

Biofilms:  
Modeling approach for *growth* and  
*spatial-separation*

Biofilme:  
Modellierungsansatz für *Wachstum* und *räumliche Separation*

BACHELORARBEIT

zur Erlangung des akademischen Grades  
Bachelor of Science (B. Sc.)  
im Fach Biophysik



eingereicht an der  
Lebenswissenschaftlichen Fakultät  
Institut für Biologie  
Humboldt-Universität zu Berlin

von  
Herrn Maxim Karnetzki  
geboren am 26.12.1992 in Berlin

eingereicht am: 09.02.2016

## Abstract

**English version:** In recent years biofilms have become one of the most exciting fields of research in biology, in particular microbiology and evolutionary biology.

A deeper understanding of this community life can lead to important findings in the investigation of bacterial infections as biofilms are known to be more resilient to antibiotics. Furthermore, general evolutionary processes that contribute to the formulation of different survival strategies and the spreading of species, can be studied as if in a time lapse (generation time, for example, *E.coli* = 20 minutes, human = 25 years). Experiments of selected papers, whose research results describe essential aspects of a biofilm, are recreated in this study with my own modeling approach, which was developed and independently programmed (in Python language). The main focus is on qualitative reproducibility of the described phenomena. The following models were adapted in this modeling approach: (i) intercellular communication, by using a cellular automata model based on reaction-diffusion (RD) equations, (ii) modeling each individual cell in the computational domain as an object using a modeling approach referred to as individual-based modeling.

Because these models are defined in a general way, a very large application area for this program is possible. It is used in this work to describe, for example, a survival competition between three different strains of *E.coli* in a cell culture or a simplified version of the mating process of Baker's yeast *Saccharomyces cerevisiae*.

### Deutsche Version:

In den letzten 15 Jahren hat sich die Biofilmerforschung zu einem der spannendsten Wissenschaftsgebiete in der Biologie, insbesondere der Mikrobiologie und der Evolutionsbiologie, entwickelt.

Ein tieferes Verständnis dieses Gemeinlebens kann zu wichtigen Erkenntnissen in der Erforschung von bakteriellen Krankheiten führen, denn Biofilme sind bekannt für ihre Antibiotikaresistenz. Des Weiteren können grundlegende evolutionäre Prozesse, die zur Herausbildung verschiedener Überlebensstrategien und Verbreitung von Arten führen, wie im Zeitraffer untersucht werden (Generationszeit z.B. *E.coli*  $\approx$  20 Minuten, Mensch  $\approx$  25 Jahre). Ausgewählte Papers, deren Forschungsergebnisse wesentliche Aspekte der bakteriellen Gemeinschaft, die den Biofilm ausmachen, beschreiben, werden mit meinem, im Rahmen dieser Studie erarbeiteten und selbstständig programmierten (in Python), Modellierungsansatz nochmals untersucht. Dabei liegt das Hauptaugenmerk auf einer qualitativen Reproduzierbarkeit der beschriebenen Phänomene. Folgende Modelle wurden in diesem Modellierungsansatz dazu verbunden: (i) unter Verwendung von zellulären Automaten mit Hilfe von Reaktions-Diffusions-Gleichungen wurde interzelluläre Kommunikation modelliert, (ii) Modellierung jeder einzelnen Zelle in der Rechenumgebung als räumliches Objekt mithilfe des sogenannten "individual based modeling"-Ansatzes.

Dadurch dass diese verbundenen Modelle sehr allgemein verfasst sind, ergibt sich ein sehr großer Anwendungsbereich für das Programm. Mein Modellierungsansatz wird in dieser Studie z.B. einem Überlebenskampf zwischen drei unterschiedlichen *E.coli* Strängen in einer Zellkultur beschreiben oder eine vereinfachte Version des Paarungsprozess von Hefezelle *Saccharomyces cerevisiae*.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theoretical Foundations</b>	<b>3</b>
2.1	Diffusion . . . . .	3
2.1.1	State-decisions . . . . .	5
2.2	Cells as Spheres . . . . .	7
2.2.1	Movement . . . . .	7
2.2.2	Border Control . . . . .	8
2.3	Growth . . . . .	8
2.3.1	Cell Division . . . . .	9
<b>3</b>	<b>Results and Discussion</b>	<b>10</b>
3.1	Computational Parameters . . . . .	10
3.2	Foundation . . . . .	11
3.3	Prisoner's Dilemma . . . . .	12
3.3.1	Discussion . . . . .	17
3.3.2	Growth Model Idea . . . . .	17
3.4	<i>Mutualism</i> . . . . .	19
3.4.1	Discussion . . . . .	20
3.5	Chemical Warfare . . . . .	21
3.5.1	Discussion . . . . .	23
<b>4</b>	<b>Conclusions and Outlook</b>	<b>25</b>

# 1 Introduction

In natural habitats, bacteria often occur in multicellular community referred to as biofilms. After attachment to a surface, bacteria can adapt to life by forming a biofilm. Two properties are often associated with surface-attached bacteria: increased synthesis of extracellular matrix and the development of antibiotic resistance [1]. The robust extracellular matrix consists of proteins, amyloid fibres, exopolysaccharides and extracellular DNA [1], [2], [3].

Therefore biofilms show pronounced stress resistance including a resilience against antibiotics that can cause serious medical problems [1], [2], [3]. These bacteria communities may also develop other properties, including increased resistance to UV light, increased rates of genetic exchange, altered biodegradative capabilities, and increased secondary metabolite production[1].

The bacterial activity is determined by various stress responses which are in return a response to gradients of nutrients, oxygen, waste products and signalling compounds that build up in growing biofilms [2]( bacteria can either grow and proliferate or enter into a stationary phase and use their remaining resources for survival).

As a consequence, biofilms can differentiate into at least two distinct layers of vegetatively growing or stationary phase cells that exhibit very different cellular physiology [2]. This includes layering of matrix production which has an impact on microscopic architecture, biophysical properties and visible morphology of the macrocolony [2]. Using selected recent studies as inspiration, this bachelor thesis describes current knowledge about the underlying networks – prominently cell-cell communication due to concentration gradients and second messengers[4]– that play a key role in spatial separation and growth [5], [6], [7], [8].

By simulating different social phenotypes/motifs of interaction that have been created by genetic manipulation of bacteria or yeast strains, I will demonstrate the versatility of my program and the underlying modeling approach. For this approach I adapted the conceptual framework for cell-cell communication proposed by Maire and Youk in 2015 [9] and connected it to an agent-based modeling approach proposed by a group at the Delft University in 2005 [10].

**Maire and Youk framework** They developed a conceptual framework for understanding the unicellular and multicellular control of gene expression for a ubiquitous class of cells that communicate through secreting and sensing a signaling molecule (see below Figure 1.1). Maire and Youk use geometric means to quantify the amounts of autonomy and collectivity of cells. Their conceptual framework ties together diverse systems, including tissues and microbes, with common principles.

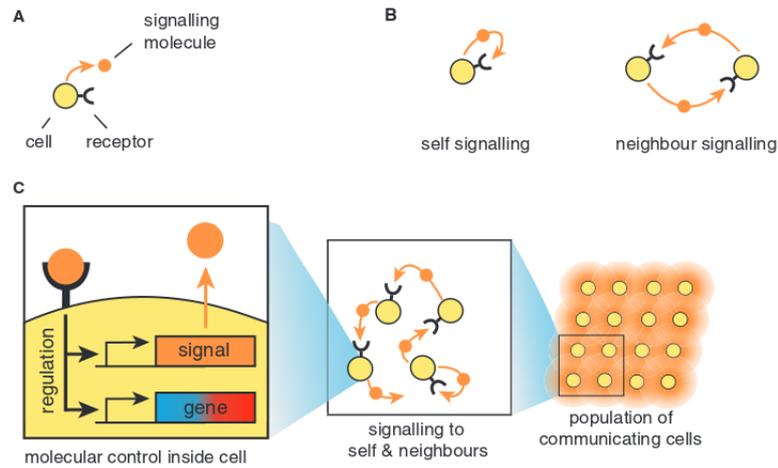


Figure 1.1: Their bottom-up approach for the framework: from molecules to populations of cells. (A) Secreted-and-sense cell. (B) Secreted-and-sense cell signal to itself (self-signaling) and signal to its neighboring cells (neighbour signaling). (C) Outline of the bottom-up approach. Image taken from [9].

The schematic overview in Figure 1.2 depicts the social phenotypes that will be simulated and their influence on the spatial structure of the biofilm starting from an initially well-mixed population. a | Antagonistic phenotypes, which for example secrete of toxins, can eliminate sensitive cells in the vicinity. b | Mutualistic cell strains tend to become entangled as their growth rates are proportional to their neighbor. This can result in spatial mixing of the cell strains and exclusion of cheating or non-interacting strains. c | Cells that secrete public goods preferentially benefit from nearby clonemates, which proliferate more rapidly than neighboring strains and thus cut other lineages off from the frontier of the biofilm.

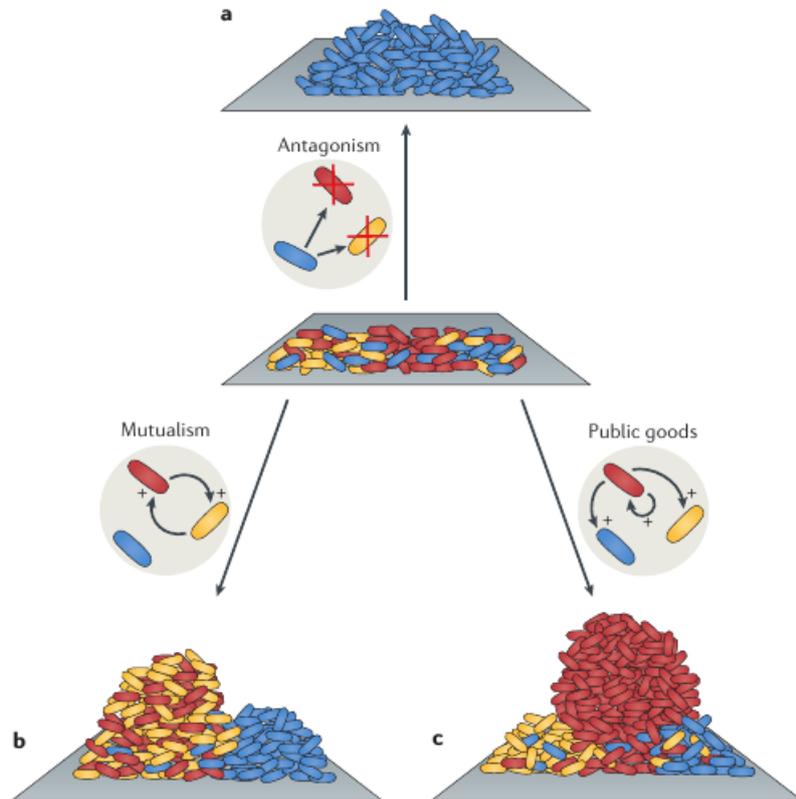


Figure 1.2: Schematic overview of social phenotypes and their influence on the spatial structure of the biofilm starting from an initially well-mixed population. a | Antagonistic phenotypes can eliminate sensitive cells in the vicinity. b | Mutualistic cell strains can become entangled. c | Cells that secrete public goods preferentially benefit from nearby clonemates. Image from [8].

## 2 Theoretical Foundations

This chapter is dedicated to the driving principles and equations that are the basis for my model.

### 2.1 Diffusion

In the paper of Maire and Youk in 2015 [9] a key element of cell-cell communication is diffusion of a signal molecule  $S$ , which the cells themselves produce, secrete and sense. This particular paper is more focused on building a conceptual framework in which a theoretical approach is build that links secrete-and sense cells genetic circuits to multicellular behavior. There are also other instances in which cells with similar attributes are artificially built [6] or observed in nature [3]e.g. quorum sensing in E.coli [4].

