Numerical approximation of bang-bang controls for the heat equation: an optimal design approach

Arnaud Munch, Francisco Periago

HAL Id: hal-00684214
https://hal.archives-ouvertes.fr/hal-00684214v3

Submitted on 2 May 2013

Numerical approximation of bang-bang controls for the heat equation: an optimal design approach

Arnaud Münch* and Francisco Periago†

February 27, 2013

Abstract

This work is concerned with the numerical approximation of null controls of minimal L^∞-norm for the linear heat equation with a bounded potential. Both the cases of internal and boundary controls are considered. Dual arguments typically allow to reduce the search of controls to the unconstrained minimization of a conjugate function with respect to the initial condition of a backward heat equation. However, as a consequence of the regularization property of the heat operator, this condition lives in a huge space that can not be approximated with robustness. For this reason the minimization is severally ill-posed. On the other hand, the optimality conditions for this problem show that the unique control v of minimal L^∞-norm has a bang-bang structure as it takes only two values: this allows to reformulate the problem as an optimal design problem where the new unknowns are the amplitude of the bang-bang control and the space-time regions where it takes its two possible values. This second optimization variable is modeled through a characteristic function. Since this new problem is not convex, we obtain a relaxed formulation of it which, in particular, lets the use of a gradient method for the numerical resolution. Numerical experiments are described within this new approach.

Keywords: Heat equation, Bang-bang control, Convex relaxation, Numerical approximation.

1 Introduction

We consider both the internal and boundary controllability problem of a linear heat equation with a bounded potential. Let us describe the problem in the distributed case for which the state equation is

$$\begin{cases} y_t - \Delta y + a\, y = v\, 1_\omega, & (x,t) \in Q_T \\ y(\sigma, t) = 0, & (\sigma, t) \in \Sigma_T, \qquad y(x,0) = y_0(x), \quad x \in \Omega. \end{cases} \tag{1}$$

Here, we denote by Ω an open and bounded set of $\mathbb{R}^N$, $N \geq 1$, with C^2 boundary Γ, $Q_T = \Omega \times (0,T)$, $\Sigma_T = \Gamma \times (0,T)$, $\omega \subset\subset \Omega$ is a non-empty open subset of Ω, 1_ω is the associated characteristic function, $q_T = \omega \times (0,T)$, $T > 0$, $y_0 \in L^2(\Omega)$, $a \in L^\infty(Q_T)$, $v \in L^\infty(q_T)$ is the *control* and y is the associated *state*.

For any $y_0 \in L^2(\Omega)$, $T > 0$ and $v \in L^\infty(q_T)$, there exists exactly one solution of (1), with $y \in C^0([0,T]; L^2(\Omega)) \cap L^2(0,T; H_0^1(\Omega))$ (see [14, 15]). The null controllability problem for (1) at time $T > 0$ amounts to find a function v in $L^\infty(q_T)$, such that the associated solution to (1) satisfies

$$y(x,T) = 0, \quad x \in \Omega. \tag{2}$$

This problem has been solved for the heat equation in the nineties: we mention the seminal contributions [8], [13] and more recently [1].

In this work, we are interested in the numerical approximation of the following optimization problem: for any $\alpha \geq 0$

$$(P_\alpha) \begin{cases} \text{Minimize} \quad J_\alpha(v) = \|v\|_{L^\infty(q_T)} \\ \text{subject to} \quad v \in \mathcal{C}_\alpha(y_0, T) \end{cases}$$

where $\mathcal{C}_\alpha(y_0, T) = \{v \in L^\infty(q_T) : y \text{ solves (1) and satisfies } \|y(\cdot, T)\|_{L^2(\Omega)} \leq \alpha\}$. Problem (P_0) corresponds to the null controllability. For any $\alpha > 0$, it is shown in [5] that the unique solution of the extremal constrained problem (P_α) is given by

$$v_\alpha = \|\varphi_\alpha\|_{L^1(q_T)} \operatorname{sign}(\varphi_\alpha) \, 1_\omega, \tag{3}$$

(a quasi bang-bang control) where $\varphi_\alpha = \varphi$ solves the backward equation

$$\begin{cases} -\varphi_t - \Delta\varphi + a\,\varphi = 0, \quad (x,t) \in Q_T, \\ \varphi(\sigma, t) = 0, \quad (\sigma, t) \in \Sigma_T, \qquad \varphi(x, T) = \varphi_{\alpha, T}(x), \quad x \in \Omega \end{cases} \tag{4}$$

and with $\varphi_{\alpha, T}$ the unique solution of the following extremal problem, dual of (P_α),

$$(D_\alpha) \begin{cases} \text{Minimize} \quad \mathcal{J}_\alpha(\varphi_{\alpha, T}) = \dfrac{1}{2}\|\varphi\|_{L^1(q_T)}^2 + \alpha\|\varphi_{\alpha, T}\|_{L^2(\Omega)} + \displaystyle\int_\Omega y_0(x)\varphi(x, 0)dx \\ \text{subject to} \quad \varphi \in L^2(\Omega). \end{cases}$$

In the case where the potential a vanishes and, in general, in space dimension $N = 1$ (see [2]), the control (3) is in fact of bang-bang type since the zero set of φ has zero Lebesgue measure. Up to best knowledge of the authors, for a non-vanishing potential a and $N > 1$ it is unknown if such a result holds. However, it is known that in this case ($a \neq 0$ and $N > 1$) the zero set of φ has empty interior (see [5, Remark 1.1]).

