



HAL
open science

Optimisation of the total population size with respect to the initial condition for semilinear parabolic equations: Two-scale expansions and symmetrisations

Idriss Mazari, Grégoire Nadin, Ana Isis Toledo Marrero

► To cite this version:

Idriss Mazari, Grégoire Nadin, Ana Isis Toledo Marrero. Optimisation of the total population size with respect to the initial condition for semilinear parabolic equations: Two-scale expansions and symmetrisations. 2021. hal-03846882v1

HAL Id: hal-03846882

<https://hal.science/hal-03846882v1>

Preprint submitted on 2 Apr 2021 (v1), last revised 12 Jul 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimisation of the total population size with respect to the initial condition for semilinear parabolic equations: Two-scale expansions and symmetrisations

Idriss Mazari*, Grégoire Nadin†, Ana Isis Toledo Marrero‡

April 2, 2021

Abstract

In this article, we propose in-depth analysis and characterisation of the optimisers of the following optimisation problem: how to choose the initial condition u_0 in order to maximise the spatial integral at a given time of the solution of the semilinear equation $u_t - \Delta u = f(u)$, under L^∞ and L^1 constraints on u_0 ? Our contribution in the present paper is to give a characterisation of the behaviour of the optimiser \bar{u}_0 when it does not saturate the L^∞ constraints, which is a key step in implementing efficient numerical algorithms. We give such a characterisation under mild regularity assumptions by proving that in that case \bar{u}_0 can only take values in the "zone of concavity" of f . This is done using two-scale asymptotic expansions. We then show how well-known isoperimetric inequalities yield a full characterisation of maximisers when f is convex. Finally, we provide several numerical simulations in one and two dimensions that illustrate and exemplify the fact that such characterisations significantly improves the computational time. All our theoretical results are in the one-dimensional case and we offer several comments about possible generalisations to other contexts, or obstructions that may prohibit doing so.

Keywords: Reaction equation, optimal control, shape optimisation, two-scale expansions.

AMS classification: 35Q92,49J99,34B15.

Acknowledgement The authors wish to warmly thank J. Pertinand for scientific conversations. I. Mazari was partially supported by the French ANR Project ANR-18-CE40-0013 - SHAPO on Shape optimisation and by the Austrian Science Fund (FWF) through the grant I4052-N32 . I. Mazari and G. Nadin were partially supported by the Project "Analysis and simulation of optimal shapes - application to lifesciences" of the Paris City Hall.

1 Introduction

1.1 Scope of the article

In this article, we propose to establish several results concerning an optimal control problem for a class of semilinear parabolic equations. Some aspects of this problem have been initially addressed

*Technische Universität Wien, Institute of Analysis and Scientific Computing, 8-10 Wiedner Hauptstrasse, 1040 Wien (idriss.mazari@tuwien.ac.at)

†CNRS, Sorbonne Universités, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France (gregoire.nadin@sorbonne-universite.fr)

‡Sorbonne Universités, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France (ana-isis.toledo.marrero@sorbonne-universite.fr)

by two of the authors in [44]. In the setting under consideration, the control variable to optimise is the initial condition. As we will see throughout the statement of the results, the form (e.g. convex or concave) of the semilinearity plays a crucial role in the analysis and calls for a detailed study of second order optimality conditions, which is our main result, Theorem 1. In the case of convex semilinearities, using rearrangement arguments, we can give a full characterisation of maximisers, see Theorem 2. Using Theorem 1, we can improve an algorithm initially developed in [44], and we display numerical results in Section 4.

Initial motivation of the paper 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. Namely, for a semilinear equation, what is the best possible initial condition ("best" being understood as maximising the integral of the solution at a certain time horizon T)? The complicated behaviour of bistable non-linearities, which are neither convex nor concave, makes the analysis of this query very intricate. The two aforementioned results, Theorems 1 and 2 enable us to show how complicated the behaviour of maximisers can be for such non-linearities. Bistable equations are of central importance in mathematical biology [43] and, very broadly speaking, model the evolution of a subgroup of a population. Among their many applications, one may mention chemical reactions [47], neurosciences [19], phase transition [35], linguistic dynamics [48], or the evolution of diseases [43]. 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 mosquito borne diseases such that dengue, and this was the main motivation in [3]. The strategy is to release a certain amount of Wolbachia carrying mosquitoes (Wolbachia is a bacterium that inhibit the transmission of mosquito borne diseases and that individuals inherit from their mother) in a population of wild mosquitoes that can potentially transmit the diseases, in order to maximise the proportion of Wolbachia carrying mosquitoes at the final time. In mathematical terms: given a time horizon T ,

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

Even without having stated it formally, we can make two observations on this problem: the first one is that, the variable of the equation being the proportion of a subgroup, we need to enforce 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.

Optimisation problems in mathematical biology Let us briefly sketch how this problem fits in the literature devoted to such optimisation and control problems for mathematical biology. Optimisation problems for reaction-diffusion equations have by now gathered a lot of attention from the mathematical community. Most of these optimisation problems are set in a stationary setting, that is, assuming that the population has already reached an equilibrium, and the main problems that have been considered often deal with the optimisation of the spatial heterogeneity [18, 22, 23, 29, 30, 38, 39, 41, 45] (we also refer to the recent surveys [26, 40]); most of these works deal with monostable nonlinearities. We also point to the recent [14] for the study of an optimal control problem for parabolic monostable equations. On the other hand, optimisation problems for bistable equations, which are the other paradigmatic class of equations in mathematical biology [43], have received a less complete mathematical treatment, but are now the topic of an intense research activity from the control point of view, see [3, 44] and the references therein. Related optimal control problems are not yet fully understood. More generally, less attention has been devoted to optimisation problem with respect to the initial condition for such semilinear evolution equations.

1.2 Mathematical setup and statement of the results

1.2.1 Statement of the problem

We work in $\Omega = (0; \pi)$. We consider a \mathcal{C}^1 function $f : [0; 1] \rightarrow \mathbb{R}$, and the associated parabolic equation

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

where u_0 is an initial condition satisfying the constraint

$$0 \leq u_0 \leq 1.$$

Since our initial motivation, as explained in the first paragraph of this introduction, is to maximise the proportion of a subgroup of a population, such an L^∞ constraint is natural. At the mathematical level, it should be noted that we could carry out the same analysis with any L^∞ constraint of the form $0 \leq u_0 \leq \kappa$ by a simple change of variable.

We define, for any $T > 0$, the functional

$$\mathcal{J}_T(u_0) := \int_{\Omega} u(T, x) dx. \quad (2)$$

The goal is to maximise \mathcal{J}_T with respect to u_0 . Since we are then again wondering how to maximise the proportion of a subgroup by controlling its distribution at the initial time, it is natural to introduce a L^1 constraint on u_0 . This constraint is encoded by a parameter $m \in (0; |\Omega|)$ which is henceforth fixed.

These considerations lead us to defining our admissible class as

$$\mathcal{A} := \left\{ u_0 \in L^\infty(\Omega), 0 \leq u_0 \leq 1 \text{ a.e.}, \int_{\Omega} u_0 = m \right\}, \quad (3)$$

and the variational problem under scrutiny throughout this paper is

$$\max_{u_0 \in \mathcal{A}} \mathcal{J}_T(u_0). \quad (\mathbf{P}_f)$$

This problem (\mathbf{P}_f) was directly addressed by two of the authors in [44], where expressions for the first and second order optimality conditions were provided. We need to recall them, to motivate and contextualise our results: if we consider $u_0 \in \mathcal{A}$ and an admissible perturbation h at u_0 (in other words, a function h_0 such that for any $\varepsilon > 0$ small enough $u_0 + \varepsilon h_0 \in \mathcal{A}$) then the first order Gâteaux-derivative of \mathcal{J}_T at u_0 in the direction h_0 is

$$\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_{\Omega} h_0(x) p(0, x) dx, \quad (4)$$

where p solves the adjoint equation

$$\begin{cases} -\partial_t p - \Delta p = f'(u)p & \text{in } (0, T) \times \Omega, \\ p(T, x) = 1 & \text{in } \Omega, \\ \frac{\partial p}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega. \end{cases} \quad (5)$$

Here u is the solution of (1) with initial condition u_0 .

