Contributions to mathematical modelling and numerical simulation of biofilms

Alejandro López Núñez.

Tese Doutoral, UDC, Ano 2018.

Directores: Eva Balsa Canto, Carlos Vázquez Cendón.

Programa de doutoramento en *Métodos matemáticos e simulación numérica en enxeñaría e ciencias aplicadas.*

UNIVERSIDADE DA CORUÑA

PhD. Thesis

Contributions to mathematical modelling and numerical simulation of biofilms

AUTOR:

Alejandro López Núñez

DIRECTORES:

Eva Balsa Canto

Carlos Vázquez Cendón

Tese presentada para a obtención do título de Doutor con Mención Internacional na Universidade da Coruña

Departamento de Matemáticas

Facultade de Informática, A Coruña (Spain)

Abril, 2018

UNIVERSIDADE DA CORUÑA

Los abajo firmantes hacen constar que son directores de la Tesis Doctoral titulada **"Contributions to the mathematical modelling and numerical simulation of biofilms"** desarrollada por Alejandro López Núñez, cuya firma también se incluye, dentro del programa de doctorado **"Métodos Matemáticos y Simulación Numérica en Ingeniería y Ciencias Aplicadas"** en el Departamento de Matemáticas (Universidade da Coruña), dando su consentimiento para su presentación y posterior defensa.

The undersigned hereby certifies that they are supervisors of the Thesis entitled **"Contributions to the mathematical modelling and numerical simulation of biofilms"** developed by Alejandro López Núñez, whose signature is also included, inside the Ph.D Program **"Mathematical Modelling and Numerical Methods in Applied Sciences and Engineering"** at the Department of Mathematics (University of A Coruña), consenting to its presentation and subsequent defence.

Abril, 2018

Directores:

Carlos Vázquez Cendón

Eva Balsa Canto

Doctorando:

Alejandro López Núñez

Funding

This research has been partially funded by the following projects:

- FPU grant (ref. FPU13/02191) from the Spanish Government program MECD-FPU, 09/16/2014-11/30/2017
- Predoctoral grant (ref. PRE/2013/293) from the Galician Government program *Plan I2C*, 12/01/2013-09/15/2014
- Spanish project with reference MTM2016-76497-R from the Spanish Ministry of Economy and Competitivity
- Spanish project with reference MTM2013-47800-C2-1-P from the Spanish Ministry of Economy and Competitivity
- Grant GRC2014/044 *Grupos de Referencia Competitiva* from the Galician Government
- FEDER funds

Financiación

Este trabajo de investigación ha sido cofinanciado por los siguientes proyectos:

- Beca FPU (ref. FPU13/02191) del programa MECD-FPU del Gobierno de España, 16/09/2014-30/11/2017
- Beca predoctoral (ref. PRE/2013/293) de Plan I2C de la Xunta de Galicia, 01/12/2013-15/09/2014
- Proyecto del Ministerio de Economía y Competitividad con referencia MTM2016-76497-R
- Proyecto del Ministerio de Economía y Competitividad con referencia MTM2013-47800-C2-1-P
- Beca GRC2014/044 Grupos de Referencia Competitiva de la Xunta de Galicia.
- Fondos FEDER

To all who wander by my side.

Acknowledgements

This PhD thesis is the result of a long and marvellous journey, achieved with the incredible assistance of quite a lot of friends, fellows, family and masters. I cannot do any less than acknowledge their support with all the gratitude of my heart.

First of all, I must mention my beloved family. My parents, *Ana* and *Juan*, my grandparents, present and absent, *Belarmina*, *Luci*, *Juan* and *Luis*, my sisters *Noelia* and *Aida*, my cousin *Candela* and my aunts *Isabel*, *Reme*, and uncle *Paco*. All of them have been fundamental, even in the distance or, maybe, specially while far away. I am what I am because of them and for that, I will always be grateful.

Second, the directors of this work, *Eva Balsa Canto* and *Carlos Vázquez Cendón*, whose guidance is invaluable. I cannot say with words how much they have changed my life both as teachers and masters, and as paragons of what a researcher should be.

Also, I am vastly grateful to professor *David L. Chopp* and to the *McCormick School of Engineering and Applied Science* from *Northwestern University*, Chicago, Illinois. His kind help and attention during my stay at his team was fundamental to the development of this PhD. Much of this work was achieved thanks to his teachings. I should not forget either his PhD students, specially *Noah*, *Josh*, *Narut* and *Tom*, among others, as they accepted me with open arms easing my time there.

Many thanks to *Maria Luisa Garzón*, not only because of her advice with the *Level Set method* but also because it was her guidance that led us to Chicago.

I would like to acknowledge the help of the researchers from the *Instituto de Investigaciones Marinas*, IIM-CSIC, from Vigo, Spain. Much of their knowledge

enriches this work, thus making their help invaluable.

As part of the FPU grant, I had the amazing opportunity to try the taste of teaching *Calculus* to the unsuspecting students from the *Faculty of Informatics* at the *University of Coruña*. As insightful as it was, it would have been quite fearful without the help of the other professors of the course, specially, *Teresa Iglesias*, *Íñigo Arregui* and *María González Taboada*.

During all these years, I have shared the laboratory with some fantastic fellows. *Jose, María, Carmen, Paula, Ana and Beatríz* deserve a special thank for their support, their ideas and the shared experiences.

Finally, my closest friends deserve a warm acknowledge. Scattered we may be, but we keep close to each other. Special thanks to my great friend *Jano* for his comfort in the hard moments and his joy in the good ones.

This journey may come to an end, but it is only the beginning of a new adventure. Farewell.

Table of Contents

List of Tables

List of Figures

Abstract

The main goal of this PhD work is the development and numerical resolution of mathematical models that simulate the dynamics of bacterial biofilm systems, paying special attention to those biofilms formed by the *Listeria monocytogenes*, a pathogen of special relevance in food safety.

A biofilm is a layer of microorganisms attached to a surface and protected by a matrix of exopolysaccharides. Biofilm structures difficult the removal of microorganisms, thus the study of the type of structures formed throughout a biofilm life cycle is key to design elimination techniques.

In the present work, we develop different models that simulate the dynamics of biofilms formed by different strains of *L. monocytogenes*. We start with a 1D model that can be used to describe the formation of flat biofilms. Afterwards, we apply the acquired knowledge to develop a 2D model capable of describing more complex structures. All this models are solved with efficient numerical methods and robust numerical techniques, such as the Level Set method, that guarantees the good behaviour of the obtained solutions.

Finally, the numerical results are compared with the experimental measurements obtained in the *Instituto de Investigaciones Marinas*, CSIC (Vigo, Spain), and the Micalis Institute, INRA (Massy, France).

Resumen

El objetivo fundamental de esta tesis consiste en el desarrollo y resolución numérica de modelos matemáticos que reproduzcan la dinámica de biopelículas, en especial, aquellas formadas por *Listeria monocytogenes*, un patógeno especialmente relevante en la seguridad alimentaria.

Una biopelícula es una capa de microorganismos adheridos a una superficie y protegidos por una matriz de expolisacáridos. La estructura de las biopelículas hace difícil la eliminación de los microorganismos, de ahí que sea fundamental estudiar el tipo de estructuras que se forman durante la vida de una biopelícula para poder diseñar nuevas técnicas de eliminación.

En el presente trabajo, desarrollamos modelos que simulan la dinámica de biopelículas der diferentes cepas de *L. monocytogenes*. Comenzamos estableciendo un modelo 1D capaz de describir la formación de biopelículas planas para, posteriormente, aplicar los conocimientos adquiridos al desarrollo de un modelo 2D capaz de describir estructuras más complejas. Estos modelos se resuelven con métodos numéricos eficientes y técnicas numéricas robustas, como el método *Level Set*, que garantizan el buen comportamiento de las soluciones.

Finalmente, los resultados obtenidos son comparados con las imágenes de microscopía procedentes de experimentos realizados en el Instituto de Investigaciones Marinas, CSIC (Vigo, España) y en el *Micalis Institute*, INRA (Massy, Francia).

Resumo

O obxectivo fundamental desta tese é o desenvolvemento e resolución numérica de modelos matemáticos que reproduzan a dinámica de biopelículas, en especial, aquelas formadas por *Listeria monocytogenes*, un patóxeno de especial relevancia na seguridade alimentaria.

Unha biopelícula é unha capa de microorganismos adheridos a unha superficie e protexidos por unha matriz de expolisacáridos. A estructura das biopelículas fai difícil a eliminación dos microorganismos, de aí que sexa fundamental estudar o tipo de estructuras que se forman ao longo da vida dunha biopelícula para poder deseñar novas técnicas de eliminación.

No presente traballo, desenvolvemos modelos que simulan a evolución de diferentes cepas de *L. monocytogenes*. Comezamos establecendo un modelo 1D capaz de describir a dinámica de cepas chás para, posteriormente, aplicar os coñecementos adquiridos ao desenvolvemento dun modelo 2D capaz de describir estruturas máis complexas. Todos estes modelos resólvense con métodos numéricos eficientes e técnicas numéricas robustas, como o método *Level Set*, que garanten o bo comportamento das solucións.

Finalmente, os resultados obtidos compáranse coas imáxes de microscopía obtidas en experimentos realizados no Instituto de Investigacions Mariñas, CSIC (Vigo, España) e no *Micalis Institute*, INRA (Massy, Francia).

Introduction

Bacterial biofilms are essential for the equilibrium of ecosystems, as well as for many industrial processes. A bacterial biofilm is a layer of prokaryotic cells anchored to a substrate layer that keeps them fed. This layer is embedded in a matrix formed by polymers, called the exopolysaccharides matrix or EPS, which keeps bacteria grouped and difficults their removal [21, 108]. In a simplified way, we could say that a biofilm is a group of micro-organisms attached to a surface. Current studies estimate that less than 0.1% of the aquatic microbial lifeforms are in a planktonic (free floating) state [20]. Therefore, biofilms constitute the preferred microbial lifeform.

The reason behind that preference is the competitive advantage achieved by the ability to adhere to surfaces and form biofilms, compared to planktonic bacteria. The latter can be easily washed out by the water flux while the former are protected from that phenomena and live in an environment where they can grow, provided nutrients are available. Biofilm physical structure also allows the distinction of biological niches that make growth and survival easier for those micro-organisms that could not compete in a fully homogeneous system. Furthermore, the microbial activity inside biofilms can change the inner environment making the biofilm more hospitable than the liquid region [86].

The fundamental layers present on a biofilm are the surface (where bacteria are attached to), the proper biofilm (formed by one or more cell species and the EPS), the media (with nutrients), and the environmental conditions that determine the biofilm development (temperature, pH, hydrodynamics, etc.).

There are beneficial biofilms, both for humans or the adequate environmental

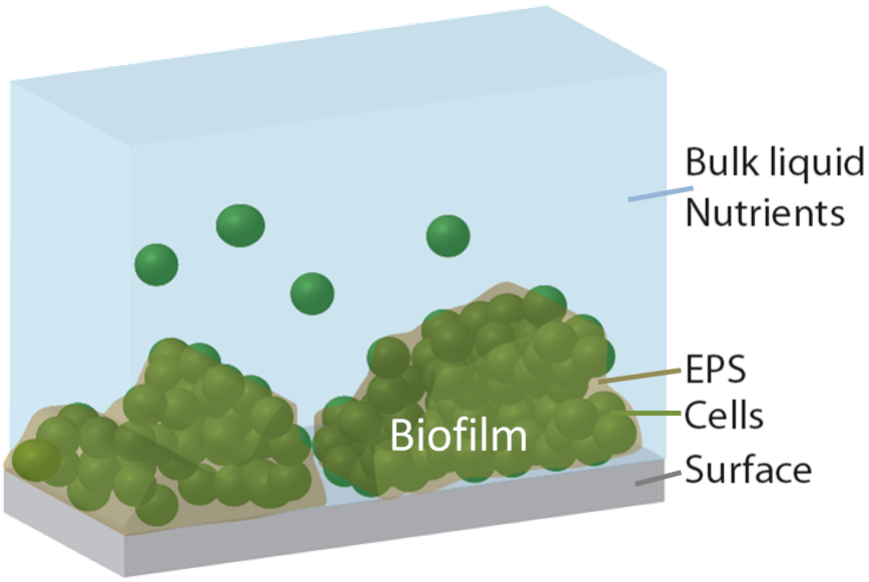

Figure 1: Fundamental parts of a bacterial biofilm system

development, as well as harmful biofilms, which usually cause health problems or poor industrial processes performance and thus economical losses. Examples [108] of the former are the biofilms used in water treatment plants (RBCs, biological reactors, etc.) or in the fermentation industry (for instance, in quick-vinegar processes), but also some naturally formed biofilms present in the subsoil (contributing to soil or groundwater decontamination), in rivers, lakes and coastal zones (colonizing rocks or suspended in the water, usually contributing to the removal of contaminants from the water), or in the roots of many plants (enhancing the availability of nutrients to plants). Naturally formed biofilms are fundamental for earth biosphere.

On the other hand, harmful biofilms appear in many situations. For instance, biofilms are a major problem in dental hygiene [61], infectious diseases or infections caused by medical implants [115]. Other examples are the biofilms that cause water

contamination or malfunctions in heat exchangers [11]. Specially relevant are the biofilms formed by pathogenic bacteria that appear in food industry because they constitute an important source of food contamination and can be a major health issue for consumers [93].

Preventing harmful biofilm formation is, in general, quite difficult, due to their ability to develop even in adverse conditions. Also, once formed, it is hard to eliminate them, because bacteria forming biofilms are much more resilient to the immune response of the host or antimicrobial agents [34, 44].

The necessity of improving the properties of beneficial biofilms or controlling the formation of the harmful ones has motivated numerous investigation work lines considering the genetic, biochemical or physical mechanisms that contribute, not only to biofilm formation, but also to its structure.

Numerous studies indicate that biofilm structure determines the magnitude of the processes taking place inside of it, such as the nutrients transference rate to inner layers, the antimicrobial agents diffusion rate or its resistance to friction. Plate counts have been widely used to analyse biofilms. However, they only consider viable culturable cells and do not inform about the structure [22]. Due to these limitations, several alternatives have been proposed.

Major emphasis has been paid to the development of various microscopy based techniques. The most successful, confocal laser scanner microscopy (CLSM), enables *in situ* and *in vivo* three-dimensional biofilm optical imaging [95]. Also, microsensors can be used to measure concentrations of the different components inside the biofilm, thus allowing the analysis of the availability of substrates of the different regions of the biofilm [116]. The advances in molecular biology and *in situ* hybridisation techniques contributed to the development of gene probes and microscopy techniques, allowing the detailed analysis of microbial communities inside the biofilm [23, 56, 97]. The combination of cellular stains for fluorescent cellular staining together with imaging techniques allow mapping viable and damaged or dead cells [99] or the distribution

of extracellular polymeric substance but also reconstructing three-dimensional structures, thus obtaining a more comprehensive study of biofilms. Remarkably CLSM together with quantitative image analysis [114] allows for the automatic determination of biofilms thickness, biovolume or roughness. This approach allows for a quantitative comparison of biofilms among different strains, species or environmental conditions (media, temperature, type of culture).

Recent works suggest several alternative work-flows and software tools for the systematic analysis of microscopy images. IMARIS (a commercial software) enables the reconstruction of 3D structures. COMSTAT [43], ISA 3D [10], or PHLIP [67] allow quantifying confocal laser scanning microscopy (CLSM) images. BIOFILMDIVER [65] permits the quantification of biofilms areal porosity, covered area, diffusion distances or the spatio-temporal population distribution from 2D images taken with epifluorescence and CLSM. Machine learning algorithms can be used to analyse scanning electron microscopy images [103].

Biofilm modelling: a brief historical perspective

Quantitative image analysis can be complemented with mathematical modelling to gain insights into the mechanisms that lead to a given biofilm structure [4]. Even the most homogeneous biofilms develop complex internal processes, intertwined between them. Mathematical modelling brings the possibility to explore different internal processes and their connections, their relative importance and the role of the environment in biofilm life cycle [105].

The inception of mathematical models of biofilms started around mid 20th century with simple empirical models [69]. Since then, mathematical models became much more complex taking into account multiple species both of bacteria and nutrients together with different microbiological, chemical and physical processes giving place to multi-dimensional PDE systems. The evolution of the mathematical models of biofilms is linked to the evolution of the experimental analysis of biofilms.

The new techniques and tools gave place to new experimental models to grow

biofilms in certain conditions in order to study specific features (such as the use of flow cells that can be helpful to study the initial stages of a biofilm), but more importantly, inspired new mathematical models [70]. For instance, confocal laser scanning microscopy allowed the visualisation of heterogeneous structures in biofilms triggering the development of three-dimensional models that were able to reproduce the three-dimensional structure of the biofilm [77]. Other examples include the models reproducing the multi-species interactions that emerged after the *in situ* visualisation of individual micro-colonies within a biofilm.

Most mathematical modelling efforts mainly focus on the exploration of different mechanisms that lead to particular biofilm structures, mostly clustered and mushrooms forming biofilms. The emphasis is paid into qualitative validations of developed models, which are able to recover experimentally observed structures. In that pursue, the new generation of biofilm models offers detailed descriptions of the formation of heterogeneous structures with clusters and mushrooms (see, for instance, the reviews by [45, 77, 105]). Thus continuous and hybrid models have been proposed that incorporate various mechanisms to describe biomass growth, spreading and detachment as well as nutrients transport and conversion.

This PhD work gives a step forward by combining CLSM image analysis, novel mathematical models and advanced numerical techniques to qualitatively and quantitatively explain *Listeria monocytogenes* biofilms behaviour.

Listeria monocytogenes

In the period 2011-2014, 183 persons were infected with an outbreak-associated strain of *Listeria monocytogenes*, resulting in 39 deaths in the USA. Such infections were linked to three different cheese brands and cantaloupe [14]. More recently, the USA Centers for Disease Control and prevention have reported a number of outbreaks from 2015 to 2017 related to the consumption of some cheese products, dairy products, packaged salads, or frozen vegetables. In Europe eight strong-evidence food-borne

outbreaks caused by *L. monocytogenes* were reported in 2011–2012 [31, 32]. *Listeriosis* affected about 2,200 people in the EU in 2015, causing 270 deaths. Besides, there was a total of 2536 reported *listeriosis* cases in 2016 [33]. *Listeria* strains were found in sandwiches, bakery products or ready-to-eat fish and meat products.

Listeria monocytogenes is a Gram-positive, food-borne pathogen that can cause systemic infections in immune compromised, pregnant or elder patients [19]. Out of the around fifteen *L. monocytogenes* serotypes reported, there are three that account for the majority of human diseases: 1/2a, 1/2b and 4b [96]. Transmission of this pathogen to humans is primarily due to consumption of contaminated food, usually contaminated by contact with unhygienic work surfaces or facilities, where biofilms are found [50, 96, 109]. *L. monocytogenes* biofilms can emerge on common food contact surfaces, such as plastic, polypropylene, rubber, stainless steel or glass [92], and are protected from a variety of environmental factors, such as temperature, salt, sugar or pH [64]. They also tolerate better biocides, thus hindering the surface decontamination process to the point that many *L. monocytogenes* strains have been isolated from food processing plants despite the programs to sanitize them [13, 87].

Miettinen et al. [63] showed that *L. monocytogenes* biofilms can persist several years in food processing plants. These facts reflect the relevance of biofilms for the food industry [64] and identifies *L. monocytogenes* as a major concern. Because of this, quantification of the biofilms formed by *L. monocytogenes* is of the highest interest. Indeed the knowledge of interstrain variations is basic for understanding biofilm behaviour and designing adequate disinfection techniques

L. monocytogenes biofilms can depict different structures: mono-layer of adhered cells, flat unstructured multi-layers, honeycomb structures or clusters [12, 15, 24, 62, 78, 82]. Recent works [37] reconstructed CLSM images to observe *L. monocytogenes* instraspecies diversity and concluded that most strains form complex honeycomb-like structures at 48h. Other works [65, 66] used quantitative image analysis to study the life cycle of biofilms formed by three *L. monocytogenes* strains, showing the presence

of, at least, three phases: first, separate clusters that evolved to interconnected clusters, honeycomb-like or flat structures and finally, a detachment phase. The timing for these phases varies significantly among strains.

Objectives

This PhD work presents the identification and numerical solution of various deterministic mathematical models intended to describe the life cycle of biofilms formed by several different strains of *Listeria monocytogenes*. Special emphasis was paid to the implementation of advanced numerical techniques for the solution of the proposed candidate models. Model predictions were compared to experimental data.

The following objectives were defined:

- To develop a one-dimensional model to describe the dynamics of flat biofilms. Mosquera et al. [66] showed that the L1A1 *L. monocytogenes* strain forms thick flat biofilms. The aim is to develop a continuous reaction-diffusion model to explore the relevant mechanisms that may lead to flat biofilms.

- To develop a two-dimensional model to describe the dynamics of flat or clustered biofilms. Bridier et al. [12] and Mosquera et al. [66] showed that *L. monocytogenes* strains might form flat, honeycomb or clustered biofilms. The aim is to propose a two-dimensional multi-species model which can account for such interstrain variability.

- To propose robust and efficient numerical techniques to solve one- and two-dimensional deterministic models with special interest on those based on the Level set method.

- To compare model predictions with quantitative image analysis data for various *L. monocytogenes* strains of relevance in the food industry.

Structure of this PhD work

In Part I we focus on the study and development of several one-dimensional models. In particular, we selected a deterministic continuous modelling framework. Deterministic reaction-diffusion models (RDM) offer the advantage of reproducibility [105]. Besides, this type of models can be solved with advanced numerical techniques to guarantee the computational efficiency required for model identification through optimization based techniques [3, 102]. Chapter 1 serves as an introduction to biofilm modelling, as well as a presentation for various candidate reaction-diffusion models to describe the *L. monocytogenes* biofilm system. The candidate models incorporate alternative mechanisms for growth, nutrient consumption and detachment. Each model was reconciled with the measured data through optimization based data fitting. The best model was selected attending to a best compromise between the number of unknown parameters and its capability to quantitatively reproduce the measurements.

Chapter 2 describes the numerical methods and techniques and offers the guidelines to build the numerical systems. The finite differences Crank-Nicolson method together with a central differences scheme in space, leads to the fully discretised nonlinear numerical system, which is solved through the linearised system resulting from the application of a Newton method. These numerical methods were selected so that the computational time is reduced and the numerical solution is efficiently calculated, as was discussed in the recent work [7]. Chapter 2 also shows the resolution of some academic tests, in order to validate the numerical methods applied.

Chapter 3 shows the numerical results of the different models, as well as a comparison between the numerical results obtained from the model and the corresponding experimental measurements, in order to validate the models and numerical methods used to solve them from a biological point of view. The focus of the discussion is the most satisfactory model, although other intermediate candidates were also analysed in the recent work [9].

In Part II a similar path is followed, this time focusing on the study and development of several two-dimensional models. We maintained the deterministic continuous modelling framework, but in this part we added damaged or dead cells as a second biological species and substituted the diffusion process by an advection process that regulates the advance of the biofilm. Chapter 4 introduces the necessity of two-dimensional biofilm modelling in order to capture the different spatial heterogeneities that a biofilm can develop. Afterwards, it presents the different studied and developed models.

Chapter 5 presents the numerical strategy used to solve the two-dimensional proposed models. As the complexity of the system increases, new methods with higher efficiency and more sophistication are required. Although we keep working in a finite differences setting, we use a modified upwind method following the strategy developed by Gibou et al. [36] together with a Level Set technique [89], WENO schemes [48], Crank-Nicolson numerical strategies and a Newton algorithm for the non-linear fully discretised models.

Finally, Chapter 6 shows the numerical results for the different models, analysing their accuracy in order to validate them from a biological point of view.

Appendix A summarises some mathematical results on the theoretical analysis of the Eberl model that appear in a recent work of Efendiev, Eberl and Zelik [28].

Appendix B presents a brief explanation on the optimisation techniques included in the AMIGO2 toolbox [5], a high level MATLAB-based toolbox which is used in the parameter calibration process.

Finally, a last chapter summarizing the main results of the thesis, the more relevant conclusions and some expected future lines of research are included.

Part I

One-dimensional models

One-dimensional models to describe flat biofilms

Chapter 1

1D biofilm modelling

1.1 Some notions on biofilm modelling

The interest in the study of biofilms started in the late 19th century. However, it is not until the publication of the pioneering work by Zobell and Anderson in 1936 [117], that their study started to receive significant attention. After that first work, the first simple mathematical models appeared. Those models were mostly empirical, consisted of simple empirical correlations derived from the patterns found in experimental data obtained under relevant conditions (see, for instance, [69]). Therefore, those models could not provide information about the internal mechanisms and their range of application was limited to the conditions in which they had been obtained.

The advent of novel experimental techniques allowed to increase the knowledge about the processes happening inside biofilms. This knowledge could be incorporated into new mechanistic models. Mechanistic biofilm modelling started in the 70s, with new models mainly focused on the role of nutrients diffusion within the biofilm. Those first models consisted of one or two one-dimensional partial differential equations (PDEs) which described the decrease of nutrients concentration inside the biofilm and the biofilm growth [39, 40, 55, 84, 85, 110]. They considered a simple geometry, a slab; a uniform biomass (viable cells) distribution and combined diffusion mechanisms with reaction terms.

In the 80s, mathematical models started to include different types of micro-organisms (multi-species) and non uniform biomass distributions inside the biofilm. These second generation models maintained the one dimensional slab geometry, but considered systems of reaction-diffusion ODE and/or PDE for both biomasses and nutrients while adding spatial patterns for the different types of nutrients and biomasses. The main goal of these models was to evaluate the total nutrients and the fluxes of metabolic products in the surface of the biofilm. [51, 83, 106, 107]

Since then, the experimental techniques used in the detailed analysis of biofilm structures and activities have improved significantly, favouring the development of experimental models that allowed the growth of biofilms in the laboratory. This fact, coupled with the reliable new numerical methods and computers, gave rise to a renewed interest in the mathematical modelling of biofilms. The possibility of studying heterogeneous structures present inside a biofilm motivated the development of mathematical models simulating the three-dimensional biofilm structure. The possibility of having *in situ* visualisations of micro-colonies inside a biofilm motivated the development of mathematical models that were able to reproduce the interaction between species.

Current models include numerous features or mechanisms that underlie complex structure dynamics. The use of these models gives a detailed description both of the structure of the biofilm and its biological behaviour, thus providing novel insights into biofilm life cycle but at the expense of computational cost, compared to the small amount of information but low cost of the first models [105].

As explained, mathematical models can be very different, from simple empirical correlations to sophisticated and computationally expensive models, or even vary from the macroscopic to the microscopic scale. Choosing the right model or building a new one requires, first, the definition of the model objective, i.e., the question we want to answer with such a model. That question will condition the mechanisms to be incorporated in the model and, thus, its complexity. Due to the flexibility that goes hand in hand with the mathematical models, it is easy to add or eliminate

processes which, in turn, makes modelling an important and versatile tool in biofilm investigation [105].

A model intended to capture detailed descriptions of all components and processes must include mass balance equations for each component in each part of the domain. Also it must take into account continuity equations for the fluid system when the biofilm is embedded in a circulating fluid and well defined boundary conditions for each and every variable. It is quite obvious that such a model would be computationally non-viable, so simplifications are in order. While applying simplifications, it is important to foresee the consequences of those simplifications, thus being another difficulty in the model design process.

The latest multidimensional models take into account the distinction between the solid domain Ω_1 (containing the biomass), and the liquid domain Ω_2 (including pores and channels). Both domains are separated by an interface Γ. In Ω_1, nutrients are consumed through biochemical reactions and are transported by diffusion, while in Ω_2 transport is due to diffusion and convection. Certain studies take also into account the possibility of decoupling processes based in the fact that nutrients consumption and diffusive and convective transport are much faster that the biofilm development processes [51].

On the other hand, equations describing the hydrodynamic processes, the transport and consumption of nutrients and biomass production are well known, whereas the modelling of biomass propagation is still an open question due to the many influencing factors both physical and environmental [75, 100, 101].

Recent models are divided into three great categories: cellular automaton, discrete models (individual model based), and continuous models. The first two, are based on the concept of *agents*. Each cell is an agent which behaves following a certain set of rules. Those rules represent cellular nutrient uptake, cellular proliferation and spatial location. When biomass is greater than a certain density threshold in one cell, certain quantity propagates randomly to a neighbouring empty cell, using different strategies when neighbours are not empty [42, 71, 74, 111]. The predictions offered

by these models are in qualitative agreement with experimental results, but they present some physical disadvantages. First of all, the predictions strongly depend on the discretisation, which makes them not invariant with respect to the coordinate system. Furthermore, there must be certain ordering of grid cells so that, when two cells try to move biomass into a shared neighbour, there is no conflict with the biomass spread mechanism. Finally, there are many different formulations for the set of propagation rules, all of them reliable, but creating qualitative differences.

In the case of continuous models all these disadvantages disappear as they do not rely that much in the discretisation and the spreading mechanism is based only in the processes added to the model. They are not free from drawbacks, either, as the computational cost can be high, although the computational cost does not depend on the number of cells considered in the system. But they generate deterministic solutions, meaning that they need to be run only one time, and they can be subject to analytical studies.

In this PhD work we propose various deterministic models to describe the dynamics of biofilms formed by several different *L. monocytogenes* strains. First, we will consider one-dimensional models which can be be efficient and reliable alternatives to describe flat biofilms (as the ones presented by L1A1 *L. monocytogenes* strain).

1.2 First model: the Eberl model

As it has been previously indicated, L1A1 *L. monocytogenes* strain develops mostly flat biofilms, so a one dimensional model seems appropriate. As the starting point for the current work, let us consider the model developed by Eberl et al. [26], a robust model from the literature validated not only numerically but also analytically. We refer to Appendix A for a summary of the analytical results, as proved in Efendiev et al. [28].

The Eberl model takes into account the following assumptions:

i. Existence of a sharp front of biomass at Γ representing the transition from the

biomass region to the bulk region.

ii. Biomass spreading is important when the biomass concentration reaches certain density.

iii. Biomass density can not exceed a maximum bound.

iv. Biomass production is due to nutrient consumption.

v. Spreading mechanism must be compatible with the hydrodinamics and the nutrients consumption and transfer model.

vi. Given certain biochemical parameters, the spatial heterogeneity of a single-species biofilm is only due to the environmental conditions.

The biomass density maximum bound postulated in *iii)* is not obtained from the reaction terms, based on *iv)*. In fact, it must be a associated to the biomass propagation process. Moreover, one can expect that limit to be obtained as a consequence of *ii)*.

The propagation mechanism will be a non-constant diffusive flux, thus avoiding an instantaneous diffusion that would contradict *ii)*. Therefore, a biomass depending diffusion coefficient that vanishes in the liquid region will be proposed.

Environmental conditions responsible for the nutrients availability are included through an accurate description of the transport processes in Ω_2, i.e., the hydrodynamic and mass transfer processes.

All things considered, the spatio-temporal variables of the model are

- $t \geq 0$: time,
- $x \in \Omega$: spatial coordinate,
- $c(t, x)$: biomass density,
- $s(t, x)$: nutrients concentration,

- Ω_1: biofilm region, $m > 0$,
- Ω_2: liquid region, $m \equiv 0$,
- $d_i(c)$: diffusion coefficients related to s $(i = 1)$, and c $(i = 2)$,
- $u(t, x)$: fluid velocity,
- $f(s, c)$: nutrients consumption rate,
- $g(s, c)$: biomass production rate,

and so, the model equations are

$$\left.\begin{aligned} &\nabla \cdot u = 0 \\ &\frac{\partial u}{\partial t} + u \cdot \nabla u = -\frac{1}{\rho}\nabla p + \nabla^2 u \qquad \text{for } x \in \Omega_2, \\ &\\ &u \equiv 0 \qquad\qquad\qquad\qquad\quad \text{for } x \in \Omega_1, \end{aligned}\right\} \tag{1.1}$$

$$\frac{\partial s}{\partial t} + u \cdot \nabla s = \nabla \cdot (d_1(c)\nabla s) - f(s, c), \tag{1.2}$$

$$\frac{\partial c}{\partial t} = \nabla \cdot (d_2(c)\nabla c) + g(s, c), \tag{1.3}$$

$$f(s, c) = \frac{k_1 sc}{k_2 + s}, \quad g(s, c) = k_3(f(s, c) - k_4 sc). \tag{1.4}$$

Equations (1.1) correspond to Navier-Stokes equations for a fluid in the bulk region Ω_2, with constant density, ρ, and constant kinematic viscosity, ν [38].

Equation (1.2) describes the nutrients diffusive and convective transport in Ω_2, and in Ω_1, where is only diffusive, and the nutrients consumption. The term $d_1(c)$ corresponds to the diffusion coefficient while the term $f(s, c)$ represents the nutrients consumption rate as described in (1.4).

Equations (1.1) and (1.2) have been widely studied while equation (1.3) is the biomass density evolution equation as proposed in [26]. Spatial propagation is due to a diffusive flux $d_2(c)\nabla c$, with $d_2(c)$ being the biomass dependent diffusion coefficient.

New biomass comes from the term $g(s, c)$, as described in (1.4), including a decay term $-k_3 k_4 c$. As the only biomass source is $g(s, c)$, postulate *iv)* is granted. Also, postulate *v)* is granted by equations (1.1) and (1.2), as long as we obtain a front with a sharp gradient separating the regions Ω_1 and Ω_2.

Since Ω_1 and Ω_2 vary with the evolution of c, it is not possible to decouple the hydrodynamic part of the model, although it can be simplified depending on the experimental conditions, as will be shown later. Parameters k_1 to k_4 are non negative and depend on the particular problem. We assume that $d_1(c)$ is positive, bounded and piecewise differentiable. The biomass diffusion coefficient, $d_2(c)$ must satisfy postulates *i-iv)*. Therefore, following the steps from [26, 68], we assume

$$d_2(c) = \left(\frac{\epsilon}{c_{max} - c} \right)^a c^b, \tag{1.5}$$

where c_{max} is the maximum biomass limit.

It can be seen that the function (1.5) is zero when $c = 0$, i.e., in the bulk, and small when c is much smaller than c_{max}. When c approaches c_{max}, the diffusion function grows giving rise to a non-negligible diffusive transport. Parameter a must be chosen so it satisfies *iii)* while parameters b and ϵ must satisfy *i)* and *ii)*.

1.2.1 Model particularisation and simplifications

Equations (1.1)-(1.5) conform a complete but complicated mathematical model, difficult to analyse or solve. A quick analysis shows some trivial solutions: $c \equiv 0$ is a solution for the model for all values of s, and u; if $s \equiv 0$ or there are not nutrients input to the system, the model will decay to the solution $c \equiv 0$ and $s \equiv 0$ for all u satisfying (1.1). These solutions do not correspond to biofilims behaviour. In the sequel we describe the modifications we will introduce into the model to accommodate the experimental system used. This will lead to the formulation of three different models regarded as M1, M2 and M3, respectively.

First of all, we are trying to simulate the dynamics of the L1A1 strain in a *batch*

system, i.e., we can assume $u \equiv 0$, thus eliminating the complexity of the Navier-Stokes equations. Secondly, let us consider as a first approach that the nutrients diffusivity in the bulk and the biofilm are of the same order of magnitude, thus making $d_1(c) \equiv d_1$, constant. Finally, we are going to solve a dimensionless model, so we introduce the dimensionless variables

$$C := c/c_{max}, \quad S := s/s_0.$$

Therefore, equation (1.2) with reaction term (1.4), constant d_1 and neglected fluid velocity u becomes

$$s_0 \frac{\partial S}{\partial t} = d_1 s_0 \Delta S - \frac{k_1 s_0 S c_{max} C}{k_2 + s_0 S},$$

or equivalently

$$\frac{\partial S}{\partial t} = d_1 \Delta S - \frac{k_1 c_{max}}{s_0} \frac{SC}{\frac{k_2}{s_0} + S}. \tag{1.6}$$

Moreover, equation (1.3) with growth term (1.4) and diffusion term (1.5) becomes

$$c_{max} \frac{\partial C}{\partial t} = \nabla \cdot \left(c_{max}^{b-a} \left(\frac{\epsilon}{1-C} \right)^a C^b c_{max} \nabla C \right) + k_3 \left(\frac{k_1 s_0 S c_{max} C}{k_2 + s_0 S} - k_4 c_{max} C \right),$$

or equivalently,

$$\frac{\partial C}{\partial t} = \nabla \cdot \left(c_{max}^{b-a} \left(\frac{\epsilon}{1-C} \right)^a C^b \nabla C \right) + k_3 k_1 \frac{SC}{\frac{k_2}{s_0} + S} - k_3 k_4 C. \tag{1.7}$$

Next, we define the functions

$$F(S, C) = K_1 \frac{SC}{K_2 + S}, \tag{1.8}$$

$$G(S, C) = K_3 \frac{SC}{K_2 + S} - K_4 C, \tag{1.9}$$

$$d_2(C) = c_{max}^{b-a} \left(\frac{\epsilon}{1-C} \right)^a C^b, \tag{1.10}$$

where the involved parameters, K_1 to K_4, are defined as

$$K_1 = c_{max}\frac{k_1}{s_0}, \quad K_2 = \frac{k_2}{s_0}, \quad K_3 = k_3 k_1, \quad K_4 = k_3 k_4,$$

with

$$k_1 = c_{max}\left(\frac{\mu_m}{Y_{XS}} + m_s\right), \quad k_2 = K_s, \quad k_3 = \frac{Y_{XS}}{c_{max}}, \quad k_4 = m_s c_{max}.$$

and the rest of the involved parameters are listed in Table 1.1. Thus, the simplified dimensionless model can be written as

$$\frac{\partial S}{\partial t} = d_1 \nabla^2 S - F(S, C), \tag{1.11}$$

$$\frac{\partial C}{\partial t} = \nabla \cdot (d_2(C)\nabla C) + G(S, C). \tag{1.12}$$

The model based on (1.11) and (1.12) is completed with initial conditions and boundary conditions. For this, we assume that bacteria adhere initially to the surface located at $x = 0$, with a given thickness of $4.5\mu m$ and that there is neither flux of bacteria nor nutrients. At time zero, a fixed nutrients concentration is fed. Therefore, boundary and initial conditions read as follows

$$C(0, x) = \begin{cases} C_0, \text{ if } 0 \leq x \leq 4.5 \times 10^{-5}, \\ 0, \text{ if } 4.5 \leq x \leq L, \end{cases} \tag{1.13}$$

$$\frac{\partial C}{\partial x}(t, 0) = 0, \ t \in [0, T], \tag{1.14}$$

$$\frac{\partial C}{\partial x}(t, L) = 0, \ t \in [0, T], \tag{1.15}$$

$$S(0, x) = 1, \ x \in [0, L], \tag{1.16}$$

$$\frac{\partial S}{\partial x}(t, 0) = 0, \ t \in [0, T], \tag{1.17}$$

$$\frac{\partial S}{\partial x}(t, L) = 0, \ t \in [0, T]. \tag{1.18}$$

Param.	Description	Used in
d_1	Nutrients diffusion coefficient (m^2/s)	M1, M2
d_2	Biomass diffusion coefficient (m^2/s)	M3
d_N	Nutrients bulk diffusion coefficient (m^2/s)	M3
μ_m	Specific growth rate $(1/s)$	M1, M2, M3
Y_{XS}	Substrate growth yield (-)	M1, M2, M3
K_s	Monod saturation constant (kg/m^3)	M1, M2
K_d	Monod decay constant (kg/m^3)	M2
k_d	Rate of activation of detachment (-)	M3
D_{min}	% of damaged or dead cells before detachment (-)	M3
N_{min}	Threshold for glucose impaired uptake (-)	M3
m_s	Maintenance coefficient $(1/s)$	M1, M2, M3
c_{max}	Maximum biomass (kg/m^3)	M1, M2, M3
s_0	Initial nutrients concentration (kg/m^3)	M1, M2, M3
C_0	Initial biomass concentration (kg/m^3)	M1, M2, M3
ϵ	Biomass diffusivity related constant (-)	M1, M2
a	Biomass diffusivity related constant (-)	M1, M2
b	Biomass diffusivity related constant (-)	M1, M2
L	Maximum length (m)	M1, M2, M3
T	Maximum time (s)	M1, M2, M3

Table 1.1: Parameters involved in the different models.

In summary, the so called Eberl model is given by equations (1.11)-(1.18), with the expressions for the involved functions given in (1.8)-(1.10). This formulation corresponds to M1. The results concerning the mathematical analysis of the Eberl model to prove the existence and uniqueness of solution are obtained in [28] and recalled in Appendix A.

1.3 Second model: non-linear detachment

As will be shown in the experimental validation in Chapter 3, the previous model can not predict L1A1 biofilm life cycle. Actually, the Eberl model predicts that cells consume all nutrients. Once the nutrients are exhausted, the cells enter a phase

of maintenance in which the thickness does not increase. Finally, the decay term dominates the growth term, and biomass depletion starts once cells maintenance is no longer possible. The maximum biofilm thickness is around the maximum found by quantitative image, although the final decay is almost imperceptible in contrast to the massive detachment observed in the experiments. In order to account for this massive final decay, in this second model we modify the decay term.

The original term K_4C is meant to represent the decay caused only by the fluid friction. However, in our case we first assumed that the decay is due to the fact that at certain moment, the nutrients are fully consumed or cannot reach the inner zones of the biofilm, so cells start to die, the EPS weakens and a massive decay happens. In the second model we change the linear decay term into a non-linear rational term, so the second model becomes

$$\frac{\partial S}{\partial t} = d_1\nabla^2 S - F(S,C), \tag{1.19}$$

$$\frac{\partial C}{\partial t} = \nabla\cdot(d_2(C)\nabla C) + G(S,C), \tag{1.20}$$

where

$$F(S,C) = K_1\frac{SC}{K_2+S}, \tag{1.21}$$

$$G(S,C) = K_3\frac{SC}{K_2+S} - K_4\frac{C}{K_d+S}, \tag{1.22}$$

$$d_2(C) = c_{max}^{b-a}\left(\frac{\epsilon}{1-C}\right)^a C^b. \tag{1.23}$$

The new parameter K_d is the decay coefficient related to the nutrients concentration, and parameters K_1 to K_4 are defined as

$$K_1 = c_{max}\frac{k_1}{s_0}, \quad K_2 = \frac{k_2}{s_0}, \quad K_3 = k_3k_1, \quad K_4 = k_3k_4,$$

with

$$k_1 = c_{max}\left(\frac{\mu_m}{Y_{XS}} + m_s\right), \quad k_2 = K_s, \quad k_3 = \frac{Y_{XS}}{c_{max}}, \quad k_4 = m_s c_{max}.$$

The rest of the involved parameters are listed in Table 1.1. The model maintains the same boundary and initial conditions (1.13)-(1.18).

In summary, the so called second model or model with non-linear detachment (M2), is given by equations (1.19)-(1.20), with the expressions for the involved functions given in (1.21)-(1.23), and the boundary conditions (1.13)-(1.18).

1.4 Third model: nutrient uptake impairment plus ageing

The second model provides a numerical solution where a significant final decay is observed. However, the decay starts rather early and the model does not achieve the expected maximum thickness. Besides, although slower than in the Eberl model, nutrients are still fully consumed while in the real biological system nutrients are still present in the bulk.

Several candidate models where tested including various mechanisms [9]. Here we regard as the third model the most successful of them. In the third model, we drop the assumption that the nutrients diffusion coefficient is constant, and instead we consider the more realistic scenario in which nutrients diffuse differently in the bulk and in the biofilm [94]. Therefore, we introduce for d_1 the dependence on C in the form:

$$d_1(C) = \begin{cases} d_N, \text{ if } C = 0, \\ d_{eff} d_N, \text{ if } C > 0, \end{cases} \tag{1.24}$$

where d_N is the glucose diffusivity in the bulk and d_{eff} is the effective glucose diffusivity within the biofilm, its value being $d_{eff} = 0.24$ as stated in [94].

Since the calibration of models M1 and M2 led to the conclusion that non-linear biomass diffusion is not relevant, a constant diffusion was considered. Furthermore, since there are plenty of nutrients in the medium, we assume that biomass growth (and therefore nutrients consumption), follows a mass action law description instead of the usual Monod formulation. This assumption was also confirmed by parameter estimation. The behaviour in the laboratory experiments showed that instead of

depleting the system of all of the nutrients, at a certain time the biomass stopped consuming them, so that a nutrients uptake impairment mechanism was introduced.

Finally, we assume that the large detachment observed in the laboratory experiments is related to cells death and the degradation of the extracellular DNA, i.e., to biofilm ageing. Therefore, as a measure of ageing, we used the measured covered area of damaged or dead cells, $CBD(t)$, in such a way that the decay starts once a given value of damaged or dead cells appears in the biofilm. Mathematically, this can be written as

$$G(S,C) = K_3SC - K_4\frac{C}{1+\exp(k_d[D_{min} - CBD(t)])}.$$

The resulting so called third model is expressed as follows

$$\frac{\partial S}{\partial t} = \nabla\cdot(d_1(C)\nabla S) - F(S,C), \tag{1.25}$$

$$\frac{\partial C}{\partial t} = d_2\nabla^2 C + G(S,C), \tag{1.26}$$

where

$$F(S,C) = K_1SC, \tag{1.27}$$

$$G(S,C) = K_3SC - K_4\frac{C}{1+\exp(k_d[D_{min} - CBD(t)])}. \tag{1.28}$$

The new parameter k_d represents the rate of activation of detachment, D_{min} is the percentage of damaged or dead cells before detachment and parameters K_1, K_3 and K_4 are defined as follows

$$K_1 = c_{max}\frac{k_1}{s_0}, \quad K_3 = k_3k_1, \quad K_4 = k_3k_4,$$

with

$$k_1 = c_{max}\left(\frac{\mu_m}{Y_{XS}} + m_s\right), \quad k_3 = \frac{Y_{XS}}{c_{max}}, \quad k_4 = m_sc_{max}.$$

The rest of the involved parameters are listed in Table 1.1. The model maintains the same boundary and initial conditions (1.13)-(1.18).

In summary, the so called third model (M3), is given by equations (1.25)-(1.26), with the expressions for the involved functions given in (1.27)-(1.28), and the boundary conditions (1.13)-(1.18).

Chapter 2

Numerical methods for the 1D case

2.1 Numerical methods

Sections 1.2.1, 1.3 and 1.4 present three different modelling approaches, all of them consisting of a system of non-linear partial differential equations, the analytical solution of which is not available. Therefore, appropriate numerical methods are required to efficiently compute the approximation of the solution. For the first model, Eberl et al. [26] proposed a combination of explicit and implicit finite difference methods that have been used to solve the system of equations in Section 1.2.1. More precisely, an explicit time-stepping was used in (1.11) while a fully implicit scheme was applied to (1.12). Note that the explicit scheme involves a stability constraint on the time step, thus limiting the use of large time steps. In both equations, a classical centred scheme is used in the spatial discretisation. The resulting non-linear system associated to the fully discretised problem was solved with a Newton-BiCGSTAB method. The main drawback of this numerical strategy comes from the use of an explicit scheme that requires a large number of time steps to obtain the concentration of nutrients and biomass at final time, thus leading to an extremely slow algorithm. Additionally, it should be noted that the finite-differences scheme results to be first order in time.

In order to avoid the stability constraint, in a previous work [60] the fully implicit scheme was applied in combination with a Newton algorithm for the non-linear fully

discretised problem. However, this approach led to very large computational times in order to properly approximate the slow non-linear diffusion processes that required small time steps.

As a more efficient alternative, in this PhD work we propose the use of a Crank-Nicolson finite difference scheme, which is second order in time and space [57], so that larger time steps can be used for the same accuracy and thus the computational time can be reduced.

For this purpose, we first consider a finite differences mesh, with N time intervals, and J spatial nodes, in terms of which we define the time and spatial steps, respectively, as

$$\Delta t = \frac{T}{N}, \quad \Delta x = \frac{L}{J-1}.$$

Moreover, we introduce the notation $S_j^n \approx S(n\Delta t, j\Delta x)$ and $C_j^n \approx C(n\Delta t, j\Delta x)$ to implement the finite differences approximations of the solution at the mesh points, with $n = 0, \ldots, N$ y $j = 0, \ldots, J - 1$.

Equations (1.11) and (1.12) in Section 1.2.1 are discretised with a Crank-Nicolson finite differences scheme in time and a central differences scheme for the second order derivative in space. The same process is applied to equations (1.19) and (1.20) in section 1.3 and to equations (1.25) and (1.26) in section 1.4.

2.1.1 Discretisation of the Eberl model

First, for $n = 0, \ldots, N$, applying Crank-Nicolson plus a central differences scheme to the spatial second derivative to equation (1.11), leads to the system of equations

$$\begin{aligned} \frac{S_j^{n+1} - S_j^n}{\Delta t} = \\ \frac{1}{2}\left[d_1 \frac{(S_{j+1}^{n+1} - 2S_j^{n+1} + S_{j-1}^{n+1})}{(\Delta x)^2} - F(S_j^{n+1}, C_j^{n+1})\right] + \\ \frac{1}{2}\left[d_1 \frac{(S_{j+1}^{n} - 2S_j^{n} + S_{j-1}^{n})}{(\Delta x)^2} - F(S_j^{n}, C_j^{n})\right], \end{aligned} \tag{2.1}$$

valid for $j = 1, \ldots, J-2$. The nodes corresponding to $j = 0$ and $j = J-1$ will be discretised taking into account the boundary conditions, as will be explaind in Section 2.2. By rearranging the terms in the previous equation, we get

$$\begin{aligned} S_j^{n+1} - S_j^n = \\ \frac{d_1 \Delta t}{2(\Delta x)^2}(S_{j+1}^{n+1} - 2S_j^{n+1} + S_{j-1}^{n+1}) - \frac{\Delta t K_1 S_j^{n+1} C_j^{n+1}}{2(K_2 + S_j^{n+1})} + \\ \frac{d_1 \Delta t}{2(\Delta x)^2}(S_{j+1}^{n} - 2S_j^{n} + S_{j-1}^{n}) - \frac{\Delta t K_1 S_j^{n} C_j^{n}}{2(K_2 + S_j^{n})}, \end{aligned} \tag{2.2}$$

or equivalently

$$\begin{aligned} -\omega S_{j+1}^{n+1} + (1+2\omega) S_j^{n+1} - \omega S_{j-1}^{n+1} + \sigma_1 \frac{S_j^{n+1} C_j^{n+1}}{K_2 + S_j^{n+1}} + \\ \left[-\omega S_{j+1}^{n} + (2\omega - 1) S_j^{n} - \omega S_{j-1}^{n} + \sigma_1 \frac{S_j^{n} C_j^{n}}{K_2 + S_j^{n}} \right] = 0, \end{aligned} \tag{2.3}$$

with $\omega = \dfrac{d_1 \Delta t}{2(\Delta x)^2}$ and $\sigma_1 = \dfrac{\Delta t K_1}{2}$.

Next, by applying a Crank-Nicolson plus a central scheme to approximate the most external spatial partial derivatives in equation (1.12), we get

$$\begin{aligned} \frac{C_j^{n+1} - C_j^n}{\Delta t} = \\ \frac{1}{2}\left[\frac{d_2(C_{j+1}^{n+1})(C_{j+1}^{n+1})_x - d_2(C_{j-1}^{n+1})(C_{j-1}^{n+1})_x}{2\Delta x} + G(S_j^{n+1}, C_j^{n+1}) \right] + \\ \frac{1}{2}\left[\frac{d_2(C_{j+1}^{n})(C_{j+1}^{n})_x - d_2(C_{j-1}^{n})(C_{j-1}^{n})_x}{2\Delta x} + G(S_j^{n}, C_j^{n}) \right]. \end{aligned}$$

As d_2 is positive, we use a backward scheme in space in the remaining spatial

derivatives in the previous equation, thus obtaining

$$\frac{C_j^{n+1}-C_j^n}{\Delta t} = \frac{1}{2}\left[\frac{d_2(C_{j+1}^{n+1})(C_{j+1}^{n+1}-C_j^{n+1})-d_2(C_{j-1}^{n+1})(C_j^{n+1}-C_{j-1}^{n+1})}{2(\Delta x)^2}+G(S_j^{n+1},C_j^{n+1})\right]+ \frac{1}{2}\left[\frac{d_2(C_{j+1}^{n})(C_{j+1}^{n}-C_j^{n})-d_2(C_{j-1}^{n})(C_j^{n}-C_{j-1}^{n})}{2(\Delta x)^2}+G(S_j^{n},C_j^{n})\right].$$

Furthermore, applying the definition of d_2, we have

$$\begin{aligned} C_j^{n+1}-C_j^n = {} & \frac{\Delta t m_{max}^{b-a}\epsilon^a}{4(\Delta x)^2}\left[\frac{(C_{j+1}^{n+1})^{b+1}-(C_{j+1}^{n+1})^bC_j^{n+1}}{(1-C_{j+1}^{n+1})^a}-\frac{(C_{j-1}^{n+1})^bC_j^{n+1}-(C_{j-1}^{n+1})^{b+1}}{(1-C_{j-1}^{n+1})^a}\right]+ \\ & \frac{\Delta t}{2}G(S_j^{n+1},C_j^{n+1})+ \\ & \frac{\Delta t m_{max}^{b-a}\epsilon^a}{4(\Delta x)^2}\left[\frac{(C_{j+1}^{n})^{b+1}-(C_{j+1}^{n})^bC_j^{n}}{(1-C_{j+1}^{n})^a}-\frac{(C_{j-1}^{n})^bC_j^{n}-(C_{j-1}^{n})^{b+1}}{(1-C_{j-1}^{n})^a}\right]+ \\ & \frac{\Delta t}{2}G(S_j^{n},C_j^{n}). \end{aligned} \tag{2.4}$$

Finally, by introducing the definition of G and rearranging terms, we have

$$\begin{aligned} & (1+\sigma_4)\,C_j^{n+1}-\sigma_3\frac{S_j^{n+1}C_j^{n+1}}{K_2+S_j^{n+1}}+ \\ & \mu\left[\frac{(C_{j+1}^{n+1})^bC_j^{n+1}-(C_{j+1}^{n+1})^{b+1}}{(1-C_{j+1}^{n+1})^a}+\frac{(C_{j-1}^{n+1})^bC_j^{n+1}-(C_{j-1}^{n+1})^{b+1}}{(1-C_{j-1}^{n+1})^a}\right]+ \\ & (\sigma_4-1)\,C_j^{n}-\sigma_3\frac{S_j^{n}C_j^{n}}{K_2+S_j^{n}}+ \\ & \mu\left[\frac{(C_{j+1}^{n})^bC_j^{n}-(C_{j+1}^{n})^{b+1}}{(1-C_{j+1}^{n})^a}+\frac{(C_{j-1}^{n})^bC_j^{n}-(C_{j-1}^{n})^{b+1}}{(1-C_{j-1}^{n})^a}\right]=0, \end{aligned} \tag{2.5}$$

with $\mu=\dfrac{\Delta t c_{max}^{b-a}\epsilon^a}{4(\Delta x)^2}$, $\sigma_3=\dfrac{\Delta t K_3}{2}$ and $\sigma_4=\dfrac{\Delta t K_4}{2}$.

2.1.2 Discretisation of the non-linear detachment model

The discretisation process of equations (1.19)-(1.20) follows the same strategy as in equations (1.11)-(1.12). In fact, as both functions (1.8) and (1.21) are the same, the nutrients equation (1.19) leads to the discretised equation

$$-\omega S_{j+1}^{n+1} + (1+2\omega)S_j^{n+1} - \omega S_{j-1}^{n+1} + \sigma_1 \frac{S_j^{n+1} C_j^{n+1}}{K_2 + S_j^{n+1}} + \left[-\omega S_{j+1}^{n} + (2\omega - 1)S_j^{n} - \omega S_{j-1}^{n} + \sigma_1 \frac{S_j^{n} C_j^{n}}{K_2 + S_j^{n}} \right] = 0, \tag{2.6}$$

with $\omega = \dfrac{d_1 \Delta t}{2(\Delta x)^2}$ and $\sigma_1 = \dfrac{\Delta t K_1}{2}$, for $j = 0, \ldots, J-1$.

On the other hand, the term $G(S, C)$ used in equations (1.12) and (1.20) differs. The discretisation is the same until equation (2.4), where the definition of G is applied and the resulting equation changes. By applying the definition of $G(S, C)$, as regarded in (1.22), equation (2.4) becomes

$$C_j^{n+1} - \sigma_3 \frac{S_j^{n+1} C_j^{n+1}}{K_2 + S_j^{n+1}} + \sigma_4 \frac{C_j^{n+1}}{K_d + S_j^{n+1}} + \mu \left[\frac{(C_{j+1}^{n+1})^b C_j^{n+1} - (C_{j+1}^{n+1})^{b+1}}{(1 - C_{j+1}^{n+1})^a} + \frac{(C_{j-1}^{n+1})^b C_j^{n+1} - (C_{j-1}^{n+1})^{b+1}}{(1 - C_{j-1}^{n+1})^a} \right] - C_j^{n} - \sigma_3 \frac{S_j^{n} C_j^{n}}{K_2 + S_j^{n}} + \sigma_4 \frac{C_j^{n}}{K_d + S_j^{n}} + \mu \left[\frac{(C_{j+1}^{n})^b C_j^{n} - (C_{j+1}^{n})^{b+1}}{(1 - C_{j+1}^{n})^a} + \frac{(C_{j-1}^{n})^b C_j^{n} - (C_{j-1}^{n})^{b+1}}{(1 - C_{j-1}^{n})^a} \right] = 0, \tag{2.7}$$

where $\mu = \dfrac{\Delta t c_{max}^{b-a} \epsilon^a}{4(\Delta x)^2}$, $\sigma_3 = \dfrac{\Delta t K_3}{2}$ and $\sigma_4 = \dfrac{\Delta t K_4}{2}$.