For my program I started from a very basic diffusion equation for the concentration  $S$  with diffusion constant  $D$ , degradation constant  $\gamma$  and radial symmetric properties, as cells are postulated to be spheres in 3-D or circles in 2-D, while the secretion is a uniform flux  $J$  over the cell's border:

$$\frac{\partial S}{\partial t} = \underbrace{\frac{1}{r^2} \cdot \frac{\partial}{\partial r} \left( D \cdot r^2 \cdot \frac{\partial S}{\partial r} \right)}_{diffusion} - \underbrace{\gamma \cdot S}_{degradation} \quad (2.1)$$

Defining source flux  $J$ :

$$J = -D \cdot \nabla S \quad (2.2)$$

$$= -D \cdot \frac{\partial S}{\partial r} \quad (2.3)$$

Boundary condition on the border of the cell with radius  $R$ :

$$J = -D \cdot \frac{\partial S}{\partial r} \Big|_{r=R} \quad (2.4)$$

Cells frequently measure the possibly varying outside concentration of  $S$  over a long time, averages these measurements, and the average concentration  $\tilde{S}$  regulates its genes (so the production of said molecule in the simplest scenario). Since the concentration usually reaches a steady-state much faster than the time taken for this averaging [11], one only needs to look at the steady-state concentration for a cell's regulation [9]. To explain this assumption I refer the Stokes-Einstein-relation:

$$D = \frac{k_B T}{6\pi\eta r}, \quad (2.5)$$

with  $D = \text{Diffusioncoefficient}$ ,  $k_B = \text{Boltzmannconstant}$ ,  $T = \text{absolute temperature}$ ,  $\eta = \text{dynamic viscosity of the solute and particle radius } r$ . It is generally assumed that  $r$  scales with the molecular weight  $M$  as  $r \sim M^{\frac{1}{3}}$ , a change by the factor of 1000 in the molecular weight causes the diffusion coefficient to only change by a factor of 10 [12]. So the dominant part that governs the diffusion coefficient is the viscosity  $\eta$  of the extracellular medium. Therefore the solving approach for equation (2.1) with  $\lambda = \sqrt{\frac{D}{\gamma}}$ , as typical traveling distance of molecule  $S$  before decay, is:

$$S(r) = \frac{S_r \cdot R}{r} \cdot e^{-\frac{r-R}{\lambda}} \quad (2.6)$$

$$\begin{aligned} \frac{\partial S}{\partial r} &= -\frac{S_r \cdot R}{r^2} \cdot e^{-\frac{r-R}{\lambda}} - \frac{S_r \cdot R}{r\lambda} \cdot e^{-\frac{r-R}{\lambda}} \\ &= -\frac{S_r \cdot R}{r} \cdot e^{-\frac{r-R}{\lambda}} \cdot \left(\frac{1}{r} + \frac{1}{\lambda}\right) \end{aligned} \quad (2.7)$$

$$\begin{aligned} J &= -D \frac{\partial S}{\partial r} \Big|_R = D \cdot S_R \cdot \left(\frac{1}{r} + \frac{1}{\lambda}\right) \\ &\Rightarrow S_R = \frac{J}{D \left(\frac{1}{r} + \frac{1}{\lambda}\right)} \end{aligned} \quad (2.8)$$

$$\text{with: } \lambda = \sqrt{\frac{D}{\gamma}}$$

$$\Rightarrow S_R = \frac{J \cdot R}{R\sqrt{D\gamma} + D} \quad (2.9)$$

$S_R$  is the secreted concentration level, which is dependent on flux  $J$  and in a straightforward manner on the cell radius  $R$  (2.9). While cell's radius will be further discussed in section *Growth*, flux  $J$  still is of importance.

If the secretion rate is chosen to be relative to the surface  $\rightarrow J = \text{const}$  and (2.9) would describe  $S_R$  fully else if the secretion rate should be relative to the volume  $V_i$  of cell  $c_i$ , flux  $J$  scales too:

$$J = \left[ \frac{\text{molecules}}{\text{time} \cdot \text{area}} \right] \Rightarrow J \cdot 4\pi R^2 = \left[ \frac{\text{molecules}}{\text{time}} \right] \quad (2.10)$$

with a new production constant per volume  $P$ :

$$P \cdot \frac{4}{3}\pi R^3 = \left[ \frac{\text{molecules}}{\text{time}} \right] \quad (2.11)$$

$$\begin{aligned} \Rightarrow J \cdot 4\pi R^2 &= P \cdot \frac{4}{3}\pi R^3 \\ \Rightarrow J &= \frac{P \cdot R}{3} \end{aligned}$$

$$\Rightarrow S_R = \frac{1}{3} \cdot \frac{P \cdot R^2}{R\sqrt{D\gamma} + D} \quad (2.12)$$

For the framework a scaling secretion rate with a constant particle number is put in place as none of the reference suggest otherwise [9], [10],[8].

So  $P$  is a settable parameter of my model. In [9] a similar approach was chosen for  $S_R$ :

$$\frac{\partial S}{\partial t} = \frac{1}{r^2} \cdot \underbrace{\frac{\partial}{\partial r} \left( D \cdot r^2 \cdot \frac{\partial S}{\partial r} \right)}_{\text{diffusion}} - \underbrace{\gamma \cdot S}_{\text{degradation}} \quad (2.13)$$

But a different source term was defined:

$$\frac{\eta}{4\pi R^2} \delta(r - R) \quad (2.14)$$

with  $\delta$  as Dirac delta function and  $\delta(r - R) = 1$  when  $r = R$  using the solving approach (2.6):

$$\Rightarrow \frac{\eta}{4\pi R^2} \frac{1}{\gamma\lambda} \cdot \frac{1}{1 + \frac{\lambda}{R}} \quad (2.15)$$

with  $\eta$  as constant secretion rate.

$$= \frac{\eta}{4\pi R^2} \cdot \frac{1}{\sqrt{D\gamma}} \cdot \frac{1}{1 + \frac{\lambda}{R}} \quad (2.16)$$

$$= \frac{\eta}{4\pi R^2} \cdot \frac{1}{\sqrt{D\gamma}} \cdot \frac{1}{1 + \frac{\sqrt{\frac{D}{\gamma}}}{R}} \quad (2.17)$$

$$S_R = \frac{\eta}{4\pi R^2} \cdot \frac{R}{R\sqrt{D\gamma} + D} \quad (2.18)$$

As one can see both solutions are very similar and by choosing  $J = \frac{\eta}{4\pi R^2}$  they are equal. With this one further understands the cell-cell communication (production and secretion).

### 2.1.1 State-decisions

The concentration level is the deciding factor for a cell's behavior and cell-cell communication. As mentioned above (see 2.1 Diffusion), the assumption is made that a cell's measurements, and then regulating its gene expression takes longer time, than the signal molecule  $S$  needs to reach a steady-state-level [9].

Therefore by assuming a simple motive in which the signal molecule  $S$  only regulates its own gene expression we have two states the cell can be in, "on" if producing or "off" if not, so two secretion levels and two concentration levels on the surface  $\tilde{S}_{on}$  and  $\tilde{S}_{off}$ . These two states are divided by a threshold concentration  $K$ .

Needless to say if one looks at only one cell  $c_i$  or if  $D \ll \gamma$  the sensed concentration is only cell  $c_i$ 's own so  $\tilde{S} = S_i$ .

To make significant comparisons, all concentration terms are divided by  $\tilde{S}_{off}$  [9]:

$$\begin{cases} K = \frac{\tilde{K}}{\tilde{S}_{off}} \\ S_{on} = \frac{S_{on}}{\tilde{S}_{off}} \\ S_{off} = 1 \end{cases} \quad (2.19)$$

From Equation (2.19) we see that  $S_{on}$  and  $K$  are freely adjustable parameters of the cell and thus of the model. By comparing its sensed concentration  $S_i$  to  $K$  cell  $c_i$  then would decide its next state. In [9] the concentration plus the feedback of the gene expression (positive  $\rightarrow$  more production by turning cells "on" or negative feedback  $\rightarrow$  less production by turning cells "off") determine cell  $c_i$ 's next state.

So for positive feedback if  $S_i > K$  then  $c_i$ 's state is set "on", and for any other

instance "off", while for negative feedback it is the exact opposite.

Now the not trivial case when there are  $N$  cells and diffusion is noteworthy.  $S_i$  is the sum of the concentration of the signal molecule that  $c_i$  secretes (denoted  $S_R$ ) and the concentration of the molecule  $S$  secreted by all the other cells (denoted  $S_{neighbors}$ ) [9]. This is calculated with equation (2.6):

$$S_i = \underbrace{S_R}_{\text{due to self}} + \underbrace{\sum_{j=1}^{N-1} S_{R,j} \cdot \frac{R_j}{r_j} \cdot \exp\left(-\frac{r_j - R_j}{\lambda}\right)}_{\text{due to neighbors}} \quad (2.20)$$

One can see that the neighbors influence is determined by how far they are from cell  $c_i$  and  $\lambda = \sqrt{\frac{D}{\gamma}}$ , which means by the set constants for diffusion  $D$  and degradation  $\gamma$ . Higher values of  $D$  in comparison to  $\gamma$  lead to further diffusion so to a stronger influence of the neighbors (term gets bigger in (2.20)), higher values of  $\gamma$  lead to quicker degradation so weaker influence of neighboring cells (smaller term in (2.20)). These observations are in line with one's expectations and intuition, and give us a tool for adjusting cell-cell communication.

Here an illustration of what cell  $c_i$  could sense:

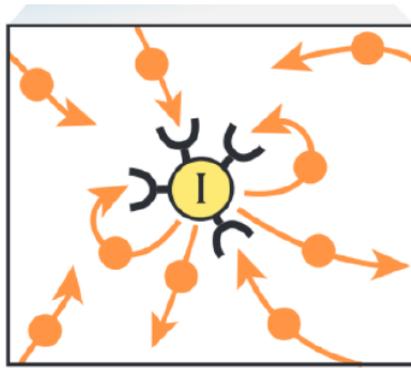


Figure 2.3: Representation of cell-cell communication. Picking any cell  $c_i$ .  $S_i$  is then the concentration of the signal molecule on cell  $c_i$  calculated after equation (2.20) with increasing  $D$  neighboring cells also increase their influence on cell  $c_i$ . Image taken from [9].

The neighbors' influence is of interest because the bigger the influence is the more the cells lose their autonomy[9]. This means that single cells form a multicellular compound in which the compound as a whole influences the individual's state e.g. quorum sensing in bacteria[7],[1],[13]. In numbers expressed this means even if  $S_{on} < K$  if all the neighbors "pitch in"  $S_i$  can be  $> K$  and therefore a different state is set than would be without the neighbors' influence. With this understanding of cell-cell communication we have a very sensibly definable, quantifiable and tunable system implemented to reproduce multicellular behavior which is not only of interest for bacteria but also might give new insight in other fields where secrete- and sense cells are of interest like breast cancer[14].

## 2.2 Cells as Spheres

As mentioned in 2.1 Diffusion the cells are implemented as spheres. The framework follows the individual-based modeling approach [10], therefore every cell  $c_i$  is modeled as an individual object  $O_i$  with starting radius  $R_i = R_0$  and coordinates  $(x_i, y_i, z_i)$ , so it occupies space. What is of interest in this section are the consequences of cell division and cell growth (this will be further explained in section *Growth*), which means movement of the cells.

### 2.2.1 Movement

In this model there are basically three reasons a cell is being moved:

1. Growth
2. Division
3. Gravity

- (1): cells get so big that they press against each other;
- (2): a new cell is created which needs space so pushing for it;
- (3): an important piece of reality and biological relevance<sup>1</sup>.

When reading "push", "press" and gravity one immediately thinks to describe this with force-vectors and in [15] a solid solution is presented, which I also implemented for this framework. The new position  $x_{i,new} = x_i - \delta x_i$ , where  $x_i$  is the coordinate of the current cell  $c$  (i.e.,  $x_1 = x$ ,  $x_2 = y$ , and  $x_3 = z$ ). The three components of the displacement(force) vector are  $\delta x_i$ :

$$\delta x_i^{(c)} = \sum_N \frac{x_i^{(n)} - x_i^{(c)}}{d^{(n)}} \left[ k \cdot (R^{(n)} + R^{(c)}) - d^{(n)} \right], \quad (2.21)$$

where  $R^{(c)}$  is the radius of the current cell,  $R^{(n)}$  is the radius of a neighboring cell,  $x_i^{(n)}$  is the coordinate of an overlapping neighboring cell, and  $d^{(n)}$  is the distance between cells:

$$d^{(n)} = \sqrt{\sum_{i=1}^3 (x_j^{(n)} - x_i^{(p)})^2} \quad (2.22)$$

The parameter  $k$  should be understood as the shove radius, so a multiplier of the cell radius with which one adjusts how close cells can be.