The main result in [44] states the following:

Theorem. [44] *There exist a solution $\bar{u}_0 \in \mathcal{A}$ of (\mathbf{P}_f) . Moreover, setting \bar{u} as the solution of (1) associated with this optimal initial data and \bar{p} as the unique solution of (5) for $u = \bar{u}$, there exists a non-negative real value \bar{c} such that*

- i) if $0 < \bar{u}_0(x) < 1$ then $\bar{p}(0, x) = \bar{c}$,*
- ii) if $\bar{p}(0, x) > \bar{c}$, then $\bar{u}_0(x) = 1$,*
- iii) if $\bar{p}(0, x) < \bar{c}$, then $\bar{u}_0(x) = 0$.*

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

$$f'(\bar{u}_0(x)) = -\bar{p}_t(0, x)/\bar{p}(0, x) \quad (6)$$

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

The characterisation of \bar{u}_0 with the help of p is almost complete here, except on the singular arc $\omega = \{0 < u_0 < 1\}$. Note first that this singular arc might have a positive measure. It was even proved in [44] that, if f is concave, then $\omega \equiv \Omega$. If f' is monotonic, equation (6) admits a unique solution and thus fully characterizes \bar{u}_0 . But for a bistable nonlinearity $f_\theta(u) = u(1-u)(u-\theta)$, equation (6) might have two solutions, one belonging to $[0; \eta]$ and the other to $(\eta; 1]$, where $\eta = \eta(\theta) \in (0; 1)$ is the unique real number such that f_θ is convex in $[0; \eta]$ and concave in $(\eta; 1]$. It is then necessary to distinguish between these two possible roots in order to completely characterize \bar{u}_0 with p .

Also, the characterisation given by this result naturally leads to a gradient descent algorithm, which is not well-posed if we are not able to characterize \bar{u}_0 on ω . This was one of the main limitation of the algorithm described in [44], and it is thus crucial to push the characterisation further in order to, somehow, close the algorithm.

1.2.2 Related works

A related problem has first been addressed by Garnier, Hamel and Roques in [21], where the authors consider a bistable reaction term $f(u) := u(1-u)(u-\theta)$, with $\theta \in (0, 1)$, over the full line $\Omega = \mathbb{R}$. In this earlier paper, the authors did not investigate (\mathbf{P}_f) , but they tried to optimize the initial datum in order to ensure the convergence to $u \equiv 1$ when $t \rightarrow +\infty$. They investigated numerically the particular case $u_0 := \mathbb{1}_{(-\alpha-\frac{m}{2}, -\alpha)} + \mathbb{1}_{(\alpha, \alpha+\frac{m}{2})}$, and proved that in some situations the initial datum associated with $\alpha = 0$ might lead to extinction (that is, $u(t, x) \rightarrow 0$ as $t \rightarrow +\infty$), while a positive $\alpha > 0$ might lead to persistence (that is, $u(t, x) \rightarrow 1$ as $t \rightarrow +\infty$). Also, numerics for more general classes of initial datum indicate that fragmentation might favor species persistence. Hence, even if the problem we consider here is a bit different, we expect the maximiser to be fragmented, that is, non-smooth, for bistable nonlinearities.

More recently, this problem was also addressed in [32], in a slightly more general form, and for the criterion $\int_\Omega |1 - u(T, x)|^2 dx$. These authors investigated in particular various conditions ensuring that the maximiser \bar{u}_0 is constant with respect to x , and, reciprocally, that the constant initial datum is a local maximiser.

1.3 Main results of the paper

The main contributions of this paper are the following:

- When u_0 does not saturate the L^∞ constraints (i.e. when the set $\omega := \{0 < u_0 < 1\}$ has positive measure), we prove in theorem 1 that any maximiser \bar{u}_0 must necessarily be in a zone of concavity of f : in ω , $f''(\bar{u}_0) \leq 0$.

- When f is convex, theorem 2 characterizes explicitly a global maximiser, using rearrangement techniques. Here, the presence of Neumann boundary conditions prohibits using in a straightforward manner the results of [11], and we need to adapt some points of the proof to this case.
- When f is a bistable non-linearity, we improve the algorithm initially introduced in [44] and display several numerical simulations. One-dimensional simulations are displayed that exemplify the fact that theorem 1 significantly improves the computational time of optimisation algorithms. We also provide two-dimensional simulations.

1.3.1 Characterisation of the singular arc

Let us first recall the expression of the second order derivative [44]:

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \iint_{(0;T) \times \Omega} f''(u(t,x)) p(t,x) h^2(t,x) dx dt, \quad (7)$$

where p solves (5) and h solves

$$\begin{cases} \frac{\partial h}{\partial t} - \Delta h = f'(u)h \text{ in } (0;T) \times \Omega, \\ h(0,x) = h_0(x), \\ \frac{\partial h}{\partial \nu} = 0 \text{ on } (0;T) \times \partial\Omega. \end{cases} \quad (8)$$

We can now state our main result:

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

$$f''(\bar{u}_0(x)) \leq 0. \quad (9)$$

Remark 1. The method we put forth is reminiscent of one that was used in [44] to study the case of a constant initial condition and to prove that such a constant u_0 was always a maximiser in the case of monostable non-linearities. 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) singular arcs, and it is a very interesting question to prove that such a property holds for any point of the singular arc $\Omega_{\bar{c}}$. We comment on the main difficulties of this approach in the conclusion and only state, for the moment, that the main problem is related to the ubiquitous problem of separation of phases in homogenisation [2].

1.3.2 Convex non-linearities and rearrangements

We present, in this section, a characterisation of maximisers when the non-linearity f is convex, using rearrangement and symmetrisation techniques. It should be noted that, since we are working with Neumann boundary conditions, it is not possible to use directly well-known parabolic isoperimetric inequalities [11, 42, 8]. We refer to [36, 11, 49] and the references therein for an introduction to parabolic isoperimetric inequalities, and only underline here that the most precise results available in the literature only encompass the case of Dirichlet boundary conditions. For Neumann boundary conditions, a large literature [16, 20, 33] is devoted to such questions. Usually, it involves a comparison of the solution with the solution of a Dirichlet, or of a mixed Dirichlet-Neumann problem, and makes use of constants appearing in relative isoperimetric inequalities. It

is unclear whether these comparison results could be used in our case. Since we are working in the one-dimensional case, a direct adaptation of the proof of [11] yields the required results.

In order to state our result, let us introduce the following notation: let \tilde{u}_0 be defined as

$$\tilde{u}_0 := \mathbf{1}_{(0;m)} \in \mathcal{A}. \quad (10)$$

Theorem 2. *Assume f is a convex, \mathcal{C}^1 function such that $f(0) = 0$. Then \tilde{u}_0 is a solution of (\mathbf{P}_f) .*

Remark 2. • It should be noted that \tilde{u}_0 appears as the solution of many other optimisation problems in population dynamics, most specifically for the monostable case [13, 23, 38] with Neumann boundary conditions.

- The maximiser $\mathbf{1}_{(0;m)}$ is clearly not unique, since $\mathbf{1}_{(1-m;1)}$ is also a maximiser for example. This provides an example of non-uniqueness of the maximiser.
- As a corollary, $\tilde{u}_0 := \mathbf{1}_{(0;m)}$ is a minimiser of \mathcal{J}_T over \mathcal{A} if f is concave.

2 Proof of Theorem 1

Heuristics and notations We recall the expression of the second order derivative of \mathcal{J} :

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

where p solves (5) and h solves (8).