2.1.3 Discretisation of the nutrients uptake impairment plus ageing model

Also, the discretisation process of equations (1.25)-(1.26) follows the same strategy as in equations (1.11)-(1.12). In this case, functions $F(S, C)$ and $G(S, C)$, as well as

the nutrients diffusion coefficient $d_1(C)$ change, so both equations must be analysed.

Let us start with equation (1.25) where the main difference is in the diffusion coefficient, $d_1(C)$, given by (1.24) which is not constant unlike the constant coefficient taken in (1.11). The term $F(S,C)$ also differs although it does not change the discretisation techniques that much. Therefore, using a Crank-Nicolson scheme for the time derivative, we can consider

$$\frac{S_j^{n+1}-S_j^n}{\Delta t} = \frac{1}{2}\left[d_1(C_j^n)\nabla^2 S_j^{n+1} - F(S_j^{n+1},C_j^{n+1})\right] + \frac{1}{2}\left[d_1(C_j^n)\nabla^2 S_j^n - F(S_j^n,C_j^n)\right],$$

where we use the diffusion coefficient at the previous time step, i.e.,

$$d_1(C_j^n) = \begin{cases} d_N, & \text{if } C_j^n \le C_{min}, \\ d_{eff}d_N, & \text{if } C_j^n > C_{min}, \end{cases} \tag{2.8}$$

for $j = 0,\ldots,J-1$. The value C_{min} is used to set a certain minimum of biomass concentration. If the numerical value of the biomass concentration is smaller than C_{min}, then we consider that it is zero, and proceed with the diffusion coefficient d_N, which is the diffusion coefficient in the bulk. Note that in our examples we choose $C_{min} = 10^{-5}$.

Next, we apply a central differences scheme to the spatial derivative, thus obtaining

$$\frac{S_j^{n+1}-S_j^n}{\Delta t} = \frac{1}{2}\left[d_1(C_j^n)\frac{(S_{j+1}^{n+1}-2S_j^{n+1}+S_{j-1}^{n+1})}{(\Delta x)^2} - F(S_j^{n+1},C_j^{n+1})\right] + \frac{1}{2}\left[d_1(C_j^n)\frac{(S_{j+1}^{n}-2S_j^{n}+S_{j-1}^{n})}{(\Delta x)^2} - F(S_j^{n},C_j^{n})\right],$$

and finally, applying (1.27) we have the following discretised equation

$$-\omega S_{j+1}^{n+1} + (1+2\omega)S_j^{n+1} - \omega S_{j-1}^{n+1} + \sigma_1 S_j^{n+1}C_j^{n+1} + \left[-\omega S_{j+1}^n + (2\omega-1)S_j^n - \omega S_{j-1}^n + \sigma_1 S_j^n C_j^n\right] = 0, \tag{2.9}$$

where $\omega \equiv \omega(C_j^n) = \dfrac{d_1(C_j^n)\Delta t}{2(\Delta x)^2}$ and $\sigma_1 = \dfrac{\Delta t K_1}{2}$.

Now, for equation (1.26) we follow a similar procedure as for equation (1.11). In equation (1.26) the diffusion coefficient is constant $d_2(C) \equiv d_2$ and the term $G(S, C)$, is

$$G(S,C) = K_3 SC - K_4 \frac{C}{1 + \exp(k_d[D_{min} - CBD(t)])},$$

where the values of $CBD(t)$ are an input to the model and, thus, are treated as known values. Applying a Crank-Nicolson finite differences scheme in time and a central differences scheme for the second order derivative in space to equation (1.26), we get

$$\begin{aligned} \frac{C_j^{n+1} - C_j^n}{\Delta t} = \\ \frac{1}{2}\left[d_2 \frac{(C_{j+1}^{n+1} - 2C_j^{n+1} + C_{j-1}^{n+1})}{(\Delta x)^2} + G(S_j^{n+1}, C_j^{n+1})\right] + \\ \frac{1}{2}\left[d_2 \frac{(C_{j+1}^{n} - 2C_j^{n} + C_{j-1}^{n})}{(\Delta x)^2} + G(S_j^{n}, C_j^{n})\right], \end{aligned}$$

for $j = 0, \ldots, J-1$. Rearranging terms, we get

$$\begin{aligned} C_j^{n+1} - C_j^n = \\ \frac{d_2 \Delta t}{2(\Delta x)^2}(C_{j+1}^{n+1} - 2C_j^{n+1} + C_{j-1}^{n+1}) + \frac{\Delta t}{2} G(S_j^{n+1}, C_j^{n+1}) + \\ \frac{d_2 \Delta t}{2(\Delta x)^2}(C_{j+1}^{n} - 2C_j^{n} + C_{j-1}^{n}) + \frac{\Delta t}{2} G(S_j^{n}, C_j^{n}), \end{aligned}$$

and, applying (1.28),

$$\begin{aligned} -\mu C_{j+1}^{n+1} + (1 + 2\mu + \sigma_{CBD}) C_j^{n+1} - \mu C_{j-1}^{n+1} - \sigma_3 S_j^{n+1} C_j^{n+1} + \\ \left[-\mu C_{j+1}^{n} + (2\mu + \sigma_{CBD} - 1) C_j^{n} - \mu C_{j-1}^{n} - \sigma_3 S_j^{n} C_j^{n}\right] = 0, \end{aligned} \tag{2.10}$$

with $\mu = \dfrac{d_2 \Delta t}{2(\Delta x)^2}$, $\sigma_3 = \dfrac{\Delta t K_3}{2}$ and $\sigma_{CBD} = \dfrac{\Delta t K_4}{2 + 2\exp(k_d[D_{min} - CBD(t)])}$.

2.2 Discretisation of the boundary conditions

The discretised equations (2.3)-(2.5) for the Alpkvist model, (2.6)-(2.7) for the second model and (2.9)-(2.10) for the third model, stand for nodes $j = 1, \ldots, J-2$. Nodes $j = 0$ and $j = J-1$ would require ghost data in order to be solved using the these equations. Instead, we are applying a left- or right-biased finite difference scheme of order 2 to discretise the boundary conditions. That way, we obtain equations that stand for nodes $j = 0$ and $j = J-1$. Therefore, let us start with equation (1.14). By applying a right-biased finite differences scheme, we obtain

$$\frac{-3C_0^n + 4C_1^n - C_2^n}{2\Delta x} = 0 \Leftrightarrow 3C_0^n - 4C_1^n + C_2^n = 0. \tag{2.11}$$

Next, by applying a left-biased finite differences scheme to equation (1.15), we obtain

$$\frac{3C_{J-1}^n - 4C_{J-2}^n + C_{J-3}^n}{2\Delta x} = 0 \Leftrightarrow 3C_{J-1}^n - 4C_{J-2}^n + C_{J-3}^n = 0. \tag{2.12}$$

In a similar way, by applying a right biased finite differences scheme to equation (1.17), we obtain

$$\frac{-3S_0^n + 4S_1^n - S_2^n}{2\Delta x} = 0 \Leftrightarrow 3S_0^n - 4S_1^n + S_2^n = 0. \tag{2.13}$$

Finally, by applying a left-biased finite differences scheme to equation (1.18), we obtain

$$\frac{3S_{J-1}^n - 4S_{J-2}^n + S_{J-3}^n}{2\Delta x} = 0 \Leftrightarrow 3S_{J-1}^n - 4S_{J-2}^n + S_{J-3}^n = 0. \tag{2.14}$$

As the boundary conditions (1.14)-(1.15) and (1.17)-(1.18) are valid for the three models, equations (2.11)-(2.14) can be used to complete the discrete system used to solve each of the models.

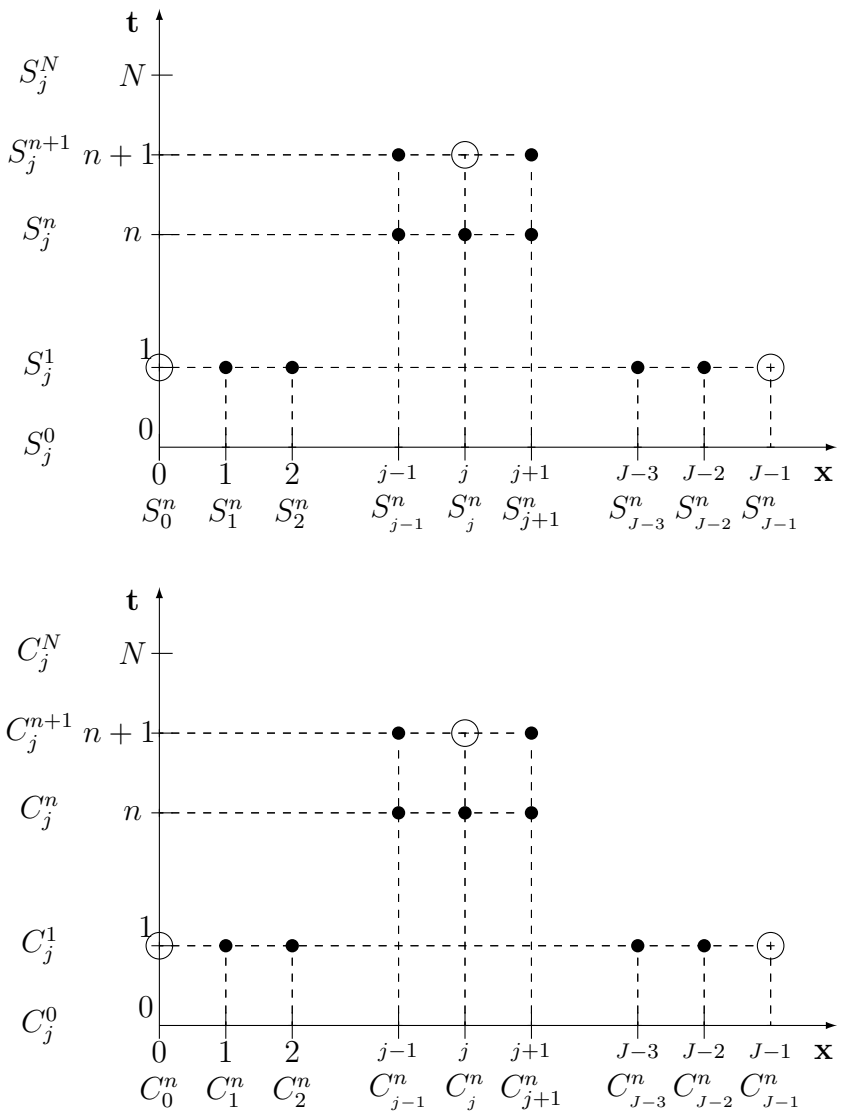

Figure 2.1: Different nodes used in the discretisation of the equations of the three models

2.3 Fully discretised problems

Once the boundary conditions are taken into account, each of the three discretised models leads to a non-linear system that must be solved through a linearisation

method. In all three cases, the non-linear system can be expressed as

$$\mathcal{F}(\mathbb{X}^{n+1}) = A(\mathbb{X}^{n+1})\mathbb{X}^{n+1} - B(\mathbb{X}^n) = 0, \tag{2.15}$$

where the definition of the function $\mathcal{F}$ depends on the particular model and boundary conditions being considered. The elements of the system are

- $\mathbb{X}^{n+1}$ is the vector of unknowns at time step n+1, the $2J$ elements of which are the components of the vectors $S^{n+1} = (S_j^{n+1})_{j=0}^{J-1}$ and $C^{n+1} = (C_j^{n+1})_{j=0}^{J-1}$, which are ordered so that the system results as simple as possible:

$$\mathbb{X}^{n+1} = \left(S_0^{n+1}, C_0^{n+1}, \ldots, S_j^{n+1}, C_j^{n+1}, \ldots, S_{J-1}^{n+1}, C_{J-1}^{n+1}\right)^T. \tag{2.16}$$

- The vector $B(\mathbb{X}^n) \equiv B$ is constant through the iterations of the Newton method when solving. Moreover, its $2J$ elements are given by each particular model equations.

- Matrix $A(\mathbb{X}^{n+1})$ is a square sparse matrix of order $2J$, the coefficients of which depend on the unknowns $\mathbb{X}^{n+1}$. As with vector B, its elements depend on each particular model equations.

As the function $\mathcal{F}$ is non-linear, in order to solve (2.15) we use a Newton method [30] at each time step n. Thus, starting from $\mathbb{X}^{n,0} = \mathbb{X}^n$, at the iteration p we compute $\mathbb{X}^{n,p+1}$ as follows:

$$J_{\mathcal{F}}(\mathbb{X}^{n,p})\Delta\mathbb{X}^{n,p} = \mathcal{F}(\mathbb{X}^{n,p}), \tag{2.17}$$

$$\mathbb{X}^{n,p+1} = \mathbb{X}^{n,p} - \Delta\mathbb{X}^{n,p}, \tag{2.18}$$

where $J_{\mathcal{F}}(\mathbb{X}^{n,p})$ denotes the jacobian matrix of $\mathcal{F}$ evaluated at $\mathbb{X}^{n,p}$. At each Newton iteration, the solution of the linear system (2.17), $\Delta\mathbb{X}^{n,p}$, is obtained using an optimized LU factorization with partial pivoting [80]. Finally, we find $\mathbb{X}^{n+1}$ through the calculation of the sequence $\mathbb{X}^{n,p}$, until achieving an index, k, so that the required relative error tolerance between two consecutive iterations is fulfilled. Then we take $\mathbb{X}^{n+1} = \mathbb{X}^{n,k}$.

2.3.1 Fully discretised problem for the Eberl model

The set of discretised equations of the Eberl model (2.3)-(2.5), (2.11)-(2.12) and (2.13)-(2.14) particularises the system of non-linear equations (2.15), so that, on one hand the independent vector B is given by

$$B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \omega S_{j+1}^n - (2\omega - 1)S_j^n + \omega S_{j-1}^n - \sigma_1 \dfrac{S_j^n C_j^n}{K_2 + S_j^n} \\ -(\sigma_4 - 1)\, C_j^n - \mu \left[\dfrac{(C_{j+1}^n)^b C_j^n - (C_{j+1}^n)^{b+1}}{(1C_{j+1}^n)^a} + \dfrac{(C_{j-1}^n)^b C_j^n - (C_{j-1}^n)^{b+1}}{(1 - C_{j-1}^n)} \right] + \sigma_3 \dfrac{S_j^n C_j^n}{K_2 + S_j^n} \\ \vdots \\ 0 \\ 0 \end{pmatrix}. \tag{2.19}$$

On the other hand, matrix A can be expressed in the form

$$A(\mathbb{X}^{n+1}) = A^c + A^{nc}(\mathbb{X}^{n+1}),$$

with A^c independent of $\mathbb{X}^{n+1}$ and $A^{nc}(\mathbb{X}^{n+1})$ holding all the dependency on $\mathbb{X}^{n+1}$ and, therefore gathering the non-linearity of the system. For simplicity, let us denote $A^{nc}(\mathbb{X}^{n+1}) \equiv A^{nc}$.

Both matrices, A^c and A^{nc}, are sparse and have a certain block structure. More precisely, next we describe their respective block structures:

- Matrix A^c

$$A^c = \left(\begin{array}{c|c|c|c|c|c|c} \gamma_1 & \gamma_2 & \gamma_3 & 0\dots & & & \\ \hline \alpha & \beta & \alpha & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha & \beta & \alpha & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha & \beta & \alpha \\ \hline & & \dots & 0 & \gamma_3 & \gamma_2 & \gamma_1 \end{array}\right), \tag{2.20}$$

where the involved 2×2 block submatrices are given by

$$\alpha = \left(\begin{array}{c|c} -\omega & 0 \\ \hline 0 & 0 \end{array}\right), \beta = \left(\begin{array}{c|c} 1+2\omega & 0 \\ \hline 0 & 1+\sigma_4 \end{array}\right),$$

$$\gamma_1 = \left(\begin{array}{c|c} 3 & 0 \\ \hline 0 & 3 \end{array}\right), \gamma_2 = \left(\begin{array}{c|c} -4 & 0 \\ \hline 0 & -4 \end{array}\right), \gamma_3 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}\right).$$

- Matrix A^{nc}

$$A^{nc} = \left(\begin{array}{c|c|c|c|c|c|c} 0 & 0 & 0 & \dots & & & \\ \hline \alpha_1^2 & \beta_1 & \alpha_1^1 & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha_j^2 & \beta_j & \alpha_j^1 & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha_{J-2}^2 & \beta_{J-2} & \alpha_{J-2}^1 \\ \hline & & & \dots & 0 & 0 & 0 \end{array}\right), \tag{2.21}$$

where the involved 2×2 block submatrices are given by

  - Block α_j^1, for $j = 1, \dots, J-2$:

$$\alpha_j^1 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & -\mu \dfrac{(C_{j+1}^{n+1})^b}{(1-C_{j+1}^{n+1})^a} \end{array}\right).$$

– Block α_j^2, for $j = 1, \ldots, J-2$:

$$\alpha_j^2 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & -\mu\dfrac{(C_{j-1}^{n+1})^b}{(1-C_{j-1}^{n+1})^a} \end{array}\right).$$

– Block β_j, for $j = 1, \cdots, J-2$:

$$\beta_j = \left(\begin{array}{c|c} \sigma_1\dfrac{C_j^{n+1}}{K_2+S_j^{n+1}} & 0 \\ \hline 0 & \mu\left[\dfrac{(C_{j+1}^{n+1})^b}{(1-C_{j+1}^{n+1})^a}+\dfrac{(C_{j-1}^{n+1})^b}{(1-C_{j-1}^{n+1})^a}\right]-\sigma_3\dfrac{S_j^{n+1}}{K_2+S_j^{n+1}} \end{array}\right).$$

Linear system for each Newton iteration

With the non-linear function $\mathcal{F}$ described, we need to build the linear system (2.17). The Jacobian matrix, $J_{\mathcal{F}}$ is evaluated in every Newton iteration at $\mathbb{X}^{n,p}$. It can be also decomposed in a constant part plus a non constant part

$$J_{\mathcal{F}}(\mathbb{X}^{n,p}) = A^c + \overline{A}(\mathbb{X}^{n,p}),$$

where A^c is the matrix (2.20). Matrix $\overline{A}(\mathbb{Y})$ is a square sparse matrix of order $2J$. More precisely, let us describe its block structure,

$$\overline{A}(\mathbb{Y}) = \left(\begin{array}{c|c|c|c|c|c|c} 0 & 0 & 0 & \dots & & & \\ \hline \overline{\alpha}_1^2 & \overline{\beta}_1 & \overline{\alpha}_1^1 & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \overline{\alpha}_j^2 & \overline{\beta}_j & \overline{\alpha}_j^1 & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \overline{\alpha}_{J-2}^2 & \overline{\beta}_{J-2} & \overline{\alpha}_{J-2}^1 \\ \hline & & & \dots & 0 & 0 & 0 \end{array}\right), \tag{2.22}$$

where the involved 2×2 block submatrices are given by

- Block $\overline{\alpha}_j^1$, for $j = 1, \dots, J-2$:

$$\overline{\alpha}_j^1 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \overline{\alpha}_{j,4}^1 \end{array}\right),$$

where

$$\overline{\alpha}_{j,4}^1 = \mu \frac{\left[b\mathbb{Y}_{2j+3}^{b-1}\mathbb{Y}_{2j+1} - (b+1)\mathbb{Y}_{2j+3}^b\right](1-\mathbb{Y}_{2j+3})^a + a(1-\mathbb{Y}_{2j+3})^{a-1}(\mathbb{Y}_{2j+3}^b\mathbb{Y}_{2j+1} - \mathbb{Y}_{2j+3}^{b+1})}{(1-\mathbb{Y}_{2j+3})^{2a}}.$$

- Block $\overline{\alpha}_j^2$, for $j = 1, \dots, J-2$:

$$\overline{\alpha}_j^2 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \overline{\alpha}_{j,4}^2 \end{array}\right),$$

where

$$\overline{\alpha}_{j,4}^2 = \mu \frac{\left[b\mathbb{Y}_{2j-1}^{b-1}\mathbb{Y}_{2j+1} - (b+1)\mathbb{Y}_{2j-1}^b\right](1-\mathbb{Y}_{2j-1})^a + a(1-\mathbb{Y}_{2j-1})^{a-1}(\mathbb{Y}_{2j-1}^b\mathbb{Y}_{2j+1} - \mathbb{Y}_{2j-1}^{b+1})}{(1-\mathbb{Y}_{2j-1})^{2a}}.$$

- Block $\overline{\beta}_j$, for $j = 1, \cdots, J-2$:

$$\overline{\beta}_j = \left(\begin{array}{c|c} \sigma_1 \dfrac{\mathbb{Y}_{2j+1} K_2}{(K_2 + \mathbb{Y}_{2j})^2} & \sigma_1 \dfrac{\mathbb{Y}_{2j}}{K_2 + \mathbb{Y}_{2j}} \\ \hline -\sigma_3 \dfrac{\mathbb{Y}_{2j+1} K_2}{(K_2 + \mathbb{Y}_{2j})^2} & \mu \left[\dfrac{\mathbb{Y}^b_{2j+3}}{(1 - \mathbb{Y}_{2j+3})^a} + \dfrac{\mathbb{Y}^b_{2j-1}}{(1 - \mathbb{Y}_{2j-1})^a} \right] - \sigma_3 \dfrac{\mathbb{Y}_{2j}}{K_2 + \mathbb{Y}_{2j}} \end{array} \right).$$

Using (2.19), (2.20), (2.21) and (2.22), we get the system of equations (2.17)-(2.18)

$$\begin{aligned} J_{\mathcal{F}}(\mathbb{X}^{n,p}) \Delta \mathbb{X}^{n,p} &= \mathcal{F}(\mathbb{X}^{n,p}), \\ \mathbb{X}^{n,p+1} &= \mathbb{X}^{n,p} - \Delta \mathbb{X}^{n,p}, \end{aligned}$$

which is solved through an LU factorization with partial pivoting for every Newton iteration, until the error tolerance is satisfied.

2.3.2 Fully discretised problem for the non-linear detachment model

Following a similar procedure as in previous section, the set of discretised equations of the second model (2.6)-(2.7), (2.11)-(2.12) and (2.13)-(2.14) particularises the system of non-linear equations (2.15), so that, on one hand the independent vector B is given by

$$B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \omega S_{j+1}^n - (2\omega - 1)S_j^n + \omega S_{j-1}^n - \sigma_1 \dfrac{S_j^n C_j^n}{K_2 + S_j^n} \\ C_j^n - \mu \left[\dfrac{(C_{j+1}^n)^b C_j^n - (C_{j+1}^n)^{b+1}}{(1C_{j+1}^n)^a} + \dfrac{(C_{j-1}^n)^b C_j^n - (C_{j-1}^n)^{b+1}}{(1 - C_{j-1}^n)} \right] + \sigma_3 \dfrac{S_j^n C_j^n}{K_2 + S_j^n} - \sigma_4 \dfrac{C_j^n}{K_d + S_j^n} \\ \vdots \\ 0 \\ 0 \end{pmatrix}. \tag{2.23}$$

On the other hand, matrix A can be expressed in the form

$$A(\mathbb{X}^{n+1}) = A^c + A^{nc}(\mathbb{X}^{n+1}),$$

with A^c independent of $\mathbb{X}^{n+1}$, and $A^{nc}(\mathbb{X}^{n+1})$ holding all the dependency on $\mathbb{X}^{n+1}$ and, therefore gathering the non-linearity of the system. Let us denote $A^{nc}(\mathbb{X}^{n+1}) \equiv A^{nc}$. Both matrices, A^c and A^{nc}, are sparse and have a certain block structure. More precisely, let us describe their respective block structures:

- Matrix A^c

$$A^c = \left(\begin{array}{c|c|c|c|c|c|c} \gamma_1 & \gamma_2 & \gamma_3 & 0 & \dots & & \\ \hline \alpha & \beta & \alpha & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha & \beta & \alpha & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha & \beta & \alpha \\ \hline & & \dots & 0 & \gamma_3 & \gamma_2 & \gamma_1 \end{array}\right), \tag{2.24}$$

where the involved 2×2 block submatrices are given by

$$\alpha = \left(\begin{array}{c|c} -\omega & 0 \\ \hline 0 & 0 \end{array}\right), \beta = \left(\begin{array}{c|c} 1+2\omega & 0 \\ \hline 0 & 1 \end{array}\right),$$

$$\gamma_1 = \left(\begin{array}{c|c} 3 & 0 \\ \hline 0 & 3 \end{array}\right), \gamma_2 = \left(\begin{array}{c|c} -4 & 0 \\ \hline 0 & -4 \end{array}\right), \gamma_3 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}\right).$$

- Matrix A^{nc}

$$A^{nc} = \left(\begin{array}{c|c|c|c|c|c|c} 0 & 0 & 0 & \dots & & & \\ \hline \alpha_1^2 & \beta_1 & \alpha_1^1 & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha_j^2 & \beta_j & \alpha_j^1 & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha_{J-2}^2 & \beta_{J-2} & \alpha_{J-2}^1 \\ \hline & & & \dots & 0 & 0 & 0 \end{array}\right), \tag{2.25}$$

where the involved 2×2 block submatrices are given by

  - Block α_j^1, for $j = 1, \dots, J-2$:

$$\alpha_j^1 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & -\mu \dfrac{(C_{j+1}^{n+1})^b}{(1 - C_{j+1}^{n+1})^a} \end{array}\right).$$

– Block α_j^2, for $j = 1, \ldots, J-2$:

$$\alpha_j^2 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & -\mu\dfrac{(C_{j-1}^{n+1})^b}{(1-C_{j-1}^{n+1})^a} \end{array}\right).$$

– Block β_j, for $j = 1, \cdots, J-2$:

$$\beta_j = \left(\begin{array}{c|c} \sigma_1\dfrac{C_j^{n+1}}{K_2+S_j^{n+1}} & 0 \\ \hline 0 & \beta_{j,4} \end{array}\right), \tag{2.26}$$

where

$$\beta_{j,4} = \mu\left[\frac{(C_{j+1}^{n+1})^b}{(1-C_{j+1}^{n+1})^a} + \frac{(C_{j-1}^{n+1})^b}{(1-C_{j-1}^{n+1})^a}\right] - \sigma_3\frac{S_j^{n+1}}{K_2+S_j^{n+1}} + \sigma_4\frac{1}{K_d+S_j^{n+1}}.$$

Linear system for each Newton iteration

With the non-linear function $\mathcal{F}$ described, we need to build the linear system (2.17). The Jacobian matrix, $J_{\mathcal{F}}$ is evaluated in every Newton iteration at $\mathbb{X}^{n,p}$. It can be decomposed in a constant part plus a non constant part

$$J_{\mathcal{F}}(\mathbb{X}^{n,p}) = A^c + \overline{A}(\mathbb{X}^{n,p}),$$

where A^c is the matrix (2.24). Matrix $\overline{A}(\mathbb{Y})$ is a square sparse matrix of order $2J$. More precisely, next we describe its block structure,

$$\overline{A}(\mathbb{Y}) = \left(\begin{array}{c|c|c|c|c|c|c} 0 & 0 & 0 & \dots & & & \\ \hline \overline{\alpha}_1^2 & \overline{\beta}_1 & \overline{\alpha}_1^1 & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \overline{\alpha}_j^2 & \overline{\beta}_j & \overline{\alpha}_j^1 & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \overline{\alpha}_{J-2}^2 & \overline{\beta}_{J-2} & \overline{\alpha}_{J-2}^1 \\ \hline & & & \dots & 0 & 0 & 0 \end{array}\right). \tag{2.27}$$

where the involved 2×2 block submatrices are given by

- Block $\overline{\alpha}_j^1$, for $j = 1, \dots, J-2$:

$$\overline{\alpha}_j^1 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \overline{\alpha}_{j,4}^1 \end{array}\right),$$

where

$$\overline{\alpha}_{j,4}^1 = \mu \frac{\left[b\mathbb{Y}_{2j+3}^{b-1}\mathbb{Y}_{2j+1} - (b+1)\mathbb{Y}_{2j+3}^b\right](1-\mathbb{Y}_{2j+3})^a + a(1-\mathbb{Y}_{2j+3})^{a-1}(\mathbb{Y}_{2j+3}^b\mathbb{Y}_{2j+1} - \mathbb{Y}_{2j+3}^{b+1})}{(1-\mathbb{Y}_{2j+3})^{2a}}.$$

- Block $\overline{\alpha}_j^2$ with $j = 1, \dots, J-2$

$$\overline{\alpha}_j^2 = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \overline{\alpha}_{j,4}^2 \end{array}\right),$$

where

$$\overline{\alpha}_{j,4}^2 = \mu \frac{\left[b\mathbb{Y}_{2j-1}^{b-1}\mathbb{Y}_{2j+1} - (b+1)\mathbb{Y}_{2j-1}^b\right](1-\mathbb{Y}_{2j-1})^a + a(1-\mathbb{Y}_{2j-1})^{a-1}(\mathbb{Y}_{2j-1}^b\mathbb{Y}_{2j+1} - \mathbb{Y}_{2j-1}^{b+1})}{(1-\mathbb{Y}_{2j-1})^{2a}}.$$

- Block $\overline{\beta}_j$, for $j = 1, \cdots, J-2$:

$$\overline{\beta}_j = \left(\begin{array}{c|c} \sigma_1 \dfrac{\mathbb{Y}_{2j+1} K_2}{(K_2 + \mathbb{Y}_{2j})^2} & \sigma_1 \dfrac{\mathbb{Y}_{2j}}{K_2 + \mathbb{Y}_{2j}} \\ \hline -\sigma_3 \dfrac{\mathbb{Y}_{2j+1} K_2}{(K_2 + \mathbb{Y}_{2j})^2} + \sigma_4 \dfrac{\mathbb{Y}_{2j+1}}{(K_2 + \mathbb{Y}_{2j})^2} & \overline{\beta}_{j,4} \end{array}\right),$$

with

$$\overline{\beta}_{j,4} = \mu \left[\frac{\mathbb{Y}^b_{2j+3}}{(1 - \mathbb{Y}_{2j+3})^a} + \frac{\mathbb{Y}^b_{2j-1}}{(1 - \mathbb{Y}_{2j-1})^a} \right] - \sigma_3 \frac{\mathbb{Y}_{2j}}{K_2 + \mathbb{Y}_{2j}} + \frac{\sigma_4}{K_2 + \mathbb{Y}_{2j}}.$$

Using (2.23), (2.24), (2.25) and (2.27) we get the system of equations (2.17)-(2.18)

$$\begin{aligned} J_{\mathcal{F}}(\mathbb{X}^{n,p}) \Delta \mathbb{X}^{n,p} &= \mathcal{F}(\mathbb{X}^{n,p}) \\ \mathbb{X}^{n,p+1} &= \mathbb{X}^{n,p} - \Delta \mathbb{X}^{n,p}, \end{aligned}$$

which is solved through an LU factorization with partial pivoting for every Newton iteration, until the error tolerance is satisfied.

2.3.3 Fully discretised problem for the nutrients uptake impairment plus ageing model

Finally, as in the previous models, the set of discretised equations of the third model (2.9)-(2.10), (2.11)-(2.12) and (2.13)-(2.14) particularises the system of non-linear

equations (2.15), so that, on one hand, the independent vector B is given by

$$B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \omega S_{j+1}^n - (2\omega - 1)S_j^n + \omega S_{j-1}^n - \sigma_1 S_j^n C_j^n \\ \mu C_{j+1}^n - (2\mu + \sigma_{CBD} - 1)C_j^n + \mu C_{j-1}^n - \sigma_3 S_j^n C_j^n \\ \vdots \\ 0 \\ 0 \end{pmatrix}. \tag{2.28}$$

On the other hand, matrix A can be expressed in the form

$$A(\mathbb{X}^{n+1}) = A^c + A^{nc}(\mathbb{X}^{n+1}),$$

with A^c independent of $\mathbb{X}^{n+1}$, and $A^{nc}(\mathbb{X}^{n+1})$ holding all the dependency on $\mathbb{X}^{n+1}$ and, therefore gathering the non-linearity of the system. Let us denote $A^{nc}(\mathbb{X}^{n+1}) \equiv A^{nc}$.

Matrix A^c is sparse and has a certain block structure. More precisely, next we describe its block structure:

$$A^c = \left(\begin{array}{c|c|c|c|c|c|c} \gamma_1 & \gamma_2 & \gamma_3 & 0 & \dots & & \\ \hline \alpha & \beta & \alpha & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha & \beta & \alpha & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha & \beta & \alpha \\ \hline & & \dots & 0 & \gamma_3 & \gamma_2 & \gamma_1 \end{array}\right), \tag{2.29}$$

where the involved 2×2 block submatrices are given by

$$\alpha = \left(\begin{array}{c|c} -\omega & 0 \\ \hline 0 & -\mu \end{array}\right), \beta = \left(\begin{array}{c|c} 1+2\omega & 0 \\ \hline 0 & 1+2\mu+\sigma_{CBD} \end{array}\right),$$
$$\gamma_1 = \left(\begin{array}{c|c} 3 & 0 \\ \hline 0 & 3 \end{array}\right), \gamma_2 = \left(\begin{array}{c|c} -4 & 0 \\ \hline 0 & -4 \end{array}\right), \gamma_3 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}\right).$$

Matrix A^{nc} is a diagonal matrix where

$$\begin{aligned} \left(A^{nc}_{0,0}, A^{nc}_{1,1}\right) &= (0,0), \\ \left(A^{nc}_{2j,2j}, A^{nc}_{2j+1,2j+1}\right) &= \left(\sigma_1 C_j^{n+1}, -\sigma_3 S_j^{n+1}\right) \text{ for } j = 1, \dots, J-2, \\ \left(A^{nc}_{2J-2,2J-2}, A^{nc}_{2J-1,2J-1}\right) &= (0,0). \end{aligned} \tag{2.30}$$

Linear system for each Newton iteration

With the non-linear function $\mathcal{F}$ described, we need to build the linear system (2.17). The Jacobian matrix, $J_{\mathcal{F}}$ is evaluated in every Newton iteration at $\mathbb{X}^{n,p}$. It can be also decomposed in a constant part plus a non constant part

$$J_{\mathcal{F}}(\mathbb{X}^{n,p}) = A^c + \overline{A}(\mathbb{X}^{n,p}),$$

where A^c is the matrix (2.29). Matrix $\overline{A}(\mathbb{Y})$ is a square sparse matrix of order $2J$. More precisely, let us describe its block structure,

$$\overline{A}(\mathbb{Y}) = \left(\begin{array}{c|c|c|c|c|c|c} 0 & 0 & 0 & \dots & & & \\ \hline 0 & \overline{\beta}_1 & 0 & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & 0 & \overline{\beta}_j & 0 & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & 0 & \overline{\beta}_{J-2} & 0 \\ \hline & & & \dots & 0 & 0 & 0 \end{array}\right), \tag{2.31}$$

where the involved 2×2 block submatrices are given by

$$\overline{\beta}_j = \left(\begin{array}{c|c} \sigma_1 \mathbb{Y}_{2j+1} & \sigma_1 \mathbb{Y}_{2j} \\ \hline -\sigma_3 \mathbb{Y}_{2j+1} & -\sigma_3 \mathbb{Y}_{2j} \end{array}\right),$$

for $j = 1, \dots, J-2$.

Using (2.28), (2.29), (2.30) and (2.31) we get the system of equations (2.17)-(2.18)

$$\begin{aligned} J_{\mathcal{F}}(\mathbb{X}^{n,p})\Delta\mathbb{X}^{n,p} &= \mathcal{F}(\mathbb{X}^{n,p}), \\ \mathbb{X}^{n,p+1} &= \mathbb{X}^{n,p} - \Delta\mathbb{X}^{n,p}, \end{aligned}$$

which is solved through an LU factorization with partial pivoting for every Newton iteration, until the error tolerance is satisfied.

2.4 Academic tests

In order to test the reliability of the numerical schemes developed to solve the one-dimensional models, we built an academic test and analysed the computed numerical results. For this purpose, let us consider the following model

$$\frac{\partial S}{\partial t} = \nabla \cdot (d_1(C)\nabla S) - F(S,C) + h_1(x,t), \tag{2.32}$$

$$\frac{\partial C}{\partial t} = \nabla \cdot (d_2(C)\nabla M) + G(S,C) + h_2(x,t), \tag{2.33}$$

with boundary conditions

$$\begin{aligned} C(0,x) &= C_0,\ t \in [0,T], && (2.34)\\ \frac{\partial C}{\partial x}(t,0) &= 0,\ t \in [0,T], && (2.35)\\ \frac{\partial C}{\partial x}(t,L) &= 0,\ t \in [0,T], && (2.36) \end{aligned}$$

$$\begin{aligned} S(0,x) &= 1,\ x \in [0,L], && (2.37)\\ \frac{\partial S}{\partial x}(t,0) &= 0,\ t \in [0,T], && (2.38)\\ \frac{\partial S}{\partial x}(t,L) &= 0,\ t \in [0,T], && (2.39) \end{aligned}$$

which is applied to models (1.11)-(1.12), (1.19)-(1.20) and (1.25)-(1.26). The additional terms h_1 and h_2 are introduced so that the system has the following exact solution

$$\begin{aligned} S(t,x) &= e^{t(x-L)}, && (2.40)\\ C(t,x) &= C_0 e^{-xt}. && (2.41) \end{aligned}$$

More precisely, in terms of the functions defined in (2.40) and (2.41), the functions h_1 and h_2 are obtained as

$$\begin{aligned} h_1 &= \frac{\partial S}{\partial t} - \nabla \cdot (d_1(C)\nabla S) + F(S,C),\\ h_2 &= \frac{\partial C}{\partial t} - \nabla \cdot (d_2(C)\nabla C) - G(S,C). \end{aligned}$$

So, taking into account the expression of $F,\ G,\ d_1$ and d_2 in the Eberl model (1.11)-(1.12), we get:

$$h_1(x,t) = (x-L)e^{t(x-L)} - d_1 t^2 e^{t(x-L)} + K_1 \frac{e^{t(x-L)} C_0 e^{-xt}}{K_2 + e^{t(x-L)}}, \tag{2.42}$$

and

$$\begin{aligned} h_2(x,t) = -C_0 x e^{-xt} - c_{max}^{b-a}\epsilon^a t^2 C_0 \frac{b e^{-xtb} + (1-b+a)e^{-xt(b+1)} - e^{-xt(b+2)}}{(1-C_0 e^{-xt})^{a+1}} +\\ K_4 C_0 e^{-xt} - K_3 \frac{e^{t(x-L)} C_0 e^{-xt}}{K_2 + e^{t(x-L)}}. \end{aligned} \tag{2.43}$$

Note that in the numerical methods, the terms h_1 and h_2 modify the original vector $B(\mathbb{X}^n)$ in (2.19), obtaining term $\hat{B}(\mathbb{X}^n)$

$$\hat{B}(\mathbb{X}^n) = B(\mathbb{X}^n) + \frac{\Delta t}{2} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ h_1(x_j, t_{n+1}) + h_1(x_j, t_n) \\ h_2(x_j, t_{n+1} + h_2(x_j, t_n)) \\ \vdots \\ 0 \\ 0 \end{pmatrix}. \tag{2.44}$$

The same strategy can be applied for the second model and the third model. Thus, if we take the expressions of terms $F,\ G,\ d_1$ and d_2 in the second model (1.19)-(1.20), then the functions h_1 and h_2 become

$$h_1(x,t) = (x-L)e^{t(x-L)} - d_1 t^2 e^{t(x-L)} + K_1 \frac{e^{t(x-L)} C_0 e^{-xt}}{K_2 + e^{t(x-L)}}, \tag{2.45}$$

and

$$h_2(x,t) = -C_0 x e^{-xt} - c_{max}^{b-a} \epsilon^a t^2 C_0 \frac{b e^{-xtb} + (1-b+a)e^{-xt(b+1)} - e^{-xt(b+2)}}{(1 - C_0 e^{-xt})^{a+1}} + K_4 \frac{C_0 e^{-xt}}{K_d + e^{t(x-L)}} - K_3 \frac{e^{t(x-L)} C_0 e^{-xt}}{K_2 + e^{t(x-L)}}. \tag{2.46}$$

Analogously, if we take the expressions of terms $F,\ G,\ d_1$ and d_2 in the third model (1.25)-(1.26), then the functions h_1 and h_2 become

$$h_1(x,t) = (x-L)e^{t(x-L)} - d_1(C) t^2 e^{t(x-L)} + K_1 e^{t(x-L)} C_0 e^{-xt}, \tag{2.47}$$

and

$$h_2(x,t) = -C_0 x e^{-xt} - d_2 t^2 C_0 x e^{-xt} + K_4 \frac{C_0 e^{-xt}}{1 + \exp(k_d[D_{min} - CBD(t)])} - K_3 e^{t(x-L)} C_0 e^{-xt}. \tag{2.48}$$

For both cases, once we have (2.45)-(2.46) and (2.47)-(2.48), the vector $B(\mathbb{X}^n)$ shall be replaced by the new term $\hat{B}(\mathbb{X}^n)$ in the numerical resolution.

Tables 2.2-2.4 show the magnitude of the relative numerical errors between the exact solution and the numerical approximation at final time simulated with the parameters shown in Table 2.1 for the three different models. The relative error is computed using the relative L^2 discrete norm, which is defined as follows

$$E = \frac{||S_e(t_{end}) - S_a(t_{end})||_{\mathrm{L}^2}}{||S_e(t_{end})||_{\mathrm{L}^2}}. \tag{2.49}$$

Function S_e represents the exact solution of the problem, which is formed by the exact nutrients, S_e, and biomass, C_e, while function S_a represents the approximated solution computed through the proposed algorithms, which is formed by the approximated nutriens, S_a, and biomass, C_e. As we have the solution only at the mesh nodes, the error formula (2.49) is computed as

$$E = \frac{\left[\sum_{i=0}^{J-1} \left\{ \left(S_{e_i}^N - S_{a_i}^N\right)^2 + \left(C_{e_i}^N - C_{a_i}^N\right)^2 \right\}\right]^{1/2}}{\left[\sum_{i=0}^{J-1} \left\{ \left(S_{e_i}^N\right)^2 + \left(C_{e_i}^N\right)^2 \right\}\right]^{1/2}}. \tag{2.50}$$

In Table 2.2 we can observe the second order convergence in time when we choose a fine enough mesh in space. Note that for 4096 elements, the error is divided by 4 in the first levels of refinement in time until a certain value.

Concerning to the order in space, we achieve in practise nearly second order convergence for small enough time steps. Actually, the error is divided by 3 in the finer meshes in time.

In the case of M2 and M3, Tables 2.3 and 2.4 show a similar behaviour with respect to the order of convergence.

Parameter	M1	M2	M3
d_1	1.9752×10^{-9}	9.2408×10^{-13}	—
d_2	—	—	4.1110×10^{-16}
d_N	—	—	1.0708×10^{-11}
μ_m	4.2556×10^{-6}	1.6826×10^{-4}	8.5001×10^{-3}
Y_{XS}	1.4094×10^{-1}	1.0134×10^{-2}	0.9851
K_s	3.2227×10^{-5}	3.1919×10^{-1}	—
K_d	—	6.1351×10^{-3}	—
k_d	—	—	311.4817
D_{min}	—	—	5.2098×10^{-2}
N_{min}	—	—	0.139
m_s	6.2528×10^{-5}	3.4251×10^{-5}	3.6649×10^{-5}
c_{max}	211.14	0.96	211.14
s_0	2.74	2.74	2.74
C_0	0.8	0.8	0.8
ϵ	7.2762×10^{-5}	6.4425×10^{-4}	—
a	2	4	—
b	2	2	—
L	1	1	1
T	1	1	1

Table 2.1: Parameter values used in the academic tests

In all cases, the here considered academic examples illustrate the convergence of the method when the time step or the spatial mesh step size tend to zero.

N \ J	512	1024	2048	4096
8	5.0614×10^{-4}	5.0079×10^{-4}	5.0028×10^{-4}	5.0022×10^{-4}
16	1.5337×10^{-4}	1.2866×10^{-4}	1.2548×10^{-4}	1.2512×10^{-4}
32	9.7334×10^{-5}	4.5100×10^{-5}	3.3360×10^{-5}	3.1595×10^{-5}
64	9.3436×10^{-5}	3.4022×10^{-5}	1.4353×10^{-5}	9.1693×10^{-6}
128	9.3336×10^{-5}	3.3324×10^{-5}	1.2308×10^{-5}	5.2411×10^{-6}
256	9.3381×10^{-5}	3.3316×10^{-5}	1.2187×10^{-5}	4.9045×10^{-6}

Table 2.2: Relative errors in the L2 norm obtained in the academic test of the Eberl model.

N \ J	512	1024	2048	4096
8	5.0578×10^{-4}	5.0046×10^{-4}	4.9996×10^{-4}	4.9991×10^{-4}
16	1.5320×10^{-4}	1.2850×10^{-4}	1.2535×10^{-4}	1.2501×10^{-4}
32	9.7121×10^{-5}	4.4816×10^{-5}	3.3148×10^{-5}	3.1471×10^{-5}
64	9.3180×10^{-5}	3.3660×10^{-5}	1.3896×10^{-5}	8.7985×10^{-6}
128	9.3081×10^{-5}	3.2965×10^{-5}	1.1791×10^{-5}	4.5613×10^{-6}
256	9.3127×10^{-5}	3.2947×10^{-5}	1.1665×10^{-5}	4.1748×10^{-6}

Table 2.3: Relative errors in the L2 norm obtained in the academic test of M2

N \ J	512	1024	2048	4096
8	4.4522×10^{-4}	4.4017×10^{-4}	4.3989×10^{-4}	4.3990×10^{-4}
16	1.4017×10^{-4}	1.1361×10^{-4}	1.1029×10^{-4}	1.0997×10^{-4}
32	9.5632×10^{-5}	4.2085×10^{-5}	2.9553×10^{-5}	2.7718×10^{-5}
64	9.2997×10^{-5}	3.3379×10^{-5}	1.3367×10^{-5}	7.9645×10^{-6}
128	9.3052×10^{-5}	3.2936×10^{-5}	1.1744×10^{-5}	4.5111×10^{-6}
256	9.3127×10^{-5}	3.2946×10^{-5}	1.1651×10^{-5}	4.2235×10^{-6}

Table 2.4: Relative errors in the L2 norm obtained in the academic test of M3.

Chapter 3

Experimental validation and calibration of 1D models

The aim of this chapter is to use mathematical models described in Chapter 1 to provide some novel insights into the mechanisms that drive the dynamics of the biofilms formed by L1A1 *L. monocytogenes* strain. For this purpose, we used a parametric identification strategy combining quantitative image analysis, cell counts, nutrient uptake tests, simulation and optimization techniques to find those parameter values that better fit the experimental data.

We base our study in the experimental set-up used in the recent work [66], in which biofilms were grown in static flow conditions, restricting the study to the hydrostatic case. According to it, L1A1 strains form almost flat and unstructured biofilms. After the initial attachment, a thin biofilm forms before 24h. The flat structure is mostly stable with a sustained increase on thickness until 96 hours after which the presence of damaged or dead cells is quite important. After 120 hours the maximum thickness greatly decreases, thus indicating a massive detachment in the last 24 h.

Biofilm dynamics were followed using Confocal Laser Scanning Microscopy (CLSM), as shown in Figure 3.1. For the purpose of visualization, appropriate markers were used to distinguish between viable and damaged or dead cells. IMARIS software was used to compute the maximum thickness (MxT) of formed biofilms from CLSM

images. MxT values were extracted for the eight replicas at each sampling time (up to 120 h) when cells were organized forming a clearly differentiable biofilm. The mean values obtained out of the eight replicas were used as the basis for parameter identification. BIOFILMDIVER [65] was used to obtain the area covered by viable and damaged or dead cells.

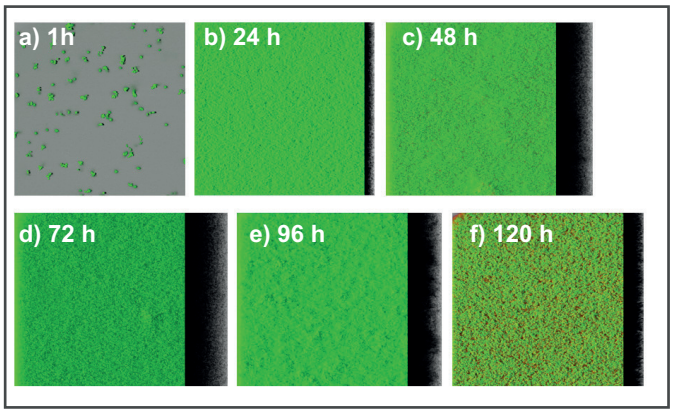

Figure 3.1: **Dynamics of L1A1 *L. monocytogenes* biofilms during life cycle.** (a–f) Present the three-dimensional reconstruction of the CLSM images captured at different times of the biofilms life cycle. Viable cells can be visualised in green whereas damaged or dead cells appear in red.

All models described in Chapter 1 were considered. All the models are discretised following the numerical methods explained in Chapter 2, and the subsequent equation system of equations has been implemented and solved in C++. The maximum height was computed by using the equation for bacteria while nutrients equation was used to compute the mean nutrient uptake. Each model incorporated different mechanisms which in turn called for various unknown parameters. Unknown parameters were estimated using data fitting techniques within the AMIGO2 toolbox [5]. The problem of calibration of model parameters is formulated as finding the unknown parameters that minimize a least squares function:

$$J_{ml}(\boldsymbol{\theta}) = \sum_{i}^{n_s} \left(mMxT_i - MxT_i(\boldsymbol{\theta})\right)^2 + \sum_{i}^{n_s} \left(mAvgN_i - AvgN_i(\boldsymbol{\theta})\right)^2, \tag{3.1}$$

subject to the dynamic constraints (the model), and bounds on the parameters $\boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U$.

In Equation (3.1), n_s represents the number of sampling times; AvgN the average nutrient concentration in the bulk liquid; mMxT and mAvgN the measured MxT and AvgN; θ corresponds to the vector of unknown model parameters. Specifities of AMIGO2 and details on the identification procedure are presented in Appendix B. The overall validation approach is shown in Figure 3.2.

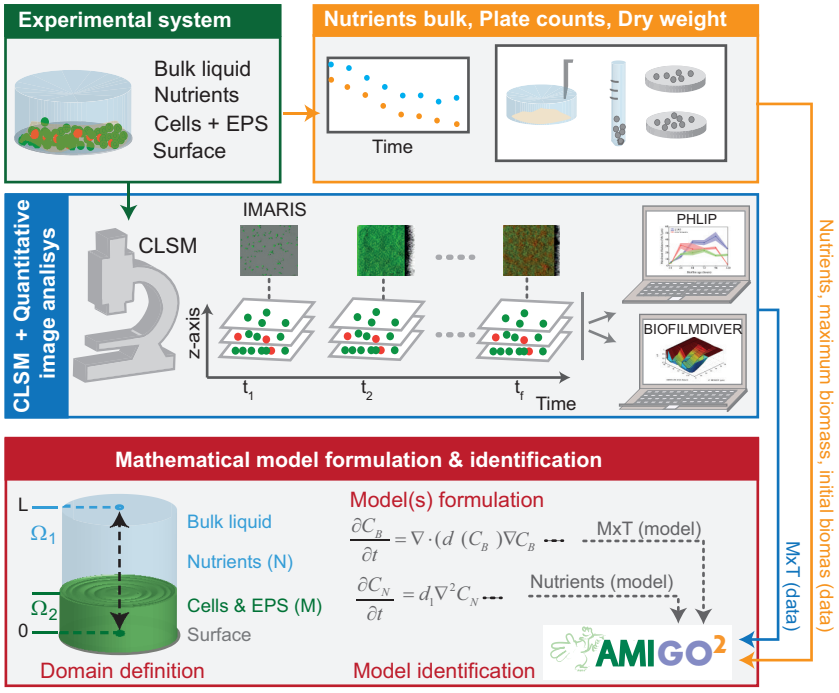

Figure 3.2: **Model identification scheme based on CLSM and nutrients consumption measurements.** Biofilms were grown under static conditions. CLSM was used to gather image stacks in several sampling times. IMARIS allowed reconstructing 3D-structures and quantifying maximum biofilm thickness throughout time. BIOFILMDIVER enabled computing biofilm covered area as a function of time and z-axis. Nutrients consumed by cells were measured at each sampling time. We defined candidate models, estimated unknown parameters and selected the most appropriate model using data fitting in the AMIGO2 toolbox.

L1A1 cells consume glucose as the primary carbon source. Therefore, the nutrients

Parameter	Model 1	Model 2	Model 3
d_1	1×10^{-12}	1×10^{-12}	—
d_2	—	—	4.1110×10^{-16}
d_N	—	—	1.0708×10^{-11}
μ_m	3.6973×10^{-3}	3.6078×10^{-1}	8.5001×10^{-3}
Y_{XS}	2.3396×10^{-1}	2.9689×10^{-1}	0.9851
K_s	9.3203×10^{-5}	7.1682×10^{2}	—
K_d	—	1.0568	—
k_d	—	—	3.1148×10^{2}
D_{min}	—	—	5.2098×10^{-2}
N_{min}	—	—	0.139
m_s	5×10^{-5}	5×10^{-5}	3.6649×10^{-5}
c_{max}	1.555×10^{1}	2.555×10^{1}	2.1114×10^{2}
s_0	2.74	2.74	2.74
C_0	3.149×10^{-3}	1.9165×10^{-3}	2.3192×10^{-4}
ϵ	4.5234×10^{-3}	4.5×10^{-3}	—
a	4	4	—
b	7.5×10^{-1}	2.4561×10^{-1}	—
N	1.8×10^{7}	1.8×10^{7}	1.8×10^{7}
J	64	64	64
L	8×10^{-5}	8×10^{-5}	8×10^{-5}
T	4.284×10^{5}	4.284×10^{5}	4.284×10^{5}

Table 3.1: Optimal parameter values found for the candidate models.

considered in the numerical resolution correspond to glucose. Experimental results show that cells consume most of the glucose in the first 24h and after that period glucose uptake stops even when there are viable cells.

Figure 3.3 presents models predictions for the optimal parameter values in Table 3.1. The figure illustrates how models M1 and M2 fail to capture the L1A1 biofilm life cycle.

M1 corresponds to the results obtained with the model proposed in Eberl et al. [26]. The optimum value of the parameters results in a maximum height quite close

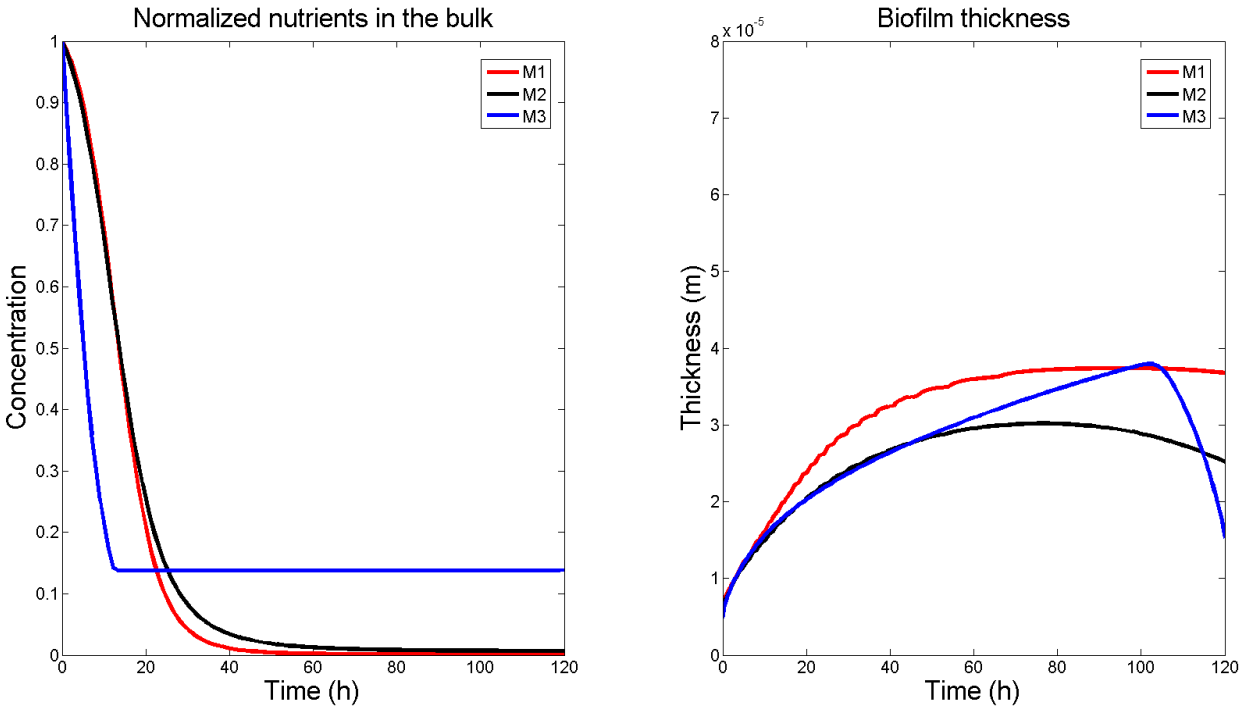

Figure 3.3: Averaged nutrients and biofilm thickness dynamics as predicted by the three candidate models

to the experimental measurements. However, the massive detachment that should happen in the final moments is not observed. Therefore, a non-linear detachment was proposed, in order to accelerate the detachment process when a certain nutrients concentration was reached. The resolution of the model for the optimal value of the parameters shows a greater detachment, much closer to the experimental data obtained for the biomass height.

