Modeling and model analysis for the size distribution of adipocytes

Léo Meyer

UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE MIPTIS

Institut Denis Poisson

THÈSE présentée par :

Léo Meyer

soutenue le : 9 octobre 2023

pour obtenir le grade de : Docteur de l'Université d'Orléans

Discipline/ Spécialité : Mathématiques

Modeling and model analysis for the size distribution of adipocytes

THÈSE dirigée par :

RIBOT Magali	Professeur, Université d'Orléans

RAPPORTEURS :

CAÑIZO A. José	Associate Professor, Université de Grenade
CHAMPAGNAT Nicolas	Directeur de recherche, INRIA Nancy - Grand Est

JURY :

ABRAHAM Romain	Professeur, Université d'Orléans
CAÑIZO A. José	Associate Professor, Université de Grenade
CHAMPAGNAT Nicolas	Directeur de recherche, INRIA Nancy-Grand Est
DOUMIC Marie	Directrice de recherche, Sorbonne Université
GABRIEL Pierre	Maître de conférences, Université de Versailles
RIBOT Magali	Professeur, Université d'Orléans
VEBER Amandine, Presidente du Jury	Directrice de recherche, Université Paris Cité
YVINEC Romain	Directeur de recherche, INRAE Nouzilly

Remerciements

Après trois ans de craies sur des tableaux, de stylos sur des feuilles et de touches de clavier à faire des mathématiques, il est l'heure des remerciements !

Commençons par ceux qui ont de loin le plus contribué au bon déroulement de cette thèse, à savoir mes encadrants : Magali Ribot et Romain Yvinec. Magali, je tiens à te remercier pour ton encadrement pendant ces trois ans, pour tes conseils mathématiques ou non, pour avoir pris le temps de m'expliquer les rouages de cette vaste machine que sont la recherche et l'enseignement supérieur. Romain, merci d'avoir accepté de co-encadrer cette thèse, merci pour ton aide et notamment les différentes pistes et conseils dans les aspects plus techniques de cette thèse et d'avoir guidé le sujet vers un angle plus probabiliste vers la fin. Je ne me prononcerai pas sur le choix entre EDP et probabilité afin de ne vexer aucun de vous deux ! Merci collectivement à vous deux pour votre encadrement en général, je crois que j'ai eu beaucoup de chance de pouvoir travailler avec vous pendant ces 3 ans, et je pense sincèrement que je ne pouvais pas rêver mieux en termes de direction de thèse. Si un jour, j'ai la chance de me retrouver à votre place, j'espère faire au moins aussi bien !

Je tiens à remercier José A. Cañizo et Nicolas Champagnat d'avoir rapporté cette thèse, particulièrement dans une période souvent chargée professionnellement. Je vous remercie également pour vos retours qui ont permis d'enrichir ce manuscrit, d'en corriger certains aspects et surtout ouvert de nouvelles questions sur la modélisation des adipocytes.

Je remercie également Romain Abraham, Marie Doumic, Pierre Gabriel et Amandine Veber d'avoir accepter de participer à ce jury de thèse ainsi que Hédi Soula qui me fait l'honneur d'être membre invité de ce jury.

Un grand merci également à l'ensemble des membres de l'Institut Denis Poisson pour les divers échanges et discussions. Un remerciement tout particulier pour Anne et son aide dans toutes les taches administratives parfois plus complexes que nécessaires.

Un petit mot aussi pour les doctorants et ex-doctorants de l'IDP que j'ai pu croiser pendant ces trois ans à Orléans : Noémie, Maxime, Rita, Alexis, Grégoire, Emile, Simon, Titouan, Sonia, Pagiel. Une salutation particulière pour mes compagnons d'aventure Ouassim et Mohammad qui ont commencé en même temps que moi et dans le même bureau ! J'ai eu l'honneur d'être leur représentant auprès de l'École Doctorale MIPTIS, j'espère avoir fait le job !

Je remercie bien sûr les divers collaborateurs qui ont participé à l'avancement de cette thèse. Tout particulièrement, le groupe de l'ANR MATIDY : Hédi, Anne-Sophie, Laurent, Aleksandra, Alois et notre cheffe Chloé Audebert. Je tiens à rappeler que cette histoire d'adipocytes commence avant tout par ma rencontre avec Chloé pour mon stage de Master 1. Je ne serais pas là sans toi ! Des remerciements également pour les membres de l'équipe MUSCA de l'INRIA et l'organisation des mathmeetings ainsi que pour les membres de l'équipe BIOS de la PRC de l'INRAE Nouzilly et pour leur accueil lors des quelques journées que j'ai passées à Nouzilly. Je remercie également Maxime Breden même si nos travaux ensemble n'ont pas encore entièrement abouti, nos échanges sont toujours fructueux et c'est grâce à toi que j'ai découvert le monde des preuves assistées par ordinateur.

Je tiens à remercier également Marie Doumic, Thierry Goudon et Erwan Hingant pour avoir accepté de suivre cette thèse lors d'un comité de suivi de thèse assez informel mais toujours riche en échanges et en pistes de recherche. J'adresse également des remerciements à mes compères du CEMRACS, Mickael Bestard et Florent Noisette ainsi que Bastien Polozzi pour son encadrement, car après tout, qui ne rêve pas de passer son mois d'août à Marseille à faire des maths !

Sur un plan moins professionnel, mais toujours mathématique, je tiens à remercier deux personnes qui ont beaucoup joué sur mon orientation vers le périlleux chemin des mathématiques. Philippe Jaouen, mon professeur de mathématiques au Lycée Saint-Esprit de Landivisiau, qui

après que je lui aie annoncé que finalement, j'allais faire un master de mathématiques, m'a dit "C'est bien tu retourne sur le droit chemin". Et Benoît Sarels, qui a accepté de m'encadrer pendant un stage de L2 et qui m'a fait comprendre qu'on peut faire des mathématiques appliquées à la biologie, à un moment où je me demandais ce que j'allais bien pouvoir faire avec une double licence Biologie et Mathématiques.

Enfin, sur un plan personnel, je remercie mes parents qui je pense ont compris assez tôt qu'ils ne comprenaient pas grand-chose à ce que je faisais, mais qui m'ont toujours soutenu. Merci d'avoir supporté huit ans d'études et d'avoir insisté pour les cours d'anglais (merci à Jos !) ce qui m'est bien utile aujourd'hui. Je ne serai pas médecin, il faudra vous contenter d'un docteur ! Un grand merci à ma grand-mère Manou qui sera ravie d'avoir un petit-fils docteur comme mon grand-père.

Le plus grand des remerciements à ma petite sœur Léna exilée dans les lointaines terres de Jacques Cartier (le Canada) qui a dû me supporter comme colocataire pendant ses deux premières années d'études. Il parait que toi aussi un jour, tu devras écrire des remerciements comme moi, il me tarde de les lire !

Un immense merci à ma famille roscovite : les Piwouis. Il y a trop de choses à dire et pas assez de lignes pour les écrire, mais vous savez combien notre amitié m'est précieuse. Parmi eux, je ne remercie pas Arthur et Tanguy qui ont plus fait pour le recul que pour l'avancée de cette thèse et ils sauront pourquoi ! Merci à Claire qui a accepté de cohabiter avec moi pendant un mois à Marseille, merci pour ton oreille attentive et pour les recaps du soir !

J'oublie certainement certaines personnes dans ces remerciements, et elles voudront bien m'en excuser. La liste est longue et il serait dommage qu'il y aie plus de remerciements que de mathématiques dans ce manuscrit.

Enfin, ce manuscrit est dédié à Grâce, qui nous a quitté bien trop tôt, et qui avait compris, bien avant moi, qu'il fallait faire des mathématiques. Sache que tu nous manques terriblement.

Contents

Chapter 1

Introduction

This thesis deals with the study and simulation of models for the adipose tissue. In particular we are interested in the size distribution of adipose cells inside the adipose tissue. As such we begin by presenting the biology of the adipose tissue as well as previous efforts at modeling. We then present our models which are based on more general classes of models originally developed for modeling polymers : the Lifshitz-Slyozov model and the Becker-Döring model.

1.1 Biological motivation : the adipose tissue

1.1.1 Generalities

The adipose tissue is part of the larger family of connective tissue. It is composed of various types of cells, the main one being adipocytes or adipose cells. Its main role is the storage of energy in the form of lipids but it also provides structures to neighboring organs as well as thermal insulation. In most species, the adipose tissue splits into two kinds: the white adipose tissue (WAT) and the brown adipose tissue (BAT). Depending on the type of tissue, different types of adipose cells are observed: white adipose cells and brown adipose cells. Those cells differ in the way they store fat and also in their role. White adipose cells (WAC), and in extent WAT, have a single lipid droplet inside their cytoplasm and are a form of energy storage. These lipid droplets give the white color to the tissue, thus its name. In contrast, brown adipose cells (BAC) can have multiple vesicles of lipids as well as multiple mitochondria. Their role is to perform adaptive thermogenesis and to release their energy content in the form of heat. BAT is present in great quantity in human infants and later regresses in adults [83]. Although they fill different roles, WAC can still convert themselves to BAC. This process is called 'beiging' or browning of WAC (from the french 'beige': a color between white and brown), where WAC turn themselves into energy-releasing adipocytes. This process has been observed to happen in response to cold exposure. This process is well documented in [114] and is reversible. In our modeling of the adipose tissue we will only consider the WAT and WAC. However the data collection is not perfect and we may assume that some of our data are actually BAC or preadipocytes or endothelial cells, ... although these additional cell types would be present in a small percentage compared to WAC.

The life of a fatty acid : from the intestine to storage

Adipocytes store energy in the form of a single or multiple triglycerides vesicles inside their cytoplasm. Those molecules are created inside the cells from fatty acid by a process called **lipogenesis**. However before fatty acids get inside the adipose cells they get ingested by the organism and end up in the intestine. From inside the intestine, micelles containing fatty acids are transported to enterocytes where they are assembled with glycerol to form triglycerides. They are then bundled together with apolipoproteins, phospholipids and cholesterol to form chylomicrons also called ultra low-density lipoproteins (ULDL). This complex protein structure allows the transport of triglycerides inside the water-based blood stream toward storage tissues such as the adipose tissue. When those chylomicrons reach the capillaries in the adipose tissue, the lipoprotein lipase dissolves the triglycerides into free fatty acid which are transported inside adipocytes. Inside the cell they are again transformed into triglycerides using glucose and acetyl-CoA. The formation of triglycerides from fatty acids is called lipogenesis. Lipogenesis is heavily regulated by a common hormone :

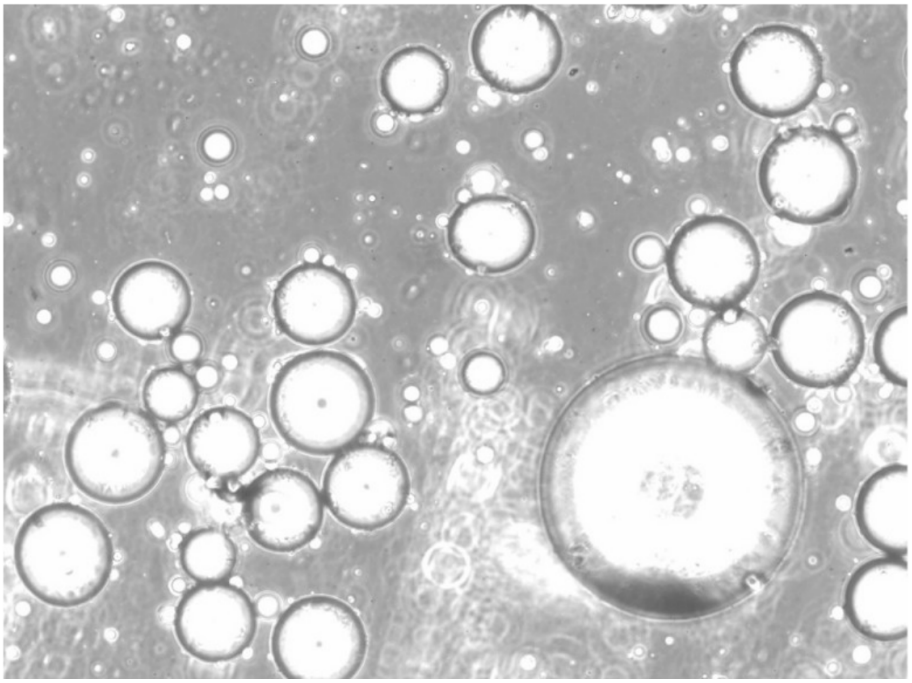


Figure 1.1: Adipocytes of different size

insulin. Evidently this implies that diabetes and other insulin-related diseases have an influence on the physiology of the adipose tissue.

The reversed biochemical reaction through which triglycerides are hydrolyzed into glycerol and fatty acids and released into the surrounding capillaries is called **lipolysis**. Lipolysis is regulated by two hormones : glucagon and norepinephrine. Both processes are illustrated in Figure 1.2 taken from [108].

1.1.2 Morphology

Localization

Being a connective tissue, the adipose tissue is often localized together with its structured organs. In humans, it is present in various places. Visceral fat is located around internal organs, and is one of the indicator of health issues such as obesity. Subcutaneous fat is located widely bellow the skin in humans where it provides minor thermal isolation as well as padding. Intermuscular fat is located in the muscular system and provides a quick access to energy storage for surrounding muscles. Marrow fat is present in bone marrow but its role beyond energy storage is not well understood.

Adipose tissue and weight

There is an obvious correlation between the adipose tissue and the weight of an individual. Nowadays fat mass is used as an indicator of health together with the Body Mass Index (BMI), although the use of BMI as a health indicator is often debated. Nonetheless with the recent surge of obesity in the human population [3], the scientific community has started tackling the study of weight loss by looking at adipose cell physiology. Many works have exposed what is usually called 'rebound weight gain' where after a substantial weight loss, individuals would regain weight faster than individuals of the same weight [89, 5]. This is explained by many factors often related to some resistance of adipose cells once they get large enough. We should however distinguish between two mechanisms that can happen during weight gain. First, hypertrophy is the increase in volume of adipose cells. This is often the first physiological change implying weight gain. Second is hyperplasia, which is the increase of the number of cells. Hypertrophy will happen until moderate obesity, at which point hyperplasia appears and the number of cells increases, hence the available storage space also increases [59]. In children, new adipose cells are created over the growth period until adulthood. This often leads to the idea that adults have a 'fixed' amount of adipose cells. However

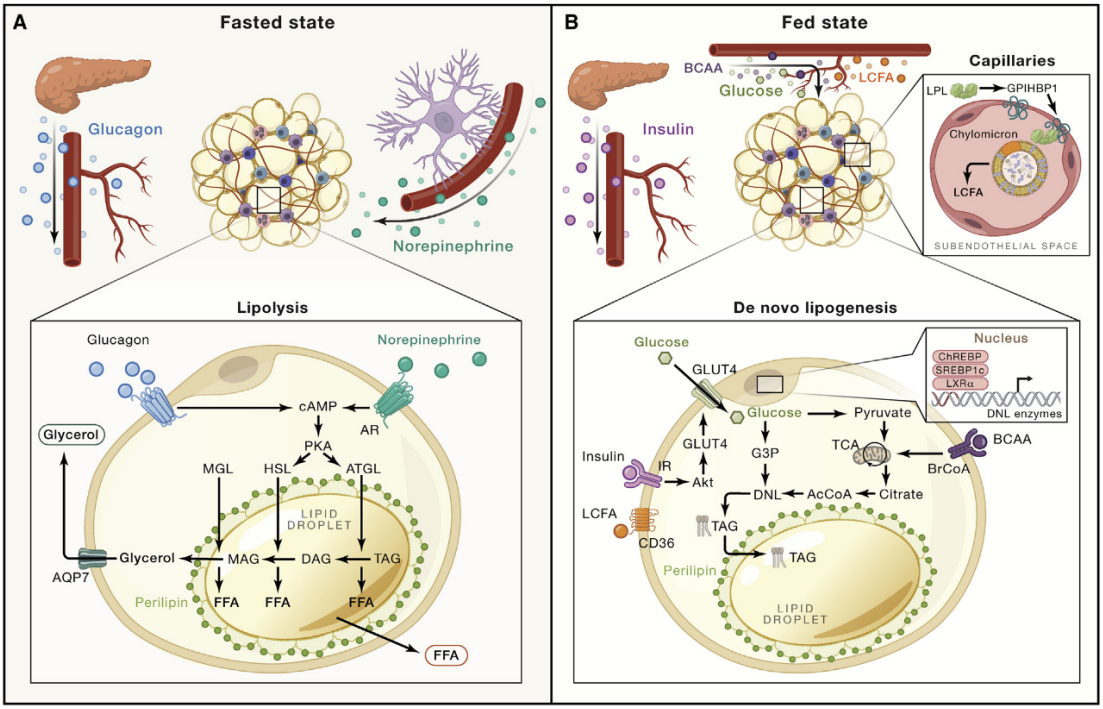


Figure 1.2: An illustration of lipogenesis and lipolysis. This figure is taken from [108].

new cells are created through the adipose tissue turnover which happens to replace naturally dying cells. Indeed a healthy individual,in regards to weight, will see small variations of its adipose cell count, but as stated before hyperplasia occurs at the later stage of obesity. Studies have yet to be performed to characterize the increase in cell count from leanness to obesity but there are many results on the reverse process. Weight loss induced by dietary changes or surgery decreases the size of cells but does not decrease the number of cells [6]. This leads to the assumption that the adipose cell count increase is a one way, non-reversible process.

As such, 'weight gain' in the sense of gaining additional adipose cells past the original amount is permanent. This implies that the most effective way to reduce fat mass in the overweight stage is a surgical operation, either through installing a gastric band to reduce food intake or liposuction to remove part of the fat mass. However, as we pointed out before, those are physical remediations to the overweight condition and do not modify the physiology of remaining adipose cells.

When we will consider the modeling of the adipose tissue, we make the simplification that the total number of cells is constant. Indeed this is not biologically relevant but this simplification is a first step towards the modeling and computations of the adipose tissue dynamic.

Size heterogeneity

Many studies have shown that the adipose tissue is heavily heterogeneous, be it in cell types, hormonal levels and cell sizes, which is our main interest [76, 65, 91]. As seen in Figures 1.3, the size of cells inside the adipose tissue follows a bimodal distribution : it has two local maxima, called modes, separated by a local minimum called the nadir. However no biological explanation has been coined down as to why adipose cells separate into two populations of small and large cells. Interestingly this separation has been observed in many species but the position and relative height of the modes are different from one species to another as well as from individual to individual. Particularly the size distribution of adipose cells is an indicator of being overweight or obese, since those individuals will tend to have a higher number of adipose cells and have larger cells.

1.1.3 Health issues

Obviously studying adipose cells involves studying related health issues. The main pathology related to the adipose tissue is obesity. This disease has seen a great increase in the human population, having tripled since 1975, even being labelled an epidemic by the World Health Organization [3].

The main symptom is the excessive amount of body fat and is characterized by a BMI index above 30. Many studies have shown that obesity has strong correlations with various afflictions, such as cardiovascular diseases [76], type 2 diabetes [42], obstructive sleep apnea, cancers [37] and osteoarthritis. The causes are various and often different from one individual to another. They include excessive intake of food, sedentary lifestyle, lack of physical exercise, genetics, ...

Concerning medical treatment, they most often consist of changing life-habits : diets and/or physical exercise, although the most efficient dietary change for long term weight loss is still unknown. Additionally, medications can help weight loss but the most efficient way to treat obesity is actually bariatric surgery. This includes gastric bypass and banding, gastrectomy or pancreatic diversion.

1.1.4 Why do we want to study the size distribution of adipose cells ?

We can now get to why we are interested in the size distribution of adipose cells. First and foremost we recall that the size distribution of adipose cells, that is to say the number of cells per size, is bimodal : it has two local maxima. This can be seen in Figure 1.3 for a rat. This type of distribution is singular for specialized cells which more often display a standard unimodal distribution centered around the mean size of a particular type of cell.

What do we mean behind 'size' of a cell ? There are various ways to define the size of a cell. We will firstly consider the radius as a way to define size. This means that we assume cells to be spheres, whereas in the adipose tissue, and since cells are mainly composed of water, we expect the cells to form a densely packed tissue, where cells will definitely not take the form of spheres. Hence, cells as sphere is a modeling simplification, but also comes from the fact that to collect data, biologists break the extra-cellular matrix that holds the tissue and cells are therefore freely swimming inside a medium, where they should resemble spheres, as seen in Figure 1.1. Further on, we will consider that the size of an adipose cell is determined by its lipid amount contained inside its vesicle. We will relate radius to the lipid amount by considering the volume of the cell compared to the volume of the vesicle, which allows us to relate both size definitions.

As stated before, there is still no biological explanation for this bimodal distribution. Moreover, the bimodal feature seems to be universal among animal species but the height and position of each maxima are individual dependent. Data collection comes with various pitfalls : the surgery for biopsy is often invasive, lack of longitudinal data, ethical considerations, truncated data, not exhaustive (no total cell counts), ...

1.1.5 What type of data do we have ?

Let us detail the type of data we have at our disposal. Mostly they come in the form of lists containing the radius or diameter of a subset of cells obtained through a biopsy. The size of each cell is measured using a Beckman Coulter Multisizer IV but due to to limitation in measurement techniques, only cell radii larger than 7.5μm or 10μm depending on experiments are measured. Hence our data may be incomplete and the chemical procedure that destroys the tissue before using the multisizer may lead to some artifacts remaining in the solution and being counted by the machine. Therefore among all the measurements, a small part may not be cells but remaining components of the tissue and also other cells (BAC, preadipocytes, endothelial cells, ...).

1.2 Modeling the adipose tissue

In this section, we go over various models for the adipose tissue. We shall see that many attempts have been made at modeling various biological features of the adipose tissue. We first present models in a wider perspective before diving in on models for the size dynamic of adipose cells and in particular the models from Jo et al. [68] and Soula et al. [118].

1.2.1 Various models for the adipose tissue

Many models have been developed to study the adipose tissue, each of them intending of capturing a particular feature of this tissue.

A common question when studying specialized cells is how they are created. In the case of adipose cells, mesenchymal stem cells, located in the bone marrow, differentiate into pre-adipocytes.

These pre-adipocytes are undifferentiated fibroblasts which can form adipose cells upon being stimulated. The exact biochemical process of the fate determination of pre-adipocytes is yet to be thoroughly described. Still some mathematical modeling was performed to capture this differentiation process. In [46], authors developed a model for the differentiation process of mesenchymal cells in adipose cells. The model introduced in [46] is composed of ODEs for the total population of mesenchymal cells and pre-adipose cells, and a transport equation structured in size with a death term for adipose cells. This model and its extension with a spatial component are able to replicate bimodal distributions.

We previously highlighted the fact that the adipose tissue is a connective tissue and therefore also provides structure to the surrounding organs or tissue, as in the visceral fat for example. The interaction between extra-cellular matrix (ECM) and adipose cells is even correlated with health issues such as obesity, as featured in [36]. The interaction between cells and their ECM has been well studied by the mathematical community and some works looked at the particular case of the adipose tissue. In [100], authors developed a model of interaction between adipose cells and collagen fibers. They are able to reproduce some features of the adipose tissue's morphology including the clustering of adipose cells and the formation of a network of fiber. They use an Individual Based Model including the forces and interaction between the cells and fibers as well as pre-adipocyte differentiation. This model was used again in [101] to investigate injury to the adipose tissue. In particular the model is able to replicate regeneration and scar formation.

Additionally biologists are interested in the biochemical processes happening inside the cell and at the scale of the tissue. In [8], authors investigated the lipid turnover in human. They found that the lipid release rate decreases with age while the intake rate is not reciprocally adjusted. They also observed that major weight loss is driven by a decrease in the intake rate of lipids and that individual with a low removal rate are more likely to keep a stable weight after weight loss. All these findings further demonstrate the role of lipid intake and release in the adipose tissue morphology and its relation to health issues. They also fitted their data with a PDE model structured in age for the population of lipids inside an individual. The methodology for data collection in this article is particularly interesting: they measure the age of lipids in an individual by looking at the ratio of radioactive carbon coming from nuclear bomb test that are incorporated inside the carbon chain of lipids.

In [71], the authors introduce a physiological based ODE model to illustrate intracellular compartmentalization of metabolic lipid processes and differential activation of lipases involved in these processes. They show that the active metabolic subdomain has to be well calibrated because it is critical for simulating data. Depending on the size of this subdomain the speed of cellular dynamic can change and thus modify the outcome of the model. A smaller domain produces faster dynamic. The differential activation of lipases is due to their role in the metabolic processes of different glycerides: tri-, di- and monoglycerides. They show that this differential activation allows them to compare more accurately to experimental data. However some of their subsequent finding has yet to be confirmed by experiments.

1.2.2 Models for the size dynamic of adipose cells

In this section we present the recent developments in modeling the distribution in size of adipose cells. The main two articles were coincidentally published in 2013 by Jo et al. [68] and Soula et al. [118]. The models introduced in this manuscript, to be described in the next section, are linked to the models from these two papers but are more closely inspired by the one in [118].

Jo et al. [68]

In this article, the authors are interested in the inference of the adipose cell size distribution dynamic using a PDE model and Bayesian inference. The model was first introduced in [66] but instead of time, the distribution depended on the fat pad mass. They perform model fitting on data from rat and are able to recover the bimodal distribution of adipose cells. Their PDE model consists in a transport-diffusion equation for the distribution in size with a source terms for creation of new cells and cell death. The size of a cell is assumed to be its radius and the distribution of cells at time t and radius r denoted $f(t,r)$ is given by :

$$\partial_t f(t,r) = b(t)\delta(r-r_0) - \partial_r(v(r)f(t,r)) + D\partial_r^2 f(t,r) - k(r)f(t,r),$$

where b is the birth rate of new cells that appear with size r_0, v is the velocity of transport, D the rate of diffusion, k the death rate and δ the Dirac delta function. The velocity v takes the form :

$$v(r) = \frac{\nu_+}{2}\left(1 + \tanh(\frac{r - r_+}{\eta_+})\right) - \frac{\nu_-}{2}\left(1 + \tanh(\frac{r - r_-}{\eta_-})\right).$$

The idea behind this choice for velocity is phenomenological : one may assume that lipid intake and release by a cell are surface limited. Hence there are some critical sizes for intake and release, denoted by r_+ and r_-, which give the half-maximal intake/release rates, denoted by $\nu_{+,-}$. The parameters $\eta_{+,-}$ are included to describe the steepness of the rates. Observe that this velocity does not depend on time, hence does not include potential rate change depending on the tissue dynamic.

Soula et al. [118]

The model introduced in [118] is the basis for all the models we later introduce in this manuscript. The authors derive a model for the size distribution of adipose cells, where the distribution depends on both the amount of lipids x inside the cell and the radius r of the cell. The same model was used in [119] to estimate the surface rate of lipolysis. The PDE describing the dynamic of the distribution is given by :

$$\partial_t f(t,x,r) = \partial_x(v(x,r,L)f) + \partial_r(R(x,r)f), \tag{1.1}$$

where L is the external amount of lipids, which also has its own dynamic given by :

$$\frac{\mathrm{d}L}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}t}\iint x f(t,x,r)\mathrm{d}x\mathrm{d}r. \tag{1.2}$$

This equation for L translates the assumption that the total amount of lipids is constant. The velocity in the direction of the amount of lipids x takes the form :

$$v(x,r,L) = a(r)\frac{L}{L+\kappa} - b(x,r),$$

The function a describes the rate of lipogenesis and b the rate of lipolysis. We go into further details in Section 1.3.3 about their construction and the shape of this velocity.

And the velocity in the direction of the radius r is given by :

$$R(x,r) = \frac{1}{\tau}\left(\frac{V_{\text{lipids}}x + V_0}{4\pi r^2} - \frac{r}{3}\right), \tag{1.3}$$

where τ is the rate at which the radius of the cell slowly adapts to the changing amount of intracellular lipids. Authors use Monte Carlo simulations to determine numerically stationary solutions of the model. They are able to recover bimodal distributions in radius.

1.3 Adaptation of the Lifshitz-Slyozov and Becker-Döring models to adipocyte size dynamic modeling

Following the work in [118], we present the main framework for the models we aim to build. First we go into details about the goal of the mathematical modeling and the purpose of our approach. This leads us to look at the Lifshitz-Slyozov model (LS). We present the model as well as various well known results from the literature then we go into details about the construction of a Lifshitz-Slyozov type model for a population of adipose cells undergoing lipogenesis and lipolysis. Unfortunately this model is not able to reproduce the biological observations but a relatively simple extension can be considered to improve it : the addition of a diffusion term. Motivated by the choice of such a diffusion term, we look into the Becker-Döring model (BD) as a 'microscopic' equivalent of the Lifshitz-Slyozov model. We also present well known results, in particular the convergence theory from the Becker-Döring model to the Lifshitz-Slyozov model, which is the main motivation for studying this model in the first place. Indeed, this convergence result provides insights into what a diffusion term for the Lifshitz-Slyozov model may look like. In turn, we construct a Becker-Döring

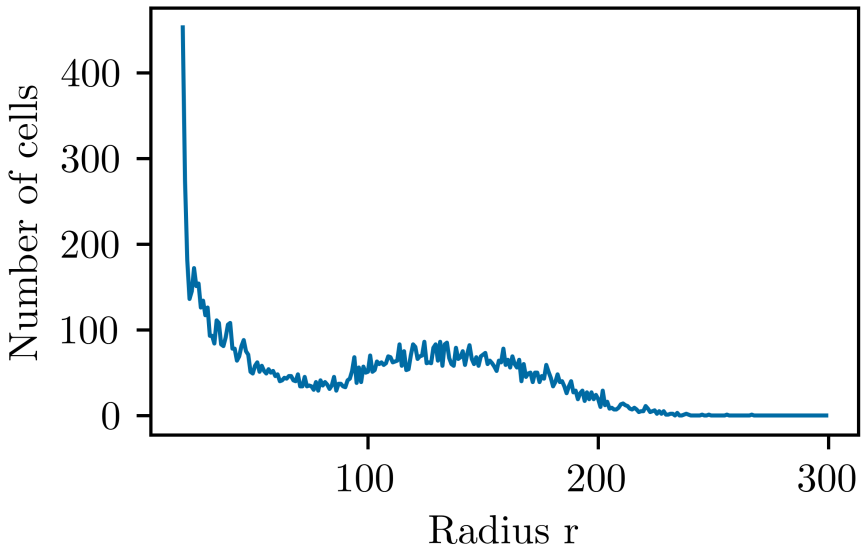


Figure 1.3: Example of collected data from a biopsy on a rat. The raw data are composed of the size of each cell, which we plot here as a histogram for clarity

model for adipose cells. This model is considered purely for its mathematical insight and has no biological assumptions in its construction other than the ones used in the Lifshitz-Slyozov model.

Next, in our approach of using the various convergence results to obtain a candidate for a diffusion term, we were unable to prove convergence toward what could be a diffusive Lifshitz-Slyozov model using the classical analytical tools. It would require to prove uniform bounds on higher order derivatives on the densities, which we are not able to obtain. This motivated us to look into stochastic equivalents to both our models. In the spirit of [74], we construct two new models : a stochastic Becker-Döring model and a stochastic diffusive Lifshitz-Slyozov model. In turn, we provide a result of convergence from one to another, with the assumption that the rates of lipogenesis and lipolysis are bounded. We then extend this result to the case of sub-linear rates which our model will verify.

1.3.1 Goals of mathematical modeling

We begin by detailing the intentions behind the modeling of the size dynamic of adipose cells as well as the objectives of the models. As stated previously, the adipose tissue is heterogeneous, particularly in cell size (radius). Nonetheless, when looking at the population of cells, the radii of cells are not randomly distributed but follow a peculiar distribution : a bimodal distribution. As seen on Figure 1.3, cell radii can take a large range of values but are concentrated around two specific values : small radii close to 25 μm and large radii close to 60 μm. The relative height and position of these modes can be different from species to species and between individuals as well. Particularly, health issues such as obesity and diabetes have an influence on the typical size of adipose cells. Hence, proper tools to study data such as the one in Figure 1.3 are needed. Such tools should be able to replicate the heterogeneity of data collected by biologists, and be based on biological considerations as much as possible.

The models presented in this manuscript aim to fill such prerequisites. As is often the case in mathematical modeling for biology, we want the models to follow a variety of simplifications of the biological truth. Additionally the mathematical analysis leads to assumptions that such models should verify.

With the intent to perform parameter estimation on the data, we first need to clarify what the output of the model is and how it relates to the available data. The data at our disposition consist of a series of measurement of cells coming from biopsy in rats. Hence the data are lists of radii of cells. The models we will consider are dynamical systems that describe the evolution of the size distribution of a population of adipose cells. We make a first simplification and consider that the available data are a sample of the adipose tissue that is stationary in time. This means that

we do not consider natural variations of the adipose tissue due to food consumption and various metabolic changes. In a healthy individual, such changes should be negligible at the scale of the whole tissue. Therefore the outputs of the model are the, eventually attractive, stationary states of the dynamical system. We also make the simplification that the number of cell is constant in time. As previously mentioned this is not true biologically but again, for a healthy individual, the variation of cell count is negligible. Last but not least, we consider that the total amount of lipids is constant in the model. This is again motivated by the idea that a healthy individual is at equilibrium in regards to lipid intake from food and lipid consumption for energy. Finally, we will make a few more simplifications in regards to biophysics, which we detail in Section 1.3.3 when we introduce the first model.

How do we build a dynamical system describing the size distribution of adipose cells ? The most natural object to consider is a partial differential equation (PDE) which describes the evolution of the distribution. Since the data we have do not contain any information about the spatial distribution, the distribution is a function depending on time and size of cells, either radius or amount of lipids. We want the model to verify two constraints coming from the biological simplifications : the total number of cells and the total amount of lipids are constant in time. The first constraint can be verified by providing accurate boundary conditions to the PDE. The second constraint leads to considering an additional unknown of the model : the external amount of lipids, available in the medium. Here the medium is a network of blood vessels and capillaries. Again since we do not have spatial data, we consider this external amount of lipids to be a common pool, available to all the cells. This may be interpreted physically as considering the lipids to be fast diffusing in the medium. Therefore the total amount of lipids is simply the sum between this external amount of lipids and the amount of lipids contained in all the cells, which can be obtained through the distribution.

We now need to choose the shape of the PDE. Continuity equations are a natural choice for describing distributions with conserved quantity. We shall firstly describe how the size of a single cell evolves based on lipogenesis and lipolysis. This in turns describes the flux of 'size' at the population scale and leads us to write a transport equation on the distribution. This transport equation is provided with boundary conditions that conserve the total number of cells as stated before.

This is the core of our model : a conservative transport PDE based on single cell dynamic and a conservation equation on the total amount of lipids in the model. This type of models fits naturally into the Lifshitz-Slyozov model, which we introduce in the next section. Now what do we expect this model to do ? Our aim is that provided some initial distribution of cells and some quantity of external lipids, the model describes the dynamic of the distribution as it stabilizes toward a stationary distribution which should resemble the data in Figure 1.3. We point out that this model does not represent what happens *in vivo*. We make various simplifications of the biological truth, but the closest biological equivalent would be to consider an individual that does not change its food intake, has a perfect balance of lipid intake and energy expenditure through lipid consumption and has a one-to-one adipocyte turnover, which is not so far from the daily life of a lab rat.

However as we shall see in the numerical simulations and by looking at previous theoretical results [18], this model has some flaws and one is especially troublesome. Its stationary solutions are nowhere near the shape of the distributions in the data, see Figure 1.3. This leads us to consider potential extensions of the model. In particular, a simple fix to our problem is to add a diffusive term in the PDE, but we have no biological insights on how to choose or construct such a diffusive term. As a first approach, we investigate simply adding a constant diffusion to the PDE. This model is studied in chapter 5, where parameter estimations are made. A second approach which is the main subject of this manuscript, is a non-constant diffusion term. To choose this term, we seek information from another model : the Becker-Döring model. This model has some close relation to the Lifshitz-Slyozov model we are interested in, particularly because it converges to it in some sense and this convergence provides a way to extend the usual Lifshitz-Slyozov model to a diffusive Lifshitz-Slyozov model. We go into greater detail as to how this is done after presenting both the Lifshitz-Slyozov and Becker-Döring models.

1.3.2 The Lifshitz-Slyozov model

The Lifshitz-Slyozov model bears its name from its creators : Ilya M. Lifshitz and Vitaly V. Slyozov. It was first introduced in [84], to study precipitation phenomena in supersaturated solid solutions. The Lifshitz-Slyozov model was also largely used in modeling polymerization processes. It describes the size dynamic of a population of clusters in a bounded domain. Those clusters are described as an amalgamation of a smaller elementary particle present in quantity u inside the medium. This elementary particle acts as the building block of the larger cluster in the model.

Considering a single cluster, let us denote x its size. The value of x can take different meanings depending on context. For example it could simply be the amount of elementary particles constituting the cluster, or the volume occupied by the cluster, its length, etc ... This cluster either grows or shrinks in size and we assume that these two phenomena depend on the size x of the cluster. Henceforth we will denote a the function describing the growth rate and b the shrinking rate of the cluster. We assume these rates to be positive. We will assume more constraining hypotheses on those rates latter on. Now the size x of the cluster grows depending on the growth rate $a(x)$ and the amount of elementary particle u and shrinks depending on the shrinking rate $b(x)$. This leads to the single cluster dynamic equation :

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = a(x)u(t) - b(x)\,. \tag{1.4}$$

Now we go from a single cluster description to the full population. We denote by t the time variable, assumed to non-negative and $f(t,x)$ the amount of clusters of size x at time t. This function f is the size distribution of the cluster population. Describing the dynamic of f is simply using the continuity equation with the velocity field $v(x,u) = \frac{\mathrm{d}x(t)}{\mathrm{d}t} = a(x)u(t) - b(x)$ which we henceforth call the velocity of the distribution. Naturally we are led to consider the following partial differential equation :

$$\partial_t f(t,x) + \partial_x(v(x,u)f(t,x)) = 0,\ t \geq 0,\ x \geq 0. \tag{1.5}$$

Now since clusters go through the process of growing and shrinking by ingesting or releasing elementary particles, the amount of elementary particles u will change in time. As such we have to provide an equation for the dynamic of u. We consider this whole process to happen in a bounded domain, hence the total amount of elementary particles should not change in time. We denote by $k(x)$ the function describing the amount of elementary particles composing a cluster of size x. Then $\int_0^\infty k(x)f(t,x)\mathrm{d}x$ is the amount of elementary particle composing the cluster. This leads to the conservation equation :

$$u(t) + \int_0^\infty k(x)f(t,x)\mathrm{d}x = \lambda,\ t \geq 0,\ x \geq 0, \tag{1.6}$$

where λ is a constant in time parameter representing the total amount of elementary particles. The function k can simply be $k(x) = x$ if we consider the size to be the amount of elementary particles, but we can also consider a measure of length $k(x) = \frac{x}{C}$, where C is the length of an elementary particle, or $k(x) = Cx^{1/3}$ if x is the volume of the cluster. Both f and u need to be provided with an initial condition, which we denote f^0 and u^0. Equation (1.5) and (1.6) lead us to the Lifshitz-Slyozov equations :

$$\begin{cases} \partial_t f(t,x) + \partial_x(v(x,u(t))f(t,x)) = 0,\ t \geq 0,\ x \geq 0, & \text{(1.7a)} \\ u(t) + \displaystyle\int_0^\infty k(x)f(t,x)\mathrm{d}x = \lambda,\ t \geq 0, & \text{(1.7b)} \\ f(0,x) = f^0(x) \text{ and } u(0) = u^0,\ x \geq 0. & \text{(1.7c)} \end{cases}$$

Since the PDE (1.7a) is a transport equation, one may need a proper boundary condition at $x = 0$ depending on the sign of v at the boundary $x = 0$. A sufficient condition for not needing any boundary condition is $a(0)\lambda - b(0) < 0$. This condition is not only technical but also describes a modeling assumption in some cases : the clusters of 'null' size are removed from the system and not considered. The Lifshitz-Slyozov model can be provided with various boundary conditions depending on modeling assumptions. For the case of adipose cell modeling, we want the equation to verify the constraint that the total number of cells is constant i.e. $\int_0^\infty f(t,x)\mathrm{d}x = m$. Hence a sufficient boundary condition is $(v(x,u(t))f(t,x))\,|_{x=0} = 0$.

Lifshitz-Slyozov-Wagner model and Ostwald ripening

There exists another version of the Lifshitz-Slyozov model called the Lifshitz-Slyozov-Wagner model. It was introduced in [131] and it is build upon the assumption that u is close to a critical value λ_{crit} and describes the Lifshitz-Slyozov model in large time. This leads to re-writing the conservation equation (1.7b) as :

$$\int_0^\infty x f(t,x)\mathrm{d}x = \lambda.$$

The PDE stays the same and multiplying it by x and integrating we are led to the following equation for u:

$$u(t)\int_0^\infty a(x)f(t,x)\mathrm{d}x = \int_0^\infty b(x)f(t,x)\mathrm{d}x.$$

In both versions of the model a peculiar phenomenon may occur : Ostwald ripening. Taking the previous description of the Lifshitz-Slyozov model using clusters, this ripening translates to the observation that large clusters will grow at the expense of small clusters. From a thermodynamics stand point, this means that large clusters are energetically favored, and in the case of crystal formation or atom clusters, it can be explained by the fact that the atoms at the center of the cluster are more stable, hence favoring the formation of large clusters. Mathematically this may be observed as a part of the distribution f 'going to infinity'. The appearance of Ostwald ripening can be characterized in the Lifshitz-Slyozov model depending on the rates a and b. Let $\lim_{x\to+\infty} \frac{b(x)}{a(x)} = \lambda_{\text{crit}}$, which represent a critical value of λ. Then in Equation (1.7), Ostwald ripening may occur when $\lambda > \lambda_{\text{crit}}$.

Remark. *We shall see in Section 1.3.3 that our model falls in the sub-critical case $\lambda < \lambda_{\text{crit}}$ since $\lambda_{\text{crit}} = +\infty$.*

As we introduced before, the original paper [84] dates back to 1961, where the rates are expressed as $a(x) = x^{1/3}$ and $b(x) = 1$. From then numerous contributions to the study of the Lifshitz-Slyozov model have been made, particularly since the year 2000. The existence and uniqueness of solutions to the Lifshitz-Slyozov model have been studied in various cases. The work of Collet and Goudon [33] presents existence and uniqueness in the case of C^1 rates with bounded derivatives. The authors use the classical theory of characteristics to show existence of mild solutions, see Definitions 1.4.2 and 1.4.3, and weak solutions, see Definition 1.4.1. Similar techniques have been used in [19], with rates that are non-Lipchitz at the boundary. Convergence result toward stationary solutions has been described in [18], where authors assume that the ratio $\frac{b}{a}$ is decreasing and using Wasserstein distances are able to show that the solution concentrates toward a Dirac mass. A convergence result for the Lifshitz-SlyozovWagner model with self similar solutions was also established in [94].

Naturally the Lifshitz-SlyozovWagner model has also seen developments in parallel to the theory developed for the Lifshitz-Slyozov model. Existence and uniqueness of weak solutions have been studied in [95] in the case of the classical rates $a(x) = x^{1/3}$ and $b(x) = 1$ and have been extended in [77] to rates that are Lipchitz continuous for large x but are are less regular near $x = 0$.

This model has been used in various contexts including biological phenomena, such as prions [79, 115], [54] or modeling in oceanography, see [62].

1.3.3 A Lifshitz-Slyozov type model for adipose cells

We now proceed with the construction of our main model, based on the Lifshitz-Slyozov model. First and foremost we recall the biological simplifications we previously introduced as well as introducing two new physical simplifications :

- The total number of cells, also called *mass* of the distribution is denoted by m and is a constant of time.
- The total amount of lipids is denoted by λ and is a constant of time.
- The amount of lipids in a cell (in nmol), denoted by x, is linearly related to its volume.

- Cells are spheres.

The fact that the amount of lipids in a cell is linearly related to its volume is a fair physical assumption, but in reality this relation may not be exact, and the change in volume due to lipid intake is probably delayed in time. We simplify the shape of a cell to be a sphere for two reasons : the volume of a sphere is easy to compute and when collecting the data, cells are stripped from the extracellular matrix and are therefore free moving inside the medium where they should take a spherical shape and be measured as such.

We base our work on [118] where the authors build a model in which the distribution depends on both the radius and the lipid amount. Our third simplification allows us to relate radii to lipids and therefore simplify the model. Going back to the model in [118], it corresponds to taking the limit $\tau \to 0$ in Equation (1.3). This relation reads as :

$$V_{\text{lipids}} x + V_0 = \frac{4}{3}\pi r(x)^3, \tag{1.8}$$

where V_{lipids} is the volume occupied by 1 nmol of lipids (in particular triglycerides), V_0 is the volume of an empty cell and $r(x)$ is the radius of a cell containing x lipids. Do observe that this simply reads as *volume of the lipid vesicle + volume of the cytoplasm and organelles = volume of the cell.*

This relation is especially important because it allows us to relate lipid amount to radius in a one to one manner, with :

$$r(x) = \left(\frac{3}{4\pi}(V_{\text{lipids}} x + V_0)\right)^{\frac{1}{3}}. \tag{1.9}$$

From now on, when we refer to the size of a cell, we intent to look at the value of x. Then, let us consider a cell of size x and we begin by describing the evolution of its size. As we previously mentioned two processes occur : intake of lipids, known as lipogenesis and release of lipids, known as lipolysis. The change in x is therefore the difference between the rate of intake and the rate of release. We construct these two rates based on biological and physical assumptions.

The intake rate is a product of three factors :

- a term for a surface limited flow $\alpha r(x)^2$, where the constant α is the rate of this flow,
- a term with a radius cutoff ρ such that for greater radius, the intake flow is greatly decreased $\dfrac{\rho^n}{r(x)^n + \rho^n}$,
- a Michaelis-Menten term for the available amount of lipids in the medium $\dfrac{L}{L+\kappa}$, where κ is the value of L at which this term is a half.

And the release rate is a product of two terms :

- a term with a basal level of release β and a surface limited flow $\gamma r(x)^2$, where the constant γ is the release equivalent of the constant α,
- a Michaelis-Menten term for the available amount of lipids in the cell $\dfrac{x}{x+\chi}$, where χ is the equivalent of κ for the release.

We point out that the construction of the release rate are based on biological observations detailed in [119]. In particular, the values of β and γ have been experimentally estimated in [119].

Remark. *We make a slight change of notation from u to L to reflect the fact that L is the external amount of lipids.*

The variation of the size x, which we denote $\dfrac{\mathrm{d}x}{\mathrm{d}t}$ can therefore be expressed as the difference between the intake and release rates as :

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \underbrace{\alpha r(x)^2 \frac{\rho^n}{r(x)^n + \rho^n} \frac{L}{L+\kappa}}_{intake} - \underbrace{(\beta + \gamma r(x)^2)\frac{x}{x+\chi}}_{release}. \tag{1.10}$$

For both rates we regroup the terms in x together as :

$$a(x) = \alpha r(x)^2 \frac{\rho^n}{r(x)^n + \rho^n}, \tag{1.11}$$

$$b(x) = (\beta + \gamma r(x)^2) \frac{x}{x + \chi}. \tag{1.12}$$

Now from a physical point of view, one may see the cells has particles moving in a state space with vector field :

$$v(x, L) = \frac{\mathrm{d}x}{\mathrm{d}t} = a(x) \frac{L}{L + \kappa} - b(x). \tag{1.13}$$

From the theory of continuity equations, this leads us to write a homogeneous transport equation for the distribution of cells $f(t, x)$, which gives the quantity of adipose cells with amount of lipid x at time t :

$$\partial_t f(t, x) + \partial_x (v(x, L) f(t, x)) = 0. \tag{1.14}$$

As per our biological simplification, the total quantity of lipids in our system, denoted by λ should be constant. There are two types of lipids in the system : the ones contained in the cells, and the lipids in the medium. This yields the following equality :

$$L(t) + \int_{\mathbb{R}_+} x f(t, x) \mathrm{d}x = \lambda, \tag{1.15}$$

which is similar to Equation (1.6) with $k(x) = x$.

In regards to boundary conditions, we want to preserve the mass of the distribution and therefore we need to impose that :

$$\int_{\mathbb{R}_+} f(t, x) \mathrm{d}x = \int_{\mathbb{R}_+} f^0(x) \mathrm{d}x = m \text{ for all } t > 0. \tag{1.16}$$

This leads to having the boundary condition $(v(x, L(t)) f(t, x))|_{x=0} = 0$ for all $t > 0$. Hence, because $v(0, L) > 0$, to conserve the mass, a Dirichlet boundary condition is sufficient :

$$f(t, x)\big|_{x=0} = 0.$$

This leads us to the Lifshitz-Slyozov model for adipose cells with initial conditions (f^0, L^0):

$$\begin{cases} \partial_t f(t, x) + \partial_x (v(x, L(t)) f(t, x)) = 0, & \text{(1.17a)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} x f(t, x) \mathrm{d}x = \lambda, & \text{(1.17b)} \\ f(t, x)\big|_{x=0} = 0, & \text{(1.17c)} \\ f(0, x) = f^0(x) \text{ and } L(0) = L^0. & \text{(1.17d)} \end{cases}$$

Stationary solutions

Since we want to observe the bimodality on stationary solutions, we need to describe them. We denote the stationary solutions by M_L and $R_L = \{x \in \mathbb{R}_+ | v(x, L) = 0\}$ the set of roots of the velocity. Now the stationary solutions should verify $\partial_t M_L = 0$ and thus :

$$\partial_x (v(x, L) M_L(x)) = 0 \text{ for all } x > 0, \tag{1.18}$$

where L solves the conservation equation $L + \int_{\mathbb{R}_+} x M_L(x) \mathrm{d}x = \lambda$. This with the boundary conditions yields :

$$v(x, L) M_L(x) = 0 \text{ a.e in } \mathbb{R}_+.$$

As stated before, M_L then takes the form of a linear combination of Dirac masses located at the roots of the velocity. This gives us the expression :

$$M_L(x) = \sum_{y \in R_L} c_y \delta_y(x). \tag{1.19}$$

Novelty of this model

The novelty of this model is the choice of the functions a and b as well as the saturation term $\frac{L}{L+\kappa}$. Classical choices for a and b are constant or rational powers of x, which often yield the existence of a single or two roots. But to obtain bimodality, we need a velocity with at least three roots : two attractive roots, which shall give the local maxima, and one repulsive which will give the nadir. In Chapter 3, we present a numerical scheme for the second order Lifshitz-Slyozov model. Since the classical model yields stationary solutions as sums of Dirac masses, we would like more continuous solutions and thus we use the second order model. However to obtain this second order model, we need to study the first order model. Therefore we are interested in studying the System (1.17). Precisely we are interested in constructing a Becker-Döring model able to reproduce the convergence theorems from the Becker-Döring model to the Lifshitz-Slyozov model.

1.3.4 Why should we concern ourselves with the Becker-Döring model and why is our first model flawed ?

Although we have tried to develop an appropriate model for the size distribution of adipose cells, the Lifshitz-Slyozov model we derived is flawed. Even if all mathematical models are approximations of reality and therefore are 'wrong' in some sense, our model also fails to fill its main mission : its stationary solutions should replicate the data in Figure 1.3. Indeed the PDE (1.17a) is a transport equation with a velocity that alternates in sign. Hence we should expect and actually we observe numerically, that this model concentrates the distribution towards Dirac masses in large times. Therefore comparison between this type of stationary solutions and the data proves to be difficult. As stated before we want to modify our initial model by adding a diffusive term. However, it is not clear if we can construct such a diffusive term using biological and/or physical assumptions. As such we turn ourselves to a model closely related to the Lifshitz-Slyozov model : the Becker-Döring model. In particular, there exists convergence results that in some limit, show that the solution to the Becker-Döring model converges in some sense toward a solution to the Lifshitz-Slyozov model. We will give details about this convergence later on. Nonetheless this convergence proves useful for choosing a diffusion for our model because terms of higher order in the convergence result take the form of a diffusion.

1.3.5 The Becker-Döring model

Becker and Döring introduced this model in 1935 [10] to study nucleation in supersaturated vapors but the equations were later popularized among mathematicians by [98]. Let us describe the Becker-Döring model for polymerization.

Let us consider a medium containing some elementary particle called monomers and these bound together to form polymers. For each $i \in \mathbb{N}^*$, the amount of polymers containing i monomers is denoted by c_i. Remark that c_1 denotes the amount of monomers. A polymer of size i denoted by p_i can gain one monomer and grow to p_{i+1} with rate a_i or loose one monomer and shrink to p_{i-1} with rate b_i :

$$p_1 + p_i \underset{b_{i+1}}{\overset{a_i}{\rightleftharpoons}} p_{i+1}.$$

Let $c = (c_i)_{i\geq 1}$ where c_i is the amount of polymers of size i. For all $i \in \mathbb{N}^*$, we denote $J_i(c)$ the right going flux of the previous reaction, which yields, assuming mass-action laws, $J_i(c) = a_i c_1 c_i - b_{i+1} c_{i+1}$. The evolution of c_i is then given by :

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = J_{i-1}(c) - J_i(c),\ \forall i \geq 2. \tag{1.20}$$

That is to say, the change of the quantity c_i is simply governed by the difference between the net amount of polymers coming in from the size $i-1$ and the net amount growing from size i. The total amount of monomers, i.e free monomers and monomers within polymers, is assumed constant, as in the Lifshitz-Slyozov model, which leads to :

$$c_1(t) + \sum_{i\geq 2} i c_i(t) = \lambda,\ \forall t \geq 0. \tag{1.21}$$

We therefore consider the system :

$$\begin{cases} \dfrac{\mathrm{d}c_i}{\mathrm{d}t} = J_{i-1}(c) - J_i(c),\ \forall i \geq 2, \\ c_1(t) + \sum\limits_{i\geq 2} i c_i(t) = \lambda,\ \forall t \geq 0, \\ c_i(0) = c_i^0,\ \forall i \geq 1. \end{cases} \tag{1.22}$$

This is the Becker-Döring model for polymerization.

Stationary solutions

Stationary solutions of the Becker-Döring equations are given by (see [9]):

$$c_i^{\mathrm{stat}} = Q_i z^i, \text{ where } Q_i = \frac{a_1 a_2 \dots a_{i-1}}{b_1 b_2 \dots b_i} \text{ and } Q_1 = 1.$$

The value of $c_1 = z$ comes down to solving $\sum\limits_{i\geq 1} i Q_i z^i = \lambda$. We denote z_s the convergence radius of the previous sum. Since $\sum\limits_{i\geq 1} i Q_i z_s^i = \lambda_s$ might be finite, this leads to a notion of subcritical and supercritical masses. We refer the reader to [9] for results in the supercritical case, since we will only investigate subcritical masses our case.

Studying the convergence to the stationary solutions is done by using an entropy method where the classical entropy is :

$$H(c) = \sum_{i\geq 1} c_i \Big(\ln \frac{c_i}{Q_i} - 1 \Big). \tag{1.23}$$

Since its introduction in 1935 by Becker and Döring, the model has seen a few major evolutions. Firstly the seminal paper [9] presents a general framework for solutions of the model as well as the study of asymptotic behavior in the case of super-critical and sub-critical initial amount of elementary particles λ. This paved the way for additional work and the characterization of Ostwald ripening for this model. Moreover this work proves that for $\lambda < \lambda_s$ the stationary solution c^{stat} minimizes the entropy and a solution c of the Becker-Döring model converges strongly toward c^{stat}. A convergence result in the subcritical case for physically relevant rates was proved in [21] and a uniform propagation of algebraic and exponential moments was proved in [23].

In regards to the relation between the Lifshitz-Slyozov and Becker-Döring models, first formal observations were made in [98] and there the Lifshitz-Slyozov model can be obtained from the Becker-Döring model in two ways : as a 'macroscopic' limit, which is how we see it in this manuscript, or as a 'large time' limit for the supercritical case. A rigorous mathematical proof of the 'macroscopic' limit, in the spirit of hydrodynamic limit, can be found in [129], while the case of the 'large time' limit is treated in [97]. In [78], a similar result is proven for the Lifshitz-Slyozov-Wagner model.

As for the Lifshitz-Slyozov model, the Becker-Döring model has been used in various contexts, including biology, such as for prion dynamic [39, 106] or oceanography [132]. However up to our knowledge, it has never been used for adipose cell dynamic. In the next section we present an adaptation of the Becker-Döring model to the modeling of adipose cell. The purpose of this model is to exploit convergence results from the Becker-Döring model to the Lifshitz-Slyozov model to obtain a diffusive term to enrich our first model (3.11).

1.3.6 A Becker-Döring type model for adipose cells

In this subpart we will build an analog to the Becker-Döring equations with a speed adapted to the modelling of adipocyte dynamic. The purpose of this construction is to investigate the classical convergence theorems from Becker-Döring to Lifshitz-Slyozov and deduce the form of a diffusion term to add in our model.

We shall now consider that an adipose cell is a bundle of smaller vesicles of typical size Λ. Hence the size of a cell can be defined by the number of vesicles it contains. For all $i \geq 0$, we denote by c_i the number of cells of size i and l the number of vesicles in the medium. Such a cell will aggregate a new vesicle with speed $a_i \frac{l\Lambda}{l\Lambda+\kappa}$ and loose a vesicle at speed b_i, following this reaction :

$$L + iL \underset{b_{i+1}}{\overset{a_i \frac{l\Lambda}{l\Lambda+\kappa}}{\rightleftharpoons}} (i+1)L.$$

Let $c = (c_i)_{i\geq 0}$. The flow $J_i(c,l)$ of the previous reaction is then given by :

$$J_i(c,l) = a_i \frac{l\Lambda}{l\Lambda+\kappa} c_i - b_{i+1} c_{i+1},\ i \geq 0.$$

Similarly, as the Lifshitz-Slyozov model, l will satisfy an equation accounting for conservation of the amount of lipids. We get the following ODE system :

$$\begin{cases} \dfrac{\mathrm{d}c_i}{\mathrm{d}t} = J_{i-1}(c,l) - J_i(c,l),\ \forall i \geq 1, & \text{(1.24a)} \\ \dfrac{\mathrm{d}c_0}{\mathrm{d}t} = -J_0(c,l), & \text{(1.24b)} \\ l(t)\Lambda + \displaystyle\sum_{i=0}^{\infty} i\Lambda c_i(t) = \lambda,\ \forall t \geq 0, & \text{(1.24c)} \\ l(0) = l^0, \quad c_i(0) = c_i^0,\ \forall i \geq 1, & \text{(1.24d)} \end{cases}$$

which is very similar to (1.22) except for the definition of the flux J_i and the difference of definition of c_1, which in the classical Becker-Döring model represented the amount of monomers while in our model is the amount of cells of size one. In our model, it is l that plays the role of the amount of monomers since it is the amount of external available vesicles. Observe also that there is no 'boundary' flux, thus the quantity $m = \sum_{i\geq 0} c_i(t)$ is constant in time. This is analogous to the conservation of the zeroth order moment of f (1.16).

The study of convergence from the Becker-Döring model to the Lifshitz-Slyozov model, is usually performed after rescaling the former, which we do in the next section. There are a variety of theorems on this convergence [78, 129, 35, 34, 110], and we prove a version of such theorems for our slightly modified Becker-Döring model fo adipose cells in appendix B.

1.3.7 Rescaled Becker-Döring model

To study convergence from the Becker-Döringmodel to the Lifshitz-Slyozov model, one needs to introduce a rescaling quantity in the Becker-Döring model. This process can be done in a variety of ways, and some may lead to different version of the Lifshitz-Slyozov model. Particularly, some rescaling procedures are well tailored to study Ostwald ripening. However in our case, we know that we are in the subcritical case, hence we perform a very simple rescaling, akin to the one in [35, 129]. We introduce the following scaling constants :

- $\bar{A}$ rescaling value of $(a_i)_{i\geq 0}$,
- $\bar{B}$ rescaling value of $(b_i)_{i\geq 1}$,
- $\bar{C}$ rescaling value of $(c_i)_{i\geq 0}$,
- $\bar{T}$ rescaling value of the time scale,
- $\bar{\lambda}$ rescaling value of λ.

We previously denoted by Λ the typical size of a vesicle. Hence it plays the role of a rescaling value and should be treated as so. Now, we introduce the rescaled variables :

$$\bar{a_i} = \frac{a_i}{\bar{A}},\ \forall i \geq 0,$$

$$\bar{b_i} = \frac{b_i}{\bar{B}},\ \forall i \geq 1,$$

$$\bar{t} = \frac{t}{\bar{T}},$$

$$\bar{c_i}(\bar{t}) = \frac{c_i(\bar{t}\bar{T})}{\bar{C}},\ \forall i \geq 0,$$

$$\bar{L}(\bar{t}) = l(\bar{t}\bar{T})\Lambda.$$

The quantity $\bar{L}$ therefore describes the total amount of lipids in the medium instead of the number of lipid vesicles. To have the proper convergence for the functions a and b later on, we relate the rescaling variables $\bar{A}$ and $\bar{B}$ with the constant in assumption (H4) :

$$\bar{A} = A \text{ and } \bar{B} = B.$$

We compute from Equation (1.24) the derivative of $\bar{c}_i$ for $i \geq 1$:

$$\begin{aligned}\frac{\mathrm{d}\bar{c}_i}{\mathrm{d}\bar{t}}(\bar{t}) &= \frac{\bar{T}}{\bar{C}}\frac{\mathrm{d}c_i}{\mathrm{d}t}(\bar{t}\bar{T}) \\ &= \frac{\bar{T}}{\bar{C}}\Big(a_{i-1}\frac{l(\bar{t}\bar{T})\Lambda}{l(\bar{t}\bar{T})\Lambda+\kappa}c_{i-1}(\bar{t}\bar{T}) - (a_i\frac{l(\bar{t}\bar{T})\Lambda}{l(\bar{t}\bar{T})\Lambda+\kappa} + b_i)c_i(\bar{t}\bar{T}) + b_{i+1}c_{i+1}(\bar{t}\bar{T})\Big) \\ &= \bar{A}\bar{T}\Big(\bar{a}_{i-1}\frac{\bar{L}(\bar{t})}{\bar{L}(\bar{t})+\kappa}\bar{c}_{i-1}(\bar{t}) - \bar{a}_i\frac{\bar{L}(\bar{t})}{\bar{L}(\bar{t})+\kappa}\bar{c}_i(\bar{t})\Big) - \bar{B}\bar{T}\Big(\bar{b}_i\bar{c}_i(\bar{t}) - \bar{b}_{i+1}\bar{c}_{i+1}(\bar{t})\Big).\end{aligned}$$

The derivative of $\bar{c}_0$ writes as :

$$\frac{\mathrm{d}\bar{c}_0}{\mathrm{d}t}(\bar{t}) = -\bar{A}\bar{T}\bar{a}_0\frac{\bar{L}(\bar{t})}{\bar{L}(\bar{t})+\kappa}\bar{c}_0(\bar{t}) + \bar{B}\bar{T}\bar{b}_1\bar{c}_1(\bar{t})$$

and the conservation equation for lipids as :

$$\bar{L}(\bar{t}) + \bar{C}\Lambda\sum_{i\geq 1} i\bar{c}_i(\bar{t}) = \lambda.$$

We now relate all the rescaling constants to a single variable $\varepsilon > 0$, such that :

$$\bar{A}\bar{T} = \bar{B}\bar{T} = \frac{1}{\varepsilon} \text{ and } \bar{C}\Lambda = \varepsilon^2.$$

Remark. *Depending on the process we are trying to model, the interpretation of this rescaling may vary. For adipose cells, this can be seen as if a lot of reactions are happening :* $\bar{A}\bar{T} = \bar{B}\bar{T} = \frac{1}{\varepsilon}$*, but each reaction step is relatively small :* $\bar{C}\Lambda = \varepsilon^2$*. In other words, the size of the individual vesicle is small.*

At last, we drop the bar above the variables and replace it with ε as superscript to show the dependency of the solution on ε.

We define $c^\varepsilon = (c_i^\varepsilon)_{i\geq 0}$ and $J_i^\varepsilon(c^\varepsilon, L^\varepsilon)$ the flow of the rescaled reaction given by :

$$J_i^\varepsilon(c^\varepsilon, L^\varepsilon) = a_i^\varepsilon\frac{L^\varepsilon}{L^\varepsilon+\kappa}c_i^\varepsilon - b_{i+1}^\varepsilon c_{i+1}^\varepsilon,\ i \geq 0. \tag{1.25}$$

Similarly as before, see Eq. (1.17b), L^ε will satisfy an equation accounting for conservation of the amount of lipids and we get the following ODE system :

$$\begin{cases} \dfrac{\mathrm{d}c_i^\varepsilon}{\mathrm{d}t} = \dfrac{1}{\varepsilon}(J_{i-1}^\varepsilon(c^\varepsilon, L^\varepsilon) - J_i^\varepsilon(c^\varepsilon, L^\varepsilon)),\ \forall i \geq 1, & \text{(1.26a)} \\ \dfrac{\mathrm{d}c_0^\varepsilon}{\mathrm{d}t} = -\dfrac{1}{\varepsilon}J_0^\varepsilon(c^\varepsilon, L^\varepsilon), & \text{(1.26b)} \\ L^\varepsilon(t) + \displaystyle\sum_{i=0}^{\infty} i\varepsilon^2 c_i^\varepsilon(t) = \lambda,\ \forall t \geq 0, & \text{(1.26c)} \\ L^\varepsilon(0) = L^{\varepsilon,0}, \quad c_i^\varepsilon(0) = c_i^{\varepsilon,0},\ \forall i \geq 1. & \text{(1.26d)} \end{cases}$$

Remark. *The rescaled Becker-Döring model already bears strong resemblance to the Lifshitz-Slyozov model. As a first observation, on may observe that Equation* (1.26a) *resembles a Finite Volume numerical scheme. Additionally, by looking at Equation* (1.25)*, and assuming that* $c_i^\varepsilon(t) \simeq c_{i+1}^\varepsilon(t)$*, one may observe that* $J_i^\varepsilon(c^\varepsilon, L^\varepsilon) \simeq v(i\varepsilon, L^\varepsilon)c_i^\varepsilon(t)$ *thus resembling the transport term in Equation* (1.17a)*. Moreover the conservation Equation* (1.26c) *can be seen as an approximation of Equation* (1.17b)*.*

1.3.8 Diffusive Lifshitz-Slyozov model or Second order Lifshitz-Slyozov model

In the case of adipose cell modeling, by looking at Figure 1.3, one can see that the cell population does not concentrate on two points but instead cells take a large range of sizes centered around two modes. Thus looking at stationary solutions of (1.17) is not particularly relevant, especially if we want to perform parameter estimation. Hence, we need a proper way to smooth those stationary distributions. This can be done by adding a diffusive term to the PDE in (1.17). In some cases, one may make additional modeling assumptions motivated by the model to describe this diffusive term. However in the case of adipose cells, there is no biological insights as to how such a term should look like. Hence we use a convergence result from the Becker-Döring model, described in Section 1.4.4, toward the Lifshitz-Slyozov model, where, by looking at higher order terms in the convergence, one can get a sense of what type of diffusion to expect. Again this choice of diffusive term is not supported by any modeling assumptions but simply by the intrinsic diffusive properties of the microscopic model, which disappear at the macroscopic scale when we retain only the first order term. Henceforth this new diffusive model shall be called the diffusive or second-order Lifshitz-Slyozov model :

$$\begin{cases} \partial_t g(t,x) + \partial_x(v(x,u(t))g(t,x)) = \dfrac{\varepsilon}{2}\partial_x^2((a(x)u(t)+b(x))g(t,x)),\ t\geq 0, x\geq 0, & (1.27)\\ u(t) + \displaystyle\int_0^\infty k(x)g(t,x)\mathrm{d}x = \lambda,\ t\geq 0, & (1.28)\\ g(0,x) = g^0(x) \text{ and } u(0) = u^0,\ x\geq 0. & (1.29) \end{cases}$$

The notation f is changed to g to make a distinction between diffusive and non-diffusive Lifshitz-Slyozov model. We retain the notation u for the general case, and will change back to L for the case of modeling the adipose cells. Here ε is the scaling parameter in the convergence result, and is assumed to be small and positive. We denote by d the diffusion rate :

$$d(x,u) = a(x)u + b(x).$$

For this equation to hold, we need d to remain strictly positive for all u and x. We will later detail the minimal assumptions we make on a and b, but for the case of adipose cell modeling the only case where the diffusion rate might equal zero is when $L = 0$ at the point $x = 0$. However, a simple domain exit argument ensures that if $L^0 > 0$, this does not happen in finite time.

In this case, continuous stationary solutions can be computed explicitly by letting $\partial_t g = 0$ for all $x \geq 0$. We denote these stationary solutions M_u (as in Maxwellians) and they verify :

$$\begin{cases} M_u(x) = \dfrac{C_{\mathrm{BC}}}{d(x,u)}\exp\left(\dfrac{2}{\varepsilon}\displaystyle\int_0^x \frac{v(y,u)}{d(y,u)}\mathrm{d}y\right),\ x\geq 0, & (1.30a)\\ u + \displaystyle\int_0^\infty k(x)M_u(x)\mathrm{d}x = \lambda. & (1.30b) \end{cases}$$

The constant $C_{\mathrm{BC}} > 0$ depends on the boundary condition imposed on the model, which the stationary solutions verify.

Up to our knowledge, one of the first observation of a relation between birth and death processes and Fokker-Planck dynamic is done in [49]. The correlation between the Becker-Döring model and the Lifshitz-Slyozov model was later established formally in [130] and [57] using Taylor expansion of the flux to second order, hence the name of this modified Lifshitz-Slyozov model. Other derivations in the spirit of diffusive numerical scheme were detailed in [129] and more in depth in [120], where a proper correlation between the two models is derived via a purely diffusive equation.

Diffusive Lifshitz-Slyozov type model for adipose cells

We described the diffusive Lifshitz-Slyozov model in its generality previously. We go into a bit more details for the case of adipose cells. The diffusion rate takes the form :

$$d(x,L) = a(x)\frac{L}{L+\kappa} + b(x).$$

Concerning boundary conditions, we consider null-flux boundary conditions which ensures the conservation of the total number of cells $m = \int_0^\infty g(t,x)\mathrm{d}x$. It goes :

$$\Big(v(x,L(t))g(t,x) - \frac{\varepsilon}{2}\partial_x(d(x,L(t))g(t,x))\Big)\Big|_{x=0} = 0. \tag{1.31}$$

This leads us to our second model for adipose cells size dynamic :

$$\begin{cases} \partial_t g(t,x) + \partial_x(v(x,L(t))g(t,x)) = \dfrac{\varepsilon}{2}\partial_x^2(d(x,L(t))g(t,x)),\ t\geq 0,\ x\geq 0, & \text{(1.32a)}\\ L(t) + \displaystyle\int_0^{+\infty} xg(t,x)\mathrm{d}x = \lambda,\ t\geq 0, & \text{(1.32b)}\\ \Big(-v(\cdot,L(t))g(t,\cdot) + \dfrac{\varepsilon}{2}\partial_x(d(\cdot,L(t))g(t,\cdot))\Big)\Big|_{x=0} = 0,\ t\geq 0, & \text{(1.32c)}\\ g(0,x) = g^0(x) \text{ and } L(0) = L^0,\ x\geq 0. & \text{(1.32d)} \end{cases}$$

All the models we have introduced in this section and the ones from the next section are recapped in Figure 1.13.

1.3.9 Stochastic models

In the next chapter, we highlight the main results of this thesis, one of which deals with stochastic equivalents of the Lifshitz-Slyozov and Becker-Döring models. This section describes the construction of such models. Henceforth all definitions hold on a given probabilistic space $(\Omega, \mathcal{F}, \mathbb{P})$. The expectation on this probability space is denoted $\mathbb{E}$. We begin by a brief recollection of some results on continuous-time Markov chains (CTMC) and stochastic differential equations (SDE).

The Poisson counting process on the real line : the most basic counting process

First and foremost we begin with a simple definition of a counting process :

Definition 1.3.1. *A counting process is a nonnegative, integer-valued, increasing stochastic process that is used to count the number of occurrences of some event as time goes by.*

In the next part and in Chapter 4, we make use of properties of the Poisson counting process. We recall its definition here, as well as some well known facts for clarity.

Definition 1.3.2. *Let $\lambda > 0$ and consider a counting process $(N(t))_{t\geq 0}$. This counting process is called a **Poisson (counting) process** or **Poisson point process** with **rate** λ if the following holds true:*

- $N(0) = 0$,
- $N(t)$ *has independent increments,*
- *the number of arrivals in any interval of length $s > 0$ has $\mathcal{P}(\lambda s)$ distribution.*

Another definition of a Poisson counting process can be written using the small o notation. It is strictly equivalent to Definition 1.3.2.

Definition 1.3.3. *Let $\lambda > 0$ and consider a counting process $(N(t))_{t\geq 0}$. This counting process is called a **Poisson (counting) process** or **Poisson point process** with **rate** λ if the following holds true:*

- $N(0) = 0$,
- $N(t)$ *has independent and stationary increments,*
- *let $s \geq 0$, we have*

$$\begin{aligned} &\mathbb{P}(N(s) = 0) = 1 - \lambda s + o(s),\\ &\mathbb{P}(N(s) = 1) = \lambda s + o(s),\\ &\mathbb{P}(N(s) \geq 2) = o(s). \end{aligned}$$

Now in some cases, one may want to consider a Poisson counting process where the rate λ depends on time. This gives rise to the **inhomogeneous Poisson counting process**.

Definition 1.3.4. *Let $\lambda : \mathbb{R} \to \mathbb{R}_+$ locally integrable and consider a counting process $(N(t))_{t\geq 0}$. This counting process is called an **inhomogeneous Poisson counting process** or **Poisson point process** with **rate** λ if the following holds true:*

- $N(0) = 0$,
- $N(t)$ *has independent increments,*
- *let $s \geq 0$, we have*

$$\mathbb{P}(N(t+s) - N(t) = 1) = \lambda(t)s + o(s),$$
$$\mathbb{P}(N(t+s) - N(t) > 1) = o(s).$$

One can show that by using a proper monotone transformation, a **unit Poisson process**, that is to say a Poisson process with rate $\lambda = 1$ can be transformed into an inhomogeneous Poisson process with rate $\lambda : \mathbb{R} \to \mathbb{R}_+$, as shown in the following lemma.

Lemma 1.3.1. *Let $(N(t))_{t\geq 0}$ be a unit Poisson process, and $\lambda : \mathbb{R} \to \mathbb{R}_+$ a locally integrable function. We denote $\Lambda(t) = \int_0^t \lambda(s)\mathrm{d}s$. Then $M(t) = N(\Lambda(t))$ is an inhomogeneous Poisson process with rate λ.*

Continuous time Markov chains

For clarity, let us consider a state space $\mathcal{S} = \mathbb{Z}$ and a stochastic process $(X_t)_{t\geq 0}$ evolving as follows. Consider a time $t \geq 0$ and a state $i \in \mathcal{S}$ such that $X_t = i$. Additionally consider a *holding time* as a random variable T_i describing the length of time spent in state i by the stochastic process. Then after T_i time has passed, the stochastic process jumps from state i to a state $j \in \mathcal{S}$ with probability $P_{i,j}$. Then the stochastic process $(X_t)_{t\geq 0}$ is a Continuous Time Markov Chain (CTMC) if it verifies the following property: *the future given the present is independent of the past.* Or in mathematical terms:

Definition 1.3.5. *A **continuous-time Markov chain** with finite or countable state space S is a family $\{X_t = X(t)\}_{t\geq 0}$ of S-valued random variables such that*

(a) The paths $t \to X(t)$ are right-continuous step functions,

(b) For any set of times $0 = t_0 < t_1 < t_2 < \dots$ and for any $i, j \in S$

$$\mathbb{P}(X(t_{k+1}) = j | X(t_k) = i, \{X(t_l) \ : \ l =< k\}) = \mathbb{P}(X(t_{k+1}) = j | X(t_k) = i) \tag{1.33}$$

Additionally let us denote $\mathbf{P}_{i,j}(t) = \mathbb{P}(X(t) = j | X(0) = i)$.

Naturally if $(X_t)_{t\geq 0}$ is a CTMC in the sense of Definition 1.3.5 it implies that the holding times $(T_i)_{i\in\mathcal{S}}$ have some good properties. In particular they should be independent of past times and only depend on the state i. Hence the holding times $(T_i)_{i\in\mathcal{S}}$ are memory-less and are exponentially distributed. Then a CTMC is only described by a transition matrix $P = (P_{i,j\in\mathcal{S}})$, which describes which jumps happen and a set of rates $\{\mu_i \ : \ i \in \mathcal{S}\}$ and the holding times are independent exponential random variables with $\mathbb{E}[T_i] = \dfrac{1}{\mu_i}$ for all $i \in \mathcal{S}$.

By letting τ_n be the time of the n-th jump of the CTMC, we can construct an embedded Markov chain. Let $X_n = X(\tau_n+)$ be the state after the n-th jump. Then $(X_n)_{n\in\mathbb{N}}$ is the embedded Markov Chain describing the state change of the CTMC and we have $P_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i)$. Hence to simulate the CTMC, one only needs to simulate the embedded Markov chain and the holding times $(T_i)_{i\in\mathcal{S}}$.

Remark. *Beware of notations ! $\mathbf{P}$ is defined in Definition 1.3.5, and $\mathbf{P}_{i,j}(t)$ is the probability of the CTMC to be in state j after t amount of time knowing that it is in state i at the current time, whereas P is the transition matrix of the embedded Markov chain which describes how the CTMC changes from one state to another when a jump occurs.*

Similarly as other stochastic processes, a CTMC has an infinitesimal generator called the transition rate matrix Q which describes the CTMC. To construct this infinitesimal generator we first consider a state $i \in \mathcal{S}$. Now the CTMC will leave this state in the next h amount of time with probability $\mu_i h$. Here h is to be understood as an 'infinitesimal' amount of time. Looking at the evolution of $\mathbf{P}$ over an 'infinitesimal' amount of time amounts to computing its derivative $\mathbf{P}'_{i,j}(0) = \lim_{h \searrow 0} \frac{\mathbf{P}_{i,j}(h) - \mathbf{P}_{i,j}(0)}{h}$. First, we need a rigorous definition for 'infinitesimal' amount of time. To do so we use the small o notation in Definition 1.3.3. From now on, we make the assumption that $P_{i,i} = 0$ et $\mathbf{P}(0) = Id$, which are usual assumptions.

For all $i \in \mathcal{S}$, let us denote by N_i a Poisson counting process with rate a_i. Recall Definition 1.3.3 and that $P_{i,j}$ is the transition probability from state i to state j. Then let $j \neq i$ and get

$$\mathbf{P}'_{i,j}(0) = \lim_{h \searrow 0} \frac{\mathbf{P}_{i,j}(h)}{h} = \lim_{h \searrow 0} \frac{\mathbb{P}(N_i(h) = 1)}{h} P_{i,j} = \mu_i P_{i,j}. \tag{1.34}$$

Similarly for $j = i$, we get

$$\mathbf{P}'_{i,i}(0) = \lim_{h \searrow 0} \frac{\mathbf{P}_{i,i}(h) - 1}{h} = \lim_{h \searrow 0} \frac{-\mathbb{P}(N_i(h) = 1)}{h} = -\mu_i. \tag{1.35}$$

Definition 1.3.6. *The matrix $Q = P'(0)$ given by* (1.34) *and* (1.35) *is called the transition rate matrix or infinitesimal generator of the CTMC.*

We now proceed with a few well known results of CTMCs. First the matrix family $(\mathbf{P}(t))_{t \in \mathbb{R}_+}$ has the semi-group property.

Lemma 1.3.2 (Chapman-Kolmogorov equation for CTMCs)**.** *The matrix family $(\mathbf{P}(t))_{t \in \mathbb{R}_+}$ verifies the following property for all $t, s \geq 0$:*

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s). \tag{1.36}$$

To compute $\mathbf{P}(t)$ we can use the Kolmogorov equations :

Proposition 1.3.1 (Kolmogorov backward and forward equations)**.** *For a CTMC with infinitesimal generator Q, the following equations are satisfied by $\mathbf{P}$:*

$$\frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t} = Q\mathbf{P}(t) \tag{1.37}$$

$$\frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t} = \mathbf{P}(t)Q \tag{1.38}$$

Equation (1.37) *is called the* ***Kolmogorov backward equation*** *and Equation* (1.38) *is called the* ***Kolmogorov forward equation****.*

Constructing a CTMC from the Becker-Döring model

Let us consider N adipose cells which size is described by a family of random variables $X^N(t) = (X^{i,N}(t))_{i=1,\dots,N} \in \mathbb{N}^N$ that evolve over time. Our intent is to use the uniform propagation of chaos and look at the limit of $X^{1,N}$ when N tends to ∞.

We assume that $(X^N(t))_{t \geq 0}$ is a CTMC. The Markov property is fair : the evolution of the size of a cell only depends on its current size and the current size of all other cells and not on what happened in the past. Indeed the dynamic of L are provided by using the empirical mean :

$$M^{X,N}(t) = \frac{1}{N} \sum_{i=1}^{N} X^{i,N}(t).$$

Then we define :

$$L^{X,N}(t) = \lambda - M^{X,N}(t).$$

Let $i \in \mathbb{N}$ and assume that the k-th cell is in state i at time t. Then $X^{k,N}$ jumps according to

$$\begin{cases} i \to i+1 \text{ at rate } a(i)\dfrac{L^{X,N}(t)}{L^{X,N}(t)+\kappa},\ i \geq 0 \\ i \to i-1 \text{ at rate } b(i),\ i \geq 1 \end{cases}$$

This means that, given $t, h > 0$, the following holds:

$$\mathbb{P}(X^{k,N}(t+h) - X^{k,N}(t) = 1 | X^{k,N}(t)) = a(X^{k,N}(t))\frac{L^{X,N}(t)}{L^{X,N}(t)+\kappa}h + o(h),$$
$$\mathbb{P}(X^{k,N}(t+h) - X^{k,N}(t) = -1 | X^{k,N}(t)) = b(X^{k,N}(t))h + o(h).$$

Hence, to describe $X^{k,N}$ we only need to consider two counting processes R^k_+ and R^k_-, where $R^k_+(t)$ counts the number of positive jumps up to time t and $R^k_-(t)$ counts the number of negative jumps. Hence the intensity of each counting process is given by :

$$\mathbb{P}(R^k_+(t+h) - R^k_+(t) = 1 | X(t)) = a(X^{k,N}(t))\frac{L^{X,N}(t)}{L^{X,N}(t)+\kappa}h + o(h),$$
$$\mathbb{P}(R^k_-(t+h) - R^k_-(t) = 1 | X(t)) = b(X^{k,N}(t))h + o(h).$$

We write :

$$X^{k,N}(t) = X^{k,N}(0) + R^k_+(t) - R^k_-(t).$$

In particular, we can write

$$R^k_+(t) = Y^k_+\Big(\int_0^t a(X^{k,N}(s))\frac{L^{X,N}(s)}{L^{X,N}(s)+\kappa}\mathrm{d}s\Big) \text{ and } R^k_-(t) = Y^k_-\Big(\int_0^t b(X^{k,N}(s))\mathrm{d}s\Big)$$

for some unit Poisson counting processes Y_+ and Y_-, independent from each other. Observe that Lemma 1.3.1 implies that R_+ and R_- are non-linear inhomogeneous Poisson processes with rates $a(X^{k,N}(t))\frac{L^{X,N}(t)}{L^{X,N}(t)+\kappa}$ and $b(X^{k,N}(t))$. This yields :

$$X^{k,N}(t) = X(0) + Y^k_+\Big(\int_0^t a(X^{k,N}(s))\frac{L^{X,N}(s)}{L^{X,N}(s)+\kappa}\mathrm{d}s\Big) - Y^k_-\Big(\int_0^t b(X^{k,N}(s))\mathrm{d}s\Big). \tag{1.39}$$

Inspired by the uniform propagation of chaos, see theorem 1.2 in [122], we define the Becker-Döring non-linear CTMC $X(t)$ by :

$$X(t) = X(0) + Y_+\Big(\int_0^t a(X(s))\frac{L^X(s)}{L^X(s)+\kappa}\mathrm{d}s\Big) - Y_-\Big(\int_0^t b(X(s))\mathrm{d}s\Big), \tag{1.40}$$

where $L^X(t) = \lambda - \mathbb{E}[X(t)]$. In some sense the Becker-Döring non-linear CTMC $X(t)$ describes the evolution of a typical stochastic cell or polymer from the Becker-Döring model.

Rescaling

To be consistent with the rescaled Becker-Döring model (1.26), we also rescale the Becker-Döring non-linear CTMC. We consider that the size of the cell $X(t)$ makes jumps of size ε following :

$$\begin{cases} i\varepsilon \to (i+1)\varepsilon \text{ at rate } \varepsilon^{-1}a(i\varepsilon)\dfrac{L^X_\varepsilon(t)}{L^X_\varepsilon(t)+\kappa},\ i \geq 0, \\ i\varepsilon \to (i-1)\varepsilon \text{ at rate } \varepsilon^{-1}b(i\varepsilon),\ i \geq 1, \end{cases}$$

where $L^X_\varepsilon(t) = \lambda - \mathbb{E}[X_\varepsilon(t)]$. This rescaled Becker-Döring non-linear CTMC verifies :

$$\begin{cases} X_\varepsilon(t) = X_\varepsilon(0) + \varepsilon Y_+\Big(\varepsilon^{-1}\displaystyle\int_0^t a(X_\varepsilon(s))\frac{L^X_\varepsilon(s)}{L^X_\varepsilon(s)+\kappa}\mathrm{d}s\Big) - \varepsilon Y_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big), & (1.41) \\ L^X_\varepsilon(t) = \lambda - \mathbb{E}[X_\varepsilon(t)]. & (1.42) \end{cases}$$

Henceforth, we assume that a solution to Equation (1.41) exists. Also observe that since $b(0) = 0$ a solution to Equation (1.41) will remain non-negative at all time.

The classical stochastic Becker-Döring model

Our version of the stochastic Becker-Döring model is a bit different from what exists in the literature. We present this model in the case of polymerization but its relation to the previous non-linear Becker-Döring CTMC is actually quite simple and extending it to the case of adipose cells is trivial. Again, the object of interest is a Markov chain in $\mathbb{N}^{\mathbb{N}}$, $C(t) = (C_i(t))_{i\in\mathbb{N}}$, where each C_i corresponds to the amount of polymers of size i. Two types of reactions can occur : a polymer of size i grows to size $i+1$ by aggregating a monomer or shrinks to size $i-1$ and releases a monomer.

The state space is given by $X_M = \{C \in \mathbb{N}^{\mathbb{N}} | \sum_{i=1}^{M} iC_i = M\}$, on which the following operators are introduced. For $i = 1$, we define :

$$R_1^+ C = (C_1 - 2, C_2 + 1, \dots, C_i, \dots),$$

$$R_2^- C = (C_1 + 2, C_2 - 1, \dots, C_i, \dots),$$

and for all $i \geq 2$, we set :

$$R_i^+ C = (C_1 - 1, \dots, C_i - 1, C_{i+1} + 1, \dots),$$

$$R_{i+1}^- C = (C_1 + 1, \dots, C_i + 1, C_{i+1} - 1, \dots).$$

The reaction R_1^+ is slightly different from other reactions since it corresponds to the polymerization of two monomers together. Its rate is therefore different : a first monomer is chosen from the C_1 available and a second is chosen from the $C_1 - 1$ remaining. Then the transition rates of the CTMC $C(t)$ are given by :

$$\begin{cases} Q(C, R_1^+ C) = a_1 C_1 (C_1 - 1), \\ Q(C, R_i^+ C) = a_i C_1 C_i, \quad i \geq 2, \\ Q(C, R_i^- C) = b_i C_i, \quad i \geq 2. \end{cases}$$

We have an alternative representation of the CTMC as a solution to :

$$\begin{cases} C_1 = C_1^0 - 2J_1(t) - \sum_{i\geq 2} J_i(t), \\ C_i = C_i^0 + J_{i-1}(t) - J_i(t), \quad i \geq 2, \end{cases}$$

where $C^0 \in X_M$ is an initial condition, and the fluxes J_i are defined using Y_i^+ and Y_{i+1}^-, two independent standard Poisson processes :

$$J_i(t) = Y_i^+ \left(\int_0^t a_i C_1(s)(C_i(s) - \delta_1^i) \mathrm{d}s \right) - Y_{i+1}^- \left(\int_0^t b_{i+1} C_{i+1}(s) \mathrm{d}s \right).$$

Various results have been shown about this model, in particular the existence of a stationary probability distribution in product-form [70]

$$\Pi(C) = B_M \prod_{i=1}^{M} \frac{(Q_i)^{C_i}}{C_i!},$$

where B_M is a normalizing constant defined recursively. Other interesting results include the law of large numbers [64] which in itself provides an existence result for solutions of the deterministic Becker-Döring model.