Let \bar{u}_0 be a solution of (\mathbf{P}_f) . We assume that the set $\Omega_{\bar{c}} := \{0 < \bar{u}_0 < 1\}$ has a non-empty interior. Let x an interior point of $\Omega_{\bar{c}}$. For any $\delta \in (0, 1/2)$, let $\Omega_{\bar{c}}^{\delta} := \{\delta \leq \bar{u}_0(x) \leq 1 - \delta\}$, so that $\Omega = \bigcap_{\delta > 0} \Omega_{\bar{c}}^{\delta}$. By assumption, there exists $\delta > 0$ such that Ω^{δ} has a non-empty interior. Let such a δ be fixed, and let x_0 be an interior point of Ω^{δ} : there exists $\ell > 0$ such that $[x_0 - \ell, x_0 + \ell] \subset \Omega_{\bar{c}}^{\delta}$. Take $h_0 \in L^{\infty}(\Omega)$ supported in $[x_0 - \ell, x_0 + \ell]$, 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 the second order optimality conditions yield

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

The issue is that this expression is almost impossible to handle in this form. Since we want to retrieve, from this inequality, an information of $f''(\bar{u}_0)$, we need to find a perturbation $h(0, \cdot)$ such that the ensuing solution h satisfies, roughly speaking,

$$h^2(t, x) \approx C \delta_{t=t_0} g(x).$$

As h is a solution of a parabolic equation, one possibility to obtain such a behaviour is to choose a highly oscillating initial condition, say $h(0, x) = \cos(kx)$ with a large integer k . This would give $h^2(t, x) \approx e^{-tk^2} \cos(kx)^2$, which, thanks to the Laplace method, does concentrate around $t = 0$ up to a proper rescaling. This however is not particularly convenient, as such a perturbation is not admissible: it is not supported in $(x_0 - \ell; x_0 + \ell)$. To overcome this difficulty, we need to truncate such highly oscillating perturbations, leading to two-scale asymptotic expansions.

Construction of the admissible perturbation Let $\theta \in \mathcal{D}((x_0 - \ell; x_0 + \ell))$ be a smooth function satisfying $0 \leq \theta \leq 1$ and $\theta \equiv 1$ on $(x_0 - \ell/4; x_0 + \ell/4)$, θ symmetric with respect to x_0 , and non-increasing on $(x_0; x_0 + \ell/4)$. We consider the following sequence of equations: let, for any $k \in \mathbb{N}$, h_k be the solution of

$$\begin{cases} \frac{\partial h_k}{\partial t} - \frac{\partial^2 h_k}{\partial x^2} = f'(u)h_k, \\ \frac{\partial h_k}{\partial \nu} = 0, \\ h_k(0, x) = \theta(x) \cos(kx). \end{cases} \quad (11)$$

In this context, it is natural [1] to look for a two-scale asymptotic expansion of h_k of the form

$$h_k(t, x) \approx h_k^0(k^2 t, x, kx) + \frac{1}{k} h_k^1(k^2 t, x, kx) + \dots \quad (12)$$

which, after a formal identification at the first and second order, gives the following equations on h_k^0 and h_k^1 :

$$\begin{cases} \frac{\partial h_k^0}{\partial s} - \frac{\partial^2 h_k^0}{\partial y^2} = 0, \\ \frac{\partial h_k^0}{\partial \nu} = 0, \\ h_k^0(0, x, y) = \theta(x) \cos(y). \end{cases} \quad (13)$$

and

$$\begin{cases} \frac{\partial h_k^1}{\partial s} - \frac{\partial^2 h_k^1}{\partial y^2} = 2 \frac{\partial^2 h_k^0}{\partial x \partial y}, \\ \frac{\partial h_k^1}{\partial \nu} = 0, \\ h_k^1(0, x, y) = 0. \end{cases} \quad (14)$$

Equation (13) can be solved explicitly, giving

$$h_k^0(s, x, y) = \theta(x) \cos(y) e^{-s}. \quad (15)$$

This, in turn, allows to solve equation (14) as

$$h_k^1(s, x, y) = -2s e^{-s} \theta'(x) \sin(y). \quad (16)$$

Proposition 1. *The asymptotic expansion (12) 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 $t \in (0; T)$,

$$\|R_k(t, \cdot)\|_{L^2(\Omega)} \leq \frac{M}{k^2}. \quad (17)$$

In particular,

$$\iint_{(0;T) \times \Omega} R_k^2 \leq \frac{M^2}{k^4}, \quad \int_0^T \|R_k\|_{L^2(\Omega)} \leq \frac{MT}{k^2}. \quad (18)$$

Proof of Proposition 1. To prove this Proposition, we write down the equation satisfied by R_k . Straightforward computations show that R_k solves

$$\partial_t R_k - \Delta R_k - f'(u)R_k := f'(u) \left(h_k^0 + \frac{1}{k} h_k^1 \right) + \frac{\partial^2 h_k^0}{\partial x^2} + 2 \frac{\partial^2 h_k^1}{\partial x \partial y} + \frac{1}{k} \frac{\partial^2 h_k^1}{\partial x^2}, \quad (19)$$

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

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

First of all, since $0 \leq u \leq 1$ and $f \in \mathcal{C}^1$, there exists $M_0 > 0$ such that

$$\|W_0\|_{L^\infty((0;T) \times \Omega)} \leq M_0. \quad (20)$$

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

Lemma 1. *There exists $\tilde{M} > 0$ such that*

$$\int_0^T \|V_{0,k}(t, \cdot)\|_{L^2(\Omega)} \leq \frac{\tilde{M}}{k^2}, \quad (21)$$

$$\int_0^T \|V_{1,k}(t, \cdot)\|_{L^2(\Omega)} dt \leq \frac{\tilde{M}}{k^2}. \quad (22)$$

$$\int_0^T \|V_{2,k}(t, \cdot)\|_{L^2(\Omega)} dt \leq \frac{\tilde{M}}{k^2}. \quad (23)$$

Proof of Lemma 1. We prove the three estimates separately. Let us recall the following consequence of the Laplace method: for any integer $m \in \mathbb{N}^*$, one has

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

Proof of (21)

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

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

We first use the explicit expressions (15)-(16) for h_k^0 and h_k^1 to obtain, using $\|\theta\|_{L^\infty} \leq 1$,

$$\begin{aligned} \|h_k^0(k^2t, \cdot, k \cdot)\|_{L^2(\Omega)}^2 &= \int_{\Omega} \theta(x)^2 \cos(kx)^2 e^{-2k^2t} dt \\ &\leq e^{-2k^2t} |\Omega|, \end{aligned} \quad (24)$$

and integrating this inequality between 0 and T gives

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

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

$$\begin{aligned} \|h_k^1(k^2t, \cdot, k \cdot)\|_{L^2(\Omega)}^2 &= 4k^4 t^2 e^{-2k^2t} \int_{\Omega} (\theta'(x))^2 \sin(x)^2 dx \\ &\leq C k^4 t^2 e^{-2k^2t} |\Omega| \cdot \|\theta'\|_{L^\infty}^2, \end{aligned} \quad (25)$$

for some constant C . Taking the square root, we get, for a constant C' ,

$$\frac{1}{k} \int_0^T \|h_k^1(k^2t, \cdot, k \cdot)\|_{L^2(\Omega)} dt \leq C' |\Omega| \cdot \|\theta'\|_{L^\infty} k \int_0^T t e^{-k^2t} dt.$$

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

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

for some constant M_1 .

Summing these two contributions gives the required result.

Proof of (22) This follows from the same arguments, by simply observing that

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

Proof of (23) We once again split the expression and estimate separately

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

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

$$\frac{\partial^2 h_k^1(k^2t, \cdot, k \cdot)}{\partial x \partial y} = -4k^2 t e^{-k^2t} \theta''(x) \cos(y).$$

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

$$\left\| \frac{\partial^2 h_k^1(k^2t, \cdot, k \cdot)}{\partial x \partial y} \right\|_{L^2(\Omega)} \leq 4k^2 t e^{-k^2t} |\Omega| \cdot \|\theta''\|_{L^\infty}$$

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

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

for some constant M_2 . The proof of the control of the second term follows along exactly the same lines. □

Let us now prove estimate (17). The equation on R_k rewrites

$$\partial_t R_k - \Delta R_k - W_0 R_k = W_0 V_{0,k} + V_{1,k} + V_{2,k}. \quad (26)$$

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

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

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

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

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

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

Hence, by Lemma 1 we get for some constant N_0 and any $t \in (0; T)$,

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

□

Back to the proof We turn back to the proof of Theorem 1.

