Galilée Doctoral School

REACTION-DIFFUSION EQUATIONS AND APPLICATIONS TO BIOLOGICAL CONTROL OF DENGUE AND INFLAMMATION

Ana Isis Toledo Marrero

A thesis submitted for the degree of Doctor of Mathematics

January 15, 2021

Jury:

Elisabeth Logack	Reporter
Lionel Roques	Reporter
Nicolas Vauchelet	Examiner
Danielle Hilhorst	Examiner
Matthieu Alfaro	Examiner
Eric Ogier-Denis	Examiner
Hatem Zaag	Advisor
Grégoire Nadin	Advisor

Abstract

This thesis is devoted to the study of two problems arising from biology and medicine. The first model is motivated by a new technique to eradicate mosquito-borne diseases such as the dengue virus. A certain number of mosquitoes, inoculated with a bacterium inhibiting mosquito-borne disease transmission to humans are released in the environment. The evolution of this subset of the mosquito population can be described by means of a reaction-diffusion equation. The problem we address here concerns the maximization of the total number of carrying individuals after a certain prescribed time, which is a quantity depending on the solution of the equation. We maximize this quantity with respect to the initial datum under certain size constraints. Existence and regularity results as well as a partial characterization of optimizers are stated by means of the study of the first and second order optimality conditions. A numerical algorithm, inspired by the classical ascent of gradient and taking advantage of the theoretical results we obtain here is described, allowing a numerical approximation of local optimizers.

On the other hand, a model describing the dynamics of immune cells and pathogenic bacteria in the gut tissues is introduced. More precisely, a reaction-diffusion system is considered with the purpose of explaining the patchy inflammatory patterns observed in patients suffering from Crohn's disease. We perform a stability analysis enabling us to identify conditions driving to the occurrence of Turing instabilities. Such instabilities could be interpreted as the patchy inflammatory patterns. Realistic parameter values for which this phenomenon arises are either computed or retrieved from the existent literature and numerical simulations are performed as well.

Keywords: Reaction-diffusion equation, control, conservation biology, optimization, Turing pattern, activator-inhibitor, inflammatory diseases.

Résumé

Cette thèse est consacrée à l'étude de deux problèmes issus de la biologie et de la médecine. Le premier est motivé par une technique de contrôle biologique pour l'éradication de l'épidémie de la dengue transmise par des moustiques. Une bactérie, dont les effets chez les moustiques inhibent la transmission de ce virus, est inoculée à un certain nombre des moustiques qui sont ensuite relâchés dans l'environnement. L'évolution de cette partie de la population porteuse de la bactérie peut être décrite par une équation de réaction-diffusion. On s'intéresse particulièrement à maximiser la population totale de moustiques porteurs de cette bactérie après un certain temps. Il s'agit d'une quantité dépendant de la solution de l'équation, que l'on maximise par rapport à la donnée initiale sous certaines contraintes. L'existence et la régularité des solutions à ce problème d'optimisation, ainsi qu'une caractérisation partielle de la donnée initiale optimale sont établies grâce à l'étude des conditions d'optimalité de premier et deuxième ordre. Un algorithme numérique, inspiré de la méthode classique de montée de gradient et tirant parti des conditions d'optimalité est décrit, permettant une approximation numérique des maxima locaux de ce problème.

D'autre part, un modèle décrivant la dynamique des cellules immunitaires et des bactéries pathogènes dans les tissus de l'intestin est introduit. Un système de réaction-diffusion est considéré, l'objectif étant d'expliquer les motif inflammatoires inégaux observés chez les patients souffrant de la maladie de Crohn. Une analyse de stabilité est réalisée et des conditions menant à l'apparition d'instabilités de Turing sont énoncées; ces instabilités pouvant être interprétées comme les patterns inflammatoires. Des valeurs réalistes des paramètres, pour lesquels ce phénomène se produit, sont calculées ou extraites de la littérature existante, des simulations numériques sont également réalisées.

Mots-clés: Équations de réaction-diffusion, contrôle, biologie de la conservation, optimisation, Turing patterns, activateur-inhibiteur, maladies inflammatoires.

Remerciements :

Ma première pensée est pour Grégoire Nadin et Hatem Zaag. Non seulement ils ont été mes référents les plus importants du point de vue scientifique, mais ils sont devenus aussi des modèles à suivre dans la vie. Leurs encadrement, soutien et encouragements constants m'ont permis de parcourir tout ce chemin, parfois difficile, beau surtout. Je leur dois tout ce que j'ai pu réussir pendant ces trois ans et demi, de cela je leur suis infiniment reconnaissante.

Un grand merci également à Eric Ogier-Denis, qui a su raviver mon amour pour la biologie et qui m'a montré la beauté de la recherche interdisciplinaire.

Merci à Lionel Roques et Elizabeth Logak d'avoir accepté d'être rapporteur·euse de cette thèse ; tous vos commentaires, observations et conseils m'ont aidée à améliorer ce manuscrit et à identifier de nouvelles directions de recherche, permettant d'élargir les applications de ces travaux.

Merci à Danielle Hilhorst, Nicolas Vauchelet et Matthieu Alfaro d'avoir consenti à faire partie de mon jury de thèse. Les travaux autour des applications des mathématiques à la biologie que vous avez présentés lors des rencontres à EDP-Normandie, à l'école CIMPA à Santiago de Cuba et à Nonlinear PDE's in Porquerolles, m'ont beaucoup inspirée ; c'est un véritable honneur de vous compter parmi les membres de mon jury.

Merci à Stéphane Mischler et à Otared Kavian, pour leur soutien dès avant mon arrivée en France. Au-delà de vos précieux conseils, enseignements, vous avez toujours été là quand j'en avais besoin, m'avez toujours fait ressentir la chaleur cubaine, si loin de Cuba. Un grand merci aussi à toute la communauté de la FSMP pour leur engagement et leur contribution au développement scientifique de l'université cubaine et des cubains.

Siento un gran agradecimiento también por mis profes de Cuba, a quienes debo el amor que siento por las matemáticas y la valentía de aventurarme a esta nueva etapa de mi vida. Especialmente todo mi agradecimiento a Sofía, Jorge y Wilfredo, quienes me enseñaron lo que es ser un buen investigador, un buen profesor, un buen amigo.

Merci aux membres du LAGA et du LJLL, les chercheur·euse·s comme le personnel administratif, qui m'ont si bien accueillie. Merci aussi pour ces discussions riches, souvent autour d'un café ou un thé, qui m'ont tant appris de la vie et la culture françaises.

Quand on est si loin de la famille, les ami·e·s deviennent tout notre monde, je mesure la chance d'avoir rencontré tant de personnes magnifiques en France, qui m'ont ouvert leur portes et leur cœur.

Merci à mes copain·e·s de bureau au LJLL, avec qui j'ai tant partagé : Martin, Hongjun, David, Cécile et Lise, qui en fait presque partie. Merci pour les conversations, le soutien, les conseils et la compréhension que j'ai toujours pu retrouver auprès de vous. Un plus grand merci encore à Idriss, avec qui j'ai eu le plaisir de faire des maths, pour sa patience et sa gentillesse.

Merci les doctorant·e·s du LJLL, certain·e·s déjà docteur·e·s, avec qui j'ai pu partager tant des beaux moments. Merci Camille pour mon premier déjeuner à l'Ardoise et à Antoine et Lydie pour leur compagnie. Merci à la team liblib, la team mange-tôt et aussi les mange-tard ; aux ami·e·s de la première heure : Anouk, Christophe, Olivier, Gontran, Gaby, Amaury, Fede, Ludo, Alex Rege, Alex Delyon, Katia, Julia, Nico, Marc, Valentin, Leo, Idriss, Cécile, Lucile, Sophie (certain·e·s me manquent déjà beaucoup) et ceux·celles rencontré·e·s ensuite : Allen (le meilleur coloc), Jules, Elise, Fatima, Nicolas, Lise, David, Cécile, Emma, Giorgia, Remi, Antoine, Matthieu, Maria, Noemi, Jesus, Ramón, Nicolás, Emilio, Agustín, Eugenio, avec qui j'espère partager encore de bons moments au labo.

A mami y papi.

"Yo dejo mi palabra en el aire sin llaves y sin velos.
Porque ella no es un arca de codicia,
ni una mujer coqueta
que trata de parecer más hermosa de lo que es.
Yo dejo mi palabra en el aire, para que todos la vean,
la palpen, la estrujen o la expriman.
Nada hay en ella que no sea yo misma;
pero en ceñirla como cilicio y no como manto
pudiera estar toda mi ciencia."

Dulce Maria Loynaz

Contents

> **Chapter 1**
>
> **Mathematical modelling and vector control techniques**

> **Chapter 2**
>
> **On the maximization problem for solutions of reaction-diffusion equations with respect to their initial data**

Chapter 3
Second order optimality conditions for optimization with respect to the initial data in reaction-diffusion equations

Partie II Crohn's disease and inflammatory patterns   56

Chapter 4
A Turing mechanism in order to explain the patchy nature of Crohn's disease

List of Figures

General introduction

> *"The isolated man does not develop any intellectual power. It is necessary for him to be immersed in an environment of other men, whose techniques he absorbs during the first twenty years of his life. He may then perhaps do a little research of his own and make a very few discoveries which are passed on to other men. From this point of view the search for new techniques must be regarded as carried out by the human community as a whole, rather than by individuals."*
>
> Alan Turing

This work has been motivated by two problems arising from biology. The first, discussed in part I, concerns an emergent technique of biological control of mosquito population aiming to stop the propagation of Dengue virus. The second, discussed in part II, suggests a mechanism for some inflammatory bowel disease known as Crohn's disease. The model has been proposed by biologists and medical doctors working in the area. Though the two topics seem to be unrelated, the mathematics behind our approach to these subjects build on the same tool: population dynamics and reaction-diffusion equations, which will make the unifying thread in this manuscript.

Before proceeding to the presentation of the main problems and results in sections 2 and 3 of this general introduction, we reserve section 1 for a review of the basic mathematical notions on which the results of this thesis are based. In section 4, an outline of the manuscript is briefly presented.

1   Basic mathematical notions

1.1   Reaction diffusion equations in population dynamics

A classical diffusion equation is a parabolic partial differential equation which has a general form obtained via conservation of mass and Fick's law:

$$\partial_t u = \sigma \Delta u; \quad t \geq 0, \quad x \in \Omega \subseteq \mathbb{R}^n, \quad u(t,x) \in \mathbb{R}.$$

Here $u(t,x)$ is a state variable and usually describes the density of the diffusive material evolving through time due to random movements and collisions of particles, and σ is the dispersal rate. As seen by Kolmogorov, Petrovskii and Piskunov in [46], assuming that diffusion is accompanied by an increase in the amount of substance at a rate that depends on the density, one obtains a semi-linear parabolic equation better known as the reaction-diffusion equation,

$$\partial_t u = \sigma \Delta u + f(u); \quad t \geq 0, \quad x \in \Omega \subseteq \mathbb{R}^n, \quad u(t,x) \in \mathbb{R}. \tag{1}$$

Seeking the well-posedness of eq. (1), an initial condition must be prescribed $u(0, x) = u_0(x)$, as well as boundary conditions if Ω is a bounded domain. In this thesis only the case of homogeneous Neumann boundary conditions will be treated, i.e. $\frac{\partial}{\partial \nu} u(t, x) = 0$ on $\partial \Omega$, where ν is the outward normal.

Provided the nonlinearity f is Lipschitz continuous, classical Banach fixed point arguments can be used to prove existence and uniqueness of a local in time weak solution of eq. (1) in the case $\Omega = \mathbb{R}^n$, see for instance [28]. Such a solution could be extended in time defining a maximal solution either globally or just in a bounded time domain. Indeed, in certain cases it is possible for issues of non-existence to arise, due to, for instance, blow-up phenomena in finite time. Methods of subsolutions and supersolutions can be also exploited to show existence for certain boundary-value problems.

A solution of eq. (1) which is time independent is called an equilibrium or stationary state. These states are important since they can have a large influence on the asymptotic behavior of solutions of semilinear parabolic equations with Neumann boundary conditions. Indeed, under certain conditions, provided that $f = f(u)$ is Lipschitz, σ is positive and Ω is bounded, it can be proved that solutions $u(t, x)$ tend to become homogeneous at an exponential rate, [75].

Reaction-diffusion equations turn out to be well-adapted to describe population dynamics in ecology. In fact, independently of the works of Kolmogorov *et al.*, Fisher introduced and studied the same type of model motivated by the study of the spatial spread of an advantageous gene as a wave, [31]. In what is known today as the Fisher-KPP equation, the reaction term is a logistic growth term of the form

$$f(u) = ru(1 - u),$$

which only depends on the population density u and the constant parameter r associated with the birth rate at low densities. This term implies that the population will grow until it reaches a limit $u = 1$ which is usually related to the carrying capacity of the environment.

The Fisher-KPP equation allows only two nonnegative constant equilibria, $u = 0$ which is unconditionally unstable (meaning that any homogeneous small initial perturbation will grow exponentially), and $u = 1$ which is unconditionally stable (meaning that for any homogeneous small initial perturbation close to 1, the solution will relax exponentially to 1), [75]. When such phenomena occur, the equations are said to be monostable; The Fisher-KPP equation is the most common monostable example, see fig. 1.

Many other nonlinearities might be considered in eq. (1). In particular bistable nonlinearities are of great interest for the results that will be stated in this manuscript. A function $f = f(u)$ is said to be bistable if there exists $\rho \in (0, 1)$ such that $f(0) = f(\rho) = f(1) = 0$, $f < 0$ on $(0, \rho)$, $f > 0$ on $(\rho, 1)$ and $f'(1) < 0$. The classical bistable term can be written as

$$f(u) = ru(1 - u)(u - \rho). \tag{2}$$

Here the equilibria $u = 0$ and $u = 1$ are both stable, while $u = \rho$ is unstable. Usually, it is assumed that the equilibrium $u = 1$ is more energetically favorable than the $u = 0$ equilibrium by assuming

$$\int_0^1 f(s) \, ds > 0.$$

In (2), this corresponds to $\rho < \frac{1}{2}$, see fig. 1.

Applied to population dynamic models, bistability might simulate invasion ($u(t, x) \to 1$ as $t \to \infty$) or extinction process ($u(t, x) \to 0$ as $t \to \infty$). The term $(u - \rho)$ implies that for low densities ($u < \rho$) population tends to decrease due, for instance, to the difficulty of finding a mate for reproduction or to the weak defense against predators. This is known as Allee effect.

Thanks to a classical comparison principle, provided that the initial data satisfies $0 \le u_0(x) \le 1$, it holds that the solution at any time t exists and satisfies $0 \le u(t, x) \le 1$. Moreover, for bistable equations there occurs a sharp transition phenomenon between propagation or extinction [23, 94], see fig. 2. More precisely,
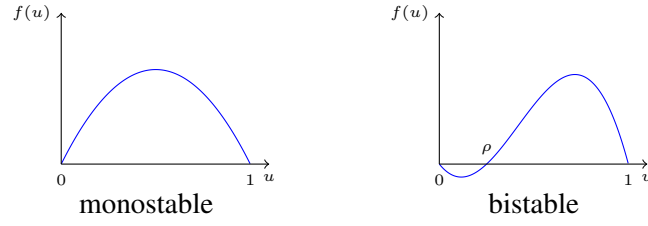
Figure 1: Monostable and bistable nonlinearities

Theorem 1 *[94] Considering initial data of the form $u_0 = \mathbb{1}_{[-L,L]}$, $L > 0$, there is $L^* > 0$ such that one of the three following possibilities holds:*

   (i) if $L < L^*$, then $u \to 0$ uniformly on $\mathbb{R}$ as $t \to \infty$;

   (ii) if $L > L^*$, then $u \to 1$ uniformly on compact subsets of $\mathbb{R}$ as $t \to \infty$;

   (iii) if $L = L^*$, then there exists a positive symmetric, non-increasing for $x > 0$ equilibrium $U = U(x)$ such that

      (a) $u(t, x) \to U(x)$ uniformly in $\mathbb{R}$ as $t \to \infty$,

      (b) $-\sigma U'' = f(U)$,

      (c) $U(0) = \rho_1 = \sup \left\{ \rho' \in (0, 1) : \int_0^{\rho'} f(s) \, ds \leq 0 \right\}, U'(0) = 0.
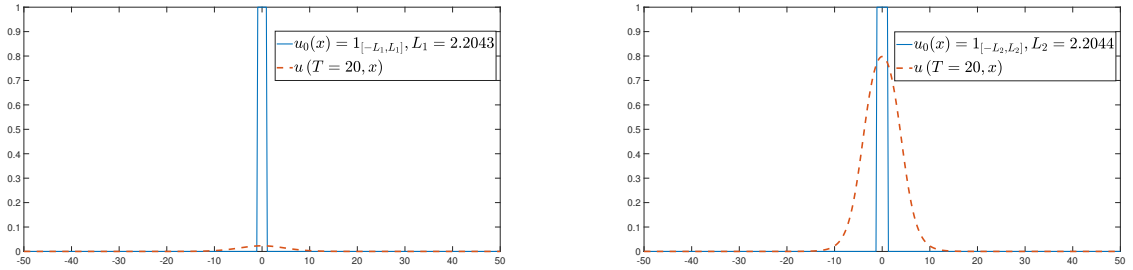
Figure 2: Example of the sharp transition phenomena enunciated in Theorem 1. Here clearly holds $L_1 < L^* < L_2$

The effect of $u_0(x)$ on the asymptotic behavior of solutions of eq. (1) will be the main interest in part I where more precise regularity estimates and properties of the solution u of the eq. (1) with bistable nonlinearity will be discussed.

1.2 Patterns in reaction-diffusion equations

Let us now go back to the case where f is not necessarily bistable, and let us also consider a general reaction-diffusion system of equations. As was mentioned above, the Laplace operator Δu stands in eq. (1) for a model of the random (Brownian) motion of individuals or particles. Intuitively, when a state is linearly stable without diffusion, one might think that adding a spreading term will do nothing but stress the stability of the system. Nevertheless, disparity in the diffusivity rates of the entities interacting in the equation might lead to unexpected instabilities. This mechanism leading to the observation of steady

states showing heterogeneous spatial patterns was first described in the context of chemical reactions by A. Turing in 1952 [87]. Ever since, it has been widely studied, see for instance [75], and now provides mathematical explanations for several patterns found in nature, see fig. 3.

Figure 3: Patterns in nature

In chapter 4 more insights on Turing Patterns will be discussed in the context of the study of pattern formation in gut tissues due to acute inflammation.

1.3 Basic notions on optimization

Since an optimization problem lays at the heart of this thesis, it is pertinent to recall some basic notions in optimization theory. For simplicity, and taking into account the problem that is going to be studied in chapter 2 of this thesis, all the classical results and notations are going to refer to a maximization problem, even when in the literature the convention is to state the results in the minimization case.

Let $(B, \|\cdot\|)$ be a Banach and consider the subset $A \subset B$ defined by the constraints of the problem, usually called the admissible set, in which we will search for the optimum element. Let $\mathcal{J} : A \to \mathbb{R}$ be the objective function, i.e. the quantity to be maximized. The maximization problem then reads

$$\sup_{a \in A \subset B} \mathcal{J}(a). \tag{3}$$

An element $\overline{a} \in B$ is called a **local maximum** of $\mathcal{J}$ over A if and only if

$$\overline{a} \in A \quad \text{and} \quad \exists \delta > 0, \quad \text{such that} \quad \forall a \in A, \quad \|\overline{a} - a\| \leq \delta \implies \mathcal{J}(\overline{a}) \geq \mathcal{J}(a).$$

The element $\overline{a}$ is called a **global maximum** if and only if

$$\overline{a} \in A \quad \text{and} \quad \forall a \in A, \quad \mathcal{J}(\overline{a}) \geq \mathcal{J}(a).$$

By definition of the supremum notion, one can always extract a sequence of $(a^n)_{n \in \mathbb{N}}$ such that

$$a^n \in A \quad \forall n \quad \text{and} \quad \lim_{n \to \infty} \mathcal{J}(a^n) = \sup_{a \in A \subset B} \mathcal{J}(a).$$

$(a^n)_{n \in \mathbb{N}}$ is called a **maximizing sequence** of $\mathcal{J}$ over A. The existence of a subsequence of $(a^n)_{n \in \mathbb{N}}$ converging to an element of K, in combination with the upper semi-continuity of $\mathcal{J}$, guarantees the existence of an optimal element.

In order to establish the classical optimality conditions for the previous maximization problem, let us recall the notion of Gâteaux differentiability. The Gâteaux derivative $\langle \nabla \mathcal{J}(a), \phi \rangle$ of $\mathcal{J}$ at $a \in A$ in the direction $\phi \in B$ is defined as

$$\langle \nabla \mathcal{J}(a), \phi \rangle := \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \frac{\mathcal{J}(a + \varepsilon \phi) - \mathcal{J}(a)}{\varepsilon} \tag{4}$$

If the limit exists for all $\phi \in B$, then it is said that $\mathcal{J}$ is **Gâteaux differentiable** at a. Moreover, if $x \mapsto \nabla \mathcal{J}(x)$ is continuous at $x = a$, then $\mathcal{J}$ is differentiable at a in the sense of Frechet. If the admissible set A is a convex, then Euler's inequality holds, i.e.

Theorem 2 *[20] Let A be a convex and $\mathcal{J}$ differentiable at $\overline{a} \in A$. If $\overline{a}$ is a local maximum of $\mathcal{J}$ on A, then*

$$\langle \nabla \mathcal{J}(\overline{a}), \overline{a} - a \rangle \leq 0, \quad \forall a \in A. \tag{5}$$

Conversely, if $\overline{a} \in A$ satisfies (5) and $\mathcal{J}$ is a concave function, then $\overline{a}$ is a global maximum of $\mathcal{J}$ on A.

Note that the inequality in (5) becomes an equality if $\overline{a}$ is an interior point of A.

When the admissible set is defined by an equality constraint, i.e. $A := \{a \in B : F(a) = 0\}$, one can define the **Lagrangian** of the problem

$$\mathcal{L}(a, \lambda) := \mathcal{J}(a) - \lambda \cdot F(a); \quad \forall a \in B; \lambda \in \mathbb{R}^+, \tag{6}$$

λ is called the Lagrangian multiplier associated to the constraint $F(a) = 0$. In terms of the Lagrangian, the first order optimality conditions are written:

Theorem 3 *[20] Consider $\mathcal{J}$ and F derivable in a neighborhood of $\overline{a}$, such that $F(\overline{a}) = 0$. If $\overline{a}$ is a local maximum of the problem*

$$\max_{a \in B; F(a) = 0} \mathcal{J}(a)$$

then, there exists $\overline{\lambda} \in \mathbb{R}$ such that

$$\frac{\partial \mathcal{L}}{\partial a}(\overline{a}, \overline{\lambda}) = \mathcal{J}'(\overline{a}) - \overline{\lambda} \cdot F'(\overline{a}) = 0, \quad and \quad \frac{\partial \mathcal{J}}{\partial \lambda}(\overline{a}, \overline{\lambda}) = F(\overline{a}) = 0.$$

Furthermore, equality constraints can be avoided thanks to the following relation

$$\max_{a \in B; F(a) = 0} \mathcal{J}(a) = \max_{a \in B} \min_{\lambda \in \mathbb{R}^+} \mathcal{L}(a, \lambda).$$

Finally, provided that the objective function is two times differentiable, the second order optimality conditions state

Theorem 4 *Assume $A = B$ and $\mathcal{J}$ two times differentiable at $\overline{a}$. If $\overline{a}$ is a local maximum of $\mathcal{J}$, then*

$$\nabla \mathcal{J}(\overline{a}) = 0, \quad and \quad \nabla^2 \mathcal{J}(\overline{a})(\phi, \phi) \leq 0, \quad \forall \phi \in B. \tag{7}$$

Reciprocally, if for all a in a neighborhood of $\overline{a}$

$$\nabla \mathcal{J}(\overline{a}) = 0, \quad and \quad \nabla^2 \mathcal{J}(a)(\phi, \phi) \leq 0, \quad \forall \phi \in B \tag{8}$$

then, $\overline{a}$ is a local maximum of $\mathcal{J}$ on A.

Optimization theory is very broad, and numerous results on existence and optimality conditions could be mentioned, see for instance [20]. However, the above simple and classical results are sufficient to understand the approach to the study of the optimization problem presented in chapters 2 and 3.

2 Optimization of *Wolbachia*-carrying mosquitoes release to slow dengue transmission

2.1 Motivation and description of the problem

The problem to which part I is devoted is motivated by a groundbreaking technique recently applied in several countries as a way to fight mosquito borne diseases. It is based on the effects of a bacterium called *Wolbachia* which prevents transmission of arboviruses from a carrying mosquito to humans. The mathematical problem being studied derives from a population rivalry generated by the forced introduction of a *Wolbachia*-carrying population within an environment colonized by wild mosquitoes, the final goal being to maximize the chances of population replacement.

The model adopted is a classical Allen-Cahn equation, which was first used as a model of the frequency of *Wolbachia*-carrying mosquitoes in [5]. The main variable $u : \mathbb{R}^+ \times \Omega \to [0, 1]$ corresponds to the proportion of *Wolbachia* carrier individuals with respect to the total mosquito population. It is assumed that this population spreads with a certain constant diffusion coefficient σ within a controlled environment Ω. The evolution equation is modeled following bistable dynamics: either *Wolbachia* population is established, which would mean that $u \to 1$ when $t \to \infty$, or it goes to extinction and $u \to 0$ when $t \to \infty$.

The model is expressed as:

$$
\begin{cases}
\partial_t u - \sigma \Delta u = \dfrac{s_h u (1 - u)(u - \rho)}{1 - s_f u - s_h u (1 - u)} & \text{in } (0, T) \times \Omega, \\[3mm]
u(0, x) = u_0(x) & \text{in } \Omega, \\[3mm]
\frac{\partial u}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega.
\end{cases}
\tag{9}
$$

The reaction term will often be identified as $f(u)$. The parameters s_h and s_f in eq. (9) are associated to the hatch and fecundity rates respectively. The value $\rho := \frac{s_f}{s_h}$ models the strength of the cytoplasmic incompatibility which is treated here as an analogue of the Allee effect. In chapter 1, the accuracy of this reaction term in the context of the reproductive cycle of *Wolbachia*-carrying mosquitoes will be explained in more detail.

The initial datum u_0 represents in this model the initial release; it plays a fundamental role, since it determines whether the *Wolbachia* carrier population will persist or disappear. The main question in the study of this subject is:

"How to release Wolbachia carrying mosquitoes with the goal of maximizing the chances
of establishing this population in the environment?"

The approach to find an answer to this question is based on the maximization of the total number of individuals with respect to u_0. This can be computed as

$$
\mathcal{J}_T(u_0) := \int_\Omega u(T, x) \, dx,
$$

where T is a fixed time and u is the solution of 9. Additionally, considering that the number of mosquitoes to be released is limited and determined by the capacity of each laboratory, the following constraint is adopted:

$$
u_0 \in \mathcal{A}_m := \left\{ u_0 : 0 \le u_0(x) \le 1, \int_\Omega u_0(x) \, dx = m \right\}
$$

where m is a fixed quantity depending on the production capacity.

2.2 State of the art

Similar problems have been already addressed in the context of fighting mosquito borne diseases by using *Wolbachia*, from an ordinary differential equation approach. In [3], for instance, a compartmental system, first introduced in [29, 30], is used to model the evolution of the densities of *Wolbachia*-infected and uninfected mosquitoes, $n_2(t)$ and $n_1(t)$, respectively. In the model,

$$\begin{cases} \dfrac{dn_1}{dt}(t) = f_1(n_1(t), n_2(t)), \\[2mm] \dfrac{dn_2}{dt}(t) = f_2(n_1(t), n_2(t)) + u(t), \end{cases} \qquad t > 0, \tag{10}$$

f_1 and f_2 describe reproductive dynamics of both populations and $u(t)$ stands here for a control which models the release of *Wolbachia*-carrying mosquitoes under certain constraints. The objective function is chosen as a least square type functional measuring the distance of the final state $(n_1(T), n_2(T))$ from the one with no wild population $(0, n_2^*)$,

$$J(u) := \frac{1}{2} n_1(T)^2 + \frac{1}{2}(n_2^* - n_2(T))_+^2. \tag{11}$$

The authors study a simplified version of this problem and prove that the best release protocol minimizing J is a single release either at the beginning or at the end of the time interval $[0, T]$. However, they underline that this may not be true for the full non-simplified problem. Similar questions are studied in [2] for a more complex model considering different stages of the reproductive cycle.

A different approach is studied in [14, 15], it is characterized by assuming periodic releases and thus a slightly different objective function including minimization of the terminal time and cultivation costs of *Wolbachia*-carrying mosquitoes in laboratory conditions.

Finally, we can also mention strategies based on feedback control techniques. In [8], a multi-release protocol guaranteeing successful invasion in finite time while keeping the control cost to a minimum is studied. More general feedback control principles for biological control of mosquito borne diseases are also presented in [7].

Returning to the question of maximizing the chance of establishment with respect to the initial datum, it is an undeniable fact that the spatial release profile is a determining factor. Hence, partial differential equations including a spatial component are fundamental. As mentioned in the previous section, common invasion processes are modeled by reaction-diffusion systems of the form:

$$\partial_t u - \Delta u = f(u). \tag{12}$$

For bistable reaction terms equation 12 exhibits a sharp threshold phenomenon (see Theorem 1) i.e. there exists a critical threshold L^* for which the solution of eq. (12), with initial datum $u_0 = \mathbb{1}_{[-L,L]}$, converges to $u = 1$ uniformly in compact sets if $L > L^*$ or tends to $u = 0$ globally uniformly in $\mathbb{R}$ if $L < L^*$. In [32], similar questions are addressed for a generalized family of initial data given by

$$u_0(x) = \mathbb{1}_{\frac{\alpha}{2} \leq |x| \leq \frac{L+\alpha}{2}} \quad \alpha \geq 0; L \geq 0.$$

The critical initial mass L^* for which a sharp threshold phenomenon occurs, just like in Theorem 1, is studied in [32] as a function of α. The results demonstrate the globally negative effects of fragmentation by proving that the function $L^* = L^*(\alpha)$ is increasing for large α. However, it highlights that for small values of α, $L^*(\alpha)$ decreases which would mean that on a local scale, fragmentation might be beneficial. This paper remains one of the first and primary motivations for this research as it clearly draws attention

to the complexity in the study of the optimal spatial distribution of initial datum even when considering relatively simple initial conditions. More recently this threshold property has been extended for a family of initial data of the form $\phi_L^\varepsilon = (\rho + \varepsilon)\mathbb{1}_{(L,L)}$, with $\varepsilon \in (0, 1 - \rho)$, [1]. Quantitative estimates on the sharp threshold value L_ε^* have also been provided in this case.

In [83], the invasion success probability (i.e. non-extinction) with respect to the initial data is studied. More precisely, theoretical and numerical lower bounds for the probability are stated by mean of reaction-diffusion models for the uncertainty quantification. The cases of single and multiple releases are studied in dimension one. However, further questions remains to be answered, for example, how to maximize the under-estimated probability of success with respect to the size of the release area, or how to optimize a release protocol in terms of the probability law of the release profile.

A different approach in the context of non-homogeneous releasing profiles includes feedback arguments, as in the following model,

$$
\begin{cases}
\partial_t u - \sigma \Delta u = f(u) + g(u)\mathbb{1}_{[0,T]\times\Omega}, & ; \quad t > 0, \ x \in \mathbb{R}^d, \\
u(0, x) = u_0(x), & ; \quad x \in \mathbb{R}^d.
\end{cases}
\tag{13}
$$

Indeed, in [9], it is proved that for $g(u) := (\mu(1 - u) - f(u))_+ \mathbb{1}_\Omega$, one can handle parameters $\mu > 0$, $T > 0$ and $\Omega \in \mathbb{R}$ such that the solution u of the system above converges to 1 as $t \to \infty$. However, conditions provided in [9] are only sufficient and may be improved for practical effects.

Finally, in [69], the effect of spatial variations in the total population density N is modeled by adding a population gradient term,

$$
\partial_t u - \Delta u - 2\frac{\nabla N \cdot \nabla u}{N} = f(u).
\tag{14}
$$

More precisely, it is proved that a heterogeneous environment inducing a strong enough population gradient can stop an invading front and converge to a stable front. In practice, supposing the total population is proportional to the environmental resources, changes on the total population might be due to environmental, geographical and other physical factors inducing changes in the carrying capacity. This result might be very useful for designing *Wolbachia*-carrying mosquitoes release protocols overriding propagation hindrance.

2.3 Presentation of the main results

Although all the previous works connect in some way with the topics that will be addressed in this manuscript, it seems that none of them have provided a geometrical characterization of an optimal initial profile in a PDE model assuming a single release. As far as that is concerned, in chapter 2 of this document it will be proved that the problem

$$
\max_{u_0 \in \mathcal{A}_m} \mathcal{J}_T(u_0) = \mathcal{J}_T(\overline{u}_0)
\tag{15}
$$

is well-posed and has at least one solution. Moreover, any solution $\overline{u}_0$ will be partially characterized by means of the adjoint state $\overline{p}$ given by

$$
\begin{cases}
-\partial_t \overline{p} - \sigma \Delta \overline{p} = f'(\overline{u})\overline{p} & \text{in } (0, T) \times \Omega, \\
\\
\overline{p}(T, x) = 1 & \text{in } \Omega, \\
\\
\frac{\partial \overline{p}}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega.
\end{cases}
\tag{16}
$$

More precisely, under certain standard regularity assumptions on f, it will be proved in chapter 2 the following

Theorem 5 *[68] There exist $\overline{u}_0 \in \mathcal{A}_m$ such that*

$$\max_{u_0 \in \mathcal{A}_m} \mathcal{J}_T(u_0) = \mathcal{J}_T(\overline{u}_0). \tag{17}$$

Moreover, setting $\overline{u}$ the solution of eq. (9) associated with this optimum initial data and $\overline{p}$ the unique solution of eq. (16), there exists a non-negative real value denoted by c such that

  i) *if $0 < \overline{u}_0(x) < 1$ then $\overline{p}(0, x) = c$,*

  ii) *if $\overline{u}_0(x) = 0$ then $\overline{p}(0, x) \leq c$,*

  iii) *if $\overline{u}_0(x) = 1$ then $\overline{p}(0, x) \geq c$.*

This result is also true for more general reaction terms $f(u)$ as long as they are sufficiently regular. Although, in the most general case the optimal element seems to be not unique, in the case of a concave $f(u)$, the uniqueness is fulfilled since it is shown that the concavity is inherited by the functional $\mathcal{J}_T$.

