Agent Based Modelling of Bacterial Culture Formation

Catherine Terrey

3A Honours Co-operative Mathematical Physics

University of Waterloo

Faculty of Science

Department of Physics and Astronomy

Agent Based Modelling of Bacterial Culture Formation

University of Waterloo

200 University Avenue West

Prepared by Catherine Terrey, 20716850

3A Honours Co-operative Mathematical Physics

Submitted on January 5, 2021

98 Emma Street

Guelph, Ontario

N1E 1T8

January 5, 2021

Brian McNamara, Department of Physics and Astronomy

Faculty of Science, University of Waterloo

200 University Avenue West, Waterloo, Ontario

N2L 3G1

Dear Dr. McNamara,

This report, entitled "Agent Based Modelling of Bacterial Culture Formation" was prepared as my Work Term Report for the University of Waterloo. This is my second work term report. The purpose of this report is to effectively model bacterial cultures using an agent based algorithm such that the simulations are similar to microscope data.

The research done in this report took place at the University of Waterloo under the Department of Applied Mathematics in Ingalls Quantitative Cell Biology Lab. Dr. Brian Ingalls lab focuses on research in mathematical biology, specifically systems biology as well as quantitative microbiology. This involves developing and simulating models of systems on a cellular scale such as chemical reaction networks or bacterial interactions.

This report was written entirely by me and has not received any previous academic credit at this or any other institution. I would like to thank Dr. Ingalls and Sara Khorasani for providing me with valuable support and data resources use in the preparation of this report. I received no other assistance.

Sincerely,

Catherine Terrey

20716850

Table of Contents

List of Figures

Abstract

Agent based models of bacterial colonies are of fundamental importance to the field of systems biology. These simulations allow for physically realistic models of forces and collisions between cells to be replicated *in silico*. An important milestone in the field would be to develop robust and widely applicable algorithms for fitting these physics explicit models. This could in principle be done by comparing microscope images taken at frequent time intervals with simulation image data. In this report, the focus is on developing physics based models of cell interactions and implementing them numerically. Various simulations including each force are compared, and the results are used to suggest improvements for the model. The software used for this report is based on BSim2.0, an agent based modeling software developed by a team at the University of Bristol. The goal is to extend the program to make the simulation output more similar to the images taken in the lab. This is done by considering new forces and interactions between cells that were not present in BSim2.0.

1 Introduction

Different species of bacteria grow in unique ways to form a variety of shapes and patterns. Replicating these configurations of cells with a mathematical model has tremendous importance to the field of systems and synthetic biology. In the past, many researchers have used different modelling techniques to achieve this goal. This report will focus on how to solve this problem using agent based modelling.

An agent based model is a method of simulating a population of cells, which we call *agents*. In an agent based model, each cell is described by a set of parameters unique to that cell. This uniqueness is what makes agent based modelling exhibit more accuracy than models which assume uniformity. In order to simulate a population of cells, each cell is placed at a location on a grid. At each time step in the simulation, cells are allowed to interact with each other and their surroundings. [1, 2].

In an agent based model the system of bacteria will change at each time step according to rules prescribed by the programmer. These rules describe the dynamics of growth, division, interaction forces and more.

An agent based model is called physics explicit if it is simulating a system based on Newton's laws of motion. This means that cells will experience realistic forces which causes them to interact with their surroundings and move about in the same way they would in real life.

The key benefit to using an agent based model is that it gives the ability to resolve the behaviour of individual cells. This property allows us to predict the spatial configuration of every cell, and to accurately model the shape of the entire colony. In addition to the high spatial resolution of agent based models, this technique allows the modeller to introduce complicated interactions between individual cells that can not be observed in a homogenized model.

A downside to using an agent based model is the computational complexity of the problem. One must update the information of every cell individually, and keep track of every interaction between cells. In a typical colony, the number of bacterial cells can be as large as 10^{10} [1], which is not feasible for a computer to handle. In order to remediate this problem, only smaller colonies are considered when using this approach.

2 Materials

The positions and orientations of each cell must be known in order to initialize a simulation. This data was collected using two different image processing programs: Oufti and CellProfiler.