For any $\alpha > 0$, the minimization of $\mathcal{J}_\alpha$ can be performed using a gradient method. Once the minimizer $\varphi_{\alpha, T}$ is determined, φ_α is computed from (4) and the control v_α of minimal L^∞-norm is then given by (3). Moreover, as a consequence of the null controllability property, the sequence $(v_\alpha)_{\alpha > 0}$ defined by (3) is uniformly bounded w.r.t. α in $L^\infty(q_T)$. However, as α goes to zero, the minimizer $\varphi_{\alpha, T}$ may be not uniformly bounded in L^2 but in a larger space, say $\mathcal{H}$ defined as the completion of $\mathcal{D}(\Omega)$ with respect to the norm $\|\varphi\|_{L^1(q_T)}$ (we refer to [23], section 4.6 for more details). Actually, for the control of minimal L^2-norm, it is shown in [16] for $N = 1$ that the set of initial data y_0 for which the corresponding minimizer of the conjugate function $\mathcal{J}_\alpha$ belongs to any negative Sobolev space is dense in $L^2(\Omega)$. We also refer to [3, 7, 20, 22] where this phenomenon is fully discussed in the L^2-case. In particular, it is seen that the (so-called HUM) control exhibits a very oscillatory behavior near the controllability time T and that the numerical minimization of $\mathcal{J}_0$ is ill-posed. Since this phenomenon is related to the regularizing property of the heat kernel, it occurs very likely for the L^∞ case as well.

The numerical approximation of bang-bang controls for the heat equation has been addressed in [9] by a way of a penalty and regularization technique: precisely, for any $k \in \mathbb{R}^+$ and any $s \in \mathbb{N}^+$, the extremal problem

$$(P_{k,s}) \begin{cases} \text{Minimize} \quad \dfrac{1}{2}\left(\|v\|_{L^s(q_T)}^2 + k^{-1}\|y(\cdot, T)\|_{L^2(\Omega)}^2\right) \\ \text{subject to} \quad v \in L^s(q_T) \end{cases}$$

is addressed in [9] for s and k^{-1} large enough. For $s = +\infty$, notice that for any α, one can find a parameter k which depend on α and on the data such that the optimal corresponding solution y satisfies exactly $\|y(\cdot, T)\|_{L^2(\Omega)} = \alpha$. In that sense, for $s = +\infty$, problems $(P_{k,s})$ and (P_α) are equivalent. Moreover, the use of the L^s-norm (s large enough) instead of the L^∞-norm avoids the problem of the non-differentiability

of the L^∞-norm and of any power of it. Numerical experiments within this approach are described in [9] (in the inner point-wise case for $N = 1$ and boundary cases for $N = 2$) and - as for the L^2-case - highlight a high oscillatory behavior of the bang-bang control near T.

In this work we propose an alternative approach to solve numerically (P_α). Precisely, we take advantage of the bang-bang structure of the control and hence consider from the very beginning the control system

$$\begin{cases} y_t - \Delta y + a\, y = [\lambda 1_{\mathcal{O}} + (-\lambda)(1 - 1_{\mathcal{O}})]1_\omega, & (x, t) \in Q_T \\ y(\sigma, t) = 0, & (\sigma, t) \in \Sigma_T, \qquad y(x, 0) = y_0(x), \quad x \in \Omega. \end{cases} \tag{5}$$

where λ denotes a positive real number and $\mathcal{O}$ a subdomain of Q_T. Notice that here we impose *a priori* that the control v is of bang-bang type, that is, it takes only two values, λ on $\mathcal{O} \cap q_T$ and $-\lambda$ on $(Q_T \backslash \mathcal{O}) \cap q_T$, respectively. λ is the amplitude of the piecewise constant control and $1_{\mathcal{O}}$ depends on (x, t) but no volume constraint nor regularity assumption are introduced on $\mathcal{O}$. In particular, discontinuities in the time evolution of $\mathcal{O}$ are allowed.

Accordingly, for any $\alpha > 0$, we consider the optimization problem

$$(BB_\alpha) \begin{cases} \text{Minimize in } (\lambda, 1_{\mathcal{O}}) : \dfrac{1}{2}\lambda^2 \\ \text{subject to } (\lambda, 1_{\mathcal{O}}) \in \mathcal{D}_\alpha(y_0, T) \end{cases}$$

where

$$\mathcal{D}_\alpha(y_0, T) = \{(\lambda, 1_{\mathcal{O}}) \in \mathbb{R}^+ \times L^\infty(q_T; \{0, 1\}) : \quad y = y(\lambda, 1_{\mathcal{O}}) \text{ solves } (5) \text{ and satisfies } \|y(\cdot, T)\|_{L^2(\Omega)} \le \alpha\}.$$

(BB_α) can be viewed as an optimal design problem where the design variable is the space-time region where the control takes its two possible values and the optimality is related to the amplitude of the bang-bang control. Also notice that in the case where the control is of bang-bang type, (BB_α) coincides with (P_α). We emphasize that this formulation makes sense and is of a practical interest even when the control of minimal L^∞-norm is not of bang-bang type. Indeed, we address directly the problem of computing bang-bang type controls with minimal amplitude which, as indicated above, is of a major interest in practice. Up to our knowledge, this perspective has not been addressed so far.

Since the space of admissible designs is not convex, we first obtain a well-posed relaxed formulation and then show how this equivalent but new formulation allows to obtain, for any $\alpha > 0$, a robust approximation of the solution of the original problem. Precisely, in Section 2 we introduce and analyze this relaxed formulation (see Theorem 2.1). In particular, we obtain that the relaxed problem is an equivalent penalty version of (P_α) and prove that there exists a minimizing sequence of bang-bang type controls (see Theorem 2.1, part 3, for precise statements). Then, we derive and discuss the first-order necessary optimality condition of the relaxed problem. From this, we recover the bang-bang structure of the control for the pure heat equation and in one space dimension.