In order to validate the nutrients model, dynamic experimental data of the nutrients concentration was obtained. The new data showed that the cells of L1A1 stopped the nutrients consumption, leaving a remnant of nutrients in the bulk. However, the optimal solutions for both M1 and M2 show a total consumption of the nutrients.

Other intermediate candidates were analysed in the recent work of Balsa-Canto et al. [9]. In this thesis, we will discuss further the most satisfactory model M3.

In light of the observed results for M1 and M2, another course of action was taken, including several new mechanisms. First of all, as nutrients seemed to be consumed too fast, in order to slow nutrients consumption, we considered a non-linear diffusion mechanism for the glucose: diffusion in the bulk should be greater than inside the biofilm. Also, to mimic the experimental behaviour, once a certain concentration of glucose is achieved, biomass should stop consuming it, so a nutrients impairing mechanism was added too (M3).

The original Monod terms were replaced for linear terms and the decay mechanism was changed. Analysing the experimental data, as explained in [9], it seems that the large detachment observed at 96 hours is related to cells death and the degradation of the extracellular DNA (eDNA) or, in other words, to biofilm ageing. Therefore, the hypothesis of the role of ageing was introduced in model 3, resulting in a decay term related to the measured covered area of damaged or dead cells, CBD(t), as as measure of ageing. The decay should start once a given value of damaged or dead cells is present in the biofilm. Mathematically, this behaviour is expressed in the second term of (1.28), a S-shaped function centred in D_{min} and with slope k_d. The term CBD(t) was obtained from the experimental data measured at certain times (see Table 3.2). Finally, we suppressed the non-linear biomass diffusion, in order to delay the moment of maximum thickness and reduce the nutrients consumption speed. The idea behind this change is due to the fact that with the non-linear coefficient (1.5), as we have initially a small concentration of biomass, the diffusion mechanism is small in the early phase and, therefore, biomass growth is limited to the starting region. The small diffusion causes biomass concentration to rapidly reach the maximum biomass bound, c_{max}, thus resulting in a large diffusion and, therefore, reaching the moment of maximum thickness too soon. By making the coefficient linear, the diffusion starts

Time (h)	CBD(t)
0	2.1601×10^{-2}
3	1.3957×10^{-2}
23	3.9539×10^{-2}
14	3.5725×10^{-2}
71	2.1329×10^{-2}
95	3.0902×10^{-2}
119	9.0586×10^{-2}

Table 3.2: Measured covered area of dead or damaged cells

earlier and the growth happens also outside the small initial region, thus resulting in slower nutrients consumption.

It can be seen in Figure 3.3 how the evolution changes drastically with these modifications. First of all, it can be noted that biomass thickness is slowed down reaching its peak around 100 hours. Nutrients consumption is also reduced and the nutrients impairing mechanism prevents biomass from consuming all the glucose in the domain. Also, it can be observed how the massive detachment happens in the final stage .

Figure 3.4 presents the best fit to the data which corresponds to model 3. Results reveal that the model is in clear agreement with the experimental data. Therefore concluding that the life cycle of L1A1 *L. monocytogenes* under the tested experimental conditions may be explained by taking into account impaired nutrients uptake and a massive detachement due to biofilm ageing.

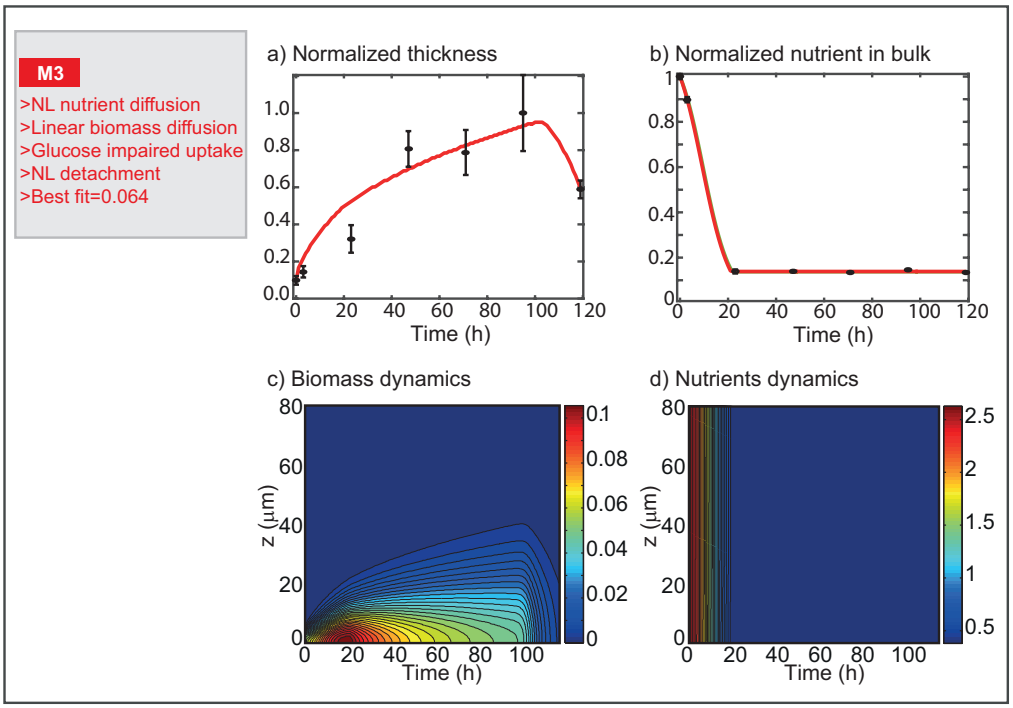

Figure 3.4: **Model 3 analysis.** (A,B) Present the best fit to the data obtained for the most successful model M3. (C,D) Show the spatio-temporal dynamics of the biomass and nutrients concentrations as predicted by model M3.

Part II

Two-dimensional multi-species models

Two-dimensional models describing biofilm evolution

Chapter 4

2D biofilm modelling

4.1 Motivation of two-dimensional models

Environmental conditions (such as nutrients availability or the hydrodynamic regime), cells life cycle, EPS production or the presence of several species, influence the structure of biofilms [47, 77].

L1A1 *L. monocytogenes* strain develops rather flat structures without much spatial heterogeneity that could be quantitative explained with 1D models. However, the presence of internal channels, inner voids or other biofilm structures (such as rugose, mushroom-shaped or clustered biofilms), can not be explained with 1D models. Thus, in order to explore those different structures we need to define 2D or 3D models. In addition, models should incorporate several components and/or species such as viable cells (from one or several microbial species), damaged or dead cells, EPS, various nutrients, etc. From now on, these models will be regarded as multi-species.

The first relevant multi-species model was presented in the work of Wanner and Gujer [107], usually referred as W-G model. The model consists of a set of 1D partial differential equations that describes the dynamics of a multi-microbial

biofilm, usually with predefined microbial distributions. With the advent of better experimental approaches to study biofilms, including Confocal Laser Microscopy, it became clear that 1D models were insufficient to represent realistic structures. This fact motivated the development of more elaborated models. Several authors suggested individual based models (see for instance, [53, 54, 73, 98, 112, 113]) or cellular automaton based models (like [46, 71]) to explain single and two-species biofilms.

In this work, a deterministic approach was selected. In particular, the model proposed by Alpkvist et al. [1] was selected as the starting point to further elaborate on the behaviour of *L. monocytogenes* biofilms. Alpkvist et al. [1] developed a continuum model of the W-G type that allowed the inclusion of multi-species as well as the study of spatial heterogeneities. When applying the model to a flat biofilm, it works as a W-G one dimensional model, although it can also be used to capture different spatial structures. The work [1] serves as an initial step to obtain a particular model that can be used to study the evolution of different strains of *L. monocytogenes* with different geometries. Our model will be able to reproduce mushroom-shaped biofilms as well as flat biofilms, biofilms with channels between different clusters or honeycomb type structures.

4.2 First model: the Alpkvist model

We begin our study of two-dimensional multi-species models by considering the model developed by Alpkvist et al. [1]. This model takes into account the active and inactive biomass of the same species and only one limiting nutrient. While referring to *active biomass* (represented by its volumetric fraction, v_1), we are considering the components of the biofilm actively reacting with the *nutrients*, whose concentration is denoted by s. By *inactive biomass* (also represented by its volumetric fraction, v_2), we refer to all other materials that do not explicitly affect the nutrients concentration such as damaged cells, dead cells or EPS.

The two dimensional domain of the problem, $\Omega = [0, L] \times [0, H]$, is represented in Figure 4.1. At each time t, the domain is decomposed into two disjoint subdomains: the biomass region, Ω_1^t, and the bulk (the region without biomass), Ω_2^t. It must be noted that the regions Ω_1^t and Ω_2^t change with the evolution of the biomasses with time. Let us denote by Γ_0^t the free boundary separating Ω_1^t and Ω_2^t. In order to track the evolution of Γ_0^t, we introduce in the model a time dependent *level set function*, ϕ, so that, at every time $t \geq 0$,

$$\Gamma_0^t = \{(x, y) \in \Omega / \phi(x, y, t) = 0\}.$$

In other words, Γ_0^t represents the level set zero of the function ϕ. Furthermore, the domains Ω_1^t and Ω_2^t can be represented in terms of the function ϕ as

$$\Omega_1^t = \{(x, y) \in \Omega / \phi(x, y, t) < 0\},$$
$$\Omega_2^t = \{(x, y) \in \Omega / \phi(x, y, t) \geq 0\}.$$

Similarly, at a fixed given distance from Γ_0^t, H_b, we define the boundary

$$\Gamma_{H_b}^t = \{(x, y) \in \Omega / \phi(x, y, t) = H_b\},$$

which is the level set H_b of the function ϕ. This is a non-physical frontier above which there is enough availability of nutrients. $\Gamma_{H_b}^t$ also represents the upper boundary of the following domain

$$\Omega_{H_b}^t = \{(x, y) \in \Omega \ / \ \phi(x, y, t) < H_b\}.$$

Finally, the left and right boundaries of the domain Ω are denoted by

$$\Gamma_L = \{(x, y) \in \Omega \ / x = 0\},$$
$$\Gamma_R = \{(x, y) \in \Omega \ / x = L\},$$

whereas, the top and bottom boundaries are denoted by

$$\Gamma_B = \{(x, y) \in \Omega \ / y = 0\},$$
$$\Gamma_H = \{(x, y) \in \Omega \ / y = H\}.$$

This model takes into account the following assumptions [1], [2]:

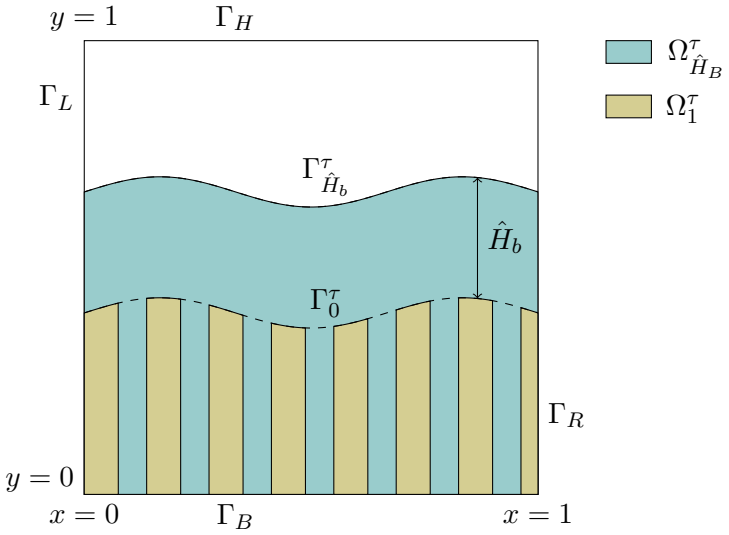

Figure 4.1: Dimensionless solving domain for the 2D case

i. The biofilm is described as a viscous fluid.

ii. Nutrients and biomasses concentrations are governed by a mass conservation equation.

iii. Active and inactive biomasses are of the same microbial species, thus implying that their density, ρ, is the same. Also, they are assumed to be incompressible, so that the density is constant.

iv. The biofilm system hosts reaction processes (production or consumption of nutrients or biomasses), transport processes (movement inside the same domain) or transference processes (movement between different domains).

v. Expansion or contraction of the biofilm is due to processes included in the model.

vi. The processes associated to the biomasses are much slower that the ones contributing to the nutrients evolution.

vii. Nutrients production or consumption is due to biotic or abiotic reaction processes.

viii. Nutrients are diluted in the media and, therefore, are present in the whole system, whereas biomass is not diluted and exists only inside the biofilm.

The description of the biofilm postulated by *i)* allows to describe in a simple way the different interactions between species. Due to *vi)*, nutrients are considered in a *quasi-steady* state.

Postulate *viii)* implies that nutrients transport is governed by a diffusion mechanism, described by Fick's Law. On the other hand, biomasses transport is regulated by and advection process. The reaction processes of the biomasses ecuations cause an increase or decrease of the biomasses concentration which, in turn, translates into an increase or decrease of the volume of the biofilm. The evolution of the biomasses concentration gives rise to a growth expansion *pressure*, p. The advective transport of biomass is regarded as a movement away from the regions with high growth expansion pressure and is described through an advective *expansive velocity field*, $\vec{u}$, dependent on the pressure as stated in Darcy's Law. As previously indicated, the evolution of Γ_0^t depends on the evolution of the biomasses concentration. More precisely, all of the level sets of ϕ are going to move following the normal velocity, f_e, which is an extension to all Ω of the normal component of $\vec{u}$, at Γ_0^t. In other words, f_e is the extension to Ω of $(\vec{u}\cdot\vec{n})_{|_{\Gamma_0^t}}$, where $\vec{n}$ is the vector normal to Γ_0^t.

Thus, the spatio-temporal variables of the model are

- $t \geq 0$: time,
- $(x, y) \in \Omega$: spatial coordinate,
- $v_i(x, y, t)$: biomasses volumetric fractions, active ($i = 1$), and inactive ($i = 2$),
- $s(x, y, t)$: nutrients concentration,

- $p(x, y, t)$: growth expansion pressure.
- $\vec{u}(x, y, t)$: expansive growth velocity field,
- $f_e(x, y, t)$: extended normal expansive velocity to Ω for the level set method,
- $\phi(x, y, t)$: level set function.

As presented in Alpkvist et al. [1], the model is given by the following set of equations

$$-D\nabla^2 s = R(v_1, v_2, s) \quad \text{in } \Omega^t_{H_b}, \tag{4.1}$$

$$-\rho\lambda\nabla^2 p = G_1(v_1, v_2, s) + G_2(v_1, v_2, s) \quad \text{in } \Omega^t_1, \tag{4.2}$$

$$\vec{u} = -\lambda\nabla p \quad \text{in } \Omega^t_1, \tag{4.3}$$

$$\partial_t\phi + f_e||\nabla\phi|| = 0 \quad \text{in } \Omega, \tag{4.4}$$

$$\rho\left(\partial_t v_i + \nabla\cdot(\vec{u}v_i)\right) = G_i(v_1, v_2, s), \text{ for } i = 1, 2, \quad \text{in } \Omega^t_1, \tag{4.5}$$

where D is the diffusion coefficient, $R(v_1, v_2, s)$ is the nutrients uptake rate and $G_i(v_1, v_2, s)$ are the biomasses growth rates ($i = 1$ for the active biomass and $i = 2$ for the inactive biomass). The second order elliptic partial differential equation 4.2 states that, when the biomass concentration increases, the pressure generated causes an expansion on the biofilm volume, whereas, when the biomass concentration decreases, the biofilm shrinks. The parameter λ is called the Darcy's parameter and depends on the viscosity of the media [25]. It can be proven [1] that p is proportional to λ^{-1}, therefore λ can be taken as one.

From equations (4.2) and (4.3) we get

$$-\lambda\nabla^2 p = -\nabla\cdot\vec{u} = \sum_{i=1}^{2}\frac{G_i(v_1, v_2, s)}{\rho}, \tag{4.6}$$

Next, by considering

$$\nabla\cdot(\vec{u}v_i) = \vec{u}\cdot\nabla v_i + v_i\nabla\cdot\vec{u},$$

and taking into account that $\lambda = 1$, equation (4.5) becomes

$$\partial_t v_i - \nabla p \cdot \nabla v_i = \frac{G_i(v_1, v_2, s)}{\rho} - v_j \sum_{j=1}^{2} \frac{G_j(v_1, v_2, s)}{\rho}, \text{ for } i = 1, 2. \tag{4.7}$$

Thus, we use this equation instead of (4.5).

4.2.1 Reaction terms and boundary conditions

For the first model, the uptake of nutrients follows a Monod expression given by:

$$R(v_1, v_2, s) = \frac{-\rho v_1}{\mathcal{Y}} \frac{\mu s}{K_s + s}, \tag{4.8}$$

where

- $\mathcal{Y}$: growth yield,
- μ: maximum growth rate,
- K_s: Monod saturation coefficient.

As previously explained, note that the inactive biomass does not contribute to the uptake of nutrients.
Biomasses growths are described by

$$G_1(v_1, v_2, s) = v_1 \rho \frac{\mu s}{K_s + s} - v_1 \rho k_d - v_1 \rho k_i, \tag{4.9}$$

$$G_2(v_1, v_2, s) = v_1 \rho k_i, \tag{4.10}$$

where the new parameters are,

- k_i: inactivation rate,
- k_d: decay rate.

Active biomass follows a growth process caused by the consumption of nutrients and is described by a Monod-like term. The parameter k_i represents the inactivation

rate, i.e., the rate at which active cells become damaged or die, whereas k_d represents the rate at which active cells are detached from the biofilm. Note that the yield coefficient does not appear in the active biomass growth. That is because we are considering that part of the consumed nutrients is used for other processes (such as EPS generation). Therefore, only a certain part of the consumed nutrients is used for the biomass growth.

In order to complete the model, boundary conditions must be imposed. First of all, at the left and right boundaries, Γ_L and Γ_R respectively, we will consider periodic boundary conditions, i.e,

$$f(0,\cdot) = f(L,\cdot),$$

for any function f. Secondly, we consider the bottom boundary of the domain, Γ_B, impermeable, so a no-flux condition is imposed, i.e,

$$\frac{\partial f}{\partial n} = 0 \text{ on } \Gamma_B.$$

In a similar manner, at the top boundary of the domain, Γ_H, a zero flux condition is imposed for the level set function. The value of H is taken big enough so that the boundary $\Gamma^t_{H_b}$ never leaves Ω. We assume that above $\Gamma^t_{H_b}$ there is enough nutrients availability. Furthermore, the concentration and flux of nutrients through $\Gamma^t_{H_b}$ must be continuous. Finally, active biomass, inactive biomass and pressure are set to be zero at Γ^t_0.

In summary, the model described by equations (4.1)-(4.4) and (4.7) is completed with the following boundary conditions

$$s = s^* \text{ on } \Gamma^t_{H_b}, \quad \partial_y s = 0 \text{ on } \Gamma_B, \tag{4.11}$$

$$p = 0 \text{ on } \Gamma^t_0, \quad \partial_y p = 0 \text{ on } \Gamma_B, \tag{4.12}$$

$$\partial_y \phi = 0 \text{ on } \Gamma_H, \quad \partial_y \phi = 0 \text{ on } \Gamma_B, \tag{4.13}$$

$$v_1 = v_2 = 0 \text{ on } \Gamma^t_0, \quad \partial_y v_1 = \partial_y v_2 = 0 \text{ on } \Gamma_B, \tag{4.14}$$

where s^* is the maximum nutrients concentration. Moreover, we consider periodic boundary conditions on Γ_L and Γ_R, and the following initial conditions:

$$\Gamma_0^0 = \{(x,y) \in \Omega / y = 0.2H + 0.05H\sin(4\pi x/L)\}, \tag{4.15}$$

$$\Omega_1^0 = \{(x,y) \in \Omega / y < 0.2H + 0.05H\sin(4\pi x/L)\}, \tag{4.16}$$

$$v_1(x,y,0) = 1, \quad v_2(x,y,0) = 0, \text{ in } \Omega_1^0. \tag{4.17}$$

4.2.2 The dimensionless Alpkvist model

For computational reasons, let us consider a dimensionless model by introducing the dimensionless spatial and temporal variables

$$X = \frac{x}{H}, \quad Y = \frac{y}{H}, \quad \tau = \frac{t}{t_D},$$

where t_D is the number of seconds per day and H is the height of the domain. The dimensionless nutrients, pressure, velocities, level set function and volumetric fractions are defined by

$$S = \frac{s}{s^*}, \quad P = \frac{\lambda t_D p}{H^2}, \quad \vec{U} = \vec{u}, \quad F_e = f_e, \quad \Phi = \phi, \quad V_i = v_i.$$

Nutrients dimensionless equation

Taking into account the previously introduced dimensionless spatial variables, we have

$$\nabla^2 s = \frac{\partial^2 s}{\partial x^2} + \frac{\partial^2 s}{\partial y^2} = \frac{s^*}{H^2}\left(\frac{\partial^2 S}{\partial X^2} + \frac{\partial^2 S}{\partial Y^2}\right) = \frac{s^*}{H^2}\nabla^2 S,$$

Therefore, equation (4.1) with the uptake term (4.8) becomes

$$-D\frac{s^*}{H^2}\nabla^2 S = -v_1\rho\frac{1}{\mathcal{Y}}\frac{\mu S}{\frac{K_s}{s^*} + S},$$

or, equivalently

$$-\nabla^2 S = -v_1\frac{\rho H^2 \mu}{\mathcal{Y} D s^*}\frac{S}{\frac{K_s}{s^*} + S}. \tag{4.18}$$

Pressure dimensionless equation

Taking into account the previously introduced dimensionless spatial variables, we have

$$\nabla^2 p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} = \frac{H^2}{\lambda t_D}\frac{1}{H^2}\left(\frac{\partial^2 P}{\partial X^2} + \frac{\partial^2 P}{\partial Y^2}\right) = \frac{1}{\lambda t_D}\nabla^2 P.$$

Therefore, equation (4.2) with the source terms (4.9) and (4.10) becomes

$$-\frac{1}{\lambda t_D}\nabla^2 P = v_1 \frac{\mu S}{\frac{K_s}{s^*} + S} - v_1 k_d,$$

or, recalling that λ is assumed to be equal to one,

$$-\nabla^2 P = v_1 \mu t_D \frac{S}{\frac{K_s}{s^*} + S} - v_1 t_D k_d. \tag{4.19}$$

Active biomass dimensionless equation

Taking into account the previously introduced dimensionless variables, we have

$$\partial_t v_1 = \frac{1}{t_D}\partial_\tau V_1, \quad \nabla v_1 = \frac{1}{H}\nabla V_1, \quad \nabla p = \frac{H}{\lambda t_D}\nabla P,$$

so that the left term of equation (4.7) particularised for $i = 1$, i.e., the active biomass, becomes

$$\partial_t v_1 - \nabla p \cdot \nabla v_1 = \frac{1}{t_D}\partial_\tau V_1 - \frac{H}{\lambda t_D}\frac{1}{H}\nabla P \cdot \nabla V_1 = \frac{1}{t_D}(\partial_\tau V_1 - \frac{1}{\lambda}\nabla P \cdot \nabla V_1).$$

On the other hand, the right term of equation (4.7) particularised for the active biomass and the growth terms (4.9) and (4.10), becomes

$$v_1 \frac{\mu s}{K_s + s} - v_1 k_d - v_1 k_i - v_1 \left(v_1 \frac{\mu s}{K_s + s} - v_1 k_d\right) =$$
$$V_1 \frac{\mu S}{\frac{K_s}{s^*} + S} - V_1 k_d - V_1 k_i - V_1 \left(V_1 \frac{\mu S}{\frac{K_s}{s^*} + S} - V_1 k_d\right).$$

Therefore, equation (4.7) particularised for $i = 1$ with the source terms (4.9) and (4.10), becomes

$$\frac{1}{t_D}(\partial_\tau V_1 - \frac{1}{\lambda}\nabla P \cdot \nabla V_1) = V_1 \frac{\mu S}{\frac{K_s}{s^*} + S} - V_1 k_d - V_1 k_i - V_1 \left(v_1 \frac{\mu S}{\frac{K_s}{s^*} + S} - V_1 k_d\right),$$

or equivalently

$$\lambda\partial_\tau V_1 - \nabla P\cdot\nabla V_1 = V_1\frac{\lambda t_D\mu S}{\frac{K_s}{s^*}+S} - V_1\lambda t_D k_d - V_1\lambda t_D k_i - V_1\left(V_1\frac{\lambda t_D\mu S}{\frac{K_s}{s^*}+S} - V_1\lambda t_D k_d\right).$$

Finally, by recalling that λ is assumed to be equal to one, we get

$$\partial_\tau V_1 - \nabla P\cdot\nabla V_1 = V_1\frac{t_D\mu S}{\frac{K_s}{s^*}+S} - V_1 t_D k_d - V_1 t_D k_i - V_1\left(V_1\frac{t_D\mu S}{\frac{K_s}{s^*}+S} - V_1 t_D k_d\right). \quad (4.20)$$

Inactive biomass dimensionless equation

Taking into account the previously introduced dimensionless variables, we have

$$\partial_t v_2 = \frac{1}{t_D}\partial_\tau V_2, \quad \nabla v_2 = \frac{1}{H}\nabla V_2, \quad \nabla p = \frac{H}{\lambda t_D}\nabla P,$$

so that the left term of equation (4.7) particularised for $i = 2$, i.e., the inactive biomass, becomes

$$\partial_t v_2 - \nabla p\cdot\nabla v_2 = \frac{1}{t_D}\partial_\tau V_2 - \frac{H}{\lambda t_D}\frac{1}{H}\nabla P\cdot\nabla V_2 = \frac{1}{t_D}(\partial_\tau V_2 - \frac{1}{\lambda}\nabla P\cdot\nabla V_2).$$

On the other hand, the right term of equation (4.7) particularised for the inactive biomass and the growth terms (4.9) and (4.10), becomes

$$v_1 k_i - v_2\left(v_1\frac{\mu s}{K_s+s} - v_1 k_d\right) = V_1 k_i - V_2\left(V_1\frac{\mu S}{\frac{K_s}{s^*}+S} - V_1 k_d\right).$$

Therefore, equation (4.7) particularised for $i = 2$ with the source terms (4.9) and (4.10), becomes

$$\frac{1}{t_D}(\partial_\tau V_1 - \frac{1}{\lambda}\nabla P\cdot\nabla V_1) = V_1 k_i - V_2\left(V_1\frac{\mu S}{\frac{K_s}{s^*}+S} - V_1 k_d\right),$$

or equivalently,

$$\lambda\partial_\tau V_1 - \nabla P\cdot\nabla V_1 = V_1\lambda t_D k_i - V_2\left(V_1\frac{\lambda t_D\mu S}{\frac{K_s}{s^*}+S} - V_1\lambda t_D k_d\right).$$

Finally, by recalling that λ is assumed to be equal to one, we get

$$\partial_\tau V_1 - \nabla P\cdot\nabla V_1 = V_1 t_D k_i - V_2\left(V_1\frac{t_D\mu S}{\frac{K_s}{s^*}+S} - V_1 t_D k_d\right). \quad (4.21)$$

The dimensionless model

Equations (4.18), (4.19), (4.20) and (4.21), with the corresponding dimensionless variables, define the dimensionless model. At every time τ, the free boundary Γ_0^τ, divides the domain $\Omega = [0, \frac{L}{H}] \times [0, 1]$ into two disjoint subdomains, the biomass region, Ω_1^τ, and the bulk, Ω_2^τ, so that $\Omega = \Omega_1^\tau \cup \Omega_2^\tau$. The boundaries of the domain are

$$\Gamma_L = \{(X, Y) \in \Omega \ / X = 0\}, \tag{4.22}$$

$$\Gamma_R = \left\{(X, Y) \in \Omega \ / X = \frac{L}{H}\right\}, \tag{4.23}$$

$$\Gamma_B = \{(X, Y) \in \Omega \ / Y = 0\}, \tag{4.24}$$

$$\Gamma_H = \{(X, Y) \in \Omega \ / Y = 1\}. \tag{4.25}$$

The free boundary Γ_0^τ corresponds to the level set zero of the level set function Φ, i.e.,

$$\Gamma_0^\tau = \{(X, Y) \in \Omega / \Phi(X, Y, \tau) = 0\}, \tag{4.26}$$

while $\Gamma_{\hat{H}_b}^\tau$ corresponds to the level set $\hat{H}_b = 0.125$, i.e.,

$$\Gamma_{\hat{H}_b}^\tau = \{(X, Y) \in \Omega / \Phi(X, Y, \tau) = \hat{H}_b\}, \tag{4.27}$$

and represents the upper boundary of the domain

$$\Omega_{\hat{H}_b}^\tau = \{(X, Y) \in \Omega / \phi(X, Y, \tau) < \hat{H}_b\}. \tag{4.28}$$

Thus, we get the dimensionless Alpkvist model, given by the set of equations:

$$-\nabla^2 S = -V_1 \frac{h_T^2 S}{K + S} \text{ in } \Omega_{\hat{H}_b}^\tau, \tag{4.29}$$

$$-\nabla^2 P = V_1 \frac{\Psi S}{K + S} - V_1 \epsilon_1 \text{ in } \Omega_1^\tau, \tag{4.30}$$

$$\vec{U} = -\nabla P \text{ in } \Omega_1^\tau, \tag{4.31}$$

$$\partial_\tau \Phi + F_e ||\nabla \Phi|| = 0 \text{ in } \Omega, \tag{4.32}$$

$$\partial_\tau V_1 - \nabla P \cdot \nabla V_1 = V_1 \left[\frac{\Psi S}{K + S} - (\epsilon_1 + \epsilon_2) - V_1 (\frac{\Psi S}{K + S} - \epsilon_1)\right] \text{ in } \Omega_1^\tau, \tag{4.33}$$

$$\partial_\tau V_2 - \nabla P \cdot \nabla V_2 = V_1 \epsilon_2 - V_2 \left(V_1 \frac{\Psi S}{K + S} - V_1 \epsilon_1\right) \text{ in } \Omega_1^\tau, \tag{4.34}$$

where the dimensionless constants are defined as

- $h_T^2 = \dfrac{H^2 \mu \rho}{\mathcal{Y} D s^*}$ (Thiele modulus),
- $K = \dfrac{K_s}{s^*}$,
- $\Psi = \mu t_D$,
- $\epsilon_1 = t_D k_d$,
- $\epsilon_2 = t_D k_I$.

The model is completed with the dimensionless boundary conditions

$$S = 1 \text{ on } \Gamma^{\tau}_{\hat{H}_b}, \quad \partial_Y S = 0 \text{ on } \Gamma_B, \tag{4.35}$$

$$P = 0 \text{ on } \Gamma^{\tau}_0, \quad \partial_Y P = 0 \text{ on } \Gamma_B, \tag{4.36}$$

$$\partial_Y \phi = 0 \text{ on } \Gamma_H, \quad \partial_Y \phi = 0 \text{ on } \Gamma_B, \tag{4.37}$$

$$V_1 = V_2 = 0 \text{ on } \Gamma^{\tau}_0, \quad \partial_Y V_1 = \partial_Y V_2 = 0 \text{ on } \Gamma_B, \tag{4.38}$$

together with periodic boundary conditions at Γ_L and Γ_R, and the following initial conditions

$$\Gamma^0_0 = \{(X,Y) \in \Omega / Y = 0.2 + 0.05 \sin(4\pi X)\}, \tag{4.39}$$

$$\Omega^0_1 = \{(X,Y) \in \Omega / Y < 0.2 + 0.05 \sin(4\pi X)\}, \tag{4.40}$$

$$V_1(X,Y,0) = 1, \quad V_2(X,Y,0) = 0, \text{ in } \Omega^0_1. \tag{4.41}$$

4.3 Second model: non-linear detachment and linear nutrients uptake

The modified Eberl model M3 considered in previous chapters led us to conclude that, at least for L1A1, nutrient uptake is impaired and biofilm dynamics is clearly affected by ageing, i.e. by the presence of inactive biomass. In this section, we

Parameter	Description
L, H	Domain length and height
H_b	Boundary layer height
ρ	Biomass Density
D	Diffusion coefficient
$\mathcal{Y}$	Growth yield
μ	Maximum growth rate
K_s	Monod saturation constant
k_d	Decay rate
k_i	Inactivation rate
t_D	Number of seconds per day
s^*	Maximum nutrients concentration

Table 4.1: Parameters involved in the 2D equations

introduce similar mechanisms in the 2D multi-species model. The aim is to obtain a model that is able to explain both flat biofilms, as the ones showed by L1A1 strain (Figure 3.1), clustered biofilms, as the ones shown by CECT5873 strain (Figure 6.6) or honeycomb profiles also discussed in the literature for various *Listeria monocytogenes* strains. Assuming that in our experimental scheme nutrients are not limited, we modified the Monod term and used a linear nutrients uptake formulation. In addition, the linear decay was substituted by the inactive cell dependent decay considered in 1D (M3). Note that the current model accounts for inactive cells, so that both equations are also coupled by the decay term.

The new reaction terms can be expressed as

$$R(v_1, v_2, s) = \frac{-\rho v_1 \mu s}{\mathcal{Y}}, \tag{4.42}$$

for the nutrients uptake, where

- $\mathcal{Y}$: growth yield,
- μ: maximum growth rate.

and

$$G_1(v_1, v_2, s) = v_1\mu s - v_1 \frac{t_D^{-1}}{1 + \exp(K_d(D_{min} - mean(v_2)))} - v_1 k_i, \quad (4.43)$$

$$G_2(v_1, v_2, s) = v_1 k_i, \quad (4.44)$$

for the biomasses growth, where the new parameters are:

- k_i: inactivation rate,
- K_d: rate of detachment activation,
- D_{min}: percentage of damaged or dead cells before detachment,
- $mean(v_2)$: mean concentration of inactive biomass.

Next, by applying (4.42), (4.43) and (4.44) to the general model (4.1)-(4.4) and (4.7), and considering the same dimensionless variables as in the Alpkvist model, we get the so called second model, which is given by the set of equations

$$-\nabla^2 S = -V_1 \hat{h}_T^2 S \text{ in } \Omega^\tau_{\hat{H}_b}, \quad (4.45)$$

$$-\nabla^2 P = V_1 \hat{\Psi} S - V_1 \mathcal{F}_{\mathcal{D}}(V_2) \text{ in } \Omega^\tau_1, \quad (4.46)$$

$$\vec{U} = -\nabla P \text{ in } \Omega^\tau_1, \quad (4.47)$$

$$\partial_\tau \Phi + F_e ||\nabla \Phi|| = 0 \text{ in } \Omega, \quad (4.48)$$

$$\partial_\tau V_1 - \nabla P \cdot \nabla V_1 = V_1 \left[\hat{\Psi} S - (\mathcal{F}_{\mathcal{D}}(V_2) + \epsilon_2) - V_1(\hat{\Psi} S - \mathcal{F}_{\mathcal{D}}(V_2)) \right] \text{ in } \Omega^\tau_1, \quad (4.49)$$

$$\partial_\tau V_2 - \nabla P \cdot \nabla V_2 = V_1 \epsilon_2 - V_2 \left(V_1 \hat{\Psi} S - V_1 \mathcal{F}_{\mathcal{D}}(V_2) \right) \text{ in } \Omega^\tau_1, \quad (4.50)$$

where

$$\mathcal{F}_{\mathcal{D}}(V_2) = \frac{1}{1 + \exp(K_d(D_{min} - mean(V_2)))}. \quad (4.51)$$

The model is completed with the same boundary conditions (4.35)-(4.38), jointly with periodic boundary conditions at Γ_L and Γ_R and the initial conditions (4.39)-(4.41), whereas the involved dimensionless constants are

$$\hat{h}_T^2 = \frac{H^2 \mu \rho}{\mathcal{Y} D}, \quad \hat{\Psi} = \mu s^* t_D, \quad \epsilon_2 = t_D k_i.$$

Chapter 5

Numerical methods for the 2D case

5.1 Overview of the global numerical strategy

In order to describe the numerical methods, let us introduce certain general notations. First of all, as we are working in a finite differences framework, let M_X and M_Y be the number of spatial intervals in directions X and Y, respectively, and let N be the number of time intervals. For the sake of simplicity, let us consider a square domain with $H = L$ and let us take $M_X = M_Y = M$. Therefore, we can define the temporal and spatial steps, respectively, as

$$\Delta\tau = \frac{T_{max}}{N}, \quad \Delta X = \Delta Y = \frac{1}{M}.$$

Moreover, let us introduce the following notations

- $X_i = i\Delta X$, $Y_j = j\Delta Y$, with $i, j = 0, \ldots, M$,
- $\tau_n = n\Delta\tau$, with $n = 0, \ldots, N$,
- $S^n_{i,j} \approx S(X_i, Y_j, \tau_n)$, with $i, j = 0, \ldots, M$ and $n = 0, \ldots, N$,
- $P^n_{i,j} \approx P(X_i, Y_j, \tau_n)$, with $i, j = 0, \ldots, M$ and $n = 0, \ldots, N$,
- $\Phi^n_{i,j} \approx \Phi(X_i, Y_j, \tau_n)$, with $i, j = 0, \ldots, M$ and $n = 0, \ldots, N$,
- $V_{1^{i,j}}^n \approx V_1(X_i, Y_j, \tau_n)$, with $i, j = 0, \ldots, M$ and $n = 0, \ldots, N$,

- $V_{2_{i,j}}^n \approx V_2(X_i, Y_j, \tau_n)$, with $i, j = 0, \ldots, M$ and $n = 0, \ldots, N$.

As we are dealing with a changing domain, it will come in handy to have $\aleph^n$, an indicator function, that identifies the nodes of the discretisation that lie inside the biofilm, $\Omega_1^{\tau_n}$ at time τ_n, and those that do not lie inside. This function is defined as

$$\aleph^n(X, Y) = \begin{cases} 1, & \text{if } \Phi(X, Y, \tau^n) < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

In a similar manner, as the solving domain of the nutrients equation , $\Omega_{\hat{H}_b}^{\tau_n}$ at time τ_n, includes part of the bulk (see Figure 4.1), the indicator function $\aleph_S^n$ identifies the nodes of the discretisation that lie inside $\Omega_{\hat{H}_b}^{\tau_n}$, and those that do not lie inside. This function is defined as

$$\aleph_S^n(X, Y) = \begin{cases} 1, & \text{if } \Phi(X, Y, \tau_n) < \hat{H}_b, \\ 0, & \text{otherwise.} \end{cases} \tag{5.2}$$

As with the rest of the involved functions, we use the notation $\aleph_{i,j}^n \approx \aleph^n(X_i, Y_j)$ and $\aleph_{S_{i,j}}^n \approx \aleph_S^n(X_i, Y_j)$, with $i, j = 0, \ldots, M$ and $n = 0, \ldots, N$.

Finally, another important feature of the problem is that we are imposing periodic boundary conditions at Γ_L and Γ_R. Therefore, we impose that

$$\mathcal{Q}_{j,M}^n = \mathcal{Q}_{0,j}^n, \quad \forall j = 0, \ldots, M, \quad \forall n = 0, \ldots, N,$$

where $\mathcal{Q}_{i,j}^n \approx \mathcal{Q}(X_i, Y_j, \tau_n)$ for any function or unknown, $\mathcal{Q}$, of the problem. In practise, this means that we are only considering the unknowns corresponding to the indices $i = 0, \ldots, M-1$.

The models presented in Chapter 4 consist of a system of coupled differential equations. It is not possible to compute the analytical solution and solving the whole system at once requires a great effort from the numerical point of view. Therefore, the idea is to solve sequentially the involved equations at each time step. Thus, the global algorithm reads as follows:

1. Initialise the variables V_1^0, V_2^0 and Φ^0.

2. Start the temporal loop (indexed by n, corresponding to time τ^n).

3. Solve the nutrients equation, i.e., equation (4.29) in the Alpkvist model and (4.45) in the second model.

4. Solve the pressure equation, i.e., equation (4.30) in the Alpkvist model and (4.46) in the second model.

5. Compute the normal extended velocity, F_e, that moves the level sets by applying equation (4.31), in the Alpkvist model, or (4.47), in the second model and the velocity extension method.

6. Advance one time step in the level set equation, i.e., equation (4.32) in the Alpkvist model and (4.48) in the second model.

7. With the new level set function, Φ^{n+1}, define the updated solving domains, $\Omega_1^{\tau_{n+1}}$ and $\Omega_{\hat{H}_b}^{\tau_{n+1}}$, the updated domain representing the bulk, $\Omega_2^{\tau_{n+1}}$ and the updated boundaries, $\Gamma_0^{\tau_{n+1}}$ and $\Gamma_{\hat{H}_b}^{\tau_{n+1}}$.

8. Advance one time step in the active biomass equation, i.e., equation (4.33) in the Alpkvist model and (4.49) in the second model.

9. Advance one time step in the inactive biomass equation, i.e., equation (4.34) in the Alpkvist model and (4.50) in the second model.

10. If $n + 1 < N$, repeat the process from step 3.

The sketch of the global algorithm is presented in Figure 5.1. The applied numerical methods vary depending on which equation is being solved, due to the particularity of each one. Those methods are described in the following sections.

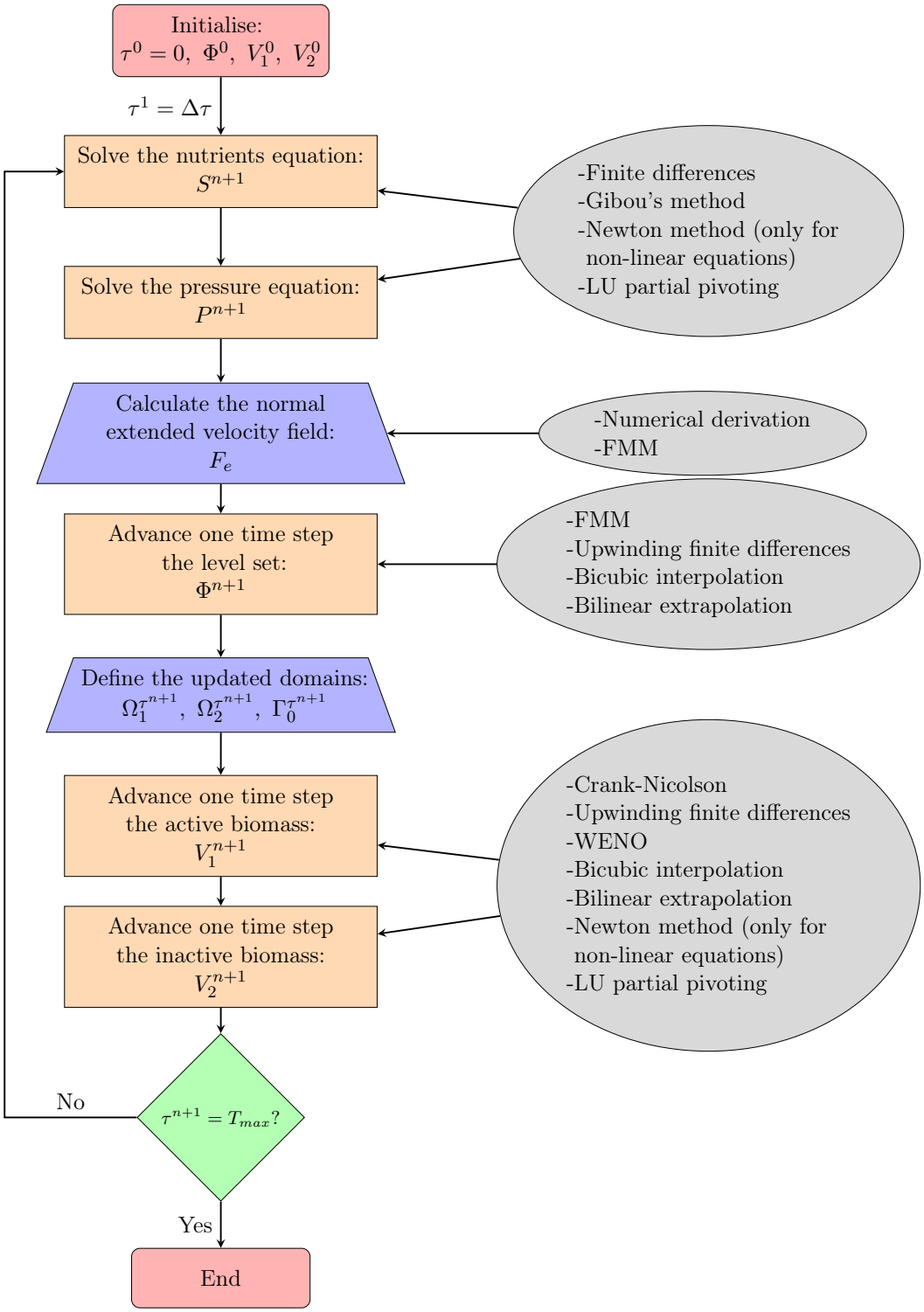

Figure 5.1: Flowchart of the global numerical algorithm in the 2D case

5.2 Nutrients equation

The semi-linear Poisson equation governing the nutrients (equation (4.29) in the Alpkvist model and (4.45) in the second model), is the first equation to be solved at each time iteration for both models. As the solving process is decoupled by solving the model sequentially, we must use the values of the active biomass from the previous iteration (or the initial condition if we are in the first iteration), in order to reach the steady-state of the nutrients. The time independence makes the resolution of this equation easier than the biomasses equations, although not entirely trivial.

Nutrients equation is solved in the domain $\Omega^{\tau}_{\hat{H}_B}$, the boundaries of which are Γ_L, Γ_R, Γ_B and $\Gamma^{\tau}_{\hat{H}_b}$ (see Figure 4.1). As stated before, the problem conditions grant enough availability of nutrients above $\Gamma^{\tau}_{\hat{H}_b}$, although in the region between $\Gamma^{\tau}_{\hat{H}_b}$ and Γ^{τ}_0 nutrients concentration is affected by the proximity to the biofilm.

In both models the spatial variables are discretised using standard finite differences, modified as in Gibou et al. [36]. However, we must proceed in a different manner depending on the model: equation (4.29) is non-linear, so that a Newton method is required, whereas equation (4.45) is linear, so that the Newton method is not required. In both cases, we end up obtaining a linear system of equations, at each Newton iteration or at each iteration step respectively, which is solved through an LU method with partial pivoting [80].

5.2.1 Gibou's Ghost Nodes method

We have already stated that the spatial discretisation of equations (4.29) and (4.45) follows a finite differences strategy, which can be problematic near the boundary $\Gamma^{\tau}_{\hat{H}_b}$, as we may need points that lie outside the computational domain. This is avoided by following Gibou's ghost nodes method [36].

In order to explain Gibou's method in a simple way, let us consider as model case the following simple one dimensional equation

$$u_{xx} = f \text{ in } \Omega = [0, L]. \tag{5.3}$$

Let be $\{x_i\}_{i=0}^{M}$ the finite differences mesh of the domain Ω, with $x_0 = 0$, $x_M = L$ and $\Delta x = x_i - x_{i-1}\ \forall i$, constant. We take $u(x_i) \approx u_i,\ f(x_i) \equiv f_i$. Using standard finite differences, the discretization of (5.3) leads to the system of equations:

$$\left[\frac{u_{i+1} - u_i}{\Delta x} - \frac{u_i - u_{i-1}}{\Delta x}\right]\frac{1}{\Delta x} = f_i, \quad \text{for } i = 1, \ldots, M-1. \tag{5.4}$$

In a regular domain, discretization (5.4) does not cause any problems. However, in our case, when approaching the boundary $\Gamma^{\tau}_{H_b}$ (or Γ^{τ}_{0} in the pressure equation), the values above the respective boundary can not be used, because they are not defined.

Let us suppose that Ω has a free boundary, Γ, that divides the domain into two disjoint subsets. As we are in the one-dimensional case, $\Gamma = \{x_I\}$. Assume that we have an index, k, such as $x_k \leq x_I \leq x_{k+1}$, and let us assume a Dirichlet boundary condition at x_I, i.e., $u(x_I) = u_I$. We can use the discretisation (5.4) in every point until reaching x_{k-1}. When solving x_k instead of using u_{k+1}, we set a value for the ghost node defined from a linear extrapolation as follows:

$$u^G_{k+1} = \frac{u_I + (\theta - 1)u_k}{\theta}, \tag{5.5}$$

where

$$\theta = \frac{|x_I - x_k|}{\Delta x} \in [0, 1].$$

Note that when θ is close to zero, equation (5.5) blows up. Therefore:

- If $\theta \leq \Delta x$, we should take $u_k = u_I$, where u_I is known.
- Otherwise, the value of the ghost node $u_{k+1} = u^G_{k+1}$ can be used in equation (5.4) for $i = k$, thus obtaining the following second order discretization

$$\left[\frac{u_I - u_k}{\theta\Delta x} - \frac{u_k - u_{k-1}}{\Delta x}\right]\frac{1}{\Delta x} = f_k.$$

Note that, as u_I is a known value, it goes directly into the independent term of the system, so that the final equation for x_k is:

$$-\left(1+\frac{1}{\theta}\right)u_k + u_{k-1} = (\Delta x)^2 f_k - \frac{u_I}{\theta}. \tag{5.6}$$

In the 2D case, the modification is applied independently to both variables in an analogous manner.

5.2.2 Numerical discretisation of the nutrients equation in the Alpkvist model

For each time step we solve the nutrients equation (4.29). More precisely, at step $n+1$ we have

$$-\nabla^2 S^{n+1} = -V_1^n \frac{h_T^2 S^{n+1}}{K + S^{n+1}}, \text{ in } \Omega_{\hat{H}_b}^{\tau_n}. \tag{5.7}$$

In order to establish a non-linear system involving the value of the nutrients at all mesh nodes, we extend the value $S \equiv 1$ on the boundary $\Gamma_{\hat{H}_b}^{\tau_n}$ to the set $\Omega \backslash \Omega_{\hat{H}_b}^{\tau_n}$. Therefore, in practise we add the following equation

$$S^{n+1} = 1, \text{ in } \Omega \backslash \Omega_{\hat{H}_b}^{\tau_n}. \tag{5.8}$$

Next, by applying a standard finite differences scheme coupled with Gibou's method, the discretisation of equations (5.7) and (5.8) leads to the system of equations

$$\aleph_{S_{i,j+1}}^n S_{i,j+1}^{n+1} + \aleph_{Si+1,j}^n S_{i+1,j}^{n+1} - \left[\frac{4+(\theta-1)\bar{\aleph}_{S_{i,j}}^n}{\theta} + \frac{(\Delta X)^2 V_{1_{i,j}}^n h_T^2}{K + S_{i,j}^{n+1}}\right] S_{i,j}^{n+1} +$$

$$\aleph_{S_{i,j-1}}^n S_{i,j-1}^{n+1} + \aleph_{S_{i-1,j}}^n S_{i-1,j}^{n+1} + \frac{4-\bar{\aleph}_{S_{i,j}}^n}{\theta} S_I^{n+1} = 0, \text{ if } \aleph_{S_{i,j}}^n = 1, \tag{5.9}$$

$$S_{i,j}^{n+1} = 1, \text{ if } \aleph_{S_{i,j}}^n = 0, \tag{5.10}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. The term $\aleph_{Si,j}^n$ is defined in (5.2), whereas $\bar{\aleph}_S^n$ denotes

$$\bar{\aleph}_{S_{i,j}}^n = \aleph_{S_{i,j+1}}^n + \aleph_{S_{i,j-1}}^n + \aleph_{S_{i+1,j}}^n + \aleph_{S_{i-1,j}}^n. \tag{5.11}$$

Moreover, note that S_I^{n+1}, the nutrients concentration on $\Gamma_{\hat{H}_b}^{\tau_n}$, is set to be one.

For the sake of simplicity, we first write the non-linear system corresponding to the case where all nodes belong to $\Omega_{\hat{H}_b}^{\tau_n}$. Then, at those nodes belonging to $\Omega \backslash \Omega_{\hat{H}_b}^{\tau_n}$, condition (5.10) is imposed, thus substituting the corresponding linearised equations from the system.

The non-linear system (5.9) is written in compact form as

$$\mathcal{F}(Z^{n+1}) = A(Z^{n+1})Z^{n+1} - B^n = 0, \tag{5.12}$$

where

- Z^{n+1} is the vector of unknowns at iteration n+1. Although equation (4.29) is time independent, the values of the nutrients must be updated every time step because the biomasses change with time. The $M(M+1)$ elements of vector Z^{n+1} are

$$Z^{n+1}(jM+i) = S_{i,j}^{n+1}, \text{ with } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.13}$$

- B^n is a vector of $M(M+1)$ elements, the values of which are

$$B^n(jM+i) = -\frac{4 - \bar{\aleph}_{S_{i,j}}^n}{\theta}, \text{ with } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.14}$$

  Note that when all nodes are in $\Omega_{\hat{H}_b}^{\tau_n}$, $B^n \equiv 0$. However, the expression (5.14) corresponds to the general case we are solving.

- $A(Z^{n+1})$ is an $M(M+1) \times M(M+1)$ sparse matrix that contains the coefficients of terms $S_{i,j}^{n+1}$ from equation (5.9).

As we are dealing with the non-linear system of equations (5.12), we propose to apply a Newton method that involves the solution of a linear system at each Newton

iteration. In order to precise the use of the Newton method, matrix $A(Z^{n+1})$ can be decomposed as the sum of two matrices, i.e.,

$$A(Z^{n+1}) = A_c^n + A_{nc}(Z^{n+1}), \tag{5.15}$$

where A_c^n contains the coefficients of the system of equations (5.12) that are independent from S and $A_{nc}(Z^{n+1})$ contains only the non-linear part of the system. Newton method requires to calculate the Jacobian matrix of the function $\mathcal{F}$ in (5.12), $J_{\mathcal{F}}(\mathbb{Y})$, which takes the form

$$J_{\mathcal{F}}(\mathbb{Y}) = A_c^n + J_{A_{nc}}(\mathbb{Y}),$$

where $J_{A_{nc}}(\mathbb{Y})$ is the Jacobian matrix related to $A_{nc}(Z^{n+1})$. Matrix A_c^n is sparse and has a certain block structure, whereas $A_{nc}(Z^{n+1})$ and $J_{A_{nc}}(\mathbb{Y})$ are diagonal matrices. More precisely, the structure of these matrices is described in the next paragraphs:

- Matrix A_c^n

$$A_c^n = \left(\begin{array}{c|c|c|c|c|c|c} \beta^1 & \alpha_0^1 & 0 & \dots & & & \\ \hline \alpha_1^2 & \beta_1^0 & \alpha_1^1 & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha_j^2 & \beta_j^0 & \alpha_j^1 & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha_{M-1}^2 & \beta_{M-1}^0 & \alpha_{M-1}^1 \\ \hline & & & \dots & 0 & \alpha_M^2 & \beta^2 \end{array}\right). \tag{5.16}$$

  There are $M+1$ blocks of size $M \times M$, that can be described as

  - α_j^1 a diagonal matrix, for $j = 0, \dots, M-1$, where

$$\alpha_j^1 = diag(\aleph_{S_{i,j+1}}^n)_{i=0}^{M-1}.$$

  - α_j^2 a diagonal for, $j = 1, \dots, M$, where

$$\alpha_j^2 = diag(\aleph_{S_{i,j-1}}^n)_{i=0}^{M-1}.$$

– β_j^0 a sparse matrix, for $j = 1, \ldots, M-1$, in the form:

$$\beta_j^0 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_{0,j}^0 & \aleph_{S_{1,j}}^n & 0 & \ldots & & & \aleph_{S_{M-1,j}}^n \\ \hline \aleph_{S_{0,j}}^n & \beta_{1,j}^0 & \aleph_{S_{2,j}}^n & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \aleph_{S_{i-1,j}}^n & \beta_{i,j}^0 & \aleph_{S_{i+1,j}}^n & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \aleph_{S_{M-3,j}}^n & \beta_{M-2,j}^0 & \aleph_{S_{M-1,j}}^n \\ \hline \aleph_{S_{0,j}}^n & & & \ldots & 0 & \aleph_{S_{M-2,j}}^n & \beta_{M-1,j}^0 \end{array}\right), \quad (5.17)$$

where

$$\beta_{0,j}^0 = -\frac{4 + (\theta - 1)(\aleph_{S_{0,j+1}}^n + \aleph_{S_{0,j-1}}^n + \aleph_{S_{1,j}}^n + \aleph_{S_{M-1,j}}^n)}{\theta},$$
$$\beta_{i,j}^0 = -\frac{4 + (\theta - 1)\bar{\aleph}_{S_{i,j}}^n}{\theta}, i = 1, \ldots, M-2,$$
$$\beta_{M-1,j}^0 = -\frac{4 + (\theta - 1)(\aleph_{S_{M-1,j+1}}^n + \aleph_{S_{M-1,j-1}}^n + \aleph_{S_{0,j}}^n + \aleph_{S_{M-2,j}}^n)}{\theta}.$$

– β^1 a sparse matrix in the form:

$$\beta^1 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_0^1 & \aleph_{S_{1,0}}^n & 0 & \ldots & & & \aleph_{S_{M-1,0}}^n \\ \hline \aleph_{S_{0,0}}^n & \beta_1^1 & \aleph_{S_{2,0}}^n & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \aleph_{S_{i-1,0}}^n & \beta_i^1 & \aleph_{S_{i+1,0}}^n & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \aleph_{S_{M-3,0}}^n & \beta_{M-2}^1 & \aleph_{S_{M-1,0}}^n \\ \hline \aleph_{S_{0,0}}^n & & & \ldots & 0 & \aleph_{S_{M-2,0}}^n & \beta_{M-1}^1 \end{array}\right), \quad (5.18)$$

where

$$\beta_0^1 = -\frac{3 + (\theta - 1)(\aleph_{S_{0,1}}^n + \aleph_{S_{1,0}}^n + \aleph_{S_{M-1,0}}^n)}{\theta},$$
$$\beta_i^1 = -\frac{3 + (\theta - 1)(\aleph_{S_{i,1}}^n + \aleph_{S_{i+1,0}}^n + \aleph_{S_{i-1,0}}^n)}{\theta}, \quad i = 1, \ldots, M-2,$$

$$\beta^1_{M-1} = -\frac{3+(\theta-1)(\aleph^n_{S_{M-1,1}} + \aleph^n_{S_{0,0}} + \aleph^n_{S_{M-2,0}})}{\theta}.$$

– β^2 a sparse matrix in the form:

$$\beta^2 = \left(\begin{array}{c|c|c|c|c|c|c} \beta^2_0 & \aleph^n_{S_{1,M}} & 0 & \dots & & & \aleph^n_{S_{M-1,M}} \\ \hline \aleph^n_{S_{0,M}} & \beta^2_1 & \aleph^n_{S_{2,M}} & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \aleph^n_{S_{i-1,M}} & \beta^2_i & \aleph^n_{S_{i+1,M}} & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \aleph^n_{S_{M-3,M}} & \beta^2_{M-2} & \aleph^n_{S_{M-1,M}} \\ \hline \aleph^n_{S_{0,M}} & & & \dots & 0 & \aleph^n_{S_{M-2,M}} & \beta^2_{M-1} \end{array}\right), \tag{5.19}$$

where

$$\begin{aligned} \beta^2_0 &= -\frac{3+(\theta-1)(\aleph^n_{S_{0,M-1}} + \aleph^n_{S_{1,M}} + \aleph^n_{S_{M-1,M}})}{\theta}, \\ \beta^2_i &= -\frac{3+(\theta-1)(\aleph^n_{S_{i,M-1}} + \aleph^n_{S_{i+1,M}} + \aleph^n_{S_{i-1,M}})}{\theta}, \quad i = 1, \dots, M-2, \\ \beta^2_{M-1} &= -\frac{3+(\theta-1)(\aleph^n_{S_{M-1,M-1}} + \aleph^n_{S_{0,M}} + \aleph^n_{S_{M-2,M}})}{\theta}. \end{aligned}$$

- Matrix $A_{nc}(Z^{n+1})$ is a diagonal matrix, where

$$A_{nc}(Z^{n+1})_{jM+i,jM+i} = -\frac{(\Delta X^2)V^n_{1_{i,j}} h^2_T}{K + S^{n+1}_{i,j}},$$

with $i = 0, \dots, M-1$ and $j = 0, \dots, M$.