2.1 Oufti

Oufti is an easy to use open source software package created to extract data from microscope images. It was published in 2016 by a research group at Yale University who needed a more efficient way to quantitatively analyze their images. A key feature of Oufti is that it has the ability to store and process a lot of data at a time. Given a microscope image, Oufti will produce a csv file that contains lots of information about the cells. This data can be processed to find values such as length, orientation and position of cells. While Oufti claims to be effective using thousands of images as input, it proved to only work well for a single image [3].

2.2 CellProfiler

CellProfiler is an alternative software package for analyzing microscope images. It was published in 2006 by a research group at Massachusetts Institute of Technology who continue to maintain the software today. CellProfiler provides more customizable features so that complicated image processing can be done by applying certain modules. Examples of these modules include identification of primary/secondary/tertiary objects, illumination correction, plotting, and much more. CellProfiler is very user friendly since there are descriptions within the program for each of these modules. Given a microscope image, CellProfiler will produce a csv file that contains lots of information such as orientation, length, position and more that requires little to no post-processing. CellProfiler is effective at processing multiple images at a time, which makes it possible to compute values such as growth rate and elongation threshold without difficulty [4].

2.3 BSim2.0

BSim2.0 is an open source package written in Java that is used to create agent based simulations of bacteria. This software allows the user to model physical and chemical interactions between capsule shaped bacteria. BSim2.0 initializes its simulations using the endpoints of each cell as well

as a series of parameters that can be varied. Given a set of initial conditions, BSim2.0 produces a series of images at particular time intervals which can be adjusted within the program to fit the need of the user. It also produces a csv file which lists the position and orientation of every cell at certain time intervals [5, 6].

2.4 Graham

The University of Waterloo built the largest supercomputer at any Canadian university in 2017 and named it Graham. Graham has over 1000 nodes, each of which contains at least 125GB of memory and has two central processing units available for use [7]. For large agent based models, multiple nodes may be used in parallel to greatly increase computational speed. The code developed for this report has been adapted for use on the Graham system.

3 Methods

BSim2.0 performs a series of computations between time steps which determines growth and motion among of the bacteria. The following is a brief description of what happens between time steps of a simulation.

Each time step every cell grows by a certain amount, which is calculated as follows. Let Δt be the time between each step of the simulation, L be the length of the cell, $L_{\max}$ be the maximum possible length of the cell, and k_{growth} be the growth rate of the cell. Then the amount by which each cell grows in a timestep is given by

$$\Delta L = k_{\text{growth}}\, \Delta t\, L \left(1 - \frac{L}{L_{\max}}\right)$$

Each cell has a different value for k_{growth}, which is sampled from a normal distribution with a mean and standard deviation specified by the user.

Every time step individual cells have the potential to divide based on their current length. If the current length exceeds some value, $L_{\text{threshold}}$, then the cell divides. For each cell, $L_{\text{threshold}}$ is sampled from a normal distribution with a mean and standard deviation specified by the user.

The growth of the bacteria causes the majority of the motion, while forces within and between bacteria account for the rest of the movement. Three forces were previously implemented by the

creators of BSim2.0, while another three forces were added for this project.

The first force previously implemented by the BSim2.0 creators is a force that keeps the cell rigid despite other forces acting on the endpoints. The endpoints of a cell are labelled $\vec{x}_1$ and $\vec{x}_2$. For the cell to stay rigid, we require that the distance d_L between the endpoints remains constant (it must match the recorded length L of the cell). Let k_{int} be the characteristic strength of this force, then the force keeping the cell rigid is defined as follows.

$$\vec{F}_L = \frac{1}{2} k_{int} (d_L - L)^2 \tag{1}$$

The next force implemented by the creators is an overlap-dependent volume exclusion force between two cells. This force pushes cells away from each other if they are overlapping. The amount of overlap is quantified by the penetration depth $d_{i,j}$ between cells i and j. Let k_{cell} be the characteristic strength of the force, then the force on the i'th cell is given by the following equation.

$$\vec{F}^i_{\text{cell}} = -\frac{2}{5} k_{\text{cell}} \sum_{j \in \text{Neighbours}} d_{i,j}^{5/2} \tag{2}$$

The last force computed by BSim2.0 is an overlap-dependent volume exclusion force between a cell and a wall. This force pushes the cell away from the wall if they are overlapping. Let k_{wall} be the characteristic strength of this force and let d_i be the penetration depth between the i'th cell and the wall. Then the force exerted on the i'th cell by the wall can be calculated from the following equation.