Proof of Theorem 1. We use the same notations x_0, θ as above. Let us now consider the initial perturbation $h_{0,k} := \theta(x) \cos(kx)$. We define

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

so that the second order derivative of \mathcal{J}_T in \bar{u}_0 rewrites as

$$\begin{aligned} d^2 \mathcal{J}_T(\bar{u}_0)[h_{0,k}, h_{0,k}] &= \int_0^T \int_{\Omega} F(t, x) h_k(t, x)^2 dx dt, \\ &= \int_0^T \int_{\Omega} F(t, x) (R_k(t, x) + V_{0,k}(t, x))^2 dx dt, \\ &= \int_0^T \int_{\Omega} F(t, x) \left(R_k(t, x)^2 + 2R_k(t, x)V_{0,k}(t, x) + V_{0,k}(t, x)^2 \right) dx dt. \end{aligned} \quad (27)$$

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

$$\|F\|_{L^\infty((0;T) \times \Omega)} \leq M_3. \quad (28)$$

Gathering (28) and (18) it follows that

$$\int_0^T \int_{\Omega} F(t, x) R_k(t, x)^2 dx dt = \mathcal{O}(k^{-4}) \quad (29)$$

and similarly, gathering (21) and (17) we obtain

$$\int_0^T \int_{\Omega} F(t, x) R_k(t, x) V_{0,k}(t, x) dx dt = \mathcal{O}(k^{-4}) \quad (30)$$

Let us now studying the term

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

Once again we split the expression. Applying Cauchy-Swchartz inequality, and the estimates in (24) and (25) it follows that the second term verify

$$\begin{aligned} \left| 2 \frac{1}{k} \int_0^T \int_{\Omega} F(t, x) h_k^0(t, x) h_k^1(t, x) dx dt \right| &\leq 2 \frac{M_3}{k} \int_0^T \|h_k^0\|_{L^2(\Omega)} \|h_k^1\|_{L^2(\Omega)} dt \\ &\leq 2M_4 k \int_0^T t e^{-2k^2 t} dt, \\ &= \mathcal{O}\left(\frac{1}{k^3}\right). \end{aligned} \quad (31)$$

The last step in the above expression follows directly from (\mathbf{I}_m) with $m = 2$. Similarly for the third term we obtain

$$\begin{aligned} \left| \frac{1}{k^2} \int_0^T \int_{\Omega} F(t, x) h_k^1(t, x)^2 dx dt \right| &\leq \frac{M_3}{k^2} \int_0^T \|h_k^1\|_{L^2(\Omega)}^2 dt \\ &\leq M_4 k^2 \int_0^T t^2 e^{-2k^2 t} dt, \\ &= \mathcal{O}\left(\frac{1}{k^4}\right). \end{aligned} \quad (32)$$

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

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

$$\int_0^T \int_{\Omega} F(t, x) h_k^0(t, x)^2 dx dt = \int_0^T e^{-2k^2 t} G(t) dt \quad (33)$$

with $G := \int_{\Omega} F(t, x) \theta(x)^2 \cos(kx)^2 dx$ is a continuous function of time as a consequence of parabolic regularity.

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

$$\int_0^T \int_{\Omega} F(t, x) h_k^0(t, x)^2 dx dt \sim \frac{1}{2k^2} G(0) \quad (34)$$

Gathering the estimates in (29), (30), (31), (32), (34) and plugging them into the second derivative of the functional \mathcal{J}_T given by (27) it follows that

$$\nabla^2 \mathcal{J}_T(\bar{u}_0)[h_{0,k}, h_{0,k}] \underset{k \rightarrow \infty}{\sim} \frac{1}{2k^2} G(0) \quad (35)$$

However, x_0 was chosen such that

$$f''(u_0(x_0)) > 0,$$

and θ as 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 $\sin(kx) \xrightarrow[k \rightarrow \infty]{} \frac{1}{2}$ one has

$$G(0) = \int_{\Omega} F(0, x)\theta(x)^2 \cos(kx)^2 dx \xrightarrow[k \rightarrow \infty]{} \frac{1}{2} \int_{\Omega} F(0, x)\theta(x)^2 dx \geq 0$$

this would mean that for k sufficiently large,

$$k^2 \nabla^2 \mathcal{J}_T(\bar{u}_0)[h_{0,k}, h_{0,k}] > 0$$

which contradicts the fact that u_0 is a maximiser of \mathcal{J}_T . The proof of the Theorem is complete. \square

3 Proof of Theorem 2

The proof follows essentially from the same arguments as in [11]. We thus only present the modifications that are in order so as to apply the methods of [11]. We define $g(u) := f(u) + cu$, with $c > \|f'\|_{L^\infty}$, so that g is increasing.

Reduction to a bang-bang maximiser. We recall that bang-bang functions are defined as characteristic functions of subsets of Ω , that is, functions only taking values 0 and 1. First, as f is convex, it follows from the same arguments as Proposition 6 of [44] that \mathcal{J}_T is convex. Hence, one can restrict to maximisers among the extremal points of \mathcal{A} . These are exactly bang-bang functions: u_0 satisfies $u_0 = 0$ or 1 almost everywhere on $(0, \pi)$.

Reduction to a periodic problem. Next, for all $t \in [0, T]$, we extend $u(t, \cdot)$ to $(-\pi, \pi)$ by symmetrisation with respect to 0, and we then extend it to \mathbb{R} by 2π -periodicity. The Neumann boundary conditions at $x = 0$ and $x = \pi$ ensure that the extended function is of class \mathcal{C}^1 , and it thus satisfies the equation on the torus:

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u + cu = g(u) & \text{in } (0; T) \times \mathbb{T}, \\ u(0, \cdot) = u_0. \end{cases} \quad (36)$$

Let us also recall some basic facts about rearrangements.

Rearrangements and comparison relationship We recall the definition of the periodic rearrangement: for any periodic function $u : \mathbb{T} \rightarrow \mathbb{R}_+$ if we identify \mathbb{T} with $[-\pi; \pi]$ there exists a unique symmetric (with respect to 0) decreasing function $u^* : \mathbb{T} \rightarrow \mathbb{R}$ that has the same distribution function as u . We recall that the distribution function of u is

$$\mu_v(t) := \text{Vol}(\{v \geq t\})$$

and that the periodic rearrangement of u is the left inverse of μ_u . Up to a translation u^* is symmetric with respect to 0 and decreasing, and we always assume that u^* is translated so as to be symmetric with respect to 0.

Proposition 2. *Let v the solution of (36) associated with the initial datum u_0^* . Then*

$$\forall r \in (0; \pi), \int_{-r}^r v(t, x) dx \geq \int_{-r}^r u^*(t, x) dx.$$

In particular, taking $t = T$ and $r = \pi$:

$$\int_{-\pi}^{\pi} v(T, x) dx \geq \int_{-\pi}^{\pi} u(T, x) dx. \quad (37)$$

As explained in the introduction, Proposition 2 follows simply by adapting minor points in the proofs of [11].