Remark. *One may go from the Becker-Döring non-linear CTMC to the stochastic Becker-Döring model by simply grouping together cells of the same size :* $C_i(t) = \sum_{k=1}^{N} \mathbb{1}_i(X^{k,N}(t))$.

Fokker-Planck PDE and corresponding SDE

Again, let us consider N adipose cells whose size is described by a family of random variables $Z_\varepsilon^N(t) = (Z_\varepsilon^{k,N}(t))_{k=1,\dots,N} \in \mathbb{N}^N$ that evolve over time. Our intent is to use the uniform propagation of chaos and look at the limit of $Z^{k,N}$ when N tends to ∞.

We define the empirical mean :

$$M^{Z,N}(t) = \frac{1}{N}\sum_{k=1}^{N} Z_\varepsilon^{k,N}(t)$$

Consider the k-th cell. The evolution of its size is given by :

$$\mathrm{d}Z_\varepsilon^{k,N}(t) = v(Z_\varepsilon^{k,N}(t), L_\varepsilon^Z(t))\mathrm{d}t + \sqrt{\varepsilon a(Z_\varepsilon^{k,N}(t))\frac{L_\varepsilon^{Z,N}(t)}{L_\varepsilon^{Z,N}(t)+\kappa}}\mathrm{d}W_t^{+,k} + \sqrt{\varepsilon b(Z_\varepsilon^{k,N}(t))}\mathrm{d}W_t^{-,k},\ t \geq 0, \tag{1.43}$$

where $L^{Z,N}(t) = \lambda - M^{Z,N}(t)$ and $W_t^{\pm,k}$ are standard Wiener processes independent from one another. This definition is inspired by the fact that equation (1.32a) is a non-linear Fokker-Planck equation which stochastic differential equation is in the form of (1.43). We split the stochastic part into two terms depending on W^+ and W^- to simplify the comparison with the Becker-Döring non-linear CTMC. Again inspired by the propagation of chaos, we define the stochastic Lifshitz-Slyozov model by :

$$\mathrm{d}Z_\varepsilon(t) = v(Z_\varepsilon(t), L_\varepsilon^Z(t))\mathrm{d}t + \sqrt{\varepsilon a(Z_\varepsilon(t))\frac{L_\varepsilon^{Z}(t)}{L_\varepsilon^{Z}(t)+\kappa}}\mathrm{d}W_t^{+} + \sqrt{\varepsilon b(Z_\varepsilon(t))}\mathrm{d}W_t^{-},\ t \geq 0, \tag{1.44}$$

where $L_\varepsilon^Z(t) = \lambda - \mathbb{E}[Z_\varepsilon(t)]$. Under some technical assumptions a solution to Equation (1.44) also verifies :

$$Z_\varepsilon(t) = Z_\varepsilon(0) + \varepsilon B_+\Big(\varepsilon^{-1}\int_0^t a(Z_\varepsilon(s))\frac{L_\varepsilon^Z(s)}{L_\varepsilon^Z(s)+\kappa}\mathrm{d}s\Big) - \varepsilon B_-\Big(\varepsilon^{-1}\int_0^t b(Z_\varepsilon(s))\mathrm{d}s\Big), \tag{1.45}$$

where $B_\pm(t) = t + W^\pm(t)$ are drifted Wiener processes and W^- and W^+ are independent Wiener processes. This is the stochastic Lifshitz-Slyozov model. Equation (1.45) is particularly useful to compare Z_ε and X_ε.

Remark. *To ensure the well-posedness of equation* (1.44)*, we need to define the functions a and b on $\mathbb{R}$. To do so we simply set $b(x) = 0$ for $x < 0$ and we know that there exists some $y < 0$ such that $a(y) = 0$, therefore we define $a(x) = 0$ for $x < y$ and extend the definition of a from $\mathbb{R}_+$ to $[y, +\infty)$.*

However the existence of a solution to equation (1.44) *that also verifies equation* (1.45) *is not guaranteed with our choice of a and b. Indeed following lemma 3.7 in [73], this can be achieved with a and b such that $\sqrt{a}$ and $\sqrt{b}$ are Lipchitz functions. This is the case for our choice of a but not for b. We shall see later on that for theorem 1.5.2 to hold we need b to be bounded which is also not the case for our choice of b. Hence we will make an additional assumption that is not relevant to adipocyte modeling but which allows for those results to hold.*

Another possibility for the well-posedness of equation (1.44)*, would have been to include a reflecting barrier condition that ensures that the process stays in $\mathbb{R}_+$. Such a method can be found in [87], with the main idea being to add the local-time term to equation* (1.44)*. This local time process tracks the cumulative amount of pushing at the boundary required to keep $Z - \varepsilon$ from reaching negative values. Writing equation* (1.44) *in its integral form this would look like :*

$$Z_\varepsilon(t) = Z_\varepsilon(0) + \int_0^t v(Z_\varepsilon(s), L_\varepsilon^Z(s))\mathrm{d}s + \int_0^t \sqrt{\varepsilon a(Z_\varepsilon(s))\frac{L_\varepsilon^Z(s)}{L_\varepsilon^Z(s)+\kappa}}\mathrm{d}W_s^+ + \int_0^t \sqrt{\varepsilon b(Z_\varepsilon(s))}\mathrm{d}W_s^- + \sqrt{\varepsilon}\mathcal{T}_t, \tag{1.46}$$

where $\mathcal{T}_t$ is the previously mentioned local-time. However using this method, it is unclear how a solution to equation (1.46) *is related to the diffusive Lifshitz-Slyozov model* (1.32) *and particularly to its boundary condition.*

Remark. *The theoretical results concern X_ε, Z_ε and Equations* (1.41) *and* (1.45). *However we are unable to directly simulate them. Hence we use Equations* (1.39) *and* (1.43) *for numerical simulations, which are systems of mean-field particles, see for example the work in [15].*

1.3.10 Constant diffusion model

In Chapter 5, we present a work where we perform parameter estimation on data from rats. We consider a different model for this work, where we use a constant diffusion rate $D > 0$ and the sizes of cells are radii instead of amounts of lipids :

$$\begin{cases} \partial_t f(t,r) + \partial_r(v(r,L(t))f(t,r)) - D\partial_r^2 f(t,r) = 0, & (1.47)\\ L(t) = \lambda - \displaystyle\int_{r_{min}}^{r_{max}} (V(r) - V_{em})\frac{4\pi r^2}{V_\ell^2} f(t,r)dr, & (1.48)\\ v(r_{min},L(t))f(t,r_{min}) - D\partial_r f(t,r_{min}) = 0, & (1.49)\\ v(r_{max},L(t))f(t,r_{max}) - D\partial_r f(t,r_{max}) = 0, & (1.50) \end{cases}$$

where v is defined by

$$v(r,L) = \frac{V_\ell}{4\pi}\left(\alpha\frac{L}{L+\kappa}\frac{\rho^3}{\rho^3+r^3} - \frac{(\beta+\gamma r^2)}{r^2}\frac{V(r)-V_{em}}{V(r)-V_{em}+V_\ell\chi}\right). \tag{1.51}$$

This model is analogous to (1.17) through the relation (1.8).

In the last section of Chapter 5 we use the same tools to perform parameter estimation on the diffusive Lifshitz-Slyozov model as well as its version with a constant diffusion given by :

$$\begin{cases} \partial_t g + \partial_x(vg) = D\partial_x^2(g), & (1.52a)\\ L(t) + \displaystyle\int_{\mathbb{R}_+} xg(t,x)\mathrm{d}x = \lambda, & (1.52b)\\ \Big(-vg + D\frac{\varepsilon}{2}\partial_x(g)\Big)\Big|_{x=0} = 0, & (1.52c)\\ g(0,x) = g^0(x) \text{ and } L(0) = L^0. & (1.52d) \end{cases}$$

The stationary solutions are then given by :

$$M_L(x) = \frac{C}{D}\exp\left(\frac{1}{D}\int_0^x v(y,L)\mathrm{d}y\right).$$

1.3.11 Recap of all the models

A visual representation of all the models is provided in Figure 1.13. We started from the classical Lifshitz-Slyozov model to build a Lifshitz-Slyozov model for adipose cells. From the relation between the classical Becker-Döring model and the classical Lifshitz-Slyozov model we introduced a Becker-Döring model for adipose cells which retains the convergence property of the classical models. From this convergence result we developed a diffusive Lifshitz-Slyozov model that more accurately fits to data. We also investigate the relevance of a similar diffusive model with constant diffusion. From the Becker-Döring model and diffusive Lifshitz-Slyozov model for adipose cells we constructed two stochastic models to investigate the convergence of order two.

1.4 Some mathematical tools

In this part, we present most of the mathematical tools we use in the proofs of this manuscript. We begin by introducing the various assumptions we use throughout this manuscript before defining solutions to our models. We then introduce classical theorems we use and the proof techniques. We then proceed with results on existence and uniqueness of those models, providing proofs in Chapter 3. We finish with the description of numerical schemes and tools for parameter estimation.

1.4.1 Assumptions

We list the assumptions we make on the rates a and b, as well as on other objects of our models.

$$a, b \in C^1(\mathbb{R}_+, \mathbb{R}_+), \tag{H1}$$

$$a(0) > 0 \text{ and } \sup_{x\in\mathbb{R}_+} |a(x)| = C_a, \tag{H2a}$$

$$|b(x)| \leq C_b x \text{ for all } x \in \mathbb{R}_+ \text{ and } \lim_{R\to\infty} \sup_{x\geq R} \frac{b(x)}{x} = 0, \tag{H2b}$$

$$\sup_{x\in\mathbb{R}_+} |a'(x)| = K_a \text{ and } \sup_{x\in\mathbb{R}_+} |b'(x)| = K_b, \tag{H3}$$

with $C_a, C_b, K_a, K_b > 0$.

These first four assumptions are the main ones for the following. Note that the choice of the lipogenesis and lipolysis rates (1.11) and (1.12), we chose when constructing our model verifies those assumptions. Concerning the rates in the Becker-Döring models, the main assumptions we make are :

$$a_i^\varepsilon \leq C_a \text{ and } b_i^\varepsilon \leq C_b i\varepsilon, \tag{H'1}$$

$$|a_i^\varepsilon - a_{i+1}^\varepsilon| \leq K_a\varepsilon \text{ and } |b_i^\varepsilon - b_{i+1}^\varepsilon| \leq K_b\varepsilon. \tag{H'2}$$

However the following assumption together with (H1)-(H3) is sufficient for (H'1) and (H'2) to hold :

$$a_i^\varepsilon = a(i\varepsilon) \text{ and } b_i^\varepsilon = b(i\varepsilon), \text{ for all } i \geq 0 \text{ and } \varepsilon > 0. \tag{H4}$$

We proceed with additional technical assumptions, which are useful in a variety of results :

$$\sup_{x\in\mathbb{R}_+} |a''(x)| < +\infty \text{ and } \sup_{x\in\mathbb{R}_+} |b''(x)| < +\infty, \tag{H5}$$

$$\sum_{i\geq 0} |c_{i+1}^{\varepsilon,0} - c_i^{\varepsilon,0}| < +\infty, \tag{H6}$$

$$\varepsilon \sum_{i\geq 0} i|c_{i+1}^{\varepsilon,0} - c_i^{\varepsilon,0}| < +\infty. \tag{H7}$$

$$\text{There exists some constant } \bar{L} > 0, \text{ such that } \inf_{\varepsilon>0} L^{\varepsilon,0} \geq \bar{L}. \tag{H8}$$

$$\text{There exists some constant } K > 0 \text{ such that } \sup_{\varepsilon>0} c_0^{\varepsilon,0} < K. \tag{H9}$$

Comments on these assumptions will be made when useful.

1.4.2 Definition of solutions

In this section we present the definitions of solutions to the deterministic models (1.17) and (1.24).

Lifshitz-Slyozov equations

The Lifshitz-Slyozov equations (1.17) are the coupling of a conservation equation and a transport equation. We give two definitions of solutions to this equation. The first is a definition of weak solutions and the second is the definition of mild solutions using characteristic curves.

Definition 1.4.1. *Given an initial condition $(f^0, L^0) \in C^0(\mathbb{R}_+) \cap L^1(\mathbb{R}_+, (1+x)\mathrm{d}x) \times \mathbb{R}_+$, a measured-valued solution to system (1.17) is composed of two functions $f \in C(0,T; \mathcal{M}^1(0,\infty) - weak-*)$ and $L \in C(0,T)$ such that for all $0 < t < T$ and for all $\varphi \in C^1([0,T]\times\mathbb{R}_+)$ the following relations hold :*

$$\int_0^T \int_{\mathbb{R}_+} (\partial_t\varphi(t,x) + v(x,L(t))\partial_x\varphi(t,x)) f(t,\mathrm{d}x) + \int_{\mathbb{R}_+} \varphi(0,x) f^0(x)\mathrm{d}x = 0,$$

$$L(t) + \int_{\mathbb{R}_+} x f(t,\mathrm{d}x) = \lambda.$$

This definition is quite classic and is used in convergence results which we present in appendix B. As for the definition of mild solutions we first proceed with the definition of characteristic curves of (1.17a).

Definition 1.4.2. *Assume $L \in C^0(\mathbb{R}_+)$ to be given. The characteristic curves associated to* (1.17a) *are solutions to :*

$$\begin{cases} \partial_s X(s;t,x) = v(X(s;t,x), L(s)), \\ X(t;t,x) = x. \end{cases} \tag{1.53}$$

Owing to assumption (H1) and the fact that $y \to \frac{y}{y+\kappa}$ is Lipchitz, the velocity v is C^1 in both x and L. Hence the characteristics are uniquely defined and form an ordered family, which leads us to consider $X_c(t) = X(t;0,0)$ the characteristic curve that is equal to 0 at time 0. This curve is key in looking at mild solutions since other characteristic curves cannot cross it and therefore stay 'to the right' of it. Then, a mild solution to system (1.17) is given by the following definition :

Definition 1.4.3. *Given a smooth initial condition f^0 and $L \in C^0(\mathbb{R}_+)$, a mild solution of* (1.17) *is a couple (f, L), given by :*

$$f(t,x) = f^0(X(0;t,x)) \exp\left(- \int_0^t \partial_x v(X(s;t,x), L(s)) \mathrm{d}s \right) \mathbb{1}_{(X_c(t),\infty)}(x).$$

and $L : \mathbb{R}_+ \to \mathbb{R}_+$ solves $L(t) + \int_{\mathbb{R}_+} x f(t,x) \mathrm{d}x = \lambda$ for all $t \geq 0$.

Remark. *Since we impose null-flux boundary conditions on this system : $v(x, L(t)) f(t,x)|_{x=0} = 0$, there is no term involving 'incoming characteristics' $\mathbb{1}_{(0,X_c(t))}(x)$.*

Remark. *One may relate the measured-valued solution to the mild solution with the following relation : let $\phi \in C_0^0(\mathbb{R}_+)$ the set of continuous functions vanishing at infinity, then*

$$< f(t, \mathrm{d}y), \phi(y) > = < f^0(\mathrm{d}y), \phi(X(t;0,y)) \mathbb{1}_{[X_c(t),+\infty)]}(y) > .$$

Becker-Döring equations

First we define the state space of (1.24) and the meaning of a solution to this system.

Definition 1.4.4. *Let $X := \{x = (x_i)_{i \geq 0} \in \mathbb{R}^{\mathbb{N}} : \sum_{i=0}^{+\infty} i x_i < \infty\}$, endorsed with the norm $\|x\|_X = \sum_{i=0}^{+\infty} i|x_i|$. Denote $x \geq 0$ if $x_i \geq 0$ for all $i \geq 0$, and $X^+ := \{x \in X : x \geq 0\}$.*

We continue with the definition of solutions to the Becker-Döring equations :

Definition 1.4.5. *Let $T > 0$. A solution (c, l) of* (1.24) *in $[0,T)$ is a couple of a function $l : [0,T) \to \mathbb{R}$ and a sequence of functions $c = (c_i)_{i \geq 0}$, $c : [0,T) \to X$ such that :*

(a) For all $t \in [0,T)$, $l(t) \geq 0$ and $c(t) \geq 0$,

(b) For all $i \geq 1$, $c_i : [0,T) \to \mathbb{R}$ is continuous and $\sup_{t \in [0,T)} \|c(t)\|_X < +\infty$,

(c) $l : [0,T) \to \mathbb{R}$ is continuous and $\sup_{t \in [0,T)} |l(t)| < +\infty$,

(d) For all $t \in [0,T)$, $\int_0^t \sum_{i=0}^{+\infty} a_i c_i(s) \mathrm{d}s < \infty$ and $\int_0^t \sum_{i=0}^{+\infty} b_i c_i(s) \mathrm{d}s < \infty$,

(e) For all $t \in [0,T)$, for all $i \geq 1$:

$$c_i(t) = c_i^0 + \int_0^t [J_{i-1}(c(s), l(s)) - J_i(c(s), l(s))] \mathrm{d}s,$$

$$c_0(t) = c_0^0 - \int_0^t J_0(c(s), l(s)) \mathrm{d}s,$$

$$l(t) = l_0 - \int_0^t \sum_{i=0}^{+\infty} J_i(c(s), l(s)) \mathrm{d}s.$$

Remark. *For a fixed $\varepsilon > 0$, solutions to the Becker-Döring model* (1.26) *also verify Definition 1.4.5, up to some rescaling by ε.*

1.4.3 Existence and uniqueness of solutions

As previously mentioned, the existence and uniqueness of solutions to the Lifshitz-Slyozov model have been shown in various cases. We recall here the main theorems as well as their sources. We begin with solutions to the Lifshitz-Slyozov model :

Theorem 1.4.1. *Let $T > 0$. Given an initial condition $(f^0, L^0) \in C^0(\mathbb{R}_+) \cap L^1(\mathbb{R}_+, (1+x)\mathrm{d}x) \times \mathbb{R}_+$, (1.7) has a unique mild solution on the interval $[0,T]$ in the sense of definition 1.4.3.*

The idea of this proof is mainly the study of the characteristic curves defined in 1.4.2. In the case of C^1 rates, see [33], a simple Cauchy-Lipschitz argument guarantees global existence and uniqueness of the characteristic curves for a given function L. Then, for each given function L, a mild solution f is given by Equation (1.4.3). A solution in the sense of Definition 1.4.3 is obtained by using Schauder fixed-point theorem on $L \in C^0(\mathbb{R}_+) : L(t) \leq \lambda,\ \forall t \in \mathbb{R}_+$. Uniqueness is obtained by looking at two different solutions L_1 and L_2 and bounding $|L_1(t) - L_2(t)|$. We point out that the use of a fixed point argument in proofs regarding the Lifshitz-Slyozov model are quite standard. The result may be extended to rates that are non-Lipchitz at the boundary as done in [19].

In regards to the Becker-Döring model, we also have the same result of existence and uniqueness of solutions :

Theorem 1.4.2. *Let $l^0 \in \mathbb{R}_+$ and $c^0 \in X^+$ such that $l^0 + \sum_{i=0}^{+\infty} i c_i^0 = \frac{\lambda}{\Lambda} < \infty$. Assume that* (H'1), (H'2) *hold true. Then there exists a unique solution (c, l) of* (1.22) *in the sense of Definition 1.4.5, with $c(0) = c^0$ et $l(0) = l^0$.*

The proof of this theorem can be found in the seminal paper by Ball, Carr and Penrose [9]. It goes as follows. We first consider a truncated model where we only consider the first n sizes. A solution of this finite dimensional system is obtained trough Cauchy-Lipchitz theorem and shown to have same properties as in Definition 1.4.5. The limit $n \to \infty$ is obtained by extracting subsequences for the c_i's and a Cauchy sequence argument for L. Uniqueness is obtained by considering two solutions c and c' and computing moments of the form $\sum_i = 1^\infty i^\gamma |c_i - c_i'|$ which are equal to zero and where the exponent γ is chosen depending on assumptions on a and b

1.4.4 Classical convergence

Having defined solutions to both models, we investigate the convergence from the Becker-Döring equations to the Lifshitz-Slyozov equations. We prove a similar result as in [129] and use a similar sketch of proof. The following result is proved in appendix B.

We begin by recalling the theorem on a 'weak' formulation fo the Becker-Döring model from [9] :

Theorem 1.4.3. *(Ball et al. [9], theorem 2.5) Let $(g_i)_{i\geq 0}$ be a given sequence. Let $(c^\varepsilon, L^\varepsilon)$ be the solution of* (1.26) *on $[0,T)$, $0 < T \leq +\infty$.*

Assume that for all $0 \leq t_1 < t_2 < T$, $\int_{t_1}^{t_2} \sum_{i=0}^{\infty} |g_{i+1} - g_i| a_i^\varepsilon c_i^\varepsilon(t)\mathrm{d}t < \infty$ and that either of the following holds :

(a) $g_i = \mathcal{O}(i)$ and $\int_{t_1}^{t_2} \sum_{i=0}^{\infty} |g_{i+1} - g_i| b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t < \infty$

(b) $\sum_{i=0}^{\infty} g_i c_i^\varepsilon(t_k) < \infty$, for $k = 1, 2$ and $g_{i+1} \geq g_i \geq 0$ for i large enough

Then :

$$\begin{aligned} \sum_{i=0}^{\infty} g_i c_i^\varepsilon(t_2) - \sum_{i=0}^{\infty} g_i c_i^\varepsilon(t_1) + \int_{t_1}^{t_2} \sum_{i=0}^{\infty} \frac{g_{i+1} - g_i}{\varepsilon} b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t \\ = \int_{t_1}^{t_2} \sum_{i=0}^{\infty} \frac{g_{i+1} - g_i}{\varepsilon} a_i^\varepsilon \frac{L^\varepsilon(t)}{L^\varepsilon(t) + \kappa} c_i^\varepsilon(t)\mathrm{d}t\,. \end{aligned} \tag{1.54}$$

Now let us define the following step functions depending on both time and space, where f^ε is a candidate for convergence to a solution of the Lifshitz-Slyozov system. Let $\Gamma_i^\varepsilon = [(i - \frac{1}{2})\varepsilon, (i + \frac{1}{2})\varepsilon[$ and we write :

$$f^\varepsilon(t,x) = \sum_{i\geq 0} \mathbf{1}_{\Gamma_i^\varepsilon}(x) c_i^\varepsilon(t).$$

Similarly we build the two space functions a^ε and b^ε as :

$$a^\varepsilon(x) = \sum_{i\geq 0} \mathbf{1}_{\Gamma_i^\varepsilon}(x) a_i^\varepsilon,$$

$$b^\varepsilon(x) = \sum_{i\geq 1} \mathbf{1}_{\Gamma_i^\varepsilon}(x) b_i^\varepsilon.$$

From theorem 1.4.3, we derive what could be considered to be a 'weak' equation for f^ε :

Proposition 1.4.1. *(Laurençot and Mischler [78],Lemma 4.1) Let $\phi \in L^\infty(\mathbb{R}_+)$. Then for every $t \geq 0$, we have the following equality :*

$$\int_0^\infty \phi(x)(f^\varepsilon(t,x) - f^\varepsilon(0,x))\mathrm{d}x = \int_0^t \int_0^\infty (\Delta_\varepsilon \phi(x) a^\varepsilon(x) \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \Delta_{-\varepsilon}\phi(x) b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t, \tag{1.55}$$

where

$$\Delta_\varepsilon \phi(x) = \frac{\phi(x+\varepsilon) - \phi(x)}{\varepsilon}.$$

Finally we are led to the classical convergence theorem. One may find different versions of this theorem in [129, 78].

Theorem 1.4.4 (Convergence towards a solution of the Lifshitz-Slyozov system)**.** *Assume the following statements :*

- *there exists $K > 0$ such that $a^\varepsilon(x) < K$ and $b^\varepsilon(x) < K(1+x)$, for all $x \in \mathbb{R}_+$;*
- *$\varepsilon \sum_{i\geq 0} c_i^{\varepsilon,0} < C_0$ where $C_0 > 0$ is a constant independent of ε,*
- *$\varepsilon \sum_{i\geq 0} (i\varepsilon)^{1+s} c_i^{\varepsilon,0} < C_s$, $\forall s < 1$.*

Additionally, suppose there exist $a \in L^\infty(0,+\infty)$, $b \in L^1(0,+\infty)$ such that $\{a^\varepsilon(.)\}$, respectively $\{b^\varepsilon(.)\}$, converges uniformly on every compact of $[0,+\infty)$ to $a(.)$, respectively to $b(.)$.

Then there exists a sequence ε_n and a solution (f,L) of (1.7) in the sense of definition 1.4.1 such that :

$$\begin{cases} f^{\varepsilon_n} \to f \ , \ \ x f^{\varepsilon_n} \to x f \ \text{in } C^0([0,+\infty[;\mathcal{M}^1(0,+\infty) - weak - *) \\ L^{\varepsilon_n} \to L \ \text{uniformly in } C^0([0,T]) \end{cases}$$

Remark. *This result is also an existence result since it shows that the solution to the Becker-Döring model $(f^\varepsilon, L^\varepsilon)$ converges to a couple of functions (f,L) that solves the Lifshitz-Slyozov model.*

The proof of this theorem relies mainly on Arzela-Ascoli theorem and density of functional spaces.

1.4.5 Construction of the diffusive Lifshitz-Slyozov model

Correlation between a birth and death process such as the Becker-Döring model and Fokker-Planck dynamic was first described in [49]. Heuristic relations to the Lifshitz-Slyozov model were derived later on in [57], [129] and later in [120]. The original idea from [49], recalled in [57], is to use a Taylor expansion of order two for the flux of the Becker-Döring model, which leads to the name second-order Lifshitz-Slyozov model. Mainly it consists in writing :

$$\frac{J_{i+1}^\varepsilon(c^\varepsilon, L^\varepsilon) - J_i^\varepsilon(c^\varepsilon, L^\varepsilon)}{\varepsilon} \simeq \partial_i J_i^\varepsilon(c^\varepsilon, L^\varepsilon) + \frac{\varepsilon}{2}\partial_i^2((a_i^\varepsilon \frac{L^\varepsilon}{L^\varepsilon+\kappa} + b_i^\varepsilon)c_i^\varepsilon) + o(\varepsilon^2),\ i \geq 0.$$

This expression has no real mathematical meaning and is more of a heuristic intuition as to where the diffusion comes from.

Then let $f^\varepsilon(t,x) = \sum_{i\geq 0} \mathbb{1}_{\Gamma_i^\varepsilon}(x) c_i^\varepsilon(t)$, where $\Gamma_i^\varepsilon = [(i-1/2)\varepsilon, (i+1/2)\varepsilon)$. It is fair to consider that f^ε is close to a solution of the non-linear Fokker-Planck equation :

$$\partial_t g + \partial_x(vg) - \frac{\varepsilon}{2}\partial_x^2(dg) = 0. \tag{1.56}$$

A second approach to relate the Becker-Döring model to the diffusive Lifshitz-Slyozov model is to observe that the Becker-Döring model is close to a numerical scheme and to add and remove terms to obtain a 'numerical' diffusion. In particular looking at proposition 1.4.1, and denoting $\bar{\Delta}_\varepsilon\phi(x) = \frac{\phi(x+\varepsilon)-\phi(x-\varepsilon)}{2\varepsilon}$ and $\Delta_\varepsilon^2\phi(x) = \frac{\phi(x+\varepsilon)-2\phi(x)+\phi(x-\varepsilon)}{\varepsilon^2}$, we can rewrite (1.55) as :

$$\begin{aligned}\int_0^\infty \phi(x)(f^\varepsilon(t,x) - f^\varepsilon(0,x))\mathrm{d}x = &\int_0^t \int_0^\infty (a^\varepsilon(x)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - b^\varepsilon(x)) f^\varepsilon(t,x)\bar{\Delta}_\varepsilon\phi(x)\mathrm{d}x\mathrm{d}t \\ &+ \frac{\varepsilon}{2}\int_0^t \int_0^\infty (a^\varepsilon(x)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - b^\varepsilon(x)) f^\varepsilon(t,x)\Delta_\varepsilon^2\phi(x)\mathrm{d}x\mathrm{d}t. \end{aligned}\tag{1.57}$$

Now formally taking the limit up to second-order term in ε leads again to the Fokker-Planck equation (1.56).

1.4.6 Strong approximation theorem for density dependent Markov chains

In this section, we detail the method we use for showing Theorems 1.5.2. These theorems bounds the L^1-norm of the difference between X_ε and Z_ε by a decreasing function of ε. The overall shape of the proof is taken from [74]. Let us consider two stochastic processes $X'_\varepsilon \in \varepsilon\mathbb{N}^{\mathbb{N}}$, $Z'_\varepsilon \in \mathbb{R}^{\mathbb{N}}$ and a family of jump directions $J \subset \mathbb{N}^{\mathbb{N}}$. Then $X'_\varepsilon(t)$ and $Z'_\varepsilon(t)$ are solutions to :

$$X'_\varepsilon(t) = X(0) + \sum_{l\in J} \varepsilon l Y_l(\varepsilon^{-1}\int_0^t f_l(X'_\varepsilon(s))\mathrm{d}s), \tag{1.58}$$

$$Z'_\varepsilon(t) = Z(0) + \sum_{l\in J} \varepsilon l B_l(\varepsilon^{-1}\int_0^t f_l(Z'_\varepsilon(s))\mathrm{d}s), \tag{1.59}$$

where f_l is the rate of the jump direction l. The Y_l are independent Poisson processes and B_l are independent drifted Wiener processes. We recall the main theorem from [74] :

Theorem 1.4.5 (Kurtz [74], Theorem 3.3)**.** *Let $T > 0$. Assume that the jump rates are bounded Lipchitz functions with $|f_l(x)| \leq C_l$ for all $x \in \mathbb{R}_+$ and all possible jump directions l and $\sum_l l f_l(x)$ is a Lipchitz function. Then if $C_l = 0$ for all but finitely many l there exists a random variable β^T independent of ε and with exponential moments such that :*

$$\sup_{t\leq T}|X'_\varepsilon(t) - Z'_\varepsilon(t)| \leq \beta^T \varepsilon \ln(\varepsilon^{-1}), \ \ a.e. \tag{1.60}$$

Remark. *In the case of adipose modeling the only two possible directions are $\{-1, 1\}$. Additionally, X'_ε and Z'_ε are density dependent while X_ε and Z_ε are also non-linear. Thus obtaining an almost sure convergence proves quite difficult and we are only able to obtain an L^1 convergence result.*

1.4.7 Numerical schemes for PDEs

This section details the numerical schemes we use. We begin by describing the Finite Volume Method (FVM) which we use for simulating the PDEs of both the Lifshitz-Slyozov and diffusive Lifshitz-Slyozov models.

Finite Volume Method

Our main goal is to simulate (1.32) and its stationary solutions. Since Equation (1.32a) is a conservative PDE, a natural direction for numerical scheme is the use of the Finite Volume Method. The theoretical results we introduce consider $x \in \mathbb{R}_+$ for all models, but as is often the case when performing numerical simulations, it is easier to work on a bounded domain with an appropriate

boundary condition at $x_{\max}$. Hence for all simulations of the Lifshitz-Slyozov models we will restrict ourselves to a bounded domain denoted $[0, x_{\max}]$, and we provide an adequate boundary condition at $x_{\max}$ that preserves the total number of cells m. For the sake of completeness, we recall the Finite Volume Method in the case of a 1D non-linear transport equation. Let us consider some distribution $f(t,x)$ on a space domain $[0, x_{\max}]$, subjected to a flow $F(f)$.

This writes as

$$\partial_t f + \partial_x F(f) = 0. \tag{1.61}$$

Now we subdivide the space domain into cells centered at points labelled x_i, for $i = 0, \dots, N$, which we denote $C_i = [x_{i-1/2}, x_{i+1/2}]$, where $x_{i+1/2}$ is the edge between cell C_i and C_{i+1}. For simplicity let us assume that all cells are of similar size $\Delta x = x_{i+1/2} - x_{i-1/2}$. The idea of the FVM is to approximate the volume average of f on each of these cells. Hence let us denote the volume average over the cell i as :

$$f_i(t) = \frac{1}{\Delta x} \int_{C_i} f(t,x) \mathrm{d}x.$$

Then integrating (1.61) over cell i, we are led to :

$$\frac{\mathrm{d}f_i(t)}{\mathrm{d}t} + \frac{1}{\Delta x} \int_{C_i} \partial_x F(f) \mathrm{d}x = 0.$$

Then let us denote $F_{i\pm1/2}$ the evaluation of the flow at the edge $x_{i\pm1/2}$ of cell i and use the divergence theorem to obtain the standard equation of the FVM :

$$\frac{\mathrm{d}f_i(t)}{\mathrm{d}t} + \frac{1}{\Delta x}(F_{i+1/2} - F_{i-1/2}) = 0.$$

Up to this last equation, computations are exact. From a numerical standpoint, the approximation will come from the choice of reconstructing the terms $F_{i\pm1/2}$. Many methods are available, including interpolation, or sometimes exact computations. However the choice is obviously made case by case. The FVM method is particularly useful in the case of PDE conserving the total mass of the distributions, i.e. $\frac{\mathrm{d}}{\mathrm{d}t}\int_E f(t,x)\mathrm{d}x = 0$. Indeed this conservation equation can simply be transformed to

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_i f_i(t) = -\sum_i \frac{F_{i+1/2} - F_{i-1/2}}{\Delta x} = \frac{F_{-1/2} - F_{N+1/2}}{\Delta x}.$$

Hence conserving the total mass of the distribution amounts to have the boundary conditions $F_{-1/2} = F_{N+1/2} = 0$.

We now have a semi-discrete numerical method for Equation (1.61) and the time derivative may be approximated using standard methods.

The time derivative is approximated using an Euler explicit scheme. Given a series of time points $t^0 < t^1 < \cdots < t^n < \cdots$, we denote f_i^n the approximation of f on cell i at time t^n. Then for all $i = 0, \dots, N$, the scheme reads :

$$\frac{f_i^{n+1} - f_i^n}{\Delta t^n} + \frac{1}{\Delta x}(F_{i+1/2}^n - F_{i-1/2}^n) = 0, \tag{1.62}$$

where $\Delta t^n = t^{n+1} - t^n$ and the superscript n on the flux simply denotes that the FVM is taken at time t^n and the scheme is therefore explicit.

Well-balanced schemes

One of our main goals since the description of our model has been to look at stationary solutions. Hence any numerical method we devise has to be able to compute stationary solutions effectively. Looking at numerical schemes, this property is related to the scheme being **well-balanced** [53, 50]. In words, this means that given an approximated stationary solution, the scheme should no deviate from this approximation. The scheme (1.62) can be written formally as an application $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ where $\Phi(f^n) = f^{n+1}$. Now let us consider $f^{\mathrm{stat}}(x)$ a stationary solution to (1.61), i.e. $\partial_x F(f^{\mathrm{stat}}) = 0$. Then for the scheme to be well-balanced means that given an approximation $(f_i^{\mathrm{stat}})_{i=0,\dots,N}$ to f^{stat}, it holds that :

$$\Phi((f_i^{\text{stat}})_{i=0,\dots,N}) = (f_i^{\text{stat}})_{i=0,\dots,N}.$$

Numerical scheme for Equation (1.17a)

As an example to the Finite Volume method, let us detail the scheme we use to obtain numerical simulations of solutions of Equation (1.17a) with a given $L : \mathbb{R}_+ \to \mathbb{R}_+$. Hence we need to provide a way to approximate the value of this flux at the boundary of each cell. To do so we use the UpWind method, which is quite standard for transport PDEs. The idea is to consider cases depending on the sign of the velocity. If the velocity is positive, then the flux is going in a left to right manner and in some sense 'information' is traveling from left to right. Hence we want to use the value in the cell on the left of the boundary to make our approximation and vice versa for the case of negative speed. Formally, this is done as follows.

Consider the approximation of the flux at the boundary $x_{i+1/2}$, which we split into two terms :

$$F_{i+1/2} = \mathcal{F}^+(v(x_{i+1/2}, L), f_i) + \mathcal{F}^-(v(x_{i+1/2}, L), f_{i+1})$$

where the positive and negative fluxes are defined as follows :

$$\mathcal{F}^+(v, f) = \begin{cases} 0 & \text{if } v < 0, \\ vf & \text{else,} \end{cases}$$

$$\mathcal{F}^-(v, f) = \begin{cases} 0 & \text{if } v > 0, \\ vf & \text{else.} \end{cases}$$

Then using the usual notation for positive and negative parts : $a^+ = \max(a, 0)$ and $a^- = \min(a, 0)$, the scheme reads as :

$$\frac{f_i^{n+1} - f_i^n}{\Delta t^n} + \frac{1}{\Delta x}((v_{i+1/2}^n)^+ f_{i+1}^n + ((v_{i+1/2}^n)^- - (v_{i-1/2}^n)^+) f_i^n - (v_{i-1/2}^n)^- f_{i-1}^n) = 0, \quad \text{for all } i = 1, \dots, N-1. \tag{1.63}$$

The scheme is provided with null-flux boundary conditions to conserve the total mass of the distribution, i.e. the amount of cells :$F_{-1/2} = F_{N+1/2} = 0$. Hence the scheme may be written as :

$$f^{n+1} = (\mathbb{I} + \frac{\Delta t^n}{\Delta x} A) f^n, \tag{1.64}$$

where $A \in \mathcal{M}_{N\times N}(\mathbb{R})$ is a symmetric matrix with non-negative eigenvalues. Naturally this scheme gives rise to a CFL condition of the form

$$\Delta t^n \leq \frac{\Delta x}{\max\limits_{i=0,\dots,N} |v_{i+1/2}^n + v_{i-1/2}^n|}$$

Now to simulate the evolution of L we need to discretize Equation (1.17b). To do so we use the method introduced in [51], which we detail in the next section. This will notably give rise to another condition on the time step Δt^n to preserve positivity of L.

Numerical scheme for Equation (1.32)

We now proceed with describing the numerical scheme we use to simulate Equation (1.32). The scheme is derived from [51] with some slight variations for boundary conditions. It uses the Finite Volume method and a clever change of variable to ease computer simulations. First we need to rewrite Equation (1.32a) in an appropriate form. Ideally we want to obtain $\partial_t g + \partial_x F(g) = 0$ for some flux term F. First let us recall that stationary solutions to (1.32a) are denoted by M_L. Once again we denote $[0, x_{\max}]$ the space domain. Then simple computations show that :

$$vg - \partial_x(dg) = dM_L \partial_x(\frac{g}{M_L}) = F(g).$$

The idea behind the scheme is to first perform a change of variable in the PDE (1.32a) such that the flux operator F becomes symmetric for the usual L^2 scalar product. This change of variable is the following : $h = \frac{g}{\sqrt{M_L}}$. Then h verifies :

$$\partial_t h + \frac{1}{\sqrt{M_L}}\partial_x(dM_L\partial_x(\frac{h}{\sqrt{M_L}})) = \partial_t h + \frac{1}{\sqrt{M_L}}\partial_x \mathbf{F}(h).$$

Observe that the correspondence between the two operators is simply $\mathbf{F}(h) = F(h\sqrt{M_L}) = F(g)$. Moreover, the boundary condition in the model (1.32) may simply be rewritten as $F(g)|_{x=0} = \mathbf{F}(h)|_{x_{\max}=0} = 0$.

Remark. *One may observe that another valid change of variable is $h = \frac{g}{M_L}$. However in this case the operator $\mathbf{F}$ is not symmetric.*

We also need to provide a boundary condition at $x_{\max}$. Since we want to conserve the total amount of cells m, and we have a null-flux boundary condition at zero, we impose a null-flux boundary condition at $x_{\max}$. This yields the bounded domain diffusive Lifshitz-Slyozov model :

$$\begin{cases} \partial_t g + \partial_x F(g) = 0, & \text{(1.65a)} \\ L(t) + \displaystyle\int_0^{x_{\max}} x g(t,x)\mathrm{d}x = \lambda, & \text{(1.65b)} \\ F(g)_{x=0,x_{\max}} = 0, & \text{(1.65c)} \\ g(0,x) = g^0(x) \text{ and } L(0) = L^0. & \text{(1.65d)} \end{cases}$$

We may now proceed with the discretization of the operator $\mathbf{F}$. The change of variable at the discrete level is written as $h_i^n = \frac{g_i^n}{\sqrt{M_i^n}}$ for all $i = 1, \dots, N$. The approximation of L at time t^n is denoted L^n. The time derivative is approximated with an Euler explicit scheme and the operator $\mathbf{F}$ using the Finite Volume method. To prevent the need for a CFL condition, we take an implicit method in h where the approximated flux $\mathbf{F}_{i+1/2}$ is taken at the next time step t^{n+1} and is denoted by $\mathbf{F}_{i+1/2}^{n+1}$. A splitting is performed between h^{n+1} and L^{n+1} : we first compute h^{n+1} using the values of h^n and L^n and then we find L^{n+1} using values of h^{n+1} and h^n. This writes as :

$$\frac{h_i^{n+1} - h_i^n}{\Delta t^n} + \frac{1}{\sqrt{M_i^n}}\frac{\mathbf{F}_{i+1/2}^{n+1} - \mathbf{F}_{i-1/2}^{n+1}}{\Delta x} = 0, \ \forall i = 1, \dots, N. \tag{1.66}$$

We are left with how to construct the approximated flux $\mathbf{F}_{i+1/2}^{n+1}$. We recall that $\mathbf{F}(h) = dM_L\partial_x(\frac{h}{\sqrt{M_L}})$ and that we aim to approximate its value at the edge of a cell. Therefore the partial derivative $\partial_x(\frac{h}{\sqrt{M_L}})$ may simply be approximated by finite difference between the two cells that share the edge in consideration. This yields :

$$\mathbf{F}_{i+1/2}^{n+1} \simeq d_{i+1/2}^n M_{i+1/2}^n \frac{\frac{h_{i+1}^{n+1}}{\sqrt{M_{i+1}^n}} - \frac{h_i^{n+1}}{\sqrt{M_i^n}}}{\Delta x}.$$

Since d is a known function, we can simply evaluate d at the edge $x_{i+1/2}$ to get $d_{i+1/2}^n$. However with the intent that the stationary solution is discretized in the same way as g, we do not have explicit values for M_L at the edges. Hence we need an approximation for $M_{i+1/2}^n$. To do so we use the geometrical mean between the values of the two neighboring cells : $M_{i+1/2}^n \simeq \sqrt{M_{i+1}^n M_i^n}$. This finally gives us our approximation of the flux $\mathbf{F}$ at the cell's edge :

$$\mathbf{F}_{i+1/2}^{n+1} = d_{i+1/2}^n \sqrt{M_{i+1}^n M_i^n} \frac{\frac{h_{i+1}^{n+1}}{\sqrt{M_{i+1}^n}} - \frac{h_i^{n+1}}{\sqrt{M_i^n}}}{\Delta x} \tag{1.67}$$

Hence the scheme can be written as :

$$\frac{h_i^{n+1} - h_i^n}{\Delta t^n} + \frac{1}{\Delta x^2}\left(d_{i+1/2}h_{i+1}^{n+1} - \left(d_{i+1/2}\sqrt{\frac{M_{i+1}^n}{M_i^n}} + d_{i-1/2}\sqrt{\frac{M_{i-1}^n}{M_i^n}}\right)h_i^{n+1} + d_{i-1/2}h_{i-1}^{n+1}\right) = 0,$$
$$\forall i = 1, \dots, N-1. \tag{1.68}$$

Now since the boundary conditions we consider take the form $F(g)|_{x=0,x_{\max}} = 0$, and it implies conservation of the mass of the distribution, we are led on the discrete level to set $F_{-1/2} = F_{N+1/2} = 0$, which is equivalent to $\mathbf{F}_{-1/2} = \mathbf{F}_{N+1/2} = 0$. Therefore the scheme is completed with the boundary terms :

$$\frac{h_0^{n+1} - h_0^n}{\Delta t^n} + \frac{1}{\Delta x^2}\left(d_{1/2}h_1^{n+1} - d_{1/2}\sqrt{\frac{M_1^n}{M_0^n}}h_0^{n+1}\right) = 0,$$

$$\frac{h_N^{n+1} - h_N^n}{\Delta t^n} + \frac{1}{\Delta x^2}\left(-d_{N-1/2}\sqrt{\frac{M_{N-1}^n}{M_N^n}}h_N^{n+1} + d_{N-1/2}h_{N-1}^{n+1}\right) = 0.$$

Hence the scheme may be written as :

$$(\mathbb{I} + \frac{\Delta t^n}{\Delta x^2}A)h^{n+1} = h^n, \tag{1.69}$$

where $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is a matrix defined as :

$$\begin{pmatrix} D_0 & d_{1/2} & 0 & \dots & 0 \\ d_{1/2} & D_1 & d_{3/2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \dots & d_{N-3/2} & D_{N-1} & d_{N-1/2} \\ 0 & \dots & 0 & d_{N-1/2} & D_N \end{pmatrix}$$

where $D_0 = d_{1/2}\sqrt{\frac{M_1^n}{M_0^n}}$, $D_N = -d_{N-1/2}\sqrt{\frac{M_{N-1}^n}{M_N^n}}$ and $D_i = d_{i+1/2}\sqrt{\frac{M_{i+1}^n}{M_i^n}} + d_{i-1/2}\sqrt{\frac{M_{i-1}^n}{M_i^n}}$, for all $i = 1, \dots, N-1$. Observe that A is indeed symmetric. In particular writing M the diagonal matrix which coefficients are the M_i^n's we can write the scheme for g as :

$$(\mathbb{I} + \frac{\Delta t^n}{\Delta x^2}M^{1/2}AM^{-1/2})g^{n+1} = g^n.$$

We point out that the terms taking the form $\frac{M_{i+1}^n}{M_i^n}$ are computed explicitly, using the definition of M_L, which might prove useful when M_L takes extreme values. We compute therefore :

$$\frac{M_{L^n}(x_{i+1})}{M_{L^n}(x_i)} = \frac{d(x_i, L^n)}{d(x_{i+1}, L^n)}\exp\left(\frac{2}{\varepsilon}\int_{x_i}^{x_{i+1}}\frac{v(y, L^n)}{d(y, L^n)}\mathrm{d}y\right),$$

and we set $\frac{M_{i+1}^n}{M_i^n} = \frac{M_{L^n}(x_{i+1})}{M_{L^n}(x_i)}$.

Once we have obtained h^{n+1} we reconstruct g by simply writing $g_i^{n+1} = h_i^{n+1}\sqrt{M_i^n}$. All we are left to do is to describe the dynamic of L at the discrete level. To do so we go back to Equation (1.7b), and its discrete counterpart reads $\lambda = L^n + \Delta x\sum_{i=0}^{N} x_i g_i^n$. Hence to keep the constraint at the next time step we need :

$$\begin{aligned} L^{n+1} - L^n &= -\Delta x\sum_{i=0}^{N} x_i(g_i^{n+1} - g_i^n) \\ &= -\Delta t^n\sum_{i=1}^{N-1} x_i(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1}) - \Delta t^n x_N F_{N-1/2}^{n+1} + \Delta t^n x_0 F_{1/2}^{n+1} \\ &= \Delta t^n\Delta x\sum_{i=0}^{N-1} F_{i+1/2}^{n+1} \\ &= \Delta t^n\Delta x\sum_{i=0}^{N-1} d_{i+1/2}^n M_{i+1/2}^n\left(\frac{g_{i+1}^{n+1}}{M_{i+1}^n} - \frac{g_i^{n+1}}{M_i^n}\right), \end{aligned}$$

which gives the update step for the dynamic of L :

$$L^{n+1} = L^n + \Delta t^n \Delta x \sum_{i=0}^{N-1} d^n_{i+1/2} M^n_{i+1/2} \left(\frac{g^{n+1}_{i+1}}{M^n_{i+1}} - \frac{g^{n+1}_i}{M^n_i} \right). \tag{1.70}$$

Although the scheme we use is implicit and does not yield a CFL condition, Equation (1.70) still imposes some condition on Δt^n to preserve positivity. In particular, let us denote $B^n \in \mathbb{R}^N$ such that :

$$B^n_i = \begin{cases} -d_{1/2} \dfrac{M^n_{1/2}}{M^n_0}, & \text{if } i = 0, \\ \dfrac{1}{M_i}(d_{i+1/2} M_{i+1/2} - d_{i-1/2} M_{i-1/2}), & \text{otherwise.} \end{cases}$$

Then the condition reads as :

$$\Delta t^n |(B^n)^T (\mathbb{I} + \frac{\Delta t^n}{\Delta x^2} M^{1/2} A M^{-1/2}) g^n| \leq L^n.$$

Again following [51] we have that since $M^{1/2} A M^{-1/2}$ is symmetric and its eigenvalues are non-negative, it is sufficient to have that, for some diagonal matrix P :

$$\Delta t^n |P^{-1} B^n| |P g^n| \leq L^n.$$

1.4.8 Stochastic simulations

In this part, we detail the methods used to perform the numerical simulations of both stochastic models. Simulations of the Becker-Döring CTMC are performed using the Gillespie algorithm and simulations of the Lifshitz-Slyozov SDE are done using the Euleur-Muruyama method. Let us also remark that we actually simulate the models with a number of cells N described by Equations (1.39) and (1.43). We point out that Equation (1.39) is rescaled by ε to give :

$$X^{k,N}_\varepsilon(t) = X(0) + \varepsilon Y_+ \Big(\varepsilon^{-1} \int_0^t a(X^{k,N}_\varepsilon(s)) \frac{L^{X,N}(s)}{L^{X,N}(s) + \kappa} \mathrm{d}s \Big) - \varepsilon Y_- \Big(\varepsilon^{-1} \int_0^t b(X^{k,N}_\varepsilon(s)) \mathrm{d}s \Big). \tag{1.71}$$

Gillespie algorithm

Let us describe the Gillespie algorithm. This algorithm was created by Doob [38] and popularized by Gillespie in [47, 48]. The algorithm generates statistically correct trajectories of a system of stochastic equations where the rates are known. It has shown great use for simulations of chemical systems of reactions, and in computational biology to simulate biochemical reactions. Let us consider a set of molecules labelled $R = (R_k)_{k=1,\dots,K}$. The amount of each molecule is denoted by X^k and we suppose that these molecules interact with each other through reactions labelled by $i = 1, \dots, I$. The rate of reaction i is denoted $a_i(X)$, and its net change of molecule abundances is denoted $\nu_i(X)$, where $X = (X^k)_{k=1,\dots,K}$.

Now the algorithm goes as follows. First we give ourselves an initial state $X(0)$ and set $t = 0$. Then we loop over the following steps :

(1) Compute $a_i(X(t))$ for all $i = 1, \dots, I$, and $a_0 = \sum_{i=1}^{I} a_i(X(t))$.

(2) Generate two unit-interval uniform random numbers r_1 and r_2. Set $\tau = \frac{1}{a_0} \ln(\frac{1}{r_1})$ and find j such that $\sum_{i=1}^{j} a_i(X(t)) \leq r_2 a_0 \leq \sum_{i=1}^{j+1} a_i(X(t))$.

(3) Update $t \leftarrow t + \tau$ and $X \leftarrow X + \nu_j(X)$ according to reaction j chosen at step 2.

(4) Go to step 1

Step 2 is the crucial one. We begin by observing that τ follows an exponential distribution with parameter a_0 and it defines the next time a reaction occurs. Having found when a reaction occurs we need to find which reaction occurs. Reaction i happens with probability $\frac{a_i}{a_0}$, hence we split the interval $[0, 1]$ into partitions of size $\frac{a_i}{a_0}$ for $i = 1, \dots, I$. Step 2 checks to which partition r_2 belongs to, and we choose the corresponding reaction to occur.

Next reaction method

Gillespie algorithm is often quite costly and there exist many optimization methods. We present the one we use : the **next reaction method** [45]. First we introduce the dependency graph for the set of reactions $\{1, \dots, I\}$. For example reaction 1 may be written as $R_1 + R_2 \to R_3$. Hence reaction 1 only affects the values of X^1,X^2 and X^3. Also, the value of a_1 only depends on the value of X^1 and X^2. The idea of the next reaction method is to use this fact and update only appropriate time at each step : we only update the reaction on which the occurring reaction has an effect. This leads us to the following definitions :

Definition 1.4.6. *For a given reaction i the set of molecules that affect the value of a_i is denoted $DependsOn(a_i)$ and the set of molecules that i affects is denoted $Affects(i)$.*

Definition 1.4.7. *Let $\mathcal{G}(V, E)$ be a graph with $V = \{1, \dots, I\}$. Let E be such that the edge (i, j) exists iff $Affects(i) \cap DependsOn(a_j) \neq \emptyset$. Then $\mathcal{G}$ is called the dependency graph for the set of reactions $\{1, \dots, I\}$.*

The idea of the next reaction method is to compute an activation time for each reaction. Then at each step the smallest reaction time is picked and the reaction occurs. Only the times of reactions adjacent to the active reaction in the dependency graph are updated.

The method goes as follows :

1 Initialize $t = 0$:

   a Compute $a_i(X(0))$ for each reaction i and set $a_{i,\text{old}} \leftarrow a_i(X(0))$.

   b For each i, generate a putative time $\tau_i \sim \mathcal{E}(a_i(X(0)))$.

   c Store the τ_i in an indexed priority queue P.

2 Let μ be the reaction with the smallest putative time τ_μ and update $t \leftarrow \tau_\mu$.

3 Update X according to reaction μ : $X \leftarrow X + \nu_\mu(X)$.

4 For each edge (μ, j) in the dependency graph $\mathcal{G}$:

   a Update $a_j(X(t))$ and set $a_{j,\text{new}} \leftarrow a_j(X(t))$.

   b If $j \neq \mu$, set $\tau_j \leftarrow \frac{a_{j,\text{old}}}{a_{j,\text{new}}}(\tau_j - t) + t$.

   c If $j = \mu$, generate $\rho \sim \mathcal{E}(a_j(X(t)))$ and set $\tau_\mu \leftarrow \rho + t$.

   d Replace τ_μ in P and $a_{j,\text{old}} \leftarrow a_{j,\text{new}}$.

5 Go to step 2.

Here $a_{j,\text{old}}$ corresponds to the rate before we update it and $a_{j,\text{new}}$ to the rate after the update at step 4.a. The step 4.b corresponds to transforming a Poisson random variable with parameter $a_{j,\text{old}}$ into a Poisson random variable with parameter $a_{j,\text{new}}$, on the condition that the first one is greater than t. This method is more efficient than the classical Gillespie algorithm because at each step, it only computes the rates a_i for the reactions that were modified by the update instead of recomputing all the rates at each step. This proves particularly efficient as the set of reactions gets bigger.

Euleur-Muruyama method for SDE

Simulations of Z_ε are done using the Euler-Muruyama method. It is an extension of the Euler method for ODEs to SDEs. As an example let us consider the following SDE :

$$dZ(t) = \mu(Z(t), t)\mathrm{d}t + \sigma(Z(t), t)dW(t),$$

with some initial condition $Z(0) = Z^0$ and where $W(t)$ is the usual Wiener process. Now we consider that the simulation should run up to some time $T > 0$. The interval $[0, T]$ is split into I sub-intervals of width $\Delta t \frac{T}{I}$, and we denote $t_0 = 0$, $t_{i+1} = t_i + \Delta t$ for $i = 1, \dots, I$. Then the approximation of Z at time t_{i+1} denoted Z_{i+1}, is given recursively by :

$$Z_{i+1} = Z_i + \mu(Z_i, t_i)\Delta t + \sigma(Z_i, t_i)\Delta W_i,$$

where $\Delta W_i = W(t_{i+1}) - W(t_i)$. At each time step, we need to generate ΔW_i, which we do using the fact that W has Gaussian increments. This gives us $\Delta W_i \sim \mathcal{N}(0, \Delta t)$. Hence, given an i.i.d. sequence of random variables $(Y_i)_{i=1,\dots,I}$ such that $Y_1 \sim \mathcal{N}(0,1)$, the method may be written as :

$$Z_{i+1} = Z_i + \mu(Z_i, t_i)\Delta t + \sigma(Z_i, t_i)\sqrt{\Delta t}Y_{i+1}.$$

Since we will make simulations on the half-line $\mathbb{R}_+$, we need to provide boundary conditions at $x = 0$. We use a reflective boundary condition which translates to taking the absolute value on the right hand side of the previous equation :

$$Z_{i+1} = |Z_i + \mu(Z_i, t_i)\Delta t + \sigma(Z_i, t_i)\sqrt{\Delta t}Y_{i+1}|.$$

1.4.9 Parameter estimation

In Chapter 5, we present a submitted work where parameter estimation on data from rats was performed. In this article, a slightly different model is used. The two differences are the fact that the size of cells is defined as the radius and that we consider a constant diffusion rate $D > 0$. The parameter estimation is performed on stationary solutions of this model. We will show that under some re-parametrization, the model is identifiable using the `Structural identifiability Toolbox` of Maple [2]. It is based on the Structural Identifiability ANalyser (SIAN) algorithm which combines differential algebra and Taylor series approaches [60, 61].

Definition 1.4.8. *Given a model $\mathcal{M} : \Theta \to A$, it is globally **identifiable** in $\theta \in \Theta$ if :*

$$\forall \bar{\theta} \in \Theta, \mathcal{M}(\theta) = \mathcal{M}(\bar{\theta}) \Rightarrow \theta = \bar{\theta}.$$

To find optimal parameter values, we use the Covariance Matrix Adaptation Estimation Strategy (CMA-ES). Given a vector of measured radii $(r_i)_{i=1,\dots,N}$, we minimize the cost function which is the opposite of the log-likelihood :

$$\mathcal{L}(\theta) = -\sum_{i=1}^{N} \log(f^{\infty}(r_i, \theta)), \tag{1.72}$$

where f^{∞} is the stationary solution to the model in radii with the set of parameters θ. The CMA-ES was developed with the goal of non-convex non-linear black-box optimization. This black-box property allows it to perform well for optimization problems where we cannot compute the gradient of the cost-function. We go into a bit of detail as to how the CMA-ES works.

CMA-ES

The CMA-ES falls in the class of genetic algorithms : from a pool of samples, the best samples are kept and are called the parents, and from these samples a new pool is generated, called the offsprings. The goal of such method is often to make sure that the offsprings are better than the parents. In the CMA-ES, this is done by updating the covariance matrix of the last generation using the best samples from the last generation.

Let us consider an objective function $h : \mathbb{R}^n \to \mathbb{R}$ which we want to minimize. The CMA-ES is an iterative process. At step k, we have a sample of candidate solutions $(x_j^k)_{j=1,\dots,\lambda}$, and we go to the next step by drawing :

$$x_j \sim m_k + \sigma_k \mathcal{N}(0, C_k), \quad j = 1, \dots, \lambda,$$

where m_k is the distribution mean, σ_k the step-size and C_k a symmetric and positive-definite covariance matrix (with $C_0 = I$). The offsprings $(x_j)_{j=1,\cdots,\lambda}$ are independent and identically distributed random variables. The next distribution mean m_{k+1} is updated using a weighted sum of the best drawn samples. Denote $(x_{j:\lambda})_{j=1,\dots,\lambda} = (x_j)_{j=1,\dots,\lambda}$ where $h(x_{1:\lambda}) \leq \cdots \leq h(x_{\mu:\lambda}) \leq h(x_{\mu+1:\lambda}) \leq \cdots$, where typically $\mu \leq \lambda/2$. Then :

$$m_{k+1} = \sum_{i=1}^{\mu} \omega_i x_i,$$

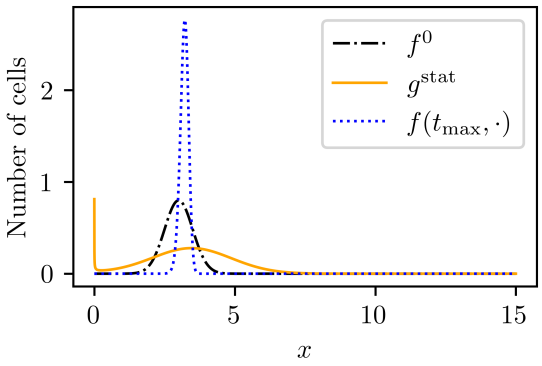


Figure 1.4: **Comparison between the diffusive Lifshitz-Slyozov model and the usual Lifshitz-Slyozov model :** $f0$ is the initial condition for both models. g^{stat} is the stationary solution of Equation (1.32a) and $f(t_{\max}, \cdot)$ the result of an UpWind scheme for the Lifshitz-Slyozov model without diffusion stopped at time $t_{\max}$ before reaching a Dirac mass.

where the weights $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_\mu$ sum to one and are defined such that $\mu_\omega = \left(\sum_{i=1}^{\mu} \omega_i^2\right)^{-1} \simeq \lambda/4$.

Then the step size is updated via the cumulative step-size adaptation. We introduce the evolution path p_k^σ of the step-size and define :

$$p_{k+1}^\sigma = (1 - c_\sigma) p_k^\sigma + \sqrt{1 - (1 - c_\sigma)^2} \sqrt{\mu_\omega} C_k^{-1/2} \frac{m_{k+1} - m_k}{\sigma_k},$$

where $c_\sigma^{-1} \simeq n/3$ and the step size is updated following :

$$\sigma_{k+1} = \sigma_k \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|p_{k+1}\|}{\mathbb{E}\left[\|\mathcal{N}(0, I)\|\right]} - 1\right)\right),$$

where d_σ is a damping parameter. Then we introduce the evolution path p_k^C of the covariance matrix and define :

$$p_{k+1}^C = (1 - c_\sigma) p_k^C + \mathbb{1}_{[0, \alpha\sqrt{n}]}(\|p_{k+1}^\sigma\|) \sqrt{1 - (1 - c_\sigma)^2} \sqrt{\mu_\omega} \frac{m_{k+1} - m_k}{\sigma_k}.$$

Finally the covariance matrix is updated following :

$$C_{k+1} = (1 - c_1 - c_\mu + c_s) C_k + c_1 p_{k+1}^C (p_{k+1}^C)^T + c_\mu \sum_{i=1}^{\mu} \omega_i \frac{x_{i:\lambda} - m_k}{\sigma_k} \left(\frac{x_{i:\lambda} - m_k}{\sigma_k}\right)^T.$$

1.5 Main results

1.5.1 New models for the size distribution of adipose cells

We have introduced a variety of models for the size distribution of adipose cells. The main model is the diffusive Lifshitz-Slyozov model (1.32) where the diffusion term is derived from the discrete Becker-Döring model (1.26) which we study in chapter 3. We are particularly interested in the stationary solutions of the diffusive Lifshitz-Slyozov model, which have an explicit formula (1.30a) and can therefore be computed exactly. We have compared the stationary solutions to the stationary solutions of the other models when they are computable. This is not the case for the classical Lifshitz-Slyozov model, but we can still study them numerically. In Figure 1.4, we show that in general the two stationary solutions are difficult to compare. For the case of adipose modeling, we observe that the velocity has two attractive zeros hence we expect the stationary distribution to divide between the two zeros. However how the solutions split between the two zeros is difficult to predict. Let us illustrate this with some general considerations.

Consider the transport PDE with a velocity v depending on x and t that has three zeros for all time $t \geq 0$ which we denote by $z_1(t) < z_2(t) < z_3(t)$. We assume that these zeros have limits when $t \to \infty$ and that, for all time, z_1 and z_3 are attractive and z_2 is to be repulsive such that $\partial_z v(z_{1,3}(t), t) < 0$ and $\partial_z v(z_2(t), t) > 0$. What we expect to happen is for the distribution to concentrate around the points $z_{1,3}^\infty = \lim_{t\to\infty} z_{1,3}(t)$. Upon this assumption, there is no observation that can be made on the relations between characteristic curves and $z_2(t)$. In particular we do not expect that if $X(s; s, x) = x > z_2(s)$ then $X(t; s, x) > z_2(t)$ for all $t > s$. This means that $z_2(t)$ does not split the ordered family of characteristics into two sub-families : the ones going to z_1^∞ and the ones going to z_3^∞. Hence characteristics may cross the 'boundary' $z_2(t)$ at any time but as times goes to infinity this 'boundary' gets harder to cross since if z_2 does not depend on time, it would be a proper boundary for the characteristics. Now this is what happens in the case of the usual Lifshitz-Slyozov model, the velocity may have up to three zeros which depend on time and the repulsive zero does not split the family of characteristics. Describing the behavior of these zeros in the case of our model is unfortunately quite difficult. Even if we are able to obtain numerical approximation of the stationary solutions, comparing it to the stationary solution of the diffusive Lifshitz-Slyozov model is difficult. Indeed for the diffusive model, the maxima of the stationary solutions are zeros of $\lim_{t\to\infty} v(\cdot, L(t)) - \varepsilon\partial_x d(\cdot, L(t))$. Hence the modes of the usual model and of the diffusive model are not located at the same values of x and moreover the addition of the diffusion in the model accentuates the fact that cells may cross the 'boundary' $z_2(t)$.

In Chapter 4, we study the non-linear Becker-Döring CTMC (1.41) and its diffusion approximation, the Lifshitz-Slyozov SDE (1.45). Both models are constructed from their deterministic equivalent and both can exhibit bimodal stationary distributions.

In Chapter 5, we perform a sensitivity analysis and parameter estimation on the constant and non-constant diffusive Lifshitz-Slyozov model, showing that in both cases we recover parameters of similar order of magnitude.

1.5.2 Convergence result

We previously introduced the classical convergence result from the Becker-Döring model to the Lifshitz-Slyozov model in Theorem 1.4.4. In this section, we introduce a new result in this theory of convergence. We make use of **tails of distributions** to compare the solutions of the Becker-Döring model and the mild solutions of the Lifshitz-Slyozov model. This new method of proof is different from the one used in the classical convergence. In particular, the use of tails of distribution allows us to reduce the non-linearity of the transport PDE by pulling the speed out of the space derivative. We point out that the study of the tails in the case of the Lifshitz-Slyozov model is not new as it was used in [22] to obtain quasi comparison principle and in [77, 19] to get refined uniqueness results. However its use in the theory of convergence from Becker-Döring to Lifshitz-Slyozov is new, up to our knowledge. We now proceed with the result.

Let $(f^\varepsilon, L^\varepsilon)$ be the solution of the Becker-Döring ODE system (1.24), and let (f, L) the mild solution of Lifshitz-Slyozov equations (1.17). We introduce notations for the tails of distributions :

$$F(t,x) = \int_x^\infty f(t,y)\mathrm{d}y, \qquad F^\varepsilon(t,x) = \int_x^\infty f^\varepsilon(t,y)\mathrm{d}y\,, \tag{1.73}$$

and introduce their difference

$$E(t,x) = F^\varepsilon(t,x) - F(t,x)\,. \tag{1.74}$$

We now state our main theorem.

Theorem 1.5.1. *Let $T > 0$. Suppose that there exists some constant $C_{\text{init}} > 0$ such that for all $\varepsilon > 0$, $\int_{\mathbb{R}_+} |E(0,x)|\mathrm{d}x \leq \varepsilon C_{\text{init}}$. Also assume that hypotheses* (H1)-(H9) *hold true. Then there exists some constant $C(T) > 0$ such that for $\varepsilon > 0$ small enough and for all $t \in [0, T]$:*

$$|L^\varepsilon(t) - L(t)| + \int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x \leq \varepsilon C(T).$$

We give a few details as to how the proof works. We begin by a simple observation, taken from [77, 19], which states that the difference between the external lipid amounts can be controlled by the difference of the tails, $|L^\varepsilon(t) - L(t)| \leq \int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. To have proper bounds on the tail,

we use Grönwall's lemma. This is done by deriving the equation for F^ε and F (Lemma 3.4.5). To do so, we need to consider apart the case $x < \varepsilon/2$ since it yields boundary terms that should be treated differently. In turn, this allows for a first bound on the integral $\int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. Next from the equation on F we obtain the equation on E (Lemma 3.4.6). We proceed with bounding the terms in the estimate of $\int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. In particular we are able to show that F^ε solves an equation similar to F up to some terms of order ε. This argument relies on uniform control of the solution of the rescaled Becker-Döring model and its increments $c^\varepsilon_{i+1} - c^\varepsilon_i$.

Hypotheses (H1) - (H4) are classical in the study of our model. However, other assumptions are less common but arise naturally from the result. Since we work with mild solutions, we need additional regularity on the rates. This is due to the fact that the higher order terms that appear, which are the ones we want to look at for choosing a diffusion, are second order derivatives. This entails assumption (H5). Naturally, we require the initial condition to be well prepared. In particular, some properties are propagated, hence hypotheses (H6) and (H7). Additionally, the assumption (H8) allows us to have strict positivity of L^ε in finite time, uniformly in ε. Considering that if this was not the case, the second order diffusive term would be ill-posed, this assumption is natural. The assumption (H9) on the initial condition $c_0^{\varepsilon,0}$ is technical and ensures that the proper boundary condition (1.17c) is satisfied for all times.

Finally the assumption on $\int_{\mathbb{R}_+} |E(0,x)|\mathrm{d}x$ is made to conclude after using Grönwall's lemma at the very end of the proof. This assumption relates both initial conditions $(c_i^{\varepsilon,0})_{i\geq 0}$ and f^0. A fair choice for the initial condition $(c_i^{\varepsilon,0})_{i\geq 0}$ is $c_i^{\varepsilon,0} = f^0(i\varepsilon)$ for all $i \geq 0$. Then we assume f^0 to decrease exponentially toward 0 at infinity and $(f^0)' \in L^1(\mathbb{R}_+, x\mathrm{d}x)$. This yields :

$$\begin{aligned}\int_{\mathbb{R}_+} |E(0,x)|\mathrm{d}x &= \int_{\mathbb{R}_+} |\int_x^\infty f^{0,\varepsilon}(y) - f^0(y)\mathrm{d}y|\mathrm{d}x \\ &\leq \int_{\mathbb{R}_+} x|f^{0,\varepsilon}(x) - f^0(x)|\mathrm{d}x \\ &\leq \sum_{i\geq 1} \int_{\Gamma_i^\varepsilon} x|f^0(i\varepsilon) - f^0(x)|\mathrm{d}x + \int_0^{\varepsilon/2} x|f^0(0) - f^0(x)|\mathrm{d}x.\end{aligned} \tag{1.75}$$

Now the assumption for f^0 to decrease exponentially toward 0 at infinity allows us to cut the previous integral in two parts. Hence there exists some constant $\mu > 0$ such that

$$\int_x^\infty y f^0(y)\mathrm{d}y \leq \int_x^\infty y\exp(-\mu y)\mathrm{d}y = \frac{1}{\mu}(x + \frac{1}{\mu})\exp(-\mu x). \tag{1.76}$$

Then taking $x = -\ln(\varepsilon)$ gives us that there exists some constant C such that $\int_{-\ln(\varepsilon)}^\infty y f^0(y)\mathrm{d}y \leq \varepsilon C$. Now we know there exists some $i \in \mathbb{N}$ such that $-\ln(\varepsilon) \in \Gamma_i^\varepsilon$ which we denote I_ε. In particular, $I_\varepsilon = \lfloor \frac{-\ln(\varepsilon)}{\varepsilon} \rceil$. This allows us to split the sum in (1.75). For $i \leq I\varepsilon + 1$ we use Taylor's expansion and the fact that $(f^0)' \in L^1(\mathbb{R}_+, x\mathrm{d}x)$ while the terms above $I_\varepsilon + 1$ are bounded using the choice of I_ε :

$$\begin{aligned}\int_{\mathbb{R}_+} |E(0,x)|\mathrm{d}x =& \sum_{i=1}^{I_\varepsilon+1} \int_{\Gamma_i^\varepsilon} x|f^0(i\varepsilon) - f^0(x)|\mathrm{d}x + \int_0^{\varepsilon/2} x|f^0(0) - f^0(x)|\mathrm{d}x \\ &+ \sum_{i\geq I_\varepsilon+2} \int_{\Gamma_i^\varepsilon} x|f^0(i\varepsilon) - f^0(x)|\mathrm{d}x \\ \leq& \sum_{i=1}^{I_\varepsilon+1} \int_{\Gamma_i^\varepsilon} x\frac{\varepsilon}{2}|(f^0)'(x)|\mathrm{d}x + \frac{\varepsilon}{2}\int_0^{\varepsilon/2} x|(f^0)'(x)|\mathrm{d}x \\ &+ o(\varepsilon^2)(\sum_{i=1}^{I_\varepsilon+1} \int_{\Gamma_i^\varepsilon} x\mathrm{d}x + \int_0^{\varepsilon/2} x\mathrm{d}x) + 2C\varepsilon \\ \leq& \frac{\varepsilon}{2}\|(f^0)'\|_{L^1(\mathbb{R}_+, x\mathrm{d}x)} + o(\varepsilon^2\ln(\varepsilon)) + 2\varepsilon C.\end{aligned}$$

Therefore if we construct $c_i^{\varepsilon,0} = f^0(i\varepsilon)$, sufficient conditions on f^0 for (H8) to hold are : f^0 decreases exponentially toward 0 at infinity and $(f^0)' \in L^1(\mathbb{R}_+, x\mathrm{d}x)$.

Finally we point out that our result uses the fact that there exists a solution to both models and proves convergence to one another, whereas the classical results only use the existence of solutions to Becker-Döring and show convergence to a measured-valued function that is a weak solution of Lifshitz-Slyozov. Hence the classical result is also an existence result for weak solutions to Lifshitz-Slyozov.

1.5.3 Stochastic convergence

Our motivation to study the Becker-Döring model and to construct a similar model for adipose cells is to get an intuition as to what form could a diffusive Lifshitz-Slyozov model take. Moreover we are particularly interested in showing some bounds on the distance between solutions of the Becker-Döring model and solutions of the diffusive Lifshitz-Slyozov model (1.32). This proved particularly difficult by using the ideas from the classical convergence and the convergence result we presented before was a first step in another direction but has not yet been fruitful. Nonetheless there also exists some results on the convergence from Becker-Döring to Lifshitz-Slyozov but in a probabilistic sense. In the spirit of these results, we look into two stochastic models for the dynamic of the size of an adipose cells, based on the Becker-Döring and Lifshitz-Slyozov models which we introduced in section 1.3.9. This leads to the following results, inspired from the classical result by Kurtz in [74]. We make the additional assumption :

$$a \text{ and } b \text{ are bounded continuous functions.} \tag{H10}$$

We state our result :

Theorem 1.5.2. *Assume X_ε and Z_ε to be solutions of Equations* (1.41) *and* (1.45) *and assumptions* (H1), (H3) *and* (H10) *to hold. Then for ε small enough and for some $T > 0$, there exists a constant β^T only dependent on T, a and b such that :*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|X_\varepsilon(t) - Z_\varepsilon(t)|\right] \leq \beta^T \varepsilon \ln(\frac{1}{\varepsilon}). \tag{1.77}$$

We make a few comments about this result. Firstly the initial result in [74] involves particles whose jump rates depend on the position of the particle itself. In our case, this is also true but additionally, through L^X and L^Z, it also depends on its expectation and therefore its law. This is the main difficulty in the proof. In particular, this prevents us to obtain almost sure convergence as in [74] and instead we obtain L^1 convergence.

We provide a few details as to how the proof works. The general outline is very similar to [74]. We first make use of a result from [72], where we bound the difference between a compensated Poisson process and a Wiener process and Levy's modulus of continuity of the Wiener process [81]. The final result is obtained by appropriate use of Jensen's inequality and Grönwall's lemma.

Then we relax assumption (H10), and consider sub-linear rates :

There exists two positive constants C_a and C_b such that

$$\text{for all } x \geq 0,\ a(x) \leq C_a(1+x) \text{ and } b(x) \leq C_b(1+x). \tag{H11}$$

1.5.4 Numerical results

In addition to the previous theoretical results, we use the previously described numerical schemes (1.62) and (1.68) to investigate the properties of our models. Firstly we recover convergence toward a stationary state numerically. Secondly we show that our model can recover the bimodality property of distributions of adipose cells. However this property is highly dependent on parameter and we show that depending on the parameter λ there is a whole continuum of stationary solutions which exhibit different types of modality. For low amounts of total lipids, only small cells are present and the distribution has a global maximum at $x = 0$. For high amounts of external lipids, the stationary solution is composed of only large cells and resembles to a Gaussian distribution centered on a large value of x. In between, we recover the bimodal property. Values of lambda that are taken to obtain these behaviors range from 0 to 15, but the associated values of L^{stat} only range from 0 to 0.1. This should be coherent with the fact that in an individual, the blood lipid amount should remain at a constant non-toxic level and that excessive lipid intake is stored

in the adipose tissue. An example of the result of the scheme is given in Figure 1.5 and the various behaviours of the stationary solutions is illustrated in Figure 1.6.

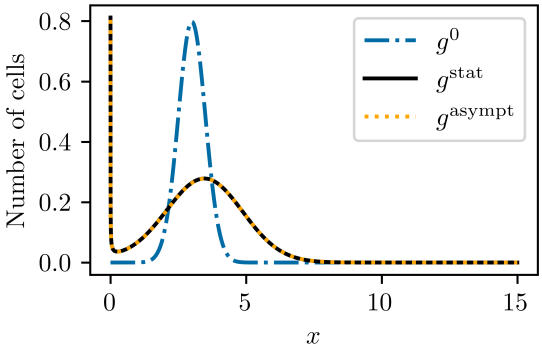


Figure 1.5: **Result of the numerical scheme :** a bimodal distribution obtained by the scheme and computed explicitly. The result of the scheme g^{asymtp} and the stationary solution computed explicitly g^{stat} are superimposed.

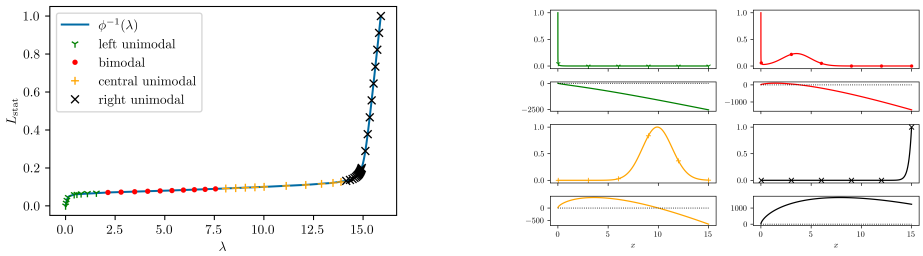
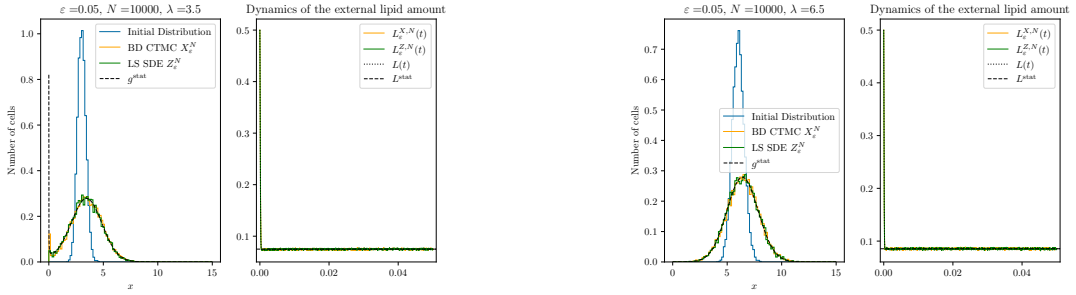


Figure 1.6: **Different types of stationary solutions :** on the left we plot the inverse of the function $\phi : L \to L + \int_0^\infty x M_L(x)\mathrm{d}x$. The inverse ϕ^{-1} is plotted because the actual value we can change in the model is λ and not L^{stat}. Some points are highlighted which corresponds to four different types of stationary solutions : left unimodal (green), bimodal (red), central unimodal(yellow) and right unimodal(black). For each color a stationary solution is plotted on the right in the same color, with its velocity below it. One may observe that for each type of stationary solutions the modes are located close to the zeros of the velocity as expected, expect for the right unimodal case, where the velocity is positive on the whole domain and therefore the distribution is transported toward the boundary $x_{\max}$ of the domain.

We are also interested in comparing the usual Lifshitz-Slyozov model to the diffusive Lifshitz-Slyozov model numerically. However this proves difficult for a few reasons. Firstly we observe that for similar parameters the stationary solution of the diffusive model does not corelate with the asymptotical approximation of an UpWind scheme for the usual Lifshitz-Slyozov model. We are unfortunately unable to compute explicitly stationary solutions to the Lifshitz-Slyozov model with our choices of rates and therefore rely on the asymptotical numerical approximation to make our comparison. We observe that the diffusive model may exhibit bimodality whereas the usual model only shows unimodality for similar parameters. This may be explained by looking at the positions of the zeros of the velocity v. Indeed looking at the characteristic curves (1.53), one can see that they reach a stationary state iff $v(X(s;t,x),L(s)) = 0$. Hence for the Lifshitz-Slyozov model, we expect the distribution to concentrate at zeros of the velocity. The case where the velocity has one zero has been investigated in [18] where they show concentration toward a Dirac mass located at the zero of the velocity. The case of multiple zeros is up to our knowledge an open question for the Lifshitz-Slyozov model. For the case of adipose modeling, we observe that the velocity has two attractive zeros hence we expect the distribution to divide between the two zeros in large time.

In regards to the stochastic simulations, we show that the stochastic models are both able to recover the different types of stationary distributions, see Figure 1.7. In the spirit of Theorem 1.5.2, we obtain a bound of order $\varepsilon \ln(\varepsilon)$ for the empirical L^1-norm between the two solutions of Equations (1.71) and (1.43), illustrated in Figure 1.8. The empirical L^1-norm is computed as

$\sup_{t\leq T} \frac{1}{N^2} \sum_{i,j=0}^{N} |X_\varepsilon^{i,N}(t) - Z_\varepsilon^{i,N}(t)|$.



(a) Bimodal stationary solution

(b) Unimodal stationary solution

Figure 1.7: Two types of modality for both stochastic models with the dynamic of L. We obtain convergence toward the corresponding stationary solution of the diffusive Lifshitz-Slyozov model.

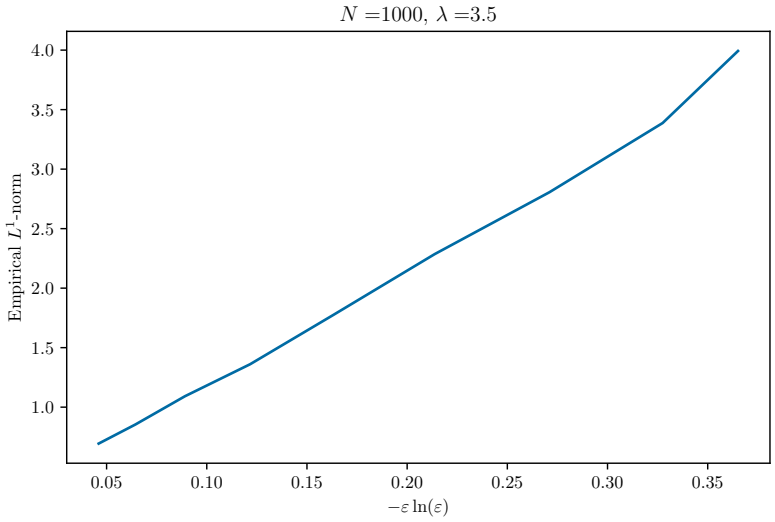


Figure 1.8: **Numerical illustration of Theorem 1.5.2 :** empirical L^1-norm between two solutions of Equations (1.71) and (1.43)

1.5.5 Parameter estimation

In Chapter 5, we present a submitted work on the parameter estimation of the constant diffusive Lifshitz-Slyozov model in collaboration with Chloé Audebert, Anne-Sophie Giacobbi, Magali Ribot, Hédi Soula and Romain Yvinec, as well as further estimations for the diffusive Lifshitz-Slyozov model.

We previously tried very basic estimation methods such as gradient descent and least square methods. These methods were often unconclusive and we seeked a more robuste method that is capable of tackling harder problems. A first method we used was the Approximate Bayesian Computation (ABC). This method uses an a priori law for each parameter and generates an a posterio law using bayesian probability. The results from the ABC were however difficult to interpret without a proper sensitivity and identifiability analysis of the model. This is done in Chapter 5.

Hence we choose to use the CMA-ES method. This method is robust in regards to non-convex, non-linear problems and does not need to compute any gradients. Since we lack theoretical results on our minimizing problem, this method is well suited because it can be used in a very general case.

The constant diffusive Lifshitz-Slyozov model is written in radii for this work instead of lipids but the overall shape of the model remains the same :

$$\begin{cases} \partial_t f(t,r) + \partial_r(v(r,L(t))f(t,r)) = D\partial_r^2(f(t,r)), & \text{(1.78a)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} (V(r)-V_0)\frac{4\pi r^2}{V_{\text{lipids}}^2} f(t,r)\mathrm{d}r = \lambda, & \text{(1.78b)} \end{cases}$$

where $V(r) = \frac{4}{3}\pi r^3$ and the velocity is defined as $v(r,L) = a(r)\frac{L}{L+\kappa} - b(r)$, which is similar to the velocity of Lifshitz-Slyozov model (1.17). The function a and r are given by :

$$a(r) = \frac{\alpha V_{\text{lipids}}}{4\pi}\frac{\rho^3}{r^3+\rho^3}\frac{L}{L+\kappa},$$

$$b(r) = \frac{\beta+\gamma r^2}{r^2}\frac{V(r)-V_0}{V(r)-V_0+V_{\text{lipids}}\chi}.$$

Remark. *One may go from Equation* (1.17a) *to Equation* (1.78a) *without the diffusion term, and vice versa via the diffeomorphism* $x \to r(x)^2$.

The unknowns of the model are $\alpha, L, \kappa, \rho, \chi$ and D. We first perform an identifiability analysis of the parameters.

The model (1.78) is re-parametrized with $\theta_1 = \frac{\alpha L}{L+\kappa}$. Using the `Structural identifiability Toolbox` of Maple [2], we show that the model is indentifiable under this re-parametrization. Then we use the CMA-ES method on synthetic data generated using stationary solutions of (1.78). We show that the method performs well even when censoring the small data points, where three out of four parameters are well estimated, see Table 1.1. Small data points censoring is investigated because depending on the machine setup during data collection the minimal radius measured is different, hence some of the data need to be censored. Then the method is used on data from 32 rats, where we show the ability of the model to recover the particular bimodal shape of the distributions and an estimation for the identifiable parameters, see Table 1.2. However the nadir, the minimum between the two modes, is not perfectly reproduced, see Figure 1.9. This is supported by the sensitivity analysis, that shows that none of the parameters has a significant influence on the nadir. Nonetheless we obtain an estimation for the surface rate of lipogenesis which is of similar order of magnitude to the surface rate of lipolysis. Additionally we show that for radii larger than 20 μm lipolysis is mainly a surface based mechanism, as in [119].

The last part of Chapter 5 is dedicated to applying the CMA-ES method to the diffusive Lifshitz-Slyozov model (1.27).

We use the same re-parametrization as in the first part, but we were unable to show that this re-parametrization is identifiable for this model. Nonetheless we recover values of parameters in the same order of magnitude as in the first part, see Table 1.3, with a slight improvement to the fitting of the nadir by censoring large data points, see Figure 1.10.

1.5.6 Recap of the main results

Let us briefly recap the main results of this manuscript chapter by chapter. Chapter 6 is dedicated to a work done during the summer school CEMRACS 2022, and does not relate to adipose cell modeling. Therefore it is presented independently in Section 1.6.