The partial characterization given by Theorem 5 is not sufficient to explicitly calculate the solution of the optimization problem (15) independently of f; however, there are some particular cases where it can be done. For instance, considering the classical Fisher-KPP reaction term $f(u) = u(1 - u)$, since it is concave, it follows that the unique maximizing element has a homogeneous profile determined by the size of Ω and the capacity constraint m.

The proof of Theorem 5 stands on compactness arguments and a priory regularity estimates on u and $\mathcal{J}$. The main difficulty arises from the fact that the set $\Omega_c := \{x \in \Omega : p(0, x) = c\}$, in which the maximizer element is different from zero or one, might be non-negligible and thus $\overline{u}_0$ might be not bang-bang, as is usually considered in papers studying related subjects (see fig. 4). However, for the proof, this issue is avoided thanks to an argument based on the Lebesgue density theorem.

| (a) | (b) |
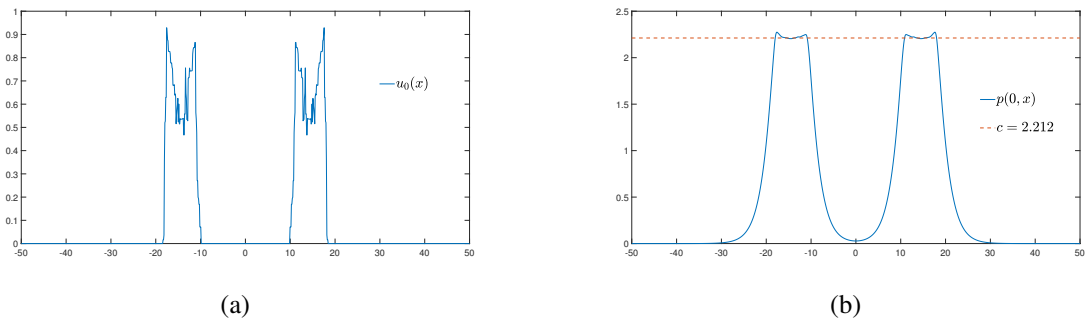
Figure 4: Singularities of a local maximizer $u_0(x)$ showed in (a) within the set Ω_c. In this case, for a bistable reaction term of the form $f(u) = u(1 - u)(u - \rho)$, the set Ω_c is non-negligible, as it is showed in figure (b).

As a numerical approach to the resolution of the optimization problem, a numerical algorithm will be presented in section 2.5. It derives from a classical gradient ascent algorithm modified such that the partial characterization given by (i)-(iii) in Theorem 5 is applied to iteratively converge to a local maximum, see the scheme in fig. 5. Numerically, the algorithm has a good performance and converges after a few iterations.
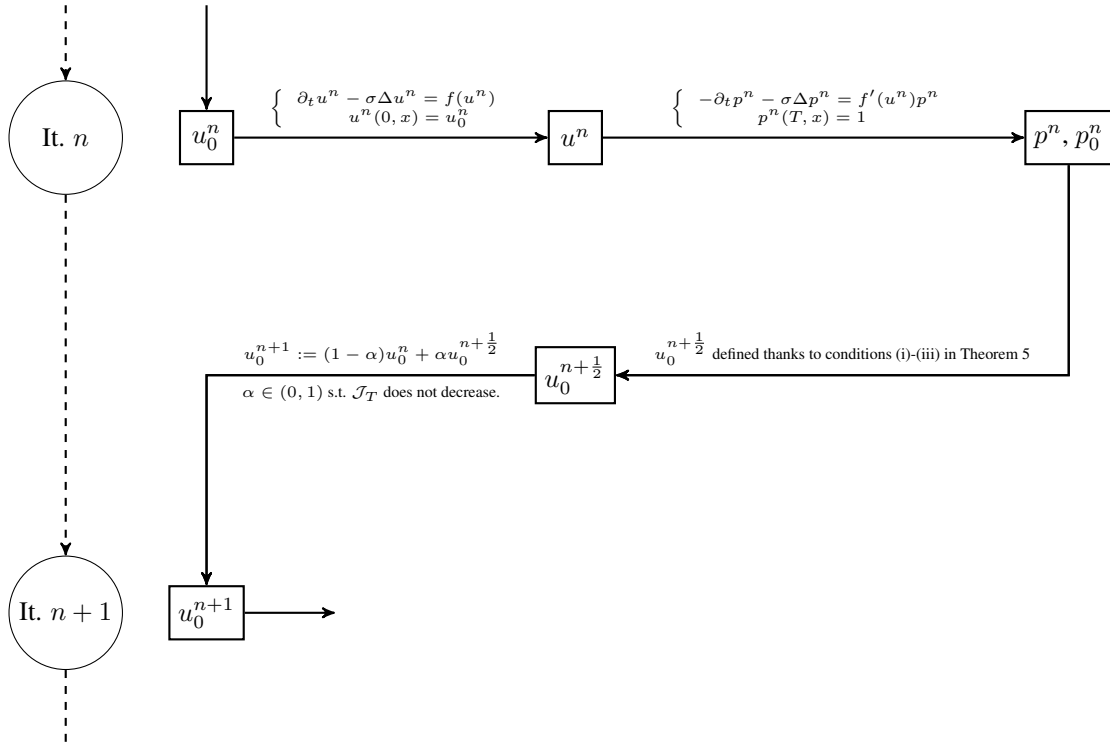
Figure 5: Scheme of the numerical algorithm, for more details see section 2.5.

The results stated in Theorem 5 rely only on the first order optimality conditions arising from the optimization problem (15). The characterization of the optimizers might be improved by exploring second order ones, though for general reaction terms this is a thorny problem. The case of bistable nonlinearity will be addressed in chapter 3, the following theorem is established:

Theorem 6 *Provided $\overline{u}_0$ is a solution of the problem (15) with bistable reaction term f, then it holds that $f''(\overline{u}_0(x)) \leq 0$ for every interior point x of the set $\Omega_{\overline{c}} := \{x \in \Omega : 0 < \overline{u}_0(x) < 1\}$.*

The previous results together with Theorem 5 allows a better characterization of any optimizer in the bistable case. Indeed, the sign of the second order derivative of f provide information about the behavior of $\overline{u}_0$ at least at any interior point of the set $\Omega_{\overline{c}}$ and this allows an improvement of the numerical algorithm. However, since we do not have much information on the regularity of $\Omega_{\overline{c}}$, this remains a partial characterization.

As an appendix to this manuscript, we include some questions related to the numerical algorithm that will be described in detail in the section 2.5 of this document. The appendix A is devoted to present some numerical simulations in the one-dimensional case and to compare the algorithm derived from our results with others well known in numerical optimization.

The appendix B shows simulations in the two-dimensional case, corroborating the adaptability of our algorithm to higher dimensional problems.

Finally, in the appendix C, we discuss through examples an alternative to adapt the algorithm output into more suitable and realistic initial data. Indeed, the regularity issues of the optimizers associated to the problem we study in the first part of this manuscript makes almost impossible the application of our results for practical fines, such as the biological control of dengue virus by using *Wolbachia*.

The strategy we propose in the appendix C is to rearrange the optimizer $\overline{u}_0$ into a piecewise constant function with the same support of $\overline{u}$ and the same mass inside each convex component, see fig. 6.
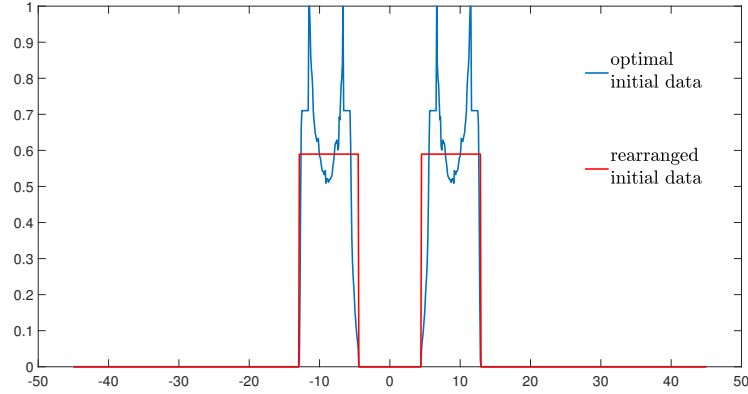
Figure 6: In blue is showed a local optimizer obtained by mean of our numerical algorithm $\overline{u}_0$, in red the associated rearranged initial data with same mass and support, (for more details see example 1 in the appendix C).

The numerical results presented in appendix C show that the value of the objective function $\mathcal{J}_T$ does not vary dramatically from the optimizer $\overline{u}_0$ to the piecewise constant distribution, in most of the cases. This seems to point out this strategy as an alternative to make our results applicable for practical purposes. However, for the moment it remains an empirical approach.

2.4 Perspectives and remain works

Several lines of research remains to be explored. For instance, getting to the bottom of the regularity properties of the optimizers remains unreached. Results in this sense might help to improve our understanding of this problem. However, the intuition and the numerical simulations seem to point out the difficulty of such aspects.

Particular cases could also be addressed in order to have a better comprehension of the general profile of the optimizers. For instance, one might study the case with ignition type nonlinearity or the limit problem assuming large diffusivity.

As in [32], one might also wonder how change the profile of optimizers when considering the same maximization problem with infinite time, i.e. $T \to \infty$. This hypothesis is slightly discussed in section 2.6.2, but still several aspects remains to be clarified.

3 A model of the inflammation due to Crohn's disease

3.1 Inflammatory bowel diseases

The term inflammatory bowel disease (IBD) is used to refer to two medical conditions that involve chronic inflammation of the gastrointestinal tract, namely: ulcerative colitis and Crohn's disease. IBD might cause gastrointestinal organ's dysfunction which can lead to persistent diarrhea, rectal bleeding, abdominal pain, poor nutrient's absorption, weight loss and fatigue. IBD is currently a growing global problem, statistics indicate that around 0.5 % of the western world population is now affected by one of these two conditions.

Ulcerative colitis was described for the first time in 1875 and Crohn's disease in 1932. Before this year they were confused with diarrheal diseases caused by infectious agents or associated to intestinal tuberculosis. There is currently a consensus on the inherited genetic predisposition of patients suffering from IBD, who develop inadequate immune responses triggered by still unknown environmental factors. Treatments are oriented to managed symptoms, but no cure is found yet. Although, patients can go through periods when the disease is quiet with few or no symptoms, the chronicity of this illness might carry changes and complications over time, [6].

Although both conditions present similarities, there are particularities that differentiate them. While ulcerative colitis is limited to the large intestine and the rectum, Crohn's disease can affect any part of the gastrointestinal tract, from the mouth to the anus. In the case of the ulcerative colitis, inflammation only affects the innermost layer of the lining of the intestine, and always starts from the rectum evolving continuously through the colon until it stops, inexplicably, at some point in the large intestine. On the contrary, Crohn's disease can leave healthy tissues between two inflamed sections but the inflammation may extend through the entire thickness of the bowel wall, see fig. 7.

(a) Crohn's disease.   (b) Ulcerative colitis

Figure 7: Inflammatory bowel diseases

Mathematical models based on cell interactions and the biophysics constraints of the diseases are particularly important for aiming to predict and prevent inflammation. All over the world multidisciplinary research institutions gather their efforts in order to better understand IBD, improve treatments and develop early diagnostic tests. The INFLAMEX excellence cluster in France (ANR-10-LABX-17) assembles high level scientists and develops interdisciplinary research in this field. As a result of the collaboration between INFLAMEX and the University Sorbonne Paris Nord, the idea of conceiving and studying a model as a starting point to understand the differences in terms of the inflammatory patterns between Crohn's disease and ulcerative colitis was developed.

In chapter 4 a model based on the interactive dynamic between pathogen bacteria (noted here as β) arriving from the intestinal flora trough the epithelial barrier, and the immune cells (noted as γ) migrating to the zone penetrated by pathogen agents is developed in detail. This model was the result of extensive and interesting discussions with E. Ogier-Denis, a researcher at the French National Institute for Scientific Research and also a member of the INFLAMEX excellence cluster. The model is given by:

$$\begin{cases} \partial_t \beta - d_b \Delta \beta &= r_b \left(1 - \frac{\beta}{b_i}\right)\beta - \frac{a\beta\gamma}{s_b + \beta} + f_e \left(1 - \frac{\beta}{b_i}\right)\gamma, \\ \\ \partial_t \gamma - d_c \Delta \gamma &= f_b \beta - r_c \gamma. \end{cases} \tag{18}$$

For the sake of simplicity, the spatial domain is assumed to be a one-dimensional interval that simulates a part of the gastrointestinal tract. Far from being realistic, the model is primarily intended to understand the irregular nature of Crohn's disease and to demonstrate that it could be the result of an interaction dynamic with a high diffusivity disparity, i.e. a Turing-type phenomena. The first part of chapter 4 will be devoted to discussing the accuracy of eq. (18) as a model of inflammation and later,

stability properties of the system will be studied, as well as an estimation of the parameters for numerical simulation. All the results presented in the aforementioned chapter have been obtained in collaboration with G. Nadin, H. Zaag and E. Ogier-Denis.

3.2 State of the art

In the scientific literature, mathematical models specifically targeting IBD dynamics are not numerous. We can cite the model studied in [93], where a system of 29 ordinary differential equations is constructed representing movements and interactions of several types of immune cells with bacteria in the lumen and other zones of the colon. Parameter values were then chosen to model existing data sets. However, from an analytical point of view, inferring qualitative properties is not easy due to the complexity of the system.

Looking for simpler models accurately simulating the inflammation process causing IBD, one can find a variety of papers globally describing, by mean of ordinary differential equations, the acute inflammatory response of the human body in various situations: [24, 18, 84, 35, 61, 48, 21]. In most cases, when the model allows it, stability properties are studied, as well as bifurcation phenomena, and phase portrait analysis. Numerical simulations are also regularly provided. Several methods oriented towards modeling the acute inflammatory response with ordinary differential equations are discussed and compared in [90].

However, since the questions that motivated this research concern the irregular nature of the inflammatory patterns due to Crohn's disease, the study of models that include spatial heterogeneities was of great interest. This drove us to the analysis of partial differential equations (PDE) based models related to inflammation phenomena. For instance, in [53, 52] models including space were introduced. The authors demonstrated how, in the dynamical behavior of bacteria interacting with immune cells, abnormalities can lead to excessive body responses and thus cause inflammation. Nevertheless, pattern formation was not addressed.

The work that most closely connects to the issues addressed here is [74]. Indeed, in this paper the authors sought to understand patterns appearing on the skin due to acute inflammation and, in order to do so, they proposed a classical chemotaxis model slightly modified to include inhibitory effects. This resulted in a variety of spatial patterns including isolated traveling pulses, rotating waves among others. However, it is important to remark that those patterns do not rely on Turing-type instabilities.

It would be pertinent also to mention numerous research works on atherosclerosis [25, 26, 17, 12, 36]. Indeed, at early stages, this medical condition is the result of acute inflammation driven lesion development. Although the agents interacting are not exactly those concerning IBD, the dynamics are closely related since both concern immune abnormal responses of the human body to the presence of a pathogen agent.

3.3 Presentation of the main results

Returning to the model (18), the main theoretical result, which will be established and proved in chapter 4 highlights the ability of this model to simulate pattern formation by stating the existence of Turing instabilities under certain conditions. This is the statement:

Proposition 7 *[70] Equation (18) has a unique positive homogeneous steady state solution* $(\overline{\beta}, \overline{\gamma})$. *Moreover, if there exist real non-negative values of the parameters* $a, r_b, r_c, s_b, f_e, f_b, b_i$ *such that the following condition holds:*

$$0 < \frac{a\kappa\overline{\beta}^2}{(s_b + \overline{\beta})^2} - r_b\theta - f_e\kappa < r_c, \tag{19}$$

then for $\frac{d_b}{d_c}$ small enough the reaction diffusion system (18) shows Turing instabilities around this steady state.

The proof of this result relies on classical techniques of stability analysis to show the existence of steady state solutions that became unstable under non-homogeneous perturbations. The condition that $\frac{d_b}{d_c}$ is small, from a biological point of view, means that the bacteria have a low relative diffusion coefficient with respect to the immune cells. This condition is realistic according to the values registered in the scientific literature. In fact, once the epithelial barrier is crossed, the pathogens do not spread very far. The infection remains local and it is the function of the immune cells to move to the damaged area, initially through the blood vessels and then through the intestinal tissues.

It is important to emphasize that Proposition 7, however, does not guarantee the existence of parameters satisfying eq. (19). In this sense, one of the main difficulties encountered during this research was to prove the non-emptiness of this set and moreover to estimate relatively accurate values from a biological point of view. The section 4.5 will be dedicated to discussing these issues. In some cases the value of the parameter will be retrieved from the existing literature, while in others they will be calculated. Numerical results can be observed in fig. 8, for the parameter values presented in table 1, for which eq. (19) holds.
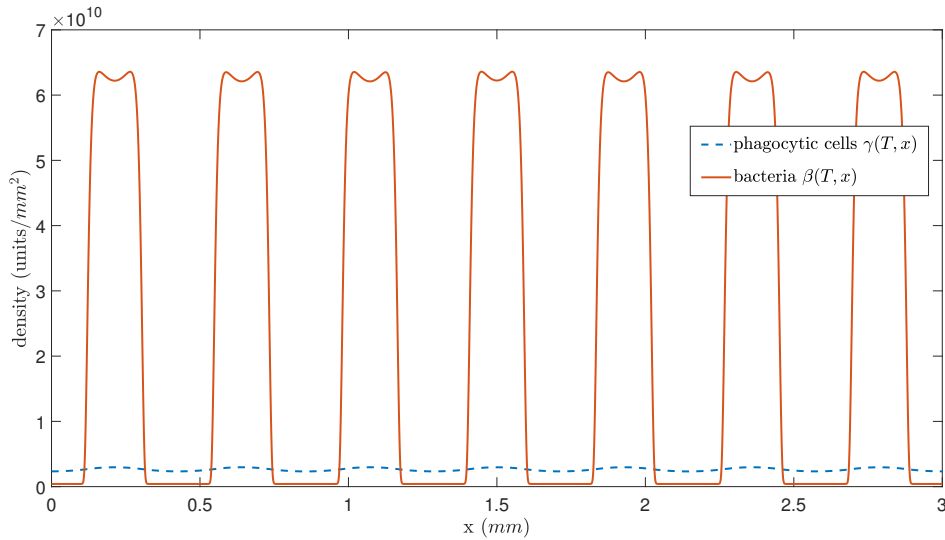
Figure 8: Bacteria (red line) and phagocytes (blue dashed line) after a time-lapse of 2 weeks with an initial bacterial infection $\beta_0(x) = 10^9 \times \mathbb{1}_{[1.495,1.505]}$ and $\gamma_0(x) = 0$

3.4 Perspectives and remain works

A natural continuation of this research would be to study the possibility of describing the continuous inflammatory pattern that characterizes the ulcerative colitis by means of the same model given by the eq. (4.1). In this case one needs to search for solutions in which the bacterial density varies abruptly from regions of high concentration to regions of low concentration, associated to the rectum and the healthy tissues respectively. This change of regime might rely on a change in the parameters values.

For instance, b_i associated with the density of bacteria in the lumen could be considered as a variable parameter dependent on x. In fact, it is known that the closer to the rectum, the greater the presence of bacteria in the lumen. Then, one could try to show that under certain conditions, the model accepts stationary wave solutions like in fig. 9.

Table 1: Assigned values for the parameters of the model (4.1)

parameter	interpretation	value	units
r_b	Reproduction rate of bacteria	0.0347	(u/min)
r_c	Intrinsic death rate of phagocytes	0.02	(u/min)
d_b	Diffusion rate of bacteria	10^{-13}	(m^2/min)
d_c	Diffusion rate of phagocytes	10^{-10}	(m^2/min)
b_i	Density of bacteria in the lumen	10^{17}	(u/m^3)
f_b	Immune response rate	0.002	(u/min)
a	Coefficient proportional to the rate of phagocytosis ($a = s_b p_c$)	0.3129	(u/min)
	it is also inversely proportional to the handling time ($a = \frac{1}{\tau}$)		
s_b	Proportionality coefficient between p_c and a	10^{15}	(u/m^3)
f_e	Related to the porosity of the epithelium	0.0856	(u/min)

Figure 9: Profile of a continuous inflammation in the gut.

Some members within the IBD medical community, suggest that the points where inflammation stops, in the case of ulcerative colitis, might often occur in places where the digestive tract folds. Consequently, it would be interesting to study the influence of the geometry of the spatial domain in the final solution. For example, two-dimensional or three-dimensional domains could be considered that include angles of twist and variations in diameter. This would also make the model more realistic.

4 Dissertation outline

The rest of the manuscript is organized as follows

- Chapter 1 is a review on the biological control technique motivating the optimization problem that is studied later in this manuscript. Although this chapter does not contain any original research, it is essential as a bridge between mathematics and biology.

- Chapter 2 corresponds to [68]. It is devoted to the study of an optimization problem on the solution of reaction-diffusion equations with respect to their initial data. It has been published in *Journal of Mathematical Modelling of Natural Phenomena*.

- Chapter 3 is a work in progress in collaboration with Grégoire Nadin and Idriss Mazari. It concerns the study of the second order optimality conditions arising from the optimization problem

presented in the previous chapter in the case of a bistable reaction term.

- Chapter 4 is devoted to the study of a mathematical model of inflammation in the context of inflammatory bowel diseases. From a mathematical point of view, the emergence of Turing instabilities is studied. The paper has already been submitted, [70].

Part I

Wolbachia control technique and optimization

This part is devoted to the study of a maximization problem for the quantity $\int_\Omega u(t,x)dx$ with respect to $u_0 := u(0,\cdot)$, where u is the solution of a given reaction-diffusion equation. In chapter 1, some insights on the technique that motivates the problem from a biological point of view, are given. The existence of a solution to the aforementioned problem as well as the first and second order optimality conditions are studied in chapters 2 and 3, where properties on the optimizers are stated. A numerical algorithm to approximate the maximizer is also provided in chapter 2, and improved thanks to the results presented in chapter 3.

Chapter 1

Mathematical modelling and vector control techniques

"Nothing in life is to be feared. It is only to be understood."

Marie Curie

1.1 General facts about dengue virus

According to the World Health Organization (WHO), dengue is a mosquito-borne viral infection which is usually present as a flu-like illness with symptoms lasting from 2 to 7 days. The disease spectrum is very wide and goes from asymptomatic or mild and self-managed cases, which represent a vast majority, to most serious clinical pictures involving complications associated with severe bleeding, organ impairment and/or plasma leakage. The latter is known as severe dengue and, although less common, has a high mortality.

Dengue is found in tropical and sub-tropical climates worldwide, mostly in urban and semi-urban areas. In the last few years, the incidence of severe dengue epidemics has increased alarmingly, evolving from being reported in only 9 countries in the seventies, to currently being endemic in more than 100 countries. Latin American, South-East Asia and Western Pacific regions are the most seriously affected, see fig. 1.1. Nevertheless, cases of local transmission have already been reported in Europe, where the risk of an outbreak of dengue is now latent.

Transmission occurs through the bites of infected females mosquitoes, mainly of the species *Aedes aegypty*, but also to a lesser extent the *Aedes albopictus*. The same species are also transmission vectors for many other viral pathogens, such as yellow fever, chikungunya, and zika. Mosquitoes are infected after stinging an infected person. An infected person is only infectious while viremic, a period of about one week. On the contrary, there is an incubation time of several days before the mosquito is capable of transmitting the virus, but after this time mosquito remains infectious for the rest of its life.

1.1.1 Vector control techniques

Since there is no unanimously accepted vaccine or specific treatment for dengue, prevention and control depend entirely on effective vector control measures. There are a number of ways mosquito-borne diseases can be controlled, for instance:
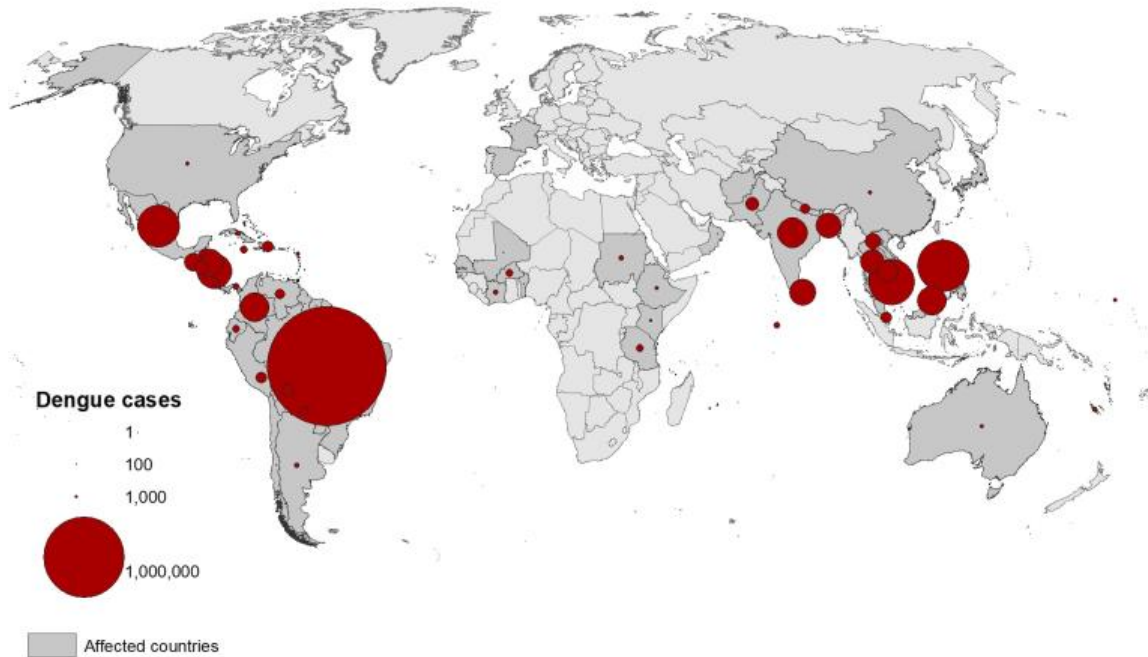
Figure 1.1: Geographical distribution of dengue cases reported worldwide, July 2020 (European Center for Disease prevention and Control).

Insecticide spraying: This method consists of spraying chemical toxins that kill mosquitoes, but also other insects. In some cases the toxins can harm human health as well. They are applied both indoors and outdoors, and require reapplication periodically. Evidence that proves that some mosquitoes can develop resistance to some insecticides has already been registered. Insecticide spraying is the most commonly used method around the world, however it is not cost-effective as a prevention tool.

Sterile insect technique (SIT): This is a technique that attempts to diminish or even eradicate mosquito population by inhibit the reproductive process. SIT uses irradiation as a method to sterilize male mosquitoes, which are then massively released in nature where they mate with female mosquitoes resulting in eggs that do not hatch. The effectiveness of this method depends on the size of the releasing and also on its periodicity, otherwise the population can rebound. For a mathematical approach to this technique the reader can see for instance [2, 7, 82].

Genetic modification: This method is characterized by the introduction of lethal genes into the target mosquito population which causes the population to decrease over time. The genetic modification can be expensive because of the need for continuous application. It is currently being used in Brazil where it has already resulted in a reduction of the mosquito population, but there is still no conclusive evidence as to whether it reduces dengue transmission. A two years trial is also recently underway in some zones in the USA. [22]

All the previously mentioned methods target the reduction of the mosquito population as a way to prevent dengue transmission. However, unless the species is completely eliminated, a prolonged period without human intervention would allow the mosquito population to rebound and return to the original

situation. Therefore, the methods are not self-sustaining in the long term and perpetually depend on the commitment of local authorities.

There is a relatively young method with a different approach to the problem of stopping mosquito-borne disease transmission. The idea is to prevent infection by making mosquito unable to transmit the pathogen agent to humans. The key element of this strategy lies in the bacterium *Wolbachia*.

Wolbachia control method

Wolbachia is natural bacterium present in up to 60 percent of insect species. However, it is not present in *Aedes aegypti*. It was discovered in 1924 but it was not until 1997 that researchers started to study the relationship between *Wolbachia* and mosquito-borne-disease. At the time, a strain of these bacteria had been shown to shorten the lifespan of *Aedes aegypti* and therefore could reduce disease transmission. Later works on this subject, [66], pointed out the crucial fact that *Wolbachia* bacterium compete with viruses like dengue, zika, chikungunya and yellow fever. As a consequence, the replication of these viruses is highly reduced in *Wolbachia*-carrying mosquitoes. Furthermore, *Wolbachia*-carrying female mosquitoes were shown to transmit the bacteria to their offspring.

This scientific revelation together with the fact that *Wolbachia* is safe for human and other animals, helped to conceive what is known today as the *Wolbachia* method to eliminate dengue: Scientists breed *Wolbachia*-carrying mosquitoes in laboratories and then release them into the wild. If this initial population is strong enough to establish and survive, they can pass on the bacterium to the next generations and one can expect that, after a long time, the wild mosquito population would be replaced by the *Wolbachia*-carrying one. In this way, dengue and other mosquito-borne-disease transmission can be prevented.

Wolbachia's hereditary property makes this method potentially self-sustaining, however several factors must be taken into account to ensure its successful application. *Wolbachia* can only be transmitted from female mosquitoes to its offspring inside the eggs, and in the carrying females it reduces the relative fecundity with respect to *Wolbachia*-free mosquitoes. Additionally, a wild female cannot have descendants with a male carrier of *Wolbachia*; this effect is known as cytoplasmic incompatibility and might initially lead the *Wolbachia*-carrying population to extinction if it is not strong enough.

Laboratory and field research on the *Wolbachia* method have been managed and coordinated by the World Mosquito Program (WMP) since the 2000s. The first *Wolbachia* releases were made in the Cairns region in Australia in 2011. The experiment lasted 10 weeks with a release per week. The results were encouraging and overwhelming about the effectiveness of this method, see [79]. In recent years, several regions have joined the WMP, which now has the collaboration of more than 10 countries. Clinical trials, risk analysis, economic impact studies, mathematical modelling studies, etc. are essential to fight dengue and other mosquito-borne-diseases on a global scale.

1.2 A model for *Wolbachia* spreading

The evolution and spreading of *Wolbachia* infection in the wild mosquito population is modelled here in terms of the frequency or density $u(t, x)$ of *Wolbachia*-carrying mosquitoes with respect to the total mosquito population, i.e.

$$u := \frac{\textit{Wolbachia}\text{-carrying individuals}}{\text{Total mosquito population}}$$

We target a wild population evolving within a bounded and regular domain Ω, and we consider an initial release at time $t = 0$ that we denote as $u_0(x)$. As explained above, the reproductive dynamics between this founding population and the wild population are characterized by three fundamental facts:

1. *Wolbachia* shows perfect maternal transmission.

2. *Wolbachia* infected females have relative fecundity $F \leq 1$ with respect to *Wolbachia*-free females.

3. Due to the cytoplasmic incompatibility the crosses between *Wolbachia* carrier male and wild females are incompatible and produce a relative hatch rate of $H \leq 1$ with respect to *Wolbachia*-free males (see table 1.1)

Table 1.1: Reproductive dynamic between *Wolbachia*-carrying mosquitoes (🦟) and wild mosquitoes (🦟).

Setting $F = 1 - s_f$ and $H = 1 - s_h$ and considering discrete generations and random mating probability, after the first reproductive cycle the frequency of *Wolbachia* infected mosquitoes is

$$U_1 = \frac{U_0(1 - s_f)}{1 - s_f U_0 - s_h U_0 (1 - U_0)}$$

This dynamic, in a continuous time model, generates a reaction term of the form

$$F(u) = \frac{s_h u (1 - u)(u - \frac{s_f}{s_h})}{1 - s_f u - s_h u (1 - u)}. \tag{1.1}$$

In the regime of strong cytoplasmic incompatibility effects ($s_h \sim 1$) and weak fecundity effects ($s_f \sim 0$) induced by *Wolbachia*, it is usual to simplify the study of F by considering the nonlinearity

$$f(u) = ru(1 - u)(u - \rho), \tag{1.2}$$

with $r = s_h$ and $\rho = \frac{s_f}{s_h}$. Indeed, with $0 \leq u \leq 1$, both functions behave similarly. Finally, the model, which has been presented initially in [5], reads

$$\begin{cases} \partial_t u - \sigma \Delta u = ru(1 - u)(u - \rho) & \text{in } \mathbb{R}_+ \times \Omega, \\ \\ u(0, x) = u_0(x) & \text{in } \Omega, \\ \\ \frac{\partial u}{\partial \nu}(t, x) = 0 & \text{for all } t \in \mathbb{R}_+, \text{ for all } x \in \partial \Omega \end{cases} \tag{1.3}$$

The model is bistable under the regime $\rho < 1$. In this case the two stable states are associated either with the *Wolbachia* population replacement of the wild population $u \equiv 1$, or with the total extinction $u \equiv 0$ of the *Wolbachia* population. There is also an unstable state $u \equiv \rho$ associated with an establishment state in which both *Wolbachia*-carrying population and wild population coexist.

Chapter 2

On the maximization problem for solutions of reaction-diffusion equations with respect to their initial data

"The popular view that scientists proceed inexorably from well-established fact to well-established fact, never being influenced by any unproved conjecture, is quite mistaken. Provided it is made clear which are proved facts and which are conjectures, no harm can result. Conjectures are of great importance since they suggest useful lines of research".

Alan Turing

This chapter corresponds to the paper [68], which has been published in the Journal of Mathematical Modelling of Natural Phenomena and is the result of a joint work with Grégoire Nadin.

