

## ON THE MAXIMIZATION PROBLEM FOR SOLUTIONS OF REACTION–DIFFUSION EQUATIONS WITH RESPECT TO THEIR INITIAL DATA

GRÉGOIRE NADIN<sup>1,\*</sup> AND ANA ISIS TOLEDO MARRERO<sup>2</sup>

**Abstract.** We consider in this paper the 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. This problem is motivated by biological conservation questions. We show the existence of a maximizer and derive optimality conditions through an adjoint problem. We have to face regularity issues since non-smooth initial data could give a better result than smooth ones. We then derive an algorithm enabling to approximate the maximizer and discuss some open problems.

**Mathematics Subject Classification.** 35B30, 35B45, 35B65, 35K15, 35K57, 35Q80, 35Q92, 35Q93, 92D25, 92D30.

Received July 5, 2019. Accepted July 27, 2020.

### 1. INTRODUCTION

#### 1.1. Statement of the problem and earlier works

We investigate in this paper 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

$$\mathcal{A}_m := \left\{ u_0 \in \mathcal{A} : \int_{\Omega} u_0 = m \right\} \quad \text{with} \quad \mathcal{A} := \left\{ u_0 \in L^1(\Omega), 0 \leq u_0 \leq 1 \right\} \quad (1.1)$$

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

$$\begin{cases} \partial_t u - \sigma \Delta u = f(u) & \text{in } (0, T) \times \Omega, \\ u(0, x) = u_0(x) & \text{in } \Omega, \\ \frac{\partial u}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega \end{cases} \quad (1.2)$$

Here,  $u$  represents the density of a population,  $\int_{\Omega} u(t, x) dx$  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.

---

*Keywords and phrases:* Reaction-diffusion - control - conservation biology - optimization.

<sup>1</sup> Sorbonne Université, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France.

<sup>2</sup> Université Paris Nord, Institut Galilée, UMR 7539, LAGA, 93430 Villetaneuse, France.

\* Corresponding author: [gregoire.nadin@upmc.fr](mailto:gregoire.nadin@upmc.fr)

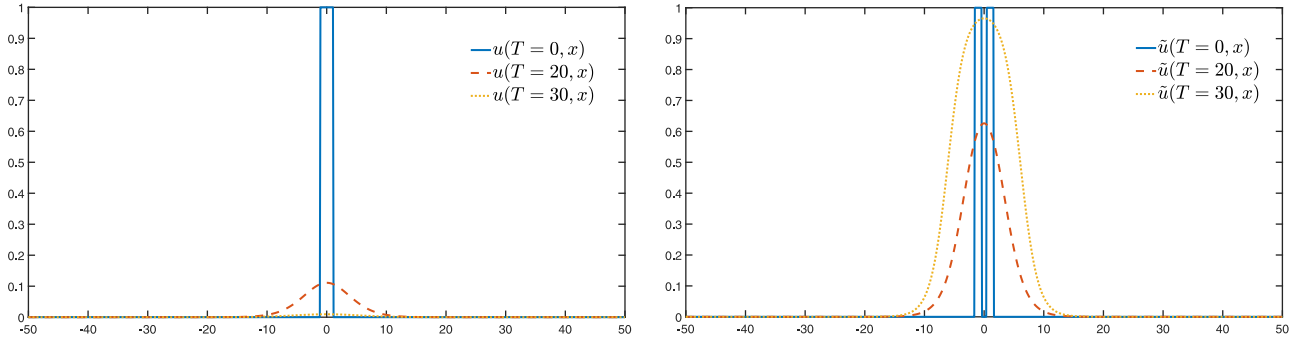


FIGURE 1. The graphs show the influence of the initial datum. On the left-hand side the initial population  $u(0, x)$  is concentrated in a single block of mass  $m = 2.22$  (solid blue line). Also is showed the solution  $u(T, x)$  of (1.2) after different times lapses:  $T = 20$  (dashed red line),  $T = 30$  (dotted yellow line); clearly the population tends to disappear as time goes by. On the right-hand side we consider an initial population with the same mass as before but distributed into two blocks slightly separated  $\tilde{u}(0, x)$ , the resulting density after same given time periods is clearly bigger, in this case the population tends to establish.

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 [7], the case of a particular initial datum  $u_0^\alpha = \mathbf{1}_{[-\frac{1}{2}-\frac{\alpha}{2}, -\frac{\alpha}{2}] \cup [\frac{\alpha}{2}, \frac{1}{2}+\frac{\alpha}{2}]}$ , 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  $\Omega = \mathbb{R}$ , for any given  $\alpha \geq 0$  it is known from [17] that there exist 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 [7] showing that one could get  $L^*(\alpha) < L^*(0)$  for  $\alpha$  small. This means that for a given initial total population  $L \in (L^*(\alpha), L^*(0))$ , the initial datum  $u_0^\alpha$  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. 1). This example shows that our present optimization problem could be difficult, since fragmented initial data could give a better total population at time  $T \gg 1$ . Hence, we expect regularity issues on a possible maximizer.

In [6] 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 Sect. 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 [15], the authors consider a bistable non-linearity  $f(t, u) = u(1-u)(u-\rho(t))$ , and the control is  $\rho$ , 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 from a target function -a travelling wave- considering a sufficiently large time.

Let us also mention [2], 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 paper 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.

## 1.2. Biological motivation

Dengue fever also known as breakbone fever and dandy fever is caused by dengue virus, which is ported 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 so the interest of finding a way to control it in absence of an effective medical treatment increases.

Manipulation of the arthropod population by introducing a maternally inherited bacterium called *Wolbachia* has been catching the attention of biologists in the last years [3, 10, 11, 16]. In the infected mosquito this bacterium prevents the development of the virus but also induces a cytoplasmic incompatibility which declines their reproduction rate when the infected population is small but became unimportant once its density becomes sufficiently large [4].

Reaction-diffusion equations have been widely used in order to describe biological phenomena of spreading and species competition, thanks to the works of [4] 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  $u : \mathbb{R}^+ \times \Omega \rightarrow [0, 1]$  of infected individuals with respect to the total population.

In such models, the reaction term  $f(u)$  is such that it reflects the positive correlation between population density and individual fitness, known as Allee effect. In the current problem, this effect is caused by the cytoplasmic incompatibility, so there exists a critical density threshold  $\rho$  under which the population of infected mosquitoes declines, but increases for large densities. 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.*  $u \rightarrow 0$  when  $t \rightarrow \infty$ , also called extinction), either the whole population get infected after a sufficiently large lapse of time (*i.e.*  $u \rightarrow 1$  when  $t \rightarrow \infty$ , 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 has been done in recent literature, from the biological point of view [8] and also from the mathematical one. In [15], that we already mentioned, it is proposed as a strategy to modify the Allee threshold  $\rho$  in order to reach an a priori given target trajectory; in practice this is possible by mean 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 [5], in fact it is proved that there exists a systematic way to choose a time  $T > 0$ , a bounded domain  $\Omega$  and a distributed control law  $g(u)$  supported in  $\Omega$ , such that for any initial value  $u_0$  the solution of the control problem

$$\begin{cases} \partial_t u - \sigma \Delta u = f(u) + g(u) \mathbb{1}_{[0, T] \times \Omega}, & ; \quad t > 0, x \in \mathbb{R}^d, \\ u(0, x) = u_0(x), & ; \quad x \in \mathbb{R}^d. \end{cases} \quad (1.3)$$

satisfies  $u(t, x) \rightarrow 1$  when  $t \rightarrow +\infty$ , 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 [1, 2].