- $k < 1$ : overlapping of cells allowed with  $1 - k$  percent  
 $k \geq 1$ : no overlapping  $k \cdot$  distance in between

For the realization of the gravitational pull a term  $F_g$  is always added to the height coordinate  $y$ .

Gravitational Term  $F_g$  for cell  $c_i$ :

$$F_g = -V_i \cdot g \quad (2.23)$$

---

<sup>1</sup>let's keep the feet on the ground

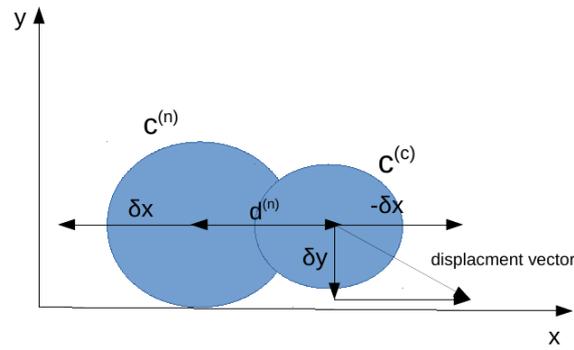


Figure 2.4: Schematic 2-dimensional example how displacement vector is calculated.

There is no directed self-movement of the cells implemented because a biofilm attaches itself to a surface[1] and cells are connected to each other by extracellular fibers(flagella, amyloid curly fibers)[2] which consequently nearly arrests single cell movement or at least renders negligible for my biofilm modeling approach.

### 2.2.2 Border Control

The computational domain is limited. In the model the cells will reach set computational domain boundaries ( usually cuboid  $[0, L_X]x[0, L_Y]x[0, L_Z]$ ). Now I present ideas how one can implement these domain boundaries:

In [15] periodic or wrapped-around boundaries in the directions parallel to the surface is chosen (directions  $x$  and  $z$ ) If the cell extends into the surface (the cell's center is at  $y < R$ ), then the cell is shifted to  $y = R$ . Furthermore, undesired edge effects can be avoided by not having edges. This means, for instance, that in the  $x$  direction, cells close to the domain boundary at  $x = L_X$  can also have particles as neighbors at the boundary opposite,  $x = 0$ . Consequently, particles shoved out at  $x = L_X + \delta x$  reenter the computational domain at  $x = \delta x$ , and those pushed out to  $x = -\delta x$  are resited at  $x = L_X - \delta x$ .

While this is an elegant way to avoid undesired edge effects, but poses problems in the implementation so instead I chose reflecting boundaries. The shifting condition is than different for any directions, it is:

e.g.  $x$  direction:

$$\max(\min(x + \delta x, L_X - R), R) \quad (2.24)$$

If the cell would be shifted outside, instead it will be placed as close to the border as possible given the cells radius  $R$ . In my opinion the undesired edge effects might be something worth studying <sup>2</sup>, but is not the focus of this bachelor's thesis.

## 2.3 Growth

The cell's growth rate is governed by a simple function. The cell radius is growing by a certain set growth rate  $g_{rate}$  so  $R_{new} = R_{old} + g_{rate}$ . As for the decision whether a cell grows this time-step  $t$  is a question of it's *state* and the social scenario the model is describing (e.g. the cell can only grow if it senses a certain amount of the

---

<sup>2</sup>see: Adhesion

signal molecule  $S$ ).

In Youk 2015 [9], the cells' growth was not of interest, while in the framework of Delft University [15],[10] growth is implemented using particle/molecule-absorption and mass-conversion with an ODE approach. In light of the computational power and duration of a bachelor thesis this more precise but complex method could not be adapted.

### 2.3.1 Cell Division

When a cell  $c_1$  reaches the critical volume  $V_1 = 2 \cdot V_0$ , with  $V_0 = \frac{4}{3} \cdot \pi \cdot r_0^3$ , while  $r_0$  is a set starting value of my choosing, cell division is initiated.

Through this division a new cell  $c_2$  is generated, while  $c_1$  is set back to the starting value  $cellradius = r_0$  and stays at its spatial coordinates.

Though for  $c_2$  a random choice function generates its spatial coordinates. The function picks a random point on the surface of a unit sphere, but it is incorrect to select spherical coordinates  $\theta$  and  $\phi$  from the distributions  $\theta$  in  $[0, 2\pi]$  and  $\phi$  in  $[0, \pi]$ , because the area element  $d\Omega = \sin \phi d\theta d\phi$  is a function of  $\phi$ , so the poles than a have slight positive bias in being selected.

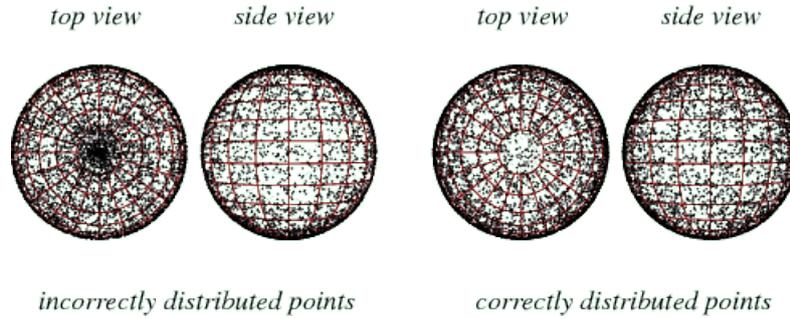


Figure 2.5: 3-dimensional representation of spherical distribution. Left side with spherical coordinates, right side intended distribution [16].

One way to achieve this intended distribution was proposed by Muller 1959 [17]. Accordingly one generates three Gaussian random numbers  $x$ ,  $y$  and  $z$ :

$$S = x^2 + y^2 + z^2 \quad (2.25)$$

with:

$$\begin{pmatrix} \frac{x}{\sqrt{S}} \\ \frac{y}{\sqrt{S}} \\ \frac{z}{\sqrt{S}} \end{pmatrix} \quad (2.26)$$

so that cell  $c_2$  coordinates are:

$$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = r_0 \cdot \begin{pmatrix} \frac{x}{\sqrt{S}} \\ \frac{y}{\sqrt{S}} \\ \frac{z}{\sqrt{S}} \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad (2.27)$$

## 3 Results and Discussion

This chapter is dedicated to the results produced by my model.

### 3.1 Computational Parameters

As in Chapter 2. *Theoretical Foundations* already mentioned the framework has several parameters that need to be set before running the model:

Table 3.1: Parameters of the model

Parameters	Description	Values
$S_{on}$	set concentration if <i>state</i> "on"	18 in arbitrary unit
$K$	threshold concentration for state-decision	10 in arbitrary unit
<i>feedback</i>	positive/negative see <i>state-decision</i>	positive
$R_0$	starting value for cell radius	1 $\mu m$
$D$	diffusion constant	7 $\frac{\mu m^2}{s}$
$\gamma$	degradation constant	7 $\frac{\mu m^2}{s}$
$L_X, L_Y, L_Z$	length of computational domain	in $\mu m$
$k$	shove factor	1 in arbitrary unit
<i>grate</i>	growth rate	0,2 $\mu m$
<i>placement</i>	probability of cell being placed for initial setup	100%
<i>stategamble</i>	probability of cell being setup with "on" <i>state</i>	50%
$x_{grid}$	sets aid-lattice size	10

Most of the parameters in Table (3.1) have been introduced in Chapter 2. *Theoretical Foundations* except the last three.

*Stategamble* is the probability for any new cell to be setup with the initial *state* "on", setting cell state occurs while initiation and cell division.

*Placement* and  $x_{grid}$  are auxiliary parameters important for the experimental domain set-up. These two can be categorized as initial cell-density setting parameters:  $x_{grid}$  divides  $L_X$ ,  $L_Y$  and  $L_Z$  in  $x_{grid}$ -equal pieces in which only one cell with the probability *placement* one cell can be placed. An 1-dimensional example:  $x_{grid} = 5$ ,  $L_X = 100.0$  possible placement sites are at: 0.0|20.0|40.0|...

After the program starts to run these two parameters have no influence on the model.

Here an illustration:

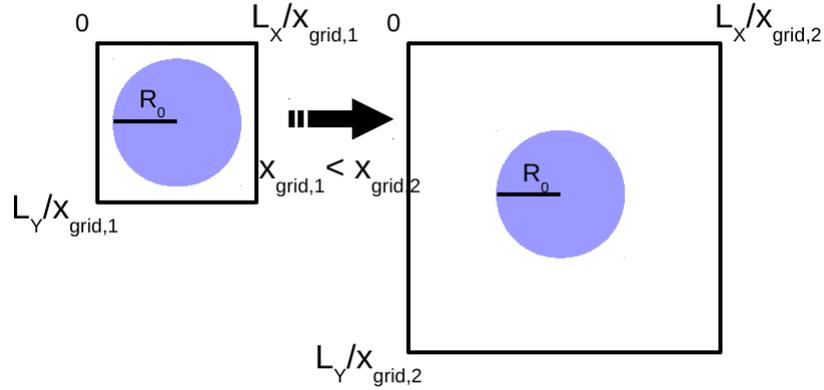


Figure 3.6: Schematic 2-dimensional samples of one panel of the lattice set by two different  $x_{grid}$ , so  $x_{grid,1} < x_{grid,2}$  but  $R_{0,1} = R_{0,2} = R_0$

## 3.2 Foundation

Here I present the core implementation based on the chapter *Theoretical Foundations*. The following sections describe derivatives of this program. Slight changes in the code are introduced for specifications of different motives. Nevertheless all core equations remain intact so a general overview is beneficial.

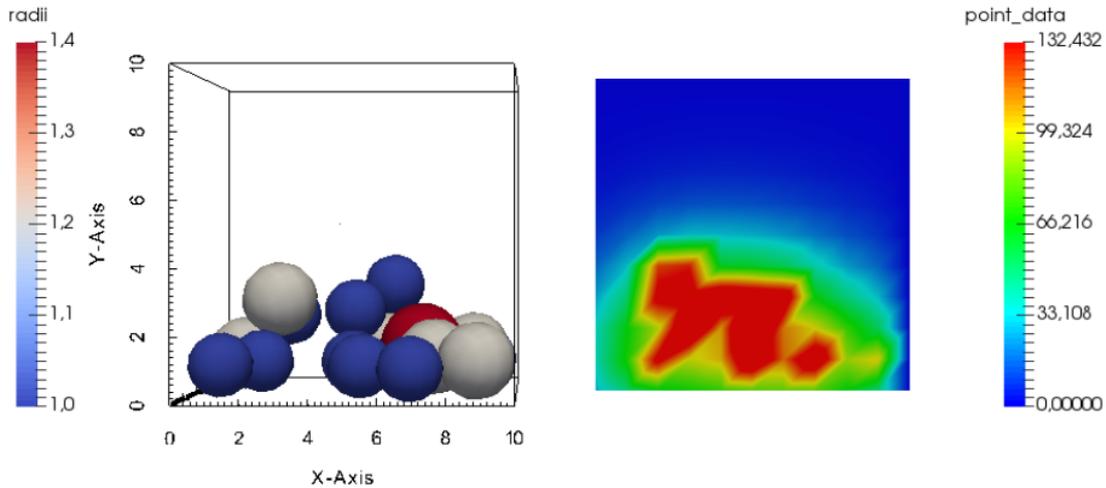


Figure 3.7: 3-dimensional representation of model data. Biofilm with  $D = 7$ ,  $S_{on} = 10.0$ ,  $K = 18.0$ ,  $\gamma = 7$  at  $t = 12$ .  $L_X = 10.0$ ,  $L_Y = 10.0$  and  $L_Z = 10.0$  set lengths of computational domain. On the left cells represented as spheres with radii  $r \in [1, 1.4]$ . On the right concentration levels of signal molecule  $S$  of points in the cube for visualization purposes. These concentration levels only schematically represent what the cells measure. Visualization program *Paraview*.