The case where the control acts on a part of the boundary in Dirichlet and Neumann forms is also considered and the same type of results is obtained. The numerical resolution of the relaxed problem is addressed in Section 3. We describe the algorithm used to solve the relaxed problem and present several numerical experiments. In particular, the approach allows to capture the oscillatory behavior of the control near T as α goes to zero.

2   Relaxation and necessary optimality conditions

2.1   The inner case

We adopt the penalty approach mentioned before and for simplicity, we still use α to denote the penalty parameter. Hence we transform (BB_α) into the following problem:

$$(T_\alpha) \begin{cases} \text{Minimize in } (\lambda, 1_\mathcal{O}): \quad J_\alpha(\lambda, 1_\mathcal{O}) = \frac{1}{2}\left(\lambda^2 + \alpha^{-1}\|y(\cdot, T)\|^2_{L^2(\Omega)}\right) \\ \text{subject to} \\ \qquad\qquad y_t - \Delta y + a\,y = \lambda\,[(2\,1_\mathcal{O} - 1)]\,1_\omega \quad \text{in} \quad Q_T, \\ \qquad\qquad y = 0 \quad \text{on} \quad \Sigma_T, \qquad y(\cdot, 0) = y_0 \quad \text{in} \quad \Omega \\ \qquad\qquad (\lambda, 1_\mathcal{O}) \in \mathbb{R}^+ \times L^\infty(q_T; \{0, 1\}). \end{cases}$$

Accordingly, we also consider the problem

$$(RT_\alpha) \begin{cases} \text{Minimize in } (\lambda, s): \quad \overline{J}_\alpha(\lambda, s) = \frac{1}{2}\left(\lambda^2 + \alpha^{-1}\|y(\cdot, T)\|^2_{L^2(\Omega)}\right) \\ \text{subject to} \\ \qquad\qquad y_t - \Delta y + a\,y = [\lambda\,(2s - 1)]\,1_\omega \quad \text{in} \quad Q_T \\ \qquad\qquad y = 0 \quad \text{on} \quad \Sigma_T, \qquad y(\cdot, 0) = y_0 \quad \text{in} \quad \Omega \\ \qquad\qquad (\lambda, s) \in \mathbb{R}^+ \times L^\infty(q_T; [0, 1]). \end{cases}$$

In (T_α), the approximate controllability condition $\|y(\cdot, T)\|_{L^2(\Omega)} \leq \alpha$ has been introduced in the cost function via a penalty technique. First, this leads to an easier numerical resolution. Second, this is motivated by the fact that a relaxation for (P_α) (and consequently also for (BB_α)) needs of an *a priori* estimate; precisely a uniform (with respect to $1_\mathcal{O}$) observability inequality for the solutions of the system (4). See [19] where such a uniform observability inequality was proved for the heat equation in 1D and for null controls of minimal L^2-norm. We claim that this type of uniform inequality is difficult to obtain in the general case considered in this work.

Also notice that for all $v \in L^\infty(q_T)$, denoting by $\lambda = \|v\|_{L^\infty(q_T)}$, the identity $v = \lambda(2s - 1)$, with $s \in L^\infty(q_T; [0, 1])$, holds. Thus, the original problem with a cost functional of the form

$$\frac{1}{2}\|v\|^2_{L^\infty(q_T)} + \frac{1}{2\alpha}\|y(\cdot, T)\|^2_{L^2(\Omega)}$$

transforms into our problem (RT_α). In this respect, we are redefining the decision variables of the optimization problem: instead of an L^∞ control function v, we are looking for a control function of the form $\lambda(2s - 1)$, with λ a positive real number and $s \in L^\infty(q_T; [0, 1])$.

From now on we consider the space $L^\infty(q_T; [0, 1])$ endowed with the usual weak-$\star$ topology. We have the following result.

THEOREM 2.1 (RT_α) *is a true relaxation of* (T_α) *in the following sense:*

1. *there exists one minimizer of* (RT_α),

2. *up to subsequences, every minimizing sequence, say* $(\lambda_n, 1_{\mathcal{O}_n})$ *of* (T_α) *converges to some* $(\lambda, s) \in \mathbb{R}^+ \times L^\infty(q_T; [0, 1])$ *such that* (λ, s) *is a minimizer for* (RT_α), *and conversely,*

3. *if* (λ, s) *is a minimizer for* (RT_α) *and if* $1_{\mathcal{O}_n}$ *converges to* s *weak-$\star$ in* $L^\infty(q_T; [0, 1])$, *then, up to a subsequence,* $(\lambda, 1_{\mathcal{O}_n})$ *is a minimizing sequence for* (T_α).