## 2. PROBLEM FORMULATION AND MAIN RESULT

We will consider in this paper a bounded, smooth, connected domain  $\Omega$ , 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 problem (1.2) has a unique solution  $u(t, x)$  and it is such that  $0 \leq u(t, x) \leq 1$ , so we can define the operator  $\mathcal{J}_T : \mathcal{A} \subset L^1(\Omega) \rightarrow \mathbb{R}$  in the following way

$$\mathcal{J}_T(u_0) = \int_{\Omega} u(T, x) dx, \quad (2.1)$$

where  $u$  is the solution of equation (1.2). We can now formulate our main result,

**Theorem 2.1.** *Let  $\Omega$  be a bounded domain and let  $f$  satisfy the hypothesis (H1), (H2) and (H3). Then there exist  $\bar{u}_0 \in \mathcal{A}_m$  such that*

$$\max_{u_0 \in \mathcal{A}_m} \mathcal{J}_T(u_0) = \mathcal{J}_T(\bar{u}_0). \quad (2.2)$$

Moreover, setting  $\bar{u}$  the solution of (1.2) associated with this optimum initial data and  $p$  the unique solution of (2.3)

$$\begin{cases} -\partial_t p - \sigma \Delta p = f'(\bar{u})p & \text{in } (0, T) \times \Omega, \\ p(T, x) = 1 & \text{in } \Omega, \\ \frac{\partial p}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega, \end{cases} \quad (2.3)$$

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

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

The existence of such a maximizer  $\bar{u}_0(x)$  corresponds with the best possible way to distribute a fixed initial mass  $m$  in a bounded domain  $\Omega$  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 2.1 give us some useful information regarding the profile of an optimal initial data; in fact it implies that any optimum can be written as

$$\bar{u}_0(x) = \mathbb{1}_{\{p_0(x) > c\}} + \gamma(x) \mathbb{1}_{\{p_0(x) = c\}}, \quad (2.4)$$

with  $0 \leq \gamma(x) \leq 1$ . In particular if the adjoint state  $p_0(x)$  is not constant in any subset of  $\Omega$ , then the optimum is  $\bar{u}_0(x) = \mathbb{1}_{\{p_0(x) > c\}}$ . In the Section 5 we will see that this result allows us to define a numerical algorithm to approximate a local maximum of  $\mathcal{J}_T$ .

### 3. PROOF OF THEOREM 2.1

We first state some results concerning the regularity of  $u$ ,  $J$  and the adjoint state  $p$  that we will later invoke.

**Lemma 3.1.** *Under the hypothesis (H1), (H2) and (H3) stated above on  $f$ , the solution  $u = u(t, x)$  of (1.2) satisfies the following estimates:*

1.  $0 \leq u(t, x) \leq 1$ , for all  $(t, x) \in [0, T] \times \Omega$ ,
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 now the other two estimates let us multiply the equation (1.2) by  $u$  and integrate on  $\Omega$ , we obtain

$$\frac{1}{2} \partial_t \int_{\Omega} u^2(t, x) dx + \sigma \int_{\Omega} (\nabla_x u(t, x))^2 dx = \int_{\Omega} u(t, x) f(u(t, x)) dx; \quad (3.1)$$

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

$$\frac{1}{2} \partial_t \|u(t)\|_{L^2(\Omega)}^2 \leq \int_{\Omega} u(t, x) f(u(t, x)) dx \leq M \|u(t)\|_{L^2(\Omega)}^2. \quad (3.2)$$

By Gronwall inequality we get

$$\|u(t)\|_{L^2(\Omega)}^2 \leq e^{2Mt} \|u_0\|_{L^2(\Omega)}^2 \leq |\Omega| e^{2MT}. \quad (3.3)$$

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 (3.1) in time we get

$$\frac{1}{2} \int_{\Omega} (u)^2(T, x) dx - \frac{1}{2} \int_{\Omega} (u_0)^2(x) dx + \sigma \int_0^T \int_{\Omega} (\nabla_x u(t, x))^2 dx = \int_0^T \int_{\Omega} u(t, x) f(u(t, x)) dx \quad (3.4)$$

and so it yields

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

Finally, choosing  $v \in H^1(\Omega)$  such that  $\|v\|_{H^1(\Omega)} \leq 1$  and multiplying and integrating again in (1.2) it holds

$$\int_{\Omega} \partial_t u(t, x) v(x) dx + \sigma \int_{\Omega} \nabla_x u(t, x) \nabla_x v(x) dx = \int_{\Omega} f(u(t, x)) v(x) dx \quad (3.5)$$

from which we can deduce that

$$\|\partial_t u(t, \cdot)\|_{H^{-1}(\Omega)} \leq M \|u\|_{L^2(\Omega)}^2 + |\sigma| \|\nabla_x u\|_{L^2(\Omega)}^2 \quad (3.6)$$

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 3.2.** *The operator  $\mathcal{J}_T$  defined in (2.1) is differentiable. Furthermore, choosing  $p$  as the unique solution of (2.3) it holds*

$$\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_{\Omega} h_0(x) p(0, x) dx, \quad (3.7)$$

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

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

it holds

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \int_0^T \int_{\Omega} f''(u(t, x)) p(t, x) h^2(t, x) \, dx \, dt, \quad (3.9)$$

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

*Proof.* Let  $h_0(x)$  be defined over  $\Omega$  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  $h_{\varepsilon}$  such that the solution  $v_{\varepsilon}$  of (1.2) with initial condition  $v_{\varepsilon}(0, x) = u_0 + \varepsilon h_0$  can be written as  $v_{\varepsilon} = u + \varepsilon h_{\varepsilon}$ , where  $h_{\varepsilon}$  is the unique solution of (3.10)

$$\begin{cases} \partial_t h_{\varepsilon} - \sigma \Delta h_{\varepsilon} = \frac{1}{\varepsilon} (f(u + \varepsilon h_{\varepsilon}) - f(u)) & \text{in } (0, T) \times \Omega, \\ h_{\varepsilon}(0, x) = h_0(x) & \text{in } \Omega, \\ \frac{\partial h_{\varepsilon}}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega. \end{cases} \quad (3.10)$$

The Gateaux derivative of  $\mathcal{J}_T$  writes then

$$\begin{aligned} \langle \nabla \mathcal{J}_T(u_0), h_0 \rangle &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon \neq 0}} \frac{\mathcal{J}_T(v_{\varepsilon}(0, \cdot)) - \mathcal{J}_T(u_0)}{\varepsilon} = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon \neq 0}} \frac{\int_{\Omega} v_{\varepsilon}(T, x) dx - \int_{\Omega} u(T, x) dx}{\varepsilon} \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon \neq 0}} \frac{\int_{\Omega} (u(T, x) + \varepsilon h_{\varepsilon}(T, x)) dx - \int_{\Omega} u(T, x) dx}{\varepsilon} \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon \neq 0}} \int_{\Omega} h_{\varepsilon}(T, x) dx. \end{aligned} \quad (3.11)$$

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

$$\left| \frac{f(u + \varepsilon h_{\varepsilon}) - f(u)}{\varepsilon} \right| \leq M |h_{\varepsilon}|, \quad \forall \varepsilon > 0, \quad (3.12)$$

then the Gronwall inequality, applied just as in the proof of the first assertion in Lemma 3.1, implies that  $h_{\varepsilon} \in L^{\infty}(0, T; L^2(\Omega))$ . As a consequence, when  $\varepsilon \rightarrow 0$  it is possible to extract a subsequence  $\varepsilon_n \rightarrow 0$  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  $\mathcal{J}_T$  writes

$$\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_{\Omega} h(T, x) dx, \quad (3.13)$$

where  $h$  is also the unique solution of the differential equation (3.8) obtained by passing in to the limit in the weak formulation of (3.10).