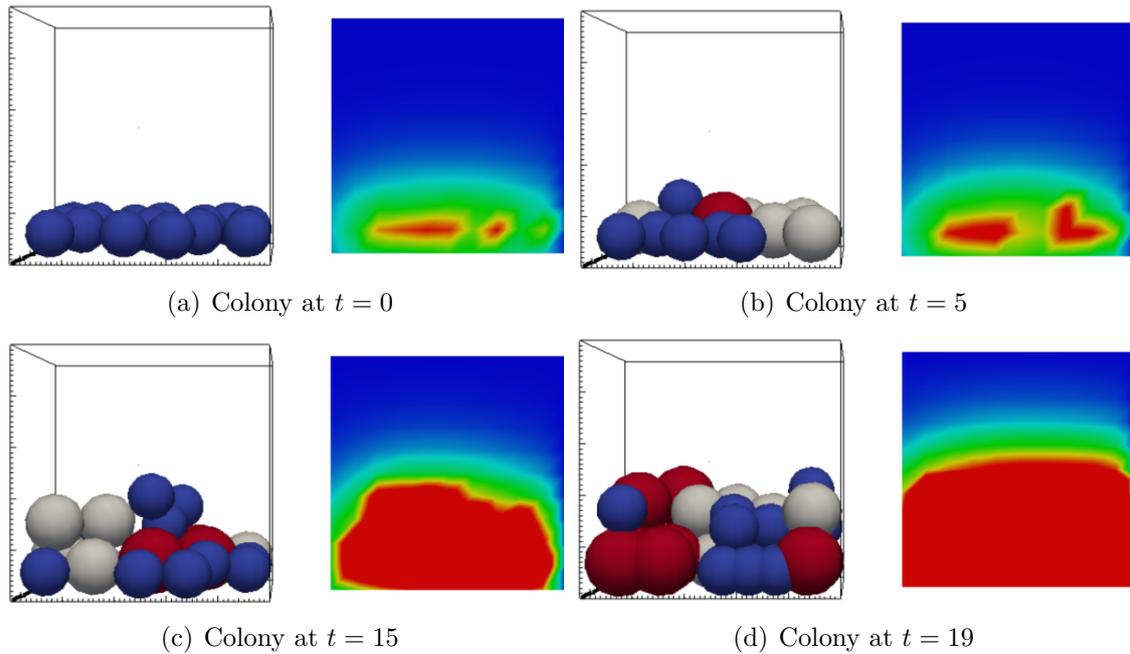


Figure 3.8: Same colony as in figure 3.7 with  $D = 7$ ,  $S_{on} = 10.0$ ,  $K = 18.0$  and  $\gamma = 7$  at different time steps  $t$ .

Figures 3.7 and 3.8 are a small scaled example of the computational domain at different time-steps  $t$ .  $S_{on} = 10.0$ ,  $K = 18.0$  and the diffusion constant  $D$  is set  $= \gamma$  so in the beginning turned the cell-cell communication is off. Diffusion constant  $D$  and degradation constant  $\gamma$  can not be  $= 0$  as they are part of denominator in equation 2.9 and because  $\lambda = \sqrt{\frac{D}{\gamma}}$  in equation 2.20, due to this only cells with the default *state* set "on" can grow. Until the concentration levels of  $S$  rise enough to turn on the other cells. In the Figures 3.7 and 3.8 it is the red zone of the right figure.

For the right side of the figures the whole cube was divided into smaller sub-cubes and the concentration in the sub-cubes is calculated by summing up all cells concentration through Equation (2.20) and their distance to the middle point of the sub-cube. This means sub-cubes can be inside of a cell and if this case happens the concentration  $S_i$  is added in equation (2.20) as if it is measured on the surface of the cell  $c_i$ . With the visualization program *Paraview* the cube was sliced at  $z = 4$ . Neighbors have influence on each other because are so close that they are touching.

### 3.3 Prisoner's Dilemma

The focus of this section is a revised version of the so-called *prisoner's dilemma*, *PD* [8], [6], [18]. It is a standard example of game theory (started in 1944) [19], [20] and has a wide range of application (Economics, Psychology, Social-science, Evolutionary Biology).

Table 3.2: Example of the payoff-matrix for the original prisoner's dilemma

	Prisoner B: cooperates	Prisoner B: defects
Prisoner A: cooperates	each serves 1 year	A: 3 years prison B : goes free
Prisoner A: defects	A: goes free B: 3 years prison	Each serves 2 years

Table 3.2 is the basic example of the *PD* [20]. Both prisoners cannot communicate and have no loyalty to each other. They are separated in two individual rooms. The game is explained to them and they will have no opportunity for revenge or reward outside the game. Cooperation means staying silent while defection means turning on the other. B can either cooperate or defect

If B cooperates, A should defect, because going free is better than serving 1 year. If B defects, A should also defect, because serving 2 years is better than serving 3 years.

Either way, A should defect. Parallel reasoning will show that B should defect. Therefore regardless of what the other decides, each prisoner gets a better result by betraying the other ("defecting"). The game shows cooperation is not beneficial for anyone and results in a social dilemma (as defecting is the logical answer but societies require cooperation).

The game was revised to apply to a broader range of examples (e.g Evolutionary Biology) by introducing a different payoff-matrix [18]. This means that the reward is now a benefit  $b$  and it introduces a cost of cooperation  $c$ , for instance by producing an extracellular molecule (invertase in budding yeast *Saccharomyces cerevisiae*[6],[21]), which benefits cooperators (producers) and defectors(non-producers).

Table 3.3: Evolutionary biologies payoff-matrix of the prisoner's dilemma

	other cooperates	other defects
payoff to cooperator	$b - c$	$-c$
payoff to defector	$b$	$0$

Even though defecting still is the dominant strategy for an individual, it can be shown in spatial PD games that spatial separation occurs and small islands of cooperators persist, but only if  $b > c$  [18]. A spatial *PD* game should be understood as a lattice of individuals that can only interact(play the game) with a small sub-population, for example, only their direct neighbors.

With this revised *PD* in the paper of Van Dyken et al. (2013) [6], it was shown that 2-dimensional spatial population expansion can promote the evolution of cooperation for  $b > c$ . They segregated the expansion process in two phases: Phase I (genetic demixing) and Phase II ("survival of the fastest"). These phases' interaction creates a force promoting high productivity strategies such as cooperation.

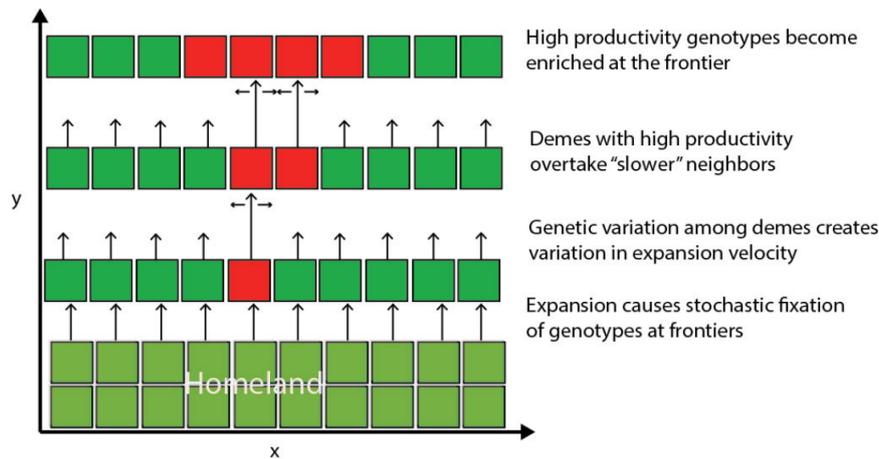


Figure 3.9: Population expanding in two spatial dimensions, with each site at coordinate  $(x,y)$  representing a subpopulation connected to nearest neighbors by dispersal according to Kimura’s stepping stone model [22]. A mixed homeland will eventually demix upon expansion into subpopulations. Cooperators (red) and defectors (green).

Spite and selfishness (defectors) actually cause a reduction in total reproductive output as these traits sweep to fixation [6]. This is known as the “Tragedy of the Commons”. However, at the frontier, expansion speed is determined by absolute fitness, generating a force of selection promoting genotypes that increase productivity. Because subpopulations with cooperators expand faster (they increase productivity) than subpopulations of defectors, cooperators become enriched at the frontier by overtaking neighboring defector’s placement.

The paper builds an experimental setup to test the effect of spatial expansion on defector/cooperator dynamics by using cooperative sucrose metabolism in haploid, vegetatively growing strains of the budding yeast *Saccharomyces cerevisiae*.

They genetically engineered two strains with different characteristics: one that produces invertase and one that does not. Yeast uses invertase in order to digest disaccharide sucrose, which can not easily be imported into the cell, by forming monosaccharides that are readily imported. In their strains [6], sucrose cannot be imported at all due to manipulation of the genes *MAL12* and *MAL22*. Because digestion occurs externally, invertase producers (“cooperators”) create a public good that is exploitable by non-producers (“defectors”), who gain a relative fitness advantage by not paying the fitness cost of production.

The defector strains still could grow in the absence of cooperators on the solid medium (YEP + 2% sucrose + 2% agar). The group assumes by consuming amino acids available in YEP, although growth is slower than for/with cooperators.

Next, the defector strain was engineered to be resistant to cycloheximide, a translation-inhibiting drug that limits growth by binding to ribosomal subunit Cyh2. This created a system that experimentally could impose a tunable “cost of cooperation” by varying the level of cycloheximide in the growth medium. Specifically, increasing the cycloheximide concentration slows the growth of cooperators but not the resistant defectors, leading to an increased “cost of cooperation”.

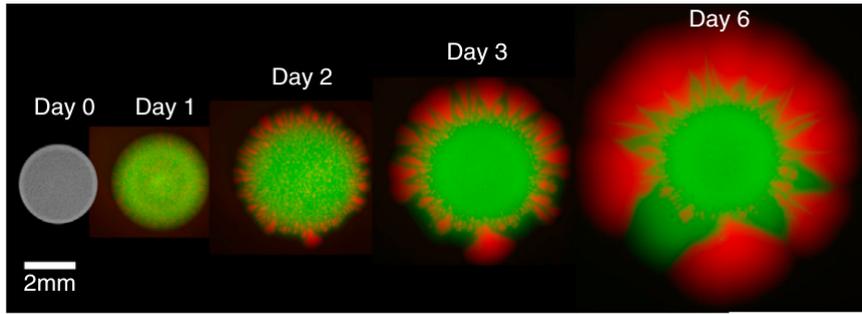


Figure 3.10: Growth of colony. Cooperators fluorescently labeled in red, defectors in green. Picture taken from [6]

Using my program I reimplemented the experimental setup by introducing two cell-types “cooperator” ( $c$ ) and “defector” ( $d$ ) and different growth-rates  $g_{rate,d_1}$ ,  $g_{rate,d_2}$  and  $g_{rate,c}$ , with  $g_{rate,d_1} < g_{rate,c} < g_{rate,d_2}$ . My cost of cooperation is defined as  $c = \frac{g_{rate,c}}{g_{rate,d_2}}$ . Instead of a concentric growing colony like in the experiment I chose a rectangular approach. Therefore the “homeland” from Figure 3.9 is now the initial randomly placed cells at the bottom of the figures.

The cell environment has been simplified by assuming that cell-colonies grow preferably in width before in height [1], [4]. This means every cell has perfect access to nutrients:  $O_2$  from above, disaccharide/YEP from the medium below. The assumption fits as long the experimental duration-time  $t$  is set smaller than the cells need to spread over the whole flat area of the computational domain. If cells far from the edge divide, internal pressure builds up that will promote a upward force that pushes cells higher.

The signal molecule  $S$  is now the concentration of invertase and the “defector” strain does not add anything to the calculation of equation (2.20), which means if a defector is too far away from cooperators it will grow with  $g_{rate,d_1}$ . Because invertase is a “public good”, a positive *feedback* is set.