$$\vec{F}^i_{\text{wall}} = -\frac{2}{5} k_{\text{wall}} d_i^{5/2} \tag{3}$$

For a wall composed of mesh elements e this force is computed by summing the contributions from each element. Let $d_{i,e}$ be the penetration depth between the cell and element e. Then the force is given by:

$$\vec{F}^i_{\text{wall}} = -\frac{2}{5} k_{\text{wall}} \sum_{e \in \text{Elements}} d_{i,e}^{5/2} \tag{4}$$

The next three forces were added this term and were based off of forces included in the work of Storck et al in 2014 [8]. Each of these forces are modelled as a spring via Hooke's law. The first force corresponds to end-to-end attraction between neighbouring cells. This force is referred to as the "filial" force. Biologically this happens due to the presence of membrane proteins and

pili on the surface of the cells [9]. The filial force has two components, a short force which pulls neighbouring endpoints of cells together and a long force which pushes the faraway endpoints apart.

The short filial force has a rest length of two times the radius of the end-cap of a capsular cell. The strength of this force is labelled $k_{f,short}$. Let $\hat{L}_{f,short,ij}$ be the vector pointing from one endpoint of a cell to the other and let $x_{f,short}$ be the distance between these two endpoints. By Hooke's law the force between cells i and j can then be computed as follows.

$$\vec{F}^{i,j}_{f,short} = k_{f,short}\hat{L}_{f,short,ij}(x_{f,short} - 2r) \tag{5}$$

The long filial force between cell i and cell j has strength $k_{f,long}$, and the rest length is $L_i + L_j$ where L_i is the length of the i'th cell (similarly for j). Let $\hat{L}_{f,long,ij}$ be the unit vector from the far endpoint of one cell to the far endpoint of the other and let $x_{f,long}$ be the distance between these two endpoints. Then the force can be computed as follows.

$$\vec{F}^{i,j}_{f,long} = k_{f,long}\hat{L}_{f,long,ij}(x_{f,long} - L_i - L_j) \tag{6}$$

The next collection of forces is four sticking spring forces between the ends of each cell. This causes cells to attract each other side-to-side rather than end-to-end. There is one force for each combination of the endpoints. Let $\vec{x}_1$ and $\vec{x}_2$ denote the endpoints of cell i and let $\vec{y}_1$ and $\vec{y}_2$ denote the endpoints of cell j. Assume the endpoints are labelled so that $\vec{x}_1$ and $\vec{y}_1$ are close together as well as $\vec{x}_2$ and $\vec{y}_2$. Then there is a long-range force along the axis from $\vec{x}_1$ to $\vec{y}_2$ and from $\vec{x}_2$ to $\vec{y}_1$, as well as a short range force from $\vec{x}_1$ to $\vec{y}_1$ and from $\vec{x}_2$ to $\vec{y}_2$. The unit vectors along each axis can be labelled $\hat{L}_{ij,11}$, $\hat{L}_{ij,12}$, $\hat{L}_{ij,21}$, and $\hat{L}_{ij,22}$. The strength of the short range force is labelled $k_{s,short}$ and the strength of the long range force is labelled $k_{s,long}$. The short range force must have a rest length of $2r$ (the same as the filial force). Suppose that L_i and L_j are the lengths of the i'th and j'th cells respectively. Then the rest length of the long range force between cells i and j should then be the minimum distance between two far endpoints given that the cells are aligned. Using pythagorean theorem this is found to be $d_{rest,long} = (4r^2 + \max(L_i, L_j))^{1/2}$. All in all, these forces

can be computed to be:

$$\vec{F}_{s,short,ij,11} = k_{s,short}\hat{L}_{ij,11}(\|\vec{x}_1 - \vec{y}_1\| - 2r) \tag{7}$$

$$\vec{F}_{s,short,ij,22} = k_{s,short}\hat{L}_{ij,22}(\|\vec{x}_2 - \vec{y}_2\| - 2r) \tag{8}$$

$$\vec{F}_{s,long,ij,12} = k_{s,long}\hat{L}_{ij,12}(\|\vec{x}_1 - \vec{y}_2\| - (4r^2 + \max(L_i, L_j))^{1/2}) \tag{9}$$

$$\vec{F}_{s,long,ij,21} = k_{s,long}\hat{L}_{ij,21}(\|\vec{x}_2 - \vec{y}_1\| - (4r^2 + \max(L_i, L_j))^{1/2}) \tag{10}$$