- Matrix $J_{A_{nc}}(\mathbb{Y})$ is a diagonal matrix, where

$$J_{A_{nc}}(\mathbb{Y})_{jM+i,jM+i} = -\frac{(\Delta X^2)V^n_{1_{i,j}} h^2_T K}{(K + \mathbb{Y}_{i,j})^2}, \tag{5.20}$$

with $i = 0, \dots, M-1$ and $j = 0, \dots, M$.

Finally, in order to set the resulting system to be solved with LU factorisation, we impose (5.10) to all nodes in $\Omega \backslash \Omega^{\tau_n}_{\dot{H}_b}$.

5.2.3 Numerical discretisation of the nutrients equation in the second model

As in the Alpkvist model, for each time step we solve the nutrients equation (4.45). However, in the second model, we changed the Monod-like nutrients uptake term by a linear term, thus eliminating the non-linearity present in (4.29). Therefore, we no longer require a Newton method to solve the resulting system of equations. At step $n+1$ we need to solve

$$-\nabla^2 S^{n+1} = -V_1^n \hat{h}_T^2 S^{n+1}, \text{ in } \Omega_{\hat{H}_B}^{\tau_n}. \tag{5.21}$$

In order to establish a linear system involving the value of the nutrients at all mesh nodes, we extend the value $S \equiv 1$ on the boundary $\Gamma_{\hat{H}_b}^{\tau_n}$ to the set $\Omega \backslash \Omega_{\hat{H}_b}^{\tau_n}$. Therefore, in practise we add the following equation

$$S^{n+1} = 1, \text{ in } \Omega \backslash \Omega_{\hat{H}_B}^{\tau_n}. \tag{5.22}$$

Next, by applying a standard finite differences scheme coupled with Gibou's method, the discretisation of equations (5.21) and (5.22) leads to the system of equations

$$\aleph_{S_{i,j+1}}^n S_{i,j+1}^{n+1} + \aleph_{Si+1,j}^n S_{i+1,j}^{n+1} - \left[\frac{4 + (\theta - 1)\bar{\aleph}_{S_{i,j}}^n}{\theta} + (\Delta X)^2 V_{1_{i,j}}^n \hat{h}_T^2 \right] S_{i,j}^{n+1} +$$

$$\aleph_{S_{i,j-1}}^n S_{i,j-1}^{n+1} + \aleph_{S_{i-1,j}}^n S_{i-1,j}^{n+1} + \frac{4 - \bar{\aleph}_{S_{i,j}}^n}{\theta} S_I^{n+1} = 0, \text{ if } \aleph_{S_{i,j}}^n = 1, \tag{5.23}$$

$$S_{i,j}^{n+1} = 1, \text{ if } \aleph_{S_{i,j}}^n = 0, \tag{5.24}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. Moreover, note that S_I^{n+1}, the nutrients at the interface $\Gamma_{\hat{H}_B}^{\tau_n}$, is set to be one.

For the sake of simplicity, we first write the linear system corresponding to the case where all nodes belong to $\Omega_{\hat{H}_b}^{\tau_n}$. Then, at those nodes belonging to $\Omega \backslash \Omega_{\hat{H}_b}^{\tau_n}$, condition (5.24) is imposed, thus substituting the corresponding linear equations from the system.

The set of equations (5.23) can be written in compact form as the linear system of equations

$$A^n Z^{n+1} = B^n, \tag{5.25}$$

where

- Z^{n+1} is the vector of unknowns at iteration n+1. As in the Alpkvist model, although equation (4.45) is time independent, the values of the nutrients must be updated every time step because the biomasses change with time. The $M(M+1)$ elements of vector Z^{n+1} are

$$Z^{n+1}(jM+i) = S_{i,j}^{n+1}, \text{ with } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.26}$$

- B^n is the independent term of the linear system, a vector of $M(M+1)$ elements, each of them being

$$B^n(jM+i) = -\frac{4 - \bar{\aleph}_{S_{i,j}}^n}{\theta}, \text{ with } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.27}$$

  Note that when all nodes are in $\Omega_{\hat{H}_b}^{\tau_n}$, $B^n \equiv 0$. However, the expression (5.27) corresponds to the general case we are solving.

- A^n is an $M(M+1) \times M(M+1)$ sparse matrix that contains the coefficients of terms $S_{i,j}^{n+1}$ from equation (5.23). The blocks of this matrix are described in the following paragraphs:

$$A^n = \left(\begin{array}{c|c|c|c|c|c|c} \beta^1 & \alpha_0^1 & 0 & \ldots & & & \\ \hline \alpha_1^2 & \beta_1^0 & \alpha_1^1 & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \alpha_j^2 & \beta_j^0 & \alpha_j^1 & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \alpha_{M-1}^2 & \beta_{M-1}^0 & \alpha_{M-1}^1 \\ \hline & & & \ldots & 0 & \alpha_M^2 & \beta^2 \end{array}\right). \tag{5.28}$$

There are $M+1$ blocks of size $M \times M$, that can be described as

- α_j^1 a diagonal matrix, for $j = 0, \ldots, M-1$, where

$$\alpha_j^1 = diag(\aleph_{S_{i,j+1}}^n)_{i=0}^{M-1}.$$

- α_j^2 a diagonal matrix, for $j = 1, \ldots, M$, where

$$\alpha_j^2 = diag(\aleph_{S_{i,j-1}}^n)_{i=0}^{M-1}.$$

- β_j^0 a sparse matrix, for $j = 1, \ldots, M-1$, in the form:

$$\beta_j^0 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_{0,j}^0 & \aleph_{S_{1,j}}^n & 0 & \ldots & & & \aleph_{S_{M-1,j}}^n \\ \hline \aleph_{S_{0,j}}^n & \beta_{1,j}^0 & \aleph_{S_{2,j}}^n & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \aleph_{S_{i-1,j}}^n & \beta_{i,j}^0 & \aleph_{S_{i+1,j}}^n & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \aleph_{S_{M-3,j}}^n & \beta_{M-2,j}^0 & \aleph_{S_{M-1,j}}^n \\ \hline \aleph_{S_{0,j}}^n & & & \ldots & 0 & \aleph_{S_{M-2,j}}^n & \beta_{M-1,j}^0 \end{array}\right), \tag{5.29}$$

where

$$\beta_{0,j}^0 = -\frac{4 + (\theta-1)(\aleph_{S_{0,j+1}}^n + \aleph_{S_{0,j-1}}^n + \aleph_{S_{1,j}}^n + \aleph_{S_{M-1,j}}^n)}{\theta} + (\Delta X)^2 V_{0,j}^n \hat{h}_T^2,$$

$$\beta_{i,j}^0 = -\frac{4 + (\theta-1)\bar{\aleph}_{S_{i,j}}^n}{\theta} + (\Delta X)^2 V_{i,j}^n \hat{h}_T^2, \quad i = 1, \ldots, M-2,$$

$$\beta_{M-1,j}^0 = -\frac{4 + (\theta-1)(\aleph_{S_{M-1,j+1}}^n + \aleph_{S_{M-1,j-1}}^n + \aleph_{S_{0,j}}^n + \aleph_{S_{M-2,j}}^n)}{\theta} + (\Delta X)^2 V_{M-1,j}^n \hat{h}_T^2.$$

- β^1 a sparse matrix, in the form:

$$\beta^1 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_0^1 & \aleph_{S_{1,0}}^n & 0 & \ldots & & & \aleph_{S_{M-1,0}}^n \\ \hline \aleph_{S_{0,0}}^n & \beta_1^1 & \aleph_{S_{2,0}}^n & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \aleph_{S_{i-1,0}}^n & \beta_i^1 & \aleph_{S_{i+1,0}}^n & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \aleph_{S_{M-3,0}}^n & \beta_{M-2}^1 & \aleph_{S_{M-1,0}}^n \\ \hline \aleph_{S_{0,0}}^n & & & \ldots & 0 & \aleph_{S_{M-2,0}}^n & \beta_{M-1}^1 \end{array}\right), \tag{5.30}$$

where

$$\begin{aligned}
\beta_0^1 &= -\frac{3+(\theta-1)(\aleph_{S_{0,1}}^n+\aleph_{S_{1,0}}^n+\aleph_{S_{M-1,0}}^n)}{\theta}+(\Delta X)^2V_{0,0}^n\hat{h}_T^2,\\
\beta_i^1 &= -\frac{3+(\theta-1)(\aleph_{S_{i,1}}^n+\aleph_{S_{i+1,0}}^n+\aleph_{S_{i-1,0}}^n)}{\theta}+(\Delta X)^2V_{i,0}^n\hat{h}_T^2,\quad i=1,\dots,M-2,\\
\beta_{M-1}^1 &= -\frac{3+(\theta-1)(\aleph_{S_{M-1,1}}^n+\aleph_{S_{0,0}}^n+\aleph_{S_{M-2,0}}^n)}{\theta}+(\Delta X)^2V_{M-1,0}^n\hat{h}_T^2.
\end{aligned}$$

- β^2 a sparse matrix, in the form:

$$\beta^2=\left(\begin{array}{c|c|c|c|c|c|c}
\beta_0^2 & \aleph_{S_{1,M}}^n & 0 & \dots & & & \aleph_{S_{M-1,M}}^n\\ \hline
\aleph_{S_{0,M}}^n & \beta_1^2 & \aleph_{S_{2,M}}^n & 0 & \dots & & \\ \hline
 & \ddots & \ddots & \ddots & \ddots & & \\ \hline
 & \dots & \aleph_{S_{i-1,M}}^n & \beta_i^2 & \aleph_{S_{i+1,M}}^n & \dots & \\ \hline
 & & & \ddots & \ddots & \ddots & \\ \hline
 & & \dots & 0 & \aleph_{S_{M-3,M}}^n & \beta_{M-2}^2 & \aleph_{S_{M-1,M}}^n\\ \hline
\aleph_{S_{0,M}}^n & & & \dots & 0 & \aleph_{S_{M-2,M}}^n & \beta_{M-1}^2
\end{array}\right),\qquad (5.31)$$

where

$$\begin{aligned}
\beta_0^2 &= -\frac{3+(\theta-1)(\aleph_{S_{0,M-1}}^n+\aleph_{S_{1,M}}^n+\aleph_{S_{M-1,M}}^n)}{\theta}+(\Delta X)^2V_{0,M}^n\hat{h}_T^2,\\
\beta_i^2 &= -\frac{3+(\theta-1)(\aleph_{S_{i,M-1}}^n+\aleph_{S_{i+1,M}}^n+\aleph_{S_{i-1,M}}^n)}{\theta}+(\Delta X)^2V_{i,M}^n\hat{h}_T^2,\ i=1,\dots,M-2,\\
\beta_{M-1}^2 &= -\frac{3+(\theta-1)(\aleph_{S_{M-1,M-1}}^n+\aleph_{S_{0,M}}^n+\aleph_{S_{M-2,M}}^n)}{\theta}+(\Delta X)^2V_{M-1,M}^n\hat{h}_T^2.
\end{aligned}$$

Finally, in order to set the resulting system to be solved with LU factorisation, we impose (5.24) to all nodes in $\Omega\backslash\Omega_{\hat{H}_b}^{\tau_n}$.

5.3 Pressure equation

Once the nutrients concentration has been updated, the linear Poisson equation governing the pressure (equation (4.30) in the Alpkvist model and (4.46) in the second model), is solved. We need the nutrients concentration at the current time step and the active biomass concentration at the previous time step in order to solve the pressure equation. The computational domain for this equation is $\Omega_1^{\tau_n}$, the boundaries of which are Γ_L, Γ_R, Γ_B and $\Gamma_0^{\tau_n}$, as seen in Figure 4.1. The pressure appears in a quasi-steady state, so the equation is time independent. Unlike equation (4.29), both equations (4.30) and (4.46) are linear, therefore no Newton method is required, thus making the numerical solution of this equation much easier and faster. We discretised the equation by using the same finite differences method as in (4.29), including Gibou's Strategy.

5.3.1 Numerical discretisation of the pressure equation in the Alpkvist model

At step $n+1$ we compute the pressure at time $\tau_{n+1} = \Delta\tau(n+1)$. That is, we solve

$$-\nabla^2 P^{n+1} = V_1^n \frac{\Psi S^{n+1}}{K + S^{n+1}} - V_1^n \epsilon_1, \text{ in } \Omega_1^{\tau_n}, \tag{5.32}$$

In order to establish a linear system involving the value of the pressure at all mesh nodes, we extend the value $P \equiv 0$ on the boundary $\Gamma_0^{\tau_n}$ to the set $\Omega_2^{\tau_n}$. Therefore, in practise we add the following equation

$$P^{n+1} = 0, \text{ in } \Omega_2^{\tau_n}. \tag{5.33}$$

Next, by applying a standard finite differences scheme coupled with Gibou's method,

we obtain the following set of equations:

$$\aleph^n_{i,j+1}P^{n+1}_{i,j+1} + \aleph^n_{i+1,j}P^{n+1}_{i+1,j} - \left[\frac{4+(\theta-1)\bar{\aleph}^n_{i,j}}{\theta}\right]P^{n+1}_{i,j} +$$

$$\aleph^n_{i,j-1}P^{n+1}_{i,j-1} + \aleph^n_{i-1,j}P^{n+1}_{i-1,j} - (\Delta X)^2 V^n_{1_{i,j}}(\epsilon_1 - \frac{\Psi S^{n+1}_{i,j}}{K+S^{n+1}_{i,j}}) = 0 \text{ if } \aleph^n_{i,j} = 1, \tag{5.34}$$

$$P^{n+1}_{i,j} = 0 \text{ if } \aleph^n_{i,j} = 0. \tag{5.35}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. The term $\aleph^n_{i,j}$ is defined in (5.1), whereas $\bar{\aleph}^n_{i,j}$ denotes

$$\bar{\aleph}^n_{i,j} = \aleph^n_{i,j+1} + \aleph^n_{i,j-1} + \aleph^n_{i+1,j} + \aleph^n_{i-1,j}. \tag{5.36}$$

Following the steps of Gibou's method, there should be a term P^{n+1}_I, the value of the pressure at the interface $\Gamma^{\tau_n}_0$, just like in (5.6). As the boundary conditions for the pressure equation states that $P^{n+1}_I = 0$, this term vanishes.

For the sake of simplicity, we first write the linear system corresponding to the case where all nodes belong to $\Omega^{\tau_n}_1$. Then, at those nodes belonging to $\Omega^{\tau_n}_2$, condition (5.35) is imposed, thus removing the corresponding linear equations from the system.

The discretised pressure equation leads to the linear system

$$A^n Z^{n+1} = B^n, \tag{5.37}$$

where

- Z^{n+1} is the vector of unknowns at iteration n+1. Although equation (4.30) is time independent, the values of the pressure must be updated every time step because the biomasses change with time. The $M(M+1)$ elements of vector Z^{n+1} are

$$Z^{n+1}(jM+i) = P^{n+1}_{i,j}, \text{ with } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.38}$$

- B^n is the independent term of the linear system, a vector of $M(M+1)$ elements, the values of which are

$$B^n(jM+i) = (\Delta X)^2 V^n_{1_{i,j}} \left(\epsilon_1 - \frac{\Psi S^{n+1}_{i,j}}{K + S^{n+1}_{i,j}} \right), \tag{5.39}$$

  for $i = 0, \ldots, M-1$ and $j = 0, \ldots, M$.

- A^n is an $M(M+1) \times M(M+1)$ sparse matrix with a certain block structure. It contains the coefficients of terms $P_{i,j}$ from equation (5.34). Described block by block, matrix A^n takes the form

$$A^n = \left(\begin{array}{c|c|c|c|c|c|c} \beta^1 & \alpha^1_0 & 0 & \ldots & & & \\ \hline \alpha^2_1 & \beta^0_1 & \alpha^1_1 & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \alpha^2_j & \beta^0_j & \alpha^1_j & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \alpha^2_{M-1} & \beta^0_{M-1} & \alpha^1_{M-1} \\ \hline & & & \ldots & 0 & \alpha^2_M & \beta^2 \end{array} \right). \tag{5.40}$$

  There are $M+1$ blocks of size $M \times M$, that can be described as:

  - α^1_j a diagonal matrix, for $j = 0, \ldots, M-1$, where

$$\alpha^1_j = diag(\aleph^n_{i,j+1})^{M-1}_{i=0}.$$

  - α^2_j a diagonal matrix, for $j = 1, \ldots, M$ where

$$\alpha^2_j = diag(\aleph^n_{i,j-1})^{M-1}_{i=0}.$$

– β_j^0 a sparse matrix, for $j = 1, \ldots, M-1$, taking the form:

$$\beta_j^0 = \begin{pmatrix} \beta_{0,j}^0 & \aleph_{1,j}^n & 0 & \ldots & & & \aleph_{M-1,j}^n \\ \aleph_{0,j}^n & \beta_{1,j}^0 & \aleph_{2,j}^n & 0 & \ldots & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & \ldots & \aleph_{i-1,j}^n & \beta_{i,j}^0 & \aleph_{i+1,j}^n & \ldots & \\ & & & \ddots & \ddots & \ddots & \\ & & \ldots & 0 & \aleph_{M-3,j}^n & \beta_{M-2,j}^0 & \aleph_{M-1,j}^n \\ \aleph_{0,j}^n & & & \ldots & 0 & \aleph_{M-2,j}^n & \beta_{M-1,j}^0 \end{pmatrix}, \tag{5.41}$$

where

$$\begin{aligned} \beta_{0,j}^0 &= -\frac{4 + (\theta - 1)(\aleph_{0,j+1}^n + \aleph_{0,j-1}^n + \aleph_{1,j}^n + \aleph_{M-1,j}^n)}{\theta}, \\ \beta_{i,j}^0 &= -\frac{4 + (\theta - 1)\bar{\aleph}_{i,j}^n}{\theta}, \quad i = 1, \ldots, M-2, \\ \beta_{M-1,j}^0 &= -\frac{4 + (\theta - 1)(\aleph_{M-1,j+1}^n + \aleph_{M-1,j-1}^n + \aleph_{0,j}^n + \aleph_{M-2,j}^n)}{\theta}. \end{aligned}$$

– β^1 a sparse matrix that takes the form:

$$\beta^1 = \begin{pmatrix} \beta_0^1 & \aleph_{1,0}^n & 0 & \ldots & & & \aleph_{M-1,0}^n \\ \aleph_{0,0}^n & \beta_1^1 & \aleph_{2,0}^n & 0 & \ldots & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & \ldots & \aleph_{i-1,0}^n & \beta_i^1 & \aleph_{i+1,0}^n & \ldots & \\ & & & \ddots & \ddots & \ddots & \\ & & \ldots & 0 & \aleph_{M-3,0}^n & \beta_{M-2}^1 & \aleph_{M-1,0}^n \\ \aleph_{0,0}^n & & & \ldots & 0 & \aleph_{M-2,0}^n & \beta_{M-1}^1 \end{pmatrix}, \tag{5.42}$$

where

$$\begin{aligned} \beta_0^1 &= -\frac{3 + (\theta - 1)(\aleph_{0,1}^n + \aleph_{1,0}^n + \aleph_{M-1,0}^n)}{\theta}, \\ \beta_i^1 &= -\frac{3 + (\theta - 1)(\aleph_{i,1}^n + \aleph_{i+1,0}^n + \aleph_{i-1,0}^n)}{\theta}, \quad i = 1, \ldots, M-2, \end{aligned}$$

$$\beta^1_{M-1} = -\frac{3+(\theta-1)(\aleph^n_{M-1,1}+\aleph^n_{0,0}+\aleph^n_{M-2,0})}{\theta}.$$

– β^2 a sparse matrix that takes the form:

$$\beta^2 = \left(\begin{array}{c|c|c|c|c|c|c} \beta^2_0 & \aleph^n_{1,M} & 0 & \dots & & & \aleph^n_{M-1,M} \\ \hline \aleph^n_{0,M} & \beta^2_1 & \aleph^n_{2,M} & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \aleph^n_{i-1,M} & \beta^2_i & \aleph^n_{i+1,M} & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \aleph^n_{M-3,M} & \beta^2_{M-2} & \aleph^n_{M-1,M} \\ \hline \aleph^n_{0,M} & & & \dots & 0 & \aleph^n_{M-2,M} & \beta^2_{M-1} \end{array}\right), \tag{5.43}$$

where

$$\begin{aligned} \beta^2_0 &= -\frac{3+(\theta-1)(\aleph^n_{0,M-1}+\aleph^n_{1,M}+\aleph^n_{M-1,M})}{\theta}, \\ \beta^2_i &= -\frac{3+(\theta-1)(\aleph^n_{i,M-1}+\aleph^n_{i+1,M}+\aleph^n_{i-1,M})}{\theta}, \quad i=1,\dots,M-2, \\ \beta^2_{M-1} &= -\frac{3+(\theta-1)(\aleph^n_{M-1,M-1}+\aleph^n_{0,M}+\aleph^n_{M-2,M})}{\theta}. \end{aligned}$$

Finally, in order to set the resulting system to be solved with LU factorisation, we impose (5.35) to all nodes in $\Omega_2^{\tau_n}$.

5.3.2 Numerical discretisation of the pressure equation in the second model

As in the Alpkvist model, for each time step we solve the pressure equation (4.46). However, in the second model, the Monod-like terms that govern the nutrients uptake mechanism (4.8) and the active biomass growth (4.9) are substituted by linear terms (4.42) and (4.43). The decay mechanism present in (4.9) is also changed, from a

constant decay to a decay dependent on the mean inactive biomass concentration (4.43). All those changes modify the pressure equation of the second model. However, this fact has no influence in the applied numerical methods. Therefore, proceeding as in Section 5.3.1, at step $n+1$ we compute the pressure at time $\tau_{n+1} = \Delta\tau(n+1)$ by solving equation

$$-\nabla^2 P^{n+1} = V_1^n \hat{\Psi} S^{n+1} - V_1^n \mathcal{F}_{\mathcal{D}}(V_2^n), \text{ in } \Omega_1^{\tau_n}, \tag{5.44}$$

where

$$\mathcal{F}_{\mathcal{D}}(V_2^n) = \frac{1}{1+\exp(K_d(D_{min} - mean(V_2^n)))}. \tag{5.45}$$

In order to establish a linear system involving the value of the pressure at all mesh nodes, we extend the value $P \equiv 0$ on the boundary $\Gamma_0^{\tau_n}$ to the set $\Omega_2^{\tau_n}$. Therefore, in practise we add the following equation

$$P^{n+1} = 0, \text{ in } \Omega_2^{\tau_n}. \tag{5.46}$$

Next, by applying a standard finite differences scheme coupled with Gibou's method, we obtain the following set of equations:

$$\aleph_{i,j+1}^n P_{i,j+1}^{n+1} + \aleph_{i+1,j}^n P_{i+1,j}^{n+1} - \left[\frac{4+(\theta-1)\bar{\aleph}_{i,j}^n}{\theta}\right] P_{i,j}^{n+1} +$$

$$\aleph_{i,j-1}^n P_{i,j-1}^{n+1} + \aleph_{i-1,j}^n P_{i-1,j}^{n+1} - (\Delta X)^2 V_{1_{i,j}}^n \left(\mathcal{F}_{\mathcal{D}}(V_2^n) - \hat{\Psi} S_{i,j}^{n+1}\right) = 0, \text{ if } \aleph_{i,j}^n = 1, \tag{5.47}$$

$$P_{i,j}^{n+1} = 0, \text{ if } \aleph_{i,j}^n = 0. \tag{5.48}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. Following the steps of Gibou's method, there should be a term P_I^{n+1}, the value of the pressure at the interface $\Gamma_0^{\tau_n}$, just like in (5.6). As the boundary conditions for the pressure equation states that $P_I^{n+1} = 0$, this term vanishes.

For the sake of simplicity, we first write the linear system corresponding to the case where all nodes belong to $\Omega_1^{\tau_n}$. Then, at those nodes belonging to $\Omega_2^{\tau_n}$, condition (5.48) is imposed, thus removing the corresponding linear equations from the system.

The discretised pressure equation leads to the linear system

$$A^n Z^{n+1} = B^n, \tag{5.49}$$

where

- Z^{n+1} is the vector of unknowns at iteration n+1. Although (4.30) is time independent, the values of the pressure must be updated every time step because the biomasses change with time. The $M(M+1)$ elements of vector Z^{n+1} are

$$Z^{n+1}(jM+i) = P_{i,j}^{n+1}, \text{ for } i = 0, \dots, M-1 \text{ and } j = 0, \dots, M. \tag{5.50}$$

- B^n is the independent term of the linear system, a vector of $M(M+1)$ elements, the values of which are

$$B^n(jM+i) = (\Delta X)^2 V_{1_{i,j}}^n \left(\mathcal{F}_{\mathcal{D}}(V_2^n) - \hat{\Psi} S_{i,j}^{n+1} \right), \tag{5.51}$$

  for $i = 0, \dots, M-1$ and $j = 0, \dots, M$.

- A^n is an $M(M+1) \times M(M+1)$ sparse matrix with a certain block structure. It contains the coefficients of terms $P_{i,j}^{n+1}$ from equation (5.47). As the changes on the pressure equation only affect to the independent term, matrix A^n is the same as in the Alpkvist model, (5.40).

It can be observed that the modifications on the second model only affect to the independent term. Therefore, from this point on, the solving strategy is the same as in the Alpkvist model. For any further details, see Section 5.3.1.

5.4 Level set equation

In a biofilm system, certain processes happen in different domains. For instance, in the models presented in Chapter 4, the evolution of biomasses concentration is reasonably studied only inside the biofilm, Ω_1^τ, and, likewise, the growth expansion

pressure only appears where the biomass is present. However, the nutrients concentration is affected not only by the presence of biomass but also by the proximity of the biofilm, thus making its domain of solution, $\Omega^{\tau}_{\hat{H}_b}$, slightly larger. Therefore, it seems necessary to have any means to differentiate numerically between the different regions of a biofilm. The *Level Set method*, helps with this task while also lowering the computational cost by reducing the size of the domain of solution.

The Level Set method [89] is a mathematical tool used to track the movement of free boundaries (in our case, Γ^{τ}_0), under a certain velocity field. There are different ways to represent the free boundary in the context of the level set framework. In our problem, we are going to use two of them: time dependent eulerian representation and time independent eulerian representation.

In the first case, the method considers such boundary as the level set zero of a higher dimensional function called the *level set function*, Φ, which is a signed distance function defined in all Ω. Thus, given a point, (X, Y), the value $\Phi(X, Y, \tau)$ is the distance between the point and the free boundary, and its sign will be positive if the point is outside the domain enclosed by Γ^{τ}_0 and negative otherwise. In our case, the free boundary does not enclose a domain but divides Ω into two subsets, Ω^{τ}_1 and Ω^{τ}_2. Therefore, we will consider the positive sign if the point is in Ω^{τ}_2, and negative if it is in Ω^{τ}_1. Another important feature about a distance function, which results useful while extending velocities or reinitialising, is that, at every point, the following condition holds:

$$||\nabla\phi|| = 1. \tag{5.52}$$

Finally, as Φ is defined in all Ω, other level sets can be tracked as well. In particular, in our model it is necessary to keep track of the level set $\hat{H}_b$ that serves as upper boundary of $\Omega^{\tau}_{\hat{H}_b}$, the domain where the nutrients equation has to be solved.

In the time independent eulerian representation, the interface Γ is represented

implicitly by the function $\Phi(\vec{X})$, called *time of crossing map*. The isocontour $z = \Phi(\vec{X})$ is the position of the interface at time z, provided that it moves with velocity $\mathcal{S}$ in the normal direction to Γ. This representation is only valid in time independent problems. In our case, it is going to be used in the Fast Marching method [16, 90]. Otherwise, we consider the time dependent Eulerian representation.

The definition of the level set equation and the numerical strategy used to solve it, is described in the following sections.

5.4.1 Definition of the level set equation and overview of the level set strategy

The level set function tracks the movement of all the level sets in Ω, in particular the movement of Γ_0^τ. Let $\vec{X}(\tau) = (X(\tau), Y(\tau))$ be a point on the interface Γ_0^τ, moving with velocity $\vec{U}$. We can split $\vec{U}$ into its normal component and its tangential component to Γ_0^τ. Note that if the normal component is zero, then the flow is tangential to the interface, meaning that the interface is stationary. This argument implies that only the normal component contributes to the motion of the interface [18]. In our model, the velocity that moves Γ_0^τ is an external advective velocity, $\vec{U}$. As we only need the normal component, we define

$$F = \vec{U} \cdot \vec{n},$$

where $\vec{n} = \dfrac{\nabla\Phi}{||\nabla\Phi||}$, the unitary normal vector to Γ_0^τ. So, the motion equation for point $\vec{X}(\tau)$ on Γ_0^τ is,

$$\frac{d\vec{X}}{d\tau}(\tau) = F(\vec{X}(\tau))\vec{n}(\vec{X}(\tau)). \tag{5.53}$$

As $\vec{X}(\tau)$ is always on Γ_0^τ, then the following equation holds:

$$\Phi(\vec{X}(\tau), \tau) = 0, \quad \forall\tau. \tag{5.54}$$

By differentiating (5.54) with respect to τ, we obtain

$$\Phi_\tau(\vec{X}(\tau), \tau) + \nabla\Phi(\vec{X}(\tau), \tau) \cdot \frac{d\vec{X}}{d\tau}(\tau) = 0,$$

and by using (5.53), we get

$$\Phi_\tau + \nabla\Phi \cdot F\vec{n} = 0,$$

Next, by using the definition of $\vec{n}$, we obtain

$$\Phi_t + F||\nabla\Phi|| = 0, \tag{5.55}$$

which is the basic level set evolution equation.

The level set evolution equation is the most complex equation of the model to be numerically solved, because it requires a series of steps as part of the level set method developed in [89].

First of all, at the beginning of the numerical solution, the level set function must be appropriately defined by using the initial expression for Γ_0^0. Then, at every time iteration, $n+1$, the following steps are executed:

1. computation of the advance normal velocity $F^{n+1} = \vec{U}^{n+1} \cdot \vec{n}$ at the level set $\Gamma_0^{\tau_n}$,
2. extension of the velocity F^{n+1} from $\Gamma_0^{\tau_n}$ to Ω, thus obtaining F_e^{n+1},
3. advance of one time step in the solution of the level set equation, (4.32) in the Alpkvist model and (4.48) in the second model, obtaining Φ^{n+1},
4. reinitialisation of Φ^{n+1} (only when needed).

Figure 5.2 summarises the steps of the method. Once we have done all steps, we get the new level set function, Φ^{n+1} at time τ^{n+1}, that defines the solving domain for the system equations, $\Omega_1^{\tau_{n+1}}$ and $\Omega_{\hat{H}_b}^{\tau_{n+1}}$

5.4.2 The Fast Marching method

Before describing of each step of the level set strategy, we introduce a brief explanation on the Fast Marching method (FMM), a useful tool for capturing the motion of

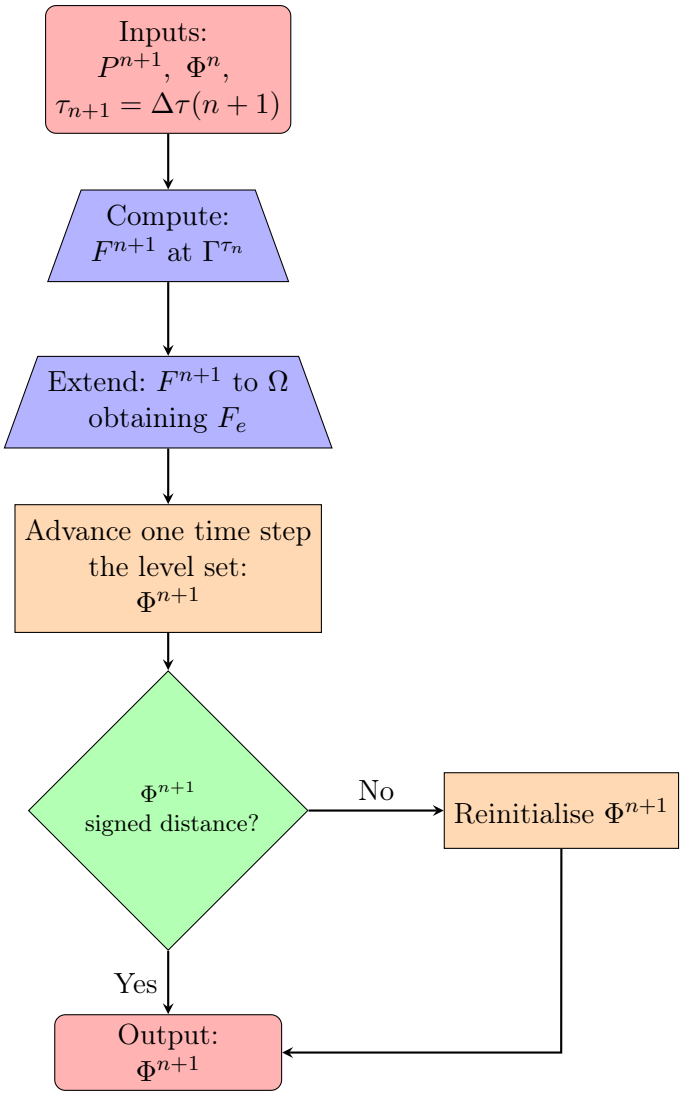

Figure 5.2: Flowchart of the level set method algorithm

interfaces developed by Sethian [90]. It is a quite novel method although actually fast when applicable. In our problem, it is used in the velocity extension step as well as in the reinitialisation process.

For the sake of the explanation, let us consider the situation of the reinitialisation step, even though this is the last step of the Level Set method. Assume that we have already advanced the level set function from time τ_n to time τ_{n+1}, i.e., we have already obtained the values of Φ^{n+1} at the mesh points, although only the zero level set moved correctly. This implies that Φ^{n+1} does not represent the values of a level set function. Therefore, condition (5.52) particularised on Φ^{n+1} is only valid near the zero level set Γ_0^{n+1}. Otherwise, only Φ^{n+1} holds the correct sign at the mesh points.

Let us denote $\Psi = \Phi^{n+1}$, saving the values of Φ^{n+1} in a new function so that we can discard Φ^{n+1} and obtain the correct reinitialised values by using Ψ as an input. We freeze temporarily the global algorithm in order to reinitialise so, for the sake of clarity, let us drop the temporal indices and denote $\tilde{\Phi} \equiv \Phi^{n+1}$, $\tilde{\Omega}_1 \equiv \Omega_1^{\tau_{n+1}}$, $\tilde{\Omega}_2 \equiv \Omega_2^{\tau_{n+1}}$ and $\tilde{\Gamma} \equiv \Gamma_0^{\tau_{n+1}}$. We want to obtain $\tilde{\Phi}$, a level set function so that

$$\tilde{\Phi}^{-1}(0) \equiv \tilde{\Gamma}, \quad ||\nabla\tilde{\Phi}|| = 1.$$

The FMM is a time independent method that uses the Eulerian time-independent representation of the level set function. As previously stated, such representation considers the interface $\tilde{\Gamma}$ represented implicitly by the function $\tilde{\Phi}(\vec{X})$, called time of crossing map, with $\mathcal{S}$ the velocity in the normal direction to $\tilde{\Gamma}$.

The basic Fast Marching Equation (FME) is

$$\mathcal{S}||\nabla\tilde{\Phi}|| = 1. \tag{5.56}$$

The velocity function $\mathcal{S}(\vec{X})$ must be monotonic and dependent only on $\vec{X}$ and not on $\tilde{\Phi}$, although these condition can be relaxed (see, for instance, [91, 16]). In the reinitialisation step, the velocity is $\mathcal{S} \equiv 1$.

In order to solve the FME, we must take into account that the information must travel outwards, i.e., from instant z' to instant z, with $z' < z$. Let be $\left\{\vec{X}_{i,j} = (X_i, Y_j) = (i\Delta X, j\Delta Y)\right\}_{i,j=0}^{M}$ the finite differences mesh of the domain Ω, with

$$\begin{aligned} X_0 = Y_0 = 0, \quad X_M = Y_M = 1, \\ \Delta X = X_i - X_{i-1} \quad \forall i, \text{ constant}, \\ \Delta Y = Y_j - Y_{j-1} \quad \forall j, \text{ constant}, \\ \Delta X = \Delta Y. \end{aligned}$$

Let us approximate $\tilde{\Phi}(\vec{X}_{i,j}) \approx \tilde{\Phi}_{i,j}$ and use the notation $\mathcal{S}(\vec{X}_{i,j}) \equiv \mathcal{S}_{i,j}$. The grid points of the mesh are sorted into three disjoint sets:

- $\mathbb{A}$ represents the set of accepted points. Initially it contains the grid points neighbouring $\tilde{\Gamma}$. The values of $\tilde{\Phi}$ at these points are calculated differently, depending on the available data.

- $\mathbb{T}$ represents the set of tentative points, which are the nodes adjacent the ones belonging to $\mathbb{A}$. For those nodes, a tentative value of $\tilde{\Phi}$ will be computed and one of those values is going to be accepted, thus moving the corresponding node from $\mathbb{T}$ to $\mathbb{A}$.

- $\mathbb{F}$ represents the set of far nodes, which are the remaining nodes of the grid. When the value from a node belonging to $\mathbb{T}$ is accepted, its neighbouring nodes are going to move from $\mathbb{F}$ to $\mathbb{T}$.

The FMM algorithm, summarised in Figure 5.3, proceeds as follows [17]:

1. Initialise grid points neighbouring the curve $\tilde{\Gamma}$, by calculating directly its $\tilde{\Phi}$ value. These are the initial points in $\mathbb{A}$.

2. Compute tentative values of $\tilde{\Phi}$ at those points adjacent to points in $\mathbb{A}$ that are not yet in $\mathbb{A}$. Those are the initial points in $\mathbb{T}$. The remaining points constitute the subset $\mathbb{F}$.

3. Choose the point $\vec{X}_{i,j}$ from $\mathbb{T}$ that satisfies

$$|\tilde{\Phi}_{i,j}| = \min_{\vec{X}_{r,s} \in \mathbb{T}} |\tilde{\Phi}_{r,s}|$$

and move $\vec{X}_{i,j}$ to A.

4. Compute tentative values at the points adjacent to $\vec{X}_{i,j}$ that are not in $\mathbb{A}$ (including points already belonging to $\mathbb{T}$).

5. Repeat until $\mathbb{F} \cup \mathbb{T} = \emptyset$.

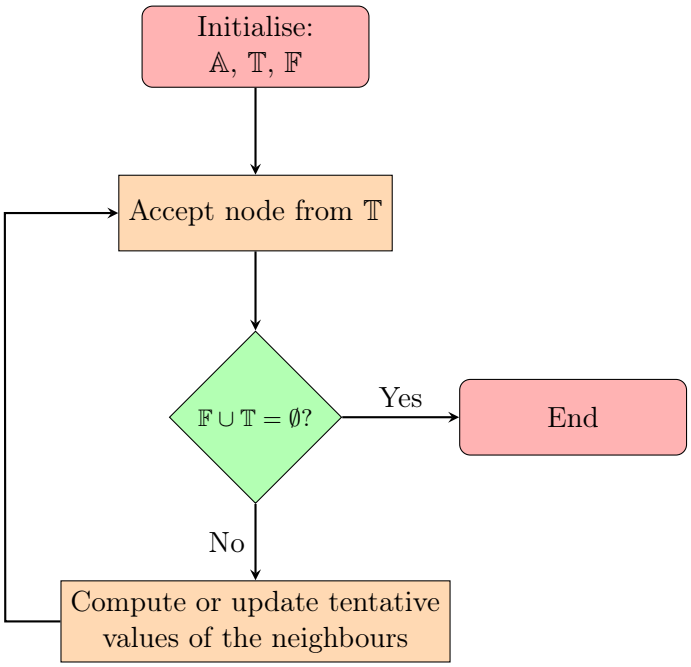

Figure 5.3: Flowchart of the Fast Marching Method

Initialising the set of accepted nodes

The available data for the initialisation of the set $\mathbb{A}$ depends on the problem and, therefore, the strategies vary. As previously stated, the situation assumed to explain

the FMM is the one observed in the reinitialisation step. Therefore, let us consider the case were we have a function Ψ that is not a level set function, but satisfies $\Psi^{-1}(0) = \tilde{\Gamma}$ and is correctly oriented, i.e., for any grid point $\vec{X}_{i,j}$, $\text{sign}(\Psi_{i,j})$ is positive in $\tilde{\Omega}_2$ and negative in $\tilde{\Omega}_1$. Moreover, we only know the values of Ψ at the gridpoints. In order to initialise $\mathbb{A}$ we proceed as follows.

First of all, let a *voxel* (denoted by $V^{i,j}$) be a square of the grid, i.e., the region bounded by grid points $\vec{X}_{i,j}, \vec{X}_{i,j+1}, \vec{X}_{i+1,j}, \vec{X}_{i+1,j+1}$ (see Figure 5.4). The points belonging to the initial accepted subset will be the corners of the voxels crossed by the curve. Locating those points is not difficult, as we only need to check the sign of Ψ at the corners of the voxel and check if it changes. Once a voxel in these conditions has been found, the values of $\tilde{\Phi}$ on the corners are computed by locating the zero level set inside the voxel, so that the distance from each corner to $\tilde{\Gamma}$ can be approximated. For this purpose, we use a bicubic interpolation algorithm. We compute a bicubic polinomial $p(X, Y)$ such as $p^{-1}(0) \approx \tilde{\Gamma}$ inside the voxel $V^{i,j}$. Now, for every corner of $V^{i,j}$ we must find the closest point (u, w) belonging to $p^{-1}(0)$ and let

$$\tilde{\Phi}_{i,j} = \frac{||\vec{X}_{i,j} - (u, w)||}{\mathcal{S}(u, w)}.$$

As previously mentioned, there are other possible situations. For instance, in some cases the curve $\tilde{\Gamma}$ is given parametrically. In other cases, we may know the analytical expression of Ψ. In the first case, the initial values may be obtained using a bisection strategy, whereas in the second case, the same previously described strategy is used to obtain the initial values.

The bicubic interpolant

As part of the FMM, the zero level set inside a particular voxel, $V^{i,j}$, must be located. However, unless we know explicitly the analytical expression of $\tilde{\Gamma}$, we need to compute an approximation. In our case, the function Ψ satisfies $\Psi^{-1}(0) \equiv \tilde{\Gamma}$. Therefore, we can use the values of Ψ at the nodes of the mesh to locate $\tilde{\Gamma}$ inside

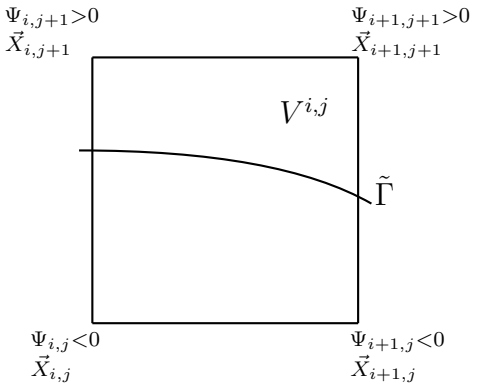

Figure 5.4: Definition of the voxel $V^{i,j}$.

the voxel $V^{i,j}$. For this purpose, we use a bicubic interpolation strategy mainly because it is arguably easily computed and it can be shown that it results to give a $C^{(1)}$ on the whole domain, thus contributing to greater stability and better results [18].

The bicubic interpolation strategy is a well known process (see for instance [88]). As we are working in a rectangular domain, we can map the voxel $V^{i,j}$ to the unit square, $\hat{V}$, where the interpolation is easier to compute (see Figure 5.5).

Finding the closest point

Let us assume that we are already in the reference voxel, $\hat{V}$, and we have already computed the bicubic polynomial, $\hat{p}$, that satisfies $\hat{p}^{-1}(0) \approx \hat{\Gamma}$, where $\hat{\Gamma}$ represents the mapping of the curve $\tilde{\Gamma}$ in $\hat{V}$. For each corner of the voxel, we want to locate the nearest point $(\hat{u}, \hat{w})$, belonging to $\hat{\Gamma}$, so lets call $\vec{X}^c = (X^c, Y^c)$ our candidate corner.

In order to compute the point $(\hat{u}, \hat{w})$ we solve the following equations:

$$\hat{p}(\hat{u}, \hat{w}) = 0, \tag{5.57}$$

$$\vec{k} \cdot [((\hat{u}, \hat{w}) - (X^c, Y^c)) \times \nabla \hat{p}(\hat{u}, \hat{w})] = 0, \tag{5.58}$$

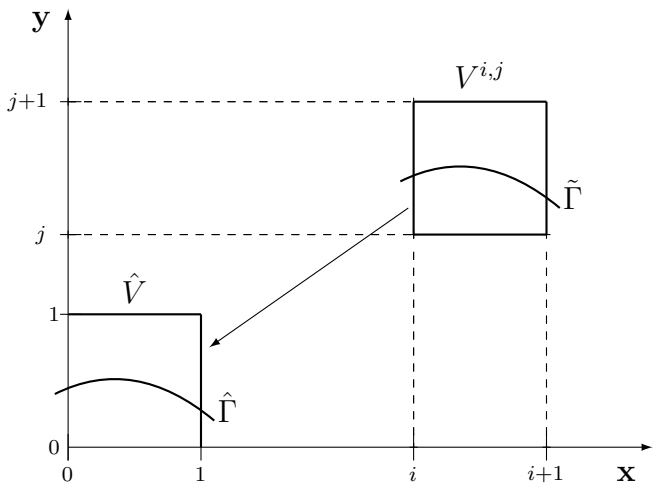

Figure 5.5: Voxel mapping.

where $\vec{k}$ represents the unitary vector in the direction normal to the plane XY. Equation (5.57) guarantees that the point $(\hat{u}, \hat{w})$ is in the interface $\hat{\Gamma}_0$ and equation (5.58) implies that $(\hat{u}, \hat{w})$ is the closest point to (X^c, Y^c).

In order to solve these equations, we follow an iterative strategy (see Figure 5.6). We will consider as the initial iterate the centre of the voxel $\hat{V}$, denoted by $\vec{X}^0 = (X^0, Y^0)$. The algorithm reads as follows:

1. We make one step in the gradient method to minimise the function p, that is

$$\vec{X}^1 = \vec{X}^0 + \alpha \nabla \hat{p}(\vec{X}^0),$$

   where

$$\alpha = \frac{-\hat{p}(\vec{X}^0)}{||\nabla \hat{p}(\vec{X}^0)||^2}$$

   Note that p is a bicubic polinomial. Therefore, its expression and gradient are available, so that α can be easily computed and, thus, $\vec{X}^1$.

2. For $\eta > 1$, starting with $(X^{\eta-1}, Y^{\eta-1})$ solve equation (5.58), obtaining the point $(\tilde{X}^\eta, \tilde{Y}^\eta)$. For this purpose, we consider the vector $\vec{m}_{\eta-1}$, the unitary vector

orthogonal to the gradient of $\hat{p}$ at $(X^{\eta-1}, Y^{\eta-1})$, i.e.,

$$\vec{m}_{\eta-1} \cdot \nabla \hat{p}(X^{\eta-1}, Y^{\eta-1}) = 0, \tag{5.59}$$

The point $(\tilde{X}^{\eta}, \tilde{Y}^{\eta})$ is the intersection of the line passing through $(X^{\eta-1}, Y^{\eta-1})$ with direction vector $\vec{m}_{\eta-1}$ with the line defined by (X^c, Y^c) and the direction vector $\nabla \hat{p}(X^{\eta-1}, Y^{\eta-1})$.

3. Repeat step 1 using $(\tilde{X}^{\eta}, \tilde{Y}^{\eta})$ as seed to obtain (X^{η}, Y^{η}).

4. If the tolerance conditions are fulfilled, set $(\hat{u}, \hat{w}) = (X^{\eta}, Y^{\eta})$. Otherwise, we repeat from step 2 until the tolerance conditions are fulfilled.

Once the point $(\hat{u}, \hat{w})$ in the reference voxel, $\hat{V}$, is found, we can go back to $V^{i,j}$ by using

$$(u, w) = (X_i, Y_j) + (\Delta X \hat{u}, \Delta Y \hat{w}).$$

Note that the solution (u, w) may not lie inside the voxel $V^{i,j}$. If that is the case, then $(\hat{u}, \hat{w})$ does not lie inside the reference voxel, $\hat{V}$. As the data used to construct $\hat{p}$ is related to $\hat{V}$, the polynomial is only accurate inside the voxel and $\hat{p}^{-1}(0)$ may be very different from the mapped curve, $\hat{\Gamma}$. Therefore, we simply discard (u, w). The correct solution will be obtained when the voxel adjacent to $V^{i,j}$ that shares the candidate corner is analysed.

Computing tentative approximations

Once the initial accepted set is obtained, we can compute the tentative approximations for the neighbouring nodes. This is achieved discretising the FME (5.56) using upwind finite differences. The upwinding direction depends on the available data. For instance, assume that we want to calculate the tentative value of $\tilde{\Phi}$ at $\vec{X}_{i,j} \in \mathbb{T}$. Let us suppose that $\vec{X}_{i+1,j}$ and $\vec{X}_{i,j+1}$ belong to $\mathbb{A}$. In such a case, equation (5.56) is discretised as

$$\mathcal{S}_{i,j} \left[\left(\frac{\tilde{\Phi}_{i+1,j} - \tilde{\Phi}_{i,j}}{\Delta X} \right)^2 + \left(\frac{\tilde{\Phi}_{i,j+1} - \tilde{\Phi}_{i,j}}{\Delta Y} \right)^2 \right]^{1/2} = 1. \tag{5.60}$$

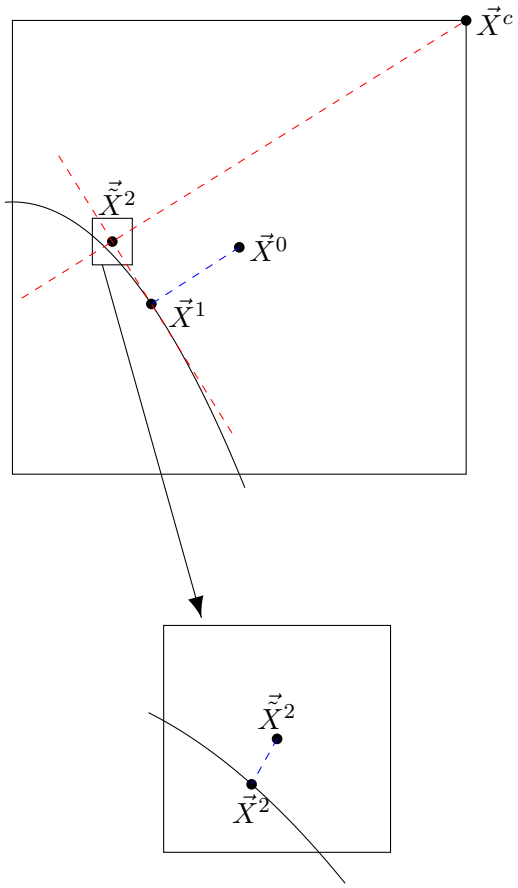

Figure 5.6: Sketch of the iterative process used to find the closest point.

which, in fact, is the only option available, as $\vec{X}_{i-1,j}$ and $\vec{X}_{i,j-1}$ belong to $\mathbb{T} \cup \mathbb{F}$.

Now, if we have more than one neighbour in the same direction providing data, we should pick the one with smaller $|\tilde{\Phi}|$ value. For instance, let us assume that $\vec{X}_{i+1,j}$, $\vec{X}_{i-1,j}$ and $\vec{X}_{i,j+1}$ belong to $\mathbb{A}$. In that case, we must compare the values $|\tilde{\Phi}_{i+1,j}|$ and $|\tilde{\Phi}_{i-1,j}|$ so that

- If $|\tilde{\Phi}_{i+1,j}| \leq |\tilde{\Phi}_{i-1,j}|$, we use discretisation 5.60.

- Otherwise, equation (5.56) is discretised as

$$\mathcal{S}_{i,j}\left[\left(\frac{\tilde{\Phi}_{i,j}-\tilde{\Phi}_{i-1,j}}{\Delta X}\right)^2+\left(\frac{\tilde{\Phi}_{i,j+1}-\tilde{\Phi}_{i,j}}{\Delta Y}\right)^2\right]^{1/2}=1. \tag{5.61}$$

If the four neighbours of $\vec{X}_{i,j}$ belong to $\mathbb{A}$, then we simply apply this argument to each variable separately.