Chapter 3 : A Lifschitz-Slyozov type model for adipocyte size dynamics : limit from Becker-Döring system and numerical simulation

- a new second order Lifshitz-Slyozov model (1.32) for adipocyte size distribution with a diffusion term derived from a discrete model,
- Becker-Döring and Lifshitz-Slyozov systems with an unusual velocity (1.11) -(1.13) with three zeros and a saturation term in L, which leads to different types of stationary solutions,
- an additional conservation law (1.16) with respect to classical systems, enforcing uncommon boundary conditions, see Eq. (2.15),

synthetic data set 1			10,000 samples			samples $> 10\mu m$		
parameter	order	true	esti. value	esti. ±20%	select. values	esti. value	esti. ±20%	select. values
θ_1	10^{-3}	9.60	9.61	7.69 - 11.53	9.58 - 9.63	9.62	7.70 - 11.54	9.59 - 9.65
ρ	10^{2}	1.50	1.50	1.20 - 1.80	1.47 - 1.53	1.49	1.19 - 1.79	1.46 - 1.52
θ_3	10^{3}	2.18	2.17	1.74 - 2.60	2.05 - 2.29	2.09	1.67 - 2.51	1.91 - 2.29
θ_4	10^{-3}	7.37	7.20	5.76 - 8.64	6.54 - 8.02	7.35	5.88 - 8.82	6.58 - 8.32
synthetic data set 2			10,000 samples			samples $> 10\mu m$		
parameter	order	true	esti. value	esti. ±20%	select. values	esti. value	esti. ±20%	select. values
θ_1	10^{-3}	9.92	9.92	7.94 - 11.90	9.90 - 9.95	9.91	7.92 - 11.89	9.86 - 9.95
ρ	10^{2}	2.00	2.00	1.60 - 2.40	1.97 - 2.03	2.01	1.61 - 2.41	1.99 - 2.05
θ_3	10^{3}	3.27	3.12	2.49 - 3.74	2.69 - 3.58	5.39	4.31 - 6.47	4.32 - 6.47
θ_4	10^{-2}	1.11	1.12	0.90 - 1.34	1.05 - 1.21	1.12	0.90 - 1.34	0.98 - 1.28

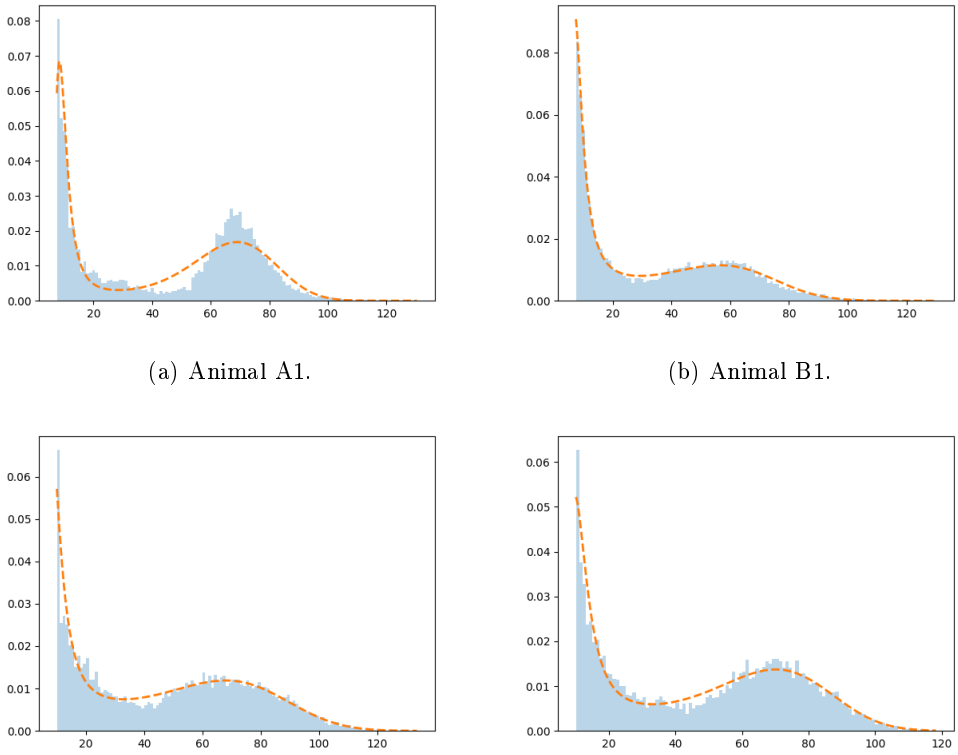
Table 1.1: **Range of selected values for the parameters.** The first three columns show the parameter names, order and true value and $\theta_1 = \frac{\alpha L}{\beta(L+\kappa)}$, $\theta_3 = V_{\text{lipids}}\chi$ and $\theta_4 = \frac{4\pi D}{V_{\text{lipids}}\beta}$. For each data set, the estimated parameter value (column "esti. value") with CMA-ES method is subject to a maximum of 20% variation (column "esti. ±20%"). From this variation, a range of values is selected for each parameter (column "selec. values") allowing a maximum error rate of 0.1% on the value of the estimated cost function $\mathcal{L}$. For each parameter 1,000 samples are generated

parameters	mean	std	RSD
θ_1	$9.6\,10^{-3}$	$2.8\,10^{-4}$	0.03
ρ	$1.57\,10^{2}$	$0.25\,10^{2}$	0.16
θ_3	$2.24\,10^{3}$	$1.07\,10^{3}$	0.47
θ_4	$8.21\,10^{-3}$	$2.58\,10^{-3}$	0.31

Table 1.2: **Parameter estimation with adipocyte size distributions measured in rats**. The first column is the parameter names. Over 32 estimations with the different animal cell size distributions, the mean is presented in the second column, the standard deviation in the third column and the fourth column is the relative standard deviation *i.e* the ratio of standard deviation over mean. The parameters are estimated with CMA-ES algorithm of `fmin2` function from `cma` Python package (with 100 initial guesses).

parameters	mean	std	RSD
$\theta_1\beta$	$2.90\,10^{-1}$	$5.00\,10^{-3}$	0.02
ρ	$1.97\,10^{2}$	$2.97\,10^{1}$	0.15
χ	$8.45\,10^{-4}$	$1.53\,10^{-3}$	1.81
ε	$9.95\,10^{-3}$	$4.11\,10^{-3}$	0.41

Table 1.3: **Results on large size censored data.** Parameter estimation with adipocyte size distributions measured in rats : mean, standard variation and relative standard variation. To compare the order of magnitude to Table 1.2, we have that β, the basal lipolysis rate, is of order 10^2 and $\theta_3 = \chi V_{\text{lipids}}$ where V_{lipids} is of order 10^6.



(a) Animal A1.

(b) Animal B1.

(c) Animal C1.

(d) Animal C5.

Figure 1.9: **Comparison model-data**. Four examples (over 32) of adipocyte radius distributions (in μm) as histograms in rat in normal physiological conditions and model output computed (dash lines) with estimated parameters (see Section 5.1.4). The parameter estimations are performed with CMA-ES algorithm of `cma` Python package by minimizing the function $\mathcal{L}$ eq. 5.18.

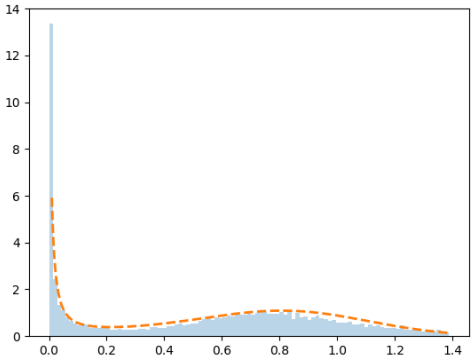


Figure 1.10: **Comparison model-data.** example of size distributions in lipids as histograms and model outputs (dash-lined) computed from the results of the CMA-ES method censoring the last 20% of the data.

- a new proof of convergence result from Becker-Döring solutions to Lifshitz-Slyozov solutions, using tails of distributions, that provides an upper bound on the speed of convergence,
- numerical results showing that bimodal distributions, as well as unimodal profiles, can be obtained asymptotically with system (1.32), according to the parameters,
- numerical results exploring the influence of parameter ε and comparing the diffusion term of system (1.32) with a time and space constant coefficient,
- numerical results showing that the second order system (1.32) provides universal asymptotic profile that does not depend on initial condition (but only on λ, m), contrary to first order system (1.17).

Chapter 4 : A stochastic approach to adipose cell modeling

- two stochastic models : the non-linear Becker-Döring CTMC (1.41) and the Lifshitz-Slyozov SDE (1.45). Both models are inspired by the deterministic Becker-Döring and Lifshitz-Slyozov models for the modeling of adipose cells,
- a new convergence result for the stochastic models, based on the work from [74], for bounded rates of lipogenesis and lipolysis,
- an extension of this convergence result to the case of sub-linear rates.

Chapter 5 : Parameter estimation and some additional numerical results

- a structural identifiability analysis on the constant diffusive Lifshitz-Slyozov model in radii (1.78) that shows that under a re-parametrization, the model is identifiable,
- a parameter estimation on the constant diffusive Lifshitz-Slyozov model (1.52) using the CMA-ES method. We are able to estimate the surface rate of lipogenesis and show that it is of similar order of magnitude as the surface rate of lipolysis.
- a parameter estimation on the diffusive Lifshitz-Slyozov model (1.32) using the CMA-ES method. We recover parameters of similar order of magnitude as in the constant diffusion case and show that this model better captures the nadir of the distribution.

In Chapter 7, we present some of the ongoing work for the modeling of adipose cells as well as perspectives of this work.

1.6 CEMRACS

Over the month of August 2022, I had the opportunity to participate in the CEMRACS Summer School. This came with the opportunity to work on a project over the course of 5 weeks. This work was done in collaboration with Olivier Bernard, Mickael Bestard, Thierry Goudon, Sebastian Minjeaud, Florent Noisette and Bastien Polizzi. We present here briefly the goals and method of this work since it is out of the scope of the modeling of adipose cells. We refer the reader to Chapter 6 for details. The objective of this work was to perform numerical simulations for a model of biofilms. Biofilms are a consortium of single cell organisms embedded in an extra-cellular matrix that proliferate on a surface. Examples of biofilms range from the microbiota inside the intestine to micro-algae on the surface or rocks. This model describes the interaction of the cells inside the biofilm with its extra-cellular matrix and a liquid medium. This three components are described by their volumic fraction and velocity. As such the volumic fractions sum to one. One of the goal of this project was to adapt a numerical scheme from [13] that is able to preserve the constraint on the sum of the volumic fraction at the numerical level. This scheme was developped for multifluid flows : different types of liquids that are in the same phase. We show that the numerical scheme we use is able to maintain the constraint at the numerical level, see Figure 1.11. We obtain the expected traveling wave pattern, see Figure 1.12, and add various substrates that potentially play a role in the growth of the biofilm.

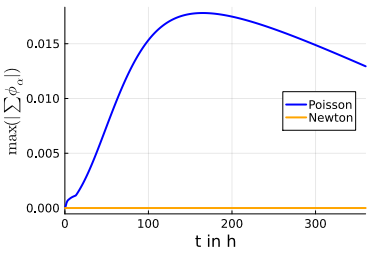


Figure 1.11: Time evolution for the maximal error within the domain on the sum of volume fractions: $E = \max_x |\sum_\alpha \phi_\alpha - 1|$.

(a) Volume fractions at $t = 120$h (b) Volume fractions at $t = 240$h (c) Volume fractions at $t = 360$h

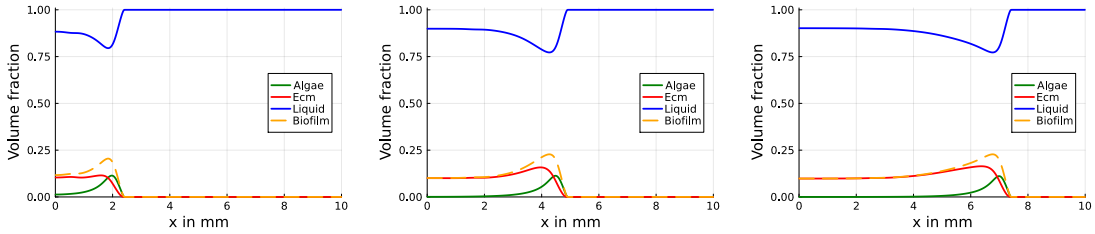


Figure 1.12: **Mixture components volume fractions for different times.**

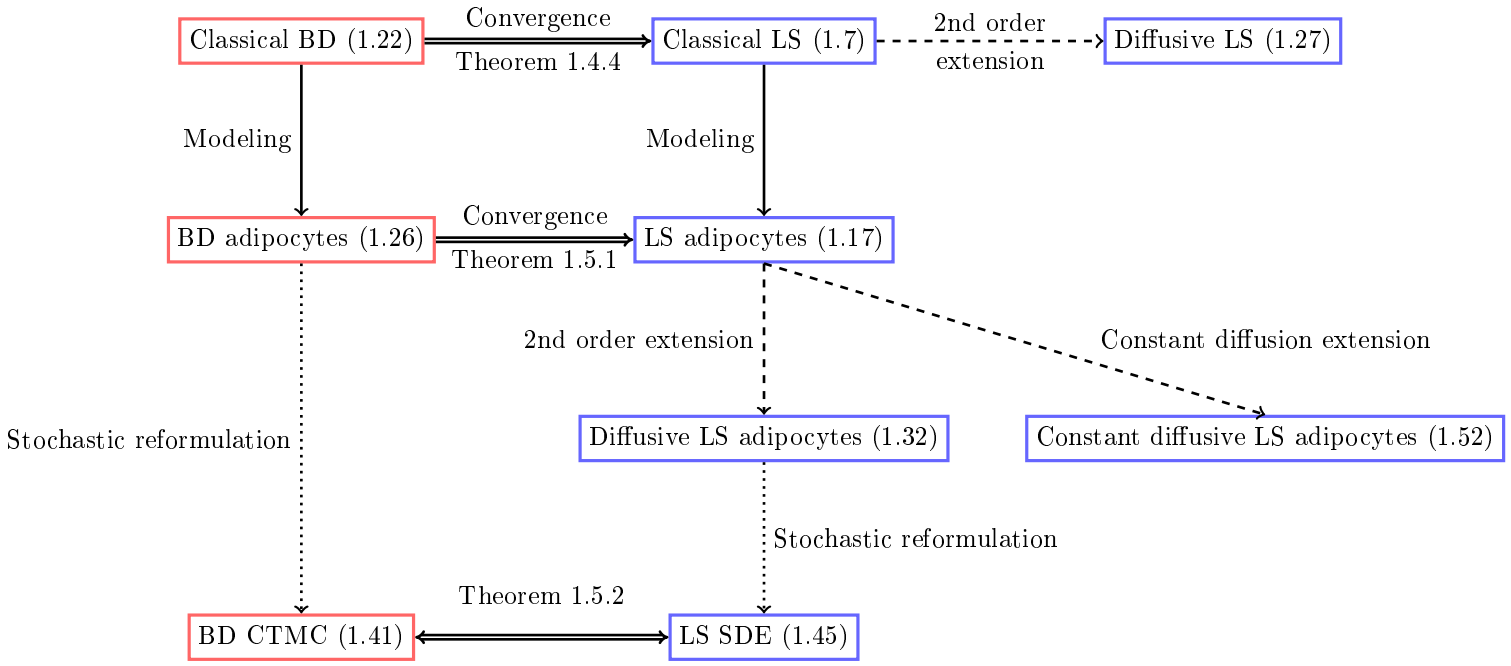


Figure 1.13: Recap of all the introduced models as well as their relations. The Becker-Döring models are in red and the Lifshitz-Slyozov models are in blue

Chapter 2

Introduction en français

Cette thèse porte sur l'étude et la simulation de modèles pour le tissu adipeux. En particulier, nous nous intéressons à la distribution en taille des cellules adipeuses à l'intérieur du tissu adipeux. À ce titre, nous commençons par présenter la biologie du tissu adipeux ainsi que les efforts antérieurs de modélisation. Nous présentons ensuite nos modèles qui sont basés sur des classes plus générales de modèles développés à l'origine pour la modélisation des polymères : le modèle de Lifshitz-Slyozov et le modèle de Becker-Döring.

2.1 Motivation biologique : le tissu adipeux

2.1.1 Généralités

Le tissu adipeux fait partie de la grande famille des tissus conjonctifs. Il est composé de différents types de cellules, dont les majoritaires sont les adipocytes ou cellules adipeuses. Son rôle principal est le stockage de l'énergie sous forme de lipides, mais il fournit également des structures aux organes voisins ainsi qu'une isolation thermique. Chez la plupart des espèces, le tissu adipeux se divise en deux types : le tissu adipeux blanc (WAT) et le tissu adipeux brun (BAT). Selon le type de tissu, on observe différents types de cellules adipeuses : les cellules adipeuses blanches et les cellules adipeuses brunes. Ces cellules se distinguent par la manière dont elles stockent les graisses et par leur rôle. Les cellules adipeuses blanches (WAC), et en particulier les WAT, possèdent une seule gouttelette lipidique à l'intérieur de leur cytoplasme et constituent une forme de stockage de l'énergie. Ces gouttelettes lipidiques donnent la couleur blanche au tissu, d'où son nom. En revanche, les cellules adipeuses brunes (BAC) peuvent avoir plusieurs vésicules de lipides ainsi que plusieurs mitochondries. Leur rôle est d'effectuer une thermogenèse adaptative et de libérer leur contenu énergétique sous forme de chaleur. Le BAT est présent en grande quantité chez les nourrissons humains et régresse ensuite chez les adultes [83]. Bien qu'elles remplissent des rôles différents, les WAC peuvent toujours se convertir en BAC. Ce processus est appelé "beiging" ou "browning" des WAC (du français "beige" : une couleur entre le blanc et le brun), où les WAC se transforment en adipocytes libérant de l'énergie. Ce processus a été observé en réponse à une exposition au froid et est bien documenté dans [114] et est réversible. Dans notre modélisation du tissu adipeux, nous ne prendrons en compte que le WAT et les WAC. Cependant, la collecte de données n'est pas parfaite et nous pouvons supposer que certaines de nos données sont en fait des BAC, des préadipocytes ou des cellules endothéliales, bien que ces types de cellules supplémentaires ne soient présents qu'en faible pourcentage par rapport aux WAC.

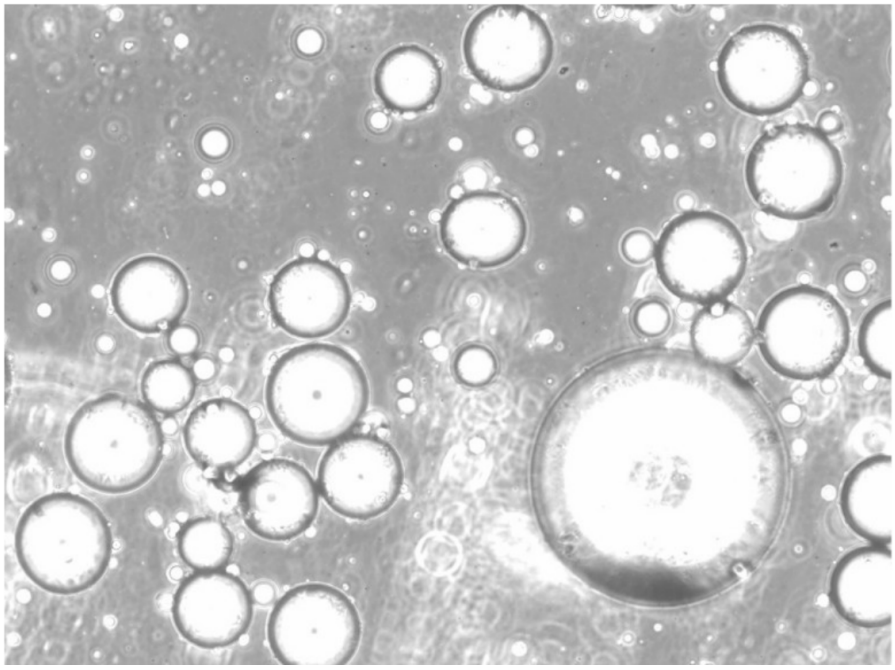


Figure 2.1: Adipocytes of different size

La vie d'un acide gras : de l'intestin au stockage

Les adipocytes stockent l'énergie sous la forme d'une ou plusieurs vésicules de triglycérides à l'intérieur de leur cytoplasme. Ces molécules sont créées à l'intérieur des cellules à partir d'acides gras par un processus appelé **lipogenèse**. Cependant, avant que les acides gras ne pénètrent dans les cellules adipeuses, ils sont ingérés par l'organisme et aboutissent dans l'intestin. De l'intérieur de l'intestin, les micelles contenant les acides gras sont transportées vers les entérocytes où elles sont assemblées avec du glycérol pour former des triglycérides. Ils sont ensuite regroupés avec des apolipoprotéines, des phospholipides et du cholestérol pour former des chylomicrons, également appelés lipoprotéines de très basse densité (ULDL). Cette structure protéique complexe permet le transport des triglycérides à l'intérieur du flux sanguin aqueux vers les tissus de stockage tels que le tissu adipeux. Lorsque ces chylomicrons atteignent les capillaires du tissu adipeux, la lipoprotéine lipase dissout les triglycérides en acides gras libres qui sont transportés dans les adipocytes. À l'intérieur de la cellule, ils sont à nouveau transformés en triglycérides à l'aide de glucose et d'acétyl-CoA. La formation de triglycérides à partir d'acides gras est appelée lipogenèse. La lipogenèse est fortement régulée par une hormone commune : l'insuline. Cela implique évidemment que le diabète et les autres maladies liées à l'insuline ont une influence sur la physiologie du tissu adipeux.

La réaction biochimique inversée par laquelle les triglycérides sont hydrolysés en glycérol et en acides gras et libérés dans les capillaires environnants est appelée **lipolyse**. La lipolyse est régulée par deux hormones : le glucagon et la norépinéphrine. Ces deux processus sont illustrés dans la figure 2.2 tirée de [108].

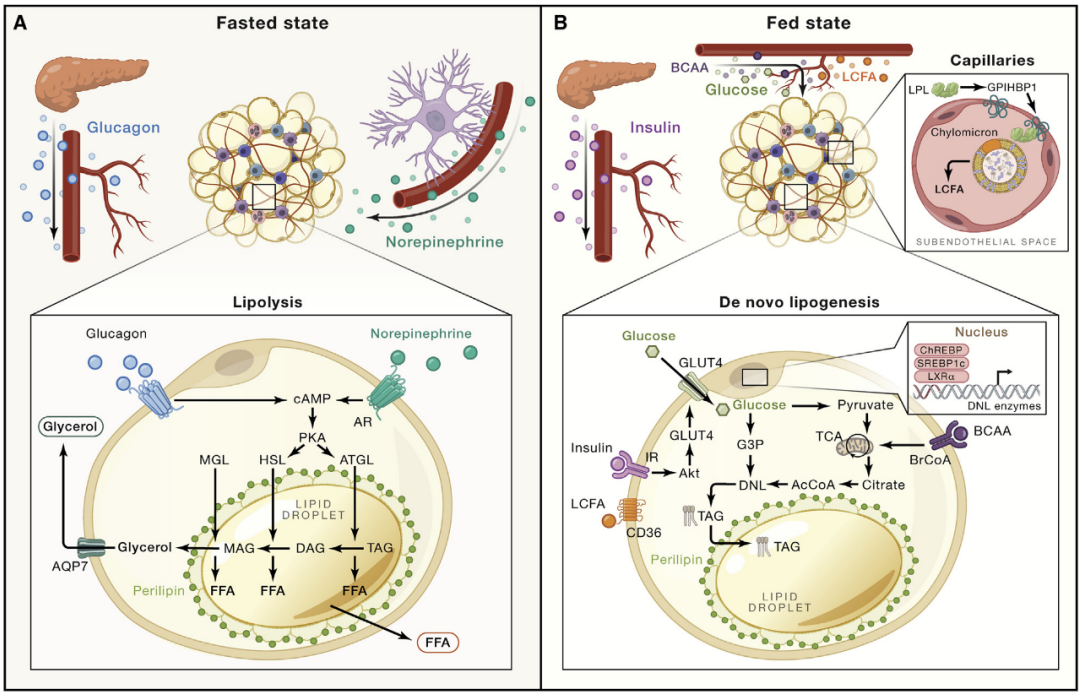


Figure 2.2: An illustration of lipogenesis and lipolysis. This figure is taken from [108].

2.1.2 Morphologie

Localisation

En tant que tissu conjonctif, le tissu adipeux est souvent localisé avec ses organes structurés. Chez l'homme, il est présent à différents endroits. La graisse viscérale est située autour des organes internes et est l'un des indicateurs de problèmes de santé tels que l'obésité. La graisse sous-cutanée est située largement sous la peau chez l'homme, où elle fournit une isolation thermique mineure ainsi qu'un rembourrage. La graisse intramusculaire est située dans le système musculaire et permet un accès rapide au stockage de l'énergie pour les muscles environnants. La graisse médullaire est présente dans la moelle osseuse, mais son rôle au-delà du stockage de l'énergie n'est pas bien compris.

Tissu adipeux et poids

Il existe une corrélation évidente entre le tissu adipeux et le poids d'un individu. De nos jours, la masse grasse est utilisée comme indicateur de santé avec l'indice de masse corporelle (IMC), bien que l'utilisation de l'IMC en tant qu'indicateur de santé soit souvent débattue. Néanmoins, avec l'augmentation récente de l'obésité dans la population humaine [3], la communauté scientifique a commencé à s'attaquer à l'étude de la perte de poids en se penchant sur la physiologie des cellules adipeuses. De nombreux travaux ont mis en évidence ce que l'on appelle généralement la "reprise de poids", c'est-à-dire qu'après une perte de poids importante, les individus reprennent du poids plus rapidement que les individus de même poids [89, 5]. Ce phénomène s'explique par de nombreux facteurs souvent liés à une certaine résistance des cellules adipeuses une fois qu'elles ont atteint une taille suffisante. Il convient toutefois de distinguer deux mécanismes susceptibles

de se produire lors d'une prise de poids. Tout d'abord, l'hypertrophie est l'augmentation du volume des cellules adipeuses. C'est souvent le premier changement physiologique impliquant une prise de poids. Le second est l'hyperplasie, c'est-à-dire l'augmentation du nombre de cellules. L'hypertrophie se produit jusqu'à l'obésité modérée, moment où l'hyperplasie apparaît et où le nombre de cellules augmente, d'où une augmentation de l'espace de stockage disponible [59]. Chez les enfants, de nouvelles cellules adipeuses sont créées au cours de la période de croissance jusqu'à l'âge adulte. Cela conduit souvent à l'idée que les adultes ont une quantité "fixe" de cellules adipeuses. Or, de nouvelles cellules sont créées par le renouvellement du tissu adipeux, qui remplace les cellules qui meurent naturellement. En effet, un individu en bonne santé, en ce qui concerne son poids, verra de petites variations de son nombre de cellules adipeuses, mais comme indiqué précédemment, l'hyperplasie se produit à un stade plus avancé de l'obésité. Aucune étude n'a encore été réalisée pour caractériser l'augmentation du nombre de cellules entre la maigreur et l'obésité, mais de nombreux résultats ont été obtenus sur le processus inverse. La perte de poids induite par des changements alimentaires ou par la chirurgie diminue la taille des cellules mais n'en diminue pas le nombre [6]. Cela conduit à l'hypothèse que l'augmentation du nombre de cellules adipeuses est un processus à sens unique et non réversible.

En tant que telle, la "prise de poids", c'est-à-dire l'augmentation du nombre de cellules adipeuses au-delà de la quantité initiale, est permanente. Cela implique que le moyen le plus efficace de réduire la masse graisseuse au stade de la surcharge pondérale est une opération chirurgicale, soit par l'installation d'un anneau gastrique pour réduire l'apport alimentaire, soit par une liposuccion pour éliminer une partie de la masse graisseuse. Toutefois, comme nous l'avons souligné précédemment, il s'agit de remèdes physiques à la surcharge pondérale, qui ne modifient pas la physiologie des cellules adipeuses restantes.

Lorsque nous examinerons la modélisation du tissu adipeux, nous simplifierons en considérant que le nombre total de cellules est constant. En effet, cela n'est pas biologiquement pertinent, mais cette simplification est un premier pas vers la modélisation et le calcul de la dynamique du tissu adipeux.

Hétérogénéité de taille

De nombreuses études ont montré que le tissu adipeux est fortement hétérogène, que ce soit au niveau des types de cellules, des niveaux hormonaux et de la taille des cellules, ce qui est notre principal intérêt [76, 65, 91]. Comme le montrent les figures 1.3, la taille des cellules à l'intérieur du tissu adipeux suit une distribution bimodale : elle présente deux maxima locaux, appelés modes, séparés par un minimum local appelé nadir. Cependant, aucune explication biologique n'a été avancée pour expliquer pourquoi les cellules adipeuses se séparent en deux populations de petites et de grandes cellules. Il est intéressant de noter que cette séparation a été observée chez de nombreuses espèces, mais que la position et la hauteur relative des modes diffèrent d'une espèce à l'autre et d'un individu à l'autre. La distribution en taille des cellules adipeuses est notamment un indicateur de surpoids ou d'obésité, puisque ces individus auront tendance à avoir un plus grand nombre de cellules adipeuses et à avoir des cellules plus grandes.

2.1.3 Questions de santé

Il est évident que l'étude des cellules adipeuses implique l'étude des problèmes de santé qui y sont liés. La principale pathologie liée au tissu adipeux est l'obésité. Cette maladie a connu une forte augmentation dans la population humaine, ayant triplé depuis 1975, et a même été qualifiée d'épidémie par l'Organisation mondiale de la santé [3]. Le principal symptôme est l'excès de graisse corporelle, qui se caractérise par un indice de masse corporelle (IMC) supérieur à 30.

De nombreuses études ont montré que l'obésité est étroitement liée à diverses afflictions, telles que les maladies cardiovasculaires [76], le diabète de type 2 [42], l'apnée obstructive du sommeil, les cancers [37] et l'arthrose. Les causes sont diverses et souvent différentes d'un individu à l'autre. Elles comprennent l'excès de nourriture, la sédentarité, le manque d'exercice physique, la génétique, etc.

En ce qui concerne le traitement médical, il consiste le plus souvent à modifier les habitudes de vie : régime alimentaire et/ou exercice physique, bien que le changement de régime alimentaire le plus efficace pour une perte de poids à long terme soit encore inconnu. En outre, des médicaments peuvent aider à la perte de poids, mais la façon la plus efficace de traiter l'obésité est en fait la chirurgie bariatrique. Celle-ci comprend le pontage et l'anneau gastriques, la gastrectomie ou la dérivation pancréatique.

2.1.4 Pourquoi voulons-nous étudier la distribution en taille des cellules adipeuses ?

Nous pouvons maintenant expliquer pourquoi nous nous intéressons à la distribution en taille des cellules adipeuses. Avant tout, nous rappelons que la distribution en taille des cellules adipeuses, c'est-à-dire le nombre de cellules par taille, est bimodale : elle présente deux maxima locaux. C'est ce que montre la Figure 1.3 pour un rat. Ce type de distribution est inhabituel pour les cellules spécialisées, qui présentent plus souvent une distribution unimodale standard centrée sur la taille moyenne d'un type particulier de cellule.

Qu'entend-on par "taille" d'une cellule ? Il existe plusieurs façons de définir la taille d'une cellule. Nous allons tout d'abord considérer le rayon comme un moyen de définir la taille. Cela signifie que nous supposons que les cellules sont des sphères, alors que dans le tissu adipeux, et puisque les cellules sont principalement composées d'eau, nous nous attendons à ce que les cellules forment un tissu très dense, où les cellules ne prendront certainement pas la forme de sphères. Le fait que les cellules soient des sphères est donc une simplification, mais vient aussi du fait que pour collecter des données, les biologistes brisent la matrice extracellulaire qui maintient le tissu et que les cellules nagent donc librement dans un milieu, où elles devraient ressembler à des sphères, comme le montre la Figure 2.1. Plus loin, nous considérerons que la taille d'une cellule adipeuse est déterminée par la quantité de lipides contenue dans sa vésicule. Nous relierons le rayon à la quantité de lipides en considérant le volume de la cellule par rapport au volume de la vésicule, ce qui nous permettra de relier les deux définitions de la taille.

Comme indiqué précédemment, il n'y a toujours pas d'explication biologique à cette distribution bimodale. De plus, la caractéristique bimodale semble être universelle parmi les espèces animales, mais la hauteur et la position de chaque maxima dépendent de l'individu. La collecte de données s'accompagne de divers écueils : la chirurgie pour la biopsie est souvent invasive, manque de données longitudinales, considérations éthiques, données tronquées, non exhaustives (pas de comptage total des cellules), ...

2.1.5 Quel type de données avons-nous ?

Détaillons le type de données dont nous disposons. Elles se présentent le plus souvent sous la forme de listes contenant le rayon ou le diamètre d'un sous-ensemble de cellules obtenues par biopsie. La taille de chaque cellule est mesurée à l'aide d'un Multisizer IV de Beckman Coulter, mais en raison des limitations des techniques de mesure, seuls les rayons cellulaires supérieurs à 7,5μm ou 10μm selon les expériences sont mesurés. Par conséquent, nos données peuvent être incomplètes et la procédure chimique qui détruit le tissu avant l'utilisation du multisizer peut conduire à ce que certains artefacts restent dans la solution et soient comptés par la machine.

Par conséquent, parmi toutes les mesures, une petite partie peut ne pas être constituée de cellules mais de composants restants du tissu et également d'autres cellules (BAC, préadipocytes, cellules endothéliales, ...).

2.2 Modélisation du tissu adipeux

Dans cette section, nous examinons différents modèles pour le tissu adipeux. Nous verrons que de nombreuses tentatives ont été faites pour modéliser diverses caractéristiques biologiques du tissu adipeux. Nous présentons d'abord les modèles dans une perspective plus large avant de nous plonger dans les modèles de la dynamique en taille des cellules adipeuses et en particulier les modèles de Jo et al. [68] et Soula et al. [118].

2.2.1 Divers modèles pour le tissu adipeux

De nombreux modèles ont été développés pour étudier le tissu adipeux, chacun d'entre eux visant à capturer une caractéristique particulière de ce tissu.

Lorsqu'on étudie des cellules spécialisées, on se demande souvent comment elles sont créées. Dans le cas des cellules adipeuses, les cellules souches mésenchymateuses, situées dans la moelle osseuse, se différencient en préadipocytes. Ces préadipocytes sont des fibroblastes indifférenciés qui peuvent former des cellules adipeuses lorsqu'ils sont stimulés. Le processus biochimique exact de la détermination du destin des préadipocytes n'a pas encore été entièrement décrit. Une modélisation mathématique a tout de même été réalisée pour appréhender ce processus de différenciation. Dans [46], les auteurs ont développé un modèle pour le processus de différenciation des cellules mésenchymateuses dans les cellules adipeuses. Le modèle introduit dans [46] est composé d'EDO pour la population totale de cellules mésenchymateuses et de cellules préadipeuses, et d'une équation de transport structurée en taille avec un terme de mort pour les cellules adipeuses. Ce modèle et son extension avec une composante spatiale sont capables de reproduire des distributions bimodales.

Nous avons précédemment souligné le fait que le tissu adipeux est un tissu conjonctif et qu'il fournit donc également une structure aux organes ou tissus environnants, comme dans la graisse viscérale par exemple. L'interaction entre la matrice extracellulaire (MEC) et les cellules adipeuses est même corrélée à des problèmes de santé tels que l'obésité, comme indiqué dans [36]. L'interaction entre les cellules et leur matrice extracellulaire a été bien étudiée par la communauté mathématique et certains travaux se sont penchés sur le cas particulier du tissu adipeux. Dans [100], les auteurs ont développé un modèle d'interaction entre les cellules adipeuses et les fibres de collagène. Ils sont capables de reproduire certaines caractéristiques de la morphologie du tissu adipeux, notamment le regroupement des cellules adipeuses et la formation d'un réseau de fibres. Ils utilisent un modèle individu-centré comprenant les forces et l'interaction entre les cellules et les fibres ainsi que la différenciation pré-adipocytaire. Ce modèle a été réutilisé dans [101] pour étudier les lésions du tissu adipeux. Le modèle est notamment capable de reproduire la régénération et la formation de cicatrices.

En outre, les biologistes s'intéressent aux processus biochimiques qui se déroulent à l'intérieur de la cellule et à l'échelle du tissu. Dans [8], les auteurs ont étudié le renouvellement des lipides chez l'homme. Ils ont constaté que le taux de libération des lipides diminue avec l'âge alors que le taux d'ingestion n'est pas ajusté réciproquement. Ils ont également observé qu'une perte de poids importante est corrélée à une diminution du taux d'absorption des lipides et que les personnes ayant un faible taux d'élimination sont plus susceptibles de conserver un poids stable après une perte de poids. Tous ces résultats démontrent le rôle de la consommation et de la libération des lipides dans la morphologie du tissu adipeux et sa relation avec les problèmes de

santé. Les chercheurs ont également adapté leurs données à un modèle PDE structuré en fonction de l'âge pour la population de lipides à l'intérieur d'un individu. La méthodologie de collecte des données dans cet article est particulièrement intéressante : ils mesurent l'âge des lipides dans un individu en examinant le ratio de carbone radioactif provenant des essais de bombes nucléaires qui sont incorporés dans la chaîne de carbone des lipides.

Dans [71], les auteurs introduisent un modèle ODE basé sur la physiologie pour illustrer la compartimentation intracellulaire des processus métaboliques des lipides et l'activation différentielle des lipases impliquées dans ces processus. Ils montrent que le sous-domaine métabolique actif doit être bien calibré car il est essentiel pour la simulation des données. Selon la taille de ce sous-domaine, la vitesse de la dynamique cellulaire peut changer et donc modifier le résultat du modèle. Un domaine plus petit produit une dynamique plus rapide. L'activation différentielle des lipases est due à leur rôle dans les processus métaboliques des différents glycérides : tri-, di- et monoglycérides. Ils montrent que cette activation différentielle leur permet de comparer plus précisément les données expérimentales. Cependant, certaines de leurs conclusions doivent encore être confirmées par des expériences.

2.2.2 Modèles pour la dynamique de la taille des cellules adipeuses

Dans cette section, nous présentons les développements récents en matière de modélisation de la distribution en taille des cellules adipeuses. Les deux principaux articles ont été publiés en 2013 par Jo et al. [68] et Soula et al. [118]. Les modèles introduits dans ce manuscrit, qui seront décrits dans la section suivante, sont liés aux modèles de ces deux articles, mais sont plus étroitement inspirés par celui de [118].

Jo et al. [68]

Dans cet article, les auteurs s'intéressent à l'inférence de la dynamique de distribution en taille des cellules adipeuses à l'aide d'un modèle PDE et de l'inférence bayésienne. Le modèle a été introduit pour la première fois dans [66] mais au lieu du temps, la distribution dépendait de la masse totale du tissu adipeux. Ils effectuent un ajustement de modèle sur des données provenant de rats et sont en mesure de retrouver la distribution bimodale des cellules adipeuses. Leur modèle PDE consiste en une équation de transport-diffusion pour la distribution en taille avec un terme source pour la création de nouvelles cellules et la mort cellulaire. La taille d'une cellule est supposée être son rayon et la distribution des cellules au temps t et au rayon r noté $f(t,r)$ est donnée par :

$$\partial_t f(t,r) = b(t)\delta(r-r_0) - \partial_r(v(r)f(t,r)) + D\partial_r^2 f(t,r) - k(r)f(t,r),$$

où b est le taux de naissance des nouvelles cellules qui apparaissent avec une taille r_0, v est la vitesse de transport, D le taux de diffusion, k le taux de mortalité et δ la fonction delta de Dirac. La vitesse v prend la forme :

$$v(r) = \frac{\nu_+}{2}\left(1+\tanh(\frac{r-r_+}{\eta_+})\right) - \frac{\nu_-}{2}\left(1+\tanh(\frac{r-r_-}{\eta_-})\right).$$

L'idée derrière ce choix de vitesse est phénoménologique : on peut supposer que l'absorption et la libération de lipides par une cellule sont limitées par la surface. Il existe donc des tailles critiques pour l'absorption et la libération, désignées par x_+ et x_-, qui donnent les taux d'absorption et libération semi-maximaux, désignés par $\nu_{+,-}$. Les paramètres $\eta_{+,-}$ sont inclus pour décrire la pente des taux. Il convient de noter que cette vitesse ne dépend pas du temps et ne tient donc pas compte des variations potentielles du taux en fonction de la dynamique du tissu.

Soula et al. [118]

Le modèle introduit dans [118] est la base de tous les modèles que nous introduisons par la suite dans ce manuscrit. Les auteurs dérivent un modèle pour la distribution en taille des cellules adipeuses, où la distribution dépend à la fois de la quantité de lipides x à l'intérieur de la cellule et du rayon r de la cellule. Le même modèle a été utilisé dans [119] pour estimer le taux de surface de la lipolyse. L'EDP décrivant la dynamique de la distribution est donnée par :

$$\partial_t f(t,x,r) = \partial_x(v(x,r,L)f) + \partial_r(R(x,r)f), \tag{2.1}$$

où L est la quantité externe de lipides, qui a aussi sa propre dynamique donnée par :

$$\frac{\mathrm{d}L}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}t}\iint x f(t,x,r)\mathrm{d}x\mathrm{d}r. \tag{2.2}$$

Cette équation pour L traduit l'hypothèse que la quantité totale de lipides est constante. La vitesse dans la direction de la quantité de lipides x prend la forme :

$$v(x,r,L) = a(r)\frac{L}{L+\kappa} - b(x,r),$$

La fonction a décrit le taux de lipogenèse et b le taux de lipolyse. Nous détaillons dans la Section 2.3.1 leur construction et la forme de cette vitesse.

La vitesse dans la direction du rayon r est donnée par :

$$R(x,r) = \frac{1}{\tau}\left(\frac{V_{\text{lipides}}x + V_0}{4\pi r^2} - \frac{r}{3}\right), \tag{2.3}$$

où τ est le taux auquel le rayon de la cellule s'adapte lentement à la quantité changeante de lipides intracellulaires. Les auteurs utilisent des simulations de Monte Carlo pour déterminer les solutions numériques stationnaires du modèle. Ils sont en mesure de récupérer des distributions bimodales du rayon.

2.3 Modèles de Lifshitz-Slyozov et Becker-Döring pour le tissu adipeux

2.3.1 Un modèle de Lifshitz-Slyozov pour les adipocytes

Nous allons maintenant procéder à la construction de notre modèle principal, basé sur le modèle de Lifshitz-Slyozov. Avant tout, nous rappelons les simplifications biologiques que nous avons introduites précédemment et nous introduisons deux nouvelles simplifications physiques :

- Le nombre total de cellules, également appelé *masse* de la distribution est noté m et est une constante du temps.
- La quantité totale de lipides est notée λ et est une constante du temps.
- La quantité de lipides dans une cellule (en nmol), notée x, est linéairement liée à son volume.
- Les cellules sont des sphères.

Le fait que la quantité de lipides dans une cellule soit linéairement liée à son volume est une hypothèse physique, mais dans la réalité cette relation peut ne pas être exacte, et le changement de volume dû à l'apport de lipides est probablement retardé dans le temps. Nous simplifions la forme d'une cellule en sphère pour deux raisons : le volume d'une sphère est facile à calculer et, lors de la collecte des données, les cellules sont débarrassées de la matrice extracellulaire et se déplacent donc librement à l'intérieur du milieu où elles devraient prendre une forme sphérique et être mesurées comme telles.

Nous basons notre travail sur [118] où les auteurs construisent un modèle dans lequel la distribution dépend à la fois du rayon et de la quantité de lipides. Notre troisième simplification nous permet de relier les rayons aux lipides et donc de simplifier le modèle. Pour revenir au modèle de [118], cela correspond à prendre la limite $\tau \to 0$ dans l'équation (2.3). Cette relation se lit comme :

$$V_{\text{lipids}} x + V_0 = \frac{4}{3}\pi r(x)^3, \tag{2.4}$$

où V_{lipids} est le volume occupé par 1 nmol de lipides (en particulier des triglycérides), V_0 est le volume d'une cellule vide et $r(x)$ est le rayon d'une cellule contenant x lipides. Cela se lit simplement comme *volume de la vésicule lipidique + volume du cytoplasme et des organites = volume de la cellule.*

Cette relation est particulièrement importante car elle nous permet de relier la quantité de lipides au rayon d'une manière bijective, avec :

$$r(x) = \left(\frac{3}{4\pi}(V_{\text{lipids}} x + V_0)\right)^{\frac{1}{3}}. \tag{2.5}$$

Dorénavant, lorsque nous ferons référence à la taille d'une cellule, nous considérons la valeur de x. Considérons donc une cellule de taille x et commençons par décrire l'évolution de sa taille. Comme nous l'avons mentionné précédemment, deux processus se produisent : l'absorption de lipides, appelée lipogenèse, et la libération de lipides, appelée lipolyse. La variation de x est donc la différence entre le taux d'absorption et le taux de libération. Nous construisons ces deux taux sur la base d'hypothèses biologiques et physiques.

Le taux d'ingestion est le produit de trois facteurs :

- un terme pour un écoulement limité en surface $\alpha r(x)^2$, où la constante α est la vitesse de cet écoulement,
- un terme avec un rayon de coupure ρ tel que pour un plus grand rayon, le taux d'absorption est fortement diminué $\dfrac{\rho^n}{r(x)^n + \rho^n}$,
- un terme de Michaelis-Menten pour la quantité de lipides disponibles dans le milieu $\dfrac{L}{L+\kappa}$, où κ est la valeur de L à laquelle ce terme est égale à $\frac{1}{2}$.

Et le taux de libération est le produit de deux termes :

- un terme avec un niveau basal de libération β et un taux limité en surface $\gamma r(x)^2$, où la constante γ est l'équivalent de la constante α,
- un terme de Michaelis-Menten pour la quantité disponible de lipides dans la cellule $\dfrac{x}{x+\chi}$, où χ est l'équivalent de κ pour l'absorption.

Nous soulignons que la construction du taux de libération est basée sur des observations biologiques détaillées dans [119]. En particulier, les valeurs de β et γ ont été estimées expérimentalement dans [119].

La variation de la taille x, que nous notons $\frac{\mathrm{d}x}{\mathrm{d}t}$, peut donc être exprimée comme la différence entre les taux d'absorption et de libération :

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \underbrace{\alpha r(x)^2 \frac{\rho^n}{r(x)^n + \rho^n} \frac{L}{L+\kappa}}_{intake} - \underbrace{(\beta + \gamma r(x)^2) \frac{x}{x+\chi}}_{release}. \tag{2.6}$$

Pour les deux taux, nous regroupons les termes en x selon :

$$a(x) = \alpha r(x)^2 \frac{\rho^n}{r(x)^n + \rho^n} \frac{L}{L+\kappa}, \tag{2.7}$$

$$b(x) = (\beta + \gamma r(x)^2) \frac{x}{x+\chi}. \tag{2.8}$$

Maintenant, d'un point de vue physique, on peut voir que les cellules ont des particules se déplaçant dans un espace d'état avec un champ vectoriel :

$$v(x, L) = \frac{\mathrm{d}x}{\mathrm{d}t} = a(x) \frac{L}{L+\kappa} - b(x). \tag{2.9}$$

A partir de la théorie des équations de continuité, cela nous amène à écrire une équation de transport homogène pour la distribution des cellules $f(t, x)$, qui donne la quantité de cellules adipeuses avec la quantité de lipide x au temps t :

$$\partial_t f(t, x) + \partial_x (v(x, L) f(t, x)) = 0. \tag{2.10}$$

Selon notre simplification biologique, la quantité totale de lipides dans notre système, notée λ, devrait être constante. Il existe deux types de lipides dans le système : ceux contenus dans les cellules et les lipides du milieu. Cela donne l'égalité suivante :

$$L(t) + \int_{\mathbb{R}_+} x f(t, x) \mathrm{d}x = \lambda. \tag{2.11}$$

En ce qui concerne les conditions aux bords, nous voulons conserver la masse de la distribution et nous devons donc imposer que :

$$\int_{\mathbb{R}_+} f(t, x) \mathrm{d}x = \int_{\mathbb{R}_+} f^0(x) \mathrm{d}x = m \text{ for all } t > 0. \tag{2.12}$$

Cela conduit à avoir la condition aux bords $(v(x, L(t)) f(t, x))|_{x=0} = 0$ pour tout $t > 0$. Ainsi, puisque $v(0, L) > 0$, pour conserver la masse, une condition aux bords de Dirichlet suffit :

$$f(t, x)\big|_{x=0} = 0.$$

Cela nous amène au modèle de Lifshitz-Slyozov pour les cellules adipeuses avec des conditions initiales (f^0, L^0) :

$$\begin{cases} \partial_t f(t,x) + \partial_x(v(x,L(t))f(t,x)) = 0, & \text{(2.13a)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} x f(t,x)\mathrm{d}x = \lambda, & \text{(2.13b)} \\ f(t,x)\big|_{x=0} = 0, & \text{(2.13c)} \\ f(0,x) = f^0(x) \text{ and } L(0) = L^0. & \text{(2.13d)} \end{cases}$$

2.3.2 Un modèle de Becker-Döring pour les adipocytes

Dans cette sous-partie nous allons construire un analogue aux équations de Becker-Döring avec une vitesse adaptée à la modélisation de la dynamique des adipocytes. Le but de cette construction est d'étudier les théorèmes de convergence classiques de Becker-Döring à Lifshitz- Slyozov et en déduire la forme d'un terme de diffusion à ajouter dans notre modèle.

Nous allons maintenant considérer qu'une cellule adipeuse est un regroupement de vésicules de taille typique Λ. Ainsi, la taille d'une cellule peut être définie par le nombre de vésicules qu'elle contient. Pour tout $i \geq 0$, on note c_i le nombre de cellules de taille i et l le nombre de vésicules dans le milieu. Une telle cellule va agréger une nouvelle vésicule à la vitesse $a_i \frac{l\Lambda}{l\Lambda+\kappa}$ et perdre une vésicule à la vitesse b_i, suivant cette réaction :

$$L + iL \underset{b_{i+1}}{\overset{a_i \frac{l\Lambda}{l\Lambda+\kappa}}{\rightleftharpoons}} (i+1)L.$$

Soit $c = (c_i)_{i\geq 0}$. Le flux $J_i(c,l)$ de la réaction précédente est alors donné par :

$$J_i(c,l) = a_i \frac{l\Lambda}{l\Lambda+\kappa} c_i - b_{i+1}c_{i+1},\ i \geq 0.$$

De même, comme le modèle de Lifshitz-Slyozov, l satisfait une équation tenant compte de la conservation de la quantité de lipides. Nous obtenons le système ODE suivant :

$$\begin{cases} \dfrac{\mathrm{d}c_i}{\mathrm{d}t} = J_{i-1}(c,l) - J_i(c,l),\ \forall i \geq 1, & \text{(2.14a)} \\ \dfrac{\mathrm{d}c_0}{\mathrm{d}t} = -J_0(c,l), & \text{(2.14b)} \\ l(t)\Lambda + \displaystyle\sum_{i=0}^{\infty} i\Lambda c_i(t) = \lambda,\ \forall t \geq 0, & \text{(2.14c)} \\ l(0) = l^0, \quad c_i(0) = c_i^0,\ \forall i \geq 1, & \text{(2.14d)} \end{cases}$$

L'étude de la convergence du modèle de Becker-Döring vers le modèle de Lifshitz-Slyozov est généralement effectuée après remise à l'échelle du premier. Il existe une variété de théorèmes sur cette convergence [78, 129, 35, 34, 110], et nous prouvons une version de ces théorèmes pour notre modèle de Becker-Döring légèrement modifié pour les cellules adipeuses en annexe B.

2.3.3 Modèle de Lifshitz-Slyozov diffusif pour les adipocytes

Nous introduisons un modèle de Lifshitz-Slyozov diffusif pour les adipocytes, motivé par les résultats de convergence de Becker-Döring vers Lifshitz-Slyozov. On note le taux de diffusion d donné par :

$$d(x,L) = a(x)\frac{L}{L+\kappa} + b(x).$$

Concernant les conditions aux bords, on considère des conditions à flux nul qui assure la conservation du nombre total de cellules $m = \int_0^\infty g(t,x)\mathrm{d}x$. Cela donne :

$$\Big(v(x,L(t))g(t,x) - \frac{\varepsilon}{2}\partial_x(d(x,L(t))g(t,x))\Big)\Big|_{x=0} = 0. \tag{2.15}$$

Cela nous amène à notre deuxième modèle de dynamique de la taille des cellules adipeuses :

$$\begin{cases} \partial_t g(t,x) + \partial_x(v(x,L(t))g(t,x)) = \dfrac{\varepsilon}{2}\partial_x^2(d(x,L(t))g(t,x)),\ t\geq 0,\ x\geq 0, & \text{(2.16a)}\\ L(t) + \displaystyle\int_0^{+\infty} xg(t,x)\mathrm{d}x = \lambda,\ t\geq 0, & \text{(2.16b)}\\ \Big(-v(\cdot,L(t))g(t,\cdot) + \dfrac{\varepsilon}{2}\partial_x(d(\cdot,L(t))g(t,\cdot))\Big)\Big|_{x=0} = 0,\ t\geq 0, & \text{(2.16c)}\\ g(0,x) = g^0(x) \text{ and } L(0) = L^0,\ x\geq 0. & \text{(2.16d)} \end{cases}$$

Tous les modèles que nous avons introduits dans cette section et ceux de la section suivante sont récapitulés dans la figure 1.13.

2.3.4 Modèle à diffusion constante

Dans le Chapitre 5, nous présentons un travail où nous effectuons une estimation de paramètres sur des données provenant de rats. Nous considérons un modèle différent pour ce travail, où nous utilisons un taux de diffusion constant $D > 0$ et les tailles des cellules sont des rayons au lieu de quantités de lipides :

$$\begin{cases} \partial_t f(t,r) + \partial_r(v(r,L(t))f(t,r)) - D\partial_r^2 f(t,r) = 0, & \text{(2.17)}\\ L(t) = \lambda - \displaystyle\int_{r_{min}}^{r_{max}} (V(r)-V_{em})\frac{4\pi r^2}{V_\ell^2} f(t,r)dr, & \text{(2.18)}\\ v(r_{min},L(t))f(t,r_{min}) - D\partial_r f(t,r_{min}) = 0, & \text{(2.19)}\\ v(r_{max},L(t))f(t,r_{max}) - D\partial_r f(t,r_{max}) = 0, & \text{(2.20)} \end{cases}$$

où v est défini par

$$v(r,L) = \frac{V_\ell}{4\pi}\left(\alpha\frac{L}{L+\kappa}\frac{\rho^3}{\rho^3+r^3} - \frac{(\beta+\gamma r^2)}{r^2}\frac{V(r)-V_{em}}{V(r)-V_{em}+V_\ell\chi}\right). \tag{2.21}$$

Ce modèle est analogue à (2.13) par la relation (2.4).

Dans la dernière section du Chapitre 5, nous utilisons les mêmes outils pour effectuer l'estimation des paramètres sur le modèle diffusif Lifshitz-Slyozov ainsi que sa version à diffusion constante donnée par :

$$\begin{cases} \partial_t g + \partial_x(vg) = D\partial_x^2(g), & \text{(2.22a)}\\ L(t) + \displaystyle\int_{\mathbb{R}_+} xg(t,x)\mathrm{d}x = \lambda, & \text{(2.22b)}\\ \Big(-vg + D\dfrac{\varepsilon}{2}\partial_x(g)\Big)\Big|_{x=0} = 0, & \text{(2.22c)}\\ g(0,x) = g^0(x) \text{ and } L(0) = L^0. & \text{(2.22d)} \end{cases}$$

Les solutions stationnaires sont alors données par :

$$M_L(x) = \frac{C}{D} \exp\left(\frac{1}{D}\int_0^x v(y,L)\mathrm{d}y\right).$$

2.4 Modèles stochastiques pour les adipocytes

2.4.1 Construction d'une chaîne de Markov en temps continue à partir du modèle de Becker-Döring

Considérons N cellules adipeuses dont la taille est décrite par une famille de variables aléatoires $X^N(t) = (X^{i,N}(t))_{i=1,\dots,N} \in \mathbb{N}^N$ qui évoluent dans le temps. Notre intention est d'utiliser la propagation uniforme du chaos et de regarder la limite de $X^{1,N}$ lorsque N tend vers $+\infty$.

Nous supposons que $(X^N(t))_{t\geq 0}$ est une chaîne de Markov en temps continue (CTMC). La propriété de Markov est alors : l'évolution de la taille d'une cellule ne dépend que de sa taille actuelle et de la taille actuelle de toutes les autres cellules et non de ce qui s'est passé dans le passé. La dynamique de L est fournie en utilisant la moyenne empirique :

$$M^{X,N}(t) = \frac{1}{N}\sum_{i=1}^{N} X^{i,N}(t).$$

Puis nous definissons :

$$L^{X,N}(t) = \lambda - M^{X,N}(t).$$

Soit $i \in \mathbb{N}$ et supposons que la k-ième cellule est dans l'état i au temps t. Alors $X^{k,N}$ saute selon

$$\begin{cases} i \to i+1 \text{ à taux } a(i)\dfrac{L^{X,N}(t)}{L^{X,N}(t)+\kappa},\ i \geq 0 \\ i \to i-1 \text{ à taux } b(i),\ i \geq 1 \end{cases}$$

Cela signifie que, étant donné $t, h > 0$, ce qui suit est vrai :

$$\mathbb{P}(X^{k,N}(t+h) - X^{k,N}(t) = 1 | X^{k,N}(t)) = a(X^{k,N}(t))\frac{L^{X,N}(t)}{L^{X,N}(t)+\kappa}h + o(h),$$

$$\mathbb{P}(X^{k,N}(t+h) - X^{k,N}(t) = -1 | X^{k,N}(t)) = b(X^{k,N}(t))h + o(h).$$

Ainsi, pour décrire $X^{k,N}$, il suffit de considérer deux processus de comptage indépendants R^k_+ et R^k_-, où $R^k_+(t)$ compte le nombre de sauts positifs jusqu'au temps t et $R^k_-(t)$ compte le nombre de sauts négatifs. Par conséquent, l'intensité de chaque processus de comptage est donnée par :

$$\mathbb{P}(R^k_+(t+h) - R^k_+(t) = 1 | X(t)) = a(X^{k,N}(t))\frac{L^{X,N}(t)}{L^{X,N}(t)+\kappa}h + o(h),$$

$$\mathbb{P}(R^k_-(t+h) - R^k_-(t) = 1 | X(t)) = b(X^{k,N}(t))h + o(h).$$

Nous écrivons :

$$X^{k,N}(t) = X^{k,N}(0) + R^k_+(t) - R^k_-(t).$$

En particulier, nous pouvons écrire

$$R_+^k(t) = Y_+^k\Big(\int_0^t a(X^{k,N}(s))\frac{L^{X,N}(s)}{L^{X,N}(s)+\kappa}\mathrm{d}s\Big) \text{ and } R_-^k(t) = Y_-^k\Big(\int_0^t b(X^{k,N}(s))\mathrm{d}s\Big)$$

pour certains processus unitaires de comptage de Poisson Y_+ et Y_-, indépendants l'un de l'autre. Observez que le lemme 1.3.1 implique que R_+ et R_- sont des processus de Poisson inhomogènes non linéaires de taux $a(X^{k,N}(t))\frac{L^{X,N}(t)}{L^{X,N}(t)+\kappa}$ et $b(X^{k,N}(t))$. Cela donne :

$$X^{k,N}(t) = X(0) + Y_+^k\Big(\int_0^t a(X^{k,N}(s))\frac{L^{X,N}(s)}{L^{X,N}(s)+\kappa}\mathrm{d}s\Big) - Y_-^k\Big(\int_0^t b(X^{k,N}(s))\mathrm{d}s\Big). \tag{2.23}$$

Inspiré par la propagation uniforme du chaos, voir théorème 1.2 dans [122], on définit la CTMC non linéaire de Becker-Döring $X(t)$ par :

$$X(t) = X(0) + Y_+\Big(\int_0^t a(X(s))\frac{L^{X}(s)}{L^{X}(s)+\kappa}\mathrm{d}s\Big) - Y_-\Big(\int_0^t b(X(s))\mathrm{d}s\Big), \tag{2.24}$$

où $L^X(t) = \lambda - \mathbb{E}[X(t)]$. Dans un certain sens, la CTMC non linéaire de Becker-Döring $X(t)$ décrit l'évolution d'une cellule ou d'un polymère stochastique typique à partir du modèle de Becker-Döring.

Rescaling

Pour être cohérent avec le modèle de Becker-Döring redimensionné (1.26), nous redimensionnons également la CTMC non linéaire de Becker-Döring. Nous considérons que la taille de la cellule $X(t)$ fait des sauts de taille ε suivant :

$$\begin{cases} i\varepsilon \to (i+1)\varepsilon \text{ at rate } \varepsilon^{-1}a(i\varepsilon)\dfrac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa},\ i\geq 0, \\ i\varepsilon \to (i-1)\varepsilon \text{ at rate } \varepsilon^{-1}b(i\varepsilon),\ i\geq 1, \end{cases}$$

où $L_\varepsilon^X(t) = \lambda - \mathbb{E}[X_\varepsilon(t)]$. Cette CTMC non linéaire de Becker-Döring remis à l'échelle vérifie :

$$X_\varepsilon(t) = X_\varepsilon(0) + \varepsilon Y_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s))\frac{L_\varepsilon^X(s)}{L_\varepsilon^X(s)+\kappa}\mathrm{d}s\Big) - \varepsilon Y_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big). \tag{2.25}$$

Dorénavant, nous supposons qu'il existe une solution à l'équation (2.25).

2.4.2 Modèle stochastique de Lifshitz-Slyozov

Considérons à nouveau N cellules adipeuses dont la taille est décrite par une famille de variables aléatoires $Z_\varepsilon^N(t) = (Z_\varepsilon^{k,N}(t))_{k=1,\dots,N} \in \mathbb{N}^N$ qui évoluent dans le temps. Notre intention est d'utiliser la propagation uniforme du chaos et de regarder la limite de $Z^{k,N}$ lorsque N tend vers ∞.

Nous définissons la moyenne empirique :

$$M^{Z,N}(t) = \frac{1}{N}\sum_{k=1}^{N} Z_\varepsilon^{k,N}(t)$$

Considérons la k-ième cellule. L'évolution de sa taille est donnée par :

$$\mathrm{d}Z_\varepsilon^{k,N}(t) = v(Z_\varepsilon^{k,N}(t), L_\varepsilon^Z(t))\mathrm{d}t + \sqrt{\varepsilon a(Z_\varepsilon^{k,N}(t))\frac{L_\varepsilon^{Z,N}(t)}{L_\varepsilon^{Z,N}(t)+\kappa}}\mathrm{d}W_t^{+,k} + \sqrt{\varepsilon b(Z_\varepsilon^{k,N}(t))}\mathrm{d}W_t^{-,k},\ t \geq 0, \tag{2.26}$$

où $L^{Z,N}(t) = \lambda - M^{Z,N}(t)$ et $W_t^{\pm,k}$ sont des processus de Wiener standard indépendants les uns des autres. Cette définition est inspirée du fait que l'équation (2.13a) est une équation de Fokker-Planck non linéaire dont l'équation différentielle stochastique est sous la forme (2.26). Nous avons scindé la partie stochastique en deux termes dépendant de W^+ et W^- pour simplifier la comparaison avec la CTMC non linéaire de Becker-Döring. Encore une fois inspirés par la propagation du chaos, nous définissons le modèle stochastique de Lifshitz-Slyozov par :

$$\mathrm{d}Z_\varepsilon(t) = v(Z_\varepsilon(t), L_\varepsilon^Z(t))\mathrm{d}t + \sqrt{\varepsilon a(Z_\varepsilon(t))\frac{L_\varepsilon^Z(t)}{L_\varepsilon^Z(t)+\kappa}}\mathrm{d}W_t^+ + \sqrt{\varepsilon b(Z_\varepsilon(t))}\mathrm{d}W_t^-,\ t \geq 0, \tag{2.27}$$

où $L_\varepsilon^Z(t) = \lambda - \mathbb{E}\left[Z_\varepsilon(t)\right]$. Sous certaines hypothèses techniques, une solution à l'équation (2.27) vérifie également :

$$Z_\varepsilon(t) = Z_\varepsilon(0) + \varepsilon B_+\Big(\varepsilon^{-1}\int_0^t a(Z_\varepsilon(s))\frac{L_\varepsilon^Z(s)}{L_\varepsilon^Z(s)+\kappa}\mathrm{d}s\Big) - \varepsilon B_-\Big(\varepsilon^{-1}\int_0^t b(Z_\varepsilon(s))\mathrm{d}s\Big), \tag{2.28}$$

où $B_\pm(t) = t + W^\pm(t)$ sont des processus de Wiener dérivés et W^- et W^+ sont des processus de Wiener indépendants. C'est le modèle stochastique Lifshitz-Slyozov. L'équation (2.28) est particulièrement utile pour comparer Z_ε et X_ε. Comme pour l'Équation (2.25), nous supposons désormais qu'une solution à l'Équation (2.28) existe.

Remark. *Les résultats théoriques concernent X_ε, Z_ε et les équations* (2.25) *et* (2.28). *Cependant, nous ne sommes pas en mesure de les simuler directement. Nous utilisons donc les équations* (2.23) *et* (2.26) *pour les simulations numériques.*

2.4.3 Récapitulatif des modèles

Une représentation visuelle de tous les modèles est fournie dans la Figure 1.13. Nous sommes partis du modèle de Lifshitz-Slyozov classique pour construire un modèle de Lifshitz-Slyozov pour les cellules adipeuses. A partir de la relation entre le modèle de Becker-Döring classique et le modèle de Lifshitz-Slyozov classique, nous avons introduit un modèle de Becker-Döring pour les cellules adipeuses qui conserve la propriété de convergence des modèles classiques. À partir de ce résultat de convergence, nous avons développé un modèle de Lifshitz-Slyozov diffusif qui s'adapte plus précisément aux données. Nous étudions également la pertinence d'un modèle diffusif similaire à diffusion constante. À partir du modèle de Becker-Döring et du modèle diffusif Lifshitz-Slyozov pour les cellules adipeuses, nous avons construit deux modèles stochastiques pour étudier la convergence d'ordre deux.

2.5 Résultats principaux

2.5.1 Nouveaux modèles pour la distribution en taille des adipocytes

Nous avons introduit une variété de modèles pour la distribution en taille des cellules adipeuses. Le modèle principal est le modèle de Lifshitz-Slyozov diffusif (2.16) où le terme de diffusion

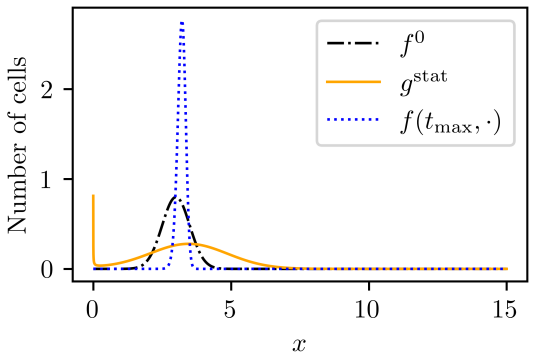


Figure 2.3: Comparaison entre le modèle diffusif Lifshitz-Slyozov et le modèle de Lifshitz-Slyozov sans diffusion.

est dérivé du modèle de Becker-Döring discret (1.26) que nous étudions au chapitre 3. Nous nous intéressons particulièrement aux solutions stationnaires du modèle diffusif Lifshitz-Slyozov qui ont une formule explicite (1.30a) et peuvent donc être calculées exactement. Nous avons comparé les solutions stationnaires aux solutions stationnaires des autres modèles lorsqu'elles sont calculables. Ce n'est pas le cas pour le modèle de Lifshitz-Slyozov classique, mais on peut quand même les étudier numériquement. Dans la Figure 2.3, nous montrons qu'en général les deux solutions stationnaires sont difficiles à comparer. Pour le cas de la modélisation adipeuse, nous observons que la vitesse a deux zéros attractifs, nous nous attendons donc à ce que la distribution stationnaire se divise entre les deux zéros. Cependant, il est difficile de prédire comment les solutions se répartissent entre les deux zéros. Illustrons cela par quelques considérations générales.

Considérons l'EDP de transport avec une vitesse v dépendant de x et t qui a trois zéros pour tout temps $t \geq 0$ que nous notons $z_1(t) < z_2(t) < z_3(t)$. Nous supposons que ces zéros ont des limites lorsque $t \to \infty$ et que, pour tous les temps, z_1 et z_3 sont attractifs et z_2 doit être répulsif tel que $\partial_z v(z_{1,3}(t), t) < 0$ et $\partial_z v(z_2(t), t) > 0$. Nous nous attendons à ce que la distribution se concentre autour des points $z_{1,3}^\infty = \lim_{t\to\infty} z_{1,3}(t)$. Dans cette hypothèse, aucune observation ne peut être faite sur les relations entre les courbes caractéristiques et $z_2(t)$. En particulier, nous ne nous attendons pas à ce que si $X(s; s, x) = x > z_2(s)$ alors $X(t; s, x) > z_2(t)$ pour tout $t > s$. Cela signifie que $z_2(t)$ ne divise pas la famille ordonnée de caractéristiques en deux sous-familles : celles allant à z_1^∞ et celles allant à z_3^∞. Par conséquent, les caractéristiques peuvent franchir la 'frontière' $z_2(t)$ à tout moment, mais à mesure que le temps tend vers l'infini, cette 'frontière' devient plus difficile à franchir car si z_2 ne dépend pas du temps, ce serait une frontière appropriée pour les caractéristiques. Or c'est ce qui se passe dans le cas du modèle de Lifshitz-Slyozov pour les adipocytes, la vitesse peut avoir jusqu'à trois zéros qui dépendent du temps et le zéro répulsif ne sépare pas la famille des caractéristiques. Décrire le comportement de ces zéros dans le cas de notre modèle est malheureusement assez difficile. Même si nous sommes capables d'obtenir une approximation numérique des solutions stationnaires, la comparer à la solution stationnaire du modèle diffusif Lifshitz-Slyozov est difficile. En effet pour le modèle diffusif, les maxima des solutions stationnaires sont des zéros de $\lim_{t\to\infty} v(\cdot, L(t)) - \varepsilon\partial_x d(\cdot, L(t))$. Ainsi les modes du modèle sans diffusion et du modèle diffusif ne se situent pas aux mêmes valeurs de x et de plus l'ajout de la diffusion dans le modèle accentue le fait que les cellules peuvent franchir la 'frontière' $z_2(t)$.

Dans Chapitre 4, nous étudions la CTMC de Becker-Döring (2.25) non-linéaire et son ap-

proximation de diffusion, la SDE de Lifshitz-Slyozov (2.28). Les deux modèles sont construits à partir de leur équivalent déterministe et les deux peuvent présenter des distributions stationnaires bimodales.

Dans le Chapitre 5, nous effectuons une analyse de sensibilité et une estimation des paramètres sur le modèle de Lifshitz-Slyozov diffusif constant et non constant, montrant que dans les deux cas nous récupérons des paramètres d'ordre de grandeur similaire.

2.5.2 Résultat de convergence

Nous avons précédemment introduit le résultat de convergence classique du modèle de Becker-Döring vers le modèle de Lifshitz-Slyozov dans le théorème 1.4.4. Dans cette section, nous introduisons un nouveau résultat dans cette théorie de la convergence. Nous utilisons des **queues de distributions** pour comparer les solutions du modèle de Becker-Döring et les solutions 'mild' du modèle de Lifshitz-Slyozov. Cette nouvelle méthode de preuve est différente de celle utilisée dans la convergence classique. En particulier, l'utilisation de queues de distribution nous permet de réduire la non-linéarité de l'EDP de transport en tirant la vitesse hors de la dérivée spatiale. Précisons que l'étude des queues dans le cas du modèle de Lifshitz-Slyozov n'est pas nouvelle puisqu'elle a été utilisée dans [22] pour obtenir le principe de quasi comparaison et in [77, 19] pour affiner résultats d'unicité. Cependant son utilisation dans la théorie de la convergence de Becker-Döring vers Lifshitz-Slyozov est nouvelle, à notre connaissance. Passons maintenant au résultat.

Soit $(f^\varepsilon, L^\varepsilon)$ la solution du système Becker-Döring ODE (2.14), et soit (f, L) la solution 'mild' du modèle de Lifshitz-Slyozov (2.13). Nous introduisons des notations pour les queues de distributions :

$$F(t,x) = \int_x^\infty f(t,y)\mathrm{d}y, \qquad F^\varepsilon(t,x) = \int_x^\infty f^\varepsilon(t,y)\mathrm{d}y\,, \tag{2.29}$$

et nous introduisons leur différence :

$$E(t,x) = F^\varepsilon(t,x) - F(t,x)\,. \tag{2.30}$$

Nous énonçons maintenant notre théorème principal :

Theorem 2.5.1. *Soit $T > 0$. Supposons qu'il existe une constante $C_{\text{init}} > 0$ telle que pour tout $\varepsilon > 0$, $\int_{\mathbb{R}_+} |E(0,x)|\mathrm{d}x \le \varepsilon C_{\text{init}}$. Supposons également que les hypothèses* (H1)-(H9) *soient vraies. Alors il existe une constante $C(T) > 0$ telle que pour $\varepsilon > 0$ assez petit et pour tout $t \in [0,T]$:*

$$|L^\varepsilon(t) - L(t)| + \int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x \le \varepsilon C(T).$$

Nous donnons quelques détails sur le fonctionnement de la preuve. Nous commençons par une observation simple, tirée de [77, 19], qui stipule que la différence entre les quantités de lipides externes peut être contrôlée par la différence des queues, $|L^\varepsilon(t) - L(t)| \le \int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. Pour avoir des limites appropriées sur la queue, nous utilisons le lemme de Grönwall. Ceci est fait en dérivant l'équation pour F^ε et F (Lemma 3.4.5). Pour ce faire, nous devons considérer à part le cas $x < \varepsilon/2$ car il donne des termes aux bords qui doivent être traités différemment. Cela permet à son tour une première borne sur l'intégrale $\int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. Ensuite à partir de l'équation sur F on obtient l'équation sur E (Lemme 3.4.6). On procède en bornant les termes dans l'estimation de $\int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. En particulier, nous pouvons montrer que F^ε résout une équation similaire à

F à quelques termes d'ordre ε près. Cet argument repose sur un contrôle uniforme de la solution du modèle de Becker-Döring redimensionné et de ses incréments $c_{i+1}^{\varepsilon} - c_i^{\varepsilon}$.

Les hypothèses (H1) - (H4) sont classiques dans l'étude de notre modèle. Cependant, d'autres hypothèses sont moins courantes mais découlent naturellement du résultat. Comme nous travaillons avec des solutions 'mild', nous avons besoin d'une régularité supplémentaire sur les taux de lipogénèse et lipolyse. Cela est dû au fait que les termes d'ordre supérieur qui apparaissent, qui sont ceux que nous voulons regarder pour choisir une diffusion, sont des dérivées du second ordre. Cela implique l'hypothèse (H5). Naturellement, nous exigeons que la condition initiale soit bien préparée. En particulier, certaines propriétés sont propagées, d'où les hypothèses (H6) et (H7). De plus, l'hypothèse (H8) nous permet d'avoir la positivité stricte de L^{ε} en temps fini, uniformément en ε. Considérant que si ce n'était pas le cas, le terme diffusif du second ordre serait mal posé, cette hypothèse est naturelle. L'hypothèse (H9) sur la condition initiale $c_0^{\varepsilon,0}$ est technique et garantit que la bonne condition aux bords (2.13c) est satisfaite à tout moment.

Enfin l'hypothèse sur $\int_{\mathbb{R}_+} |E(0,x)| \mathrm{d}x$ est faite pour conclure après avoir utilisé le lemme de Grönwall à la toute fin de la preuve. Cette hypothèse relie les conditions $(c_i^{\varepsilon,0})_{i\geq 0}$ et f^0. Un choix logique pour la condition initiale $(c_i^{\varepsilon,0})_{i\geq 0}$ est $c_i^{\varepsilon,0} = f^0(i\varepsilon)$ pour tout $i \geq 0$. Afin de vérifier l'hypothèse sur la condition initiale, nous supposons que f^0 diminue exponentiellement vers 0 à l'infini et $(f^0)' \in L^1(\mathbb{R}_+, x\mathrm{d}x)$. Cela donne :

$$\begin{aligned}\int_{\mathbb{R}_+} |E(0,x)| \mathrm{d}x &= \int_{\mathbb{R}_+} |\int_x^{\infty} f^{0,\varepsilon}(y) - f^0(y) \mathrm{d}y| \mathrm{d}x \\ &\leq \int_{\mathbb{R}_+} x |f^{0,\varepsilon}(x) - f^0(x)| \mathrm{d}x \\ &\leq \sum_{i\geq 1} \int_{\Gamma_i^{\varepsilon}} x |f^0(i\varepsilon) - f^0(x)| \mathrm{d}x + \int_0^{\varepsilon/2} x |f^0(0) - f^0(x)| \mathrm{d}x.\end{aligned} \tag{2.31}$$

Maintenant, l'hypothèse que f^0 diminue exponentiellement vers 0 à l'infini nous permet de couper l'intégrale précédente en deux parties. Il existe donc une constante $\mu > 0$ telle que

$$\int_x^{\infty} y f^0(y) \mathrm{d}y \leq \int_x^{\infty} y \exp(-\mu y) \mathrm{d}y = \frac{1}{\mu}(x + \frac{1}{\mu}) \exp(-\mu x). \tag{2.32}$$

Alors prendre $x = -\ln(\varepsilon)$ nous donne qu'il existe une constante C telle que $\int_{-\ln(\varepsilon)}^{\infty} y f^0(y) \mathrm{d}y \leq \varepsilon C$. Nous savons maintenant qu'il existe des $i \in \mathbb{N}$ tels que $-\ln(\varepsilon) \in \Gamma_i^{\varepsilon}$ que nous notons I_{ε}. En particulier, $I_{\varepsilon} = \lfloor \frac{-\ln(\varepsilon)}{\varepsilon} \rceil$. Cela nous permet de diviser la somme (2.31). Pour $i \leq I_{\varepsilon} + 1$ on utilise le développement de Taylor et le fait que $(f^0)' \in L^1(\mathbb{R}_+, x\mathrm{d}x)$ alors que les termes au-dessus de $I_{\varepsilon} + 1$ sont bornés en utilisant le choix de I_{ε} :

$$\begin{aligned}\int_{\mathbb{R}_+}|E(0,x)|\mathrm{d}x &= \sum_{i=1}^{I_\varepsilon+1}\int_{\Gamma_i^\varepsilon} x|f^0(i\varepsilon)-f^0(x)|\mathrm{d}x + \int_0^{\varepsilon/2} x|f^0(0)-f^0(x)|\mathrm{d}x\\ &\quad+\sum_{i\geq I_\varepsilon+2}\int_{\Gamma_i^\varepsilon} x|f^0(i\varepsilon)-f^0(x)|\mathrm{d}x\\ &\leq \sum_{i=1}^{I_\varepsilon+1}\int_{\Gamma_i^\varepsilon} x\frac{\varepsilon}{2}|(f^0)'(x)|\mathrm{d}x + \frac{\varepsilon}{2}\int_0^{\varepsilon/2} x|(f^0)'(x)|\mathrm{d}x\\ &\quad+o(\varepsilon^2)(\sum_{i=1}^{I_\varepsilon+1}\int_{\Gamma_i^\varepsilon} x\mathrm{d}x + \int_0^{\varepsilon/2} x\mathrm{d}x) + 2C\varepsilon\\ &\leq \frac{\varepsilon}{2}\|(f^0)'\|_{L^1(\mathbb{R}_+,x\mathrm{d}x)} + o(\varepsilon^2\ln(\varepsilon)) + 2\varepsilon C.\end{aligned}$$

Donc si on construit $c_i^{\varepsilon,0} = f^0(i\varepsilon)$, les conditions suffisantes sur f^0 pour (H8) sont : f^0 diminue exponentiellement vers 0 à l'infini et $(f^0)' \in L^1(\mathbb{R}_+, x\mathrm{d}x)$.

Enfin, nous soulignons que notre résultat utilise le fait qu'il existe une solution aux deux modèles et prouve la convergence de l'un vers l'autre, alors que les résultats classiques n'utilisent que l'existence de solutions à BD et montrent la convergence vers une fonction à valeurs mesure qui est une solution faible de Lifshitz-Slyozov. Donc le résultat classique est aussi un résultat d'existence pour les solutions faibles de Lifshitz-Slyozov.

2.5.3 Convergence stochastique

Notre motivation pour étudier le modèle de Becker-Döring et construire un modèle similaire pour les cellules adipeuses est d'avoir une intuition sur la forme que pourrait prendre un modèle diffusif Lifshitz-Slyozov. De plus, nous nous intéressons particulièrement à montrer certaines bornes sur la distance entre les solutions du modèle de Becker-Döring et les solutions du modèle diffusif de Lifshitz-Slyozov (2.16). Cela s'est avéré particulièrement difficile en utilisant les idées de la convergence classique et le résultat de convergence que nous avons présenté précédemment était un premier pas dans une autre direction mais n'a pas encore été fructueux. Néanmoins il existe aussi des résultats sur la convergence de Becker-Döring vers Lifshitz-Slyozov mais dans un sens probabiliste. Dans l'esprit de ces résultats, nous nous intéressons à deux modèles stochastiques de la dynamique de la taille d'une cellule adipeuse, basés sur les modèles Becker-Döring et Lifshitz-Slyozov que nous avons introduits dans la section 2.4. Ceci conduit aux résultats suivants, inspirés du résultat classique de Kurtz in [74]. Nous faisons l'hypothèse supplémentaire :

a et b sont des fonctions continues bornées. (H10)

Nous énonçons notre résultat :

Theorem 2.5.2. *Supposons que X_ε et Z_ε soient les solutions des équations* (2.25) *et* (2.28) *et que les hypothèses* (H1), (H3) *et* (H10) *soient vérifiées. Alors pour ε assez petit et pour $T > 0$, il existe une constante β^T dépendant uniquement de T, a et b telle que :*

$$\sup_{t\in[0,T]}\mathbb{E}\left[|X_\varepsilon(t)-Z_\varepsilon(t)|\right] \leq \beta^T\varepsilon\ln(\frac{1}{\varepsilon}). \tag{2.33}$$

Nous faisons quelques commentaires sur ce résultat. Premièrement, le résultat initial dans [74] implique des particules dont les taux de saut dépendent de la position de la particule elle-même.

Dans notre cas, c'est aussi vrai mais en plus, à travers L^X et L^Z, cela dépend aussi de son espérance et donc de sa loi. C'est la principale difficulté de la preuve. En particulier, cela nous empêche d'obtenir une convergence presque sûre comme dans [74] et nous obtenons à la place une convergence L^1.

Nous donnons quelques détails sur le fonctionnement de la preuve. Le schéma général est très similaire à [74]. Nous utilisons d'abord un résultat de [72], où nous bornons la différence entre un processus de Poisson compensé et un processus de Wiener et le module de continuité de Levy du processus de Wiener [81]. Le résultat final est obtenu par une utilisation appropriée de l'inégalité de Jensen et du lemme de Grönwall.

Ensuite, nous assouplissons l'hypothèse (H10) et considérons les taux sous-linéaires :

Il existe deux constantes positives C_a et C_b telles que

$$\text{pour tout } x \geq 0,\ a(x) \leq C_a(1+x) \text{ et } b(x) \leq C_b(1+x). \quad \text{(H11)}$$

Ensuite, nous montrons le résultat suivant :

Theorem 2.5.3. *Supposons que X_ε et Z_ε soient des solutions des équations* (2.25) *et* (2.28) *et des hypothèses* (H1), (H3) *et* (H11) *pour tenir. Alors pour ε assez petit et pour certains $T > 0$, il existe une constante β^T dépendant uniquement de T, a et b telle que :*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|X_\varepsilon(t) - Z_\varepsilon(t)|\right] \leq \beta^T \varepsilon \ln(\frac{1}{\varepsilon}). \tag{2.34}$$

La preuve de ce résultat est très similaire à celle du théorème 2.5.2. La principale différence est de diviser l'étude de la norme L^1 en fonction d'un temps d'atteinte d'une grande taille x : $\tau_x = \inf\{t \geq 0 | X_\varepsilon(t) > x \text{ ou } Z_\varepsilon(t) > x\}$. Le cas $t < \tau_x$ est traité de la même manière que le premier résultat, puisque dans ce cas les taux peuvent être bornés en utilisant l'hypothèse (H11). Pour le cas $t \geq \tau_x$, nous utilisons des bornes brutes qui sont ensuite raffinées par un choix approprié de la taille x qui dépend de ε.

2.5.4 Résultats numériques

En plus des résultats théoriques précédents, nous utilisons les schémas numériques (1.62) et (1.68) décrits précédemment pour étudier les propriétés de nos modèles. Dans un premier temps on retrouve numériquement la convergence vers un état stationnaire. Dans un deuxième temps, nous montrons que notre modèle peut récupérer la propriété de bimodalité des distributions des cellules adipeuses. Cependant cette propriété est fortement dépendante des paramètres et nous montrons qu'en fonction du paramètre λ il existe tout un continuum de solutions stationnaires qui présentent différents types de modalité. Pour de faibles quantités de lipides totaux, seules de petites cellules sont présentes et la distribution a un maximum global à $x = 0$. Pour de grandes quantités de lipides externes, la solution stationnaire est composée uniquement de grandes cellules et ressemble à une distribution gaussienne centrée sur une grande valeur de x. Entre les deux, nous récupérons la propriété bimodale. Les valeurs de lambda prises pour obtenir ces comportements vont de 0 à 15, mais les valeurs associées de L^{stat} ne vont que de 0 à $0,1$. Ceci doit être cohérent avec le fait que chez un individu, la quantité de lipides sanguins doit rester à un niveau constant non toxique et qu'un apport excessif de lipides est stocké dans le tissu adipeux. Un exemple du résultat du schéma est donné sur la Figure 2.4 et les différents comportements des solutions stationnaires sont illustrés sur la Figure 2.5.

Nous nous intéressons également à comparer numériquement le modèle de Lifshitz-Slyozov usuel au modèle diffusif Lifshitz-Slyozov. Cependant, cela s'avère difficile pour plusieurs raisons.

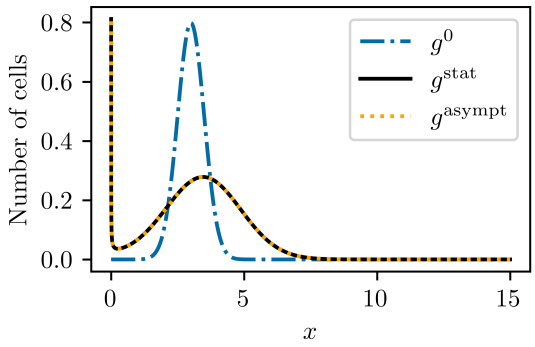


Figure 2.4: **Résultat du schéma numérique :** une distribution bimodale obtenue par le schéma et calculée explicitement. Le résultat du schéma g^{asympt} et la solution stationnaire calculée explicitement g^{stat} sont superposés.

Premièrement, nous observons que pour des paramètres similaires, la solution stationnaire du modèle diffusif ne correspond pas à l'approximation asymptotique d'un schéma UpWind pour le modèle de Lifshitz-Slyozov usuel. Nous sommes malheureusement incapables de calculer explicitement des solutions stationnaires au modèle de Lifshitz-Slyozov avec nos choix de taux et nous nous appuyons donc sur l'approximation numérique asymptotique pour faire notre comparaison. Nous observons que le modèle diffusif peut présenter une bimodalité alors que le modèle habituel montre une unimodalité pour des paramètres similaires. Cela peut s'expliquer en regardant les positions des zéros de la vitesse v. En effet en regardant les courbes caractéristiques (1.53), on peut voir qu'elles atteignent un état stationnaire ssi $v(X(s;t,x),L(s))=0$. Par conséquent, pour le modèle de Lifshitz-Slyozov , nous nous attendons à ce que la distribution se concentre sur les zéros de la vitesse. Le cas où la vitesse a un zéro a été étudié dans [18] où ils montrent une concentration vers une masse de Dirac située au zéro de la vitesse. Le cas des zéros multiples est à notre connaissance une question ouverte pour le modèle de Lifshitz-Slyozov. Pour le cas de la modélisation adipeuse, nous observons que la vitesse a deux zéros attractifs, nous nous attendons donc à ce que la distribution se divise entre les deux zéros dans un temps long.