Proof. Let us first prove that the functional $\overline{J}_\alpha(\lambda, s)$ is continuous. Assume that $(\lambda_n, s_n) \in \mathbb{R}^+ \times L^\infty(q_T; [0, 1])$ satisfies

$$\begin{cases} \lambda_n \to \lambda \\ s_n \rightharpoonup s \quad \text{weak} - \star \text{ in } L^\infty(q_T; [0, 1]) \end{cases} \qquad \text{as } n \to \infty.$$

Since $[\lambda_n (2s_n (x,t) - 1)] \rightharpoonup [\lambda (2s (x,t) - 1)]$ weak-$\star$ in $L^\infty (q_T; [0,1])$ (in particular, also weakly in $L^2 (q_T)$), the solution y^n of the system

$$\begin{cases} y_t^n - \Delta y^n + a\, y^n = [\lambda_n (2s_n (x,t) - 1)] \, 1_\omega, & (x,t) \in Q_T, \\ y^n (\sigma, t) = 0, \quad (\sigma, t) \in \Sigma_T, \qquad y^n (x,0) = y_0 (x), \quad x \in \Omega \end{cases}$$

satisfies

$$\begin{cases} y^n \rightharpoonup y & \text{weakly in } L^2 \left(0, T; H_0^1 (\Omega)\right) \\ y_t^n \rightharpoonup y_t & \text{weakly in } L^2 \left(0, T; H^{-1} (\Omega)\right), \end{cases}$$

where $y = y (x,t)$ solves

$$\begin{cases} y_t - \Delta y + a\, y = [\lambda (2s (x,t) - 1)] \, 1_\omega, & (x,t) \in Q_T, \\ y (\sigma, t) = 0, \quad (\sigma, t) \in \Sigma_T, \qquad y (x,0) = y_0 (x), \quad x \in \Omega. \end{cases}$$

By Aubin's lemma, up to a subsequence still labeled by n, $y^n \to y$ strongly in $L^2 \left(0, T; L^2 (\Omega)\right)$.

Hence, up to a subsequence,

$$y^n (t, \cdot) \to y (t, \cdot) \quad \text{strongly in } L^2 (\Omega) \text{ and a.e. } t \in [0, T]. \tag{6}$$

Since $y^n(t)$ are continuous functions, convergence (6) in fact holds for all $t \in [0, T]$. In particular, $\overline{J}_\alpha (\lambda_n, s_n) \to \overline{J}_\alpha (\lambda, s)$ as $n \to \infty$. Moreover, $\overline{J}_\alpha (\lambda, s)$ is clearly coercitive. As a consequence, problem (RT_α) has a solution.

Statements 2. and 3. are a straightforward consequence of the continuity of $\overline{J}_\alpha (\lambda, s)$ and of the density of the space $L^\infty (q_T; \{0, 1\})$ in $L^\infty (q_T; [0, 1])$ (see [12, Prop. 7.2.14, p. 289]). □

Next, we analyze the first-order necessary optimality condition for the relaxed problem (RT_α).

THEOREM 2.2 *The functional $\overline{J}_\alpha$ as defined in problem (RT_α) is Gâteaux differentiable and its directional derivative at (λ, s) in the admissible direction $\left(\widehat{\lambda}, \widehat{s}\right)$ is given by*

$$\frac{\partial \overline{J}_\alpha(\lambda, s)}{\partial(\lambda, s)} \cdot (\widehat{\lambda}, \widehat{s}) = \widehat{\lambda} \left(\lambda - \int_{q_T} p(2s - 1) \, dx \, dt\right) - 2\lambda \int_{q_T} p\widehat{s} \, dx dt \tag{7}$$

where $p \in C \left([0, T]; L^2 (\Omega)\right) \cap L^2 \left(0, T; H_0^1(\Omega)\right)$ solves the adjoint equation

$$\begin{cases} -p_t - \Delta p + a\, p = 0, & (x,t) \in Q_T \\ p(\sigma, t) = 0, \quad (\sigma, t) \in \Sigma_T, \qquad p(x,T) + \alpha^{-1} y(x,T) = 0, \quad x \in \Omega, \end{cases} \tag{8}$$

where y is the solution of the heat equation in problem (RT_α).

Proof. Let $\left(\widehat{\lambda}, \widehat{s}\right) \in \mathbb{R}^+ \times L^\infty (q_T; [0,1])$ be an admissible direction, i.e., for ε small enough, $\left(\lambda + \varepsilon\widehat{\lambda}, s + \varepsilon\widehat{s}\right) \in \mathbb{R}^+ \times L^\infty (q_T; [0,1])$. Denote by $y_{\left(\lambda + \varepsilon\widehat{\lambda}, s + \varepsilon\widehat{s}\right)}$ the solution of the state law as defined in (RT_α) associated with the perturbation $\left(\lambda + \varepsilon\widehat{\lambda}, s + \varepsilon\widehat{s}\right)$. Thanks to the linearity of the heat equation it is easy to see that

$$y_{\left(\lambda + \varepsilon\widehat{\lambda}, s + \varepsilon\widehat{s}\right)} = y_{(\lambda, s)} + \varepsilon\widehat{y} + \varepsilon^2 \widetilde{y}$$

where $y_{(\lambda, s)}$ is the state associated with the control (λ, s), $\widehat{y}$ is a solution to

$$\begin{cases} \widehat{y}_t - \Delta\widehat{y} + a\widehat{y} = \left[2\lambda\widehat{s} + \widehat{\lambda} (2s - 1)\right] 1_\omega, & (x,t) \in Q_T \\ \widehat{y} (\sigma, t) = 0, \quad (\sigma, t) \in \Sigma_T, \qquad \widehat{y} (x,0) = 0, \quad x \in \Omega, \end{cases} \tag{9}$$

and $\widetilde{y}$ solves

$$\begin{cases} \widetilde{y}_t - \Delta\widetilde{y} + a\,\widetilde{y} = 2\widehat{\lambda}\widehat{s}1_\omega, & (x,t) \in Q_T \\ \widetilde{y}(\sigma,t) = 0, & (\sigma,t) \in \Sigma_T, \qquad \widetilde{y}(x,0) = 0, \quad x \in \Omega. \end{cases}$$