The final force added to BSim2.0 is an attractive force between cells and walls. The force is applied to the cell if either endpoint is within range of the wall. Let $k_{s,wall}$ denote the strength of this force. The rest length should then be the radius r of a cell end-cap. If the distance between this endpoint is $d_{i,wall}$ and the unit vector pointing from the endpoint perpendicularly towards the wall is given by $\hat{L}_{i,wall}$, the force can be written as the following.

$$\vec{F}_{s,wall,i} = k_{s,wall}\hat{L}_{i,wall}(d_{i,wall} - r) \tag{11}$$

The following is a diagram displaying each of the forces present in the simulation.

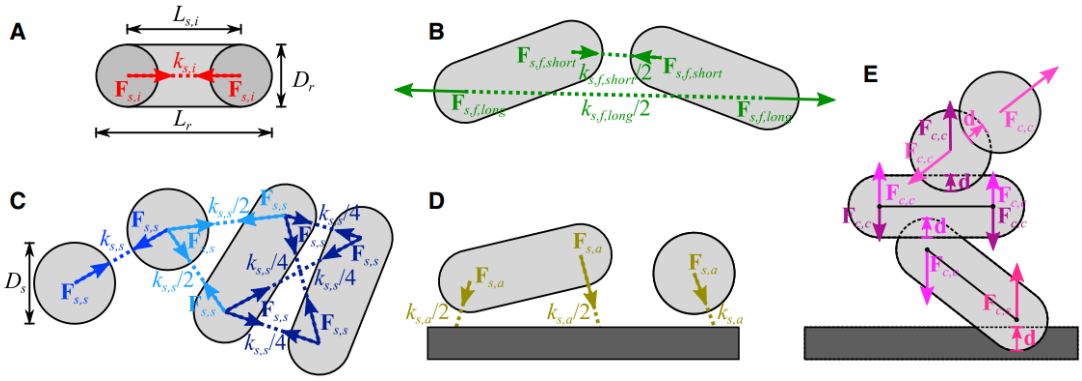


Figure 1: Diagram of each force present in the simulation. A) shows the internal force present within a cell. B) shows the filial forces which affect cells end-to-end. C) shows the side-to-side attractive forces. D) shows attractive forces between cells and walls. E) shows the overlap-dependent exclusion forces. The diagram is taken from Storck et al, 2014 [8] in which the constants have been scaled in a slightly different way.

Together there are the internal forces in a cell $\vec{F}_{\text{int}}$ and the forces exerted by nearby objects on a cell $\vec{F}_{\text{ext}}$. In this simulation the bacteria do not move if these forces add to zero. This is different from the typical situation in which $\sum \vec{F} = 0$ implies that objects have no acceleration.

This is caused by hydrostatic effects. Microscopic objects in a fluid will obey a particular limit of the Navier-Stokes equation called the Stokes equation. In this limit, the drag force on the cell is $\vec{F}_{drag} = -\beta\vec{v}$, where $\vec{v}$ is the velocity of the cell. If the cell's acceleration is zero, instead of $\vec{F}_{\text{ext}} + \vec{F}_{\text{int}} = 0$, we have that $\vec{v} = \beta^{-1}(\vec{F}_{\text{ext}} + \vec{F}_{\text{int}})$. So when the collision forces and attractive forces cancel with the internal forces, the bacteria stop moving. This is why there is no drift in the simulations unless a background flow field $\vec{u}$ is added externally.

4 Results and Analysis

Many of the simulations used to test this software were initialized based off of data taken from microscope images collected by Sara Khorasani in Dr. Ingalls' lab. One such image is displayed in the following figure.

