on the quantity itself. ODEs are sometimes "simple" to analyse, and an analytic solution can be found by mathematically manipulating the equation and re-expressing the problem as an integration problem. However, most of the time, the only way to solve such equations is through numerical means, especially when more than one interdependent quantity is of interest. In that case, a system of ODEs needs to be solved through numerical methods. This will be elaborated upon later.

An initial value problem (IVP)

$$\begin{cases} \frac{dy(x)}{dx} = f(x, y(x)) \\ y(x_0) = y_0 \end{cases}$$

is made of an ODE and an initial condition. An eventual solution to an IPV is well-defined if it exists and if it is unique. The uniqueness clause is often satisfied thanks to an initial condition or the value at any fixed time value when the change variable is time. However, often it is more interesting to study the collection of ODEs with all the possible initial conditions. Often ODEs arising from real-world problems result in a range of initial conditions leading to the same value. i.e. for time going to infinity, the limit value is the same. However, a different range of initial conditions may lead to a different limit value. This limit value is called a "fixed point". A fixed point is called "asymptotically stable" if values in its vicinity tend to move closer to the fixed point. It is called "asymptotically unstable" if values in its vicinity move away from it. There are also cases where a fixed point is stable in one (or more) direction(s) and unstable in another direction(s). There is an example from population dynamics that is easy to visualise. If a small group of animals of an endangered species are placed in a new area for the first time, the group might die out due to multiple factors (small gene pool, not enough numbers to deter predators or for other successful cooperative behaviours to work, and so on). If the group is big enough, its number of individuals may rise and fall, but over the long term, it would stabilise over a value that the ecosystem can sustain. This behaviour is an example of different outcomes depending on the initial condition, where a range of initial condition values shares the same outcome.

3.1.2 Numerical Integration

Most initial value problems are not solvable analytically. Solving the problem numerically involves discretising the solution space, starting from the initial condition and approximating the following step given a current value and eventually a past value. Given the IVP,

$$\begin{cases} \frac{dy(t)}{dt} = f(t, y(t)) \\ y\,(0) = y_0 \end{cases}$$

the simplest method to solve such a problem is the Euler method, discovered in the 18^{th} century, which defines all steps to depend on the previous one in the following way,

$$y_{n+1} = y_n + \Delta t \cdot f(t, y_n)$$

given a step size Δt and $y_n := y(n \cdot \Delta t)$. Despite not knowing the exact solution, it is possible to have an estimate of the error of the numerical approximation. For the Euler method, the local (per step) error is proportional to the square of the step size: $\mathcal{O}\left(\Delta t^t wo\right)$. The global error is proportional to the step size: $\mathcal{O}\left(\Delta t\right)$. There are several issues affecting numerical methods. First of all, the numerical simulation might be unstable, which means that small, artificial oscillations (those arising from the numerical solver and not from the problem) appear after some number of steps, and they add up and get amplified by several orders of magnitude after a few additional steps. In order to counter this, it is possible to calculate the slope at the current and previous steps and take a weighted average. Methods that employ this technique are called "multistep methods", and by using more than one point, they can dampen the artificial oscillations. A simple two-step method:

$$y_{n+1} = y_n + \frac{f\left(t_n, y_n\right) + f\left(t_{n-1}, y_{n-1}\right)}{2}$$

Second, numerical solutions do not necessarily have a small global error, even if every step has a small local error. In order to reduce the global error, it is possible to reduce the step size and perform more steps, but this could be computationally expensive: there is a trade-off between computational cost and accuracy. It may be appropriate to write the next step implicitly using the old and new steps for some problems. This approach significantly increases the computational cost of one step, but it might be worth it by significantly decreasing the number of steps. In order to decrease the global error, it is also possible to decrease the local error. Mathematicians C. Runge and W. Kutta, at the beginning of the 20^{th}, developed numerical methods where, by taking intermediate steps, it was possible to reduce the local error as well as the global error. The so-called Runge-Kutta 4 (RK4) method given step size Δt and $t_{n+1} = t_n + \Delta t$ is as follows [16]:

$$\begin{aligned}
k_1 &= f\left(t_n, y_n\right), \\
k_2 &= f\left(t_n + \frac{\Delta t}{2}, y_n + \Delta t \cdot \frac{k_1}{2}\right), \\
k_3 &= f\left(t_n + \frac{\Delta t}{2}, y_n + \Delta t \cdot \frac{k_2}{2}\right), \\
k_4 &= f\left(t_n + \Delta t, y_n + \Delta t \cdot k_3\right),
\end{aligned}$$

$$y_{n+1} = y_n + \frac{1}{6}\Delta t \cdot \left(k_1 + k_2 + k_3 + k_4\right)$$

This method has a local error of $\mathcal{O}\left(\Delta t^5\right)$ and a global error of $\mathcal{O}\left(\Delta t^4\right)$, therefore allowing for a bigger step size and therefore a lower computational cost. The RK4 method, however, has some constraints regarding the differentiability of the right-hand side of the IVP. More on the theory of numerical methods can be found at [16].

3.1.3 Slow/Fast Systems

Many physical phenomena have various processes that interact with each other, but occur at different speeds. There is a way in the context of ODEs to model how one process occurs way faster than another. Let a 2D ODE system be:

$$\begin{aligned}
\frac{dx}{d\tau} &= f(x, y), \\
\frac{dy}{d\tau} &= \epsilon g(x, y),
\end{aligned} \tag{1}$$

where $0 < \epsilon \ll 1$ is very small. The dynamics for y are very slow as the y ODE has very small values. Hence, y is the slow variable and x is the fast variable. System 1 is the so-called fast system, but by re-scaling time $\tau = \epsilon t$ to time t we get the so-called slow system:

$$\begin{aligned}
\epsilon \frac{dx}{dt} &= f(x, y), \\
\frac{dy}{dt} &= g(x, y),
\end{aligned} \tag{2}$$

Variable t is the slow variable while variable τ is the fast variable. By taking the limit $\epsilon \to 0$ the fast 1 and slow 2 systems, respectively, are transformed to:

$$\begin{aligned} \frac{dx}{dt} &= f(x,y), \\ \frac{dy}{dt} &= 0, \end{aligned} \tag{3}$$

and

$$\begin{aligned} 0 &= f(x,y), \\ \frac{dy}{d\tau} &= g(x,y), \end{aligned} \tag{4}$$

The slow-system 4 is addressed by solving $f(x,y) = 0$ to find an expression $x = h(y)$ s.t. $f(h(y), y) = 0$. The fast system 3 gives constant slow variable y from $y' = 0$. Therefore only fast variable changes and the problem is reduced to a one-dimensional ODE: $x' = f_y(x)$, given constant y. The solution to the 1D ODE from the slow system may be a sufficient simplification of the system. However, sometimes there are jumps where solutions temporarily leave the slow manifold and have some fast dynamics until they eventually reach another slow manifold.

An example of a slow/fast system is the hysteresis shown in Figure 1. If cell density is treated as a variable, it would be the slow variable, while the state would be the fast variable. As cell density slightly increases from small values, y adjusts much faster to follow the green curve. Once the concentration reaches the threshold, the fast system is the only visible one as y increases very fast until it reaches the other branch of the green curve seen in Figure 1. More on Slow/Fast systems at [17].

3.1.4 Bifurcation Theory

It is also possible to analyse a broader class of ODEs by having variable parameters and initial conditions. We consider ODE $f_r(x)$ where function f depends on a parameter $r \in \mathbb{R}$ in addition to variable $x \in \mathbb{R}^n$. This branch of mathematics is called bifurcation theory. Mathematically speaking, bifurcation analysis studies the change in the topology of the solution space given a change in parameters. In higher dimensions, such analyses are harder to perform. There are three types of bifurcations in one dimension: Saddle-node bifurcation, Transcritical bifurcation and Pitch-fork bifurcation. Here, we introduce some examples of such bifurcations where only one bifurcation parameter is involved.