A straightforward computation shows that

$$\frac{\partial\overline{J}_\alpha(\lambda,s)}{\partial(\lambda,s)} \cdot (\hat{\lambda},\hat{s}) = \lim_{\varepsilon\to 0}\frac{\overline{J}_\alpha\left(\lambda + \varepsilon\widehat{\lambda}, s + \varepsilon\widehat{s}\right) - \overline{J}_\alpha(\lambda,s)}{\varepsilon} = \lambda\widehat{\lambda} + \alpha^{-1}\int_\Omega y_{(\lambda,s)}(\cdot,T)\,\widehat{y}(\cdot,T)\,dx. \tag{10}$$

On the other hand, taking into account the initial condition $\widehat{y}(\cdot,0) = 0$ and the final condition $p(\cdot,T) = -\alpha^{-1}y(\cdot,T)$, from the weak form of the system (9) it is deduced that

$$\alpha^{-1}\int_\Omega y_{(\lambda,s)}(\cdot,T)\,\widehat{y}(\cdot,T)\,dx = -\widehat{\lambda}\int_{q_T} p\,(2s-1)\,dxdt - 2\lambda\int_{q_T} p\,\widehat{s}\,dxdt$$

for $p \in \mathcal{W}(0,T) = \left\{v \in L^2\left(0,T;H_0^1(\Omega)\right) : v_t \in L^2\left(0,T;H^{-1}(\Omega)\right)\right\}$ solution of (8). Replacing this expression into (10) we obtain (7). $\qquad\square$

COROLLARY 2.1 *Let* $(\lambda^\star, s^\star) \in \mathbb{R}^+ \times L^\infty(q_T;[0,1])$ *be an optimal solution of* (RT_α). *Then,*

$$s^\star(x,t) = \begin{cases} 0 & if \quad p(x,t) < 0 \\ 1 & if \quad p(x,t) > 0 \end{cases} \tag{11}$$

and $\lambda^\star = \|p\|_{L^1(q_T)}$. *Consequently, if* $N = 1$ *or if the potential* $a = 0$ *for* $N > 1$, *then* $s^\star$ *is a characteristic function and therefore problem* (T_α) *is well-posed, i.e., the control is of bang-bang type.*

Proof. Let $(\lambda^\star, s^\star) \in \mathbb{R}^+ \times L^\infty(q_T;[0,1])$ be an optimal solution of (RT_α). From (7) it follows that

$$(\lambda - \lambda^\star)\left(\lambda^\star - \int_{q_T} p(2s^\star - 1)\,dx\,dt\right) - 2\lambda^\star\int_{q_T} p(s - s^\star)\,dxdt \geq 0 \tag{12}$$

for all $(\lambda,s) \in \mathbb{R}^+ \times L^\infty(q_T;[0,1])$. In particular, if $\lambda = \lambda^\star$, then

$$\int_{q_T} p\,s^\star dxdt \geq \int_{q_T} p\,s\,dxdt \quad \forall s \in L^\infty(q_T;[0,1]).$$

A standard localization argument (see for instance [21, pages 67-69]) shows that this variational inequality is equivalent to the point-wise variational inequality

$$p(x,t)\,s^\star(x,t) \geq p(x,t)\,s(x,t) \quad \forall s \in L^\infty(q_T;[0,1]), \text{ for a.e. } (x,t) \in q_T.$$

From this we easily obtain (11). Now consider the case of the pure heat equation, i.e., $a = 0$. Using the fact that thanks to the analyticity of p the zero set of p has zero Lebesgue measure, we conclude that $s^\star$ is a characteristic function. The same holds if $a \neq 0$ and $N = 1$ (see [2]).

Finally, if we put $s = s^\star$ in (12), then

$$\lambda^\star = \int_{q_T} p\,(2s^\star - 1)\,dxdt = \|p\|_{L^1(q_T)},$$

where the last equality is a consequence of (11). Notice that from this last equality and (11), the optimal control $\lambda^\star(2s^\star - 1)$ has exactly the structure given by (3). $\qquad\square$

Remark 1 *Notice that even in the case where* (T_α) *is well-posed, the relaxed formulation* (RT_α) *remains very useful at the numerical level. Precisely, since the admissibility set for* (RT_α) *is convex (contrary to what happens in* (T_α)*), it is allowed to make variations in this space and therefore we may implement a gradient algorithm to solve* (RT_α)*, and consequently also* (T_α)*. Moreover, Theorem 2.1, part 3, provides a constructive way of computing a minimizing sequence of bang-bang type controls for the general case of the heat equation with a potential in dimension* $N > 1$.

2.2 The boundary case

We address the situation where the control acts on a part of the boundary. Recall that $\Sigma_T := \Gamma \times (0, T)$ with $\Gamma = \partial\Omega$. We note $\Sigma_0 = \Gamma_0 \times (0, T)$ for any $\Gamma_0 \subset \Gamma$.

2.2.1 Dirichlet-type controls

For a fixed $\alpha > 0$, we focus on the system

$$\begin{cases} y_t - \Delta y + a\,y = 0, & (x, t) \in Q_T \\ y(\sigma, t) = f(\sigma, t)\,1_{\Sigma_0}, & (\sigma, t) \in \Sigma_T, \qquad y(x, 0) = y_0(x), \quad x \in \Omega \end{cases} \tag{13}$$

and look for the control f, with support in Σ_0, which satisfies

$$\|y(\cdot, T)\|_{H^{-1}(\Omega)} \le \alpha. \tag{14}$$

It is important to notice that we have moved from the L^2-norm for the final state to the $H^{-1}-$norm because, as observed in ([14, p. 217]), if $f \in L^2(\Sigma_0)$ and $y_0 \in L^2(\Omega)$, then, in general, $y(\cdot, T) \notin L^2(\Omega)$.