Proof of Proposition 2. Let us consider $u_0 \in \mathcal{A}$, that we symmetrise and extend to \mathbb{T} by periodicity. For any $\tau > 0$ integrate the associated equation (36) over the level set $\{u(t, \cdot) > \tau\}$ (this is why we needed to reduce ourselves to the case of periodic boundary conditions- Neumann boundary conditions may lead to possible trouble if this level set intersects the boundary of the domain). Since we are working in the torus and since we have the classical isoperimetric inequality in this context, it follows from standard arguments [11, 42] that, setting

$$k_u = \int_{u(t, \cdot) > \tau} u$$

we have

$$\frac{\partial k_u}{\partial t} - 4 \frac{\partial^2 k_u}{\partial \tau^2} + ck_u \leq \int_{\{u(t, \cdot) > \tau\}} g(u) \leq \int_0^{\mu_u(t, \tau)} g(u) \leq \int_0^{\mu_u(t, \tau)} g(u^*)$$

The last inequality comes from the fact that g is convex and increasing with respect to v . The conclusion follows in the same way as in [11]. Namely, one can prove that k_v is a solution of the equation related to this inequation, and thus the parabolic maximum principle yields $k_u \leq k_v$, which is equivalent to the property we seek to prove. \square

Conclusion. Assume that u_0 is a bang-bang maximiser of problem (\mathbf{P}_f) . Symmetrise it and extend it by periodicity. Consider the symmetric decreasing rearrangement u_0^* of its extension. Then $\int_{-\pi}^{\pi} v(T, x) dx \geq \int_{-\pi}^{\pi} u(T, x) dx$ by Proposition 2, where v is the solution of the periodic equation (36) associated with the initial datum u_0^* . Clearly, $v(t, \cdot)$ and $u(t, \cdot)$ are symmetric with respect to $x = 0$ for all time $t > 0$. Hence, $\int_0^{\pi} v(T, x) dx \geq \int_0^{\pi} u(T, x) dx$. Also, one easily remarks that v restricted to $(0, \pi)$ is the solution of the parabolic equation with Neumann boundary conditions (1), associated with the initial datum u_0^* restricted to $(0, \pi)$. On the other hand, as u_0 is bang-bang, one has $u_0^* = \mathbf{1}_{(-m, m)}$. Hence, $\mathbf{1}_{(0, m)}$ increases the criterion in (\mathbf{P}_f) . Thus, it is a solution of (\mathbf{P}_f) .

4 Numerical analysis in the bistable framework

As we explained in the introduction, the behaviour of optimisers vary wildly depending on the shape of the reaction term f . To exemplify this phenomenon, we rely on the bistable non-linearity that motivated [44]. Consider $f(u) := u(1 - u)(u - \theta)$, with $\theta \in (0, 1)$.

When considering the optimisation problem (\mathbf{P}_f) , 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.

Let us first recall the main principles of the numerical algorithm introduced in [44] and explain the difficulty related to $\{p = c\}$ further. 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 (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

$$|\{p_0^n > c^n\}| < m, |\{p_0^n \geq c^n\}| > m, |\{p_0^n = c^n\}| > 0. \quad (38)$$

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 of u_0^{n+1} on ω_n . Defining $\omega_n := \{p_0^n = c^n\}$ and discretising equation (5) on ω_n we obtain, with an explicit finite difference scheme

$$-\left(\frac{p_0^n(dt, x) - c^n}{dt}\right) = f'(u_0^{n+1})c^n \quad (39)$$

and the value on u_0^{n+1} on ω_n must be a root of (39). 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,

$$f''(\mu_1^n) > 0, f''(\mu_2^n) < 0. \quad (40)$$

In [44] this difficulty is overcome by examining the two different possibilities and choosing the best one, which significantly lessens the performance of the algorithm, but Theorem 1 allows to overcome this difficulty by choosing directly the root μ_2^n , which is in the "concavity" zone of f .

4.1 Comparison of different numerical methods in the one-dimensional case

In this section, we want to study an example in order to compare the performances of our numerical algorithm with other well known optimisation algorithms to solve general nonlinear problems under constraints. More precisely, we will consider the following numerical methods:

- *Method 1*: The numerical algorithm introduced in [44], which we improve using theorem 1, which will be referred to as *our algorithm*.
- *Method 2*: The *interior-point* method, which is used to solve optimisation problems with linear equality and inequality constraints by applying Newton's method to a sequence of equality constrained problems. For a more detailed description of this method see for instance [15].
- *Method 3*: The *sequential quadratic programming* (SQP), which solves a sequence of optimisation sub-problems, each of which optimizes a quadratic model of the objective function subject to a linearisation of the constraints, see for instance [46].
- *Method 4*: The *simulated annealing* method, which is a probabilistic technique to approximate global optimisation in a large search space. See for instance [28] for more details on this technique.

We used 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.

Setting the data Let us consider $\Omega = [-50, 50]$, and $m = 13$; thus the admissible set is defined as follows:

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

Note that this set is defined by two inequalities and an equality constraint. We aim at maximising the quantity $\mathcal{J}_T(u_0) := \int_{\Omega} u(T, x) dx$ for $T = 25$. We consider a bistable reaction term $f(u) := u(1 - u)(u - 0.25)$.

In order to compare the performance of the four algorithms under the same conditions, we consider the same discretisation of Ω . Moreover, the solution of the equation is systematically computed by the Crank-Nicolson method, and for the initialisation we will consider the same u_0^0 given by a single block of mass 13. The value of the objective function at each iteration is numerically approximated by the rectangle rule. In particular, for the initialisation we have $\mathcal{J}_{25}(u_0^0) = 29.42$.

The results of the simulations are shown in Fig.1 and Table 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 1: Comparing algorithms

Algorithm	Objective function $\mathcal{J}_{25}(\bar{u}_0)$	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 optimisers 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 maximiser but only to a local one.

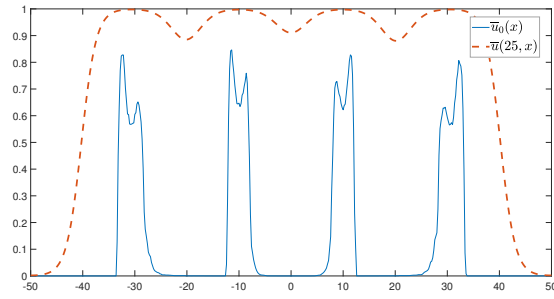
4.2 Numerical simulations in the two-dimensional case

We only considered in the present paper the one-dimensional case. We now display some numerical results obtained in dimension 2, for which new patterns might arise.

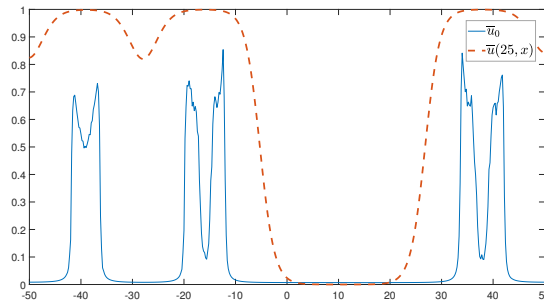
To solve the reaction-diffusion equation in the two-dimensional case, we consider the alternating direction implicit method (ADI) which is a classic method to solve parabolic and elliptic partial differential equations in two or three dimensions. As in the one-dimensional case, the algorithm and routines were coded in MATLAB.

We consider a square domain $\Omega = [-10, 10] \times [-10, 10]$, discretised uniformly by squares of side $dx = 0.4$. We fix $T = 30$ and a bistable reaction term

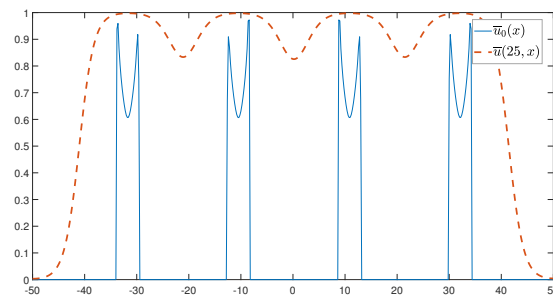
$$f(u) = u(1 - u)(u - 0.25).$$



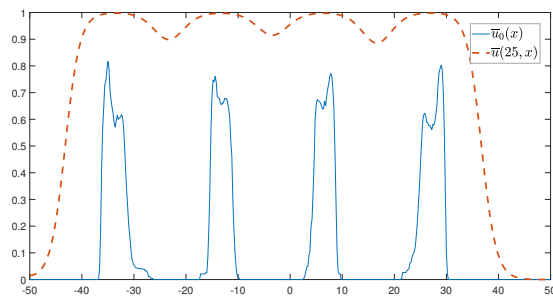
(a) Our algorithm



(b) Interior-point method



(c) Sequential quadratic programming



(d) Simulated annealing

Figure 1: Optimum found by means of the four different numerical algorithms.