A Saddle-Node bifurcation occurs when there is an appearance/disappearance of two fixed points, one stable and one unstable. We denote fixed points as x^*. The 1D ODE

$$\dot{x} = r - x^2 := f_r(x)$$

has fixed points if $r - x^2 = 0$. If there are fixed points we can find their stability by calculating the derivative of $f_r(x)$. If $r > 0$, there are two fixed points: $x^* = \sqrt{r}$ (stable) and $x^* = -\sqrt{r}$ (unstable). If $r = 0$ there is only one fixed point: $x^* = 0$. If $r < 0$ there are no real fixed points.

A Transcritical bifurcation occurs when one stable fixed point and one unstable fixed point swap stability. Given 1D ODE

$$\dot{x} = x \cdot (r - x) =: f_r(x),$$

if $r > 0$, there are two fixed points: $x^* = 0$ (unstable) and $x^* = r$ (stable). If $r = 0$ there is only one fixed point: $x^* = 0$. If $r < 0$ there are again two fixed points: $x^* = 0$ (stable) and $x^* = r$ (unstable). So fixed points $x^* = 0$ and r exchange stability at $r = 0$.

A Pitch-fork bifurcation happens when a Saddle-node and Transcritical bifurcation coincide: a stable/unstable fixed point turns into unstable/stable, and two stable/unstable fixed points appear, respectively. Or vice-versa. Given 1D ODE

$$\dot{x} = x \cdot \left(r - x^2\right) =: f_r(x),$$

if $r > 0$, there are three fixed points: $x^* = \sqrt{r}$ (stable), $x^* = 0$ (unstable), and $x^* = -\sqrt{r}$ (stable). If $r = 0$ there is only one fixed point: $x^* = 0$. If $r < 0$ there is only one fixed point and it is stable: $x^* = 0$. Two fixed points disappear and $x^* = 0$ exchanges stability at $r = 0$.

In the context of the switch transition between the on and off state, there could or could not be hysteresis; see section 1.2.2. Without hysteresis, there is only one fixed point representing the activation state. Depending on the bacterial population concentration value, either the fixed point value is low, representing an "ON" state, or high, representing an "OFF" state. Both are not possible. Alternatively, if there is hysteresis, there is a range of concentration values where both the "OFF" state and the "ON" are feasible. In this situation, there are two stable fixed points (as well as a "hidden" unstable one). The potential transition from no hysteresis to hysteresis, which depends on some model parameters, leads to a Saddle-Node bifurcation. In Figure 1 there are two saddle-node bifurcations in correspondence of the two dotted vertical lines: in the left one, two fixed points appear, and in the right one, two fixed points disappear. More on bifurcation theory at [18].

3.1.5 Parameter Fitting

Having a functioning model is not enough to explain this phenomenon. As seen in section 3.1.4, parameters significantly influence the outcome of the model simulation, so they need to reflect reality. Parameters [19] are chosen to minimise the error, according to some predefined metric, between the model simulation with the chosen parameters and experimental data. This paper uses least squares as an error metric: the chosen parameres minimise the residual sum of squares between the model simulation and the real data. Each step of the optimisation process requires a new model simulation. Therefore the whole process is of high computational complexity, especially when it involves many parameter values. Additionally, the complexity rises quickly when the number of parameters increases, and there is a guarantee to find local minima but not global minima in the optimisation process. For this reason, there could be a different set of significantly different parameters that give an equivalent or lower least square error.

3.1.6 Stochastic Process

A stochastic process [20] is a collection of random variables $X := \{X_t : t \in T\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ that describes the evolution in time of some random phenomena. T is the index set, and it represents time. Stochastic processes are useful to model phenomena when variability plays a big part in predicting the outcomes, as the model includes uncertainty. An equivalent deterministic model outputs the expected value of the stochastic model result.

A subset of stochastic processes is Markov processes, where the following value of the process only depends on the current value and not on any past value. This way, the process is much simpler to analyse, although there is no closed-form solution most of the time.

3.2 Data

Dr Martin Schuster and his team from the department of microbiology at the Oregon State University provided the experimental data. Several experiments were repeated with bacterium *P. aeruginosa* and mutants of the bacterium. A colony of bacteria is grown in 1ml of solution. The density of bacterial cells in solution was measured by optical density at a wavelength of 600nm (OD_{600}). At the same time, the synthetase molecules were linked to a Green Fluorescent Protein (GFP) tag to know its evolution over time. The bacteria start at a low concentration; consequently, the activation level is low. As cell concentration increases, synthetase molecules (representing activation) rapidly increase around a specific cell concentration value. Bi-stable behaviour is evident in some mutants. As a first step in developing the mathematical model, we chose one mutant where the anti-activator proteins qslA and qteE were absent. More information on the experimental data can be found at [21].

Note that OD_{600} and GFP are used as a proxy of cell density and synthetase density, respectively. However, the exact values of cell and synthetase density are unknown. According to Kim et al [22] 1.0 OD_{600} corresponds to $2.04 \cdot 10^8$ CFU/ml (Colony Forming Unit) and 2.085 mg/ml with well defined linear relationships of $Y(\text{CFU/ml}) = 2.0 \cdot 10^8 \cdot [\text{OD}_{600}] + 4.0 \cdot 10^6$ and $Y\ (\text{mg/ml}) = 2.0087 \cdot [\text{OD}_{600}] + 0.0764$. These relationships produce artifacts for small cell concentrations, so we relaxed the relationship to: $Y(\text{CFU/ml}) = 2.0 \cdot 10^8 \cdot [\text{OD}_{600}]$ and $Y\ (\text{mg/ml}) = 2.0 \cdot [\text{OD}_{600}]$. We use the GFP as a proxy value for the synthetase concentration. This doesn't distort our work because the relationship is approximately linear, and we could obtain the synthetase concentration value from a rescaling of the

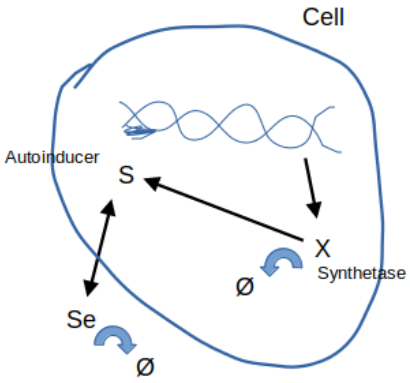

Figure 3: Simplified diagram of the main processes of the regulatory network showing the dynamics of the autoinducer and the synthetase within the *las* system.

model, if needed. With the GFP value we are still able to obtain some qualitative and quantitative dynamics of the molecules of interest.

Lastly, we are more interested in the synthetase concentration within cells than the whole experimental space (synthetase molecules stay within cells). For this reason, the GFP value was divided by the OD_{600} value to account for population growth.

3.3 First Model: Global Perspective, Colony Level

The aim is to construct a mathematical model that explains quorum sensing and replicates bi-stable behaviour. We focus exclusively on the *las* system and aim to include noise in the model; as such, we take the least amount of terms needed to explain quorum sensing. Fujimoto [14] showed that it is possible to model the positive feedback loop in the *las* system by merely considering the synthetase, the intra-cellular autoinducer and the extracellular autoinducer.

In the beginning, we are interested in looking at the colony level rather than the cell level but the colony level. Therefore the autoinducer and synthetase levels are taken to be average values over the whole colony rather than cell-by-cell levels. Figure 3 shows the processes that were included. Dealing with averaged values means ignoring local stochastic effects. Therefore a deterministic model is appropriate for this setting. In the simplified scheme, there is a time dynamic and a spatial dynamic. To avoid overcomplicating things, we chose not to have a PDE model.