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 (3.8) for  $u = w, v$  respective solutions of (1.2) with  $w_0, v_0$  as initial conditions, it is then easy to

check that

$$|\langle \nabla \mathcal{J}_T(w_0), h_0 \rangle - \langle \nabla \mathcal{J}_T(v_0), h_0 \rangle| = \left| \int_{\Omega} (h_w - h_v)(T, x) dx \right| \leq \sqrt{|\Omega|} \| (h_w - h_v)(T, \cdot) \|_{L^2(\Omega)}. \quad (3.14)$$

Multiplying the equation on  $h_w - h_v$  by  $h_w - h_v$  and integrating on  $\Omega$  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

$$\begin{aligned} \frac{1}{2} \partial_t \int_{\Omega} (h_w - h_v)^2(t, x) dx &\leq \int_{\Omega} (f'(\delta_{w,v} + v)h_w - f'(v)h_v)(h_w - h_v)(t, x) dx \\ &\leq \int_{\Omega} (C\delta_{w,v} + f'(v))h_w - f'(v)h_v)(h_w - h_v)(t, x) dx \\ &\leq C \int_{\Omega} h_w \delta_{w,v}(h_w - h_v)(t, x) dx + M \int_{\Omega} (h_w - h_v)^2(t, x) dx; \end{aligned} \quad (3.15)$$

the second and third inequalities in (3.15) follows 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

$$\partial_t \| (h_w - h_v) \|_{L^2(\Omega)}^2 \leq CL \|\delta_{w_0, v_0}\|_{L^2(\Omega)}^2 + M \| (h_w - h_v) \|_{L^2(\Omega)}^2, \quad (3.16)$$

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

$$\| (h_w - h_v)(t, x) \|_{L^2(\Omega)} \leq \frac{CL}{M} \|\delta_{w_0, v_0}\|_{L^2(\Omega)} e^{Mt}. \quad (3.17)$$

Combining (3.17) and (3.14) 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,

$$|\langle \nabla \mathcal{J}_T(w_0), h_0 \rangle - \langle \nabla \mathcal{J}_T(v_0), h_0 \rangle| \leq \sqrt{|\Omega|} \frac{CL}{M} \|\delta_{w_0, v_0}\|_{L^2(\Omega)} e^{MT}. \quad (3.18)$$

Multiplying (3.8) by the solution  $p$  of (2.3) with  $\bar{u} = u$  and integrating by parts we can rewrite (3.13) as

$$\langle \nabla \mathcal{J}_T(u_0), h_0 \rangle = \int_{\Omega} h(T, x) dx = \int_{\Omega} h_0(x) p(0, x) dx \quad (3.19)$$

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  $v_{\varepsilon}$ .

From the regularity hypothesis on  $f$  and the estimation (3.17) on  $h$ ,  $v_{\varepsilon}$  and  $u_{\varepsilon}$  it follows that  $k_{\varepsilon} \in L^{\infty}(0, T; L^2(\Omega))$ . A passage to the limit when  $\varepsilon \rightarrow 0$  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  $k$ ,

$$\begin{cases} \partial_t k - \sigma \Delta k = f'(u)k + f''(u)h & \text{in } (0, T) \times \Omega, \\ k(0, x) = 0 & \text{in } \Omega, \\ \frac{\partial k}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega. \end{cases} \quad (3.20)$$

An analysis similar to that yielding (3.19) shows that

$$\int_{\Omega} k(T, x) dx = \int_0^T \int_{\Omega} f''(u) h^2 p \, dx \, dt, \quad (3.21)$$

in this case we multiply (3.20) by  $p$  and we integrate in space and time. Finally, we have proved that when  $\varepsilon \rightarrow 0$ , it holds

$$\begin{aligned} \mathcal{J}_T(u_0 + \varepsilon h_0) &= \int_{\Omega} u(T, x) dx + \varepsilon \int_{\Omega} h(T, x) dx + \frac{\varepsilon^2}{2} \int_{\Omega} k(T, x) dx + o(\varepsilon^2) \\ &= \mathcal{J}_T(u_0) + \varepsilon \langle \nabla \mathcal{J}_T(u_0), h \rangle + \frac{\varepsilon^2}{2} \int_0^T \int_{\Omega} f''(u) h^2 p \, dx \, dt + o(\varepsilon^2), \end{aligned}$$

which establishes the formula

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \int_0^T \int_{\Omega} f''(u(t, x)) h^2(t, x) p(t, x) \, dx \, dt. \quad (3.22)$$

□

**Lemma 3.3.** *Let  $\Omega$  be a bounded domain, then the solution  $p$  of the equation (2.3) is such that*

1.  $p_t, p_{tt}, \nabla p, \nabla^2 p, \nabla p_t, \nabla^2 p_t \in L_{loc}^q((0, T) \times \Omega)$  for all  $1 \leq q < \infty$ ,
2.  $\forall c \in \mathbb{R}$ , for almost every  $x \in \{p(0, \cdot) = c\}$ , one has

$$f'(\bar{u}_0(x)) = -p_t(0, x)/p(0, x).$$

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

$$\begin{cases} -\partial_t \bar{p} = M \bar{p} & \text{in } (0, T), \\ \bar{p}(T) = 1, \end{cases} \quad (3.23)$$

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

$$p \in L^\infty(0, T; L^\infty(\Omega)) \quad (3.24)$$

Second, we know from the classical  $L^q$  regularity theory for parabolic equations (see for instance Theorem 9.1, in chapter IV of [13]) that, as  $0 \leq u \leq 1$ , one has  $p_t, \nabla p, \nabla^2 p \in L_{loc}^q((0, T) \times \Omega)$  for all  $1 \leq q < \infty$ . Similarly, one has  $u_t, \nabla u, \nabla^2 u \in L_{loc}^q((0, T) \times \Omega)$  for all  $1 \leq q < \infty$ .

Next, let us define  $\phi := p_t$ , deriving on (2.3) we obtain that it is the only solution of the equation

$$\begin{cases} -\partial_t \phi - \Delta \phi = G, & \text{in } (0, T) \times \Omega, \\ \phi(T, x) = 0, & \text{in } \Omega, \\ \frac{\partial \phi}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial \Omega, \end{cases} \quad (3.25)$$

where  $G = f''(u) u_t p + f'(u) \phi$ . Due to the previous estimates, one has  $G \in L_{loc}^q((0, T) \times \Omega)$  for all  $1 \leq q < \infty$ . Hence, again the  $L^q$  regularity theory for parabolic equations yields  $\phi_t, \nabla \phi, \nabla^2 \phi \in L_{loc}^q((0, T) \times \Omega)$  for all



$1 \leq q < \infty$ . This means in particular that  $p_t \in L_{loc}^q((0, T), W_{loc}^{2,q}(\Omega))$  for all  $1 \leq q < \infty$ . Taking  $q$  large enough, the Morrey inequality thus yields  $p \in C_{loc}^{0,\alpha}((0, T), W_{loc}^{2,q}(\Omega))$  for all  $1 \leq q < \infty$  and  $\alpha \in (0, 1)$ .

Now, as  $p(0, \cdot) \in W_{loc}^{2,1}(\Omega)$ , we know (see for example [9]) that for almost every  $x \in \{p(0, \cdot) = c\}$ , one has  $\Delta p(0, x) = 0$ . Moreover, as  $p_{tt} \in L_{loc}^q((0, T) \times \Omega)$  for all  $1 \leq q < \infty$ , one has  $p_t \in C_{loc}^{0,\alpha}((0, T), L_{loc}^q(\Omega))$  and, in particular,  $p_t(0, \cdot) \in L_{loc}^q(\Omega)$ . We eventually derive from (2.3) that for almost every  $x \in \{p(0, \cdot) = c\}$ , one has

$$-p_t(0, x) = f'(\bar{u}(0, x))p(0, x).$$

□

Now we proceed with proof of Theorem 2.1.

*Proof.* This proof falls naturally into two parts, firstly 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  $\bar{u}_0 \in \mathcal{A}_m$  defined as the limit of a maximizing sequence in  $\mathcal{A}_m$ .