2.1 Introduction

2.1.1 Statement of the problem and earlier works

We investigate in this chapter the following optimization problem: given $T > 0$, we want to maximize the functional $\mathcal{J}_T(u_0) := \int_\Omega u(T, x) dx$ among all possible initial data $u_0 \in \mathcal{A}_m$, where

$$\mathcal{A}_m := \left\{ u_0 \in \mathcal{A} : \int_\Omega u_0 = m \right\} \quad \text{with} \quad \mathcal{A} := \left\{ u_0 \in L^1(\Omega), \ 0 \le u_0 \le 1 \right\} \tag{2.1}$$

and $u = u(t, x)$ is the solution of the reaction-diffusion equation

$$\begin{cases} \partial_t u - \sigma \Delta u = f(u) & \text{in } (0, T) \times \Omega, \\\\ u(0, x) = u_0(x) & \text{in } \Omega, \\\\ \frac{\partial u}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega \end{cases} \tag{2.2}$$

24

Here, u represents the density of a population, and $\int_\Omega u(t, x)dx$ is thus the total population at time t. Given an initial total population m, we thus want to place it in such a way that the total population at time T is maximized.

This is a very natural problem but, as far as we know, it has never been addressed. Let us just mention three papers that investigate similar questions.

In [32], the case of a particular initial datum $u_0^\alpha = \mathbb{1}_{[-\frac{L}{2}-\frac{\alpha}{2}, -\frac{\alpha}{2}] \cup [\frac{\alpha}{2}, \frac{L}{2}+\frac{\alpha}{2}]}$, with $\alpha \geq 0$, has been investigated for a bistable non-linearity $f(u) = u(1-u)(u-\rho)$, at infinite horizon $T = +\infty$. In that case, when $\Omega = \mathbf{R}$, for any given $\alpha \geq 0$ it is known from [94] that there exists a critical mass $L^*(\alpha) > 0$ such that for any $L < L^*(\alpha)$ the solution goes to 0 and for $L > L^*(\alpha)$ it converges to 1. The authors provided numerics [32] showing that one could get $L^*(\alpha) < L^*(0)$ for α small. This means that for a given initial total population $L \in (L^*(\alpha), L^*(0))$, the initial datum u_0^α associated with two blocks separated by a small gap will converge to 1, while the initial datum u_0^0 associated with a single block will converge to 0 (fig. 2.1). This example shows that our present optimization problem could be difficult, since fragmented initial data could give a better total population at time $T >> 1$. Hence, we expect regularity issues on a possible maximizer.
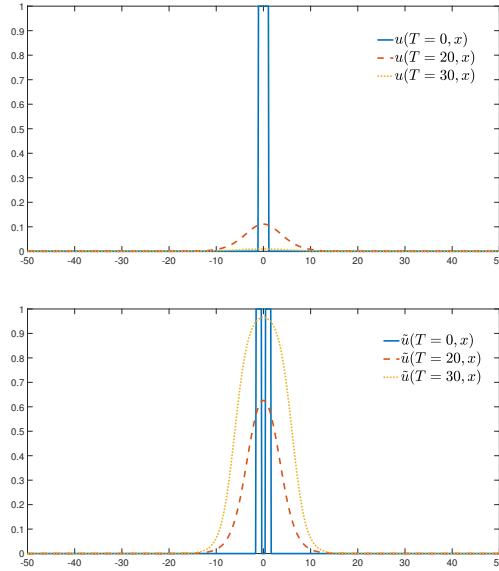
Figure 2.1: The graphs show the influence of the initial datum. On the top the initial population $u(0, x)$ is concentrated in a single block of mass $m = 2.22$ (solid blue line). Also is showed the solution $u(T, x)$ of eq. (2.2) after different times lapses: $T = 20$ (dashed red line), $T = 30$ (dotted yellow line); clearly the population tends to disappear as time goes by. On the bottom we consider an initial population with the same mass as before but distributed into two blocks slightly separated $\widetilde{u}(0, x)$, the resulting density after same given time periods is clearly bigger, in this case the population tends to establish.

In [13] a similar problem as the present one is investigated, with a more complex cost, but for a concave non-linearity f, which will latter appear to be quite restrictive in our case (see section 2.4.1), and with a global control at every time $t \in (0, T)$. First order optimality conditions are heuristically derived, but the authors do not investigate it further in order to determine the optimal control.

Lastly, in [85], the authors consider a bistable non-linearity $f(t, u) = u(1-u)(u-\rho(t))$, and the control is ρ, which is assumed to belong to $[0, 1]$ for all $t \geq 0$. The authors prove that with such a

control, one could get arbitrarily close to a target function -a travelling wave- considering a sufficiently large time.

Let us also mention [3], where a similar model is investigated. In this paper, a particular bistable non-linearity is considered, and the authors optimize the L^2 distance to 1 at time T for several releases at various times. They prove the existence of an optimizer, compute the first order derivative, and then consider a toy model (with $f \equiv 0$) and the particular case where u_0 lies in the class of additions of Gaussian-type functions, for which they optimize on the centers of the Gaussian functions numerically.

The main contributions of this research are the following. First, we show that there exists a maximizer for the functional $\mathcal{J}_T$. Second, we establish some optimality conditions for this maximizer arising from the study of the adjoint state. This allows us to provide a numerical algorithm to approximate this optimal distribution in practice.

Before getting into the statement of our results, let us briefly comment on the biological motivations of this work.

2.1.2 Biological motivation

Dengue fever also known as breakbone fever and dandy fever is caused by dengue virus, which is transported and transmitted by mosquitoes of the genus known as *Aedes*, the two most prominent species being *A. aegypti* and *A. albopictus*. Nowadays the progress of this virus is increasing and with it, the interest of finding a way to control it in absence of an effective medical treatment.

Manipulation of the arthropod population by introducing a maternally inherited bacterium called *Wolbachia* has recently caught the attention of biologists [4, 37, 39, 91]. In the infected mosquito, this bacterium prevents the development of the virus but also induces a cytoplasmic incompatibility which reduces their reproduction rate when the infected population is small but became non-important once its density becomes sufficiently large [5].

Reaction-diffusion equations have been widely used in order to describe biological phenomena of spreading and species competition, thanks to the works of [5] the dynamic between infected and non-infected mosquitoes population can be described using a Lotka-Volterra system. It has been rigorously shown that it may be studied by mean of a single reaction diffusion equation on the proportion u : $\mathbb{R}^+ \times \Omega \to [0, 1]$ of infected individuals with respect to the total population.

In such models, the reaction term $f(u)$ reflects the positive correlation between population density and individual fitness, known as the Allee effect. In the current problem, this effect is caused by the cytoplasmic incompatibility, so there exists a critical density threshold ρ under which the population of infected mosquitoes declines, but above which it increases. In fact, we have $f(u) < 0$ if $0 < u < \rho$ and $f(u) > 0$ if $\rho < u < 1$. Hence, there is a bistable mechanism, either the infected population disappears (i.e. $u \to 0$ when $t \to \infty$, also called extinction), or the whole population get infected after a sufficiently large lapse of time (i.e. $u \to 1$ when $t \to \infty$, also called invasion).

Of important and practical interest is the study of sufficient conditions on the different parameters of the problem in order to reach the invasion state once the deliberately infected mosquito population gets released in the environment. Different approaches to this problem have been done in recent literature, from the biological point of view [33] and also from the mathematical one. In [85], that we already mentioned, it is proposed as a strategy to modify the Allee threshold ρ in order to reach an a priori given target trajectory; in practice this is possible by means of manipulation of different biological factors which affect directly the mosquito population like increasing or decreasing natural predator's population or affecting carrying capacity of the environment. A similar problem is studied in [9]. In fact it is proved that there exists a systematic way to choose a time $T > 0$, a bounded domain Ω and a distributed control

law $g(u)$ supported in Ω, such that for any initial value u_0 the solution of the control problem

$$\begin{cases} \partial_t u - \sigma \Delta u = f(u) + g(u) \mathbb{1}_{[0,T] \times \Omega}, & ; \quad t > 0, \ x \in \mathbb{R}^d, \\ u(0, x) = u_0(x), & ; \quad x \in \mathbb{R}^d. \end{cases} \tag{2.3}$$

satisfies $u(t, x) \to 1$ when $t \to +\infty$, for any x in $\mathbb{R}^d$.

In practice, the process of infecting mosquitoes manually can be laborious and expensive; so it is usual that institutions has a limited amount of resource and it would be suitable to know which is the best way to use it. If we assume that we have a fixed mass of infected mosquitoes to be released on the environment, it is crucial to find out how to distribute them in order to maximize the effect of this infected founding population after some time T, see for example the works [3] and [2].

2.2 Problem formulation and main result

We will consider a bounded, smooth, connected domain Ω, and we make the following standard assumptions on the reaction term f

(H1) $f \in C^1(\Omega)$,

(H2) f' Lipschitz-continuous,

(H3) $f(0) = f(1) = 0$.

Under the above assumptions, the eq. (2.2) has a unique solution $u(t, x)$ and it is such that $0 \le u(t, x) \le 1$, so we can define the operator $\mathcal{J}_T : \mathcal{A} \subset L^1(\Omega) \to \mathbf{R}$ in the following way

$$\mathcal{J}_T(u_0) = \int_\Omega u(T, x) dx, \tag{2.4}$$

where u is the solution of eq. (2.2). We can now formulate our main result,

Theorem 8 *Let Ω be a bounded domain and let f satisfy the hypothesis (H1), (H2) and (H3). Then there exists $\overline{u}_0 \in \mathcal{A}_m$ such that*

$$\max_{u_0 \in \mathcal{A}_m} \mathcal{J}_T(u_0) = \mathcal{J}_T(\overline{u}_0). \tag{2.5}$$

Moreover, setting $\overline{u}$ the solution of eq. (2.2) associated with this optimum initial data and $\overline{p}$ the unique solution of eq. (2.6)

$$\begin{cases} -\partial_t \overline{p} - \sigma \Delta \overline{p} = f'(\overline{u})\overline{p} & \text{in } (0, T) \times \Omega, \\[2mm] \overline{p}(T, x) = 1 & \text{in } \Omega, \\[2mm] \frac{\partial \overline{p}}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega, \end{cases} \tag{2.6}$$

then there exists a non-negative real value noted by c such that

   i) *if $0 < \overline{u}_0(x) < 1$ then $\overline{p}(0, x) = c$,*

   ii) *if $\overline{u}_0(x) = 0$ then $\overline{p}(0, x) \le c$,*

   iii) *if $\overline{u}_0(x) = 1$ then $\overline{p}(0, x) \ge c$.*

The existence of such a maximizer $\overline{u}_0(x)$ corresponds with the best possible way to distribute a fixed initial mass m in a bounded domain Ω in order to maximize the total mass at $t = T$. Any way, the issue of uniqueness is still an open problem.

The second part of the Theorem 8 give us some useful information regarding the profile of an optimal initial data; in fact it implies that any optimizer can be written as

$$\overline{u}_0(x) = \mathbb{1}_{\{\overline{p}_0(x)>c\}} + \gamma(x)\mathbb{1}_{\{\overline{p}_0(x)=c\}}, \tag{2.7}$$

with $0 \leq \gamma(x) \leq 1$. In particular if the adjoint state $\overline{p}_0(x)$ is not constant in any subset of Ω, then the optimum are $\overline{u}_0(x) = \mathbb{1}_{\{\overline{p}_0(x)>c\}}$. In the section 2.5 we will see that this result allows us to define a numerical algorithm to approximate a local maximum of $\mathcal{J}_T$.

2.3 Proof of Theorem 8

We first state some results concerning the regularity of any solution of eq. (2.2), the functional $\mathcal{J}$ and the adjoint state p, that we will more generally define as the solution of the equation

$$\begin{cases} -\partial_t p - \sigma\Delta p = f'(u)p & \text{in } (0,T) \times \Omega, \\[2mm] p(T,x) = 1 & \text{in } \Omega, \\[2mm] \frac{\partial p}{\partial \nu}(t,x) = 0 & \text{for all } t \in (0,T), \text{ for all } x \in \partial\Omega, \end{cases} \tag{2.8}$$

for any u solution of eq. (2.2).

Lemma 9 *Under the hypothesis (H1), (H2) and (H3) stated above on f, the solution $u = u(t,x)$ of eq. (2.2) satisfies the following estimates:*

1. $0 \leq u(t,x) \leq 1$, a.e. in $[0,T] \times \Omega$,

2. $u \in L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$,

3. $\partial_t u \in L^2(0,T;H^{-1}(\Omega))$.

Proof: The first assertion is a straightforward consequence of the maximum principle and the properties of f. In fact, since $0 \leq u_0(x) \leq 1$ and $f(0) = f(1) = 0$ we have that $U = 1$ is a super-solution and $U = 0$ is a sub-solution, so we get the result.

In order to prove the other two estimates let us multiply the eq. (2.2) by u and integrate on Ω. We obtain

$$\frac{1}{2}\partial_t \int_\Omega u^2(t,x)dx + \sigma \int_\Omega (\nabla_x u(t,x))^2 dx = \int_\Omega u(t,x)f(u(t,x))dx; \tag{2.9}$$

thus, choosing M such that $|f'(u)| \leq M, \forall u \in \mathcal{A}$ we have

$$\frac{1}{2}\partial_t \|u(t)\|^2_{L^2(\Omega)} \leq \int_\Omega u(t,x)f(u(t,x))dx \leq M\|u(t)\|^2_{L^2(\Omega)}. \tag{2.10}$$

By the Gronwall inequality we get

$$\|u(t)\|^2_{L^2(\Omega)} \leq e^{2Mt}\|u_0\|^2_{L^2(\Omega)} \leq |\Omega|e^{2MT}. \tag{2.11}$$

Note that the constant in the right-hand side is independent of t, so we have actually proved that $u \in L^\infty(0, T; L^2(\Omega))$.

Integrating eq. (2.9) in time we get

$$\frac{1}{2} \int_\Omega (u)^2(T, x) dx - \frac{1}{2} \int_\Omega (u_0)^2(x) dx + \sigma \int_0^T \int_\Omega (\nabla_x u(t, x))^2 dx = \int_0^T \int_\Omega u(t, x) f(u(t, x)) dx \quad (2.12)$$

and so it yields

$$\sigma \|\nabla_x u\|_{L^2(0, T; L^2(\Omega))}^2 = \int_0^T \int_\Omega u(t, x) f(u(t, x)) dx + \frac{1}{2} \|u_0\|_{L^2(\Omega)}^2 - \frac{1}{2} \|u(T, \cdot)\|_{L^2(\Omega)}^2$$

$$\leq M \|u\|_{L^\infty(0, T; L^2(\Omega))} + \|u_0\|_{L^2(\Omega)}^2.$$

Finally, choosing $v \in H^1(\Omega)$ such that $\|v\|_{H^1(\Omega)} \leq 1$ and multiplying and integrating again in eq. (2.2) it holds

$$\int_\Omega \partial_t u(t, x) v(x) dx + \sigma \int_\Omega \nabla_x u(t, x) \nabla_x v(x) dx = \int_\Omega f(u(t, x)) v(x) dx \quad (2.13)$$

from which we can deduce that

$$\|\partial_t u(t, \cdot)\|_{H^{-1}(\Omega)} \leq M \|u\|_{L^2(\Omega)}^2 + |\sigma| \|\nabla_x u\|_{L^2(\Omega)}^2 \quad (2.14)$$

and therefore, thanks to the estimate on the L^2 norm of $\nabla_x u$ we obtain $\partial_t u \in L^2(0, T; H^{-1}(\Omega))$. $\square$

Lemma 10 *The operator $\mathcal{J}_T$ defined in (2.4) is differentiable. Furthermore, if for any $u_0 \in L^1(\Omega)$ we consider the solution u of eq. (2.2) and the unique solution p of eq. (2.8), it holds*

$$\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_\Omega h_0(x) p(0, x) dx, \quad (2.15)$$

for any increment $h_0 \in L^2(\Omega)$ such that $u_0 + \varepsilon h_0$ remains in the admissible set $\mathcal{A}_m$ for $|\varepsilon|$ small enough.

Moreover, defining h as the solution of the following equation

$$\begin{cases} \partial_t h - \sigma \Delta h = f'(u) h & \text{in } (0, T) \times \Omega, \\ h(0, x) = h_0(x) & \text{in } \Omega, \\ \frac{\partial h}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega, \end{cases} \quad (2.16)$$

it holds

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \int_0^T \int_\Omega f''(u(t, x)) p(t, x) h^2(t, x) \, dx \, dt, \quad (2.17)$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $L^2(\Omega)$.

Proof: Let $h_0(x)$ be defined over Ω such that $h_0 \in L^2(\Omega)$ and $u_0 + \varepsilon h_0$ is admissible, that is $0 \leq u_0 + \varepsilon h_0 \leq 1$ for any $|\varepsilon|$ small enough and $\int_\Omega h_0(x) dx = 0$. Then there exist h_ε such that the solution v_ε of eq. (2.2) with initial condition $v_\varepsilon(0, x) = u_0 + \varepsilon h_0$ can be written as $v_\varepsilon = u + \varepsilon h_\varepsilon$, where h_ε is the unique solution of eq. (2.18)

$$\begin{cases} \partial_t h_\varepsilon - \sigma \Delta h_\varepsilon = \frac{1}{\varepsilon} \left(f(u + \varepsilon h_\varepsilon) - f(u) \right) & \text{in } (0, T) \times \Omega, \\ h_\varepsilon(0, x) = h_0(x) & \text{in } \Omega, \\ \frac{\partial h_\varepsilon}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega. \end{cases} \quad (2.18)$$

The Gateaux derivative of $\mathcal{J}_T$ is written

$$
\begin{aligned}
\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle &= \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \frac{\mathcal{J}_T(v_\varepsilon(0, \cdot)) - \mathcal{J}_T(u_0)}{\varepsilon} = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \frac{\int_\Omega v_\varepsilon(T, x)dx - \int_\Omega u(T, x)dx}{\varepsilon} \\
&= \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \frac{\int_\Omega (u(T, x) + \varepsilon h_\varepsilon(T, x))\, dx - \int_\Omega u(T, x)dx}{\varepsilon} \\
&= \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \int_\Omega h_\varepsilon(T, x)dx.
\end{aligned}
\tag{2.19}
$$

By the Lipschitz continuity of f there exists a positive constant M such that

$$
\left| \frac{f(u + \varepsilon h_\varepsilon) - f(u)}{\varepsilon} \right| \leq M|h_\varepsilon|, \quad \forall \varepsilon > 0,
\tag{2.20}
$$

then the Gronwall inequality, applied as in the proof of the first assertion in Lemma 9, implies that $h_\varepsilon \in L^\infty(0, T; L^2(\Omega))$. As a consequence, when $\varepsilon \to 0$ it is possible to extract a subsequence $\varepsilon_n \to 0$ such that $h_{\varepsilon_n} \overset{*}{\rightharpoonup} h$ in $L^\infty(0, T; L^2(\Omega))$ and after possible another extraction it satisfies that $h_{\varepsilon_n}(T, x) \rightharpoonup h(T, x)$ weakly in $L^2(\Omega)$. We can then conclude that the Gateaux derivative of $\mathcal{J}_T$ can be written

$$
\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_\Omega h(T, x)dx,
\tag{2.21}
$$

where h is also the unique solution of eq. (2.16) obtained by passing in to the limit in the weak formulation of eq. (2.18).

If we show the continuity of this operator $u_0 \mapsto \nabla \mathcal{J}_T(u_0)$, the differentiability of $\mathcal{J}_T$ follows. Let h_w, h_v be the solution to eq. (2.16) for $u = w, v$ respective solutions of eq. (2.2) with w_0, v_0 as initial conditions, it is then easy to check that

$$
|\langle \nabla \mathcal{J}_T(w_0), h_0 \rangle - \langle \nabla \mathcal{J}_T(v_0), h_0 \rangle| = |\int_\Omega (h_w - h_v)(T, x)dx| \leq \sqrt{|\Omega|}\|(h_w - h_v)(T, \cdot)\|_{L^2(\Omega)}. \tag{2.22}
$$

Multiplying the equation on $h_w - h_v$ by $h_w - h_v$ and integrating on Ω we get

$$
\frac{1}{2}\partial_t \int_\Omega (h_w - h_v)^2(t, x)dx + \sigma \int_\Omega (\nabla_x(h_w - h_v))^2(t, x)dx = \int_\Omega (f'(w)h_w - f'(v)h_v)(h_w - h_v)(t, x)dx
$$

Noting $\delta_{w,v}(t, x) = (w - v)(t, x)$ it holds that

$$
\begin{aligned}
\frac{1}{2}\partial_t \int_\Omega (h_w - h_v)^2(t, x)dx &\leq \int_\Omega (f'(\delta_{w,v} + v)h_w - f'(v)h_v)(h_w - h_v)(t, x)dx \\
&\leq \int_\Omega ((C\delta_{w,v} + f'(v))h_w - f'(v)h_v)(h_w - h_v)(t, x)dx \\
&\leq C \int_\Omega h_w \delta_{w,v}(h_w - h_v)(t, x)dx + M \int_\Omega (h_w - h_v)^2(t, x)dx;
\end{aligned}
\tag{2.23}
$$

the second and third inequalities in (2.23) follow from the regularity of f. The constants C and M are such that $|f''| \leq C$ and $|f'| \leq M$.

By the Gronwall Lemma, for some real $L > 0$ it holds that $\int_\Omega \delta_{w,v}^2(t, x)dx \leq L\|\delta_{w_0,v_0}\|_{L^2(\Omega)}$. Together with the Cauchy-Schwartz's inequality and the fact that h_w and h_v are bounded, this result allows us to get

$$
\partial_t \|(h_w - h_v)\|_{L^2(\Omega)}^2 \leq CL\|\delta_{w_0,v_0}\|_{L^2(\Omega)}^2 + M\|(h_w - h_v)\|_{L^2(\Omega)}^2,
\tag{2.24}
$$

from which we deduce a L^2-bound to $h_w - h_v$

$$
\|(h_w - h_v)(t, x)\|_{L^2(\Omega)} \leq \frac{CL}{M}\|\delta_{w_0,v_0}\|_{L^2(\Omega)}e^{Mt}.
\tag{2.25}
$$

Combining (2.25) and (2.22) yields the continuity of $u_0 \mapsto \nabla \mathcal{J}_T(u_0)$ and hence the differentiability of $\mathcal{J}_T$ in a larger sense,

$$|\langle \nabla \mathcal{J}_T(w_0), h_0 \rangle - \langle \nabla \mathcal{J}_T(v_0), h_0 \rangle| \leq \sqrt{|\Omega|} \frac{CL}{M} \|\delta_{w_0, v_0}\|_{L^2(\Omega)} e^{MT}. \tag{2.26}$$

Multiplying eq. (2.16) by the solution p of eq. (2.8) with $\overline{u} = u$ and integrating by parts we can rewrite (2.21) as

$$\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_\Omega h(T, x) dx = \int_\Omega h_0(x) p(0, x) dx \tag{2.27}$$

and consequently $\nabla \mathcal{J}_T(u_0) = p(0, x)$; this p is often called the adjoint state of h.

Let us now find an expression for the second order derivative of $\mathcal{J}_T$. We set $v_\varepsilon = u + \varepsilon h + \frac{\varepsilon^2}{2} k_\varepsilon$ and we write the differential equation satisfied by v_ε.

From the regularity hypothesis on f and the estimation (2.25) on h, v_ε and u_ε it follows that $k_\varepsilon \in L^\infty(0, T; L^2(\Omega))$. A passage to the limit when $\varepsilon \to 0$ implies, after extraction, the existence of a subsequence $k_{\varepsilon_n} \in L^\infty(0, T; L^2(\Omega))$ such that $k_{\varepsilon_n} \overset{*}{\rightharpoonup} k$ in $L^\infty(0, T; L^2(\Omega))$. By mean of a Taylor expansion we deduce the differential equation satisfied by k,

$$\begin{cases} \partial_t k - \sigma \Delta k = f'(u)k + f''(u)h & \text{in } (0, T) \times \Omega, \\[2mm] k(0, x) = 0 & \text{in } \Omega, \\[2mm] \frac{\partial k}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial \Omega. \end{cases} \tag{2.28}$$

An analysis similar to that yielding (2.27) shows that

$$\int_\Omega k(T, x) dx = \int_0^T \int_\Omega f''(u) h^2 p \, dx \, dt, \tag{2.29}$$

in this case we multiply eq. (2.28) by p and we integrate in space and time. Finally, we have proved that when $\varepsilon \to 0$, it holds

$$\mathcal{J}_T(u_0 + \varepsilon h_0) = \int_\Omega u(T, x) dx + \varepsilon \int_\Omega h(T, x) dx + \frac{\varepsilon^2}{2} \int_\Omega k(T, x) dx + o(\epsilon^2)$$

$$= \mathcal{J}_T(u_0) + \varepsilon \langle \nabla \mathcal{J}_T(u_0), h \rangle + \frac{\varepsilon^2}{2} \int_0^T \int_\Omega f''(u) h^2 p \, dx \, dt + o(\epsilon^2),$$

which establishes the formula

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \int_0^T \int_\Omega f''(u(t, x)) h^2(t, x) p(t, x) \, dx \, dt. \tag{2.30}$$

$\square$

Lemma 11 *Let Ω be a bounded domain, then the solution p of eq. (2.8) is such that*

1. *$\partial_t p, \partial_{tt} p, \nabla p, \nabla^2 p, \nabla(\partial_t p), \nabla^2(\partial_t p) \in L^q_{loc}\big((0, T) \times \Omega\big)$ for all $1 \leq q < \infty$,*

2. *$\forall c > 0$, for almost every $x \in \{p(0, \cdot) = c\}$, one has*

$$f'\big(u(0, x)\big) = -\partial_t p(0, x)/p(0, x).$$

Proof: Let us start by proving that p is bounded. In fact, if we consider $\hat{p}$ the solution of the ordinary differential equation

$$\begin{cases} -\frac{d\hat{p}}{dt} = M\hat{p} & \text{in} \quad (0,T), \\ \\ \hat{p}(T) = 1, \end{cases} \tag{2.31}$$

with $M = \|f'\|_{L^\infty}$, it is clear that $0 \le p(t,x) \le \hat{p}(t)$, which is a consequence of the maximum principle. Since we know explicitly that $\hat{p} = e^{M(T-t)} \in L^\infty(0,T)$, then it holds

$$p \in L^\infty(0,T;L^\infty(\Omega)) \tag{2.32}$$

Second, we know from the classical L^q regularity theory for parabolic equations (see for instance Theorem 9.1, in chapter IV of [50]) that, as $0 \le u \le 1$, one has $\partial_t p, \nabla p, \nabla^2 p \in L^q_{loc}\big((0,T) \times \Omega\big)$ for all $1 \le q < \infty$. Similarly, one has $\partial_t u, \nabla u, \nabla^2 u \in L^q_{loc}\big((0,T) \times \Omega\big)$ for all $1 \le q < \infty$.

Next, let us define $\phi := \partial_t p$, deriving on eq. (2.8) we obtain that it is the only solution of the equation

$$\begin{cases} -\partial_t \phi - \Delta\phi = G, & \text{in } (0,T) \times \Omega, \\ \\ \phi(T,x) = 0, & \text{in } \Omega, \\ \\ \frac{\partial\phi}{\partial\nu}(t,x) = 0 & \text{for all } t \in (0,T), \text{ for all } x \in \partial\Omega, \end{cases} \tag{2.33}$$

where $G = f''(u)\partial_t u\, p + f'(u)\phi$. Due to the previous estimates, one has $G \in L^q_{loc}\big((0,T) \times \Omega\big)$ for all $1 \le q < \infty$. Hence, again the L^q regularity theory for parabolic equations yields $\partial_t \phi, \nabla\phi, \nabla^2\phi \in L^q_{loc}\big((0,T) \times \Omega\big)$ for all $1 \le q < \infty$. This means in particular that $\partial_t p \in L^q_{loc}\big((0,T), W^{2,q}_{loc}(\Omega)\big)$ for all $1 \le q < \infty$. Taking q large enough, the Morrey inequality thus yields $p \in C^{0,\alpha}_{loc}\big((0,T), W^{2,q}_{loc}(\Omega)\big)$ for all $1 \le q < \infty$ and $\alpha \in (0,1)$.

Now, as $p(0,\cdot) \in W^{2,1}_{loc}(\Omega)$, we know (see for example [34]) that for almost every $x \in \{p(0,\cdot) = c\}$, one has $\Delta p(0,x) = 0$. Moreover, as $\partial_{tt} p \in L^q_{loc}\big((0,T) \times \Omega\big)$ for all $1 \le q < \infty$, one has $\partial_t p \in C^{0,\alpha}_{loc}\big((0,T), L^q_{loc}(\Omega)\big)$ and, in particular, $\partial_t p(0,\cdot) \in L^q_{loc}(\Omega)$. We eventually derive from eq. (2.8) that for almost every $x \in \{p(0,\cdot) = c\}$, one has

$$-\partial_t p(0,x) = f'\big(u(0,x)\big)p(0,x).$$

$\square$

Now we proceed with proof of Theorem 8.

Proof: This proof falls naturally into two parts, first we set the existence of a maximal element and then we characterize it.

Step 1: Existence of a maximal element

The basic idea of this part of the proof is to establish the existence of a supremum element in the set $\{\mathcal{J}_T(u_0) : u_0 \in \mathcal{A}_m\}$ and then to show that it is reached for some element $\overline{u}_0 \in \mathcal{A}_m$ defined as the limit of a maximizing sequence in $\mathcal{A}_m$.

From the first estimate on Lemma 9 it follows that $\mathcal{J}_T$ is bounded. We note also that $\mathcal{J}_T$ is a continuous operator thanks to the results in Lemma 10,

$$|\mathcal{J}_T(u_0) - \mathcal{J}_T(v_0)| \le |\Omega|^{\frac{1}{2}} e^{MT} \|u_0 - v_0\|_{L^2(\Omega)}. \tag{2.34}$$

It must exist, therefore, a supremum element in the set of images of $\mathcal{J}_T$, and so a maximizing sequence u_0^n in $\mathcal{A}_m$, which means

$$\lim_n \mathcal{J}_T(u_0^n) = \sup_{\{\mathcal{A}_m\}} \mathcal{J}_T(u_0). \tag{2.35}$$

Since $0 \le u_0^n(x) \le 1$, it is clear that $u_0^n \in L^\infty(\Omega)$ and so after an extraction, we can state that $u_0^n \overset{*}{\rightharpoonup} \overline{u}_0$ weakly in $L^\infty(\Omega)$, for some $\overline{u}_0 \in L^\infty(\Omega)$ i.e.

$$\int_\Omega u_0^n(x)\varphi(x)dx \to \int_\Omega \overline{u}_0(x)\varphi(x)dx, \quad \forall \varphi \in L^1(\Omega). \tag{2.36}$$

Choosing $\varphi = 1$ in (2.36), we get that $\overline{u}_0$ is still in $\mathcal{A}_m$.

Now, to each $u_0^n, n = 1, 2, \dots$ we can associate the solution of eq. (2.2) with initial datum u_0^n

$$\begin{cases} \partial_t u^n - \sigma \Delta u^n = f(u^n) & \text{in } (0, T) \times \Omega, \\[2mm] u^n(0, x) = u_0^n(x) & \text{in } \Omega, \\[2mm] \frac{\partial u^n}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega, \end{cases} \tag{2.37}$$

in a weak sense this means that $\forall \varphi \in C^\infty(0, T) \times \Omega$ the following holds

$$\langle u^n(T), \varphi(T) \rangle - \int_0^T \langle u^n(t), \partial_t\varphi(t) \rangle dt - \sigma \int_0^T \langle u^n(t), \Delta\varphi(t) \rangle = \int_0^T \langle f(u^n(t)), \varphi(t) \rangle dt + \langle u_0^n, \varphi(0) \rangle. \tag{2.38}$$

Thanks to the first assertion in Lemma 9, we can deduce that $u^n(T, x) \in L^2(\Omega)$ for all $n = 1, 2, \dots$ and consequently the existence of an element $\widetilde{u} \in L^2(\Omega)$ such that, after extraction, $u^n(T, x) \rightharpoonup \widetilde{u}$ in $L^2(\Omega)$, i.e.

$$\langle u^n(T), \varphi \rangle \to \langle \widetilde{u}, \varphi \rangle, \quad \forall \varphi \in L^2(\Omega). \tag{2.39}$$

Since $u^n \in L^\infty(0, T; L^2(\Omega))$, after possibly another extraction, there exist $\overline{U} \in L^\infty(0, T; L^2(\Omega))$ such that $u^n \overset{*}{\rightharpoonup} \overline{U}$ in $L^\infty(0, T; L^2(\Omega))$, i.e.

$$\int_0^T \langle u^n(t), \varphi(t) \rangle dt \longrightarrow \int_0^T \langle \overline{U}(t), \varphi(t) \rangle dt, \quad \forall \varphi \in L^\infty(0, T; L^2(\Omega)). \tag{2.40}$$

Again, from the second assertion in Lemma 9, it follows the existence of a subsequence still noted by $\partial_t u^n$ and $v \in L^2(0, T; H^{-1}(\Omega))$ such that $\partial_t u^n \overset{*}{\rightharpoonup} v$ in $L^2(0, T; H^{-1}(\Omega))$, i.e.

$$\int_0^T \langle \partial_t u^n(t), \varphi(t) \rangle dt \to \int_0^T \langle v(t), \varphi(t) \rangle dt \quad \forall \varphi \in L^2(0, T; H^1(\Omega)). \tag{2.41}$$

We can easily prove that $\partial_t \overline{U} = v$. In fact, from the weak definition of partial derivative the following equality must holds for all $\varphi \in C_c^\infty(0, T) \times \Omega$

$$\int_0^T \langle \partial_t u^n(t), \varphi(t) \rangle dt = -\int_0^T \langle u^n(t), \partial_t\varphi(t) \rangle dt; \tag{2.42}$$

a simple passage to the limit implies the desired result and consequently that $\overline{U} \in H^1(0, T; L^2(\Omega))$,

$$\int_0^T \langle v(t), \varphi(t) \rangle dt = -\int_0^T \langle \overline{U}(t), \partial_t\varphi(t) \rangle dt. \tag{2.43}$$

Now choosing $\varphi \in H^1(0, T; L^2(\Omega))$ such that $\varphi(0, x) = 0$ for all $x \in \Omega$, after an integration by parts we get

$$\int_0^T \langle \partial_t u^n(t), \varphi(t) \rangle dt = \langle u^n(T), \varphi(T) \rangle - \int_0^T \langle u^n(t), \partial_t\varphi(t) \rangle dt \tag{2.44}$$

and so passing to the limit and integrating by parts again we obtain

$$
\begin{aligned}
\int_0^T \langle v, \varphi(t)\rangle dt &= \langle \widetilde{u}, \varphi(T)\rangle - \int_0^T \langle \overline{U}, \partial_t\varphi(t)\rangle dt \\
&= \langle \widetilde{u}, \varphi(T)\rangle - \langle \overline{U}(T), \varphi(T)\rangle + \int_0^T \langle \partial_t\overline{U}, \varphi\rangle dt.
\end{aligned}
$$

This equality together with (2.43) implies that $\widetilde{u}(x) = \overline{U}(T, x)$ almost everywhere in Ω. Similarly choosing φ adequately we prove that $\overline{u}_0(x) = \overline{U}(0, x)$ almost everywhere in Ω.