We model the diffusion process by having two different variables representing the extracellular autoinducer concentration and intra-cellular autoinducer concentration. For simplification, both the insides of the cells and the extracellular medium are considered homogeneous. We consider variables s_e, s and x representing extracellular autoinducer concentration, mean intracellular autoinducer concentration and mean synthetase concentration. Their values are expressed in nMol / mL. The model is then the following:

$$\begin{aligned}
\frac{d}{dt}s_e &= \rho_v c_{\text{sec}}\left(s - s_e\right) - \gamma_{ex} s_e \\
\frac{d}{dt}s &= c_{\text{sec}}\left(s_e - s\right) + A_{syn} x \\
\frac{d}{dt}x &= V_{\text{min}} + \left(V_{\text{max}} - V_{\text{min}}\right)\frac{s^n}{s^n + K^n} - \gamma_x x
\end{aligned}$$

Parameter ρ_v is the ratio of intracellular volume over extracellular volume. c_{sec} represents the ease with which autoinducer molecules pass through the cell membrane. γ_{ex} is the autoinducer degradation rate. A_syn is the autoinducer synthesis rate from synthetase. V_{min} is the synthetase basal production rate. V_{max} is the synthetase maximum production rate. K tunes the production of synthetase given autoinducer. γ_x is the synthetase degradation rate.

The experiment is undergone in 1ml of watery solution. The units of measurement of the various variables and parameters are:

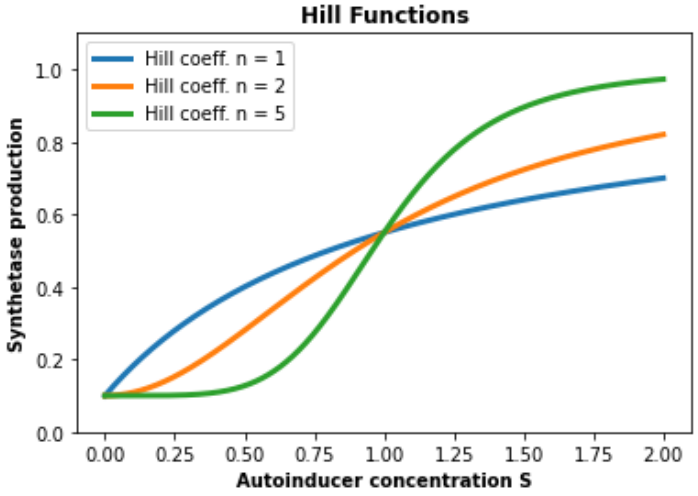

Figure 4: Production functions with several Hill coefficients. $V_{min} = 0.1$, $V_{max} = 1$, $K = 1$. The figure shows how increasing the Hill coefficient affects the induction process.

- $[s_i]$, $[s_e]$, $[x_i]$, $[K]$ = mMol / mL
- $[\rho_v] = \emptyset$
- $[\gamma_{ex}]$, $[\gamma_X]$, $[c_{sec}]$, $[A_{syn}]$ = 1/s
- $[V_{max}]$, $[V_{min}]$ = mMol / mL $\cdot$ s

We derive the degradation terms for the synthetase $(-\gamma_x x)$ and the autoinducer $(-\gamma_{ex} s_e)$, as well as the autoinducer synthesis rate $(+A_{syn} x)$ from the Law of Mass Action that governs the rate of a reaction given the concentration of the reagents when at a molecular level the reactions occur through stochastic processes. The degradation rate for the intracellular autoinducer is negligible and hence is not considered. We model the diffusion term $(c_{\text{sec}}\,(s - s_e))$ by assuming that the diffusion is proportional to the change in concentration between the two sides of the cell boundary. The Ordinary Differential Equation (ODE) for the extracellular autoinducer has an extra parameter ρ_v that accounts for the fact that the autoinducer molecules entering the extracellular medium come not from one single cell but a number of them. ρ_v is defined as the ratio of intracellular volume over extracellular volume. Since cells have (on average) fixed volume but are growing in number, ρ_v changes with the population concentration, and we treat it separately as a dynamic parameter. The last term of the ODE system is the synthetase production term: $V_{\text{min}} + (V_{\text{max}} - V_{\text{min}}) \frac{s^n}{s^n + K^n} := Prod(S)$. When the promoter region is empty, it results in the synthetase production at basal rate V_{min}. If the autoinducer is bound to the promoter region, on the other hand, it results in synthetase production at a higher rate V_{max}. The production of synthetase then depends on the frequency with which the promoter region is bound, which depends on the concentration of autoinducer molecules, as the binding is a random event. Literature models such phenomena with a Hill function. Figure 4 depicts how the Hill coefficient changes the Production function. For high Hill coefficients the switch takes time to happen, but then occurs fast. It is known from the literature that Hill Coefficients bigger than one lead to more than one fixed point, while a coefficient equal to one leads to only one fixed solution. Bi-stability requires two stable fixed points. Therefore we chose the Hill coefficient to be two. Choosing bigger values would have quantitatively affected the model, but not qualitatively. Furthermore, it is a parameter difficult to estimate from data.

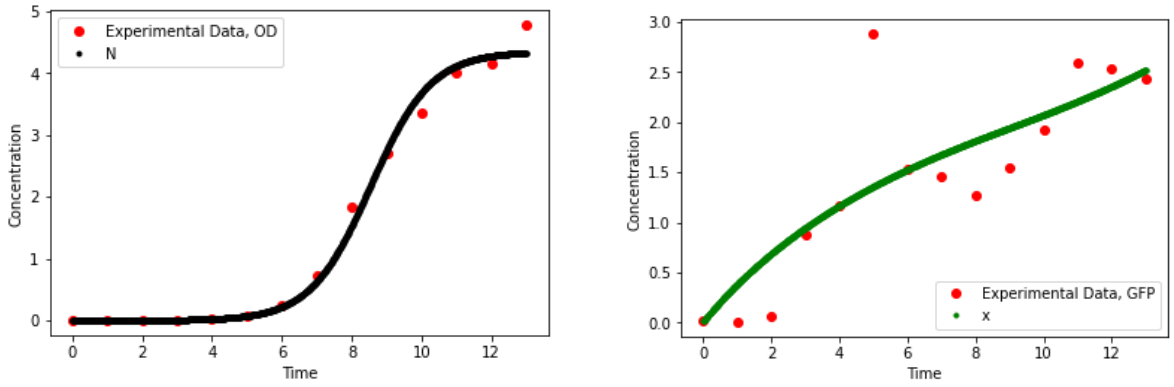

Figure 5: Left – Bacterial population dynamics. Right – Synthetase concentration dynamics. The parameters used are found through least-squares fitting: $c_{sec} = 20.7$, $\gamma_{ex} = 0.787$, $A_{syn} = 0.973$, $V_{min} = 0.379$, $V_{max} = 1.99$, $K = 0.450$, $\gamma_x = 0.161$, $r = 1.16$, $k_{carr} = 4.36$. The plots show the dynamics for the real (dotted) and simulated data (full line).

Lastly, to estimate ρ_v, the cell population number is needed. Therefore, the system must turn from a 3D ODE system to a 4D one. The added ODE for the population dynamics is a simple logistic equation. The final model is then:

$$\begin{aligned}
\frac{d}{dt}s_e &= \rho_v c_{\text{sec}}\left(s - s_e\right) - \gamma_{ex} s_e \\
\frac{d}{dt}s &= c_{\text{sec}}\left(s_e - s\right) + A_{syn} x \\
\frac{d}{dt}x &= V_{\text{min}} + \left(V_{\text{max}} - V_{\text{min}}\right)\frac{s^2}{s^2 + K^2} - \gamma_x x \\
N' &= rN\left(1 - \frac{N}{k_{carr}}\right) \\
\rho_v &= \frac{N V_{cell}}{V_{tot}}
\end{aligned}$$

Simulation We run the simulation using the Python function scipy.integrate.odeint, which generally solves systems of first-order ODEs. The results are shown in Figure 5.