Due to my assumption for the cell environment, interesting parameters for variation of the behavior of my model are in equation (2.20) so  $D$ ,  $\gamma$ ,  $S_{on}$  and  $K$ . As mentioned in 2. *Theoretical Foundations*  $\gamma$  and  $D$  basically counteract each other, therefore varying one gives a good impression of their influence. For  $K$  and  $S_{on}$  the same reasoning applies.

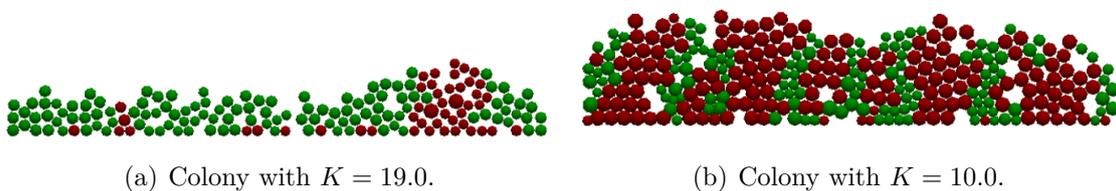


Figure 3.11: Colonies with  $D = 7$ ,  $S_{on} = 18.0$ ,  $\gamma = 7$  and  $L_x = 100$  but different  $K$ , at time  $t = 30$ . Cost of cooperation = 0.5. Cooperators labeled in red, defectors in green.

In Figure 3.11a, one can clearly see that my model can reproduce the spatial separation for cells growing on a flat medium. Note that the right side with a higher

concentration of cooperators grows faster and expands further than the left side with a high concentration of defectors which is consistent with the findings in [6]. My rectangular growth approach can be seen as small cut out section of a concentric growing culture.

For comparison in Figure 3.11b  $K$  is significantly smaller than  $S_{on}$  meaning cells stay more easily “on” and the expansion is faster but no genetic demixing emerges. This experiment has its faults because varying only  $K$  might give cooperators an advantage. Defectors will always be turned “off” if  $K > S_{on}$  and the neighbors’ influence is weaker. A single active cooperator cell will also turn “off”, which makes little sense in this scenario because each individual cell strives to survive. Therefore varying  $\gamma$  and with it the typical travel distance of the invertase because  $\lambda = \sqrt{\frac{D}{\gamma}}$  is more appropriate.

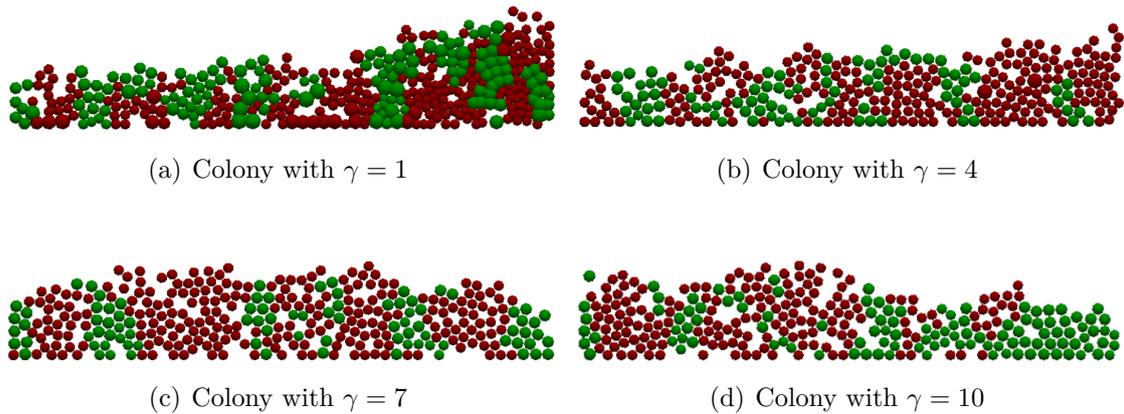


Figure 3.12: Colonies with  $D = 7$ ,  $S_{on} = 18.0$ ,  $L_x = 100$  and  $K = 18.0$  but different  $\gamma$ , at time  $t = 30$ . Cost of cooperation = 0.5. Cooperators labeled in red, defectors in green.

In Figure 3.12  $\gamma$  was varied while  $S_{on} = K$ . This crucially changes the setting since having one active cooperator benefits its surrounding close neighbors. This sole cooperator now stays “on”, because there is no reason for an individual to stop producing an essential molecule that it needs for survival.

One can see that even though the cooperators grow half as fast as the defectors, they start to dominate the cell colony for higher values of  $\gamma$ . This can be explained with the aforementioned effect that cooperators provide high productivity and this promotes them to be enriched at the frontier.

From Equation (2.20) one knows the bigger  $\gamma$  becomes the neighbors’ influence on the cells’ state-decision reduces. By comparing Figures 3.12(a-d) one can see this effect clearly for the defectors since they do not produce invertase: the higher  $\gamma$  is the defectors’  $S_i$  reduces and they are more likely grow with  $g_{rate,d_1}$  so they lose in expansion speed and become less dense and more scattered between cooperators, which grow with  $g_{rate,c}$  (with the exception of the right side in Figure 3.12d but that results from the stochastic nature of the initial cell-placement).

### 3.3.1 Discussion

It is important to note that two different simulation approaches come to a similar conclusion. My program runs with a realistic approach: concentration-gradients which the cell senses and then the internal regulatory mechanisms decide.

*PD* simulations are more general and simplify the problem by decision patterning<sup>3</sup>, which means every cell plays the game in every time-step with its neighbors and following set rules a cells decision is made(cooperate or not) and it reaps the accompanied benefit.

Using *PD* is useful to qualitatively show that, under the condition  $b > c$ , cooperation promotes spatial separation of non-cooperators and cooperators. For such cases *PD* simulation is better than my model because less computational power and time is needed for the simulations, but at a cost of losing information of the underlying complexity of the bacterial colony.

Next is the question of what would happen if the experiment duration  $t$  is bigger than the spreading time to occupy the whole flat area of the computational domain. A biofilm is bound to grow upward [1] and a concentration gradient will start to develop [2] e.g for  $O_2$  from above and nutrients below: the highest cells have a high  $O_2$ - and low nutrient-concentration while the lowest cells have high nutrient- but low  $O_2$ -concentration. This can cause a response of the biofilm by cells physiologically differentiating depending of their spatial location[2] and will be further discussed in Chapter 4.*Outlook*.

To conclude, my program is consistent with the findings in VanDyke et al. 2013 [6]. It also shows that the diffusion distance of the public good molecule (so availability) is an important factor for spatial separation.

### 3.3.2 Growth Model Idea

This experimental set-up can also be reinterpreted for the growth of biofilms.

Biofilms are a multicellular community characterized by a robust extracellular matrix of proteins, amyloid fibres, exopolysaccharides and extracellular DNA (here abbreviated to ExMat). This matrix has two interesting characteristics for competition between different strains and the growth of a biofilm. First the volume of ExMat that one cell can produce is bigger than if the cell grows and divides, but at a cost of cooperation (in my model it is implemented with a slower growing rate). This means the producing strain can advance even faster in space as their volume grows faster.

Secondly the ExMat of one strain can possibly kill cells of other strains [24].

Autoinducers are recognized as molecules that bacteria use for communication [4],[13] and their concentration level is known to play a key role in regulating the production of the extracellular matrix [1], [2].

The concentration level of autoinducers indicate to bacteria cells the current cell density of its' surroundings. Changes in cell density can result in regulation of gene expression in response to it, a phenomenon defining quorum sensing [13]. Instead of assuming the signal molecule is a public good like invertase, it is now an autoinducer for the bacterial "cooperator" strain.

This changes the experiment. In the beginning I now have two planktonic bacterial strains that are attached to a surface (one that produces an autoinducer and if the

---

<sup>3</sup>for more and detailed information on this subject with biological reference [18],[23]

condition is right ExMat and one that does not). The non-producer strain has a growth rate  $g_{rate,non}$  while the producing strain has two. One for being it the “off” state denoted  $g_{rate,off}$  and one for begin “on” denoted ( $g_{rate,on}$ ). If the producing strain is “on” it produces the ExMat which has a negative effect on the growth rate. The relationship for the rates is as follows:  $g_{rate,non} \approx g_{rate,off} = 2 \cdot g_{rate,on}$ . ExMat production is implemented by adding a sphere of the radius  $r = 0.4$  to the system at a random location next to the producing cell. The biofilm grows in height so gravitational pull is added in the force calculation.

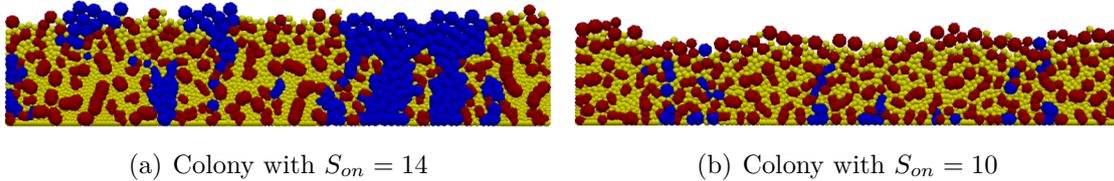


Figure 3.13: Simulation of growth model with different  $S_{on}$ . Yellow the ExMat, blue non-producers and red ExMat producing cells

In Figure 3.13 simulations with different  $S_{on}$  are shown producing ExMat seems to give the cells a spatial advantage because of their larger volume. By comparing the simulations it seems that a reasonable low  $S_{on}$  concentration (qualitative testing showed not lower than  $S_{on} = 10$ ) leads to even better results in spatial occupancy. This agrees with my intuition because when the producing cells are more dense before producing ExMat it becomes more probable that cells of the same strain are pushed towards the growing frontier.

The experiment was done in a similar fashion by Xavier et al. in 2007 [24] using their model from 2005 [10]. They show that with an  $O_2$ -gradient ExMat-producing cells have a competitive advantage for nutrient-access similar to plants competition for optimal light exposure.

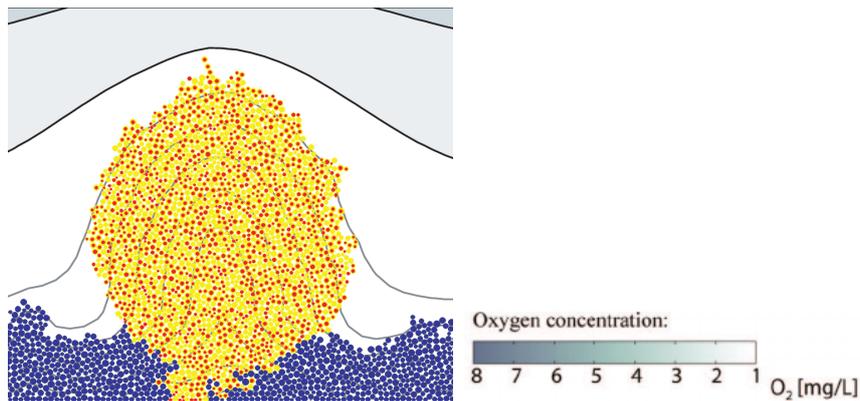


Figure 3.14: Colony with  $O_2$ -gradient after  $t=20$  days. Yellow the ExMat, blue non-producers and red ExMat producing cells.

Previously mentioned in 3.3.1 *Discussion* and now clearly shown in Figure 3.14 concentration gradients for nutrients and  $O_2$  have a major impact on the biofilm growth and spatial expansion. By comparing Figure 3.13 to Figure 3.14 one clearly sees the current limit of my program and its possible predictions. I could not implement a satisfying method for the calculation of molecule gradients other than for

the signal molecule  $S$ . This will be further discussed in in Chapter 4. *Conclusions and Outlook*.