Example 1 The algorithm is initialised with a ball of full density spotted in the middle of the domain Ω , the mass is fixed to $m = 6\pi$, see Fig.2 (a). After 27 iterations, the algorithm converges to the local optimum showed in Fig.2(b). The evolution of the objective function \mathcal{J}_{30} through iterations is showed in Fig.3

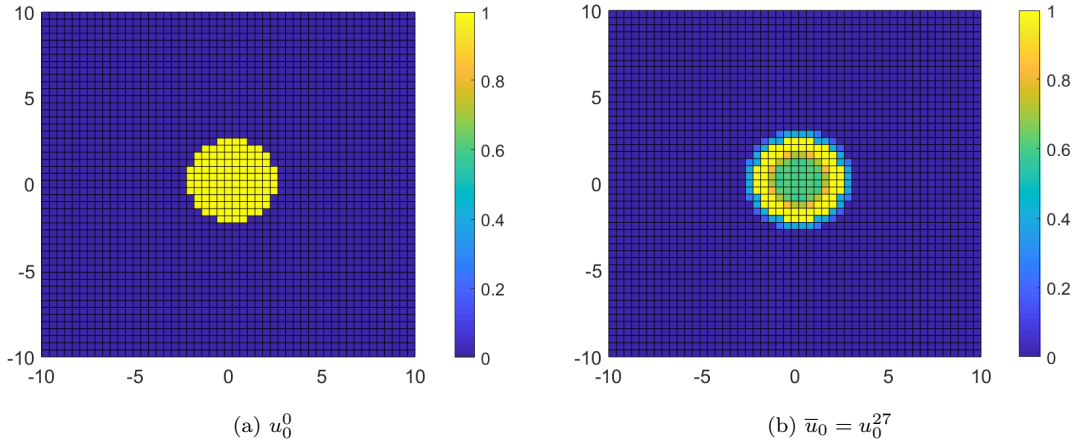


Figure 2: In the left-hand side, we show the input of the algorithm, given by the ball of radius $r = \sqrt{6}$ centered at the origin. In the right-hand side, we display the local optimum found by the numerical algorithm after 27 iterations, which looks radial but which is no longer a bang-bang distribution it does not only take values 1 and 0.

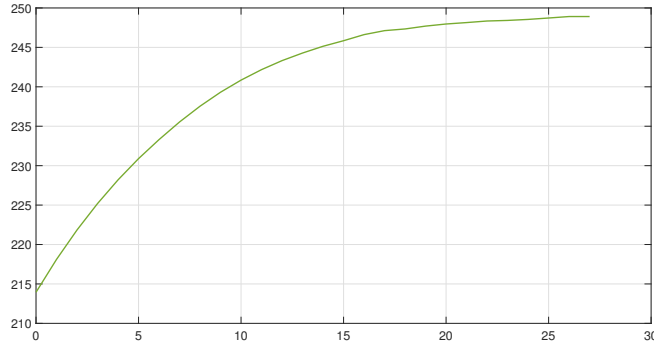


Figure 3: Evolution of the objective function from the initialisation $\mathcal{J}_{30}(u_0^0) = 213.9$ to the last iteration $\mathcal{J}_{30}(u_0^{27}) = 248.9$.

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 equation (5), 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.4.

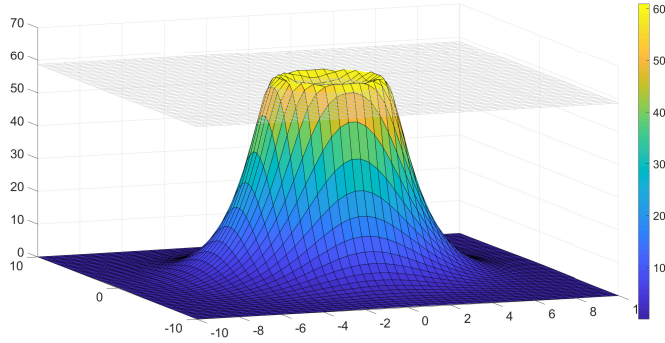


Figure 4: The figure shows the surface given by the solution $\bar{p}(0, x)$ of the adjoint problem defined by the eq. 5 associated to the initial data \bar{u}_0 found by our algorithm. The plane colored in gray, is associated to the value \bar{c} described in Theorem 1.2.1 and thus for every $x \in \Omega$ such that $p_0(\mathbf{x}) = \bar{c}$, one has $0 < \bar{u}_0(x) < 1$, see Fig.2 (b).

Example 2 In this case we refine the discretisation and take $dx = 0.27$. We consider an initial data which is a stripe of mass $m = 16.21$ and full density dividing our domain into two equal regions of zero density, see Fig.5 (a). The algorithm converges after 37 iterations and the local optimum is displayed in Fig.5 (b). The corresponding variations of the objective function is showed in Fig.6

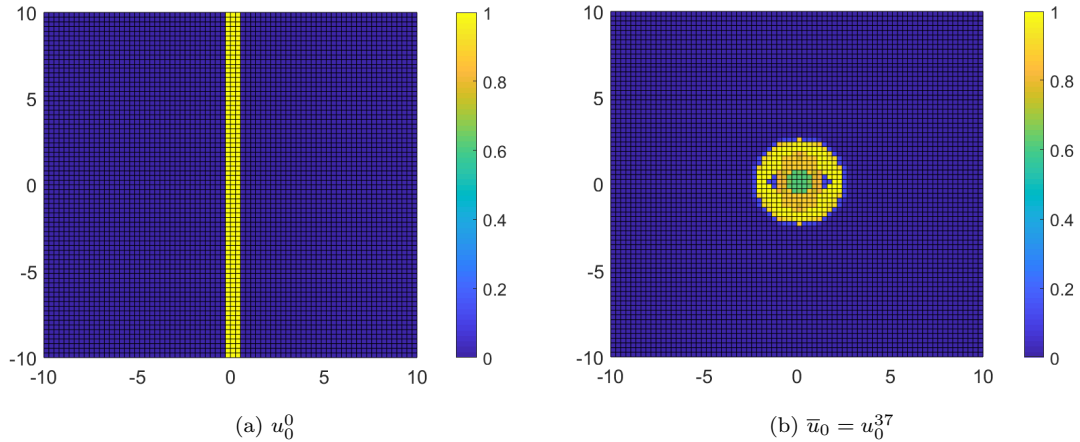


Figure 5: In the left-hand side is showed the input of the algorithm, given by the stripe of width $r = 0.81$ centered at the origin. In the right-hand side, the local optimum found by the numerical algorithm after 37 iterations.

Even though the initial mass is clearly not enough to guarantee convergence to 1, the value of the objective function increases 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

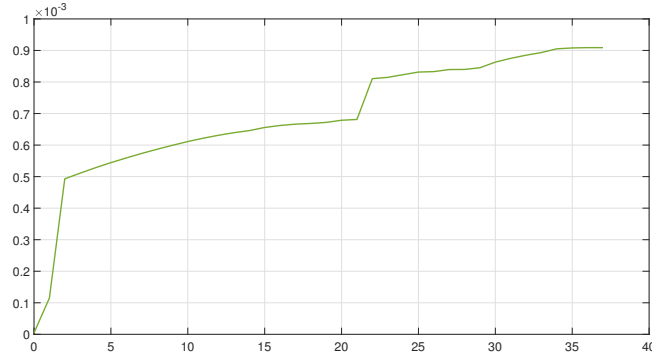


Figure 6: Evolution of the objective function from the initialisation $\mathcal{J}_{30}(u_0^0) = 4.5 \times 10^{-6}$ to the last iteration $\mathcal{J}_{30}(u_0^{37}) = 0.9 \times 10^{-3}$.

in the middle of regions of full density, which exemplifies the phenomenon described for the one-dimensional case in [21].