We do not know the experimental values of cell number and synthetase concentration for either figure, but only OD_{600} and GFP, which we use as proxies. As Figure 5 shows, the simulations well follow the experimental data.

3.4 Second Model: Local Perspective

The next step involves modelling quorum sensing at the cell scale. In order to do so, we cannot ignore stochastic events as in the previous model. Furthermore, to account for cell variability (which is necessary for quorum sensing), it is not possible to look at the average concentration value of autoinducer and synthetase over all cells; instead, it is required to look at the cell-by-cell case. One way to include randomness is by adding noise to the system, e.g. Gaussian Noise as in the following ODE system as done by Fujimoto et al. [14]:

$$\begin{aligned}
\frac{d}{dt}s_e &= \rho_v c_{\text{sec}}\left(\bar{s} - s_e\right) - \gamma_{ex} s_e \\
\frac{d}{dt}s_i &= c_{\text{sec}}\left(s_e - s_i\right) + A_{syn} x_i \\
\frac{d}{dt}x_i &= V_{\text{min}} + \left(V_{\text{max}} - V_{\text{min}}\right)\frac{s_i^2}{s_i^2 + K^2} - \gamma_x x_i + \eta_x
\end{aligned}$$

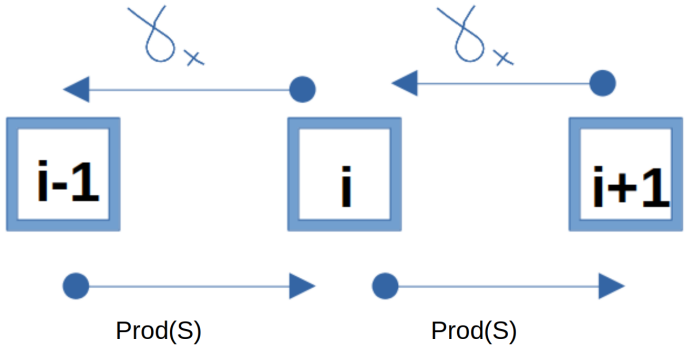

Figure 6: A Markov process modelling the synthetase dynamics. The state value represents the total number of synthetase molecules. A state change depends only on the current state, and the degradation and production rates determine the waiting time for a state jump.

where the parameters are defined as before, η_x is some Gaussian noise that needs to be re-scaled with population size, $\bar{s}$ and $\bar{x}$ are the averaged-out intra-cellular autoinducer and synthetase over all cells, respectively, s_e is the extracellular autoinducer and s_i and x_i are the intracellular autoinducer and synthetase concentration for cell i. A formal definition for $\bar{s}$ and $\bar{x}$ is $\bar{s} := \sum_{i=1}^{N_{cells}} s_i$ and $\bar{x} := \sum_{i=1}^{N_{cells}} x_i$ A simulation over billions of cells is impractical, but having a small number of "cells" is sufficient to account for heterogeneity.

An attempt to construct a more complex model involves writing a stochastic process for synthetase molecules rather than an ODE. While the deterministic model had continuous dynamics, a stochastic process is discreet. For this reason, the (discreet) absolute amount of synthetase molecules X_t^i is needed rather than the (continuous) concentration value $x_i(t)$. There are two dynamics: production and degradation. Then X_t follows the following branching process:

$$\begin{aligned} X_t \to X_t + 1 \quad &\text{at rate} \quad \zeta\left(V_{\min} + (V_{\max} - V_{\min})\frac{s_j^2}{s_j^2 + K^2}\right) \\ X_t \to X_t - 1 \quad &\text{at rate} \quad \gamma_x X_t \end{aligned}$$

where parameters V_{min}, V_{max}, K, γ_x and variable s_i are as previously defined. X_t is the number of molecules rather than the concentration of molecules inside cell j. For this reason, we need a conversion factor ζ that converts a concentration value into an absolute number. However, since we use GFP as a proxy for synthetase concentration, it is not fully apt to view ζ as a way to convert a concentration value to an absolute number. Instead, ζ represents a parameter controlling the importance of random effects for the system. Random effects are dominant for small values of ζ, while for big values of ζ random effects are negligible, and a deterministic model sufficiently describes the system.

A Markov process for synthetase dynamics can then be expressed as the scheme in Figure 6. Given state j (given population is j), the rate of loss of one synthetase molecule to degradation during time Δt is $\gamma_x \cdot j \cdot \Delta t + \mathcal{O}(\Delta t^2)$, while the synthetase molecule increase rate is $\text{Prod}(j)\,\Delta t + \mathcal{O}(\Delta t^2)$, where $\text{Prod}(j)$ is given by:

$$\text{Prod}(j) = \zeta\left(V_{\min} + (V_{\max} - V_{\min})\frac{s_j^2}{s_j^2 + K^2}\right) \tag{5}$$

The possible events are:

$$\begin{aligned} P(X_{t+\Delta t} = i+1 | X_t = i) &= \Delta t \cdot \text{Prod}(i) + \mathcal{O}(\Delta t^2) \\ P(X_{t+\Delta t} = i-1 | X_t = i) &= \Delta t \cdot \gamma_x i + \mathcal{O}(\Delta t^2) \\ P(X_{t+\Delta t} = i | X_t = i) &= 1 - \Delta t \cdot (\text{Prod}(i) + \gamma_x i) + \mathcal{O}(\Delta t^2) \end{aligned} \tag{6}$$

One way to simulate this stochastic process would be through Monte-Carlo simulations, which repeat pseudo-random simulations a sufficient number of times and use the results to estimate a probability function.

Another approach is to find the expression of the so-called Master equations, ODEs that express the dynamics of the probability to be in some state, for any state.

Let $p_i(t) := P(X_t = i)$. Then an expression for $p_i(t + \Delta t)$ (from the law of total probability) is:

$$p_i(t+\Delta t) = p_i(t) \cdot (1 - \Delta t \cdot (\text{Prod}(i) + \gamma_x i)) + p_{i+1}(t)\,\Delta t \cdot \gamma_x (i+1) + p_{i-1}(t)\,\Delta t \cdot \text{Prod}(i-1) + O(\Delta t^2)$$

We can re-express this equation and take the limit to get the expression of the master equation:

$$\lim_{\Delta t \to 0} \frac{p_i(t+\Delta t) - p_i(t)}{\Delta t} =$$

$$\dot{p}_i(t) = -(\text{Prod}(i) + \gamma_x i)\,p_i(t) + \gamma_x (i+1)\,p_{i+1}(t) + \text{Prod}(i-1)\,p_{i-1}(t) \tag{7}$$

$$\text{Prod}(j) = \zeta\left(V_{\min} + (V_{\max} - V_{\min})\frac{s_j^2}{s_j^2 + K^2}\right) \tag{8}$$

We set $\text{Prod}(i-1) := 0$ so that $\dot{p}_0$ is defined.

We now have an infinitely dimensional system of ODEs that we can simulate. There is now an infinite amount of ODEs to simulate, but we can overcome this issue, as later explained.