Finally, if we have data available only in one direction, we must discard the derivative in the other direction. For instance, suppose that $\vec{X}_{i+1,j}$ belongs to $\mathbb{A}$ but $\vec{X}_{i-1,j}$, $\vec{X}_{i,j+1}$ and $\vec{X}_{i,j-1}$ belong to $\mathbb{T}\cup\mathbb{F}$. In that case, equation (5.56) is discretised as

$$\mathcal{S}_{i,j}\left|\frac{\tilde{\Phi}_{i+1,j}-\tilde{\Phi}_{i,j}}{\Delta X}\right|=1. \tag{5.62}$$

Now, if we recall that the information travels from nodes with time of crossing z' to nodes with time of crossing $z>z'$, then if $\vec{X}_{i+1,j}\in\mathbb{A}$, the following must be satisfied

$$\tilde{\phi}_{i+1,j}\leq\tilde{\phi}_{i,j}.$$

Therefore, (5.62) results in

$$\tilde{\Phi}_{i,j}=\tilde{\Phi}_{i+1,j}+\frac{\Delta X}{\mathcal{S}_{i,j}}, \tag{5.63}$$

which is the solution of the FME (5.56) if the data available is located only in one direction. When the data available is present in both direction, solving the FME is a bit more complicated. Let us go back to the discretised FME (5.60). In this equation, the terms $\mathcal{S}_{i,j}$, $\tilde{\Phi}_{i+1,j}$ and $\tilde{\Phi}_{i,j+1}$ are known, whereas the unknown is $\tilde{\Phi}_{i,j}$. Rearranging terms in equation (5.60) we obtain the quadratic equation

$$\left(\frac{\Delta X}{\Delta Y}+\frac{\Delta Y}{\Delta X}\right)\tilde{\Phi}_{i,j}^2-2\left(\frac{\Delta X}{\Delta Y}\tilde{\Phi}_{i,j+1}+\frac{\Delta Y}{\Delta X}\tilde{\Phi}_{i+1,j}\right)+\left(\frac{\Delta X}{\Delta Y}\tilde{\Phi}_{i,j+1}^2+\frac{\Delta Y}{\Delta X}\tilde{\Phi}_{i+1,j}^2-\frac{\Delta X\Delta Y}{\mathcal{S}_{i,j}}\right)=0, \tag{5.64}$$

the discriminant of which is

$$d=4\left(\frac{\Delta X^2+\Delta Y^2}{\mathcal{S}_{i,j}^2}-(\tilde{\Phi}_{i+1,j}-\tilde{\Phi}_{i,j+1})^2\right). \tag{5.65}$$

Depending on the value of d, equation (5.64) may have no solution ($d < 0$), one solution ($d = 0$) or two solutions ($d > 0$). The case where $d = 0$ does not require further discussion, but the others do. Since $\mathcal{S}$ is monotonic, let us assume that $\mathcal{S} > 0$ (the negative case is analogous). Then, if $d < 0$

$$\frac{\sqrt{\Delta X^2 + \Delta Y^2}}{\mathcal{S}_{i,j}} < |\tilde{\Phi}_{i+1,j} - \tilde{\Phi}_{i,j+1}|. \tag{5.66}$$

The left term in (5.66) represents the elapsed time to travel from $\vec{X}_{i+1,j}$ to $\vec{X}_{i,j+1}$, whereas the right term represents the assigned elapsed time, as $\tilde{\Phi}$ is the time of crossing map (see Figure 5.7). Therefore, the elapsed time is smaller that the currently assigned elapsed time. This usually happens when the interface crosses $\vec{X}_{i+1,j}$ and $\vec{X}_{i,j+1}$ almost simultaneously. Suppose that $\tilde{\Phi}_{i+1,j} < \tilde{\Phi}_{i,j+1}$, i.e., the interface arrives at $\vec{X}_{i+1,j}$ before arriving at $\vec{X}_{i,j+1}$. Then, as the time window is small, we simply use the value of $\vec{X}_{i+1,j}$ and discard the other direction. The argument is analogous if $\tilde{\Phi}_{i,j+1} < \tilde{\Phi}_{i+1,j}$. Therefore, if $d < 0$, we simply take

$$\tilde{\Phi}_{i,j} = \min(\tilde{\Phi}_{i+1,j} + \frac{\Delta X}{\mathcal{S}_{i,j}}, \tilde{\Phi}_{i,j+1} + \frac{\Delta Y}{\mathcal{S}_{i,j}}). \tag{5.67}$$

If $d > 0$, then the solutions of equation (5.64) are

$$\tilde{\Phi}_{i,j} = \frac{\left(\frac{\Delta Y}{\Delta X}\tilde{\Phi}_{i+1,j} + \frac{\Delta X}{\Delta Y}\tilde{\Phi}_{i,j+1}\right) \pm \sqrt{\frac{\Delta X^2 + \Delta Y^2}{\mathcal{S}_{i,j}^2} - (\tilde{\Phi}_{i+1,j} - \tilde{\Phi}_{i,j+1})^2}}{\frac{\Delta X}{\Delta Y} + \frac{\Delta Y}{\Delta X}} \tag{5.68}$$

Let us suppose that $\tilde{\Phi}_{i,j+1} = 0$. Then (5.68) is

$$\tilde{\Phi}_{i,j} = \frac{\left(\frac{\Delta Y}{\Delta X}\tilde{\Phi}_{i+1,j}\right) \pm \sqrt{\frac{\Delta X^2 + \Delta Y^2}{\mathcal{S}_{i,j}^2} - \tilde{\Phi}_{i+1,j}^2}}{\frac{\Delta X}{\Delta Y} + \frac{\Delta Y}{\Delta X}} \tag{5.69}$$

Provided that we want $\tilde{\phi}_{i+1,j} \leq \tilde{\phi}_{i,j}$, the root that gives the appropriate solution is the positive root. Therefore, if $d > 0$ the solution of (5.64) is the positive root in (5.68).

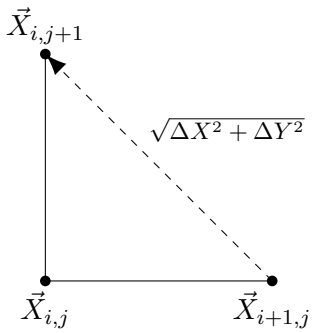

Figure 5.7: Elapsed time between nodes.

Accepting new nodes

Once the tentative values have been computed, the next step in the FMM algorithm is to find the node $\vec{X}_{i,j}$ that satisfies

$$|\tilde{\Phi}_{i,j}| = \min_{\vec{X}_{r,s} \in \mathbb{T}} |\tilde{\Phi}_{r,s}|.$$

Once it has been found, the node is moved from $\mathbb{T}$ to $\mathbb{A}$, and the value $\tilde{\Phi}_{i,j}$ is accepted. After that, the tentative values at the neighbours of $\vec{X}_{i,j}$ must be computed, at those neighbours in $\mathbb{F}$, or updated, at those neighbours in $\mathbb{T}$. The neighbours that already belong to $\mathbb{A}$ maintain their $\tilde{\Phi}$ value unchanged.

In order to speed up the process of accepting new nodes, it comes in handy to have the set $\mathbb{T}$ ordered in a binary tree structure attending to the tentative values of $\tilde{\Phi}$. Each node of the tree represents a grid point belonging to $\mathbb{T}$. Given a node of the tree, the absolute tentative value associated to it must be smaller than the absolute tentative value of its children (see Figure 5.8). That way, the accepted node is always the top node.

When a tentative value is computed, if the corresponding node was moved from $\mathbb{F}$ to $\mathbb{T}$, then it must be placed at the bottom of the tree and moved upwards to its right position. If the node was already in $\mathbb{T}$, as its tentative value was updated, we may

need to move it upwards (if its absolute tentative value is smaller than before), or downwards (if its absolute tentative value has increased), to its right position.

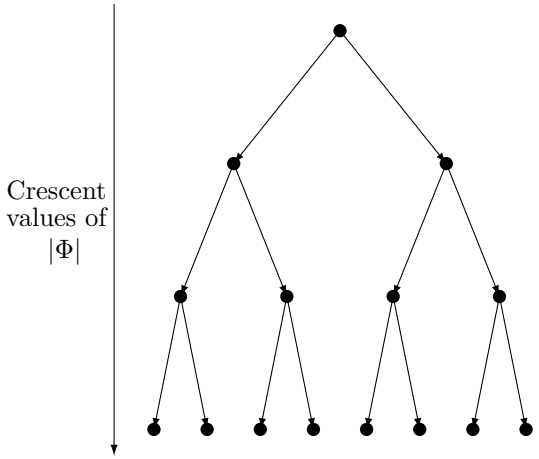

Figure 5.8: Binary tree ordering.

5.4.3 Initialization of the level set function

The initial expression of the interface, Γ_0^0, is given by the initial condition (4.39). However, as we do not know the initial expression of the level set function, Φ^0, we must build it accordingly.

Let us recall that the level set function is a signed distance function. This is not required arbitrarily, it is recommended for stability purposes. Therefore, to compute the initial level set function, one has to take care of two things: the sign and the distance.

a) The sign

Provided that we know at least a discretisation of the initial interface Γ_0^0, we can use a topological strategy for determining the sign at any point: counting interface crossings. First, we must pick a given node with known sign. Then,

to determine the sign of a target point, we draw a path following gridlines from the target to the given node, and count the number of transversal interface crossings. If the resulting number is even, then the target node is in the same side as the given node so their signs are the same. If it is odd, they are in opposite sides, so that the signs are opposite as well. One must take care to count only transverse crossings and note that for, some points, zero crossings can happen.

Although this is quite simple, the particular problem can make the task even easier. In our case, as we explicitly know the analytical expression of the initial Γ_0^0, we can simply take the sign as negative if the grid point is in Ω_2^0 and positive if it is in Ω_2^0 . Note that if a point is on Γ_0^0, its level set value is zero.

b) The distance

Let us assume again that we have a discretisation of the initial interface $\Gamma_0^0 \approx \{(X_{c_k}, Y_{c_k})\}_{k=0}^K$, ordered clockwise. For each point belonging to the mesh, $\vec{X}_{i,j}$, we need to compute the distance to Γ_0^0. For this purpose, first we compute the distance, $d_k^{i,j}$, to every line segment of the discretised curve, defined by (X_{c_k}, Y_{c_k}) and $(X_{c_{k+1}}, Y_{c_{k+1}})$. Then, we take the minimum of those distances as the absolute value of the level set function at $\vec{X}_{i,j}$, i.e,

$$|\Phi_{i,j}| = \min_k \{d_k^{i,j}\}.$$

Despite its simplicity, one must be cautious when computing those values. For instance, note that $d_k^{i,j}$ are distances to a *line segment* and not to a line. Furthermore, if the curve is closed, the distance to the segment defined by (X_{c_K}, Y_{c_K}) and (X_{c_0}, Y_{c_0}) must be computed. However, in our case the curve is not closed.

As a side note, there are other possible initial situations. The best case scenario is to know the initial expression of the level set function, Φ, although this is the least common one. In this case, nothing needs to be done. In other cases, the available data

is a discretisation of the initial interface Γ_0^0, and we may proceed as described above. Finally, the least common situation is to know the values of the level set function at the grid points surrounding the interface. In this case, we can use the FMM to easily and adequately compute the level set function at the rest of the domain. This process is the same as in the reinitialisation step, and will be described later in this work.

5.4.4 Velocity extension

As indicated in Section 5.4.1, at each time iteration we need to compute the advance velocity, F^{n+1}, and its appropriate extension to Ω, F_e^{n+1}, before advancing one time step the level set equation.

The velocity that moves the level sets should come from the expansive growth of the biofilm, as expressed in equation (4.31), in the Alpkvist model, and (4.47), in the second model. This velocity can be computed after obtaining the updated values of the pressure P^{n+1}.

However, two problems arise. The first one is that P^{n+1}, and therefore $\vec{U}^{n+1}$, is known only in Ω_1^n, whereas Φ^{n+1} has to be computed in Ω, the solving domain for the level set equation. The second problem is that even if $\vec{U}^{n+1}$ were known in Ω, we can only ensure that the only level set that is going to move correctly is $\Gamma_0^{\tau_n}$, whereas the rest may go offtrack. Therefore, we can only use $\vec{U}^{n+1}$ in points belonging to $\Gamma_0^{\tau_n}$. We need to extend accordingly $\vec{U}^{n+1}$ to all Ω, so that if $||\nabla\Phi^n|| = 1$ then $||\nabla\Phi^{n+1}|| = 1$. We will use the values of $\vec{U}^{n+1}$ at $\Gamma_0^{\tau_n}$ to obtain such an extension.

Finally, as explained in Section 5.4.1, we are only interested in the normal component of $\vec{U}^{n+1}$. Therefore, we restrict ourselves to the computation of $F^{n+1} = \vec{U}^{n+1} \cdot \vec{n}$ at $\Gamma_0^{\tau_n}$ and the extension F_e^{n+1} to Ω.

Just as a reminder, note that at this point of the solving strategy, Φ^n and P^{n+1} are

known at grid points, whereas F^{n+1} and F_e^{n+1} are unknown.

Definition of the velocity extension equation

The velocity extension equation arises from the fact that we want the level set function to move in such a way that it keeps being a signed distance function. Assume that we have F, a scalar velocity, that preserves the level set function, Φ as a signed distance function, i.e.,

$$||\nabla\Phi|| = 1.$$

Then, we get

$$\frac{d}{d\tau}(\nabla\Phi \cdot \nabla\Phi) = 2\nabla\Phi_\tau \cdot \nabla\Phi = 0,$$

Combining this identity with equation (5.55), we obtain

$$2\nabla(-F||\nabla\Phi||) \cdot \nabla\Phi = 0. \tag{5.70}$$

Finally, since Φ is a signed distance function, $||\nabla\Phi|| = 1$, so that equation (5.70) becomes

$$\nabla F \cdot \nabla\Phi = 0. \tag{5.71}$$

Note that $\nabla\Phi$ is normal to the interface, and ∇F is orthogonal to lines of constant F, so that F is constant along the lines which are normal to the interface. Therefore, we have here a propagation of information problem to which makes sense to apply the FMM. The strategy of the FMM applied to (5.71) is the same as the one applied for the FME (5.56) with certain specific issues [17].

Initialising the set of accepted nodes

In Section 5.4.2 we described the initialisation of the set of accepted nodes, $\mathbb{A}$, for the case where the function that needed to be extended was the level set function Φ.

That was rather simple, because at those grid points located in the vicinity of the interface, $\Gamma_0^{\tau_n}$, we only needed to find the closest point belonging to the interface and compute the distance.

In many problems, initialising the values of the speed function F^{n+1} is not much more complicated. If $\vec{X}_{i,j}$ is a point neighbouring $\Gamma_0^{\tau_n}$, we just need to locate the closest point, $\vec{X}_I$, and simply set $F_{e_{i,j}}^{n+1} = F^{n+1}(\vec{X}_I)$. However, we have another difficulty here: we do not know the values of F at the interface, as they depend on the values of the gradient of the pressure, which is a magnitude that we only know at certain grid points.

This can be overcame by using the following strategy:

1. find a voxel, $V^{i,j}$, crossed by the interface, $\Gamma_0^{\tau_n}$.
2. at every corner of the voxel, $\vec{X}_c$, find the closest point in the interface, $\vec{X}_I$, by using the same strategy as described in Section 5.4.2. At this point the value of the pressure, $P^{n+1}(\vec{X}_I)$, is going to be zero because of the boundary conditions (4.36) of the problem.
3. at $\vec{X}_I$ follow the direction normal to the interface, i.e., the gradient of Φ^n and find the point located at $(\Phi^n)^{-1}(-2\Delta x)$. Let us denote it $\vec{X}_{-2I}$.
4. locate the voxel where that point is enclosed, $V^{p,q}$, and compute the bicubic interpolation of the pressure in the voxel, $\tilde{p}_{p,q}$. Approximate the value of the pressure at that point by using the bicubic polinomial, $P^{n+1}(\vec{X}_{-2I}) \approx \tilde{p}_{p,q}(\vec{X}_{-2I})$.
5. finally, recall that at $\Gamma_0^{\tau_n}$, $F^{n+1} = -\nabla P^{n+1} \cdot \nabla \Phi^n$, which is the negative normal derivative of the pressure, $F^{n+1} = -\dfrac{\partial P^{n+1}}{\partial \vec{n}}$. We can approximate the value of the speed function at $\vec{X}_I$ as

$$F^{n+1}(\vec{X}_I) \approx -\frac{P^{n+1}(\vec{X}_I) - P^{n+1}(\vec{X}_{-2I})}{2\Delta X} \approx \frac{\tilde{p}_{p,q}(\vec{X}_{-2I})}{2\Delta X}.$$

Once we have $F^{n+1}(\vec{X}_I)$ we simply take $F_e^{n+1}(\vec{X}_c) = F^{n+1}(\vec{X}_I)$.

When we are computing the bicubic interpolation of the pressure inside the voxel $V^{p,q}$, if we are too close to the level set zero, then we may need to use values of the pressure outside the biofilm to compute the bicubic polynomial. If this is the case, an extrapolation technique is required for the absent values, as it greatly improves the result of the interpolation. This case should not be that usual, and even when it happens, the number of absent values should be small.

The extrapolation method

In order to obtain the extrapolated values needed to compute the bicubic polynomial, we solve equation

$$\Delta P^{n+1} = P_{xx}^{n+1} + P_{yy}^{n+1} = 0, \tag{5.72}$$

in the discrete domain defined by the nodes used in the bicubic interpolation, $\mathcal{D} = \{X_{p-1}, X_p, X_{p+1}, X_{p+2}\} \times \{Y_{q-1}, Y_q, Y_{q+1}, Y_{q+2}\}$.

Equation (5.72) is discretised using backward or forward finite differences, depending on which node we are discretising at, and always using values at nodes belonging to $\mathcal{D}$. In particular,

- If we are extrapolating at nodes (X_{p-1}, Y_η) and (X_p, Y_η), for $\eta = q-1, \ldots, q+2$, then we use forward finite differences scheme to approximate P_{xx}^{n+1}

$$P_{xx_{\tau,\eta}}^{n+1} \approx \frac{P_{\zeta+2,\eta}^{n+1} - 2P_{\zeta+1,\eta}^{n+1} + P_{\zeta,\eta}^{n+1}}{\Delta X^2}, \quad \text{for } \zeta = p-1, p \text{ and } \eta = q-1, \ldots, q+2.$$

- If we are extrapolating at nodes (X_{p+1}, Y_η) and (X_{p+2}, Y_η), for $\eta = q-1, \ldots, q+2$, then we use backward finite differences scheme to approximate P_{xx}^{n+1}

$$P_{xx_{\tau,\eta}}^{n+1} \approx \frac{P_{\zeta-2,\eta}^{n+1} - 2P_{\zeta-1,\eta}^{n+1} + P_{\zeta,\eta}^{n+1}}{\Delta X^2}, \quad \text{for } \zeta = p+1, p+2 \text{ and } \eta = q-1, \ldots, q+2.$$

- If we are extrapolating at nodes (X_ζ, Y_{q-1}) and (X_ζ, Y_q), for $\zeta = p-1, \ldots, p+2$, then we use forward finite differences scheme to approximate P^{n+1}_{yy}

$$P^{n+1}_{yy_{\tau,\eta}} \approx \frac{P^{n+1}_{\zeta,\eta+2} - 2P^{n+1}_{\zeta,\eta+1} + P^{n+1}_{\zeta,\eta}}{\Delta Y^2}, \quad \text{for } \zeta = p-1, \ldots, p+2 \text{ and } \eta = q-1, q.$$

- If we are extrapolating at nodes (X_ζ, Y_{q+1}) and (X_ζ, Y_{q+2}), for $\zeta = p-1, \ldots, p+2$, then we use backward finite differences scheme to approximate P^{n+1}_{yy}

$$P^{n+1}_{yy_{\tau,\eta}} \approx \frac{P^{n+1}_{\zeta,\eta-2} - 2P^{n+1}_{\zeta,\eta-1} + P^{n+1}_{\zeta,\eta}}{\Delta Y^2}, \quad \text{for } \zeta = p-1, \ldots, p+2, \text{ and } \eta = q+1, q+2.$$

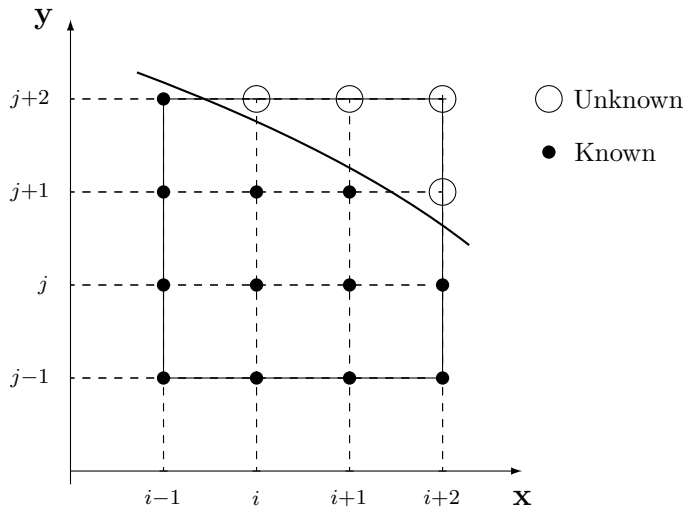

Figure 5.9: Example of missing values on $\mathcal{D}$ that need extrapolation.

Equation (5.72) is posed in the domain $\mathcal{D}$, where the values of P^{n+1} at certain nodes of $\mathcal{D}$ are known. These values are used as an input. Assume that the unknown values of P^{n+1} are those related to the nodes $\vec{X}_{p+2,q+1}, \vec{X}_{p,q+2}, \vec{X}_{p+1,q+2}$ and $\vec{X}_{p+2,q+2}$, while the rest are known values, as depicted in Figure 5.9. The discretised equations for

those nodes are

$$\Delta P^{n+1}(\vec{X}_{p+2,q+1}) \approx \frac{P^{n+1}_{p,q+1} - 2P^{n+1}_{p+1,q+1} + P^{n+1}_{p+2,q+1}}{\Delta X^2} + \frac{P^{n+1}_{p+2,q-1} - 2P^{n+1}_{p+2,q} + P^{n+1}_{p+2,q+1}}{\Delta Y^2},$$
$$\Delta P^{n+1}(\vec{X}_{p,q+2}) \approx \frac{P^{n+1}_{p+2,q+2} - 2P^{n+1}_{p+1,q+2} + P^{n+1}_{p,q+2}}{\Delta X^2} + \frac{P^{n+1}_{p,q} - 2P^{n+1}_{p,q+1} + P^{n+1}_{p,q+2}}{\Delta Y^2},$$
$$\Delta P^{n+1}(\vec{X}_{p+1,q+2}) \approx \frac{P^{n+1}_{p-1,q+2} - 2P^{n+1}_{p,q+2} + P^{n+1}_{p+1,q+2}}{\Delta X^2} + \frac{P^{n+1}_{p+1,q} - 2P^{n+1}_{p+1,q+1} + P^{n+1}_{p+1,q+2}}{\Delta Y^2},$$
$$\Delta P^{n+1}(\vec{X}_{p+2,q+2}) \approx \frac{P^{n+1}_{p,q+2} - 2P^{n+1}_{p+1,q+2} + P^{n+1}_{p+2,q+2}}{\Delta X^2} + \frac{P^{n+1}_{p+2,q} - 2P^{n+1}_{p+2,q+1} + P^{n+1}_{p+2,q+2}}{\Delta Y^2}.$$

By rearranging terms and considering $\Delta X = \Delta Y$, the discretised equations become

$$P^{n+1}_{p+2,q+1} = P^{n+1}_{p+1,q+1} + P^{n+1}_{p+2,q} - \frac{1}{2}(P^{n+1}_{p,q+1} + P^{n+1}_{p+2,q-1}), \tag{5.73}$$
$$2P^{n+1}_{p,q+2} - 2P^{n+1}_{p+1,q+2} + P^{n+1}_{p+2,q+2} = 2P^{n+1}_{p,q+1} - P^{n+1}_{p,q}, \tag{5.74}$$
$$-P^{n+1}_{p,q+2} + P^{n+1}_{p+1,q+2} = P^{n+1}_{p+1,q+1} - \frac{1}{2}(P^{n+1}_{p-1,q+2} + P^{n+1}_{p+1,q}), \tag{5.75}$$
$$-2P^{n+1}_{p+2,q+1} + P^{n+1}_{p,q+2} - 2P^{n+1}_{p+1,q+2} + 2P^{n+1}_{p+2,q+2} = -P^{n+1}_{p+2,q}. \tag{5.76}$$

The system of equations (5.73)-(5.76) can be written in matrix form as

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & -2 & 1 \\ 0 & -1 & 1 & 0 \\ -2 & 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} P^{n+1}_{p+2,q+1} \\ P^{n+1}_{p,q+2} \\ P^{n+1}_{p+1,q+2} \\ P^{n+1}_{p+2,q+2} \end{pmatrix} = \begin{pmatrix} P^{n+1}_{p+1,q+1} + P^{n+1}_{p+2,q} - \frac{1}{2}(P^{n+1}_{p,q+1} + P^{n+1}_{p+2,q-1}) \\ 2P^{n+1}_{p,q+1} - P^{n+1}_{p,q} \\ P^{n+1}_{p+1,q+1} - \frac{1}{2}(P^{n+1}_{p-1,q+2} + P^{n+1}_{p+1,q}) \\ -P^{n+1}_{p+2,q} \end{pmatrix}, \tag{5.77}$$

which is solved by using a LU algorithm with partial pivoting, thus obtaining the extrapolated values of the pressure. With those values, and the rest of the known values associated to nodes belonging to $\mathcal{D}$, the bicubic polynomial can be obtained.

Computing the tentative values

Once the initial values are obtained, the tentative values must be calculated and ordered in a tree structure following the crescent values of $|\Phi^n|$, as explained in

Section 5.4.2. In order to compute the tentative values, we need to solve equation (5.71), particularised for $|\Phi^n|$, and with unknown speed function F_e^{n+1}. For this purpose, we propose an upwind finite differences scheme for the spatial discretisation. The upwinding direction depends on the available data and follows the direction of increasing $|\Phi^n|$ values.

Let us assume that we need to compute the tentative values of the velocity at point $\vec{X}_{i,j}$, that is, $F_{e_{i,j}}^{n+1}$. Suppose that $\vec{X}_{i+1,j}$ and $\vec{X}_{i,j+1}$ belong to $\mathbb{A}$. Then, equation (5.71) results in the following discretised equation

$$\left(\frac{F_{e_{i+1,j}}^{n+1} - F_{e_{i,j}}^{n+1}}{\Delta X}\right)\left(\frac{\Phi_{i+1,j}^n - \Phi_{i,j}^n}{\Delta X}\right) + \left(\frac{F_{e_{i,j+1}}^{n+1} - F_{e_{i,j}}^{n+1}}{\Delta Y}\right)\left(\frac{\Phi_{i,j+1}^n - \Phi_{i,j}^n}{\Delta Y}\right) = 0, \quad (5.78)$$

which is the only option available, as $\vec{X}_{i-1,j}$ and $\vec{X}_{i,j-1}$ belong to $\mathbb{T} \cup \mathbb{F}$. The values of $\Phi_{i,j}^n$, $\Phi_{i+1,j}^n$, $\Phi_{i,j+1}^n$, $F_{e_{i+1,j}}^{n+1}$ and $F_{e_{i,j+1}}^{n+1}$. are known, whereas $F_{e_{i,j}}^{n+1}$ is the unknown. Therefore, by rearranging terms and recalling that $\Delta X = \Delta Y$, we obtain

$$F_{e_{i,j}}^{n+1} = \frac{F_{e_{i+1,j}}^{n+1}(\Phi_{i+1,j}^n - \Phi_{i,j}^n) + F_{e_{i,j+1}}^{n+1}(\Phi_{i,j+1}^n - \Phi_{i,j}^n)}{\Phi_{i+1,j}^n - 2\Phi_{i,j}^n + \Phi_{i,j+1}^n}. \quad (5.79)$$

Now, assume that we have more than one grid point accepted in one of the directions. In this case, we simply choose the point with smaller $|\Phi^n|$ value. For instance, let us suppose that $\vec{X}_{i+1,j}$, $\vec{X}_{i-1,j}$ and $\vec{X}_{i,j+1}$ belong to $\mathbb{A}$. Then, we compare $|\Phi_{i+1,j}^n|$ and $|\Phi_{i-1,j}^n|$ so that

- if $|\Phi_{i+1,j}^n| \leq |\Phi_{i-1,j}^n|$, we use discretisation (5.78).
- otherwise, equation (5.71) is discretised as

$$\left(\frac{F_{e_{i,j}}^{n+1} - F_{e_{i-1,j}}^{n+1}}{\Delta X}\right)\left(\frac{\Phi_{i,j}^n - \Phi_{i-1,j}^n}{\Delta X}\right) + \left(\frac{F_{e_{i,j+1}}^{n+1} - F_{e_{i,j}}^{n+1}}{\Delta Y}\right)\left(\frac{\Phi_{i,j+1}^n - \Phi_{i,j}^n}{\Delta Y}\right) = 0, \quad (5.80)$$

Note that, for the case where the four neighbours are in $\mathbb{A}$, we simply apply this argument to each variable separately.

Finally, if we have available data only in one direction, we must discard the derivative in the other direction. For instance, assume that $\vec{X}_{i+1,j}$ belongs to $\mathbb{A}$ but $\vec{X}_{i-1,j}$, $\vec{X}_{i,j+1}$ and $\vec{X}_{i,j-1}$ belong to $\mathbb{T} \cup \mathbb{F}$. In this case, equation (5.71) is discretised as

$$\left(\frac{F^{n+1}_{e_{i+1,j}} - F^{n+1}_{e_{i,j}}}{\Delta X} \right) \left(\frac{\Phi^{n}_{i+1,j} - \Phi^{n}_{i,j}}{\Delta X} \right) = 0, \tag{5.81}$$

thus, $F^{n+1}_{e_{i,j}} = F^{n+1}_{e_{i+1,j}}$

These expressions are used to compute the values of the tentative points and also while updating the tentative values. Once a tentative value is computed or updated, the point must be placed at its right position in the tree, by moving it upwards or downwards, as explained in Section 5.4.2.

5.4.5 The level set equation

With the appropriate velocity values computed at the whole domain, we can start solving the level set equation, (4.32) in the Alpkvist model and (4.48) in the second model. As both equations are similar, so from now on we will refer to (4.32).

At every time step we need to solve

$$\partial_\tau \Phi + F_e ||\nabla \Phi|| = 0 \text{ in } \Omega. \tag{5.82}$$

Remember that the velocity extension comes from the necessity of having a velocity that preserves Φ as a signed distance function in all Ω, which is the computational domain where we solve the level set equation. Note that, we move not only the level set zero but also the rest of the level sets.

Equation (4.32) is fairly simple and straightforward, we simply apply an upwinding scheme for the spatial derivatives determining its direction by using the sign of the velocity and an explicit scheme for the time derivative. The construction of the velocity guarantees that the solution of the level set is reliable, provided that we

discretise taking into account the direction from which the information is travelling. Therefore, by applying an explicit scheme for the time derivative we obtain

$$\frac{\Phi^{n+1}-\Phi^{n}}{\Delta\tau} = -F_e^{n+1}||\nabla\Phi^n||. \tag{5.83}$$

Next, by applying an upwinding strategy in space that takes into account the sign of F, we get

$$\Phi_{i,j}^{n+1} = \Phi_{i,j}^{n} - \Delta\tau F_{e_{i,j}}^{n}\Big[\max(\text{sign}(F_{e_{i,j}}^{n})D_{-X}\Phi_{i,j}^{n}, -\text{sign}(F_{e_{i,j}}^{n})D_{+X}\Phi_{i,j}^{n})^2 + \max(\text{sign}(F_{e_{i,j}}^{n})D_{-Y}\Phi_{i,j}^{n}, -\text{sign}(F_{e_{i,j}}^{n})D_{+Y}\Phi_{i,j}^{n})^2\Big]^{1/2}, \tag{5.84}$$

where

$$D_{-X}\Phi_{i,j}^{n} = \frac{\Phi_{i,j}^{n}-\Phi_{i-1,j}^{n}}{\Delta X}, \quad D_{+X}\Phi_{i,j}^{n} = \frac{\Phi_{i+1,j}^{n}-\Phi_{i,j}^{n}}{\Delta X},$$
$$D_{-Y}\Phi_{i,j}^{n} = \frac{\Phi_{i,j}^{n}-\Phi_{i,j-1}^{n}}{\Delta Y}, \quad D_{+Y}\Phi_{i,j}^{n} = \frac{\Phi_{i,j+1}^{n}-\Phi_{i,j}^{n}}{\Delta Y}.$$

Once equation (5.84) has been solved at each grid point, we obtain the values of Φ^{n+1}. With these values we must define the new solving domain Ω_1^{n+1}, which is

$$\Omega_1^{\tau_{n+1}} = \{(X,Y)\in\Omega \ / \ \Phi^{n+1}(X,Y) < 0\}.$$

Similarly, the domain $\Omega_2^{\tau_{n+1}}$ is defined as

$$\Omega_2^{\tau_{n+1}} = \{(X,Y)\in\Omega \ / \ \Phi^{n+1}(X,Y) \geq 0\},$$

whereas, the new boundary is

$$\Gamma_0^{n+1} = \partial\Omega_1^{\tau_{n+1}} \cap \partial\Omega_2^{\tau_{n+1}}.$$

Finally, the domain $\Omega_{\hat{H}_b}^{\tau_{n+1}}$ is defined as

$$\Omega_{\hat{H}_b}^{\tau_{n+1}} = \{(X,Y)\in\Omega \ / \ \Phi^{n+1}(X,Y) < \hat{H}_b\},$$

whereas $\Gamma_{\hat{H}_b}^{\tau_{n+1}}$ is

$$\Gamma_{\hat{H}_b}^{\tau_{n+1}} = \{(X,Y)\in\Omega \ / \ \Phi^{n+1}(X,Y) = \hat{H}_b\}.$$

5.4.6 Reinitialisation

There are occasions where a well behaved problem comes to a point where the solution becomes unstable, thus causing the level set function to stop from being a signed distance function, or even leading to results far from acceptable. It is in these situations when reinitialisation comes in handy: we just need to reset the level set before everything breaks and continue working. Setting the condition for a reinitialisation depends on each problem, although a good starting point is to maintain $||\nabla\Phi||$ not far away from 1. We must proceed with certain flexibility, as being too strict can lead into reinitialising too often.

Even though reinitialisation helps bringing back an stray level set function, it is not a desirable procedure. This is because reinitialising affects the level set and tends to smoothen it, so doing it too often can lead to incorrect solutions too. If possible, not reinitialising at all is for the best. That is another reason to have a well extended velocity function because a well behaved velocity field maintains Φ as a signed distance.

The reinitialisation procedure has already been explained in Section 5.4.2: just apply the Fast Marching Method to equation $||\nabla\Phi|| = 1$, which is the FME (5.56) particularised for $F = 1$.

As a side note, in those problems when the initial available data are the values of the level set function at grid points surrounding the interface, we use these values to solve $||\nabla\Phi^0|| = 1$ as if we were reinitialising in order to obtain the initial values of Φ^0.

5.4.7 A useful tip: modification of Gibou's method

In section 5.2.1 we needed to know the position of the the point x_I in order to compute $\theta = \dfrac{|x_I - x_k|}{\Delta x}$. However, as we only know points of the discretization, if x_I is not one

of those points, then it is unknown. In our problem we have a level set function Φ, which is a signed distance function, where the level set zero is the interface. Therefore,

$$|x_I - x_k| = |\Phi(x_k)|.$$

The level set function measures the distance between x_k and the boundary, which can be used to substitute the value $x_I - x_k$ that may not be possible to obtain. In a 1D scheme, the previous expression is an equality. In higher dimensions it is only an approximation, although a reliable one.

5.5 Active biomass equation

As indicated in Section 5.1, the next step in the numerical strategy is to solve the equation governing the active biomass (equation (4.33) in the Alpkvist model and (4.49) in the second model), in the recently computed domain $\Omega_1^{\tau_{n+1}}$, which is bounded by Γ_L, Γ_R, Γ_B and $\Gamma_0^{\tau_{n+1}}$.

The active biomass equation is an advection equation. The advection process is due to the growth expansion pressure and is the source of biomass diffusion. Note that this is a non-linear equation, so a Newton method is applied. This is the main difference between the numerical strategies used in the active biomass and inactive biomass equations. We also applied Crank-Nicolson for time discretisation and an upwinding technique for the spatial discretisation, with the upwinding direction determined by the sign of the pressure derivatives.

Special care must be taken while solving the active biomass equation. More precisely, note that due to the application of the Crank-Nicolson method we have an explicit part in which both the pressure and active biomass spatial derivatives in the previous step are involved, that is ∇P^n and ∇V_1^n. These derivatives, as well as the term ∇P^{n+1} used in the implicit part, are computed by using a WENO scheme.

5.5.1 WENO method

Essentially non Oscillatory schemes (also known as ENO schemes), are finite differences schemes developed in 1987 by Harten et al. [41]. The development of these schemes arose from the need of finding a high order approximation that was accurate for piecewise continuous functions, in the sense of reducing the magnitude of the oscillation originated in the regions with discontinuities through the refinement of the mesh [48]. In 1994, Liu et al. [58] developed the Weighted Essentially Non Oscillatory method (also known as WENO method), which proposes a convex combination of the different ENO schemes, thus taking into account the smoothness of the function in each subinterval. The idea under the WENO schemes is that they should behave similar to central schemes where the solution is regular, and similar to ENO schemes near the singularities. In order to present the WENO scheme, first we briefly introduce the ENO schemes.

Let f be Lipschitz-continuous in $\Omega = [0, L] \times [0, H]$, a 2-D domain, with piecewise smooth derivatives, the singularities of which are isolated. Let $\{\vec{x}_{i,j} = (x_i, y_j)\}_{i,j=0}^{M}$ be a finite differences mesh in Ω, with $x_0 = y_0 = 0$, $x_M = L$ and $y_M = H$. Let be $\Delta x = x_i - x_{i-1}$ and $\Delta y = y_j - y_{j-1}$ for $i, j = 0, \ldots, M$, the spatial steps with $\Delta x = \Delta y$, constant. We use the notation $f(\vec{x}_{i,j}) \approx f_{i,j}$, and define

$$\Delta_x^+ f_{i,j} = f_{i+1,j} - f_{i,j}, \quad \Delta_x^- f_{i,j} = f_{i,j} - f_{i-1,j},$$
$$\Delta_y^+ f_{i,j} = f_{i,j+1} - f_{i,j}, \quad \Delta_y^- f_{i,j} = f_{i,j} - f_{i,j-1}.$$

The value $f_x(\vec{x}_{i,j})$ can be approximated through an ENO scheme by using the set of left nodes, $\{\vec{x}_{k,j}\}_{k=i-3}^{i+2}$ for a fixed index j, and applying one of the following approximations

$$f_{x,i,j}^{-,0} = \frac{1}{3}\frac{\Delta_x^+ f_{i-3,j}}{\Delta x} - \frac{7}{6}\frac{\Delta_x^+ f_{i-2,j}}{\Delta x} + \frac{11}{6}\frac{\Delta_x^+ f_{i-1,j}}{\Delta x}, \tag{5.85}$$

$$f_{x,i,j}^{-,1} = -\frac{1}{6}\frac{\Delta_x^+ f_{i-2,j}}{\Delta x} + \frac{5}{6}\frac{\Delta_x^+ f_{i-1,j}}{\Delta x} + \frac{1}{3}\frac{\Delta_x^+ f_{i,j}}{\Delta x}, \tag{5.86}$$

$$f_{x,i,j}^{-,2} = \frac{1}{3}\frac{\Delta_x^+ f_{i-1,j}}{\Delta x} + \frac{5}{6}\frac{\Delta_x^+ f_{i,j}}{\Delta x} - \frac{1}{6}\frac{\Delta_x^+ f_{i+1,j}}{\Delta x}. \tag{5.87}$$

Each option corresponds to a subset of 4 points of the set $\{\vec{x}_{k,j}\}_{k=i-3}^{i+2}$. More precisely, $f_{x,i,j}^{-,s}$ is associated to the subset $N_{s,j} = \{\vec{x}_{k,j}\}_{k=i+s-3}^{i+s}$, with $s = 0, 1, 2$. The choice of which estimation should be taken is based on the relative smoothness of the function in each subset.

As previously mentioned, WENO method takes a convex combination of the three approximations:

$$f_{x,i,j}^{-} = \sum_{s=0}^{2} \omega_{s,j} f_{x,i,j}^{-,s},$$

where the coefficients $\omega_{s,j}$ are the weights associated to the particular subsets $N_{s,j}$.

For a fixed index i, the value $f_y(\vec{x}_{i,j})$ can be approximated in an analogous way by applying the previously explained techniques to the set of left nodes $\{\vec{x}_{i,k}\}_{k=j-3}^{j+2}$.

Analogously, for a fixed j, the value $f_x(\vec{x}_{i,j})$ can be approximated through the WENO method by using the set of right nodes $\{\vec{x}_{k,j}\}_{k=i-2}^{i+3}$.

The choice of the set to be used depends on the particular problem. In our case, when using a WENO approximation to compute ∇P^n, and ∇P^{n+1} in the biomasses equations, we choose the left-biased WENO scheme. However, near the boundaries, when there is not enough data to apply the left-biased WENO scheme, we use the right-biased one, provided that there are enough available data to use it. If neither the left- or right-biased WENO can be applied, the derivatives are approximated following the extrapolation-interpolation technique we explain in the following section.

For the calculation of the gradient of the active biomass, ∇V_1^n, (or the inactive biomass in the case of equations (4.34) and (4.50)), the selection on which set is to be used for the WENO method is made attending to the sign of ∇P^n: if $P_x^n > 0$, then we use a left-biased WENO to compute $V_{1_x}^n$; otherwise, we use a right-biased

WENO. An analogous argument is used for $V_{1_y}^n$.

5.5.2 Pressure and biomass derivatives near the boundaries

As explained in the previous section, WENO scheme uses the value of the function at certain set of nodes to compute the derivatives. However, the values of the pressure and active biomass are not available outside the domain $\Omega_1^{\tau_{n+1}}$. Therefore, when computing the values of the derivatives in the vicinity of the boundaries, we must use other strategies.

At Γ_L and Γ_R we can use the periodic boundary conditions, wrapping the domain, to get the values needed for the WENO method. At Γ_B and $\Gamma_0^{\tau_{n+1}}$, instead of using a WENO scheme, we apply an alternative strategy that involves the extrapolation technique described in Section 5.4.4, as well as the bicubic interpolation described in Section 5.4.2. For the sake of the explanation, we will restrict ourselves to the case of the derivatives of the pressure, although the strategy is the same for the derivatives of the active biomass. Therefore, suppose that we want to compute the gradient of the pressure, ∇P^n, at the grid point $\vec{X}_{i,j}$, and that we are close to Γ_B or $\Gamma_0^{\tau_{n+1}}$. We proceed as follows:

1. We extrapolate the values of the pressure by using the extrapolation technique described in Section 5.4.4, thus obtaining a set of accepted values at the domain $\mathcal{D} = \{X_{i-1}, X_i, X_{i+1}, X_{i+2}\} \times \{Y_{j-1}, Y_j, Y_{j+1}, Y_{j+2}\}$.

2. With those values, we compute the bicubic polynomial, $p^{i,j}$, that approximates the value of the pressure as described in Section 5.4.2.

3. We calculate the spatial derivatives of $p^{i,j}$, namely $p_X^{i,j}$ and $p_Y^{i,j}$, and evaluate them at the point (X_i, Y_j), thus obtaining the values $p_X^{i,j}(X_i, Y_j)$ and $p_Y^{i,j}(X_i, Y_j)$.

4. Finally, we set $P_{X_{i,j}}^n = p_X^{i,j}(X_i, Y_j)$ and $P_{Y_{i,j}}^n = p_Y^{i,j}(X_i, Y_j)$.

5.5.3 Numerical discretisation of the active biomass equation in the Alpkvist model

At each iteration we must advance one time step in the active biomass equation. More precisely, at step $n+1$, corresponding to time $\tau_{n+1} = \Delta\tau(n+1)$, we need to compute the values $V_{1_{i,j}}^{n+1}$, the approximations of the active biomass at the mesh nodes. For this purpose, first we apply a Crank-Nicolson scheme for time discretisation, so that equation (4.33) is approximated by

$$\begin{aligned}\frac{V_1^{n+1} - V_1^n}{\Delta\tau} = \\ \frac{1}{2}\left[\nabla P^{n+1}\cdot\nabla V_1^{n+1} + V_1^{n+1}\left(\frac{\Psi S^{n+1}}{K+S^{n+1}} - (\epsilon_1+\epsilon_2) - V_1^{n+1}(\frac{\Psi S^{n+1}}{K+S^{n+1}} - \epsilon_1)\right)\right] + \\ \frac{1}{2}\left[\nabla P^{n}\cdot\nabla V_1^{n} + V_1^{n}\left(\frac{\Psi S^{n}}{K+S^{n}} - (\epsilon_1+\epsilon_2) - V_1^{n}(\frac{\Psi S^{n}}{K+S^{n}} - \epsilon_1)\right)\right], \text{ in } \Omega_1^{\tau_{n+1}}.\end{aligned} \tag{5.88}$$

In order to establish a non-linear system involving the value of the biomass at all mesh nodes, we extend the value $V_1 \equiv 0$ on the boundary $\Gamma_0^{\tau_{n+1}}$ to the set $\Omega_2^{\tau_{n+1}}$. Therefore, we add the following equation

$$V_1^{n+1} = 0, \text{ in } \Omega_2^{\tau_{n+1}}. \tag{5.89}$$

As mentioned in Section 5.1, the active biomass equation is the next-to-last to be solved at each iteration. Therefore $P^n, P^{n+1}, S^n, S^{n+1}$ and V_1^n are known at the mesh points. Thus, in order to compute $\nabla P^{n+1}, \nabla P^n$ and ∇V_1^n we use the methods explained in Sections 5.5.1 and 5.5.2. Finally, in order to discretise ∇V_1^{n+1} we use an upwind strategy. Therefore, the discretisation of equation (5.88) results in:

$$\begin{aligned}V_{1_{i,j}}^{n+1} - \frac{1}{2}\Delta\tau\Bigg[&\max(P_X^{n+1},0)\frac{V_{1_{i+1,j}}^{n+1} - V_{1_{i,j}}^{n+1}}{\Delta X} + \min(P_X^{n+1},0)\frac{V_{1_{i,j}}^{n+1} - V_{1_{i-1,j}}^{n+1}}{\Delta X} + \\ &\max(P_Y^{n+1},0)\frac{V_{1_{i,j+1}}^{n+1} - V_{1_{i,j}}^{n+1}}{\Delta Y} + \min(P_Y^{n+1},0)\frac{V_{1_{i,j}}^{n+1} - V_{1_{i,j-1}}^{n+1}}{\Delta Y} + \\ &V_{1_{i,j}}^{n+1}\left(\frac{\Psi S_{i,j}^{n+1}}{K+S_{i,j}^{n+1}} - (\epsilon_1+\epsilon_2)\right) - (V_{1_{i,j}}^{n+1})^2\left(\frac{\Psi S_{i,j}^{n+1}}{K+S_{i,j}^{n+1}} - \epsilon_1\right)\Bigg] -\end{aligned}$$

$$\frac{1}{2}\Delta\tau\left[\nabla P^n_{i,j}\cdot\nabla V^n_{1_{i,j}}+V^n_{1_{i,j}}\left(\frac{\Psi S^n_{i,j}}{K+S^n_{i,j}}-(\epsilon_1+\epsilon_2)\right)-(V^n_{1_{i,j}})^2\left(\frac{\Psi S^n_{i,j}}{K+S^n_{i,j}}-\epsilon_1\right)\right]-V^n_{1_{i,j}}=0,\ \text{if}\ \aleph^{n+1}_{i,j}=1, \tag{5.90}$$

whereas, equation (5.89) results in

$$V^{n+1}_{1_{i,j}}=0,\ \text{if}\ \aleph^{n+1}_{i,j}=0, \tag{5.91}$$

for $i=0,\ldots,M-1$, $j=0,\ldots,M$. For the purpose of clarity, let us introduce the following notation

$$P^+_{X_{i,j}}=\max(P^{n+1}_X,0),\quad P^+_{Y_{i,j}}=\max(P^{n+1}_Y,0), \tag{5.92}$$

$$P^-_{X_{i,j}}=\min(P^{n+1}_X,0),\quad P^-_{Y_{i,j}}=\min(P^{n+1}_Y,0), \tag{5.93}$$

$$\alpha_1=\frac{\Delta\tau}{2\Delta X}, \tag{5.94}$$

$$\theta^k_{i,j}\equiv\theta(S^k_{i,j})=\frac{\Psi S^k_{i,j}}{K+S^k_{i,j}}-\epsilon_1,\ \text{with}\ k=n,n+1, \tag{5.95}$$

$$\sigma^k_{i,j}\equiv\sigma(S^k_{i,j})=\frac{\Psi S^k_{i,j}}{K+S^k_{i,j}}-\epsilon_1-\epsilon_2,\ \text{with}\ k=n,n+1, \tag{5.96}$$

$$g^n_{i,j}\equiv g(P^n_{i,j},V^n_{1_{i,j}})=\nabla P^n_{i,j}\cdot\nabla V^n_{1_{i,j}}+V^n_{1_{i,j}}\sigma^n_{i,j}-(V^n_{1_{i,j}})^2\theta^n_{i,j}, \tag{5.97}$$

$$\delta^{n+1}_{i,j}=1+\alpha_1P^+_{X_{i,j}}-\alpha_1P^-_{X_{i,j}}+\alpha_1P^+_{Y_{i,j}}-\alpha_1P^-_{Y_{i,j}}-\alpha_1\Delta XV^{n+1}_{1_{i,j}}\sigma^{n+1}_{i,j}. \tag{5.98}$$

Then, the system of equations (5.90) and (5.91) can be written in equivalent form

$$-\alpha_1P^+_{X_{i,j}}V^{n+1}_{1_{i+1,j}}-\alpha_1P^+_{Y_{i,j}}V^{n+1}_{1_{i,j+1}}+\delta^{n+1}_{i,j}V^{n+1}_{1_{i,j}}+\alpha_1\Delta X(V^n_{1_{i,j}})^2\theta^{n+1}_{i,j}+\alpha_1P^-_{X_{i,j}}V^{n+1}_{1_{i-1,j}}+\alpha_1P^-_{Y_{i,j}}V^{n+1}_{1_{i,j-1}}-\alpha_1g^n_{i,j}-V^n_{1_{i,j}}=0,\ \text{if}\ \aleph^{n+1}_{i,j}=1, \tag{5.99}$$

$$V^{n+1}_{1_{i,j}}=0,\ \text{if}\ \aleph^{n+1}_{i,j}=0, \tag{5.100}$$

for $i=0,\ldots,M-1$, $j=0,\ldots,M$.

For the sake of simplicity, we first write the non-linear system corresponding to the case where all nodes belong to $\Omega_1^{\tau^{n+1}}$. Then, at those nodes belonging to $\Omega_2^{\tau^{n+1}}$ condition (5.100) is imposed, thus removing the corresponding linearised equations from the system.

The non-linear system of equations (5.99) is written in compact form as

$$\mathcal{F}(Z^{n+1}) = A(Z^{n+1})Z^{n+1} - B^n = 0, \tag{5.101}$$

where

- Z^{n+1} is the vector of unknowns at iteration n+1, the $M(M+1)$ elements of which are

$$Z^{n+1}(jM+i) = V_{1_{i,j}}^{n+1}, \text{ with } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.102}$$

- B^n is a vector of $M(M+1)$ elements, the values of which are

$$B^n(jM+i) = \alpha_1 g_{i,j}^n + V_{1_{i,j}}^n, \text{ with } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.103}$$

- $A(Z^{n+1})$ is an $M(M+1) \times M(M+1)$ sparse matrix that contains the coefficients of terms $V_{1_{i,j}}^{n+1}$ from equation (5.99).

Just like we did in Section 5.2.2, as it is necessary to solve the non-linear system of equations, we propose to apply a Newton method that involves the solution of a linear system at each Newton iteration. In order to precise the use of the Newton method, matrix $A(Z^{n+1})$ can be decomposed as the sum of two matrices, i.e.,

$$A(Z^{n+1}) = A_c^n + A_{nc}(Z^{n+1}), \tag{5.104}$$

where A_c^n contains the coefficients of the system of equations (5.101) that are independent of V_1 and $A_{nc}(Z^{n+1})$ contains only the non-linear part of the equation. Newton method requires to calculate the Jacobian matrix of the function $\mathcal{F}$ in (5.101), $J_{\mathcal{F}}(\mathbb{Y})$ which takes the form of

$$J_{\mathcal{F}}(\mathbb{Y}) = A_c^n + J_{A_{nc}}(\mathbb{Y}), \tag{5.105}$$

where, $J_{A_{nc}}(\mathbb{Y})$ is the Jacobian matrix related to $A_{nc}(Z^{n+1})$. Matrix A_c^n is sparse and has a certain block structure, whereas $A_{nc}(Z^{n+1})$ and $J_{A_{nc}}(\mathbb{Y})$ are diagonal matrices. More precisely, these matrices are described in next paragraphs. First of all, matrix A_c^n is

$$A_c^n = \left(\begin{array}{c|c|c|c|c|c|c} \beta^1 & \kappa_0^1 & 0 & \dots & & & \\ \hline \kappa_1^2 & \beta_1^0 & \kappa_1^1 & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \kappa_j^2 & \beta_j^0 & \kappa_j^1 & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \kappa_{M-1}^2 & \beta_{M-1}^0 & \kappa_{M-1}^1 \\ \hline & & & \dots & 0 & \kappa_M^2 & \beta^2 \end{array}\right), \tag{5.106}$$

with $M+1$ blocks of size $M \times M$, that can be described as

- κ_j^1 a diagonal matrix, for $j = 0, \dots, M-1$, where

$$\kappa_j^1 = diag(-\alpha_1 P^+_{Y_{i,j}})_{i=0}^{M-1}.$$

- κ_j^2 a diagonal matrix, for $j = 1, \dots, M$, where

$$\kappa_j^2 = diag(\alpha_1 P^-_{Y_{i,j}})_{i=0}^{M-1}.$$

- β_j^0 a sparse matrix, for $j = 1, \dots, M-1$, in the form:

$$\beta_j^0 = \left(\begin{array}{c|c|c|c|c|c|c} \delta_{0,j}^{n+1} & -\alpha_1 P^+_{X_{0,j}} & 0 & \dots & & & \alpha_1 P^-_{X_{0,j}} \\ \hline \alpha_1 P^-_{X_{1,j}} & \delta_{1,j}^{n+1} & -\alpha_1 P^+_{X_{1,j}} & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha_1 P^-_{X_{i,j}} & \delta_{i,j}^{n+1} & -\alpha_1 P^+_{X_{i,j}} & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha_1 P^-_{X_{M-2,j}} & \delta_{M-2,j}^{n+1} & -\alpha_1 P^+_{X_{M-2,j}} \\ \hline -\alpha_1 P^+_{X_{M-1,j}} & & & \dots & 0 & \alpha_1 P^-_{X_{M-1,j}} & \delta_{M-1,j}^{n+1} \end{array}\right). \tag{5.107}$$

- β^1 a sparse matrix in the form

$$\beta^1 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_0^1 & -\alpha_1 P^+_{X_{0,0}} & 0 & \dots & & & \alpha_1 P^-_{X_{0,0}} \\ \hline \alpha_1 P^-_{X_{1,0}} & \beta_1^1 & -\alpha_1 P^+_{X_{1,0}} & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha_1 P^-_{X_{0,i}} & \beta_i^1 & -\alpha_1 P^+_{X_{0,i}} & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha_1 P^-_{X_{M-2,0}} & \beta^1_{M-2} & -\alpha_1 P^+_{X_{M-2,0}} \\ \hline -\alpha_1 P^+_{X_{M-1,0}} & & & \dots & 0 & \alpha_1 P^-_{X_{M-1,0}} & \beta^1_{M-1} \end{array}\right), \tag{5.108}$$

where

$$\beta_i^1 = \delta_{0,i}^{n+1} + \alpha_1 P^-_{Y_{0,i}}, \quad i = 0, \dots, M-1.$$

- β^2 a sparse matrix in the form:

$$\beta^2 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_0^2 & -\alpha_1 P^+_{X_{0,M}} & 0 & \dots & & & \alpha_1 P^-_{X_{0,M}} \\ \hline \alpha_1 P^-_{X_{1,M}} & \beta_1^2 & -\alpha_1 P^+_{X_{1,M}} & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha_1 P^-_{X_{M,i}} & \beta_i^2 & -\alpha_1 P^+_{X_{M,i}} & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha_1 P^-_{X_{M-2,M}} & \beta^2_{M-2} & -\alpha_1 P^+_{X_{M-2,M}} \\ \hline -\alpha_1 P^+_{X_{M-1,M}} & & & \dots & 0 & \alpha_1 P^-_{X_{M-1,M}} & \beta^2_{M-1} \end{array}\right), \tag{5.109}$$

where

$$\beta_i^2 = \delta_{M,i}^{n+1} - \alpha_1 P^+_{Y_{M,i}}, \;\; i = 0, \dots, M-1.$$

Secondly, the diagonal matrix $A_{nc}(Z^{n+1})$ is

$$A_{nc}(Z^{n+1})_{(jM+1)} = \alpha_1 \Delta X V^n_{1_{i,j}} \theta^{n+1}_{i,j}, \tag{5.110}$$

for $i = 0, \ldots, M-1$ and $j = 0, \ldots, M$, whereas the Jacobian of A_{nc}, $J_{A_{nc}}(\mathbb{Y})$, is

$$J_{A_{nc}}(\mathbb{Y})_{(jM+1)} = 2\alpha_1 \Delta X \mathbb{Y}_{i,j} \theta_{i,j}^{n+1}, \tag{5.111}$$

with $i = 0, \ldots, M-1$ and $j = 0, \ldots, M$.

Finally, in order to set the resulting system to be solved with LU factorisation, we impose (5.100) to all nodes in $\Omega_2^{\tau_{n+1}}$.

5.5.4 Numerical discretisation of the active biomass equation in the second model

As in the Alpkvist model, at each time iteration we must advance one time step in the active biomass equation. However, in the second model, the nutrients consumption and active biomass growth mechanisms are changed using the knowledge gained from the experimental validation of the 1-D models. The original Monod-like terms present in (4.9) are substituted with linear terms, as can be observed in (4.43). The decay mechanism is also changed, thus being a function depending on the mean concentration of inactive biomass. In order to decrease the difficulty of the problem, instead of considering the mean inactive biomass of the current iteration, we use the mean inactive biomass from the previous iteration. However, this does not affect to the precision of the numerical results obtained as the biomass evolution dynamics are slow enough.