En ce qui concerne les simulations stochastiques, nous montrons que les modèles stochastiques sont tous deux capables de récupérer les différents types de distributions stationnaires, voir Figure 2.6. Dans l'esprit du Théorème 2.5.3, on obtient une borne d'ordre $\varepsilon\ln(\varepsilon)$ pour la norme empirique L^1 entre les deux solutions des Équations (1.71) et (2.26), illustrés dans la Figure 2.7. La norme empirique L^1 est calculée comme $\sup_{t\leq T}\frac{1}{N^2}\sum_{i,j=0}^{N}|X_\varepsilon^{i,N}(t)-Z_\varepsilon^{i,N}(t)|$.

2.5.5 Estimation de paramètres

Dans le Chapitre 5, nous présentons un travail soumis sur l'estimation des paramètres du modèle diffusif constant Lifshitz-Slyozov en collaboration avec Chloé Audebert, Anne-Sophie Giacobbi, Magali Ribot, Hédi Soula et Romain Yvinec, ainsi comme estimations supplémentaires pour le modèle diffusif Lifshitz-Slyozov.

Nous avons précédemment essayé des méthodes d'estimation très basiques telles que la descente de gradient et les méthodes des moindres carrés. Ces méthodes n'étaient souvent pas concluantes et nous avons cherché une méthode plus robuste capable de résoudre des problèmes

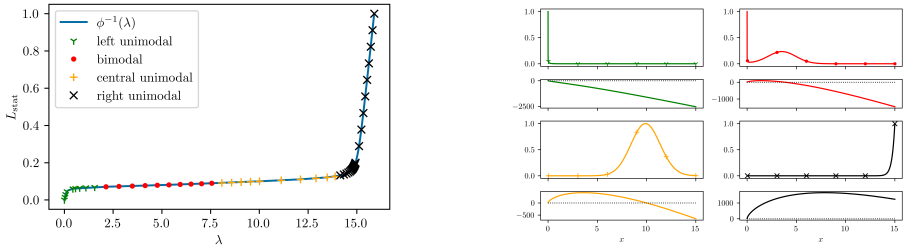


Figure 2.5: **Différents types de solutions stationnaires :** à gauche on trace l'inverse de la fonction $\phi \ : L \to L + \int_0^\infty x M_L(x)\mathrm{d}x$. L'inverse ϕ^{-1} est tracé car la valeur réelle que nous pouvons modifier dans le modèle est λ et non L^{stat}. Certains points sont mis en évidence qui correspondent à quatre types différents de solutions stationnaires : unimodal gauche (vert), bimodal (rouge), unimodal central (jaune) et unimodal droit (noir). Pour chaque couleur, une solution stationnaire est tracée à droite dans la même couleur, avec sa vitesse en dessous. On peut observer que pour chaque type de solutions stationnaires les modes sont situés près des zéros de la vitesse comme attendu, sauf pour le cas unimodal droit, où la vitesse est positive sur tout le domaine et donc la distribution est transportée vers la frontière $x_{\max}$ du domaine.

plus difficiles. Une première méthode que nous avons utilisée était le calcul bayésien approximatif (ABC). Cette méthode utilise une loi a priori pour chaque paramètre et génère une loi a posteriori en utilisant la probabilité bayésienne. Les résultats de l'ABC étaient cependant difficiles à interpréter sans une analyse appropriée de sensibilité et d'identifiabilité du modèle. Ceci est fait dans Chapter 5.

Nous avons donc choisi d'utiliser la méthode CMA-ES. Cette méthode est robuste en ce qui concerne les problèmes non convexes et non linéaires et n'a pas besoin de calculer de gradients. Comme nous manquons de résultats théoriques sur notre problème de minimisation, cette méthode est bien adaptée car elle peut être utilisée dans un cas très général.

Le modèle diffusif constant Lifshitz-Slyozov est écrit en rayons pour ce travail au lieu de lipides mais la forme générale du modèle reste la même :

$$\begin{cases} \partial_t f(t,r) + \partial_r(v(r,L(t))f(t,r)) = D\partial_r^2(f(t,r)), & \text{(2.35a)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} (V(r) - V_0)\frac{4\pi r^2}{V_{\text{lipids}}^2} f(t,r)\mathrm{d}r = \lambda, & \text{(2.35b)} \end{cases}$$

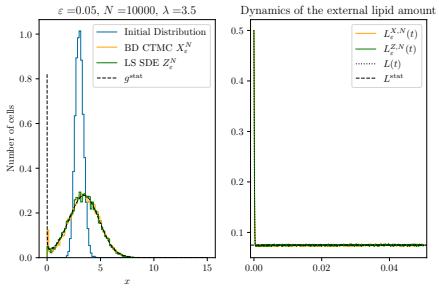
où $V(r) = \frac{4}{3}\pi r^3$ et la vitesse est définie comme $v(r,L) = a(r)\frac{L}{L+\kappa} - b(r)$, qui est similaire à la vitesse de Lifshitz-Slyozov model (2.13). Les fonctions a et r sont données par :

$$a(r) = \frac{\alpha V_{\text{lipids}}}{4\pi}\frac{\rho^3}{r^3+\rho^3}\frac{L}{L+\kappa},$$

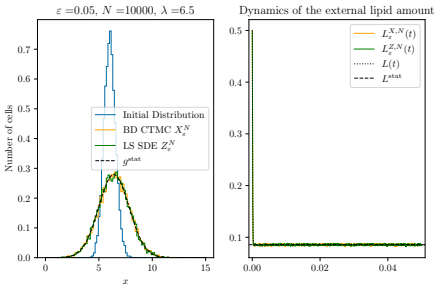
$$b(r) = \frac{\beta + \gamma r^2}{r^2}\frac{V(r) - V_0}{V(r) - V_0 + V_{\text{lipids}}\chi}.$$

Remark. *On peut passer de l'Equation* (2.13a) *à l'Equation* (2.35a) *sans le terme de diffusion, et inversement via le difféomorphisme* $x \to r(x)^2$ *à constante multiplicative près.*

Les inconnues du modèle sont $\alpha, L, \kappa, \rho, \chi$ et D. Nous effectuons d'abord une analyse d'identifiabilité des paramètres.



(a) Solution stationnaire bimodale



(b) Solution stationnaire unimodale

Figure 2.6: Deux types de modalité pour les deux modèles stochastiques avec la dynamique de L. On obtient une convergence vers la solution stationnaire correspondante du modèle diffusif de Lifshitz-Slyozov.

Le modèle (2.35) est reparamétré avec $\theta_1 = \frac{\alpha L}{L+\kappa}$. En utilisant la `Structural identifiability Toolbox` de Maple [2], nous montrons que le modèle est identifiable sous cette reparamétrisation. Ensuite, nous utilisons la méthode CMA-ES sur des données synthétiques générées à l'aide de solutions stationnaires de (2.35). Nous montrons que la méthode fonctionne bien même en censurant les petits points de données, où trois paramètres sur quatre sont bien estimés, voir Table 2.1. La censure des petits points de données est étudiée car, selon la configuration de la machine lors de la collecte des données, le rayon minimal mesuré est différent, d'où la nécessité de censurer certaines données. Ensuite, la méthode est utilisée sur des données de 32 rats, où nous montrons la capacité du modèle à récupérer la forme bimodale particulière des distributions et une estimation des paramètres identifiables, voir Table 2.2. Cependant le nadir, le minimum entre les deux modes, n'est pas parfaitement reproduit, voir Figure 2.8. Ceci est étayé par l'analyse de sensibilité, qui montre qu'aucun des paramètres n'a d'influence significative sur le nadir. Néanmoins, nous obtenons une estimation du taux de surface de lipogenèse qui est d'un ordre de grandeur similaire au taux de surface de lipolyse. De plus, nous montrons que pour des rayons supérieurs à 20 μm, la lipolyse est principalement un mécanisme basé sur la surface, comme in [119].

La dernière partie du Chapitre 5 est consacrée à l'application de la méthode CMA-ES au modèle diffusif Lifshitz-Slyozov (2.35).

Nous utilisons la même re-paramétrisation que dans la première partie, mais nous n'avons pas pu montrer que cette re-paramétrisation est identifiable pour ce modèle. Néanmoins, nous récupérons des valeurs de paramètres dans le même ordre de grandeur que dans la première partie, voir Table 2.3, avec une légère amélioration de l'ajustement du nadir en censurant les grands points de données, voir Figure 2.9.

2.5.6 Résumé des résultats principaux

Résumons brièvement les principaux résultats de ce manuscrit chapitre par chapitre. Le chapitre 6 est consacré à un travail réalisé lors de l'école d'été CEMRACS 2022, et ne concerne pas la modélisation des cellules adipeuses. Il est donc présenté indépendamment dans la Section 2.6.

synthetic data set 1			10,000 samples			samples $> 10\mu m$		
parameter	order	true	esti. value	esti. ±20%	select. values	esti. value	esti. ±20%	select. values
θ_1	10^{-3}	9.60	9.61	7.69 - 11.53	9.58 - 9.63	9.62	7.70 - 11.54	9.59 - 9.65
ρ	10^{2}	1.50	1.50	1.20 - 1.80	1.47 - 1.53	1.49	1.19 - 1.79	1.46 - 1.52
θ_3	10^{3}	2.18	2.17	1.74 - 2.60	2.05 - 2.29	2.09	1.67 - 2.51	1.91 - 2.29
θ_4	10^{-3}	7.37	7.20	5.76 - 8.64	6.54 - 8.02	7.35	5.88 - 8.82	6.58 - 8.32
synthetic data set 2			10,000 samples			samples $> 10\mu m$		
parameter	order	true	esti. value	esti. ±20%	select. values	esti. value	esti. ±20%	select. values
θ_1	10^{-3}	9.92	9.92	7.94 - 11.90	9.90 - 9.95	9.91	7.92 - 11.89	9.86 - 9.95
ρ	10^{2}	2.00	2.00	1.60 - 2.40	1.97 - 2.03	2.01	1.61 - 2.41	1.99 - 2.05
θ_3	10^{3}	3.27	3.12	2.49 - 3.74	2.69 - 3.58	5.39	4.31 - 6.47	4.32 - 6.47
θ_4	10^{-2}	1.11	1.12	0.90 - 1.34	1.05 - 1.21	1.12	0.90 - 1.34	0.98 - 1.28

Table 2.1: **Plage de valeurs sélectionnées pour les paramètres.** Les trois premières colonnes indiquent les noms des paramètres, l'ordre et la valeur vraie. Pour chaque ensemble de données, la valeur estimée du paramètre (colonne "esti. value") avec la méthode CMA-ES est sujette à une variation maximale de 20% (colonne "esti. ±20%"). À partir de cette variation, une plage de valeurs est sélectionnée pour chaque paramètre (colonne "select. values") permettant un taux d'erreur maximum de 0.1% sur la valeur de la fonction de coût estimée $\mathcal{L}$. Pour chaque paramètre, 1000 d'échantillons sont générés

parameters	mean	std	RSD
θ_1	$9.6\,10^{-3}$	$2.8\,10^{-4}$	0.03
ρ	$1.57\,10^{2}$	$0.25\,10^{2}$	0.16
θ_3	$2.24\,10^{3}$	$1.07\,10^{3}$	0.47
θ_4	$8.21\,10^{-3}$	$2.58\,10^{-3}$	0.31

Table 2.2: **Estimation des paramètres avec les distributions de taille des adipocytes mesurées chez le rat**. La première colonne contient les noms des paramètres. Sur 32 estimations avec les différentes distributions de taille des cellules animales, la moyenne est présentée dans la deuxième colonne, l'écart-type dans la troisième colonne et la quatrième colonne est l'écart-type relatif *c'est-à-dire le rapport de l'écart-type sur la moyenne. Les paramètres sont estimés avec l'algorithme CMA-ES de la fonction* `fmin2` *du package Python* `cma` *(avec 100 estimations initiales).*

parameters	mean	std	RSD
$\theta_1\beta$	$2.90\,10^{-1}$	$5.00\,10^{-3}$	0.02
ρ	$1.97\,10^{2}$	$2.97\,10^{1}$	0.15
χ	$8.45\,10^{-4}$	$1.53\,10^{-3}$	1.81
ε	$9.95\,10^{-3}$	$4.11\,10^{-3}$	0.41

Table 2.3: **Résultats sur des données censurées de grande taille.** Estimation des paramètres avec les distributions de taille des adipocytes mesurées chez le rat : moyenne, variation standard et variation standard relative. Pour comparer l'ordre de grandeur à Table 2.2, nous avons que β est d'ordre 10^2 et $\theta_3 = \chi V_{\text{lipids}}$ où V_{lipides} est de l'ordre de 10^6.

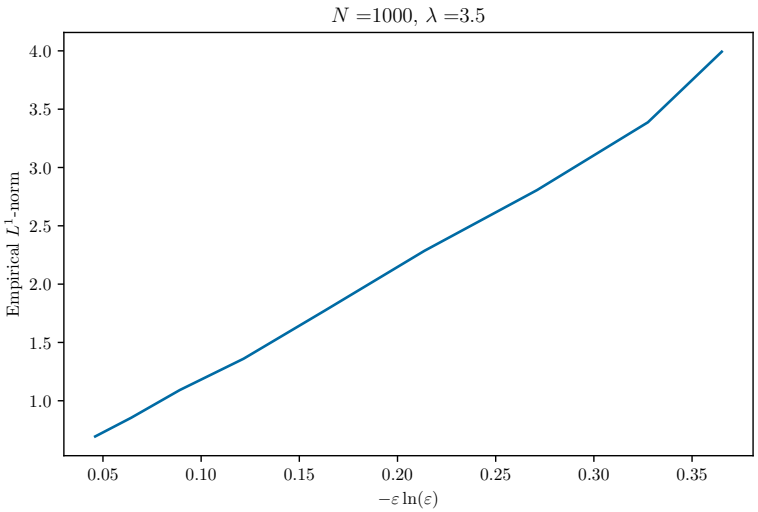


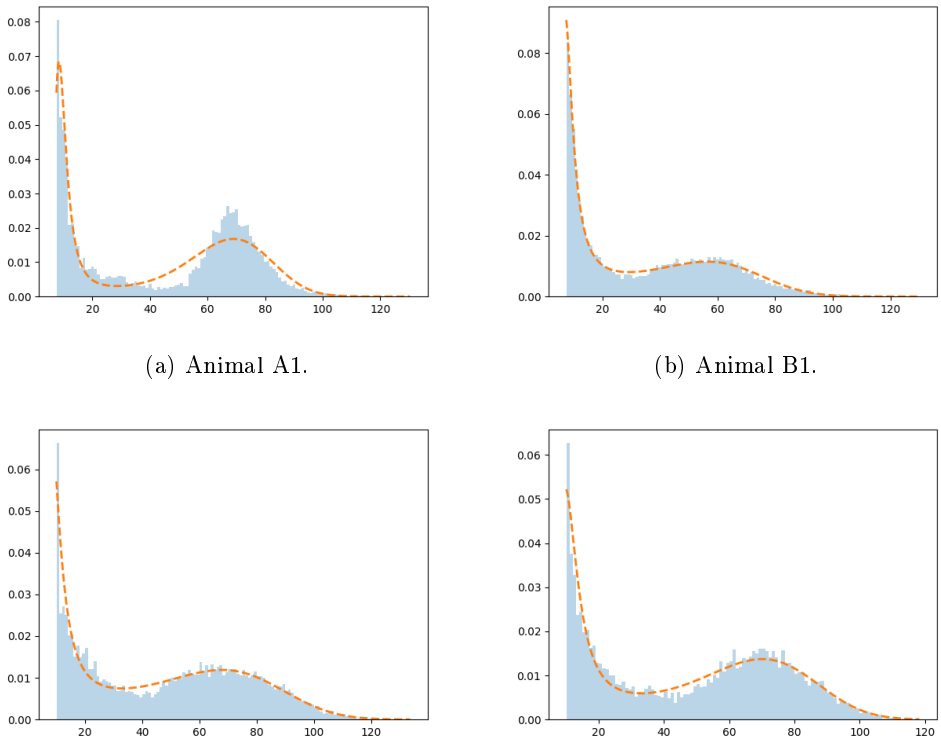
Figure 2.7: **Illustration numérique du théorème 2.5.3 :** norme L^1 empirique entre deux solutions d'équations (1.71) et (2.26)

Chapter 3 : A Lifschitz-Slyozov type model for adipocyte size dynamics : limit from Becker-Döring system and numerical simulation

- un nouveau modèle de Lifshitz-Slyozov du second ordre (2.16) pour la distribution en taille des adipocytes avec un terme de diffusion dérivé d'un modèle discret,
- deux modèles Becker-Döring et Lifshitz-Slyozov avec une vitesse inhabituelle (2.7) -(2.9) avec trois zéros et un terme de saturation en L, ce qui conduit à différents types de solutions stationnaires,
- une loi de conservation supplémentaire (2.12) par rapport aux systèmes classiques, imposant des conditions aux bords inhabituelles, voir Eq. (2.15),
- un nouveau résultat de preuve de convergence de solutions de Becker-Döring vers des solutions de Lifshitz-Slyozov, en utilisant des queues de distributions, qui fournit une borne supérieure sur la vitesse de convergence,
- des résultats numériques montrant que des distributions bimodales, ainsi que des profils unimodaux, peuvent être obtenus asymptotiquement avec system (2.16), selon les paramètres,
- résultats numériques explorant l'influence du paramètre ε et comparant le terme de diffusion du système (2.16) avec un coefficient constant de temps et d'espace,
- résultats numériques montrant que le système du second ordre (2.16) fournit un profil asymptotique universel qui ne dépend pas de la condition initiale (mais seulement de λ, m), contrairement au système du premier ordre (2.13).

Chapter 4 : A stochastic approach to adipose cell modeling

- deux modèles stochastiques : la CTMC deBecker-Döring (2.25) non linéaire et la SDE de Lifshitz-Slyozov (2.28). Les deux modèles sont inspirés des modèles déterministes Becker-



(a) Animal A1.

(b) Animal B1.

(c) Animal C1.

(d) Animal C5.

Figure 2.8: **Comparaison des données du modèle**. Quatre exemples (sur 32) de distributions de rayons d'adipocytes (en μm) sous forme d'histogrammes chez le rat dans des conditions physiologiques normales et la sortie du modèle calculée (lignes pointillées) avec des paramètres estimés (voir Section 5.1.4) . Les estimations des paramètres sont effectuées avec l'algorithme CMA-ES du package Python `cma` en minimisant la fonction $\mathcal{L}$ eq. 5.18.

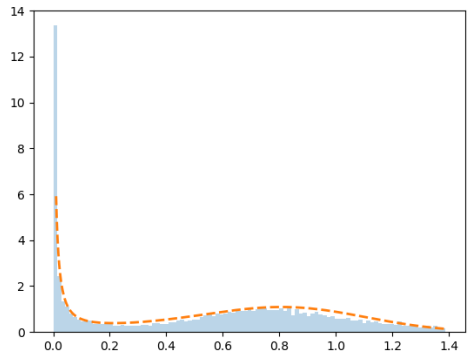


Figure 2.9: **Comparison model-data.** exemple de distributions de taille dans les lipides sous forme d'histogrammes et de sorties de modèle (lignes pointillées) calculées à partir des résultats de la méthode CMA-ES censurant les derniers 20 % des données.

Döring et Lifshitz-Slyozov pour la modélisation des cellules adipeuses,

- un nouveau résultat de convergence pour les modèles stochastiques, basé sur les travaux de [74], pour des taux bornés de lipogenèse et de lipolyse,
- une extension de ce résultat de convergence au cas des taux sous-linéaires.

Chapter 5 : Parameter estimation and some additional numerical results

- une analyse d'identifiabilité structurale sur le modèle de Lifshitz-Slyozov diffusif constant en rayons (2.35) qui montre que sous une re-paramétrisation, le modèle est identifiable,
- une estimation de paramètre sur le modèle diffusif constant Lifshitz-Slyozov (2.22) à l'aide de la méthode CMA-ES. Nous sommes en mesure d'estimer le taux de surface de la lipogenèse et de montrer qu'il est d'un ordre de grandeur similaire au taux de surface de la lipolyse.
- une estimation de paramètre sur le modèle diffusif Lifshitz-Slyozov (2.16) en utilisant la méthode CMA-ES. Nous récupérons des paramètres du même ordre de grandeur que dans le cas de la diffusion constante et montrons que ce modèle capture mieux le nadir de la distribution.

Dans le Chapitre 7, nous présentons une partie des travaux en cours pour la modélisation des cellules adipeuses ainsi que les perspectives de ces travaux.

2.6 CEMRACS

Au cours du mois d'août 2022, j'ai eu l'opportunité de participer à la Summer School du CEMRACS. Cela s'est accompagné de l'opportunité de travailler sur un projet pendant 5 semaines. Ce travail a été réalisé en collaboration avec Olivier Bernard, Mickael Bestard, Thierry Goudon, Sébastien Minjeaud, Florent Noisette et Bastien Polizzi. Nous présentons ici succinctement les objectifs et la méthode de ce travail puisqu'il sort du cadre de la modélisation des cellules

adipeuses. Nous renvoyons le lecteur au chapitre 6 pour plus de détails. L'objectif de ce travail était de réaliser des simulations numériques pour un modèle de biofilms. Les biofilms sont un consortium d'organismes unicellulaires intégrés dans une matrice extracellulaire qui prolifèrent sur une surface. Les exemples de biofilms vont du microbiote à l'intérieur de l'intestin aux micro-algues à la surface ou sur les rochers. Ce modèle décrit l'interaction des cellules à l'intérieur du biofilm avec sa matrice extra-cellulaire et un milieu liquide. Ces trois composants sont décrits par leur fraction volumique et leur vitesse. Ainsi, les fractions volumiques totalisent un. L'un des objectifs de ce projet était d'adapter un schéma numérique de [13] capable de préserver la contrainte sur la somme de la fraction volumique au niveau numérique. Ce schéma a été développé pour des écoulements multifluides : différents types de liquides qui sont dans la même phase. Nous montrons que le schéma numérique que nous utilisons est capable de maintenir la contrainte au niveau numérique, voir Figure 1.11. Nous obtenons le modèle d'onde progressive attendu, voir Figure 1.12, et ajoutons divers substrats qui jouent potentiellement un rôle dans la croissance du biofilm.

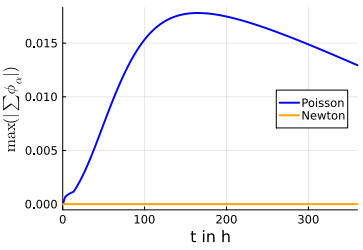


Figure 2.10: Evolution temporelle de l'erreur maximale dans le domaine sur la somme des fractions volumiques : $E = \max_x |\sum_\alpha \phi_\alpha - 1|$.

(a) Fraction volumique à t = 120h (b) Fraction volumique à t = 240h (c) Fraction volumique à $t = 360$h

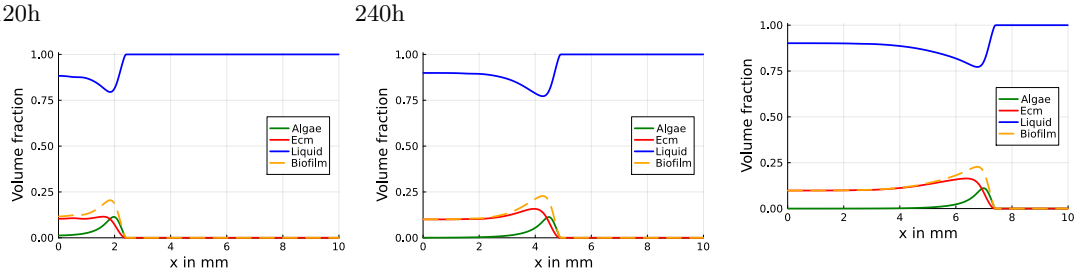


Figure 2.11: Fractions volumiques des composants du mélange pour différents temps.

Chapter 3

A Lifschitz-Slyozov type model for adipocyte size dynamics : limit from Becker-Döring system and numerical simulation

3.1 Brief introduction

In this chapter we introduce the work titled "A Lifschitz-Slyozov type model for adipocyte size dynamics : limit from Becker-Döring system and numerical simulation"[90]. We briefly go over what the article contains. The main theoretical result is the proof of theorem 1.5.1. We also present some numerical result using the well-balanced scheme introduced in section 1.4.7. The article is slightly edited for ease of reading and continuity within the manuscript. This work is submitted and small editions were made for ease of reading in the manuscript.

Abstract

Biological data show that the size distribution of adipose cells follows a bimodal distribution. In this work, we introduce a Lifshitz-Slyozov type model, based on a transport partial differential equation, for the dynamics of the size distribution of adipose cells. We prove a new convergence result from the related Becker-Döring model, a system composed of several ordinary differential equations, toward mild solutions of the Lifshitz-Slyozov model using distribution tail techniques. Then, this result allows us to propose a new advective-diffusive model, the second-order diffusive Lifshitz-Slyozov model, which is expected to better fit the experimental data. Numerical simulations of the solutions to the diffusive Lifshitz-Slyozov model are performed using a well-balanced scheme and compared to solutions to the transport model. Those simulations show that both bimodal and unimodal profiles can be reached asymptotically depending on several parameters. We put in evidence that the asymptotic profile for the second-order system does not depend on initial conditions, unlike for the transport Lifshitz-Slyozov model.

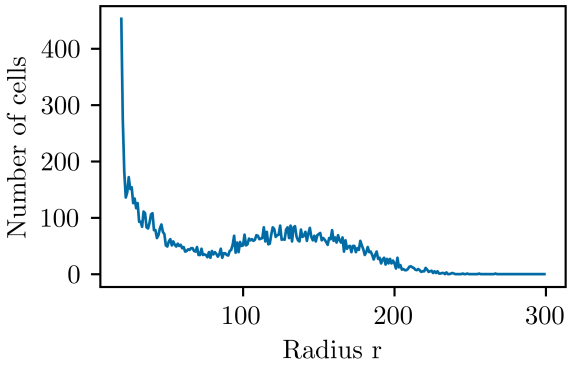


Figure 3.1: Example of the distribution of adipocytes in a rat biopsy. Credits : H. Soula

3.2 Introduction

White adipose tissue is mainly composed of cells, called adipocytes, which store lipids in the body under the form of triglyceride droplets. Experiments in most animals [66, 67, 118] show that the size distribution of adipocytes follows a striking bimodal distribution with a large peak for small adipocytes around the minimal radius, see Fig. 3.1. The changes in volume of an adipocyte are governed by two opposite phenomena : lipogenesis, that is to say size increase by triglyceride intake, and lipolysis, that is to say size decrease through the hydrolyze of triglycerides and the excretion of fatty acids. Modeling the dynamics of size evolution of adipocytes is of great interest in order to study metabolic disorders, such as obesity or type 2 diabetes. Correlation between such diseases and the size and metabolism of adipose cells has been well established in the biological literature. Indeed, in [128], authors show that the size of an adipose cell has a strong correlation with its insulin sensitivity. As such, large cells are less sensitive, therefore a higher body weight leads to greater risks of type 2 diabetes. This study also shows that adipose tissue are very heterogeneous in terms of size of cells. Those findings have also been described in [80], where the authors show that the adipose tissue is composed of cells that are different both molecularly and phenotypically.

Some computational models have also been used to provide insights into the adipose tissue physiology. In [71], the authors use an ODE model to investigate the role of lipases in the biochemistry of lipids. They are able to show that determining the active metabolic subdomain in the tissue is the key for accurate simulations, as well as the different activation rates of lipases for diglyceride and triglyceride breakdown. The rate of lipid turnover has also been studied in [8], where a decrease of the lipid release rate is correlated with the age of the individual. Finally, there are strong links between the adipose tissue and its extracellular matrix, and in case of obesity, one may observe tissue fibrosis such as described in [36]. In [100, 101], adipose tissue is modeled by a 2D agent-based model which takes into account the mechanical interactions between adipose cells and fibers forming the extra-cellular matrix. The authors study the spatial distribution of adipocytes under the form of lobules or in the case of tissue regeneration.

However, only a few mathematical models have been proposed in order to describe the size dynamics of adipose cells and no previous work has tackled a mechanistic understanding of the bimodal feature of adipocyte size distribution.

A first model has been derived by Jo, Periwal et al. in [68, 67] using a PDE for the adipose cell growth with a phenomenological cell growth rate. They are able to recover the bimodal feature

of distributions as well as to perform some curve fitting on biological data. In [118] and later in [119], the authors describe the velocity of size change of adipocytes by biological considerations for lipogenesis and lipolysis, leading thus to some transport PDE models. They obtain bimodal distributions by using stochastic variations of the parameters. In [46], the authors perform the analysis and numerical simulations for a size-structured model describing the evolution of a set of adipocytes, including the creation of new adipocytes through differentiation processes from mesenchymal cells and preadipocytes, and accounting for a size velocity inversely proportional to the total surface of adipocytes. Finally, in [104], authors use an ODE model to investigate interplay and feedback loop between inflammatory response of bigger adipose cells and the immune system, which may lead to type 2 diabetes. The size of adipocytes is updated at each time step according to some probability of swelling and by a factor depending on the surplus of calories intake. However they do not concern themselves with the size distribution but with the whole tissue inflammation and the body weight dynamic.

3.2.1 Transport equation for adipocyte size evolution

Following the work in [118], we first describe intake and release of lipids trough the cellular membrane, thus describing how the size of an adipose cell evolves. This will in turn allow us to build a model based on continuity equations.

Our first assumption will be the correlation between the amount of storage in an adipose cell and its radius. Cells shall be considered as spheres of a certain radius r, and the amount of lipids in the cell is denoted by x. Let us denote by $r(x)$ the radius of a cell containing x amount of lipids, by V_0 the volume of an empty cell and by V_{lipids} the molar volume of triglycerides.

We express the total volume of the cell in two different ways and we obtain the following relation:

$$V_{\text{lipids}}x + V_0 = \frac{4}{3}\pi r(x)^3,$$

which leads to :

$$r(x) = \left(\frac{3}{4\pi}(V_{\text{lipids}}x + V_0)\right)^{\frac{1}{3}}. \tag{3.1}$$

We also denote by L the amount of external lipids in the medium.

Henceforth, x will be considered as the size of our cell. Its variation $\frac{\mathrm{d}x}{\mathrm{d}t}$ depends on two flows : the intake of lipids by the cell from the medium and the release of lipids in the medium. As those two flows go through the membrane of the cell, they should be surface limited. We will also consider fast diffusion of the lipids in the medium so that the amount of lipids available for each cell is the same.

The intake term is a product of three factors :

- a term for a surface limited flow $\alpha r(x)^2$, where the constant α is the rate of this flow ;
- a Hill-like term with a radius cutoff ρ to describe resistance toward indefinite intake of lipids $\dfrac{\rho^n}{r(x)^n + \rho^n}$;
- a term that accounts for the available amount of lipids in the medium, in the form of a Michaelis-Menten term $\dfrac{L}{L+\kappa}$ with a saturation effect when the amount of external lipids L is large, with κ giving the order of magnitude of the threshold.

The release is a product of two terms :

- a term with a basal level of release β and a surface limited flow $\gamma r(x)^2$, where the constant γ is the release equivalent of the constant α ;
- a Michaelis-Menten term for the available amount of lipids in the cell $\dfrac{x}{x+\chi}$, where χ is the equivalent of κ for the release.

The variation $\dfrac{\mathrm{d}x}{\mathrm{d}t}$ can therefore be expressed as the difference between intake and release as :

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \underbrace{\alpha r(x)^2 \frac{\rho^n}{r(x)^n+\rho^n}\frac{L}{L+\kappa}}_{\text{intake}} - \underbrace{(\beta+\gamma r(x)^2)\frac{x}{x+\chi}}_{\text{release}}. \tag{3.2}$$

We can now build a transport equation for the distribution $f(t,x)$ of adipose cells by amount of lipids $x \geq 0$ at time t. According to Eq. (3.2), the transport velocity will be given by :

$$v(x,L) = a(x)\frac{L}{L+\kappa} - b(x), \tag{3.3}$$

where

$$a(x) = \alpha r(x)^2 \frac{\rho^n}{r(x)^n+\rho^n} \tag{3.4}$$

and

$$b(x) = (\beta+\gamma r(x)^2)\frac{x}{x+\chi}. \tag{3.5}$$

Consequently, the function f satisfies the following transport equation :

$$\partial_t f(t,x) + \partial_x(v(x,L)f(t,x)) = 0, \qquad x \geq 0, \qquad t > 0. \tag{3.6}$$

We now need to describe the behaviour of the available amount of lipids in the medium L. As per our assumption, the total quantity of lipids in our system, denoted by λ, should be constant. There are two types of lipids in the system : the ones contained in the cells, and the lipids in the medium and we therefore have the following equality :

$$L(t) + \int_{\mathbb{R}_+} x f(t,x)\mathrm{d}x = \lambda. \tag{3.7}$$

Another hypothesis is that the number of adipocytes does not change in time. Thus, in regards to boundary conditions, we want to preserve the total population number and therefore we impose that :

$$\int_{\mathbb{R}_+} f(t,x)\mathrm{d}x = \int_{\mathbb{R}_+} f^0(x)\mathrm{d}x = m \text{ for all } t > 0. \tag{3.8}$$

This leads for Eq. (3.6) to boundary condition

$$(v(x,L(t))f(t,x))\big|_{x=0} = 0\,, \quad \text{for all } t > 0\,. \tag{3.9}$$

Notice that by Eqs. (3.4)-(3.5), we have $b(0) = 0$ and $a(0) > 0$. Hence, the boundary conditions (3.9) is equivalent to the Dirichlet boundary condition :

$$f(t,x)\big|_{x=0} = 0 \text{ for all } t > 0. \tag{3.10}$$

To sum up, the transport model for adipose cells with initial conditions (f^0, L^0), that will be called first-order Lifshitz-Slyozov model in the following, reads as:

$$\begin{cases} \partial_t f(t,x) + \partial_x(v(x,L(t))f(t,x)) = 0, & \text{(3.11a)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} x f(t,x)\mathrm{d}x = \lambda, & \text{(3.11b)} \\ (v(x,L(t))f(t,x))\big|_{x=0} = 0, & \text{(3.11c)} \\ f(0,x) = f^0(x) \text{ and } L(0) = L^0. & \text{(3.11d)} \end{cases}$$

3.2.2 New models for adipose tissue dynamics

In this subsection, we present the various models under consideration in this article. Starting from the description of lipogenesis and lipolysis as done in Eq. (3.4) and (3.5) following [118], we build size-structured PDE model following the framework of Becker-Döring system [10] and Lifshitz-Slyozov equations [84] initially derived for polymerization.

The aim of this model is to reproduce the adipocyte size distributions observed experimentally and their bimodal structure. However, the transport equation (3.6) possesses asymptotic solutions as a linear combination of Dirac masses centered on the zeros of the asymptotic speed. To recover the bimodality, we are therefore lead to introduce a diffusion term in the equation : we can either add a diffusion term with a constant rate with no real biological meaning, or we can compute a time and space dependent diffusion term coming from the discrete nature behind the Lifshitz-Slyozov formalism [57, 129]. For that purpose, we come back to a Becker-Döring system of ODEs giving the evolution with respect to time of the number of adipocytes with discrete sizes and from this, we derive a second order Lifshitz-Slyozov equation with a diffusion term.

Therefore, in the following, we will consider three different models for the size distribution of an adipocyte population, namely

- the ODE system (3.13) with discrete sizes, a variant of the Becker-Döring model.
- the previously published transport equation (3.11), also called first order Lifshitz-Slyozov equation,
- the transport-diffusion equation (3.17), the second order Lifshitz-Slyozov equation.

In all three models, the lipogenesis and lipolysis rate will be given by Eq. (3.4) and (3.5), respectively.

Note that we impose in all three models two conservation laws: (i) the conservation of the total amount of lipids and (ii) the conservation of the total number of adipocytes. Therefore, all these models have a constant population number and are coupled with a lipid conservation equation which ensures that the sum of the lipids in the external medium and the lipids inside the cells is constant.

A brief insight in Becker-Döring and Lifshitz-Slyozov equations

Becker-Döring equations have been introduced in [10] to model polymers undergoing aggregation and fragmentation. The Lifshitz-Slyozov model was introduced in [84] and first used for nucleation in super-saturated solid solutions and polymerisation processes. A rigourous treatment of the mathematical properties of the Becker-Döring equations was given by [9]. The relation between Becker-Döring equations and Lifshitz-Slyozov model goes back to [98]. For a detailed review of both models, see [58] and references therein.

Let us explain briefly the idea of these models for polymers. We denote by c_i, the amount of polymers containing i monomers for $i \in \mathbb{N}^*$ and hence c_1 stands for the amount of monomers. A polymer of size i denoted p_i can gain one monomer and grow to p_{i+1} with rate a_i or lose one monomer and shrink to p_{i-1} with rate b_i. We may write as a system of ODEs for the time evolution of the number of polymers c_i, one for each size i. Furthermore, the total amount of monomers, i.e free monomers and monomers within polymers, is assumed constant, which leads to a conservation equation. Stationary solutions of the Becker-Döring equations can be easily computed, and long time behavior has been characterized by [9].

Another possibility for the modeling of polymerisation-fragmentation processes is to describe continuously the size of polymers through a variable $x \in \mathbb{R}$. The distribution of polymers of size x at time t is therefore denoted by $f(t,x)$ and the quantity of monomers at time t is denoted by $L(t)$. The distribution is classically transported as in Eq.(3.6) with speed $v(x,L) = a(x)L(t) - b(x)$ where $a(x)$ is the rate of polymerisation for size x and $b(x)$ is the rate of depolymerisation for size x. As previously, the total amount of monomers is conserved. Depending on the sign of $a(0)L(t) - b(0)$, boundary conditions should be provided for the system, see [35] for example.

After an adapted rescaling, it has been shown in various papers [34, 35, 78, 92, 110, 129] that the solutions to Becker-Döring system tend to the solutions to Lifshitz-Slyozov model. Formally, the limit up to second order can be considered and gives rise to an advection-diffusion equation as computed in [57, 129]. Existence of solutions is widely known for both models, see the seminal paper [9] for the Becker-Döring model and [19, 33] for the Lifshitz-Slyozov model.

Remark that Becker-Döring and Lifshitz-Slyozov equations have already been used in various contexts, for example modeling of biological phenomena, such as prions [39, 79, 115], [106], [54] or modeling in oceanography, see [132, 62].

A Becker-Döring model for adipose cells

Now, let us explain how we adapt this formalism to derive new models for adipocyte size dynamics. The purpose of this construction is to investigate the classical convergence theorems from Becker-Döring to Lifshitz-Slyozov and deduce the form of a diffusion term to add in our model.

We mention the main differences with the classical Becker-Döring and Lifshitz-Slyozov systems for polymerisation. First, velocity (3.2), arising from biological considerations, possesses three zeros for a well-chosen range of parameters which leads to bimodal asymptotic distributions, whereas classical choices for a and b are constant or power laws of x, which yields the existence of a single positive root. See also [17] for a polymerisation-fragmentation model without diffusion giving rise to bimodal asymptotics. Second, in our model, external lipids L cannot be assimilated to monomers c_1 and the conservation law (3.7) is therefore not the same as in the usual polymerisation models. Moreover, the saturation term $\dfrac{L}{L+\kappa}$ is not common in polymerisation modeling. Finally, our model conserves the total population number due to the boundary condition (3.11c), which adds an additional conservation law compared to the classical Becker-Döring model.

We shall now consider that an adipose cell is a bundle of smaller vesicles of typical size Λ. Hence the size of a cell can be defined by the number of vesicles it contains. We denote by c_i the number of cells of size i and by l the number of vesicles in the medium. A cell will aggregate a new vesicle with speed $a_i M(l)$,where $M(l) = \frac{l\Lambda}{l\Lambda+\kappa}$ following Eq. (3.2), and loose a vesicle at speed b_i, following this reaction :

$$\Lambda + i\Lambda \underset{b_{i+1}}{\overset{a_i M(l)}{\rightleftharpoons}} (i+1)\Lambda.$$

Following the rescaling procedure described in the introduction, we define $c^\varepsilon = (c_i^\varepsilon)_{i\geq 0}$ and

$J_i^\varepsilon(c^\varepsilon, L^\varepsilon)$ the flow of the previous reaction given by :

$$J_i^\varepsilon(c^\varepsilon, L^\varepsilon) = a_i^\varepsilon \frac{L^\varepsilon}{L^\varepsilon + \kappa} c_i^\varepsilon - b_{i+1}^\varepsilon c_{i+1}^\varepsilon,\ i \geq 0, \tag{3.12}$$

where a_i^ε (resp. b_i^ε) are discrete counterpart of the continous function a defined at Eq.(3.4) (resp. b defined at Eq.(3.5)), see Sec.3.3 for more details.

Similarly as before, see Eq. (3.7), L^ε will satisfy an equation accounting for conservation of the amount of lipids and we get the following ODE system :

$$\begin{cases} \dfrac{dc_i^\varepsilon}{\mathrm{d}t} = \dfrac{1}{\varepsilon}(J_{i-1}^\varepsilon(c^\varepsilon, L^\varepsilon) - J_i^\varepsilon(c^\varepsilon, L^\varepsilon)),\ \forall i \geq 1, & \text{(3.13a)} \\ \dfrac{dc_0^\varepsilon}{\mathrm{d}t} = -\dfrac{1}{\varepsilon} J_0^\varepsilon(c^\varepsilon, L^\varepsilon), & \text{(3.13b)} \\ L^\varepsilon(t) + \displaystyle\sum_{i=0}^{\infty} i\varepsilon^2 c_i^\varepsilon(t) = \lambda,\ \forall t \geq 0, & \text{(3.13c)} \\ L^\varepsilon(0) = L^{\varepsilon,0}, \quad c_i^\varepsilon(0) = c_i^{\varepsilon,0},\ \forall i \geq 1, & \text{(3.13d)} \end{cases}$$

which is similar to Becker-Döring equations except for the definition of the flux J_i^ε (saturing fluxes of monomers), and the minimal size is 0 and not 1. Observe also that there is no 'boundary' flux, thus the quantity

$$m = \varepsilon \sum_{i \geq 0} c_i^\varepsilon(t) \text{ is constant in time.} \tag{3.14}$$

This is the discrete analogous to the previous conservation (3.8) of the zeroth order moment of f.

A solution to the previous system exists according to Theorem 3.3.1 recalled in Sec.3.3. Now let us define the following step functions depending on both time and space :

$$f^\varepsilon(t,x) = \sum_{i \geq 0} \mathbb{1}_{\Gamma_i^\varepsilon}(x) c_i^\varepsilon(t),$$

where $\Gamma_i^\varepsilon = [(i - \frac{1}{2})\varepsilon, (i + \frac{1}{2})\varepsilon[$, and $(c_i^\varepsilon)_{i \geq 0}$ is a solution to (3.13).

Convergence of function f^ε when $\varepsilon \to 0$ towards a solution f of the Lifshitz-Slyozov equation (3.11) is a classical result, see Theorem 3.3.3 recalled in Sec.3.3. In the present work, we prove that a similar convergence result hold in a stronger topology, and with a control of the speed of convergence, of order at least ε.

To that, we introduce the tail distributions :

$$F(t,x) = \int_x^\infty f(t,y)\mathrm{d}y, \qquad F^\varepsilon(t,x) = \int_x^\infty f^\varepsilon(t,y)\mathrm{d}y.$$

The main analytical result of this article is the following theorem, whose more rigorous statement will be specified later in Theorem 3.4.1:

Theorem (Convergence of tails of distributions)**.** *Under some hypotheses detailed in Sec.3.3 and 3.4, there exists some constant $K > 0$ and some time $T > 0$ such that for all $t \in (0,T)$ and for ε small enough :*

$$|L^\varepsilon(t) - L(t)| + \int_{\mathbb{R}_+} |F^\varepsilon(t,x) - F(t,x)|\mathrm{d}x \leq \varepsilon K\,.$$

This result provides a new approach for looking into convergence from Becker-Döring to Lifshitz-Slyozov. Contrary to more classical results where convergence towards a weak solution is achieved using Ascoli-Arzela's Theorem, this theorem yields convergence towards mild solutions and gives a bound of order ε on the speed of this convergence.

A second order Lifshitz-Slyozov model

Another goal in this article is to derive a new model with a diffusive term from Becker-Döring system (3.13). One can see this diffusive term as a second order term emerging from the convergence theorem 3.4.1. There are various ways to yield this term, see for example [129], [110], [35]. The derivation of the diffusive term will be detailed in Section 3.5, but we present the model here for the sake of completeness.

The so-called second order Lifshitz-Slyozov model therefore takes the form of a transport-diffusion equation, with a diffusive term which depends both on x and $L(t)$, i.e. :

$$\partial_t g + \partial_x(vg) = \frac{\varepsilon}{2}\partial_x^2(dg), \qquad \forall x \geq 0,$$

where

$$d : (x, L) \in \mathbb{R}_+ \times \mathbb{R}_+ \to d(x, L) = a(x)\frac{L}{L+\kappa} + b(x). \tag{3.15}$$

We need to complement this PDE with adapted boundary conditions. Since we want the conservation of the zeroth order moment denoted by $\int_{\mathbb{R}_+} g(t,x)\mathrm{d}x = m$, we need to impose the following null-flux boundary condition :

$$(-vg + \frac{\varepsilon}{2}\partial_x(dg))\Big|_{x=0} = 0. \tag{3.16}$$

Therefore, we consider the following system, which consists of the previous PDE and boundary conditions, complemented by previous constraint (3.7) and initial conditions for g and L :

$$\begin{cases} \partial_t g + \partial_x(vg) = \dfrac{\varepsilon}{2}\partial_x^2(dg), & \text{(3.17a)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} xg(t,x)\mathrm{d}x = \lambda, & \text{(3.17b)} \\ \Big(- vg + \dfrac{\varepsilon}{2}\partial_x(dg)\Big)\Big|_{x=0} = 0, & \text{(3.17c)} \\ g(0,x) = g^0(x) \text{ and } L(0) = L^0. & \text{(3.17d)} \end{cases}$$

We provide interesting numerical evidence of stationary solutions of the advection-diffusion model (3.17) following a bimodal distribution. The numerical simulations are performed using a well-balanced scheme developped in [51]. We also demonstrate that to observe a bimodal asymptotics, parameters should be taken into an adapted parameter range.

3.2.3 Outline of the article

In Section 3.3, we will give some preliminary results on the existence of solutions to systems (3.13) and (3.11). In Section 3.4, we will show the convergence theorem thanks to the tail of distributions technique. Then, in Section 3.5, we derive formally the second-order Lifshitz-Slyozov model, that is to say system (3.17) and we give the expression for its stationary solutions. In Section 3.6, we display some numerical results and we show that bimodality of the stationary solution can be observed in well-chosen parameter range. Finally, we discuss our results in Section 3.7.

3.3 Preliminary results

In this section, we give the main already-known results of existence of solutions to systems (3.13) and (3.11) and convergence of solutions to system (3.13) towards (3.11). Proofs have been easily adapted to our framework.

3.3.1 Existence results on Becker-Döring system

We consider first the Becker-Döring system (3.13) for fixed ε.

We assume that there exist some strictly positive constants A, B, C_a, C_b, K_a, K_b and δ such that for all $i \geq 0$:

$$a_i^\varepsilon \leq C_a \text{ and } b_i^\varepsilon \leq C_b i\varepsilon, \tag{H'1}$$

$$|a_i^\varepsilon - a_{i+1}^\varepsilon| \leq K_a\varepsilon \text{ and } |b_i^\varepsilon - b_{i+1}^\varepsilon| \leq K_b\varepsilon. \tag{H'2}$$

We define the state space for Eq. (3.13) by

$$X := \{x = (x_i)_{i\geq 0} \in \mathbb{R}_+^{\mathbb{N}} : \sum_{i=0}^{+\infty} i x_i < \infty\},$$

endorsed with the norm $\|x\|_X = \sum_{i=0}^{+\infty} i|x_i|$. We denote $x \geq 0$ if $x_i \geq 0$ for all $i \geq 0$, and $X^+ := \{x \in X : x \geq 0\}$. We give the following definition of solution to Eq. (3.13):

Definition 3.3.1. *Let $T > 0$ and $\varepsilon > 0$. A solution $(c^\varepsilon, L^\varepsilon)$ of (3.13) in $[0,T)$ is a couple of a function $L^\varepsilon : [0,T) \to \mathbb{R}$ and a sequence of functions $c^\varepsilon = (c_i^\varepsilon)_{i\geq 0}$, $c_i^\varepsilon : [0,T) \to X$ such that :*

(i) For all $t \in [0,T)$, $L^\varepsilon(t) \geq 0$ and $c^\varepsilon(t) \in X^+$,

(ii) For all $i \geq 1$, $c_i^\varepsilon : [0,T) \to \mathbb{R}$ is continuous and $\sup_{t\in[0,T)} \|c^\varepsilon(t)\|_X < +\infty$,

(iii) $L^\varepsilon : [0,T) \to \mathbb{R}$ is continuous and $\sup_{t\in[0,T[} |L^\varepsilon(t)| < +\infty$,

(iv) For all $t \in [0,T)$, $\int_0^t \sum_{i=0}^{+\infty} a_i^\varepsilon c_i^\varepsilon(s)\mathrm{d}s < \infty$ and $\int_0^t \sum_{i=0}^{+\infty} b_i^\varepsilon c_i^\varepsilon(s)\mathrm{d}s < \infty$,

(v) For all $t \in [0,T)$, for all $i \geq 1$:

$$c_i^\varepsilon(t) = c_i^{\varepsilon,0} + \frac{1}{\varepsilon}\int_0^t [J_{i-1}^\varepsilon(c^\varepsilon(s), L^\varepsilon(s)) - J_i^\varepsilon(c^\varepsilon(s), L^\varepsilon(s))]\mathrm{d}s,$$
$$c_0^\varepsilon(t) = c_0^{\varepsilon,0} - \frac{1}{\varepsilon}\int_0^t J_0^\varepsilon(c^\varepsilon(s), L^\varepsilon(s))\mathrm{d}s,$$
$$L^\varepsilon(t) = L^{\varepsilon,0} - \varepsilon\int_0^t \sum_{i=0}^{+\infty} J_i^\varepsilon(c^\varepsilon(s), L^\varepsilon(s))\mathrm{d}s$$

Well-posedness of solutions to (3.13) as defined at Def.3.3.1 can be shown by finite dimensional approximation, using the method developped in [9] :

Theorem 3.3.1. *Let $T > 0$ and $\varepsilon > 0$. Let $L^{\varepsilon,0} \in \mathbb{R}_+$ et $c^{\varepsilon,0} \in X^+$ such that $L^{\varepsilon,0} + \sum_{i=0}^{+\infty} i\varepsilon^2 c_i^{\varepsilon,0} = \lambda < \infty$. Assume that (H'1), (H'2) hold true. Then there exists a unique solution $(c^\varepsilon, L^\varepsilon)$ to Becker-Döring system (3.13) in the sense of Def.3.3.1 which satisfies initial conditions $c^\varepsilon(0) = c^{\varepsilon,0}$ and $L^\varepsilon(0) = L^{\varepsilon,0}$.*

The uniqueness and conservation properties of the solution are obtained using the following proposition that will be needed later on, see Sec.3.5. In particular, the following proposition states that any solution of the Becker-Döring system (3.13) preserves the first two moments for all times, and provides the starting point to compute any admissible moments for the solution of the Becker-Döring system. In [9], we can find the following Theorem 2.5 that we reproduce here for the reader's convenience :

Proposition 3.3.1. *Let $(\phi_i)_{i\geq 0}$ be a given sequence. Let $(c^\varepsilon, L^\varepsilon)$ be the solution of (3.13) on $[0,T)$, $0 < T \leq +\infty$.*

Assume that for all $0 \leq t_1 < t_2 < T$, $\int_{t_1}^{t_2} \sum_{i=0}^{\infty} |\phi_{i+1} - \phi_i| a_i^\varepsilon c_i^\varepsilon(t)\mathrm{d}t < \infty$ and that either of the following holds :

(a) $\phi_i = \mathcal{O}(i)$ and $\int_{t_1}^{t_2} \sum_{i=0}^{\infty} |\phi_{i+1} - \phi_i| b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t < \infty$ or

(b) $\sum_{i=0}^{\infty} \phi_i c_i^\varepsilon(t_k) < \infty$, for $k = 1, 2$ and $\phi_{i+1} \geq \phi_i \geq 0$ for i large enough.

Then :

$$\sum_{i=0}^{\infty} \phi_i c_i^\varepsilon(t_2) - \sum_{i=0}^{\infty} \phi_i c_i^\varepsilon(t_1) + \int_{t_1}^{t_2} \sum_{i=0}^{\infty} \frac{\phi_{i+1} - \phi_i}{\varepsilon} b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t = \int_{t_1}^{t_2} \sum_{i=0}^{\infty} \frac{\phi_{i+1} - \phi_i}{\varepsilon} a_i^\varepsilon \frac{L^\varepsilon(t)}{L^\varepsilon(t) + \kappa} c_i^\varepsilon(t)\mathrm{d}t.$$

3.3.2 Lifschitz-Slyozov system and classical convergence result

Even though we have precise forms for the intake and release functions, for the sake of generality we make the following assumptions on functions a and b occurring in Eq. (3.3) :

$$a, b \in C^1(\mathbb{R}_+, \mathbb{R}_+), \tag{H1}$$

$$a(0) > 0 \text{ and } \sup_{x\in\mathbb{R}_+} |a(x)| = C_a, \tag{H2a}$$

$$|b(x)| \leq C_b x \text{ for all } x \in \mathbb{R}_+ \text{ and } \lim_{R\to\infty} \sup_{x\geq R} \frac{b(x)}{x} = 0, \tag{H2b}$$

$$\sup_{x\in\mathbb{R}_+} |a'(x)| = K_a \text{ and } \sup_{x\in\mathbb{R}_+} |b'(x)| = K_b, \tag{H3}$$

with $C_a, C_b, K_a, K_b > 0$. We first define measured-valued solutions to the Lifschitz-Slyozov system (3.11), following [33]

Definition 3.3.2. *Given an initial condition $(f^0, L^0) \in C^0(\mathbb{R}_+) \cap L^1(\mathbb{R}_+, (1+x)\mathrm{d}x) \times \mathbb{R}_+$, a measured-valued solution to system (3.11) is composed of two functions $f \in C(0,T; \mathcal{M}^1(0,\infty) - weak - *)$ and $L \in C(0,T)$ such that for all $0 < t < T$ and for all $\varphi \in C^1([0,T] \times \mathbb{R}_+)$ the following relations hold:*

$$\int_0^T \int_{\mathbb{R}_+} (\partial_t \varphi(t,x) + v(x, L(t)) \partial_x \varphi(t,x)) f(t, \mathrm{d}x) + \int_{\mathbb{R}_+} \varphi(0,x) f^0(x)\mathrm{d}x = 0,$$

$$L(t) + \int_{\mathbb{R}_+} x f(t, \mathrm{d}x) = \lambda.$$

Now, let us state the convergence of solutions to Becker-Döring system towards solutions to Lifschitz-Slyozov system. In order to compare solutions to Becker-Döring system to solutions to Lifschitz-Slyozov system, we need to define the following piecewise constant functions. Let $\Gamma_i^\varepsilon = [(i-\frac{1}{2})\varepsilon, (i+\frac{1}{2})\varepsilon)$ and c_i^ε be solutions to (3.13), then we define

$$\begin{cases} f^\varepsilon(t,x) = \sum_{i\geq 0} \mathbb{1}_{\Gamma_i^\varepsilon}(x) c_i^\varepsilon(t)\,, & \text{(3.18a)} \\ a^\varepsilon(x) = \sum_{i\geq 0} \mathbb{1}_{\Gamma_i^\varepsilon}(x) a_i^\varepsilon\,, & \text{(3.18b)} \\ b^\varepsilon(x) = \sum_{i\geq 1} \mathbb{1}_{\Gamma_i^\varepsilon}(x) b_i^\varepsilon\,, & \text{(3.18c)} \end{cases}$$

where we assume that :

$$a_i^\varepsilon = a(i\varepsilon) \text{ and } b_i^\varepsilon = b(i\varepsilon), \text{ for all } i \geq 0 \text{ and } \varepsilon > 0. \tag{H4}$$

Given our definitions in Eq. (3.18), from Proposition 3.3.1 and with $\phi_i = \int_{\Gamma_i^\varepsilon} \phi(x)\mathrm{d}x$, we deduce the following proposition, that is the starting point to study the convergence of the solution of the Becker-Döring system (3.13) towards solution of the Lifshitz-Slyozov equation (3.11).

Proposition 3.3.2. *Let $\phi \in L^\infty(\mathbb{R}_+)$. Then for every $t \geq 0$, we have the following equality :*

$$\int_0^\infty \phi(x)(f^\varepsilon(t,x) - f^\varepsilon(0,x))\mathrm{d}x = \int_0^t \int_0^\infty (\Delta_\varepsilon \phi(x) a^\varepsilon(x) \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \Delta_{-\varepsilon}\phi(x) b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t,$$

where

$$\Delta_\varepsilon \phi(x) = \frac{\phi(x+\varepsilon) - \phi(x)}{\varepsilon}. \tag{3.19}$$

Finally, we obtain the following convergence theorem from the Becker-Döring equations to the Lifshitz-Slyozov equations, as in [129]:

Theorem 3.3.2. *Consider an initial condition $(L^{\varepsilon,0}, (c_i^{\varepsilon,0})_{i\geq 0})$ and the corresponding solution $(L^\varepsilon, (c_i^\varepsilon)_{i\geq 0})$ in the sense of Definition 3.3.1. We assume that there exists a constant $K > 0$ and $0 < s \leq 1$ such that :*

- $L^{\varepsilon,0} + \varepsilon^2 \sum_{i\geq 0} i c_i^{\varepsilon,0} = \lambda$,
- $\varepsilon \sum_{i\geq 0} c_i^{\varepsilon,0} < K$,
- $\varepsilon \sum_{i\geq 0} (i\varepsilon)^{1+s} c_i^{\varepsilon,0} < K$.

We also assume hypotheses H1 - H4 to hold. Then there exists a sequence ε_n and a solution (f, L) to (3.11) in the sense of Def. 3.3.2 such that :

$$\begin{cases} f^{\varepsilon_n} \to f\,, \;\; x f^{\varepsilon_n} \to xf \text{ in } C^0([0,+\infty[; \mathcal{M}^1(0,+\infty) - weak - *), \\ L^{\varepsilon_n} \to L \text{ uniformly in } C^0([0,T]). \end{cases}$$

Now, let us consider the existence of mild solutions to (3.11). For that purpose, we first define the characteristic curves.

Assume $L \in C^0(\mathbb{R}_+)$ to be given. The characteristic curves associated to (3.11) are solutions to :

$$\begin{cases} \partial_s X(s;t,x) = v(X(s;t,x), L(s)), \\ X(t;t,x) = x. \end{cases}$$

Since v is C^1 in both x and L, the characteristics are uniquely defined and form an ordered family. We denote $I_{t,x}$ their maximal time interval and by $X_c(t) = X(t;0,0)$ the characteristic curve that is equal to 0 at time 0. Then, a mild solution to system (3.11) is given by the following definition :

Definition 3.3.3. *Given a smooth initial condition f^0 and $L \in C^0(\mathbb{R}_+)$, a mild solution of*

$$\begin{cases} \partial_t f + \partial_x(v(x, L(t))f) = 0, \\ (v(x, L(t))f(t,x))\big|_{x=0} = 0, \\ f(0,x) = f^0(x), \end{cases}$$

is given by :

$$f(t,x) = f^0(X(0;t,x)) \exp\left(- \int_0^t \partial_x v(X(s;t,x), L(s)) ds \right) \mathbb{1}_{(X_c(t),\infty)}(x).$$

A couple (f, L) is said to be a solution of (3.11) iff f is a mild solution associated to L and $L : \mathbb{R}_+ \to \mathbb{R}_+$ solves $L(t) + \int_{\mathbb{R}_+} x f(t,x) \mathrm{d}x = \lambda$ for all $t \geq 0$.

Remark. *Since we impose null-flux boundary conditions on this system : $v(x, L(t))f(t,x)|_{x=0} = 0$, there is no term involving 'incoming characteristics' $\mathbb{1}_{(0,X_c(t))}(x)$.*

We follow the proofs in [33] and [19] and we obtain in a straightforward way the expected existence and uniqueness result :

Theorem 3.3.3. *Given an initial condition $(f^0, L^0) \in C^0(\mathbb{R}_+) \cap L^1(\mathbb{R}_+, (1+x)\mathrm{d}x) \times \mathbb{R}_+$ and assuming hypotheses (H1)-(H3), Lifschitz-Slyozov system (3.11) has a unique solution on the interval $[0,T]$ in the sense of Def.3.3.3.*

Note that the mild solution given by Theorem 3.3.3 is also a weak solution in the sense of Definition 3.3.2, see [19], and under hypotheses (H1)-(H3) both definitions coincide.

3.4 A new convergence result from Becker-Döring to Lifschitz-Slyozov equations

In this part of our work we introduce a different way to see the convergence from the Becker-Döring equations to the Lifschitz-Slyozov equations. Using tail distributions allows to reduce the non linearity of our system by pulling the speed of advection outside of the space derivative. Tail distributions were also found to be useful to obtain a quasi comparison principle in [22] and to obtain refined uniqueness properties in [77, 19]. The main idea is to use results on the tail of the distributions to show convergence. Finally, we note that out result uses the fact that a solution to system (3.11) exists while the previous result also shows existence of solution of (3.11), by showing a convergence to a measure valued function which turns out to be a solution of (3.11).

Let $(f^\varepsilon, L^\varepsilon)$ be the solution of the Becker-Döring ODE system (3.13) and Eq. (3.18), and let (f, L) the mild solution of Lifshitz-Slyozov equations (3.11). We recall the tail distribution definition,

$$F(t,x) = \int_x^\infty f(t,y)\mathrm{d}y, \qquad F^\varepsilon(t,x) = \int_x^\infty f^\varepsilon(t,y)\mathrm{d}y\,, \tag{3.20}$$

and introduce their difference

$$E(t,x) = F^\varepsilon(t,x) - F(t,x)\,. \tag{3.21}$$

We introduce the following additional hypotheses to use in our main theorem :

$$\sup_{x\in\mathbb{R}_+} |a''(x)| < +\infty \text{ and } \sup_{x\in\mathbb{R}_+} |b''(x)| < +\infty, \tag{H5}$$

$$\sum_{i\geq 0} |c_{i+1}^{\varepsilon,0} - c_i^{\varepsilon,0}| < +\infty, \tag{H6}$$

$$\varepsilon \sum_{i\geq 0} i|c_{i+1}^{\varepsilon,0} - c_i^{\varepsilon,0}| < +\infty. \tag{H7}$$

$$\text{There exists some constant } \bar{L} > 0, \text{ such that } \inf_{\varepsilon>0} L^{\varepsilon,0} \geq \bar{L}. \tag{H8}$$

$$\text{There exists some constant } K > 0 \text{ such that } \sup_{\varepsilon>0} c_0^{\varepsilon,0} < K. \tag{H9}$$

We now state our main theorem.

Theorem 3.4.1. *Let $T > 0$. Suppose that there exists some constant $C_{\text{init}} > 0$ such that for all $\varepsilon > 0$, $\int_{\mathbb{R}_+} |E(0,x)|\mathrm{d}x \leq \varepsilon C_{\text{init}}$. Also assume that hypotheses* (H1)-(H9) *hold true. Then there exists some constant $C(T) > 0$ such that for $\varepsilon > 0$ small enough and for all $t \in [0,T]$:*

$$|L^\varepsilon(t) - L(t)| + \int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x \leq \varepsilon C(T).$$

The proof proceeds as follows. Taking inspiration from [77, 19], we first note that owing to the total population number conservation, the lipid terms can be controlled by the tail, $|L^\varepsilon(t) - L(t)| \leq \int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. The control on the tail relies on a Grönwall's lemma argument. For that purpose, we derive the equation followed by $F^\varepsilon(t,x)$ (Lemma 3.4.5). We point out that the case $x < \varepsilon/2$ has to be treated separately due to remaining boundary terms. This allows us to give a first estimate on the integral $\int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$. We then make use of the mild solution formulation to derive the partial differential equation followed by F and in turn the one followed by E (Lemma 3.4.6). The proof follows by bounding the terms in the estimate on $\int_{\mathbb{R}_+} |E(t,x)|\mathrm{d}x$, and in particular we show that $F^\varepsilon(t,x)$ satisfies the same equation as F up to an order ε (Lemma 3.4.7). To this end, the key argument relies on refined estimates of the difference between the first order derivative of $F^\varepsilon(t,x)$ and its discrete analog. This estimate needs uniform control on the solutions c_i^ε of the Becker-Döring system and their increments $c_{i+1}^\varepsilon - c_i^\varepsilon$ (Subsection 3.4.1, Lemmas 3.4.1 to 3.4.4), which is new, up to our knowledge.

Hypotheses (H1) - (H4) are classical in the study of our model. However, other assumptions are less common but arise naturally from the result. Contrary to the classical convergence result, we work with mild solutions of the Lifshitz-Slyozov system. Hence, we need proper bounds on

second order terms. We shall see in Section 3.5 that those terms lead us to the second order Lifshitz-Syozov model. Nonetheless, those terms involve second order derivatives of both a and b which leads us to hypothesis (H5). Hypotheses (H6) and (H7) simply tell us that the initial condition for the Becker-Döring system must have finite zeroth order moment and first moment increments. Lemma 3.4.4 shows that this property propagates in time. Additional assumptions have to be made to obtain our main theorem. The assumption (H8) on the initial condition $L^{\varepsilon,0}$ is necessary since it leads to strict positivity of L^ε in finite time, uniformly in ε. The assumption (H9) on the initial condition $c_0^{\varepsilon,0}$ is technical and ensures that the proper boundary condition (3.11c) is satisfied for all times. Finally the assumption on $\int_{\mathbb{R}_+} |E(0,x)|\mathrm{d}x$ is made to conclude after using Grönwall's lemma at the very end of the proof. This assumption relates both initial conditions $(c_i^{\varepsilon,0})_{i\geq 0}$ and f^0. A fair choice for the initial condition $(c_i^{\varepsilon,0})_{i\geq 0}$ is $c_i^{\varepsilon,0} = f^0(i\varepsilon)$ for all $i \geq 0$. Then the assumption is verified as long as $(f^0)' \in L^1(\mathbb{R}_+, x\mathrm{d}x)$.

In all this section, we assume that hypotheses (H1) - (H9) hold true.

3.4.1 Preliminary results on Becker-Döring system

We start with a lemma that allows to control the lipid term away from 0, in the lines of previous results from [19].

Lemma 3.4.1. *A solution $(L^\varepsilon, c^\varepsilon)$ of (3.13) with $\lambda > 0$ verifies that there exists $C > 0$ such that for all $t > 0$,*

$$\inf_{\varepsilon>0} L^\varepsilon(t) \geq \bar{L}\exp(-Ct). \tag{3.22}$$

Proof. For all $t > 0$, we have, using the three first equations of system (3.13) :

$$\begin{aligned}\frac{\mathrm{d}L^\varepsilon(t)}{\mathrm{d}t} = -\varepsilon\sum_{i\geq 0} J_i^\varepsilon(c^\varepsilon(t), L^\varepsilon(t)) &= -\varepsilon\sum_{i\geq 0}\left(a_i^\varepsilon \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} c_i^\varepsilon(t) - b_{i+1}^\varepsilon c_{i+1}\varepsilon(t)\right) \\ &\geq -\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}\varepsilon\sum_{i\geq 0} a_i^\varepsilon c_i^\varepsilon(t),\end{aligned}$$

and thus, because $\sup_{x\in\mathbb{R}_+}|a(x)| = C_a$, and $\frac{L^\varepsilon}{L^\varepsilon+\kappa} \leq \frac{1}{\kappa}L^\varepsilon$:

$$\frac{\mathrm{d}L^\varepsilon(t)}{\mathrm{d}t} \geq -\frac{C_a}{\kappa}L^\varepsilon\int_{-\varepsilon/2}^{+\infty} f^\varepsilon(t,x)\mathrm{d}x$$

and by conservation of the moment (3.14), $\int_{-\varepsilon/2}^{+\infty} f^\varepsilon(t,x)\mathrm{d}x = \varepsilon\sum_{i\geq 0} c_i^\varepsilon(t) = m$,

$$\frac{\mathrm{d}L^\varepsilon(t)}{\mathrm{d}t} \geq -\frac{C_a m}{\kappa}L^\varepsilon.$$

We conclude by Grönwall's lemma and using Hypothesis (H8). □

We next state a lemma adapted from [35] that allows to obtain pointwise estimates of the density f^ε near the boundary, through the uniform propagation of exponential moments. For $x \in \mathbb{R}_+$ and $t > 0$, let

$$H^\varepsilon(t,x) = \sum_{i\geq 0} c_i^\varepsilon(t)\,\mathrm{e}^{-ix}\ .$$

Lemma 3.4.2. *Let $x \in \mathbb{R}_+^*$. Then there exist some constants $\tilde{K} > 0$ and $\varepsilon^* > 0$ such that for all $0 < \varepsilon < \varepsilon^*$:*

$$H^\varepsilon(t,x) \leq H^\varepsilon(0,x) + \tilde{K} \text{ for all } t > 0,$$

and in particular :

$$\text{for all } i \geq 0, \sup_{0<\varepsilon<\varepsilon^*} \sup_{t\in[0,T]} c_i^\varepsilon(t) \leq \bar{c}_i < +\infty. \tag{3.23}$$

Proof. Using Lemma 3.4.1, and the assumption (H8) on $L^\varepsilon(0)$, we have that $\inf_{\varepsilon>0} \inf_{t\in(0,T]} L^\varepsilon(t) \geq \bar{L}\exp(-CT)$. Thus we can find a constant $c > 0$ such that :

$$\inf_{\varepsilon>0} \inf_{t\in(0,T]} \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} \geq c.$$

Now we choose $\delta > 0$ such that $c > 2\delta$. Using Taylor's expansion, we have $a(i\varepsilon) = a(0) + i\varepsilon a'(0) + \mathcal{O}((i\varepsilon)^2)$. Then with hypotheses (H2b) and (H3) and for ε small enough, we find that $a(i\varepsilon) \geq \frac{3}{4}\left(a(0) - i\varepsilon K_a\right) > 0$. Therefore we have that for ε small enough :

$$\forall i \leq \frac{1}{\sqrt{\varepsilon}},\ a(i\varepsilon) \geq \frac{a(0)}{2}. \tag{3.24}$$

In turn, by hypotheses (H2a), (H2b) and (H4), we have that for ε small enough and for all $i \leq \dfrac{1}{\sqrt{\varepsilon}}$:

$$\frac{b_i^\varepsilon}{a_i^\varepsilon} = \frac{b(i\varepsilon)}{a(i\varepsilon)} \leq 2C_b \frac{\sqrt{\varepsilon}}{a(0)} \xrightarrow[\varepsilon\to 0]{} 0.$$

Let $x \in \mathbb{R}_+^*$. Hence, one can find $\varepsilon^* > 0$ such that :

$$\sup_{\varepsilon<\varepsilon^*} \sup_{i\leq\frac{1}{\sqrt{\varepsilon}}} \left|\frac{b_i^\varepsilon}{a_i^\varepsilon}\right| \leq \delta\, \mathrm{e}^{-x}.$$

This gives us that for ε^* small enough, $\varepsilon < \varepsilon^*$ and $i \leq \frac{1}{\sqrt{\varepsilon}}$:

$$\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \frac{b_i^\varepsilon}{a_i^\varepsilon}\mathrm{e}^x \geq 2\delta - \delta = \delta. \tag{3.25}$$

Now we proceed with the bound on H^ε using Eq.(3.13) and (3.12) :

$$\begin{aligned}
\varepsilon\partial_t H^\varepsilon(t,x) &= (\mathrm{e}^{-x}-1)\sum_{i\geq 0} J_i^\varepsilon(c)\,\mathrm{e}^{-ix} \\
&= (\mathrm{e}^{-x}-1)\left[\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} a_0^\varepsilon c_0^\varepsilon(t) + \sum_{i\geq 1} a_i^\varepsilon\left(\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \frac{b_i^\varepsilon}{a_i^\varepsilon}\mathrm{e}^x\right) c_i^\varepsilon(t)\,\mathrm{e}^{-ix}\right].
\end{aligned}$$

Now we split the sum on the right depending on $\frac{1}{\sqrt{\varepsilon}}$ with ε small enough as before. Note that since $x > 0$, we have that $(\mathrm{e}^{-x}-1) < 0$. The first sum is treated using (3.25) and the bound (3.24) :

$$\varepsilon\partial_t H^\varepsilon(t,x) \le (\mathrm{e}^{-x}-1)\left[2\delta a_0^\varepsilon c_0^\varepsilon(t) + \frac{a(0)}{2}\delta\sum_{i=1}^{\lfloor\frac{1}{\sqrt{\varepsilon}}\rfloor} c_i^\varepsilon(t)\,\mathrm{e}^{-ix} - \mathrm{e}^{x}\sum_{i\ge\lfloor\frac{1}{\sqrt{\varepsilon}}\rfloor+1} b_i^\varepsilon c_i^\varepsilon(t)\,\mathrm{e}^{-ix}\right].$$

The term in c_0^ε and the first sum are combined and using our choice of δ, it yields :

$$2\delta a_0^\varepsilon c_0^\varepsilon(t) + \frac{a(0)}{2}\delta\sum_{i=1}^{\lfloor\frac{1}{\sqrt{\varepsilon}}\rfloor} c_i^\varepsilon(t)\,\mathrm{e}^{-ix} \ge \frac{a(0)}{2}\delta\left(H^\varepsilon(t,x) - \sum_{i\ge\lfloor\frac{1}{\sqrt{\varepsilon}}\rfloor+1} c_i^\varepsilon(t)\,\mathrm{e}^{-ix}\right).$$

Hence :

$$\varepsilon\partial_t H^\varepsilon(t,x) \le (1-\mathrm{e}^{-x})\quad\left[\frac{a(0)}{2}\delta\left(-H^\varepsilon(t,x) + \sum_{i\ge\lfloor\frac{1}{\sqrt{\varepsilon}}\rfloor+1} c_i^\varepsilon(t)\,\mathrm{e}^{-ix}\right) + \mathrm{e}^{x}\sum_{i\ge\lfloor\frac{1}{\sqrt{\varepsilon}}\rfloor+1} b_i^\varepsilon c_i^\varepsilon(t)\,\mathrm{e}^{-ix}\right].$$

Observe that for ε small enough depending on x, for all $i \ge \lfloor\frac{1}{\sqrt{\varepsilon}}\rfloor$, we have :

$$(\delta\frac{a(0)}{2} + \mathrm{e}^{x}\, b_i^\varepsilon)\,\mathrm{e}^{-ix} \le K(C_a+C_b)(1+i\varepsilon)\,\mathrm{e}^{-ix} \le K\varepsilon,$$

which leads to :

$$\varepsilon\partial_t H^\varepsilon(t,x) \le \frac{a(0)}{2}\delta(\mathrm{e}^{-x}-1)H^\varepsilon(t,x) + (1-\mathrm{e}^{-x})Km.$$

We conclude by using Grönwall's lemma and $\tilde{K} = \frac{2Km}{\delta a(0)}$ and (3.23) follows immediately. □

A direct consequence of Lemma 3.4.2 is the following refined estimate on c_0^ε which shows that at the limit $\varepsilon \to 0$, the density f^ε vanishes at the boundary, in agreement with the limiting boundary condition (3.11c):

Lemma 3.4.3. *There exist constants $C_1, C_2 > 0$ such that for ε small enough and for all $t \in (0,T]$:*

$$c_0^\varepsilon(t) \le \mathrm{e}^{-\frac{C_1}{\varepsilon}t}\, c_0^{\varepsilon,0} + \varepsilon C_2. \tag{3.26}$$

Proof. As in the proof of Lemma 3.4.2, there exists ε small enough such that :

$$\frac{\mathrm{d}c_0^\varepsilon(t)}{\mathrm{d}t} = \frac{1}{\varepsilon}(b_1^\varepsilon c_1^\varepsilon(t) - a_0^\varepsilon\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}c_0^\varepsilon(t)) \le C_b\bar{c_1} - \frac{a(0)}{\varepsilon}\delta c_0^\varepsilon(t),$$

thanks to hypothesis (H'1). Now applying Grönwall's lemma, we obtain :

$$c_0^\varepsilon(t) \le \mathrm{e}^{-\frac{a(0)\delta t}{\varepsilon}}\, c_0^\varepsilon(0) + C_b\bar{c_1}\frac{\varepsilon}{\delta a(0)}(1-\mathrm{e}^{-\frac{a(0)}{\varepsilon}\delta t}),$$

which gives the desired result. □

We end this section by a last lemma that will be useful to control the first spatial derivative of F^ε.

Lemma 3.4.4. *We have, for all $T > 0$, and for ε small enough,*

$$\sup_{t\le T}\sum_{i\ge 0}|c_{i+1}^\varepsilon - c_i^\varepsilon|(t) < \infty, \tag{3.27}$$

$$\sup_{t\le T}\varepsilon\sum_{i\ge 0} i|c_{i+1}^\varepsilon - c_i^\varepsilon|(t) < \infty. \tag{3.28}$$

Proof. Let $u_i = c_{i+1}^\varepsilon - c_i^\varepsilon$ and let's estimate its time derivative. Then, for all $i \ge 1$, we have from Eq.(3.13) and Eq.(3.12):

$$\begin{aligned}
\frac{\mathrm{d}u_i}{\mathrm{d}t} =& \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}\left(\frac{a_{i-1}^\varepsilon}{\varepsilon}u_{i-1} - \frac{a_{i-1}^\varepsilon - a_i^\varepsilon}{\varepsilon}c_i^\varepsilon - \frac{a_i^\varepsilon}{\varepsilon}u_i - \frac{a_{i+1}^\varepsilon - a_i^\varepsilon}{\varepsilon}c_{i+1}^\varepsilon\right)\\
&- \left(\frac{b_i^\varepsilon}{\varepsilon}u_i - \frac{b_{i+1}^\varepsilon + b_i^\varepsilon}{\varepsilon}c_{i+1}^\varepsilon - \frac{b_{i+1}^\varepsilon}{\varepsilon}u_{i+1} + \frac{b_{i+1}^\varepsilon - b_{i+2}^\varepsilon}{\varepsilon}c_{i+2}^\varepsilon\right)\\
=& \frac{J_{i-1}^\varepsilon(u,L^\varepsilon) - J_i^\varepsilon(u,L^\varepsilon)}{\varepsilon} + \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}\left(\frac{a_{i-1}^\varepsilon - a_i^\varepsilon}{\varepsilon}u_i - \left(\frac{a_{i-1}^\varepsilon - a_i^\varepsilon}{\varepsilon} + \frac{a_{i+1}^\varepsilon - a_i^\varepsilon}{\varepsilon}\right)c_{i+1}^\varepsilon\right)\\
&+ \frac{b_{i+1}^\varepsilon - b_i^\varepsilon}{\varepsilon}u_{i+1} + \left(\frac{b_{i+1}^\varepsilon - b_i^\varepsilon}{\varepsilon} + \frac{b_{i+1}^\varepsilon - b_{i+2}^\varepsilon}{\varepsilon}\right)c_{i+2}^\varepsilon.
\end{aligned}$$

We multiply the previous expression for $i \ge 1$ by $\mathrm{sign}(u_i)$ on both sides, which gives :

$$\begin{aligned}
\frac{\mathrm{d}|u_i|}{\mathrm{d}t} \le& \frac{J_{i-1}^\varepsilon(|u|,L^\varepsilon) - J_i^\varepsilon(|u|,L^\varepsilon)}{\varepsilon} + \frac{|a_{i-1}^\varepsilon - a_i^\varepsilon|}{\varepsilon}|u_i| + \frac{|a_{i-1}^\varepsilon - 2a_i^\varepsilon + a_{i+1}^\varepsilon|}{\varepsilon}c_{i+1}^\varepsilon\\
&+ \frac{|b_{i+1}^\varepsilon - b_i^\varepsilon|}{\varepsilon}|u_{i+1}| + \frac{|b_{i+2}^\varepsilon - 2b_{i+1}^\varepsilon + b_i^\varepsilon|}{\varepsilon}c_{i+2}^\varepsilon.
\end{aligned}$$

Hence, thanks to hypotheses (H3) and (H5) and Lemma 3.4.2, there exists ε small enough such that for $i \ge 1$:

$$\frac{\mathrm{d}|u_i|}{\mathrm{d}t} \le \frac{J_{i-1}^\varepsilon(|u|,L^\varepsilon) - J_i^\varepsilon(|u|,L^\varepsilon)}{\varepsilon} + \|a'\|_\infty|u_i| + \varepsilon\|a''\|_\infty c_{i+1}^\varepsilon + \|b'\|_\infty|u_{i+1}| + \varepsilon\|b''\|_\infty c_{i+2}^\varepsilon. \tag{3.29}$$

Now, for $i = 0$, we obtain :

$$\frac{\mathrm{d}u_0}{\mathrm{d}t} = -\frac{1}{\varepsilon}J_0^\varepsilon(u,L^\varepsilon) - \frac{a_1^\varepsilon - a_0^\varepsilon}{\varepsilon}\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}c_1^\varepsilon + \frac{a_0^\varepsilon}{\varepsilon}\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}c_0^\varepsilon + \frac{b_2^\varepsilon - b_1^\varepsilon}{\varepsilon}c_2^\varepsilon - \frac{b_1^\varepsilon}{\varepsilon}c_1^\varepsilon.$$

Since $b_0^\varepsilon = b(0) = 0$, we can treat the remaining terms in b as before by adding and removing the right terms for free. The terms with c_1^ε are bounded using Lemma 3.4.2 and the one with c_0^ε using Lemma 3.4.3. Hence, there exists ε small enough such that :

$$\frac{\mathrm{d}|u_0|}{\mathrm{d}t} \le -\frac{1}{\varepsilon}J_0^\varepsilon(|u|,L^\varepsilon) + \|a'\|_\infty\bar{c_1} + \frac{a(0)}{\varepsilon}\,\mathrm{e}^{-\frac{C_1}{\varepsilon}t}\,c_0^{\varepsilon,0} + a(0)C_2 + \|b'\|_\infty|u_1| + \varepsilon\|b''\|_\infty c_2^\varepsilon.$$

We sum the previous estimates for all $i \ge 0$ and we get :

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i\ge 0}|u_i| \le (\|a'\|_\infty + \|b'\|_\infty)\sum_{i\ge 0}|u_i| + (\|a''\|_\infty + \|b''\|_\infty)\varepsilon\sum_{i\ge 0}c_i^\varepsilon + \|a'\|_\infty\bar{c_1} + \frac{a(0)}{\varepsilon}\,\mathrm{e}^{-\frac{C_1}{\varepsilon}t}\,c_0^{\varepsilon,0} + a(0)C_2.$$

We integrate the previous inequality over $[0,t]$, for $0<t<T$:

$$\sum_{i\geq 0}|u_i|(t) \leq \sum_{i\geq 0}|u_i|(0) + (\|a'\|_\infty + \|b'\|_\infty)\int_0^t \sum_{i\geq 0}|u_i|(s)\mathrm{d}s \\ + (\|a''\|_\infty + \|b''\|_\infty)mT + \frac{a(0)}{C_1}c_0^{\varepsilon,0} + (a(0)C_2 + \|a'\|_\infty \bar{c_1})T.$$

And finally Grönwall's lemma yields :

$$\sum_{i\geq 0}|u_i|(t) \leq C_u(T)$$

with

$$C_u(T) = \left(\sum_{i\geq 0}|u_i|(0) + (\|a''\|_\infty + \|b''\|_\infty)mT + \frac{a(0)}{C_1}c_0^{\varepsilon,0} + (a(0)C_2 + \|a'\|_\infty \bar{c_1})T\right) \\ \times \exp((\|a'\|_\infty + \|b'\|_\infty)T),$$

which gives Eq. (3.27).

Using the definition (3.13c) of λ, estimate (3.29) and hypothesis (H'1), we obtain the following inequalities :

$$\begin{aligned}\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i\geq 1} i|u_i| &\leq \varepsilon\sum_{i\geq 1} i\frac{J_{i-1}^\varepsilon(|u|) - J_i^\varepsilon(|u|)}{\varepsilon} + \varepsilon(\|a'\|_\infty + \|b'\|_\infty)\sum_{i\geq 1} i|u_i| + (\|a''\|_\infty + \|b''\|_\infty)\varepsilon^2\sum_{i\geq 0} ic_i^\varepsilon \\ &\leq \sum_{i\geq 0} J_i^\varepsilon(|u|) + (\|a'\|_\infty + \|b'\|_\infty)\varepsilon\sum_{i\geq 1} i|u_i| + (\|a''\|_\infty + \|b''\|_\infty)\lambda \\ &\leq C_a\sum_{i\geq 0}|u_i| + (\|a'\|_\infty + \|b'\|_\infty)\varepsilon\sum_{i\geq 1} i|u_i| + (\|a''\|_\infty + \|b''\|_\infty)\lambda.\end{aligned}$$

Integrating over $[0,t]$ and using the previous bound on $\sum_{i\geq 0}|u_i|$, we conclude using Grönwall's lemma :

$$\varepsilon\sum_{i\geq 1} i|u_i|(t) \leq \left(\varepsilon\sum_{i\geq 1} i|u_i|(0) + C_aC_u(T)T + (\|a''\|_\infty + \|b''\|_\infty)\lambda T\right)\exp((\|a'\|_\infty + \|b'\|_\infty)T),$$

which yields Eq. (3.28). □

3.4.2 Proof of theorem 3.4.1

In this section, we make use of the lemmas from the previous section. As such, from then on, ε is taken small enough to apply those lemmas. We first derive the equation satisfied by the tail distribution F^ε defined at Eq.(3.20). Recall that operator Δ_ε is defined at Eq.(3.19).