From the first estimate on Lemma 3.1 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 3.2,

$$|\mathcal{J}_T(u_0) - \mathcal{J}_T(v_0)| \leq |\Omega|^{\frac{1}{2}} e^{MT} \|u_0 - v_0\|_{L^2(\Omega)}. \quad (3.26)$$

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

$$\lim_n \mathcal{J}_T(u_0^n) = \sup_{\{\mathcal{A}_m\}} \mathcal{J}_T(u_0). \quad (3.27)$$

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

$$\int_{\Omega} u_0^n(x) \varphi(x) dx \rightarrow \int_{\Omega} \bar{u}_0(x) \varphi(x) dx, \quad \forall \varphi \in L^1(\Omega). \quad (3.28)$$

Choosing  $\varphi = 1$  in (3.28), we get that  $\bar{u}_0$  is still in  $\mathcal{A}_m$ .

Now, to each  $u_0^n, n = 1, 2, \dots$  we can associate the solution of the problem (1.2) with initial datum  $u_0^n$

$$\begin{cases} \partial_t u^n - \sigma \Delta u^n = f(u^n) & \text{in } (0, T) \times \Omega, \\ u^n(0, x) = u_0^n(x) & \text{in } \Omega, \\ \frac{\partial u^n}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega, \end{cases} \quad (3.29)$$

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

$$\langle u^n(T), \varphi(T) \rangle - \int_0^T \langle u^n(t), \partial_t \varphi(t) \rangle dt - \sigma \int_0^T \langle u^n(t), \Delta \varphi(t) \rangle dt = \int_0^T \langle f(u^n(t)), \varphi(t) \rangle dt + \langle u_0^n, \varphi(0) \rangle. \quad (3.30)$$

Thanks to the first assertion of Lemma 3.1, 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  $\tilde{u} \in L^2(\Omega)$  such that, after extraction,  $u^n(T, x) \rightharpoonup \tilde{u}$  in  $L^2(\Omega)$ , *i.e.*

$$\langle u^n(T), \varphi \rangle \rightarrow \langle \tilde{u}, \varphi \rangle, \quad \forall \varphi \in L^2(\Omega). \quad (3.31)$$

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

$$\int_0^T \langle u^n(t), \varphi(t) \rangle dt \longrightarrow \int_0^T \langle \bar{U}(t), \varphi(t) \rangle dt, \quad \forall \varphi \in L^\infty(0, T; L^2(\Omega)). \quad (3.32)$$

Again, from the second assertion in Lemma 3.1, 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  $\partial_t u^n \rightharpoonup^* v$  in  $L^2(0, T; H^{-1}(\Omega))$ , *i.e.*

$$\int_0^T \langle \partial_t u^n(t), \varphi(t) \rangle dt \rightarrow \int_0^T \langle v(t), \varphi(t) \rangle dt \quad \forall \varphi \in L^2(0, T; H^1(\Omega)). \quad (3.33)$$

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

$$\int_0^T \langle \partial_t u^n(t), \varphi(t) \rangle dt = - \int_0^T \langle u^n(t), \partial_t \varphi(t) \rangle dt; \quad (3.34)$$

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

$$\int_0^T \langle v(t), \varphi(t) \rangle dt = - \int_0^T \langle \bar{U}(t), \partial_t \varphi(t) \rangle dt. \quad (3.35)$$

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

$$\int_0^T \langle \partial_t u^n(t), \varphi(t) \rangle dt = \langle u^n(T), \varphi(T) \rangle - \int_0^T \langle u^n(t), \partial_t \varphi(t) \rangle dt \quad (3.36)$$

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

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

This equality together with (3.35) implies that  $\tilde{u}(x) = \bar{U}(T, x)$  almost everywhere in  $\Omega$ . Similarly choosing  $\varphi$  adequately we prove that  $\bar{u}_0(x) = \bar{U}(0, x)$  almost everywhere in  $\Omega$ .

Finally, let us define the set

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

The estimates in Lemma 3.1 implies that  $u^n$  is bounded in  $W$  for every element of the subsequence. Moreover, thanks to the Aubin-Lions lemma [14], the set  $W$  embeds compactly into  $L^2([0, T]; L^2(\Omega))$  which ensures the existence of a subsequence still noted as  $u^n$  which is Cauchy in  $L^2([0, T]; L^2(\Omega))$ . Then necessarily  $u^n \rightarrow \bar{U}$  strongly in  $L^2([0, T]; L^2(\Omega))$  and thus

$$\int_0^T \langle f(u^n(t)), \varphi \rangle dt \rightarrow \int_0^T \langle f(\bar{U}(t)), \varphi \rangle dt, \quad (3.37)$$

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

Now we can pass to the limit in (3.30), gathering (3.28), (3.31) and (3.37) to obtain that  $\forall \varphi \in C^\infty(0, T) \times \Omega$  it holds

$$\langle \bar{U}(T), \varphi(T) \rangle - \int_0^T \langle \bar{U}(t), \partial_t \varphi(t) \rangle dt - \int_0^T \langle \bar{U}(t), \Delta \varphi(t) \rangle = \int_0^T \langle f(\bar{U}(t)), \varphi(t) \rangle dt + \langle \bar{u}_0, \varphi(0) \rangle, \quad (3.38)$$

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

$$\begin{cases} \partial_t \bar{U} - \sigma \Delta \bar{U} = f(\bar{U}) & \text{in } (0, T) \times \Omega, \\ \bar{U}(0, x) = \bar{u}_0(x) & \text{in } \Omega, \\ \frac{\partial \bar{U}}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega. \end{cases} \quad (3.39)$$

Now, we get the following equalities from (3.27) and (3.31) choosing  $\varphi = 1$

$$\sup_{\{\mathcal{A}_m\}} \mathcal{J}_T(u) = \lim_n \mathcal{J}_T(u_0^n) = \lim_n \int_\Omega u^n(T, x) dx = \int_\Omega \bar{U}(T, x) dx = \mathcal{J}_T(\bar{u}_0), \quad (3.40)$$

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

**Step 2:** Characterization of the maximal element

We first prove (i).

Let  $\mu$  be the Lebesgue measure. We define the set  $\mathcal{S} = \{x \in \Omega : 0 < \bar{u}_0(x) < 1\}$  and we suppose that  $\mu(\mathcal{S}) \neq 0$ , otherwise there exists a set  $E$  such that  $\bar{u}_0 = \mathbf{1}_E$  almost everywhere. We note that  $\mathcal{S}$  can be written as  $\mathcal{S} = \cup_{k=1}^\infty \mathcal{S}_k$  where  $\mathcal{S}_k = \{x \in \Omega : \frac{1}{k} < \bar{u}_0(x) < 1 - \frac{1}{k}\}$ . 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

$$v_0(x) = \bar{u}_0(x) + \varepsilon \frac{\mu(B(x^*, r))}{\mu(B(x^*, r) \cap \mathcal{S}_k)} \mathbf{1}_{B(x^*, r) \cap \mathcal{S}_k} - \varepsilon \frac{\mu(B(y^*, r))}{\mu(B(y^*, r) \cap \mathcal{S}_k)} \mathbf{1}_{B(y^*, r) \cap \mathcal{S}_k}, \quad (3.41)$$

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

$$\lim_{r \rightarrow 0^+} \frac{\mu(B(x, r))}{\mu(B(x, r) \cap \mathcal{S}_k)} \rightarrow 1. \quad (3.42)$$

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

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

so  $v_0 \in \mathcal{A}_m$ . We shall now use the fact that  $\bar{u}_0$  is a maximizing element in  $\mathcal{A}_m$ ; gathering (3.13) and (3.19) we have

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

$$\begin{aligned}
&= \int_{\Omega} \left( \frac{\mu(B(x^*, r)) \mathbf{I}_{B(x^*, r) \cap \mathcal{S}_k}}{\mu(B(x^*, r) \cap \mathcal{S}_k)} - \frac{\mu(B(y^*, r)) \mathbf{I}_{B(y^*, r) \cap \mathcal{S}_k}}{\mu(B(y^*, r) \cap \mathcal{S}_k)} \right) p(0, x) dx \\
&= \frac{\mu(B(x^*, r))}{\mu(B(x^*, r) \cap \mathcal{S}_k)} \int_{B(x^*, r) \cap \mathcal{S}_k} p(0, x) dx - \frac{\mu(B(y^*, r))}{\mu(B(y^*, r) \cap \mathcal{S}_k)} \int_{B(y^*, r) \cap \mathcal{S}_k} p(0, x) dx.
\end{aligned}$$

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  $x^*, y^* \in \mathcal{S}_k$  it holds

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

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  $k \rightarrow +\infty$  we have the result for almost every  $x \in \mathcal{S}$ .