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 stripe of full density dividing the domain into two regions of zero density, like in the example 2, see Fig.7 (a).

The corresponding local optimum found by the numerical algorithm is showed in Fig.7 (b). As in the previous case, the optimiser 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.8 reflect that, in this case, the fixed mass is sufficient to trigger an increase in total density over time.

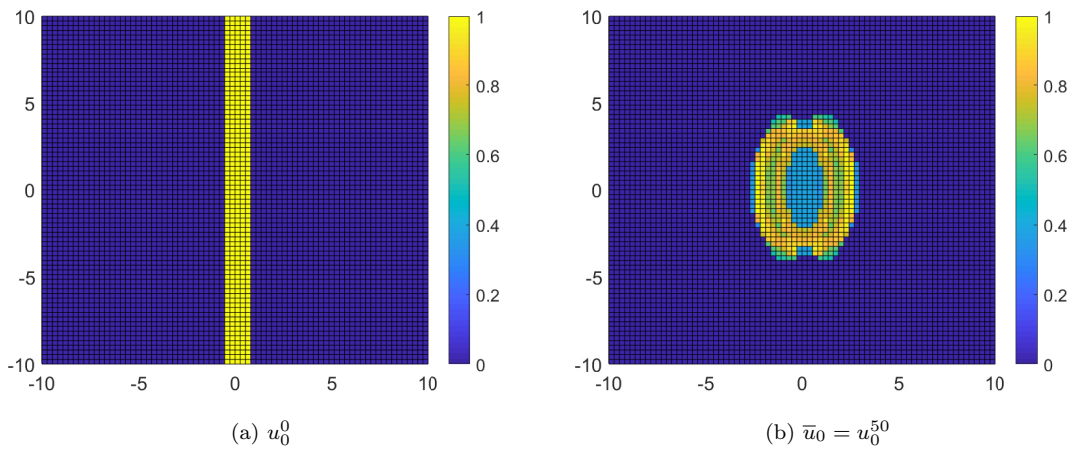


Figure 7: In the left-hand side is showed the input of the algorithm, given by the stripe of width $r = 1.6$ centered at the origin. In the right-hand side, the local optimum found by the numerical algorithm after 50 iterations.

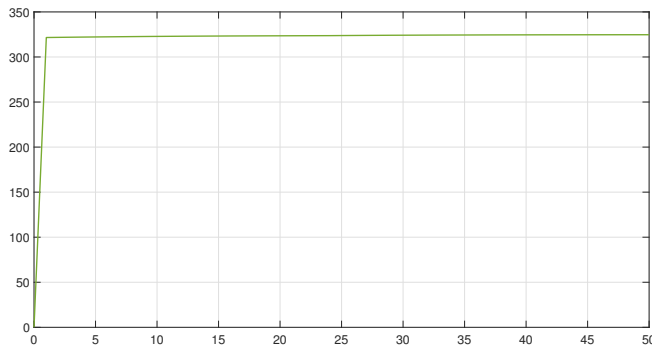


Figure 8: Evolution of the objective function from the initialisation $\mathcal{J}_{30}(u_0^0) = 3 \times 10^{-5}$ to the last iteration $\mathcal{J}_{30}(u_0^{50}) = 324$.

This example point out the non suitability of stripe type initial distributions. Note from Fig.8 that, despite the considerable increase of the initial mass with respect to the example 2, the values of the objective function $\mathcal{J}_{30}(u_0^0)$ associated to the stripe is of the order of 10^{-5} , which is very low compared with the value associated to the final distribution.

Lastly, let us mention that a possible approximation of the maximiser was discussed in the Appendix of [34]. Namely, in this thesis the author replaced the maximiser \bar{u}_0 by its mean on each of the connected components of $\{\bar{u}_0 > 0\}$. This gives pretty good results in several cases. It would be good to manage to quantify analytically the difference of criterion between this approximated initial datum and the global maximiser.

5 Conclusion, open problems and possible extensions

We make, in this conclusion, several concluding remarks and comments about possible generalisations and extensions of the results presented in this paper. For each of them, we try to present the arguments that have led us to the conclusion that other approaches were necessary in general.

5.1 Regarding the regularity of the singular arc

One of the main drawbacks of Theorem 1 is the regularity assumption on the singular arc. Namely, we obtain, for a maximiser u_0 of \mathcal{J}_T , the characterisation $f''(u_0) \leq 0$ only on the interior of the singular arc $\{0 < u_0 < 1\}$. The method presented in this paper (two-scale expansions) strongly relies on the smoothness properties of the cut-off function θ .

It may be tempting, in the one dimensional case, to overcome this difficulty arguing as in [37]: if we simply assume that the singular arc ω is measurable, but has positive measure, one can show that for every $K \in \mathbb{N}$ there exists $h_K \in L^2(\Omega)$, supported in ω , that writes

$$h_K = \sum_{j \geq K} \alpha_{j,K} \cos(j \cdot), \|h_K\|_{L^2} = 1. \quad (41)$$

Using the fact that h_K only has high Fourier modes, one may hop for a two-scale expansion of the form

$$h_K \approx \sum_{j \geq K} \alpha_{j,K} \left(h_k^0(k^2 t, x, kx) + \frac{1}{k} h_k^1(k^2 t, x, kx) \right).$$

This is however *a priori* prohibited by the problem of separation of phase: to obtain such a description, one needs well-separated phases, in the sense of [2]. In the context of Fourier series, this would require, at the very least, that h_K should write as a lacunary Fourier series (typically, $h_K = \sum_{j=0}^{\infty} \alpha_{j,K} \cos(K^j x)$). However, for such lacunary Fourier series, Zygmund's theorem (see [25] for instance) prohibits that they have compact support, so that admissible perturbations can not have this structure. This is a major drawback, and it is unclear whether or not one may be able to overcome this difficulty via a similar approach, or if an entirely new strategy needs to be devised.

Another approach would be to prove some regularity on ω ensuring that almost every of its point lie in its interior. This is satisfied for example if \bar{u}_0 is Riemann integrable (since, due to Lebesgue's characterisation of Riemann integrable functions, almost every point is a continuity point of \bar{u}_0). Riemann integrability is satisfied by BV functions. Unfortunately, we were not able to push the regularity further than L^∞ .

5.2 The singular arc in higher dimensions

We believe it plausible that we may adapt the methods of Theorem 1 to obtain a characterisation of the singular arc analogous to that of Theorem 1 in the case $\Omega = \prod_{i=1}^N [0; a_i]$, $a_i > 0$. To do so, the main difference with our proof would be to replace the initial perturbation $\theta(x) \cos(kx)$ with $\prod_{i=1}^N \theta(x) \cos(kx_i)$.

5.3 Rearrangement inequalities for other types of boundary conditions

In this work, we mostly dealt with the case of Neumann boundary conditions in the one-dimensional case. We ought to note two things: first, the proof of theorem 1 should hold in the case of Dirichlet or of Robin boundary conditions, provided the functions $\cos(k\cdot)$, in the proof, are replaced with the Dirichlet or Robin eigenfunctions of the laplacian in the interval. Second, regarding theorem 2, the same type of results can be obtained in a straightforward manner for the case of Dirichlet boundary conditions, by applying directly [11]. The case of Robin boundary conditions may be encompassed by using the recent Talenti inequalities obtained in this case in [6]. Addressing the problem on the full line \mathbb{R} is more tricky, since in this case even the existence of a maximiser is unclear. We plan on investigating such matters in future works.

References

- [1] G. Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6):1482–1518, 1992.
- [2] G. Allaire and M. Briane. Multiscale convergence and reiterated homogenisation. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 126(2):297–342, 1996.
- [3] L. Almeida, Y. Privat, M. Strugarek, and N. Vauchelet. Optimal releases for population replacement strategies: application to *wolbachia*. *SIAM J. Math. Anal.*, 51(4):3170–3194, 2019.
- [4] A. Alvino, P. Lions, and G. Trombetti. A remark on comparison results via symmetrization. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 102(1-2):37–48, 1986.
- [5] A. Alvino, P.-L. Lions, and G. Trombetti. Comparison results for elliptic and parabolic equations via symmetrization: a new approach. *Differential Integral Equations*, 4(1):25–50, 1991.