Lemma 3.4.5. *For all $x \geq \frac{\varepsilon}{2}$ and $t \geq 0$:*

$$\partial_t F^\varepsilon(t,x) = -\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x}(a^\varepsilon(y)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - a(x)\frac{L(t)}{L(t)+\kappa})f^\varepsilon(t,y)\mathrm{d}y - a(x)\frac{L(t)}{L(t)+\kappa}\Delta_{-\varepsilon}F^\varepsilon(t,x) + \frac{1}{\varepsilon}\int_{x}^{x+\varepsilon}(b^\varepsilon(y)-b(x))f^\varepsilon(t,y)\mathrm{d}y + b(x)\Delta_\varepsilon F^\varepsilon(t,x) \quad (3.30)$$

and for all $x < \frac{\varepsilon}{2}$ and $t \geq 0$:

$$\partial_t F^\varepsilon(t,x) = \frac{1}{\varepsilon}\int_{-\frac{\varepsilon}{2}}^{x} a^\varepsilon(y)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}f^\varepsilon(t,y)\mathrm{d}y - \frac{1}{\varepsilon}\int_{\frac{\varepsilon}{2}}^{x+\varepsilon} b^\varepsilon(y)f^\varepsilon(t,y)\mathrm{d}y. \quad (3.31)$$

Remark. *The function f^ε is defined on $[-\frac{\varepsilon}{2},+\infty[$ whereas f is defined on $\mathbb{R}_+$. However we will only concern ourselves with $x \in \mathbb{R}_+$ in the following subsections. Hence we will treat the case $x < \varepsilon/2$ independently to accommodate for boundary terms that might be left off from f^ε. We also point out that :*

$$\int_0^{+\infty} f^\varepsilon(t,x)\mathrm{d}x = m - \frac{\varepsilon}{2}c_0^\varepsilon(t) \quad (3.32)$$

Owing to Lemma 3.4.2, the right hand side is bounded and tends to m as $\varepsilon \to 0$.And for the first order we have an exact computation :

$$L^\varepsilon(t) + \int_0^{+\infty} x f^\varepsilon(t,x)\mathrm{d}x = \lambda \quad (3.33)$$

Proof. For all $x \in \mathbb{R}_+$ and $t \in [0,T]$, it comes directly from the definitions (3.19) and (3.20) that the following equations hold true :

for all $x \geq 0$ and $t \geq 0$,

$$\Delta_\varepsilon F^\varepsilon(t,x) = -\frac{1}{\varepsilon}\int_{x}^{x+\varepsilon} f^\varepsilon(t,y)\mathrm{d}y, \quad (3.34)$$

and for all $x \geq \frac{\varepsilon}{2}$ and $t \geq 0$,

$$\Delta_{-\varepsilon} F^\varepsilon(t,x) = -\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x} f^\varepsilon(t,y)\mathrm{d}y. \quad (3.35)$$

Denote $H_x = \mathbb{1}_{[x,+\infty)}$. First observe that :

$$\Delta_\varepsilon H_x(y) = \frac{1}{\varepsilon}\left(\mathbb{1}_{[x,+\infty)}(y+\varepsilon) - \mathbb{1}_{[x,+\infty)}(y)\right) = \frac{1}{\varepsilon}\mathbb{1}_{[x-\varepsilon,x)}(y),$$

$$\Delta_{-\varepsilon} H_x(y) = \frac{1}{\varepsilon}\mathbb{1}_{[x,x+\varepsilon)}(y).$$

Then we use Proposition 3.3.2 for the Heaviside function. It yields that for all $x \geq \frac{\varepsilon}{2}$:

$$\begin{aligned}
\partial_t F^\varepsilon(t,x) &= \int_{\mathbb{R}_+} H_x(y)\partial_t f^\varepsilon(t,y)\mathrm{d}y \\
&= \int_{\mathbb{R}_+} \left(\Delta_\varepsilon H_x(y)a^\varepsilon(y)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \Delta_{-\varepsilon}H_x(y)b^\varepsilon(y)\right) f^\varepsilon(t,y)\mathrm{d}y \\
&= \int_{x-\varepsilon}^{x} \frac{1}{\varepsilon}a^\varepsilon(y)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} f^\varepsilon(t,y)\mathrm{d}y - \int_x^{x+\varepsilon}\frac{1}{\varepsilon}b^\varepsilon(y)f^\varepsilon(t,y)\mathrm{d}y \\
&= \frac{1}{\varepsilon}\int_{x-\varepsilon}^{x}(a^\varepsilon(y)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - a(x)\frac{L(t)}{L(t)+\kappa})f^\varepsilon(t,y)\mathrm{d}y - a(x)\frac{L(t)}{L(t)+\kappa}\Delta_{-\varepsilon}F^\varepsilon(t,x) \\
&- \frac{1}{\varepsilon}\int_x^{x+\varepsilon}(b^\varepsilon(y)-b(x))f^\varepsilon(t,y)\mathrm{d}y + b(x)\Delta_\varepsilon F^\varepsilon(t,x)
\end{aligned}$$

The case for $x < \frac{\varepsilon}{2}$ follows from simple computation, using that $\int_{-\varepsilon/2}^{+\infty} f^\varepsilon(t,x)\mathrm{d}x = m$ by conservation of the moment (3.14) :

$$\begin{aligned}
\partial_t F^\varepsilon(t,x) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\frac{-\varepsilon}{2}}^{+\infty} f^\varepsilon(t,y)\mathrm{d}y - \frac{\mathrm{d}}{\mathrm{d}t}\int_{\frac{-\varepsilon}{2}}^{x} f^\varepsilon(t,y)\mathrm{d}y = -(x+\frac{\varepsilon}{2})\frac{\mathrm{d}}{\mathrm{d}t}c_0^\varepsilon(t) \\
&= \frac{1}{\varepsilon}\int_{\frac{-\varepsilon}{2}}^{x} a^\varepsilon(y)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}f^\varepsilon(t,y)\mathrm{d}y - \frac{1}{\varepsilon}\int_{\frac{\varepsilon}{2}}^{x+\varepsilon} b^\varepsilon(y)f^\varepsilon(t,y)\mathrm{d}y.
\end{aligned}$$

□

We then derive an upper bound for the time derivative of $\int_{\mathbb{R}_+}|E(t,x)|\mathrm{d}x$, where E is defined at Eq.(3.21).

Lemma 3.4.6. *For all $t \geq 0$, we have :*

$$\begin{aligned}
\frac{d}{\mathrm{d}t}\int_{\mathbb{R}_+}|E(t,x)|\mathrm{d}x \leq (\|a'\|_\infty + \|b'\|_\infty)\int_{\mathbb{R}_+}|E(t,x)|\mathrm{d}x + \int_{\mathbb{R}_+}|\partial_t F^\varepsilon(t,x) + v(x,L(t))\partial_x F^\varepsilon(t,x)|\mathrm{d}x \\
+ \frac{\varepsilon}{2}a(0)\frac{L(t)}{L(t)+\kappa}c_0^\varepsilon(t).
\end{aligned} \tag{3.36}$$

Proof. From the definitions of the tail distributions in Eq. (3.20), the following equations hold true :

$$\partial_x F(t,x) = -f(t,x)\,,$$

$$\partial_x F^\varepsilon(t,x) = -f^\varepsilon(t,x), \text{ a.e. in } \mathbb{R}_+\,.$$

By Def.3.3.3, we have :

$$F(t,x) = \int_{\max(x,X_c(t))}^{+\infty} f^0(X(0;t,y))\partial_y X(0;t,y)\mathrm{d}y = \begin{cases} \int_0^{+\infty} f^0(y)\mathrm{d}y & \text{if } x \leq X_c(t), \\ \int_{X(0;t,x)}^{+\infty} f^0(y)\mathrm{d}y & \text{if } x \geq X_c(t). \end{cases}$$

Therefore if $x \leq X_c(t)$, then $\partial_t F(t,x) = 0 = \partial_x F(t,x)$. And if $x \geq X_c(t)$, the following expressions hold :

$$\partial_t F(t,x) = -f^0(X(0;t,x))\partial_t X(0;t,x),$$

$$\partial_x F(t,x) = -f^0(X(0;t,x))\partial_x X(0;t,x) = -f(t,x).$$

By properties of characteristics we have : $\partial_t X(0;t,x) + v(x,L)\partial_x X(0;t,x) = 0$ and thus :

$$\partial_t F(t,x) + v(x,L(t))\partial_x F(t,x) = -f^0(X(0;t,x))(\partial_t X(0;t,x) + v(x,L)\partial_x X(0;t,x)) = 0.$$

We then compute :

$$\begin{aligned}\partial_t E(t,x) &= \partial_t F^\varepsilon(t,x) - \partial_t F(t,x)\\ &= -v(x,L(t))\partial_x(F^\varepsilon - F)(t,x) + \partial_t F^\varepsilon(t,x) + v(x,L(t))\partial_x F^\varepsilon(t,x).\end{aligned}$$

We integrate the previous equality, we use the definition (3.3) of v with hypothesis (H3) and we find :

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}_+}|E|\mathrm{d}x &= -\int_{\mathbb{R}_+} v\partial_x|E|\mathrm{d}x + \int_{\mathbb{R}_+}\mathrm{sign}(E)(\partial_t F^\varepsilon + v(x,L(t))\partial_x F^\varepsilon)\mathrm{d}x\\ &= -[v(x,L(t))|E(t,x)|]_0^{+\infty} + \int_{\mathbb{R}_+}\partial_x v|E|\mathrm{d}x + \int_{\mathbb{R}_+}\mathrm{sign}(E)(\partial_t F^\varepsilon + v\partial_x F^\varepsilon)\mathrm{d}x\\ &\leq v(0,L(t))|E(t,0)| + (\|a'\|_\infty + \|b'\|_\infty)\int_{\mathbb{R}_+}|E|\mathrm{d}x + \int_{\mathbb{R}_+}|\partial_t F^\varepsilon + v\partial_x F^\varepsilon|\mathrm{d}x\\ &= (\|a'\|_\infty + \|b'\|_\infty)\int_{\mathbb{R}_+}|E|\mathrm{d}x + \int_{\mathbb{R}_+}|\partial_t F^\varepsilon + v\partial_x F^\varepsilon|\mathrm{d}x + \frac{\varepsilon}{2}a(0)\frac{L(t)}{L(t)+\kappa}c_0^\varepsilon(t).\end{aligned}$$

The last equality is obtained since $b(0) = 0$ and $|E(t,0)| = |\int_{\mathbb{R}_+}(f^\varepsilon - f)(t,x)\mathrm{d}x| = \frac{\varepsilon}{2}c_0^\varepsilon(t)$, see Eq.(3.32). □

Thanks to the equation on F^ε given by Lemma 3.4.5, we control the second term in Eq. (3.36) in the next lemma.

Lemma 3.4.7. *There exist some constants $C_1, C_2 > 0$ independent of ε such that for all $t \in [0,T]$:*

$$\int_{\mathbb{R}_+}|(\partial_t F^\varepsilon + v\partial_x F^\varepsilon)(t,x)|\mathrm{d}x \leq \varepsilon C_1 + C_2|L^\varepsilon(t) - L(t)|. \tag{3.37}$$

Proof. First by construction of both a^ε and b^ε, and the fact that a and b are lipshitz continuous, one has for all $x, y \in \mathbb{R}_+$ such that $|y - x| \leq \varepsilon$:

$$|\frac{a^\varepsilon(y) - a(x)}{\varepsilon}| \leq 2\|a'\|_\infty$$

and similarly for b^ε.

Then using equation (3.30) and definition (3.3) of v, we find the following estimate for all $x \geq \frac{\varepsilon}{2}$:

$$|\partial_t F^\varepsilon + v\partial_x F^\varepsilon| \le 2\|a'\|_\infty \int_{x-\varepsilon}^{x} f^\varepsilon + \frac{1}{\varepsilon}\frac{a(x)}{\kappa}|L^\varepsilon(t) - L(t)| \int_{x-\varepsilon}^{x} f^\varepsilon + 2\|b'\|_\infty \int_{x}^{x+\varepsilon} f^\varepsilon$$
$$+ a(x)\frac{L(t)}{L(t)+\kappa}|\partial_x F^\varepsilon - \Delta_{-\varepsilon} F^\varepsilon| + b(x)|\partial_x F^\varepsilon - \Delta_\varepsilon F^\varepsilon|.$$

Using equation (3.31), hypotheses (H'1) and (H2b) and Lemma 3.4.2, we find for all $x < \frac{\varepsilon}{2}$:

$$\begin{aligned}
|\partial_t F^\varepsilon + v\partial_x F^\varepsilon| &\le |\frac{1}{\varepsilon}\int_{-\frac{\varepsilon}{2}}^{x} a^\varepsilon(y)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} f^\varepsilon(t,y)\mathrm{d}y - a(x)\frac{L(t)}{L(t)+\kappa} f^\varepsilon(t,x)| \\
&\quad + |\frac{1}{\varepsilon}\int_{\frac{\varepsilon}{2}}^{x+\varepsilon} b^\varepsilon(y) f^\varepsilon(t,y)\mathrm{d}y - b(x) f^\varepsilon(t,x)| \\
&\le |\frac{x+\frac{\varepsilon}{2}}{\varepsilon} a(0)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} c_0^\varepsilon(t) - a(x)\frac{L(t)}{L(t)+\kappa} c_0^\varepsilon(t)| + |\frac{x+\frac{\varepsilon}{2}}{\varepsilon} b_1^\varepsilon c_1^\varepsilon(t) - b(x)c_0^\varepsilon(t)| \\
&\le a(0)c_0^\varepsilon(t)\frac{x+\frac{\varepsilon}{2}}{\varepsilon}\frac{1}{\kappa}|L^\varepsilon(t) - L(t)| + |a(x) - \frac{x+\frac{\varepsilon}{2}}{\varepsilon}a(0)|c_0^\varepsilon(t) + C_b\varepsilon c_1^\varepsilon(t) + C_b\frac{\varepsilon}{2}c_0^\varepsilon(t) \\
&\le a(0)\bar{c_0}\frac{1}{\kappa}|L^\varepsilon(t) - L(t)| + (a(0) + \frac{\varepsilon}{2}\|a'\|_\infty)\bar{c_0} + C_b\bar{c_1}\varepsilon + C_b\frac{\varepsilon}{2}\bar{c_0}.
\end{aligned}$$

Now we can integrate $|\partial_t F^\varepsilon + v\partial_x F^\varepsilon|$ over $\mathbb{R}_+$ using the two previous estimates. Note that using Fubini's theorem and Eq.(3.32), we have

$$\begin{aligned}
\int_{\frac{\varepsilon}{2}}^{+\infty}\int_{x-\varepsilon}^{x} f^\varepsilon(y)\mathrm{d}y\mathrm{d}x &= \int_{-\frac{\varepsilon}{2}}^{+\infty} f^\varepsilon(t,y)\int_{\max(y,\varepsilon/2)}^{y+\varepsilon}\mathrm{d}x\mathrm{d}y = \varepsilon\int_{\frac{\varepsilon}{2}}^{+\infty} f^\varepsilon(t,y)\mathrm{d}y + \frac{\varepsilon^2}{2}c_0^\varepsilon(t) \\
&= \int_0^{+\infty} f^\varepsilon(t,y)\mathrm{d}y \le \varepsilon m.
\end{aligned}$$

Therefore, we get :

$$\begin{aligned}
\int_{\mathbb{R}_+} |\partial_t F^\varepsilon + v\partial_x F^\varepsilon|\mathrm{d}x &\le \varepsilon(2\|a'\|_\infty + 2\|b'\|_\infty|)m + \frac{\|a\|_\infty}{\kappa}|L^\varepsilon(t) - L(t)|m \\
&\quad + \|a\|_\infty\frac{L(t)}{L(t)+\kappa}\int_{\frac{\varepsilon}{2}}^{+\infty}|\partial_x F^\varepsilon - \Delta_{-\varepsilon}F^\varepsilon|\mathrm{d}x + \int_{\frac{\varepsilon}{2}}^{+\infty} b(x)|\partial_x F^\varepsilon - \Delta_\varepsilon F^\varepsilon|\mathrm{d}x \\
&\quad + \frac{\varepsilon}{2}(a(0)\bar{c_0}\frac{1}{\kappa}|L^\varepsilon(t) - L(t)| + (a(0) + \frac{\varepsilon}{2}\|a'\|_\infty)\bar{c_0} + C_b\bar{c_1}\varepsilon + C_b\frac{\varepsilon}{2}\bar{c_0}).
\end{aligned} \tag{3.38}$$

We now compute the difference between the continuous and discrete derivatives on F^ε. We denote by $\lfloor x \rceil$ the nearest integer function with the upper-rounding convention : $\lfloor 0.5 \rceil = 1$. Figure 3.2 shows a representation of the cells Γ_{i-1}^ε, Γ_i^ε and Γ_{i+1}^ε as well as an example of the result of $\lfloor \frac{x}{\varepsilon} \rceil$ for $x \in \Gamma_i^\varepsilon$.

Let us first compute the integral of $|\partial_x F^\varepsilon - \Delta_{-\varepsilon}F^\varepsilon|$, using Eq.(3.35) and the fact that f^ε is

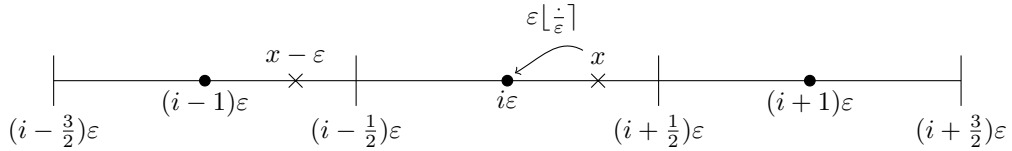


Figure 3.2: Representation of the cells Γ^ε_{i-1}, Γ^ε_i and Γ^ε_{i+1} and of the value $\varepsilon\lfloor\frac{x}{\varepsilon}\rceil$.

constant equal to $c^\varepsilon_i(t)$ on the cells Γ^ε_i :

$$\begin{aligned}\int_{\frac{\varepsilon}{2}}^{+\infty}|\partial_x F^\varepsilon-\Delta_{-\varepsilon}F^\varepsilon|\mathrm{d}x &= \int_{\frac{\varepsilon}{2}}^{+\infty}|f^\varepsilon(t,x)-\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x}f^\varepsilon(t,y)\mathrm{d}y|\mathrm{d}x\\ &= \int_{\frac{\varepsilon}{2}}^{+\infty}\frac{1}{\varepsilon}|\int_{x-\varepsilon}^{x}(f^\varepsilon(t,x)-f^\varepsilon(t,y))\mathrm{d}y|\mathrm{d}x\\ &= \int_{\frac{\varepsilon}{2}}^{+\infty}\frac{1}{\varepsilon}|\int_{x-\varepsilon}^{\varepsilon(\lfloor\frac{x}{\varepsilon}\rceil-\frac{1}{2})}(f^\varepsilon(t,x)-f^\varepsilon(t,y))\mathrm{d}y|\mathrm{d}x\\ &= \int_{\frac{\varepsilon}{2}}^{+\infty}|c^\varepsilon_{\lfloor\frac{x}{\varepsilon}\rceil}(t)-c^\varepsilon_{\lfloor\frac{x}{\varepsilon}\rceil-1}(t)|\frac{\varepsilon(\lfloor\frac{x}{\varepsilon}\rceil-\frac{1}{2})-(x-\varepsilon)}{\varepsilon}\mathrm{d}x.\end{aligned}$$

Now, observe that for all $x\in\Gamma^\varepsilon_i$ one has $c^\varepsilon_{\lfloor\frac{x}{\varepsilon}\rceil}(t)-c^\varepsilon_{\lfloor\frac{x}{\varepsilon}\rceil-1}(t)=c^\varepsilon_i(t)-c^\varepsilon_{i-1}(t)$, which gives :

$$\begin{aligned}\int_{\frac{\varepsilon}{2}}^{+\infty}|\partial_x F^\varepsilon-\Delta_{-\varepsilon}F^\varepsilon|\mathrm{d}x &= \sum_{i\geq 1}|c^\varepsilon_i(t)-c^\varepsilon_{i-1}(t)|\int_{\Gamma^\varepsilon_i}\frac{\varepsilon(\lfloor\frac{x}{\varepsilon}\rceil-\frac{1}{2})-(x-\varepsilon)}{\varepsilon}\mathrm{d}x\\ &= \frac{\varepsilon}{2}\sum_{i\geq 1}|c^\varepsilon_i(t)-c^\varepsilon_{i-1}(t)|.\end{aligned}$$

Hence, according to Lemma 3.4.4 there exists a constant $C(T)>0$ independent of ε such that :

$$\int_{\frac{\varepsilon}{2}}^{+\infty}|\partial_x F^\varepsilon-\Delta_{-\varepsilon}F^\varepsilon|\mathrm{d}x\leq\varepsilon C(T). \tag{3.39}$$

We proceed similarly for the term $\int_{\mathbb{R}_+}b(x)|\partial_x F^\varepsilon-\Delta_\varepsilon F^\varepsilon|\mathrm{d}x$:

$$\begin{aligned}\int_{\frac{\varepsilon}{2}}^{\infty}b(x)|f^\varepsilon(t,x)-\frac{1}{\varepsilon}\int_{x}^{x+\varepsilon}f^\varepsilon(t,y)\mathrm{d}y|\mathrm{d}x &= \int_{\frac{\varepsilon}{2}}^{\infty}\frac{b(x)}{\varepsilon}|\int_{x}^{x+\varepsilon}(f^\varepsilon(t,x)-f^\varepsilon(t,y))\mathrm{d}y|\mathrm{d}x\\ &= \int_{\frac{\varepsilon}{2}}^{\infty}\frac{b(x)}{\varepsilon}|\int_{\varepsilon(\lfloor\frac{x}{\varepsilon}\rceil+\frac{1}{2})}^{x+\varepsilon}(f^\varepsilon(t,x)-f^\varepsilon(t,y))\mathrm{d}y|\mathrm{d}x\\ &= \int_{\frac{\varepsilon}{2}}^{\infty}|c^\varepsilon_{\lfloor\frac{x}{\varepsilon}\rceil+1}(t)-c^\varepsilon_{\lfloor\frac{x}{\varepsilon}\rceil}(t)|b(x)\frac{x-\varepsilon(\lfloor\frac{x}{\varepsilon}\rceil-\frac{1}{2})}{\varepsilon}\mathrm{d}x\\ &= \sum_{i\geq 1}|c^\varepsilon_{i+1}(t)-c^\varepsilon_i(t)|\int_{\Gamma^\varepsilon_i}b(x)\frac{x-\varepsilon(\lfloor\frac{x}{\varepsilon}\rceil-\frac{1}{2})}{\varepsilon}\mathrm{d}x.\end{aligned}$$

Owing to hypothesis (H2b), we simply bound the last integral as follows :

$$\int_{\Gamma_i^\varepsilon} b(x)\frac{x-\varepsilon(\lfloor\frac{x}{\varepsilon}\rceil-\frac{1}{2})}{\varepsilon}\mathrm{d}x \leq C_b\frac{\varepsilon^2}{2}(i+\frac{1}{6}).$$

Hence, according to Lemma 3.4.4, there exists a constant $C(T)>0$ independent of ε such that :

$$\int_{\frac{\varepsilon}{2}}^{+\infty} b(x)|\partial_x F^\varepsilon - \Delta_x F^\varepsilon|\mathrm{d}x < \frac{\varepsilon}{6}C(T)(1+\frac{\varepsilon}{2}). \tag{3.40}$$

We conclude from Eqs. (3.38)-(3.39)-(3.40) by regrouping together terms not depending on ε.

□

We now proceed with the proof of Theorem 3.4.1.

Proof of Theorem 3.4.1. Let $T>0$ and consider $t\in(0,T]$. We begin by integrating (3.36) over $[0,t]$, using Lemma 3.4.7 :

$$\begin{aligned}\int_{\mathbb{R}_+}|E(t,x)|\mathrm{d}x &\leq \int_{\mathbb{R}_+}|E(0,x)|\mathrm{d}x + (\|a'\|_\infty+\|b'\|_\infty)\int_0^t\int_{\mathbb{R}_+}|E(s,x)|\mathrm{d}x\mathrm{d}s\\ &+\int_0^t\int_{\mathbb{R}_+}|\partial_t F^\varepsilon(s,x)+v(x,L(s))\partial_x F^\varepsilon(s,x)|\mathrm{d}x\mathrm{d}s+\frac{\varepsilon}{2}a(0)T\bar{c_0}\\ &\leq \int_{\mathbb{R}_+}|E(0,x)|\mathrm{d}x + (\|a'\|_\infty+\|b'\|_\infty)\int_0^t\int_{\mathbb{R}_+}|E(s,x)|\mathrm{d}x\mathrm{d}s\\ &+\varepsilon TC_1 + C_2\int_0^t|L^\varepsilon(s)-L(s)|\mathrm{d}s+\frac{\varepsilon}{2}a(0)T\bar{c_0}.\end{aligned}$$

Then observe that :

$$\int_{\mathbb{R}_+} xf^\varepsilon(t,x)dx = \int_{\mathbb{R}_+} x\sum_{i\geq 0}\mathbb{1}_{\Gamma_i^\varepsilon}(x)c_i^\varepsilon(t)\mathrm{d}x = \sum_{i\geq 0}\int_{\Gamma_i^\varepsilon} x\mathrm{d}x c_i^\varepsilon(t) = \sum_{i\geq 0} i\varepsilon^2 c_i^\varepsilon(t).$$

Using conservation equations (3.13c) and (3.7) and Fubini's theorem, this leads to the bound :

$$|L^\varepsilon(t)-L(t)| = |\int_0^\infty x(f^\varepsilon(t,x)-f(t,x))\mathrm{d}x| = |\int_0^\infty (F^\varepsilon(t,x)-F(t,x))\mathrm{d}x| \leq \int_{\mathbb{R}_+}|E(t,x)|\mathrm{d}x,$$

which finally yields :

$$\begin{aligned}|L^\varepsilon(t)-L(t)|+\int_{\mathbb{R}_+}|E(t,x)|\mathrm{d}x &\leq 2\int_{\mathbb{R}_+}|E(0,x)|\mathrm{d}x + 2(\|a'\|_\infty+\|b'\|_\infty)\int_0^t\int_{\mathbb{R}_+}|E(s,x)|\mathrm{d}x\mathrm{d}s\\ &+2\varepsilon TC_1+2C_2\int_0^t|L^\varepsilon(s)-L(s)|\mathrm{d}s+\varepsilon a(0)T\bar{c_0}.\end{aligned}$$

By the assumption on $E(0,x)$ and Grönwall's lemma, we finally conclude that :

$$|L^\varepsilon(t)-L(t)|+\int_{\mathbb{R}_+}|E(t,x)|\mathrm{d}x \leq \varepsilon\Big(2C_{\text{init}}+2C_1T+a(0)T\bar{c_0}\Big)\exp(2(\|a'\|_\infty+\|b'\|_\infty+C_2)T).$$

□

3.5 Derivation of second order model and stationary solutions

3.5.1 A second-order Lifshitz-Slyozov model with diffusion

Up to this point we have studied a Lifshitz-Slyozov model, that is to say a transport PDE. However, this model leads to stationary solutions which are combinations of Dirac masses centered at the zeros of velocity v. Hence, since we aim at obtaining asymptotically bimodal distributions, we would like to add a diffusion term to our model in order to smooth the stationary solutions. Unfortunately and up to our knowledge, no biological argument can be found to explain such a diffusive term or to give a proper way of deriving it from biological considerations. Nonetheless, one can see this diffusive term as a second order term emerging from the preceding convergence result, see for example [129], [110], [35].

We follow here the derivation of the diffusive term presented in [129] and we use the notation introduced at Section 3.4.

Now, from Proposition 3.3.2, we can add and subtract the appropriate terms in ϕ to get :

$$\begin{aligned}
\int_0^\infty (f^\varepsilon(t,x) - f^\varepsilon(0,x))\phi(x)\mathrm{d}x &= \int_0^t \int_0^\infty (\Delta_\varepsilon \phi(x) a^\varepsilon(x) \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \Delta_{-\varepsilon}\phi(x) b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t \\
&= \int_0^t \int_0^\infty \frac{\phi(x+\varepsilon)-\phi(x-\varepsilon)}{2\varepsilon}(a^\varepsilon(x)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t \\
&+ \frac{\varepsilon}{2}\int_0^t \int_0^\infty \frac{\phi(x+\varepsilon)-2\phi(x)+\phi(x-\varepsilon)}{\varepsilon^2} a^\varepsilon(x)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t \\
&+ \frac{\varepsilon}{2}\int_0^t \int_0^\infty \frac{\phi(x+\varepsilon)-2\phi(x)+\phi(x-\varepsilon)}{\varepsilon^2} b^\varepsilon(x) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t \\
&= \int_0^t \int_0^\infty \bar{\Delta}_\varepsilon \phi(x)(a^\varepsilon(x)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t \\
&+ \frac{\varepsilon}{2}\int_0^t \int_0^\infty \Delta_\varepsilon^2 \phi(x)(a^\varepsilon(x)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} + b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t
\end{aligned}$$

where $\bar{\Delta}_h\phi(x) = \frac{\phi(x+h)-\phi(x-h)}{2h}$ and $\Delta_h^2\phi(x) = \frac{\phi(x+h)-2\phi(x)+\phi(x-h)}{h^2}$.

This leads us to study the PDE (3.17).

3.5.2 Stationary solutions for the second-order Lifshitz-Slyozov model

In this section, we present the stationary solutions of system (3.11) without diffusion and system (3.17) with diffusion. We notice that stationary solutions are very different in nature from one model to the other. Eq.(3.11) does not yield nontrivial smooth stationary functions, and we rather expect stationary solutions to be linear combinations of Dirac masses, located at roots of the asymptotic velocity.

We can compute explicitly stationary solutions of system (3.17), namely :

$$\partial_t g = 0 \Longleftrightarrow \partial_x(vg) - \frac{\varepsilon}{2}\partial_x^2(dg) = 0.$$

Together with boundary conditions (3.16) this leads to stationary solutions denoted by $M_{L_{\text{stat}}}$, depending on stationary $L_{\text{stat}} \in \mathbb{R}_+$ under the form :

$$M_{L_{\text{stat}}}(x) = \frac{C(m, L_{\text{stat}})}{d(x, L_{\text{stat}})} \exp\Big(\frac{2}{\varepsilon}\int_0^x \frac{v(y, L_{\text{stat}})}{d(y, L_{\text{stat}})}\mathrm{d}y\Big), \tag{3.41}$$

where the constant $C(m, L_{\text{stat}})$ is determined in order to satisfy $\int_{\mathbb{R}_+} M_{L_{\text{stat}}}(t, x)\mathrm{d}x = m$, that is to say

$$C(m, L_{\text{stat}}) = \frac{m}{\displaystyle\int_{\mathbb{R}_+} \frac{1}{d(x, L_{\text{stat}})} \exp\Big(\frac{2}{\varepsilon}\int_0^x \frac{v(y, L_{\text{stat}})}{d(y, L_{\text{stat}})}\mathrm{d}y\Big)\mathrm{d}x}$$

and L_{stat} solves the constraint equation

$$L_{\text{stat}} + \int_{\mathbb{R}_+} x M_{L_{\text{stat}}}(x)\mathrm{d}x = \lambda. \tag{3.42}$$

Note that function $\Phi : L \to L + \int_{\mathbb{R}_+} x M_L(x)\mathrm{d}x$ is continuous on $\mathbb{R}_+$. Moreover, straightforward computations show that thanks to expression (3.41) and expressions for a and b that $\Phi(0) = 0$ and $\Phi \underset{L\to+\infty}{\to} +\infty$. Therefore, for all $\lambda \geq 0$, there exists at least one value for L which satisfies Eq.(3.42). Regarding uniqueness of stationary solutions, it would need to prove strict monotonicity of Φ, which is so far an open question. However, we may observe numerically that the application $\Phi : L \to L + \int_0^{x_{\max}} x M_L(x)\mathrm{d}x$ seems strictly non-decreasing, see Figure 3.3.

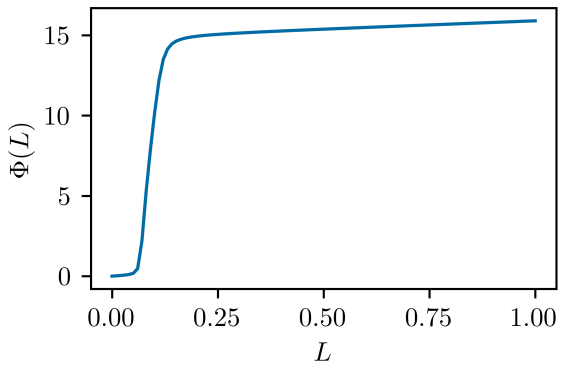


Figure 3.3: Plot of function $L \to \Phi(L)$ for $L \in [10^{-12}, 1]$ with functions a and b defined at Eq.(3.4) and (3.5) and parameters given at Table 3.1.

Remark. *In other modeling contexts, one may choose different functions a and b such that existence of stationary solutions may not be true for all value of λ. For example, $a(x) = 1$ and $b(x) = x^s$ with $s \leq 1$ implies $\lim_{L\to 0^+} \Phi(L) = \lambda_0 > 0$. Hence for values of λ such that $\lambda < \lambda_0$, the system might not have smooth stationary solutions, see section 3.6.2 and Figures 3.14 and 3.15.*

In the following section, we will present some numerical simulations for system (3.17) and control that stationary solutions $M_{L_{\text{stat}}}$ follow a bimodal distribution for well-chosen parameters.

3.6 Numerical simulations

In this part, we use a finite volume well-balanced scheme introduced in [51] to approximate time dependent solutions to Eq. (3.17). Afterwards, we explore numerically the solutions to system (3.17) for various sets of parameters. We also compare the Lifshitz-Slyozov diffusive equation with the transport equation (3.11) and with the transport equation (3.11) with a constant diffusive term. We will finally explore the case when $0 \leq \lambda < \lambda_0 = \lim_{L \to 0^+} \Phi(L)$ mentioned previously in the remark of Sec.3.5.2.

Note that in this section, unlike the previous ones, we are working on a bounded domain $x \in [0, x_{\max}]$ rather than on $\mathbb{R}^+$.

3.6.1 A well-balanced numerical scheme for system (3.17)

In the following, we will need to compute some approximations for the stationary solutions $M_{L_{\text{stat}}}$ since we need them in the well-balanced scheme, see later on. Moreover, it will enable us to compare the asymptotic profiles with the stationary solutions in the numerical tests.

Let us recall that stationary solutions $M_{L_{\text{stat}}}$ are defined by an explicit expression given at Eq. (3.41) with L_{stat} satisfying constraint equation (3.42). Therefore, to compute this stationary solution, a simple dichotomy method is implemented to find the solution to $\Phi(L) = \lambda$, since the application Φ is increasing in the range of L that interests us, see Figure 3.3. We use the trapezoidal rule for the computation of the integrals.

Since we are interested in a conservative PDE, we use a finite volume scheme. We also aim at capturing correctly stationary solutions and for that purpose, we implement a well-balanced scheme introduced in [51]. Let us detail the scheme here.

The scheme is based on a change of variables in the PDE (3.17) to obtain a symmetric operator. This will allow simpler calculations down the line. Denote D_L the spatial operator in the PDE, i.e. :

$$D_L g = \partial_x F(g; x, L) = \partial_x \Big(- v(x, L) g + \partial_x (d(x, L) g) \Big).$$

We recall that the stationary solution associated with the value L is given by :

$$M_L(x) = \frac{C(m, L)}{d(x, L)} \exp\Big(\int_0^x \frac{v(y, L)}{d(y, L)} \mathrm{d}y \Big). \tag{3.43}$$

This stationary solution satisfies $D_L M_L = 0$ and we can rewrite the operator D_L in the following way :

$$D_L g = \partial_x \Big(d(x, L) M_L \partial_x (\frac{g}{M_L}) \Big).$$

Then we perform the change of variable $h = \frac{g}{\sqrt{M_L}}$ and introduce the new operator $\tilde{D}_L$, which is symmetric for the L^2 inner product :

$$\tilde{D}_L h = \frac{1}{\sqrt{M_L}} D_L (h \sqrt{M_L}) = \frac{1}{\sqrt{M_L}} \partial_x \Big(d(x, L) M_L \partial_x (\frac{h}{\sqrt{M_L}}) \Big).$$

Note that we use an implicit discretization in time in order to avoid a constraining time step for the diffusion operator.

Given a mesh of size $\Delta x > 0$ in space, we discretize the interval $[0, x_{\max}]$ and consider N cells $\mathcal{C}_j = [x_{j-1/2}, x_{j+1/2}]$, $1 \leq j \leq N$ centered at point x_j, with $x_j = j \Delta x$ and $x_{j+1/2} = (j+1/2)\Delta x$. We also introduce a time step $\Delta t > 0$ and the discretization times $t_n = n \Delta t$, $n \in \mathbb{N}$.

We denote by h_j^n an approximation of the average of function h on cell $\mathcal{C}_j$ at time t_n, that is to say $h_j^n \sim \frac{1}{\Delta x}\int_{\mathcal{C}_j} h(t_n,x)\mathrm{d}x$. We also define $M_{L^n,j}$ as an approximation of stationary solution M_{L^n} defined at Eq. (3.43) at point x_j with $L = L^n$, and $D_{j+1/2}^n$ as an approximation of diffusion coefficient $d(x_{j+1/2}, L^n)$ at point $x_{j+1/2}$ with $L = L^n$, see expression (3.15).

We denote by $F_{j+1/2}^n$ an approximation of flux $d(x,L)M_L\partial_x(\frac{h}{\sqrt{M_L}})$ at the boundary $x_{j+1/2}$ of cell $\mathcal{C}_j$ at time t_n.

We therefore discretize Eq.(3.17a) as follows :

$$\begin{aligned}\frac{h_{j+1}^n - h_j^n}{\Delta t} &= \frac{1}{\Delta x\sqrt{M_{L^n,j}}}(F_{j+1/2}^n - F_{j-1/2}^n)\\ &= \frac{1}{\Delta x\sqrt{M_{L^n,j}}}\Big(D_{j+1/2}^n\sqrt{M_{L^n,j+1}M_{L^n,j}}\frac{h_{j+1}^{n+1}/\sqrt{M_{L^n,j+1}} - h_j^{n+1}/\sqrt{M_{L^n,j}}}{\Delta x}\\ &\qquad - D_{j-1/2}^n\sqrt{M_{L^n,j}M_{L^n,j-1}}\frac{h_j^{n+1}/\sqrt{M_{L^n,j}} - h_{j-1}^{n+1}/\sqrt{M_{L^n,j-1}}}{\Delta x}\Big).\end{aligned}$$

Regarding boundary conditions, we want to preserve the zeroth-order moment $\int_0^{x_{\max}} g(t,x)\mathrm{d}x = \int_0^{x_{\max}} g^0(x)\mathrm{d}x$, which implies to use the following null-flux boundary conditions :

$$-v(x,L(t))g(t,x) + \frac{\varepsilon}{2}\partial_x(d(x,L(t))g(t,x))|_{x=0,x_{\max}} = 0.$$

In practice, those boundary conditions are implemented using ghost cells centered at points $-\Delta x$ and $x_{\max}+\Delta x$, and by setting the null-flux conditions $F_{-1/2}^n = F_{N+1/2}^n = 0$.

In order to update the value of L, we differentiate Eq.(3.17b) with respect to time and we discretize the equation $\partial_t L = -\int_{\mathbb{R}_+} x\partial_t g(t,x)\mathrm{d}x$, which gives:

$$L^{n+1} = L^n - \Delta x\sum_{i=1}^N x_i(g_i^{n+1} - g_i^n).$$

This update leads to a restriction on the time step Δt to preserve positivity of L^{n+1} as seen in [51].

3.6.2 Numerical results

The previous numerical scheme enables us to explore the properties of system (3.17) as a model for adipocyte distribution evolution in time. Table 3.1 presents the value of most parameters for the simulations. Unless stated otherwise, these parameters shall be fixed for this section. Concerning values of parameters, a few of them are chosen in accordance with biological observations. V_{lipids} and r_0 have fixed given values. The value of γ and β are taken from [119]. Values of other parameters are chosen as to observe bimodal distributions. We refer the reader to chapter 5 for further investigation into the values of those parameters.

Parameter	Value	Unit	Description	Related equation
α	0.7	nmol h^{-1} μm^{-1}	Lipogenesis surface limited flow rate	Eq.(3.4)
ρ	200	μm	Lipogenesis saturation in radius cutoff	Eq.(3.4)
n	3	$\emptyset$	Lipogenesis saturation in radius power	Eq.(3.4)
κ	0.01	$\emptyset$	Lipogenesis saturation in external lipid constant	Eq.(3.4)
β	1	nmol h^{-1}	Lipolysis basal flow rate	Eq.(3.5)
γ	0.27	nmol h^{-1} μm^{-1}	Lipolysis surface limited flow rate	Eq.(3.5)
χ	0.01	$\emptyset$	Lipolysis saturation in internal lipid constant	Eq.(3.5)
V_{lipids}	10^6	μm^3	Molar volume of triglycerides	Eq.(3.1)
r_0	6	μm	Radius of an adipocyte without lipid	Eq.(3.1)
ε	0.05	$\emptyset$	Diffusion scaling parameter	Eq.(3.17a)
$x_{\max}$	15	nmol	Maximal lipid size of an adipocyte	Sec. 3.6.1
N	10^4	$\emptyset$	Number of discretization points	Sec. 3.6.1

Table 3.1: Values of parameters for the model

Asymptotic behaviour of the second order Lifshitz-Slyozov system (3.17)

To begin with, we check that the asymptotic profile obtained with the time evolution of the solution thanks to the previous described scheme coincides with the stationary solution of Sec.3.5.2.

First, one may assume that given an initial condition (g^0, L^0), the asymptotic behaviour of the system is governed by the two parameters m and λ. This means that given two initial conditions (g_1^0, L_1^0) and (g_2^0, L_2^0) such that $m_1 = m_2$ and $\lambda_1 = \lambda_2$, the stationary solutions are equal. In Figure 3.4, both initial conditions are Gaussian functions centered at $x_1 = 1$ and $x_2 = 3$ with $m_1 = m_2$ and initial values L_1^0 and L_2^0 are chosen so that $\lambda_1 = \lambda_2$. We indeed observe that the asymptotic profile is the same for these two initial conditions.

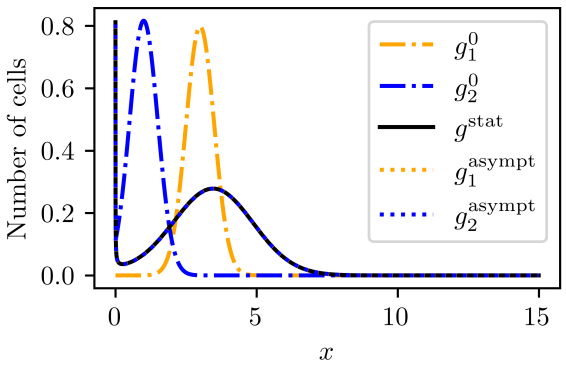


Figure 3.4: Asymptotic profiles g_1^{asympt} and g_2^{asympt}, from two initial conditions (g_1^0, L_1^0) and (g_2^0, L_2^0) such that $m_1 = m_2$ and $\lambda_1 = \lambda_2$. Parameters of the system are given at Table 3.1. Note that the stationary solution and the two asymptotic distributions are superimposed.

Bimodality vs unimodality

Since the main aim of the model we develop and study in this paper is to represent bimodality of the distribution on the stationary solution, we check if we effectively find some parameter

ranges for which we observe this behaviour. In particular, we investigate the dependency with respect to λ. Note that, since λ is defined by expression (3.7), we change λ by changing the initial conditions L^0 and g^0, in the case of time evolution of the system, or by changing the value of L^{stat} when considering stationary solutions.

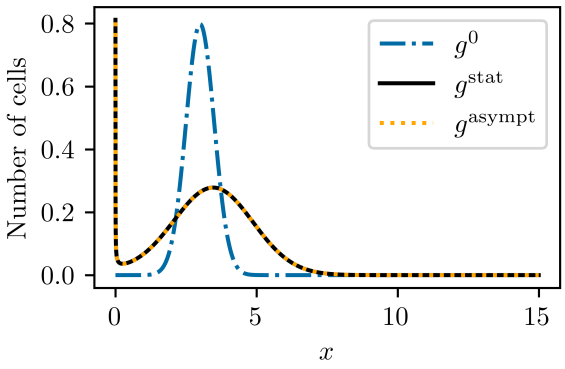


Figure 3.5: Distributions of adipocytes with respect to size, i.e. amount of lipids, starting with $\lambda = 3.5$ and initial distribution $g^0(x) = C \exp\left(-\frac{1}{2}\left(\frac{x-3}{0.5}\right)^2\right)$ (in dashed blue line) : asymptotic profile (dotted yellow line) and stationary solution (black full line) both present bimodality. Parameters of the system are given at Table 3.1.

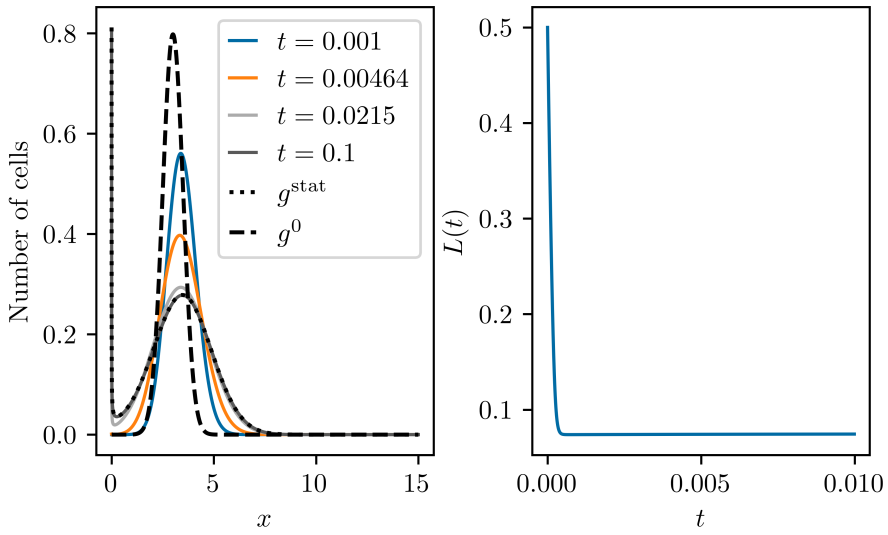


Figure 3.6: On the left : time evolution of the size distribution with respect to size in the bimodal case; on the right : time evolution of the external lipid concentration. We observe that the asymptotic profile coincides with the computed stationary solution. Parameters of the system are given at Table 3.1.

In Figure 3.5, we plot densities of adipocytes as a function of size x. It shows the result of

the scheme starting from a Gaussian initial condition $g^0(x) = C\exp\left(-\frac{1}{2}\left(\frac{x-3}{0.5}\right)^2\right)$ plotted in dashed blue line and L^0 such that $\lambda = 3.5$. The value of C is determined such that $m = 1$. The stationary solution is denoted $g^{\text{stat}} = M_{L_{\text{stat}}}$- in black full line - and the final result of the scheme at time $t = t_{\max}$ is denoted $g(t_{\max}, \cdot)$ and represented in dotted yellow line. $t_{\max}$ is determined such that the relative difference between the size distribution $g(t_{\max}, \cdot)$ and the stationary solution $M_{L_{\text{stat}}}$ is less than 5×10^{-5}. We can observe that bimodality is obtained for the stationary solution as well as for the asymptotic profile of the adipocyte size distribution and that there is a good correspondence between the two functions. Up to some numerical error of order 10^{-12}, both the initial number of cells $m = \int_0^{x_{\max}} g^0(x)\mathrm{d}x$ and the initial amount of lipids λ are conserved, as expected. In Fig.3.6, we plot the time evolution of the solution : on the left, adipocyte density is displayed as a function of x for various times and on the right, the evolution with respect to time of external lipid concentration L is plotted. We observe that L tends to a stationary value and g to a stationary profile with bimodality as expected.

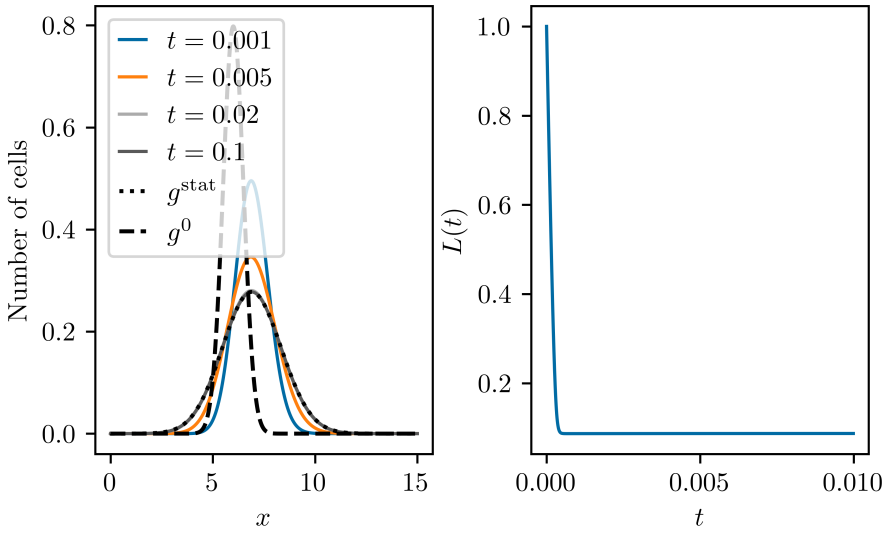


Figure 3.7: On the left : distributions of adipocytes with respect to size, i.e. amount of lipids, starting with $\lambda = 7$ and initial distribution $g^0(x) = C\exp\left(-\frac{1}{2}\left(\frac{x-6}{0.5}\right)^2\right)$ (in dashed blue line). On the right : time evolution of the external lipid concentration. Asymptotic profile (dotted yellow line) and stationary solution (black full line) both present unimodality. Parameters of the system are given at Table 3.1.

Now we may investigate the behaviour of the stationary solutions and the asymptotic profiles with respect to λ. A first crucial information is that depending on λ, different types of modality can be observed. Figure 3.7 presents a case where the stationary solution is unimodal, obtained with initial conditions $g^0(x) = C\exp\left(-\frac{1}{2}\left(\frac{x-6}{0.5}\right)^2\right)$, where C is choosen such that $m = 1$, and L^0 such that $\lambda = 7$. We remark that L also tends to a stationary value and densities converge towards a stationary distribution with unimodality. Biologically, we can relate this to the fact

that if the amount of lipids in the system is higher, cells have a tendency to put into storage the maximum amount of lipids and thus cells are bigger in average. From a mathematical point of view, since the optima can be linked to the velocity zeros, this means that for bigger λ, two of the zeros of speed V - and therefore two optima - disappear and thus only one zero remains giving rise to a unimodal profile.

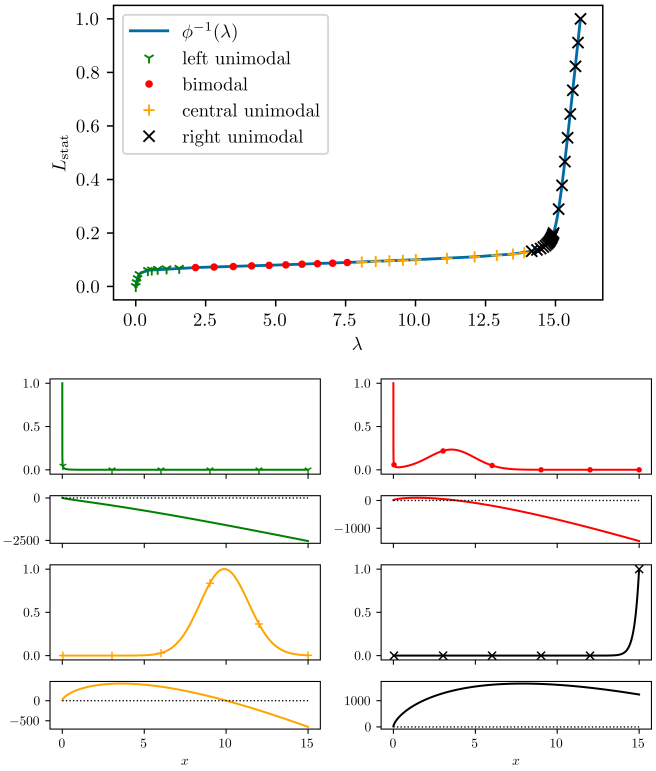


Figure 3.8: On top : plot of function $\lambda \to \Phi^{-1}(\lambda)$, the inverse function of the one displayed at Fig.3.3, and type of modality of the stationary solution with respect to the value of λ. On bottom : Normalized stationary solutions for different values of λ, the stationary velocity as a function of x is represented bellow each solution. Top left : left unimodal; top right : bimodal ; bottom left : central unimodal ; bottom right : left unimodal.

More generally, we can investigate the profile modality with respect to the value of λ using the computation of stationary solutions. In Figure 3.8 on top, we present the type of modality of the stationary solutions as a function of λ. Left (resp. right) unimodality is labeled in green Y (resp. in black x) when a single mode concentrated on the left (resp. right) of the domain is observed. Central unimodal stationary solution is labeled in yellow + when the unique mode is concentrated inside the domain. Bimodality is labeled in red dot.

In Fig.3.8 on bottom, a plot for each of the 4 types of modality is presented. On the top, we represent the stationary solution $M_{L_{\text{stat}}}$ with respect to x and on the bottom, we plot the stationary velocities with respect to x. Four different values of L_{stat} corresponding to various λ are considered, namely $L_{\text{stat}} = 0.05$ and $\lambda = 0.191$ for the left unimodality (top left, in blue), $L_{\text{stat}} = 0.075$ and $\lambda = 3.52$ for the left bimodality (top right, in green), $L_{\text{stat}} = 0.1$ and $\lambda = 9.96$ for the central unimodality (bottom left, in yellow) and $L_{\text{stat}} = 0.2$ and $\lambda = 14.9$ for the left

unimodality (bottom right, in black). For a biological interpretation, left modality is observed when the amount of lipids is too low and thus cells are of relative small sizes. Right modality is a consequence of the amount of lipids being too large and represents the whole cell population approaching its maximal volume. A mathematical interpretation is given by again considering the zeros of the velocity with an influence on the optima of the profile. Left (resp. right) modality is reached when zeros disappear and/or go outside the domain from the left (resp. from the right). The first mode in the bimodal case can also be localized at 0, the smallest zero of the velocity being outside the domain (on the left).

Influence of ε and comparison with a constant diffusion rate D

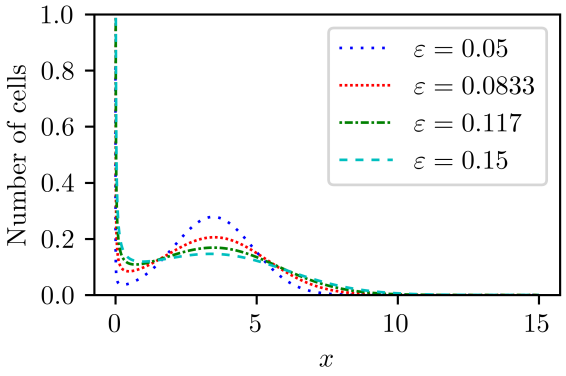


Figure 3.9: Different stationary solutions depending on the value of ε. We observe that bimodality holds for values of ε small enough. Parameters of the system are given at Table 3.1.

In this part, we explore the influence of parameter ε on the shape of stationary solutions. We can observe in Figure 3.9 that higher values of ε smoothen the two maxima of the solution, as expected. For smaller values of ε, the nadir (i.e. the local minimum between the two maxima) gets sharper and for very small ε this may result numerically in taking very small time and space steps. This is easily interpreted as the fact that when $\varepsilon = 0$, we consider the classical Lifshitz-Slyozov system where stationary solutions are sums of Dirac masses which is difficult to obtain numerically without a dedicated scheme.

The choice we made for the diffusion rate is supported by the convergence results from the Becker-Döring to Lifshitz-Syozov model and the behaviour of second order terms. However this choice is not motivated by biological observation. Hence one may make the assumption that the diffusion rate is constant in both time and space. This unfortunately results in quite different results as shown in Figure 3.10. We point out that to obtain bimodality some parameters need to be readjusted in this case. Hence comparing the solutions of the system under consideration (3.17) and the solutions with constant diffusion rate proves to be difficult because the behaviour of stationary solutions is heavily dependent on the choice of parameters.

We still can make a few comments about the resulting solutions. The constant diffusion rate tends to smoothen the first maximum whereas in the non-constant case, the diffusion is relatively close to zero, leading to a sharper maximum. Our investigation of the available data for adipose cell distribution leads us to believe that non-constant diffusion rates have better chances of making the model fit with the data. We also point out that in the case of constant diffusion, each type of modality, as previously described, is obtainable.

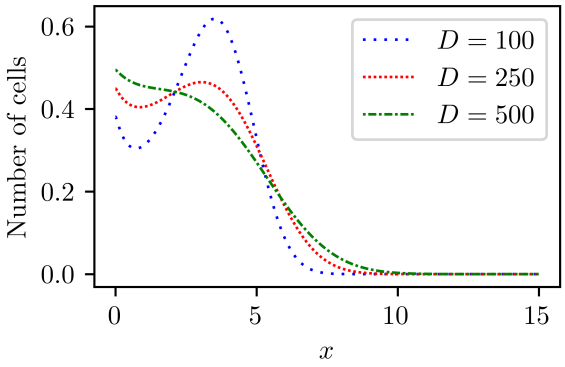


Figure 3.10: Different stationary solutions depending on the value of diffusion rate D taken as constant in space and time. Parameters of the system are given at Table 3.1.

Comparison with the first order model

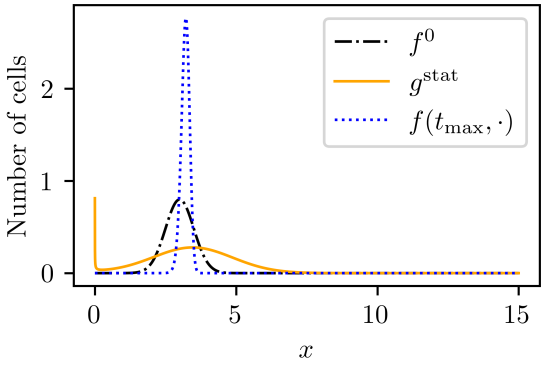


Figure 3.11: Numerical solution for the first order Lifshitz-Slyozov model (3.11) (in dotted blue line) compared to the stationary solution of the Lifshitz-Slyozov diffusive model (3.17) (in orange plain line) with the same parameters and same initial condition (displayed in black dashed -dotted line). The solution to the first order Lifshitz-Slyozov model is expected to converge to a Dirac mass and is displayed for a time before reaching the asymptotic profile.

Stationary solutions for the first order Lifshitz-Slyozov model are not so easily computed theoretically. Nonetheless we can explore these solutions numerically as asymptotic profiles of the solutions of system (3.11). For that purpose, we use a standard upwind scheme for transport equations, since the velocity is known. Figure 3.11 presents the result of an upwind scheme for the Lifshitz-Slyozov model with the same initial conditions and parameters as in Fig. 3.6. We expect singular stationary state for the first order Lifshitz-Slyozov model. We may interpret stationary state that concentrates at two points as a degenerate bimodal solution. Using the same parameters as in Fig. 3.6, we can see on Fig. 3.12 that the solution concentrates to a singular Dirac mass and that in this case we cannot recover bimodality, unlike the case of second-order Lifshitz-Slyozov model, see Fig.3.11. We also point out that the asymptotic values of L are different in both cases.

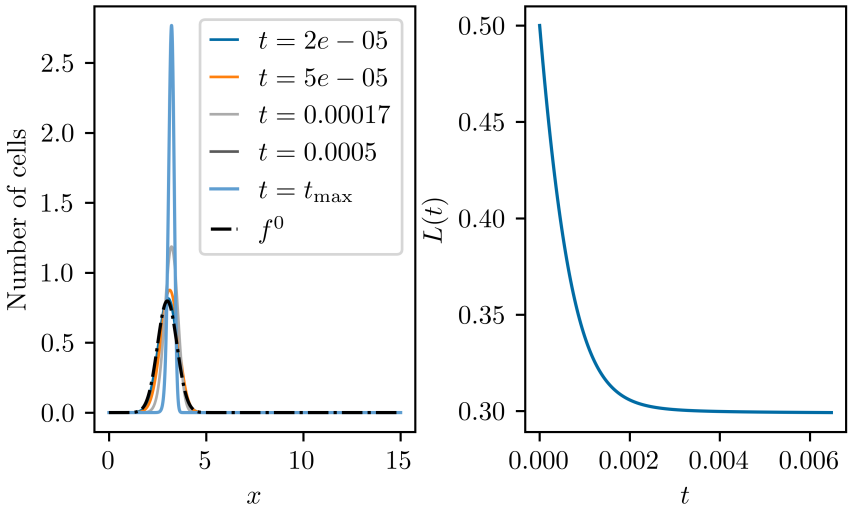


Figure 3.12: Numerical solution for the first order Lifshitz-Slyozov model (3.11) with same parameters and initial data as in Fig. 3.6. On the left : time evolution of the size distribution with respect to size; on the right : time evolution of the external lipid concentration. The solution to the first order Lifshitz-Slyozov model is expected to converge to a Dirac mass and is displayed for a time before reaching the asymptotic profile.

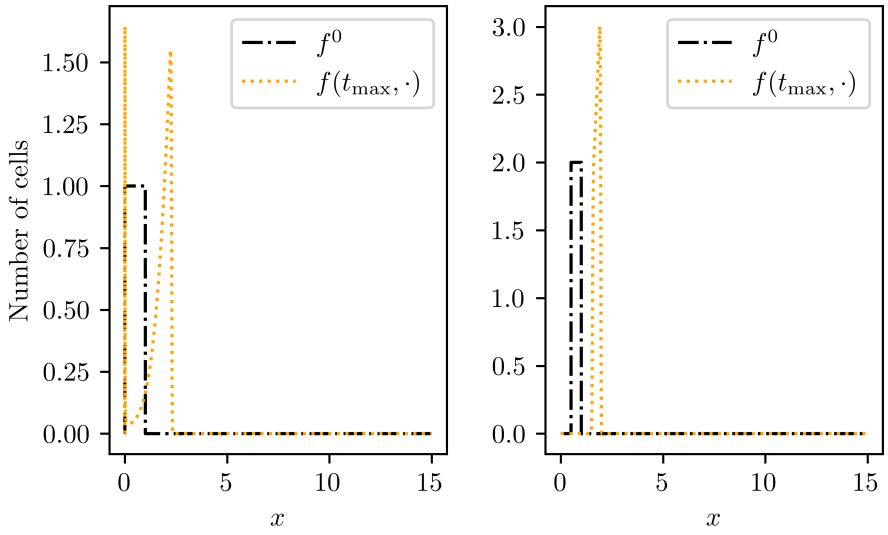


Figure 3.13: Asymptotic profiles for the first order Lifshitz-Slyozov model (3.11) with $m = 1$ and $\lambda = 2$. Left : $f^0(x) = C\mathbb{1}_{[\Delta x,1]}(x)$. Right : $f^0(x) = C\mathbb{1}_{[0.5,1]}(x)$. The difference in the initial conditions leads to different profiles. To observe bimodality the parameter β was changed to $\beta = 100$ in both cases.

By changing initial conditions and the parameter β to $\beta = 100$, we can nonetheless obtain a bimodal solution for the first order Lifshitz-Slyozov model (3.11) as seen in Fig. 3.13 on the left. However, by changing the initial condition f^0, we can see on Fig. 3.13 on the right that we do not obtain the same asymptotic solutions. This leads us to believe that in the case of the first order Lifshitz-Slyozov model the asymptotic solutions depend on the initial condition g^0 and not only on m and λ, unlike for second order Lifshitz-Slyozov model (3.17).

The case $\lambda < \Phi(0)$

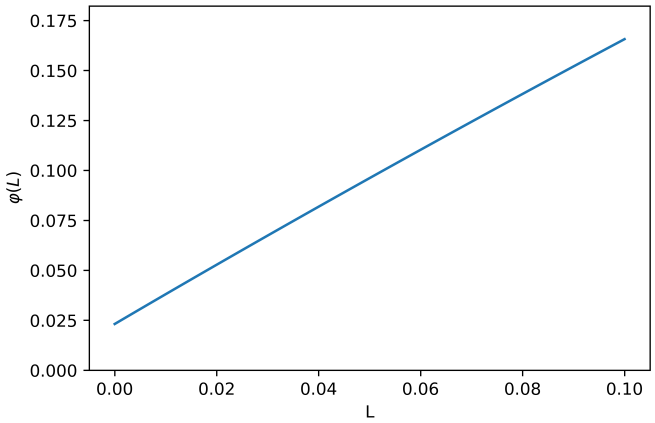


Figure 3.14: Plot of function $L \to \Phi(L)$ with $a(x) = 1$ and $b(x) = (x+1)^{2/3}$. In that case, $\lim_{L \to 0^+} \Phi(L) \sim 0.025 > 0$ and the existence of smooth stationary solutions for values of λ such that $\lambda < \lambda_0$ is not guaranteed.

As explained in the remark of Sec.3.5.2, for different choices of functions a and b than those of the adipocyte model, we may find situations where $\lim_{L \to 0^+} \Phi(L) = \lambda_0 > 0$. In this subsection, we explore the evolution of a solution for a value of λ such that $0 < \lambda < \lambda_0$, that is to say in a case when no smooth stationary solution exists. An example of choice for a and b is $a(x) = 1$ and $b(x) = (x+1)^{2/3}$ and in Fig.3.14, the function $L \to \Phi(L)$ is displayed in that case.

We show in Figure 3.15 the time evolution of the density profile (on the left) and of the external lipid concentration L (on the right) computed numerically in a case where $\lambda < \lambda_0$. We observe that, as expected, L tend to 0 asymptotically and that the adipocyte density seems to converge towards a Dirac mass centered at 0. Numerical simulations prove difficult because of the constraint on Δt to enforce the stability of the numerical scheme. More precisely, this constraint induces that Δt should be bounded above by L^n. Hence as the computation time increases, we observe that the value of L^n tends to zero, as the solution gets closer to the asymptotic profile and therefore that the time step eventually gets smaller than machine precision. In this case, the scheme fails to conserve both λ and m.

3.7 Conclusion

Our work provides a new approach for looking into convergence from Becker-Döring to Lifshitz-Slyozov, and numerical results indicating that the second order Lifshitz-Slyozov model is better

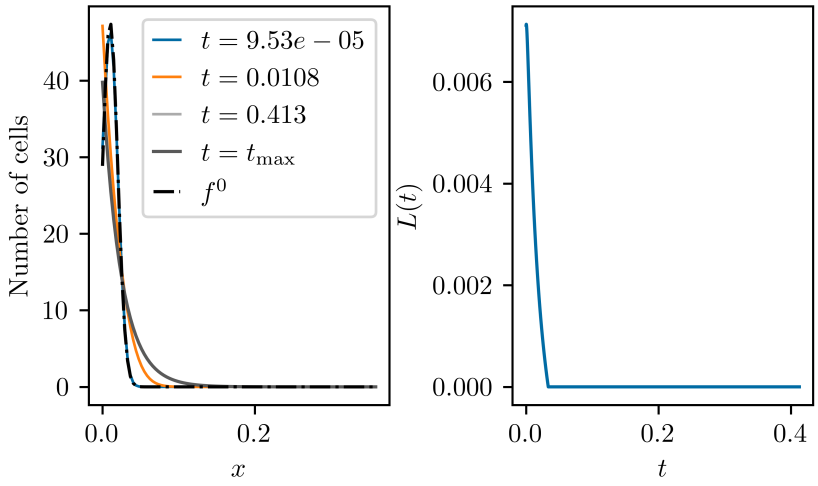


Figure 3.15: Case when $a(x) = 1$ and $b(x) = (x+1)^{2/3}$ and $\lambda < \lambda_0$. On the left : time evolution of the size distribution with respect to size; on the right : time evolution of the external lipid concentration.

suited to model adipocyte size distribution than previous approach relying on first order Lifshitz-Slyozov model.

The originality of this study lies in the following points :

- a new second order Lifshitz-Slyozov model (3.17) for adipocyte size distribution with a diffusion term derived from a discrete model,
- Becker-Döring and Lifshitz-Slyozov systems with an unusual velocity (3.3) -(3.5) with three zeros and a saturation term in L, which leads to different types of stationary solutions,
- an additional conservation law (3.8) with respect to classical systems, enforcing uncommon boundary conditions, see Eq. (3.10) and (3.16),
- a new proof of convergence result from Becker-Döring solutions to Lifshitz-Slyozov solutions, using tails of distributions, that provides an upper bound on the speed of convergence.
- numerical results showing that bimodal distributions, as well as unimodal profiles, can be obtained asymptotically with system (3.17), according to the parameters,
- numerical results exploring the influence of parameter ε and comparing the diffusion term of system (3.17) with a time and space constant coefficient.
- numerical results shows that the second order system (3.17) provides universal asymptotic profile that does not depend on initial condition (but only on λ, m), contrary to first order system (3.11).

We believe that the distribution tail approach could be further investigated to show convergence towards the solutions to the second order Lifshitz-Slyozov equation. The asymptotic behavior of solutions to the second order Lifshitz-Slyozov model will be investigated in future works.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Chapter 4

A stochastic approach to adipose cell modeling

This chapter is dedicated to the study of the two stochastic models : the non-linear Becker-Döring CTMC (1.41) and the Lifshitz-Slyozov SDE (1.45). In this chapter we assume that solutions to both stochastic models exist. We begin by recalling these two equations and we provide some insight on their relation to the deterministic Becker-Döring model and diffusive Lifshitz-Slyozov model. This is done by using Kolmogorov backward equation for the Becker-Döring model and showing that the the law of X_ε is a solution of the Becker-Döring model. Showing a rigorous relation between the Lifshitz-Slyozov SDE and the diffusive Lifshitz-Slyozov SDE is an ongoing work but we detail some heuristic that support this relation. let us also mentioned [55], where mean-field equations are derived for systems of stochastic particles.

Then we prove theorem 1.5.2 where we assume that the rates are bounded. The proof follows the method in [74], but we obtain a bound in L^1-norm which involves additional steps. We then extend this results to the case of sub-linear rates which our initial choice of a and b verifies. The proof involves splitting the L^1-norm depending on the hitting time of some size x. Before this hitting time we proceed exactly as in the bounded case and we use crude bounds for times larger than the hitting time. Then the proof concludes with choosing the size x depending on ε to obtain the final bound.

As mentioned before a different approach for extending the first theorem would be to consider a bounded domain, where cells have a maximal size $x_{\max}$. This is actually what we do in the numerical schemes for the deterministic Lifshitz-Slyozov models. However for stochastic models, if we consider a bounded domain the diffusion process Z_ε is not a good approximation of X_ε. Indeed given a density-dependent Markov chain that visits the boundary with non-negligible probability, then the diffusion approximation is valid only up to the first visit of the boundary, and is ill-posed afterwards. In [14], the authors show that a good approximation of a density-dependent Markov chain for the case of a bounded domain is a jump-diffusion process. This jump-diffusion process is build such that there is a correction of the trajectory when the process hits the boundary and leaves it. However our aim when constructing these stochastic models is that they retain some relation to the deterministic models. By considering a bounded model and the jump-diffusion process from [14], it is unclear how this process relates to the diffusive Lifshitz-Slyozov model.

In the last section of this chapter, we present some numerical simulations of the stochastic models (1.71) and (1.43). As we mentioned in section 1.3.9, we make an heuristic assumption that when N gets large, the uniform propagation of chaos ensures that in some sense $X_\varepsilon^{1,N} \to X_\varepsilon$

and $Z_\varepsilon^{1,N} \to \mathbb{Z}_\varepsilon$. We recall the results from [122] that establishes this convergence for birth and death processes. An analogous result for our non-linear models is an ongoing work.

4.1 Stochastic models

We recall that the Becker-Döring CTMC $X_\varepsilon(t)$ is given by a solution to :

$$X_\varepsilon(t) = X_\varepsilon(0) + \varepsilon Y_+\Big(\varepsilon^{-1}\int_0^t a(X(t))\frac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa}\mathrm{d}s\Big) - \varepsilon Y_-\Big(\varepsilon^{-1}\int_0^t b(X(t))\mathrm{d}s\Big), \tag{4.1}$$

where Y_+ and Y_- are independent unit Poisson processes and $L_\varepsilon^X(t) = \lambda - \mathbb{E}\left[X_\varepsilon(t)\right]$.

The Lifshitz-Slyozov diffusion approximation $Z_\varepsilon(t)$ is given by a solution to :

$$Z_\varepsilon(t) = Z_\varepsilon(0) + \varepsilon B_+\Big(\varepsilon^{-1}\int_0^t a(Z_\varepsilon(s))\frac{L_\varepsilon^Z(s)}{L_\varepsilon^Z(s)+\kappa}\mathrm{d}s\Big) - \varepsilon B_-\Big(\varepsilon^{-1}\int_0^t b(Z_\varepsilon(s))\mathrm{d}s\Big). \tag{4.2}$$

where B_+ and b_- are independent Brownian motions with $\mathbb{E}\left[B_\pm(t)\right] = \mathrm{Var}(B_\pm(t)) = t$ and $L_\varepsilon^Z(t) = \lambda - \mathbb{E}\left[Z_\varepsilon(t)\right]$.

Remark. *In this chapter, we assume that there exists a solution to both equations* (4.1) *and* (4.2).

4.1.1 Relation with the deterministic models

Becker-Döring CTMC and deterministic Becker-Döring model

How do we relate equation (4.1) to the deterministic Becker-Döring model ? The first thing is to construct the transition rate matrix Q. In our case this matrix depends on time and is given by :

$$Q_{i,i-1}(t) = a(i-1)\frac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa}, \quad Q_{i,i}(t) = -(a(i)\frac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa} + b(i)), \quad Q_{i,i+1}(t) = b(i+1),$$
$$i = 1, \ldots, +\infty \tag{4.3}$$

$$Q_{0,0}(t) = -a(0)\frac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa}, \quad Q_{0,1}(t) = b(1). \tag{4.4}$$

Now the law of X is given by the probabilities $\mathbf{P}_j(t) = \sum_{i=0}^{+\infty}\mathbf{P}_{i,j}(t)$, where $\mathbf{P}_{i,j}(t) = \mathbb{P}(X_\varepsilon(t) = j|X(0) = i)$. Denoting $\overrightarrow{\mathbf{P}} = (\mathbf{P}_j)_{j=1,\ldots,+\infty}$ and $\mathbf{1}$ the infinite dimensional vector filled with 1, we get that $\overrightarrow{\mathbf{P}} = \mathbf{P}\cdot\mathbf{1}$. Hence with **Kolmogorov backward equation** we get an ODE on $\overrightarrow{\mathbf{P}}$:

$$\frac{\mathrm{d}\overrightarrow{\mathbf{P}}(t)}{\mathrm{d}t} = Q(t)\overrightarrow{\mathbf{P}}(t). \tag{4.5}$$

And finally we obtain the system :

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{P}_i(t)}{\mathrm{d}t} = a(i-1)\dfrac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa}\mathbf{P}_{i-1}(t) - (a(i)\dfrac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa} + b(i))\mathbf{P}_i(t) + b(i+1)\mathbf{P}_{i+1}(t),\ \forall j \geq 1 \\ \dfrac{\mathrm{d}\mathbf{P}_0(t)}{\mathrm{d}t} = -a(0)\dfrac{L_\varepsilon^X(t)}{L_\varepsilon^X(t)+\kappa}\mathbf{P}_0(t) + b(1)\mathbf{P}_1(t) \\ L_\varepsilon^X(t) + \displaystyle\sum_{i=0}^{\infty} i\mathbf{P}_i(t) = \lambda,\ \forall t \geq 0 \\ \displaystyle\sum_{i\ \geq 0} \mathbf{P}_i(t) = 1 \end{cases}$$

Hence the law of $X(t)$ given by $\overrightarrow{\mathbf{P}}(t)$ solves the Becker-Döring equations with the number of cells rescaled to 1 (this is simply done by dividing each equation by the number of cells).

Lifshitz-Slyozov SDE and deterministic Lifshitz-Slyozov model

How do we relate equation 4.2 to the diffusive Lifshitz-Slyozov model ? First a comment on the choice of B_+ and B_-. The splitting of the stochastic part into two terms is purely technical and with the aim of comparing X_ε to Z_ε. Intuitively we split the stochastic part, in two parts : the diffusion coming from positive jumps and the diffusion coming from negative jumps. The relation between the stochastic Lifshitz-Slyozov model and the deterministic Lifshitz-Slyozov model is a heuritic : consider $g(t,x)$ the probability density of Z_ε. This density solves the Fokker-Planck equation (1.32a) and the conservation equation(3.17b). Obviously the total amount of cells is rescaled to equal 1. Here we assume existence of a solution to both the stochastic and deterministic diffusive Lifshitz-Slyozov models. A rigorous correspondence between the two models and an existence result are ongoing works.

4.2 A convergence result for the stochastic models

We begin by recalling the assumptions that hold for this chapter :

$$a, b \in C^1(\mathbb{R}_+, \mathbb{R}_+), \tag{H1}$$

$$\sup_{x\in\mathbb{R}_+} |a'(x)| = K_a \text{ and } \sup_{x\in\mathbb{R}_+} |b'(x)| = K_b, \tag{H3}$$

$$a \text{ and } b \text{ are bounded functions.} \tag{H10}$$

We denote $\|\cdot\|_\infty$ the norm in $L^\infty(R_+, R_+)$ and $\|a\|_\infty = C_a$ and $\|b\|_\infty = C_b$. Also observe that $\mu : x \to \frac{x}{x+\kappa}$ is a bounded C^1 function and $\|\mu'\|_\infty = \frac{1}{\kappa}$.

4.2.1 Preliminary results

The proof of theorem 1.5.2 involves the use of various result from probability which we recap here. The overall shape of the proof is similar to the one found in [73] although we obtain an L^1 bound which involves additional steps.

Lemma 4.2.1. *Let $W(t)$ be a Wiener process. Then a unit Poisson process $Y(t)$ can be constructed on the same probability space as $W(t)$ such that :*

$$\sup_{t\geq 0}\frac{|Y(t)-t-W(t)|}{\ln(max(2,t))}=K<\infty, \tag{4.6}$$

and $\mathbb{E}\left[\exp(\lambda K)\right]<\infty$ for $\lambda>0$.

Proof. This lemma is a direct implication of the Komlós–Major–Tusnády approximation found in [72]. □

Lemma 4.2.2. *Let $W(t)$ be a Wiener process and define :*

$$\omega=\sup_{u,v<1}\frac{W(u)-W(v)}{\sqrt{|u-v|\ln(\frac{1}{|u-v|})}}. \tag{4.7}$$

Then ω is almost surely finite and $\mathbb{E}\left[\exp(\lambda\omega)\right]<\infty$ for $\lambda>0$.

Proof. The finiteness of ω is derived from the definition of the Levy modulus of continuity and proved in [81]. The bounded exponential moments is proved in [43]. □

4.2.2 Introductory lemmas

We assume for simplicity that $X_\varepsilon(0)=Z_\varepsilon(0)$.

Now we compare $X_\varepsilon(t)$ and $Z_\varepsilon(t)$ using (4.1) and (4.2) and by adding and removing appropriate terms in their difference :

$$\begin{aligned}
X_\varepsilon(t)-Z_\varepsilon(t)=&\varepsilon Y_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s))\mu(L_\varepsilon^X(s))\mathrm{d}s\Big)-\varepsilon Y_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big)\\
&-\varepsilon B_+\Big(\varepsilon^{-1}\int_0^t a(Z_\varepsilon(s))\mu(L_\varepsilon^Z(t))\mathrm{d}s\Big)+\varepsilon B_-\Big(\varepsilon^{-1}\int_0^t b(Z_\varepsilon(s))\mathrm{d}s\Big)\\
=&\varepsilon Y_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s))\mu(L_\varepsilon^X(s))\mathrm{d}s\Big)-\varepsilon B_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s))\mu(L_\varepsilon^X(s))\mathrm{d}s\Big)\\
&+\varepsilon B_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big)-\varepsilon Y_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big)\\
&+\varepsilon B_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s)\mu(L_\varepsilon^X(s))\mathrm{d}s\Big)-\varepsilon B_+\Big(\varepsilon^{-1}\int_0^t a(Z_\varepsilon(s))\mu(L_\varepsilon^Z(t))\mathrm{d}s\Big)\\
&+\varepsilon B_-\Big(\varepsilon^{-1}\int_0^t b(Z_\varepsilon(s))\mathrm{d}s\Big)-\varepsilon B_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big)
\end{aligned} \tag{4.8}$$

The terms of the form $Y-B$ will be bounded using lemma 4.2.1. To deal with the term of the form $B-B$, we recall that $W_{+,-}(t)=B_{+,-}(t)-t$ are standard Brownian motions. Hence we need to bound the terms in $W-W$ using lemma 4.2.2, and the remaining terms are easily bounded by using assumptions on both a and b.

Let $T>0$. For the sake of clarity, we make use of the following notations :

$$\gamma_\varepsilon(t)=\varepsilon^{-1}|X_\varepsilon(t)-Z_\varepsilon(t)| \text{ and } \gamma_\varepsilon=\sup_{t\in[0,T]}\gamma_\varepsilon(t) \tag{4.9}$$

$$\delta_\varepsilon^T = \sup_{0\le t\le T} |Y_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s))\mu(L_\varepsilon^X(s))\mathrm{d}s\Big) - B_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s))\mu(L_\varepsilon^X(s))\mathrm{d}s\Big)| \\ + \sup_{0\le t\le T} |B_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big) - Y_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big)| \quad (4.10)$$

Furthermore the terms not depending on $Y_{+,-}$ and $W_{+,-}$ are easily bounded using assumption (H3) :

$$\begin{aligned}\varepsilon^{-1}|\int_0^t a(X_\varepsilon(s))\mu(L_\varepsilon^X(s))\mathrm{d}s - \int_0^t a(Z_\varepsilon(s))\mu(L_\varepsilon^Z(s))\mathrm{d}s| \\ \le K_a\int_0^t \gamma_\varepsilon(s)\mathrm{d}s + \frac{C_a}{\kappa}\varepsilon^{-1}\int_0^t |L_\varepsilon^X(s) - L_\varepsilon^Z(s)|\mathrm{d}s \quad (4.11)\\ \le K_a\int_0^t \gamma_\varepsilon(s)\mathrm{d}s + \frac{C_a}{\kappa}\int_0^t \mathbb{E}[\gamma_\varepsilon(s)]\mathrm{d}s,\end{aligned}$$

$$\varepsilon^{-1}|\int_0^t b(X_\varepsilon(s))\mathrm{d}s - \int_0^t b(Z_\varepsilon(s))\mathrm{d}s| \le K_b\int_0^t \gamma_\varepsilon(s)\mathrm{d}s. \quad (4.12)$$

This leads us to:

$$\begin{aligned}\gamma_\varepsilon(t) \le& \delta_\varepsilon^T + K_a\int_0^t \gamma_\varepsilon(s)\mathrm{d}s + \frac{C_a}{\kappa}\int_0^t \mathbb{E}[\gamma_\varepsilon(s)]\mathrm{d}s + K_b\int_0^t \gamma_\varepsilon(s)\mathrm{d}s \\ &+ \sup_{0\le t\le T} |W_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s))\mu(L_\varepsilon^X(s))\mathrm{d}s\Big) - W_+\Big(\varepsilon^{-1}\int_0^t a(Z_\varepsilon(s))\mu(L_\varepsilon^Z(s))\mathrm{d}s\Big)| \quad (4.13)\\ &+ \sup_{0\le t\le T} |W_-\Big(\varepsilon^{-1}\int_0^u b(X_\varepsilon(s))\mathrm{d}s\Big) + W_-\Big(\varepsilon^{-1}\int_0^u b(Z_\varepsilon(s))\mathrm{d}s\Big)|\end{aligned}$$

We proceed with the bounds for the terms of the form $W - W$.

Lemma 4.2.3. *Let $T > 0$ and $0 \le t \le T$. Define*

$$\psi_C(x) = \begin{cases} x\ln\left(\dfrac{C}{x}\right) & \text{if } 0 \le x < C, \\ \dfrac{C}{\mathrm{e}} & \text{else.}\end{cases} \quad (4.14)$$

Then there exists two independent random variable $\omega_+ > 0$ and $\omega_- > 0$ also independent of ε such that :

$$|W_+\Big(\varepsilon^{-1}\int_0^t a(X_\varepsilon(s)\mu(L_\varepsilon^X(s))\mathrm{d}s\Big) - W_+\Big(\varepsilon^{-1}\int_0^t a(Z_\varepsilon(s))\mu(L_\varepsilon^Z(s))\mathrm{d}s\Big)| \\ \le \omega_+\sqrt{\psi_{\varepsilon^{-1}C_a}(K_a\gamma_\varepsilon + \frac{C_a}{\kappa}\,\mathbb{E}\,[\gamma])}, \quad (4.15)$$

$$|W_-\Big(\varepsilon^{-1}\int_0^t b(X_\varepsilon(s))\mathrm{d}s\Big) + W_-\Big(\varepsilon^{-1}\int_0^t b(Z_\varepsilon(s))\mathrm{d}s\Big)| \le \omega_-\sqrt{\psi_{\varepsilon^{-1}C_b}(K_b\gamma_\varepsilon)}. \quad (4.16)$$

Proof. First we can bound the difference between the two amount of external lipids by the expected value of γ_ε :

$$|L_\varepsilon^X(t) - L_\varepsilon^Z(t)| = |\,\mathbb{E}[X_\varepsilon(t) - Z_\varepsilon(t)]| \leq \varepsilon\,\mathbb{E}[\gamma_\varepsilon].$$

Also observe that assumption (H10), all the time integrals are bounded by $\varepsilon^{-1}C_aT$ or $\varepsilon^{-1}C_bT$.

We introduce the two following constants (derived from the modulus of continuity of the Wiener process).

$$\omega_+ = \sup_{u,v\leq\varepsilon^{-1}C_aT} \frac{|W_+(u) - W_+(v)|}{\sqrt{|u-v|\ln\left(\dfrac{\varepsilon^{-1}C_aT}{|u-v|}\right)}},$$

$$\omega_- = \sup_{u,v\leq\varepsilon^{-1}C_bT} \frac{|W_-(u) - W_-(v)|}{\sqrt{|u-v|\ln\left(\dfrac{\varepsilon^{-1}C_bT}{|u-v|}\right)}}.$$

From lemma 4.2.2, ω_+ and ω_- are finite and have finite exponential moments i.e.

$$\mathbb{E}\left[\exp(\lambda\omega_{-,+})\right] < \infty,$$

for some $\lambda > 0$. Moreover there laws are independent of ε. Indeed let

$$\omega = \sup_{u,v\leq C} \frac{|W_+(u) - W_+(v)|}{\sqrt{|u-v|\ln\left(\frac{C}{|u-v|}\right)}},$$

with $C > 0$. Now changing the notation with $u' = \frac{u}{C}$ and $v' = \frac{v}{C}$, one has :

$$\frac{|W_+(u) - W_+(v)|}{\sqrt{|u-v|\ln\left(\dfrac{C}{|u-v|}\right)}} = \frac{|W_+(Cu') - W_+(Cv')|}{\sqrt{C|u'-v'|\ln\left(\dfrac{1}{|u'-v'|}\right)}}.$$

Using the self similarity of the Wiener process $W(u') \sim \frac{W(Cu')}{\sqrt{C}}$ we obtained that

$$\omega = \sup_{u,v\leq 1} \frac{|W_+(u) - W_+(v)|}{\sqrt{|u-v|\ln\left(\frac{1}{|u-v|}\right)}}.$$

Hence ω is independent of the choice of C and we also have independence of $\omega_{+,-}$ and ε.

We proceed with the bound on W_+. Let $u, v \geq 0$ such that $u, v \leq \varepsilon^{-1}C_aT$. In our case, we also expect u and v to verify $|u - v| \leq K_aT\gamma_\varepsilon + \frac{C_a}{\kappa}T\,\mathbb{E}[\gamma]$. However this bound may not be smaller than $\varepsilon^{-1}C_aT$, which would allow us to use the fact that the function $x \to \sqrt{x\ln\left(\dfrac{C}{x}\right)}$ is increasing on $[0, C]$. Nonetheless we proceed as follow. We have that the function $x \to \sqrt{\psi_C(x)}$ is non-decreasing and $x\ln\left(\frac{C}{x}\right) \leq \psi_C(x)$ for all $x \geq 0$. Hence for $u, v \leq \varepsilon^{-1}C_aT$ such that $|u - v| \leq K_aT\gamma_\varepsilon + \frac{C_a}{\kappa}T\,\mathbb{E}[\gamma]$ one has :

$$
\begin{aligned}
|W_+(u) - W_+(v)| &\leq \frac{|W_+(u) - W_+(v)|}{\sqrt{|u-v| \ln\left(\frac{\varepsilon^{-1} C_a T}{|u-v|}\right)}} \sqrt{\psi_{\varepsilon^{-1} C_a T}(|u-v|)} \\
&\leq \omega_+ \sqrt{\psi_{\varepsilon^{-1} C_a T}(K_a T \gamma_\varepsilon + \frac{C_a}{\kappa} T \, \mathbb{E}\left[\gamma\right])} \\
&= \omega_+ \sqrt{T} \sqrt{\psi_{\varepsilon^{-1} C_a}(K_a \gamma_\varepsilon + \frac{C_a}{\kappa} \, \mathbb{E}\left[\gamma\right])}
\end{aligned}
$$

Note that in the case $u = v$ the bound is simply 0.

This yields (4.15) by letting $u = \varepsilon^{-1} \int_0^t a(X_\varepsilon(s)) \mu(L_\varepsilon^X(s)) \mathrm{d}s$ and $v = \varepsilon^{-1} \int_0^t a(Z_\varepsilon(s)) \mu(L_\varepsilon^Z(s)) \mathrm{d}s$. We proceed exactly in the same way for W_- to get (4.16).