3.4.1 Slow-fast System

The master equations to simulate are in terms of intracellular autoinducer s_j. To get their expressions, we could construct an additional stochastic process for the intracellular and extracellular autoinducer. However, a three-dimensional stochastic model is quite hard to analyse. Keeping the two ODEs for intracellular and extracellular autoinducer is also not necessary. It is possible to assume autoinducer dynamics (production, degradation and diffusion) are significantly faster than the synthetase dynamics, as executed in [14]. Hence, they reach equilibrium much faster than the synthetase. The system can be assumed to be a slow-fast system.

$$\begin{aligned}
\epsilon\frac{d}{dt}s_e &= \rho_v c_{\text{sec}}(\bar{s} - s_e) - \gamma_{ex} s_e \\
\epsilon\frac{d}{dt}s_i &= c_{\text{sec}}(s_e - s_i) + A_{syn} x_i \\
\dot{p}_i(t) &= -(\text{Prod}(i) + \gamma_x i)\,p_i(t) + \gamma_x(i+1)\,p_{i+1}(t) + \text{Prod}(i-1)\,p_{i-1}(t) \\
\text{Prod}(j) &= \zeta\left(V_{\min} + (V_{\max} - V_{\min})\frac{s_j^2}{s_j^2 + K^2}\right)
\end{aligned}$$

for slow time t. We may obtain the fast system by rescaling to the fast time: $\tau = t/\epsilon$, with $\epsilon \ll 1$. However, we focus on the autoinducer ODEs in the slow system:

$$\begin{aligned}
\epsilon\frac{d}{dt}s_e &= \rho_v c_{\text{sec}}(\bar{s} - s_e) - \gamma_{ex} s_e \\
\epsilon\frac{d}{dt}s_i &= c_{\text{sec}}(s_e - s_i) + A_{syn} x_i
\end{aligned}$$

For $\epsilon \ll 1$, we re-express these equations as:

$$\begin{aligned}
s_e(\rho_v c_{\text{sec}} + \gamma_{ex}) &= \rho_v c_{\text{sec}} \bar{s} \\
c_{\text{sec}} s_i &= c_{\text{sec}} s_e + A_{syn} x_i
\end{aligned}$$

Averaging out both sides of the second equation gives:

$$c_{\text{sec}}\bar{s} = c_{\text{sec}} s_e + A_{syn}\bar{x}$$

Re-inputting the value in the first equation results in:

$$s_e = \frac{\rho_v A_{syn}}{\gamma_{ex}}\bar{x}$$

We can re-input this expression in the original equation 2 to have an expression solely dependent on x.

$$s_i = \frac{\rho_v A_{syn}}{\gamma_{ex}} \bar{x} + \frac{A_{syn}}{c_{sec}} x_i$$

Note that the stochastic process tracks the dynamics of X_i, the number of synthetase molecules in cell i, and not x_i, the concentration of synthetase molecules in cell i. Therefore a more helpful expression is:

$$s_i = \frac{\rho_v A_{syn}}{\zeta \gamma_{ex}} \bar{X} + \frac{A_{syn}}{\zeta c_{sec}} X_i \tag{9}$$

3.4.2 Complete Model

Simulation Preparation With the Markov Process and the approximation to the Master Equations, the values to be found are the probability functions $p_i(t)$'s: the probability of there being i synthetase molecules in some cell. The states i's are merely variables. With this line of thought, X_i is simply i. On the other hand, $\bar{X}$ is not computable. Therefore $\bar{X}$ needs to be approximated by:

$$\bar{X}(t) = \sum_{j=0}^{\infty} j \cdot p_j(t)$$

Infinitely to Finitely Dimensional Model As mentioned before, there are infinitely many master equations, as the possible states go from 0 to ∞. This system of ODEs is not possible to simulate. However, for high states, the degradation term dominates the production term. This is to be expected; otherwise, it would imply that the synthetase molecules grow unbounded. A simple way to estimate at which point the degradation term dominates the production term is to look at the Markov process as a flow between states. Then, we choose state $M > 0$, and for $t \gg 1$ the flow from state $M-1$ to M is $p_{M-1}(t) \cdot \text{Prod}(M-1)$. An upper bound of this expression is $p_{M-1}(t) \cdot V_{max}$. The flow from state M to $M-1$ is $\gamma_x M p_M(t)$. For $t \gg 1$ the flows are in equilibrium, and $p_{M-1} > p_M$ if:

$$M > \frac{\zeta \cdot V_{max}}{\gamma_x}$$

Therefore by taking

$$M = 1.5 \cdot \frac{\zeta \cdot V_{max}}{\gamma_x}$$

we make sure that states above that value are big enough that to reach them, there would be the need for an unlikely "flow against the current". We cut all master equations for states bigger than M so that we have a $(M+1)$ - dimensional system by retaining the master equations for states 0 to M. We re-express X_{mean} as

$$\bar{X}(t) = \sum_{j=0}^{M} j \cdot p_j(t) \tag{10}$$

Initial Conditions The system needs an initial condition to be well-defined. To avoid artificial simulation artefacts, we set the model to start with 0 synthetase molecules in the cells. Mathematics-wise this means $p_0(0) = 1$, and $p_i(0) = 0$ for lal other i's.

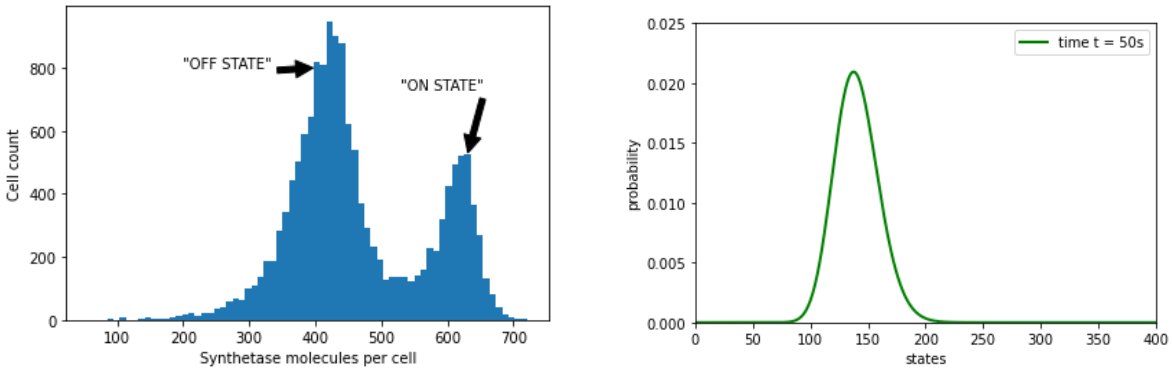

Figure 7: Left – A histogram counting the number of bacteria per synthetase molecule count. The data is taken six hours from the beginning of the experiment for the *P. aeruginosa* strain missing the *qslA* and *qteE* genes; see section 3.2. Right – We simulated the stochastic model using the fitted parameters from experimental data: $c_{sec} = 20.7$, $\gamma_{ex} = 0.787$, $A_{syn} = 0.973$, $V_{min} = 0.379$, $V_{max} = 1.99$, $K = 0.450$, $\gamma_x = 0.161$, $r = 1.16$, $k_{carr} = 4.36$. The graph represents the probability curve for cell i to have several synthetase molecules corresponding to the state value. For implementation purposes, we performed a variable rescaling. It is not appropriate to perform a quantitative comparison between experimental and modelling results, so the time scales and the synthetase molecule numbers are different. However, it is possible to compare the two qualitatively, but our first simulation (right) does not show the bistable behaviour seen in the experimental data (left).