Under some technical assumptions, some positive results concerning the existence of a solution for the approximate controllability problem (13)-(14) are obtained in [5].

Similarly to the inner situation, we consider the optimization problem

$$(B_\alpha) \quad \begin{cases} \text{Minimize in } (\lambda, 1_{\mathcal{O}}): & J_\alpha(\lambda, 1_{\mathcal{O}}) = \frac{1}{2}\left(\lambda^2 + \alpha^{-1}\|y(\cdot, T)\|_{H^{-1}(\Omega)}^2\right) \\ \text{subject to} \\ \qquad y_t - \Delta y + a\,y = 0 \quad \text{in} \quad Q_T \\ \qquad y = \lambda\,[2\,1_{\mathcal{O}} - 1]\,1_{\Sigma_0} \quad \text{on} \quad \Sigma_T, \quad y(\cdot, 0) = y_0 \quad \text{in} \quad \Omega \\ \qquad (\lambda, 1_{\mathcal{O}}) \in \mathbb{R}^+ \times L^\infty(\Sigma_0; \{0, 1\}) \end{cases}$$

with $y_0 \in L^2(\Omega)$. We also consider the new problem

$$(RB_\alpha) \quad \begin{cases} \text{Minimize in } (\lambda, s): & \overline{J}_\alpha(\lambda, s) = \frac{1}{2}\left(\lambda^2 + \alpha^{-1}\|y(\cdot, T)\|_{H^{-1}(\Omega)}^2\right) \\ \text{subject to} \\ \qquad y_t - \Delta y + a\,y = 0 \quad \text{in} \quad Q_T \\ \qquad y = \lambda\,[2s - 1]\,1_{\Sigma_0} \quad \text{on} \quad \Sigma_T, \quad y(\cdot, 0) = y_0 \quad \text{in} \quad \Omega \\ \qquad (\lambda, s) \in \mathbb{R}^+ \times L^\infty(\Sigma_0; [0, 1]). \end{cases}$$

Then, we have:

THEOREM 2.3 *(RB_α) is a relaxation of (B_α) in the same terms as stated in Theorem 2.1.*

Before proving this result, we recall some results concerning the properties of the solution to the following non-homogeneous system:

$$\begin{cases} y_t - \Delta y + a\,y = 0, & (x, t) \in Q_T \\ y(\sigma, t) = f(\sigma, t), & (\sigma, t) \in \Sigma_T, \qquad y(x, 0) = y_0(x), \quad x \in \Omega, \end{cases} \tag{15}$$

with $f \in L^\infty(\Sigma_T)$ and $y_0 \in L^2(\Omega)$. Following [14, pp. 208-221] or [15, Vol. II, p.86], a weak solution of (15) is a function $y \in L^2(Q_T)$ which satisfies

$$\int_{Q_T} y(-\varphi_t - \Delta\varphi + a\,\varphi)dxdt = \int_\Omega y_0(x)\varphi(x, 0)dx - \int_{\Sigma_T} f\partial_\nu\varphi\,d\Sigma_T$$

for all $\varphi \in X^1(Q_T) = \{v \in H^{2,1}(Q_T) : v = 0 \quad \text{on} \quad \Sigma_T \quad \text{and} \quad v(x,T) = 0, x \in \Omega\}$. As usual, $\partial_\nu \varphi$ denotes the derivative of φ in the direction of the outward unit normal vector to Γ. Then, it is proved (see also [5, Prop. 5.1]) that there exists a unique weak solution of system (15) which has the regularity

$$y \in H^{1/2,1/4}(Q_T) = L^2\left(0,T;H^{1/2}(\Omega)\right) \cap H^{1/4}\left(0,T;L^2(\Omega)\right).$$

Moreover, the estimate

$$\|y\|_{H^{1/2,1/4}(Q_T)} \leq c\left(\|f\|_{L^\infty(\Sigma_T)} + \|y_0\|_{L^2(\Omega)}\right)$$

holds. In particular,

$$\|y\|_{L^2(0,T;L^2(\Omega))} \leq c\left(\|f\|_{L^\infty(\Sigma_T)} + \|y_0\|_{L^2(\Omega)}\right). \tag{16}$$

On the other hand, standard arguments show that the weak solution of (15) also satisfies $y_t \in L^2\left(0,T;H^{-2}(\Omega)\right)$ and

$$\|y_t\|_{L^2(0,T;H^{-2}(\Omega))} \leq c\left(\|f\|_{L^\infty(\Sigma_T)} + \|y_0\|_{L^2(\Omega)}\right). \tag{17}$$

Finally, we also notice that $y \in C\left([0,T];H^{-1}(\Omega)\right)$ (see [15, Vol. I, Th. 3.1]). As a consequence, the cost functionals J_α and $\overline{J}_\alpha$ above are well-defined.