□

Terms of the form $Y - B$ are treated with lemma 4.2.1 in the following lemma :

Lemma 4.2.4. *Assume the processes Y_+ and Y_- to be constructed as in lemma 4.2.1. Let $T > 0$. Then there exists some random variable $K > 0$ with exponential moments such that :*

$$
\begin{aligned}
|Y_+\Big(\varepsilon^{-1} \int_0^t a(X_\varepsilon(s)) \mu(L_\varepsilon^X(s)) \mathrm{d}s\Big) - B_+\Big(\varepsilon^{-1} \int_0^t a(X_\varepsilon(s)) \mu(L_\varepsilon^X(s)) \mathrm{d}s\Big)| \\
\leq K \ln(\varepsilon^{-1} C_a T \vee 2) \quad (4.17)
\end{aligned}
$$

$$
|B_-\Big(\varepsilon^{-1} \int_0^t b(X_\varepsilon(s)) \mathrm{d}s\Big) - Y_-\Big(\varepsilon^{-1} \int_0^t b(X_\varepsilon(s)) \mathrm{d}s\Big)| \leq K \ln(\varepsilon^{-1} C_b T \vee 2) \quad (4.18)
$$

Proof. Following lemma 4.2.4, we introduce the random variable :

$$
K = \sup_{t \geq 0} \frac{|Y(t) - B(t)|}{\ln(\max(t, 2))}.
$$

Let $T > 0$ and $u \geq 0$ such that $u \leq \varepsilon^{-1} C_a T$. Then

$$
\begin{aligned}
|Y_+(u) - B_+(u)| \leq & \frac{|Y_+(u) - B_+(u)|}{\ln(\max(u, 2))} \ln(\max(u, 2)) \\
\leq & \frac{|Y_+(u) - B_+(u)|}{\ln(\max(u, 2))} \ln(\max(\varepsilon^{-1} C_a T, 2)) \\
\leq & K \ln(\max(\varepsilon^{-1} C_a T, 2)).
\end{aligned}
$$

We proceed similarly for Y_- and B_- to get (4.18). □

We apply the bounds (4.15) and (4.16) to (4.13) :

$$
\begin{aligned}
\gamma_\varepsilon(t) \leq & \delta_\varepsilon^T + \frac{C_a}{\kappa} \int_0^t \mathbb{E}[\gamma_\varepsilon(s)] \mathrm{d}s + (K_a + K_b) \int_0^t \gamma_\varepsilon(s) \mathrm{d}s \\
& + \omega_+ \sqrt{T} \sqrt{\psi_{\varepsilon^{-1} C_a}(C_a \gamma_\varepsilon + \frac{C_a}{\kappa} \, \mathbb{E}\left[\gamma\right])} \\
& + \omega_- \sqrt{T} \sqrt{\psi_{\varepsilon^{-1} C_b}(K_b \gamma)}
\end{aligned} \quad (4.19)
$$

4.2.3 Proof of theorem 1.5.2

We can proceed with the proof of theorem 1.5.2, which we recall :

Theorem 4.2.1. *Assume X_ε and Z_ε to be solutions of Equations* (4.1) *and* (4.2) *and for assumptions* (H1), (H3) *and* (H10) *to hold. Then for ε small enough and for some $T > 0$, there exists a constant β^T only dependant on T, a and b such that :*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|X_\varepsilon(t) - Z_\varepsilon(t)|\right] \leq \beta^T \varepsilon \ln(\frac{1}{\varepsilon}) \tag{4.20}$$

Proof. We begin from (4.19) by taking the expected value and applying the Cauchy-Schwartz inequality and Fubini theorem :

$$\begin{aligned}\mathbb{E}\left[\gamma_\varepsilon(t)\right] \leq \mathbb{E}\left[\delta_\varepsilon^T\right] + (\frac{C_a}{\kappa} + K_a + K_b)\int_0^t \mathbb{E}[\gamma_\varepsilon(s)]\mathrm{d}s \\ + \sqrt{\mathbb{E}\left[\omega_+^2\right]T}\sqrt{\mathbb{E}\left[\psi_{\varepsilon^{-1}C_a}(K_a\gamma_\varepsilon + \frac{C_a}{\kappa}\mathbb{E}\left[\gamma\right])\right]} \\ + \sqrt{\mathbb{E}\left[\omega_-^2\right]T}\sqrt{\mathbb{E}\left[\psi_{\varepsilon^{-1}C_b}(K_b\gamma)\right]}.\end{aligned}$$

Again we use a key property of the functions $\psi_C(x)$: it is concave on $\mathbb{R}_+$. This allows us to apply Jensen's lemma. Then :

$$\begin{aligned}\mathbb{E}\left[\gamma_\varepsilon(t)\right] \leq \mathbb{E}\left[\delta_\varepsilon^T\right] + (\frac{C_a}{\kappa} + K_a + K_b)\int_0^t \mathbb{E}[\gamma_\varepsilon(s)]\mathrm{d}s \\ + \sqrt{\mathbb{E}\left[\omega_+^2\right]T}\sqrt{\psi_{\varepsilon^{-1}C_a}((K_a + \frac{C_a}{\kappa})\mathbb{E}\left[\gamma\right])} \\ + \sqrt{\mathbb{E}\left[\omega_-^2\right]T}\sqrt{\psi_{\varepsilon^{-1}C_b}(K_b\,\mathbb{E}\left[\gamma\right])}.\end{aligned}$$

We know need a proper bound on $\psi_C(x)$. Since it is a concave function, it is bellow its tangent functions. This yields that for all $x, y > 0$:

$$\psi_C(y) \leq \ln\left(\frac{C}{x}\right)y + x.$$

In particular, in our case, we have that for ε small enough $\varepsilon^{-1}T(C_a + C_b) > 1$. Hence we can choose $x = 1$ in the previous bound, which leads to, for all $y \geq 0$:

$$\psi_{\varepsilon^{-1}C_a}(y) \leq \ln\left(\varepsilon^{-1}C_a\right)y + 1,$$

and similarly :

$$\psi_{\varepsilon^{-1}C_b}(y) \leq \ln\left(\varepsilon^{-1}C_b\right)y + 1.$$

And when taking the square root : $\sqrt{Cy+1} \leq 1 + \sqrt{Cy}$. Finally this yields :

$$\begin{aligned}\mathbb{E}\left[\gamma_\varepsilon(t)\right] \leq \mathbb{E}\left[\delta_\varepsilon^T\right] + (\frac{C_a}{\kappa} + K_a + K_b)\int_0^t \mathbb{E}[\gamma_\varepsilon(s)]\mathrm{d}s \\ + \sqrt{(K_a + \frac{C_a}{\kappa})T\,\mathbb{E}\left[\omega_+^2\right]} + \sqrt{K_b T\,\mathbb{E}\left[\omega_-^2\right]} \\ + \left(\sqrt{(K_a + \frac{C_a}{\kappa})T\,\mathbb{E}\left[\omega_+^2\right]}\sqrt{\ln(\varepsilon^{-1}C_a)} + \sqrt{K_b T\,\mathbb{E}\left[\omega_-^2\right]}\sqrt{\ln(\varepsilon^{-1}C_b)}\right)\sqrt{\mathbb{E}[\gamma_\varepsilon]}.\end{aligned}$$

We make use of Grönwall's lemma to get:

$$\begin{aligned}\mathbb{E}\left[\gamma_\varepsilon(t)\right] \leq \mathrm{e}^{(\frac{C_a}{\kappa}+K_a+K_b)T}\Big[\,\mathbb{E}\left[\delta_\varepsilon^T\right] \\ + \sqrt{(K_a + \frac{C_a}{\kappa})T\,\mathbb{E}\left[\omega_+^2\right]} + \sqrt{K_b T\,\mathbb{E}\left[\omega_-^2\right]} \\ + \left(\sqrt{(K_a + \frac{C_a}{\kappa})T\,\mathbb{E}\left[\omega_+^2\right]}\sqrt{\ln(\varepsilon^{-1}C_a)} + \sqrt{K_b T\,\mathbb{E}\left[\omega_-^2\right]}\sqrt{\ln(\varepsilon^{-1}C_b)}\right)\sqrt{\mathbb{E}[\gamma_\varepsilon]}\Big].\end{aligned}$$

This leads us to investigate solutions to the inequality with $x = \sqrt{\mathbb{E}\left[\gamma\right]}$:

$$x^2 \leq c + bx,$$

where $b, c > 0$. Hence x lies between the roots of the polynomial $c + by - y^2$ and its square is bounded by the square of the biggest root in absolute value. Simple computations show that then $x^2 \leq b^2 + 2c$. To tidy up some of the notation let us denote :

$$C_1^T = \mathrm{e}^{(\frac{C_a}{\kappa}+K_a+K_b)T},$$

$$C_2^T = \sqrt{(K_a + \frac{C_a}{\kappa})T\,\mathbb{E}\left[\omega_+^2\right]} + \sqrt{K_b T\,\mathbb{E}\left[\omega_-^2\right]},$$

$$C_3 = \max(C_a, C_b).$$

This leads to :

$$\mathbb{E}\left[\gamma\right] \leq 2C_1^T\Big(\mathbb{E}\left[\delta_\varepsilon^T\right] + C_2^T\Big) + (C_1^T C_2^T)^2\ln(\varepsilon^{-1}C_3). \tag{4.21}$$

According to lemma 4.2.4, there exists a random variable $K > 0$ such that for ε small enough $\delta_\varepsilon^T \leq K\ln(\varepsilon^{-1})$. Furthermore K has exponential moments which means $\mathbb{E}\left[\delta_\varepsilon^T\right] \leq \mathbb{E}\left[K\right]\ln(\varepsilon^{-1})$ for ε small enough. The expected values of K, ω_- and ω_+ are bounded. This yields :

$$\mathbb{E}\left[\gamma\right] \leq 2C_1^T C_2^T + (C_1^T C_2^T)^2\ln(C_3) + (2C_1^T\,\mathbb{E}\left[K\right] + (C_1^T C_2^T)^2)\ln(\varepsilon^{-1}). \tag{4.22}$$

Finally, we regroup all terms not depending on ε under the notation β^T to conclude the proof and get for ε small enough :

$$\mathbb{E}\left[\gamma\right] \leq \beta^T\ln(\varepsilon^{-1}).$$

□

4.3 Simulations

We perform simulations of both stochastic models, using the methods introduced in 1.4.8. We recall that the Next reaction method is an exact simulation while the Euler-Muruyama method is an approximation. We also recall that the models we simulate are (1.71) and (1.43), whose solutions are a set of particles $(X_\varepsilon^{k,N})_{k=1,\dots,N}$ or $(Z_\varepsilon^{k,N})_{k=1,\dots,N}$

In both cases, the choice of a stopping time for the method is tricky. Indeed even if we have reached a stationary distribution, the set of particles may still change. With the intent of comparing the simulation of both stochastic model to the stationary solutions of the diffusive Lifshitz-Slyozov model, we fixed a final time equivalent to the one reached by the numerical scheme used to simulate the diffusive Lifshitz-Slyozov model.

4.3.1 A first basic simulation and choosing an initial condition

We begin by a simple simulation with $N = 10000$ and $\varepsilon = 0.05$. We expected the approximation to get better as $N \to \infty$ but increasing the number of particles greatly increases the run time of the program. The choice of N is made as a tradeoff between good approximation and run time. This simulation is presented in figure 4.1. We also need to choose an initial condition for both models. For simplicity we choose the same initial condition $(X_{k,\varepsilon}^0)_{k=1,\dots,N} = (Z_{k,\varepsilon}^0)_{k=1,\dots,N}$. The choice of this initial condition is non-trivial since it needs to verify that $L_\varepsilon^{X,0} + \frac{1}{N}\sum_{k=1}^{N} X_{k,\varepsilon}^0 = \lambda$. Moreover we have that $X_{k,\varepsilon}^0 \in \varepsilon\mathbb{N}$ for all $k = 1, \dots, N$. Hence this amounts to generating a random set of integers n_k such that :

$$\sum_{k=1}^{N} n_k = N\frac{\lambda - L_\varepsilon^{X,0}}{\varepsilon} = S.$$

One also needs to choose $L_\varepsilon^{X,0}$, which we do in such a way that S is also an integer - this means ε divides $\lambda - L_\varepsilon^{X,0}$. Since some n_k may be equal to 0, generating an initial condition comes down to choosing a random partition of S. Choosing uniformly a partition of S gets very costly as N gets bigger thus we use a simpler method, which does not yield a uniform draw of a partition but is skewed toward the average $\frac{\lambda - L_\varepsilon^{X,0}}{\varepsilon}$:

1 Initialization : let $n_k = 0$ for all $k = 1, \dots, N$.

2 While $\sum_{k=1}^{N} n_k < S$, do :

  a Generate $i \sim \mathcal{U}([1, \dots, N])$.

  b Update $n_i \leftarrow n_i + 1$.

Looking at figure 4.1, we observe that for both model we recover a bimodal distribution that is overall correlated with the stationary distribution of the diffusive Lifshitz-Slyozov model. However, it seams that the first maxima, located at or close to zero, is underestimated. This may be explained by a simple fact : the stationary solution for the diffusive Lifshitz-Slyozov model is continuous whereas the set of particles for both stochastic models are in some sense discrete. The stochastic Becker-Döring model lives on $\varepsilon\mathbb{N}$ hence depending on the choice of ε the approximation close to zero may get better as ε tends to zero.

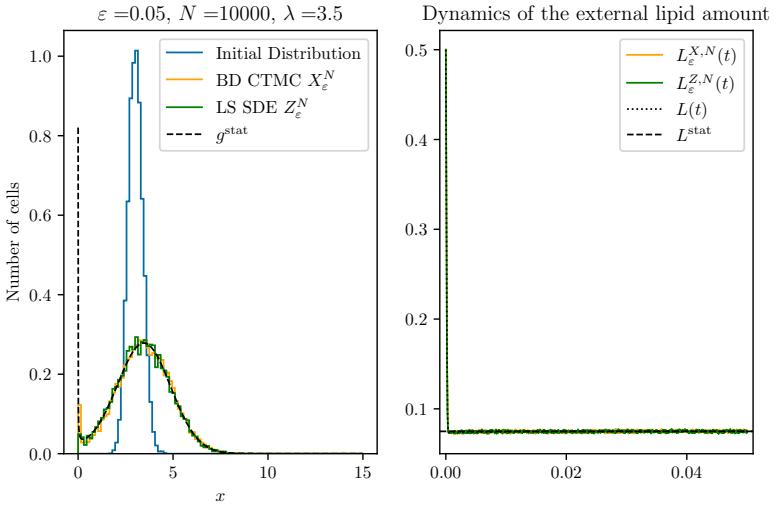


Figure 4.1: Simulation of both stochastic models, with $N = 10000$, $\varepsilon = 0.05$ and $\lambda = 3.5$. On the left side, we compare the stationary distribution of both stochastic model with the explicit stationary solution of the diffusive Lifshitz-Slyozov model. On the right, we compare the dynamics of the external amount of lipids for both stochastic models as well as the result obtained with the scheme described in section 1.4.7. The green and orange lines are superimposed on the right side.

4.3.2 Can we recover the whole range of stationary distribution ?

As we have seen in the previous chapter, there exists a whole range of stationary solutions (at least numerically) to the diffusive Lifshitz-Slyozov model. We show in figure 4.2 that both stochastic models are able to capture this range of stationary distributions. We also show that the shape of the initial condition has no influence on the stationary distribution, as is the case for the diffusive Lifshitz-Slyozov model.

4.3.3 Influence of N and ε and a numerical proof of theorem 1.5.2

In this section, we investigate the relation between the number of particles N and ε and their common influence on the distribution of particles. We are also interested in seeing if we can recover the theoretical result from theorem 1.5.2 at the numerical level. Since we use models with a fixed number of particles N, we need a way to compare two population of particles. This can be done in a variety of ways, but since the theoretical result involves the L^1-norm, we need an empirical approximation of this norm for two set of particles. We do so with the formula :

$$d((X_{k,\varepsilon}(t))_{k=1,\dots,N}, (Z_{k,\varepsilon}(t))_{k=1,\dots,N}) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|X_{i,\varepsilon}(t) - Z_{j,\varepsilon}(t)|. \tag{4.23}$$

For clarity we abbreviate $d((X_{k,\varepsilon}(t))_{k=1,\dots,N}, (Z_{k,\varepsilon}(t))_{k=1,\dots,N}) = d(X_\varepsilon(t), Z_\varepsilon(t))$. Our hope is that there is some constant β such that as $\varepsilon \to 0$ we have $d(X_\varepsilon(t), Z_\varepsilon(t)) \leq \beta\varepsilon\ln(\varepsilon^{-1})$. This result is presented in figure 4.3. We choose a set of time points $t_1, \dots, t_K$ and we compute $\max_{k=1,\dots,K} d(X_\varepsilon(t_k), Z_\varepsilon(t_k))$ for different values of $\varepsilon > 0$. Observe that even for small amounts

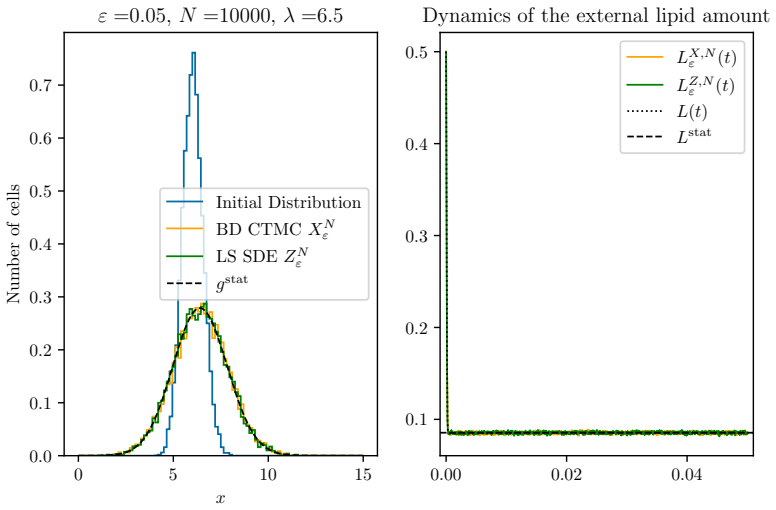


Figure 4.2: A simulation of both models with a unimodal stationary solution

of particles we have similar values of $d(X_\varepsilon(t), Z_\varepsilon(t))$. However for small values of N the approximation of the diffusive Lifshitz-Slyozov model by both stochastic model is worse as seen in figure 4.4

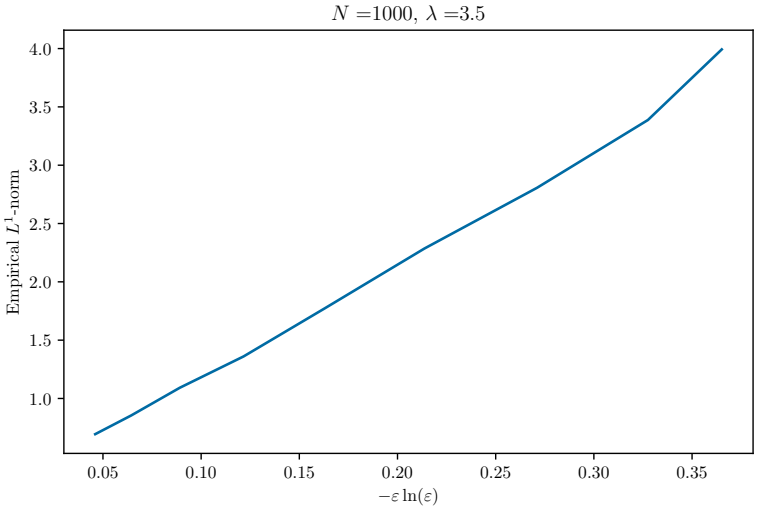


Figure 4.3: Maximum empirical L^1-norm of the two stochastic model. We obtain that this norm is of magnitude $\varepsilon \ln(\varepsilon)$ at least for small ε.

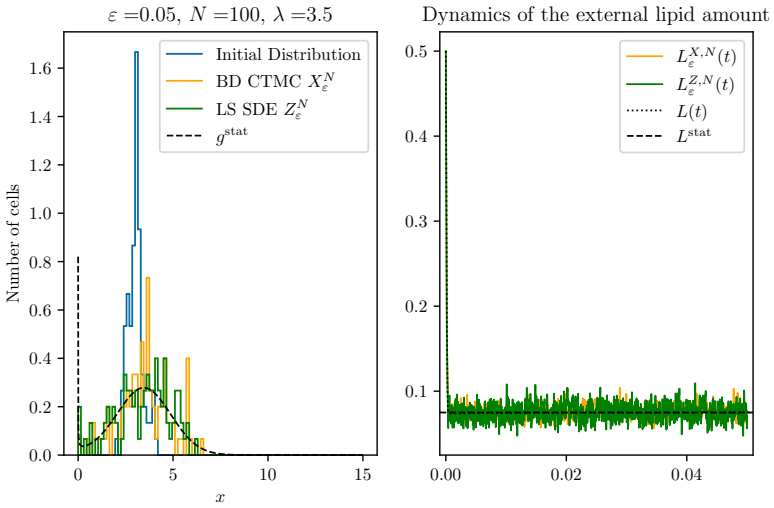


Figure 4.4: Simulation of both stochastic model for $N = 100$.

Chapter 5

Parameter estimation and some additional numerical results

In this chapter, we present the work done in collaboration with Chloé Audebert, Anne-Sophie Giacobbi, Magali Ribot, Hédi Soula and Romain Yvinec with the model in radii (1.78). This work is submitted under the title "Mathematical modeling of adipocyte size distributions: identifiability and parameter estimation from rat data" [44].In the second section, we introduce similar results for the model in lipids which as a non-constant rate of diffusion. The work presented in the first section is submitted and small editions were made for ease of reading in the manuscript.

5.1 Mathematical modeling of adipocyte size distributions: identifiability and parameter estimation from rat data

5.1.1 Introduction

Pathologies related to obesity are characterized by an important accretion of adipose tissue which is mainly composed of adipose cells, called adipocytes. The adipocytes are designed to regulate energy homeostasis by storing energy in form of lipids. During an excess of energy, adipocytes compensate with two mechanisms: hypertrophy (increase in size) and hyperplasia (increase in number)[40]. Adipocyte size variations are very large with radii ranging from 10μm to more than 100μm, corresponding to 3 orders of magnitude in volume. In addition, cell size distribution among a tissue is not unimodal but presents two peaks: one for small adipocytes (radius below 30μm) and one for large adipocytes (above 80μm) [88]. A bimodal distribution of cell sizes is striking. Indeed, most cells in the population are small adipocytes, which do not contribute significantly to the storing capacity. There is no scientific consensus on the functional importance of this bimodality. However, cell size has been associated with metabolic properties dysfunction that may be linked to obesity-related pathologies [127, 116, 88, 85] or to play a role in the development of those diseases [30].

Few mathematical models have been proposed for adipocyte size dynamics in various health conditions. In [66, 67, 68, 82], the authors consider partial differential equation models that describe adipocyte size distribution dynamics. They have assumed a size-dependent rate described by an imposed function where the associated parameters are difficult to relate to physiological processes. The adipocyte modeling in [86] is based on three compartments and has been developed to describe small, medium and large adipocytes. The cell size evolution depends on lipid

fluxes that are related to protein concentration controlling lipotoxicity – a cellular dysfunction due to lipid accumulation in non-adipose tissue. All these models provide studies of the adipose tissue growth dynamic and its bimodality through cell hyperplasia and/or hypertrophy, but the mechanisms governing lipid fluxes involved in adipocyte hypertrophy have not been considered. Furthermore, model parameters lack biological meaning.

A detailed model of cell hypertrophy based on lipid exchanges has been proposed in [118]. Adipocyte bimodal distributions have been explained based on mathematical analyses. Individual-based Monte Carlo techniques were performed to solve the model. However, this approach is computationnally costly so parameter estimation using biological measurements is very difficult. A similar simplified model, accounting only for lipolysis (deflation), compares well with distributions obtained from fasting rats [119].

The paper is organized as follows. Based on [118, 119], we formulate the mathematical model in section 5.1.2. It is based on partial differential equations, to describe stationary adipocyte size distribution. The contribution of our work is to have a diffusion term in the partial differential equation describing the cell size fluctuations like in [67]. Through parameter estimation, we aim at comparing the distribution obtained with the model to cell size distribution measured in rats before any manipulation [119, 63]. To perform parameter estimation, we first conduct an identifiability analysis in order to select model parameters that can be uniquely estimated with the available data. Using these selected parameters, we carry out a study on synthetic data (generated with model equations). The model identifiability and the parameter estimation on synthetic data are presented in section 5.1.3. Once the parameter estimation problem is verified, in section 5.1.4 we perform parameter estimation using adipocyte size distributions measured in 32 healthy rats [119, 63]. The estimated parameters are presented and then commented through a sensitivity analysis. We conclude this paper with some discussions in section 5.1.5.

5.1.2 Mathematical model for adipocyte size distributions

Model construction

Based on Soula et al. [118] work, we introduce a new model for adipocyte size distribution that we aim at fitting on experimental measurements. We first briefly recall the main hypotheses of the model in [118]. To represent adipocyte size density, the variation of the content of lipids ℓ and variation of radius r to adapt to lipid content are described by,

$$\begin{cases} \dfrac{d\ell}{dt} = \mathcal{T}(r, \ell, L(t)), & (5.1) \\ \dfrac{dr}{dt} = \mathcal{R}(r, \ell), & (5.2) \end{cases}$$

where the term $L(t)$ represents the extracellular amount of lipids at time t. These two equations refer to evolution with different characteristic times: the first equation is a rapid evolution of fatty acid content whereas the second is a slower variation of radius to adapt to cell lipid content.

We first assume a quasi steady state for equation (5.2) to describe a faster adaptation to lipid content. The relation between the lipid content ℓ and the radius r of a cell is then given by $\mathcal{R}(r, \ell) = 0$, leading to

$$\ell = \frac{V(r) - V_{em}}{V_\ell}, \quad V(r) = \frac{4}{3}\pi r^3, \tag{5.3}$$

with V_{em} the volume of the cell with no lipid, V_ℓ the conversion constant: the volume taken by 1 $nmol$ of triglyceride, and the cell volume $V(r)$ is assumed to be spherical. Second, similarly to [67], we introduce a constant diffusion term D to represent cell size fluctuations.

With the above mentioned assumptions, we can re-write the main equation in [118], replacing ℓ by (5.3) and keeping only the radius variable. We then consider the cell size density f expressed as a function of time $t \in \mathbb{R}_+$ and radius $r \in [r_{min}, r_{max}]$, and we introduce the following system:

$$\begin{cases} \partial_t f(t,r) + \partial_r(v(r,L(t))f(t,r)) - D\partial_r^2 f(t,r) = 0, & (5.4) \\ L(t) = \lambda - \displaystyle\int_{r_{min}}^{r_{max}} (V(r) - V_{em}) \frac{4\pi r^2}{V_\ell^2} f(t,r) dr, & (5.5) \\ v(r_{min}, L(t)) f(t, r_{min}) - D\partial_r f(t, r_{min}) = 0, & (5.6) \\ v(r_{max}, L(t)) f(t, r_{max}) - D\partial_r f(t, r_{max}) = 0, & (5.7) \end{cases}$$

where v is defined by

$$v(r,L) = \frac{V_\ell}{4\pi} \left(\alpha \frac{L}{L + \kappa} \frac{\rho^3}{\rho^3 + r^3} - \frac{(\beta + \gamma r^2)}{r^2} \frac{V(r) - V_{em}}{V(r) - V_{em} + V_\ell \chi} \right). \quad (5.8)$$

The total amount of lipids λ is assumed to be constant over time and the second term of the right-hand side of (5.5) describes the intracellular amount of lipids at time t contained within all cells. The transport function v describes the exchange of lipids within the population of cells [118]. The lipid exchanges are based on two biochemical processes: lipogenesis – cell store lipids – and lipolysis – release of lipids in the extracellular environment. Lipogenesis depends on a surface-limited rate α, and it increases with the extracellular amount of lipids L with a saturation effect depending on the value of κ. The parameter ρ is a cell size threshold above which lipogenesis slows down. Lipolysis activity includes a basal rate β and a surface-limited rate γ. The term $\frac{V(r)-V_{em}}{V(r)-V_{em}+V_\ell \chi} = \frac{\ell}{\ell+\chi}$ is small when cells contain few lipids and becomes close to one for larger lipid content through parameter χ.

We assume that in the measurements at the time of the biopsy the adipose tissue is at equilibrium, thus we neglect the recruitment of new cells. In addition, it has been shown that the life time of a human adipocyte is around 10 years [7], so the cell death is not taken into account. It gives the boundary conditions (5.6)-(5.7). The total number of cells is then constant and we assume the density integral is 1 between r_{min} and r_{max}, which leads to

$$\forall t \geq 0, \quad \int_{r_{min}}^{r_{max}} f(t,r) dr = 1. \quad (5.9)$$

Table 5.1 reports the details on model variables and parameters. The parameter values of V_{em}, V_l, β and γ are known from literature [118, 119] and will be fixed. We choose the values of r_{min} and r_{max} as the boundary values of the measured radii in the considered adipose tissue.

Stationary solution

In model (5.4)-(5.7), the number of adipocytes is fixed and the total amount of lipids is constant, thus we expect the size distribution to reach a steady state [111]. The mathematical study of the asymptotic behavior is not the purpose of this work.

We denote by f^∞ and L^∞ a stationary density of cell size and the extracellular amount of lipids respectively. A stationary solution verifies $\partial_t f^\infty(r) = 0$. With the boundary conditions (5.6)-(5.7) and assuming $D \neq 0$, we obtain the following system:

$$\begin{cases} \partial_r f^\infty(r) = \dfrac{1}{D} v(r, L^\infty) f^\infty(r), & (5.10) \\ L^\infty = \lambda - \displaystyle\int_{r_{min}}^{r_{max}} (V(r) - V_{em}) \frac{4\pi r^2}{V_\ell^2} f^\infty(r) dr. & (5.11) \end{cases}$$

Table 5.1: **Description of model variables and parameters.** Parameter units and known values are summed up in the second column and a description of each variable is given in the third column.

name	value (unit)	description
t	- (h)	time
r	$\in [7.5, 150]$ (μm)	adipocyte radius [119, 63]
$L(t)$	- $(nmol)$	extracellular amount of lipids at time t
$f(t,r)$	-	cell density at time t with respect to radius r
V_{em}	$\frac{4\pi}{3}6^3$ (μm^3)	volume of an empty adipocyte (zero lipid) [4]
V_ℓ	$1.091\,10^6$ $(\mu m^3.nmol^{-1})$	volume taken by 1 $nmol$ of triglyceride [118]
α	- $(nmol.\mu m^{-2}.h^{-1})$	surface-limited rate in lipogenesis
κ	- $(nmol)$	constant of the limiting term in lipogenesis
ρ	- (μm)	cell size threshold of the Hill function in lipogenesis
β	31.25 $(nmol.h^{-1})$	basal lipolysis rate [119]
γ	0.27 $(nmol.\mu m^{-2}.h^{-1})$	surface-limited rate in lipolysis [119]
χ	- $(nmol)$	constant of the limiting term in lipolysis
D	- $(\mu m^2.h^{-1})$	diffusion coefficient for size fluctuations
λ	- $(nmol)$	total amount of lipids

We note that assuming $f^\infty(r)$ is known for all $r \in [r_{min}, r_{max}]$, then L^∞ is determined by the equation (5.11) and only depends on the unknown parameter λ. In parameter identifiability analysis and parameter estimation we assume that the cell size distribution is observed. So to simplify the dependency on parameters we consider L to be a parameter instead of λ. We thus replace L^∞ by a parameter L, and it leads to the following simplified model,

$$\begin{cases} (f^\infty)'(r) = \dfrac{1}{D} v(r) f^\infty(r), & (5.12) \\ \displaystyle\int_{r_{min}}^{r_{max}} f^\infty(r)dr = 1, & (5.13) \\ v(r) = \dfrac{V_\ell}{4\pi}\left(\alpha \dfrac{L}{L+\kappa}\dfrac{\rho^3}{\rho^3+r^3} - \dfrac{(\beta+\gamma r^2)}{r^2}\dfrac{V(r)-V_{em}}{V(r)-V_{em}+V_\ell\chi}\right), & (5.14) \end{cases}$$

where the unknown parameters to be estimated are α, L, κ, ρ, χ and D.

Given those parameters, we can compute a stationnary solution of model (5.12)-(5.14) and we have for $r \in [r_{min}, r_{max}]$,

$$f(r) = \frac{\exp\left(\displaystyle\int_{r_{min}}^{r} \frac{1}{D} v(s)ds\right)}{\displaystyle\int_{r_{min}}^{r_{max}} \exp\left(\int_{r_{min}}^{r} \frac{1}{D} v(s)ds\right)dr}. \quad (5.15)$$

This solution can be computed numerically and when possible, the integrals are computed explicitly otherwise a trapezoid rule is used. Typically, in the computation, a radius step of 0.1 μm is considered and an interpolation is applied to compute f at any radius.

The model can represent a bimodal distribution of cell size

We first study the impact of the diffusion parameter that is the main change with respect to model in [118]. Figure 5.1 shows solutions computed numerically with the equation (5.15) for a given set of parameters. The model is able to qualitatively reproduce a bimodal distribution of cell size as measured in rats. Upon investigation of equations (5.12)-(5.14), it is immediate that the number of extremal points of f, and their locations, will depend only on the parameters that appear in the velocity v (5.14). However, variations in the value of the diffusion parameter also impact the size distribution: increasing the diffusion reduces the difference between the height of the two peaks and the density value at the nadir (lowest point between the two peaks) increases with diffusion.

In the model of Soula et al. [118], an individual-based Monte Carlo technique (20,000 cells) has been performed leading to a large computational time. It was then very hard to perform quantitative comparison with measurements. The proposed model enables a fast computation of the cell size distribution by computing directly a stationary solution with equation (5.15). It is now possible to perform quantitative comparison with measured size distribution and estimate parameters.

Prior to this parameter estimation, we study which parameters are likely to be estimated with the available data through model parameter identifiability analysis and parameter estimation on synthetic data.

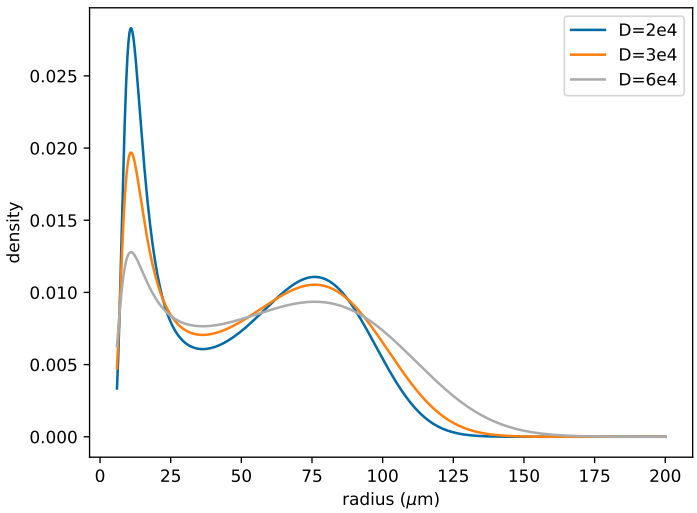


Figure 5.1: **Computed stationary solutions** from eq. 5.15 with three values for diffusion parameter. The other parameters are fixed to values reported in Table 5.1 and $L = 3\,nmol$, $\alpha = 0.29\,nmol.\mu m^{-2}.h^{-1}$, $\kappa = 0.001\,nmol$, $\rho = 200\,\mu m$, $\chi = 0.0035\,nmol$.

5.1.3 Model identifiability and parameter estimation

Parameter identifiability analysis

We perform an identifiability analysis of the unknown parameters of the model: α, L, κ, ρ, χ and D. We define a parameterized model $\mathcal{M}(\theta)$ derived from equations (5.12)-(5.14) and study its parameter identifiability which is an intrinsic property: from [32], the model $\mathcal{M}$ is said to be globally identifiable in $\theta \in \Theta$ if

$$\forall \tilde{\theta} \in \Theta, \mathcal{M}(\theta) = \mathcal{M}(\tilde{\theta}) \Rightarrow \theta = \tilde{\theta}.$$

The parametric structure of model (5.12)-(5.14) is complex in the sense that it includes non-linear functions in which some parameters are combined in a product. This might result in redundancies in the model – only a smaller set of unknown parameters can be estimated – or in a non-identifiable model [24].

To study the parametric structure of the model, we first set the observed outputs,

$$x_1 = f^{\infty},\ x_2 = r$$

and we introduce the following quantities to re-parameterize the model:

$$\theta_1 = \frac{\alpha L}{\beta(L+\kappa)},\ \theta_2 = \rho^3,\ \theta_3 = V_\ell \chi \text{ and } \theta_4 = \frac{4\pi D}{V_\ell \beta}. \tag{5.16}$$

We obtain the system parameterized by $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ the vector of unknown quantities (assumed to be strictly positive),

$$\left\{ \begin{array}{rcl} \dfrac{dx_1}{dr} & = & \dfrac{1}{\theta_4}\left(\theta_1 \dfrac{1}{1+\frac{x_2^3}{\theta_2}} - \dfrac{1+\frac{\gamma}{\beta}x_2^2}{x_2^2}\dfrac{\frac{4}{3}\pi x_2^3 - V_{em}}{\frac{4}{3}\pi x_2^3 - V_{em} + \theta_3}\right)x_1, \\ \dfrac{dx_2}{dr} & = & 1. \end{array} \right. \tag{5.17}$$

We recall that the values of V_{em}, β and γ are known (see Table 5.1).

We investigate the identifiability of unknown parameters using the `Structural identifiability Toolbox` of Maple [2]. It is based on the Structural Identifiability ANalyser (SIAN) algorithm which combines differential algebra and Taylor series approaches [60, 61]. From an input ODE model, a polynomial equations system is generated and the associated Gröbner basis is computed to assess the identifiability. This method ranks parameters in three categories: *globally identifiable*, *locally but not globally identifiable* and *non-identifiable*. A parameter θ_k is said to be locally identifiable if there is a finite set of possible values for θ_k given the observation. When a parameter is neither locally nor globally identifiable, it is called non-identifiable.

Applied to the system (5.17), SIAN algorithm returns that all the quantities $\theta_k, k \in \{1, \ldots, 4\}$ are globally identifiable. Going back to the model parameters in equations (5.12)-(5.14), the parameters V_ℓ, β are known and the function $\rho \mapsto \rho^3$ is bijective so assuming the cell size distribution is observed, the set of identifiable quantities is

$$\left\{ \frac{\alpha L}{L+\kappa}, \rho, \chi, D \right\}.$$

We notice that we need at least the values of (L, κ), (L, α) or (α, κ) to uniquely estimate α, κ or L respectively. Only a combination of these values can be uniquely retrieved when a size distribution $f(r)$ is given for all $r \in [r_{min}, r_{max}]$.

Parameter estimation procedure

Thanks to the parameter identifiability analysis, we know which parameters or parameter combinations we can expect to estimate from size distribution. We now need a procedure to estimate these parameters and we want to verify this procedure on a benchmark case: synthetic data.

Minimization algorithm To define a procedure to estimate model parameters, we first introduce a cost function. We want to minimize this function to compare the model output and the measurements. Then, we choose an algorithm to minimize this function.

Let θ be the parameter vector to be estimated. We denote by N the number of measured radii for the considered observation. Given the vector of measured radii, $(r_i)_{i=1,\dots,N}$, we estimate θ by minimizing the cost function defined as follows,

$$\mathcal{L}(\theta) = -\sum_{i=1}^{N} \log(f(r_i, \theta)) \tag{5.18}$$

where $f(r_i, \theta)$ is the value of a density f, solution of the model, computed at (measured) radius r_i with the parameter vector θ.

To find the optimal parameter values, we use the Covariance Matrix Adaptation Estimation Strategy (CMA-ES) algorithm [56]. In this algorithm, from initial parameters, new possible solutions are sampled with a multivariate normal distribution. The covariance matrix depends on a step-size control introduced to enhance the exploration of parameter space. A weighted combination of the best candidates is then selected according to the value of the cost function (5.18) and it is updated with the covariance matrix. These steps are repeated until termination criteria are reached. At each generation, this method takes into account recombination, mutation and selection of the possible candidates as an evolution algorithm.

Estimation of $(\theta_1, \rho, \theta_3, \theta_4)$ is performed with CMA-ES using cell size distribution as observation (we replace $\theta_2 = \rho^3$ by ρ). The vector of parameters is also scaled to have components of similar order of magnitude (scaling factors are $[\theta_1\, 10^2, \rho\, 10^{-3}, \theta_3\, 10^{-4}, \theta_4\, 10^2]$). Finally, to test the impact of the initial guess on the algorithm results, we perform 100 runs of CMA-ES with different initial parameters, we report the mean and standard deviation of these runs.

In order to run the CMA-ES algorithm, we used `cma` Python package [1]. The `fmin2` function of this package is used with default parameters and an initial standard deviation of 0.05 (in each coordinate). The files to run parameter estimation are available on `https://plmlab.math.cnrs.fr/audebert/adipocyte_size_modeling`.

Parameter estimation on synthetic data We first estimate parameters with data generated with the model (synthetic data). To generate such data, we compute the solution of the model for chosen parameters with equation (5.15). Then, from the obtained density, $10,000$ samples are drawn leading to a first synthetic data set. To mimic the true measurements we also consider a second type of synthetic data where on the $10,000$ samples only radii greater than $10\mu m$ are observed. With this procedure, we want to assess the impact of missing data on the parameter estimation. To quantify the precision of the parameter estimation we compute a relative error defined by $\mathcal{E} = |p - p_e|/p$, with p_e the parameter estimated value and p the true value of the parameter (chosen to generate synthetic data).

Two different parameter vectors are used to obtain synthetic data sets (*synthetic data set 1* and *synthetic data set 2*). The second column of Table 5.2 sums up the chosen parameter values (true). The parameter estimation is performed for both synthetic data sets without and with missing observations (Table 5.2 columns 3 to 8).

Table 5.2: **Results of parameter estimation procedure performed on synthetic data sets without and with missing data.** First two columns display the parameter names and true values for both synthetic data sets. Columns 3 and 4 present the estimated parameters for complete data sets (10,000 samples), it shows the average over 100 estimations with different initial guesses and standard deviations. The fifth column sums up the difference between true parameter and its estimation with a relative error in percentage. The three last columns present the same values for the same data sets with missing observations: only radii over $10\mu m$ are observed (samples $> 10\mu m$). All estimations are performed with CMA-ES algorithm of `fmin2` function from `cma` Python package. We choose the default parameters and an initial standard deviation of 0.05 (in each coordinate). The parameters are scaled to have similar sensitivity ($[\theta_1 .10^2, \rho .10^{-3}, \theta_3 .10^{-4}, \theta_4 .10^2]$).

synthetic data set 1		10,000 samples			samples $> 10\mu m$		
parameter	true	esti. value	std	rel. err.	esti. value	std	rel. err.
θ_1	$9.60\,10^{-3}$	$9.61\,10^{-3}$	$1\,10^{-11}$	0.2%	$9.62\,10^{-3}$	$2\,10^{-11}$	0.3%
ρ	$1.50\,10^{2}$	$1.50\,10^{2}$	$1\,10^{-6}$	0.2%	$1.49\,10^{2}$	$2\,10^{-6}$	0.8%
θ_3	$2.18\,10^{3}$	$2.17\,10^{3}$	$5\,10^{-5}$	0.6%	$2.09\,10^{3}$	$2\,10^{-4}$	4.2%
θ_4	$7.37\,10^{-3}$	$7.2\,10^{-3}$	$2\,10^{-10}$	2.3%	$7.35\,10^{-3}$	$4\,10^{-10}$	0.3%
synthetic data set 2		10,000 samples			samples $> 10\mu m$		
parameter	true	esti. value	std	rel. err.	esti. value	std	rel. err.
θ_1	$9.92\,10^{-3}$	$9.92\,10^{-3}$	$1\,10^{-11}$	0.04%	$9.91\,10^{-3}$	$1\,10^{-10}$	0.1%
ρ	$2.00\,10^{2}$	$2\,10^{2}$	$1\,10^{-6}$	0.2%	$2.01\,10^{2}$	$5\,10^{-6}$	0.6%
θ_3	$3.27\,10^{2}$	$3.12\,10^{2}$	$2\,10^{-5}$	4.8%	$5.39\,10^{2}$	$4\,10^{-4}$	65%
θ_4	$1.11\,10^{-2}$	$1.12\,10^{-2}$	$2\,10^{-10}$	1.7%	$1.12\,10^{-2}$	$1\,10^{-9}$	1.2%

Columns 3 and 4 in Table 5.2 display the average and the standard deviation of the estimated parameter values over the 100 runs. We note that the differences between the 100 estimations can be neglected, showing that the initial guess has no impact on the estimation.

In both synthetic data cases, when the estimation is performed with the complete data set, the estimated parameter values are similar to the true values with relative errors smaller than 5% (Table 5.2 column 5).

One can notice a difference between the two data sets when the estimation is performed with missing observations in the data. The last three columns of Table 5.2 show that depending on the considered data set more or less information is lost when we observe only cells with radii larger than a threshold. In *synthetic data set 1*, the impact on the parameter estimation is relatively small and relative errors remain below 5%. In *synthetic data set 2*, we are able to correctly estimate the values of θ_1, ρ and θ_4 but the information about parameter θ_3 seems lost, and the relative error increases to 65%.

The number of observed cells is reduced in these data sets and not in the same way in each set. On synthetic data we know exactly the percentage of information that is missing. In *synthetic data set 1* when we remove samples larger than $10\,\mu m$, 15% of the observation is missing, whereas in *synthetic data set 2* we remove 28% of the initial distribution. This difference may explain the poor estimation of θ_3 in *synthetic data set 2* with missing observations.

The parameter θ_3 is related to parameter χ that drives the lipolysis mechanism in the model (size reduction). If limited observations exist on small radius, we can expect that this parameter is difficult to estimate.

Table 5.3: **Range of selected values for the parameters.** The first three columns show the parameter names, order and true value. For each data set, the estimated parameter value (column "esti. value") with CMA-ES method is subject to a maximum of 20% variation (column "esti. ±20%"). From this variation, a range of values is selected for each parameter (column "selec. values") allowing a maximum error rate of 0.1% on the value of the estimated cost function $\mathcal{L}$. For each parameter 1,000 samples are generated

synthetic data set 1				10,000 samples			samples $> 10\mu m$	
parameter	order	true	esti. value	esti. ±20%	select. values	esti. value	esti. ±20%	select. values
θ_1	10^{-3}	9.60	9.61	7.69 - 11.53	9.58 - 9.63	9.62	7.70 - 11.54	9.59 - 9.65
ρ	10^{2}	1.50	1.50	1.20 - 1.80	1.47 - 1.53	1.49	1.19 - 1.79	1.46 - 1.52
θ_3	10^{3}	2.18	2.17	1.74 - 2.60	2.05 - 2.29	2.09	1.67 - 2.51	1.91 - 2.29
θ_4	10^{-3}	7.37	7.20	5.76 - 8.64	6.54 - 8.02	7.35	5.88 - 8.82	6.58 - 8.32
synthetic data set 2				10,000 samples			samples $> 10\mu m$	
parameter	order	true	esti. value	esti. ±20%	select. values	esti. value	esti. ±20%	select. values
θ_1	10^{-3}	9.92	9.92	7.94 - 11.90	9.90 - 9.95	9.91	7.92 - 11.89	9.86 - 9.95
ρ	10^{2}	2.00	2.00	1.60 - 2.40	1.97 - 2.03	2.01	1.61 - 2.41	1.99 - 2.05
θ_3	10^{3}	3.27	3.12	2.49 - 3.74	2.69 - 3.58	5.39	4.31 - 6.47	4.32 - 6.47
θ_4	10^{-2}	1.11	1.12	0.90 - 1.34	1.05 - 1.21	1.12	0.90 - 1.34	0.98 - 1.28

From estimated parameter values to parameter intervals The identifiability analysis ensures that the minimization problem should have only one solution and the estimation procedure computes this solution. Here, we want to compute intervals of parameter values for which the cost function remains close to its minimum. Our approach follows the strategy of ABC method where parameters are sampled from a prior distribution and are then selected according to a criterion based on the evaluation of the model output [123].

To sample a parameter θ_i, a new parameter $\bar{\theta}_i$ is first generated uniformly in $[0.8\hat{\theta}_i, 1.2\hat{\theta}_i]$ where $\hat{\theta}_i$ is the estimated parameter value obtained with the CMA-ES algorithm. Then, the cost function is computed with parameter $\bar{\theta}_i$ while the other parameters are fixed at their estimated values. The parameter is selected if the cost function is below 0.1% of $\mathcal{L}(\hat{\theta})$. This threshold was set to investigate the parameter space with small changes on cell size distribution. Note that the parameter sampling is performed one at a time. This strategy is repeated until 1,000 replicates are selected per parameter.

Table 5.3 shows for each parameter the considered range of values and the selected intervals for each synthetic data set. For synthetic data sets without missing observations, the range of values selected by the procedure is reduced in comparison with the initial one and contains the true parameter. This analysis gives an information on the range of accepted values for each parameter. We note that, in *synthetic data set 1*, the model output seems less sensitive to parameter θ_4 that has the largest range of selected values. In *synthetic data set 2* the largest range of selected values is for parameter θ_3.

In data sets with missing observations, the selected ranges are not impacted for *synthetic data set 1* (small difference for θ_3). In *synthetic data set 2*, the loss of information about small cells leads to the selection of the total initial interval for parameter θ_3 (±20% of the estimated value) and an important increase of the selected range for θ_4 (almost twice the length) compared to the case without missing observations. As observed in section 5.1.2, parameter D (hence θ_4) controls the relative heights of both modes in the cell size distribution. This can explains that data sets with missing observations on small sizes lead to higher uncertainty on θ_4. These results are in agreement with the computed relative errors of the previous paragraph (Table 5.2).

Table 5.4: **Parameter estimation with adipocyte size distributions measured in rats**. The first column is the parameter names. Over 32 estimations with the different animal cell size distributions, the mean is presented in the second column, the standard deviation in the third column and the fourth column is the relative standard deviation *i.e* the ratio of standard deviation over mean. The parameters are estimated with CMA-ES algorithm of `fmin2` function from `cma` Python package (with 100 initial guesses).

parameters	mean	std	RSD
θ_1	$9.6\,10^{-3}$	$2.8\,10^{-4}$	0.03
ρ	$1.57\,10^{2}$	$0.25\,10^{2}$	0.16
θ_3	$2.24\,10^{3}$	$1.07\,10^{3}$	0.47
θ_4	$8.21\,10^{-3}$	$2.58\,10^{-3}$	0.31

5.1.4 Application to adipocyte size distribution measured in rats

Measurements of adipocyte size distribution

The measured cell size distributions used to perform parameter estimation come from previous experiments [119] and data from [63], but this part of the experiment has not been published. Here, only adipocyte size distributions of animals in normal physiological conditions are considered.

We use two data sets of size distribution in retroperitoneal adipose tissue for a total of 32 male Wistar rats (20 rats METAJ, aged between 20 and 24 months, Charles River, L'Arbresle, France and, 12 rats EMPA, 12-week-old, Le Genest-Saint-Isle, France). Cell size distributions were measured with Beckman Coulter Multisizer IV (Beckman Coulter, Villepinte, France) [88]. Due to limitation in measurement techniques, only cell radii larger than $7.5\mu m$ for the first experiment and $10\mu m$ in the second were measured. Each animal cell size distribution is composed of a minimum of 6,000 cell radii.

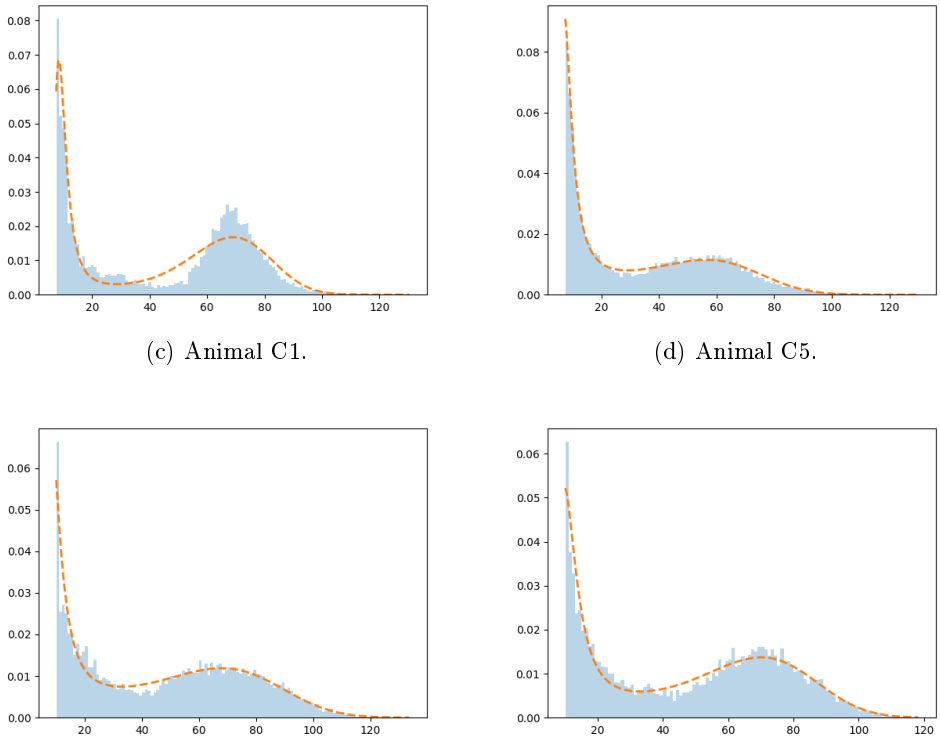
Parameter estimation with measured data

The estimation procedure validated on synthetic data is now applied to measured size distributions. Parameter estimation is performed with CMA-ES algorithm with radius distributions measured for 32 rats in the same experimental conditions. Figure 5.2 shows four examples of model-data fitting (the model fitting results of the 32 rats are available on `https://plmlab.math.cnrs.fr/audebert/adipocyte_size_modeling`). These results show the ability of the model to reproduce different types of cell size distribution. The height of each peak is not always correctly captured. This could be related to the loss of information due to missing observation for small cells in experimental data. In addition, the nadir is always underestimated by the model. We hypothesize that we are missing a process in the model to properly capture this point. However, the overall size distribution obtained with the model is in good agreement with the measured one.

Table 5.4 shows the mean, standard deviation and relative standard deviation (RSD) of the estimated parameter values obtained in the 32 rats. The RSD are relatively small for θ_1 and ρ, showing that the size distribution of adipocytes for rats in the same experimental conditions can be characterized with parameters in the same value ranges. The variability in the population is larger for parameter θ_3 and θ_4 (larger RSD). However, the previous analysis on synthetic data showed that less confidence in the estimation is expected for these parameters, especially θ_3.

For each animal, accepted parameter ranges are also computed following the procedure described in section 5.1.3 (Table 5.5). Figure 5.3 displays for each parameter the estimated value

(a) Animal A1. (b) Animal B1.



(c) Animal C1. (d) Animal C5.

Figure 5.2: **Comparison model-data**. Four examples (over 32) of adipocyte radius distributions (in μm) as histograms in rat in normal physiological conditions and model output computed (dash lines) with estimated parameters (see section 5.1.4). The parameter estimations are performed with CMA-ES algorithm of `cma` Python package by minimizing the function $\mathcal{L}$ eq. 5.18.

for each animal with the range of selected values (dots and bars). As expected, the parameter ranges are larger for parameters θ_3 and θ_4 compared to parameters θ_1 and ρ. Figure 5.3 also shows the mean (dash red line) and the standard deviation (gray area) over the rat population for each estimated parameter. It enables to compare the amplitude of the range of accepted values for each parameter for each animal with the variability within the population. We can see that for each parameter the range of accepted values is always smaller than the standard deviation in the population. It shows that the largest standard deviation within the population obtained for θ_3 and θ_4 (Table 5.4) should not be attributed to less confidence in the estimations.

The range of selected values of parameter θ_3 in rats population is between 1070 and 4429 μm^3. From this range of values, we can compute a range of radii for which the lipolysis term becomes mainly a surface based mechanism (i.e. $(V(r) - V_{em})/(V(r) - V_{em} + \theta_3) > 0.95$). We find radii in the range $17.2 - 27.3\,\mu m$.

Similarly, for lipogenesis, the parameter θ_1 is estimated within the rats population between 0.0092 and $0.010 \mu m^{-2}$. We remind that this quantity is a combination of parameters : $\theta_1 = \frac{\alpha L}{\beta(L+\kappa)}$ and parameter β is known [119]. We then obtain an estimation of $\frac{\alpha L}{(L+\kappa)}$ between 0.29 and $0.31\,nmol.\mu m^{-2}.h^{-1}$. In the case of high available lipids, L is large and we can assume $\frac{L}{L+\kappa} \sim 1$.

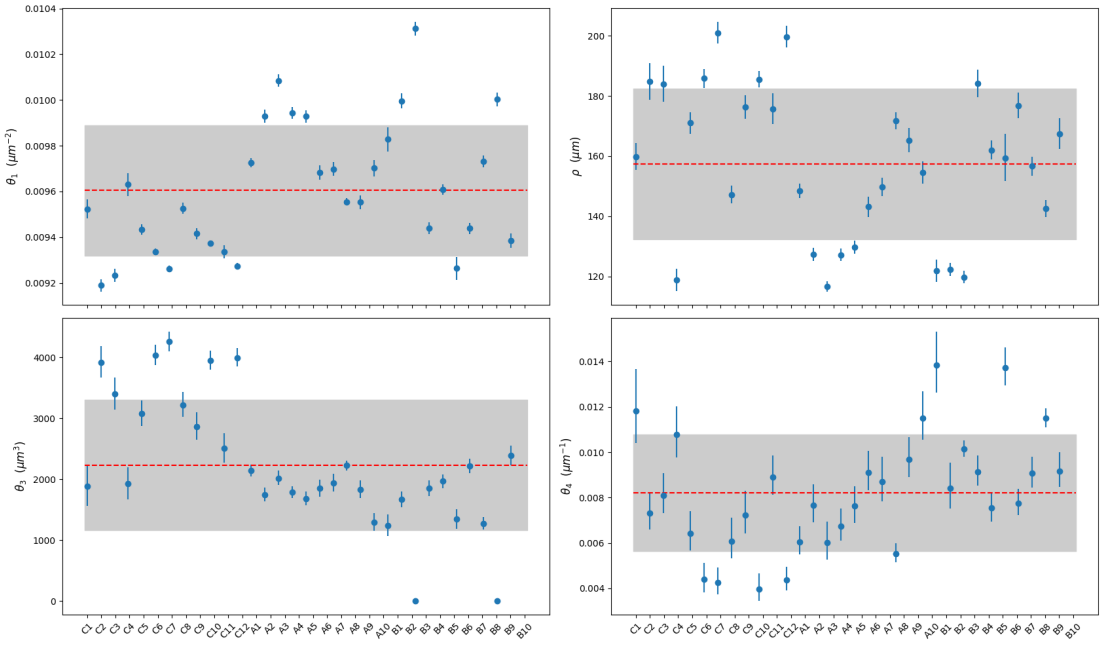


Figure 5.3: **Group variability and range of selected values.** Upper left and right figures display the results for parameters θ_1 and ρ. Lower left and right figures show the results for parameters θ_3 and θ_4. For each estimated parameter the average over the population is shown with dash red line and the gray area is one standard deviation around the average computed over the population (values are reported in Table 5.4). For each parameter, the estimated value for each animal is displayed with dots and the bar represents the range of selected values. These ranges consist in values of the parameter (assuming the 3 others are fixed) for which the maximal cost function is 0.1% of the obtained cost function with the estimation (see section 5.1.3). All numerical values are reported in Table 5.5.

Under this assumption, the parameter α is estimated between 0.29 and 0.31 $nmol.\mu m^{-2}.h^{-1}$. An alternative case is for low L, then we can assume $\frac{L}{L+\kappa} \sim L$ and the estimated values of θ_1 provide an estimation for αL.

The cell size threshold ρ of the Hill function in lipogenesis term is estimated in the range $115 - 204\mu m$. Above this threshold, the term $\rho^3/(\rho^3 + r^3)$ is smaller than 0.5 and limits the growth of the cell.

Sensitivity analysis

In order to investigate the differences between model output and measured cell size distribution, a sensitivity analysis is performed. Sensitivity analysis is a local analysis and quantifies how sensitive the model output is to parameter changes. We choose to apply the Sobol' method [117]. The sensitivity indices are based on the decomposition of the output variance at each cell size point.

The first order index measures the singular effect of a parameter on the model output. It represents the contribution part of the parameter alone in the variability of model output. A high value of this index indicates a high contribution of the parameter, which means that the

Table 5.5: **Parameter estimation results on measured adipocyte radius distribution in 32 rats.** First column is the animal identification. Estimation is performed with CMA-ES algorithm of `fmin2` function from `cma` Python package by minimizing the cost function $\mathcal{L}$, see eq. (5.18). The second to fourth columns show each parameter estimated value for each rat averaged over 100 runs with different initial guesses and the standard deviations are in brackets. For each estimated parameter, considering a maximum change of 20% of its estimated value, 1,000 samples are selected with a maximal error rate of 0.1% of the cost function value. The range of selected values of each parameter is given in the last four columns. These ranges consist in values of the parameter (assuming the other are fixed) for which the maximal cost function is 0.1% of the obtained cost function with the estimation. One can note that animals *B3* and *B9* have a value of θ_3 that is estimated to be zero ($10^{-12}/10^{-13}$). Indeed, these animals show particular cell size distributions with a very large number of small cells which can be due to a measurement artifact.

	estimated values				selected ranges			
animal	$\theta_1\,10^{-3}$ (std 10^{-11})	$\rho\,10^{2}$ (std 10^{-6})	$\theta_3\,10^{3}$ (std 10^{-4})	$\theta_4\,10^{-3}$ (std 10^{-10})	$\theta_1\,10^{-3}$	$\rho\,10^{2}$	$\theta_3\,10^{3}$	$\theta_4\,10^{-3}$
C1	9.52 (3.12)	1.60 (2.94)	1.89 (2.31)	11.8 (7.95)	9.49 - 9.56	1.56 - 1.64	1.59 - 2.21	10.5 - 13.5
C2	9.19 (1.05)	1.85 (2.16)	3.92 (1.07)	7.31 (3.43)	9.17 - 9.21	1.79 - 1.90	3.70 - 4.16	6.67 - 8.10
C3	9.23 (1.33)	1.84 (2.54)	3.40 (1.50)	8.10 (4.82)	9.21 - 9.26	1.79 - 1.89	3.17 - 3.65	7.40 - 8.99
C4	9.63 (2.27)	1.19 (1.63)	1.92 (1.52)	10.8 (5.75)	9.59 - 9.68	1.15 - 1.22	1.70 - 2.16	9.88 - 11.9
C5	9.43 (1.75)	1.71 (1.99)	3.08 (1.34)	6.43 (3.25)	9.41 - 9.45	1.68 - 1.74	2.90 - 3.27	5.75 - 7.31
C6	9.34 (1.41)	1.86 (1.92)	4.04 (1.13)	4.39 (2.33)	9.32 - 9.35	1.83 - 1.89	3.89 - 4.19	3.88 - 5.08
C7	9.26 (1.02)	2.01 (1.89)	4.26 (0.91)	4.26 (2.02)	9.25 - 9.27	1.98 - 2.04	4.12 - 4.42	3.80 - 4.86
C8	9.53 (2.15)	1.47 (1.86)	3.22 (1.36)	6.09 (3.29)	9.50 - 9.55	1.45 - 1.50	3.04 - 3.41	5.42 - 7.01
C9	9.42 (1.9)	1.76 (2.25)	2.87 (1.49)	7.24 (3.65)	9.39 - 9.44	1.73 - 1.80	2.68 - 3.08	6.52 - 8.20
C10	9.37 (1.81)	1.86 (2.19)	3.95 (1.35)	3.97 (2.25)	9.36 - 9.39	1.83 - 1.88	3.81 - 4.10	3.50 - 4.58
C11	9.34 (1.69)	1.76 (2.46)	2.51 (1.81)	8.90 (5.35)	9.31 - 9.36	1.71 - 1.80	2.30 - 2.73	8.19 - 9.76
C12	9.27 (0.95)	2.00 (1.80)	4.00 (0.83)	4.37 (1.78)	9.26 - 9.29	1.96 - 2.03	3.87 - 4.14	3.98 - 4.89
A1	9.73 (1.5)	1.48 (1.04)	2.14 (0.60)	6.06 (1.77)	9.71 - 9.74	1.46 - 1.50	2.05 - 2.23	5.55 - 6.67
A2	9.93 (2.23)	1.27 (1.12)	1.75 (0.74)	7.65 (2.87)	9.90 - 9.96	1.25 - 1.29	1.65 - 1.85	6.98 - 8.50
A3	10.1 (2.76)	1.17 (1.01)	2.02 (0.79)	6.01 (2.48)	10.1 - 10.1	1.15 - 1.18	1.92 - 2.13	5.35 - 6.86
A4	9.94 (2.01)	1.27 (1.03)	1.79 (0.60)	6.74 (2.17)	9.92 - 9.97	1.25 - 1.29	1.70 - 1.88	6.18 - 7.44
A5	9.93 (1.91)	1.30 (1.05)	1.68 (0.63)	7.62 (2.58)	9.90 - 9.95	1.28 - 1.32	1.59 - 1.79	6.98 - 8.42
A6	9.68 (1.44)	1.43 (1.15)	1.85 (0.62)	9.11 (2.70)	9.65 - 9.71	1.40 - 1.46	1.73 - 1.98	8.41 - 9.97
A7	9.70 (1.81)	1.50 (1.38)	1.93 (0.75)	8.72 (3.22)	9.67 - 9.72	1.47 - 1.53	1.81 - 2.07	7.92 - 9.68
A8	9.55 (0.98)	1.72 (1.07)	2.23 (0.49)	5.53 (1.43)	9.54 - 9.57	1.69 - 1.74	2.16 - 2.30	5.20 - 5.94
A9	9.55 (1.54)	1.65 (1.66)	1.83 (0.78)	9.70 (3.63)	9.53 - 9.58	1.62 - 1.69	1.71 - 1.97	8.98 - 10.1
A10	9.70 (1.71)	1.54 (1.35)	1.3 (0.77)	11.5 (4.09)	9.67 - 9.74	1.51 - 1.58	1.17 - 1.44	10.5 - 12.6
B1	9.83 (2.16)	1.22 (1.36)	1.24 (0.94)	13.8 (6.65)	9.78 - 9.88	1.19 - 1.25	1.09 - 1.40	12.7 - 15.2
B2	10.0 (2.31)	1.22 (1.06)	1.67 (0.70)	8.42 (2.80)	9.97 - 10.0	1.20 - 1.24	1.56 - 1.79	7.61 - 9.45
B3	10.3 (1.42)	1.20 (0.74)	$9.24\,10^{-16}$ ($1.52\,10^{-8}$)	10.1 (9.13)	10.3 - 10.3	1.18 - 1.22	0.76 - 1.09 10^{-15}	9.83 - 10.5
B4	9.44 (1.04)	1.84 (1.43)	1.85 (0.62)	9.14 (2.53)	9.41 - 9.46	1.80 - 1.88	1.74 - 1.97	8.60 - 9.81
B5	9.61 (1.16)	1.62 (1.06)	1.97 (0.61)	7.54 (2.35)	9.59 - 9.63	1.59 - 1.65	1.87 - 2.07	7.02 - 8.19
B6	9.26 (2.11)	1.59 (2.15)	1.34 (0.87)	13.7 (6.14)	9.22 - 9.31	1.52 - 1.67	1.21 - 1.49	13.0 - 14.5
B7	9.44 (0.93)	1.77 (1.2)	2.22 (0.55)	7.76 (2.44)	9.42 - 9.46	1.73 - 1.81	2.11 - 2.33	7.27 - 8.34
B8	9.73 (1.69)	1.57 (1.22)	1.28 (0.73)	9.07 (3.23)	9.71 - 9.76	1.54 - 1.59	1.19 - 1.37	8.51 - 9.72
B9	10 ($1.17\,10^{5}$)	1.43 ($1.25\,10^{5}$)	$1.74\,10^{-6}$ (17.4)	11.5 ($8.22\,10^{4}$)	9.97 - 10.0	1.40 - 1.45	1.43 - 2.05 10^{-6}	11.1 - 11.9
B10	9.38 (1.52)	1.67 (1.98)	2.39 (0.81)	9.17 (3.68)	9.36 - 9.41	1.63 - 1.72	2.25 - 2.53	8.55 - 9.93

model output is highly sensitive to this parameter. The total order index enables to include the effects depending on parameter interactions (higher order indices).

The model output is the cell size distribution f computed with equation (5.15) for radii from $7.5\mu m$ to $140\mu m$. To study the influence of the estimated parameters, each parameter θ_i is uniformly distributed in a range of $\pm 1\%$ of estimated mean over the population of rats (Table 5.4). The change of $\pm 1\%$ in parameters values is chosen such that the adipocyte size distributions computed with these parameters are bimodal. Then, Saltelli algorithm is performed to explore the parameter space leading to the generation of $n(2d+2)$ parameter samples with a Monte-Carlo approach [109, 117]. We choose $n = 2048$ and $d = 4$ the number of parameters. The sensitivity analysis is performed using the `SALib` Python Library [117, 109, 20, 96].

Figure 5.4(a) shows cell size distributions ranges computed with parameters from the sampling design. With these small perturbations, a large variability is found between the cell size distributions around the two modes. The first mode of the adipocyte size density is represented

by cells with radii from 7.5 to $10\mu m$. Regarding large adipocytes, the higher densities present a high variability and correspond to adipocyte size values from 50 to $120\mu m$. These results illustrate the heterogeneity of cell sizes that can be obtained with the model with small changes in parameters.

Then, Sobol' indices are computed to determine which parameters are most influential on the cell size dynamic. The first-order indices are displayed for several radii and each parameter in Figure 5.4(b). The results indicate that parameter θ_1 explains the most the variations of cell sizes with a first-order sensitivity index between 0.6 and 1 for all radii. Interestingly, for the cells with radii around $40\mu m$, the index of θ_1 decreases and we notice that ρ index increases (index equals 0.36). It shows that parameter ρ around this point explains the variability of the model output up to 36%. The impacts of θ_3 and θ_4 are almost negligible on cells size distribution. From $r = 90\mu m$, the results show that the influence of θ_1 decreases whereas ρ becomes more influential and explains up to 18% of the output variability. The total-order sensitivity indices are also computed (not shown) and are similar to first-order indices, revealing that parameter interactions have a negligible influence on the adipocyte size distributions.

The sensitivity analysis suggests that the cell size dynamics in rats is mainly driven by the parameters depending on lipogenesis, and especially by θ_1 which represents the combination of the unknown parameters (α, κ, L).

Parameters θ_3 and θ_4, associated with lipolysis (through χ) and diffusion (D) respectively, have a negligible impact on the cell size dynamic along all cell sizes. This result confirms the difficulty to identify these parameters in practice and are in agreement with the largest ranges of selected parameter values. In addition, this study highlights the fact that the nadir is difficult to capture since we observe an opposite change in the parameter sensitivity around this radius. With this study we are able to explain the results of parameter estimation on the measured data.

5.1.5 Discussion

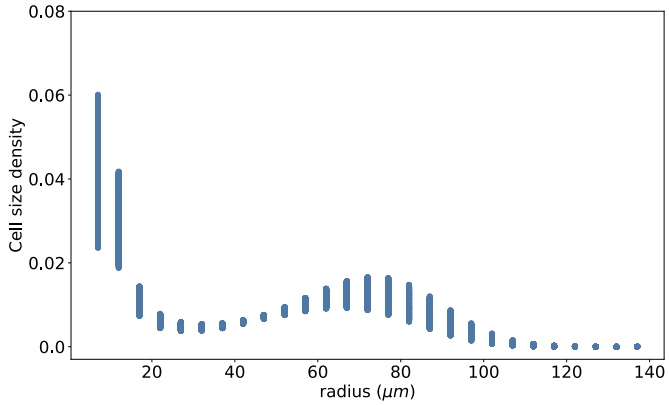
We presented a mathematical model to describe adipocytes cell size distribution, based on a partial differential equation and including lipid exchanges. With the formulation of a stationary solution we were able to solve numerically and efficiently this model. Prior to the estimation of parameter with measurements we analyzed which parameter can be identifiable and how reliable are the estimations.

The identifiability of unknown parameters was studied with a re-parameterized form of the model. We showed that only four quantities can be uniquely identified and that three of our parameters of interest are related. These three parameters cannot be identified separately with an observation of the cell size distribution only. However, we can identify the threshold radius ρ involved in lipogenesis, the lipolysis threshold χ as well as the diffusion coefficient D that describes cell size fluctuations.

The model calibration on synthetic data sets showed, in practice, an accurate estimation of the parameters. When we considered data sets with missing observations (similar to the measurements) we found that three over the four quantities can be correctly estimated.

The model parameters were estimated on 32 adipocyte size distributions measured in rats. With these estimated parameters, the overall distribution of cell size was captured. However, the nadir part of the distribution as well as the height of the modes were not perfectly reproduced. It is possible that the model is missing some aspect of the adipocyte size dynamics that would help to better capture the nadir. This is supported by the sensitivity analysis, that showed that the nadir part was not sensitive specifically to one of the four considered parameters. Therefore, it makes this part of the distribution difficult to fit. In addition, in the presented model, the diffusion parameter D via θ_4 affects linearly both lipogenesis and lipolysis. It would be interesting

(a) Cell size distributions



(b) Estimation of first-order Sobol' indices

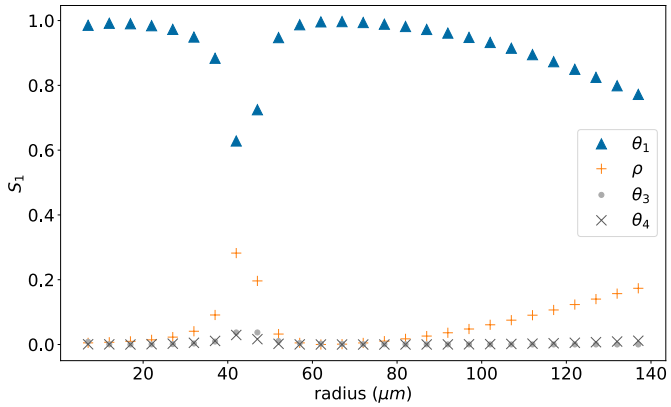


Figure 5.4: (a) A sample of cell size distributions. The parameter sampling design is constructed using Saltelli algorithm where each parameter is uniformly distributed in a range of values corresponding to $\pm 1\%$ of the mean of its estimated value in rats (Table 5.4). A number of $20,480$ samples giving bimodal distributions are generated to estimate the Sobol' indices. (b) Estimation of first-order Sobol' indices for θ_1, ρ, θ_3 and θ_4 using a Monte-Carlo based approach [117, 109, 20, 96].

to change this modeling assumption with a more complex diffusion process, impacting differently lipogenesis and lipolysis. For instance, considering a size dependent diffusion coefficient could improve the agreement between the model outputs and the observations.

We also think that our assumption regarding the normalization of the cell size distribution (it integrates to 1 between r_{min} and r_{max}) affects the fits (especially the height of the 2 modes). However, we have no background knowledge about the total number of adipocytes in the distribution. In addition, we know that the data collection does not include cells with a radius below a certain threshold. In [66], a formulation has been proposed to approximate the total cell number in a fat pad but to do this estimation, we need to have the fat pad mass which is not the case in our experimental data. An other way to solve this issue would be to introduce a parameter that quantifies the total number of cells. However with an additional parameter, we will lose parameter identifiability. Then, we might need to fix other unknown quantities, so this solution

only shifts the problem.

Nevertheless, we have estimated parameter values for 32 rats. We found a larger variability between rats in the estimated values of θ_3 and θ_4 (Figure 5.3). However, the sensitivity analysis showed that the model is less sensitive to these parameters (Figure 5.4). For θ_1 and ρ, the estimated values were more robust within the population leading us to believe that θ_1, ρ are less individual-specific parameters. However they could change if the estimation is performed with another species. This result suggests lipolysis (driven by χ) is more an individual-dependent process than lipogenesis (driven by θ_1 and ρ) that is more constant within the population.

Recruitment of new cells via adipogenesis was not included in our model. Since we were looking at the distribution of size at one specific time, this mechanism can be neglected. However, if one wants to represent longitudinal adipocyte size distributions specially in case of diet changes, this process should be considered. This will have an impact on the cell size distribution, especially for small cells, as suggested in [119]. Moreover, it is known that past diets affect the adipocyte size regulation and may be irreversible [66, 118]. Indeed, past diets could lead to a larger number of cells in the tissue. However, in the presented model, the number of cells is not explicitly considered. This assumption should be modified to take into account longitudinal size distributions and to be able to compare animals with different diets. In past works [66, 67, 68, 82], the authors have considered partial differential equation models that take into account a recruitment rate of new cells. Our model could be extended with this extra term for adipogenesis modeling.

We believe that the presented framework can be adapted to estimate model parameters with adipocyte size distribution in other species than rats and in different health conditions. It may enable to establish links between the mathematical model parameters and health conditions based on adipocyte size distribution observations. The final purpose is to be able to characterize and potentially classify the different obesity-related pathologies.

5.2 Similar methods for the diffusive Lifshitz-Slyozov model

In this section we perform similar parameter estimation from our diffusive Lifshitz-Slyozov model (1.32) to the same data. Since our model is in lipids, we use equation (1.8) to transform the data from radii to lipids. Given a vector of measured lipid amounts $(x_i)_{i=1,\dots,N}$, we estimate a set of parameters θ by minimizing the cost function :

$$\mathcal{L}(\theta) = -\sum_{i=1}^{N} \log(M(x_i, \theta)), \tag{5.19}$$

where $M(x, \theta)$ is the stationary solution of the diffusive Lifshitz-Slyozov model, in the form of equation (1.30a). The dependency on L is 'lost' in some sense because we see L as a parameter. Similarly, as in the previous study, the set of parameters to estimate is given by $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ where :

$$\theta_1 = \alpha \frac{L}{L + \kappa}, \quad \theta_2 = \rho, \quad \theta_3 = \chi, \quad \theta_4 = \varepsilon. \tag{5.20}$$

This set of parameters is slightly different from the one used in the previous section. This is due in part to the model being written in lipids and to the fact that the stationary solutions are different when considering a model with non-constant rate of diffusion.

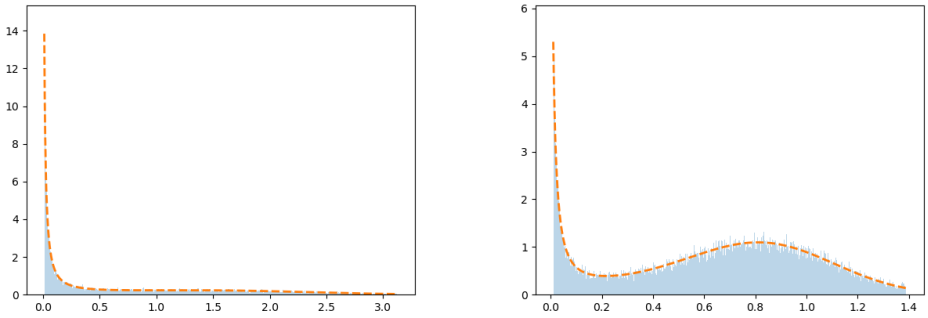


Figure 5.5: Examples of fitted models on synthetic data.

5.2.1 Estimations on synthetic data

In the same spirit as the previous study, we begin by studying how the CMA-ES method behaves on synthetic data. First we fix a set of target parameters θ_{target} for which we generate a stationary distribution. Then we take 10000 samples from this distribution to create a synthetic data set. Then we use the CMA-ES method to estimate the parameters of this synthetic data set. We perform 100 runs with different initial parameter sets and take the mean. Our choice of parameters is up to some rescaling similar to the one in the previous work, however we are unable to show that this set of parameters is identifiable. We also perform the estimation on censored data, meaning that in lipids we only consider data above 0.03 nmol. Both results are displayed in Table 5.6.

We can see that when considering the whole sample, the CMA-ES method is able to recover both θ_1 and ρ with acceptable errors. However the values of χ and ε are more difficult to estimate. This further supports the assumption that the set of parameters $\theta_1, \rho, \chi, \varepsilon$ may not be identifiable. Examples of fitted models for the whole sample are presented in Figure 5.5. However when considering censored samples with only lipid amounts above 0.003 nmol, the method performs poorly. From our observations, it seems that the first mode of the distribution is actually critical to the parameter estimation. For this reason, we choose not to censor small sizes when fitting with real data.

5.2.2 Estimations on biological data

From the study we previously made on synthetic data, it seems that for the diffusive Lifshitz-Slyozov model the set of identifiable parameters is either different from the one for the constant diffusive model or empty. Nonetheless we proceed with the method and compare it with the constant diffusion model (5.21). Since the model in the previous work is in radii, it can be tricky to compare stationary solutions and results from parameter estimation. Hence we compare our results for the diffusive Lifshitz-Slyozov model to a similar model in lipids, with a constant diffusion rate :

$$\begin{cases} \partial_t h + \partial_x(vh) = D\partial_x^2 h, & \text{(5.21a)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} xh(t,x)\mathrm{d}x = \lambda. & \text{(5.21b)} \end{cases}$$

For this version of the model the stationary solutions are given by :

synthetic data set 1		10,000 samples			samples > 0.003 nmol		
parameter	true	esti. value	std	rel. err.	esti. value	std	rel. err.
θ_1	$3.0\,10^{-1}$	$3.02\,10^{-1}$	$3\,10^{-9}$	0.6%	$3.28\,10^{-1}$	$8.70\,10^{-3}$	9.4%
ρ	$1.50\,10^{2}$	$1.44\,10^{2}$	$8\,10^{-6}$	4%	$1.11\,10^{2}$	$0.23\,10^{2}$	25.6%
χ	$2.05\,10^{-3}$	$1.64\,10^{-3}$	$7\,10^{-10}$	20%	$1.34\,10^{-2}$	$4.84\,10^{-2}$	>100%
ε	$5.0\,10^{-2}$	$5.48\,10^{-2}$	$1\,10^{-8}$	10%	$1.41\,10^{-1}$	$5.58\,10^{-2}$	>100%
synthetic data set 2		10,000 samples			samples > 0.003 nmol		
parameter	true	esti. value	std	rel. err.	esti. value	std	rel. err.
θ_1	$3.10\,10^{-1}$	$3.04\,10^{-1}$	$5\,10^{-9}$	2%	$3.64\,10^{-1}$	$2.87\,10^{-8}$	17.3%
ρ	$2.00\,10^{2}$	$2.10\,10^{2}$	$9\,10^{-6}$	5%	$1.49\,10^{2}$	$1.50\,10^{-5}$	25.0%
χ	$3.00\,10^{-4}$	$3.03\,10^{-4}$	$5\,10^{-10}$	1%	$6.41\,10^{-18}$	$4.87\,10^{-18}$	>100%
ε	$1.0\,10^{-1}$	$8.01\,10^{-2}$	$2\,10^{-8}$	20%	$2.53\,10^{-1}$	$8.04\,10^{-8}$	>100%

Table 5.6: **Results of parameter estimation procedure performed on synthetic data sets without and with missing data.** First two columns display the parameter names and true values for both synthetic data sets. Columns 3 and 4 present the estimated parameters for complete data sets (10,000 samples), it shows the average over 100 estimations with different initial guesses and standard deviations. The fifth column sums up the difference between true parameter and its estimation with a relative error in percentage. The three last columns present the same values for the same data sets with missing observations: only lipid amounts over 0.003 nmol are observed (samples 0.003 nmol).

parameters	mean	std	RSD
θ_1	$2.83\,10^{-1}$	$3.95\,10^{-3}$	0.014
ρ	$2.51\,10^{2}$	$0.40\,10^{2}$	0.15
χ	$3.89\,10^{-3}$	$1.75\,10^{-3}$	0.45
ε	$8.68\,10^{-3}$	$2.71\,10^{-3}$	0.31

Table 5.7: Parameter estimation with adipocyte size distributions measured in rats : mean, standard variation and relative standard variation.

$$\begin{cases} M_L(x) = \dfrac{C}{D}\exp\left(\dfrac{1}{D}\displaystyle\int_0^x v(y,L)\mathrm{d}y\right), & \text{(5.22a)} \\ L + \displaystyle\int_{\mathbb{R}_+} xM_L(x)\mathrm{d}x = \lambda, & \text{(5.22b)} \end{cases}$$

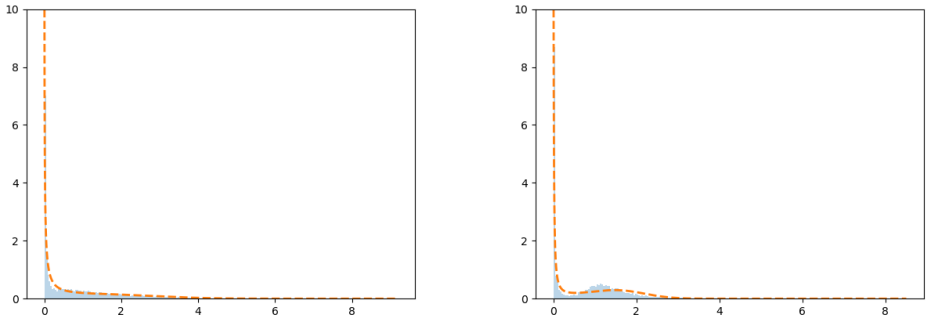
where C is a constant that ensures $\int_{\mathbb{R}_+} M_L(x)\mathrm{d}x = m$. In practice, we take $m = 1$.

Censoring the large data points

We begin by fitting the model on the whole set of data points for an individual. Unfortunately when looking at the histograms of the data in lipids, the bimodal feature of the distribution is less visible than when plotting in radii. The mean over all 32 animals is presented in Table 5.7.

From what we observe on Figure 5.6, we often recover the shape of data for small cells and large cells but we are not able to recover the nadir, the local minimum between the two maxima.

To compensate for this, we choose to censor the data and remove the larger size in the data that corresponds to 20% of the total weight. The goal of this censoring is to concentrate



(a) Animal C1 (b) Animal CA12

Figure 5.6: **Comparison model-data :** 2 examples of size distributions in lipids as histograms and model outputs (dash-lined) computed from the results of the CMA-ES method. On the left the fit reproduces the whole distribution. On the right the method does not reproduce the nadir and the position of the second mode.

the estimation method on data points that are important to the feature we want to recover : bimodality.

Values for the estimated parameters are presented in Table 5.8. We recover values of similar order of magnitude for θ_1, ρ and ε. However the value of χ is different in this case, and has a very large standard deviation. This is due to the fact that the result of the estimation of χ is very different from individual to individual and is either of the same order of the not-censored case or of order 10^{-15} which can be considered to be a numerical zero. Data for which we obtain χ to be a numerical zero are the ones that do not display a 'strong' bimodality, in the sense that they are almost strictly decreasing. Examples of both cases are presented in Figure 5.7.

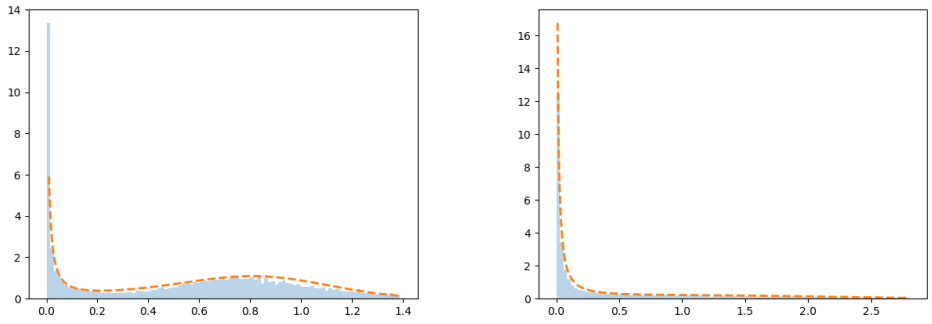
Hence we believe the model presents some troubles to capture the heterogeneity of biological data. However, with some data preparation and appropriate estimation method, one may recover good fitting between the model and the data.

parameters	mean	std	RSD
θ_1	$2.90\,10^{-1}$	$5.00\,10^{-3}$	0.02
ρ	$1.97\,10^{2}$	$2.97\,10^{1}$	0.15
χ	$8.45\,10^{-4}$	$1.53\,10^{-3}$	1.81
ε	$9.95\,10^{-3}$	$4.11\,10^{-3}$	0.41

Table 5.8: **Results on large size-censored data.** Parameter estimation with adipocyte size distributions measured in rats : mean, standard variation and relative standard variation.