Let us now prove (ii).

Lets define the set  $\mathcal{S}^0 = \{x \in \Omega : \bar{u}_0(x) = 0\}$  and for every  $k = 1, 2, \dots$  the set  $\mathcal{S}_k^0 = \{x \in \Omega : 0 \leq \bar{u}_0(x) < \frac{1}{k}\}$ , we note that  $\mathcal{S}^0 = \bigcap_{k=1}^{\infty} \mathcal{S}_k^0$ . We assume that  $\mu(\mathcal{S}^0) > 0$ , otherwise  $\bar{u}_0 > 0$  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 (3.42) and  $0 < \varepsilon < \frac{1}{2k}$ , it holds

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

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

$$\begin{aligned}
0 &\geq \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left[ \mathcal{J}_T \left( \bar{u}_0 + \varepsilon \left( \frac{\mu(B(x^*, r)) \mathbf{I}_{B(x^*, r) \cap \mathcal{S}_k^0}}{\mu(B(x^*, r) \cap \mathcal{S}_k^0)} - \frac{\mu(B(y^*, r)) \mathbf{I}_{B(y^*, r) \cap \mathcal{S}_k}}{\mu(B(y^*, r) \cap \mathcal{S}_k)} \right) \right) - \mathcal{J}_T(\bar{u}_0) \right] \\
&= \int_{\Omega} \left( \frac{\mu(B(x^*, r)) \mathbf{I}_{B(x^*, r) \cap \mathcal{S}_k^0}}{\mu(B(x^*, r) \cap \mathcal{S}_k^0)} - \frac{\mu(B(y^*, r)) \mathbf{I}_{B(y^*, r) \cap \mathcal{S}_k}}{\mu(B(y^*, r) \cap \mathcal{S}_k)} \right) p(0, x) dx \\
&= \frac{\mu(B(x^*, r))}{\mu(B(x^*, r) \cap \mathcal{S}_k^0)} \int_{B(x^*, r) \cap \mathcal{S}_k^0} p(0, x) dx - \frac{\mu(B(y^*, r))}{\mu(B(y^*, r) \cap \mathcal{S}_k)} \int_{B(y^*, r) \cap \mathcal{S}_k} p(0, x) dx
\end{aligned}$$

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  $r \rightarrow 0$  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  $k \rightarrow +\infty$ .

Similarly, we can prove (iii). Remark that  $\underline{p} \equiv 0$  is a sub-solution of equation (2.3) and thus necessarily  $p \geq 0$ . Since  $c$  is in the range of  $p$ , then it must be non-negative as well. This way we end with the proof of the Theorem 2.1.  $\square$

#### 4. THE $u_0$ -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  $\Omega$ . 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 the equation (1.2) 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

$$\begin{cases} \partial_t u = f(u) & \text{in } (0, T), \\ u(0) = \frac{m}{|\Omega|}. \end{cases} \quad (4.1)$$

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 4.1.** *Let  $u_0$  be the constant distribution defined as  $u_0 := \frac{m}{|\Omega|}$ ,  $\forall x \in \Omega$  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

$$\begin{cases} -\partial_t p = f'(u)p & \text{in } (0, T) \times \Omega, \\ p(T, x) = 1 & \text{in } \Omega. \end{cases} \quad (4.2)$$

It is easy to check that  $p(t, x) = f(u(T))/f(u(t))$ , which is also homogeneous in space. Consequently,  $p_0 = f(u(T))/f\left(\frac{m}{|\Omega|}\right)$  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

$$u_0 := \frac{m}{|\Omega|} > \rho, \quad (4.3)$$

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 (4.2) 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

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle = \int_0^T \int_{\Omega} f''(u(t)) p(t) h^2(t, x) dx dt \leq -\delta \int_0^T e^{-M(T-t)} \int_{\Omega} h^2(t, x) dx dt. \quad (4.4)$$

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

$$\begin{cases} \partial_t h - \sigma \Delta h = f'(u)h & \text{in } (0, T) \times \Omega, \\ h(0, x) = h_0(x) & \text{in } \Omega, \end{cases} \quad (4.5)$$

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  $\delta$ ,  $M$  and  $T$ , it holds that

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0 \rangle \leq -C \|h_0\|_{L^2(\Omega)}^2, \quad (4.6)$$

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 2.1 is necessarily  $c \equiv p_0$  otherwise  $u_0(x)$  is either null or totally saturated over the domain  $\Omega$  which would imply that  $u_0 \notin \mathcal{A}_m$ . Hence, the set  $\{p_0 = c\}$  coincides with the whole domain  $\Omega$ .

Let us now show the second part of the Proposition 4.1. 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.*

$$\langle \nabla^2 \mathcal{J}_T(u_0), h^k(x) \rangle \leq 0. \quad (4.7)$$

In particular, we consider  $h_0^k(x) = \cos(kx)$ ,  $k = 1, 2, \dots$ , 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

$$\begin{cases} \partial_t h^k - \Delta h^k = f'(u)h^k & \text{in } (0, T) \times (0, \pi), \\ h^k(0, x) = h_0^k(x) = \cos(kx) & \text{in } \Omega, \\ \frac{\partial h^k}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \{0, \pi\}. \end{cases} \quad (4.8)$$

The solution of (4.8) is explicitly given by

$$h^k(t, x) = e^{-k^2 t} \cos(kx) \frac{f(u(t))}{f(u_0)} \quad (4.9)$$

and consequently from (3.9) and thanks to the Laplace method it follows that

$$\langle \nabla^2 \mathcal{J}_T(u_0), h_0^k \rangle = \frac{f(u(T))}{f^2(u_0)} \int_0^T f''(u(t)) f(u(t)) e^{-2k^2 t} \int_0^\pi \cos^2(kx) dx dt \quad (4.10)$$

$$= \frac{\pi}{2} \frac{f(u(T))}{f^2(u_0)} \int_0^T f''(u(t)) f(u(t)) e^{-2k^2 t} dt \quad (4.11)$$

$$\underset{k \rightarrow \infty}{\sim} \frac{\pi}{4} p_0 \frac{f''(u_0)}{k^2}. \quad (4.12)$$

Gathering (4.12) and (4.7) we get that necessarily  $f''(u_0) \leq 0$ , which completes the prove.  $\square$

#### 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 4.1 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 4.2.** *Let  $u$  be the solution of the differential equation (1.2). If the reaction term  $f(u)$  is concave, then the functional  $\mathcal{J}_T(u_0)$  defined by (2.1) is also concave.*

*Proof.* Let  $\alpha, \beta \in [0, 1]$  be such that  $\alpha + \beta = 1$ , if we call  $u$  the solution of (1.2) with initial condition  $\tilde{u}_0(x) = \alpha u_0^1(x) + \beta u_0^2(x)$  and we set  $\hat{u}(t, x) = \alpha u^1(t, x) + \beta u^2(t, x)$  where  $u^i$  is the solution of (1.2) with initial condition  $u_0^i(x)$ ,  $i \in \{1, 2\}$ , because of the concavity of  $f$  we have

$$\partial_t \hat{u} - \Delta \hat{u} = \alpha f(u^1) + \beta f(u^2) \leq f(\alpha u^1 + \beta u^2) = f(\hat{u}), \quad (4.13)$$

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

$$\mathcal{J}_T(\alpha u_0^1 + \beta u_0^2) = \mathcal{J}_T(\tilde{u}_0) = \int_{\Omega} u(T, x) dx \geq \int_{\Omega} \hat{u}(T, x) dx = \alpha \mathcal{J}_T(u_0^1) + \beta \mathcal{J}_T(u_0^2), \quad (4.14)$$