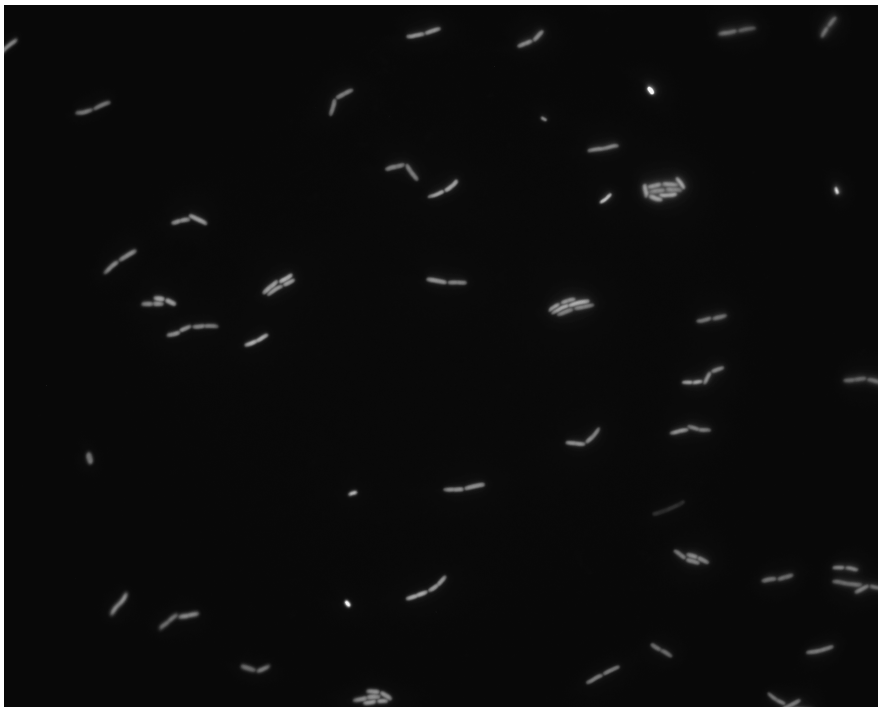
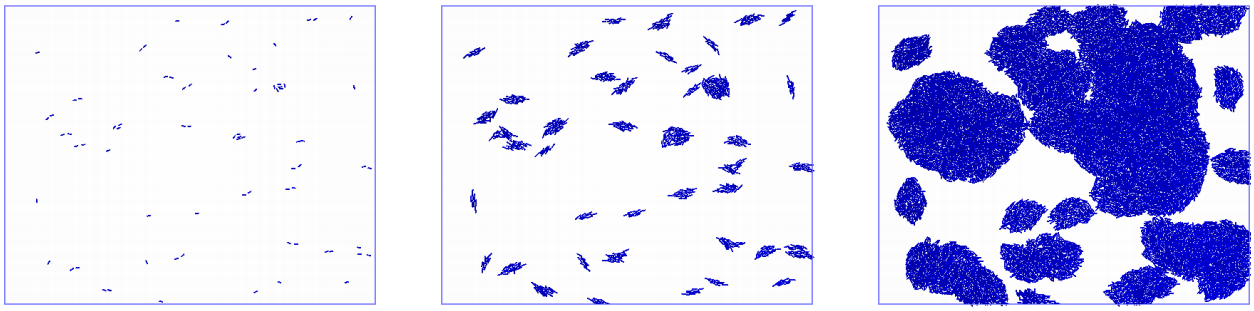


Figure 2: Image taken of E. coli K-12 colony.

The following figures demonstrate how implementing each of these forces affects the simulation. The first set of images is collected from an early simulation in which only the forces in the original BSim2.0 example simulation were implemented. The bacteria grow in round colonies.



(a) Time $t = 0$ (b) Time $t = 30$ (c) Time $t = 60$

Figure 3: Snapshots of simulation with only the original BSim2.0 forces incorporated.

The second set of images was collected from the current software in which all of the forces are available. Only the end-to-end forces between the bacteria have been activated to demonstrate the structural changes that happen.

(a) Time $t = 0$ (b) Time $t = 40$ (c) Time $t = 80$

Figure 4: Snapshots of simulation with new filial forces incorporated. For this simulation, $k_{f,short} = 2, k_{f,long} = 0.01$.

In the third set of images, all of the forces are active.

(a) Time $t = 0$ (b) Time $t = 30$ (c) Time $t = 60$

Figure 5: Snapshots of simulation with all forces incorporated. For this simulation $k_{f,short} = 2, k_{f,long} = 0.01, k_{s,short} = k_{s,long} = 2$.