Finally, let us define the set

$$
W := \{u \in L^2([0, T]; H^1(\Omega)) \mid \partial_t u \in L^2([0, T]; H^{-1}(\Omega))\}.
$$

The estimates in Lemma 9 imply that u^n is bounded in W for every element of the subsequence. Moreover, thanks to the Aubin-Lions lemma [55], the set W embeds compactly into $L^2([0, T]; L^2(\Omega))$ which ensures the existence of a subsequence still denoted as u^n which is Cauchy in $L^2([0, T]; L^2(\Omega))$. Then necessarily $u^n \to \overline{U}$ strongly in $L^2([0, T]; L^2(\Omega))$ and thus

$$
\int_0^T \langle f(u^n(t)), \varphi\rangle dt \to \int_0^T \langle f(\overline{U}(t)), \varphi\rangle dt, \tag{2.45}
$$

which follows from the Lipschitz continuity of f and the Cauchy-Schwartz inequality.

Now we can pass to the limit in (2.38), gathering (2.36), (2.39) and (2.45) to obtain that $\forall \varphi \in C^\infty(0, T) \times \Omega$ it holds that

$$
\langle \overline{U}(T), \varphi(T)\rangle - \int_0^T \langle \overline{U}(t), \partial_t\varphi(t)\rangle dt - \int_0^T \langle \overline{U}(t), \Delta\varphi(t)\rangle = \int_0^T \langle f(\overline{U}(t)), \varphi(t)\rangle dt + \langle \overline{u}_0, \varphi(0)\rangle, \tag{2.46}
$$

which means that $\overline{U}$ is a weak solution to the problem

$$
\begin{cases}
\partial_t\overline{U} - \sigma\Delta\overline{U} = f(\overline{U}) & \text{in } (0, T) \times \Omega, \\[2mm]
\overline{U}(0, x) = \overline{u}_0(x) & \text{in } \Omega, \\[2mm]
\frac{\partial\overline{U}}{\partial\nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega.
\end{cases} \tag{2.47}
$$

Now, we get the following equalities from (2.35) and (2.39) choosing $\varphi = 1$

$$
\sup_{\{\mathcal{A}_m\}} \mathcal{J}_T(u) = \lim_n \mathcal{J}_T(u_0^n) = \lim_n \int_\Omega u^n(T, x)dx = \int_\Omega \overline{U}(T, x)dx = \mathcal{J}_T(\overline{u}_0), \tag{2.48}
$$

which in fact means that $\overline{u}_0$ is a maximizing element of $\mathcal{J}_T$ in $\mathcal{A}_m$.

Step 2: Characterization of the maximal element

We first prove (i).

Let μ be the Lebesgue measure. We define the set $\mathcal{S} = \{x \in \Omega : 0 < \overline{u}_0(x) < 1\}$ and we suppose that $\mu(\mathcal{S}) \neq 0$, otherwise there exists a set E such that $\overline{u}_0 = \mathbb{1}_E$ almost everywhere. We note that $\mathcal{S}$ can be written as $\mathcal{S} = \cup_{k=1}^\infty \mathcal{S}_k$ where $\mathcal{S}_k = \{x \in \Omega : \frac{1}{k} < \overline{u}_0(x) < 1 - \frac{1}{k}\}$. Let us fix a sufficiently large k such that $\mu(\mathcal{S}_k) > 0$ and consider two points x^*, y^* in this set. For $\varepsilon \in \mathbb{R}$ and $r \in \mathbb{R}_+$, we define

$$
v_0(x) = \overline{u}_0(x) + \varepsilon\frac{\mu(B(x^*, r))}{\mu(B(x^*, r) \cap \mathcal{S}_k)}\mathbb{1}_{B(x^*, r)\cap\mathcal{S}_k} - \varepsilon\frac{\mu(B(y^*, r))}{\mu(B(y^*, r) \cap \mathcal{S}_k)}\mathbb{1}_{B(y^*, r)\cap\mathcal{S}_k}, \tag{2.49}
$$

here $\mathbb{1}_C$ is the characteristic function of the set C. Thanks to the Lebesgue Density Theorem, for almost every $x \in \mathcal{S}_k$,

$$\lim_{r \to 0^+} \frac{\mu(B(x,r))}{\mu(B(x,r) \cap \mathcal{S}_k)} \to 1. \tag{2.50}$$

In particular, we can choose r small enough and $x^*, y^* \in \mathcal{S}_k$ such that $\frac{\mu(B(x^*,r))}{\mu(B(x^*,r) \cap \mathcal{S}_k)} < 2$ and $\frac{\mu(B(y^*,r))}{\mu(B(y^*,r) \cap \mathcal{S}_k)} < 2$ and $|\varepsilon| < \frac{1}{2k}$ as well, then it is clear that $0 < v_0 < 1$ and so v_0 is still in $\mathcal{S}$. We note also that

$$\int_\Omega v_0(x) dx = \int_\Omega u_0(x) + \varepsilon \frac{\mu(B(x^*,r))}{\mu(B(x^*,r) \cap \mathcal{S}_k)} \int_\Omega \mathbb{1}_{B(x^*,r) \cap \mathcal{S}_k} dx - \varepsilon \frac{\mu(B(y^*,r))}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \int_\Omega \mathbb{1}_{B(y^*,r) \cap \mathcal{S}_k} dx$$
$$= m + \varepsilon\mu(B(x^*,r)) - \varepsilon\mu(B(y^*,r))$$
$$= m,$$

so $v_0 \in \mathcal{A}_m$. We shall now use the fact that $\overline{u}_0$ is a maximizing element in $\mathcal{A}_m$; gathering (2.21) and (2.27) we have

$$0 = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\mathcal{J}_T \left(\overline{u}_0 + \varepsilon \left(\frac{\mu(B(x^*,r))\mathbb{1}_{B(x^*,r) \cap \mathcal{S}_k}}{\mu(B(x^*,r) \cap \mathcal{S}_k)} - \frac{\mu(B(y^*,r))\mathbb{1}_{B(y^*,r) \cap \mathcal{S}_k}}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \right) \right) - \mathcal{J}_T(\overline{u}_0) \right]$$
$$= \int_\Omega \left(\frac{\mu(B(x^*,r))\mathbb{1}_{B(x^*,r) \cap \mathcal{S}_k}}{\mu(B(x^*,r) \cap \mathcal{S}_k)} - \frac{\mu(B(y^*,r))\mathbb{1}_{B(y^*,r) \cap \mathcal{S}_k}}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \right) p(0,x) dx$$
$$= \frac{\mu(B(x^*,r))}{\mu(B(x^*,r) \cap \mathcal{S}_k)} \int_{B(x^*,r) \cap \mathcal{S}_k} p(0,x) dx - \frac{\mu(B(y^*,r))}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \int_{B(y^*,r) \cap \mathcal{S}_k} p(0,x) dx.$$

Here we can multiply the whole equality by $\frac{1}{\mu(B(x^*,r))}$ to obtain

$$0 = \frac{1}{\mu(B(x^*,r) \cap \mathcal{S}_k)} \int_{B(x^*,r) \cap \mathcal{S}_k} p(0,x) dx - \frac{1}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \int_{B(y^*,r) \cap \mathcal{S}_k} p(0,x) dx$$

then, making r goes to zero we get that for almost every $x^*, y^* \in \mathcal{S}_k$,

$$0 = p(0,x^*) - p(0,y^*).$$

We have finally obtained the existence of a constant $c \in \mathbb{R}$ such that $p(0,x) = c$ almost everywhere in $\mathcal{S}_k$. The same statement holds for every k large enough, so with $k \to +\infty$ we have the result for almost every $x \in \mathcal{S}$.

Let us now prove (ii).

Lets define the set $\mathcal{S}^0 = \{x \in \Omega : \overline{u}_0(x) = 0\}$ and for every $k = 1, 2, \ldots$ the set $\mathcal{S}_k^0 = \{x \in \Omega : 0 \leq \overline{u}_0(x) < \frac{1}{k}\}$, we note that $\mathcal{S}^0 = \cap_{k=1}^\infty \mathcal{S}_k^0$. We assume that $\mu(\mathcal{S}^0) > 0$, otherwise $\overline{u}_0 > 0$ almost everywhere and we pass to (iii). Choosing $x^* \in \mathcal{S}_k^0$ and $y^* \in \mathcal{S}_k$ defined as above; r sufficiently small such that $\frac{\mu(B(x^*,r))}{\mu(B(x^*,r) \cap \mathcal{S}_k^0)} < 2$ and $\frac{\mu(B(y^*,r))}{\mu(B(y^*,r) \cap \mathcal{S}_k)} < 2$ as in (2.50) and $0 < \varepsilon < \frac{1}{2k}$, it holds

$$0 \leq v_0(x) = \overline{u}_0(x) + \varepsilon \frac{\mu(B(x^*,r))}{\mu(B(x^*,r) \cap \mathcal{S}_k^0)} \mathbb{1}_{B(x^*,r) \cap \mathcal{S}_k^0} - \varepsilon \frac{\mu(B(y^*,r))}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \mathbb{1}_{B(y^*,r) \cap \mathcal{S}_k} \leq 1,$$

and similarly to the previous case $\int_\Omega v_0(x) dx = m$, so $v_0 \in \mathcal{A}_m$. Since $\overline{u}_0$ is a maximizing element in

$\mathcal{A}_m$ and ε is strictly positive, we get

$$0 \geq \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\mathcal{J}_T \left(\overline{u}_0 + \varepsilon \left(\frac{\mu(B(x^*,r)) \mathbb{1}_{B(x^*,r) \cap \mathcal{S}_k^0}}{\mu(B(x^*,r) \cap \mathcal{S}_k^0)} - \frac{\mu(B(y^*,r)) \mathbb{1}_{B(y^*,r) \cap \mathcal{S}_k}}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \right) \right) - \mathcal{J}_T(\overline{u}_0) \right]$$

$$= \int_\Omega \left(\frac{\mu(B(x^*,r)) \mathbb{1}_{B(x^*,r) \cap \mathcal{S}_k^0}}{\mu(B(x^*,r) \cap \mathcal{S}_k^0)} - \frac{\mu(B(y^*,r)) \mathbb{1}_{B(y^*,r) \cap \mathcal{S}_k}}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \right) p(0,x) dx$$

$$= \frac{\mu(B(x^*,r))}{\mu(B(x^*,r) \cap \mathcal{S}_k^0)} \int_{B(x^*,r) \cap \mathcal{S}_k^0} p(0,x) dx - \frac{\mu(B(y^*,r))}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \int_{B(y^*,r) \cap \mathcal{S}_k} p(0,x) dx$$

again, we can multiply the inequality by $\frac{1}{\mu(B(x^*,r))} > 0$ and obtain

$$0 \geq \frac{1}{\mu(B(x^*,r) \cap \mathcal{S}_k^0)} \int_{B(x^*,r) \cap \mathcal{S}_k^0} p(0,x) dx - \frac{1}{\mu(B(y^*,r) \cap \mathcal{S}_k)} \int_{B(y^*,r) \cap \mathcal{S}_k} p(0,x) dx.$$

Passing to the limit when $r \to 0$ we get

$$0 \geq p(0,x^*) - p(0,y^*),$$

from where $c \geq p(0,x^*)$ for almost every $x^* \in \mathcal{S}_k^0$ and every k large enough. We have done the proof of (ii) making $k \to +\infty$.

Similarly, we can prove (iii). We remark that from the strong maximum principle and the Hopf's Lemma follows that $p > 0$ in $\overline{\Omega}$. Since c is in the range of p, then it must be positive as well. This way we end with the proof of the Theorem 8. $\qquad\square$

2.4 The u₀-constant case

We will restrict ourselves in this section to the study of the case where the initial mass m is distributed homogeneously over the bounded domain Ω. We thus consider $u_0 := \frac{m}{|\Omega|}$ with $0 < m < |\Omega|$, which is the only constant initial distribution that belongs in $\mathcal{A}_m$. In this case the solution of eq. (2.2) is homogeneous in space for every $t \in [0,T]$, meaning that $u(t,x) = u(t)$ for all $x \in \Omega$. More precisely u satisfies the ordinary differential equation

$$\begin{cases} \partial_t u = f(u) & \text{in } (0,T), \\ \\ u(0) = \dfrac{m}{|\Omega|}. \end{cases} \tag{2.51}$$

We also assume that the reaction term $f(u)$ satisfies (H1), (H2), (H3) and the following additional hypothesis

(H4) $\exists \rho \in [0,1]$ and $\delta > 0$ such that $\forall x \geq \rho : f(x) > 0$ and $f''(x) < -\delta; f \in C^2([0,1])$.

Proposition 12 *Let u_0 be the constant distribution defined as $u_0 := \frac{m}{|\Omega|}, \quad \forall x \in \Omega$ and f satisfying (H1)-(H3). Then the following assertions holds:*

    i.) If (H4) is satisfied and if $\frac{m}{|\Omega|} \geq \rho$, then u_0 is a local maximizer of $\mathcal{J}_T$ in the L^2-norm.

    ii.) In dimension one, if u_0 is a local maximizer of $\mathcal{J}_T$ in the L^2-norm, then $f''(u_0) \leq 0$.

Proof: As seen previously, the derivative of the target operator $\mathcal{J}_T(u_0)$ on the admissible set $\mathcal{A}_m$ writes $\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_\Omega h_0(x)p_0(x)$ for every zero mean value function $h_0 \in L^2(\Omega)$ and $p_0(x) = p(0,x)$ being the adjoint state, which is characterized by

$$
\begin{cases}
-\partial_t p = f'(u)p & \text{in } (0,T) \times \Omega, \\[2mm]
p(T,x) = 1 & \text{in } \Omega.
\end{cases} \tag{2.52}
$$

It is easy to check that $p(t,x) = f\left(u(T)\right)/f\left(u(t)\right)$, which is also homogeneous in space. Consequently, $p_0 = f\left(u(T)\right)/f\left(\frac{m}{|\Omega|}\right)$ is constant and $\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = p_0 \int_\Omega h_0(x)dx = 0$ which means that $u_0 := \frac{m}{|\Omega|}$ is a critical point.

Let us now to check that, provided that the initial mass is large enough, the second order optimality conditions on this critical point are satisfied. We suppose that

$$
u_0 := \frac{m}{|\Omega|} > \rho, \tag{2.53}
$$

then, since $f(u_0)$ is positive, $u(t)$ stay increasing in time implying that $\rho < u(t) < 1$ for every $t > 0$ and consequently from (H4) we get $f''(u(t)) < -\delta$. Besides, from eq. (2.52) follows that $p(t) \geq e^{-M(T-t)}$ where M is such that $f'(u(t)) \geq M, \forall t > 0$. Gathering those estimates we obtain

$$
\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \int_0^T \int_\Omega f''(u(t))p(t)h^2(t,x)dxdt \leq -\delta \int_0^T e^{-M(T-t)} \int_\Omega h^2(t,x)dx\, dt. \tag{2.54}
$$

As shown in a previous section, $h(t,x)$ satisfies the equation

$$
\begin{cases}
\partial_t h - \sigma \Delta h = f'(u)h & \text{in } (0,T) \times \Omega, \\[2mm]
h(0,x) = h_0(x) & \text{in } \Omega,
\end{cases} \tag{2.55}
$$

from which we can deduce $\|h(t)\|_{L^2(\Omega)}^2 \leq \|h_0\|_{L^2(\Omega)}^2 e^{2Mt}$.

Finally, for a certain positive constant C depending only on δ, M and T, it holds that

$$
\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle \leq -C\|h_0\|_{L^2(\Omega)}^2, \tag{2.56}
$$

which ensures that the second order optimality conditions on this critical point are fulfilled and concludes the proof of the first assertion.

Note that in this case, the constant c derived from the Theorem 8 is necessarily $c \equiv p_0$ otherwise $u_0(x)$ is either null or totally saturated over the domain Ω which would imply that $u_0 \notin \mathcal{A}_m$. Hence, the set $\{p_0 = c\}$ coincides with the whole domain Ω.

Let us now show the second part of the Proposition 12. We suppose that u_0 is a local maximizer in the L^2-norm, then for any sufficiently small perturbation $h_0(x)$ the second order optimality condition holds, i.e.

$$
\langle \nabla^2 \mathcal{J}_T(u_0), h^k(x) \rangle \leq 0. \tag{2.57}
$$

In particular, we consider $h_0^k(x) = \cos(kx)$, $k = 1, 2, \ldots$, for the sake of simplicity $\Omega = (0, \pi)$ and we assume a diffusion coefficient $\sigma = 1$. Then we can explicitly calculate the second order derivative of our target operator $\mathcal{J}_T$, which depends on the solution $h^k(t,x)$ of the differential equation

$$
\begin{cases}
\partial_t h^k - \Delta h^k = f'(u)h^k & \text{in } (0,T) \times (0,\pi), \\[2mm]
h^k(0,x) = h_0^k(x) = \cos(kx) & \text{in } \Omega, \\[2mm]
\frac{\partial h^k}{\partial \nu}(t,x) = 0 & \text{for all } t \in (0,T), \text{ for all } x \in \{0, \pi\}.
\end{cases} \tag{2.58}
$$

The solution of eq. (2.58) is explicitly given by

$$h^k(t,x) = e^{-k^2 t} \cos(kx) \frac{f(u(t))}{f(u_0)} \qquad (2.59)$$

and consequently from (2.17) and thanks to the Laplace method it follows that

$$\left\langle \nabla^2 \mathcal{J}_T(u_0), h_0^k \right\rangle \quad = \quad \frac{f(u(T))}{f^2(u_0)} \int_0^T f''(u(t)) f(u(t)) e^{-2k^2 t} \int_0^\pi \cos^2(kx) dx\, dt \qquad (2.60)$$

$$= \quad \frac{\pi}{2} \frac{f(u(T))}{f^2(u_0)} \int_0^T f''(u(t)) f(u(t)) e^{-2k^2 t} dt \qquad (2.61)$$

$$\underset{k \to \infty}{\sim} \quad \frac{\pi}{4} p_0 \frac{f''(u_0)}{k^2}. \qquad (2.62)$$

Gathering (2.62) and (2.57) we get that necessarily $f''(u_0) \leq 0$, which completes the prove. □

2.4.1 The case of a concave non-linearity

In this section we consider concave non-linearities. We have in mind in particular the well known Fisher-KPP equation where the reaction term $f(u) = ru(1-u)$ and the system is monostable. We remark the fact that this particular f satisfies (H1)-(H4) so the Proposition 12 applies for homogeneously distributed initial data. In what follows we will prove that a constant initial distribution is in fact the optimal distribution.

We start by showing that the functional $\mathcal{J}_T(u_0)$ inherits the concavity from the reaction term. In a general framework we have the following result :

Proposition 13 *Let u be the solution of eq. (2.2). If the reaction term $f(u)$ is concave, then the functional $\mathcal{J}_T(u_0)$ defined by (2.4) is also concave.*

Proof: Let $\alpha, \beta \in [0,1]$ be such that $\alpha + \beta = 1$, if we call u the solution of eq. (2.2) with initial condition $\widetilde{u}_0(x) = \alpha u_0^1(x) + \beta u_0^2(x)$ and we set $\hat{u}(t,x) = \alpha u^1(t,x) + \beta u^2(t,x)$ where u^i is the solution of eq. (2.2) with initial condition $u_0^i(x)$, $i \in \{1,2\}$, because of the concavity of f we have

$$\partial_t \hat{u} - \Delta \hat{u} = \alpha f(u^1) + \beta f(u^2) \leq f\left(\alpha u^1 + \beta u^2\right) = f(\hat{u}), \qquad (2.63)$$

so thanks to the maximum principle $\hat{u}(t,x) \leq u(t,x)$, and therefore

$$\mathcal{J}_T\left(\alpha u_0^1 + \beta u_0^2\right) = \mathcal{J}_T(\widetilde{u}_0) = \int_\Omega u(T,x) dx \geq \int_\Omega \hat{u}(t,x) dx = \alpha \mathcal{J}_T(u_0^1) + \beta \mathcal{J}_T(u_0^2), \qquad (2.64)$$

which means that $\mathcal{J}_T$ is concave. □

This concavity property in the Fisher-KPP case ensures that if $\overline{u}_0$ is a critical point then it is a maximizer for $\mathcal{J}_T$. As straightforward consequence of the Proposition 12 we have that $\overline{u}_0(x) \equiv \frac{m}{|\Omega|}$ is a global maximum for $\mathcal{J}_T$. Explicitly, the solution can be written as

$$u(t,x) = \frac{\kappa_0 e^{rt}}{1 + \kappa_0 e^{rt}}, \quad \kappa_0 := \frac{m}{|\Omega| - m}, \quad p(t,x) = \frac{e^{r(T-t)}}{e^{2r \int_t^T u(s) ds}}. \qquad (2.65)$$

and in consequence the maximum value of the functional is $\mathcal{J}_T(\overline{u}_0) = \frac{m e^{rT}}{1 - \overline{u}_0(1 - e^{rT})}$.

2.5 Numerical Algorithm

The aim of this section is to describe an algorithm to find approximately an optimal distribution provided that the space Ω, the mass m and the time T are prescribed. In order to achieve this goal, the first order optimality conditions (*i*)-(*iii*) on Theorem 8 will be crucial. The strategy, which is basically inspired by gradient descent optimization algorithms, will be to find a maximizing sequence $u_0^1, u_0^2, u_0^3 \ldots$ which converges to the optimal element $\overline{u}_0(x)$.

From now on we shall make the assumption that $\Omega \subset \mathbb{R}$ is an interval. Let us recall that the question we study can be seen as an optimization problem under constraints

$$\max_{u_0 \in \mathcal{A}} \quad \mathcal{J}_T(u_0) \tag{2.66}$$

$$s.t. \quad \int_\Omega u_0(x)dx = m. \tag{2.67}$$

We can then define the Lagrangian function

$$\mathcal{L}(u_0, \lambda) := \mathcal{J}_T(u_0) - \lambda \left(\int_\Omega u_0(x)dx - m \right),$$

and the associated problem

$$\max_{u_0 \in \mathcal{A}} \min_{\lambda \in \mathbb{R}_+} \mathcal{L}(u_0, \lambda) \tag{2.68}$$

where $\lambda \in \mathbb{R}_+$ is the Lagrangian multiplier.

As already proved, $\mathcal{J}_T$ is differentiable so $\mathcal{L}$ is also differentiable, therefore any critical point $(\overline{u}, \overline{\lambda})$ must satisfy

$$\left\langle \partial_{u_0} \mathcal{L}(\overline{u}_0, \overline{\lambda}), h_0 \right\rangle = \int_\Omega \left(\overline{p}_0(x) - \overline{\lambda} \right) h_0(x)dx \geq 0, \quad \forall h_0 \in L^2(\Omega) : \overline{u}_0 + h_0 \in \mathcal{A}, \tag{2.69}$$

$$\partial_\lambda \mathcal{L}(\overline{u}_0, \overline{\lambda}) = \int_\Omega \overline{u}_0(x)dx - m = 0, \tag{2.70}$$

where $\overline{p}_0(x) = \overline{p}(0, x)$ is the solution of eq. (2.8) whose reaction term depends on $\overline{u}(t, x)$, which solves eq. (2.2) with $\overline{u}_0(x)$ as initial datum.

We choose as the first element $u_0^0(x)$ a single block of mass m with maximal density, symmetrically distributed in Ω. Let us precise now how to define $u_0^{n+1}(x)$ from $u_0^n(x)$ at the *n-th* iteration of our algorithm. Numerically solving the differential equation (2.2) with $u_0(x) = u_0^n(x)$ we can compute $u^n(t, x)$ and then $p^n(t, x)$ as the solution of

$$\begin{cases} -\partial_t p^n - \sigma \Delta p^n = f'(u^n)p^n & \text{in } (0, T) \times \Omega, \\[2mm] p^n(T, x) = 1 & \text{in } \Omega, \\[2mm] \frac{\partial p^n}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega. \end{cases} \tag{2.71}$$

Let us assume that the new element will be set as $u_0^{n+1}(x) = u_0^n(x) + h_0^n(x)$ for some optimal increment $h_0^n(x)$. It is suitable to increase the value $\mathcal{L}(u_0^n)$; we should then move into the positive gradient direction, that is

$$\left\langle \partial_{u_0} \mathcal{L}(u_0^n, \lambda), h_0^n \right\rangle = \int_\Omega (p_0^n(x) - \lambda)h_0^n(x)dx \geq 0. \tag{2.72}$$

For some value λ^n, that we will clarify later, if $p_0^n(x) > \lambda^n$ then it is necessary that $h_0^n(x)$ be positive and as large as possible, in order to improve as much as possible the current value of the functional $\mathcal{J}_T$.

Since $0 \leq u_0^{n+1}(x) \leq 1$, the largest $h_0^n(x)$ is such that $u_0^{n+1}(x) = 1$. Analogously, if $p_0^n(x) < \lambda^n$ then $h_0^n(x)$ need to be negative and as small as possible, thus $u_0^{n+1}(x) = 0$.

The definition of $u_0^{n+1}(x)$ in the subset $\Omega_{p,\lambda^n} := \{x \in \Omega : p_0^n(x) = \lambda^n\}$ is not a straightforward fact. We propose an alternative based on a discretized in time version of eq. (2.71) in the set $\Omega_{p,\lambda}$ around $t = 0$. We use a semi-implicit finite difference scheme with time step dt

$$-\left(\frac{p^n(dt, x) - \lambda^n}{dt}\right) = f'(u_0^{n+1}(x))\lambda^n, \tag{2.73}$$

then $u_0^{n+1}(x)$ for $x \in \Omega_{p,\lambda}$ can be chosen as a solution of eq. (2.73). Depending on the form of the reaction term f, eq. (2.73) might have several roots which mean that this definition is not well posed. Numerically this issue can be overcome by considering all the roots and keeping the one with the maximal $\mathcal{J}_T$. In particular, for a bistable dynamic the function f' has at most two roots over the interval $(0, 1)$, which means that for every node in $\Omega_{p,\lambda}$ we have at most two possible values for u_0^{n+1}. In order to simplify the computation time in each iteration we have considered just two possibilities, either we assign to u_0^{n+1} the value of the smallest root of eq. (2.73) for all the nodes in $\Omega_{p,\lambda}$, or we assign to all of them the value of the greatest root.

The issue of computing the value assigned to λ^n in each iteration is tackled numerically by a bisection method. Indeed, starting by $\lambda^{n,0} = \frac{\max(p_0^n(x)) - \min(p_0^n(x))}{2}$ we search for the smallest λ^n such that $\int_\Omega u_0^{n+1}(x)dx \leq m$. Note that the dependence of $u_0^{n+1}(x)$ on λ^n is implicitly given by its definition; i.e

$$u_0^{n+1}(x) = 1, \qquad \forall x : p_0^n(x) > \lambda^n, \tag{2.74}$$
$$u_0^{n+1}(x) = 0, \qquad \forall x : p_0^n(x) < \lambda^n, \tag{2.75}$$
$$u_0^{n+1}(x) \text{ solution of eq. (2.73)}, \qquad \forall x : p_0^n(x) = \lambda^n. \tag{2.76}$$

Remark that defining $u_0^{n+1}(x)$ this way, is compatible with the characterization given at Theorem 8 of any optimal element.

In practice, iterating this algorithm might fall into an infinite loop due to the fact that the functional value $\mathcal{J}_T(u_0^n)$ is not increasing in general. In order to overcome this issue we define an intermediate state $u_0^{n+\frac{1}{2}}$ exactly as in (2.74-2.76):

$$u_0^{n+\frac{1}{2}}(x) = 1, \qquad \forall x : p_0^n(x) > \lambda^n, \tag{2.77}$$
$$u_0^{n+\frac{1}{2}}(x) = 0, \qquad \forall x : p_0^n(x) < \lambda^n, \tag{2.78}$$
$$u_0^{n+\frac{1}{2}}(x) \text{ solution of eq. (2.73)}, \qquad \forall x : p_0^n(x) = \lambda^n; \tag{2.79}$$

λ^n chosen as the smallest value such that $\int_\Omega u_0^{n+\frac{1}{2}}(x)dx \leq m$. Then we define u_0^{n+1} as the best convex combination of u_0^n and $u_0^{n+\frac{1}{2}}$, i.e.

$$u_0^{n+1} := (1 - \theta^n)u_0^n + \theta^n u_0^{n+\frac{1}{2}}, \tag{2.80}$$
$$\theta^n := \arg\max_{\theta \in [0,1]} \mathcal{J}_T\left((1 - \theta)u_0^n + \theta u_0^{n+\frac{1}{2}}\right). \tag{2.81}$$

This way to define u_0^{n+1} guarantees the monotonicity of the algorithm. Although this transformation seems to violate the optimality conditions set on Theorem 8, once the algorithm converges, the limit distribution satisfies it, but this is not a straightforward fact, so we prove it as follows.

Claim: If the numerical algorithm described above converges after K iterations, i.e.

$$\forall n > K, \quad u_0^n(x) = u_0^{n+1}(x), \forall x \in \Omega, \tag{2.82}$$

then the following statements holds:

    i) if $0 < u_0^n(x) < 1$ then $p_0^n = \lambda^n$,

    ii) if $u_0^n(x) = 0$ then $p_0^n \leq \lambda^n$,

    iii) if $u_0^n(x) = 1$ then $p_0^n \geq \lambda^n$.

Proof: From the definition of u_0^{n+1} trough the convex combination of u_0^n and $u_0^{n+\frac{1}{2}}$ and as a consequence of (2.82) it holds that

$$\theta^n \left(u_0^{n+\frac{1}{2}} - u_0^n \right) = 0, \quad \text{for every} \quad n \geq K. \tag{2.83}$$

From this equality we deduce that for all $n \geq K$ one of the following two possibilities must stand,

    (a) $u_0^{n+\frac{1}{2}} = u_0^n$;

    (b) $\theta^n = 0$.

If (a) stands, then by the definition of $u_0^{n+\frac{1}{2}}$, the optimality conditions set on Theorem 8 necessarily holds for $c = \lambda^n$. Relatively less intuitive is the fact that the optimality conditions also holds in the (b) case. Indeed, for every $\mu \in [0, 1]$ we have

$$\mathcal{J}_T((1 - \mu)u_0^n + \mu u_0^{n+\frac{1}{2}}) \leq \mathcal{J}_T((1 - \theta^n)u_0^n + \theta^n u_0^{n+\frac{1}{2}}); \tag{2.84}$$

using a Taylor expansion in both sides we get

$$(\mu - \theta^n)\langle \nabla \mathcal{J}_T(u_0^n), u_0^{n+\frac{1}{2}} - u_0^n \rangle \leq 0 \tag{2.85}$$

in particular for $\mu > \theta^n$ we obtain $\langle \nabla \mathcal{J}_T(u_0^n), u_0^{n+\frac{1}{2}} - u_0^n \rangle \leq 0$. Now we use the explicit formula for the derivative of $\mathcal{J}_T$ established in Lemma 10

$$\begin{aligned}
\langle \nabla \mathcal{J}_T(u_0^n), u_0^{n+\frac{1}{2}} - u_0^n \rangle &= \int_\Omega p_0^n (u_0^{n+\frac{1}{2}} - u_0^n)dx \\
&= \int_\Omega (p_0^n - \lambda^n)(u_0^{n+\frac{1}{2}} - u_0^n)dx \\
&= \int_{\{p_0^n > \lambda^n\}} (p_0^n - \lambda^n)(u_0^{n+\frac{1}{2}} - u_0^n)dx + \int_{\{p_0^n < \lambda^n\}} (p_0^n - \lambda^n)(u_0^{n+\frac{1}{2}} - u_0^n)dx \\
&= \int_{\{p_0^n > \lambda^n\}} (p_0^n - \lambda^n)(1 - u_0^n)dx + \int_{\{p_0^n < \lambda^n\}} (p_0^n - \lambda^n)(0 - u_0^n)dx \\
&\geq 0.
\end{aligned} \tag{2.86}$$

Together (2.86) and (2.85) imply that $\langle \nabla \mathcal{J}_T(u_0^n), u_0^{n+\frac{1}{2}} - u_0^n \rangle = 0$ or equivalently

$$\int_{\{p_0^n > \lambda^n\}} (p_0^n - \lambda^n)(1 - u_0^n)dx + \int_{\{p_0^n < \lambda^n\}} (p_0^n - \lambda^n)(0 - u_0^n)dx = 0 \tag{2.87}$$
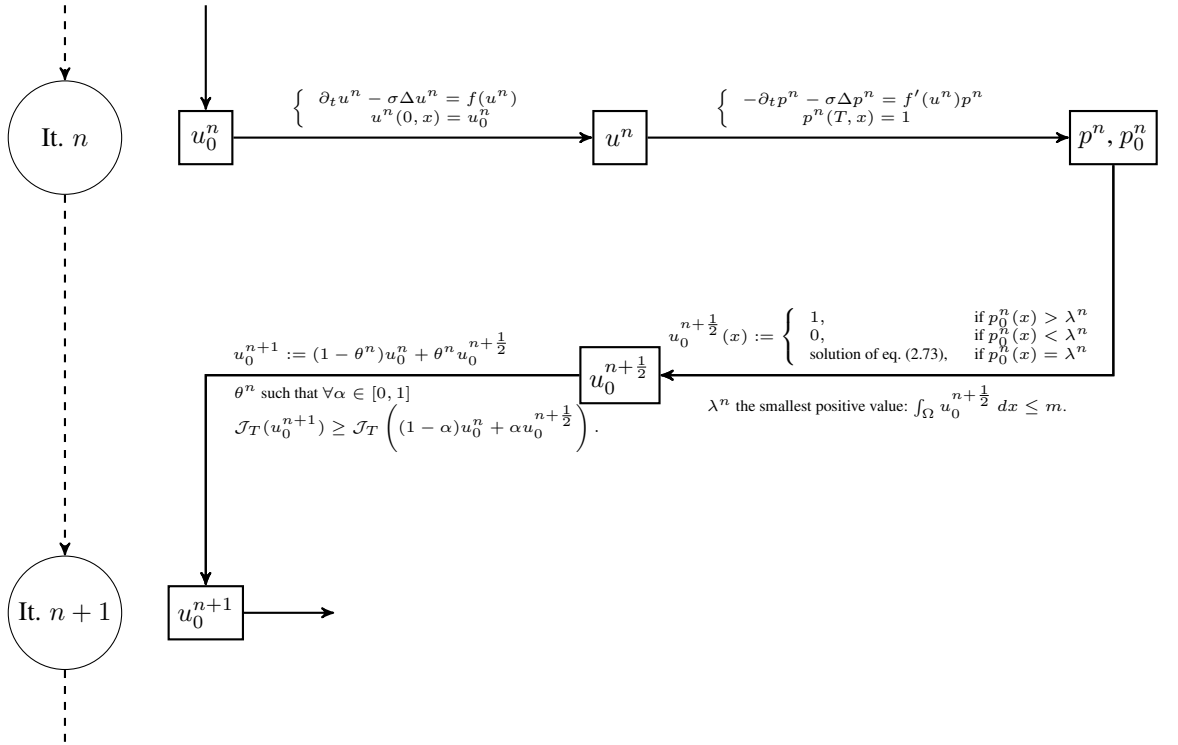
Figure 2.2: Scheme of the numerical algorithm.

from which we deduce (*i*)-(*iii*). □

This mechanism not only improves the convergence but also makes it easy to identify; in fact if at the nth-iteration the best θ for the convex combination is $\theta = 0$ then it means that the algorithm has converged, i.e. $u_0^{n+1} = u_0^n$ and then we can stop iterating.