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

This concavity property in the Fisher-KPP case ensures that if  $\bar{u}_0$  is a critical point then it is a maximizer for  $\mathcal{J}_T$ . As straightforward consequence of the Proposition 4.1 we have that  $\bar{u}_0(x) \equiv \frac{m}{|\Omega|}$  is a global maximum for  $\mathcal{J}_T$ . Explicitly, the solution writes

$$u(t, x) = \frac{\kappa_0 e^{rt}}{1 + \kappa_0 e^{rt}}, \quad \kappa_0 := \frac{m}{|\Omega| - m}, \quad p(t, x) = \frac{e^{r(T-t)}}{e^{2r \int_t^T u(s) ds}}. \quad (4.15)$$

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

## 5. NUMERICAL ALGORITHM

The aim of this section is to describe an algorithm to find approximately an optimal distribution provided that the space  $\Omega$ , the mass  $m$  and the time  $T$  are prescribed. In order to achieve this goal, the first order optimality conditions (i)–(iii) on Theorem 2.1 will be crucial. The strategy, which is basically inspired by gradient descent optimization algorithms, will be to find a maximizing sequence  $u_0^1, u_0^2, u_0^3, \dots$  which converges to the optimal element  $\bar{u}_0(x)$ .

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

$$\max_{u_0 \in \mathcal{A}} \mathcal{J}_T(u_0) \quad (5.1)$$

$$s.t. \quad \int_{\Omega} u_0(x) dx = m. \quad (5.2)$$

We can then consider the associated problem

$$\min_{\lambda \in \mathbb{R}_+} \max_{u_0 \in \mathcal{A}} \mathcal{L}(u_0, \lambda) = \mathcal{J}_T(u_0) - \lambda \left( \int_{\Omega} u_0(x) dx - m \right), \quad (5.3)$$

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 must satisfy

$$\langle \partial_{u_0} \mathcal{L}(\bar{u}_0, \bar{\lambda}), h_0 \rangle = \int_{\Omega} (\bar{p}_0(x) - \bar{\lambda}) h_0(x) dx \geq 0, \quad \forall h_0 \in L^2(\Omega) : \bar{u}_0 + h_0 \in \mathcal{A}, \quad (5.4)$$

$$\partial_{\lambda} \mathcal{L}(\bar{u}_0, \bar{\lambda}) = \int_{\Omega} \bar{u}_0(x) dx - m = 0, \quad (5.5)$$

where  $\bar{p}_0(x) = \bar{p}(0, x)$  is the solution of (2.3) whose reaction term depends on  $\bar{u}(t, x)$ , which solves (1.2) with  $\bar{u}_0(x)$  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  $\Omega$ . 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 (1.2) 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

$$\begin{cases} -\partial_t p^n - \sigma \Delta p^n = f'(u^n) p^n & \text{in } (0, T) \times \Omega, \\ p^n(T, x) = 1 & \text{in } \Omega, \\ \frac{\partial p^n}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, T), \text{ for all } x \in \partial\Omega. \end{cases} \quad (5.6)$$

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

$$\langle \partial_{u_0} \mathcal{L}(u_0^n, \lambda), h_0^n \rangle = \int_{\Omega} (p_0^n(x) - \lambda) h_0^n(x) dx \geq 0. \quad (5.7)$$

For some value  $\lambda^n$ , 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 the equation (5.6) in the set  $\Omega_{p, \lambda}$  around  $t = 0$ . We use a semi-implicit finite difference scheme with time step  $dt$

$$-\left( \frac{p^n(dt, x) - \lambda^n}{dt} \right) = f'(u_0^{n+1}(x)) \lambda^n, \quad (5.8)$$

then  $u_0^{n+1}(x)$  for  $x \in \Omega_{p, \lambda}$  can be chosen as a solution of (5.8). Depending on the form of the reaction term  $f$ , the equation (5.8) might have several roots which mean that this definition is not well posed. Numerically this issue can be overcome by considering all the possibilities and keeping the one which have a better image by  $\mathcal{J}_T$ . 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 two possible values for  $u_0^{n+1}$  at most. 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 the equation (5.8) for all the nodes in  $\Omega_{p, \lambda}$ , either we assign to all of them the value of the greatest root.

The issue of computing the value assigned to  $\lambda^n$  in each iteration is tackled numerically by 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  $\lambda^n$  such that  $\int_{\Omega} u_0^{n+1}(x) dx \leq m$ . Note that the dependence of  $u_0^{n+1}(x)$  on  $\lambda^n$  is implicitly given by its definition; *i.e*

$$u_0^{n+1}(x) = 1, \quad \forall x : p_0^n(x) > \lambda^n, \quad (5.9)$$

$$u_0^{n+1}(x) = 0, \quad \forall x : p_0^n(x) < \lambda^n, \quad (5.10)$$

$$u_0^{n+1}(x) \text{ solution of (5.8)}, \quad \forall x : p_0^n(x) = \lambda^n. \quad (5.11)$$

Remark that defining  $u_0^{n+1}(x)$  this way, is compatible with the characterization given at Theorem 2.1 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 (5.9–5.11):

$$u_0^{n+\frac{1}{2}}(x) = 1, \quad \forall x : p_0^n(x) > \lambda^n, \quad (5.12)$$

$$u_0^{n+\frac{1}{2}}(x) = 0, \quad \forall x : p_0^n(x) < \lambda^n, \quad (5.13)$$

$$u_0^{n+\frac{1}{2}}(x) \text{ solution of (5.8),} \quad \forall x : p_0^n(x) = \lambda^n; \quad (5.14)$$

$\lambda^n$  being 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.*

$$u_0^{n+1} := (1 - \theta^n)u_0^n + \theta^n u_0^{n+\frac{1}{2}}, \quad (5.15)$$

$$\theta^n := \arg \max_{\theta \in [0,1]} \mathcal{J}_T \left( (1 - \theta)u_0^n + \theta u_0^{n+\frac{1}{2}} \right). \quad (5.16)$$

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 2.1, 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.*

$$\forall n > K, \quad u_0^n(x) = u_0^{n+1}(x), \forall x \in \Omega, \quad (5.17)$$

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  $u_0^{n+1}$  through the convex combination of  $u_0^n$  and  $u_0^{n+\frac{1}{2}}$  and as a consequence of (5.17) it holds that

$$\theta^n \left( u_0^{n+\frac{1}{2}} - u_0^n \right) = 0, \quad \text{for every } n \geq K. \quad (5.18)$$

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 2.1 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

$$\mathcal{J}_T((1 - \mu)u_0^n + \mu u_0^{n+\frac{1}{2}}) \leq \mathcal{J}_T((1 - \theta^n)u_0^n + \theta^n u_0^{n+\frac{1}{2}}); \quad (5.19)$$

using a Taylor expansion in both sides we get

$$(\mu - \theta^n) \langle \nabla \mathcal{J}_T(u_0^n), u_0^{n+\frac{1}{2}} - u_0^n \rangle \leq 0 \quad (5.20)$$

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 3.2

$$\begin{aligned}
\langle \nabla \mathcal{J}_T(u_0^n), u_0^{n+\frac{1}{2}} - u_0^n \rangle &= \int_{\Omega} p_0^n (u_0^{n+\frac{1}{2}} - u_0^n) dx \\
&= \int_{\Omega} (p_0^n - \lambda^n) (u_0^{n+\frac{1}{2}} - u_0^n) dx \\
&= \int_{\{p_0^n > \lambda^n\}} (p_0^n - \lambda^n) (u_0^{n+\frac{1}{2}} - u_0^n) dx + \int_{\{p_0^n < \lambda^n\}} (p_0^n - \lambda^n) (u_0^{n+\frac{1}{2}} - u_0^n) dx \quad (5.21) \\
&= \int_{\{p_0^n > \lambda^n\}} (p_0^n - \lambda^n) (1 - u_0^n) dx + \int_{\{p_0^n < \lambda^n\}} (p_0^n - \lambda^n) (0 - u_0^n) dx \\
&\geq 0.
\end{aligned}$$