At time step $n+1$, corresponding to time $\tau_{n+1} = \Delta\tau(n+1)$, we need to compute the values V_1^{n+1}, the approximation of the active biomass at the mesh nodes. For this purpose, by applying a Crank-Nicolson scheme to equation (4.49), we obtain

$$\frac{V_1^{n+1} - V_1^n}{\Delta\tau} =$$
$$\frac{1}{2}\left[\nabla P^{n+1} \cdot \nabla V_1^{n+1} + V_1^{n+1}\left(\hat{\Psi} S^{n+1} - (\mathcal{F}_{\mathcal{D}}(V_2^n) + \epsilon_2) - V_1^{n+1}\left(\hat{\Psi} S^{n+1} - \mathcal{F}_{\mathcal{D}}(V_2^n)\right)\right)\right] +$$

$$+\frac{1}{2}\left[\nabla P^n\cdot\nabla V_1^n+V_1^n\left(\hat{\Psi}S^n-(\mathcal{F}_{\mathcal{D}}(V_2^n)+\epsilon_2)-V_1^n\left(\hat{\Psi}S^n-\mathcal{F}_{\mathcal{D}}(V_2^n)\right)\right)\right],\ \text{in}\ \Omega_1^{\tau_{n+1}},\tag{5.112}$$

where

$$\mathcal{F}_{\mathcal{D}}(V_2^n)=\frac{1}{1+\exp(K_d(D_{min}-mean(V_2^n)))}.$$

In order to establish a linear system involving the value of the biomass at all mesh nodes, we extend the value $V_1\equiv 0$ on the boundary $\Gamma_0^{\tau_{n+1}}$ to the set $\Omega_2^{\tau_{n+1}}$. Therefore, we add the following equation

$$V_1^{n+1}=0,\ \text{in}\ \Omega_2^{\tau_{n+1}}.\tag{5.113}$$

As mentioned in Section 5.1 the active biomass equation is the next-to-last to be solved at every iteration. This fact still holds in the second model and, therefore, $P^n, P^{n+1}, S^n, S^{n+1}$ and V_1^n are all known values. We have also established that we are using V_2^n to compute the value $mean(V_2^n)$, so this is also a known value. In order to compute $\nabla P^{n+1}, \nabla P^n$ and ∇V_1^n we use the methods explained in Sections 5.5.1 and 5.5.2. Finally, in order to discretise ∇V_1^{n+1} we use an upwind strategy, so equation (5.112) results in:

$$\begin{aligned}V_{1_{i,j}}^{n+1}-\frac{1}{2}\Delta\tau\Bigg[&\max(P_X^{n+1},0)\frac{V_{1_{i+1,j}}^{n+1}-V_{1_{i,j}}^{n+1}}{\Delta X}+\min(P_X^{n+1},0)\frac{V_{1_{i,j}}^{n+1}-V_{1_{i-1,j}}^{n+1}}{\Delta X}+\\&\max(P_Y^{n+1},0)\frac{V_{1_{i,j+1}}^{n+1}-V_{1_{i,j}}^{n+1}}{\Delta Y}+\min(P_Y^{n+1},0)\frac{V_{1_{i,j}}^{n+1}-V_{1_{i,j-1}}^{n+1}}{\Delta Y}+\\&V_{1_{i,j}}^{n+1}\left(\hat{\Psi}S_{i,j}^{n+1}-(\mathcal{F}_{\mathcal{D}}(V_2^n)+\epsilon_2)\right)-(V_{1_{i,j}}^{n+1})^2\left(\hat{\Psi}S_{i,j}^{n+1}-\mathcal{F}_{\mathcal{D}}(V_2^n)\right)\Bigg]-\\&\frac{1}{2}\Delta\tau\left[\nabla P_{i,j}^n\cdot\nabla V_{1_{i,j}}^n+V_{1_{i,j}}^n\left(\hat{\Psi}S_{i,j}^n-(\mathcal{F}_{\mathcal{D}}(V_2^n)+\epsilon_2)\right)-\right.\\&\left.(V_{1_{i,j}}^n)^2\left(\hat{\Psi}S_{i,j}^n-\mathcal{F}_{\mathcal{D}}(V_2^n)\right)\right]-V_{1_{i,j}}^n=0,\ \text{if}\ \aleph_{i,j}^{n+1}=1,\end{aligned}\tag{5.114}$$

whereas, equation (5.113) becomes

$$V_{1_{i,j}}^{n+1}=0,\ \text{if}\ \aleph_{i,j}^{n+1}=0,\tag{5.115}$$

for $i=0,\ldots,M-1$, $j=0,\ldots,M$. For the purpose of clarity, we use the notation

defined in (5.92)-(5.98) with the following modifications to (5.95) and (5.96):

$$\theta_{i,j}^k \equiv \theta(S_{i,j}^k) = \hat{\Psi} S_{i,j}^k - \mathcal{F}_{\mathcal{D}}(V_2^n), \text{ with } k = n, n+1, \tag{5.116}$$

$$\sigma_{i,j}^k \equiv \sigma(S_{i,j}^k) = \hat{\Psi} S_{i,j}^k - \mathcal{F}_{\mathcal{D}}(V_2^n) - \epsilon_2, \text{ with } k = n, n+1, \tag{5.117}$$

Then, the system of equations (5.114) and (5.115) can be written in equivalent form

$$\begin{aligned} -\alpha_1 P_{X_{i,j}}^+ V_{1_{i+1,j}}^{n+1} - \alpha_1 P_{Y_{i,j}}^+ V_{1_{i,j+1}}^{n+1} + \delta_{i,j}^{n+1} V_{1_{i,j}}^{n+1} + \\ \alpha_1 \Delta X (V_{1_{i,j}}^n)^2 \theta_{i,j}^{n+1} + \alpha_1 P_{X_{i,j}}^- V_{1_{i-1,j}}^{n+1} + \alpha_1 P_{Y_{i,j}}^- V_{1_{i,j-1}}^{n+1} - \\ \alpha_1 g_{i,j}^n - V_{1_{i,j}}^n = 0, \text{ if } \aleph_{i,j}^{n+1} = 1, \end{aligned} \tag{5.118}$$

$$V_{1_{i,j}}^{n+1} = 0, \text{ if } \aleph_{i,j}^{n+1} = 0, \tag{5.119}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. It can be noted that the system of equations (5.114) and (5.115) holds a similar structure as the system of equations (5.99) and (5.115). Therefore, from here onward, we follow the same procedure as described for the Alpkvist model, obtaining a non-linear system of equations, which is solved through a Newton method. For further details see Subsection 5.5.3.

5.6 Inactive biomass equation

The last step in the numerical strategy is to calculate the inactive biomass, which is governed by equation (4.34) in the Alpkvist model, and equation (4.50) in the second model. The inactive biomass equation is solved in the domain $\Omega_1^{\tau_{n+1}}$, bounded by Γ_L, Γ_R, Γ_B and $\Gamma_0^{\tau_{n+1}}$.

As in the active biomass, equations (4.34) and (4.50) are also advection equations. However, in this case we do not need to deal with a non-linearity, thus the Newton method is not required. Apart from this particularity, we followed the same strategy as with the active biomass, applying a Crank-Nicolson scheme for time discretisation and an upwinding techniques for the spatial discretisation. The explicit part of the Crank-Nicolson scheme also involves computing the gradient of the pressure and inactive

biomass from the previous step which is achieved by using a WENO scheme in the inner parts of the domain, whereas near the boundaries, the gradient is computed as explained in Section 5.5.2.

5.6.1 Numerical discretisation of the inactive biomass equation in the Alpkvist model

At each iteration we must advance one time step in the inactive biomass equation. More precisely, at step $n+1$, corresponding to time $\tau_{n+1} = \Delta\tau(n+1)$, we need to compute the values V_2^{n+1}, the approximations of the active biomass at mesh nodes. In order to do so, first we apply a Crank-Nicolson scheme for time discretisation, so that equation (4.34) is approximated by

$$\begin{aligned}\frac{V_2^{n+1} - V_2^n}{\Delta\tau} = \\ \frac{1}{2}\left[\nabla P^{n+1}\cdot\nabla V_2^{n+1} + V_1^{n+1}\epsilon_2 - V_2^{n+1}\left(V_1^{n+1}\frac{\Psi S^{n+1}}{K+S^{n+1}} - V_1^{n+1}\epsilon_1\right)\right] + \\ \frac{1}{2}\left[\nabla P^{n}\cdot\nabla V_2^{n} + V_1^{n}\epsilon_2 - V_2^{n}\left(V_1^{n}\frac{\Psi S^{n}}{K+S^{n}} - V_1^{n}\epsilon_1\right)\right], \text{ in } \Omega_1^{\tau_{n+1}}. \end{aligned} \tag{5.120}$$

In order to establish a linear system involving the value of the inactive biomass at all mesh nodes, we extend the value $V_2 \equiv 0$ on the boundary $\Gamma_0^{\tau_{n+1}}$ to the set $\Omega_2^{\tau_{n+1}}$. Therefore, we add the following equation

$$V_2^{n+1} = 0, \text{ in } \Omega_2^{\tau_{n+1}}. \tag{5.121}$$

As explained in Section 5.1, inactive biomass equation is the last to be solved at every iteration. Therefore, $P^n, P^{n+1}, V_1^n, V_1^{n+1}, S^n, S^{n+1}$ and V_2^n are all known values. Thus, in order to compute $\nabla P^{n+1}, \nabla P^n$ and ∇V_2^n we use the methods explained in Sections 5.5.1 and 5.5.2. Finally, in order to discretise ∇V_2^{n+1} we use an upwind strategy.

Therefore, equation (5.120) results in:

$$V_{2_{i,j}}^{n+1} - \frac{1}{2}\Delta\tau \left[\max(P_X^{n+1}, 0) \frac{V_{2_{i+1,j}}^{n+1} - V_{2_{i,j}}^{n+1}}{\Delta X} + \right.$$
$$\min(P_X^{n+1}, 0) \frac{V_{2_{i,j}}^{n+1} - V_{2_{i-1,j}}^{n+1}}{\Delta X} + \max(P_Y^{n+1}, 0) \frac{V_{2_{i,j+1}}^{n+1} - V_{2_{i,j}}^{n+1}}{\Delta Y} +$$
$$\left. \min(P_Y^{n+1}, 0) \frac{V_{2_{i,j}}^{n+1} - V_{2_{i,j-1}}^{n+1}}{\Delta Y} + V_{1_{i,j}}^{n+1}\epsilon_2 - V_{2_{i,j}}^{n+1} V_{1_{i,j}}^{n+1} \left(\frac{\Psi S_{i,j}^{n+1}}{K + S_{i,j}^{n+1}} - \epsilon_1 \right) \right] -$$
$$\frac{1}{2}\Delta\tau \left[\nabla P_{i,j}^n \cdot \nabla V_{2_{i,j}}^n + V_{1_{i,j}}^n \epsilon_2 - V_{2_{i,j}}^n V_{1_{i,j}}^n \left(\frac{\Psi S_{i,j}^n}{K + S_{i,j}^n} - \epsilon_1 \right) \right] - V_{2_{i,j}}^n = 0, \text{ if } \aleph_{i,j}^{n+1} = 1, \tag{5.122}$$

whereas, equation (5.121) becomes

$$V_{2_{i,j}}^{n+1} = 0, \text{ if } \aleph_{i,j}^{n+1} = 0, \tag{5.123}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. For the purpose of clarity, we use the notation defined in (5.92)-(5.98) with the following modifications to (5.97) and (5.98):

$$g_{i,j}^n \equiv g(P_{i,j}^n, V_{1_{i,j}}^n, V_{2_{i,j}}^n) = \nabla P_{i,j}^n \cdot \nabla V_{2_{i,j}}^n + V_{1_{i,j}}^n \epsilon_2 - V_{2_{i,j}}^n V_{1_{i,j}}^n \theta_{i,j}^n, \tag{5.124}$$
$$\delta_{i,j}^{n+1} = 1 + \alpha_1 P_{X_{i,j}}^+ - \alpha_1 P_{X_{i,j}}^- + \alpha_1 P_{Y_{i,j}}^+ - \alpha_1 P_{Y_{i,j}}^- + \alpha_1 \Delta X V_{1_{i,j}}^{n+1} \theta_{i,j}^{n+1}. \tag{5.125}$$

Then, the system of equations (5.122) and (5.123) can be written in equivalent form

$$-\alpha_1 P_{X_{i,j}}^+ V_{2_{i+1,j}}^{n+1} - \alpha_1 P_{Y_{i,j}}^+ V_{2_{i,j+1}}^{n+1} + \delta_{i,j}^{n+1} V_{2_{i,j}}^{n+1} +$$
$$\alpha_1 P_{X_{i,j}}^- V_{2_{i-1,j}}^{n+1} + \alpha_1 P_{Y_{i,j}}^- V_{2_{i,j-1}}^{n+1} - \alpha_1 g_{i,j}^n - \alpha_1 \Delta X V_{1_{i,j}}^{n+1} \epsilon_2 - V_{2_{i,j}}^n = 0, \text{ if } \aleph_{i,j}^{n+1} = 1, \tag{5.126}$$
$$V_{2_{i,j}}^{n+1} = 0, \text{ if } \aleph_{i,j}^{n+1} = 0, \tag{5.127}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$.

For the sake of simplicity, we first write the linear system corresponding to the case where all nodes belong to $\Omega_1^{\tau_{n+1}}$. Then, at those nodes belonging $\Omega_2^{\tau_{n+1}}$ condition (5.127) is imposed, thus removing the corresponding linear equations from the system.

The linear system (5.126) is written in compact form as

$$A^n Z^{n+1} = B^n, \tag{5.128}$$

where

- Z^{n+1} is the vector of unknowns at iteration n+1, the $M(M+1)$ elements of which are

$$Z^{n+1}(jM+i) = V_{2_{i,j}}^{n+1}, \text{ for } i = 0, \ldots, M-1 \text{ and } j = 0, \ldots, M. \tag{5.129}$$

- B^n is the independent term of the linear system, a vector of $M(M+1)$ elements, the values of which are

$$B(jM+i) = \alpha_1 g_{i,j}^n + \alpha_1 V_{1_{i,j}}^{n+1} \epsilon_2 + V_{2_{i,j}}^n, \tag{5.130}$$

  with $i = 0, \ldots, M-1$ and $j = 0, \ldots, M$.

- A^n is an $M(M+1) \times M(M+1)$ sparse matrix with certain block structure. It contains the coefficients of terms $V_{2_{i,j}}^{n+1}$ from equation (5.126). Described block by block, matrix A^n takes the form

$$A^n = \left(\begin{array}{c|c|c|c|c|c|c} \beta^1 & \kappa_0^1 & 0 & \ldots & & & \\ \hline \kappa_1^2 & \beta_1^0 & \kappa_1^1 & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \kappa_j^2 & \beta_j^0 & \kappa_j^1 & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \kappa_{M-1}^2 & \beta_{M-1}^0 & \kappa_{M-1}^1 \\ \hline & & & \ldots & 0 & \kappa_M^2 & \beta^2 \end{array}\right), \tag{5.131}$$

  There are $M+1$ blocks, each one of size $M \times M$, that can be described as

  - κ_j^1 a diagonal matrix, for $j = 0, \ldots, M-1$, where

$$\kappa_j^1 = diag(-\alpha_1 P_{Y_{i,j}}^+)_{i=0}^{M-1}.$$

– κ_j^2 a diagonal matrix, for $j = 1, \ldots, M$, where

$$\kappa_j^2 = diag(\alpha_1 P^-_{Y_{i,j}})_{i=0}^{M-1}.$$

– β_j^0 a sparse matrix, for $j = 1, \ldots, M-1$, taking the form:

$$\beta_j^0 = \left(\begin{array}{c|c|c|c|c|c|c} \delta_{0,j}^{n+1} & -\alpha_1 P^+_{X_{0,j}} & 0 & \ldots & & & \alpha_1 P^-_{X_{0,j}} \\ \hline \alpha_1 P^-_{X_{1,j}} & \delta_{1,j}^{n+1} & -\alpha_1 P^+_{X_{1,j}} & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \alpha_1 P^-_{X_{i,j}} & \delta_{i,j}^{n+1} & -\alpha_1 P^+_{X_{i,j}} & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \alpha_1 P^-_{X_{M-2,j}} & \delta_{M-2,j}^{n+1} & -\alpha_1 P^+_{X_{M-2,j}} \\ \hline -\alpha_1 P^+_{X_{M-1,j}} & & & \ldots & 0 & \alpha_1 P^-_{X_{M-1,j}} & \delta_{M-1,j}^{n+1} \end{array}\right), \tag{5.132}$$

– β^1 a sparse matrix that takes the form:

$$\beta^1 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_0^1 & -\alpha_1 P^+_{X_{0,0}} & 0 & \ldots & & & \alpha_1 P^-_{X_{0,0}} \\ \hline \alpha_1 P^-_{X_{1,0}} & \beta_1^1 & -\alpha_1 P^+_{X_{1,0}} & 0 & \ldots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \ldots & \alpha_1 P^-_{X_{0,i}} & \beta_i^1 & -\alpha_1 P^+_{X_{0,i}} & \ldots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \ldots & 0 & \alpha_1 P^-_{X_{M-2,0}} & \beta_{M-2}^1 & -\alpha_1 P^+_{X_{M-2,0}} \\ \hline -\alpha_1 P^+_{X_{M-1,0}} & & & \ldots & 0 & \alpha_1 P^-_{X_{M-1,0}} & \beta_{M-1}^1 \end{array}\right), \tag{5.133}$$

where

$$\beta_i^1 = \delta_{0,i}^{n+1} + \alpha_1 P^-_{Y_{0,i}}, \quad i = 0, \ldots, M-1.$$

– β^2 a sparse matrix that takes the form

$$\beta^2 = \left(\begin{array}{c|c|c|c|c|c|c} \beta_0^2 & -\alpha_1 P^+_{X_{0,M}} & 0 & \dots & & & \alpha_1 P^-_{X_{0,M}} \\ \hline \alpha_1 P^-_{X_{1,M}} & \beta_1^2 & -\alpha_1 P^+_{X_{1,M}} & 0 & \dots & & \\ \hline & \ddots & \ddots & \ddots & \ddots & & \\ \hline & \dots & \alpha_1 P^-_{X_{M,i}} & \beta_i^2 & -\alpha_1 P^+_{X_{M,i}} & \dots & \\ \hline & & & \ddots & \ddots & \ddots & \\ \hline & & \dots & 0 & \alpha_1 P^-_{X_{M-2,M}} & \beta^2_{M-2} & -\alpha_1 P^+_{X_{M-2,M}} \\ \hline -\alpha_1 P^+_{X_{M-1,M}} & & & \dots & 0 & \alpha_1 P^-_{X_{M-1,M}} & \beta^2_{M-1} \end{array}\right), \tag{5.134}$$

where

$$\beta_i^2 = \delta_{M,i}^{n+1} - \alpha_1 P^+_{Y_{M,i}}, \quad i = 0, \dots, M-1.$$

Finally, in order to set the resulting system to be solved with LU factorisation, we impose (5.127) to all nodes in $\Omega_2^{\tau_{n+1}}$.

5.6.2 Numerical discretisation of the inactive biomass equation in the second model

As in the Alpkvist model, at each time iteration we must advance one time step in the active biomass equation. The modification of the nutrients consumption and the decay mechanism presented in Section 4.3 also affects the inactive biomass equation. This fact, however, does not alter the numerical strategy followed to solve the inactive biomass equation. Therefore, proceeding as in subsection 5.6.1, at step $n+1$, corresponding to time $\tau^{n+1} = \Delta\tau(n+1)$, we need to compute the values V_2^{n+1}, the approximation of the inactive biomass at the mesh nodes. In order to do so, by

applying a Crank-Nicolson scheme for time discretisation to (4.50), we obtain

$$\frac{V_2^{n+1}-V_2^n}{\Delta\tau} = \frac{1}{2}\left[\nabla P^{n+1}\cdot\nabla V_2^{n+1} + V_1^{n+1}\epsilon_2 - V_2^{n+1}\left(V_1^{n+1}\hat{\Psi}S^{n+1} - V_1^{n+1}\mathcal{F}_{\mathcal{D}}(V_2^n)\right)\right] + \frac{1}{2}\left[\nabla P^{n}\cdot\nabla V_2^{n} + V_1^{n}\epsilon_2 - V_2^{n}\left(V_1^{n}\hat{\Psi}S^{n} - V_1^{n}\mathcal{F}_{\mathcal{D}}(V_2^n)\right)\right], \text{ in } \Omega_1^{\tau_{n+1}}, \quad (5.135)$$

where

$$\mathcal{F}_{\mathcal{D}}(V_2^n) = \frac{1}{1+\exp(K_d(D_{min} - mean(V_2^n)))}.$$

In order to establish a linear system involving the value of the biomass at all mesh nodes, we extend the value $V_2 \equiv 0$ on the boundary $\Gamma_0^{\tau_{n+1}}$ to the set $\Omega_2^{\tau_{n+1}}$. Therefore, we add the following equation

$$V_2^{n+1} = 0, \text{ in } \Omega_2^{\tau_{n+1}}. \quad (5.136)$$

As explained in Section 5.1, the inactive biomass equation is the last to be solved at every iteration. Therefore, $P^n, P^{n+1}, V_1^n, V_1^{n+1}, S^n, S^{n+1}$ and V_2^n are all known values. In order to decrease the difficulty of the problem, instead of considering the mean inactive biomass of the current iteration, we will use the mean inactive biomass from the previous iteration. Therefore $mean(V_2^n)$ is also a known value. For the computation of $\nabla P^{n+1}, \nabla P^n$ and ∇V_2^n we use the methods explained in Sections 5.5.1 and 5.5.2. Finally, in order to discretise ∇V_2^{n+1} we use an upwind strategy. The discretisation of (5.135) results in:

$$\begin{aligned} V_{2_{i,j}}^{n+1} - \frac{1}{2}\Delta\tau\Bigg[&\max(P_X^{n+1},0)\frac{V_{2_{i+1,j}}^{n+1} - V_{2_{i,j}}^{n+1}}{\Delta X} + \\ &\min(P_X^{n+1},0)\frac{V_{2_{i,j}}^{n+1} - V_{2_{i-1,j}}^{n+1}}{\Delta X} + \max(P_Y^{n+1},0)\frac{V_{2_{i,j+1}}^{n+1} - V_{2_{i,j}}^{n+1}}{\Delta Y} + \\ &\min(P_Y^{n+1},0)\frac{V_{2_{i,j}}^{n+1} - V_{2_{i,j-1}}^{n+1}}{\Delta Y} + V_{1_{i,j}}^{n+1}\epsilon_2 - V_{2_{i,j}}^{n+1}V_{1_{i,j}}^{n+1}\left(\hat{\Psi}S_{i,j}^{n+1} - \mathcal{F}_{\mathcal{D}}(V_2^n)\right)\Bigg] - \\ &\frac{1}{2}\Delta\tau\left[\nabla P_{i,j}^n\cdot\nabla V_{2_{i,j}}^n + V_{1_{i,j}}^n\epsilon_2 - V_{2_{i,j}}^n V_{1_{i,j}}^n\left(\hat{\Psi}S_{i,j}^n - \mathcal{F}_{\mathcal{D}}(V_2^n)\right)\right] - V_{2_{i,j}}^n = 0, \text{ if } \aleph_{i,j}^{n+1} = 1, \end{aligned} \quad (5.137)$$

whereas, equation (5.136) becomes

$$V_{2_{i,j}}^{n+1} = 0, \text{ if } \aleph_{i,j}^{n+1} = 0. \tag{5.138}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. For the purpose of clarity, we use the notation defined in (5.92)-(5.94), (5.124) and (5.125) with the following modification to (5.95):

$$\theta_{i,j}^{k} \equiv \theta(S_{i,j}^{k}) = \hat{\Psi} S_{i,j}^{k} - \mathcal{F}_{\mathcal{D}}(V_2^n), \text{ with } k = n, n+1. \tag{5.139}$$

Then, the system of equations (5.137) and (5.138) can be written in equivalent form:

$$\begin{aligned} -\alpha_1 P_{X_{i,j}}^{+} V_{2_{i+1,j}}^{n+1} - \alpha_1 P_{Y_{i,j}}^{+} V_{2_{i,j+1}}^{n+1} + \delta_{i,j}^{n+1} V_{2_{i,j}}^{n+1} + \\ \alpha_1 P_{X_{i,j}}^{-} V_{2_{i-1,j}}^{n+1} + \alpha_1 P_{Y_{i,j}}^{-} V_{2_{i,j-1}}^{n+1} - \alpha_1 g_{i,j}^{n} - \alpha_1 \Delta X V_{1_{i,j}}^{n+1} \epsilon_2 - V_{2_{i,j}}^{n} = 0, \text{ if } \aleph_{i,j}^{n+1} = 1 \quad (5.140) \\ V_{2_{i,j}}^{n+1} = 0, \text{ if } \aleph_{i,j}^{n+1} = 0. \quad (5.141) \end{aligned}$$

for $i = 0, \ldots, M-1$, $j = 0, \ldots, M$. It can be noted that equation (5.140) holds a similar structure as equation (5.126). Therefore, from here onward, we follow the same procedure as described for the Alpkvist model, thus obtaining a linear system of equations. For further details see subsection 5.6.1.

Chapter 6

Biological validation of the 2D models

This chapter focuses on the validation of the 2D models in the qualitative description of bacterial biofilms. First, the developed numerical methods will be tested using the Alpkvist model to predict mushroom-shaped biofilms. Later the second model will be used to describe various biofilms formed by *L. monocytogenes* strains.

6.1 Validation of the numerical methods

Many bacterial species, such as the largely studied *Pseudomonas aeruginosa*, form mushroom-shaped structures. Biofilms are composed of mushroom-shaped microcolonies, in which cells are embedded in the exopolymeric matrix (EPS), and channels separate the microcolonies.

The beauty and complexity of the mushroom-shaped biofilms led to the development of sophisticated mathematical models and numerical approaches intended to find the biological mechanisms underlying their formation. In fact, the case of mushroom-shaped biofilms is typically used as a benchmark for the validation of biofilm models (see, for example, [1, 2, 25, 27, 76] to name a few).

The model considered here, based on the Alpkvist model (Section 4.2), consists of two species, regarded as active and inert biomass. Active biomass corresponds to viable bacteria while inert biomass corresponds to a mixture of dead or damaged cells plus EPS. Active biomass has a single nutrient at disposal.

Model parameters, presented in Table 6.1, were taken from the literature ([1, 74, 79]). Initial conditions are defined to simulate the presence of two interconnected microcolonies, often called stalks, as typically found at the initial times in mushroom-type bacterial biofilms. Initially, there is no inert biomass. These initial conditions are formulated as follows:

$$\Gamma_0^0 = \{(X,Y) \in \Omega / Y = 0.2 + 0.05\sin(4\pi X)\}, \tag{6.1}$$

$$\Omega_1^0 = \{(X,Y) \in \Omega / Y < 0.2 + 0.05\sin(4\pi X)\}, \tag{6.2}$$

$$V_1(X,Y,0) = 1, \quad V_2(X,Y,0) = 0, \text{ in } \Omega_1^0. \tag{6.3}$$

Parameter	Value	Unit
L, H	3×10^{-4}	m
H_b	0.125H	m
ρ	60	kg/m^3
D	2×10^{-9}	m^2/s
$\mathcal{Y}$	0.1	-
μ	1×10^{-5}	s^{-1}
K_s	1×10^{-5}	kg/m^3
k_d	2×10^{-6}	s^{-1}
k_i	1×10^{-6}	s^{-1}
t_D	86400	s
s^*	4×10^{-3}	kg/m^3

Table 6.1: Parameters values used in the Alpkvist model

Figure 6.1 presents the evolution of the biofilm structure for the given conditions and parameters. Results illustrate the formation of mushrooms. Mushrooms grow

in height with time, while the presence of damaged cells increases. Inert biomass appears in the deeper parts of the mushrooms. Mushroom-type structures have been linked to sequential processes involving motile and non-motile bacteria or to a limited substrate regime (see for instance, [35, 52, 81]).

The use of the level set method guarantees the robustness and good behaviour of the numerical resolution. Figure 6.2 shows the evolution of the level sets in the Alpkvist model. It can be seen how the extended velocity keeps the function as a signed distance at every moment. The correct resolution of the level set equation is fundamental to obtain the appropriate solving domains at each time step which, in turn, gives rise to the correct resolution of the problem.

Finally, Figure 6.3 shows the evolution of nutrients and pressure. The growth of the expansive pressure narrows the biofilm in the inner layer, whereas widens the biofilm in the most external parts. The lack of nutrients in the inner parts, due to their fully consumption, contributes to the inactivation of the cells, while at the top layers the nutrients availability allows the growth of the biofilm.

6.2 Validation of the second 2D model

Several recent works considered the spatial characterisation of *L. monocytogenes* biofilms at particular given times using microscopy and image analysis. In their work, Rieu et al. [82] presented a comparison of biovolume and mean thickness for *L. monocytogenes* biofilms under batch and flow conditions using CLSM and PHLIP. Bridier et al. [12] quantified maximum thickness, biovolume and roughness using PHLIP for biofilms formed by 10 *L. monocytogenes* strains at 24 h. Similarly Guilbaud et al. [37] studied the structural diversity of biofilms formed by 96 *L. monocytogenes* strains at 48 h using also mean thickness, biovolume and roughness. Mosquera et al. [66] considered the dynamics of the structure of different strains

throughout their life cycle.

These works will be the basis for the proposed model qualitative model validation.

6.2.1 Flat *L. monocytogenes* biofilms

As already discussed in previous chapters, visual inspection of the biofilms formed by the *L. monocytogenes* strain L1A1 (see Figure 3.1) reveals thick almost flat biofilms. The images at earlier times show the presence of small clusters, which eventually joined to each other in a honeycomb pattern. Later this pattern disappears to form an unstructured flat mass.

L1A1 is not the only strain in forming flat structures. Guilbaud et al. [37] observed that *L. monocytogenes* strains without flagellum developed flat structures at 48 h. Some examples include CIP 82110 (avirulent type strain from rabbit), LM 6298 (from soil) or H6 (from a healthy 30-year-old woman) strains. Pilchová et al. [78] studied structures at 1, 3 and 6 days and suggested that the EGD-e strain isolated from listeriosis outbreaks could skip the honeycomb-like pattern during the life cycle, evolving to flat structures.

Model parameters are adapted to the *L. monocytogenes* species taking into account model calibration results obtained in Chapter 3. Parameter values are presented in Table 6.2.

As a first approximation we solve the 2D multi-species Alpkvist model by modifying the parameters accordingly to the best set obtained with the 1D M3 model. Figure 6.4 shows the results achieved. Remarkably the model predicts a flat biofilm. However, the dynamics are unrealisticaly fast both for nutrients consumption and biofilm growth.

Parameter	Value	Unit
L, H	3×10^{-4}	m
H_b	0.125H	m
ρ	17	kg/m^3
D	2.568×10^{-12}	m^2/s
$\mathcal{Y}$	0.98	-
μ	Case dependent	s^{-1}
K_d	311.5	-
D_{min}	5.21×10^{-2}	-
k_i	Case dependent	s^{-1}
t_D	86400	s
s^*	2.74	kg/m^3

Table 6.2: Parameters values used in the simulations of M2 for *L. monocytogenes* strains

We now consider the modified 2D multi-species model M2. Since the modified Eberl Model 3 led to conclude that a significant concentration of inert biomass was required for L1A1 to form flat biofilms, k_i was selected to be one order of magnitude higher than the reference used in the simulation of the Alpkvist model ($k_i = 10^{-5}s^{-1}$). The growth rate was chosen to be $\mu = 4.25 \times 10^{-5}s^{-1}$ a standard value for *L. monocytogenes* species. As for the initial conditions, two joint microcolonies were considered.

In an initial attempt, the local concentration of inert biomass was used as a reference for detachment. The results obtained reproduced the initial growth stages of the biofilm accurately. However, when the inactive biomass exceeded the threshold for detachment, an instantaneous massive detachment was produced in the deeper layers of the biofilm. This effect resulted not only on unrealistic results but in numerical instabilities. The detachment mechanism was reformulated to consider the mean inert biomass as for the modified Eberl model 3. In this way, the detachment process starts when there is a given mean inert biomass in the biofilm and cells detach in all areas of the biofilm.

Figure 6.5 shows how model 2 rapidly evolves to a dense flat biofilm. The dynamics observed are much slower than the predicted by the 2D multi-species Alpkvist model. Also the maximum height achieved is in good agreement with the observed experimentaly. Inert biomass appears stratified in the biofilm with higher concentrations in deeper layers as typically observed experimentally (Mosquera et al. [66]). At later stages, inert biomass is abundant and can be readily observed on the surface of the biofilm forming low-density layers with scattered higher concentration spots. These profiles can be directly compared to those depicted for L1A1 in Chapter 3 and CIP 82110, LM6298 or H6 in Guilbaud et al. [37].

6.2.2 Clustered and honeycomb *L. monocytogenes* biofilms

Bridier et al. [12] observed that several *L. monocytogenes* strains produced rough biofilms containing several small clusters of variable thickness in static cultures at 24h. Mosquera et al. [66] also reported the presence of clusters at early stages of biofilm formation for the strains CECT 4032 and CECT 5873. At later stages, these clusters may join to each other and form honey-comb patterns with small channels or even voids such as those observed by Guilbaud et al. [37], Picholva et al. [78] or Mosquera et al. [66].

Figure 6.6 presents the CLSM reconstructions at different times for the strain CECT5873. The figures illustrate the presence of multiple microcolonies or small clusters at the initial time. Those clusters evolve to form larger clusters which eventually disappear to form an unstructured flatter biofilm once the amount of inert biomass is substantial.

To test the capacity of the second model to describe both types of structures, values for growth and inactivation rates as well as the initial conditions were modified. Figure 6.7 presents a case in which two joint microcolonies evolve into thicker

clusters which are separated by channels. For this simulation $\mu = 8.5 \times 10^{-3} s^{-1}$ and $k_i = 8.85 \times 10^{-8}$ were used. Remark that growth rate is two orders of magnitude larger than the one used in the previous example while the inactivation rate is significantly lower.

Results reveal that the larger growth rate contributed to forming thick clusters while the lower inactivation rate contributed to slowing down the detachment. As an overall effect, dense clusters are observed in which channels are formed due to the slow superficial detachment. Again inert biomass tends to accumulate inside the clusters, as it can also be seen in the CLSM reconstructions (Figure 6.6).

Figure 6.8 presents the formation of honeycomb structures with hollow voids. The hypothesis is that initially separated microcolonies may eventually join each other while leaving channels or voids in deeper layers. To simulate for this scenario the following initial conditions were selected:

$$\Gamma_0^0 = \Gamma_{0,a}^0 \cup \Gamma_{0,b}^0, \quad (6.4)$$

$$\Gamma_{0,a}^0 = \{(X,Y) \in \Omega / X < 0.5, Y = -0.3565 + 0.6065 * \sin(2\pi X)\}, \quad (6.5)$$

$$\Gamma_{0,b}^0 = \{(X,Y) \in \Omega / X >= 0.5, Y = -0.3565 + 0.6065 * \sin(2\pi (X - 0.5))\}, \quad (6.6)$$

$$\Omega_0^0 = \Omega_{0,a}^0 \cup \Omega_{0,b}^0, \quad (6.7)$$

$$\Omega_{0,a}^0 = \{(X,Y) \in \Omega / X < 0.5, Y < -0.3565 + 0.6065 * \sin(2\pi X)\}, \quad (6.8)$$

$$\Omega_{0,b}^0 = \{(X,Y) \in \Omega / X >= 0.5, Y < -0.3565 + 0.6065 * \sin(2\pi (X - 0.5))\}, \quad (6.9)$$

that would correspond to two separated microcolonies. Again only the growth and the inactivation rates were modified: $\mu = 8.5 \times 10^{-4} s^{-1}$ and $k_i = 10^{-6}$.

Figure 6.8 presents the results showing how the microcolonies rapidly evolve to large clusters which eventually connect to each other while showing void channels in the deeper layers of the biofilms.

All in all, results show that the proposed 2D multi-species model M2 is able to qualitatively predict the interstain variability found in *L. monocytogenes* biofilms. Initial attachment (initial conditions), and detachment rate (which relates to inactive biomass), condition the development of flat, clustered or honeycomb type structures.

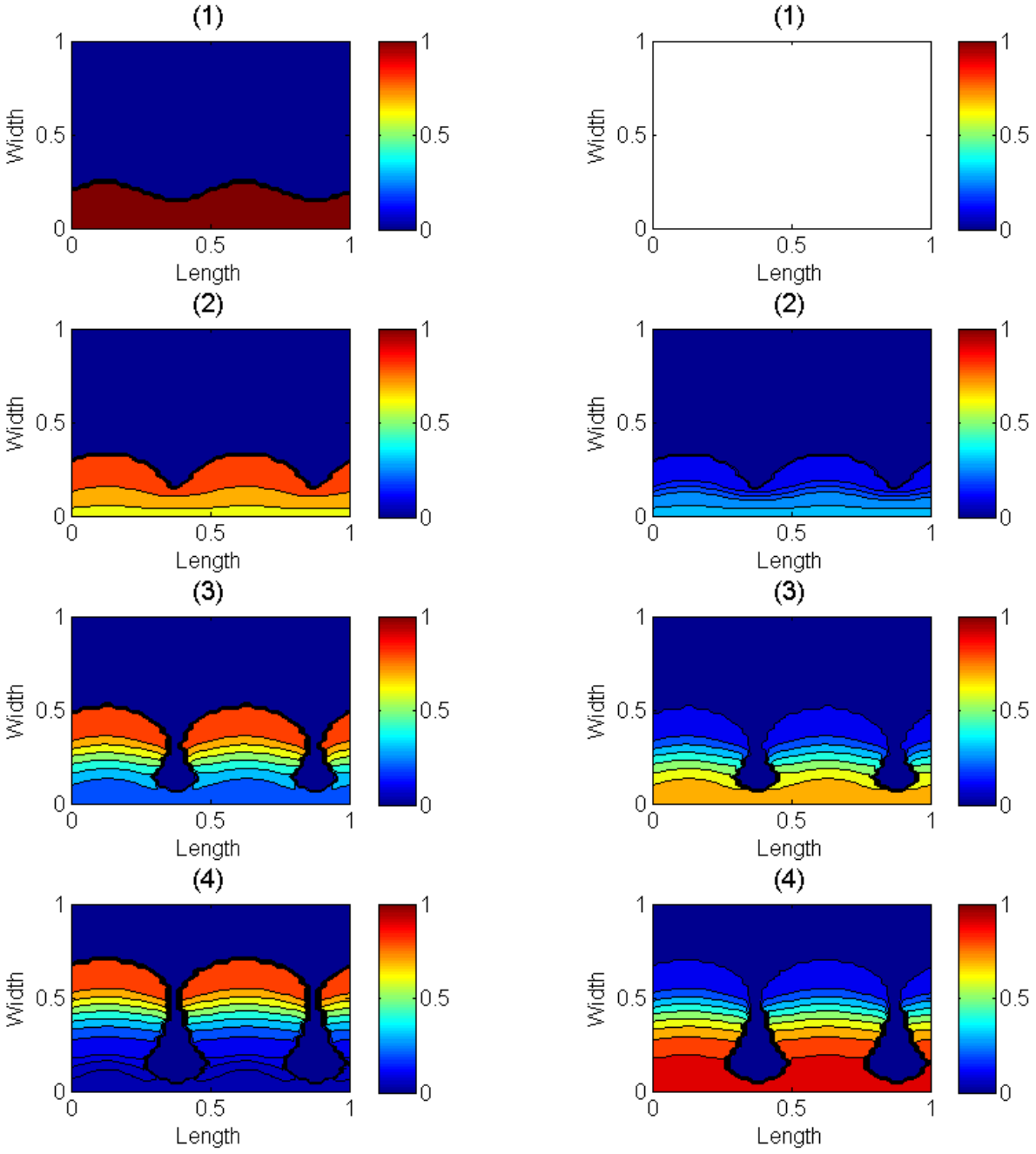

Figure 6.1: **Numerical results obtained for the Alpkvist model.** In the left column the active biomass concentration is shown, whereas the right corresponds to the inert biomass concentration. Figures (1) correspond to the initial conditions and (2), (3) and (4) correspond to the biofilm structure at 3.2 days, 9.6 days and 16 days, respectively.

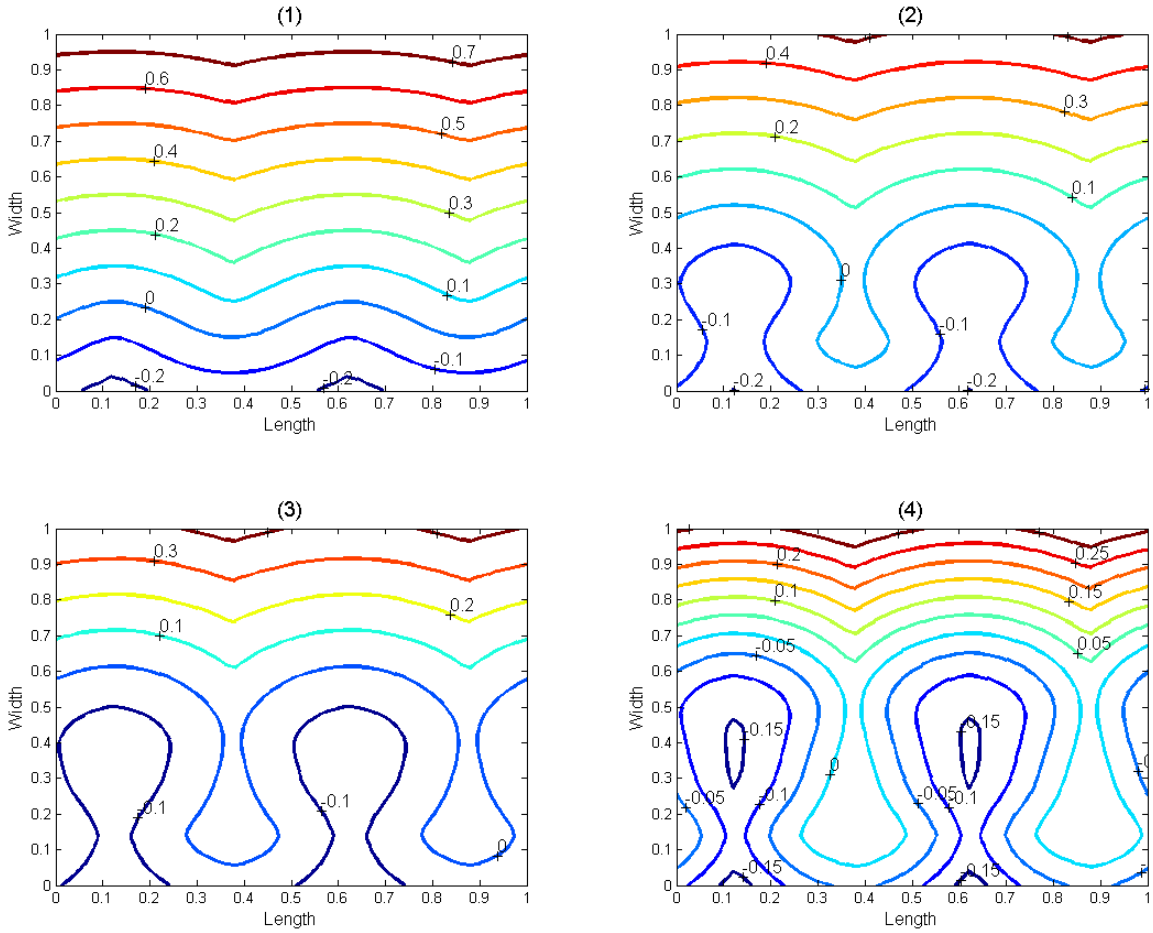

Figure 6.2: **Level set function evolution for the Alpkvist model.** Frame (1) corresponds to the initial conditions and (2), (3) and (4) correspond to the level set function at 3.2, 9.6 days and 16 days, respectively.

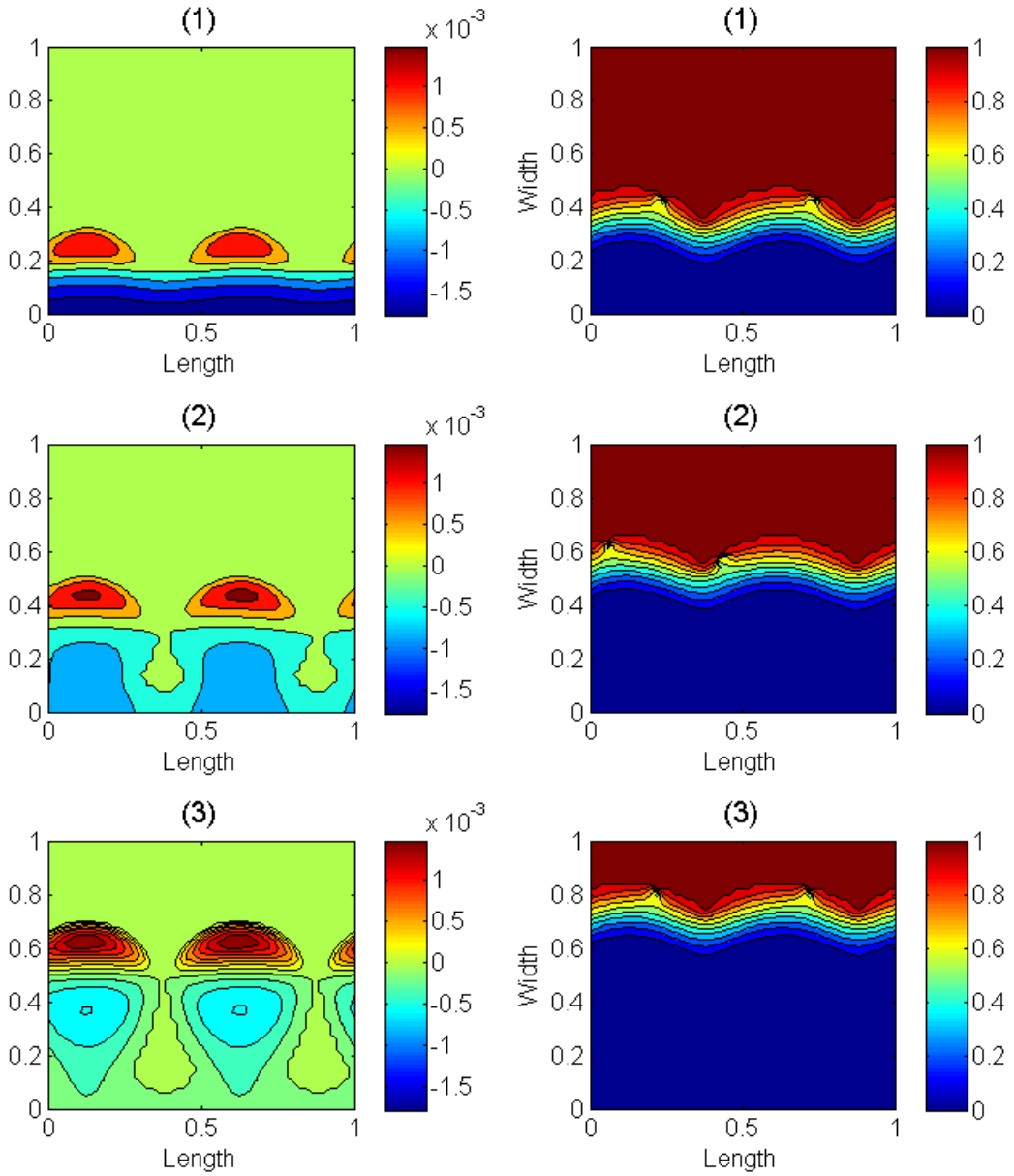

Figure 6.3: **Pressure and nutrients obtained for the Alpkvist model.** In the left column the pressure concentration is shown, whereas the right corresponds to the nutrients concentration. Figures (1), (2) and (3) correspond to the results obtained at 3.2 days, 9.6 days and 16 days, respectively.

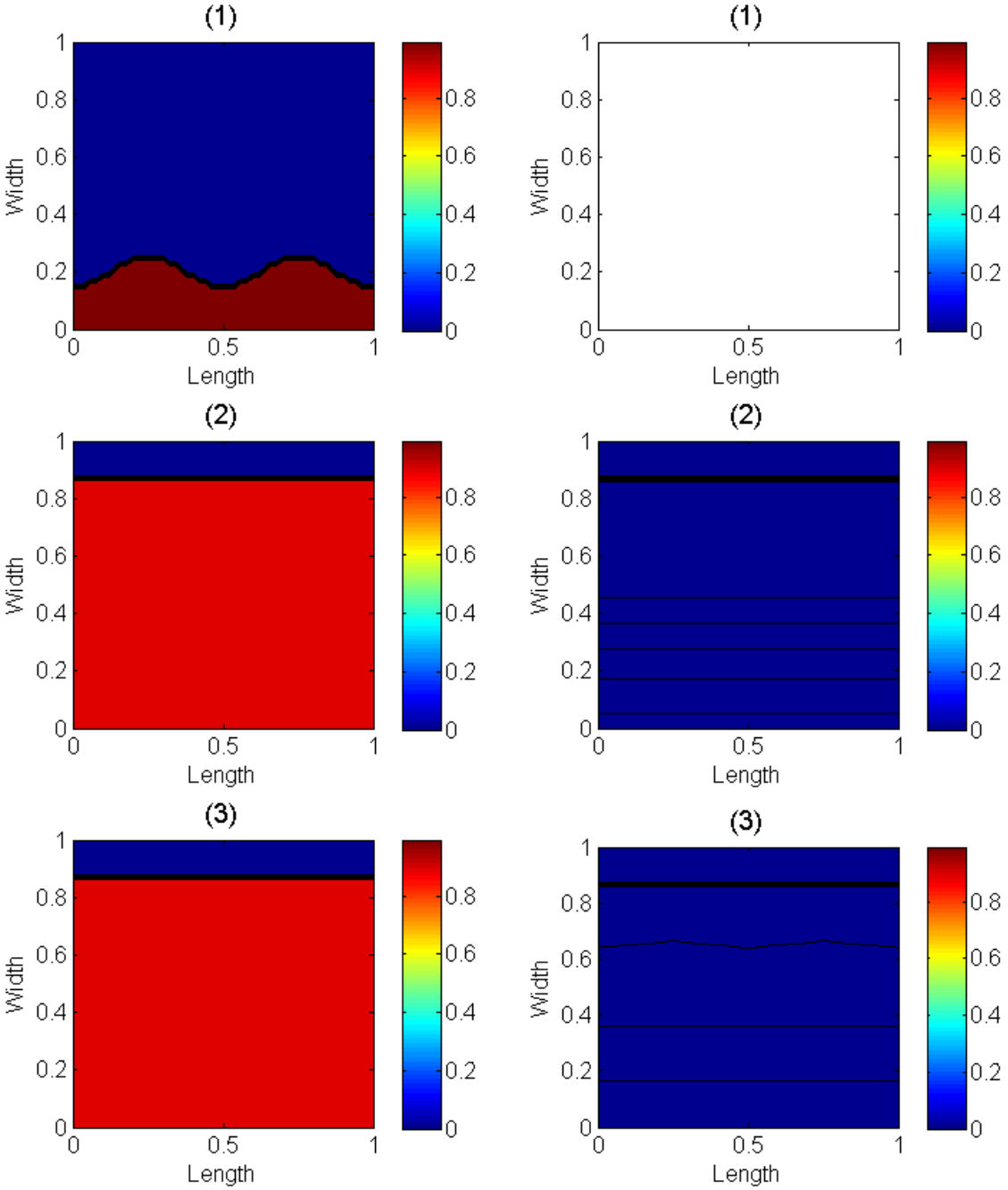

Figure 6.4: **Flat biofilms as predicted by the Alpkvist model.** Active biomass dynamic is shown in the left column, while the right column shows the dynamic of inert biomass. Figures (1) present the initial conditions, (2) and (3) present biomass at 5.76h and 9.6h, respectively.

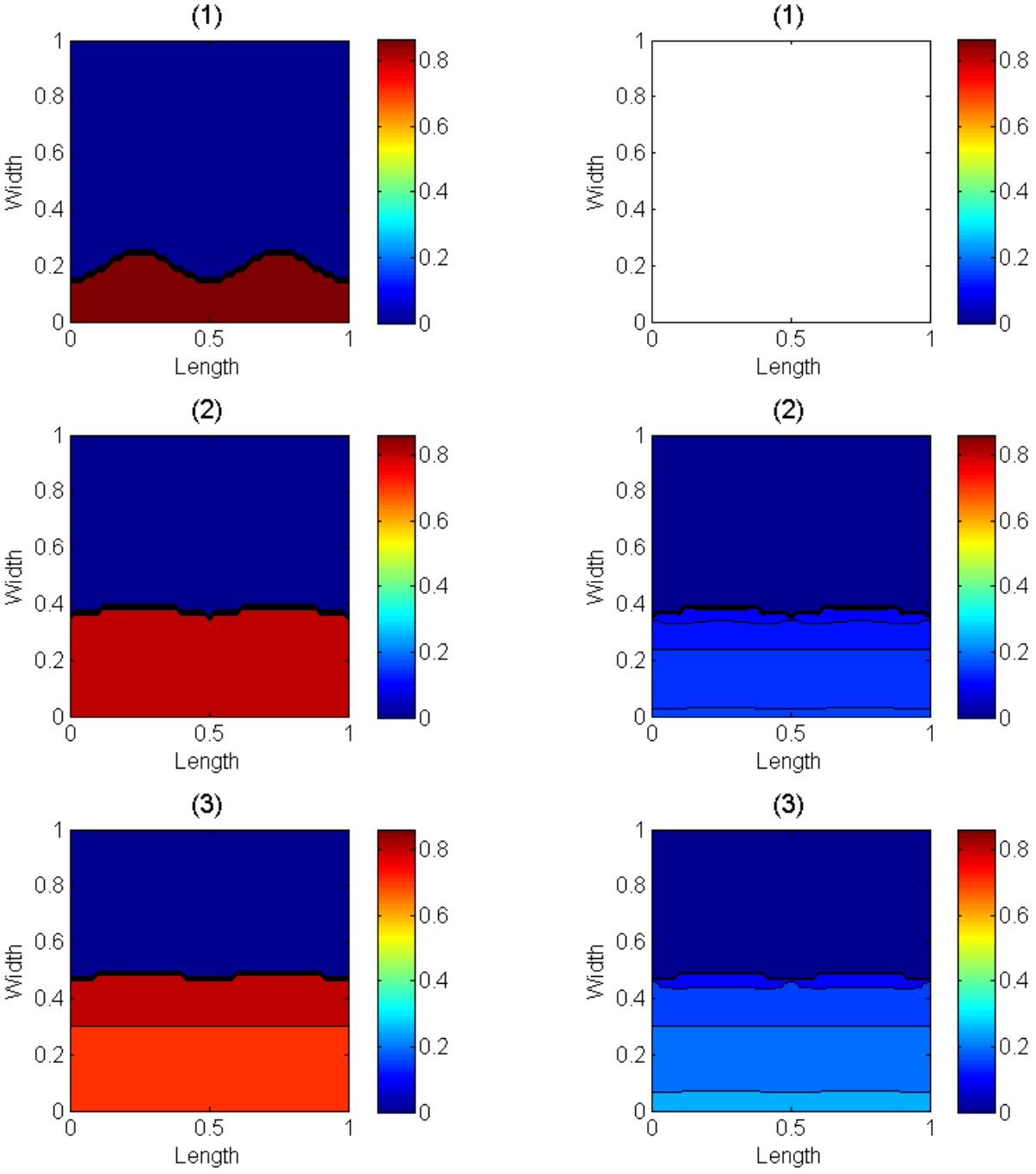

Figure 6.5: **Flat biofilms as predicted by the second model.** Active biomass dynamic is shown in the left column, while the right column shows the dynamic of inert biomass. (1) Present the initial conditions, (2) and (3) present biomass at 5.76h and 9.6h, respectively.

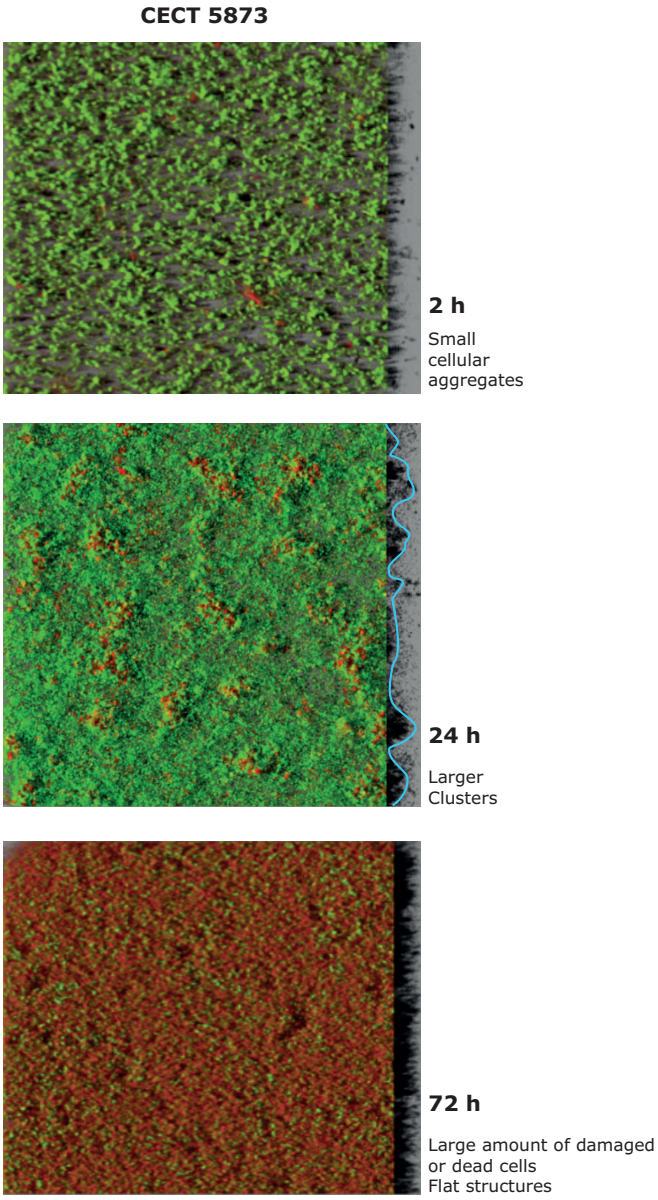

Figure 6.6: Dynamics of CECT 5873 strain *L. monocytogenes* biofilms during the growth phase of the life cycle as obtained by CLSM and image reconstruction.