### 3.4 Mutualism

*Mutualism* can be understood as interspecies cooperation [8], [7], [25]. This means that species A produces something or has a waste product that benefits species B and vice versa. In the paper of Momeni et al. from 2012 [7] they show both theoretically and experimentally, using the colony expansion of non-mating but cross-feeding strains of budding yeast *Saccharomyces cerevisiae* (the strains exchange amino acids) as a model system, that through strong mutualism genetic demixing of the strains does not occur in the spatial expansion process. The cross-feeding is genetically engineered so that strain  $Lys^+A^-$  overproduces lysine ( $L^+$ ) and leaks it into the medium, but cannot produce adenine ( $A^-$ ) and its partner strain  $Lys^-A^+$ .

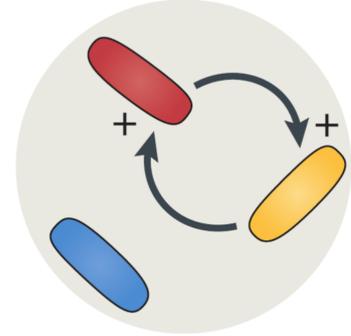


Figure 3.15: Social phenotype of *Mutualism*

There is also a non-cooperative wild-typ for comparison. overproduces and leaks adenine ( $A^+$ ) but cannot produce lysine ( $Lys^-$ ).

To recreate such behavior a new parameter is introduced in my model called mutualfactor  $\mu$ . If  $j$  is a mutualist it is multiplied in Equation (2.20) to the added term. In my simulation the mutualists are “cooperator” strains for themselves but also benefit from the concentration level of the other strains’ production to the extend of  $\mu \in [0, 1]$ .

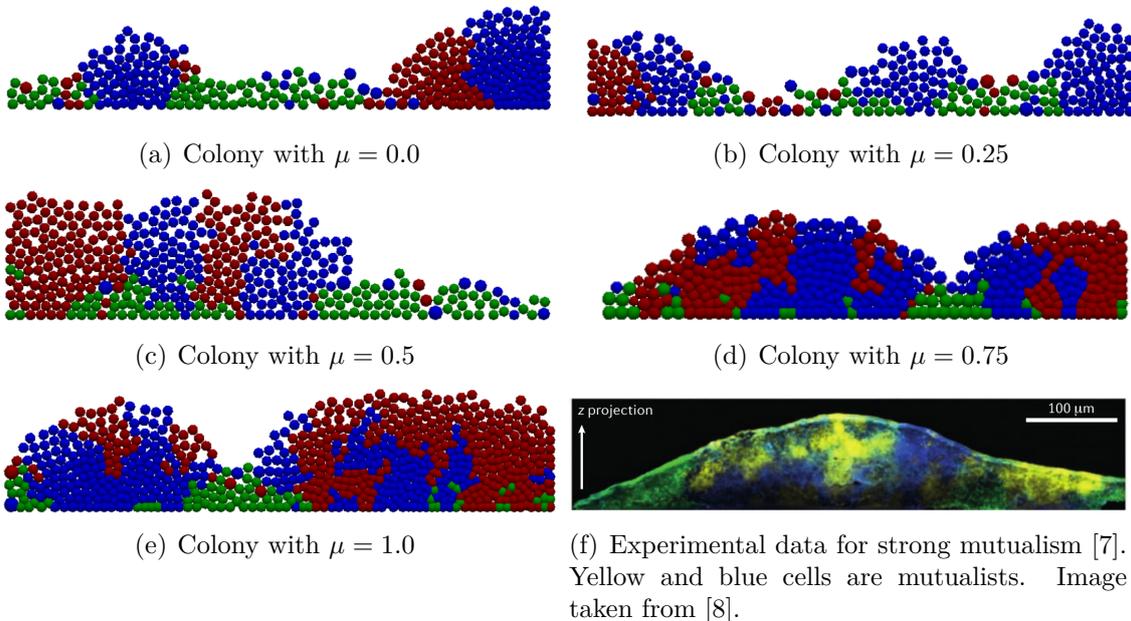


Figure 3.16: Simulations with different mutualfactor  $\mu$  and the cell ratio  $R = 1:1:1$ . Simulations at  $t = 19$  with parameters  $D = 7$ ,  $S_{on} = 16.0$ ,  $K = 18.0$ ,  $\gamma = 7$ . Red and blue colored cells are mutualists, while green cells are non-mutualists (“lone wolfs”).

Figures 3.16 (a-e) show my simulations for the two mutualist strains plus a cell lineage of non-producers. One strain ( $Lys^+A^-$ ) senses the concentration of  $A$  multiplied with the factor  $\mu$  and its own concentration, while other strain ( $Lys^-A^+$ ) senses the concentration of  $Lys$  multiplied with the factor  $\mu$  and its own.

One can see in these figures how varying  $\mu$  changes the spatial structure of the biofilm. The bigger the value of  $\mu$  becomes the colony of the mutualists becomes more intermixed, while the non-cooperator strain is excluded.

Weak mutualism is negligible relative to competition between the strains and genetic demixing happens with resulting spatial separation (see Figures 3.15(a,b)). My model is able to depict by straying from the experiment because the mutualist strains are implemented as cooperators for themselves, even though the overproduction of one amino and not producing another is not increasing productivity of the strain overall.

Figure 3.16(f) shows experimental data for strong mutualism (image taken from the paper of Momeni et al. in 2012 [7]). The simulations in Figure 3.16(d,e) have a striking resemblance to the experimental data.

### 3.4.1 Discussion

The population in Figure 3.15e only looks like to be smaller than Figure 3.15d. There is a scaling problem caused not only by my lack of picture manipulation skills, but also comes from the force calculation method.

For every time step the force calculation calculates the next position of each cell individually, and ideally places the cell shove radius distanced from all the neighboring cells.

This is the source of the problem because running the method just once will not result in perfectly distanced cells<sup>4</sup>. In order to solve this general problem of my implementation I used the explicit Euler method.

It is a numeric solution for initial value problem of simple ODEs:

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0.$$

Choose a value  $h$  for the size of every step and set  $t_n = t_0 + nh$ .

Now, one step of the Euler method from  $t_n$  to  $t_{n+1} = t_n + h$  is

$y_{n+1} = y_n + hf(t_n, y_n)$ . The calculated value of  $y_n$  is an approximation of the exact solution to the ODE at time  $t_n$ :  $y_n \approx y(t_n)$ . The smaller the step size  $h$  is chosen, one needs more calculation time but the approximated values are closer to the exact solution. In Figure 3.17 one can see a simple example.

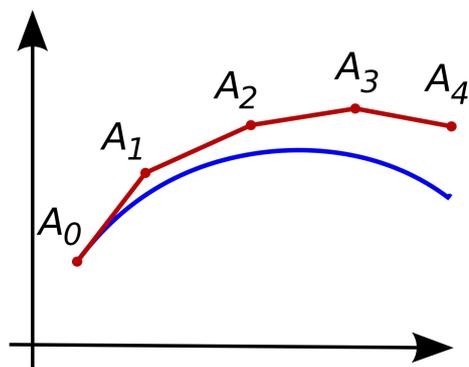


Figure 3.17: Euler method in red for unknown blue curve.

Even though I am not solving an actual ODE, this approximation method can be used.

<sup>4</sup>when one cell is moved at a time then all the others have to be moved accordingly to it, and then the next cell is moved and changes the arrangement so that the cell from before should also be moved again and so on...

- (1)  $y_n$  is one of the spatial coordinates (x,y,z).
- (2)  $f(t_n, y_n)$  is the force calculated for this axis as a distance the coordinate has to be displaced respectively denoted  $f_x, f_y, f_z$ .
- (3)  $h$  is factor to make the step size smaller  $h = \frac{1}{10 \cdot \max(f_x, f_y, f_z)}$ .

Now by iteration the force calculation method and then changing the spatial coordinates of the cells in this manner will approximate the ideal solution of shove radius distanced cells (more iterations  $\hat{=}$  smaller step size  $\hat{=}$  closer to exact solution but the program takes longer). Therefore this calculation is actually the most time consuming method in my program. Normally I have my program set to 40 iterations per actual time step. There is only a slight difference in the how stacked the cells are if the iteration number is greater than 40, but the runtime is getting higher (with 40 iterations a simulation took 40 minutes; with 100 iterations 2 hours). While the runtime for simulations with a low iteration number take mere seconds, it is hard to make any meaningful observations from the visualization of the simulated biofilm because all the cells are so stacked.

The observed simulations with high  $\mu$  values have striking resemblance to the experimental data of the group [7]. The group also observed genetic demixing if the mutualism is tuned down like in my model with low  $\mu$  values. Both observations can be explained with the same reasoning as discussed in Section 3.3 *Prisoner's Dilemma*.

High productivity is favored in the growing frontier so the non-cooperators are excluded. Depending on the  $\mu$  value cells from the other strain become either excluded ( $0 \approx \mu \ll 1$ ) or included ( $0.5 \lesssim \mu \approx 1$ ) in the growing frontier.

The lower the  $\mu$ -factor becomes the strains seem to be more like non-cooperators or competition and a clear boundary between the cells becomes visible(see Figure 3.16(a,b)).

More testing is needed for a quantitative estimation range of values for the specific social phenotype(cooperate and intermixed or demixed and competitors) and the influence of other parameters should be tested because in the simulation for  $\mu = 1$  the different mutualist cell strains are treated identical.

To conclude this section my model qualitatively can recreate existing experimental and theoretical predictions [25], [7] for mutualism.

## 3.5 Chemical Warfare

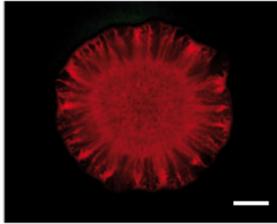
Biofilms are characterized by a complex community structure, which is shaped by competition between strains for resources such as nutrients and space [4],[8], different forms of mutualism, cooperation and cheating [4], and interference/antagonism through the production of toxins [5], [26]. This section is about toxins in microbial colonies and biofilms.

In the paper of Weber et al. 2014 [5] three E.coli strains (toxin sensitive(denoted by S), toxin resistant(denoted by R) and toxin producing strain (denoted by T)) were genetically engineered to determine if longterm coexistence is possible and what parameters influence this coexistence.

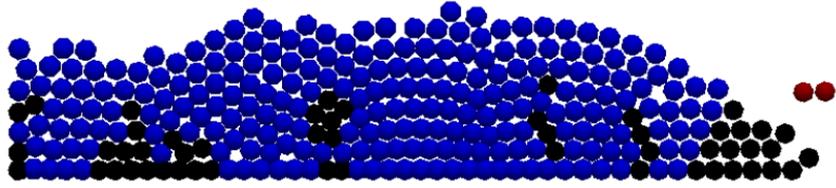
They designed three different ecological scenarios by altering strain growth rates through the expression of the fluorescent protein mCherry. A cyclic scenario with  $g_{rate,S} > g_{rate,R} > g_{rate,T}$ , a hierarchical scenario with  $g_{rate,R} > g_{rate,S} \geq g_{rate,T}$  and a intermediate scenario with  $g_{rate,S} \approx g_{rate,R} > g_{rate,T}$ . The ratio of the strains is also

altered. These changes together try to depict different survival strategies that exist in microbial colonies.

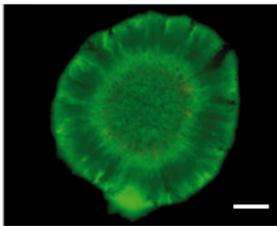
The grow rate relations of the scenarios and cell types are easily translated into my program. The toxin becomes now the signal molecule and it is set to a deadly toxic range of  $r_{tox} = 2 * R_{toxic Cell}$  [5] so that means if a sensitive cell (S) is exposed to that concentration level it dies.