Complete Stochastic Model Putting equations 3.4, 9 and 10 together, the model we need to simulate is:

$$\begin{aligned}
\dot{p}_i(t) &= -\left(\mathrm{Prod}(i) + \gamma_x i\right) p_i(t) + \gamma_x (i+1)\, p_{i+1}(t) + \mathrm{Prod}(i-1)\, p_{i-1}(t) \\
\mathrm{Prod}(i) &= \zeta \left(V_{\min} + (V_{\max} - V_{\min}) \frac{s_i^2(t)}{s_i^2(t) + K^2} \right) \\
s_i(t) &= \frac{\rho_v A_{syn}}{\zeta \cdot \gamma_{ex}} \bar{X}(t) + \frac{A_{syn}}{\zeta \cdot c_{sec}} i \\
\bar{X}(t) &= \sum_{j=0}^{M} j \cdot p_j(t) \\
p_i(0) &= \begin{cases} 1 & \text{if } i = 0 \\ 0 & \text{if } i \neq 0 \end{cases}
\end{aligned}$$

3.4.3 Preliminary Results: no Bistability.

We simulated the model using a two-step method, where single steps follow the standard fourth-order Runge Kutta method. A first simulation of the model 7, using parameters estimated through fitting to the data (L_2 regression), does not show any bi-stable behaviour.

This failure could result from many inaccuracies: from experimental error to model oversimplification, unsatisfactory parameter estimation, and inaccurate simulations. We perform a bifurcation analysis to understand why this is the case and whether our model can predict bi-stable behaviour. We hope to find a set of parameters that can reproduce the bi-stable behaviour shown in the data. As literature shows [23], the deterministic model 3.3 is a limit case of the stochastic model. Hence, we go back to the deterministic model, which is simpler to analyse and see whether the model can result in bi-stable behaviour.

3.5 Bifurcation Analysis.

To make the analysis more manageable, we take one further assumption (as in Fujimoto et al. [14]): we choose s_e to be a parameter (hence fixed). The aim is to keep all parameters except K fixed.

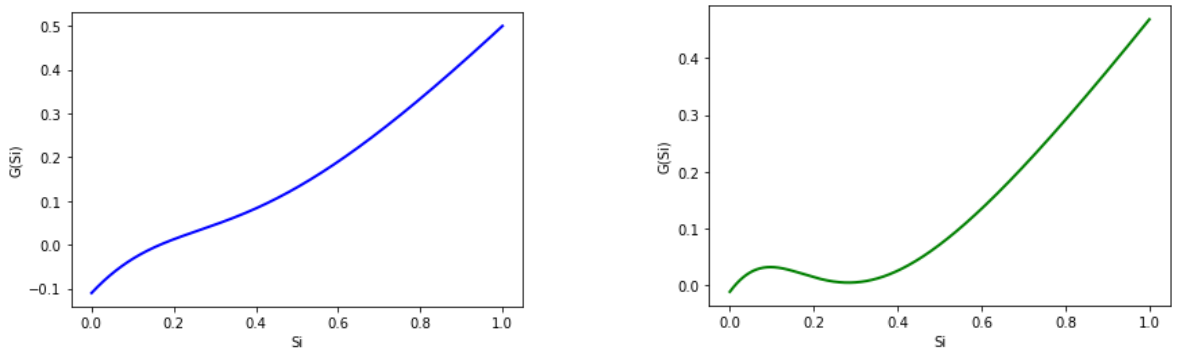

Figure 8: Left – An example of a $G(s_i)$ function that is monotonically increasing, leading to one steady state only. Right – An example of a $G(s_i)$ function that has a local maximum and minimum, leading to two stable steady states.

The first step is to determine the stationary point for average intra-cellular autoinducer s and average (intra-cellular) synthetase x:

$$0 = c_{\text{sec}}\left(s_e - s\right) + A_{syn}x$$
$$0 = V_{\min} + \left(V_{\max} - V_{\min}\right)\frac{s^2}{s^2 + K^2} - \gamma_x x$$

Putting them together gives the condition for the stationary point:

$$s_e = s - \frac{A_{syn}}{c_{sec}\gamma_x}\left(V_{\min} + \left(V_{\max} - V_{\min}\right)\frac{s^2}{s^2 + K^2}\right) =: G(s)$$

Mathematically, mono-stability occurs if the horizontal line $c_{sec}s_e$ intersect the right hand side function $G\left(s\right)$ only and exactly once. We know it intersects at least once because $G\left(s\right)$ is a continuous function and

$$G\left(0\right) = 0$$
$$\lim_{s\to\infty} G\left(s\right) = \infty$$

Further analysis shows that this fixed point would be a stable fixed point. As seen in 8, $G\left(s\right)$ can have one local minimum, or none, depending on the chosen parameter set. If a local minimum exists, then for a range of values of s_e, three intersections of s_e with $G\left(s\right)$ exist, which result (after further analysis) in three fixed points, one unstable and two stable ones. The problem then changes into finding if and with which parameter values a local minimum for $G\left(s\right)$ for $s \geq 0$ exists.

Derivations A necessary condition for local minima/maxima (they go hand-in-hand) is that:

$$\frac{d}{ds}G(s) = 0 \tag{11}$$
$$1 - \frac{A_{syn}}{\gamma_x c_{sec}}\left(V_{\max} - V_{\min}\right)\frac{2K^2 s}{\left(s^2 + K^2\right)^2} = 0 \tag{12}$$
$$\frac{K^2 s}{\left(s^2 + K^2\right)^2} = \frac{\gamma_x c_{sec}}{2A_{syn}\left(V_{\max} - V_{\min}\right)} \tag{13}$$

Define:

$$F\left(s\right) := \frac{K^2 s}{\left(s^2 + K^2\right)^2}$$

The problem changes once more: bistability is possible if the original function G(s) has local minima/maxima. G(s) has local minima/maxima if:

$$F\left(s\right) = \frac{\gamma_x c_{sec}}{2A_{syn}\left(V_{\max} - V_{\min}\right)}$$

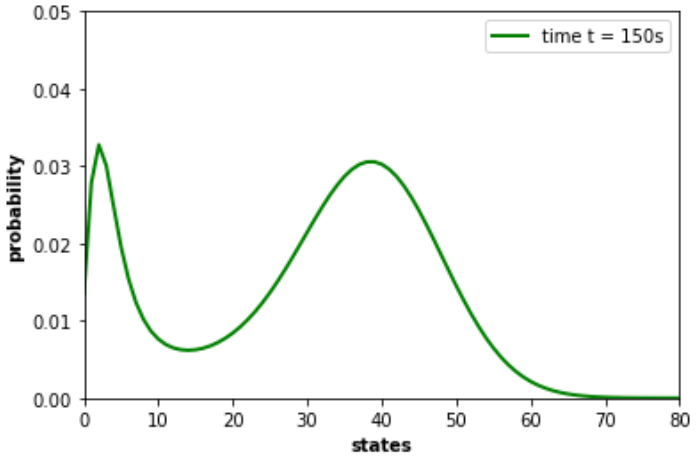

Figure 9: A simulation showing two demarcated peaks representing bistability in the bacterial activation. The parameters used are: $c_{sec} = 14.6$, $\gamma_{ex} = 1.29$, $A_{syn} = 0.818$, $V_{min} = 0.0742$, $V_{max} = 1.99$, $K = 0.0403$, $\gamma_x = 1.29$, $r = 1.16$, $carr_k = 4.34$.

has a (real & positive) solution. The right-hand side of the equation is a constant. F(s) has a bell shape: $F(0) = 0$, $F(s) \to \infty$ for $s \to \infty$ and F(s) is positive for all values other than 0. If:

$$\max F(s) > \frac{\gamma_x c_{sec}}{2A_{syn}(V_{\max} - V_{\min})}$$

then equation (number) is satisfied. The new problem then is to find whether (and for which values) is the maximum of F(s) bigger than the constant right-hand side of (number). Once again, we find the maximum of F(s) from differentiation. After calculations, we find F(s) to have a unique maximum at $\hat{s} = \frac{\sqrt{3}}{3}K$. Inserting the value back in F(s) gives

$$\max F(s) =$$
$$F(\hat{s}) = \frac{3\sqrt{3}}{16K} > \frac{\gamma_x c_{sec}}{2A_{syn}(V_{\max} - V_{\min})}$$

Which we can re-express as

$$K < \frac{3\sqrt{3}A_{syn}(V_{\max} - V_{\min})}{8\gamma_x c_{sec}}$$

If the condition is satisfied, there is bistability. Otherwise, there is only one stable state, and equality gives the bifurcation value.