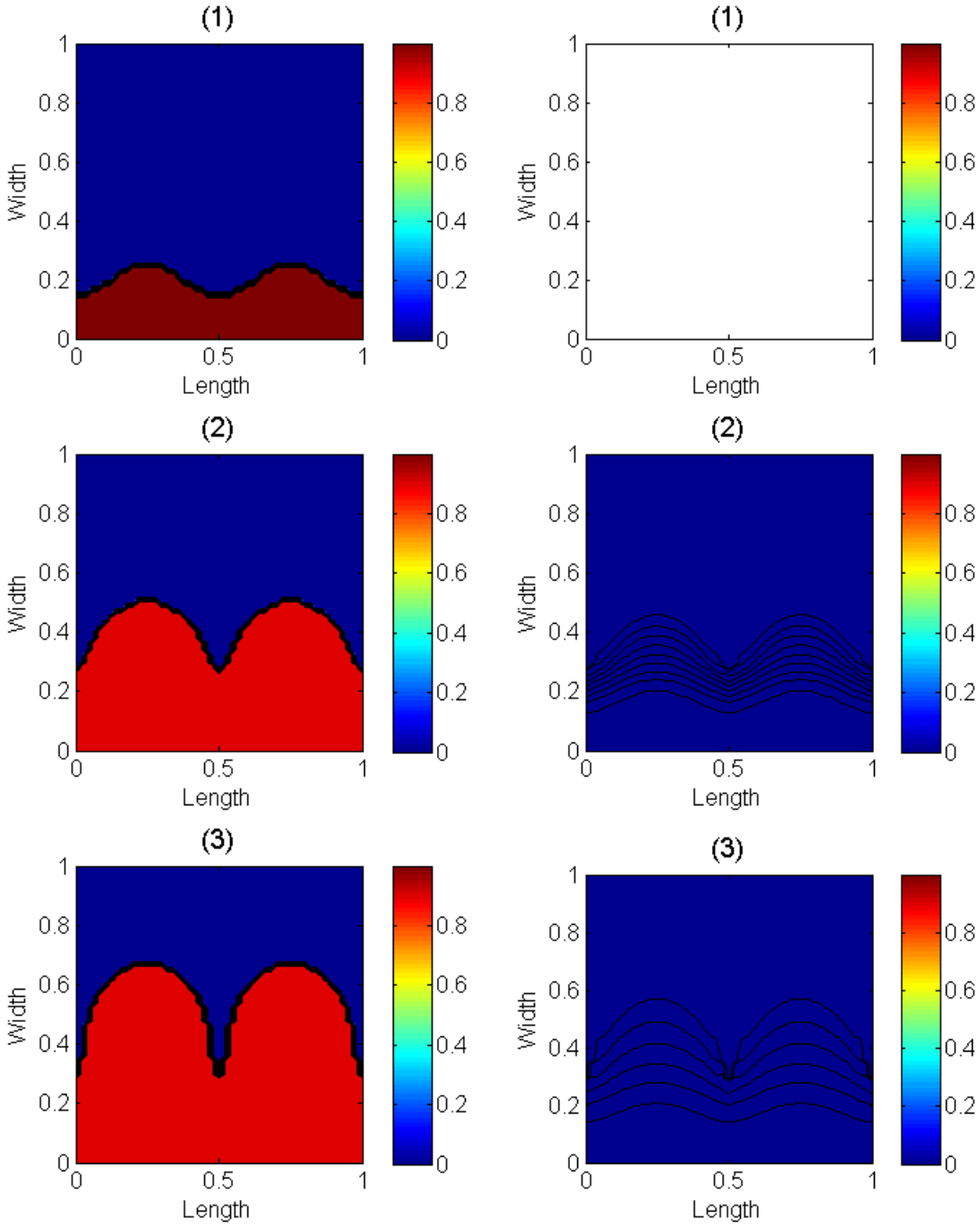

Figure 6.7: **Simulation of clustered biofilms as obtained by the second model.** In the left column the active biomass concentration is shown, whereas the right corresponds to the inactive biomass concentration. The instants shown are (1) initial situation, (2) 5.76h, (3) 9.6h.

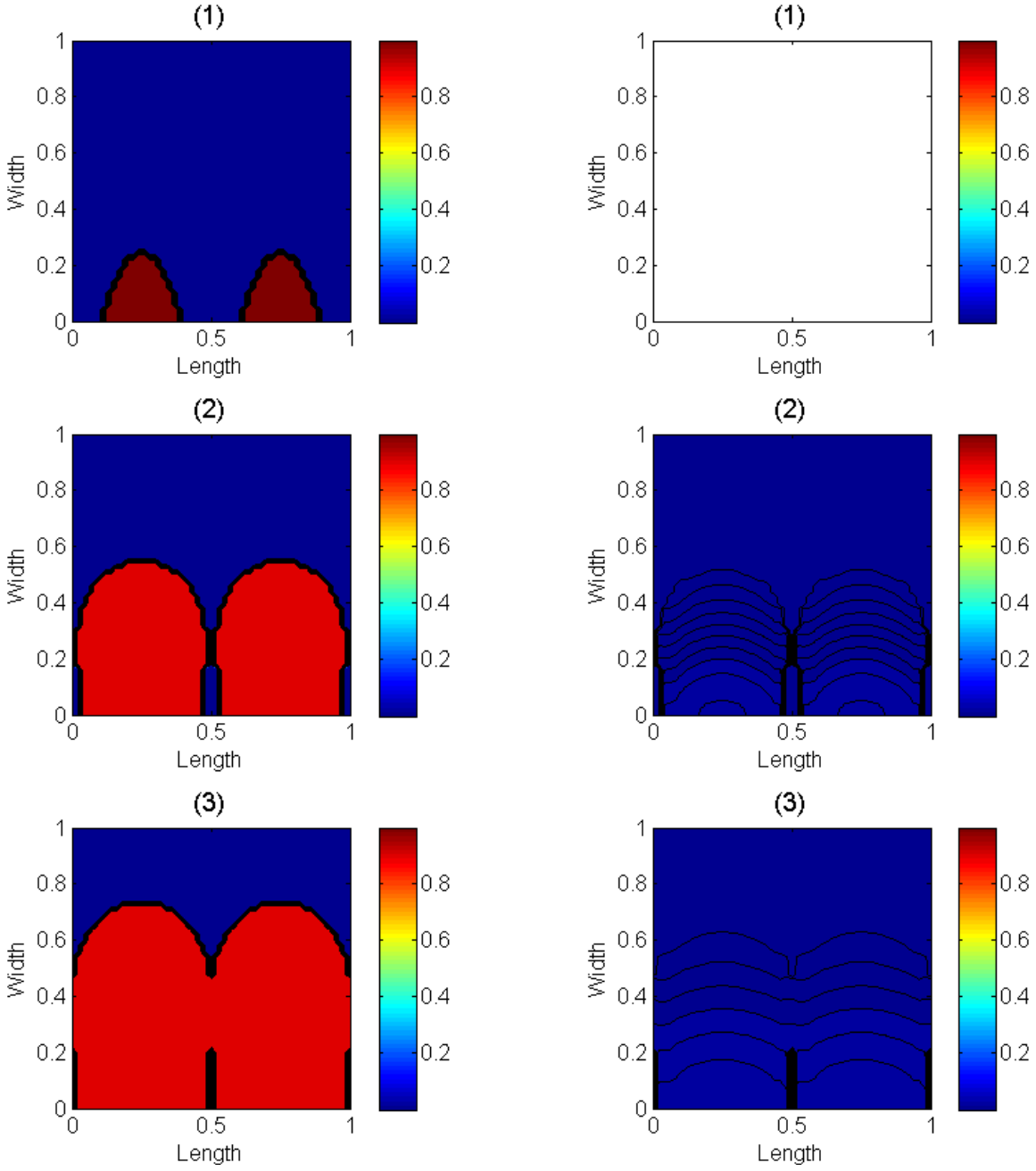

Figure 6.8: **Honeycomb structures as obtained by the second model.** In the left column the active biomass concentration is shown, whereas the right corresponds to the inactive biomass concentration. The instants shown are (1) initial situation, (2) 5.76h, (3) 9.6h.

Appendix A

Mathematical analysis of the Eberl model

In this appendix we summarise some results on the mathematical analysis that appear in the recent work [28]. This work is a comprehensive mathematical study of the Eberl Model (1.11)-(1.12) that basically presents a complete set of demonstrations validating the existence and uniqueness of solution for various types of boundary condition, as well as some qualitative properties related to the solution.

Following the procedures of Efendiev, Eberl and Zelik [28], we have the following model in a bounded spatial domain $\Omega \subset \mathbb{R}^n$ with piecewise regular boundary $\Gamma = \partial\Omega$, where $n \geq 1$:

$$\begin{cases} \partial_t C = d_1 \Delta_x C - K_1 \dfrac{CM}{K_2 + C}, & \text{in } \Omega, \\ \partial_t M = d_2 \nabla_x \cdot \left(\dfrac{M^b}{(1-M)^a} \nabla_x M \right) - K_4 M + K_3 \dfrac{CM}{K_2 + C}, & \text{in } \Omega, \\ C = 1, \quad M = 0, & \text{in } \partial\Omega, \\ C = C_0, \quad M = M_0, & \text{in } t = 0, \end{cases} \tag{A.1}$$

where Δ_x denotes the spatial laplacian $x = (x^1, \ldots, x^n)$, ∇_x represents the gradient/divergence operator, $C(t,x)$ is the nutrients concentration and $M(t,x)$ is the biomass density at point x in time t. In Table A.1 are listed the rest of the problem data:

Data	Desecription
$d_1 > 0$	Nutrients diffusion coefficient
$d_2 > 0$	Biomass diffusion coefficient
$K_1 \geq 0$	Maximum specific consumption rate
$K_2 > 0$	Monod half saturation constant (relative to C_0)
$K_3 \geq 0$	Maximum specific growth rate
$K_4 \geq 0$	Biomass decay rate
a	Biomass spreading parameter
b	Biomass spreading parameter

Table A.1: Parameters involved in (A.1)

Nutrients evolution equation is a semi-linear standard reaction-diffusion equation. In can be noted that the diffusion coefficient in the biomass equation is zero in absence of biomass and presents a singularity when biomass reaches $M = 1$. Both equations are coupled through the Monod Term, thus having a degenerate equations system.

In [28] an auxiliary non-degenerate system is introduced. The system depends on a parameter but the estimations obtained are independent of said parameter. Through that result, existence and uniqueness of the problem solution are obtained, under certain hypothesis on the initial conditions. Besides, certain qualitative properties related to stability are obtained.

It is worth noticing that the unknowns of (A.1) are non-dimensional magnitudes: C scaled with respect to the nutrients concentration in the liquid region and M with

respect to the maximum biomass density.
The following *a priori* hypothesis on the initial data (C_0, M_0) are required to obtain the results:

$$C_0, M_0 \in L^\infty(\Omega),\ 0 \leq C_0(x) \leq 1,\ 0 \leq M_0(x) \leq 1,\ x \in \Omega, \tag{A.2}$$

and we look for solutions (C, M) to problem (A.1) verifying the following conditions, according to the expected model properties:

$$C(t), M(t) \in L^\infty(\Omega),\ 0 \leq C(t,x) \leq 1,\ 0 \leq M(t,x) \leq 1, \tag{A.3}$$

where $(C(t), M(t)) = (C(t,.), M(t,.))$. For the sake of simplicity, let us define the set of acceptable initial conditions:

$$\mathbb{V} := \left\{(C_0, M_0) \in L^1(\Omega) \times L^1(\Omega) : (C_0, M_0) \text{ satisfies (A.2)}\right\},$$

and a set of more regular initial conditions based on the following hypothesis:

$$\begin{cases} C_0 \in L^\infty(\Omega) \cap H^1(\Omega),\ 0 \leq C_0(x) \leq 1,\ C_{0|\partial\Omega} = 1, \\ M_0 \in L^\infty(\Omega),\ F(M_0) \in H_0^1(\Omega), \\ M_0 \geq 0,\ ||M_0||_{L^\infty(\Omega)} < 1, \end{cases} \tag{A.4}$$

with

$$F(u) := \int_0^u \frac{v^b}{(1-v)^a} dv,\ 0 \leq u < 1.$$

In particular, let us consider the set:

$$\mathbb{V}_{smooth} := \{(C_0, M_0) \in L^\infty(\Omega) \times L^\infty(\Omega) : (C_0, M_0) \text{ satisfies (A.4) }\}.$$

It can be shown that $\mathbb{V} = [\mathbb{V}_{smooth}]_{L^1(\Omega)\times L^1(\Omega)}$, where $[\quad]_V$ denotes the closure operator in V.
We start stating some particular results for (A.1). Then, we will treat only the case with Neumann boundary conditions. This is a more complex problem, because the diffusion mechanism is not strong enough to guarantee the upper bound $M(t,x) \leq 1$.

In fact, depending on the values of the parameters from Table 2.1, it can be shown that given any non trivial M_0, exists a finite time at which the singularity $M \equiv 1$ is reached.

A.1 Auxiliary problem

For $(C_0, M_0) \in \mathbb{V}$ y $R > 1$, we want to find $(C_R(t), M_R(t))$ solutions of the following auxiliary problem:

$$\begin{cases} \partial_t C = d_1 \Delta_x C - K_1 \dfrac{CM}{K_2 + C}, & \text{in } \Omega, \\ \partial_t M = d_2 \nabla_x (f_R(M) \nabla_x M) - K_4 + K_3 \dfrac{CM}{K_4 + C}, & \text{in } \Omega, \\ C = 1, \quad M = 0, & \text{in } \partial\Omega, \\ C = C_0, \quad M = M_0, & \text{in } t = 0, \end{cases} \tag{A.5}$$

with

$$f_R(z) := \begin{cases} (z + 1/R)^b/(1 - z)^a, \text{ if } z \leq 1 - 1/R, \\ R^a, \text{ if } z > 1 - 1/R. \end{cases}$$

The estimations of the solutions of (A.5) are shown in [28]. With those solutions, one can pass to the limit $R \to +\infty$ in the auxiliary problem to construct a Lipschitz continuous semi-group associated with the degenerate problem (A.1).

A.2 Existence and uniqueness of the solution of problem (A.1)

Using the estimates and results of the auxiliary problem, the following result on existence of the solution of the degenerated problem (A.1) with smooth data is obtained.

Theorem 1. *Let be $(C_0, M_0) \in \mathbb{V}_{smooth}$, then there exists a solution (C, M) of the*

problem (A.1) (in the sense of distributions) that verifies:

$$\begin{cases} C, M \in L^\infty(\mathbb{R}_+ \times \Omega) \cap C([0,\infty), L^2(\Omega)), \\ C, F(M) \in L^\infty(\mathbb{R}_+, H^1(\Omega)) \cap C([0,\infty), L^2(\Omega)), \\ 0 \leq S(t,x), M(t,x) \leq 1, ||M||_{L^\infty(\mathbb{R}_+ \times \Omega)} < 1. \end{cases} \tag{A.6}$$

Using additional estimations satisfied by the solutions, the following theorem establishes uniform Lipschitz continuity properties with respect of the initial data of solutions that satisfy (A.6).

Theorem 2. *Let be (C_1, M_1) and $C_2, M_2)$ two solutions of (A.1) satisfying (A.6), then the following holds*

$$\begin{aligned} || C_1(t,.) - C_2(t,.) ||_{L^1(\Omega)} + || M_1(t,.) - M_2(t,.) ||_{L^1(\Omega)} \quad \leq \\ \exp(Kt)(|| C_1(0,.) - C_2(0,.) ||_{L^1(\Omega)} + || M_1(0,.) - M_2(0,.) ||_{L^1(\Omega)}), \end{aligned} \tag{A.7}$$

where K is a generic constant. In particular, the uniqueness of solution is granted for problem (A.1).

As a consequence of the existence and uniqueness theorems, Theorems 1 and 2, respectively, the model (A.1) generates a semi-group over $\mathbb{V}_{smooth}$, $\mathcal{S}_t$, that is uniformly Lipschitz continuous (in $L^1(\Omega) \times L^1(\Omega)$), defined as:

$$\mathcal{S}_t : \mathbb{V}_{smooth} \to \mathbb{V}_{smooth}, \ \mathcal{S}_t(C_0, M_0) := (C(t), M(t)),$$

where $(C(t), M(t)) = (C(t,.), M(t,.))$ is the solution of (A.1). That solution can be obtained as the limit in L^2 of the solution of the auxiliary problem when $R \to +\infty$.

Using the fact that $\mathbb{V}_{smooth}$ is dense in $\mathbb{V}$ and the property (A.7) we can extend in a unique way the semi-group $\mathcal{S}_t$ to $\mathbb{V}$ verifying (A.7). Said extension is obtained as follows:

$$\mathcal{S}_t(C_0, M_0) := L^1(\Omega) - \lim_{k \to \infty} \mathcal{S}_t(C_0^k, M_0^k),$$

where

$$(C_0^k, M_0^k) \in \mathbb{V}_{smooth}, \ \forall k,$$

and

$$(C_0, M_0) = L^1(\Omega) \times L^1(\Omega) - \lim_{k\to\infty}(C_0^k, M_0^k).$$

The extension gives the solution of problem (A.1) with initial conditions in $\mathbb{V}$.

Theorem 3 (Existence of solution in the sense of distributions)**.** *Let be* $(C_0, M_0) \in \mathbb{V}$ *and* $(C(t), M(t)) := C_t(C_0.M_0)$*. Then*

$$S, M \in L^\infty(\mathbb{R}_+ \times \Omega) \cap C([0,\infty), L^1(\Omega)). \tag{A.8}$$

More over, the following is hold

$$mes\{x \in \Omega : M(t,x) = 1\} = 0, \forall t > 0, \tag{A.9}$$

where mes$\{V\}$ *denotes the n-dimensional Lebesgue measure in* $\mathbb{R}^n$*, and*

$$||C(t)||^2_{H^1(\Omega)} + ||\partial_t C(t)||^2_{H^{-1}(\Omega)} + ||F(M(t))||^2_{H^1(\Omega)} + ||M(t)||^2_{H^s(\Omega)} + ||\partial_t M(t)||_{H^{-1}(\Omega)} \leq C\frac{t^\kappa + 1}{t^\kappa},\ t > 0,$$

with $0 < s < (b+1)^{-1}$*,* $\kappa > 0$ *and* $C > 0$ *given constants independent of* $(C_0, M_0) \in \mathbb{V}$*, and the functions* $(C(t), M(t))$ *are solution of (A.1) in the sense of distributions.*

Corollary 1. *Let be* $(C_0, M_0) \in \mathbb{V}$ *and let be* $(C_R(t), M_R(t))$ *solution of the auxiliary problem (A.5), then the solution of problem (A.1) can be calculated as*

$$C(t) = L^1(\Omega) - \lim_{R\to\infty} C_R(t),\ M(t) = L^1(\Omega) - \lim_{R\to\infty} M_R(t). \tag{A.10}$$

In other words, for each $(C_0, M_0) \in \mathbb{V}$*, the solutions of the auxiliary problem (A.5) converge when* $R \to +\infty$ *to the solution of (A.1) built in Theorem 3.*

A.3 Other boundary conditions

Suppose now that we have the boundary Γ divided into two subsets, i.e., that there exist $\Gamma_D \subset \Gamma$ and $\Gamma_N \subset \Gamma$ so that $\Gamma = \Gamma_N \cup \Gamma_D \cup \partial\Gamma$, $\partial\Gamma = \partial\Gamma_D = \partial\Gamma_N$, $\Gamma_N \cap \Gamma_D = \emptyset$.

Then, we consider homogeneous Dirichlet conditions in boundary Γ_D and homogeneous Neumann conditions in boundary Γ_N, i.e.:

$$M_{|\Gamma_D} = 0,\ \partial_n M_{|\Gamma_N} = 0.$$

We have basically the following two possible cases:

1. $\Gamma_D \neq \emptyset$, which corresponds to mixed conditions if $\Gamma_N \neq \emptyset$ or pure Dirichlet otherwise.

2. $\Gamma_D = \emptyset$, which corresponds to pure Neumann conditions.

For the first case, analogous results can be obtained, similar to the ones established for the case with pure Dirchlet conditions, just by repeating the demonstrations [28], and therefore a solution for problem (A.1) with mixed conditions can be found. In the second case, pure Neumann conditions, demonstrations of the existence of solutions can not be repeated, because it can not be guaranteed that, whatever the initial data $(C_0, M_0) \in \mathbb{V}$ are, the following holds

$$M(t,x) < 1, \text{ for almost all } (t,x).$$

Moreover, as said in [28], if the following is not fulfilled

$$K_3/(1+K_2) - K_4 > 0, \tag{A.11}$$

it can be demonstrated that the biomass density decays exponentially in time, which guarantees trivially that the maximum biomass bound is not reached. But, if (A.11), is fulfilled, the following result is obtained.

Proposition 1. *Let (A.11) be satisfied, then there exist initial data $(C_0, M_0) \in \mathbb{V}$, such that*

$$0 \leq C_0 < 1, \text{ and } 0 \leq M_0 < 1,$$

and there exist $T = T(C_0, M_0)$, such that

$$\langle M(t) \rangle < 1, \forall t < T, \text{ and } \lim_{t \to T^-} \langle M(t) \rangle = 1. \tag{A.12}$$

Proposition 2. *There exist positive values for* $K_1,\ K_2,\ K_3,\ K_4,\ d_1$ *and* d_2 *such as any solution of (A.1) with pure Neumann boundary conditions reaches the biomass bound in finite time.*

Remark 1. *It can be proven that condition (A.12) determines the existence interval for the solution* $(C(t), M(t))$*, i.e., problem (A.1) with pure Neumann boundary conditions for* M *has a local solution in time for every* $(C_0, M_0) \in \mathbb{V}$ *with* $\langle M_0 \rangle < 1$.

Appendix B

Theoretical background of the parameter calibration using AMIGO2 toolbox

As happens in the present work, mathematical models often consist of sets of different equations with the presence of many parameters which are not accessible through experimentation. Model identification and optimization is at the core of model building and requires the computation of unknown parameters by means of experimental data fitting [104] . Parametric identification is, however, a bottleneck in the modelling process due to the, usually, ill-posed and multimodal nature of the optimization problems, so it is necessary the use of suitable optimization methods to avoid local solutions.

AMIGO2 [5] is the first multi-platform environment that automatises the solution of the parametric identification processes, fully covering the iterative identification of dynamic models, allowing the use of optimality principles for predicting biological behaviour and dealing with the optimal control of biological systems using constrained multi-objective dynamic optimisation. In this appendix, an explanation on the identification process as solved by AMIGO2 is presented, following Balsa-Canto et al.[6].

B.1 Parameter identification iterative procedure

The predictive character of a mathematical model often relies on determining the non-measurable parameters that represent the greatest influence over the response of the model. These parameters can be usually estimated by fitting the model to the experimental data, i.e, through parameter estimation or identification. However, it is a challenging problem, specially when applied to the parametric identification of non-linear dynamic models.

In what regards to parametric identification, AMIGO2 covers the numerical steps summarised in Figure B.1 and described as follows:

- Simulation process where the system dynamics are solved for different parameter values under different schemes. It is a useful step to analyse the model tendencies *a priori* and for the validation step *a posteriori*.
- Global ranking of the parameters that helps to decide which parameters are the most relevant to the output. The global ranking may be used to make decision like to reformulate the model or to select which parameters we need to estimate.
- Parameter estimation formulated as a non-linear optimisation problem whose objective is to find unknown model parameters so that the measure of the distance between the model predictions and the experimental data is minimised. It is necessary the use of global optimisation methods to guarantee somehow that the best possible solution is located.
- Statistical best fit post analysis.
- Practical identifiability analysis to evaluate the possibility of assigning unique values to the parameters.
- Optimal experimental design via dynamic optimization in order to design dynamic experiments with the aim of maximising data quality and quantity for the purpose of model calibration.

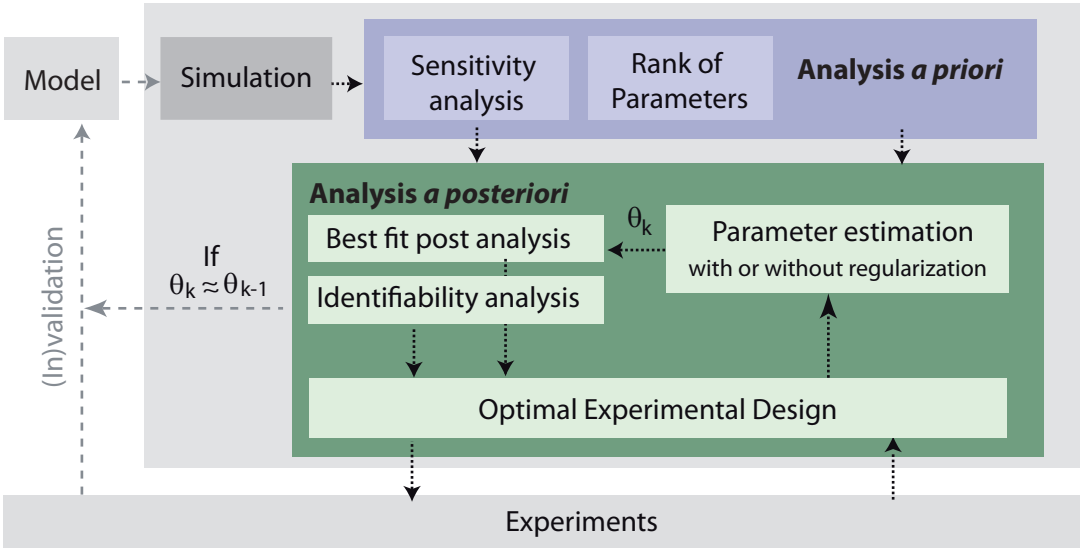

Figure B.1: AMIGO2 iterative procedure. Source [6]

B.1.1 Elements for parametric identification

The model

The mathematical model will consist, on one hand, on the set of differential equations, f, describing the system dynamics, a general deterministic non-linear dynamic model

$$f(\dot{x}, x, u, \theta, t) = 0, \tag{B.1}$$

and, on the other hand, the observation function, g^{ϵ}, describing the relationship among the states in the model and the available measured quantities

$$y^{\epsilon} = g^{\epsilon}(x, u, \theta, t), \tag{B.2}$$

where x are the state variables, y^{ϵ} is the vector of observables, u is the vector of the manipulable variables, θ is the vector of model parameters and t the set of admissible parameters that may be fixed by physical, chemical or biological considerations.

The experimental scheme and data

The experimental scheme collects the information related to the way experimental data are obtained. The experimental data consist on matrices of values corresponding to individual measurements obtained under the conditions stated by the experimental scheme ϵ. This data and the model predictions are represented as

$$y = [y_1, y_2, \ldots, y_d, \ldots, y_{n_d}], \qquad \tilde{y} = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_d, \ldots, \tilde{y}_{n_d}],$$

where d represents a certain experimental condition and n_d represents the total number of such conditions (the total number of data). Note that is also desirable to provide information about the type and quantity of noise in the experimental data. In the *homoscedastic* case the noise can be constant or known for all data, while in the *heteroscedastic* case it can be unknown and dependent on the data.

B.1.2 Parameter estimation

The parameter estimation problem is to find model unknown parameters so as to minimise a given measure of the distance among the model predictions and the experimental data.

Distance measure

There are different ways to measure the distance between the experimental data and the model predictions, depending on the available information of the experiment. The most well known is the generalised least squares

$$J_{glsq}(\theta) = \sum_{d=1}^{n_d} q_d \left(y_d(\theta) - \tilde{y}_d\right)^2, \tag{B.3}$$

where $\{q_d\}_{d=1}^{n_d} \geq 0$ are the weighting coefficients fixed *a priori*. This coefficients determine the confidence in the experimental data and their importance in the model performance with regards to each type of measurement, experiment and sampling time.

Instead of using the least-squares function, when there is information about the nature of the experimental noise, the maximum (log-)likelihood function can be used

$$J_{llk} = ln\left(\Pi(\tilde{y}|\theta)\right). \tag{B.4}$$

This function, looks for the value of the parameters that give the highest probability to the measured data. Function θ is the probability density function and conditions the type of cost function.

In the Gaussian homoscedastic case, where the variance is known or constant, the cost function is similar to the generalised least squares taking as weights the inverse of the variance of the experimental data [104]

$$J_{lsq}(\theta) = \sum_{d=1}^{n_d} \frac{\left(y_d(\theta) - \tilde{y}_d\right)^2}{\sigma_d^2}. \tag{B.5}$$

The estimation of this variances requires a great amount of a prior experiments and involve multiple identification problems. Therefore, although not realistic, many applications select for all measurements a constant variance $\sigma_d = \sigma, \forall d = 1, ..., n_d$. In this case, the error must be carefully analysed.

Finally, in the Gaussian heteroscedastic case, the corresponding log-likelihood function is

$$J_{llk}(\theta) = \sum_{d=1}^{n_d} \log|y_d(\theta)| + \frac{\left(y_d(\theta) - \tilde{y}_d\right)^2}{\sigma^2(a, b, y_d(\theta))}, \tag{B.6}$$

where the variance depends on what is being measured and is assumed to be

$$\sigma^2(a, b, y(\theta)) = |ay(\theta)|^b,$$

with $a > 0$ and $0 \leq b \leq 2$. The case where $b = 2$ is the standard deviation proportional to the output.

Single shooting v multiple shooting

As described above, the parameter estimation problem is formulated as a non-linear optimisation problem where the objective is to find the set of model unknowns to

minimize a certain cost function subject to the system dynamics and the possible bounds on the unknowns. The numerical solution of this problem calls for a non-linear programming method (NLP, an outer iterative method to generate values for the unknowns and the inital conditions) and the initial value problem solver (IVP, an iterative method to solve the differential equations).
In the single shooting approach, the IVP is solved from the initial conditions to the final time for all the iterates generated by the NLP. On the other hand, the multiple shooting approach discretises the process into several shooting interval solving an IVP in every shooting so that at least one experimental data may be found. In this case, the initial conditions for the different intervals are also to be computed during optimisation, therefore, further constraints on the parameter estimation problem are required so as to guarantee that the optimal solution is smooth. In this work the single shooting approach was selected due for convenience to handle PDE based systems.

B.1.3 Practical identifiability analysis

There are two types of practical identifiability analysis: *a priori*, which anticipates de quality of the selected experimental scheme in terms of the expected uncertainty of the parameters; and *a posteriori*, which assesses the quality of the parameter estimates after the model calibration in terms of the confidence region.

Note that *a priori* a maximum experimental error is assumed while *a posteriori* the experimental error may be estimated through the available experimental data manipulation or after the parameter estimation using the residuals.

The simplest approach to perform such analyses is to draw contours of the cost J_{lsq} or J_{llk} by pairs of parameters, that help detect typical identifiability problems. Another possibility is the Crammér-Rao inequality [59] which establishes a relationship between the Fisher Information Matrix ($\mathcal{F}$) and the covariance matrix (C) for the

case that the estimator is asymptotically unbiased

$$C \geq \mathcal{F}(\theta^*),$$

where θ^* an optimum value for the parameters. The confidence interval for a certain parameter θ_i^* is

$$t_{\alpha/2}^{\gamma}\sqrt{\boldsymbol{C}_{ii}},$$

where $t_{\alpha/2}^{\gamma}$ is given by the Students t-distribution, γ the number of degrees of freedom and α the $(1-\alpha)100\%$ confidence interval selected by the user.

A final alternative for a most robust analysis are Monte-Carlo [8], [49] sampling and bootstrap methods that can robustly quantify the expected uncertainty of the parameters or the confidence region. The underlying idea is to stimulate the possibility of performing hundreds of replicates of the same experimental scheme for a certain experimental error. The model calibration problem is solved for each replicate and the cloud of solutions is recorded in a matrix. It should be noted that an efficient global optimization method is required to avoid convergence to local solutions.

With the confidence intervals, a decision on the need to perform further experiments to improve the quality of the parameter estimates can be taken, thus improving the predictive capabilities of the model.

B.2 Numerical methods

B.2.1 Initial value problem solvers

AMIGO2 offers the possibility of handling ODE based models with different numerical techniques, and also the possibility of linking your own numerical approaches for model simulation or optimization. In this work we exploited that possibility and

linked our numerical implementation of the Eberl model to AMIGO2 for the quantitative validation of the model. For the case of linked models there is the possibility of computing parametric sensitivities by means of finite differences schemes. In our case we used 5 points FD formulae.

B.2.2 Non-linear programming solvers

The aim of the optimisation methods is to generate a sequence of solutions that, eventually, converges to the the minimum of the cost function. The methods are classified as local methods, able to handle convex non-linear problems, and global methods, able to handle non-linear non-convex multimodal problems.

Local methods

The direct-search methods use the value of the cost function at several points in the vicinity of the current iterate to generate new iterates. Its major disadvantage is the slow convergence. As an alternative, indirect methods use the gradient or the gradient and Hessian information to increase convergence speed. In the context of least squares minimization, the most widely used is the Levenberg-Marquardt method, a combination of the steepest descent with the Newton method for the least squares cost function. The two major advantages of the indirect local method are

- Convergence to a minimum is guaranteed by the fact that the gradient of the cost function evaluated at the optimum is zero, and the Hessian is positive definite.
- The methods are highly efficient when started close to the solution.

The local methods have been also used in combination with the single and multiple shooting approaches, although the non-linear character of the dynamic biological models tends to lead to the presence of several suboptimal solutions so local methods may end up in such a solution. Multiple shooting can evade some local minima

by allowing for discontinuous trajectories while searching the global minimum but convergence to the global solution can not be guaranteed. Moreover, in the presence of a bad fit, there is no way of knowing if that happens due to a wrong model or as a consequence of local convergence.

Global methods

Global methods are an alternative to local methods. Suitable techniques for the solution of multi-modal optimization problems have been developed [72]. The successful methodologies combine effective mechanisms of exploration of the search space and exploitation of the previous knowledge obtained by the search. Depending on how the search is performed and the information exploited, there are deterministic methods, stochastic methods and hybrid methods.

Global deterministic methods take advantage of the structure of the problem and guarantee convergence to the global minimum for some particular problems. Although promising and powerful, there are still limitations to their application such as the rapid increase of computational cost with the size of the system and the number of its parameters.

Global stochastic methods do not require any assumptions about the structure of the problem and make use of pseudorandom sequences to determine search directions toward the global optimum, leading to an increasing probability of finding the global optimum during the runtime of the algorithm. The main advantage is that they rapidly arrive to the proximity of solution. Some of these strategies have been applied to parameter estimation problems in the context of system biology.

As already stated, stochastic methods can locate the vicinity of global solution quite rapidly, but the computational cost associated with the refinement of the solution is very large. Hybrid methods aim to surmount this difficulty. They speed up the methodologies while retaining their robustness. An excellent hybrid method is the Scatter Search method [29].

B.2.3 Numerical methods in AMIGO2

AMIGO2 offers state-of-the-art methods to cover all numerical tasks: simulation, sensitivity analysis, and optimization. Figure B.2 summarize the currently available numerical methods used in the non-linear solvers.

Models considered in this work resulted in multimodal parameter estimation problems. Therefore we used a global metaheuristics eSS (Egea et al. [29]) which combines the Scatter search approach with a deterministic local method, in this case the Nelder-Mead method.

Local solvers	Algorithm	Description
Direct	fminsearch	Nelder-Mead. J.C. Lagarias, J.A. Reeds, M. Wright, P. E. Wright, Convergence properties of the Nelder-Mead Simplex Method in Low Dimensions, SIAM J Opt , 9(1):.112-147, 1998.
	dhc	Dynamic hill climbing method. de la Maza & D. Yuret. Dynamic hill climbing. AI Expert, 9(3):26(31, 1994.
	nomad	Pattern search method. M. A. Abramson. Pattern Search Algorithms for Mixed Variable General Constrained Optimization Problems. PhD , Rice University, 2002
Indirect	lsqnonlin	Included in MATLAB Optimization toolbox. Levenberg-Marquardt and trust-region-reflective methods.
	n2fb (32bits) dn2fb (64bits)	Least-squares method. J.E. Dennis, D. M. Gay, and R. E. Welsch. An adaptive non-linear least-squares algorithm. ACM Trans Math Soft, 7(3):348-368, 1981.
	nl2sol	Least-squares method. J.E. Dennis, D. M. Gay, and R. E. Welsch. An adaptive non-linear least-squares algorithm. ACM Trans Math Soft, 7(3):348-368, 1981.
	fmincon	SQP (Sequential Quadratic Programming). Included in MATLAB Optimization toolbox.
	solnp	Interior point SQP. Y. Ye. Interior algorithms for linear, quadratic and linearly constrained non-linear programming. PhD , Stanford University, 1987.
	ipopt (32bits)	Large scale interior point. A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math Prog., 06(1):25-57,2006.
Multistart	multi_local	N starts of any of the available local solvers: to analyse multimodality

Global solvers	Algorithm	Description
	DE	Population based di erential evolution. Storn R, Price K. Di erential Evolution – a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces.J Global Optim, 1 1:341-359, (1997)
	SRES	Evolutionary search method. Runarsson T, Yao X. Stochastic ranking for constrained evolutionary optimization. IEEE Trans Evol Comp, 564:284-294, (2000)

Hybrid solvers	Algorithm	Description
	hyp_global_local	All possible combinations of Global stochastic methods with the above mentioned local solvers Balsa-Canto et al: Dynamic Optimization of Single- and Multi-Stage Systems Using a Hybrid Stochastic-Deterministic Method.Ind Eng Chem Res,44(5):1514-1523, 2005. Rodriguez-Fernandez et al. A hybrid approach for e cient and robust parameter estimation in biochemical pathways. Biosyst, 83:248-265, (2006)
	eSS	Egea JA, Rodriguez-Fernandez M, Banga JR, Martí R. Scatter Search for Chemical and Bio-Process Optimization. J Glob Opt, 37(3):481-503, (2007)

Multi-objective	Algorithm	Description
	wsm_solver	Weighted sum method with any of the available solvers
	NSGAII	K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-II. IEEE Transaction on Evolutionary Computation, 6(2):181–197, 2002.

Other optimizers	Algorithm	Description
	my_solver	Users can test their own optimizers within the tool

Figure B.2: AMIGO^2 numerical methods for optimization. Source [6]

Conclusions

The main objective of this work has been the study and statement of some mathematical models for biofilms, so that they can be successfully applied to describe the dynamics of the formation of biofilms by several strains of *Listeria monocytogenes*, a pathogenic bacterium of serious concern in the food industry. These models are based on systems of linear and non-linear partial differential equations jointly with appropriate boundary and initial conditions, the solutions of which require the use of efficient and accurate numerical methods. Once a model has been selected and numerically solved, the computed results that reflect the model predictions are compared with those obtained in *ad hoc* experiments in laboratory. In some cases the parameters are optimised so that the matching with experimental results is obtained.

In this methodological setting, in Part I of this work we have first addressed one-dimensional biofilm models starting from the one proposed by Eberl et al. [26]. One-dimensional models are suited to describe flat biofilms as those formed by several *L. monocytogenes* strains. In order to obtain a more realistic representation of the life cycle followed by one particular strain, L1A1, we introduce a second model that considers non-linear detachment and a third one, including nutrient uptake impairment and ageing. As the three models are posed in terms of systems of linear and non-linear time-dependent partial differential equations without analytical expression for their respective solutions, we propose a set of suitable numerical techniques to approximate their solutions. These techniques mainly consist of a Crank-Nicolson for

the time discretisation and appropriate finite differences schemes for the discretsation in space, which are combined with Newton method for the involved non-linear terms when they appear. The numerical methods have been implemented from scratch in C++ and suitable academic examples with known analytical solution have been used to validate their implementation and illustrate their performance and order of convergence in practice. In a next step, the experimental validation of the models was addressed through a calibration of the parameters using the data fitting technique included in AMIGO2 toolbox. From these procedures, we summarize several conclusions that came up:

- Although in Eberl model the maximum height is close to the experimental results, the final massive detachment that should occur is not observed.
- In view of the results of Eberl model, a non-linear detachment is introduced in model 2, obtaining a greater final detachment closer to experimental values. However, in both models an unrealistic total consumption of nutrients takes place.
- In order to overcome the drawbacks of the previous models, a third model is set up. This model replaces the non-linear detachment by a linear term that considers biofilm ageing and includes nutrient impairment uptake. Once the parameters are calibrated to the batch experiment of L1A1, the conclusion is that this third new model results to be the one that better reflects the life cycle of the selected biofilm.

Secondly, in Part II of the thesis we have addressed biofilm models involving two spatial dimensions and several species. As previously indicated in this document, the motivation of the so called two dimensional models comes from the fact that the presence of internal tunnels, inner voids or other biofilm structures developed during the life cycle of L1A1 or other structures including clusters or mushrooms formed by other *L. monocytogenes* strains cannot be represented by means of one dimensional models. The selection of a multi-species model allows describing the role of viable,

damage or dead cells, together with several nutrients in the biofilms life cycle.

Thus, in the setting of two-dimensional models we start from the model proposed by Alpkvist and Klapper [1] and then we propose a new model that takes into account the conclusions observed in the analysis of one-dimensional models. More precisely, this new model incorporates the non-linear detachment and linear nutrients uptake to better reflect flat biofilms produced by L1A1 strain, clustered biofilms as those formed by L1A1, CECT4032 or CECT5873 strains at initial stages of their life cycle and honeycomb profiles observed in various *L. Monocytogenes* strains at different stages of their life cycle. As the complexity of the biofilm models increases in the case of two spatial dimensions, the set of suitable numerical techniques becomes more sophisticated. These techniques include Crank-Nicolson discretisation in time; central finite differences scheme in space modified with Gibou's method [36] for the nutrients and pressure equations; an upwinding finite differences scheme in space for the biomasses equations; the level set method to track the evolution of the numerical domains and a Newton method when non-linear terms appear. All of these numerical techniques are described in Chapter 5, the larger of this document, and have been implemented form scratch in C++. The step of qualitative validation of the two-dimensional models is based on several works in the literature.

- Concerning to the validation of flat *Listeria monocytogenes* life cycle, we first took advantage of some calibrated parameters in the more successful one-dimensional model. Next, we observed that although Alpkvist model reproduces the expected flat structures of L1A1, the dynamics of nutrients consumption and biofilm growth result to be unrealistically fast.

- The consideration of the new proposed two-dimensional model including a detachment mechanism that considers the mean inert biomass leads to a more realistic model, so that the observed dynamics are much slower than in the Alpkvist model.

- Furthermore, the manipulation of initial conditions and parameters in the second two-dimensional model successfully allows to represent the development of clustered or honeycomb structures, which are observed under certain conditions in the literature.

In a general sense, the present work aims to be a contribution to the modelling and numerical simulation of biofilms. Thus, we have studied some relevant biofilm models in the literature and analysed their calibration to the experimental results in the laboratory. This analysis motivated the introduction of new one-dimensional and two-dimensional models that become more realistic to represent the experimental results we observed and the expected dynamics of particular biofilms formed by *L. monocytogenes* strains. Also, this work can be understood as another modest step in the comprehension of the mechanisms of life cycles in certain biofilms by using the relevant tools of mathematical modelling and numerical simulation from Applied Mathematics.

Of course, the study opens the possibility of new research steps that were not covered here. For example, among them are the mathematical analysis of these complex models to obtain existence and uniqueness of solutions in the appropriate functional spaces; the rigorous proof of convergence of some of the involved numerical methods for specific or global problems; the adaptation of the models to include mechanisms of biofilm elimination by biocides or antibiotics, cell motility or distinction between cells and EPS; the optimisation of the numerical resolution in order to speed it up so a parameter calibration of the two-dimensional models can be achieved; etc.

Resumen extenso

Las biopelículas bacterianas son esenciales tanto para el equilibrio de los ecosistemas como para muchos procesos industriales. Una biopelícula bacteriana es una capa de células procariotas ancladas a una superficie que les proporciona sustrato, mantieniéndolas alimentadas. Esta capa de células está integrada en una matriz formada por polímeros, llamada matriz de exopolisacáridos o EPS, que mantiene las bacterias agrupadas y dificulta su eliminación [21, 108]. De forma simplificada, podemos decir que una biopelícula es una agrupación de microorganismos anclados a una superficie. Los estudios actuales estiman que menos de un 0.1% de la vida acuática microbiana se encuentra en estado planctónico (flotando libremente) [20]. Por ende, las biopelículas constituyen la forma de vida preferida de las bacterias.

La razón de esta preferencia es la ventaja competitiva que supone la habilidad de adherirse a superficies y formar biopelículas, en comparación con las bacterias en estado planctónico. Éstas últimas pueden ser barridas con facilidad por el flujo del agua, mientras que las primeras están protegidas de ese fenómeno y viven en un entorno donde pueden crecer, siempre que haya suficiente disponibilidad de nutrientes. La estructura física de una biopelícula también permite la aparición de diferentes nichos biológicos que facilitan el crecimiento y la supervivencia de microorganismos que no podrían competir en un sistema completamente homogéneo. Más aún, la actividad microbiana en el interior de una biopelícula puede cambiar el entorno interno, haciendo que la biopelícula sea más hospitalaria que la región líquida. Las partes fundamentales de una biopelícula son la superficie (a la que se adhieren las

bacterias), la biopelícula propiamente dicha (formada por una o más especies de microorganismos y los EPS), el medio (con los nutrientes) y las condiciones del entorno que determinan el desarrollo de la biopelícula (temperatura, pH, hidrodinámica, etc.).

Existen biopelículas beneficiosas, tanto para los humanos como para el desarrollo adecuado del medio ambiente, así como biopelículas perjudiciales, que suelen causar problemas de salud o bajos rendimientos en procesos industriales y, por tanto, perdidas económicas. Ejemplos [108] de las primeras son las biopelículas empleadas en las plantas de tratamiento de aguas (RBC, reactores biológicos, etc.) o en la industria fermentera (por ejemplo, en procesos de *quick-vinegar*), pero también algunas biopelículas creadas de forma natural presentes en el subsuelo (contribuyendo a la descontaminación del suelo o de las aguas subterráneas), en ríos, lagos y zonas costeras (colonizando rocas o suspendidas en el agua, normalmente contribuyendo a la eliminación de contaminantes en el agua), o en las raices de muchas plantas (aumentando la disponibilidad de nutrientes para las plantas). Las biopelículas creadas de forma natural son fundamentales para la biosfera terrestre.

Por otra parte, las biopelículas perjudiciales aparecen en muchas situaciones. Por ejemplo, las biopelículas son un problema considerable en la higiene dental [61]. También son la causa de infecciones en implantes médicos y de complicaciones en enfermedades infecciosas [115]. Otros ejemplos son las biopelículas que causan contaminación del agua o mal funcionamiento de los intercambiadores de calor [11]. Especialmente relevantes son las biopelículas de bacterias patógenas que aparecen en la industria alimentaria, pues constituyen una fuente importante de contaminación alimentaria y pueden ser un grave problema de salud para los consumidores [93].

Prevenir la formación de biopelículas es, en general, bastante difícil, debido a su habilidad para desarrollarse incluso en condiciones adversas. Más aún, una vez formadas, son difíciles de eliminar, pues las bacterias que están formando biopelículas son mucho

más resistentes a la respuesta inmune del huésped o a agentes antimicrobianos [34, 44].

La necesidad de mejorar las propiedades de las biopelículas beneficiosas o controlar la formación de aquellas perjudiciales ha motivado numerosas lineas de investigación que tienen en cuenta los mecanismos genéticos, bioquímicos o físicos que contribuyen, no solo a la formación de biopelículas, si no también a su estructura.

Numerosos estudios indican que la estructura de una biopelícula determina la magnitud de los procesos que tienen lugar en ella, como por ejemplo la tasa de transferencia de nutrientes hacia las capas interiores, la tasa de difusión de agentes microbianos o su resistencia a la fricción. El recuento de placas ha sido ampliamente utilizado para analizar biopelículas. Sin embargo, sólo incluye células cultivables viables y no dan información acerca de la estructura [22]. Debido a esas limitaciones, varias alternativas han sido propuestas.

Se ha prestado especial énfasis al desarrollo de varias técnicas microscópicas. La más exitosa, la microscopía confocal láser de barrido (CLSM), permite realizar escaneo optico tridimensional *in situ* e *in vivo* [95]. Por otra parte, los microsensores pueden ser utilizados para medir concentraciones de los diferentes componentes dentro de una biopelícula, permitiendo así el análisis de la disponibilidad de nutrientes en las diferentes regiones de la biopelícula [116]. Los avances en biología molecular y en técnicas de hibridación *in situ* contribuyeron al desarrollo de sondas genéticas y técnicas de microscopía, permitiendo el análisis detallado de comunidades microbianas en el interior de la biopelícula [23, 56, 97].

La combinación de tintes celulares utilizados en técnicas de tinción celular fluorescente junto con ténicas de obtención de imágenes permiten localizar células viables y células dañadas o muertas [99] o la distribución de la sustancia polimérica extracelular, pero también la reconstrucción de estructuras tridimensionales, obteniendo así

un estudio de las biopelículas más exhaustivo. Cabe destacar que la técnica de CLSM junto con el análisis cuantitativo de imágenes [114] permite determinar de forma automática el grosor de una biopelícula, su volumen o la rugosidad. Este enfoque permite una comparación cuantitativa de biopelículas de diferentes cepas y especies o con diferentes condiciones ambientales (medio, temperatura, tipo de cultivo).

Algunos trabajos recientes sugieren varias estrategias de trabajo y herramientas de *software* para el análisis sistemático de imágenes de microscopía. IMARIS (un *software* comercial) permite la reconstrucción de estructuras 3D. COMSTAT [43], ISA 3D [10], o PHLIP [67] permiten cuantificar las imágenes obtenidas mediante CLSM. BIOFILMDIVER [65] permite la cuantificación de la porosidad de una biopelícula, el área cubierta, las distancias de difusión o la distribución de población espacio-temporal mediante imágenes 2D tomadas por epifluorescencia y CLSM. Algoritmos tipo *machine learning* pueden ser empleados para analizar las imágenes obtenidas por microscopía electrónica de barrido [103].

El análisis de imagen cuantitativo puede complementarse con el modelado matemático para ganar conocimiento acerca de los mecanismos que dan lugar a una determinada estructura en una biopelícula. Incluso las biopelículas más homogéneas desarrollan procesos internos complejos, entrelazados entre sí. El modelado matemático da la posibilidad de explorar diferentes procesos internos y sus conexiones, su importancia relativa y el rol que tiene el medio ambiente en el ciclo vital de una biopelícula [105]. Ese es, precisamente, el paso que da el presente trabajo, al combinar el análisis de imágenes obtenidas mediante CLSM con nuevos modelos matemáticos y técnicas numéricas avanzadas para explicar cuantitativamente y cualitativamente el comportamiento de biopelículas de *Listeria monocytogenes*.

Las biopelículas de *L. monocytogenes* pueden llegar a desarrollar diferentes estructuras: mono-capas de células adheridas, multi-capas planas desestructuradas, o

estructuras en panal o *clusters* [12, 15, 24, 62, 78, 82]. Trabajos recientes [37] reconstruyeron imágenes obtenidas por CLSM para observar la diversidad intraespecie de *L. monocytogenes*, concluyendo que la mayoría de cepas forman estructuras tipo panal complejas a las 48h. Otros trabajos, [65, 66] usan análisis cuantitativo de imágenes para estudiar el ciclo vital de biopelículas formadas por tres cepas de *L. monocytogenes*, mostrando la presencia de, al menos, tres fases: una fase inicial de clusters separados que evolucionan a clusters interconectados, estructuras tipo panal o estructuras planas, y una fase final en la que las células se separan de la biopelícula. La duración de estas fases varía significativamente entre cepas.

En una primera aproximación, en este trabajo proponemos modelos 1D continuos de reacción-difusión que permiten el análisis de la cepa L1A1 de *L. monocytogenes*. Dicha cepa forma biopelículas planas [66], por lo que parece apropiado utilizar modelos 1D para explorar los mecanismos relevantes que causen este tipo de estructuras. Los modelos aquí presentados han sido resueltos numéricamente empleando técnicas numéricas eficientes y robustas, que consisten fundamentalmente en un método Crank-Nicolson para la discretización en tiempo y esquemas en diferencias finitas apropiados para la discretización en espacio, todo ello junto con un método de Newton allí donde aparecen términos no lineales. Los métodos numéricos han sido implementados desde cero en C++ y validados numéricamente mediante tests académicos apropiados.

El estudio se basa en la disposición experimental utilizada en el trabajo reciente [66] que considera biopelículas que crecen en condiciones de flujo estático, restringiendo el estudio al caso hidrostático. De acuerdo con dicho estudio, las cepas de L1A1 desarrollan biopelículas prácticamente planas y desestructuradas. Tras una sedimentación inicial, se forma una biopelícula delgada antes de 24 horas. La estructura plana es estable con un aumento sostenido de su grosor hasta llegar a las 96 horas, momento en el cual la presencia de células dañadas o muertas es importante. Tras 120 horas,

la altúra máxima de la biopelícla decrece fuertemente, indicando un desprendimiento masivo en las últimas 24 horas.

Las células de L1A1 consumen glucosa como fuente primaria de carbono. Por tanto, los datos de nutrientes considerados en la resolución numérica corresponden a glucosa. Los datos experimentales muestran que las células consumen la mayor parte de las reservas de glucosa en las primeras 24 horas, tras lo cual el consumo de nutrientes cesa, incluso aunque aun se encuentren en el sistema celulas viables.

Empezamos nuestro estudio utilizando como punto de partida el modelo presentado por Eberl et al. [26] y validado por resultados de análisis matemático [28],

$$\frac{\partial S}{\partial t} = d_1 \nabla^2 S - K_1 \frac{SC}{K_2 + S}, \tag{1}$$

$$\frac{\partial C}{\partial t} = \nabla \cdot \left(c_{max}^{b-a} \left(\frac{\epsilon}{1-C} \right)^a C^b \nabla C \right) + K_3 \frac{SC}{K_2 + S} - K_4 C. \tag{2}$$

$$C(0,x) = \begin{cases} C_0, \text{ si } 0 \leq x \leq 4.5 \times 10^{-5}, \\ 0, \text{ si } 4.5 \leq x \leq L, \end{cases} \tag{3}$$

$$\frac{\partial C}{\partial x}(t,0) = 0, \ t \in [0,T], \tag{4}$$

$$\frac{\partial C}{\partial x}(t,L) = 0, \ t \in [0,T], \tag{5}$$

$$\tag{6}$$

$$S(0,x) = 1, \ x \in [0,L], \tag{7}$$

$$\frac{\partial S}{\partial x}(t,0) = 0, \ t \in [0,T], \tag{8}$$

$$\frac{\partial S}{\partial x}(t,L) = 0, \ t \in [0,T]. \tag{9}$$

siendo S la concentración de nutrientes y C la concentración de biomasa. El modelo, denominado M1, junto con condiciones de contorno e iniciales apropiadas se utiliza

para tratar de reproducir resultados que describan la dinámica de la cepa L1A1 de *Listeria monocytogenes*. La calibración paramétrica del modelo se realizó a través de técnicas de ajuste de datos incluidas en la herramienta AMIGO2 [5]. Los resultados obtenidos, Figura 3.3, muestran que la altura máxima alcanzada es muy cercana a la obtenida en los experimentos de laboratorio. Sin embargo, los nutrientes se consumen prácticamente en su totalidad, mientras que el desprendimiento masivo esperado en las últimas 24 horas no se aprecia.

Para tratar de obtener el desprendimiento masivo que aparece en las etapas finales de la cepa L1A1, introdujimos un termino de desprendimiento no lineal, en contraste con el término lineal presente en el modelo de Eberl, que representaba un desprendimiento por fricción que no ocurre en los experimentos de laboratorio utilizados como datos. Así pues, obtuvimos el modelo denominado M2, representado por

$$\frac{\partial S}{\partial t} = d_1 \nabla^2 S - K_1 \frac{SC}{K_2 + S}, \tag{10}$$

$$\frac{\partial C}{\partial t} = \nabla \cdot \left(c_{max}^{b-a} \left(\frac{\epsilon}{1 - C} \right)^a C^b \nabla C \right) + K_3 \frac{SC}{K_2 + S} - K_4 \frac{C}{K_d + S}, \tag{11}$$

que se completa con condiciones de contorno e iniciales (3)-(9) y una calibración paramétrica mediante AMIGO2. En los resultados obtenidos, Figura 3.3, se puede observar que se produce un desprendimiento significativo. Sin embargo, dicho desprendimiento empieza demasiado pronto y la altura máxima del biofilm no se llega a conseguir. Además, la concentración de nutrientes desaparece, como ocurre en el modelo de Eberl, aunque algo más despacio.