Comparison with the constant diffusion Lifshitz-Slyozov model (5.21)

The results of the CMA-ES method with the constant diffusion model (5.21) are presented in Table 5.9. We use the CMA-ES method on the whole data set or as before with a censored data set where we remove points above 80% of the total mass. The orders of magnitude between the two choices are the same but the model fits are not comparable as presented in Figure 5.8. In average, the fitted models are better for the censored data sample. Nonetheless, when compared



(a) Animal A3 (b) Animal C2

Figure 5.7: **Comparison model-data.** 2 examples of size distributions in lipids as histograms and model outputs (dash-lined) computed from the results of the CMA-ES method censoring the last 20% of the data. The left figure exhibits a 'strong' bimodality and we are able to properly reproduce the nadir, while on the right, the bimodality is harder to see and parameter χ is estimated to be zero.

with Figure 5.7, the non-constant diffusive Lifshitz-Slyozov model has a better reproducibility of data.

However the fitting of the constant diffusive model for the censored sample is relatively better in average than its non-censored version, but not as good as the non-constant diffusive Lifshitz-Slyozov model.

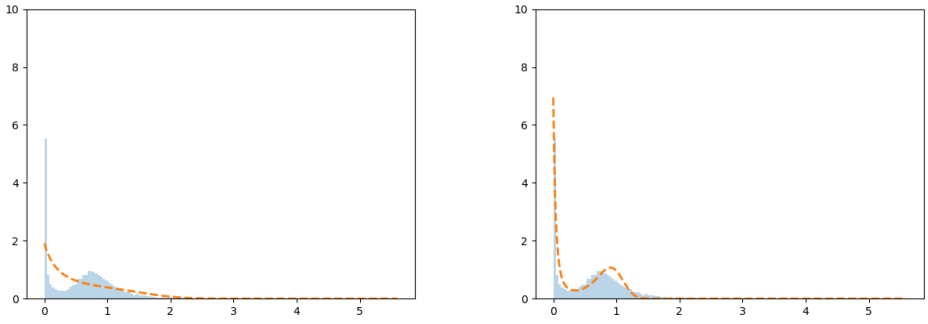
	whole data set			data $< 80\%$		
parameters	mean	std	RSD	mean	std	RSD
θ_1	$2.79\,10^{-1}$	$9.90\,10^{-4}$	0.035	$2.83\,10^{-1}$	$2.48\,10^{-3}$	0.008
ρ	$3.31\,10^{2}$	$3.22\,10^{1}$	0.097	$2.53\,10^{2}$	$3.46\,10^{1}$	0.14
χ	$5.36\,10^{-17}$	$3.37\,10^{-17}$	1	$4.54\,10^{-17}$	$2.81\,10^{-17}$	0.62
D	3.86	1.23	0.32	1.68	$5.32\,10^{-1}$	0.32

Table 5.9: **Results for the constant diffusive Lifshitz-Slyozov model.** Parameter estimation with adipocyte size distributions measured in rats for the constant diffusive Lifshitz-Slyozov model (5.21).

Conclusion

In the submitted work presented in Section 5.1, we introduced a version of the Lifshitz-Slyozov model for adipose cells in radii with constant diffusion. We investigate the parameter identifiability and estimate parameter values with CMA-ES method. We identified a set of 4 parameters that are identifiable and validate the method on synthetic data. We obtained an estimate for the surface rate of lipogenesis in line with the literature.

In Section 5.2 we use the same method on the diffusive Lifshitz-Slyozov model (1.32). We show that the model still performs well on synthetic data. However when applied to the data from rats, we obtained that censoring large sizes actually increases the fit between model and data. We recover values of parameters in the same order of magnitude as the first study. To be able



(a) Animal A3

(b) Animal A3 with censored data points

Figure 5.8: **Comparison model-data for the constant diffusive model.** 2 examples of size distribution in lipids as histograms and model output (dash-lined) computed from the results of the CMA-ES method. The left figure is the result for the whole data sample while the right is the result with data points larger than 80% of the total mass censored.

to compare with the constant diffusion model, we introduce a constant diffusive Lifshitz-Slyozov model in lipids and performer parameter estimation with this model. We obtained that the non-constant diffusive Lifshitz-Slyozov model performs better overall than the constant diffusive model.

Chapter 6

Numerical schemes for mixture theory models with filling constraint: application to biofilm ecosystems

6.1 Abstract

This work was done during the summer school CEMRACS 2022 on Transport in Physic, Biology and Urban traffic. The goal of this project was to apply a numerical scheme from mixture theory to a biofilm model. This models describes a biofilm as three phase system: cells, extra-cellular matrix and liquid, where equation for volume part and velocity is provided for each phase. Obviously the three volume parts should sum to one, which leads to a constraint equation on the system. Validating this constraint at the numerical level proves difficult which is why we introduce the use of numerical schemes from mixture theory that were developed with this numerical constraint in mind. We show that we are able to recover the typical behavior of an algae biofilm while validating the constraint at all time point. We also introduce the inclusion of substrates to the system which complicates numerical computations but still validates the constraint. This work is submitted [12] and small editions were made for ease of reading in the manuscript.

6.2 Introduction

There are many physical cases of flows composed of different gas or liquids interacting together. For example, tissue bodies and tumors can be described as a set of interacting viscoelastic materials. Powder-snow avalanches can be described as a mixture of fluid phases. Similarly, the rheology of the gut microbiota and its interactions with chyme (a mixture of partially digested food and water) and the host can be modeled using mixture theory [75]. Complex flows can also be found in many engineering applications involving multiphase systems such as boiling water in nuclear reactors. Therefore, the framework of mixture theory is a common tool to model and study complex flows.

Mathematical models based on mixture theory take the form of systems of partial derivative equations, coupled with algebraic constraints. The theoretical analysis of such systems and the characterization of the qualitative properties of the solutions are extremely complicated [11, 52, 69, 113]. Thus, it is important to develop efficient numerical methods able to accurately capture the solutions [16, 28, 29, 41].

In this article, we are interested in applying mixture models to describe biofilm dynamics. Indeed, mixture theory revealed a powerful approach to represent microbial biofilms where a consortium of cells is embedded in a polymeric structure [29, 102, 103].

In mixture theory, the unknowns of the model are requested to satisfy certain constraints. As far as the continuous equations are considered, several equivalent formulations of these constraints can be derived and used to bring out the properties of the model. However, the preservation of these constraints by a numerical scheme is a challenge and, once a discretization setup has been adopted, it is not clear that all the formulations of the constraints remain equivalent. This issue can induce a loss of stability and accuracy, and eventually a dramatic loss of key physical properties of the simulated flows. Thus, we adapt and extend the numerical scheme proposed in [13] in order to preserve these constraints. The numerical scheme will be tested and illustrated with a multiphasic model representing the development of a photosynthetic biofilm, with the application for biofuel, protein, or drug production.

The paper is organized as follows. The first section is dedicated to the mixture theory framework with a presentation of the simplified model used to test our numerical scheme. The second section details the numerical scheme and its properties. The third section presents the results and comparison with standard numerical schemes.

6.3 Mixture theory framework: application to biofilms

6.3.1 Mixture theory framework

The mixture theory framework [107], also known as mixture mechanics or continuum mechanics for fluid dynamics, enables describing multi-phasic systems at the mesoscopic scale which is an intermediary scale between microscopic and macroscopic scales. It was introduced in the 1960s by Truesdell [124, 126, 125] and generalizes Navier-Stokes equations to multi-phasic systems. The mixture theory framework assumes that each component of the mixture might be present at every point in space and at any time. Moreover, the system's physical properties (ex. viscosity, incompressibility, ...) are naturally included.

Consider a mixture of k components indexed by $\alpha \in [\![1, k]\!]$. Each component is locally described by its volumetric mass density ρ_α, its volume fraction ϕ_α, and its local velocity v_α. The volume fraction represents the relative volume occupied by a component in an elementary normalized piece of volume. Thus, assuming that there is no vacuum they satisfy the algebraic

constraint

$$\sum_{\alpha=1}^{k} \phi_\alpha = 1. \tag{6.1}$$

The mixture dynamic depends on mass transfers which are modeled through mass balance equations (6.2a) and the local forces applied to the system which are accounted for through momentum balance equations (6.2b). Thus, for each component the state variables satisfy the equations:

$$\partial_t(\rho_\alpha \phi_\alpha) + \nabla_x \cdot (\rho_\alpha \phi_\alpha v_\alpha) = \Gamma_\alpha, \tag{6.2a}$$

$$\partial_t(\rho_\alpha \phi_\alpha v_\alpha) + \nabla_x \cdot (\rho_\alpha \phi_\alpha v_\alpha \otimes v_\alpha) + \nabla_x \pi_\alpha + \phi_\alpha \nabla_x P = \nabla_x \cdot (\phi_\alpha \tau_\alpha) + F_\alpha + \phi_\alpha \rho_\alpha \boldsymbol{g} + \Gamma_\alpha v_\alpha, \tag{6.2b}$$

where Γ_i is the mass exchange term, π_α is the elastic tensor, P is the common pressure, τ_α the viscous stress tensor, F_α the friction forces, and $\boldsymbol{g}$ the gravity force. Depending on the considered application some forces can be neglected and some others might be added.

Depending on the targeted application one can add for each component an extra equation for the evolution of the density ρ_α. Nevertheless, liquids are weakly compressible, especially when pressure variations are small. Therefore, in most cases, for liquids the component densities ρ_α can be assumed constant. When all the component volumetric mass densities are assumed constant, the mass balance equations (6.2a) are equivalent to $\partial_t(\phi_\alpha) + \nabla_x \cdot (\phi_\alpha v_\alpha) = \Gamma_\alpha/\rho_\alpha$. Then summing these equations for each phase leads to the pseudo incompressibility constraint:

$$\nabla_x \cdot \left(\sum_\alpha \phi_\alpha v_\alpha \right) = \sum_\alpha \frac{\Gamma_\alpha}{\rho_\alpha}. \tag{6.3}$$

This means that the local divergence of the averaged mixture velocity is equal to the local volume variation induced by mass exchanges.

The elastic tensor π_α can be interpreted as the internal pressure of the component. There are several ways to model this term depending on the nature of the component. When the component α represents particles, as in [13], there is a close-packing limit. This property can be enforced by using an appropriate expression for π_α as

$$\pi_\alpha = \gamma_\alpha \frac{\phi_\alpha^{\beta_\alpha}}{\phi_\alpha^\star - \phi_\alpha}, \quad \text{with} \quad \gamma_\alpha > 0, \quad \text{and} \quad \beta_\alpha > 1, \tag{6.4}$$

where $0 < \phi_\alpha^\star < 1$ is the so-called close-packing volume fraction limit. When the component α represents softer material like living tissues it can take the form of standard pressure law:

$$\pi_\alpha = \gamma_\alpha \left(\frac{\phi_\alpha}{\phi_\alpha^\star} \right)^{\beta_\alpha}, \quad \text{with} \quad \gamma_\alpha > 0, \quad \text{and} \quad \beta_\alpha \geqslant 1, \tag{6.5}$$

where $0 < \phi_\alpha^\star < 1$ is a threshold, see [29, 105]. More complex laws, based on the Flory–Huggins theory:

$$\pi_\alpha = -\gamma_\alpha \big(\ln(1 - \phi_\alpha) + \phi_\alpha + \phi_\alpha^2 \big), \quad \text{with} \quad \gamma_\alpha > 0, \tag{6.6}$$

enable accounting for colligative properties at low concentrations, see [31].

The viscous stress tensor τ_α is defined by

$$\tau_\alpha = \mu_\alpha \phi_\alpha \big(\nabla v_\alpha + {}^t\nabla v_\alpha - \tfrac{2}{3} (\nabla \cdot v_\alpha) \mathrm{Id} \big), \tag{6.7}$$

where the constant $\mu > 0$ stands for the component dynamic viscosity and ${}^t\nabla v_\alpha$ stands for the transpose of the velocity differential matrix.

The friction force F_α is induced by the difference in the relative speed of the mixture components:

$$F_\alpha = \sum_{\alpha' \neq \alpha} f_{\alpha,\alpha'}(v_{\alpha'} - v_\alpha) \tag{6.8}$$

with $f_{\alpha,\alpha'}$ the friction force law between the components pair α and α'. As a first approximation, it can be assumed that $f_{\alpha,\alpha'}$ is a strictly positive constant. However, the friction between two components should vanish when one of them disappears. Thus, a more realistic alternative is to consider that friction depends on the local composition and use instead $f_{\alpha,\alpha'}(\phi_\alpha \phi_{\alpha'})^{r_{\alpha,\alpha'}}$. Nevertheless, the total momentum conservation principle enforces that

$$\sum_\alpha F_\alpha = 0.$$

Dissolved components, like substrate, can be included. A dissolved component p within a phase α is described through its concentration θ_p. In addition to the transport by the phase, it can also diffuse within the phase at a rate D_p. Thus, the mass balance equations for a dissolved component within the phase α writes:

$$\partial_t(\rho_\alpha \phi_\alpha \theta_p) + \nabla_x \cdot (\rho_\alpha \phi_\alpha \theta_p v_\alpha) - \nabla_x \cdot (\rho_\alpha \phi_\alpha D_p \nabla_x \theta_p) = \Gamma_p. \tag{6.9}$$

where again the source term Γ_p represents the mass exchange associated to component p.

6.3.2 Mixture model for biofilm

We focus on a simplified 1D model for biofilms. Biofilms are made of microorganisms $\mathcal{A}$ (microalgae, bacteria, or a consortium of both) and an extra-cellular matrix $\mathcal{E}$. The biofilm is usually immersed in water $\mathcal{L}$. Therefore, according to mixture theory framework, see section 6.3.1, each component $\alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$ is described through three macroscopic variables: the mass density ρ_α, the volume fraction ϕ_α, and the velocity v_α. By definition, the volume fractions satisfy at any time the algebraic volume-filling constraint (6.1) which reads in this case: $\phi_\mathcal{A} + \phi_\mathcal{E} + \phi_\mathcal{L} = 1$. In the one-dimensional case, the mass balance equations (6.2a) writes:

$$\partial_t \rho_\alpha \phi_\alpha + \partial_x \rho_\alpha \phi_\alpha v_\alpha = \Gamma\alpha, \qquad \alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}. \tag{6.10}$$

In this context, the volumetric mass densities ρ_α can be assumed to be constant. Thus, the mixture averaged velocity satisfies the pseudo incompressibility constraint (6.11) which writes here:

$$\partial_x\big(\phi_\mathcal{A} v_\mathcal{A} + \phi_\mathcal{E} v_\mathcal{E} + \phi_\mathcal{L} v_\mathcal{L}\big) = \frac{\Gamma\mathcal{A}}{\rho_\mathcal{A}} + \frac{\Gamma\mathcal{E}}{\rho_\mathcal{E}} + \frac{\Gamma\mathcal{L}}{\rho_\mathcal{L}}. \tag{6.11}$$

For biofilms, there are various biological processes to be taken into account. The main processes are growth, extra-cellular matrix excretion, and death. These reactions are schematically represented in Table 6.1. The parameters η_α are pseudo-stoichiometric coefficients that quantify how much a reactant (ex. liquid, algae, substrate, ...) or a product (ex. algae, extra-cellular matrix, ...) is consumed or produced when a reaction occurs. The functions ψ_i are the reaction

	Biological reaction representation		
Name	Reactant(s)	Rate	Product(s)
Growth	$\eta_{\mathcal{L}}\mathcal{L} + \eta_{\mathcal{S}}\mathcal{S}$	$\xrightarrow{\psi_g}$	$\mathcal{A}$
Excretion	$\mathcal{A}$	$\xrightarrow{\psi_e}$	$\mathcal{E}$
Death	$\mathcal{A}$	$\xrightarrow{\psi_d}$	$\eta_{\mathcal{E}}\mathcal{E} + (1-\eta_{\mathcal{E}})\mathcal{L}$

Table 6.1: Schematic representation of the biochemical reactions considered in the model.

rates. They describe the speed at which reactions take place as a function of the local composition of the mixture. The source terms read as follows:

$$\Gamma\mathcal{A} = \psi_g - \psi_e - \psi_d, \qquad \Gamma\mathcal{E} = \psi_e + \eta_{\mathcal{E}}\psi_d, \qquad \Gamma\mathcal{L} = (1-\eta_{\mathcal{E}})\psi_d - \eta_{\mathcal{L}}\psi_g.$$

The growth is mainly induced by substrate ($\mathcal{S}$) assimilation and liquid ($\mathcal{L}$) absorption. However, as a first approximation, we assume that the substrate is in excess. Thus, the growth rate ψ_g takes the form $\psi_g = \mu_g \rho_{\mathcal{A}} \phi_{\mathcal{A}} \phi_{\mathcal{L}}$, where μ_g is the maximal growth rate. The extra-cellular matrix excretion ψ_e and the death rate ψ_d are assumed to be proportional to the quantity of microalgae, thus $\psi_e = \mu_e \rho_{\mathcal{A}} \phi_{\mathcal{A}}$ and $\psi_d = \mu_d \rho_{\mathcal{A}} \phi_{\mathcal{A}}$ respectively. Nevertheless, biofilms are very complex ecosystems and the biological processes are very simplified here. Thus, a model extension accounting for substrate and oxygen is presented in section 6.6.3.

In the one-dimensional case and neglecting the gravity, for $\alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$ the momentum balance equations simplify into:

$$\partial_t(\rho_\alpha \phi_\alpha v_\alpha) + \partial_x\big(\rho_\alpha \phi_\alpha v_\alpha^2\big) + \partial_x \pi_\alpha = -\phi_\alpha \partial_x P + \frac{4}{3}\partial_x(\mu_\alpha \phi_\alpha \partial_x v_\alpha) + F_\alpha + \Gamma_\alpha v_\alpha.$$

To keep the model as simple as possible, let us assume that the elastic tensor takes the form of a pressure law, see equation (6.5), for the tissues (ie. algae and extra-cellular matrix). Since the liquid phase is not elastic this term is null for the liquid, namely $\pi_{\mathcal{L}} = 0$. Similarly, let us assume that the friction forces are constant and symmetric. Thus, in the expression (6.8) for F_α, the term $f_{\alpha,\alpha'}$ for $(\alpha, \alpha') \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}^2$ and $\alpha \neq \alpha'$ are constant and such that $f_{\alpha,\alpha'} = f_{\alpha',\alpha}$.

The model is supplemented by boundary conditions. Let $\Omega = [0, L]$ be the domain and $\partial\Omega$ its boundary. In 1D, the domain should correspond to a biofilm core drilling in the orthogonal axis of the support where the biofilm develops. The velocities at the bottom of the domain, which corresponds to the surface on which the biofilm develops, vanish $v_\alpha(0) = 0$, $\alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$. However, the velocity on the top must satisfy a constraint induced by the incompressibility constraint (6.3). Indeed, the integration over the whole domain of equation (6.3) combined with the null velocity at the bottom leads to

$$\big(\phi_{\mathcal{A}} v_{\mathcal{A}} + \phi_{\mathcal{E}} v_{\mathcal{E}} + \phi_{\mathcal{L}} v_{\mathcal{L}}\big)(x = L) = \int_0^L \left(\frac{\Gamma_{\mathcal{A}}}{\rho_{\mathcal{A}}} + \frac{\Gamma_{\mathcal{E}}}{\rho_{\mathcal{E}}} + \frac{\Gamma_{\mathcal{L}}}{\rho_{\mathcal{L}}}\right) dx$$

To enforce this condition, let assume that on the top, the velocities are given by $v_{x=L} = \int_0^L \left(\frac{\Gamma_{\mathcal{A}}}{\rho_{\mathcal{A}}} + \frac{\Gamma_{\mathcal{E}}}{\rho_{\mathcal{E}}} + \frac{\Gamma_{\mathcal{L}}}{\rho_{\mathcal{L}}}\right) dx$.

Remark. *Although there is no biophysical reason to impose the equality between the top velocities, this assumption remains acceptable in this context. Indeed, our focus concerns the biofilm development and the final time considered prevents the biofilm to reach the top of the domain. Therefore, in our context, the hypothesis that all top velocity are equals should not affect the dynamics of the biofilm growth.*

6.3.3 Synthesis of model equations

According to the previous section the PDE system under consideration writes:

$$\phi_{\mathcal{A}} + \phi_{\mathcal{E}} + \phi_{\mathcal{L}} = 1, \tag{6.12a}$$

$$\partial_t \phi_{\mathcal{A}} + \partial_x(\phi_{\mathcal{A}} v_{\mathcal{A}}) = \frac{\Gamma\mathcal{A}}{\rho_{\mathcal{A}}}, \tag{6.12b}$$

$$\partial_t \phi_{\mathcal{E}} + \partial_x(\phi_{\mathcal{E}} v_{\mathcal{E}}) = \frac{\Gamma\mathcal{E}}{\rho_{\mathcal{E}}}, \tag{6.12c}$$

$$\partial_t \phi_{\mathcal{L}} + \partial_x(\phi_{\mathcal{L}} v_{\mathcal{L}}) = \frac{\Gamma\mathcal{L}}{\rho_{\mathcal{L}}}, \tag{6.12d}$$

$$\partial_t(\rho_{\mathcal{A}}\phi_{\mathcal{A}} v_{\mathcal{A}}) + \partial_x\big(\rho_{\mathcal{A}}\phi_{\mathcal{A}} v_{\mathcal{A}}^2\big) + \partial_x \pi_{\mathcal{A}} = -\phi_{\mathcal{A}}\partial_x P + \frac{4}{3}\partial_x\Big(\mu_{\mathcal{A}}\phi_{\mathcal{A}}\partial_x v_{\mathcal{A}}\Big) + F_{\mathcal{A}} + \Gamma\mathcal{A} v_{\mathcal{A}}, \tag{6.12e}$$

$$\partial_t(\rho_{\mathcal{E}}\phi_{\mathcal{E}} v_{\mathcal{E}}) + \partial_x\big(\rho_{\mathcal{E}}\phi_{\mathcal{E}} v_{\mathcal{E}}^2\big) + \partial_x \pi_{\mathcal{E}} = -\phi_{\mathcal{E}}\partial_x P + \frac{4}{3}\partial_x\Big(\mu_{\mathcal{E}}\phi_{\mathcal{E}}\partial_x v_{\mathcal{E}}\Big) + F_{\mathcal{E}} + \Gamma\mathcal{E} v_{\mathcal{E}}, \tag{6.12f}$$

$$\partial_t(\rho_{\mathcal{L}}\phi_{\mathcal{L}} v_{\mathcal{L}}) + \partial_x\big(\rho_{\mathcal{L}}\phi_{\mathcal{L}} v_{\mathcal{L}}^2\big) = -\phi_{\mathcal{L}}\partial_x P + \frac{4}{3}\partial_x\Big(\mu_{\mathcal{L}}\phi_{\mathcal{L}}\partial_x v_{\mathcal{L}}\Big) + F_{\mathcal{L}} + \Gamma\mathcal{L} v_{\mathcal{L}}, \tag{6.12g}$$

where the sources terms $(\Gamma_\alpha)_\alpha$, the elastic tensors $(\pi_\alpha)_\alpha$ and the drag forces $(F_\alpha)_\alpha$ are given by:

$$\Gamma\mathcal{A} = \psi_g - \psi_e - \psi_d, \qquad \Gamma\mathcal{E} = \psi_e + \eta_{\mathcal{E}}\psi_d, \qquad \Gamma\mathcal{L} = (1-\eta_{\mathcal{E}})\psi_d - \eta_{\mathcal{L}}\psi_g, \tag{6.13a}$$

$$\psi_g = \mu_g \rho_{\mathcal{A}}\phi_{\mathcal{A}}\phi_{\mathcal{L}}, \qquad \psi_e = \mu_e \rho_{\mathcal{A}}\phi_{\mathcal{A}}, \qquad \psi_d = \mu_d \rho_{\mathcal{A}}\phi_{\mathcal{A}}, \tag{6.13b}$$

$$\pi_\alpha = \gamma_\alpha\left(\frac{\phi_\alpha}{\phi_\alpha^\star}\right)^{\beta_\alpha}, \qquad \alpha \in \{\mathcal{A},\mathcal{E}\}, \tag{6.13c}$$

$$F_\alpha = \sum_{\alpha'\neq\alpha} f_{\alpha,\alpha'}(v_{\alpha'} - v_\alpha), \qquad f_{\alpha,\alpha'} = f_{\alpha',\alpha} \qquad \alpha \in \{\mathcal{A},\mathcal{E},\mathcal{L}\}. \tag{6.13d}$$

The system (6.12) is supplemented with the boundary conditions

$$v_\alpha(x=0) = 0, \qquad \text{and} \qquad v_\alpha(x=L) = \int_0^L \left(\frac{\Gamma_{\mathcal{A}}}{\rho_{\mathcal{A}}} + \frac{\Gamma_{\mathcal{E}}}{\rho_{\mathcal{E}}} + \frac{\Gamma_{\mathcal{L}}}{\rho_{\mathcal{L}}}\right)\mathrm{d}x,$$

for all $\alpha \in \{\mathcal{A},\mathcal{E},\mathcal{L}\}$.

The initial data for the volume fraction can be chosen arbitrarily provided they are biologically relevant. However, to enforce the algebraic constraint on the sum over all the volume fractions (6.1), the velocities have to satisfy the incompressibility constraint (6.11) at all times and therefore the initial velocities must verify this constraint as well. Thus, the initial velocities are defined through a pressure P computed using the incompressibility constraint, see section 6.4.3, by $v_\alpha^0 = \tilde{v}_\alpha^0 - \frac{\partial_x P}{\rho_\alpha}$ where $\tilde{v}_\alpha^0$ is the initial desired velocity. Here, the system is assumed to be initially at rest so $\tilde{v}_\alpha^0 = 0$ for all the phases.

Most of the parameters come from [13] or [102, 103]. The viscosity coefficient for microalgae and the extra-cellular-matrix are taken from [99]. All the parameter values are gathered in table 6.2.

6.4 Numerical scheme

In this section, we are interested in the numerical approximation of the PDE system (6.12). Nevertheless, the general principles and in particular the treatment of the pseudo incompressibility constraint remain valid in a more general context. In such PDE systems, the pressure is

Symbol	Name	Value	Unit
μ_g	Microalgae maximal growth rate	2	1/day
μ_e	Microalgae maximal ECM excretion rate	0.4	1/day
μ_d	Microalgae maximal death rate rate	0.2	1/day
μ_r	Microalgae maximal respiration rate rate	0.2	1/day
$\eta_{\mathcal{A}}$	Microalgae pseudo-stoichiometric coefficient	1.0	∅
$\eta_{\mathcal{L}}$	Liquid pseudo-stoichiometric coefficient	0.96	∅
$\eta_{\mathcal{S}}$	Substrate pseudo-stoichiometric coefficient	$8.67 \cdot 10^{-2}$	∅
$\eta_{\mathcal{C}}$	Inorganic carbon pseudo-stoichiometric coefficient	0.146	∅
$\eta_{\mathcal{O}}$	Oxygen pseudo-stoichiometric coefficient	0.106	∅
$\eta_{\mathcal{E}}$	Liquid pseudo-stoichiometric coefficient for death	0.90	∅
$K_{\mathcal{I}}$	Light parameter	0.1	∅
τ	Light absorption coefficient for the biofilm	$2.5 \cdot 10^4$	m^{-1}
$\mathcal{I}_{\mathrm{surf}}$	Light intensity at the surface	100	$\mu\mathrm{mol\,m^{-2}s^{-1}}$
$\mathcal{I}_{\mathrm{opt}}$	Optimal light intensity	100	$\mu\mathrm{mol\,m^{-2}s^{-1}}$
$K_{\mathcal{S}}$	Substrate half saturation coefficient	$6.2 \cdot 10^{-8}$	kg/L
$K_{\mathcal{C}}$	Inorganic carbon half saturation coefficient	$4.4 \cdot 10^{-6}$	kg/L
$K_{\mathcal{O}}$	Oxygen threshold for growth	$3.2 \cdot 10^{-5}$	kg/L
$n_{\mathcal{O}}$	Oxygen exponent for growth	14	∅
$K_{\mathcal{I}}$	Light coefficient for Haldane law	0.1	∅
K_r	Oxygen half saturation coefficient	$1.0 \cdot 10^{-6}$	kg/L
$\theta_{\mathrm{in},\mathcal{S}}$	Input concentration for substrate	$4 \cdot 10^{-5}$	kg/L
$\theta_{\mathrm{in},\mathcal{C}}$	Input concentration for inorganic carbon	$10 \cdot 10^{-5}$	kg/L
$\theta_{\mathrm{in},\mathcal{O}}$	Input concentration for oxygen	$7.2 \cdot 10^{-6}$	kg/L
$D_{\mathcal{S}}$	Diffusion coefficient for substrate	$1.47 \cdot 10^{-4}$	m^2/day
$D_{\mathcal{C}}$	Diffusion coefficient for inorganic carbon	$1.80 \cdot 10^{-4}$	m^2/day
$D_{\mathcal{O}}$	Diffusion coefficient for oxygen	$1.98 \cdot 10^{-4}$	m^2/day
$\rho_{\mathcal{A}}$	Microalgae volumetric mass density	1050	$\mathrm{kg/m^3}$
$\rho_{\mathcal{A}}$	Extra-cellular matrix volumetric mass density	1050	$\mathrm{kg/m^3}$
$\rho_{\mathcal{L}}$	Liquid volumetric mass density	1025	$\mathrm{kg/m^3}$
$\phi^\star_{\mathcal{A}}$	Microalgae close packing threshold	0.75	∅
$\gamma_{\mathcal{A}}$	Microalgae viscoelastic tensor coefficient	$1.2 \cdot 10^{-9}$	$\mathrm{kg\,m^{-1}day^{-1}}$
$\beta_{\mathcal{A}}$	Microalgae viscoelastic tensor exponent	1	∅
$\phi^\star_{\mathcal{E}}$	Extra-cellular matrix close packing threshold	0.75	∅
$\gamma_{\mathcal{E}}$	Extra-cellular matrix viscoelastic tensor coefficient	$1.2 \cdot 10^{-9}$	$\mathrm{kg\,m^{-1}day^{-1}}$
$\beta_{\mathcal{E}}$	Extra-cellular matrix viscoelastic tensor exponent	1	∅
$\mu_{\mathcal{L}}$	Liquid viscosity	10^{-3}	Pa s
$\mu_{\mathcal{A}}$	Microalgae viscosity	0.25	Pa s
$\mu_{\mathcal{E}}$	Extra-cellular matrix viscosity	0.75	Pa s
$f_{\mathcal{A},\mathcal{E}}$	Friction coefficient between $\mathcal{A}$ and $\mathcal{E}$	20	$\mathrm{kg\,m^{-3}day^{-1}}$
$f_{\mathcal{A},\mathcal{L}}$	Friction coefficient between $\mathcal{A}$ and $\mathcal{L}$	20	$\mathrm{kg\,m^{-3}day^{-1}}$
$f_{\mathcal{E},\mathcal{L}}$	Friction coefficient between $\mathcal{E}$ and $\mathcal{L}$	20	$\mathrm{kg\,m^{-3}day^{-1}}$

Table 6.2: Model parameters. The parameters come from [13, 102, 99].

defined through the volume filling constraint (6.1), namely $\phi_{\mathcal{A}} + \phi_{\mathcal{E}} + \phi_{\mathcal{L}} = 1$ for the considered model. The treatment of this constraint and thus the definition of the pressure is always an issue and requires specific treatment. To this end, the momentum equations are treated using a projection correction method inspired by the numerical method introduced by Chorin [25, 26, 27] and Temam [121] for incompressible viscous flows. In a nutshell, the momentum equation is decomposed using a time splitting to separate the contribution of the pressure as follows:

$$\partial_t(\rho_\alpha \phi_\alpha v_\alpha) + \partial_x\big(\rho_\alpha \phi_\alpha v_\alpha^2\big) + \partial_x \pi_\alpha = \frac{4}{3}\partial_x\Big(\mu_\alpha \phi_\alpha \partial_x v_\alpha\Big) + F_\alpha + \Gamma_\alpha v_\alpha, \tag{6.14a}$$

$$\partial_t(\rho_\alpha \phi_\alpha v_\alpha) + \phi_\alpha \partial_x P = 0. \tag{6.14b}$$

for $\alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$.

6.4.1 Projection correction method

Let us start with the presentation of the time discretization. Let $T \in \mathbb{R}^+$ be the final time and $(t_n)_{n\geqslant 0}$ a subdivision of $[0, T]$ such that $t_n = \sum_{k=0}^{n} \Delta t_k$. Consider $\alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$ a phase and its associated volume fraction ϕ_α and velocity v_α. Then, $\phi_\alpha^n(x)$ and $v_\alpha^n(x)$ denote, respectively, their approximation at time t_n. To shorten the notations, let us drop the space variable x and denote $\delta t = \Delta t_{n+1}$. Assuming that all the quantities are known at time t_n, the approximated solution at time $t_{n+1} = t_n + \delta t$ is computed using the following steps:

(a) Update the volume fractions according to the mass balance equations (6.12b)-(6.12d):

$$\phi_\alpha^{n+1} = \phi_\alpha^n - \frac{\delta t}{\rho_\alpha}\partial_x(\phi_\alpha^n v_\alpha^n) + \frac{\delta t}{\rho_\alpha}\Gamma_\alpha.$$

(b) Update the momentum equations without the contribution of the pressure term by solving the following system:

$$\begin{aligned}\phi_\alpha^{n+1} v_\alpha^{n+\frac{1}{2}} - \phi_\alpha^n v_\alpha^n \\ = \frac{\delta t}{\rho_\alpha}\left(-\partial_x\Big(\phi_\alpha^n (v_\alpha^n)^2\Big) - \partial_x \pi_\alpha^n + \frac{4}{3}\partial_x\Big(\mu_\alpha \phi_\alpha^{n+1}\partial_x v_\alpha^{n+\frac{1}{2}}\Big) + F_\alpha\Big(\phi^{n+1}, v^{n+\frac{1}{2}}\Big) + \Gamma_\alpha^n v_\alpha^n\right).\end{aligned} \tag{6.15}$$

(c) Compute the pressure using the incompressibility constraint (6.3). This step is detailed in subsection 6.4.3.

(d) Update the velocity using the pressure with:

$$v_\alpha^{n+1} = v_\alpha^{n+\frac{1}{2}} - \frac{\delta t}{\rho_\alpha}\partial_x P^{n+1}.$$

6.4.2 1D space discretization

Following [13], the space is discretized using staggered grids. This enables avoidance of any odd/even decoupling in the stencil of the discrete version of the system. Moreover, the use of staggered grids also allows to have or deduce naturally the quantity of interest (e. g. deduce the pressure gradient on the velocity mesh grid). Let $(x_i)_{i\in[\![0,I]\!]}$ be a regular subdivision of the domain Ω such that $x_i = i\Delta x$ with $\Delta x = \frac{L}{I}$ the mesh step. Let also define the mesh cell centers: $x_{i+\frac{1}{2}} = \big(i + \frac{1}{2}\big)\Delta x$ for $i \in [\![0, I-1]\!]$. The model variables are located:

- at the mesh cell centers for the volume fraction and the pressure: $\phi_{\alpha,i+\frac{1}{2}}$, $P_{i+\frac{1}{2}}$ for $0 \leqslant i \leqslant I-1$
- at the mesh cell edges for the velocities: $v_{\alpha,i}$ for $0 \leqslant i \leqslant I$.

Figure 6.1 gives an example of the staggered grids with the localization of model variables.

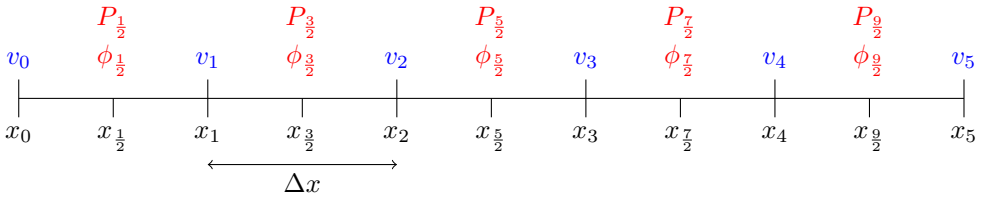


Figure 6.1: Regular staggered grid in one dimension for 5 mesh cells with the volume fractions and the velocities locations. The pressure P and the phase volume fractions $(\phi_\alpha)_\alpha$ are located at the mesh cell centers $\left(x_{i+\frac{1}{2}}\right)_{0\leqslant i\leqslant 4}$.

Model unknowns are discretized using a finite volume scheme. The transport terms in the mass balance equations (6.12b)-(6.12d) are written:

$$\phi_{\alpha,i+\frac{1}{2}}^{n+1} = \phi_{\alpha,i+\frac{1}{2}}^{n} - \frac{\delta t}{\rho_\alpha \Delta x}\Big(\mathcal{F}_{i+1}(\phi_\alpha^n, v_\alpha^n) - \mathcal{F}_i(\phi_\alpha^n, v_\alpha^n)\Big) + \frac{\delta t}{\rho_\alpha}\Gamma_{\alpha,i+\frac{1}{2}} \tag{6.16}$$

where $\mathcal{F}_i$ represents the numerical mass flux at the interface x_i, which is a function of the neighboring cells. There are multiple relevant choices for the definition of the numerical flux. For the sake of simplicity, to ensure stability and since it is well adapted to staggered grids, it is convenient to use upwind numerical flux. Thus, the discrete mass flux is defined by $\mathcal{F}_i(\phi, v) = \mathcal{F}^+\left(\phi_{i-\frac{1}{2}}, v_i\right) + \mathcal{F}_i^-\left(\phi_{i+\frac{1}{2}}, v_i\right)$ with

$$\mathcal{F}^+(\phi, v) = \begin{cases} 0 & \text{if } v \leqslant 0, \\ \phi v & \text{if } v > 0, \end{cases} \qquad \text{and} \qquad \mathcal{F}^-(\phi, v) = \begin{cases} \phi v & \text{if } v < 0, \\ 0 & \text{if } v \geqslant 0. \end{cases}$$

All the volume fractions are updated using equation (6.16). Thus, the volume-filling constraint enforcement is not guaranteed and depends on the strategy used to compute the pressure, see sections 6.4.3 and 6.5.2.

Remark. *To update the volume fractions and ensure volume-filling constraint enforcement another strategy consists to use equation* (6.16) *for all the components except one (usually the liquid) which is computed using the algebraic volume-filling constraint* (6.1)*:* $\phi_{\alpha'} = 1 - \sum_{\alpha\neq\alpha'} \phi_\alpha$ *as done in [28, 29, 102, 103].*

For the momentum balance equation, following [13], the transport term is also discretized using an upwind strategy based on the material velocity v, that is the momentum flux is defined by

$$\mathcal{G}_{i+\frac{1}{2}} = \frac{v_{\alpha,i}^n}{2}\left(\mathcal{F}^+\left(\phi_{i-\frac{1}{2}}, v_i\right) + \mathcal{F}^+\left(\phi_{i+\frac{1}{2}}, v_{i+1}\right)\right) + \frac{v_{\alpha,i+1}^n}{2}\left(\mathcal{F}^-\left(\phi_{i+\frac{1}{2}}, v_i\right) + \mathcal{F}^-\left(\phi_{i+\frac{3}{2}}, v_{i+1}\right)\right).$$

The other terms of equation (6.15) are discretized using standard approximations. Remark that interpolation on the dual mesh is required only for the zeroth order terms like the momentum

supply induced by mass exchanges or friction forces. For these terms, the approximation of the volume fraction on the dual mesh is obtained by approximating the volume fractions using the values in the neighboring cells: $\phi_i = \frac{1}{2}\Big(\phi_{i-\frac{1}{2}} + \phi_{i+\frac{1}{2}}\Big)$. Therefore, dropping the α for readability, equation (6.15) is discretized as follows:

$$\begin{aligned}\phi_i^{n+1} v_i^{n+\frac{1}{2}} - \frac{4\delta t}{3\Delta x^2}\mu\Big(\phi_{i+\frac{1}{2}}^{n+1} v_{i+1}^{n+\frac{1}{2}} - \Big(\phi_{i+\frac{1}{2}}^{n+1} + \phi_{i-\frac{1}{2}}^{n+1}\Big) v_i^{n+\frac{1}{2}} + \phi_{i-\frac{1}{2}}^{n+1} v_{i-1}^{n+\frac{1}{2}}\Big) - \delta t F\Big(\phi_i^{n+1}, v_i^{n+\frac{1}{2}}\Big) \\ = \phi_i^n v_i^n - \frac{\delta t}{\Delta x \rho}\Big(\mathcal{G}_{i+\frac{1}{2}} + \pi\Big(\phi_{i+\frac{1}{2}}^n\Big) - \mathcal{G}_{i-\frac{1}{2}} - \pi\Big(\phi_{i-\frac{1}{2}}^n\Big)\Big) + \delta t \Gamma(\phi_i^n) v_i^n.\end{aligned}$$

Remark. *In this projection step, the viscosity and the friction are treated implicitly. For the viscosity, this treatment enables the relaxation of the CFL constraint and avoids numerical instabilities.*

Remark. *Like in [29, 102, 103] the computation of friction forces requires a specific treatment. Indeed, the friction forces depend on the difference between the phase velocities, and when a phase vanishes the velocity can not be deduced from the momentum (ie. ϕv). In the considered applications, areas of pure liquid or biofilm are important so the adaptation of the initial data to avoid phase vanishing is irrelevant. To overcome this difficulty, a strategy consists to treat these terms implicitly so the velocity can be directly computed using the above equation. However, this is costly because it imposes to solve at each time step a linear system of size: number of phases$\times$mesh grid size.*

Finally, the space discretization of the correction step is given by: $v_\alpha^{n+1} = v_\alpha^{n+\frac{1}{2}} - \frac{\delta t}{\rho_\alpha \Delta x}\Big(P_{i+\frac{1}{2}}^{n+1} - P_{i-\frac{1}{2}}^{n+1}\Big)$

6.4.3 Pressure approximation

Let us detail the third step of the projection correction method. This is the key step to enforce the algebraic constraint on the sum over all the volume fractions (6.1). The standard strategy consists in plugging the time discrete version of equation (6.14b): $\phi_\alpha^{n+1} v_\alpha^{n+1} = \phi_\alpha^{n+1} v_\alpha^{n+\frac{1}{2}} - \frac{\delta t}{\rho_\alpha}\phi_\alpha^{n+1}\partial_x P^{n+1}$ into the incompressibility constraint (6.11) to obtain the following equation on the pressure:

$$\partial_x\left(\sum_\alpha \phi_\alpha^{n+1} v_\alpha^{n+\frac{1}{2}} - \frac{\delta t}{\rho_\alpha}\phi_\alpha^{n+1}\partial_x P^{n+1}\right) = \sum_\alpha \frac{\Gamma_\alpha}{\rho_\alpha} \tag{6.17}$$

Thus the pressure can be obtained by solving a non-linear and inhomogeneous Poisson equation. As mentioned above, this strategy relies on the use of the continuous version of the incompressibility constraint. Therefore, there is no guarantee that the algebraic volume-filling constraint will be fulfilled at the discrete level.

To enforce the algebraic volume filling constraint, we adapt the strategy proposed in [13], which consists in using the fully discretized mass balance equations to deduce the appropriate discrete incompressibility constraint. To this end, let us assume that the constraint $\sum_\alpha \phi_{\alpha,i+\frac{1}{2}}^n = 1$ is satisfied for all times $(t_n)_{n\geqslant 0}$ and in all the grid mesh cells. Thus, the sum of the equations (6.16) over the phases leads to

$$\sum_\alpha \frac{1}{\rho_\alpha}\Big(\mathcal{F}_{i+1}(\phi_\alpha^n, v_\alpha^n) - \mathcal{F}_i(\phi_\alpha^n, v_\alpha^n)\Big) = \Delta x \sum_\alpha \frac{\Gamma_{\alpha,i+\frac{1}{2}}}{\rho_\alpha}. \tag{6.18}$$

Then, as in the standard strategy, an equation on the pressure or its gradient can be deduced by using the time discrete version of equation (6.14b). Since in the correction step, the volume fractions remain unchanged, the time discrete version of equation (6.14b) simplifies into $v_\alpha^{n+1} = v_\alpha^{n+\frac{1}{2}} - \frac{\delta t}{\rho_\alpha}\partial_x P^{n+1}$. Injecting this relation into equation (6.18) gives:

$$\sum_\alpha \frac{1}{\rho_\alpha}\left(\mathcal{F}_{i+1}\left(\phi_\alpha^n, v_\alpha^{n+\frac{1}{2}} - \frac{\delta t}{\rho_\alpha}\partial_x P^{n+1}\right) - \mathcal{F}_i\left(\phi_\alpha^n, v_\alpha^{n+\frac{1}{2}} - \frac{\delta t}{\rho_\alpha}\partial_x P^{n+1}\right)\right) = \Delta x \sum_\alpha \frac{\Gamma_{\alpha,i+\frac{1}{2}}}{\rho_\alpha}. \tag{6.19}$$

Consequently, to ensure that the algebraic volume-filling constraint is met, the pressure must be the solution of the non-linear equation (6.19). The solution can be approximated using Newton's methods. In practice, although this method is more expensive than the standard approach its cost remains reasonable. Indeed, the Jacobian matrix is explicitly known and the solution at the previous time step reveals to be a good initial guess so only very few iterations are necessary to converge. Both strategies are compared in subsection 6.5.2.

6.5 Numerical results

The aim of the paper is to present and test a numerical method able to simulate mixture models for biofilms by guaranteeing the preservation of the algebraic volume filling constraint. Another challenge when one wants to go towards the applications, relies on the difficulty to calibrate the parameters of the model. Many parameters are, up to our knowledge, not available in the current literature and very difficult to extrapolate from experimental data. For example, in [28, 29, 102, 103] the elastic tensor (ie. π_α) settings are calibrated so that the biofilm front velocity matches observations. Consequently, any modification of the model requires recalibration. To avoid such difficulties, subsection 6.5.1 presents numerical simulations based on the numerical scheme presented in section 6.4, but assumes that the viscosity can be neglected, which enables reusing parameters from [28, 29, 102, 103] for the elastic tensors. Secondly, subsection 6.5.2 presents comparisons between the two strategies to approximate the pressure, still neglecting the viscosity. Finally, subsection 6.6.1 presents the dynamic of the full model including viscosity and recalibration of the elastic tensors.

Initially, the mixture is only made of microalgae and liquid and the volume fractions are set by

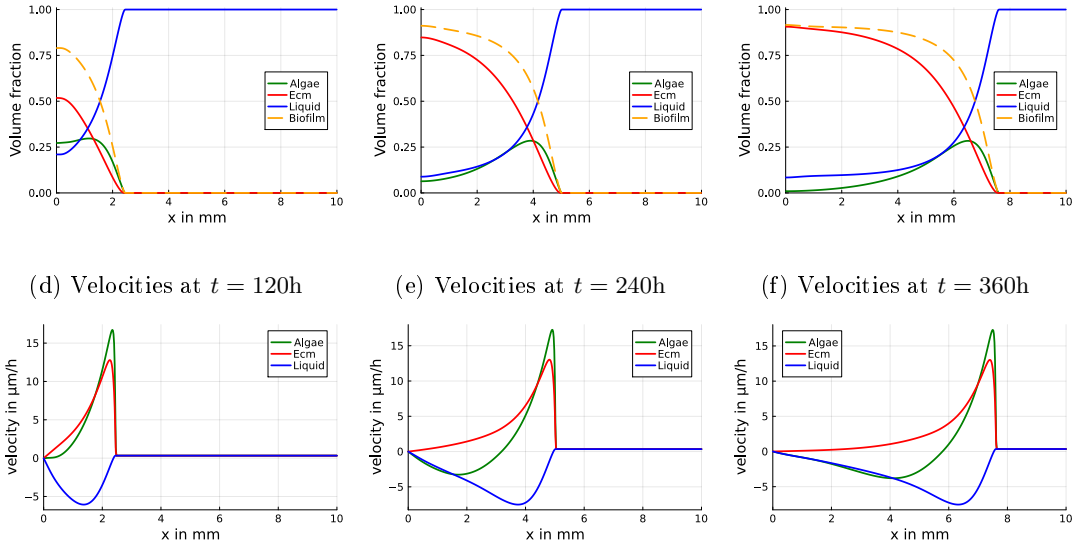
$$\phi_{\mathcal{A}}^0(x) = \max\{0, 0.05(x-0.1)(x+0.1)\}, \quad \phi_{\mathcal{E}}^0 = 0, \quad \text{and} \quad \phi_{\mathcal{L}}^0 = 1 - \phi_{\mathcal{A}}. \tag{6.20}$$

As mentioned in subsection 6.3.3, the system is assumed to be at rest. Thus, the initial velocities are defined by $v_\alpha^0 = -\frac{1}{\rho_\alpha}\partial_x P$ where the pressure P is determined according to the strategy presented in subsection 6.4.3 to enforce the algebraic constraint on the sum over all the volume fractions (6.1).

6.5.1 Biofilm dynamic without viscosity

Figure 6.2 presents the numerical results for different times of system (6.12) where the viscous terms are neglected, namely $\mu_\alpha = 0$ for $\alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$. The simulation is made using the numerical scheme presented in section 6.4 and using the strategy based on the adaptation of [13] for the computation of the pressure, see subsection 6.4.3. In these figures, the left side corresponds to the surface where the biofilm sticks and develops and the right side corresponds to the side covered by the liquid, where nutrients are brought.

(a) Volume fractions at $t = 120$h (b) Volume fractions at $t = 240$h (c) Volume fractions at $t = 360$h



(d) Velocities at $t = 120$h (e) Velocities at $t = 240$h (f) Velocities at $t = 360$h

Figure 6.2: Mixture components volume fractions (first row) and velocities (second row) for different times.

According to Figures 6.2a, 6.2b and 6.2c, there is front propagation corresponding to the biofilm (dashed orange curve) development within the liquid. As in [102], two areas can be distinguished within the biofilm. For example in Figure 6.2c, on the left side, namely for $x \in [0, \sim 6]\mu$m, the biofilm is mainly made of extra-cellular matrix (ie. $\mathcal{E}$), whereas on the biofilm front, namely for $x \in [\sim 6, 7.6]\mu$m, the biofilm is mainly made of microalgae. On the opposite, the right side, namely for $x > 7.6$, is made of pure water.

For the biofilm components, the velocities are positive near the front, which is expected and explains the biofilm expansion. Otherwise, the liquid velocity is negative in the biofilm region, which means that the liquid is drained into the biofilm due to its consumption for the biofilm growth.

6.5.2 Volume filling constraint validation

Let us compare the two strategies presented in subsection 6.4.3 to enforce the algebraic volume-filling constraint, that is computing the pressure P either as the solution of discretization the linear equation (6.17) (standard strategy), or as the solution of the non-linear equation (6.19) (adapted strategy). To this aim, as in subsection 6.5.1, the system (6.12) without the viscous terms is simulated, but using the standard strategy to enforce the volume filling constraint. The results for the mixture components volume fractions at $t = 360h$ are presented in Figure 6.3a and can be compared to Figure 6.2c. According to these plots, the results are comparable. Similarly, the shape of the pressure gradients curves are also similar, see Figure 6.3b. Nevertheless, according to the dotted purple curve in Figure 6.3b, there is a significant discrepancy close to the biofilm front (ie. at $x = 7.6mm$) in the pressure gradients between the two correction strategies. Note that the front is the active part of the biofilm. Namely, it is in this area that the source terms are the largest and lead to significant changes in mixture composition. Therefore, it is

expected that the effect of the pressure gradient is observable notably there.

Besides, the pressure can be interpreted as the Lagrange multiplier associated with the volume-filling constraint. Thus, it is important to compare how these strategies enable enforcing at the discrete level the volume filling constraint (6.1). To this end, Figure 6.3c represents the sum of the volume fractions within the domain a time $t = 360h$. According to this plot, the strategy adapted from [13] enables ensuring the volume filling constraint, whereas the standard strategy does not. Numerically, the maximal error on the volume-filling constraint for the standard strategy is $1.007 \cdot 10^{-3}$, whereas with the adapted strategy, it is $5.107 \cdot 10^{-15}$, namely the order of magnitude of the precision used in Newton's method. Moreover, with this adapted method, the error remains negligible throughout the simulation whereas, with the standard strategy, it varies over time, see Figure 6.7 in Appendix 6.8.2.

(a) Phases volume fraction (b) Pressure gradients (c) Sum of volume fractions

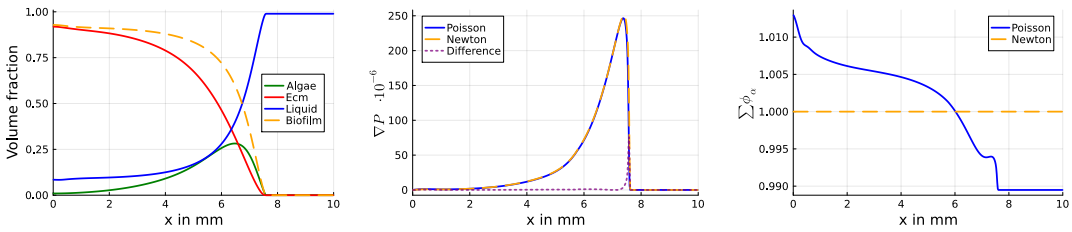


Figure 6.3: Mixture components volume fractions (left), pressure gradients (center) and the sum of the volume fractions (right) at time $t = 360h$. In figures 6.3b and 6.3c the blue curve represents correspond to a simulation made using the standard strategy for the pressure gradient computation (ie. solving a Poisson equation) and the dashed orange curve represents the results obtained using the strategy adapted from [13] which require the resolution of a non-linear equation (ie. Newthon's method).

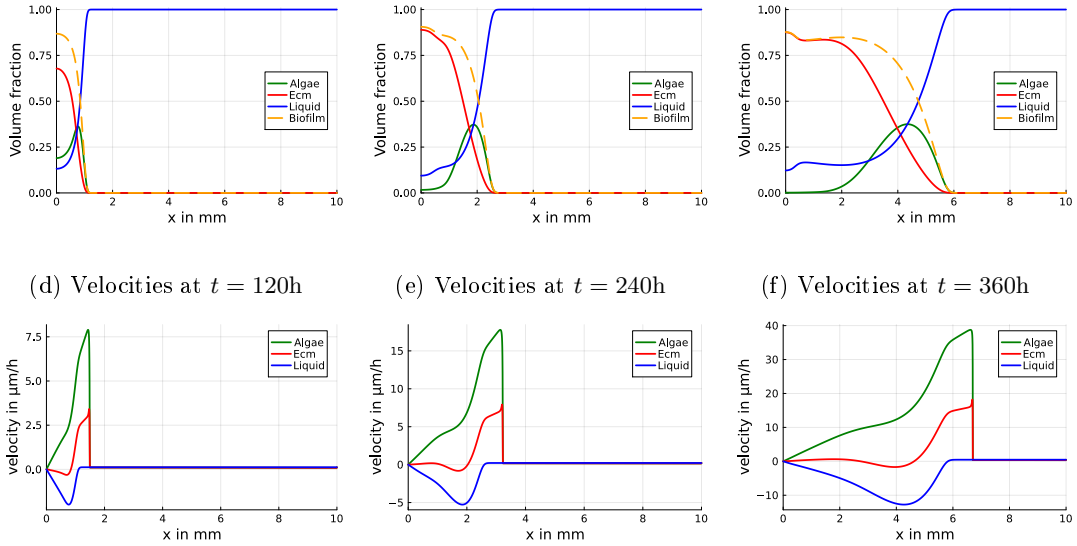
6.6 Model extensions

Following insights coming from [102], this section presents various relevant extensions of the model and their numerical simulations.

6.6.1 Including the viscosity

Adding the viscous terms for the components requires recalibrating the model parameters. Indeed, the viscosity is a measure of the component's resistance to deformation. Therefore, when accounting for the viscosity, the parameters associated with the component's ability to deform must be adapted. In particular, the elastic tensors for the microalgae and the extra-cellular matrix must be recalibrated. Moreover, up to our knowledge, there is no direct measurement of the parameters and they are calibrated, see [28, 29, 102], such that the biofilm front velocity matches experimental measurements, see [112]. However, such calibration is extremely complex because the biofilm front velocity depends also on many other parameters like the growth or death rate. Nevertheless, to get the right order of magnitude of the biofilm front velocity the elastic tensor coefficients must be significantly increased: multiplied by $9 \cdot 10^7$ so is set to $\gamma_{\mathcal{A}} = \gamma_{\mathcal{E}} = 4.5 \cdot 10^{-3}$kg/m/day.

(a) Volume fractions at $t = 120$h (b) Volume fractions at $t = 240$h (c) Volume fractions at $t = 360$h



(d) Velocities at $t = 120$h (e) Velocities at $t = 240$h (f) Velocities at $t = 360$h

Figure 6.4: Mixture components volume fractions (first row) and velocities (second row) for different times. In this simulation, the viscosity is included and the elastic tensors for the biological phases are multiplied $9 \cdot 10^7$ so the biofilm front velocity matches experimental measurements. The simulation is made using 2048 mesh cells for the space grid.

Figure 6.4 represents the time dynamic of mixture components when accounting for the contribution of viscosity. The global dynamic remains comparable to the dynamic observed in Figure 6.2. In particular, there is still a biofilm traveling front. Again, there are two areas within the biofilm: the back which is mainly made of an extra-cellular matrix (for $x \in [0, 5]$ in subfigure 6.4c), and the front which is mainly made of microalgae (for $x \in [3.97, 6.7]\mu$m in subfigure 6.6c). Nevertheless, a major discrepancy is that the microalgae remain more located at the front when including the viscosity. This is particularly visible at $t = 240$h when comparing Figure 6.2b and Figure 6.4b. In addition, at $t = 240$h, the velocities order of magnitude close to the front is larger when including the viscosity. However, the interpretation of this observation is tricky. Indeed, the shift in the elastic tensors for the biological phases imposes the use of very refined mesh grids to properly capture the biofilm dynamic. Thus, it would be of particular interest to design and use well-balanced numerical scheme able to preserve the biofilm front structure. For more details about the numerical convergence of the scheme, see appendix 6.8.3.

6.6.2 Including light intensity

A microalga is a photosynthetic organism. Thus, microalgae require light to grow. When microalgae develop within a biofilm, the upper layers overshadow the lower layers. Following [28, 29, 102] to account for these mechanisms, the microalgae growth rate becomes $\psi_g = \mu_g \rho_{\mathcal{A}} \phi_{\mathcal{A}} \phi_{\mathcal{L}} f_I$, where f_I accounts for the effect of light on growth. This term depends on the rescaled received light intensity I and takes the form of the Haldane law:

$$f_I = \frac{2(1 + K_{\mathcal{I}})\mathcal{I}}{\mathcal{I}^2 + 2K_{\mathcal{I}}\mathcal{I} + 1}. \tag{6.21}$$

The rescaled light intensity is the ratio between the received light and the optimal light intensity $\mathcal{I}_{\text{opt}}$, namely:

$$\mathcal{I}(t,x) = \frac{\mathcal{I}_{\text{surf}}}{\mathcal{I}_{\text{opt}}} \exp\left(-\int_x^L \tau\big(1-\phi_{\mathcal{L}}(t,y)\big)dy\right), \tag{6.22}$$

where I_{surf} is the light intensity at the surface of the tank (ie. $x = L$) and τ the attenuation coefficient of the biofilm, assuming that microalgae and extra-cellular matrix have the same attenuation rate. The parameter values associated to the light are gathered in Table 6.2.

(a) Volume fractions at $t = 120$h (b) Volume fractions at $t = 240$h (c) Volume fractions at $t = 360$h

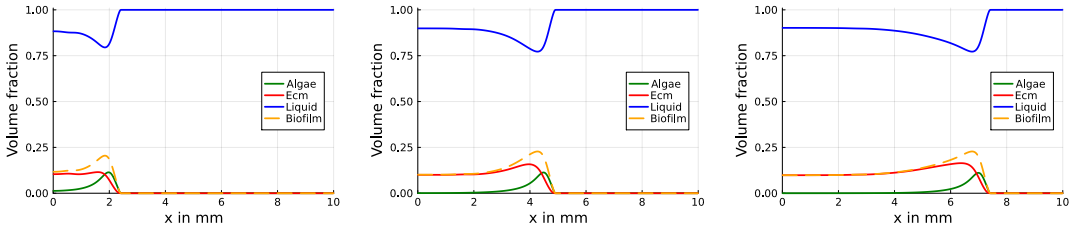


Figure 6.5: Mixture components volume fractions for different times. In this simulation, the effective microalgae growth rate (ie. ψ_g) accounts for the contribution of light intensity through Haldane's law (6.21) and light attenuation induced by biofilm layers, see equation (6.22)

Figure 6.5 represents the time dynamic of mixture components when accounting for the contribution of light. The global dynamic is comparable to the dynamic observed in Figure 6.2. In particular, the biofilm front position travels at the same speed, and, again, there are two areas within the biofilm: the back which is mainly made of an extra-cellular matrix (for $x \in [0,6]$ in subfigure 6.5c) and the front which is mainly made of microalgae (for $x \in [6,7.6]\mu$m in subfigure 6.5c). However, as expected, the volume fraction of biofilm is lower. Indeed, taking into account the effect of light reduces the growth in the shadowed areas and thus the biomass of microalgae. The extra-cellular matrix is also reduced since it is made from microalgae excretion and dead organisms.

6.6.3 Including light intensity and solutes

Following [102], let us now include three dissolved components: the substrate ($\mathcal{S}$), the inorganic carbon ($\mathcal{C}$), and the oxygen ($\mathcal{O}$). As mentioned in subsection 6.3.1, the dynamic for dissolved components is modeled using a convection-diffusion reaction equation (6.9).

In a nutshell, the substrate represents the nitrate which is a nutrient of primary importance for the growth of autotrophic organisms like microalgae. Besides, roughly speaking, photosynthesis is the assimilation of inorganic carbon using light energy by autotrophic organisms. Photosynthesis releases oxygen. Thus, including these components is of primary interest. Taking into account these compounds also allows us to include the process of respiration. Basically, respiration is the opposite mechanism of photosynthesis and its consideration allows us to better describe the dynamic of thick biofilms. Indeed, the process of respiration becomes non-negligible in the absence of light, namely in the biofilm's inner layers.

As for the light, the contributions of the dissolved components to the photosynthesis process are accounted for in the growth through the multiplication by functions f_p, $p \in \{\mathcal{S},\mathcal{C},\mathcal{O}\}$ which represent how the growth is modified by the local concentration of these components. On the

one hand, limited contribution in high concentration regimes of the substrate and the inorganic carbon is modeled using Monod's law: $f_p = \frac{\theta_p}{K_p+\theta_p}$. On the other hand, the inhibition induced by high oxygen concentration is modeled by the sigmoidal function $f_{\mathcal{O}} = \frac{1}{1+\left(\frac{\theta_{\mathcal{O}}}{K_{\mathcal{O}}}\right)^{n_{\mathcal{O}}}}$. Thus, including the contribution of the dissolved components and the light intensity, the algae growth rate becomes: $\psi_g = \mu_g \rho_{\mathcal{A}} \phi_{\mathcal{A}} \phi_{\mathcal{L}} f_I f_{\mathcal{S}} f_{\mathcal{C}} f_{\mathcal{O}}$.

The respiration process is modeled by $\psi_r = \mu_r \phi_{\mathcal{A}} \frac{\theta_{\mathcal{O}}}{K_r+\theta_{\mathcal{O}}}$ where μ_r is the maximal respiration rate and K_r the half-saturation constant for the oxygen.

The modification of the microalgae growth rate and the inclusion of the respiration process requires to adapt the source terms for the phases as follows:

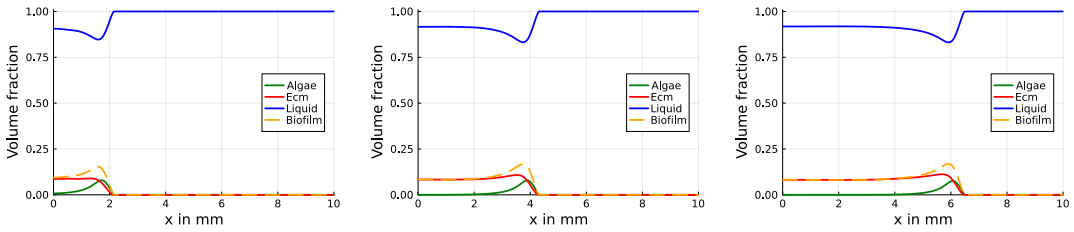
$$\Gamma\mathcal{A} = \psi_g - \psi_e - \psi_d - \psi_r, \qquad \Gamma\mathcal{E} = \psi_e + \eta_{\mathcal{E}}\psi_d, \qquad \Gamma\mathcal{L} = (1-\eta_{\mathcal{E}})\psi_d + \eta_{\mathcal{L}}(\psi_r - \psi_g).$$

As for a phase, the source terms for a dissolved component is the sum of the pseudo-stoichiometric coefficients multiplied by the reaction rates. Thus, for the dissolved components, the source terms are

$$\Gamma\mathcal{S} = -\eta_{\mathcal{S}}\psi_g, \qquad \Gamma\mathcal{C} = -\eta_g^{\mathcal{C}}\psi_g + \eta_r^{\mathcal{C}}\psi_r, \qquad \Gamma\mathcal{O} = \eta_g^{\mathcal{O}}\psi_g - \eta_r^{\mathcal{O}}\psi_r.$$

The external supply for the dissolved components is modeled through Dirichlet boundary conditions at the top of the bioreactor, namely at $x = L$. Otherwise, the no flux boundary condition at the bottom of the bioreactor is modeled using the Neumann boundary condition: $\partial_x \theta_p|_{x=0} = 0$ for $p \in \{\mathcal{S}, \mathcal{C}, \mathcal{O}\}$.

(a) Volume fractions at $t = 120$h (b) Volume fractions at $t = 240$h (c) Volume fractions at $t = 360$h



(d) Solutes concentration at $t = 120$h (e) Solutes concentration at $t = 240$h (f) Solutes concentration at $t = 360$h

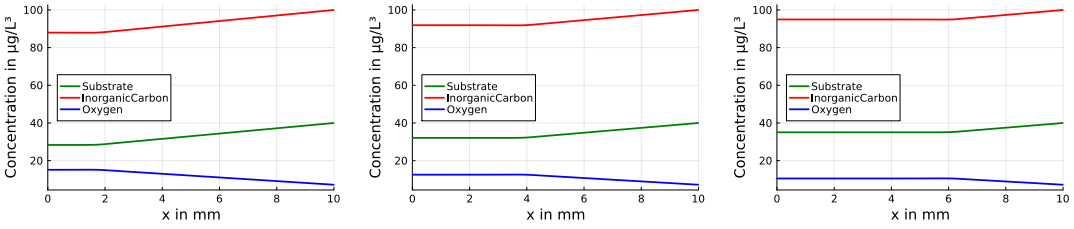


Figure 6.6: Mixture components volume fractions (first row) and velocities (second row) for different times.

The parameter values associated with the inclusion of the dissolved components are gathered in Table 6.2.

Numerically, the transport and reaction terms in the mass balance equations for the solutes are treated similarly to the other components. The diffusion terms are treated implicitly to ensure stability without constraining the CFL condition.

Figure 6.6 represents the time dynamic of mixture components when accounting for the contribution of light and solutes. The global dynamic is comparable to the dynamic observed in Figure 6.5. In particular, the biofilm front position travels at a comparable speed. Again there are two areas within the biofilm: the back which is mainly made of an extra-cellular matrix (for $x \in [0, 6.5]$ in subfigure 6.6c), and the front which is mainly made of microalgae (for $x \in [5.5, 6.6]\mu$m in subfigure 6.6c). However, the biofilm front velocity is slightly slower here. This can be explained by the fact that the lack or excess of solutes in the active part of the biofilm slightly reduces its growth. Indeed, for example, at $t = 360h$, within the biofilm area, the concentration of substrate is reduced by 12.4% and the concentration of inorganic carbon is reduced by 5.1% relatively to the input values (ie. θ_{in}). Besides, the concentration of oxygen is increased by 46.3% relatively to $\theta_{in,\mathcal{C}}$. These discrepancies are larger at the beginning and tend to decrease over time, see Figure 6.9 in the supplementary material. These results are in good agreement with the results presented in [102, 103].

6.7 Conclusions and perspectives

This article proposes an adaptation of the numerical scheme presented in [13] able to enforce the volume filling constraint in mixture models including mass exchanges. As in [13] the strategy consists in deducing the discrete version of the incompressibility constraint from the discretized mass balance equations. Numerical simulations show that this method enables the enforcement of the total volume filling constraint at the discrete level.

In addition, on the modeling side, previous models from the literature are enriched by the inclusion of viscous terms. These terms are essential to properly model biofilms in their fluidic environment especially when there is a mixing of the surrounding fluid. In this context, this work has allowed us to highlight the importance of designing well-balanced numerical scheme able to efficiently capture the biofilm dynamic when including the viscosity. Indeed, including the viscosity requires to recalibrate model parameters; in particular the elastic tensors need to be strongly rescaled in order to recover realistic front features. However, with these parameters, the numerical set up is more demanding to reach convergence. This difficulty leads to consider further the design of a specific well-balanced scheme for the problem. To this end, the use of well-balanced numerical schemes able to preserve the equilibrium at the biofilm front can be considered.

Finally, in further works, it would be interesting to include additional biological features. Among others, biofilms are generally multi-species. The framework of mixture theory is well adapted to incorporate different species and such extensions are affordable if the interaction between the species and their metabolisms is known. To make the model even more realistic and predictive its calibration on experimental data is also particularly interesting. In conclusion, real-life biofilms are 3D and therefore the extension and implementation of the numerical method in 2D and 3D should be considered.

Acknowledgement

- This work was supported by Conseil Regional de Bourgogne Franche-Comté (France).

6.8 Appendix

6.8.1 Spatiotemporal equilibrium

The spatiotemporal equilibrium states for the system (6.12) correspond to the state solution where the source terms of all phases vanish, namely: $\Gamma_i = 0$ for $i \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$. In particular, $\Gamma_\mathcal{E} = 0$ induces $\psi_e + \eta_\mathcal{E}\psi_d = 0$ which lead to $\phi_\mathcal{A} = 0$. Thus, the only spatiotemporal stationary state is the null state, namely $\phi_i = 0$ for $i \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$.

6.8.2 Time dynamic of the volume fraction sum

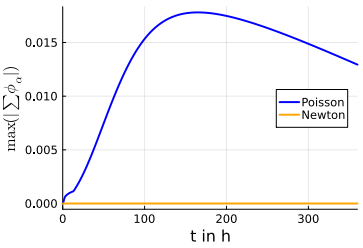


Figure 6.7: Time evolution for the maximal error within the domain on the sum of volume fractions: $E = \max_x |\sum_\alpha \phi_\alpha - 1|$.

6.8.3 Numerical convergence analysis for the model including viscosity

Numerical experiments have shown that numerical parameters need to be significantly reduced to reach convergence when the viscosity is included, and re-estimating the elastic tensor accordingly. Indeed, as mentioned in section 6.6.1, when including the viscosity, the elastic tensor coefficients must be rescaled and multiplied by $9 \cdot 10^7$ to obtain realistic front velocities for the biofilm. Figure 6.8 shows the convergences of the numerical scheme in both cases: with and without the viscous term. As expected, numerical convergence is obtained in both cases. Nevertheless, as presented in Figure 6.8 the convergences rate is lower when the viscous term is included. This explains at least partially why the numerical parameters need to be significantly reduced to reach acceptable precision for the application considered when the viscosity is included

6.8.4 Relative variation of solutes concentration

Figure 6.9 shows the relative variation of solutes concentration for different times associated to the simulation presented in subsection 6.6.3. In this figure, we observe that the variations relatively to the input concentration are larger at the beginning (ie. $t = 120$h) than at the end (ie. $t = 360$h).

(a) Numerical convergence for the model without the viscous term.

(b) Numerical convergence for the model incluging the viscous term.

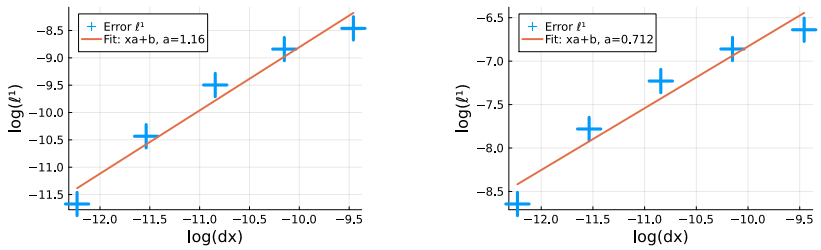


Figure 6.8: Numerical analysis of the convergence of the scheme. The left side correspond to the case without the viscous term, namely $\mu_\alpha = 0$ for $\alpha \in \{\mathcal{A}, \mathcal{E}, \mathcal{L}\}$ and the right side correspond to the case with the viscous term and using very large values for the elastic tensor coefficients.

(a) $t = 120$h (b) $t = 240$h (c) $t = 360$h

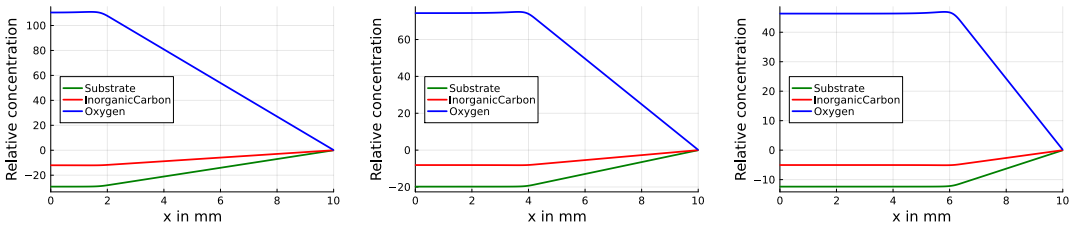


Figure 6.9: Solutes relative concentration for different times

Chapter 7

Perspectives

In this chapter, we lay down a few perspectives of the work described in this manuscript. Some of these perspectives are ongoing works while others are yet to be explored.

7.1 Uniqueness of stationary solutions

We are interested in showing that the diffusive Lifshitz-Slyozov model has a unique stationary solutions, given by the values of λ and $m = \int_{\mathbb{R}_+} f^0(x)\mathrm{d}x$. We previously introduced the map $\phi : L \to L + \int_{\mathbb{R}_+} xM_L(x)\mathrm{d}x$, which we use to find stationary solutions by solving $\phi(L) = \lambda$. Let us recall that $M_L(\cdot)$ is the stationary distribution associated to L defined by :

$$M_L(x) = \frac{C}{d(x,L)} \exp(\frac{2}{\varepsilon}\int_0^x \frac{v(y,L)}{d(y,L)}\mathrm{d}y),$$

where $C > 0$ is a re-normalizing constant defined by : Œ

$$C = \frac{m}{\int_{\mathbb{R}_+} \frac{1}{d(x,L)} \exp(\frac{2}{\varepsilon}\int_0^x \frac{v(y,L)}{d(y,L)}\mathrm{d}y)\mathrm{d}x}.$$

Existence of a solution to $\phi(L) = \lambda$ is obtained by first noting that ϕ is a continuous function and $\lim_{L\to+\infty} \phi(L) = +\infty$ and at $L = 0$ we have that :

$$M_0(x) = \frac{C}{b(x)} \exp(-\frac{2}{\varepsilon}x),$$

where the constant C is actually zero since $b(x) \leq C_b x$ and, letting $\delta > 0$, we have :

$$\int_\delta^1 \frac{1}{b(x)} \exp(-\frac{2}{\varepsilon}x)\mathrm{d}x \geq \frac{\exp(-\frac{2}{\varepsilon})}{C_b} \int_\delta^1 \frac{1}{x}\mathrm{d}x \xrightarrow[\delta\to 0^+]{} +\infty$$

Therefore $M_0(\cdot) \equiv 0$. Hence $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous surjective and there exists a solution to $\phi(L) = \lambda$ for all $\lambda \geq 0$.

However uniqueness is harder to prove. The case for Dirichlet boundary conditions was treated in [57], however in our case, with null-flux boundary conditions their method is unsuccessful. Other methods were looked into such as showing monotony or fixed point methods, all without success. Thus we turn ourselves to a relatively novel mathematical idea : computer

assisted proof. This work is currently underway with Maxime Breden (CMAP, Ecole Polytechnique). The key point is we want to apply Banach fixed-point theorem, and thus we want to show that $|\frac{\mathrm{d}\phi(L)}{\mathrm{d}L}| \leq c < 1$. We use methods coming from interval arithmetic that are able to perform exact computation and for a range of values of L will give us bounds on the value of $|\frac{\mathrm{d}\phi(L)}{\mathrm{d}L}|$. Let us describe briefly what interval arithmetic entails and how we apply it to our problem.

Consider the numbers π and e. Both are transcendental numbers and their numerical evaluation is therefore an approximation. The idea behind interval arithmetic is to replace numbers with intervals that contain them. Hence let us denote $\pi \in [3.14, 3.15]$ and $\mathrm{e} \in [2.71, 2.72]$. Now imagine one wants to compute $\pi^2 + \mathrm{e}^2$. Hence we need to define the square function on intervals as well as the sum. This is done easily : $([a,b])^2 = [\mathbb{1}_{a>0}a^2, \max(a^2, b^2)]$ and $[a,b] + [c,d] = [a+c, b+d]$. Observe that we have taken two decimal digit as precision for our intervals, thus when computing the upper bound $\max(a^2, b^2)$ we need to use an upper approximation while using a lower approximation for lower bounds.

Which finally yields :

$$\pi^2 + \mathrm{e}^2 \in ([3.14, 3.15])^2 + ([2.71, 2.72])^2 = [9.85, 9.93] + [7.34, 7.40] = [17.19, 17.33].$$

An evaluation using Python gives $\pi^2 + \mathrm{e}^2 \simeq 17.26$ which is in our interval. These interval arithmetic can be used for numerical computations which will give theoretical bounds on a computed object. Hence our method works as follow : we use interval arithmetic to show that for a bounded range $L \in [L_{\min}, L_{\max}]$ there exists some constant $0 \leq c < 1$ such that $|\frac{\mathrm{d}\phi(L)}{\mathrm{d}L}| \leq c$. A problem arises when $L \to 0$ and $L \to \infty$ due to numerical computation of diverging integrals. Hence we need to show uniqueness of a solution to $\phi(L) = \lambda$ on $[0, L_{\min}]$ with the hope that $L_{\min}$ is not too small, as well as on $[L_{\max}, +\infty)$. Progress have been made in both directions and we have already observe that the numerical results hold for a wide range of L, particularly close to zero.

7.2 Additional theoretical results and connection between stochastic and deterministic models

On a more theoretical side, an interesting research direction would be to properly study solutions of the diffusive Lifshitz-Slyozov model. A first approach has been derived in [34], where the authors derive a different version of the diffusive Lifshitz-Slyozov model. Particularly, they obtain a diffusion rate that is constant in time but not in space. This allows them to transform the diffusion term to a constant rate diffusion via a change of variable. However in our case, this method does not seem to work. Hence showing existence and uniqueness of solutions to the diffusive Lifshitz-Slyozov model is tricky and still an open question.

Another parallel direction would be to provide both stochastic models with an existence and uniqueness result. Our hope is that in both cases the sketch of the proof would not be far from the deterministic approach : fix L and show that both X_ε and Z_ε exists, then use Schauder fixed-point theorem on L to show uniqueness. Uniqueness could be obtained in a similar way as in the deterministic case.

We previously mentioned the uniform propagation of chaos that leads us to writing Equations (1.41) and (1.45). Such a result is prove in [122] for birth and death processes. We recall here the theorem from [122] adapted to our modeling of adipose cells :

Theorem 7.2.1 (Uniform propagation of chaos)**.** *Let $T > 0$. Assume that the rates a and b are Lipchitz functions as well as some additional technical assumptions. Then there exists a coupling*

and a constant $K > 0$ such that :

$$\sup_{t \leq T} \mathbb{E}\left[|X_\varepsilon^{1,N}(t) - X_\varepsilon(t)|\right] \leq \frac{K}{\sqrt{N}}. \tag{7.1}$$

A rigorous proof of this theorem for the stochastic models of adipose cells is currently in the work and the technical assumptions are omitted for ease of reading. Nonetheless we refer the reader to [122] for a rigorous proof for birth and death processes which is out of the scope of this manuscript. The proof for adipose cell stochastic models should not deviate too far from the proof in [122].

Concerning equation (1.39) and (1.43), a result in the spirit of the Law of Large Numbers (LLN) on the empirical measure $\mu_N(t, \mathrm{d}x) = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_{i,\varepsilon}(t)}(\mathrm{d}x)$ could be a direction to show existence of a solution to the diffusive Lifshitz-Slyozov model by looking at the limit $\lim_{N \to \infty} \mu_N(t, \mathrm{d}x)$.

In [18], the author show that in under the asumption that the ratio $\frac{b}{a}$ is C^1 and strictly decreasing, the mild-solution to the Lifshitz-Slyozov model convergences exponentially in time toward a Dirac mass located at the unique solution of $\lambda = mx + \frac{b(x)}{a(x)}$. They use an entropy method and Wasserstien distances to obtain this results. The assumption that $\frac{b}{a}$ is strictly decreasing is key in their argument. This result could be extend to the case of a non-monotonous ratio of rate, but this involves an appropriate choice of an entropy function which is still an open question.

Additionally, the gradient-flow structure of both the Becker-Döring and Lifshitz-Slyozov model has been well established [92, 93, 111]. This structure may be a starting point for new theoretical results, including the ones we previously mentioned.

7.3 Model extensions

On the side of modeling, various extensions can be made to the model. Their pertinence is obviously relative to our ability to collect and study biological data that would support these extensions.

In the spirit of [46], the addition to the model of the dynamics of pre-adipocytes is an interesting research direction particularly for modeling the effect of weight gain on the total amount of cells. Let us denote by $s(t)$ the amount of stem cells at time t and $p(t)$ the amount of pre-adipocytes at time t. We consider two functions ϕ_s and ϕ_p that describe the dynamics of m and p respectively, which leads to the system :

$$\begin{cases} \dfrac{\mathrm{d}m(t)}{\mathrm{d}t} = \phi_m(m(t), p(t), g(t, \cdot)) & \text{(7.2a)} \\ \dfrac{\mathrm{d}p(t)}{\mathrm{d}t} = \phi_p(m(t), p(t), g(t, \cdot)) & \text{(7.2b)} \\ \partial_t g + \partial_x(vg) = \dfrac{\varepsilon}{2}\partial_x^2(dg) - \gamma g, & \text{(7.2c)} \\ L(t) + \displaystyle\int_{\mathbb{R}_+} x g(t, x) \mathrm{d}x = \lambda. & \text{(7.2d)} \end{cases}$$

The functions g and L are the distribution of cells and the external lipid amount respectively. The constant γ is the death rate of adipose cells. A possible choice for ϕ_m and ϕ_p is described in [46]. As mentioned in the introduction, the rate of differentiation of stem cells into pre-adipocytes and then adipose cells is comparable to the death rate γ in a healthy individual.

However as an individual gains weight, we expect the body to adapt and to increase the count of adipose cells. This raises a modeling question as to how to model this change in behavior. Pre-adipocytes differentiate and become adipose cells of size zero, thus the boundary condition at $x = 0$ should take this into account. This extended model is more complicated from a theoretical point of view, and stationary solutions may not be as specific as in our case. Additionally since the numerical scheme we used for the diffusive Lifshitz-Slyozov model is well-balanced and uses the exact formula of stationary solutions, we may need to developed new numerical methods for this extended model.

Another direction is to consider that the total amount of lipids changes over time $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$. An obvious choice would be a periodic function to simulate food intake at specific time points. In this case, depending on the limit $\lim_{t \to +\infty} \lambda(t)$, stationary solutions may not have an explicit formula. As in the previous extension, this would mean introducing new numerical methods for simulations. New theoretical results are also obviously needed. Additionally, in [100] and [101] authors developed an agent based model of adipose cells to study tissue morphology. This agent based model could yield an additional way of looking at size distributions of adipose cells and an explicit connection between this model and our Lifshitz-Slyozov model has yet to be investigated. A possible approach would be to add a spacial position in the tissue to the model. This increases the complexity of the model and poses new modeling questions on the spacial distribution of adipose cells inside the tissue.

In regards to parameter estimation, possible extensions often come with the need for additional data : comparison between different species, temporal analysis of weight gain and weight loss from longitudinal data, etc ... Moreover, the various extensions proposed for our model may be able to better reproduce the data.

Appendix A

Existence and uniqueness of solutions

A.1 The Becker-Döring model

In this part, we prove theorem 1.4.2. The proof follows the method developped in [9]. We begin the proof by introducing the following truncated system where we only consider the first n sizes. We show global existence and uniqueness of solutions to this system and their positiveness. Then we take the limit as $n \to \infty$.

$$\begin{cases} \dfrac{dc_i^n}{\mathrm{d}t} = J_{i-1}(c^n, l^n) - J_i(c^n, l^n)\,, \ \forall\, 1 \leq i < n \\ \dfrac{dc_0^n}{\mathrm{d}t} = -J_0(c^n, l^n) \\ \dfrac{dc_n^n}{\mathrm{d}t} = J_{n-1}(c^n, l^n) \\ \dfrac{dl^n}{\mathrm{d}t} = -\sum\limits_{i=0}^{n-1} J_i(c^n, l^n) \end{cases} \tag{A.1}$$

Lemma A.1.1. *For any $l_0 \in \mathbb{R}_+$ and $c^0 \in \mathbb{R}_+^n$, the system (A.1) has a unique solution $(c^n, l^n) \in C^1(\mathbb{R}_+^n) \times C^1(\mathbb{R}_+)$, such that for all $0 \leq i \leq n$, $c_i^n(t) \geq 0$, $l^n(t) \geq 0$ and $l^n(t) + \sum\limits_{i=1}^{n} i c_i^n(t) = l_0 + \sum\limits_{i=1}^{n} i c_i^0$*

Proof of Lemma A.1.1. By Cauchy-Lipshitz's theorem, we know that (A.1) has a unique maximal solution.

We simply need to show that the solution stays positive for all time t. To do so we introduce the following system, where we add $\varepsilon > 0$ to each ODE :

$$\frac{dc_i^{\varepsilon,n}}{\mathrm{d}t} = J_{i-1}(c^{\varepsilon,n}, l^{\varepsilon,n}) - J_i(c^{\varepsilon,n}, l^{\varepsilon,n}) + \varepsilon \tag{A.2}$$

$$\frac{dc_0^{\varepsilon,n}}{\mathrm{d}t} = -J_0(c^{\varepsilon,n}, l^{\varepsilon,n}) + \varepsilon \tag{A.3}$$

For $l^{\varepsilon,n}$ the derivative takes the form :

$$\frac{dl^{\varepsilon,n}}{\mathrm{d}t} = -\sum_{i=0}^{n-1} J_i(c^{\varepsilon,n}, l^{\varepsilon,n})$$

By Cauchy-Lipshitz, we also now that this system has a unique maximal solution $(c^{\varepsilon,n}, l^{\varepsilon,n}) \in C^1([0,T))^{\mathbb{N}} \times C^1([0,T))$

Let $\tau = \inf\{t > 0 | \exists i, c_i^{\varepsilon,n}(t) = 0\}$. τ is the first hitting time of $\{0\}$ of the solution $(c^{\varepsilon,n}, l^{\varepsilon,n})$. First consider that this time is reached for one of the $c_i^{\varepsilon,n}$, where $i \geq 1$. Then :

$$\frac{dc_i^{\varepsilon,n}}{\mathrm{d}t}(\tau) = a_{i-1}\frac{l^{\varepsilon,n}(\tau)\Lambda}{l^{\varepsilon,n}(\tau)\Lambda + \kappa}c_{i-1}^{\varepsilon,n}(\tau) + b_{i+1}c_{i+1}^{\varepsilon,n}(\tau) + \varepsilon$$

By positivity of a_i and b_i for all $i \geq 0$, $\frac{dc_i^{\varepsilon,n}}{\mathrm{d}t}(\tau) > 0$. Similarly in the case $i = 0$:

$$\frac{dc_0^{\varepsilon,n}}{\mathrm{d}t}(\tau) = b_1 c_1^{\varepsilon,n}(\tau) + \varepsilon$$

Hence $\frac{dc_0^{\varepsilon,n}}{\mathrm{d}t}(\tau) > 0$. Now if $\tau = 0$ then the solution is positive in some neighborhood of 0. Otherwise if $\tau > 0$, this contradicts the fact that $c_i^{\varepsilon,n}(\tau) = 0$.