In any case, given K is within the bi-stability range, we can find the solutions s^{1*}, s^{2*} to $F(s) =$ const. We then get a range for $s_e \in \left(G\left(s^{2*}\right), G\left(s^{1*}\right)\right)$

such that bi-stability occurs.

3.6 Bi-stability

The bifurcation analysis applies to the deterministic model, not the stochastic one. We should not expect results to be exact, as the randomness makes the bifurcation value "blurred". Thanks to the bifurcation analysis, we can choose a set of parameters that show bistable behaviour 9. The dynamics show that after starting with no synthetase molecules, the number of synthetase molecules stabilises very fast at a peak around some equilibrium value. After more time, a second peak appears due to some "tunnelling behaviour", see section 3.8.

The deterministic model finds that the population is either in one state or another. However, due to random effects (local non-homogeneity of cell distribution, factors not counted in the model), some cells "jump" from the lower state to the higher one and stabilise. This effect is possible uniquely thanks to the inclusion of stochastic effects.

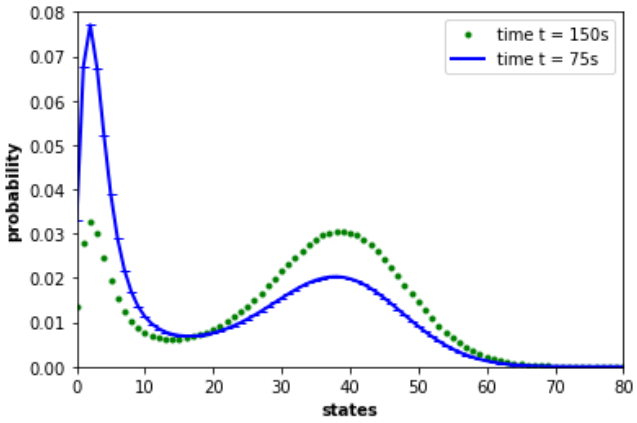

Figure 10: Simulation of the stochastic model. As time passes, the first peak corresponding to the OFF state loses area to the second peak corresponding to the ON state. The parameters used are: $c_{sec} = 14.6$, $\gamma_{ex} = 1.29$, $A_{syn} = 0.818$, $V_{min} = 0.0742$, $V_{max} = 1.99$, $K = 0.0403$, $\gamma_x = 1.29$, $r = 1.16$, $carr_k = 4.34$.

3.7 Alternative Model: Self-Sensing.

In order to capture the self-sensing behaviour, we slightly changed the model. Since we are following an isolated cell's behaviour, no other cells affect the extracellular synthetase concentration: s_e is constant. The slow-fast system then tracks only the intra-cellular autoinducer and synthetase dynamics. The slow-fast system analysis gives:

$$s_i(t) = s_e + \frac{A_{syn}}{\zeta \cdot c_{sec}} i$$

where the other model components remain unchanged, the difference is that s_e is a parameter. After computing a similar analysis than for the quorum sensing situation, the complete model becomes:

$$\begin{aligned}
\dot{p}_i(t) &= -\left(\mathrm{Prod}(i) + \gamma_x i\right) p_i(t) + \gamma_x (i+1)\, p_{i+1}(t) + \mathrm{Prod}(i-1)\, p_{i-1}(t) \\
\mathrm{Prod}(i) &= \zeta \left(V_{\min} + (V_{\max} - V_{\min}) \frac{s_i^2(t)}{s_i^2(t) + K^2} \right) \\
s_i(t) &= s_e + \frac{A_{syn}}{\zeta \cdot c_{sec}} i \\
\bar{X}(t) &= \sum_{j=0}^{M} j \cdot p_j(t) \\
p_i(0) &= \begin{cases} 1 & \text{if } i = 0 \\ 0 & \text{if } i \neq 0 \end{cases}
\end{aligned}$$

This model aims to see whether changing this parameter affects the dynamics: i.e. whether for low s_e values there is no activation, while for bigger s_e values bi-stability can occur.

3.8 Tunnelling Behaviour

Some parameter values show a tunnelling behaviour. Given that two probability peaks coexist, we observe that the first peak loses area to the second peak, as seen in Figure 10. We expect this behaviour to derive from cell heterogeneity, in the form of spatial structures in the bacterial colony, or simply due to variability of parameters within the cell population. We can expect that a setting with hysteresis leads to cells not activating all at once at the deterministic threshold. Because of the cell heterogeneity and white noise, a first group of cells start activating before the archetypical average cell is expected to activate, as predicted in the deterministic model. Furthermore, some cells may still be inactivated even if the deterministic model expects all bacteria to have jumped to the ON state. Hence, as extracellular autoinducer concentration increases, different cohorts of bacteria start jumping from the OFF to the ON state.

4 Discussion and Conclusion

In this thesis, we presented several approaches that capture aspects of quorum sensing and replicate bi-stable behaviour.

We focused on the LasI positive feedback loop that produces the autoinducer molecule. We started with a global outlook and did not include noise at first. We constructed an ODE system, taking the modelling approach of Fujimoto et al. [14]. The ODE system modelled the expectation value of the dynamics of synthetase and autoinducer. To express the value of a dynamic parameter, we included a logistic ODE for population growth. We used experimental data to choose sensible parameters through least-squares parameter fitting. We simulated the deterministic model with the scipy.optimise.odeint python package.

Similarly to Weber and Buceta [15], we were also interested in studying how noise affected quorum sensing, so we assumed heterogeneity among bacteria, and focused on a single cell i. We kept the ODEs expressing the autoinducer dynamics and constructed a stochastic process for the degradation and production of the synthetase molecules. We needed to account for the conversions from a concentration value to a molecule number. We analysed the system from a fast/slow perspective. After recomputation, we found an explicit expression for the intracellular autoinducer concentration in terms of the average intracellular autoinducer concentration, and the average synthetase concentration. Contrary to Weber and Buceta [15], who simulated their stochastic model via the Gillespie algorithm, we chose to derive master equations from the stochastic process and simulate them. The master equations gave an infinitely dimensional system, but we cut away all master equations after some state such that the dynamics remained unchanged. We simulated the model using a two-step method, where single steps follow the standard fourth-order Runge Kutta method. We estimated the dynamics and steady-state of the probability function that outputted the number of synthetase molecules present in cell i. Despite the experimental data showing bistable behaviour with hysteresis, we did not initially see a hysteresis in the model simulations. We supposed the parameter estimation method employed did not accurately fit the parameters to the experimental data. Hence, we performed a bifurcation analysis to find whether bistability was theoretically possible within our model. We found a condition for the parameters in the deterministic model such that there was bistability, and we used this result in the stochastic model to choose parameters that gave bistability. We also observed the interesting "tunnelling behaviour". For some parameter sets, before the stochastic model simulation reaches the steady-state, probability peaks coexisted, representing the two stable states. As the simulation time increased, we observed that the first peak lost area to the second peak. We could expect this in a setting with hysteresis: given a heterogeneous population, for several factors, cells do not activate all at the deterministic threshold. Hence, as extracellular autoinducer concentration increases, different cohorts of bacteria start jumping from the OFF to the ON state. We did not include spatial effects explicitly in the model. However, this tunnelling behaviour showed that the stochastic model can capture bacterial population heterogeneity, in addition to white noise.

Lastly, we constructed a model for the case where self-sensing leads to the activation dynamics rather than quorum sensing. For that, we fixed the extracellular autoinducer concentration and similarly repeated the analysis of the quorum sensing case. We could also reproduce activation in this situation.