Although the convergence of the algorithm has not been proved, the simulations show good results. In most of the cases convergence occurs after a few iterations and the limit is always an element of the admissible set $\mathcal{A}_m$ (see fig. 2.3). In a few cases the algorithm falls into a quasi-stationary state, in these cases the optimum seems to be very irregular which might be the cause of the slow convergence. For a general picture of the algorithm (see fig. 2.2).

2.5.1 On the issue of symmetry

The fact of choosing a symmetrically distributed density for the initialization of the algorithm strongly induces the symmetric feature over the searching space of solutions. Although this choice can be interpreted as a bias to the search space, it can actually be theoretically justified.

Without lost of generality, consider $\Omega = (0, a)$. As the solution satisfies Neumann boundary conditions, any optimal density distribution $\overline{u}_0$ defined over Ω is associated with a symmetric distribution $\overline{u}_{0_s}$ defined over $\Omega_s = (-a, a)$. Reciprocally, any maximizer $\overline{v}_0$ in the class of symmetric initial data on $\Omega_s = (-a, a)$ induces a solution satisfying a Neumann boundary condition at $x = 0$. Hence, $\overline{v}_0$ restricted to $(0, a)$ is also a maximizer for the problem set on $(0, a)$. Hence, there is a bijection between the maximizers on $(0, a)$ and the maximizers in the class of symmetric functions on $(-a, a)$.

2.5.2 Numerical simulations in the bistable case

For the numerical simulations we have coded the algorithm in a MATLAB routine. At each iteration we solve the differential equations for u_n and p_n by using a forward Euler's scheme in time and a finite difference approximation of the Laplace term in space. We consider a domain in space $\Omega = (-50, 50)$ with $dx = 0.1$ and a time space $t \in [0, T]$ for a given T and dt chosen such that $dt = \frac{dx^2}{3\sigma}$ which respect the CFL condition and the stability condition for this scheme.

The simulations show that the algorithm described above converges after a few iterations and successfully increases the values of $\mathcal{J}_T(u_0)$ in comparison with the trivial single block distribution (see fig. 2.3 (a)); we can also observe singularities which are associated with the values verifying $p(0, x) = \lambda$; this behavior will be discussed later on section 2.6 (see fig. 2.4).

2.6 Discussion

2.6.1 Possible generalizations

We have considered here the cost function $\mathcal{J}_T(u_0) = \int_\Omega u(T, x)dx$. Other costs are possible, such as, for example, $I_T(u_0) = -\int_\Omega \|1 - u(T, x)\|^2 dx$, where we put a minus in front of the cost so that we still want to maximize this function. More generally, assume that we want to maximize a cost function

$$I_T(u_0) := \int_\Omega F\big(u(T, x)\big)dx,$$

where F is Lipschitz-continuous over $[0, 1]$. In this case, the reader could easily verify that our method is still valid, the only change being that the condition at $t = T$ for the adjoint p becomes

$$p(T, x) = F'\big(u(T, x)\big).$$

The reader could also check that Dirichlet or Robin boundary conditions on $\partial\Omega$ could also be addressed with our method. The case of unbounded domains is more tedious. If, for example, $\Omega = \mathbb{R}$, then a concentration-compactness theorem should be used when trying to prove the existence of a maximizer $u_0 \in \mathcal{A}_m$. We leave such a generalization for a possible future work.

2.6.2 Letting $T \to +\infty$

Assume that we have as much time as needed, and that we want to optimize the initial datum u_0 in order to promote invasion, that is, convergence to 1. Such a problem is not well-posed, since many initial data should give the convergence to 1 at large time. Hence, the set of maximizing initial data could be quite large. But still a way of reaching it would be useful.

A natural ansatz is the limit of u_0^T with $T > 0$ if it exists, where u_0^T is a maximizer of $\mathcal{J}_T$. Let (u^T, p^T) the solutions associated with u_0^T. Consider a limit, up to extraction, u_0^∞ of u_0^T as $T \to +\infty$, for the L^∞ weak star convergence. Let u be the solution on $(0, \infty) \times \Omega$ associated with u_0^∞, which is indeed the limit of u^T.

Next, define $\widetilde{p}^T := m^T p^T$, where m^T is a positive constant chosen so that $\int_\Omega \widetilde{p}^T(0, x)dx = 1$. We know from Theorem 8 that there exists a constant c^T such that

i) if $0 < u_0^T(x)$ then $\widetilde{p}^T(0, x) \geq c^T$,

ii) if $u_0^T(x) < 1$ then $\widetilde{p}^T(0, x) \leq c^T$.
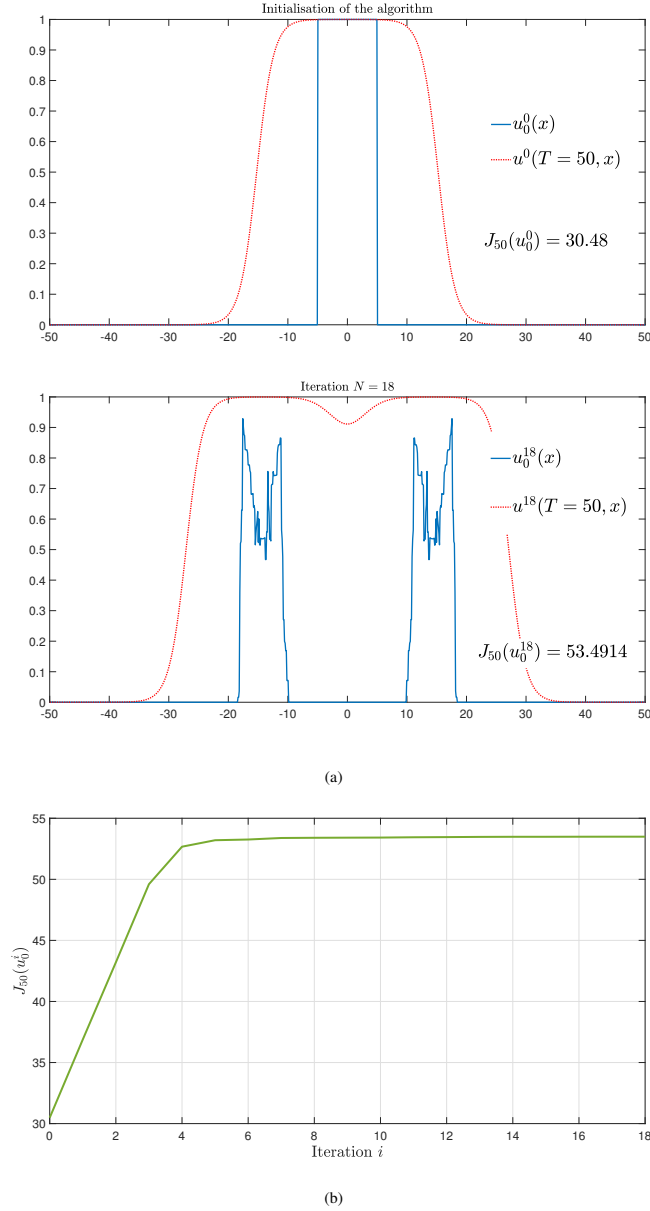
Figure 2.3: Considering a fixed mass $m = 10$, this figure shows the initial data associated with the first and the last iteration of the algorithm and the corresponding solutions of eq. (2.2) with reaction term $f(u) = u(1-u)(u-0.35)$ (a). We also show the evolution line of the operator $\mathcal{J}_{50}(u_0^i)$ from the first iteration to the last one (b). Note that the limit reached after 18 iterations is an initial data separated in two blocs and shows singularities as a consequence of the definition of the initial solution within the set $\Omega_{p,\lambda^{18}}$.

Parabolic regularity yields that the solution $\widetilde{p}^T$ converges in $W_{q,loc}^{1,2}\big((0,\infty)\times\Omega\big)$ for all $q \in (1,\infty)$
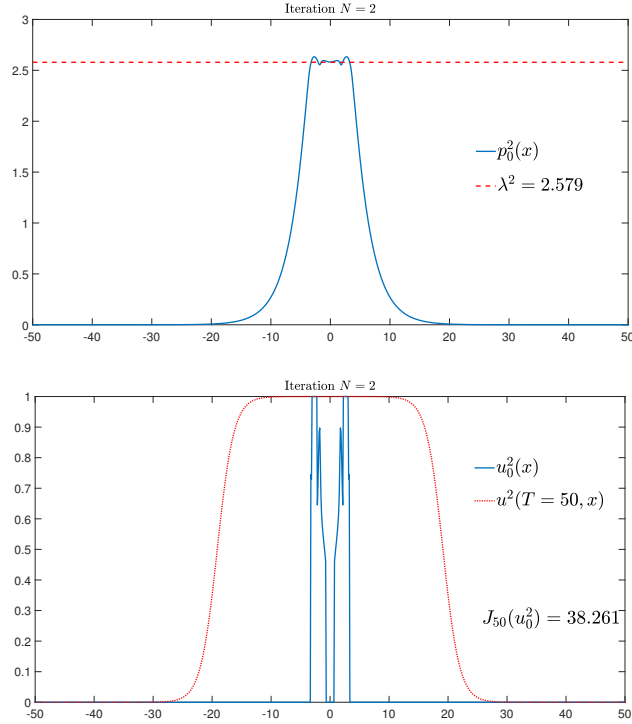
Figure 2.4: The figure corresponds to the 2nd and last iteration of the algorithm for a given final time $T = 50$ and a fixed mass $L = 4$. First we show the adjoint state p which is the solution of eq. (2.71) and its associated value λ^2 mentioned in Theorem 8. Note that in this case the set Ω_{p,λ^2} is not negligible so the associated u_0^2 showed at the bottom present singularities arising from the solution of eq. (2.73) within this set.

as $T \to +\infty$ to a solution p of the backward equation

$$\begin{cases} -\partial_t p - \sigma \Delta p = f'(u)p & \text{in } (0, \infty) \times \Omega, \\ \\ p(t, x) > 0 & \text{in } (0, \infty) \times \Omega, \\ \\ \frac{\partial p}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, \infty), \text{ for all } x \in \partial\Omega, \end{cases} \qquad (2.88)$$

Indeed, we know from Proposition 2.7 of [40] that such a solution is unique, up to normalization, which is indeed given here by $\int_\Omega p(0, x)dx = 1$. The following partial characterization of u_0^∞ follows:

i) if $0 < u_0^\infty(x)$ then $p(0, x) \geq c$,

ii) if $u_0^\infty(x) < 1$ then $p(0, x) \leq c$,

where c is indeed the limit of c^T.

Of course, such a partial characterization is mostly theoretical, since there is no way of constructing p numerically, except by approximating it as the limit of the functions $\widetilde{p}^T$. Note that this adjoint function does not depend on the cost function, which is satisfying since, as we expect convergence to 1 at large time, the shape of the cost function should not play any role.

Chapter 3

Second order optimality conditions for optimization with respect to the initial data in reaction-diffusion equations

"Life is an unfoldment, and the further we travel the more truth we can comprehend. To understand the things that are at our door is the best preparation for understanding those that lie beyond."

Hypatia

This chapter corresponds to an in progress research in collaboration with Idriss Mazari and Grégoire Nadin.

3.1 Introduction

3.1.1 Scope of the chapter

In this chapter, we propose to establish several results concerning an optimal control problem for a class of semilinear parabolic equations. In this problem, the control variable is the initial condition. As we will see throughout the statement of the results, the form (e.g. convex or concave) of the non-linearity plays a crucial role in the analysis, and our main theorem is a fine study of second-order optimality conditions for such problems.

The origin of this paper is the study of an optimal control problem that arises naturally in mathematical biology and that deals with bistable reaction-diffusion equations, whose complicated behaviour, the non-linearity being neither convex nor concave, makes the analysis very intricate. Let us briefly sketch how this problem fits in the literature devoted to such optimisation and control problems for mathematical biology.

Optimization problems for reaction-diffusion equations have by now garnered a lot of attention from the mathematical community. Most of these optimization problems are set in a stationary setting, that is, assuming that the population has already reached an equilibrium, and the main problems considered often deal with the optimization of the spatial heterogeneity [16, 42, 44, 58, 62, 63, 65, 71, 59] and we refer for instance to the surveys [51, 64], and that they deal with monostable equations. On the other hand, optimization problems for bistable equations, which are the other paradigmatic class of equations,

have received a less complete mathematical treatment (we cite [3], where a bistable ODE is considered) and are not yet fully understood. Furthermore, less attention has been devoted to optimization problem for evolution equations, and the aim of this chapter is to provide a detailed analysis of second order optimality conditions for the optimization of a criterion depending on the solution of a bistable equation with respect to the initial condition.

Let us explain the main motivation behind this paper. Bistable equations are of central importance in mathematical biology [67] and, very broadly speaking, model the evolution of a subgroup of a population. Among their many applications, one may mention chemical reactions [75], neurosciences [27], phase transition [60], linguistic dynamics [76], or the evolution of diseases [67]. The last interpretation is of particular relevance to us, given that this model is used to design optimal strategies in order to control the spread of several diseases including the dengue, and this was the main motivation in [3]. The strategy is to release a certain amount of mosquitoes carrying *Wolbachia* (i.e. a bacterium inhibiting the ability of transmitting the disease to humans) in a population of wild mosquitoes that can potentially transmit the disease, the objective is to maximize the *Wolbachia* carrier population in the final time. Put more mathematically: given a time horizon T,

How should we arrange the initial population in order to maximize the population size at T?

Even without having stated it formally, we can make two observations: the first one is that, the variable of the equation being the proportion of a subgroup, this will lead to pointwise (L^∞) constraints. The second one is that we naturally have to add an L^1 constraint for modelling reasons. Both these constraints can in practice be very complicated to handle.

The intricate nature of bistable non-linearities makes the analysis of optimality conditions extremely complicated. In this chapter, we provide a detailed analysis of second order optimality conditions, which, as will be explained later, are very useful for numerical simulations. Our argument relies on a two-scale expansion method which we believe is new in this context and could be adapted to derive optimality conditions for other relevant problems in optimization.

3.1.2 Mathematical setup

We work in a smooth, bounded connected domain $\Omega \subset \mathbb{R}^n$. Let us first state our equation, our problem and our results, we will later specify how these results can be used to study the optimisation problem for bistable non-linearities. We consider a bounded $\mathscr{C}^2$ function $f : [0; 1] \to \mathbb{R}$, and the associated parabolic equation

$$\begin{cases} \partial_t u - \sigma \Delta u = f(u) & \text{in } \mathbb{R}_+ \times \Omega, \\[2mm] u(0, x) = u_0(x) & \text{in } \Omega, \\[2mm] \frac{\partial u}{\partial \nu}(t, x) = 0 & \text{in } \mathbb{R}_+ \times \partial\Omega, \end{cases} \tag{3.1}$$

where u_0 is an initial condition satisfying the constraint

$$0 \le u_0 \le 1.$$

It is standard to see that this equation admits a unique solution u, and we can hence define, for any $T > 0$, the functional

$$\mathcal{J}_T(u_0) := \int_\Omega u(T, x) dx. \tag{3.2}$$

The goal is to maximise $\mathcal{J}_T$ with respect to u_0, and it is thus natural, in view of the motivations laid out in the first subsection of the Introduction, to introduce a L^1 constraint on u_0, which will be encoded

by a parameter $m \in (0; |\Omega|)$ and the following admissible class:

$$\mathcal{A}_m := \left\{ u_0 \in L^\infty(\Omega), 0 \le u_0 \le 1 \text{ a.e. }, \int_\Omega u_0 = m \right\}. \tag{3.3}$$

The variational problem is then

$$\max_{u_0 \in \mathcal{A}_m} \mathcal{J}_T(u_0). \tag{P}$$

An expression for the first and second order optimality conditions have been given in [68]. In order to state our results, we need to recall these optimality conditions: if we consider $u_0 \in \mathcal{A}_m$ and an admissible perturbation h_0 at u_0 (in other words, a function h_0 such that for any $\varepsilon > 0$ small enough $u_0 + \varepsilon h_0 \in \mathcal{A}_m$) then the first order Gâteaux-derivative of $\mathcal{J}_T$ at u_0 in the direction h_0 is

$$\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_\Omega h_0(x) p(0, x) dx, \tag{3.4}$$

where p solves the adjoint equation

$$\begin{cases} -\partial_t p - \sigma \Delta p = f'(u) p & \text{in } (0, T) \times \Omega, \\\\ p(T, x) = 1 & \text{in } \Omega, \\\\ \frac{\partial p}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial \Omega. \end{cases} \tag{3.5}$$

Here u is the solution of (3.1) with initial condition u_0.

The main result in [68] states the following:

Theorem 14 *[68] There exist $\overline{u}_0 \in \mathcal{A}_m$ solution of (**P**). Moreover, setting $\overline{u}$ as the solution of (3.1) associated with this optimum initial data and $\overline{p}$ the unique solution of (3.5) for $u = \overline{u}$, there exists a non-negative real value $\overline{c}$ such that*

i) if $0 < \overline{u}_0(x) < 1$ then $\overline{p}(0, x) = \overline{c}$,

ii) if $\overline{p}(0, x) > \overline{c}$, then $\overline{u}_0(x) = 1$,

iii) if $\overline{p}(0, x) < \overline{c}$, then $\overline{u}_0(x) = 0$.

Moreover, for almost every $x \in \{\overline{p}(0, \cdot) = \overline{c}\}$, one has

$$f'\big(\overline{u}_0(x)\big) = -\partial_t \overline{p}(0, x)/\overline{p}(0, x) \tag{3.6}$$

and the left-hand side belongs to $L^p_{loc}(\Omega)$.

If f' is monotonic, equation (3.6) admits a unique solution and thus Theorem 14 fully characterizes $\overline{u}_0$. But for the model that initially motivated this problem, a bistable nonlinearity $f(u) = u(1-u)(u-\rho)$ was considered [5]. In such case, equation (3.6) might have several solutions and thus we need further information in order to characterize $\overline{u}_0$.

Our first result deals with ways to specify the behaviour of u_0 on that normal set, and states that, roughly speaking, u_0 must always be in a zone of "concavity" of f. Let us recall the expression of the second order derivative:

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \int_0^T \int_\Omega f''(u(t, x)) p(t, x) h^2(t, x) \, dx \, dt, \tag{3.7}$$

where h stands for the solution of the following equation

$$\begin{cases} \partial_t h - \sigma \Delta h = f'(u)h & \text{in } (0,T) \times \Omega, \\[2mm] h(0,x) = h_0(x) & \text{in } \Omega, \\[2mm] \frac{\partial h}{\partial \nu}(t,x) = 0 & \text{for all } t \in (0,T), \text{ for all } x \in \partial \Omega, \end{cases} \qquad (3.8)$$

We can now state our main result:

Theorem 15 *Let $\overline{u}_0$ be a solution of* (**P**)*. If the set $\Omega_{\overline{c}} := \{x \in \Omega : 0 < \overline{u}_0(x) < 1\}$ has a positive measure then, for any interior point x of $\Omega_{\overline{c}}$, there holds*

$$f''(\overline{u}_0(x)) \le 0. \qquad (3.9)$$

Remark 16 *The method we put forth is reminiscent of one that was used in [68] to study the case of a constant initial condition. Here, working with interior point greatly complexifies the situation and calls for two-scale asymptotic expansions.*

The main drawback to our approach is that it can not cover the case of singular (e.g. Cantor-like) abnormal sets, and it is a very interesting question to prove that such a property holds for any point of the abnormal set $\Omega_{\overline{c}}$. Another possibility would be to prove a priori regularity estimates on this abnormal set, but this seems to be highly challenging.

Remark 17 *In the case of a monostable non-linearity, we refer to [68] where it was proved that a constant initial condition was always the unique solution of the optimization problem.*

Bistable reaction-diffusion equations   Let us now briefly explain how this result applies to the bistable reaction-diffusion equation. To model the evolution of the subpopulation mosquitoes, we use a bistable non-linearity, that is a function $f : [0;1] \to \mathbb{R}$ such that

1. f is $\mathscr{C}^2$ on $[0,1]$,

2. There exists $\rho \in (0;1)$ such that 0, ρ and 1 are the only three roots of f in $[0,1]$, This parameter ρ accounts for the aforementioned Allee effect.

3. $f'(0), f'(1) < 0$ and $f'(\rho) > 0$,

The fact that the set $\{p = c\}$ may have a positive measure or, in other words, that a solution may not be the characteristic function of a set, leads to several difficulties in terms of numerical methods, because standard gradient methods or fixed-point algorithms fail to capture what this so-called "singular arc" should be replaced with. Explaining this difficulty further is a key to stating our results. Supposing we put forth an iterative procedure and we are given the initial condition at the n-step u_0^n, we construct $u_0^{n+1} = u_0^n + h_0^n$, where h_0^n maximises (3.4) and is an admissible perturbation. Since the adjoint at the n-th step p_0^n may have level sets of positive measure, one can not directly apply the bathtub principle and choose h_0^n as the difference of characteristic functions of two level sets of p_0^n and we must thus describe what happens on the singular arc, that is, on the level set $\{p_0^n = c^n\}$ where c_n is chosen so that

$$\mu(p_0^n > c^n) < m \,, \mu(p_0^n \ge c^n) > m \,. \qquad (3.10)$$

We first define, in this case, $u_0^{n+1} = 1$ on $\{p_0^n > c^n\}$, $u_0^{n+1} = 0$ on $\{p_0^n < c^n\}$, and it remains to fix the value on $\omega_n := \{p_0^n = c^n\}$. Discretising Equation (3.5) on ω_n we obtain, with an explicit finite difference scheme

$$-\left(\frac{p_0^n(dt,x) - c^n}{dt}\right) = f'(u_0^{n+1})c^n \qquad (3.11)$$

and the value on u_0^{n+1} on ω_n must be a root of (3.11). However, for bistable non-linearities, this equation may have two roots, say μ_1^n and μ_2^n. In this case, these two roots can be distinguished through the convexity of f. In other words, if we have two roots, up to relabelling,

$$f''(\mu_1^n) > 0 \, , f''(\mu_2^n) < 0. \tag{3.12}$$

In [68] this difficulty is overcome by examining the two different possibilities and choosing the best one, but our Theorem allows to overcome this difficulty by choosing directly the root that is in the "concavity" zone of f.

3.2 Proof of Theorem 15

3.2.1 Strategy of proof

We remind to the reader that

$$\left\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \right\rangle = \int_0^T \int_\Omega f''(u(t,x)) p(t,x) h^2(t,x) \, dx \, dt.$$

Let x an interior point of $\Omega_{\bar{c}}$. For all $\delta \in (0, 1/2)$, let $\Omega_{\bar{c}}^\delta := \{\delta \leq \bar{u}_0(x) \leq 1 - \delta\}$, so that $\Omega = \cap_{\delta > 0} \Omega_{\bar{c}}^\delta$. There exists $\delta > 0$ such that x is an interior point of $\Omega_{\bar{c}}^\delta$, that is, there exists $\ell > 0$ such that $[x - \ell, x + \ell] \subset \Omega_{\bar{c}}^\delta$. Take $h_0 \in L^\infty(\Omega)$ supported in $[x - \ell, x + \ell]$, such that $\int_\Omega h_0 = 0$. Then, for $\epsilon > 0$ small enough, $\bar{u}_0 + \epsilon h_0$ is an admissible initial datum and thus $\mathcal{J}_T(\bar{u}_0 \pm \epsilon h_0) \leq \mathcal{J}_T(\bar{u}_0)$. Developing up to order 2, this yields

$$= \int_0^T \int_\Omega f''(u(t,x)) p(t,x) h^2(t,x) \, dx \, dt \leq 0.$$

The issue now is that this expression is almost impossible to manipulate. We will thus construct a particular h_0, which is strongly oscillating around x, in order to concentrate this integral around $t = 0$ using a Laplace method.

3.2.2 Construction and properties of the admissible perturbation

For simplicity, we will assume $\sigma = 1$. Let us consider the equation

$$\begin{cases} \dfrac{\partial h_k}{\partial t} - \dfrac{\partial^2 h_k}{\partial x^2} = f'(\bar{u}) h_k \, , \\[2mm] \dfrac{\partial h_k}{\partial \nu} = 0 \, , \\[2mm] h_k(0, x) = \theta(x) \sin(kx). \end{cases} \tag{3.13}$$

We look for an asymptotic expansion of h_k of the form

$$h_k(t, x) = h_k^0(k^2 t, x, kx) + \frac{1}{k} h_k^1(k^2 t, x, kx) + \dots \tag{3.14}$$

which, declaring the new variables $s = k^2 t$ and $y = kx$ and after identification at the first and second order, gives the following equations on h_k^0 and h_k^1:

$$\begin{cases} \dfrac{\partial h_k^0}{\partial s} - \dfrac{\partial^2 h_k^0}{\partial y^2} = 0 \, , \\[2ex] \dfrac{\partial h_k^0}{\partial \nu} = 0 \, , \\[2ex] h_k^0(0, x, y) = \theta(x) \sin(y). \end{cases} \tag{3.15}$$

and

$$\begin{cases} \dfrac{\partial h_k^1}{\partial s} - \dfrac{\partial^2 h_k^1}{\partial y^2} = 2 \dfrac{\partial^2 h_k^0}{\partial x \partial y}, \\[2ex] \dfrac{\partial h_k^1}{\partial \nu} = 0 \, , \\[2ex] h_k^1(0, x, y) = 0. \end{cases} \tag{3.16}$$

Equation (3.15) can be solved explicitly, giving

$$h_k^0(s, x, y) = \theta(x) \sin(y) e^{-s}. \tag{3.17}$$

This, in turn, allows to solve Equation (3.16) as

$$h_k^1(s, x, y) = s e^{-s} \theta'(x) \cos(y). \tag{3.18}$$

Proposition 18 *The asymptotic expansion* (3.14) *is valid in* $L^2(\Omega)$ *in the following sense: there exists* $M > 0$ *such that, if we define*

$$R_k := h_k(t, x) - h_k^0(k^2 t, x, kx) - \frac{1}{k} h_k^1(k^2 t, x, kx)$$

then, for any $t \in (0; T)$,

$$\|R_k(t, \cdot)\|_{L^2(\Omega)} \leq \frac{M}{k^2}. \tag{3.19}$$

Proof: To prove this Proposition, we write down the equation satisfied by R_k. Straightforward computations show that R_k solves

$$\partial_t R_k - \Delta R_k - f'(\overline{u}) R_k := f'(\overline{u}) \left(h_k^0 + \frac{1}{k} h_k^1 \right) + \frac{\partial^2 h_k^0}{\partial x^2} + 2 \frac{\partial^2 h_k^1}{\partial x \partial y} + \frac{1}{k} \frac{\partial^2 h_k^1}{\partial x^2}, \tag{3.20}$$

and all the functions on the right hand side are evaluated at (k^2t, x, kx) (we dropped this for notational convenience). We now introduce the following notations:

$$
\begin{cases}
W_0 := f'(\overline{u}), \\
\\
V_{0,k}(t,x) := h_k^0(k^2t, x, kx) + \frac{1}{k}h_k^1(k^2t, x, kx), \\
\\
V_{1,k} := \frac{\partial^2 h_k^0}{\partial x^2}(k^2t, x, kx), \\
\\
V_{2,k} := 2\frac{\partial^2 h_k^1}{\partial x \partial y} + \frac{1}{k}\frac{\partial^2 h_k^1}{\partial x^2}.
\end{cases}
$$

First of all, there exists $M_0 > 0$ such that

$$\|W_0\|_{L^\infty((0;T)\times\Omega)} \le M_0. \tag{3.21}$$

We gather the main estimates on source terms in the following Lemma:

Lemma 19 *There exists $\widetilde{M} > 0$ such that*

$$\int_0^T \|V_{0,k}(t,\cdot)\|_{L^2(\Omega)} \le \frac{\widetilde{M}}{k^2}, \tag{3.22}$$

$$\int_0^T \|V_{1,k}(t,\cdot)\|_{L^2\Omega)}dt \le \frac{\widetilde{M}}{k^2}. \tag{3.23}$$

$$\int_0^T \|V_{2,k}(t,\cdot)\|_{L^2\Omega)}dt \le \frac{\widetilde{M}}{k^2}. \tag{3.24}$$

Proof: We mainly use the Laplace method. Let us recall the following consequence of the Laplace method, for any integer $m \in \mathbb{N}$, one has

$$\int_0^T t^{m-1}e^{-k^2t}dt \underset{k\to\infty}{\sim} \frac{(m-1)!}{k^{2m}}. \tag{$\mathbf{I}_m$}$$

Proof of (3.22)

By the triangle inequality we get, for any $t \in (0;T)$,

$$\|V_{0,k}(t,\cdot)\|_{L^2} \le \|h_k^0(k^2t, \cdot, k\cdot)\|_{L^2(\Omega)} + \frac{1}{k}\|h_k^1(k^2t, \cdot, k\cdot)\|_{L^2(\Omega)}$$

We first use the explicit expressions (3.17)-(3.18) for h_k^0 and h_k^1 to obtain

$$
\begin{aligned}
\|h_k^0(k^2t, \cdot, k\cdot)\|_{L^2(\Omega)}^2 &= \int_\Omega \theta(x)^2 \sin(kx)^2 e^{-2k^2t}dt \\
&\le e^{-2k^2t}|\Omega|,
\end{aligned} \tag{3.25}
$$

and integrating this inequality between 0 and T gives

$$\int_0^T \|h_k^0(k^2t, x, k\cdot)\|_{L^2(\Omega)}dt \le \frac{M_0}{k^2}.$$

In the same way we have, for any $t \in (0; T)$,

$$\|h_k^1(k^2 t, \cdot, k\cdot)\|_{L^2(\Omega)}^2 = k^4 t^2 e^{-2k^2 t} \int_\Omega (\theta'(x))^2 \sin(x)^2 dx$$
$$\leq k^4 t^2 e^{-2k^2 t} |\Omega|. \tag{3.26}$$

Taking the square root, we get

$$\frac{1}{k} \int_0^T \|h_k^1(k^2 t, \cdot, k\cdot)\|_{L^2(\Omega)} dt \leq |\Omega| k \int_0^T t e^{-k^2 t} dt.$$

Using $(\mathbf{I}_m)$ with $\lambda = 2$ gives

$$\int_0^T \frac{1}{k} \|h_k^1(k^2 t, \cdot, k\cdot)\|_{L^2(\Omega)} dt \leq \frac{M_1}{k^3}$$

for some constant M_1.

Summing these two contributions gives the required result.