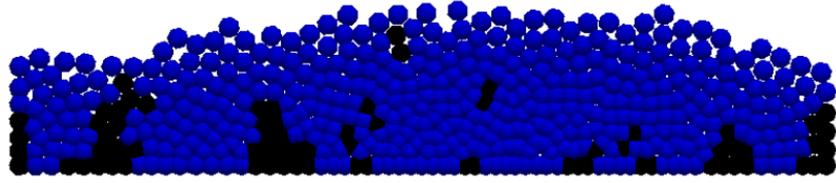
(a) Colony in cyclic scenario (S/R/T)



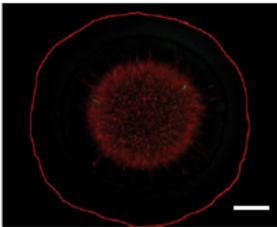
(b) Colony in cyclic scenario (S/R/T)



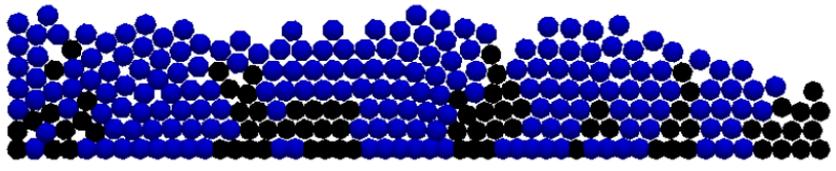
(c) Colony in hierarchical scenario (S/R/T)



(d) Colony in hierarchical scenario (S/R/T)



(e) Colony in intermediate scenario (S/R/T)



(f) Colony in intermediate scenario (S/R/T)

Figure 3.18: The three different scenarios with and the ratio  $R = (S/R/T) = 1:1:1$ . Left side experimental data from Weber et al. 2014 [5]. Right side simulations at  $t = 19$  with  $D = 7$ ,  $S_{on} = 10.0$ ,  $\gamma = 7$  and deadly toxin range  $r_{tox} = 2 * R_{toxic Cell}$ .

In Figure 3.18 the ratio of the different E.coli strains is the same. The computational domain is viewed from above that means the colony advances on flat ground. Therefore no gravitational force is added in the force calculation. This results in Figure 3.18(b) right side where two sensitive cells remain to this time step. On the right one sees the experimental data of the Weber group [5] obtained after 48h incubation time, in all three scenarios the sensitive strain becomes extinct. By comparing the figures for the different scenarios one can assume that strains with a higher growth rate are able to expand quicker and occupy more space.

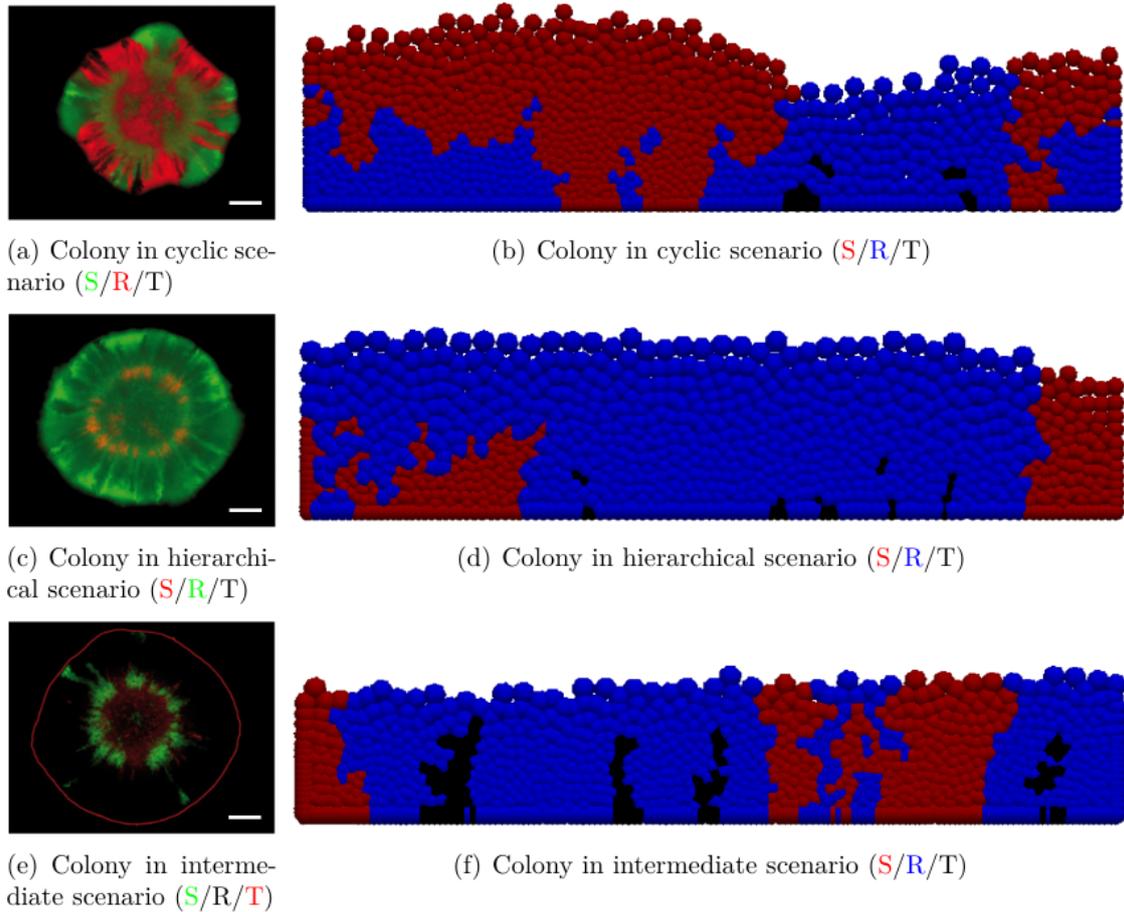


Figure 3.19: The three different scenarios with ratio  $R=(S/R/T)=1:1:0.1$ . Left side experimental data from Weber et al. 2014 [5]. Right side simulations at  $t = 19$  with  $D = 7$ ,  $S_{on} = 10.0$ ,  $\gamma = 7$  and deadly toxin range  $r_{tox} = 2 * R_{toxic Cell}$ .

In the second part of the experiment the ratio of the cells is changed to  $(S/R/T)=(1:1:0.1)$ . In Figure 3.19 one can see the resulting structure of the colonies. By reducing the number of toxic cells the probability of death caused by the toxin also lowers therefore it is possible for the sensitive strain to expand (see Figure 3.19). By comparing the scenarios differences in the colony structure are evident. As the only growth rates differ between the scenarios it is reasonable to assume that growth rates have a substantial impact on the forming of a colony structure.

The R strain clearly dominates in more hierarchical scenarios ( Figure 3.18 and 3.19(c-f)). Therefore, toxin resistance seems to be a more effective survival strategy than either rapid growth or toxin production in case of a hierarchical competition network.

### 3.5.1 Discussion

The Weber group also did a modeling approach that is consistent with their experimental findings [5]. It is a coarse-grain model that did not treat individual bacterial cells as agents. Instead, each agent was a colonized lattice site which then represented the bacterial strain that locally dominated this area (patch) of the colony. Therefore their model can be seen as a mesoscopic description of the bacterial ex-

pansion process, while my modeling approach is more microscopic. For a big cell population size the coarse-grain model is the better choice as it depicts a simplified representation of the colony but as general aspects are conserved, they can be studied on a larger scale.

Whether a bacterial strain manages to survive a range expansion and to populate the colony's expanding front seems to depend on two aspects: first, on its ability to form initial clusters from which outward expansion may emerge; second, on the stability of the arising sectors of high cell strain number to the annihilation of neighboring cells (see the differences in the S strain population). Both of these aspects are subjected to random genetic drift and may prevent the establishment of stable cell strain compartments in a simulation (see the S strain in Figure 3.18).

One can see not only strength of numbers, the growth rate differences, or competition strategy alone determine the success of a cell strains long term survival. The right balance between these factors has to be struck. Therefore modeling approaches that can recreate current knowledge and than possibly make predictions for on still unknown factors are of importance.

To conclude, my program's output is consistent with the experimental and modeling data from Weber et al. [5] even though the modeling approaches are on different scales.

In Appendix A one can see a simulation with  $t \in [0, 19]$ .

## 4 Conclusions and Outlook

In this work, I presented the versatility of my program. It can qualitatively recreate many of the current findings in the field of microbial community life.

With slight changes and a creative mind this modeling approach can serve as a handy tool in ones first steps to analyze new social phenotypes in biofilms or other macrocolonies.

Regrettably there was no time to implement an adequate method to build concentration gradients for environmental molecules (like  $O_2$  or nutrient uptake). This would not change my current findings but consequently narrows down the range of conclusions and predictions one can make with my model.

For the future of my program there are two possible steps.

The first step is finding more natural instances of cell-cell communication which my current program can possibly tackle e.g. a simplified version yeast mating process. The second step is to thoroughly recreate the natural habitat of biofilms. As already mentioned in Section 3.3 concentration gradients of essential environmental molecules are missing from my program and these have a significant influence on the biofilm structure [2] , competition [24] and the growth [24].

**Mating yeast:** Wolfgang Giese suggested the idea to implement the mating process of budding yeast *Saccharomyces cerevisiae* [27]. Now follows a basic description of the mating process and yeast life cycle:

Yeast cells can be found in nature in a diploid as well as a haploid life cycle. Normally in their diploid phase, one diploid yeast cells can form into four haploid spores by meiosis when nutrients are limited (in the beginning encased by a solid and protected structure called ascus).

In the haploid phase two mating types ( $MATa$  or  $MAT\alpha$ ) exist, which can either grow and divide or enter the sexual life cycle. These mating types communicate via pheromones that are secreted into the extracellular medium. The pheromones can be sensed by cell-surface receptors of the complementary mating type. In order to mate, yeast cells arrest their cell cycle if the concentration level of the complementary mating factor is big enough, then typically cells grow towards the mating partner, since yeast can not actively move. Due to this growth the cell shape changes during mating. The shape is commonly referred to as a shmoo. These mating partners of opposite type can form a zygote and this zygote enters the diploid life cycle again.

The mating process is precisely coordinated and involves complex signaling and communication principles that have been conserved throughout the evolution of eukaryotic cells.

I implemented a simplified version of this mating process. By assigning two cell types which decide if they enter the sexual life cycle (state “on”) or stay haploid (state “off”). This decision is made by comparing the concentration level of the com-

plementary mating types' pheromone, at the cells spatial coordinates, to a threshold  $K$ . The two states have distinct characteristics. In the “on” state cells do not grow anymore and the program checks if a complementary cell type which is also “on” is close enough for them to fuse (the fusion range is set to  $1.2 \cdot R_{Cell}$  (simplified idea for shmooing)). In the “off” state cells grow and divide.

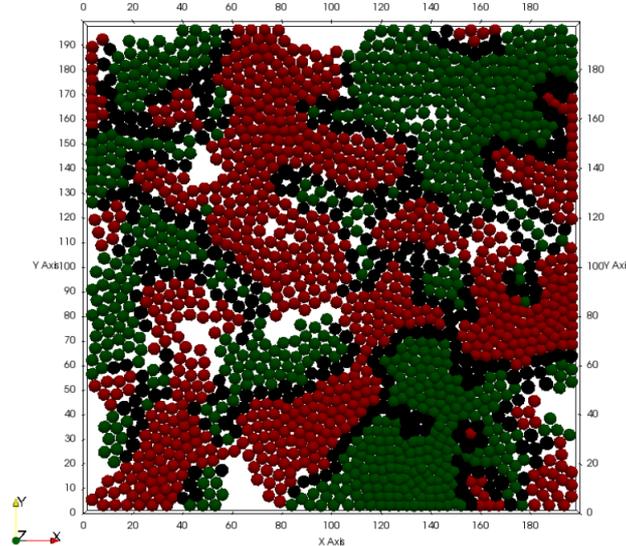


Figure 4.20: Simulation of yeast mating at  $t = 19$  with  $K = S_{on} \approx 0.27$ ,  $D_a \approx D_\alpha = 360 \frac{\mu m^2}{s}$  and  $\gamma \ll D$ . Diploid cells black, haploid  $MATa$ -cells in green and haploid  $MAT\alpha$ -cells in red

In Figure 4.20 one can see the current output of my implementation. Diploid as well as haploid yeast cells are capable of mitosis, giving rise to two new cells in a process called budding. Both(diploid and haploid) cells exhibit a preferred budding direction. Usually, diploid cells follow a bipolar(the bud creation site changes e.g. once on the right side than on the left side) and haploid cells an axial budding pattern(e.g. keep building the bud on the right side). This is not implemented up to date but is the next step along the way. After this the introduction of Bar1 might follow. Bar1 is a molecule produced by  $MATa$  cells that cleaves the  $\alpha$ -factor [28] and that Bar1 promotes higher growth rates in subpopulations of only  $MATa$  cells [27].