Research has made many attempts to simulate quorum sensing. Our work simulated quorum sensing in *P. aeruginosa* using a small model, as done by Fujimoto et al. [14]. We studied how noise affects the system by writing some dynamics through a stochastic process. We simulated the stochastic model via the master equation approximation, rather than through the Gillespie algorithm, as done by Weber and Buceta (2013) [15]. As performed by Fujimoto et al., we found a parameter set where hysteresis occurs through a bifurcation analysis. Our study could be continued by studying how experimental data from different *P. aeruginosa* strains affect the bifurcation range. For example, a strain lacking a specific gene may result in a parameter set that facilitates hysteresis, meaning that the gene negatively affects hysteresis in the wild-type. Another step in the same direction is to do a more thorough parameter estimation study.

5 Acknowledgements

I want to thank my supervisor, Prof. Johannes Müller, for helping me and being supportive throughout my work on the thesis. I want to thank my parents, who have supported me in my studies, and, in general, all the people who bring joy to my life.

6 Appendix

All the code can be found at: https://github.com/maeh98/Thesis

References

[1] Iglewski BH. Pseudomonas. *Medical Microbiology*. Ed. by Baron S. Galveston (TX): University of Texas Medical Branch at Galveston, 1996. Chap. 27. URL: https://www.ncbi.nlm.nih.gov/books/NBK8326/.

[2] Centers for Disease Control et al. *Antibiotic resistance threats in the United States*. Tech. rep. 2019. DOI: http://dx.doi.org/10.15620/cdc:82532.

[3] Daniel J. Smith et al. "Targeting iron uptake to control Pseudomonas aeruginosa infections in cystic fibrosis". In: *European Respiratory Journal* 42.6 (2013), pp. 1723–1736. ISSN: 0903-1936. DOI: 10.1183/09031936.00124012. eprint: https://erj.ersjournals.com/content/42/6/1723.full.pdf. URL: https://erj.ersjournals.com/content/42/6/1723.

[4] Karen L. Visick et al. "Vibrio fischeri lux genes play an important role in colonization and development of the host light organ". In: *Journal of Bacteriology* 182.16 (2000), pp. 4578–4586. DOI: 10.1128/JB.182.16.4578-4586.2000. eprint: https://journals.asm.org/doi/pdf/10.1128/JB.182.16.4578-4586.200. URL: https://journals.asm.org/doi/abs/10.1128/JB.182.16.4578-4586.2000.

[5] Colleen T. O'Loughlin et al. "A quorum-sensing inhibitor blocks ¡i¿Pseudomonas aeruginosa¡/i¿ virulence and biofilm formation". In: *Proceedings of the National Academy of Sciences* 110.44 (2013), pp. 17981–17986. DOI: 10.1073/pnas.1316981110. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.1316981110. URL: https://www.pnas.org/doi/abs/10.1073/pnas.1316981110.

[6] Jasmine Lee and Lianhui Zhang. "The hierarchy quorum sensing network in Pseudomonas aeruginosa". In: *Protein & cell* 6 (Sept. 2014). DOI: 10.1007/s13238-014-0100-x.

[7] Marie Allesen-Holm et al. "A characterization of DNA release in Pseudomonas aeruginosa cultures and biofilms". In: *Molecular Microbiology* 59.4 (2006), pp. 1114–1128. DOI: https://doi.org/10.1111/j.1365-2958.2005.05008.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2958.2005.05008.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2958.2005.05008.x.

[8] Burkhard A. Hense and Martin Schuster. "Core Principles of Bacterial Autoinducer Systems". In: *Microbiology and Molecular Biology Reviews* 79.1 (2015), pp. 153–169. DOI: 10.1128/MMBR.00024-14. eprint: https://journals.asm.org/doi/pdf/10.1128/MMBR.00024-14. URL: https://journals.asm.org/doi/abs/10.1128/MMBR.00024-14.

[9] Geoffrey W. Hoffmann. "Neurons with hysteresis?" In: *Computer Simulation in Brain Science*. Ed. by Rodney M. J.Editor Cotterill. Cambridge University Press, 1988, pp. 74–87. DOI: 10.1017/CBO9780511983467.006.

[10] Rosemary J Redfield. "Is quorum sensing a side effect of diffusion sensing?" In: *Trends in Microbiology* 10.8 (2002), pp. 365–370. ISSN: 0966-842X. DOI: https://doi.org/10.1016/S0966-842X(02)02400-9. URL: https://www.sciencedirect.com/science/article/pii/S0966842X02024009.

[11] Burkhard A. Hense et al. "Does efficiency sensing unify diffusion and quorum sensing?" In: *Nature Reviews Microbiology* 5 (2007), pp. 230–239.

[12] Magnus Fagerlind et al. "The Role of Regulators in the Expression of Quorum-Sensing Signals in Pseudomonas aeruginosa". In: *Journal of molecular microbiology and biotechnology* 6 (Feb. 2003), pp. 88–100. DOI: 10.1159/000076739.

[13] Gabriel Dilanji, Jessica Langebrake, and Stephen Hagen. "Quorum Activation at a Distance: Spatiotemporal Patterns of Gene Regulation from Diffusion of an Autoinducer Signal". In: *Journal of the American Chemical Society* 134 (Feb. 2012), pp. 5618–26. DOI: 10.1021/ja211593q.

[14] Koichi Fujimoto and Satoshi Sawai. "A Design Principle of Group-level Decision Making in Cell Populations". In: *PLoS computational biology* 9 (June 2013), e1003110. DOI: 10.1371/journal.pcbi.1003110.

[15] Marc Weber and Javier Buceta. "Dynamics of the quorum sensing switch: Stochastic and non-stationary effects". In: *BMC systems biology* 7 (Jan. 2013), p. 6. DOI: 10.1186/1752-0509-7-6.

[16] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. *Numerical Mathematics (Texts in Applied Mathematics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN: 3540346589.

[17] "Deterministic Slow-Fast Systems". In: *Noise-Induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach*. London: Springer London, 2006, pp. 17–49. ISBN: 978-1-84628-186-0. DOI: 10.1007/1-84628-186-5_2. URL: https://doi.org/10.1007/1-84628-186-5_2.

[18] Jack K. Hale. *Dynamics and Bifurcations by Jack K. Hale, Hüseyin Kocak.* eng. 1st ed. 1991. Texts in Applied Mathematics, 3. New York, NY: Springer New York, 1991. ISBN: 1-4612-4426-9.

[19] K. Godfrey. "Identification of parametric models from experimental data [Book Review]". In: *IEEE Transactions on Automatic Control* 44.12 (1999), pp. 2321–2322. DOI: 10.1109/TAC.1999.811220.

[20] J.L. Doob. *Stochastic Processes*. Wiley publications in statistics. Wiley, 1962. URL: https://books.google.de/books?id=7Bu8jgEACAAJ.

[21] Parker Smith and Martin Schuster. "Antiactivators prevent self-sensing in *Pseudomonas aeruginosa* quorum sensing". In: *Proceedings of the National Academy of Sciences* 119.25 (2022), e2201242119. DOI: 10.1073/pnas.2201242119. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.2201242119. URL: https://www.pnas.org/doi/abs/10.1073/pnas.2201242119.

[22] Dong-Ju Kim et al. "Relation of microbial biomass to counting units for Pseudomonas aeruginosa". In: *African Journal of Microbiology Research* 6 (2012), pp. 4620–4622.

[23] "CHAPTER 8 The Deterministic Limit of Stochastic Theory". In: *Stochastic Systems*. Ed. by George Adomian. Vol. 169. Mathematics in Science and Engineering. Elsevier, 1983, pp. 225–242. DOI: https://doi.org/10.1016/S0076-5392(08)61933-5. URL: https://www.sciencedirect.com/science/article/pii/S0076539208619335.