- [6] A. Alvino, C. Nitsch, and C. Trombetti. A talenti comparison result for solutions to elliptic problems with robin boundary conditions. *arXiv: Analysis of PDEs*, 2019.
- [7] A. Alvino, G. Trombetti, and P. Lions. On optimization problems with prescribed rearrangements. *Nonlinear Analysis: Theory, Methods & Applications*, 13(2):185–220, Feb. 1989.
- [8] A. Alvino, G. Trombetti, and P.-L. Lions. Comparison results for elliptic and parabolic equations via schwarz symmetrization. *Annales de l’Institut Henri Poincare (C) Non Linear Analysis*, 7(2):37–65, Mar. 1990.
- [9] X. Bai, X. He, and F. Li. An optimization problem and its application in population dynamics. *Proc. Amer. Math. Soc.*, 144(5):2161–2170, May 2016.
- [10] X. Bai, X. He, and F. Li. An optimization problem and its application in population dynamics. *Proc. Amer. Math. Soc.*, 144(5):2161–2170, 2016.
- [11] C. Bandle. *Isoperimetric Inequalities and Applications*. Monographs and studies in mathematics. Pitman, 1980.
- [12] N. Barton and M. Turelli. Spatial waves of advances with bistable dynamics: cytoplasmic and genetic analogues of Allee effects. *The American Naturalist*, 78(3):E48–E75, 2011.
- [13] H. Berestycki, F. Hamel, and L. Roques. Analysis of the periodically fragmented environment model : I – species persistence. *Journal of Mathematical Biology*, 51(1):75–113, 2005.
- [14] J. BINTZ and S. LENHART. Optimal resources allocation for a diffusive population model. *Journal of Biological Systems*, 28(04):945–976, Dec. 2020.
- [15] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [16] M. Bramanti. Symmetrization in parabolic neumann problems. *Applicable Analysis*, 40(1):21–39, Jan. 1991.
- [17] R. S. Cantrell and C. Cosner. The effects of spatial heterogeneity in population dynamics. *J. Math. Biol.*, 29(4):315–338, 1991.
- [18] F. Caubet, T. Deheuvels, and Y. Privat. Optimal location of resources for biased movement of species: the 1D case. *SIAM Journal on Applied Mathematics*, 77(6):1876–1903, 2017.
- [19] J. Evans. Nerve axon equations: Iv the stable and unstable impulse. 1975.
- [20] V. Ferone and A. Mercaldo. Neumann problems and steiner symmetrization. *Communications in Partial Differential Equations*, 30(10):1537–1553, Sept. 2005.
- [21] J. Garnier, L. Roques, and F. Hamel. Success rate of a biological invasion in terms of the spatial distribution of the founding population. *Bulletin of Mathematical Biology*, 74(2):453–473, 2012.
- [22] J. Inoue, , and K. Kuto. On the unboundedness of the ratio of species and resources for the diffusive logistic equation. *Discrete & Continuous Dynamical Systems - B*, 22(11):0–0, 2017.
- [23] C.-Y. Kao, Y. Lou, and E. Yanagida. Principal eigenvalue for an elliptic problem with indefinite weight on cylindrical domains. *Math. Biosci. Eng.*, 5(2):315–335, 2008.
- [24] S. Kesavan. *Symmetrization and Applications*. WORLD SCIENTIFIC, Apr. 2006.

- [25] O. Kovrizhkin. A version of the uncertainty principle for functions with lacunary fourier transforms. *Journal of Mathematical Analysis and Applications*, 288(2):606–633, Dec. 2003.
- [26] K.-Y. Lam, S. Liu, and Y. Lou. Selected topics on reaction-diffusion-advection models from spatial ecology. preprint.
- [27] J. Lamboley, A. Laurain, G. Nadin, and Y. Privat. Properties of optimizers of the principal eigenvalue with indefinite weight and Robin conditions. *Calculus of Variations and Partial Differential Equations*, 55(6), Dec. 2016.
- [28] M. Locatelli. Simulated annealing algorithms for continuous global optimization: Convergence conditions. *Journal of Optimization Theory and Applications*, 104:121–133, 2000.
- [29] Y. Lou. *Some Challenging Mathematical Problems in Evolution of Dispersal and Population Dynamics*, pages 171–205. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [30] Y. Lou, K. Nagahara, and E. Yanagida. Maximizing the total population with logistic growth in a patchy environment. *Submitted*, 2020.
- [31] Y. Lou and E. Yanagida. Minimization of the principal eigenvalue for an elliptic boundary value problem with indefinite weight, and applications to population dynamics. *Japan J. Indust. Appl. Math.*, 23(3):275–292, 2006.
- [32] Y. P. M. Duprez, R. Hélie and N. Vauchelet. Optimization of spatial control strategies for population replacement, application to wolbachia. *Peprint*, 2021.
- [33] C. Maderna, S. Salsa, and C. Pucci. Symmetrization in neumann problems. *Applicable Analysis*, 9(4):247–256, Dec. 1979.
- [34] J. I. T. Marrero. *Reaction-diffusion equations and applications to biological control of dengue and inflammation*. PhD thesis, 2021.
- [35] A. D. Masi, P. A. Ferrari, and J. L. Lebowitz. Reaction-diffusion equations for interacting particle systems. *Journal of Statistical Physics*, 44(3-4):589–644, Aug. 1986.
- [36] I. Mazari. Quantitative estimates for parabolic optimal control problems under l^∞ and l^1 constraints in the ball, 2021.
- [37] I. Mazari, G. Nadin, and Y. Privat. In preparation, 2021.
- [38] I. Mazari, G. Nadin, and Y. Privat. Optimal location of resources maximizing the total population size in logistic models. *arXiv: Analysis of PDEs*, 2019.
- [39] I. Mazari, G. Nadin, and Y. Privat. Shape optimization of a weighted two-phase dirichlet eigenvalue. Preprint, 2020.
- [40] I. Mazari, G. Nadin, and Y. Privat. Some challenging optimisation problems for logistic diffusive equations and numerical issues. *to appear in "Handbook of Numerical Analysis"*, 2020.
- [41] I. Mazari and D. Ruiz-Balet. A fragmentation phenomenon for a non-energetic optimal control problem: optimisation of the total population size in logistic diffusive models. *arXiv: Optimization and Control*, 2020.
- [42] J. Mossino and J. M. Rakotoson. Isoperimetric inequalities in parabolic equations. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, Ser. 4, 13(1):51–73, 1986.

- [43] J. D. Murray. *Mathematical Biology*. Springer Berlin Heidelberg, 1993.
- [44] G. Nadin and A. I. Toledo Marrero. On the maximization problem for solutions of reaction-diffusion equations with respect to their initial data. accepted for publication in *Journal of Mathematical Modelling of Natural Phenomena*, 2020.
- [45] K. Nagahara and E. Yanagida. Maximization of the total population in a reaction–diffusion model with logistic growth. *Calculus of Variations and Partial Differential Equations*, 57(3):80, Apr 2018.
- [46] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, NY, USA, second edition, 2006.
- [47] B. Perthame. *Parabolic Equations in Biology*. Springer International Publishing, 2015.
- [48] K. Prochazka and G. Vogl. Quantifying the driving factors for language shift in a bilingual region. *Proceedings of the National Academy of Sciences*, 114(17):4365–4369, Mar. 2017.
- [49] J.-M. Rakotoson. *Réarrangement Relatif*. Springer Berlin Heidelberg, 2008.
- [50] G. Talenti. Elliptic equations and rearrangements. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, Ser. 4, 3(4):697–718, 1976.
- [51] J. L. Vazquez. Symétrisation pour $u_t = \Delta\varphi(u)$ et applications. *C. R. Acad. Sci. Paris Sér. I Math.*, 295, 1982.