We have shown that at leat one of the $c_i^{\varepsilon,n}$ is strictly positive. As such if there exists a time τ such that $l^{\varepsilon,n}(\tau) = 0$ then one has :

$$\frac{dl^{\varepsilon,n}}{\mathrm{d}t}(\tau) \geq \sum_{i=0}^{n-1} b_{i+1}c_{i+1}^{\varepsilon,n}(\tau) > 0$$

Then either $\tau = 0$ and the solution is positive on some neighborhood of 0 or $\tau > 0$ and this contradicts $l^{\varepsilon,n}(\tau) = 0$.

Using Arzelà-Ascoli's theorem, we show that there exists a sub-sequence of $(c^{\varepsilon,n}, l^{\varepsilon,n})$ which converges uniformly toward a solution (c^n, l^n) of (A.1). Bounds on the solution itself are simple using $l^{\varepsilon,n}(t) + \sum_{i=1}^{n} i c_i^{\varepsilon,n}(t) = l_0 + \sum_{i=1}^{n} i c_i^0$.

Let $i \geq 2$. We choose n large enough so that $i < n$.

$$\begin{aligned}
|\frac{dc_i^{\varepsilon,n}}{\mathrm{d}t}| &= |a_{i-1}\frac{l^{\varepsilon,n}\Lambda}{l^{\varepsilon,n}\Lambda + \kappa}c_{i-1}^{\varepsilon,n} - (b_i + a_i\frac{l^{\varepsilon,n}\Lambda}{l^{\varepsilon,n}\Lambda + \kappa})c_i^{\varepsilon,n} + b_{i+1}c_{i+1}^{\varepsilon,n} + \varepsilon| \\
&\leq \left(\frac{a_{i-1}}{i-1} + \frac{a_i}{i} + \frac{b_i}{i} + \frac{b_{i+1}}{i+1}\right)(||c^0||_X + l_0) + \varepsilon \\
&\leq C_i + \varepsilon < (\frac{3C_a}{2} + 4C_b)(||c^0||_X + l_0) + \varepsilon
\end{aligned}$$

For $i = 1$, we get :

$$\begin{aligned}
|\frac{dc_1^{\varepsilon,n}}{\mathrm{d}t}| &= |a_0\frac{l^{\varepsilon,n}\Lambda}{l^{\varepsilon,n}\Lambda + \kappa}c_0^{\varepsilon,n} - (b_1 + a_1\frac{l^{\varepsilon,n}\Lambda}{l^{\varepsilon,n}\Lambda + \kappa})c_1^{\varepsilon,n} + b_2c_2^{\varepsilon,n} + \varepsilon| \\
&\leq a_0 m + (a_1 + b_1 + \frac{b_2}{2})(||c^0||_X + l_0) + \varepsilon \\
&\leq C_1 + \varepsilon
\end{aligned}$$

And finally for $i = 0$:

$$\begin{aligned}|\frac{dc_0^{\varepsilon,n}}{\mathrm{d}t}| &= |a_0 \frac{l^{\varepsilon,n}\Lambda}{l^{\varepsilon,n}\Lambda + \kappa} c_0^{\varepsilon,n} - b_1 c_1^{\varepsilon,n} + \varepsilon| \\ &\leq a_0 m + b_1(||c^0||_X + l_0) + \varepsilon \\ &\leq C_0 + \varepsilon\end{aligned}$$

For $l^{\varepsilon,n}$ we have the following inequality :

$$\begin{aligned}|\frac{dl^{\varepsilon,n}}{\mathrm{d}t}| &= |\sum_{i=0}^{n-1} J_i| \\ &\leq C_0 + \sum_{i=1}^{n-1}(a_i \frac{l^{\varepsilon,n}\Lambda}{l^{\varepsilon,n}\Lambda + \kappa} c_i^{\varepsilon,n} + b_{i+1} c_{i+1}^{\varepsilon,n}) \\ &\leq C_0 + (C_a + C_b)m + C_b\delta(||c^0||_X + l_0)\end{aligned}$$

Thus assuming $\varepsilon << 1$, one has sufficient conditions to apply Arzelà-Ascoli's theorem. Having found a solution to (A.1), we now show it is a global solution by using simple bounds in time. For $i \geq 1$:

$$0 \leq c_i^n(t) \leq \frac{1}{i}(l^n(0) + \sum_{i=1}^{n} i c_i(0))$$

For $i = 0$, we have the conservation of the number of cells : $\frac{d}{\mathrm{d}t}\sum_{i=1}^{n} c_i(t) = 0$. Hence :

$$0 \leq c_0^n(t) \leq \sum_{i=1}^{n} c_i(0)$$

And finally for l^n :

$$0 \leq l^n(t) \leq l^n(0) + \sum_{i=1}^{n} i c_i(0)$$

Therefore the solution (c^n, l^n) is global. The last part is a straightforward computation :

$$\frac{d}{\mathrm{d}t}\Big(l^n(t) + \sum_{i=0}^{n} i c_i^n(t)\Big) = -\sum_{i=0}^{n-1} J_i(c^n, l^n) + \sum_{i=1}^{n-1} i(J_{i-1}(c^n, l^n) - J_i(c^n, l^n)) - nJ_{n-1}(c^n, l^n) = 0$$

□

Proof of theorem 1.4.2. The proof goes as follows : we show that the series of functions $(c_i^n)_{n\geq 1}$ are equicontinuous, then we use Arzela-Ascoli's theorem to get uniform convergence on a compact subset of $\mathbb{R}_+$. We proceed similarly for l^n.

Let $c^0 \in X^+$, $l_0 \in \mathbb{R}_+$ and note $c^{0,n} = (c_1^0, \ldots, c_n^0)$. Denote (c^n, l^n) the solution of (A.1), such that $c_i^n(0) = c_i^0$, for $i \geq 0$. For $j > n$, we note $c_j^n = 0$, and as such $(c_i^n)_{i\geq 0} \in X_+$.

The bounds on c^n and l^n are straightforward, as seen in the previous proof.

We need to bound uniformly in n the derivatives $\frac{dc_i^n}{\mathrm{d}t}$ and $\frac{dl^n}{\mathrm{d}t}$. The cases for $i = 1, 0$ are treated separatly because the bounds are slightly different.

Let $i \geq 2$. We choose n large enough so that $i < n$.

$$\begin{aligned}
|\frac{dc_i^n}{\mathrm{d}t}| &= |a_{i-1}\frac{l^n\Lambda}{l^n\Lambda+\kappa}c_{i-1}^n - (b_i + a_i\frac{l^n\Lambda}{l^n\Lambda+\kappa})c_i^n + b_{i+1}c_{i+1}^n| \\
&\leq \left(\frac{a_{i-1}}{i-1} + \frac{a_i}{i} + \frac{b_i}{i} + \frac{b_{i+1}}{i+1}\right)(||c^0||_X + l_0) \\
&\leq C_i < (\frac{3C_a}{2} + 4C_b)(||c^0||_X + l_0)
\end{aligned}$$

For $i = 1$, we get :

$$\begin{aligned}
|\frac{dc_1^n}{\mathrm{d}t}| &= |a_0\frac{l^n\Lambda}{l^n\Lambda+\kappa}c_0^n - (b_1 + a_1\frac{l^n\Lambda}{l^n\Lambda+\kappa})c_1^n + b_2c_2^n| \\
&\leq a_0 m + (a_1 + b_1 + \frac{b_2}{2})(||c^0||_X + l_0) \\
&\leq C_1
\end{aligned}$$

And finally for $i = 0$:

$$\begin{aligned}
|\frac{dc_0^n}{\mathrm{d}t}| &= |a_0\frac{l^n\Lambda}{l^n\Lambda+\kappa}c_0^n - b_1c_1^n| \\
&\leq a_0 m + b_1(||c^0||_X + l_0) \\
&\leq C_0
\end{aligned}$$

For l^n we have the following inequality :

$$\begin{aligned}
|\frac{dl^n}{\mathrm{d}t}| &= |\sum_{i=0}^{n-1} J_i| \\
&\leq C_0 + \sum_{i=1}^{n-1}(a_i\frac{l^n\Lambda}{l^n\Lambda+\kappa}c_i^n + b_{i+1}c_{i+1}^n) \\
&\leq C_0 + (C_a + C_b)m + C_b\delta(||c^0||_X + l_0)
\end{aligned}$$

Hence the sequence of functions $(c_i^n)_{n\geq 1}$ are equicontinuous on $[0, +\infty)$ for each $i \geq 0$ as well as $(l^n)_{n\geq 1}$. We can apply Arzela-Ascoli theorem : by a Cantor diagonalization argument, we can extract a subsequence n_k such that for every $i \geq 0$ there exists a continuous function $c_i : [0, +\infty[\to \mathbb{R}$ with $c_i^{n_k} \to c_i$ as $k \to +\infty$, uniformly on every compact of $[0, +\infty)$ and extract a subsequence $l^{n_k} \to l$ as $k \to \infty$ uniformly on every compact of $[0, \infty)$.

We can now pass to the limit in (A.1) :

$$c_i^{n_k}(t) = c_i^0 + \int_0^t \Big(J_{i-1}(c^{n_k}(s), l^{n_k}(s)) - J_i(c^{n_k}(s), l^{n_k}(s))\Big)\mathrm{d}s$$

And by uniform convergence of $c_i^{n_k}$ and l^{n_k} :

$$c_i(t) = c_i^0 + \int_0^t \Big(J_{i-1}(c(s), l(s)) - J_i(c(s), l(s))\Big)\mathrm{d}s \tag{A.4}$$

For l^{n_k}, we have :

$$l^{n_k}(t) = l_0 - \int_0^t \sum_{i=0}^{n_k-1} (a_i \frac{l^{n_k}(s)\Lambda}{l^{n_k}(s)\Lambda + \kappa} c_i^{n_k}(s) - b_{i+1} c_{i+1}^{n_k}(s)) \mathrm{d}s \tag{A.5}$$

To obtain convergence, the sum is split in two parts : $\sum_{i=0}^{n_k-1} = \sum_{i=0}^{P} + \sum_{i=P+1}^{n_k-1}$. The first converges since we have already shown convergence of the c_i's. Convergence of the second sum is obtain with the following bound :

$$\sum_{i=P+1}^{n_k-1} |a_i \frac{l^{n_k}(s)\Lambda}{l^{n_k}(s)\Lambda + \kappa} c_i^{n_k}(s) - b_{i+1} c_{i+1}^{n_k}(s)| \leq C_a \sum_{i=P+1}^{n_k-1} c_i(s) + \sup_{i>P} \frac{b_i}{i} \sum_{i=P+1}^{n_k-1} i c_i(s)$$

Both terms on the right go to zero as $P \to \infty$ thanks to the fact that the zeroth and first order moments of a solution are bounded and the hypothesis on b.

Then by the dominated convergence theorem :

$$l(t) = l_0 + \int_0^t \sum_{i=0}^{\infty} J_i(c(s), l(s)) \mathrm{d}s \tag{A.6}$$

We have shown that (c, l) is a solution of (1.24) and (iv) is a direct consequence of the use of the dominated convergence theorem. □

A.2 Lifshitz-Slyozov model

The goal of this section is to prove theorem 1.4.1. The proof in itself relies on the description of the characteristic curves for the transport equation and a fixed point argument for the conservation equation. We refer the viewer to [33] for the proof in the case of the classical Lifshitz-Slyozov equations.

We consider the following system of equations :

$$\begin{cases} \partial_t f + \partial_x (v(x, L) f) = 0 \\ L(t) + \int_{\mathbb{R}_+} x f(t, x) \mathrm{d}x = \rho \end{cases} \tag{A.7}$$

where v is of the form : $v(x, L) = a(x) \dfrac{L}{L + \kappa} - b(x)$. We impose null-flux boundary conditions and thus :

$$f(t, x = 0) = 0, \text{ for all } t > 0 \tag{A.8}$$

Let us recall the definition of characteristic curves :

Definition A.2.1 (Characteristic curves)**.** *The characteristics associated to* (A.7) *are solutions to :*

$$\begin{cases} \partial_s X(s; t, x) = v(X(s; t, x), L(s)) \\ X(t; t, x) = x \end{cases} \tag{A.9}$$

We denote $I_{t,x}$ its maximal interval. Here we assume $L \in C^0(\mathbb{R}_+)$ to be given. Since v is C^1 in both x and L the characteristics are uniquely defined.

Property A.2.1. *By simple calculations we have the following formulas :*

- $X(t;s,X(s;t,x))=x$
- $J(s;t,x)=\partial_x X(s;t,x)=\exp(\int_s^t -\partial_x v(X(\sigma;t,x),L(\sigma))d\sigma)$
- $\partial_t X(s;t,x)=-v(x,L(t))J(s;t,x)$

To use the characteristics formulation for a solution to the Lifshotz-Slyozov model, we need to describe the life-time of these characteristics. Therefore we consider the entry-time of those characteristics in Ω_T:

$$\sigma_t(x)=inf I_{t,x} \tag{A.10}$$

Now two cases arise : either a characteristic goes back to a positive x at time $s=0$ or they go back to the boundary $x=0$ in positive time $s>0$. Since the family of characteristics is an ordered family, we can split them according to the characteristic that reachs the boundary $x=0$ at time $s=0$:

$$X_c(t)=inf\{x>0|\sigma_t(x)=0\}=X(t;0,0) \tag{A.11}$$

Lemma A.2.1. *Let $(t,x)\in\Omega_T^*$. Then if $\sigma_t(x)=0$, $X(s;t,x)>0$ for all $0<s<t$.*

Proof. The proof is a direct consequence of the definitions of $I_{t,x}$, $\sigma_t(x)$ and the fact that solutions are defined on $\mathbb{R}_+^*$. □

Lemma A.2.2. *There exists some constant $K>0$ such that for all $s>t$, $X(s;t,x)\leq x\exp(K(s-t))$.*

Proof of the preceding lemma.

$$\begin{aligned}X(s;t,x)=&X(t;t,x)+\int_t^s v(X(\sigma;t,x),L(\sigma))d\sigma\\ \leq&x+\int_t^s a(X(\sigma;t,x))d\sigma\leq x+C_a\int_t^s X(\sigma;t,x)d\sigma\end{aligned}$$

We conclude by Grönwall's lemma :

$$X(s;t,x)\leq x\exp(C_a(s-t))$$

□

Lemma A.2.3. *For all $t\in(0,T)$, and $s\in[0,t)$, the map $x\to X(t;s,x)$ is a strictly increasing $\mathcal{C}^1$-diffeomorphism from $(0,+\infty)$ to $(X(t;s,0)),+\infty)$*

Proof. Since the vector field v is C^1 on $\mathbb{R}_+$ in both variable, the associated flow $x\to X(t;s,x)$ is a C^1-diffeomorphism onto its image, by Cauchy-Lipshitz theory. Since its derivative is strictly positive it is also strictly increasing. We conclude by showing that $\lim_{x\to\infty}X(t;s,x)=+\infty$. By lemma A.2.2, we have :

$$x=X(s;t,X(t;s,x))\leq X(t;s,x)\exp(C_a(s-t))\leq X(t;s,x)C(T)$$

Taking the limit on both sides as $x\to+\infty$ yields the result. □

Definition A.2.2 (Mild solution). *Given a smooth initial condition f^0 and $L \in C^0(\mathbb{R}_+)$ a mild solutions of*

$$\begin{cases} \partial_t f + \partial_x(v(x, L(t))f) = 0 \\ f(0,x) = f^0(x) \end{cases}$$

is given by :

$$f(t,x) = f^0(X(0;t,x))J(0;t,x)\mathbb{1}_{(X_c(t),\infty)}(x) \tag{A.12}$$

Definition A.2.3 (Solutions to (A.7)). *A couple (f, L) is said to be a solution of* (A.7) *iif f is a mild solution in the sense of A.2.2 associated to L and $L : \mathbb{R}_+ \to \mathbb{R}_+$ solves $L(t) + \int_{\mathbb{R}_+} xf(t,x)\mathrm{d}x = \rho$.*

Remark. *Since we impose null-flux boundary conditions on this system : $v(x, L(t))f(t,x)|_{x=0} = 0$, there is no term involving $\mathbb{1}_{(0,X_c(t))}(x)$.*

Lemma A.2.4 (Conservation of moments).

$$\sup_{t\in[0,T)} \int_{\mathbb{R}_+} (1+x)f(t,x)\mathrm{d}x < +\infty \tag{A.13}$$

Proof. The proof relies on lemma A.2.3. For $t \in [0,T)$, one has :

$$\begin{aligned} \int_{\mathbb{R}_+} f(t,x)\mathrm{d}x &= \int_{\mathbb{R}_+} f^0(X(0;t,x))J(0;t,x)\mathbb{1}_{(X_c(t),\infty)}(x)\mathrm{d}x \\ &= \int_{\mathbb{R}_+} f^0(x)\mathrm{d}x \end{aligned}$$

Similarly :

$$\begin{aligned} \int_{\mathbb{R}_+} xf(t,x)\mathrm{d}x &= \int_{\mathbb{R}_+} xf^0(X(0;t,x))J(0;t,x)\mathbb{1}_{(X_c(t),\infty)}(x)\mathrm{d}x \\ &= \int_{\mathbb{R}_+} X(T;0,x)f^0(x)\mathrm{d}x \\ &\leq \exp(KT)\int_{\mathbb{R}_+} xf^0(x)\mathrm{d}x \end{aligned}$$

In both cases, $f^0 \in L^1(\mathbb{R}_+, (1+x)\mathrm{d}x)$ which concludes the proof.

□

Definition A.2.4 (Fixed-point map). *We denote $\mathcal{T}$ the following map :*

$$\mathcal{T} : L \in C^0([0,T]) \to \mathcal{T}L(t) = [\rho - \int_{\mathbb{R}_+} xf(t,x)\mathrm{d}x]_+$$

where f is a solution to (A.2.2) *associated with L, and $[.]_+$ simply denotes the positive part. We also denote $\mathcal{B}$ the set of eligible L functions :*

$$\mathcal{B} = \{L \in C^0([0,T]) \text{ s.t. } \forall t > 0,\ 0 \leq L(t) \leq \rho\}$$

Property A.2.2. *$\mathcal{T}$ is continuous*

We will need the following lemmas for the proof.

Lemma A.2.5. *Let $L_1, L_2 \in \mathcal{B}$ and X_1, X_2 their respective characteristics. Let us denote $z = X_2 - X_1$. Then there exists some constants $K_1, K_2 > 0$ such that for $t \leq t_1 \leq t_2$:*

$$\begin{aligned} |z(t_2;t,x)| \leq\ & |z(t_1;t,)| \exp(K_1(t_2-t_1)) \\ & + K_2 \int_{t_1}^{t_2} (1 + x\exp(K(\sigma - t)))|L_2(\sigma) - L_1(\sigma)| \exp(K_1(t_2-\sigma))d\sigma \end{aligned} \tag{A.14}$$

And in particular, there exists $C, K > 0$ such that :

$$|z(s;0,x)| \leq C\exp(KT)T\|L_2 - L_1\|_{C^0([0,T])}(1+x) \tag{A.15}$$

Also there exists some constants $K_1, K_2 > 0$ such that for $t_2 \leq t_1 \leq t$:

$$\begin{aligned} |z(t_2;t,x)| \leq\ & |z(t_1;t,x)| \exp(K_1(t_1-t_2)) \\ & + K_2(1+x)\exp(Kt - K_1t_2) \int_{t_2}^{t_1} |L_2(\sigma) - L_1(\sigma)| \exp((K_1 - K)\sigma)d\sigma \end{aligned} \tag{A.16}$$

And in particular, there exists $C, K > 0$ such that :

$$|z(0;t,0)| \leq C\exp(KT)T\|L_2 - L_1\|_{C^0([0,T])} \tag{A.17}$$

Proof of the preceding lemma. Consider the times $t \leq t_1 \leq t_2$. Let $t_1 \leq \sigma \leq s$. Then :

$$\begin{aligned} |v(X_2(\sigma;t,x), L_2(\sigma)) - v(X_1(\sigma;t,x), L_1(\sigma))| & \\ \leq |v(X_2(\sigma;t,x), L_2(\sigma)) - v(X_1(\sigma;t,x), L_2(\sigma))| & \\ + |v(X_1(\sigma;t,x), L_2(\sigma)) - v(X_1(\sigma;t,x), L_1(\sigma))| & \\ \leq K_1|z(\sigma;t,x)| + K_2(1 + X_1(\sigma;t,x))|L_2(\sigma) - L_1(\sigma)| & \end{aligned}$$

We have the following equality :

$$z(t_2;t,x) = z(t_1;t,x) + \int_{t_1}^{t_2} (v(X_2(\sigma;t,x), L_2(\sigma)) - v(X_1(\sigma;t,x), L_1(\sigma)))d\sigma$$

Then by Grönwall's lemma:

$$\begin{aligned} |z(t_2;t,x)| \leq & |z(t_1;t,x)| + \int_{t_1}^{t_2} K_1|z(\sigma;t,x)|d\sigma \\ & + \int_{t_1}^{t_2} K_2(1 + X_1(\sigma;t,x))|L_2(\sigma) - L_1(\sigma)|d\sigma \\ \leq & |z(t_1;t,x)| \exp(K_1(t_2-t_1)) \\ & + K_2 \int_{t_1}^{t_2} (1 + X_1(\sigma;t,x))|L_2(\sigma) - L_1(\sigma)| \exp(K_1(t_2-\sigma))d\sigma \\ \leq & |z(t_1;t,x)| \exp(K_1(t_2-t_1)) \\ & + K_2 \int_{t_1}^{t_2} (1 + x\exp(K(\sigma - t)))|L_2(\sigma) - L_1(\sigma)| \exp(K_1(t_2-\sigma))d\sigma \end{aligned}$$

Then, let $t = t_1 = 0$.

$$\begin{aligned}|z(s;0,x)| \leq & |z(0;0,x)| \exp(K_1 s) + K_2 \int_0^s (1 + x\exp(K\sigma))|L_2(\sigma) - L_1(\sigma)| \exp(K_1(s-\sigma))d\sigma \\ \leq & x\exp(K_1 s) + K_2 \exp(K_1 s)(1+x)\int_0^s |L_2(\sigma) - L_1(\sigma)|d\sigma \\ \leq & C\exp(K_1 T)T\|L_2 - L_1\|_{C^0([0,T])}(1+x)\end{aligned}$$

The proof of (A.15) is done similarly but applied tp $z(t-s;t,x)$. (A.16) follows from previous bounds. □

Proof. Consider a sequence $(L_n)_{n\in\mathbb{N}}$ in $\mathcal{B}$ such that it converges to $L \in C^0([0,T])$ for the uniform norm. We want to show that :

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |\int_{\mathbb{R}_+} x f_n(t,x)\mathrm{d}x - \int_{\mathbb{R}_+} x f(t,x)\mathrm{d}x| = 0$$

$$\begin{aligned}|\int_{\mathbb{R}_+} x f_n(t,x)\mathrm{d}x - & \int_{\mathbb{R}_+} x f(t,x)\mathrm{d}x| \\ & = |\int_{\mathbb{R}_+} x f^0(X_n(0;t,x))J_n(0;t,x)\mathrm{d}x - \int_{\mathbb{R}_+} x f^0(X(0;t,x))J(0;t,x)\mathrm{d}x| \\ & = |\int_{X_n(0;t,0)}^{\infty} X_n(t;0,y)f^0(y)\mathrm{d}y - \int_{X(0;t,0)}^{\infty} X(t;0,y)f^0(y)\mathrm{d}y| \\ & \leq \int_{X(0;t,0)}^{\infty} |z_n(t;0,y)|f^0(y)\mathrm{d}y + |\int_{X_n(0;t,0)}^{X(0;t,0)} X_n(t;0,y)f^0(y)\mathrm{d}y|\end{aligned}$$

The first integral is treated using (A.15) and thus converges toward 0 uniformly in t as $n \to \infty$. For the second integral, we proceed as follow. Let y lie between $X_n(0;t,0)$ and $X(0;t,0)$. Then :

$$\begin{aligned}|X_n(t;0,y)| &= |X_n(t;0,y) - X_n(t;0,X_n(0;t,0))| \\ &= |\int_y^{X_n(0;t,0)} \partial_x X_n(t;0,\zeta)d\zeta| \\ &= |\int_y^{X_n(0;t,0)} \exp(\int_0^t \partial_x v(X_n(\sigma;0,\zeta),L_n(\sigma))d\sigma)d\zeta \\ &\leq \exp((K_1+K_2)T)|y - X_n(0;t,0)| \\ &\leq \exp((K_1+K_2)T)|X(0;t,0) - X_n(0;t,0)| \\ &= \exp((K_1+K_2)T)|z_n(0;t,0)| \\ &\leq \exp((K_1+K_2)T)C\exp(KT)T\|L_n - u\|_{C^0([0,T])}\end{aligned}$$

Hence for y belonging to the segment between $X_n(0;t,0)$ and $X(0;t,0)$, $|X_n(t;0,y)|$ converges uniformly in t and y towards 0 as $n \to \infty$. This concludes the proof of the required convergence. □

Lemma A.2.6. *The image of $\mathcal{B}$ through $\mathcal{T}$ is a precompact subset of $C^0([0,T])$.*

In order to prove this lemma, we need to control the derivability of the first moment of f. This is done using a weak formulation of the transport problem and showing the following property :

Property A.2.3. *Let f be the mild solution to* (A.7). *Then f belongs to $C^1([0,T);$ weak-*-$L^1(0,\infty;(1+x)\mathrm{d}x))$ and for every h locally bounded such that $h' \in L^\infty(0,\infty)$:*

$$\int_{\mathbb{R}_+} h(x)f(t,x)\mathrm{d}x = \int_{\mathbb{R}_+} h(x)f^0(x)\mathrm{d}x + \int_0^t \int_{\mathbb{R}_+} v(x,L(s))h'(x)f(s,x)dxds \tag{A.18}$$

Proof. Let $\varphi \in C_c^1([0,T)\times[0,\infty))$. Denote $\Psi = -(\partial_t\varphi + v\partial_x\varphi)$. Then by the definition of a mild solution, Fubini theorem and lemma A.2.3, one has :

$$\begin{aligned}\int_0^T \int_{\mathbb{R}_+} (\partial_t\varphi + v\partial_x\varphi)f(s,x)dxds &= -\int_0^T \int_{X_c(s)}^\infty \Psi(s,x)f^0(X(0;s,x))J(0;s,x)dxds \\ &= -\int_{\mathbb{R}_+} (\int_0^T \Psi(s,X(s;0,x)))f^0(x)\mathrm{d}x\end{aligned}$$

And by the definition of characteristic curves, for all $s \in (\sigma_t(x),T)$:

$$\partial_s\varphi(s,X(s;t,x)) = -\Psi(s,X(s;t,x))$$

Hence $\varphi(0,x) = \int_0^T \Psi(s,X(s;t,x))\mathrm{d}s$. And thus :

$$\int_0^T \int_{\mathbb{R}_+} (\partial_t\varphi + v\partial_x\varphi)f(s,x)dxds = \int_{\mathbb{R}_+} \varphi(0,x)f^0(x)\mathrm{d}x$$

We conclude using a series of regularizaton argument. First the previous equality can be extend to functions $\varphi(t,x) = g(t)h(x)$ with $g \in C_c^1(0,T)$ and $h \in C_c^0([0,\infty))$ with $h' \in L^\infty([0,\infty))$ using $f \in L^\infty([0,T),L^1(\mathbb{R}_+))$, lemma A.2.3 and the fact that the velocity is locally bounded. Again the result can be extended to function h locally bounded such that $h' \in L^\infty([0,\infty))$ using $f \in L^\infty([0,T),L^1(\mathbb{R}_+,(1+x)\mathrm{d}x))$ and the sublinearity of the velocity :

$$\int_0^T g'(t)\int_{\mathbb{R}_+} h(x)f(t,x)\mathrm{d}x\mathrm{d}t + \int_0^T g'(t)\int_{\mathbb{R}_+} v(x,L(t))h'(x)f(t,x)\mathrm{d}x\mathrm{d}t = 0$$

Hence the map $t \to \int_{\mathbb{R}_+} h(x)f(t,x)\mathrm{d}x$ has a bounded time derivative, which leads to (A.18). This also implies that the first moment of f as a time derivative by choosing $h(x) = x$.

□

Proof of lemma A.2.6. Let $u \in \mathcal{B}$. Then applying Stampacchia's theorem to $[.]_+$ and $\rho - \int_{\mathbb{R}_+} xf(t,x)\mathrm{d}x$ allows us to write :

$$\frac{d}{\mathrm{d}t}\mathcal{T}(u)(t) = \begin{cases} 0 & \text{if } \int_{\mathbb{R}_+} xf(t,x)\mathrm{d}x > \rho \\ -\dfrac{d}{\mathrm{d}t}(\int_{\mathbb{R}_+} xf(t,x)\mathrm{d}x) & \text{else} \end{cases} \tag{A.19}$$

Then by proposition A.2.3 and lemma A.2.3 the derivative $\dfrac{d}{\mathrm{d}t}\mathcal{T}(u)(t)$ is uniformly bounded independently of u. As such the image of $\mathcal{B}$ through $\mathcal{T}$ is a compact for the uniform topology. □

We are know ready to apply Schauder's fixed point theorem and proof the following theorem :

Theorem A.2.1. *Under the previous assumptions,* (A.7) *has a solution on the interval* $[0,T]$.

Proof. By Schauder's fixed point theorem, we have a fixed point $u \in \mathcal{B}$. The only thing left is to prove that u satisfies the relation in (A.7) instead of the one involving $[.]_+$.

Assume there exists a time t_* such that :

$$\int_0^\infty xf(t_*,x)\mathrm{d}x = \rho \text{ and } \int_0^\infty xf(t,x)\mathrm{d}x = \rho \text{ for } t < t_* \tag{A.20}$$

Since $\mathcal{T}(c) = c$, this implies that $c(t_*) = 0$. Then

$$\frac{d}{\mathrm{d}t}\int_0^\infty xf(t_*,x)\mathrm{d}x = \int_0^\infty v(x,L(t_*))f(t_*,x)\mathrm{d}x = -\int_0^\infty b(x)f(t_*,x)\mathrm{d}x$$

which is non-positive since $b(x)f(t_*,x) \geq 0$. Since it cannot vanish without contradicting (A.20), we have

$$\frac{d}{\mathrm{d}t}\int_0^\infty xf(t_*,x)\mathrm{d}x < 0$$

Now if $t_* = 0$, then $\int_0^\infty xf(t,x)\mathrm{d}x < \rho$ on a neighborhood of 0. Else if $t_* > 0$, then in some neighborhood of t_*, $t \to \int_0^\infty xf(t,x)\mathrm{d}x$ is non-increasing which contradicts (A.20). Hence we obtain $\int_0^\infty xf(t,x)\mathrm{d}x < \rho$ for al time $t > 0$. Thus completing the proof.

□

Lemma A.2.7 (Uniqueness)**.** *Let* (f_1, L_1) *and* $(f_2, L2)$ *be two solutions of* (A.7) *in the sense of definition A.2.3 with initial conditions* (f_1^0, L_1^0) *and* (f_2^0, L_2^0)*. Then :*

$$\|L_2 - L_1\|_{C^0([0,T])} \leq C(|\rho_2 - \rho_1| + \|f_2^0 - f_1^0\|_{L^1(R_+,xdx)})$$

Proof. Using the mild solution expression and the conservation equation, one gets :

$$\begin{aligned} |L_2(t) - L_1(t)| \leq & |\rho_2 - \rho_1| + \int\limits_{X_1(0;t,0)}^{\infty} |z(t;0,y)| f_2^0(y)\mathrm{d}y \\ & + \int\limits_{X_1(0;t,0)}^{\infty} X_1(t;0,y)|f^0,2(y) - f_1^0(y)|\mathrm{d}y \\ & + \Big| \int\limits_{X_2(0;t,0)}^{X_1(0;t,0)} X_2(t;0,y) f_2^0(y)\mathrm{d}y \Big| \end{aligned} \tag{A.21}$$

In the first integral on the right, we first use (A.14) to get :

$$z(t;0,y) \leq K_2 \exp(K_1 t)(1+y)\int_0^t |L_2(s) - L_1(s)| \exp((K - K_1)s)\mathrm{d}s$$

hence

$$\int_{X_1(0;t,0)}^{\infty} |z(t;0,y)| f_2^0(y)\mathrm{d}y \leq K\exp((K_1+K_2)T)(\int_{\mathbb{R}_+}(1+y)f_2^0(y)\mathrm{d}y)\int_0^t |L_2(s)-L_1(s)|\mathrm{d}s$$

Using A.2.2 we boudn the second integral on the right :

$$\int_{X_1(0;t,0)}^{\infty} X_1(t;0,y)|f^0,2(t;0,y)-f_1^0(t;0,y)|\mathrm{d}y \leq \exp(KT)\int_0^{\infty} y|f^0,2(y)-f_1^0(y)|\mathrm{d}y$$

The last integral is bounded using a method used previously :

$$\begin{aligned}
|\int_{X_2(0;t,0)}^{X_1(0;t,0)} X_2(t;0,y)f_2^0(y)\mathrm{d}y| &= |\int_{X_2(0;t,0)}^{X_1(0;t,0)} (X_2(t;0,y)-X_2(t;0,X_2(0;t,0)))f_2^0(y)\mathrm{d}y| \\
&\leq |\int_{X_2(0;t,0)}^{X_1(0;t,0)}\int_{X_2(0;t,0)}^{y} \frac{\partial X_2}{\partial_x}(t;0,\xi)f_2^0(y)d\xi\mathrm{d}y| \\
&\leq \exp(K_1t)|z(0;t,0)|\int_0^{\infty} f_2^0(y)\mathrm{d}y
\end{aligned}$$

We have the final bound using (A.16) :

$$\begin{aligned}
&|\int_{X_2(0;t,0)}^{X_1(0;t,0)} X_2(t;0,y)f_2^0(y)\mathrm{d}y| \\
&\leq \exp((K_1+K)t)K_2\|f_2^0\|_{L^1(\mathbb{R}_+,\mathrm{d}x)}\int_0^t |L_2(s)-L_1(s)|\exp((K_1-K)s)\mathrm{d}s\|f_2^0\|_{L^1(\mathbb{R}_+,\mathrm{d}x)} \\
&\leq \kappa(T)\int_0^t |L_2(s)-L_1(s)|\exp((K_1-K)s)\mathrm{d}s
\end{aligned}$$

Henc we can write :

$$|L_2(t)-L_1(t)| \leq |\rho_2-\rho_1| + \bar{K}(T)\int_0^t |L_2(s)-L_1(s)|\mathrm{d}s + \kappa(T)\|f_2^0-f_1^0\|_{L^1(\mathbb{R}_+,xdx)} \tag{A.22}$$

The proof ends using Grönwall's lemma.

□

Appendix B

Classical convergence result

In appendix A, we showed that the Becker-Döring system (1.24) has a unique continuous solution (c, l) in $C^1([0,T], X_+) \times C([0,T], \mathbb{R}_+)$. We will now show that after some rescaling procedures, this solution can be related to the solution of the Lifshitz-Slyozov equations. A classical proof can be found in [129]. We will use similar techniques,the main difference lying in the definition of the speed and the hypotheses on functions a and b.

B.1 Classical technique

From then on, we assume hypothesis H4 to hold for $(a_i)_{i\geq 0}$ and $(b_i)_{i\geq 1}$. Let us recall the rescaled Becker-Döring model div

$$\begin{cases} \dfrac{dc_i^\varepsilon}{\mathrm{d}t} = \dfrac{1}{\varepsilon}(J_{i-1}^\varepsilon(c^\varepsilon, L^\varepsilon) - J_i^\varepsilon(c^\varepsilon, L^\varepsilon)),\ \forall i \geq 1, & \text{(B.1a)} \\ \dfrac{dc_0^\varepsilon}{\mathrm{d}t} = -\dfrac{1}{\varepsilon} J_0^\varepsilon(c^\varepsilon, L^\varepsilon), & \text{(B.1b)} \\ L^\varepsilon(t) + \displaystyle\sum_{i=0}^{\infty} i\varepsilon^2 c_i^\varepsilon(t) = \lambda,\ \forall t \geq 0, & \text{(B.1c)} \\ L^\varepsilon(0) = L^{\varepsilon,0}, \quad c_i^\varepsilon(0) = c_i^{\varepsilon,0},\ \forall i \geq 1. & \text{(B.1d)} \end{cases}$$

We define the following step functions depending on both time and space, where f^ε is a candidate for convergence to a solution of the Lifshitz-Slyozov system. Let $\Gamma_i^\varepsilon = [(i-\frac{1}{2})\varepsilon, (i+\frac{1}{2})\varepsilon[$ and we write :

$$f^\varepsilon(t,x) = \sum_{i\geq 0} \mathbf{1}_{\Gamma_i^\varepsilon}(x) c_i^\varepsilon(t)$$

Similarly we build the two space functions a^ε and b^ε as :

$$a^\varepsilon(x) = \sum_{i\geq 0} \mathbf{1}_{\Gamma_i^\varepsilon}(x) a_i^\varepsilon$$

$$b^\varepsilon(x) = \sum_{i\geq 1} \mathbf{1}_{\Gamma_i^\varepsilon}(x) b_i^\varepsilon$$

We begin by proving the following theorem :

Theorem B.1.1. *Let $(g_i)_{i\geq 0}$ be a given sequence. Let $(c^\varepsilon, L^\varepsilon)$ be the solution of (B.1) on $[0,T)$, $0<T\leq+\infty$.*

Assume that for all $0\leq t_1<t_2<T$, $\int_{t_1}^{t_2}\sum_{i=0}^{\infty}|g_{i+1}-g_i|a_i^\varepsilon c_i^\varepsilon(t)\mathrm{d}t<\infty$ and that either of the following holds :

(a) $g_i=\mathcal{O}(i)$ and $\int_{t_1}^{t_2}\sum_{i=0}^{\infty}|g_{i+1}-g_i|b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t<\infty$

(b) $\sum_{i=0}^{\infty}g_i c_i^\varepsilon(t_k)<\infty$, for $k=1,2$ and $g_{i+1}\geq g_i\geq 0$ for i large enough

Then :

$$\begin{aligned}\sum_{i=0}^{\infty}g_i c_i^\varepsilon(t_2)-\sum_{i=0}^{\infty}g_i c_i^\varepsilon(t_1)+\int_{t_1}^{t_2}\sum_{i=0}^{\infty}\frac{g_{i+1}-g_i}{\varepsilon}b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t\\ =\int_{t_1}^{t_2}\sum_{i=0}^{\infty}\frac{g_{i+1}-g_i}{\varepsilon}a_i^\varepsilon\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}c_i^\varepsilon(t)\mathrm{d}t\end{aligned}\tag{B.2}$$

Proof. First, since $(c^\varepsilon, L^\varepsilon)$ is a solution of (B.1), we have for all $m\geq 1$:

$$\begin{aligned}\sum_{i=m}^{n}g_i c_i^\varepsilon(t_2)-\sum_{i=m}^{n}g_i c_i^\varepsilon(t_1)=\int_{t_1}^{t_2}\sum_{i=m}^{n}\frac{g_{i+1}-g_i}{\varepsilon}J_i^\varepsilon(t)\mathrm{d}t\\ +\frac{g_m}{\varepsilon}\int_{t_1}^{t_2}J_{m-1}^\varepsilon(t)\mathrm{d}t-\frac{g_{n+1}}{\varepsilon}\int_{t_1}^{t_2}J_n^\varepsilon(t)\mathrm{d}t\end{aligned}\tag{B.3}$$

By reorganizing some terms we get :

$$\begin{aligned}\sum_{i=m}^{n}g_i c_i^\varepsilon(t_2)-\sum_{i=m}^{n}g_i c_i^\varepsilon(t_1)+\int_{t_1}^{t_2}\sum_{i=m}^{n}\frac{g_{i+1}-g_i}{\varepsilon}b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t\\ =\frac{g_m}{\varepsilon}\int_{t_1}^{t_2}J_{m-1}^\varepsilon(t)\mathrm{d}t-\frac{g_{n+1}}{\varepsilon}\int_{t_1}^{t_2}J_n(t)\mathrm{d}t\\ +\int_{t_1}^{t_2}\sum_{i=m}^{n}\frac{g_{i+1}-g_i}{\varepsilon}a_i^\varepsilon\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}c_i^\varepsilon(t)\mathrm{d}t\end{aligned}\tag{B.4}$$

By properties of a solution of (B.1) we know that $L^\varepsilon(t_2)-L^\varepsilon(t_1)=\int_{t_1}^{t_2}\sum_{i=0}^{\infty}J_i^\varepsilon(s)\mathrm{d}s$. Thus we have that $\lim_{n\to\infty}\int_{t_1}^{t_2}J_n^\varepsilon(t)\mathrm{d}t=0$. Let $g_i=1$ and take the limit as $n\to\infty$ in (B.4), we obtain :

$$\sum_{i=m}^{\infty}(c_i^\varepsilon(t_2)-c_i^\varepsilon(t_1))=\int_{t_1}^{t_2}J_{m-1}^\varepsilon(t)\mathrm{d}t$$

Replacing m by $n+1$ we have the following limits in the case 1 for $k=1,2$:

$$|g_{n+1}|\sum_{i=n+1}^{\infty}c_i^\varepsilon(t_k)\leq C(n+1)\sum_{i=n+1}^{\infty}c_i^\varepsilon(t_k)\leq C\sum_{i=n+1}^{\infty}ic_i^\varepsilon(t_k)$$

Or in the case 2:

$$|g_{n+1}| \sum_{i=n+1}^{\infty} c_i^\varepsilon(t_k) \leq C \sum_{i=n+1}^{\infty} g_i c_i^\varepsilon(t_k)$$

Both inequalities lead to $\lim_{n\to+\infty} |g_{n+1}| \sum_{i=n+1}^{\infty} c_i^\varepsilon(t_k) = 0$ and consequently to :

$$\lim_{n\to\infty} |g_{n+1}| \int_{t_1}^{t_2} J_n^\varepsilon(t)\mathrm{d}t = 0 \tag{B.5}$$

Since $\dfrac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa}$ is bounded and $\int_{t_1}^{t_2} \sum_{i=m}^{\infty} |g_{i+1}-g_i| a_i^\varepsilon c_i^\varepsilon(t)\mathrm{d}t < \infty$, we have :

$$\lim_{n\to\infty} \int_{t_1}^{t_2} \sum_{i=m}^{n} |g_{i+1}-g_i| a_i^\varepsilon \frac{L^\varepsilon(t)}{L^\varepsilon(t)+T_\theta} c_i^\varepsilon(t)\mathrm{d}t = \int_{t_1}^{t_2} \sum_{i=m}^{\infty} |g_{i+1}-g_i| a_i^\varepsilon \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} c_i^\varepsilon(t)\mathrm{d}t < \infty \tag{B.6}$$

Convergence is obtained using (B.4), (B.5), (B.6) and the monotone or domianted convergence theorem based on the choise of (g_i):

$$\begin{aligned} \sum_{i=m}^{\infty} g_i c_i^\varepsilon(t_2) - \sum_{i=m}^{\infty} g_i c_i^\varepsilon(t_1) + \int_{t_1}^{t_2} \sum_{i=m}^{\infty} \frac{g_{i+1}-g_i}{\varepsilon} b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t \\ = \frac{g_m}{\varepsilon} \int_{t_1}^{t_2} J_{m-1}^\varepsilon(t)\mathrm{d}t + \int_{t_1}^{t_2} \sum_{i=m}^{\infty} \frac{g_{i+1}-g_i}{\varepsilon} a_i^\varepsilon \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} c_i^\varepsilon(t)\mathrm{d}t \end{aligned} \tag{B.7}$$

We conclude by letting $m = 1$ and adding the terms for c_0^ε.

□

Using the previous theorem, we show the following proposition :

Proposition B.1.1. *Let $\phi \in W^{1,+\infty}_{loc}(\mathbb{R}_+)$, positive such that $\partial_x\phi \in L^{+\infty}(\mathbb{R}_+)$. We denote $\Delta_h\phi(x) = \dfrac{\phi(x+h)-\phi(x)}{h}$. Then for every $t \geq 0$, we have the following equality :*

$$\int_0^{\infty} \phi(x)(f^\varepsilon(t,x) - f^\varepsilon(0,x))\mathrm{d}x = \int_0^t \int_0^{\infty} (\Delta_\varepsilon\phi(x)a^\varepsilon(x)\frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \Delta_{-\varepsilon}\phi(x)b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t \tag{B.8}$$

Proof. We use the previous theorem with the choice $g_i = \int_{\Gamma_i^\varepsilon} \phi(x)\mathrm{d}x$. Since $\partial_x\phi \in L^{+\infty}(\mathbb{R}_+)$, we are using case 1 :

$$\begin{aligned} \sum_{i=0}^{\infty} \int_{\Gamma_i^\varepsilon} \phi(x)\mathrm{d}x (c_i^\varepsilon(t) - c_i^0) + \int_0^t \sum_{i=0}^{\infty} \int_{\Gamma_{i+1}^\varepsilon} \frac{\phi(x)-\phi(x-\varepsilon)}{\varepsilon}\mathrm{d}x\, b_{i+1}^\varepsilon c_{i+1}^\varepsilon(t)\mathrm{d}t \\ = \int_0^t \sum_{i=0}^{\infty} \int_{\Gamma_i^\varepsilon} \frac{\phi(x+\varepsilon)-\phi(x)}{\varepsilon}\mathrm{d}x\, a_i^\varepsilon \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} c_i^\varepsilon(t)\mathrm{d}t \end{aligned} \tag{B.9}$$

We conclude using Fubini's theorem :

$$\int_0^\infty (f^\varepsilon(t,x) - f^\varepsilon(0,x))\phi(x)\mathrm{d}x = \int_0^t \int_0^\infty (\Delta_\varepsilon \phi(x) a^\varepsilon(x) \frac{L^\varepsilon(t)}{L^\varepsilon(t)+\kappa} - \Delta_{-\varepsilon}\phi(x) b^\varepsilon(x)) f^\varepsilon(t,x)\mathrm{d}x\mathrm{d}t$$

□

We now show convergence of $(f^\varepsilon, L^\varepsilon)$ towards a solution of the Lifshitz-Slyozov system. The proof of the following theorem is based on compacity argument of the sequence of $(f^\varepsilon, L^\varepsilon)$ and bounded estimates on moments of order greater than one of f^ε.

First we prove the following lemma :

Lemma B.1.1. *Assume the same statements of the preceding theorem. Then :*

$$\sup_{\varepsilon>0}\sup_{t\in[0,T]} \int_0^\infty (1+x+x^{1+s}) f^\varepsilon(t,x)\mathrm{d}x < +\infty \tag{B.10}$$

Proof. We use theorem (B.1.1) with $g_i = \varepsilon$ and obtain :

$$\int_0^\infty f^\varepsilon(t,x)\mathrm{d}x = \varepsilon \sum_{i\geq 0} c_i^\varepsilon(t) = \varepsilon \sum_{i\geq 0} c_i^{\varepsilon,0} < C_0$$

For the first moment we use the conservation of the quantity of lipids :

$$\int_0^\infty x f^\varepsilon(t,x)\mathrm{d}x < \lambda$$

For the moments of order $1+s$ where $s<1$, we proceed as follows.

Let $e^\varepsilon(x) = \sum_{i\geq 0} \mathbf{1}_{[i\varepsilon,(i+1)\varepsilon[}(x) i\varepsilon$. Hence for all $r>0$, we have :

$$\int_0^\infty (\frac{x}{2})^r f^\varepsilon(t,x)\mathrm{d}x \leq \int_0^\infty (e^\varepsilon(x))^r f^\varepsilon(t,x)\mathrm{d}x \leq \int_0^\infty x^r f^\varepsilon(t,x)\mathrm{d}x$$

As previously we use theorem (B.1.1)) with $g_i = \varepsilon (e^\varepsilon(i\varepsilon))^{1+s}$:

$$\begin{aligned}
\int_0^\infty (e^\varepsilon(x))^{1+s} f^\varepsilon(t,x)\mathrm{d}x &= \varepsilon \sum_{i\geq 0} (i\varepsilon)^{1+s} c_i^\varepsilon(t) \\
&= \varepsilon \sum_{i\geq 0} (i\varepsilon)^{1+s} c_i^{\varepsilon,0} - \int_0^t \sum_{i=1}^\infty ((\varepsilon i+\varepsilon)^{1+s} - (\varepsilon i)^{1+s}) b_{i+1}^\varepsilon c_{i+1}^\varepsilon(\tau) d\tau \\
&+ \int_0^t \sum_{i=0}^\infty ((\varepsilon i+\varepsilon)^{1+s} - (\varepsilon i)^{1+s}) a_i^\varepsilon \frac{L^\varepsilon(\tau)}{L^\varepsilon(\tau)+\kappa} c_i^\varepsilon(\tau) d\tau - \int_0^t \varepsilon^{1+s}\varepsilon^\gamma b_1^\varepsilon c_1^\varepsilon(\tau) d\tau
\end{aligned} \tag{B.11}$$

The first term on the right is less than C_s by the assumption and the second term is negative. Since we want to pass to the limit on $\varepsilon \searrow 0$ we can assume $\varepsilon < 1$. Then for $\varepsilon < 1$ and $s < 1$ we have :

$$(\varepsilon i+\varepsilon)^{1+s} - (\varepsilon i)^{1+s} \leq \varepsilon(1+s)(1+i\varepsilon)$$

We can now give a proper estimate on the last term :

$$\begin{aligned}\int_0^t \sum_{i=0}^{\infty}((\varepsilon i+\varepsilon)^{1+s}-(\varepsilon i)^{1+s})a_i^\varepsilon \frac{L^\varepsilon(\tau)}{L^\varepsilon(\tau)+\kappa}c_i^\varepsilon(\tau)d\tau &\leq \int_0^t \sum_{i\geq 0}\varepsilon(1+s)(1+i\varepsilon)a_i^\varepsilon c_i^\varepsilon(\tau)d\tau \\ &\leq K(1+s)\int_0^t \sum_{i\geq 0}\varepsilon(1+i\varepsilon)c_i^\varepsilon(\tau)d\tau \\ &= K(1+s)\int_0^t\int_0^\infty (1+e^\varepsilon(x))f^\varepsilon(\tau,x)dxd\tau \\ &\leq K(1+s)\int_0^t\int_0^\infty (1+x)f^\varepsilon(\tau,x)dxd\tau \\ &< \infty\end{aligned} \tag{B.12}$$

□

Since $\sup_{\varepsilon>0}\sup_{t\in[0,T]} L^\varepsilon(t) < +\infty$ we know that L^ε converges to u in $L^\infty([0,T])$. We will know show that convergence is also true in $C^0([0,T])$.

Lemma B.1.2. *Assume the same statements of the preceding theorem. Then $(L^\varepsilon)_{\varepsilon>0}$ is equicontinuous.*

Proof. First let $\varphi = x$ in proposition B.1.1 :

$$L^\varepsilon(t+h)-L^\varepsilon(t) = -\int_t^{t+h}\int_0^{+\infty}\Big(a^\varepsilon(x)\frac{L^\varepsilon(\tau)}{L^\varepsilon(\tau)+\kappa}-b^\varepsilon(x)\Big)f^\varepsilon(\tau,x)\mathrm{d}xd\tau \tag{B.13}$$

We have an easy bound for the integral on the right :

$$\begin{aligned}&\Big|\int_t^{t+h}\int_0^{+\infty} a^\varepsilon(x)\frac{L^\varepsilon(\tau)}{L^\varepsilon(\tau)+\kappa}-b^\varepsilon(x)\Big)f^\varepsilon(\tau,x)\mathrm{d}xd\tau\Big| \\ &\qquad\leq K\int_t^{t+h}\int_0^{+\infty} f^\varepsilon(\tau,x)dxd\tau + \int_t^{t+h}\int_0^{+\infty}|b^\varepsilon(x)f^\varepsilon(\tau,x)|dxd\tau\end{aligned}$$

By conservation of the moment of order 0, the first integral is bounded by $const.\times Kh$. We treat the second integral as follows :

$$\begin{aligned}\int_t^{t+h}\int_0^{+\infty}|b^\varepsilon(x)f^\varepsilon(\tau,x)|dxd\tau &\leq \int_t^{t+h}\int_0^{+\infty}(B+br(x)^2)|f^\varepsilon(\tau,x)|\mathrm{d}xd\tau \\ &\leq B\int_t^{t+h}\int_0^{+\infty}|f^\varepsilon(\tau,x)|dxd\tau \\ &+ const.\int_t^{t+h}\int_0^{+\infty}(V_lx+V_0)^{\frac{2}{3}}|f^\varepsilon(\tau,x)|dxd\tau \\ &< C(T)h\end{aligned} \tag{B.14}$$

The last inequality uses the preceding lemma and the bound on the first moment.
Hence $|L^\varepsilon(t+h)-L^\varepsilon(t)| \xrightarrow[h\to 0]{} 0$ uniformly in ε et $t\in[0,T]$.

□

Proof of theorem 1.4.4. Firstly by Arzela-Ascoli's theorem and the lemma B.1.2 the sequence of functions $(L^\varepsilon)_{\varepsilon>0}$ converges uniformly to u in $C^0([0,T])$.

We will now show convergence of the sequence $(f^\varepsilon)_{\varepsilon>0}$.

Let $\varphi \in C_0^\infty([0,+\infty[)$ such that $\text{supp}(\varphi) \subset [0,R]$ where $0 < R < +\infty$. Denote $\varphi^\varepsilon(x) = \sum_{i\geq 0} \mathbf{1}_{[i\varepsilon,(i+1)\varepsilon[}(x)\varphi(i\varepsilon)$.

Hence $|\varphi^\varepsilon(x) - \varphi(x)| \leq \varepsilon||\varphi'||_{L^\infty}$, so φ^ε converges uniformly to φ.

We can us φ^ε in property B.1.1 and we choose $\varepsilon < \delta$, such that $\phi^\varepsilon(x) = 0$ on $[0,\varepsilon[$. We then get :

$$\int_0^\infty \varphi^\varepsilon(x)(f^\varepsilon(t,x) - f^\varepsilon(0,x))\mathrm{d}x = \int_0^t \int_0^\infty f^\varepsilon(\tau,x)\Big(a^\varepsilon(x)\frac{L^\varepsilon(\tau)}{L^\varepsilon(\tau)+\kappa}\Delta_\varepsilon\varphi^\varepsilon(x) - b^\varepsilon(x)\Delta_{-\varepsilon}\varphi^\varepsilon(x)\Big)\mathrm{d}x d\tau \quad \text{(B.15)}$$

First observe that $\Delta_\varepsilon\varphi^\varepsilon \xrightarrow[\varepsilon\to 0]{} \varphi'$ uniformly. By using this in conjonction with lemma B.1.1 and the equation (B.15), the sequence $(\int_0^\infty \varphi^\varepsilon(x)f^\varepsilon(t,x)\mathrm{d}x)_{\varepsilon>0}$ is equicontinuous on $[0,T]$. Again by Arzela-Ascoli's theorem it belongs to a compact subset of $C^0([0,T])$. Moreover since $\varphi^\varepsilon \xrightarrow[\varepsilon\to 0]{}$ φ uniformly on $\mathbb{R}_+$, the sequence $(\int_0^\infty \varphi(x)f^\varepsilon(t,x)\mathrm{d}x)_{\varepsilon>0}$ is also in a compact subset of $C^0([0,T])$. By density, it stays true for every $\varphi \in C_0^0(\mathbb{R}_+^*)$.

Therefore by separability of $C_0^0(\mathbb{R}_+^*)$ and Cantor's diagonal argument we can extract a subsquence (still denoted ε) such that for every $\varphi \in C_0^0(\mathbb{R}_+^*)$ and for every T we have :

$$\int_0^\infty \varphi(x)f^\varepsilon(t,x)\mathrm{d}x \xrightarrow[\varepsilon\to 0]{} \int_0^\infty \varphi(x)f(t,\mathrm{d}x) \quad \text{(B.16)}$$

uniformly on $[0,T]$, where $f \in C^0([0,+\infty[;\mathcal{M}^1(0,+\infty) - weak - *)$. Moreover by lemma B.1.1 we have :

$$(1+|x|^{1+s})f \in L^\infty([0,T);\mathcal{M}^1(0,+\infty)) \quad \text{(B.17)}$$

We can now pass to the limit in (B.15) and get the main result of the theorem :

$$\begin{aligned}\int_0^\infty \varphi(x)f(t,\mathrm{d}x) = &\int_0^\infty \varphi(x)f(0,\mathrm{d}x) \\ &+ \int_0^t \int_0^\infty \Big(a(x)\frac{L(\tau)}{L(\tau)+\kappa} - b(x)\Big)\varphi'(x)f(\tau,\mathrm{d}x)d\tau\end{aligned} \quad \text{(B.18)}$$

We conclude by showing that the conservation equation also passes to the limit. We have that $e^\varepsilon(x) \xrightarrow[\varepsilon\to 0]{} x$ uniformly and by existence of moments of order strictly less than two we obtain :

$$\lim_{\varepsilon\to 0}\Big(L^\varepsilon(t) + \int_0^\infty xf^\varepsilon(t,x)\mathrm{d}x\Big) = \lambda = L(t) + \int_0^\varepsilon xf(t,\mathrm{d}x) \quad \text{(B.19)}$$

Hence we have shown that (f,L) is a measure-valued solution of the Lifshitz-Slyozov system. □

Bibliography

[1] URL `https://github.com/CMA-ES/pycma`.

[2] URL `https://maple.cloud/app/6509768948056064`.

[3] Obesity. https://www.who.int/health-topics/obesity.

[4] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular biology of the cell. 4th edn.* Garland Science, 2002.

[5] J. W. Anderson, S. Vichitbandra, W. Qian, and R. J. Kryscio. Long-term weight maintenance after an intensive weight-loss program. *Journal of the American College of Nutrition*, 18(6):620–627, 1999.

[6] D. P. Andersson, D. Eriksson Hogling, A. Thorell, E. Toft, V. Qvisth, E. Näslund, A. Thörne, M. Wirén, P. Löfgren, J. Hoffstedt, et al. Changes in subcutaneous fat cell volume and insulin sensitivity after weight loss. *Diabetes care*, 37(7):1831–1836, 2014.

[7] E. Arner, P. O. Westermark, K. L. Spalding, T. Britton, M. Rydén, J. Frisén, S. Bernard, and P. Arner. Adipocyte turnover: relevance to human adipose tissue morphology. *Diabetes*, 59(1):105–109, 2010.

[8] P. Arner, S. Bernard, L. Appelsved, K.-Y. Fu, D. P. Andersson, M. Salehpour, A. Thorell, M. Rydén, and K. L. Spalding. Adipose lipid turnover and long-term changes in body weight. *Nature Medicine*, 25(9):1385–1389, 2019.

[9] J. M. Ball, J. Carr, and O. Penrose. The Becker-Döring cluster equations: Basic properties and asymptotic behaviour of solutions. *Communications in mathematical physics*, 104(4): 657–692, 1986.

[10] R. Becker and W. Döring. Kinetische behandlung der keimbildung in übersättigten dämpfen. *Annalen der physik*, 416(8):719–752, 1935.

[11] H. Beirão da Veiga. Diffusion on viscous fluids. existence and asymptotic properties of solutions. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 10(2):341–355, 1983.

[12] O. Bernard, M. Bestard, T. Goudon, L. Meyer, S. Minjeaud, F. Noisette, and B. Polizzi. Numerical schemes for mixture theory models with filling constraint: application to biofilm ecosystems. 2023.

[13] F. Berthelin, T. Goudon, and S. Minjeaud. Multifluid flows: A kinetic approach. *Journal of Scientific Computing*, 66(2):792–824, 2016.

[14] E. Bibbona and R. Sirovich. Strong approximation of density dependent Markov chains on bounded domains, July 2017.

[15] D. Burkholder, E. Pardoux, and A. Sznitman. Ecole d'ete de probabilites de. 1991.

[16] C. Calgaro, E. Creusé, and T. Goudon. Modeling and simulation of mixture flows: Application to powder–snow avalanches. *Computers & Fluids*, 107:100–122, 2015.

[17] V. Calvez, N. Lenuzza, M. Doumic, J.-P. Deslys, F. Mouthon, and B. Perthame. Prion dynamics with size dependency–strain phenomena. *Journal of Biological Dynamics*, 4(1): 28–42, 2010.

[18] J. Calvo, M. Doumic, and B. Perthame. Long-time asymptotics for polymerization models. *Communications in Mathematical Physics*, 363(1):111–137, 2018.

[19] J. Calvo, E. Hingant, and R. Yvinec. The initial-boundary value problem for the Lifshitz–Slyozov equation with non-smooth rates at the boundary. *Nonlinearity*, 34(4):1975, 2021.

[20] F. Campolongo, A. Saltelli, and J. Cariboni. From screening to quantitative sensitivity analysis. a unified approach. *Computer physics communications*, 182(4):978–988, 2011.

[21] J. Cañizo, A. Einav, and B. Lods. Trend to equilibrium for the becker–döring equations: an analogue of cercignani's conjecture. *Analysis & PDE*, 10(7):1663–1708, 2017.

[22] J. A. Canizo, A. Einav, and B. Lods. Uniform moment propagation for the Becker–Döring equations. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 149(4): 995–1015, 2019.

[23] J. A. Canizo, A. Einav, and B. Lods. Uniform moment propagation for the becker–döring equations. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 149(4): 995–1015, 2019.

[24] E. A. Catchpole and B. J. Morgan. Detecting parameter redundancy. *Biometrika*, 84(1): 187–196, 1997.

[25] A. J. Chorin. The numerical solution of the navier-stokes equations for an incompressible fluid. *Bulletin of the American Mathematical Society*, 73(6):928–931, 1967.

[26] A. J. Chorin. Numerical solution of the navier-stokes equations. *Mathematics of computation*, 22(104):745–762, 1968.

[27] A. J. Chorin. On the convergence of discrete approximations to the navier-stokes equations. *Mathematics of computation*, 23(106):341–353, 1969.

[28] F. Clarelli, C. Di Russo, R. Natalini, and M. Ribot. Mathematical models for biofilms on the surface of monuments. In *Applied and industrial mathematics in Italy III*, pages 220–231. World Scientific, 2010.

[29] F. Clarelli, C. Di Russo, R. Natalini, and M. Ribot. A fluid dynamics model of the growth of phototrophic biofilms. *Journal of mathematical biology*, 66(7):1387–1408, 2013.

[30] K. Clement. Bariatric surgery, adipose tissue and gut microbiota. *International Journal of Obesity*, 35(3):S7–S15, 2011.

[31] N. Cogan and J. P. Keener. The role of the biofilm matrix in structural development. *Mathematical Medicine and Biology*, (21):147–166, 2004.

[32] D. J. Cole, B. J. Morgan, and D. Titterington. Determining the parametric structure of models. *Mathematical biosciences*, 228(1):16–30, 2010.

[33] J.-F. Collet and T. Goudon. On solutions of the Lifshitz-Slyozov model. *Nonlinearity*, 13 (4):1239, 2000.

[34] J. G. Conlon and A. Schlichting. A non-local problem for the Fokker-Planck equation related to the Becker-Döring model. *Discrete and Continuous Dynamical Systems*, 39(4): 1821–1889, 2019.

[35] J. Deschamps, E. Hingant, and R. Yvinec. Quasi steady state approximation of the small clusters in Becker–Döring equations leads to boundary conditions in the Lifshitz–Slyozov limit. *Communications in Mathematical Sciences*, 15(5):1353–1384, 2017.

[36] A. Divoux and K. Clement. Architecture and the extracellular matrix: The still unappreciated components of the adipose tissue. *Obesity reviews*, 12(5):e494–e503, 2011.

[37] C. L. Donohoe, J. Lysaght, J. O'Sullivan, and J. V. Reynolds. Emerging concepts linking obesity with the hallmarks of cancer. *Trends in Endocrinology & Metabolism*, 28(1):46–62, 2017.

[38] J. L. Doob. Markoff chains—denumerable case. *Transactions of the American Mathematical Society*, 58:455–473, 1945.

[39] M. Doumic, T. Goudon, and T. Lepoutre. Scaling limit of a discrete prion dynamics model. *Communications in Mathematical Sciences*, 7(4):839–865, 2009.

[40] R. Drolet, C. Richard, A. Sniderman, J. Mailloux, M. Fortier, C. Huot, C. Rhéaume, and A. Tchernof. Hypertrophy and hyperplasia of abdominal adipose tissues in women. *International journal of obesity*, 32(2):283–291, 2008.

[41] D. Dutykh, C. Acary-Robert, and D. Bresch. Mathematical modeling of powder-snow avalanche flows. *Studies in Applied Mathematics*, 127(1):38–66, 2011.

[42] L. Fang, F. Guo, L. Zhou, R. Stahl, and J. Grams. The cell size and distribution of adipocytes from subcutaneous and visceral fat is associated with type 2 diabetes mellitus in humans. *Adipocyte*, 4(4):273–279, 2015.

[43] X. Fernique. Intégrabilité des vecteurs gaussiens. *Comptes Rendus de l'Académie des sciences de Paris Serie A*, 270:1698–1699, 1970.

[44] A.-S. Giacobbi, L. Meyer, M. Ribot, R. Yvinec, H. Soula, and C. Audebert. Mathematical modeling of adipocyte size distributions: identifiability and parameter estimation from rat data. 2023.

[45] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The journal of physical chemistry A*, 104(9):1876–1889, 2000.

[46] J. Gilleron, T. Goudon, F. Lagoutière, H. Martin, B. Mauroy, P. Millet, M. Ribot, and C. Vaghi. Modeling and analysis of adipocytes dynamic with a differentiation process. *ESAIM: Proceedings and Surveys*, 67:210–241, 2020.

[47] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

[48] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[49] F. C. Goodrich. Nucleation rates and the kinetics of particle growth II. The birth and death process. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 277(1369):167–182, 1964.

[50] L. Gosse. A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms. *Computers & Mathematics with Applications*, 39 (9-10):135–159, 2000.

[51] T. Goudon and L. Monasse. Fokker–Planck Approach of Ostwald Ripening: Simulation of a Modified Lifshitz–Slyozov–Wagner System with a Diffusive Correction. *SIAM Journal on Scientific Computing*, 42(1):B157–B184, 2020.

[52] T. Goudon and A. Vasseur. On a model for mixture flows: Derivation, dissipation and stability properties. *Archive for Rational Mechanics and Analysis*, 220:1–35, 2016.

[53] J. M. Greenberg and A.-Y. LeRoux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM Journal on Numerical Analysis*, 33(1):1–16, 1996.

[54] M. L. Greer, L. Pujo-Menjouet, and G. F. Webb. A mathematical analysis of the dynamics of prion proliferation. *Journal of theoretical biology*, 242(3):598–606, 2006.

[55] S. Grosskinsky and W. Jatuviriyapornchai. Derivation of mean-field equations for stochastic particle systems. *Stochastic Processes and their Applications*, 129(4):1455–1475, 2019.

[56] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, 9(2):159–195, 2001.

[57] S. Hariz and J. F. Collet. A modified version of the Lifshitz-Slyozov model. *Applied mathematics letters*, 12(1):81–85, 1999.

[58] E. Hingant and R. Yvinec. Deterministic and stochastic Becker–Döring equations: Past and recent mathematical developments. In *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology*, pages 175–204. Springer, 2017.

[59] J. Hirsch and B. Batchelor. Adipose tissue cellularity in human obesity. *Clinics in endocrinology and metabolism*, 5(2):299–311, 1976.

[60] H. Hong, A. Ovchinnikov, G. Pogudin, and C. Yap. Sian: software for structural identifiability analysis of ode models. *Bioinformatics*, 35(16):2873–2874, 2019.

[61] H. Hong, A. Ovchinnikov, G. Pogudin, and C. Yap. Global identifiability of differential models. *Communications on Pure and Applied Mathematics*, 73(9):1831–1879, 2020.

[62] G. A. Jackson and A. B. Burd. Aggregation in the marine environment. *Environmental science & technology*, 32(19):2805–2814, 1998.

[63] M. Jacquier, F. Crauste, C. O. Soulage, and H. A. Soula. A predictive model of the dynamics of body weight and food intake in rats submitted to caloric restrictions. *PLoS One*, 9(6):e100073, 2014.

[64] I. Jeon. Existence of gelling solutions for coagulation-fragmentation equations. *Communications in mathematical physics*, 194:541–567, 1998.

[65] M. Jernås, J. Palming, K. Sjöholm, E. Jennische, P.-A. Svensson, B. G. Gabrielsson, M. Levin, A. Sjögren, M. Rudemo, T. C. Lystig, et al. Separation of human adipocytes by size: hypertrophic fat cells display distinct gene expression. *The FASEB journal*, 20(9): 1540–1542, 2006.

[66] J. Jo, O. Gavrilova, S. Pack, W. Jou, S. Mullen, A. E. Sumner, S. W. Cushman, and V. Periwal. Hypertrophy and/or hyperplasia: Dynamics of adipose tissue growth. *PLoS computational biology*, 5(3):e1000324, 2009.

[67] J. Jo, Z. Shreif, and V. Periwal. Quantitative dynamics of adipose cells. *Adipocyte*, 1(2): 80–88, 2012.

[68] J. Jo, Z. Shreif, J. R. Gaillard, M. Arroyo, S. W. Cushman, and V. Periwal. Mathematical models of adipose tissue dynamics. In *The Mechanobiology of Obesity and Related Diseases*, pages 11–34. Springer, 2013.

[69] A. V. Kazhikhov and S. Smagulov. The correctness of boundary-value problems in a diffusion model of an inhomogeneous liquid. In *Soviet Physics Doklady*, volume 22, page 249, 1977.

[70] F. P. Kelly. *Reversibility and Stochastic Networks*. Cambridge University Press, 2011.

[71] J. Kim, G. M. Saidel, and S. C. Kalhan. A computational model of adipose tissue metabolism: Evidence for intracellular compartmentation and differential activation of lipases. *Journal of theoretical biology*, 251(3):523–540, 2008.

[72] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV'-s, and the sample DF. I. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 32:111–131, 1975.

[73] T. G. Kurtz. Limit theorems and diffusion approximations for density dependent Markov chains. *Stochastic Systems: Modeling, Identification and Optimization, I*, pages 67–78, 1976.

[74] T. G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and their Applications*, 6(3):223–240, 1978.

[75] S. Labarthe, B. Polizzi, T. Phan, T. Goudon, M. Ribot, and B. Laroche. A mathematical model to investigate the key drivers of the biogeography of the colon microbiota. *Journal of Theoretical Biology*, 462:552–581, 2019.

[76] S. Laforest, A. Michaud, G. Paris, M. Pelletier, H. Vidal, A. Géloën, and A. Tchernof. Comparative analysis of three human adipocyte size measurement methods and their relevance for cardiometabolic risk. *Obesity*, 25(1):122–131, 2017.

[77] P. Laurençot. Weak solutions to the Lifshitz-Slyozov-Wagner equation. *Indiana University mathematics journal*, pages 1319–1346, 2001.

[78] P. Laurençot and S. Mischler. From the becker–döring to the lifshitz–slyozov–wagner equations. *Journal of statistical physics*, 106(5-6):957–991, 2002.

[79] P. Laurençot and C. Walker. Well-posedness for a model of prion proliferation dynamics. *Journal of Evolution Equations*, 7(2):241–264, 2007.

[80] K. Y. Lee, Q. Luong, R. Sharma, J. M. Dreyfuss, S. Ussar, and C. R. Kahn. Developmental and functional heterogeneity of white adipocytes within a single fat depot. *The EMBO Journal*, 38(3):e99291, 2019.

[81] P. Lévy. *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, 1954.

[82] Y. Li, V. Periwal, S. W. Cushman, and K. G. Stenkula. Adipose cell hypertrophy precedes the appearance of small adipocytes by 3 days in c57bl/6 mouse upon changing to a high fat diet. *Adipocyte*, 5(1):81–87, 2016.

[83] M. E. Lidell, M. J. Betz, O. D. Leinhard, M. Heglind, L. Elander, M. Slawik, T. Mussack, D. Nilsson, T. Romu, P. Nuutila, et al. Evidence for two types of brown adipose tissue in humans. *Nature medicine*, 19(5):631–634, 2013.

[84] I. M. Lifshitz and V. V. Slyozov. The kinetics of precipitation from supersaturated solid solutions. *Journal of physics and chemistry of solids*, 19(1-2):35–50, 1961.

[85] M. Lönn, K. Mehlig, C. Bengtsson, and L. Lissner. Adipocyte size predicts incidence of type 2 diabetes in women. *The FASEB journal*, 24(1):326–331, 2010.

[86] J. MacKellar, S. W. Cushman, and V. Periwal. Waves of adipose tissue growth in the genetically obese zucker fatty rat. *PloS one*, 5(1):e8197, 2010.

[87] H. P. McKean Jr. A. skorohod's stochastic integral equation for a reflecting barrier diffusion. *Journal of Mathematics of Kyoto University*, 3(1):85–88, 1963.

[88] T. McLaughlin, A. Sherman, P. Tsao, O. Gonzalez, G. Yee, C. Lamendola, G. Reaven, and S. Cushman. Enhanced proportion of small adipose cells in insulin-resistant vs insulin-sensitive obese individuals implicates impaired adipogenesis. *Diabetologia*, 50:1707–1715, 2007.

[89] D. E. McNay and J. R. Speakman. High fat diet causes rebound weight gain. *Molecular metabolism*, 2(2):103–108, 2013.

[90] L. Meyer, M. Ribot, and R. Yvinec. A lifschitz-slyozov type model for adipocyte size dynamics: limit from becker-döring system and numerical simulation. 2023.

[91] L. K. Meyer, T. P. Ciaraldi, R. R. Henry, A. C. Wittgrove, and S. A. Phillips. Adipose tissue depot and cell size dependency of adiponectin synthesis and secretion in human obesity. *Adipocyte*, 2(4):217–226, 2013.

[92] B. Niethammer. A scaling limit of the Becker-Döring equations in the regime of small excess density. *Journal of Nonlinear Science*, 14(5):453–468, 2004.

[93] B. Niethammer. Macroscopic limits of the Becker-Döring equations. *Communications in Mathematical Sciences*, 2(Supplemental Issue):85–92, 2004.

[94] B. Niethammer and R. L. Pego. Non-self-similar behavior in the lsw theory of ostwald ripening. *Journal of statistical physics*, 95:867–902, 1999.

[95] B. Niethammer and R. L. Pego. On the initial-value problem in the Lifshitz–Slyozov–Wagner theory of Ostwald ripening. *SIAM Journal on Mathematical Analysis*, 31(3):467–485, 2000.

[96] A. B. Owen. On dropping the first sobol'point. In *International conference on Monte Carlo and quasi-Monte Carlo methods in scientific computing*, pages 71–86. Springer, 2020.

[97] O. Penrose. The Becker-Döring equations at large times and their connection with the LSW theory of coarsening. *Journal of statistical physics*, 89:305–320, 1997.

[98] O. Penrose, J. L. Lebowitz, J. Marro, M. H. Kalos, and A. Sur. Growth of clusters in a first-order phase transition. *Journal of Statistical Physics*, 19:243–267, 1978.

[99] B. W. Peterson, Y. He, Y. Ren, A. Zerdoum, M. R. Libera, P. K. Sharma, A.-J. van Winkelhoff, D. Neut, P. Stoodley, H. C. van der Mei, and H. J. Busscher. Viscoelasticity of biofilms and their recalcitrance to mechanical and chemical challenges. *FEMS microbiology reviews*, 39(2):234–245, 03 2015.

[100] D. Peurichard, F. Delebecque, A. Lorsignol, C. Barreau, J. Rouquette, X. Descombes, L. Casteilla, and P. Degond. Simple mechanical cues could explain adipose tissue morphology. *Journal of Theoretical Biology*, 429:61–81, Sept. 2017.

[101] D. Peurichard, M. Ousset, J. Paupert, B. Aymard, A. Lorsignol, L. Casteilla, and P. Degond. Extra-cellular matrix rigidity may dictate the fate of injury outcome. *Journal of theoretical biology*, 469:127–136, 2019.

[102] B. Polizzi, O. Bernard, and M. Ribot. A time-space model for the growth of microalgae biofilms for biofuel production. *Journal of Theoretical Biology*, 432:55 – 79, 2017. ISSN 0022-5193.

[103] B. Polizzi, A. Fanesi, F. Lopes, M. Ribot, and O. Bernard. Understanding photosynthetic biofilm productivity and structure through 2d simulation. *PLoS Computational Biology*, 18(4):e1009904, 2022.

[104] V. Prana, P. Tieri, M. C. Palumbo, E. Mancini, and F. Castiglione. Modeling the effect of high calorie diet on the interplay between adipose tissue, inflammation, and diabetes. *Computational and mathematical methods in medicine*, 2019, 2019.

[105] L. Preziosi and A. Tosin. Multiphase modelling of tumour growth and extracellular matrix interaction: mathematical tools and applications. *Journal of mathematical biology*, 58: 625–656, 2009.

[106] S. Prigent, A. Ballesta, F. Charles, N. Lenuzza, P. Gabriel, L. M. Tine, H. Rezaei, and M. Doumic. An efficient kinetic model for assemblies of amyloid fibrils and its application to polyglutamine aggregation. *PloS one*, 7(11):e43273, 2012.

[107] K. Rajagopal and L. Tao. *Mechanics of mixtures*. Series on Advances in Mathematics for Applied Sciences. World Scientific, Singapore, 1995.

[108] A. Sakers, M. K. De Siqueira, P. Seale, and C. J. Villanueva. Adipose-tissue plasticity in health and disease. *Cell*, 185(3):419–446, 2022.

[109] A. Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer physics communications*, 145(2):280–297, 2002.

[110] A. Schlichting. Macroscopic limit of the Becker–Döring equation via gradient flows. *ESAIM: Control, Optimisation and Calculus of Variations*, 25:22, 2019.

[111] A. Schlichting. The exchange-driven growth model: basic properties and longtime behavior. *Journal of Nonlinear Science*, 30(3):793–830, 2020.

[112] P. J. Schnurr and D. G. Allen. Factors affecting algae biofilm growth and lipid production: A review. *Renewable and Sustainable Energy Reviews*, 52:418–429, 2015.

[113] P. Secchi. On the motion of viscous fluids in the presence of diffusion. *SIAM Journal on Mathematical Analysis*, 19(1):22–31, 1988.

[114] M. Shao, Q. A. Wang, A. Song, L. Vishvanath, N. C. Busbuso, P. E. Scherer, and R. K. Gupta. Cellular origins of beige fat cells revisited. *Diabetes*, 68(10):1874–1885, 2019.

[115] G. Simonett and C. Walker. On the solvability of a mathematical model for prion proliferation. *Journal of mathematical analysis and applications*, 324(1):580–603, 2006.

[116] T. Skurk, C. Alberti-Huber, C. Herder, and H. Hauner. Relationship between adipocyte size and adipokine expression and secretion. *The Journal of Clinical Endocrinology & Metabolism*, 92(3):1023–1033, 2007.

[117] I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation*, 55(1-3):271–280, 2001.

[118] H. A. Soula, H. Julienne, C. O. Soulage, and A. Geloen. Modelling adipocytes size distribution. *Journal of theoretical biology*, 332:89–95, 2013.

[119] H. A. Soula, A. Geloen, and C. O. Soulage. Model of adipose tissue cellularity dynamics during food restriction. *Journal of Theoretical Biology*, 364:189–196, 2015.

[120] G. Stoltz and P. Terrier. A mathematical justification of the finite time approximation of becker-döring equations by a Fokker-Planck dynamics. *arXiv preprint arXiv:1810.01462*, 2018.

[121] R. Temam. Sur l'approximation de la solution des équations de navier-stokes par la méthode des pas fractionnaires (i). *Archive for Rational Mechanics and Analysis*, 32:135–153, 1969.

[122] M.-N. Thai. Birth and death process in mean field type interaction. *arXiv preprint arXiv:1510.03238*, 2015.

[123] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf. Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6(31):187–202, 2009.

[124] C. Truesdell. Sulle basi della termomeccanica. I, II. *Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. (8)*, 1957.

[125] C. Truesdell. *Rational thermodynamics*. McGraw-Hill Book Co., New York-London-Sydney, 1969. A course of lectures on selected topics, With an appendix on the symmetry of the heat-conduction tensor by C. C. Wang.

[126] C. Truesdell and R. Toupin. The classical field theories. In *Handbuch der Physik, Bd. III/1*. Springer, Berlin, 1960. With an appendix on tensor fields by J. L. Ericksen.

[127] V. Van Harmelen, T. Skurk, K. Röhrig, Y. Lee, M. Halbleib, I. Aprath-Husmann, and H. Hauner. Effect of bmi and age on adipose tissue cellularity and differentiation capacity in women. *International journal of obesity*, 27(8):889–895, 2003.

[128] O. Varlamov, R. Somwar, A. Cornea, P. Kievit, K. L. Grove, and C. T. Roberts Jr. Single-cell analysis of insulin-regulated fatty acid uptake in adipocytes. *American Journal of Physiology-Endocrinology and Metabolism*, 299(3):E486–E496, 2010.

[129] A. Vasseur, F. Poupaud, J.-F. Collet, and T. Goudon. The Becker–Döring System and Its Lifshitz–Slyozov Limit. *SIAM Journal on Applied Mathematics*, 62(5):1488–1500, 2002.

[130] J. J. L. Velázquez. The Becker–Döring equations and the Lifshitz–Slyozov theory of coarsening. *Journal of statistical physics*, 92:195–236, 1998.

[131] C. Wagner. Theorie der alterung von niederschlägen durch umlösen (Ostwald-reifung). *Zeitschrift für Elektrochemie, Berichte der Bunsengesellschaft für physikalische Chemie*, 65(7-8):581–591, 1961.

[132] O. Wurl, E. Wurl, L. Miller, K. Johnson, and S. Vagle. Formation and global distribution of sea-surface microlayers. *Biogeosciences*, 8(1):121–135, 2011.

Léo MEYER

Modélisation et analyse des modèles pour la distribution en taille des adipocytes

Résumé :
Le tissu adipeux est un tissu conjonctif responsable du stockage de l'énergie sous forme de gouttelettes lipidiques à l'intérieur des cellules adipeuses. Ces cellules, également appelées adipocytes, changent de taille de façon dynamique en fonction de deux processus : le stockage des acides gras du sang à l'intérieur de la cellule, appelé lipogenèse, et la libération active des acides gras dans le sang pour une utilisation énergétique, appelée lipolyse. Cette thèse porte sur la modélisation de la dynamique de ces changements de taille pour une population de cellules adipeuses.
Une caractéristique frappante des données de distribution de taille des adipocytes, recueillies sur diverses espèces animales, est leur bimodalité bien conservée : les distributions ont deux maxima locaux. Nous partons d'une loi de conservation scalaire non linéaire qui décrit la dynamique des échanges de masse entre les cellules (macroparticules) et les lipides (monomères), inspirée du système de Lifschitz-Slyozov dans la théorie du grossissement. Ce premier modèle est exclu pour expliquer les données, car son état stationnaire ne peut pas être une distribution bimodale lisse. Deux extensions diffusives sont alors envisagées. La première est heuristique et consiste à ajouter un opérateur de diffusion constant. La seconde, appelée modèle de Lifshitz-Slyozov de second ordre, implique un opérateur de diffusion dépendant du temps et de la taille qui est motivé par la version discrète du modèle de Lifshitz-Slyozov, à savoir le système ODE infini donné par les équations de Becker-Döring. Nous prouvons un nouveau résultat de convergence des équations de Becker-Döring vers les équations de Lifshitz-Slyozov qui soutient l'extension du second ordre. Nous introduisons également deux modèles stochastiques non linéaires équivalents respectivement aux équations de Becker-Döring et aux équations de Lifshitz-Slyozov diffusives du second ordre, qui nous permettent de montrer un résultat de convergence du premier modèle vers le second. La simulation numérique de tous ces modèles est réalisée grâce à des méthodes numériques adaptées, notamment un schéma 'well-balanced' et l'algorithme de Gillespie. Les deux extensions diffusives du modèle de Lifshitz-Slyozov sont capables de retrouver la bimodalité à partir de leurs solutions stationnaires, et dans les deux cas, en effectuant une estimation des paramètres sur des données provenant de rats, nous retrouvons les paramètres impliqués dans la lipogenèse et la lipolyse avec des ordres de grandeur similaires.
Cette thèse contient également un travail réalisé lors de l'école d'été CEMRACS 2022. Nous avons utilisé un schéma numérique issu de la cinétique des écoulements multifluides pour étudier la croissance d'un biofilm. Le biofilm est vu comme trois composants interagissant ensemble : les cellules, la matrice extracellulaire, et le milieu environnant. Chaque composant est décrit par sa fraction volumique et sa vitesse, ce qui conduit à la contrainte naturelle que la somme des trois fractions volumiques soit égale à un. Ce nouveau schéma permet de valider cette contrainte sur le système au niveau numérique. Nous obtenons le phénomène d'onde progressive attendu et ajoutons divers substrats qui jouent potentiellement un rôle dans la croissance du biofilm.
Mots clefs : modélisation mathématique, équations aux dérivées partielles, modèles stochastiques, estimation de paramétres

Modeling and model analysis for the size distribution of adipocytes

Abstract :
The adipose tissue is a connective tissue responsible for the storage of energy in the form of lipid droplets inside the adipose cells. These cells, also called adipocytes, change size dynamically depending on two processes : the storageof fatty acids from the blood inside the cell, called lipogenesis, and the active release of the fatty acids into the blood for energy usage, called lipolysis. This thesis deals with the modeling of the dynamics of these size changes for a population of adipose cells.
A striking feature of the adipocyte size distribution data, collected on various animal species, is their well conserved bimodality : the distributions always have two local maxima. We start from a nonlinear scalar conservation law that describes the dynamics of mass exchange between cells (macro–particles) and lipids (monomers), inspired by the Lifschitz–Slyozov system in the theory of coarsening. This first model is ruled-out to explain the data, because its stationary state cannot be a smooth bimodal distribution. Two diffusive extensions are then considered. The first one is heuristic and consider adding a constant diffusion operator. The second one, called second-order Lifshitz-Slyozov model, involves a time and size-dependent diffusion operator that is motivated from the discrete version of the Lifshitz-Slyozov model, namely the infinite ODE system given by the Becker-Döring equations. We prove a new convergence result from the Becker-Döring equations to the Lifshitz-Slyozov equations that supports the second-order extension. We also introduce two nonlinear stochastic models equivalent to respectively the Becker-Döring equations and the second-order diffusive Lifshitz-Slyozov equation, that allow us to show a convergence result from the former model to the latter model. Numerical simulation of all these models is performed thanks to adapted numerical methods, including a well-balanced scheme and Gillespie algorithm. Both diffusive extensions of the Lifshitz-Slyozov model are able to reproduce bimodality from their stationary solutions, and in both cases, performing parameters estimation on data from rats, we recover parameters involved in lipogenesis and lipolysis with similar orders of magnitudes.
This thesis also contains a work done during the summer school CEMRACS 2022. We used a numerical scheme from multifluid flow kinetics to study the growth of a biofilm. The biofilm is seen as three components ineteracting together : the cells, the extra-cellular matrix, and the surrounding medium. Each components is described by its volumic fraction and its velocity, which leads to the natural constraint that the three volume fraction sum to one. This new scheme allows the validation of this constraint on the system at the numerical level. We obtain the expected traveling wave pattern and add various substrate that potentially play a role in the growth of the biofilm.
Keywords : mathematical modeling, partial differential equations, stochastic models, parameters estimation

Institut Denis Poisson
rue de Chartres B.P. 6759
45067 Orléans CEDEX 2