Proof of Theorem 2.3. The proof follows the same lines as in Theorem 2.1. The only difference is that in this case, from estimates (16) and (17), we have the convergences

$$\begin{cases} y^n \rightharpoonup y & \text{weakly in } L^2\left(0,T;L^2(\Omega)\right) \\ y_t^n \rightharpoonup y_t & \text{weakly in } L^2\left(0,T;H^{-2}(\Omega)\right), \end{cases}$$

where $y = y(x,t)$ solves

$$\begin{cases} y_t - \Delta y + ay = 0, & (x,t) \in Q_T \\ y(\sigma,t) = \lambda\left(2s(\sigma,t) - 1\right)1_{\Sigma_0}, & (\sigma,t) \in \Sigma_T, \qquad y(x,0) = y_0(x), \quad x \in \Omega. \end{cases}$$

But again by Aubin's lemma, up to a subsequence $y^n \to y$ strongly in $L^2\left(0,T;H^{-1}(\Omega)\right)$. The rest of the proof runs as in Theorem 2.1. □

Remark 2 *We notice that if the initial condition $y_0 \in L^\infty(\Omega)$, then the solution of system (15) satisfies $y \in L^\infty(Q_T)$ (see [14, p.221]). In particular, $y(\cdot,T) \in L^2(\Omega)$ and therefore the $H^{-1}(\Omega)-$norm in the costs J_α and $\overline{J}_\alpha$ may be replaced by the $L^2(\Omega)-$norm.*

From now on, we assume that y_0 belongs to $L^\infty(\Omega)$ and then replace in J_α and $\overline{J}_\alpha$ the term $\alpha^{-1}\|y(\cdot,T)\|^2_{H^{-1}(\Omega)}$ by $\alpha^{-1}\|y(\cdot,T)\|^2_{L^2(\Omega)}$. This leads to a stronger underlying controllability condition. Similarly to Theorem 2.2 and Corollary 2.1 we have:

THEOREM 2.4 *For $y_0 \in L^\infty(\Omega)$ the functional $\overline{J}_\alpha$ as defined above is Gâteaux differentiable and its directional derivative at (λ,s) in the admissible direction $\left(\widehat{\lambda},\widehat{s}\right)$ is given by*

$$\frac{\partial \overline{J}_\alpha(\lambda,s)}{\partial(\lambda,s)} \cdot (\hat{\lambda},\hat{s}) = \hat{\lambda}\left(\lambda + \int_{\Sigma_0} \partial_\nu p(2s-1)\,d\Sigma_0\right) + 2\lambda \int_{\Sigma_0} \partial_\nu p\,\hat{s}\,d\Sigma_0 \tag{18}$$

where p solves the backward equation (8).

Remark 3 *As shown in [5, Lemma 4.1 and Remark 5.1], if $y_0 \in L^\infty(\Omega)$, then $\partial_\nu p \in L^1(\Sigma_0)$ and hence the integrals in (18) make sense.*

COROLLARY 2.2 *Let $(\lambda^\star,s^\star) \in \mathbb{R}^+ \times L^\infty(\Sigma_0;[0,1])$ be an optimal solution of (RB_α). Then,*

$$s^\star(\sigma,t) = \begin{cases} 0 & \text{if} \quad \partial_\nu p(\sigma,t) < 0 \\ 1 & \text{if} \quad \partial_\nu p(\sigma,t) > 0 \end{cases} \tag{19}$$

and $\lambda^\star = \|\partial_\nu p\|_{L^1(\Sigma_0)}$. As a consequence, if $N = 1$ or if the potential $a = 0$ for $N > 1$, then $s^\star$ is a characteristic function and therefore problem (B_α) is well-posed, i.e., the control is of bang-bang type.

2.2.2   Neumann-type controls

Consider the system

$$\begin{cases} y_t - \Delta y + a\, y = 0 & (x,t) \in Q_T \\ \partial_\nu y = g \quad \Sigma_T, \qquad y(\cdot,0) = y_0(x), \quad x \in \Omega. \end{cases} \tag{20}$$

For $y_0 \in L^2(\Omega)$ and $g \in L^2\left(0,T;H^{-1/2}(\Gamma)\right)$ the system (20) has a unique solution $y \in L^2\left(0,T;H^1(\Omega)\right) \cap C\left([0,T];L^2(\Omega)\right)$ which satisfies the variational formulation

$$\frac{d}{dt} \int_\Omega y(x,t)\, v(x)\, dx + \int_\Omega [\nabla y(x,t)\, \nabla v(x) + a(x,t)\, y(x,t)\, v(x)]\, dx = \; <g(t),v>_\Gamma \quad \forall v \in H^1(\Omega),$$

where $< \cdot, \cdot >_\Gamma$ stands for the duality product in $H^{1/2}(\Gamma)$ (see [15, 21]). Moreover,

$$\|y\|_{C([0,T];L^2(\Omega))} + \|y\|_{L^2(0,T;H^1(\Omega))} \leq C\left(\|y_0\|_{L^2(\Omega)} + \|g\|_{L^2(0,T;H^{-1/2}(\Gamma))}\right).$$

With the same notation as in the preceding section, we consider the two problems

$$(NB_\alpha) \begin{cases} \text{Minimize in } (\lambda, 1_\mathcal{O}): \quad J_\alpha(\lambda, 1_\mathcal{O}) = \frac{1}{2}\left(\lambda^2 + \alpha^{-1}\|y(\cdot,T)\|^2_{L^2(\Omega)}\right) \\ \text{subject to} \\ \qquad\qquad y_t - \Delta y + a\, y = 0 \quad \text{in} \quad Q_T \\ \qquad\qquad \partial_\nu y = \lambda\, [2\, 1_\mathcal{O} - 1]\, 1_{\Sigma_0} \quad \text{on} \quad \Sigma_T, \quad y(\cdot,0) = y_0 \quad \text{in} \quad \Omega \\ \qquad\qquad (\lambda, 1_\mathcal{O}) \in \mathbb{R}^+ \times L^\infty\left(\Sigma_T;\{0,1\}\right) \end{cases}$$

and

$$(RNB_\alpha) \begin{cases} \text{Minimize in } (\lambda, s): \quad \overline{J}_\alpha(\lambda, s) = \frac{1}{2}\left(\lambda^2 + \alpha^{-1}\|y(\cdot,T)\|^2_{L^2(\Omega)}\right) \\ \text{subject to} \\ \qquad\qquad y_t - \Delta y + a\, y = 0 \quad \text{in} \quad Q_T \\ \qquad\qquad \partial_\nu y = \lambda\, [2s - 1]\, 1_{\Sigma_0} \quad \text{on} \quad \Sigma_T, \quad y(\cdot,0) = y_0 \quad \text{in} \quad \Omega \\ \qquad\qquad (\lambda, s) \in \mathbb{R}^+ \times L^\infty\left(\Sigma_T;[0,1]\right). \end{cases}$$