Así pues, tras analizar algunos candidatos más que incorporaban diferentes mecanismos [9], obtuvimos un nuevo modelo exitoso, denotado por M3. En él, cambiamos el coeficiente de difusión de nutrientes que deja de ser constante para depender de la concentración de biomasa, de forma que los nutrientes se difundan de manera diferente

en la región líquida y en la biopelícula. Así,

$$d_1(C) = \begin{cases} d_N, \text{ si } C = 0, \\ d_{eff} d_N, \text{ si } C > 0, \end{cases} \tag{12}$$

Al mismo tiempo, consideramos el coeficiente de difusión de la biomasa constante y un crecimiento de biomasa descrito por una ley de masas en lugar del habitual coeficiente de tipo Monod. Este cambio en el crecimiento de biomasa se debe a que consideramos que hay suficiente disponibilidad de nutrientes en el medio. Para asegurar que los nutrientes dejan de consumirse cuando se alcanza cierta concentración de biomasa, se inlcuye un mecanismo de bloqueo de consumo de nutrientes. Finalmente, asumimos un mecanismo de desprendimiento debido al envejecimiento del ADN extracelular. Todo ello da lugar al siguiente modelo 1D

$$\frac{\partial S}{\partial t} = \nabla \cdot (d_1(C) \nabla S) - K_1 SC, \tag{13}$$

$$\frac{\partial C}{\partial t} = d_2 \nabla^2 C + K_3 SC - K_4 \frac{C}{1 + \exp(k_d [D_{min} - CBD(t)])}, \tag{14}$$

que se completa con condiciones de contorno e iniciales (3)-(9) y una calibración paramétrica mediante AMIGO2. El término $CBD(t)$ se obtuvo mediante mediciones experimentales a ciertos tiempos y representa el area cubierta por células muertas o dañadas.

El modelo 1D obtenido prermite reproducir de manera fiable la dinamica de la cepa L1A1 con gran acuerdo entre los datos numéricos y los datos experimentales, tanto cualitativa como cuantitativamente. En la Figura 3.3 se puede observar el cambio drástico en la evolución de la biopelícula con los cambios realizados para el modelo 3. En primer lugar, el crecimiento de biomasa se ralentiza, alcanzando su máximo en torno a las 100 horas. Por otro lado, el consumo de nutrientes se reduce y el mecanismo de bloqueo evita que desaparezcan todos los nutrientes del sistema. También se puede observar el desprendimiento masivo que tiene lugar en la etapa final.

Los modelos 1D ofrecen buenos resultados para el caso de la cepa L1A1, pero no permiten reproducir otras estructuras complejas. Es por ello que es necesario desarrollar modelos en dimensiones superiores. En este aspecto, este trabajo presenta un modelo 2D modificando adecuadamente el modelo propuesto por Alpkvist y Klapper [1], de forma que se incluyen las nociones aprendidas del caso 1D. Así, obtenemos el modelo 2D adimensional

$$-\nabla^2 S = -V_1 \hat{h}_T^2 S \text{ en } \Omega^\tau_{\hat{H}_b}, \tag{15}$$

$$-\nabla^2 P = V_1 \hat{\Psi} S - V_1 \mathcal{F}_{\mathcal{D}}(V_2) \text{ en } \Omega^\tau_1, \tag{16}$$

$$\vec{U} = -\nabla P \text{ en } \Omega^\tau_1, \tag{17}$$

$$\partial_\tau \Phi + (\vec{U}_e \cdot \vec{n})||\nabla \Phi|| = 0 \text{ en } \Omega, \tag{18}$$

$$\partial_\tau V_1 - \nabla P \cdot \nabla V_1 = V_1 \left[\hat{\Psi} S - (\mathcal{F}_{\mathcal{D}}(V_2) + \epsilon_2) - V_1(\hat{\Psi} S - \mathcal{F}_{\mathcal{D}}(V_2)) \right] \text{ en } \Omega^\tau_1, \tag{19}$$

$$\partial_\tau V_2 - \nabla P \cdot \nabla V_2 = V_1 \epsilon_2 - V_2 \left(V_1 \hat{\Psi} S - V_1 \mathcal{F}_{\mathcal{D}}(V_2) \right) \text{ en } \Omega^\tau_1, \tag{20}$$

donde

$$\mathcal{F}_{\mathcal{D}}(V_2) = \frac{1}{1 + \exp(K_d(D_{min} - mean(V_2)))}, \tag{21}$$

siendo S la concentración de nutrientes, P la presión de expansión, $\vec{U}$ la velocidad de expansión, Φ la función *level set*, V_1 la concentración de biomasa activa y V_2 la concentración de biomasa inerte. Estas ecuaciones se resuelven en los dominios evolutivos Ω^τ_1 (región con biomasa) y $\Omega^\tau_{\hat{H}_b}$ (región de consumo de nutrientes), en los que se establecen condiciones de contorno apropiadas que, junto con la asunción de condiciones iniciales y los parámetros derivados del caso 1D, permiten obtener resultados de relevancia, que reproducen cualitativamente las fases iniciales de diversas cepas de *L. monocytogenes*.

El modelo 2D entraña una complejidad superior, por lo que en su resolución están involucradas técnicas numéricas avanzadas. En particular, este trabajo hace énfasis en el método *Level Set*, un método para seguir la evolución de fronteras libres y que permite, de forma sencilla, trabajar con la misma malla, pero con dominios

evolutivos. Además, otras técnicas numéricas eficientes han sido empleadas: método Crank-Nicolson para la discretización temporal, diferencias finitas modificadas con el método de Gibou [36] o diferencias finitas descentradas para la discretización espacial, método de Newton para trabajar con los términos no lineales, etc. Todas estas técnicas numéricas han sido implementadas desde cero en C++.

La validación cualitativa del modelo 2D se basa en diferentes trabajos presentes en la literatura:

- Para validar los resultados relativos a cepas planas de *L. monocytogenes*, empleamos lo aprendido en el caso 1D, incluyendo los parámetros calibrados correspondientes.

- El modelo de Alpkvist proporcionó como resultado una dinámica plana, pero demasiado rápida, que no se correspondía con el comportamiento esperado. El modelo modificado ralentiza esa dinámica dando lugar a resultados mucho más realistas.

- La manipulación de las condiciones iniciales y de los parámetros involucrados permite reproducir la dinámica de otras cepas de *L. monocytogenes*, dando lugar a estructuras con canales, agrupaciones o de tipo panal.

El presente trabajo se establece como una contribución novedosa al modelado y la simulación numérica de biopelículas. Por ello, hemos estudiado algunos modelos de biopelículas relevantes presentes en la literatura y analizado su calibración con los resultados experimentales obetnidos en el laboratorio. Es este análisis lo que motiva la introducción de nuevos modelos 1D y 2D que son capaces de reperesentar de forma más realista los resultados experimentales observados y las dinámicas esperadas de varios biofilms formados por cepas de *L. monocytogenes*. Del mismo modo, este trabajo puede entenderse como un modesto paso en la comprensión de los mecanismos que intervienen en el ciclo vital de ciertas biopelículas gracias al uso

de herramientas relevantes del modelado matemático y simulación numérica.

Así pues, el resultado de este trabajo es, partiendo de un modelo 1D presente en la literatura [26] y validado por resultados de análisis matemático [28], la obtención de un modelo 1D novedoso que refleja la evolución cualitativa y cuantitativa de la cepa L1A1 con gran precisión. Del mismo modo, aplicando los conocimientos adquiridos del caso unidimensional hemos obtenido un modelo 2D que permite reproducir las estructuras formadas en las fases de crecimiento por diferentes cepas de de *L. monocytogenes*, todo ello mediante técnicas numéricas refinadas y de gran eficiencia.

Este trabajo se estructura en dos partes.

En la Parte I nos centramos en el estudio y desarrollo de varios modelos unidimensionales. En particular, nos enfocamos en modelos continuos deterministas. Los modelos deterministas de reacción-difusión (RDM) tienen como ventaja la posibilidad de ser reproducidos [105]. Además, este tipo de modelos puede ser resuelto con técnicas numéricas avanzadas que garantizan la eficiencia computacional requerida para una identificación de modelo mediante técnicas de optimización[3, 102]. El Capítulo 1 sirve, por un lado, como introducción al modelado de biopelículas y, por otro, como presentación de varios posibles modelos de reacción-difusión para describir el sistema formado por biopelículas de *L. monocytogenes*. Los modelos candidatos incorporan diversos mecanismos de crecimiento, consumo de nutrientes y desprendimiento. Cada modelo se ajusta a los datos medidos con un ajuste de datos mediante técnicas de optimización. La elección del modelo óptimo se realiza atendiendo al mejor compromiso entre el número de parámetros desconocidos y su capacidad para reproducir cuantitativamente las mediciones experimentales.

El Capítulo 2 describe los métodos y técnicas numéricas y muestra los pasos necesarios para construir los sistemas numérios. El método de Crank-Nicolson en

diferencias finitas junto con un esquema centrado en espacio, resulta en el sistema numérico no lineal completamente discretizado, que se resuelve a través del sistema lineal obtenido tras aplicar un método de Newton. Estos métodos numéricos fueron elegidos de forma que el tiempo computacional sea reducido y la solución numérica sea eficientemente calculada, tal y como se discute en el trabajo reciente [7]. El Capítulo 2 muestra también la resolución de algunos tests académicos que permiten validar los métodos aplicados.

El Capítulo 3 muestra los resultados numéricos de los distintos modelos, así como una comparación entre los resultados numéricos obtenidos con los modelos y las mediciones experimentales correspondientes, de forma que se validan los modelos y los métodos numéricos utilizados desde un punto de vista biológico. El foco de la discusión es el modelo más satisfactorio, aunque otros candidatos intermedios fueron analizados también en el trabajo reciente [9].

La Parte II sigue un camino similar, enfocado en el estudio y desarrollo de varios modelos bidimensionales. Del mismo modo que en el caso 1D, trabajamos con modelos continuos deterministas, pero añadiendo las celulas dañadas o muertas como segunda especie biológica y sustituyendo el proceso de difusión por un proceso de advección que regula la evolución de la biopelícula. El Capítulo 4 sirve como explicación de la necesidad de usar modelos bidimensionales en el modelado de biopelículas para capturar las diferentes heterogeneidades espaciales que una biopelícula puede desarrollarse. Tras esto, presenta los diferentes modelos estudiados y desarrollados.

El Capítulo 5 muestra la estrategia numérica usada en la resolución de los modelos bidimensionales propuestos. Dado el aumento en la complejidad del sistema, son necesarios nuevos métodos numéricos con mayor eficiencia y sofisticación. Aunque nos mantenemos en el marco de las diferencias finitas, utilizamos un método *upwind*

modificado mediante la estrategia propuesta por Gibou et al. [36] junto con el método de *Level Set* [89], esquemas WENO [48], esquemas Crank-Nicolson y un algoritmo de Newton para los sistemas completamente discretos no lineales.

Por último, el Capítulo 6 muestra los resultados numéricos obtenidos en los diferentes modelos, analizando su precisión de forma que sean validados desde un punto de vista biológico.

El Anexo A resume algunos resultados del análisis teórico del modelo de Eberl que aparecen el el trabajo reciente de Efendiev, Eberl y Zelik [28].
El Anexo B presenta una explicación breve de las técnicas de optimización incluidas en la herramienta AMIGO2 [5], una herramienta de alto nivel para MATLAB utilizada en el proceso de calibración de parámetros.

Finalmente, se incluye un apartado de Conclusiones que aglutina los resultados más importantes de la tesis, las conclusiones más relevantes y la lineas futuras a seguir.

Resumo extenso

As biopelículas bacterianas son esenciais tanto para o equilibrio dos ecosistemas como para moitos procesos industriais. Unha biopelícula bacteriana é unha capa de células procariotas ancoradas a unha superficie que lles proporciona substrato, manténdoas alimentadas. Esta capa de células está integrada nunha matriz formada por polímeros, chamada matriz de exopolisacáridos ou EPS, que manten as bacterias agrupadas e dificulta a súa eliminación[21, 108]. Dun xeito sinxelo, podemos dicir que unha biopelícula é unha agrupación de microorganismos ancorados a unha superficie. Os estudos actuais estiman que menos dun 0.1% da vida acúatica microbiana atópase en estado planctónico (flotando libremente) [20]. Polo tanto, as biopelículas constitúen a forma de vida preferida das bacterias.

A razón desta preferencia é a vantaxe competitiva que supón a habilidade de adherirse a superficies e formar biopelículas, en comparación coas bacterias en estado planctónico. Estas últimas poden ser varridas con facilidade polo fluxo da auga, mentres que as primieras están protexidas dese fenómeno e viven nunha contorna onde poden crecer, sempre que haxa suficiente dispoñibilidade de nutrientes. A estrutura física dunha biopelícula tamén permite a aparición de diferentes nichos biolóxicos que facilitan o crecemento e a supervivencia de microorganismos que non poderían competir nun sistema completamente homoxéneo. Máis aínda, a actividade microbiana no interior dunha biopelícula pode cambiar a contorna interna, facendo que a biopelícula sexa máis hospitalaria que a rexión líquida. As partes fundamentáis dunha biopelícula son a superficia (á que se adhiren as

bacterias), a biopelícula propiamente dita (formada por unha ou máis especies de microorganismos e os EPS), o medio (cos nutrientes) e as condicións da contorna que determinan o desenvolvemento da biopelícula (temperatura, pH, hidrodinámica, etc.).

Existen biopelículas beneficiosas, tanto para os humanos como para o desenvolvemento axeitado do medio ambiente, así como biopelículas prexudiciais, que adoitan causar problemas de saúde ou baixos rendementos en procesos industriais e, polo tanto, perdas económicas. Exemplos [108] das primeiras son as biopelículas empregadas nas plantas de tratamento de augas (RBC, reactores biolóxicos, etc.) ou na industria fermenteira (por exemplo, en procesos de *quick-vinegar*), pero tamén algunhas biopelículas creadas de xeito natural presentes no subsolo (contribuíndo á descontaminación do solo ou en augas subterráneas), en ríos, lagos e zonas costeiras (colonizando as rochas ou suspendidas na auga, normalmente contribuíndo á eliminación de contaminantes na auga), ou nas raíces de moitas plantas (aumentando a disponibilidade de nutrientes para as plantas). As biopelículas creadas de xeito natural son fundamentáis para a biosfera terrestre.

Por outra banda, as biopelículas prexudiciáis aparecen en moitas situacións. Por exemplo, as biopelículas son un problema considerable na hixiene dental [61]. Tamén son a causa de infeccións en implantes médicos e de complicacións en enfermidades infecciosas [115]. Outros exemplos son as biopelículas que causan contaminación da auga ou mal funcionamento dos intercambiadores de calor [11]. Especialmente relevantes son as biopelículas de bacterias patóxenas que aparecen na industria alimentaria, pois constitúen unha fonte importante de contaminación alimentaria e poden ser un grave problema de saúde para os consumidores [93].

Previr a formación de biopelículas é, en xeral, bastante difícil, debdo á súa habilidade para desenvolverse incluso en condicións adversas. Máis aínda, unha vez formadas, son difíciles de eliminar, xa que as bacterias que están formando biopelículas son

moito máis resistentes á resposta inmune do hóspede ou a axentes antimicrobianos [34, 44].

A necesidade de mellorar as propiedades das biopelículas beneficiosas ou controlar a formación daquelas prexudiciais motivou numerosas liñas de investigación que teñen en conta os mecanismos xenéticos, bioquímicos ou físicos que contribuen, non só á formación de biopelículas, se non tamén á súa estrutura.

Numerosos estudos indican que a estrutura dunha biopelícula determina a magnitude dos procesos que teñen lugar nela, como por exemplo a taxa de transferencia de nutrientes cara ás capas interiores, a taxa de difusión de axentes microbianos ou a súa resistencia á fricción. O reconto de placas foi amplamente empregado para analizar biopelículas. Con todo, só inclúe células cultivables viables e non da información da estrutura [22]. Debido a esas limitacións, varias alternativas foron propostas.

Prestouse especial énfase ao desenvolvemento de varias técnicas microscópicas. A máis exitosa, a microscopía confocal láser de varrido (CLSM), permite realizar escaneo óptico tridimensional *in situ* e *in vivo* [95]. Por outra banda, os microsensores poden ser empregados para medir concentracións dos diferentes componentes dentro dunha biopelícula, permitindo así a análise da dispoñibilidade de nutrientes nas diferentes rexións da biopelícula [116]. Os avances en bioloxía molecular e nas técnicas de hibridación *in situ* contribuiron ao desenvolvemento de sondas xenéticas e ténicas de microscpía, permitindo a análise detallada de comunidades microbianas no interior da biopelícula [23, 56, 97].

A combinación de tinturas celulares empregadas en ténicas de tinguidura celular fluorescente xunto coas técnicas de obtención de imaxes permiten localizar células viables e células danadas ou mortas [99] ou a distribución da substancia polimérica extracelular, pero tamén a reconstrucción de estruturas tridimensionais, obtendo

así un estudo das biopelículas máis exhaustivo. Cabe desetacar que a técnica de CLSM xunto coa análise cuantitativa de imáxes [114] permite determinar dun xeito automático o grosor dunha biopelícula, o seu volumen ou a rugosidade. Este enfoque permite unha comparación cualitativa de biopelículas de diferentes cepas e especies ou con diferentes condicións ambientais (medio, temperatura, tipo de cultivo).

Algúns traballos recentes suxiren varias estratexias de traballo e ferramentas de *software* para a análise sitemática de imáxes de microscopía. IMARIS (un *software* comercial) permite a reconstrucción de estruturas 3D. COMSTAT [43], ISA 3D [10], ou PHLIP [67] permiten cuantificar as imaxes obtidas mediante CLSM. BIOFILM-DIVER [65] permite a cuantificación da porosidade dunha biopelícula, a área cuberta, as distancias de difusión ou a distribución de poboación espacio-temporal mediante imáxes 2D tomadas por epifluorescencia e CLSM. Algoritmos tipo *machine learning* poden ser empregados para analizar as imaxes obtidas por microscopía electrónica de varrido [103].

A análise de imáxe cuantitativa podese complementar co modelado matemático para gañar coñecemento acerca dos mecanismos que dan lugar a unha determinada estrutura nunha biopelícula. Incluso as biopelículas máis homoxéneas desenvolven procesos internos complexos, entretecidos entre si. O modelado matemático dá a posibilidade de explorar diferentes procesos internos e as súas conexións, a súa importancia relativa e o rol que ten o medio ambiente no ciclo vital dunha biopelícula [105]. Ese é, precisamente, o paso que dá o presente traballo, ao combinar a análise de imáxes obtidas mediante CLSM con novos modelos matemáticos e técnicas numéricas avanzadas para explicar cuantitativa e cualitativamente o comportamento de biopelículas de *Listeria monocytogenes*.

As biopelículas de *L. monocytogenes* poden chegar a desenvolver diferentes estruturas: mono-capas de células adheridas, multi-capas chás desestruturadas, ou

estruturas en panal ou *clusters* [12, 15, 24, 62, 78, 82]. Traballos recentes [37] reconstruiron imáxes obtidas por CLSM para observar a diversidade intraespecie de *L. monocytogenes*, concluindo que a maioría das cepas forman estruturas tipo panal complexas ás 48h. Outros traballos, [65, 66] usan a análise cuantitativa de imáxes para estudar o ciclo vital de biopelículas formadas por tres cepas de textitL. monocytogenes, amosando a presenza de, polo menos, tres fases: unha fase inicial de clusters separados que avolucionan a clusters intercoñectados, estruturas tipo panal ou estruturas chás, e unha fase final na que as células sepáranse da biopelícula. A duración destas fases varía significativamente entre cepas.

Nunha primeira aproximación, neste traballo propomos modelos 1D continuos de reacción-difusión que permiten a análise da cepa L1A1 de *L. monocytogenes*. A devandita cepa forma biopelículas chás [66], polo que semella axeitado empregar modelos 1D para explorar os mecanismos relevantes que causen este tipo de estruturas. Os modelos aquí presentados foron resoltos numéricamente empregando técnicas numéricas eficientes e robustas, que consisten fundamentalmente nun método Crank-Nicolson para a discretización en tempo e esquemas en diferenzas finitas axeitados para a discretización en espazo, todo iso xunto cun método de Newton alí onde aparecen termos non lineais. Os métodos numéricos foron implementados desde cero en C++ e validados numéricamente mediate tests académicos axeitados.

O estudo baséase na disposición experimental empregada no traballo recente [66] que considera biopelículas que crecen en condicións de fluxo estático, restrinxindo o estudo ao caso hidrostático. Atendendo ao devandito estudo, as cepas de L1A1 desenvolven biopelículas practicamente chás e desestruturadas. Despois dunha sedimentación inicial, fórmase unha biopelícula delgada antes de 24 horas. A estrutura chá é estable, cun aumento sostido do seu grosor ata chegar ás 96 horas, momento no cal a presenza de células danadas our mortas é imporante. Tras 120 horas, a altura máxima da biopelícula decrece fortemente, indicando un desprendemento masivo nas

derradeiras 24 horas.

As células de L1A1 consomen glucosa como fonte primaria de carbono. Polo tanto, os datos de nutrientes considerados na resolución numérica corresponde a glucosa. Os datos experimentais amosan que as células consome a maior parte das reservas de glucosa nas primeras 24 horas, tras o cal o consumo de nutrientes cesa, mesmo que aínda se atopen no sistema células viables.

Comezamos o noso estudo empregando como punto de partida o modelo presentado por Eberl et al. [26] e validado por resultados da análise matemática [28],

$$\frac{\partial S}{\partial t} = d_1 \nabla^2 S - K_1 \frac{SC}{K_2 + S}, \tag{1}$$

$$\frac{\partial C}{\partial t} = \nabla \cdot \left(c_{max}^{b-a} \left(\frac{\epsilon}{1-C} \right)^a C^b \nabla C \right) + K_3 \frac{SC}{K_2 + S} - K_4 C. \tag{2}$$

$$C(0,x) = \begin{cases} C_0, \text{ se } 0 \leq x \leq 4.5 \times 10^{-5}, \\ 0, \text{ se } 4.5 \leq x \leq L, \end{cases} \tag{3}$$

$$\frac{\partial C}{\partial x}(t,0) = 0, \ t \in [0,T], \tag{4}$$

$$\frac{\partial C}{\partial x}(t,L) = 0, \ t \in [0,T], \tag{5}$$

$$\tag{6}$$

$$S(0,x) = 1, \ x \in [0,L], \tag{7}$$

$$\frac{\partial S}{\partial x}(t,0) = 0, \ t \in [0,T], \tag{8}$$

$$\frac{\partial S}{\partial x}(t,L) = 0, \ t \in [0,T]. \tag{9}$$

sendo S a concentración de nutrientes e C a concentración de biomasa. O modelo, denominado M1, xunto coas condicións de contorna e iniciais axeitadas, emprégase para tratar de reproducir resultados que describan a dinámica da cepa L1A1 de

Listeria monocytogenes. A calibración paramétrica do modelo realizouse ao través de técnicas de axuste de datos incluidas na ferramenta AMIGO2 [5]. Os resultados obtidos, Figura 3.3, amosan que a altura máxima alcanzada é moi próxima á obtida nos experimentos de laboratorio. Con todo, os nutrientes consómense practicamente na súa totalidade, mentres que o desprendemento masivo esperado nas derradeiras 24 horas non se aprecia.

Para tratar de obter o desprendemento masivo que aparece nas estapas finais da cepa L1A1, introducimos un termo de desprendemento non lineal, en contraste co termo lineal presente no modelo de Eberl, que representaba un desprendemento por fricción que non ocorre nos experimentos de laboratorio empregados como datos. Así pois, obtivemos o modelo denominado M2, representado por

$$\frac{\partial S}{\partial t} = d_1 \nabla^2 S - K_1 \frac{SC}{K_2 + S}, \tag{10}$$

$$\frac{\partial C}{\partial t} = \nabla \cdot \left(c_{max}^{b-a} \left(\frac{\epsilon}{1-C} \right)^a C^b \nabla C \right) + K_3 \frac{SC}{K_2 + S} - K_4 \frac{C}{K_d + S}, \tag{11}$$

que se completa con condicións de contorna e iniciais (3)-(9) e unha calibración paramétrica mediante AMIGO2. Nos resultados obtidos, Figura 3.3, pódese observar que se produce un desprendemento significativo. Con todo, o devandito desprendemento comeza moi cedo e a altura máxima do biofilm non se chega a conseguir. Ademáis, a concentración de nutrientes desaparece, como ocorre no modelo de Eberl, ainda que algo máis amodo.

Así pois, tras analizar algúns candidatos máis que incorporaban diferentes mecanismos [9], obtivemos un novo modelo exitoso, denotado M3. Nel, cambiamos o coeficiente de difusión de nutrientes que deixa de ser constante para depender da concentración de biomasa, de xeito que os nutrientes difúndense de maneira diferente na rexión líquida e na biopelícula. Así,

$$d_1(C) = \begin{cases} d_N, \text{ se } C = 0, \\ d_{eff} d_N, \text{ se } C > 0, \end{cases} \tag{12}$$

Ao mesmo tempo, consideramos o coeficiente de difusión da biomasa constante e un crecemento da biomasa descrito por unha lei de masa no lugar do habitual coeficiente de topo Monodo. Este cambio no crecemento da biomasa débese a que consideramos que hai suficiente dispoñibilidade de nutrientes no medio. Para asegurar que os nutrientes deixan de consumirse cando se alcanza certa concentración de biomasa, inclúese un mecanismo de bloque de consumo de nutrientes. Finalmente, asumimos un mecanismo de desprendemento debido ao envellecemento do ADN extracelular. Todo iso dá lugar ao seguinte modelo 1D

$$\frac{\partial S}{\partial t} = \nabla \cdot (d_1(C)\nabla S) - K_1 SC, \tag{13}$$

$$\frac{\partial C}{\partial t} = d_2 \nabla^2 C + K_3 SC - K_4 \frac{C}{1 + \exp(k_d[D_{min} - CBD(t)])}, \tag{14}$$

que se completa con condicións de contorna e iniciais (3)-(9) e unha calibración paramétrica mediante AMIGO2. O termo $CBD(t)$ foi obtido mediante medicións experimentais a certos tempos e representa a área cuberta por células mortas ou danadas.

O modelo 1D obtido permite reproducir dun xeito fiable a dinámica da cepa L1A1 con gran acordo entre os datos numéricos e os datos experimentais, tanto cualitativa como cuantitativamente. Na figura 3.3 pódese observar o cambio drástico na evolución da biopelícula cos cambios realizados para o modelo 3. No primeiro lugar, o crecemento da biomasa retárdase, alcanzando o seu máximo ao redor das 100 horas. Por outra banda, o consumo de nutrientes redúcese e o mecanismo de bloqueo evita que desaparezan todos os nutrientes do sistema. Tamén se pode observar o desprendemento masivo que ten lugar na etapa final.

Os modelos 1D ofrecen bos resultados para o caso da cepa L1A1, pero non permiten reproducir outras estruturas complexas. É por iso que é necesario desenvolver modelos en dimensións superiores. Neste aspecto, este traballo presenta un modelo 2D modificando axeitadamente o modelo proposto por Alpvist e Klapper [1], de forma que

se inclúen as nocións aprendidas no caso 1D. Así, obtemos o modelo 2D adimensional

$$-\nabla^2 S = -V_1 \hat{h}_T^2 S \text{ en } \Omega^\tau_{\hat{H}_b}, \quad (15)$$

$$-\nabla^2 P = V_1 \hat{\Psi} S - V_1 \mathcal{F}_{\mathcal{D}}(V_2) \text{ en } \Omega^\tau_1, \quad (16)$$

$$\vec{U} = -\nabla P \text{ en } \Omega^\tau_1, \quad (17)$$

$$\partial_\tau \Phi + (\vec{U_e} \cdot \vec{n})||\nabla \Phi|| = 0 \text{ en } \Omega, \quad (18)$$

$$\partial_\tau V_1 - \nabla P \cdot \nabla V_1 = V_1 \left[\hat{\Psi} S - (\mathcal{F}_{\mathcal{D}}(V_2) + \epsilon_2) - V_1(\hat{\Psi} S - \mathcal{F}_{\mathcal{D}}(V_2))\right] \text{ en } \Omega^\tau_1, \quad (19)$$

$$\partial_\tau V_2 - \nabla P \cdot \nabla V_2 = V_1 \epsilon_2 - V_2 \left(V_1 \hat{\Psi} S - V_1 \mathcal{F}_{\mathcal{D}}(V_2)\right) \text{ en } \Omega^\tau_1, \quad (20)$$

onde

$$\mathcal{F}_{\mathcal{D}}(V_2) = \frac{1}{1 + \exp(K_d(D_{min} - mean(V_2)))}, \quad (21)$$

sendo S a concentración de nutrientes, P a presión de expansión, $\vec{U}$ a velocidade de expansión Φ a función *level set*, V_1 a concentración de biomasa activa e V_2 a concentración de biomasa inerte. Estas ecuacións resólvense nos dominios evolutivos Ω^τ_1 (rexión con biomasa) e $\Omega^\tau_{\hat{H}_b}$ (rexión de consumo de nutrientes), nos que se establecen condicións de contorna axeitadas que, xunto coa asunción de condicións iniciais e parámetros derivados do caso 1D, permiten obter resultados de relevancia que reproducen cualitativamente as fases iniciais de diversas cepas de *L. monocytogenes*.

O modelo 2D entraña unha complexidade superior, polo que na súa resolución están involucradas técnicas numéricas avanzadas. En particular, este traballo fai énfase no método *Level Set*, un método para seguir a evolución de fronteiras libres e que permite, dun xeito sinxelo, traballar coa mesma malla, pero con dominios evolutivos. Ademais, outras técnicas numéricas eficientes foron empregadas: método Crank-Nicolson para a discretización temporal, diferencias finitas modificadas co método de Gibou [36] ou diferencias finitas descentradas para a discretización espacial, método de Newton para traballar cos termos no lineais, etc. Todas estas técnicas numéricas foron implementadas desde cero en C++.

A validación cualitativa do modelo 2D baséase en diferentes traballos presentes na literatura:

- Para validar os resultados relativos a cepas chás de *L. monocytogenes*, empregamos o aprendido no caso 1D, incluíndo os parámetros calibrados correspondentes.
- O modelo de Alpkvist proporcionou como resultado unha dinámica chá, pero demasiado rápida, que non se correspondía co comportamento esperado. O modelo modificado retarda esa dinámica dando lugar a resultados moito máis realistas.
- A manipulación das condicións iniciais e dos parámetros involucrados permite reproducir a dinámica de outras cepas de *L. monocytogenes* dando lugar a estruturas con canles, agrupacións ou de tipo panal.

O presente traballo estabécese como unha contribución nova ao modelado e a simulación numérica de biopelículas. Por iso, estudamos algúns modelos de biopelículas relevantes presentes na literatura e analizado a súa calibración cos resultados experimentais obtidos no laboratorio. É esta análise a que motiva a introducción de novos modelos 1D e 2D que son capaces de reperesentar dun xeito máis realista os resultados experimentais observados e as dinámicas esperadas de varios biofilms formados por cepas de *L. monocytogenes*. Do mesmo xeito, este traballo podese entender como un modesto paso na compresión dos mecanismos que interveñen no ciclo vital de certas biopelículas grazas ao uso de ferramentas relevantes do modelado matemático e simulación numérica.

Así pois, o resultado deste traballo é, partindo dun modelo 1D presente na literatura [26] e validado por resultados da análise matemática [28], a obtención dun modelo 1D novo que reflicte a evolución cualitativa e cuantitativa da cepa L1A1 con gran precisión. Do mesmo xeito, aplicando os coñecementos adquiridos do caso unidimensional, obtivemos un modelo 2D que premite reproducir as estruturas

formadas nas fases de crecemento por diferentes cepas de *L. monocytogenes*, todo iso mediante técnicas numéricas refinadas e de gran eficiencia.

Este traballo estruturase en dúas partes

Na Parte I centrámonos no estudo e desenvolvemento de varios modelos unidimensionais. En particular, enfocámonos nos modelos continuos deterministas. Os modelos deterministas de reacción-difusión (RDM) teñen como vantaxe a posibilidade de ser reproducidos [105]. Ademais, este tipo de modelos pode ser resolto con técnicas numéricas avanzadas que garanten a eficiencia computacional requerida para unha identificación do modelo mediante técnicas de optimización [3, 102]. O Capítulo 1 serve, por unha banda, como introducción ao modelado de biopelículas e, por outra, como presentación de varios pobles modelos de reacción-difusión para describir o sistema formado por biopelíclas de *L. monocytogenes*. Os modelos candidatos incorporan diversos mecanismos de crecemento, consumo de nutrientes e desprendemento. CAda modelo axustase aos datos medidos cun axuste de datos mediante técnicas de optimización. A elección do modelo óptimo realízase atendendo ao mellor compromismo entre o núimero de parámetros descoñecidos e a súa capacidades para reproducir cuantitativamente as medicións experimentais.

O Capítulo 2 describe os métodos e técnicas numéricas e amosa os pasos necesarios para construir os sistémas numéricos. O método de Crank-Nicolson en diferenzas finitas xunto cun esquema centrado en espacio, resulta no sistema numérico non lineal completamente discretizado, que se resolve a través do sistema lineal obtido tras aplicar un método de Newton. Estes métodos numéricos foron elixidos para que o tempo computaciónal sexa reducido e a solución numérica sexa eficiemtemente calculada, tal e como se discute no traballo recente [7]. O capítulo 2 amosa tamén a resolución dalgúns tests académicos que permiten validar os métodos aplicados.

O Capítulo 3 amosa os resultados numéricos dos distintos modelos, así como unha comparación entre os resultados numéricos obtidos cos modelos e as medicións experimentais correspondentes, de xeito que se validen os modelos e os métodos numéricos empregados desde o punto de vista biolóxico. O foco da discusión é o modelo máis satisfactorio, ainda que outros candidatos intermedios foron tamén analizados no traballo recente [9].

A Parte II segue un camiño similar, enfocado no estudo e desenvolvemento de varios modelos bidimensionais. Do mesmo xeito que no caso 1D, traballamos con modelos continuos deterministas, pero engadindo as celulas danadas o mortas como segunda especie biolóxica e substituíndo o proceso de difusión por un proceso de advección que regula a evolución da biopelícula. O Capítulo 4 serve como explicación da necesidade de empregar modelos bidimensionais no modelado de biopelículas para capturar as diferentes heteroxeneidades espaciais que unha biopelícula pode desenvolver. Tras isto, presenta os diferentes modelos estudados e desenvolvidos.

O Capítulo 5 amosa a estratexia numérica empregada na resolución dos modelos bidimensionais propostos. Dado o aumento na complexidade do sistema, son necesarios novos métodos numéricos con maior eficiencia e sofisticación. Ainda que nos mantemos no marco das diferenzas finitas, empregamos un método *upwind* modificado mediante a estratexia proposta por Gibou et al. [36] xunto co método de *Level Set* [89], esquemas WENO [48], esquemas Crank-Nicolson e un algoritmo de Newton para os sistemas completamente discretos non lineais.

Para rematar, o Capítulo 6 amosa os resultados numéricos obtidos nos diferentes modelos, analizando a súa precisión de xeito que sexan validados desde o punto de vista biolóxico.

O Anexo A resume alguns resultados da análise teórica do modelo de Eberl que

aparecen no traballo recente de Efendiev, Eberl e Zelik [28].
O Anexo B presenta unha explicación breve das técnicas de optimización incluídas na ferramenta AMIGO2 [5], unha ferramenta de alto nivel para MATLAB empregada no proceso de calibración de parámetros.

Finalmente, inclúese un apartado de Conclusións que aglutina os resultados máis importantes da tese, as conclusión máis relevantes e as liñas futuras a seguir.

Bibliography

[1] Alpkvist E, Klapper I. A Multidimensional Multispecies Continuum Model for Heterogeneous Biofilm Development. *Bulletin of Mathematical Biology*, 69 (2007), 765-789.

[2] Alpkvist E, Picioreanu C, van Loosdrecht M C M, Heyden A. Three-dimensional biofilm model with individual cells and continuum EPS matrix. *Biotechnology and Bioengineering*, 94 (2006), 961–979.

[3] Balsa-Canto E, Alonso A, Banga J. An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Systems Biology*, 4:11 (2010).

[4] Balsa-Canto E, Alonso A A, Arias-Méndez A, García M R, López-Núñez A, Mosquera-Fernández M, Vázquez C, Vilas C. Modeling and Optimization Techniques with Applications in Food Processes, Bio-processes and Bio-systems. In: Higueras I., Roldán T., Torrens J. (eds) Numerical Simulation in Physics and Engineering. *SEMA SIMAI Springer Series*, 9 (2016).

[5] Balsa-Canto E, Henriques D, Gabor A, Banga J R. AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. *Bioinformatics* 32 (2016), 3357–3359.

[6] Balsa-Canto E, Henriques D, Gabor A, Banga J R. AMIGO2, theoretical background. *AMIGO2 toolbox documentation*, sites.google.com/site/amigo2toolbox/doc

[7] Balsa-Canto E, López-Núñez A, Vázquez C. Numerical methods for a non-linear reaction-diffusion system modelling a batch culture of biofilm. *Applied Mathematical Modelling*, 41 (2017), 164–179.

[8] Balsa-Canto E, Rodriguez-Fernandez M, Alonso A A, Banga J R. Computational design of optimal dynamic experiments in systems biology: a case study in cell signaling. In Cánovas M, Iborra J L and Manjón A, editors, *Understanding and Exploiting Systems Biology in Bioprocesses and Biomedicine*, (2006), 103-117. Fundación Caja Murcia.

[9] Balsa-Canto E, Vilas C, López-Núñez A, Mosquera-Fernández M, Briandet R, Cabo M L, Vázquez C. Modeling reveals the role of aging and glucose uptake impairment in L1A1 *Listeria monocytogenes* biofilm life cycle, *Frontiers In Microbiology*, 8 (2017), 2118.

[10] Beyenal H, Lewandowski Z, Harkin G. Quantifying biofilm structure: facts and fiction. *Biofouling*, 20 (2004),1-23.

[11] Bott T.R. Fouling of heat exchangers. *Elsevier Science*, 1995.

[12] Bridier A, Dubois-Brissonnet F, Boubetra A, Thomas V, Briandet R. The biofilm architecture of sixty opportunistic pathogens deciphered using a high throughput CLSM method. *Journal of Microbiological Methods*, 82 (2010), 64–70.

[13] Carpentier B, Cerf P. Persistence of *Listeria monocytogenes* in food industry equipment and premises. *International Journal of Food Microbiology*. 145 (2011), 1–8.

[14] Center for Disease Control and Prevention (CDC), 2014. National Center for Emerging and Zoonotic Infectious Diseases (NCEZID), Division of Foodborne,Waterborne, and Environmental Diseases (DFWED). (Available at http//www.cdc.gov/Listeria/outbreaks/ index.html. Acceded 10 March 2014).

[15] Chae M S, Schraft H. Comparative evaluation of adhesion and biofilm formation of different *Listeria monocytogenes* strains. *International Journal of Food Microbiology*, 62 (2000), 103–111.

[16] Chopp L D. Some improvements of the Fast Marching Method. *SIAM Journal on Scientific Computing*, 23 (2001), 230-244.

[17] Chopp L D. Another look at velocity extensions in the Level Set method. *SIAM Journal on Scientific Computing*, 31 (2009), 3255-3273.

[18] Chopp L D. Numerical Methods for Moving Interfaces. *ESAM 449 course notes*, 2017.

[19] Cossart P, Lebreton A. A trip in the "new microbiology" with the bacterial pathogen *Listeria monocytogenes*. *FEBS Letters*, 588 (2014), 2437–2445.

[20] Costerton J W, Lewandowski Z, Caldwell D E, Korber D R, Lappin-Scott H M. Microbial biofilms. *Annual Review Microbiology*, 49 (1995), 711-745.

[21] Costerton J W, Cheng K J, Geesey G G, Ladd T I, Nickel J C, Dasgupta M, Thomas J M. Bacterial biofilms in nature and disease. *Annual Review of Microbioliogy*, 41 (1987), 435–464.

[22] Daims H, Wagner M. Quantification of uncultured microorganisms by fluorescence microscopy and digital image analysis. *Applied Microbiology and Biotechnology*, 75 (2007). 237-248.

[23] DeBeer D, Stoodley P, Lewandowski Z. Measurement of local diffusion coefficients in biofilms by microinjection and confocal microscopy. *Biotechnology and Bioengineering*, 53 (1997), 151-158.

[24] Djordjevic D, Wiedmann M, McLandsborough L. Microtiter plate assay for assessment of *Listeria monocytogenes* biofilm formation. *Applied Environmental Microbiology*, 68 (2002), 2950–2958.

[25] Dockery J D, Klapper I. Finger formation in biofilm layers. *SIAM Journal of Applied Mathematics*, 62 (2001), 853–869.

[26] Eberl H J, Parker D F, van Loosdrecht M C M. A New Deterministic Spatio-Temporal Continuum Model for Biofilm Development. *Journal of Theoretical Medicine*, 3 (2001), 161-175.

[27] Eberl H J, Picioreanu C, van Loosdrecht M C M, Heijnen J J. A three-dimensional numerical study on the correlation of spatial structure, hydrodynamic conditions, and mass transfer and conversion in biofilms. *Chemical Engineering Science*, 55 (2000), 6209-6222.

[28] Efendiev M A , Eberl H J, and Zelik S V. Existence and longtime behavior of solutions of a nonlinear reaction-diffusion system arising in the modeling of biofilms. *Nonlinear Diffusive Systems and Related Topics*, 1258 (2002), 49-71.

[29] Egea J A, Balsa-Canto E, Garcia M S G, Banga J R. Dynamic optimization of nonlinear processes with an enhance scatter search. *Industrial and Engineering Chemistry Research*, 48 (2009), 4388-4401.

[30] Epperson J F. *An introduction to numerical methods and analysis*. Edición revisada. John Wiley & Sons, 2007.

[31] European Food Safety Authority (EFSA), European Centre for Disease Prevention and Control (ECDC). The European Union summary report on trends and sources of zoonoses and zoonotic agents and food-borne outbreaks in 2011. *EFSA Journal*, 11 (2013), 3129.

[32] European Food Safety Authority (EFSA), European Centre for Disease Prevention and Control (ECDC). The European Union summary report on trends and sources of zoonoses and zoonotics agents and food-borne outbreaks in 2012. *EFSA Journal*, 12 (2014), 3547.

[33] European Food Safety Authority (EFSA), European Centre for Disease Prevention and Control (ECDC). European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. *EFSA Journal*, 15 (2017). Data on case numbers come from the European Surveillance System (TESSy), data on food come from EFSA zoonoses database.

[34] Fux C A, Costerton J W, Stewart P S, Stoodley P. Survival strategies of infectious biofilms. *Trends in Microbiology*, 13 (2005), 34–40.

[35] Ghanbari A, Dehghany J, Schwebs T, Müsken M, Häussler S, Meyer-Hermann M. Inoculation density and nutrient level determine the formation of mushroom-shaped structures in Pseudomonas aeruginosa biofilms. *Scientific Reports*, 6 (2016), Artilce number 32097.

[36] Gibou F, Fedkiw R, Cheng L T, Kang M. A second order accurate symmetric discretization of the Poisson equation on irregular domains. *Journal of Computational Physics*, **176** (1), 205–227.

[37] Guilbaud M, Piveteau P, Desvaux M, Brisse S, Briandet R. Exploring the diversity of *Listeria monocytogenes* biofilm architecture by high-throughput confocal laser scanning microscopy and the predominance of the honeycomb-like morphotype. *Applied and Environmental Microbiology*, 81 (2015), 1804–1810.

[38] Gurtin M E. *An Introduction to Continuum Mechanics*. Academic Press, 1981.

[39] Harremoës P. The significance of pore diffusion to filter dentrification. *Journal Water Pollution Control Federation*, 48 (1976), 377-388.

[40] Harris N P, Hansford G S. A study of substrate removal in a microbial film reactor. *Water Research*, 10 (1976), 935-943.

[41] Harten A, Engquist B, Osher S, Chakraverthy S. Uniformly high-order accurate essentially non-oscillatory schemes III. *Journal of Computational Physics*, **71**, 231-303, 1987

[42] Hermanowicz S W. Two-dimensional simulations of biofilm development: effects of external environmental conditions. *Water Science and Technology*, 39 (1999), 107-114.

[43] Heydorn A, Nielsen A T, Hentzer M, Sternberg C, Givskov M, Ersboll B K, et al. Quantification of biofilm structures by the novel computer program COMSTAT. *Microbiology*, 146 (2000), 2395-2407.

[44] Hoiby N, Bjarnsholt T, Givsko M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. *International Journal of Antimicrobial Agents* 35 (2010), 322–332.

[45] Horn H, Lackner S. Modeling of biofilm systems: a review. *Advances in Biochemical Engineering/Biotechnology*, 146 (2014), 53-76.

[46] Hunt S, Hamilton M, Sears J, Harkin G, Reno J. 2003. A computer investigation of chemically mediated detachment in bacterial biofilms. *Microbiology*, 149 (5), 1155-1163.

[47] James G, Beaudette L, Costerton J. Interspecies bacterial interactions in biofilms. *Journal of Industrial Microbiology*, 15 (1995), 257–262.

[48] Jiang G-S, Peng D. Weighted ENO Schemes for Hamilton-Jacobi Equations. *SIAM Journal of Scientific Computing*, **21**, 2126-2143, 2000.

[49] Joshi M, Seidel-Morgenstern A, Kremling A. Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. *Metabolic Engineering*, 8 (2006), 447-455.

[50] Kathariou S. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. *Journal of Food Protection*, 65 (2002), 1811–1829.

[51] Kissel J C, McCarty P L, Street R L. Numerical Simulation of Mixed-Culture Biofilm. *Journal of Environmental Engineering*, 110 (1984), 393-411.

[52] Klausen M, Aases-Jorgensen A, Molin S, Tolker-Nielsen T. Involvement of bacterial migration in the development of complex multicellular structures in Pseudomonas aeruginosa biofilms. *Molecular Microbiology*, 50 (2003), 61–68.

[53] Kreft J, Picioreanu C, Wimpenny J, Loosdrecht M. 2001. Individual-based modelling of biofilms. *Microbiology*, 147 (11), 2897-2912.

[54] Kreft J, Wimpenny J. 2001. effect of EPS on biofilm structure and function as revealed by an individual-based model of biofilm growth. *Water, Science and Technology Journal*, 43 (6), 135-141.

[55] LaMotta E J. Internal diffusion and reaction in biological films. *Environmental Science and Technology*, 10 (1976), 765-769.

[56] Lawrence J R, Wolfaardt G M, Korber D R. Determination of diffusion coefficients in biofilms by confocal laser microscopy. *Applied Environmental Microbiology*, 60 (1994), 1166-1173.

[57] LeVeque R J , *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*. SIAM, 2007.

[58] Liu X D, Osher S, Chan T. Weighted Essentially Nonoscillatory Schemes. *Journal of Computational Physics*, **115 (1)**, 200-212, 1994.

[59] Ljung L, System identification: Theory for the user. *Prentice Hall*, New Jersey, 1999.

[60] López-Núñez A., Mathematical modelling in biofilms (in Spanish). Master Thesis, University of A Coruña, 2013.

[61] Marsh P D. Dental plaque as a biofilm and a microbial community - implications for health and disease. *BMC Oral Health*, 6 (2006).

[62] Marsh E J, Luo H L, Wang H. A three-tiered approach to differentiate *Listeria monocytogenes* biofilm-forming abilities. *FEMS Microbiology Letters*, 228 (2003), 203–210.

[63] Miettinen M K, Björkroth KJ, Korkeala H J. Characterization of *Listeria monocytogenes* from an ice cream plant by serotyping and pulsed-field gel electrophoresis. *International Journal of Food Microbiology*, 46 (1999), 187–192.

[64] Møretrø T, Langsrud S. *Listeria monocytogenes*: biofilm formation and persistence in food-processing environments. *Biofilms*, 1 (2004), 107–121.

[65] Mosquera-Fernández M, Rodríguez-López P, Cabo M L, Balsa-Canto E. Numerical spatio-temporal characterization of *Listeria monocytogenes* biofilms. *International Journal of Food Microbiology*, 182 (2014), 26–36.

[66] Mosquera-Fernández M, Sanchez-Vizuete P, Briandet R, Cabo M L, Balsa-Canto E. Quantitative image analysis to characterize the dynamics of*Listeria monocytogenes* biofilms. *International Journal of Food Microbiology*, 236 (2016), 130–137.

[67] Mueller L N, de Brouwer J F C, Almeida J S, Stal L J, Xavier J B. Analysis of a marine phototrophic biofilm by confocal laser scanning microscopy using the new image quantification software PHLIP. *BMC Ecology*, 6:1 (2006).

[68] Murray J D. *Mathematical Biology*. Springer, 1993.

[69] National Research Council. Trickling filters in sewage treatment at military installations. *Sewage Works Journal*, 18 (1946).

[70] Noguera R D, Okabe S, Picioreanu C. Biofilm modeling: Present status and future directions. *Water Science and Technology*, 39 (1999), 273-278.

[71] Noguera R D, Pizarro G, Stahl D A, Rittmann B E. Simulation of multispecies biofilm development in three dimensions. *Water Science and Technology*, 39 (1999), 123-130 .

[72] Pardalos P M, Romeijna H E, Tuyb H. Recent developments and trends in global optimization. *Journal of Computational and Applied Mathematics*, 124 (2000), 209-228.

[73] Picioreanu C, Kreft J, van Loosdrecht M. Particle-based multidimensional multispecies biofilm model. *Applied and Environmental Microbiology*, 70 (2004), 3024-3040.

[74] Picioreanu C, van Loosdrecht M C M, Heijnen J J. A new combined differential-discrete cellular automaton approach for biofilm modelling: Application for growth in gel beads. *Biotechnology and Bioengineering*, 57 (1998), 718-731.

[75] Picioreanu C, van Loosdrecht M C M, Heijnen J J. Mathematical Modelling of Biofilm Structure with a Hybrid Differential-Discrete Cellular Automaton Approach. *Biotechnology and Bioengineering*, 58 (1998), 101-116.

[76] Picioreanu C, van Loosdrecht M C M, Heijnen J J. Effect of diffusive and convective substrate transport on biofilm structure formation: A two-dimensional modeling study. *Biotechnology and Bioengineering*, 69 (2000), 504–515.

[77] Picioreanu C, Xavier J B, van Loosdrecht M C M. Advances in mathematical modeling of biofilm structure. *Biofilms 1* (2004), 1-12.

[78] Pilchová T, Hernould M, Prévost H, Demnerová K, Pazlarová J, Tresse O. Influence of food processing environments on structure initiation of static biofilm of *Listeria monocytogenes*. *Food Control*, 35 (2014), 366–372.

[79] Pritchett L. Analysis of a one dimensional biofilm model. PhD thesis, Montana State University, 2000.

[80] Quarteroni A, Sacco R, Saleri F. Numerical Mathematics. *Springer-Verlag*, New York, 2000.

[81] Rice S A, Tan C H, Mikkelsen P J, Kung V, Woo J, Tay M, Hauser A, McDougald D, Webb S A, Kjelleberg S. The biofilm life cycle and virulence of Pseudomonas aeruginosa are dependent on a filamentous prophage. *ISME Journal*, 3 (2009), 271–282.

[82] Rieu A, Briandet R, Habimana O, Garmyn D, Guzzo J, Piveteau P. *Listeria monocytogenes* EGD-e biofilms: no mushrooms but a network of knitted chains. *Applied and Environmental Microbiology*, 74 (2008), 4491–4497.

[83] Rittmann B E, Manem J. Development and experimental evaluation of steady-state multi-species biofilm models. *Biotechnology and Bioengineering*, 39 (1992), 914-922.

[84] Rittmann B E, McCarty P L. Model of steady-state-biofilm kinetics. *Biotechnology and Bioengineering*, 22 (1980), 2343-2357.

[85] Rittmann B E, McCarty P L. Substrate flux into biofilms of any thickness. *Journal of Environmental Engineering*, 108 (1981), 831-849.

[86] Rittmann B E, McCarty P L. Environmental Biotechnology. *Principles and Applications*, McGraw Hill, 2001

[87] Rodríguez-López P, Saá-Ibusquiza P, Mosquera-Fernández M, López-Cabo M. *Listeria monocytogenes*-carrying consortia in food industry. Composition, subtyping and numerical characterisation of mono-species biofilm dynamics on stainless steel.*International Journal of Food Microbiology*, 206 (2015), 84–95.

[88] Russell, W. Polynomial interpolation schemes for internal derivative distributions on structured grids. *Applied Numerical Mathematics*, 17 (1995), 129-171.

[89] Sethian J A. Level Set Methods and Fast Marching Methods. *Cambridge University Press*, 2nd Edition, 1999.

[90] Sethian J A. Fast Marching Methods. *SIAM Review*, **41**(2), 1999.

[91] Sethian J A, Vladimirsky A. Ordered upwind methods for static Hamilton-Jacobi equations: theory and algorithms. *SIAM Journal on Numerical Analytics*, 41 (2003), 325-363.

[92] Silva S, Teixeira P, Oliveira R, Azeredo J.Adhesion to and viability of *Listeria monocytogenes* on food contact surfaces.*Journal of Food Protection*, 71 (2008), 1379–1385.

[93] Srey S, Jahid I K, Ha S-D. Biofilm formation in food industries: A food safety concern. *Food Control*, 31 (2013) 572-585.

[94] Stewart P S. A review of experimental measurements of effective diffusive permeabilities and effective diffusion coefficients in biofilms. *Biotechnology and Bioengineering*, 59 (1998), 261–272.

[95] Stoodley P, Boyle J D, DeBeer D, Lappin-Scott H M. Evolving perspectives of biofilm structures. *Biofouling*, 14 (1999), 75–90.

[96] Swaminathan B, and Gerner-Smidt P. The epidemiology of human listeriosis. *Microbes and Infections*, 9 (2007), 1236–1243.

[97] Silyn-Roberts G, Lewis G. A technique in confocal laser microscopy for establishing biofilm coverage and thickness. *Water Science and Tecnology*, 36 (1997), 117-124.

[98] Tack I L, Logist F, Noriega-Fernández E, Van Impe J F. An individual-based modeling approach to simulate the effects of cellular nutrient competition on *Escherichia coli* K-12 MG1655 colony behavior and interactions in aerobic structured food systems. *Food Microbiology*, 45(2015), Pt B:179-88.

[99] Tawakoli P N, Al-Ahmad A, Hoth-Hanning W, Hanning M, Hanning C. Comparison of different live/dead stainings for detection and quantification of adherent microorganisms in the initial oral biofilm. *Clinical Oral Investigations*, 17 (2013), 841-850.

[100] Van Loosdrecht M C M,Eikelboom D, Gjaltema A, Mulder A, Tijhuis L, Heijnen J J. Biofilm Structures. *Water Science and Tecnology*,32 (1995) 35-43.

[101] Van Loosdrecht M C M, Picioreanu C, Heijnen J J. A more unifying hypothesis for biofilm structures. *FEMS Microbiology Ecology*, 24 (1997), 181-183.

[102] Vilas C, Arias-Méndez A, Garcia M R, Alonso A A, Balsa-Canto E. Towards predictive food process models: a protocol for parameter estimation. *Critical Reviews in Food Science and Nutrition*, 58 (2017), 436-449.

[103] Vyas N, Sammons R L, Addison O, Dehghani H, Walmsley A D. A quantitative method to measure biofilm removal efficiency from complex biomaterial surfaces using SEM and image analysis. *Science Reports*, 6:32694 (2016)

[104] Walter E, Pronzato L. Identification of Parametric Models from Experimental Data. *Springer, Masson*, 1997.

[105] Wanner O, Eberl H J, Morgenroth E, Noguera D R, Picioreanu C, Rittmann B E, van Loosdrecht M C M. *Mathematical Modelling of Biofilms*. Technical report, IWA Task Group on Biofilm Modelling, 2006.

[106] Wanner O, Gujer W. Competition in biofilms. *Water Science and Technology*, 17 (1984), 27-44.

[107] Wanner O, Gujer W. A multispecies biofilm model. *Biotechnology and bioengineering*, 1986. 28 (33), 314-328.

[108] Widerer P A, Characklis W G. Structure and function of biofilms. In: *Structure and function of biofilms*. John Wiley & Sons. Eds. Widerer P A, Characklis W G, 1989, pp 5-17.

[109] Wilks S A, Michels H T, Keevil C W. Survival of *Listeria monocytogenes* scott a on metal surfaces: Implications for cross-contamination. *International Journal of Food Microbiology*, 111 (2006), 93–98.

[110] Williamson K, McCarty P L. A model of substrate utilization by bacterial films. *Journal Water Pollution Control Federation*, 48 (1976), 9-24.

[111] Wimpenny J W T, Colasanti R, A unifying hypothesis for the structure of microbial biofilms based on cellular automaton models. *FEMS Microbiology Ecology*, 22 (1997), 1-16.

[112] Xavier J, Picioreanu C, van Loosdrecht M. 2004. A modelling study of the activity and structure of biofilms in biological reactors. *Biofilms*, 1 (3), 377-391.

[113] Xavier J, Picioreanu C, van Loosdrecht M. 2005. A framework for multidimensional modelling of activity and structure of multispecies biofilms. *Environmental Microbiology*, 7 (8), 1085-1103.

[114] Yang X, Beyenal H, Harkin G, Lewandowski Z. Quantifying biofilms structure using image analysis. *Journal of Microbiological Methods*, 39 (2000), 109–119.

[115] Zaborowska M, Tillander J, Branemark R, Hagberg L, Thomsen P, Trobos M. Biofilm formation and antimicrobial susceptibility of staphylococci and enterococci from osteomyelitis associated with percutaneous orthopaedicimplants. *Journal of Biomedical Matererials Research Part B*, (2017), 105B:2630–2640.

[116] Zhang T C, Bishop P L. Structure, activity and composition of biofilms. *Water Science and Technology*, 29 (1994), 335-344.

[117] Zobell C E, Anderson D Q. Observations on the multiplication of bacteria in different volumes of stored seawater and the influence of oxygen tension and solid surfaces. *The Biological Bulletin*, 71 (1936), 342.