Finally, to study the chain-forming mechanism many single-cell simulations were run. Shown in the figure is one such simulation.

(a) Time $t = 0$ (b) Time $t = 40$ (c) Time $t = 80$

Figure 6: Snapshots of simulation with all forces incorporated but only one initial cell.

5 Discussion

Initially the results from the simulations showed round colonies forming around individual cells. This is due to the lack of any forces keeping cells aligned with one another. When end-to-end attractive forces are included, the cells begin to form chains. The microscope images taken from the lab show cells forming colonies composed of layered chains. This behaviour is most prominent in the simulation including all of the forces, especially when the constants associated with the side-to-side forces are made large. In the simulations shown above, the bacteria in a chain slip against one another and the chain curves left or right. This is because the end-to-end forces are too weak to form long chains for this choice of $k_{f,short}$ and $k_{f,long}$. If chains didn't slip the bacteria would form long unbreaking lines, which is not what happens in real life. Therefore there is a critical value for

$k_{f,short}$ and $k_{f,long}$ so that chains will still break and form side-to-side bonds, but not curve too much. This should result in non-curving colonies which are multiple cells thick.

6 Conclusions and Recommendations

The next step to improving this model will be to verify that there are no unintended effects from changing the strength of each force. Specifically, the side-to-side forces should be more closely examined.

It will also be important to quantitatively measure how similar the simulation output is to the microscope data. The simulation output files can be directly compared to the image processing data using a script. The simulations could then be run in batches in order to fit each of the parameters in the model empirically. Since this is a very high dimensional problem, many combinations of parameters have to be tested. This is computationally expensive, so these calculations should be performed on a computing cluster such as Graham.

Currently the growth rate and cell division threshold length have not been numerically fitted to data. The growth rate constant determines the conversion between simulation time and real time, so it must be found first. The cell division threshold fixes the maximum length a cell reaches before it divides, and so it fixes the characteristic length scale of the simulation. Therefore it is important to measure these two values before trying to fit other parameters.

References

[1] Antonio P Garcia. "A first approach to individual-based modeling of the bacterial conjugation dynamics". Masters' Thesis. Universidad Politecnica de Madrid, 2011.

[2] Antonio Prestes García and Alfonso Rodríguez-Patón. "A Preliminary Assessment of Three Strategies for the Agent-Based Modeling of Bacterial Conjugation". In: *Advances in Intelligent Systems and Computing* 375 (Jan. 2015), pp. 1–9. DOI: 10.1007/978-3-319-19776-0_1.

[3] Ahmad Paintdakhi et al. "Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis". In: *Molecular Microbiology* 99.4 (2016), pp. 767–777. ISSN: 13652958. DOI: 10.1111/mmi.13264.

[4] Anne E. Carpenter et al. "CellProfiler: Image analysis software for identifying and quantifying cell phenotypes". In: *Genome Biology* 7.10 (2006). ISSN: 14747596. DOI: 10.1186/gb-2006-7-10-r100.

[5] Antoni Matyjaszkiewicz et al. "BSim 2.0: An Advanced Agent-Based Cell Simulator". In: *ACS Synthetic Biology* 6.10 (2017), pp. 1969–1972. ISSN: 21615063. DOI: 10.1021/acssynbio.7b00121.

[6] Antoni Matyjaszkiewicz et al. "BSim 2.0: Supplementary Information". In: (2017), pp. 1–26. URL: https://github.com/bsim-bristol/bsim..

[7] *Graham*. Nov. 2020. URL: https://docs.computecanada.ca/wiki/Graham.

[8] Tomas Storck et al. "Variable cell morphology approach for individual-based modeling of microbial communities". In: *Biophysical Journal* 106.9 (2014), pp. 2037–2048. ISSN: 15420086. DOI: 10.1016/j.bpj.2014.03.015. URL: http://dx.doi.org/10.1016/j.bpj.2014.03.015.

[9] Saul Burdman et al. "Aggregation in Azospirillum brasilense: Effects of chemical and physical factors and involvement of extracellular components". In: *Microbiology* 144.7 (1998), pp. 1989–1999. ISSN: 13500872. DOI: 10.1099/00221287-144-7-1989.