Together (5.21) and (5.20) imply that  $\langle \nabla \mathcal{J}_T(u_0^n), u_0^{n+\frac{1}{2}} - u_0^n \rangle = 0$  or equivalently

$$\int_{\{p_0^n > \lambda^n\}} (p_0^n - \lambda^n) (1 - u_0^n) dx + \int_{\{p_0^n < \lambda^n\}} (p_0^n - \lambda^n) (0 - u_0^n) dx = 0 \quad (5.22)$$

from which we deduce (i)-(iii).  $\square$

This mechanism not only improves the convergence but also makes it easy to identify; in fact if at the  $n$ th-iteration the best  $\theta$  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 have 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. 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 Figure 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  $\bar{u}_0$  defined over  $\Omega$  is associated with a symmetric distribution  $\bar{u}_{0_s}$  defined over  $\Omega_s = (-a, a)$ . Reciprocally, any maximizer  $\bar{v}_0$  in the class of symmetric initial data on  $\Omega_s = (-a, a)$  induces a solution satisfying a Neumann boundary condition at  $x = 0$ . Hence,  $\bar{v}_0$  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  $(-a, a)$ .

### 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.

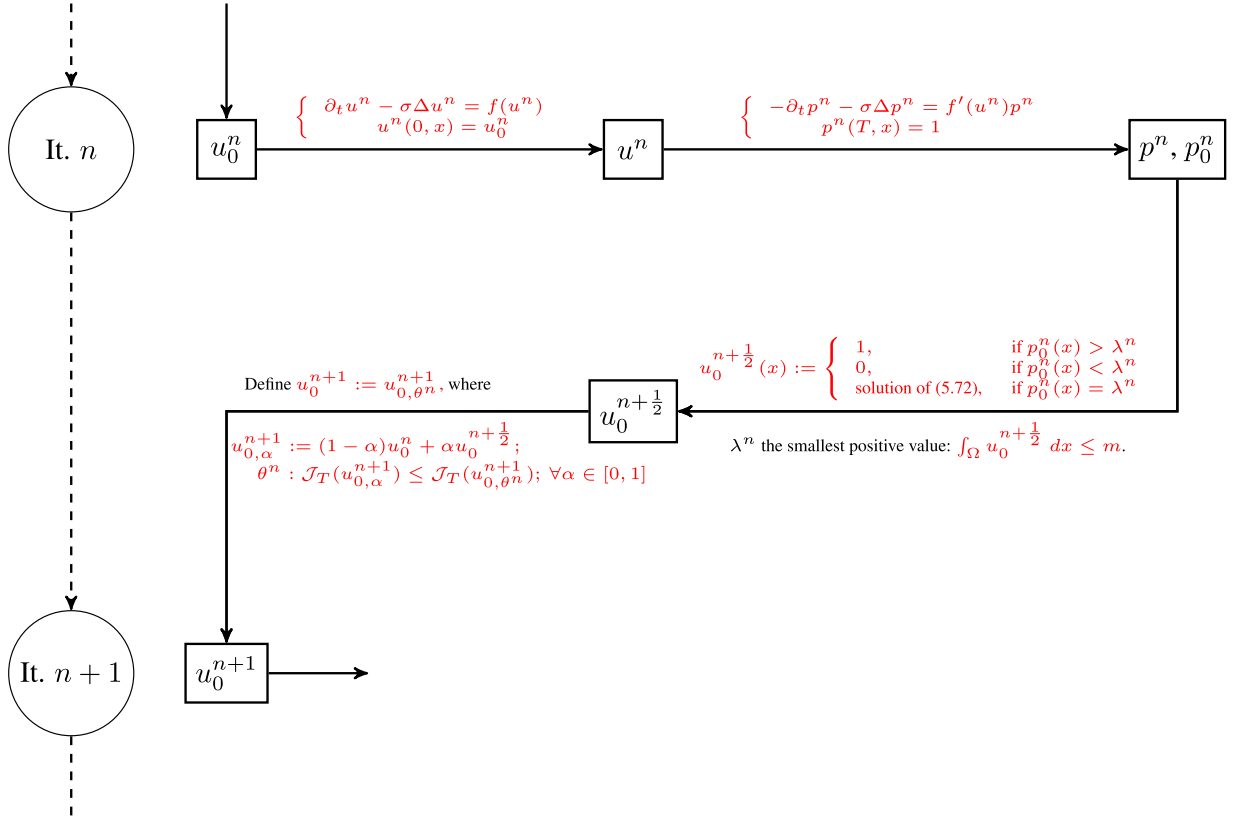


FIGURE 2. Scheme of the numerical algorithm.

The simulations show that the algorithm described above converges after a few iterations and increase successfully the values of  $\mathcal{J}_T(u_0)$  in comparison with the trivial single block distribution (Fig. 3a); we can also observe singularities which are associated with the values verifying  $p(0, x) = \lambda$ ; this behavior will be discussed later on Section 6 (Fig. 4).

## 6. DISCUSSION

### 6.1. Possible generalizations

We have considered in this paper 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

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

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

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

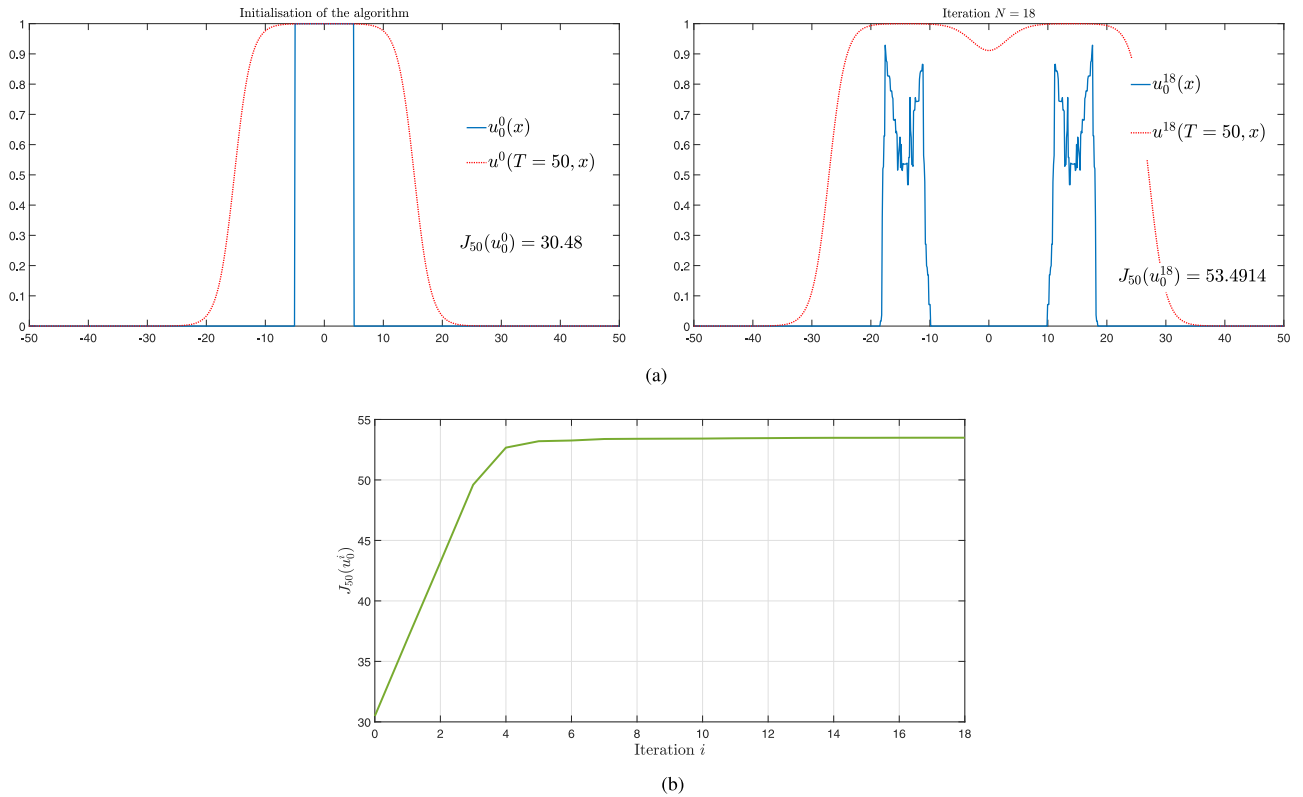


FIGURE 3. Considering a fixed mass  $m = 10$ , this figure shows the initial data associated with the first and the last iteration of the algorithm and the corresponding solutions of equation (1.2) (a). We also show the evolution line of the operator  $\mathcal{J}_{50}(u_0^i)$  from the first iteration 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  $\Omega_{p, \lambda^{18}}$ .