Proof of (3.23) This follows from the same arguments, by simply observing that

$$V_{1,k}(t, x) = e^{-k^2 t} \theta''(x) \sin(kx).$$

Proof of (3.24) We once again split the expression and estimate separately

$$\int_0^T \left\| \frac{\partial^2 h_k^1(k^2 t, \cdot, k\cdot)}{\partial x \partial y} \right\|_{L^2(\Omega)} dt \text{ and } \frac{1}{k} \int_0^T \left\| \frac{\partial^2 h_k^1(k^2 t, \cdot, k\cdot)}{\partial x^2} \right\|_{L^2(\Omega)} dt.$$

We first observe that for any $t \in (0; T)$, we have

$$\frac{\partial^2 h_k^1(k^2 t, \cdot, k\cdot)}{\partial x \partial y} = -k^2 t e^{-k^2 t} \theta''(x) \sin(kx).$$

In particular, for any $t \in (0; T)$

$$\left\| \frac{\partial^2 h_k^1(k^2 t, \cdot, k\cdot)}{\partial x \partial y} \right\|_{L^2(\Omega)} \leq k^2 t e^{-k^2 t} |\Omega|$$

so that the Laplace method $(\mathbf{I}_m)$ with $\lambda = 2$ gives the bound

$$\int_0^T \left\| \frac{\partial^2 h_k^1(k^2 t, \cdot, k\cdot)}{\partial x \partial y} \right\|_{L^2(\Omega)} dt \leq \frac{M_2}{k^2}$$

for some constant M_2. The last term is controlled in exactly the same way and gives a $\mathcal{O}(k^{-2})$ bound. $\square$

Let us now prove Estimate (3.19). The equation on R_k rewrites

$$\partial_t R_k - \Delta R_k - W_0 R_k = W_0 V_{0,k} + V_{1,k} + V_{2,k}. \tag{3.27}$$

Multiplying the equation by R_k and integrating by parts in space gives

$$\frac{1}{2} \partial_t \int_\Omega R_k^2 + \int_\Omega |\nabla R_k|^2 - \int_\Omega W_0 R_k^2 \leq \|R_k\|_{L^2(\Omega)} \left(M_0 \|V_{0,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} \right).$$

In other words, bounding W_0 by M_0 and defining $g(t) := \|R_k(t, \cdot)\|^2_{L^2(\Omega)}$ we obtain

$$\frac{1}{2}g'(t) \leq M_0 g(t) + \sqrt{g(t)} \left(M_0 \|V_{0,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} \right).$$

Furthermore, $R_k(0, \cdot) = 0$. We thus obtain, by the Gronwall Lemma, for any $t \in (0; T)$,

$$\sqrt{g(t)}e^{-M_0 t} \leq \int_0^t e^{-M_0 s} \left(M_0 \|V_{0,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} \right).$$

Hence, by Lemma 19 we get for some constant N_0 and any $t \in (0; T)$,

$$\|R_k(t, \cdot)\|_{L^2(\Omega)} \leq N_0 \int_0^T \left(\|V_{0,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} + \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} \right) \leq \frac{N_0 \widetilde{M}}{k^2}.$$

$\square$

3.2.3   Proof of Theorem 15

As a consequence, for the initial perturbation $h_{0,k} := \theta(x)\sin(kx)$, and defining

$$F(t, x) := f''(\overline{u}(t, x))\overline{p}(t, x),$$

the second order derivative of $\mathcal{J}_T$ in $\overline{u}_0$ can be written as

$$\begin{aligned}
d^2 \mathcal{J}_T(\overline{u}_0)[h_{0,k}, h_{0,k}] &= \int_0^T \int_\Omega F(t, x)h_k(t, x)^2 \, dx \, dt, \\
&= \int_0^T \int_\Omega F(t, x)(R_k(t, x) + V_{0,k}(t, x))^2 \, dx \, dt, \\
&= \int_0^T \int_\Omega F(t, x) \left(R_k(t, x)^2 + 2R_k(x, t)V_{0,k}(t, x) + V_{0,k}(t, x)^2 \right) \, dx \, dt.
\end{aligned}$$
(3.28)

From the assumptions on f and the estimates on u and p, it easy to see that

$$\|F\|_{L^\infty((0;T)\times\Omega)} \leq M_3.$$
(3.29)

Gathering (3.29) and (3.19) it follows that

$$\int_0^T \int_\Omega F(t, x)R_k(t, x)^2 \, dx \, dt = \mathcal{O}(k^{-4})$$
(3.30)

and similarly, gathering (3.22) and (3.19) we obtain

$$\int_0^T \int_\Omega F(t, x)R_k(t, x)V_{0,k}(t, x) \, dx \, dt = \mathcal{O}(k^{-4})$$
(3.31)

Let us now studying the term

$$\begin{aligned}
\int_0^T \int_\Omega F(t, x)V_{0,k}(t, x)^2 \, dx \, dt &= \int_0^T \int_\Omega F(t, x) \left(h_k^0(t, x) + \frac{1}{k}h_k^1(t, x) \right)^2 \\
&= \int_0^T \int_\Omega F(t, x) \left(h_k^0(t, x)^2 + 2\frac{1}{k}h_k^0(t, x)h_k^1(t, x) + \frac{1}{k^2}h_k^1(t, x)^2 \right) dxdt
\end{aligned}$$

Once again we split the expression. Applying Cauchy-Swchartz inequality, and the estimates in (3.25) and (3.26) it follows that the second term verify

$$\left| 2\frac{1}{k}\int_0^T \int_\Omega F(t,x)h_k^0(t,x)h_k^1(t,x)dxdt \right| \leq 2\frac{M_3}{k}\int_0^T \|h_k^0\|_{L^2(\Omega)}\|h_k^1\|_{L^2(\Omega)}dt$$
$$\leq 2M_4k\int_0^T te^{-2k^2t}dt,$$
$$= \mathscr{O}(k^{-3}). \tag{3.32}$$

M_4 is a constant depending only on $|\Omega|$ and M_3. The last step in the above expression follows directly from $(\mathbf{I}_m)$ with $\lambda = 2$. Similarly for the third term we obtain

$$\left| \frac{1}{k^2}\int_0^T \int_\Omega F(t,x)h_k^1(t,x)^2dxdt \right| \leq \frac{M_3}{k^2}\int_0^T \|h_k^1\|_{L^2(\Omega)}^2dt$$
$$\leq M_4k^2\int_0^T t^2e^{-2k^2t}dt,$$
$$= \mathscr{O}(k^{-4}). \tag{3.33}$$

In this case we apply $(\mathbf{I}_m)$ for $m = 3$.

Finally, let us study the first term which can be written as

$$\int_0^T \int_\Omega F(t,x)h_k^0(t,x)^2dxdt = \int_0^T e^{-2k^2t}G(t,x)\,dt \tag{3.34}$$

with $G := \int_\Omega F(t,x)\theta(x)^2\sin(kx)^2dx$ is a $\mathscr{C}^1$ function of time.

It follows from the Laplace method that, when $k \to \infty$ one have

$$\int_0^T \int_\Omega F(t,x)h_k^0(t,x)^2dxdt \sim \frac{1}{2k^2}G(0) \tag{3.35}$$

Gathering the estimates in (3.30),(3.31), (3.32), (3.33), (3.35) and plugging them into the second derivative of the functional $\mathcal{J}_T$ given by (3.28) it follows that

$$d^2\mathcal{J}_T(\overline{u}_0)[h_{0,k},h_{0,k}] \underset{k\to\infty}{\sim} \frac{1}{2k^2}G(0) \tag{3.36}$$

We now argue by contradiction, and we assume that there exists a point x_0 such that

$$f''(\overline{u}_0(x_0)) > 0,$$

and θ is henceforth a bump function around x_0. Since $p > 0$ this implies that $F(0,x_0) > 0$. Let us check now that as a consequence, $G(0)$ has a sign for k large enough. Indeed, since $\sin(kx) \underset{k\to\infty}{\rightharpoonup} \frac{1}{2}$ one has

$$G(0) = \int_\Omega F(0,x)\sin(kx)^2dx \underset{k\to\infty}{\longrightarrow} \frac{1}{2}\int_\Omega F(0,x)dx \geq 0$$

this would mean that for k sufficiently large,

$$d^2\mathcal{J}_T(\overline{u}_0)[h_{0,k},h_{0,k}] \geq 0$$

which contradicts the fact that $\overline{u}_0$ is a maximizer of $\mathcal{J}_T$. The proof is then completed.

Part II

Crohn's disease and inflammatory patterns

Crohn's disease is an inflammatory bowel disease (IBD) that is not well understood. In particular, unlike other IBDs, the inflamed parts of the intestine compromise deep layers of the tissue and are not continuous but separated and distributed through the whole gastrointestinal tract, displaying a patchy inflammatory pattern. In this part of the manuscript we introduce a toy-model which might explain the appearance of such patterns. We consider a reaction-diffusion system involving bacteria and immune cells, and prove that under certain conditions this system might reproduce an activator-inhibitor dynamic leading to the occurrence of Turing-type instabilities. We also propose a set of parameters for which the system exhibits such phenomena and compare it with realistic parameters found in the literature.

Chapter 4

A Turing mechanism in order to explain the patchy nature of Crohn's disease

"I like crossing the imaginary boundaries people set up between different fields - it's very refreshing."

Maryam Mirzakhani

The results presented in this chapter came from a multi-disciplinary collaboration with Grégoire Nadin, Eric Ogier-Denis and Hatem Zaag and has been submitted for publication, [70].

4.1   Introduction

Ulcerative colitis and Crohn's disease represent the two main types of inflammatory bowel disease (IBD). Both are relapsing diseases and may present similar symptoms including long-term inflammation in the digestive system, however they are very different: Ulcerative colitis affects only the large intestine and the rectum whereas Crohn's disease can affect the entire gastrointestinal tract from the mouth to the anus (see [43] for example). Typical presentations of Crohn's disease include the discontinuous involvement of various portions of the gastrointestinal tract and the development of complications including strictures, abscesses, or fistulas that compromise deep layers of the tissue, while ulcerative colitis remains superficial but present no healthy areas between inflamed spots.

There is consensus now that IBD results from an unsuitable response of a deficient mucosal immune system to the indigenous flora and other luminal antigens due to alterations of the functions of the epithelial barrier. Our goal was to propose a simplified mathematical model simulating the immune response triggering inflammation. In the particular case of Crohn's disease, we seek to understand the patchy inflammatory patterns that differentiate patients suffering from this illness from those who have been diagnosed with ulcerative colitis. For recent articles on IBD discussing the differences of patterns between Crohn's disease and Ulcerative colitis, see for instance [43, 57, 11].

IBD can be seen as an example of the acute inflammatory response of body tissues caused by harmful stimuli such as the presence of pathogenic germs or damaged cells. This protective response is also associated with the origin of other well-known diseases such as rheumatoid arthritis, the inflammatory phase in diabetic wounds or tissue inflammation, and has been extensively studied. Today it is still of central interest for researchers and, although several models have been proposed in order to understand

the causes that lead to acute inflammation, the mathematical approach to this topic remains a relatively new field of research. A more complete review on the subject is provided in [89, 90].

Among the mathematical works on inflammation we can refer to many models based on ordinary differential equations [21, 24, 35, 48, 53, 61, 77, 78, 93]. Most of the authors take into account pro-inflammatory and anti-inflammatory mediators but also pathogens and other more or less realistic physiological variables. Depending on the parameters and the initial data, these models manage to reproduce a variety of scenarios that can be observed experimentally and clinically; for example the case in which the host can eliminate the infection and also other situations in which the immune system cannot keep the disease under control or where the existence of oscillatory solutions determines a chronic cycle of inflammation. Most of the conclusions in the referenced papers are the result of stability studies of the equilibrium states and numerical analyses of the simulations by phase portraits methods. In addition, in [21, 48, 78, 93] a sensitivity analysis of the variables to the parameters of the models is performed in order to adjust the numerical results to experimental data and achieve greater biological fidelity.

Several authors have also considered spatial heterogeneity in order to model the inflammatory response, we can mention [25, 26, 41] in the particular case of atherogenesis, [52, 74] in the tissue inflammation context and [17, 84] for the acute inflammatory response. The main variables of the models introduced in the mentioned works vary according to the dynamics that the authors wish to describe: the density of phagocytic cells, pro-inflammatory cytokines, anti-inflammatory mediators and bacteria are some standard quantities that are often taken into account. As in the ordinary differential equations approach, the stability of the systems is systematically studied, in [17, 19, 25, 53] a vast analysis of all possible scenarios is performed depending on the values of the model parameters, the authors provide biological interpretation of such behavior as well as numerical simulations; furthermore, in [26] the existence of travelling waves solutions is proved to be at the origin of a chronic inflammatory response.

A different approach is presented in [74]. The model introduced in this paper aims to explain mathematically the patterns observed in the skin due to acute inflammation in the absence of specific pathogenic stimuli. By analyzing the stability of homogeneous and non-homogeneous states, sufficient conditions leading to the existence of such pattern solutions are obtained; several numerical examples are given as well. Similarly, in [52] the authors claim that the instability of the uniform steady distribution of phagocytic cells might trigger non-uniform cell density distributions which is potentially dangerous since tissue damage may occur in regions of high cell concentration. In this sense some sufficient conditions are given in order to prevent the existence of such kind of unstable states, these conditions primarily involve the phagocyte random motility coefficient and a chemotaxis coefficient included in the model.

As suggested by *in vitro* studies, phagocytic cells (big eaters) may move following a chemotactic impulse generated by the presence of pathogens germs, for this reason most of the authors cited above include the effect of chemotaxis by means of the classical term first introduced by Patlak in 1953 and Keller and Segel in 1970 [45, 73]. Nevertheless, there is no consensus on this assumption, as noted in [52], *in vivo* observations more often show that the phagocytes seem to move within an infected lesion randomly. This is the case in the models introduced in [19, 25, 26].

In this research we propose a mechanism leading to patterns which does not rely on chemotactism. We think the inflammatory response could be modeled by an activator-inhibitor system. Such systems are known to produce the Turing mechanism, that is, periodic stationary solutions. This could possibly explain the patchy nature of Crohn's disease.

4.2 The model

We propose here a reaction-diffusion system modelling the dysfunctional immune response that triggers IBD. As mentioned in the introduction, this kind of system has attracted much interest as a prototypical

model for pattern formation. In this case we refer in particular to inflammatory patterns.

Roughly speaking, the first line of defense of the mucosal immune system is the epithelial barrier which is a polarized single layer covered by mucus in which commensal microbes are embedded. Lowered epithelial resistance and increased permeability of the inflamed and non-inflamed mucosa is systematically observed in patients with Crohn's disease and ulcerative colitis, hence the epithelial barrier becomes leaky and luminal antigens gain access to the underlying mucosal tissue. In a healthy gut, the immune response by means of intestinal phagocytes eliminates the external agents limiting the inflammatory response in the gut. Unfortunately, in a disease-state, the well-controlled balance of the intestinal immune system is disturbed at all levels. This dysfunctional mechanism contributes to acute and chronic inflammatory processes. Indeed, an excessive amount of immune cells migrating to the damaged zone can engage the permeability of the epithelial barrier and thus might allow further infiltration of microbiota which aggravate inflammation. This complex network triggers the initiation of an inflammatory cascade that causes ulcerative colitis and Crohn's diseases (see fig. 4.1).

For the sake of simplicity in this model we will consider just two components varying in time and space: 1. The density of non-resident bacteria leaking into the intestinal tissue through the epithelial barrier noted as β, also refereed as microbiota, pathogens or antigens and 2. The density of immune cells γ which we will often refer to as phagocytic cells. Also, by simplicity we model a portion of the digestive tube as an interval $\Omega \subset \mathbb{R}$ of the real axis, which will be very large. The model reads:

$$
\begin{cases}
\partial_t \beta - d_b \Delta \beta &= r_b \left(1 - \frac{\beta}{b_i}\right) \beta - \frac{a\beta\gamma}{s_b + \beta} + f_e \left(1 - \frac{\beta}{b_i}\right) \gamma, \\
\partial_t \gamma - d_c \Delta \gamma &= f_b \beta - r_c \gamma.
\end{cases}
\tag{4.1}
$$

We will consider Neumann boundary conditions and initial data $\beta(0, x) = \beta_0(x)$ and $\gamma(0, x) = \gamma_0(x)$ for all $x \in \Omega$.

During the immune response there is a first stage where the non-resident phagocytes, i.e. immune cells, migrate from the vasculature into the intestinal mucosa and a second stage where they move to the damaged zone and fight the bacteria. This first stage results from a transport movement through the blood vessels and it is almost instantaneous compared to the second one, so we omit it in this simplified model.

Another main assumption is to consider that immune cells and bacteria move randomly through the damaged tissue and the epithelial barrier. As mentioned in the introduction, it is generally accepted that diffusion provides an adequate description of molecular spreading but, in the case of phagocytic cells, chemotaxis is claimed to be crucial establishing the direction of movement in the sense of the pathogen gradient. However, there are *in vivo* experiments that corroborate our hypothesis [52] and several authors have made similar assumptions [19, 25, 26]. Nevertheless, by neglecting chemotaxis in our model we do not claim that it is an unimportant phenomenon, instead, this assumption must be seen as a simplification and an idealization of the physiological mechanism we seek to describe.

The coefficients $d_b > 0$ and $d_c > 0$ are the diffusion rates of bacteria and phagocytes, respectively. The parameter $r_b > 0$ is associated with the reproduction rate of bacteria.

In healthy conditions the number of bacteria within the lumen remains almost constant, and they are not able to penetrate the epithelial barrier. We associate this quantity to the parameter $b_i > 0$. We remark that this parameter b_i is in some sense a carrying capacity; in fact, in the total absence of the epithelial barrier, the maximum amount of bacteria in the colon would not be greater than $\beta = b_i$. This is the reason why we add the logistic term $1 - \frac{\beta}{b_i}$ in the first equation, [88, 75].

The parameter $f_b > 0$ is associated with the immune response rate of the organism sending cells to fight bacteria in the damaged zones. In others words, phagocytes appear as soon as the presence of pathogens is detected.
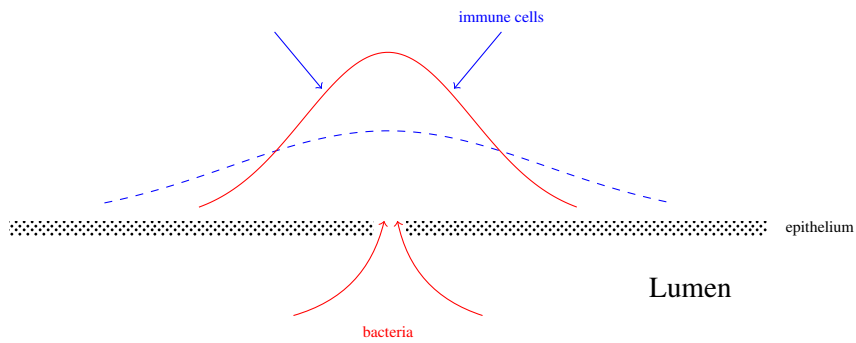
(a) Bacteria (red line) break through the epithelium (dotted zone); phagocytes (blue dashed line) are recruited in order to neutralize them.
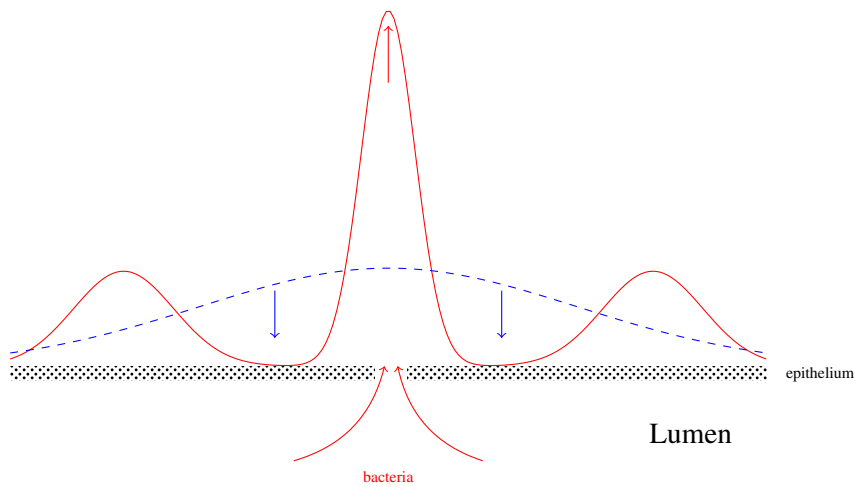
(b) Phagocytes spread rapidly through blood vessels. A high spot of bacteria remains with a lateral inhibition by phagocytes.
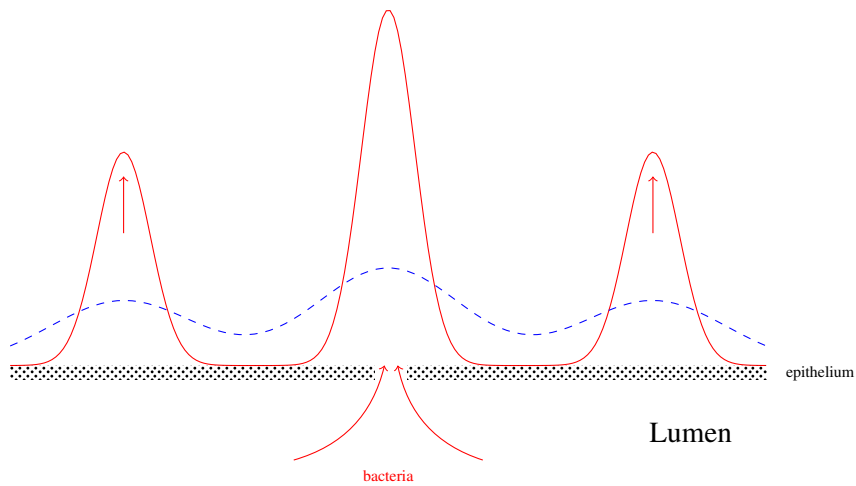
(c) Other spots appear.

Figure 4.1: Initiation of the inflammatory process.

The term $-\frac{a\beta\gamma}{s_b+\beta}$ with $a > 0$ and $s_b > 0$ corresponds to the effect of the immune system on the pathogen agents. In particular $\frac{a\beta}{s_b+\beta}$ is the phagocytosis rate or intake rate. It suggests that the attack rate of immune cells on bacteria varies with the density of the pathogen. This functional response term takes into account the rate p_c at which phagocytes encounter a bacterium per unit of bacteria density, which is $p_c := \frac{a}{s_b}$ and the average time τ that it takes a phagocyte to neutralize a bacterium (or handling time) which can be computed as $\tau := \frac{1}{a}$. Experiments presented in [54, 81] reflect this dynamic. In the mathematical literature this type of term is often referred as a Holling Type II functional response, see [38, 75].

We consider $f_e > 0$ as a measure of the negative effect of the phagocyte's concentration for the epithelial resistance, and therefore it has a positive impact on the bacteria density i.e. the larger the epithelial gap, the more bacteria there are, the more immune cells there are drifting to the damaged zone and the more porous is the epithelium and so on.

Finally, a self-regulation function of anti-inflammatory cells limits their life-time, so immune cells have an intrinsic death rate which is noted in the model as $r_c > 0$.

4.3 On Turing Patterns

Since one of our main interest is to explain patchy inflammatory bowel patterns often observed in patients suffering from Crohn's disease, we seek to demonstrate that the model we propose may present Turing-type instabilities under certain conditions. This denomination is due to Alan Turing who was the first to describe spatial patterns caused by the effects of diffusion in his article on morphogenesis theory published in 1952, [86].

Roughly speaking, a Turing system consist of an activator that must diffuse at a much slower rate than an inhibitor to produce a pattern. We recall that diffusion causes areas of high concentration to spread out to areas of low concentration. In such kinds of systems the activator component must increase the production of itself while the inhibitor restrains the production of both. Turing's analysis shows that in certain regimes those systems are unstable to small perturbations, leading to the growth of large scale patterns.

In the model we previously introduced, the bacteria are the activator and the immune cells the inhibitor. Indeed, the bacteria reproduce at a certain rate r_b and immune cells neutralize bacteria by phagocytosis (Holling-type term) and self-regulate their own life-time r_c. In practice, we should look for steady state solutions of eq. (4.1) which are linearly unstable, i.e. such that there are perturbations for which the linearized system has exponentially growing solutions in time. To be sure that a Turing-type phenomena is occurring it is important to exclude the cases where the corresponding growth modes are unbounded, that is solutions with infinitely high frequencies and also the cases in which solutions blow up or go to extinction [75].

In section 4.4.2 we study the conditions leading to the observation of Turing phenomena in our model.

4.4 Results

4.4.1 Non-negativity property and boundedness

We begin by establishing some elementary properties in the model to guarantee eq. (4.1) accuracy as a population dynamics model. In other words, it is important that whenever the initial data have a reasonable biological meaning, the solution of the differential equation inherits that property. We start by a non-negativity property:

Proposition 20 *Provided that the initial condition $(\beta_0(x), \gamma_0(x))$ is non-negative the solutions of eq. (4.1) remain non-negative for every $t > 0$.*

Similarly, we establish a boundedness property associated with the carrying capacity of the population environment:

Proposition 21 *If $\beta_0(x) < b_i$ then for all $t > 0$ one has $\beta(t,x) < b_i$. Moreover, if $\gamma_0(x)$ is bounded, then $\gamma(t, \cdot)$ remains bounded in the L^2-norm in Ω for every $t > 0$.*

4.4.2 Stability analysis

Let us study now the steady states of the model and their stability properties. The equation (4.1) has two non-negative homogeneous steady states. One of them is the trivial solution $(\beta, \gamma) = (0, 0)$ associated with the absence of bacteria and immune cells. The other one, that we denote $(\beta, \gamma) = (\overline{\beta}, \overline{\gamma})$, satisfies:

$$0 = (r_b + f_e\kappa)\left(1 - \frac{\overline{\beta}}{b_i}\right) - \frac{a\kappa\overline{\beta}}{s_b + \overline{\beta}}, \tag{4.2}$$

where $\kappa := \frac{f_b}{r_c}$ and $\overline{\gamma} = \kappa\overline{\beta}$.

We remark that $(0, 0)$ is unstable. Indeed, the linearized matrix around this steady state has negative determinant and thus an eigenvalue with positive real part. For the non-trivial equilibrium point $(\overline{\beta}, \overline{\gamma})$ the stability analysis is less straightforward. The following proposition establishes the conditions leading to the stability of this steady state.

Proposition 22 *Consider the O.D.E system associated with eq. (4.1) with non-negative real parameters $a, r_b, r_c, f_b, f_e, b_i$ and s_b,*

$$\begin{cases} \partial_t\beta &= r_b\left(1 - \dfrac{\beta}{b_i}\right)\beta - \dfrac{a\beta\gamma}{s_b + \beta} + f_e\left(1 - \dfrac{\beta}{b_i}\right)\gamma \\ \partial_t\gamma &= f_b\beta - r_c\gamma \end{cases}. \tag{4.3}$$

This system has a unique positive steady state solution $\big(\beta(t), \gamma(t)\big) = (\overline{\beta}, \overline{\gamma})$ which is linearly stable if and only if

$$\frac{a\kappa\overline{\beta}^2}{(s_b + \overline{\beta})^2} - r_b\frac{\overline{\beta}}{b_i} - f_e\kappa < r_c. \tag{4.4}$$

We conjecture that the model might show some unexpected behavior around this steady state which could be at the origin of patchy inflammatory patterns. Hence, let us focus on conditions leading the formation of Turing patterns for the reaction diffusion system (4.1), that is perturbations around the steady state $(\overline{\beta}, \overline{\gamma})$ such that the linearized system has exponential growth in time and for which the corresponding growth modes are bounded. The following proposition establishes the necessary conditions for the occurrence of such phenomenon.

Proposition 23 *Consider eq. (4.1) and its unique positive homogeneous steady state solution $(\overline{\beta}, \overline{\gamma})$; assume that there exist real non-negative values of the parameters $a, r_b, r_c, s_b, f_e, f_b, b_i$ such that the following condition holds:*

$$0 < \frac{a\kappa\overline{\beta}^2}{(s_b + \overline{\beta})^2} - r_b\frac{\overline{\beta}}{b_i} - f_e\kappa < r_c \tag{4.5}$$

Then for $\frac{d_b}{d_c}$ small enough the reaction diffusion system (4.1) shows Turing instabilities around this steady state.

4.5   Parameters of the model

In this section we want to estimate the values of the parameters of the model and to prove the non emptiness of the parameter set defined by (4.5). As long as it is possible we will rely on values obtained from real observations or *in vitro* experiments. However, in some cases the exact values are unknown due to the difficulty of measuring them *in vivo* or even *in vitro*.

Let us start with an estimation of the reproduction rate of the bacteria, represented in our model as r_b. The bacterium's generation time, which is the time it gets to the population to double the number of individuals, might vary from 12 minutes to several hours depending on temperature, nutrients, culture medium, among others factors. For E. Coli, for instance, it is around 20 minutes in standard conditions, [47]. We can then consider that the evolution of bacteria population is given by $\partial_t b = r_b b$ and so $r_b = \frac{ln(2)}{20}$ measured in bacteria per minute. That gives us an approximate value $r_b = 3.47 * 10^{-2} \ u/min$ which is in the estimated range of values given in [52] for this parameter.

Similarly, it is known that in healthy conditions, phagocytes have, in average, a half-life of two days [49], and so from $\partial_t c = -r_c c$ we get $r_c = \frac{ln(2)}{2880}$ cells per minute which means that the death rate of phagocytes is ideally of the order of $10^{-4} \ u/min$, which coincides with that considered in [92] for immune cells in diabetic wounds or in [52] for bacterial infection causing tissue inflammation. However, there is no consensus, some authors assume this parameter to be of the order of $10^{-3} u/min$ in the inflammatory response framework [18] or even of the order of $10^{-6} u/min$ in the case of early atherosclerosis [17]. For such parameters, corresponding to a healthy organism, we do not expect to observe Crohn's disease. Indeed, the mechanism we describe below occurs with $r_c = 2 \times 10^{-2} \ u/min$ (see table 4.1). For $r_c = 10^{-3} \ u/min$, the range of parameters for which a Turing pattern occurs is quite narrow (see fig. 4.3).

The diffusion coefficient of immune cells might also vary according to the type of cell and the part of the body where they act. In the consulted literature the value of this parameter varies from 10^{-12} m^2/min to $10^{-10} \ m^2/min$ depending on the context [19, 26, 52, 80]. In the absence of experimental data providing more precise information about the order of this parameter in the particular case of bacterial infection in the intestinal track, we consider this coefficient to remain within this range in damaged areas of the intestine.

Although there is not precise information concerning the diffusion rate of bacteria through the epithelial barrier, it is known that in aqueous solutions like the lumen, the diffusion rate might vary from 10^{-11} m^2/min to $10^{-8} \ m^2/min$ depending on the type of bacteria. However, in a non-liquid framework, which is the case of bacteria penetrating through the epithelial barrier, motility should be reduced.

We will now roughly compute a value for the parameter a, we suppose that there is a significant density of bacteria in a certain position $x = x_0$, and we study the time evolution of the population within this point. If β is large enough, the term $1 - \frac{\beta}{b_1}$ is negligible, moreover the term $-\frac{a\beta\gamma}{s_b+\beta}$ tends to approach $-a\gamma$, so we can approximately write

$$\partial_t \beta(t, x_0) = -a\gamma(t, x_0). \tag{4.6}$$

Let us now define τ as the average time it takes a phagocyte to neutralize a bacterium, which is around 3 minutes in the *in vitro* observations, which implies that

$$\beta(t + \tau, x_0) = \beta(t, x_0) - \gamma(t, x_0) \tag{4.7}$$

and consequently $\partial_t \beta(t, x_0) \approx -\frac{\gamma(t,x_0)}{\tau}$. Replacing this into eq. (4.6) we conclude that a is of the order of $\frac{1}{\tau}$ units per minute. An underlying assumption here is that one phagocyte is needed in order to neutralize one bacterium. If two phagocytes were needed, it would give rise to a quadratic term γ^2 for example. This assumption is based on *in vitro* observations, but it might become unrelevant in particular
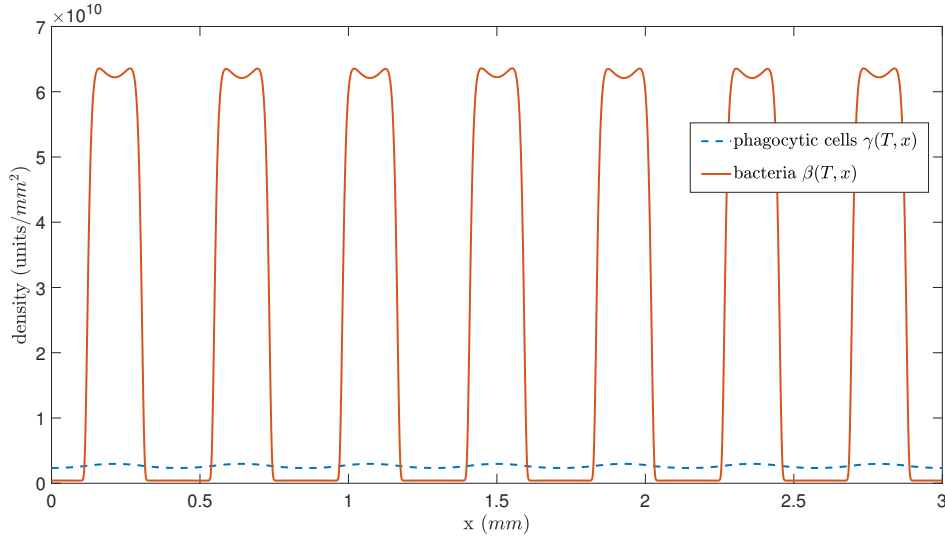
Figure 4.2: Bacteria (red line) and phagocytes (blue dashed line) after a time-lapse of 2 weeks with an initial bacterial infection $\beta_0(x) = 10^9 \times \mathbb{1}_{[1.495,1.505]}$ and $\gamma_0(x) = 0$. As explained in the section 4.3, here the bacteria behave as the activator and the immune cells as the inhibitor in the sense of the Turing mechanism. The fact that macrophage cells spread faster than bacteria plays a fundamental role in the emergence of spatial patterns, as it is stated in the Proposition 23.

frameworks or it might depend on the type of phagocytes. We stick to it here in order to simplify the analysis.

The density of bacteria in the lumen is approximately $b_i = 10^{17} \ u/m^3$. At the positive equilibrium stage $(\overline{\beta}, \overline{\gamma})$, which is associated to an inflammatory phase, we suppose that around 30% of the total density of bacteria within the lumen might penetrate the epithelial barrier without going out of control. Therefore, we set $\overline{\beta} = 0.3 \times b_i$ units of bacteria. Even though we have no exact data concerning the density of immune cells in the damaged zone, the *in vitro* experiments suggest that during the inflammation stage it is around ten times less than the bacteria density, this is quite natural considering that the size of a phagocyte is much larger than the size of a bacterium. Hence, we set the hypothesis that $\kappa = \frac{1}{10}$ which means that at the equilibrium point, $\overline{\beta} = 10\overline{\gamma}$ and consequently $\overline{\gamma} = 3 \cdot 10^{-2} \times b_i$. Taking this into account from the equilibrium condition we have that $f_b = 10^{-1} r_c$ measured in units per minute.

The parameter f_e is finally computed so that (4.2) holds at the equilibrium state.

4.6 Numerical simulations

We perform some numerical simulations in MATLAB by mean of a semi-implicit scheme to solve the system of equations (4.1). The results are shown in fig. 4.2. We have considered the parameter values presented in the table 4.1 which were estimated in the previous section. For these values, the condition (4.5) associated to a Turing phenomenon occurrence established in the Proposition 23 is verified. However, there is a whole family of parameters verifying (4.5), as shown in fig. 4.3.