The significance of mating yeast is that it is a well orchestrated process. It exhibits a lot of different elements of communication and signal-transmission processes which are also important for both unicellular organisms and cell tissues. So therefore better understanding and describing this process may contribute to solving fundamental questions of cell biology.

**Implement environment:** To implement an environment my program needs a solid method that can calculate the concentration of these environmental molecules for every time step at every position or at least for every cell. This method will basically depict the consumption of the cells. Based on the concentration levels different stress is on each cell and this can lead to differentiation of the cells' physiology caused by its spatial location in the biofilm [2].

To solve such diffusion equations, most groups [7],[10] use a finite difference time-domain

method (Crank, 1979 [29]) with no-flow ( $\frac{\partial S}{\partial y}$ ) boundary conditions applied to the bottom surfaces of the simulation domain, a constant influx from above and periodic boundary conditions applied to the vertical sides of the domain. For implementation of this approach the time given for my bachelor thesis was too short. I have to point out that Equation 2.20 for the signal molecule  $S$  is a simplified version of this method.

It is of importance to me to mention that in all the presented papers the bacteria or yeast strains are genetically altered to fit certain assumptions, of course this is needed to be actually able to validate ones hypotheses and prove the theoretical approach but makes the experiment artificial.

The future of the field I see in more complex experimental setups with a mix of the wildtypes of bacteria which then possibly make up the inner workings of a human mouth or the digestive tract. A bacteria *Helicobacter pylori* infection for instance is known to increase the risk of gastric carcinoma [30]. Begin able to model the interactions would possibly help prevent the infection all together or give a better insight.

All in all my view on bacteria as single celled organisms has changed. The complex structure and multicellular behavior bacteria can exhibit in biofilms and the new properties (like resistances) they gain is material enough for research to continue for decades.

# Bibliography

- [1] George O’Toole, Heidi B Kaplan and Roberto Kolter, Biofilm formation as microbial development. *Annual Reviews in Microbiology* **54** (1), (2000), 49–79.
- [2] Diego O Serra and Regine Hengge, Stress responses go three dimensional—the spatial order of physiological differentiation in bacterial macrocolony biofilms. *Environmental microbiology* **16** (6), (2014), 1455–1471.
- [3] Luanne Hall-Stoodley, J William Costerton and Paul Stoodley, Bacterial biofilms: from the natural environment to infectious diseases. *Nature reviews microbiology* **2** (2), (2004), 95–108.
- [4] Bonnie L Bassler and Richard Losick, Bacterially speaking. *Cell* **125** (2), (2006), 237–246.
- [5] Markus F Weber, Gabriele Poxleitner, Elke Heibisch, Erwin Frey and Madeleine Opitz, Chemical warfare and survival strategies in bacterial range expansions. *Journal of The Royal Society Interface* **11** (96), (2014), 20140172.
- [6] J David Van Dyken, Melanie JI Müller, Keenan ML Mack and Michael M Desai, Spatial population expansion promotes the evolution of cooperation in an experimental prisoner’s dilemma. *Current Biology* **23** (10), (2013), 919–923.
- [7] Babak Momeni, Kristen A Briley, Matthew W Fields and Wenying Shou, Strong inter-population cooperation leads to partner intermixing in microbial communities. *eLife* **2**, (2013), e00230.
- [8] Carey D Nadell, Knut Drescher and Kevin R Foster, Spatial structure, cooperation and competition in biofilms. *Nature Reviews Microbiology* .
- [9] Théo Maire and Hyun Youk, Molecular-level tuning of cellular autonomy controls the collective behaviors of cell populations. *Cell systems* **1** (5), (2015), 349–360.
- [10] Joao B Xavier, Cristian Picioreanu and Mark Van Loosdrecht, A framework for multidimensional modelling of activity and structure of multispecies biofilms. *Environmental microbiology* **7** (8), (2005), 1085–1103.
- [11] Thomas Gregor, David W Tank, Eric F Wieschaus and William Bialek, Probing the limits to positional information. *Cell* **130** (1), (2007), 153–164.
- [12] Francis Crick, Diffusion in Embryogenesis. *Nature* **225**, (1970), 421.

- [13] Michael G Surette, Melissa B Miller and Bonnie L Bassler, Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production. *Proceedings of the National Academy of Sciences* **96** (4), (1999), 1639–1644.
- [14] Pasquale Sansone, Gianluca Storci, Simona Tavolari, Tiziana Guarnieri, Catia Giovannini, Mario Taffurelli, Claudio Ceccarelli, Donatella Santini, Paola Paterini, Kenneth B Marcu *et al.*, IL-6 triggers malignant features in mammospheres from human ductal breast carcinoma and normal mammary gland. *The Journal of clinical investigation* **117** (12), (2007), 3988–4002.
- [15] Cristian Picioreanu, Jan-Ulrich Kreft and Mark CM van Loosdrecht, Particle-based multidimensional multispecies biofilm model. *Applied and environmental microbiology* **70** (5), (2004), 3024–3040.
- [16] Eric W. Weisstein, "Sphere Point Picking." . Technical report, A Wolfram Web Resource. (2016).
- [17] Mervin E Muller, A note on a method for generating points uniformly on  $n$ -dimensional spheres. *Communications of the ACM* **2** (4), (1959), 19–20.
- [18] Michael Doebeli and Christoph Hauert, Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecology Letters* **8** (7), (2005), 748–766.
- [19] William Poundstone, Prisoner's Dilemma: John von Neuman, Game Theory, and the Puzzle of the Bomb (1992).
- [20] J von Neumann, Oskar Morgenstern *et al.*, Theory of games and economic behavior, volume 60. Princeton university press Princeton (1944).
- [21] Oskar Hallatschek, Pascal Hersen, Sharad Ramanathan and David R Nelson, Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences* **104** (50), (2007), 19926–19930.
- [22] Motoo Kimura and George H Weiss, The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49** (4), (1964), 561.
- [23] Robert Axelrod, The evolution of cooperation (1984).
- [24] Joao B Xavier and Kevin R Foster, Cooperation and conflict in microbial biofilms. *Proceedings of the National Academy of Sciences* **104** (3), (2007), 876–881.
- [25] Melanie JI Müller, Beverly I Neugeboren, David R Nelson and Andrew W Murray, Genetic drift opposes mutualism during spatial population expansion. *Proceedings of the National Academy of Sciences* **111** (3), (2014), 1037–1042.
- [26] Benjamin Kerr, Margaret A Riley, Marcus W Feldman and Brendan JM Bohannan, Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature* **418** (6894), (2002), 171–174.

- [27] Wolfgang Giese, The choreography of yeast mating - modelling intercellular communication, cell polarity and morphogenesis. Ph.D. thesis, Humboldt Universitaet (2016).
- [28] Naama Barkai, Mark D Rose and Ned S Wingreen, Protease helps yeast find mating partners. *Nature* **396** (6710), (1998), 422–423.
- [29] John Crank, The mathematics of diffusion. Oxford university press (1979).
- [30] Julie Parsonnet, Gary D Friedman, Daniel P Vandersteen, Yuan Chang, Joseph H Vogelmann, Norman Orentreich and Richard K Sibley, Helicobacter pylori infection and the risk of gastric carcinoma. *New England Journal of Medicine* **325** (16), (1991), 1127–1131.

# Appendix A

**Methods:** For Visualization I used Paraview Version 4.4 and the python library of VTK-tools. The library turns lists into data sets that Paraview can interpret. The code is written in python language 2.7.

I will build a repository on GitHub. There one can find the programs newest implementation and all the different implementations used for the simulations. The link will be <https://github.com/Nexia23/Bachelor>.

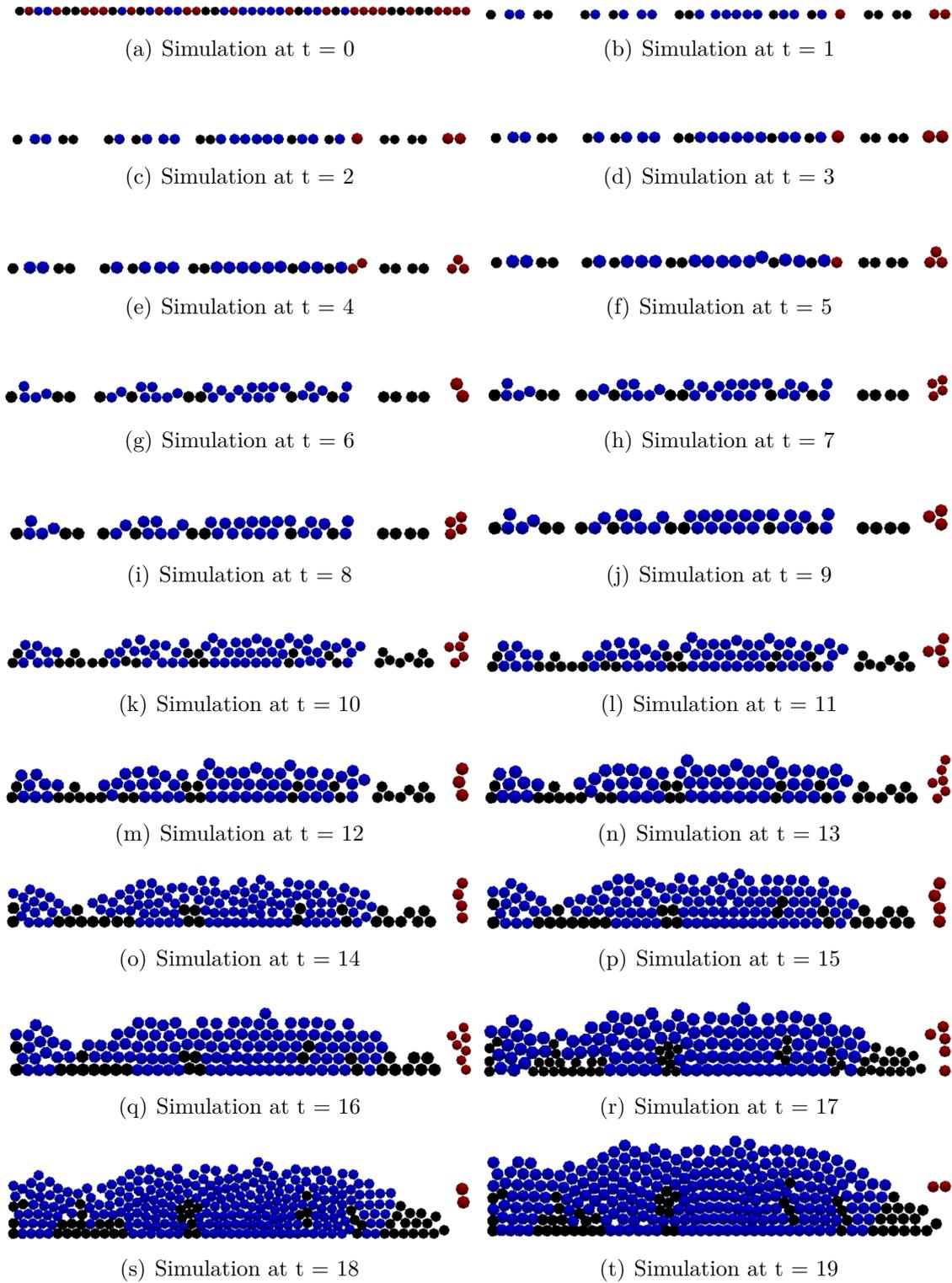


Figure 21: Simulation of cyclic scenario with ratio  $R = (S/R/T) = (1:1:1)$ . With parameters  $D = 7$ ,  $S_{on} = 10.0$ ,  $\gamma = 7$  and deadly toxin range  $r_{tox} = 2 * R_{toxic Cell}$ .

# Selbstständigkeitserklärung

Hiermit versichere ich, Maxim Karnetzki, dass ich meine Abschlussarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

.....  
(Datum und Unterschrift)