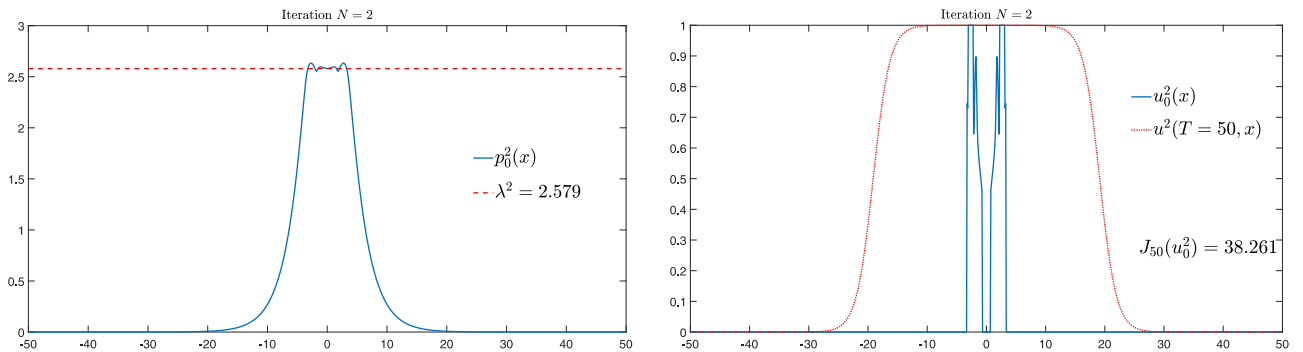


FIGURE 4. The figure corresponds to the 2nd and also last iteration of the algorithm for a given final time  $T = 50$  and a fixed mass  $L = 4$ . At the left-hand side we show the adjoint state  $p$  which is the solution of the equation (5.6) and its associated value  $\lambda^2$  mentioned in Theorem 2.1. Note that in this case the set  $\Omega_{p, \lambda^2}$  is not negligible so the associated  $u_0^2$  showed on the right side present singularities arising from the solution of (5.8) within this set.

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.

## 6.2. Letting $T \rightarrow +\infty$

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,  $u_0^\infty$  of  $u_0^T$  as  $T \rightarrow +\infty$ , for the  $L^\infty$  weak star convergence. Let  $u$  be the solution on  $(0, \infty) \times \Omega$  associated with  $u_0^\infty$ , which is indeed the limit of  $u^T$ .

Next, define  $\tilde{p}^T := m^T p^T$ , where  $m^T$  is a positive constant chosen so that  $\int_\Omega \tilde{p}^T(0, x) dx = 1$ . We know from Theorem 2.1 that there exists a constant  $c^T$  such that

- i) if  $0 < u_0^T(x)$  then  $\tilde{p}^T(0, x) \geq c^T$ ,
- ii) if  $u_0^T(x) < 1$  then  $\tilde{p}^T(0, x) \leq c^T$ .

Parabolic regularity yields that the solution  $\tilde{p}^T$  converges in  $W_{q,loc}^{1,2}((0, \infty) \times \Omega)$  for all  $q \in (1, \infty)$  as  $T \rightarrow +\infty$  to a solution  $p$  of the backward equation

$$\begin{cases} -\partial_t p - \sigma \Delta p = f'(u)p & \text{in } (0, \infty) \times \Omega, \\ p(t, x) > 0 & \text{in } (0, \infty) \times \Omega, \\ \frac{\partial p}{\partial \nu}(t, x) = 0 & \text{for all } t \in (0, \infty), \text{ for all } x \in \partial\Omega, \end{cases} \quad (6.1)$$

Indeed, we know from Proposition 2.7 of [12] 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  $u_0^\infty$  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  $\tilde{p}^T$ . Note that this adjoint function does not depend on the cost function anymore, which is satisfying since, as we expect convergence to 1 at large time, the shape of the cost function should not play any role.

## REFERENCES

- [1] L. Almeida, M. Duprez, Y. Privat and N. Vauchelet, Mosquito population control strategies for fighting against arboviruses. *Math. Biosci. Eng.* **16** (2019) 6274–6297.
- [2] L. Almeida, Y. Privat, M. Strugarek and N. Vauchelet, Optimal releases for population replacement strategies, application to Wolbachia. *SIAM J. Math. Anal.* **51** (2019) 3170–3194.
- [3] L. Alphey, A. McKemey, D. Nimmo, M. Neira, R. Lacroix, K. Matzen and C. Beech, Genetic control of aedes mosquitoes. *Pathogens Glob. Health* **107** (2013) 170–179.
- [4] N. Barton and M. Turelli, Spatial waves of advances with bistable dynamics: cytoplasmic and genetic analogues of Allee effects. *Am. Natural.* **78** (2011) E48–E75.
- [5] P.-A. Bliman and N. Vauchelet, establishing traveling wave in bistable reaction-diffusion system by feedback. *IEEE Control Syst. Lett.* **1** (2017) 62–67.
- [6] C. Camacho, B. Zou and M. Briani, On the dynamics of capital accumulation across space. *Eur. J. Oper. Res.* **186** (2008) 451–465.
- [7] J. Garnier, L. Roques and F. Hamel, Success rate of a biological invasion in terms of the spatial distribution of the founding population. *Bull. Math. Biol.* **74** (2012) 453–473.

- [8] A. Hancock, P. Sinkins and H. Charles, Strategies for introducing Wolbachia to reduce transmission of mosquito-borne diseases. *PLoS Negl. Trop. Dis.* **5** (2011) 1–10.
- [9] A. Henrot and M. Pierre, Vol. 48 of *Variation et optimisation de formes. Une analyse géométrique* [A geometric Analysis] (2005).
- [10] A. Hoffmann *et al.*, Successful establishment of Wolbachia in *Aedes* populations to suppress dengue transmission. *Nature* **7361** (2011) 454–457.
- [11] H. Hughes and N. Britton, Modeling the use of Wolbachia to control dengue fever transmission. *Bull. Math. Biol.* **75** (2013) 796–818.
- [12] J. Húska, Harnack inequality and exponential separation for oblique derivative problems on Lipschitz domains. *J. Differ. Equ.* **226** (2006) 541–557.
- [13] O.A. Ladyzenskaja, V.A. Solonnikov and N.N. Uralceva, *Linear and quasilinear equations of parabolic type* (1968).
- [14] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod; Gauthier-Villars, Paris (1969).
- [15] E. Trelat, J. Zhu and E. Zuazua, Allee optimal control of a system in ecology. *Math. Models Methods Appl. Sci.* (2018) 1665–1697.
- [16] T. Walker *et al.*, The wMel Wolbachia strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature* **7361** (2011) 450–453.
- [17] A. Zlatos, Sharp transition between extinction and propagation of reaction. *J. Am. Math. Soc.* **19** (2005) 251–263.