For the simulations we have considered an initial datum with no phagocytes presence and a tiny spot of bacteria concentrated in the middle of the domain Ω. This might be understood as a slight leak of bacteria from the lumen through the epithelium. The activator-inhibitor dynamics generated by the body's immune response to the presence of bacteria and the contrast in the propagation rates of the two
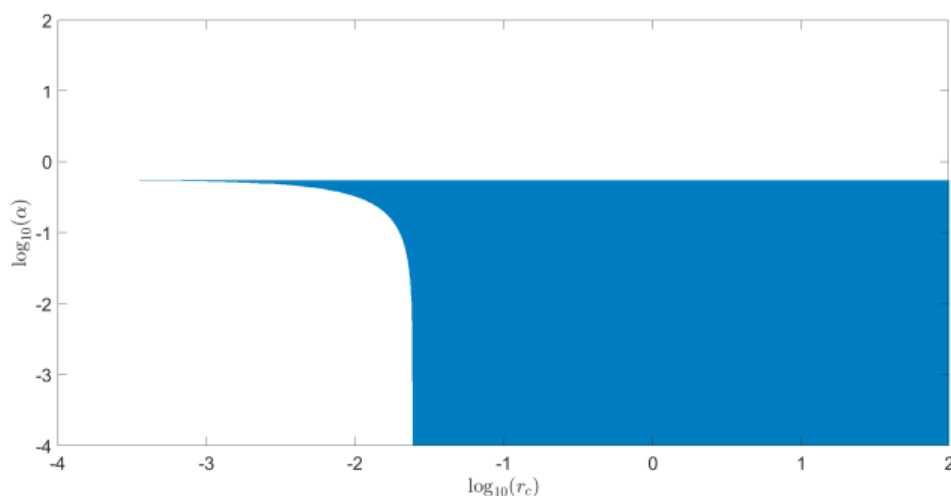
Figure 4.3: Region (blue) defined by the parameters r_c and a that verify condition (4.5) leading to Turing patterns observation

actors of the system is the reason why the patterns emerge in fig. 4.2 after a certain time. This behavior is definitely associated with a Turing phenomenon.

We remark that the values we assign to the diffusion coefficients remains within the range estimated in the previous section. However, from the mathematical point of view what is really important in order to ensure the conditions leading to the Turing patterns is the smallness of the ratio $\delta = \frac{d_b}{d_c}$. To change those values by preserving δ only represents a spatial rescaling that does not affect the pattern formation.

4.7 Proof of the results

Proof of the Proposition 20

Proof: Assume first that $\inf_\Omega \beta_0 > 0$ and $\inf_\Omega \gamma_0 > 0$. Consider $\bar{t} > 0$ the first instant when either $\beta(t, x)$ or $\gamma(t, x)$ became non-positive, then for some $x^* \in \overline{\Omega}$ one has $\beta(\bar{t}, x^*)\gamma(\bar{t}, x^*) = 0$. The Hopf lemma and the Neumann boundary conditions exclude that this minimum is reached at the boundary.

Table 4.1: Assigned values for the parameters of the model (4.1)

parameter	interpretation	value	units
r_b	Reproduction rate of bacteria	0.0347	(u/min)
r_c	Intrinsic death rate of phagocytes	0.02	(u/min)
d_b	Diffusion rate of bacteria	10^{-13}	(m^2/min)
d_c	Diffusion rate of phagocytes	10^{-10}	(m^2/min)
b_i	Density of bacteria in the lumen	10^{17}	(u/m^3)
f_b	Immune response rate	0.002	(u/min)
a	Coefficient proportional to the rate of phagocytosis ($a = s_b p_c$)	0.3129	(u/min)
	it is also inversely proportional to the handling time ($a = \frac{1}{\tau}$)		
s_b	Proportionality coefficient between p_c and a	10^{15}	(u/m^3)
f_e	Related to the porosity of the epithelium	0.0856	(u/min)

Hence $x^* \in \Omega$.

If $\gamma(\bar{t}, x^*) = 0$, then as $\beta \geq 0$ over $(0, t^*) \times \Omega$, one has

$$\partial_t \gamma - d_c \Delta \gamma \geq -r_c \gamma \quad \text{in } (0, t^*) \times \Omega.$$

The strong parabolic maximum principle then yields $\gamma \equiv 0$ on this set, which contradicts the positivity of the initial condition.

Next, if $\beta(\bar{t}, x^*) = 0$ and $\gamma(\bar{t}, x^*) > 0$, the equation satisfied by β reads:

$$0 \geq \partial_t \beta(\bar{t}, x^*) - d_b \Delta \beta(\bar{t}, x^*) = f_e \gamma(\bar{t}, x^*) > 0$$

a contradiction.

We get the result for general initial data by approximation.

$\square$

Proof of the Proposition 21

Proof: The argument of this proof is similar to the one used to prove the non-negativity property. Indeed, consider $\bar{t}$ the first instant when β rises the value b_i, then there exists $x^* \in \Omega$ such that $\beta(\bar{t}, x^*) = b_i$ and one has $\partial_t \beta(\bar{t}, x^*) \geq 0$. Nevertheless from the equation associated with β one conclude that $\partial_t \beta(\bar{t}, x^*) = -\frac{a\beta(\bar{t},x^*)\gamma(\bar{t},x^*)}{s_b+\beta(\bar{t},x^*)} < 0$ from the positivity property. So we get a contradiction which implies that for all $t > 0$ one has necessarily $\beta(t, x) < b_i$.

The boundedness of γ in the L^2-norm follows directly from the boundedness of β and γ_0. In fact multiplying by γ in the second equality of eq. (4.1), integrating by parts and applying Holder inequality one gets that

$$\frac{1}{2}\partial_t \|\gamma\|_{L^2(\Omega)}^2 + d_c\|\nabla\gamma\|_{L^2(\Omega)}^2 \leq f_b\|\beta\|_{L^2(\Omega)}\|\gamma\|_{L^2(\Omega)} - r_c\|\gamma\|_{L^2(\Omega)}^2 \tag{4.8}$$

from where we deduce

$$\partial_t \|\gamma\|_{L^2(\Omega)} \leq f_b\|\beta\|_{L^2(\Omega)} - r_c\|\gamma\|_{L^2(\Omega)}. \tag{4.9}$$

Keeping the notation $\kappa := \frac{f_b}{r_c}$ introduced in the section 4.4.2, this last inequality implies that

$$\|\gamma\|_{L^2(\Omega)} \leq \max\{\kappa b_i, \|\gamma_0\|_{L^2(\Omega)}\} \tag{4.10}$$

which completes the proof. $\square$

In order to simplify the notations in the proofs of propositions 22 and 23 we will denote $\theta := \frac{\bar{\beta}}{b_i}$ and we will also keep the notation $\kappa := \frac{f_b}{r_c}$.

Proof of the Proposition 22

Proof: The existence of such a positive steady state follows from the analysis of (4.2). Let us define $F(\beta) = (r_b + f_e\kappa)\left(1 - \frac{\beta}{b_i}\right) - \frac{a\kappa\beta}{s_b+\beta}$. From the positivity of the parameters of the model we have that $F(0) > 0$ and $F(b_i) < 0$, this means that there are at least one positive value $\bar{\beta} \in (0, b_i)$ that satisfies $F(\bar{\beta}) = 0$ or equivalently (4.2). Moreover, since the derivative of F is strictly negative we deduce that it has at most one root which leads to the uniqueness of $\bar{\beta}$.

Let us now study the conditions leading to the stability of this steady state. In order to simplify the notations we will define $\theta := \frac{\bar{\beta}}{b_i}$. We also define $\mathbf{M}$ as the matrix of the linearized system around this positive steady state $(\bar{\beta}, \bar{\gamma})$

$$\mathbf{M} := \begin{pmatrix} r_b(1-2\theta) - \dfrac{as_b\kappa\overline{\beta}}{(s_b+\overline{\beta})^2} - f_e\kappa\theta & -\dfrac{a\overline{\beta}}{s_b+\overline{\beta}} + f_e(1-\theta) \\ f_b & -r_c \end{pmatrix}.$$

We compute the determinant and the trace of this matrix

$$tr(\mathbf{M}) = \dfrac{a\kappa\overline{\beta}^2}{(s_b+\overline{\beta})^2} - r_b\theta - f_e\kappa - r_c,$$

$$det(\mathbf{M}) = r_c r_b\theta + f_b f_e\theta + \dfrac{af_b s_b\overline{\beta}}{(s_b+\overline{\beta})^2}.$$

From the positivity of the parameters of the model it is clear that the determinant of $\mathbf{M}$ is positive, therefore in order to have linear stability around $(\overline{\beta},\overline{\gamma})$ it is necessary and sufficient to impose the negativity of the trace of $\mathbf{M}$ which is equivalent to the condition (4.4).

$\square$

Proof of the Proposition 23

Proof: We linearize the system around $(\overline{\beta},\overline{\gamma})$. For the sake of simplicity we keep the notation $\beta(t,x),\gamma(t,x)$ for the linearized variables

$$\begin{cases} \partial_t\beta - d_b\Delta\beta = \left(r_b(1-2\theta) - \dfrac{a\kappa s_b\overline{\beta}}{(s_b+\overline{\beta})^2} - f_e\kappa\theta\right)\beta + \left(-\dfrac{a\overline{\beta}}{s_b+\overline{\beta}} + f_e(1-\theta)\right)\gamma \\ \partial_t\gamma - d_c\Delta\gamma = f_b\beta - r_c\gamma \end{cases} \tag{4.11}$$

We are seeking in particular for solutions with exponential growth in time, so we consider that

$$\beta(t,x) = e^{\lambda t}B(x) \; ; \; \gamma(t,x) = e^{\lambda t}C(x) \tag{4.12}$$

with $\lambda > 0$. This means that $B(x)$ and $C(x)$ should satisfy the fallowing problem

$$\begin{cases} -d_b\Delta B(x) = \left(r_b(1-2\theta) - \dfrac{a\kappa s_b\overline{\beta}}{(s_b+\overline{\beta})^2} - f_e\kappa\theta - \lambda\right)B(x) + \left(-\dfrac{a\overline{\beta}}{s_b+\overline{\beta}} + f_e(1-\theta)\right)C(x) \\ -d_c\Delta C(x) = f_b B(x) + (-r_c - \lambda)C(x) \end{cases} \tag{4.13}$$

or equivalently that they are eigenfunctions associated with the positive eigenvalue λ. We consider in particular Fourier modes of the form

$$B(x) = Be^{i\xi x} \; ; \; C(x) = Ce^{i\xi x},$$

and we replace it in eq. (4.13) to obtain the following homogeneous linear system of equations

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} r_b(1-2\theta) - \dfrac{a\kappa s_b\overline{\beta}}{(s_b+\overline{\beta})^2} - f_e\kappa\theta - \lambda - d_b\xi^2 & -\dfrac{a\overline{\beta}}{s_b+\overline{\beta}} + f_e(1-\theta) \\ f_b & -r_c - \lambda - d_c\xi^2 \end{pmatrix} \begin{pmatrix} B \\ C \end{pmatrix}$$

Let us call $\mathbf{M}_{\lambda,\xi}$ the matrix associated to the previous linear system. It can be written in terms of ξ, λ and the matrix $\mathbf{M}$ introduced before in the proof of the Proposition 22

$$\mathbf{M}_{\lambda,\xi} = \begin{pmatrix} \mathbf{M}_{(1,1)} - \lambda - d_b\xi^2 & \mathbf{M}_{(1,2)} \\ \mathbf{M}_{(2,1)} & \mathbf{M}_{(2,2)} - \lambda - d_c\xi^2 \end{pmatrix}.$$

We recall that

$$\mathbf{M}_{(1,1)} = r_b(1 - 2\theta) - \frac{a s_b \kappa \overline{\beta}}{(s_b + \overline{\beta})^2} - f_e \kappa \theta, \tag{4.14}$$

$$\mathbf{M}_{(1,2)} = -\frac{a \overline{\beta}}{s_b + \overline{\beta}} + f_e(1 - \theta), \tag{4.15}$$

$$\mathbf{M}_{(2,1)} = f_b, \tag{4.16}$$

$$\mathbf{M}_{(2,2)} = -r_c. \tag{4.17}$$

In other words we look for a certain λ with positive real part and ξ^2 for which $\det(\mathbf{M}_{\lambda,\xi}) = 0$. The determinant of $\mathbf{M}_{\lambda,\xi}$ is a quadratic polynomial function in λ

$$det(\mathbf{M}_{\lambda,\xi}) = \lambda^2 + a_1\lambda + a_2 \tag{4.18}$$

with coefficients

$$
\begin{aligned}
a_1 &= -tr(\mathbf{M}) + (d_b + d_c)\xi^2, \\
a_2 &= \det(\mathbf{M}) - (\mathbf{M}_{(1,1)}d_c + \mathbf{M}_{(2,2)}d_b)\xi^2 + d_b d_c \xi^4.
\end{aligned}
$$

Since the right-hand side inequality in (4.5) ensures that $tr(\mathbf{M}) < 0$, we conclude that $a_1 > 0$. Hence, the polynomial associated to $\det(\mathbf{M}_{\lambda,\xi})$ can have a positive root λ if and only if $a_2 < 0$. The term a_2 is itself a quadratic polynomial in ξ^2 with positive second order coefficient. For the sake of simplicity we will define $\delta := \frac{d_b}{d_c}$, and we will study the sign of $\frac{a_2}{d_b d_c}$ which has roots explicitly given by

$$\Lambda_{\pm} = \frac{\mathbf{M}_{(1,1)} + \delta\mathbf{M}_{(2,2)}}{2 * d_b}\left[1 \pm \sqrt{1 - \frac{4\det(\mathbf{M})\delta}{(\mathbf{M}_{(1,1)} + \delta\mathbf{M}_{(2,2)})^2}}\right]. \tag{4.19}$$

In the regime δ small enough the Taylor expansion gives us the following approximate values

$$\Lambda_{-} = -\frac{\det(\mathbf{M})}{d_c\mathbf{M}_{(1,1)}} \tag{4.20}$$

$$\Lambda_{+} = \frac{\mathbf{M}_{(1,1)}}{d_c\delta} \tag{4.21}$$

The left-hand side inequality in (4.5) guarantees that Λ_+ is positive and since δ can be as small as desired, $\Lambda_+ >> 1$ and the interval (Λ_-, Λ_+) where a_2 is negative is large enough.

In other words, there exists a positive real λ and Fourier modes for which $\det(\mathbf{M}_{\lambda,\xi^2}) = 0$ and consequently we can find exponential growth in time solutions to the linearized system around the steady state $(\overline{\beta}, \overline{\gamma})$. However, the Fourier modes for which this condition holds are bounded.

We have showed the existence of perturbations such that the linearized system has exponential growth in time. The frequency of the perturbations cannot be infinity and from Proposition 20 and 21 neither extinction nor blow-up are possible. Hence, we have finally proved the formation of Turing Patterns. Those patterns, which can be observed in the Figure 4.2, correspond to a non-constant steady state. Such solutions are known to be stable with respect to periodic perturbations. But their stability with respect to more general perturbations, such as compactly supported ones, is known to be a tough question and remains unaddressed in this research.

$$\square$$

4.8 Conclusion

This work remains a simplified approach to the question of modelling the inflammatory response in Crohn's disease. We have made several hypotheses with the aim of globally understanding the biological mechanism behind the abnormal body reaction leading to the disease but staying relatively simple in terms of the number of variables and equations.

Though we have tried to consider parameter values true to medical and biological observations, we highlight the qualitative results over quantitative ones. In this sense, obtaining a Turing mechanism through our model might explain the patchy inflammatory patterns observed in patients suffering from Crohn's disease and must be interpreted as another step in the aim to fully understand this illness and its causes.

It remains a question concerning the ulcerative colitis since it has several common factors that relate it to Crohn's disease but also others that set them apart. It might be interesting to study the possibility of modelling the ulcerative colitis by mean of the same system of equations in a different parameter regime and eventually finding responses to help doctors with early diagnosis or treatments.

Appendix A

1D Numerical simulations

We aim in this section to compare, through an example, the performance of the numerical algorithm we propose and describe in section 2.5, with other well known optimization algorithms to solve general nonlinear problems under constraints. More precisely, we will consider the following numerical methods:

- *Method 1* The numerical algorithm presented in chapter 2 and described in detail in section 2.5, that we will refer to from now on as *our algorithm*.

- *Method 2* The *interior-point* method, which is used to solve optimization problems with linear equality and inequality constraints by applying Newton's method to a sequence of equality constrained problems. For more detailed description of this method see for instance [10].

- *Method 3* The *sequential quadratic programming* (SQP), which solves a sequence of optimization sub-problems, each of which optimizes a quadratic model of the objective function subject to a linearization of the constraints, see for instance [72]

- *Method 4* The *simulated annealing* method, which is a probabilistic technique to approximate global optimization in a large search space. See for instance [56] for more details on this technique.

We will use the MATLAB platform to perform the simulations. Methods 2 and 3 are already coded in the MATLAB function `"fmincon"` while methods 1 and 4 were coded for the experiment.

A.1 Setting the data

In what follows we will keep the notation introduced in chapter 2. Let us consider $\Omega = [-50, 50]$, and $m = 13$ such that the admissible set is defined as follows:

$$\mathcal{A}_{13} = \left\{ u_0 \in L^1(\Omega) : 0 \le u_0(x) \le 1, \text{ and } \int_\Omega u_0(x)dx = 13 \right\}.$$

Note that this set is defined by two inequalities and an equality constraint. We aim to maximize the quantity $\mathcal{J}_T(u_0) := \int_\Omega u(T, x)dx$ for $T = 25$. We recall that u is the solution of the following reaction-diffusion equation with a bistable reaction term, and u_0 as initial datum,

$$\begin{cases} \partial_t u - \Delta u = u(1 - u)(u - 0.25) & \text{in } \mathbb{R}_+ \times \Omega, \\ \\ u(0, x) = u_0(x) & \text{in } \Omega, \\ \\ \frac{\partial u}{\partial \nu}(t, x) = 0 & \text{in } \mathbb{R}_+ \times \partial\Omega. \end{cases} \tag{A.1}$$

71

In order to compare the performance of the four algorithms under the same conditions, we consider the same discretization of Ω. Moreover, the solution of the equation is systematically computed by means of the Crank-Nicolson method, and for the initialization we will consider the same u_0^0 given by a single block of mass 13, as described in fig. A.1.

The value of the objective function at each iteration is numerically approximated by the rectangle rule. In particular for the initialization we have $\mathcal{J}_{25}(u_0^0) = 29.42$.
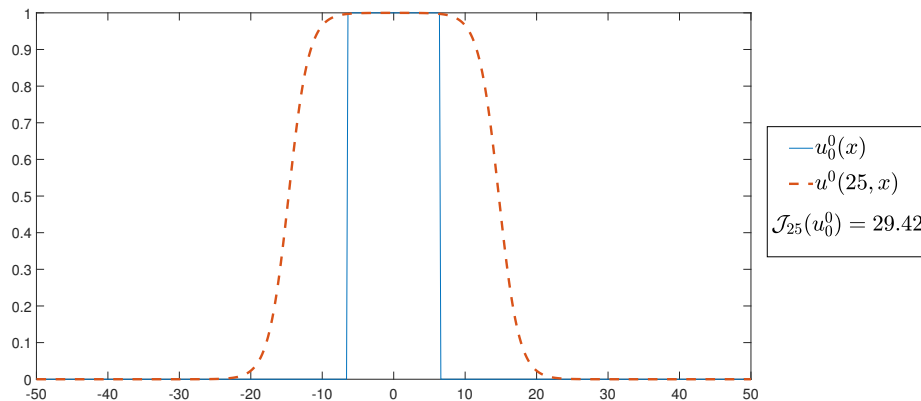

Figure A.1: Initialization of the numerical algorithm with $u_0^0 = \mathbb{1}_{[-6.5, 6.5]}$.

A.2 Results

The results of the simulations are shown in fig. A.2 and table A.1. For this example, our algorithm turns out to be faster than other well-known algorithms. Moreover, the evaluation of the objective function differs in less than 1% with respect to the best result obtained with the sequential quadratic programming method which takes more than twice the run-time of our algorithm.

Table A.1: Comparing algorithms

| Algorithm	Objective function $\mathcal{J}_{25}(\overline{u}_0)$	Run-time (in seconds)
our algorithm	77.9864	452
interior point	65.6175	676
Sequential quadratic programming	78.7672	1342
simulated annealing	77.6238	4148

Though the solution given by the sequential quadratic programming method is clearly more regular than the others, the profile of the local optimizers found by simulated annealing and by our algorithm, do not seem to be far from this profile. It is important, however, to highlight that since uniqueness is not guaranteed in general, one can not ensure that the algorithms have converged to a global maximizer but only to a local one.
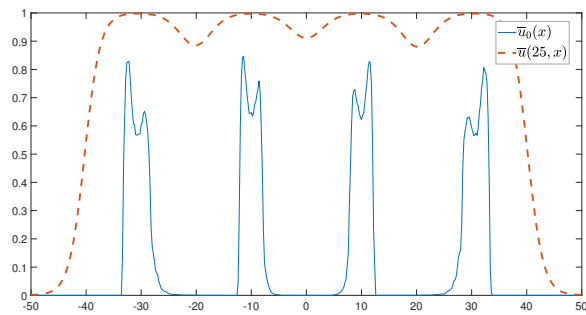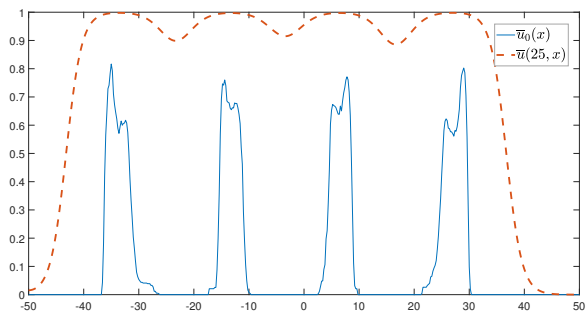
(a) Our algorithm
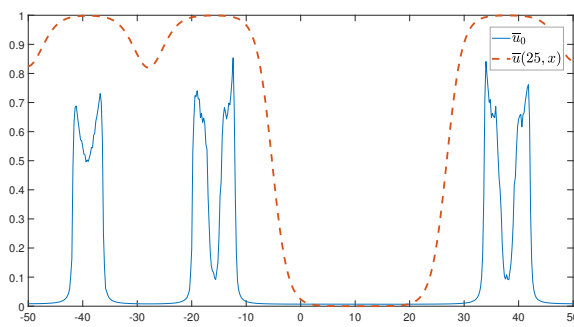
(b) Interior-point method
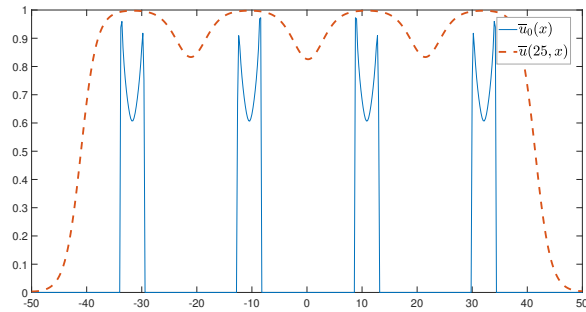
(c) Sequential quadratic programming

(d) Simulated annealing

Figure A.2: Optimum found by means of the four different numerical algorithms.

Appendix B

2D Numerical simulations

In this section we present simulations showing the performance of the algorithm, introduced and described in section 2.5, in the two-dimensional case. To solve the reaction-diffusion equation, we consider the alternating direction implicit method (ADI) which is a classic method to solve parabolic and elliptic partial differential equations in two or three dimensions. As in the previous appendix, the algorithm and routines were coded in MATLAB.

We consider a square domain $\Omega = [-10, 10] \times [-10, 10]$, discretized uniformly by squares of side $dx = 0.4$. We fixed $T = 30$ and assume a diffusion coefficient $\sigma = 1$ and a bistable reaction term

$$f(u) = u(1 - u)(u - 0.25).$$

B.1 Example 1

The algorithm is initialized with a ball of full density spotted in the middle of the domain Ω, the mass is fixed to $m = 6\pi$, see fig. B.1(a). After 27 iterations, the algorithm converges to the local optimum showed in fig. B.1(b). The evolution of the objective function $\mathcal{J}_{30}$ through iterations is showed in fig. B.2

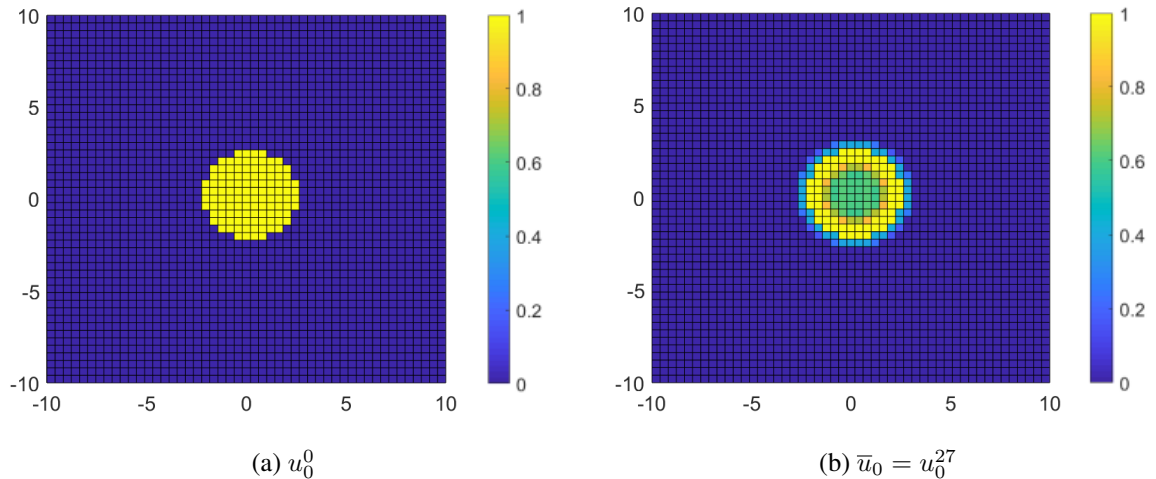(a) u_0^0 　　　　　　　　　　　　　　　　　(b) $\overline{u}_0 = u_0^{27}$

Figure B.1: In the left-hand side is showed the input of the algorithm, given by the ball of ratio $r = \sqrt{6}$ centered at the origin. In the right-hand side, the local optimum found by the numerical algorithm after 27 iterations, which remains radial but which is no longer a bang-bang distribution.
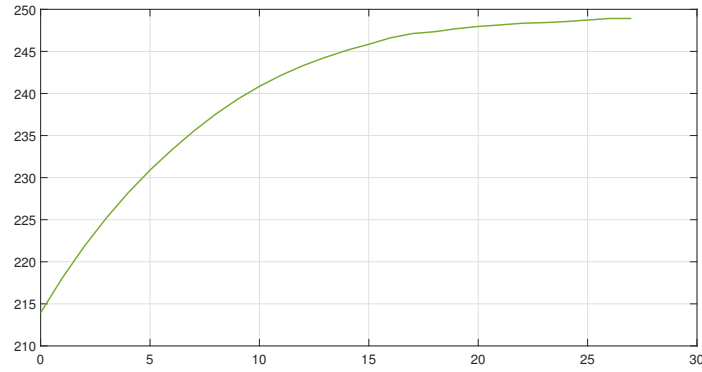
Figure B.2: Evolution of the objective function from the initialization $\mathcal{J}_{30}(u_0^0) = 213.9$ to the last iteration $\mathcal{J}_{30}(u_0^{27}) = 248.9$.

One might see that the local optimum found by the numerical algorithm is no longer a bang-bang function but a circular ball with less mass in the middle and a slightly bigger ratio. Looking at the adjoint state defined as the solution of the eq. (2.8), associated to this initial data, one might see that the area in the middle of the circle corresponds to a set where the adjoint state remains constant, see fig. B.3.
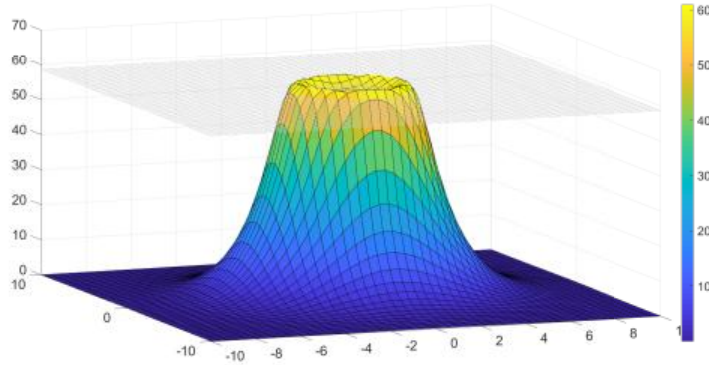

Figure B.3: The figure shows the surface given by the solution $\overline{p}(0, \mathbf{x})$ of the adjoint problem defined by the eq. (2.8) associated to the initial data $\overline{u}$ found by our algorithm. The plane colored in gray, is associated to the value c described in Theorem 8 and thus for every $\mathbf{x} \in \Omega$ such that $p_0(\mathbf{x}) = c$, one has $0 < \overline{u}_0(\mathbf{x}) < 1$, see fig. B.1 (b).

B.2 Example 2

In this case we refine the grid fixing $dx = 0.27$ and we consider an initial data which is a band of mass $m = 16.21$ and full density dividing our domain into two equal regions of zero density, see fig. B.4 (a). The algorithm converges after 37 iterations and the local optimum is showed in fig. B.4 (b). The corresponding variations of the objective function is showed in fig. B.5

Even though the initial mass is clearly not enough to guarantee convergence to 1, the value of the objective function increase considerably through the iterative process. Also, the geometry of the local optimum found by the algorithm is interesting since it shows regions of zero density in the middle of

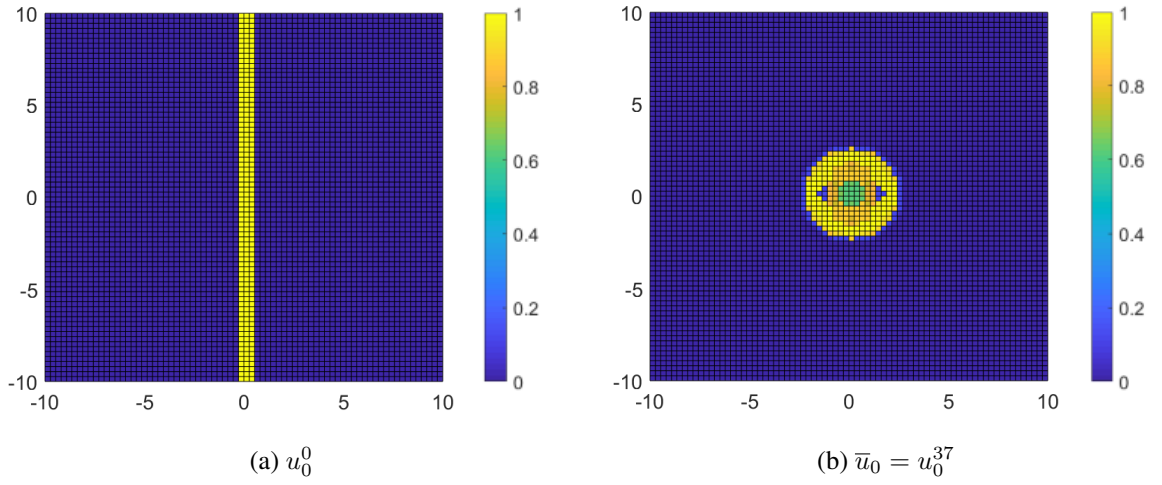(a) u_0^0 (b) $\overline{u}_0 = u_0^{37}$

Figure B.4: In the left-hand side is showed the input of the algorithm, given by the stripe of width $r = 0.81$ centered at the origin. In the right-hand side, the local optimum found by the numerical algorithm after 37 iterations.
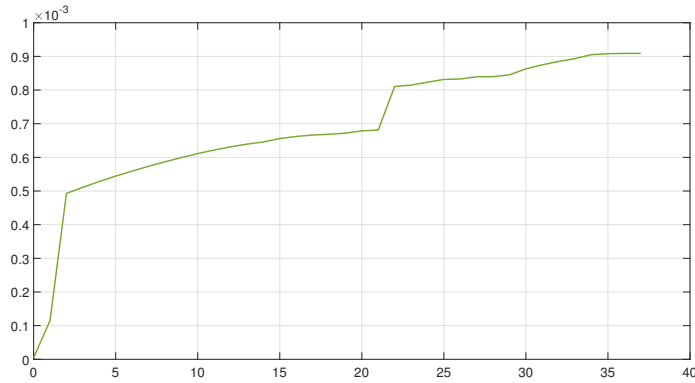

Figure B.5: Evolution of the objective function from the initialization $\mathcal{J}_{30}(u_0^0) = 4.5 \times 10^{-6}$ to the last iteration $\mathcal{J}_{30}(u_0^{37}) = 0.9 \times 10^{-3}$.

regions of full density, which exemplify the phenomenon described for the one-dimensional case in [32].

B.3 Example 3

For this example we keep the settings of the previous one, but we consider a higher initial mass $m = 27$. The geometry of the initial distribution is a band of full density dividing the domain into two regions of zero density, like in the example 2, see fig. B.6 (a).

The corresponding local optimum found by the numerical algorithm is showed in fig. B.6 (b). As in the previous case, the optimizer reflects a low density zone ringed by a high density region. This gap is clearly filled by diffusion as time evolves. The values of the objective function in fig. B.7 reflect that, in this case, the fixed mass is sufficient to trigger an increase in total density over time.

This example point out the non suitability of band type initial distributions. Note from fig. B.7 that, despite the considerable increase of the initial mass with respect to the example 2, the values of the

(a) u_0^0 (b) $\overline{u}_0 = u_0^{50}$
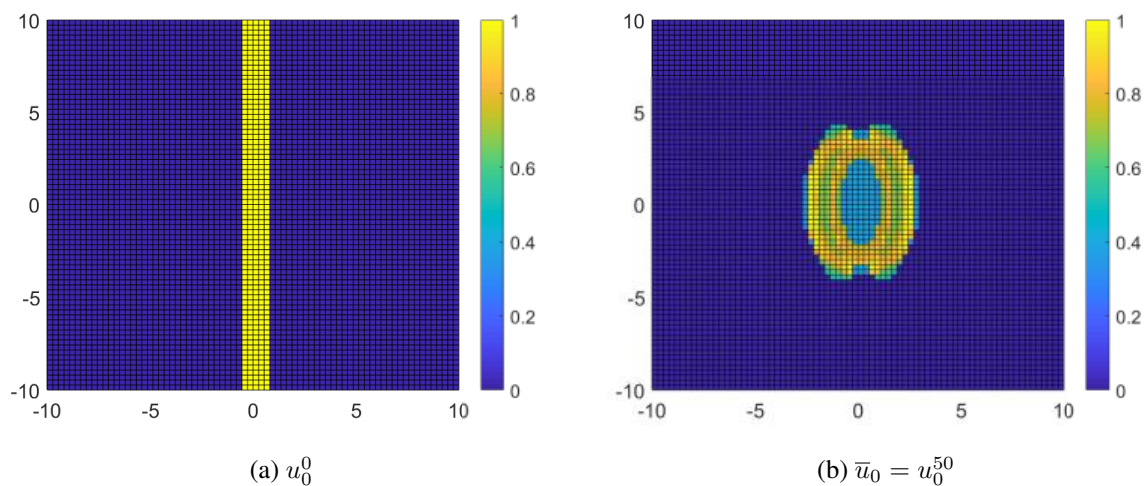
Figure B.6: In the left-hand side is showed the input of the algorithm, given by the stripe of width $r = 1.6$ centered at the origin. In the right-hand side, the local optimum found by the numerical algorithm after 50 iterations.

Figure B.7: Evolution of the objective function from the initialization $\mathcal{J}_{30}(u_0^0) = 3 \times 10^{-5}$ to the last iteration $\mathcal{J}_{30}(u_0^{50}) = 324$.

objective function $\mathcal{J}_{30}(u_0^0)$ associated to the band is of the order of 10^{-5}, which is very low compared with the value associated to the final distribution.

Appendix C

An alternative for practical applications
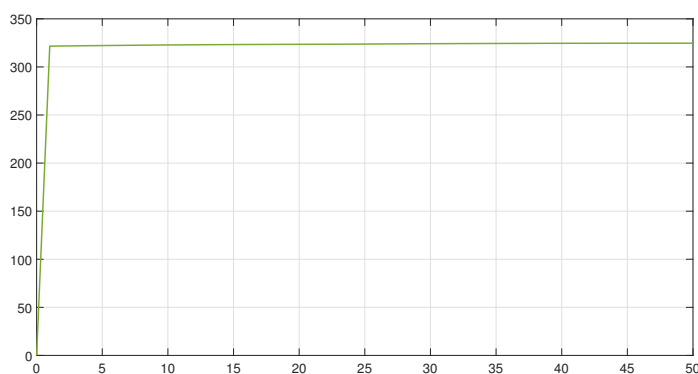
As mentioned in the introduction and in the chapters 2 and 3, the solutions to the optimization problem we study in the first part of this thesis might be very irregular, which makes its application to the biological control of the dengue virus inadequate and unrealistic. In this section we present an alternative that transform the solutions given by the numerical algorithm described in section 2.5 into something suitable for implementation in the field.

The idea is to consider a new initial distribution $\hat{u}_0$ which have the same mass and support as the one given by the algorithm $\overline{u}_0$, but which is constant in every convex component of its support. So we define:

$$\hat{u}_0(x) = \sum_k \frac{\int_{\Omega^k} \overline{u}_0(x)\,dx}{|\Omega^k|} \mathbb{1}_{\Omega^k}(x), \tag{C.1}$$

where $\Omega^1, \Omega^2, ...,$ are the convex components of Ω.

In what follow we show some numerical simulations in the one-dimensional case, applying this alternative. As in the previous examples, we consider a bistable reaction term $f(u) = u(1-u)(u-\rho)$ and diffusion coefficient $\sigma = 1$.

C.1   Example 1

In this case the domain Ω is the interval $[-45, 45]$ discretized uniformly by a grid of size $dx = 0.09$. The fixed mass is $m = 10$, and we set our problem for $T = 25$ and $\rho = 0.35$. The results are shown in fig. C.1. The variation of the objective function, for this problem, represents a decrease of less than 1.5% and thus the initial data $\hat{u}_0$ is a good alternative for practical applications.

C.2   Example 2

For the second example we consider $m = 15$, $\Omega = [-50, 50]$ with a discretization step-size of $dx = 0.1$, $T = 25$ and $\rho = 0.25$. The results are shown in fig. C.2.

In this case the variation of the objective function is of 1.7 % from the local optimum given by our numerical algorithm to the rearranged initial distribution. The modification of $\overline{u}_0$ into $\hat{u}_0$ is, in this example, also a good compromise for practical applications.

Figure C.1: The solid lines represent the initial data obtained by our numerical algorithm (blue), and the alternative initial data defined in eq. (C.1) (red). The dashed lines represent their corresponding solutions of the reaction diffusion model at time $T = 25$. The objective function varies from $\mathcal{J}_T(\overline{u}_0) = 32.1284$ to $\mathcal{J}_T(\hat{u}_0) = 31.6509$.

Figure C.2: The solid lines represent the initial data obtained by our numerical algorithm (blue), and the alternative initial data defined in eq. (C.1) (red). The dashed lines represent their corresponding solutions of the reaction diffusion model at time $T = 25$. $\mathcal{J}_T(\overline{u}_0) = 83.9227$ and $\mathcal{J}_T(\hat{u}_0) = 82.482$

C.3 Example 3

Despite the good results shown by most of the simulations we have performed, there are cases where this alternative does not provide a good competitor in terms of the objective function evaluation.

Indeed, considering the same domain Ω and space discretization from the previous example, $\rho = 0.3$, the initial mas $m = 13$ and the final time $T = 50$ the variation of the objective function from the optimal initial data $\overline{u}_0$ to the rearranged initial data $\hat{u}_0$ is of 21%.
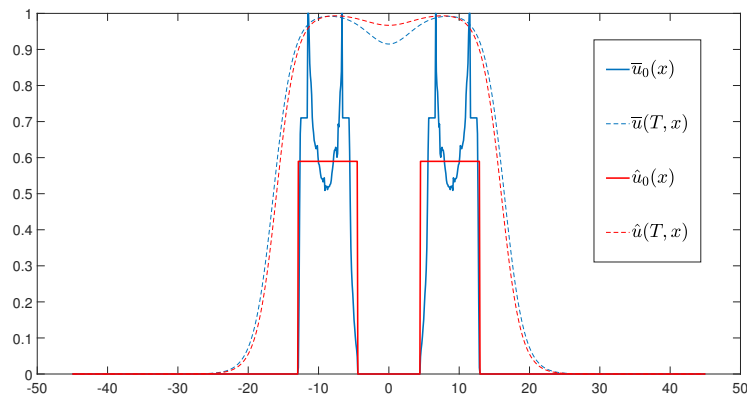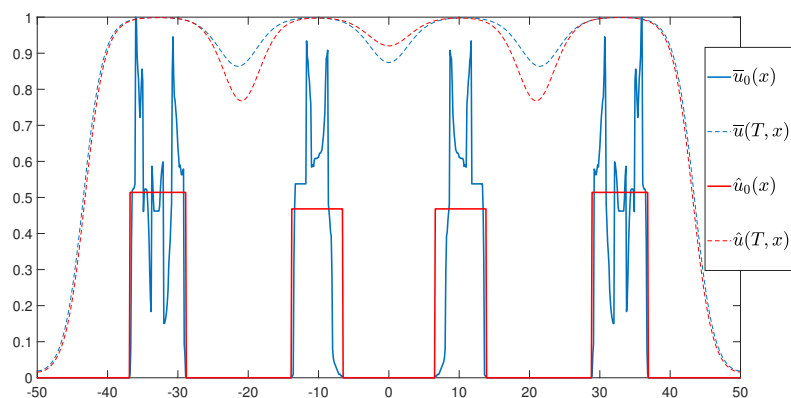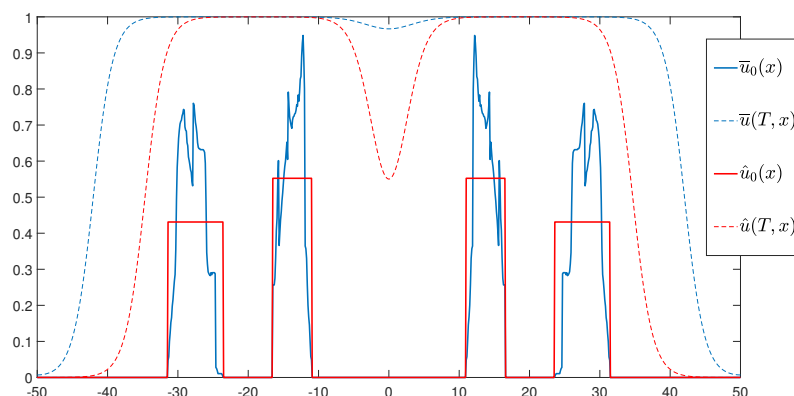

Figure C.3: The solid lines represent the initial data obtained by our numerical algorithm (blue), and the alternative initial data defined in eq. (C.1) (red). The dashed lines represent the corresponding solutions of the reaction diffusion model at time $T = 25$. $\mathcal{J}_T(\overline{u}_0) = 83.7864$ and $\mathcal{J}_T(\hat{u}_0) = 66.1989$

We believe that this rearrangement could be a good strategy to transform the theoretical results into a more realistic release protocol for the problem of biological control of dengue virus by using *Wolbachia*. However, a study of the particularities of each scenario would be essential to avoid situations in which such modifications of the optimal initial distribution does not provide a good option, as seen in the example 3.

Author's bibliography

[68]   Grégoire Nadin and Ana Isis Toledo Marrero. "On the maximization problem for solutions of reaction–diffusion equations with respect to their initial data". In: *Mathematical Modelling of Natural Phenomena* 15 (2020), p. 71. DOI: 10.1051/mmnp/2020030.

[70]   Grégoire Nadin, Eric Ogier-Denis, Ana Isis Toledo, and Hatem Zaag. *A Turing mechanism in order to explain the patchy nature of Crohn's disease*. submitted for publication. 2020. arXiv: 2007.13587 [math.AP].

Bibliography

[1] Matthieu Alfaro, Arnaud Ducrot, and Grégory Faye. "Quantitative Estimates of the Threshold Phenomena for Propagation in Reaction-Diffusion Equations". In: *SIAM Journal on Applied Dynamical Systems* 19.2 (Jan. 2020), pp. 1291–1311. DOI: 10.1137/19m1292187.

[2] Luís Almeida, Michel Duprez, Yannick Privat, and Nicolas Vauchelet. "Mosquito population control strategies for fighting against arboviruses". In: *Mathematical Biosciences and Engineering* 16.6 (2019), pp. 6274–6297. DOI: 10.3934/mbe.2019313.

[3] Luís Almeida, Yannick Privat, Martin Strugarek, and Nicolas Vauchelet. "Optimal releases for population replacement strategies, application to Wolbachia". In: *SIAM Journal on Mathematical Analysis* 51.4 (2019), pp. 3170–3194. DOI: 10.1137/18M1189841.

[4] Luke Alphey et al. "Genetic control of Aedes mosquitoes". In: *Pathogens and global health* 107 (June 2013), pp. 170–9. DOI: 10.1179/2047773213Y.0000000095.

[5] N. Barton and M. Turelli. "Spatial waves of advances with bistable dynamics: cytoplasmic and genetic analogues of Allee effects". In: *The Americal Naturalist* 78(3) (2011), E48–E75.

[6] Daniel C Baumgart and Simon R Carding. "Inflammatory bowel disease: cause and immunobiology". In: *The Lancet* 369.9573 (2007), pp. 1627 –1640. DOI: https://doi.org/10.1016/S0140-6736(07)60750-8.

[7] Pierre-Alexandre Bliman. "Feedback Control Principles for Biological Control of Dengue Vectors". In: *European Control Conference ECC19*. The last version of this manuscript has been accepted for publication in the Proceedings of European Control Conference ECC19. A complete version of the report appeared in arXiv, and is available at: http://arxiv.org/abs/1903.00730. Naples, Italy, June 2019, p. 6.

[8] Pierre-Alexandre Bliman, M. Soledad Aronna, Flávio C. Coelho, and Moacyr A. H. B. da Silva. "Ensuring successful introduction of Wolbachia in natural populations of Aedes aegypti by means of feedback control". In: *Journal of Mathematical Biology* (Aug. 2017). DOI: 10.1007/s00285-017-1174-x.

[9] Pierre-Alexandre Bliman and Nicolas Vauchelet. "Establishing Traveling Wave in Bistable Reaction-Diffusion System by Feedback". In: *IEEE Control Systems Letters* 1.1 (2017), pp. 62–67. DOI: 10.1109/LCSYS.2017.2703303.

[10] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[11] J. Tielbeek C. Puylaert and J. Stoker. "Crohn's disease - role of MRI". In: *Radiology assistant* (2016-02-17).

[12] Vincent Calvez et al. "Mathematical and numerical modeling of early atherosclerotic lesions". In: *ESAIM: Proceedings* 30 (Aug. 2010), pp. 1–14. DOI: 10.1051/proc/2010002.

[13] Carmen Camacho, Benteng Zou, and Maya Briani. "On the dynamics of capital accumulation across space". In: *European J. Oper. Res.* 186.2 (2008), pp. 451–465. DOI: 10.1016/j.ejor.2007.02.031.

[14] Doris E. Campo-Duarte, Daiver Cardona-Salgado, and Olga Vasilieva. "Establishing wMelPop Wolbachia Infection among Wild Aedes aegypti Females by Optimal Control Approach". In: *Applied Mathematics & Information Sciences* 11.4 (July 2017), pp. 1011–1027. DOI: 10.18576/amis/110408.

[15] Doris E. Campo-Duarte, Olga Vasilieva, Daiver Cardona-Salgado, and Mikhail Svinin. "Optimal control approach for establishing wMelPop Wolbachia infection among wild Aedes aegypti populations". In: *Journal of Mathematical Biology* 76.7 (Feb. 2018), pp. 1907–1950. DOI: 10.1007/s00285-018-1213-2.

[16] Fabien Caubet, Thibaut Deheuvels, and Yannick Privat. "Optimal location of resources for biased movement of species: the 1D case". In: *SIAM Journal on Applied Mathematics* 77.6 (2017), pp. 1876–1903.

[17] Alexander Chalmers, Anna Cohen, Christina Bursill, and Mary Myerscough. "Bifurcation and dynamics in a mathematical model of early atherosclerosis : How acute inflammation drives lesion development". In: *Journal of mathematical biology* 71 (Mar. 2015). DOI: 10.1007/s00285-015-0864-5.

[18] Carson Chow et al. "The acute inflammatory response in diverse shock states". In: *Shock (Augusta, Ga.)* 24 (Aug. 2005), pp. 74–84. DOI: 10.1097/01.shk.0000168526.97716.f3.

[19] N. Cónsul, S. M. Oliva, and M. Pellicer. "A PDE approach of inflammatory phase dynamics in diabetic wounds". In: *Publ. Mat.* 58.2 (2014), pp. 265–293.

[20] *Conception optimale de structures*. Springer Berlin Heidelberg, 2006. DOI: 10.1007/978-3-540-36856-4.

[21] Judy Day et al. "A reduced mathematical model of the acute inflammatory response II. Capturing scenarios of repeated endotoxin administration". In: *Journal of Theoretical Biology* 242.1 (2006), pp. 237 –256. DOI: https://doi.org/10.1016/j.jtbi.2006.02.015.

[22] H. Diaz, A.A. Ramirez, A. Olarte, and C. Clavijo. "A model for the control of malaria using genetically modified vectors". In: *Journal of Theoretical Biology* 276.1 (2011), pp. 57 –66. DOI: https://doi.org/10.1016/j.jtbi.2011.01.053.

[23] Y. Du and H. Matano. "Convergence and sharp thresholds for propagation in nonlinear diffusion problems". In: *Journal of the European Mathematical Society* (2010).

[24] Joe Dunster, Helen Byrne, and J King. "The Resolution of Inflammation: A Mathematical Model of Neutrophil and Macrophage Interactions". In: *Bulletin of mathematical biology* 76 (July 2014). DOI: 10.1007/s11538-014-9987-x.

[25] Nader EL Khatib and Stéphane Génieys. "Atherosclerosis Initiation Modeled as an Inflammatory Process". In: *Mathematical Modelling of Natural Phenomena* (Jan. 2007). DOI: 10.1051/mmnp:2008022.

[26] Nader EL Khatib, Stéphane Génieys, and Bogdan Kazmierczak. "Reaction-Diffusion Model of Atherosclerosis Development". In: *Journal of mathematical biology* 65 (Aug. 2011), pp. 349–74. DOI: 10.1007/s00285-011-0461-1.

[27] J. Evans. "Nerve axon equations: IV The stable and unstable impulse". In: 1975.

[28] L. C. Evans. *Partial Differential Equations*. Ed. by American Mathematical Society. 2nd. Vol. 19. American Mathematical Society, 2010.

[29] József Z. Farkas and Peter Hinow. "Structured and Unstructured Continuous Models for Wolbachia Infections". In: *Bulletin of Mathematical Biology* 72.8 (Mar. 2010), pp. 2067–2088. DOI: 10.1007/s11538-010-9528-1.

[30] Andrew Fenton, Karyn N. Johnson, Jeremy C. Brownlie, and Gregory D. D. Hurst. "Solving the Wolbachia Paradox: Modeling the Tripartite Interaction between Host, Wolbachia, and a Natural Enemy". In: *The American Naturalist* 178.3 (Sept. 2011), pp. 333–342. DOI: 10.1086/661247.

[31] R. A. FISHER. "THE WAVE OF ADVANCE OF ADVANTAGEOUS GENES". In: *Annals of Eugenics* 7.4 (June 1937), pp. 355–369. DOI: 10.1111/j.1469-1809.1937.tb02153.x.

[32] J. Garnier, L. Roques, and F. Hamel. "Success rate of a biological invasion in terms of the spatial distribution of the founding population". In: *Bulletin of Mathematical Biology* (2012).

[33] A. Hancock, P. Sinkins, and H. Charles. "Strategies for introducing Wolbachia to reduce transmission of mosquito-borne diseases". In: *PLoS Negl Trop Dis* 5(4) (2011), pp. 1–10.

[34] A. Henrot and M. Pierre. *Variation et optimisation de formes*. Vol. 48. Une analyse géométrique. [A geometric Analysis]. 2005.

[35] Meagan Herald. "General Model of Inflammation". In: *Bulletin of mathematical biology* 72 (Oct. 2009), pp. 765–79. DOI: 10.1007/s11538-009-9468-9.

[36] A. Hidalgo, L. Tello, and E. F. Toro. "Numerical and analytical study of an atherosclerosis inflammatory disease model". In: *Journal of Mathematical Biology* 68.7 (May 2013), pp. 1785–1814. DOI: 10.1007/s00285-013-0688-0.

[37] A. Hoffmann et al. "Successful establishment of Wolbachia in Aedes populations to suppress dengue transmission". In: *Nature* 7361(476) (2011), pp. 454–457.

[38] C. S. Holling. "The Functional Response of Predators to Prey Density and its Role in Mimicry and Population Regulation". In: *Memoirs of the Entomological Society of Canada* 97.S45 (1965), 5–60. DOI: 10.4039/entm9745fv.

[39] H. Hughes and N. Britton. "Modeling the use of Wolbachia to control dengue fever transmission". In: *Bull. Math. Biol.* 75 (2013), pp. 796–818.

[40] Juraj Húska. "Harnack inequality and exponential separation for oblique derivative problems on Lipschitz domains". In: *J. Differential Equations* 226.2 (2006), pp. 541–557. DOI: 10.1016/j.jde.2006.02.008.

[41] Akif Ibragimov, Catherine Mcneal, L. Ritter, and Jay Walton. "A mathematical model of atherogenesis as an inflammatory response". In: *Mathematical medicine and biology : a journal of the IMA* 22 (Jan. 2006), pp. 305–33. DOI: 10.1093/imammb/dqi011.

[42] Jumpei Inoue, and Kousuke Kuto. "On the unboundedness of the ratio of species and resources for the diffusive logistic equation". In: *Discrete & Continuous Dynamical Systems - B* 22.11 (2017), pp. 0–0. DOI: 10.3934/dcdsb.2020186.

[43] A. S. Cheifetz J. D. Feuerstein. "Crohn Disease: Epidemiology, Diagnosis, and Management". In: *Mayo Clinic proceedings* 92.7 (2017), pp. 1088–1103.

[44] C-Y. Kao, Y. Lou, and E. Yanagida. "Principal eigenvalue for an elliptic problem with indefinite weight on cylindrical domains". In: *Math. Biosci. Eng.* 5.2 (2008), pp. 315–335. DOI: 10.3934/mbe.2008.5.315.

[45] E. F. Keller and L. A. Segel. "Initiation of slime mold aggregation viewed as an instability." In: *J. Theor. Biol.* 26 (1970), pp. 399–415.

[46] A. Kolmogoroff, I. Petrovsky, and N. Piscounoff. "Study of the Diffusion Equation with Growth of the Quantity of Matter and its Application to a Biology Problem". In: *Dynamics of Curved Fronts*. Ed. by Pierre Pelcé. San Diego: Academic Press, 1988, pp. 105 –130. DOI: `https://doi.org/10.1016/B978-0-08-092523-3.50014-9`.

[47] T. Korem et al. "Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples". In: *Science* 349.6252 (July 2015), pp. 1101–1106. DOI: `10.1126/science.aac4812`.

[48] Rukmini Kumar, Gilles Clermont, Yoram Vodovotz, and Carson C. Chow. "The dynamics of acute inflammation". In: *Journal of Theoretical Biology* 230.2 (2004), pp. 145 –155. DOI: `https://doi.org/10.1016/j.jtbi.2004.04.044`.

[49] Marie-Thérése Labro. "Interference of Antibacterial Agents with Phagocyte Functions: Immunomodulation or "Immuno-Fairy Tales"?" In: *Clinical Microbiology Reviews* 13.4 (Oct. 2000), pp. 615–650. DOI: `10.1128/cmr.13.4.615`.

[50] O.A. Ladyzenskaja, V.A. Solonnikov, and N.N. Uralceva. *Linear and quasilinear equations of parabolic type*. 1968.

[51] K-Y Lam, S. Liu, and Y. Lou. *Selected topics on reaction-diffusion-advection models from spatial ecology*. preprint.

[52] DA Lauffenburger and CR Kennedy. "Localized bacterial infection in a distributed model for tissue inflammation". In: *Journal of mathematical biology* 16.2 (1983), 141—163. DOI: `10.1007/bf00276054`.

[53] Douglas A. Lauffenburger and Clinton R. Kennedy. "Analysis of a lumped model for tissue inflammation dynamics". In: *Mathematical Biosciences* 53.3 (1981), pp. 189 –221. DOI: `https://doi.org/10.1016/0025-5564(81)90018-3`.

[54] Peter C. J. Leijh, Maria T. van den Barselaar, Ivonne Dubbeldeman-Rempt, and Ralph van Furth. "Kinetics of intracellular killing of Staphylococcus aureus and Escherichia coli by human granulocytes". In: *European Journal of Immunology* 10.10 (Oct. 1980), pp. 750–757. DOI: `10.1002/eji.1830101005`.

[55] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod; Gauthier-Villars, Paris, 1969, pp. xx+554.

[56] M. Locatelli. "Simulated Annealing Algorithms for Continuous Global Optimization: Convergence Conditions". In: *Journal of Optimization Theory and Applications* 104 (2000), pp. 121–133.

[57] Serena Longo et al. "New Insights into Inflammatory Bowel Diseases from Proteomic and Lipidomic Studies". In: *Proteomes* 8.3 (2020). DOI: `10.3390/proteomes8030018`.

[58] Y. Lou. "Some Challenging Mathematical Problems in Evolution of Dispersal and Population Dynamics". In: *Tutorials in Mathematical Biosciences IV: Evolution and Ecology*. Ed. by Avner Friedman. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 171–205. DOI: `10.1007/978-3-540-74331-6_5`.

[59] Y. Lou, K. Nagahara, and E. Yanagida. "Maximizing the total population with logistic growth in a patchy environment". In: *Submitted* (2020).

[60] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. "Reaction-diffusion equations for interacting particle systems". In: *Journal of Statistical Physics* 44.3-4 (Aug. 1986), pp. 589–644. DOI: 10.1007/bf01011311.

[61] H. Mayer, K. S. Zaenker, and U. an der Heiden. "A basic mathematical model of the immune response". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 5.1 (1995), pp. 155–161. DOI: 10.1063/1.166098. eprint: https://doi.org/10.1063/1.166098.

[62] Idriss Mazari, Grégoire Nadin, and Yannick Privat. "Optimal location of resources maximizing the total population size in logistic models". In: *arXiv: Analysis of PDEs* (2019).

[63] Idriss Mazari, Grégoire Nadin, and Yannick Privat. "Shape optimization of a weighted two-phase Dirichlet eigenvalue". Preprint. 2020.

[64] Idriss Mazari, Grégoire Nadin, and Yannick Privat. "Some challenging optimisation problems for logistic diffusive equations and numerical issues". In: *to appear in "Handbook of Numerical Analysis"* (2020).

[65] Idriss Mazari and Domènec Ruiz-Balet. "A fragmentation phenomenon for a non-energetic optimal control problem: optimisation of the total population size in logistic diffusive models." In: *arXiv: Optimization and Control* (2020).

[66] Luciano A. Moreira et al. "A Wolbachia Symbiont in Aedes aegypti Limits Infection with Dengue, Chikungunya, and Plasmodium". In: *Cell* 139.7 (Dec. 2009), pp. 1268–1278. DOI: 10.1016/j.cell.2009.11.042.

[67] James D. Murray. *Mathematical Biology*. Springer Berlin Heidelberg, 1993. DOI: 10.1007/978-3-662-08542-4.

[69] Grégoire Nadin, Martin Strugarek, and Nicolas Vauchelet. "Hindrances to bistable front propagation: application to Wolbachia invasion". working paper or preprint. Jan. 2017.

[71] K. Nagahara and E. Yanagida. "Maximization of the total population in a reaction–diffusion model with logistic growth". In: *Calculus of Variations and Partial Differential Equations* 57.3 (2018), p. 80. DOI: 10.1007/s00526-018-1353-7.

[72] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. second. New York, NY, USA: Springer, 2006.

[73] Clifford S. Patlak. "Random walk with persistence and external bias". In: *The Bulletin of Mathematical Biophysics* 15.3 (Sept. 1953), pp. 311–338. DOI: 10.1007/bf02476407.

[74] Kevin. Penner, Bard. Ermentrout, and David. Swigon. "Pattern Formation in a Model of Acute Inflammation". In: *SIAM Journal on Applied Dynamical Systems* 11.2 (2012), pp. 629–660. DOI: 10.1137/110834081. eprint: https://doi.org/10.1137/110834081.

[75] Benoît Perthame. *Parabolic Equations in Biology*. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-19500-1.

[76] Katharina Prochazka and Gero Vogl. "Quantifying the driving factors for language shift in a bilingual region". In: *Proceedings of the National Academy of Sciences* 114.17 (Mar. 2017), pp. 4365–4369. DOI: 10.1073/pnas.1617252114.

[77] Angela Reynolds et al. "A reduced mathematical model of the acute inflammatory response: I. Derivation of model and analysis of anti-inflammation". In: *Journal of Theoretical Biology* 242.1 (2006), pp. 220 –236. DOI: https://doi.org/10.1016/j.jtbi.2006.02.016.

[78] Anirban Roy et al. "A mathematical model of acute inflammatory response to endotoxin challenge". In: (Jan. 2007).

[79]   Peter A. Ryan et al. "Establishment of wMel Wolbachia in Aedes aegypti mosquitoes and reduction of local dengue transmission in Cairns and surrounding locations in northern Queensland, Australia". In: *Gates Open Research* 3 (Sept. 2019), p. 1547. DOI: 10.12688/gatesopenres.13061.1.

[80]   Douglas F. Stickle, Douglas A. Lauffenburger, and Ronald P. Daniele. "The Motile Response of Lung Macrophages: Theoretical and Experimental Approaches Using the Linear Under-Agarose Assay". In: *Journal of Leukocyte Biology* 38.3 (Sept. 1985), pp. 383–401. DOI: 10.1002/jlb.38.3.383.

[81]   Thomas P. Stossel. "Quantitative studies of phagocytosis". In: *The Journal of Cell Biology* 58.2 (Aug. 1973), pp. 346–356. DOI: 10.1083/jcb.58.2.346.

[82]   Martin Strugarek, Herve Bossin, and Yves Dumont. "On the use of the sterile insect technique or the incompatible insect technique to reduce or eliminate mosquito populations". working paper or preprint. May 2018.

[83]   Martin Strugarek, Nicolas Vauchelet, and Jorge Zubelli. "Quantifying the Survival Uncertainty of Wolbachia-infected Mosquitoes in a Spatial Model *". working paper or preprint. Aug. 2016.

[84]   Joshua Sullivan and Ivan Yotov. "Mathematical and Numerical Modeling of Inflammation". In: 2006.

[85]   E. Trelat, J. Zhu, and E. Zuazua. "Allee optimal control of a system in ecology". In: *Math. Models Methods Appl. Sci.* 9 (2018), pp. 1665–1697.

[86]   A. Turing. "The chemical basis of morphogenesis". In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237.641 (Aug. 1952), pp. 37–72. DOI: 10.1098/rstb.1952.0012.

[87]   A. M. Turing. "The chemical basis of morphogenesis". In: *Bulletin of Mathematical Biology* 52.1-2 (Jan. 1990), pp. 153–197. DOI: 10.1007/bf02459572.

[88]   P. F. Verhulst. *Recherches mathématiques sur la loi d'accroissement de la population*. 1845.

[89]   Yoram Vodovotz. "Deciphering the Complexity of Acute Inflammation Using Mathematical Models". In: *Immunologic research* 36 (Feb. 2006), pp. 237–45. DOI: 10.1385/IR:36:1:237.

[90]   Yoram Vodovotz, Gilles Clermont, Carson Chow, and Gary An. "Mathematical models of the acute inflammatory response". In: *Current opinion in critical care* 10 (Nov. 2004), pp. 383–90. DOI: 10.1097/01.ccx.0000139360.30327.69.

[91]   T. Walker et al. "The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations". In: *Nature* 7361(476) (2011), pp. 450–453.

[92]   Helen V. Waugh and Jonathan A. Sherratt. "Modeling the effects of treating diabetic wounds with engineered skin substitutes". In: *Wound Repair and Regeneration* 15.4 (July 2007), pp. 556–565. DOI: 10.1111/j.1524-475x.2007.00270.x.

[93]   Katherine Wendelsdorf, Josep Bassaganya-Riera, Raquel Hontecillas, and Stephen Eubank. "Model of colonic inflammation: Immune modulatory mechanisms in inflammatory bowel disease". In: *Journal of Theoretical Biology* 264.4 (2010), pp. 1225 –1239. DOI: https://doi.org/10.1016/j.jtbi.2010.03.027.

[94]   A. Zlatos. "Sharp transition between extinction and propagation of reaction". In: *Journal of the American Mathematical Society* 19.1 (2005), pp. 251–263.

Abstract

This thesis is devoted to the study of two problems arising from biology and medicine. The first model is motivated by a new technique to eradicate mosquito-borne diseases such as the dengue virus. A certain number of mosquitoes, inoculated with a bacterium inhibiting mosquito-borne disease transmission to humans are released in the environment. The evolution of this subset of the mosquito population can be described by means of a reaction-diffusion equation. The problem we address here concerns the maximization of the total number of carrying individuals after a certain prescribed time, which is a quantity depending on the solution of the equation. We maximize this quantity with respect to the initial datum under certain size constraints. Existence and regularity results as well as a partial characterization of optimizers are stated by means of the study of the first and second order optimality conditions. A numerical algorithm, inspired by the classical ascent of gradient and taking advantage of the theoretical results we obtain here is described, allowing a numerical approximation of local optimizers.

On the other hand, a model describing the dynamics of immune cells and pathogenic bacteria in the gut tissues is introduced. More precisely, a reaction-diffusion system is considered with the purpose of explaining the patchy inflammatory patterns observed in patients suffering from Crohn's disease. We perform a stability analysis enabling us to identify conditions driving to the occurrence of Turing instabilities. Such instabilities could be interpreted as the patchy inflammatory patterns. Realistic parameter values for which this phenomenon arises are either computed or retrieved from the existent literature and numerical simulations are performed as well.

Keywords: Reaction-diffusion equation, control, conservation biology, optimization, Turing pattern, activator-inhibitor, inflammatory diseases.

Résumé

Cette thèse est consacrée à l'étude de deux problèmes issus de la biologie et de la médecine. Le premier est motivé par une technique de contrôle biologique pour l'éradication de l'épidémie de la dengue transmise par des moustiques. Une bactérie, dont les effets chez les moustiques inhibent la transmission de ce virus, est inoculée à un certain nombre des moustiques qui sont ensuite relâchés dans l'environnement. L'évolution de cette partie de la population porteuse de la bactérie peut être décrite par une équation de réaction-diffusion. On s'intéresse particulièrement à maximiser la population totale de moustiques porteurs de cette bactérie après un certain temps. Il s'agit d'une quantité dépendant de la solution de l'équation, que l'on maximise par rapport à la donnée initiale sous certaines contraintes. L'existence et la régularité des solutions à ce problème d'optimisation, ainsi qu'une caractérisation partielle de la donnée initiale optimale sont établies grâce à l'étude des conditions d'optimalité de premier et deuxième ordre. Un algorithme numérique, inspiré de la méthode classique de montée de gradient et tirant parti des conditions d'optimalité est décrit, permettant une approximation numérique des maxima locaux de ce problème.

D'autre part, un modèle décrivant la dynamique des cellules immunitaires et des bactéries pathogènes dans les tissus de l'intestin est introduit. Un système de réaction-diffusion est considéré, l'objectif étant d'expliquer les motif inflammatoires inégaux observés chez les patients souffrant de la maladie de Crohn. Une analyse de stabilité est réalisée et des conditions menant à l'apparition d'instabilités de Turing sont énoncées; ces instabilités pouvant être interprétées comme les patterns inflammatoires. Des valeurs réalistes des paramètres, pour lesquels ce phénomène se produit, sont calculées ou extraites de la littérature existante, des simulations numériques sont également réalisées.

Mots-clés: Équations de réaction-diffusion, contrôle, biologie de la conservation, optimisation, Turing patterns, activateur-inhibiteur, maladies inflammatoires.