The same type of arguments as the ones used in the two preceding cases lets prove that (RNB_α) is a relaxation of (NB_α). Also, a direct computation shows that the functional $\overline{J}_\alpha$ is Gâteaux differentiable and its directional derivative at (λ, s) in the admissible direction $\left(\widehat{\lambda}, \widehat{s}\right)$ is given by

$$\frac{\partial \overline{J}_\alpha(\lambda, s)}{\partial(\lambda, s)} \cdot \left(\widehat{\lambda}, \widehat{s}\right) = \widehat{\lambda}\left(\lambda + \int_{\Sigma_0} p\,(2s - 1)\, d\Sigma_0\right) + 2\lambda \int_{\Sigma_0} \widehat{s}\, p\, d\Sigma, \tag{21}$$

where p solves the system

$$\begin{cases} -p_t - \Delta p + a\, p = 0 \quad \text{in} \quad Q_T \\ \partial_\nu p = 0 \quad \text{on} \quad \Sigma_T, \qquad p(\cdot,T) = \alpha^{-1} y(\cdot,T) \quad \text{in} \quad \Omega. \end{cases} \tag{22}$$

By using (21), we obtain that if $(\lambda^\star, s^\star)$ is a solution of (RNB_α), then

$$s^\star(\sigma, t) = \begin{cases} 0 & \text{if} \quad p(\sigma, t) > 0 \\ 1 & \text{if} \quad p(\sigma, t) < 0 \end{cases}$$

and $\lambda^\star = \|p\|_{L^1(\Sigma_0)}$.

3 Numerical experiments for $N = 1$ and $N = 2$

We first consider the 1D case with $\Omega = (0,1)$ and highlight the oscillatory behavior of the null bang-bang controls near T. Then, we consider a 2D situation with $\Omega = (0,1) \times (0,1)$ and compare our approach with the dual method introduced in [9]. Preliminary, we briefly discuss the gradient algorithms and the approximation used to solve the different extremal problems.

3.1 Algorithm and numerical approximation

Let us provide some details on the inner situation and in the 1D case for which we take $\Omega = (0,1)$.

First notice that we may remove the constraint $\lambda \in \mathbb{R}^+$ since, if (λ, s) solves (RT_α), then $(-\lambda, 1 - s)$ is also a solution. Hence, the expression (7) provides the following iterative descent algorithm :

$$
\begin{cases}
(\lambda^0, s^0) \quad \text{given in} \quad \mathbb{R} \times L^\infty(Q_T, [0,1]), \\
\lambda^{n+1} = \lambda^n - a_n \left(\lambda^n - \int_{q_T} p^n (2s^n - 1) \, dx \, dt \right), \quad n \geq 0, \\
s^{n+1} = P_{[0,1]}(s^n + b_n \lambda^n p^n), \quad n \geq 0
\end{cases}
\tag{23}
$$

where p^n solves (8), $P_{[0,1]}(x) = \max(0, \min(1, x))$ denotes the projection of any x onto $[0,1]$ and a_n, b_n denote the optimal descent step which is obtained as the solution of the extremal problem :

$$
\min \overline{J}_\alpha(\lambda^{n+1}(a), s^{n+1}(b)) \quad \text{over} \quad a, b \in \mathbb{R}^+.
\tag{24}
$$

Problem (24) is solved by using line search techniques. The gradient algorithm is stopped as soon as

$$
\left| \lambda^n - \int_{q_T} p^n (2s^n - 1) \, dx \, dt \right| \leq \sigma |\lambda^n|
\tag{25}
$$

for some given tolerance $\sigma > 0$ small enough.

As for the numerical discretization, we use the two-step Gear scheme (also used in [4] in the same context) for the time integration coupled with a P_1 finite element approach for the spatial approximation. Precisely, for large integer N_x, we consider the N_x points $x_i \in [0,1]$ such that $x_1 = 0$, $x_i < x_{i+1}$ and $x_{N_x} = 1$. We note for $i \in \mathcal{I} = \{1, \ldots, N_x - 1\}$, $\Delta x_i = x_{i+1} - x_i$ and $\Delta x = \max_{i \in \mathcal{I}} \Delta x_i$. We note by $\mathcal{P}_{\Delta x}$ the corresponding partition of $\overline{\Omega} = [0,1]$ and by $\mathcal{P}_{\Delta t}$ the corresponding partition of $[0,T]$, obtained in the same way. Finally, set $h = (\Delta x, \Delta t)$ and $\mathcal{Q}_h$ the quadrangulation of Q_T associated to h so that $\overline{Q_T} = \bigcup_{K \in \mathcal{Q}_h} K$.

The following (conformal) finite element approximation of $L^2(0, T; H_0^1(0,1))$ is introduced:

$$
X_{0h} = \{ \varphi_h \in C^0(\overline{Q_T}) : \varphi_h|_K \in (\mathbb{P}_{1,x} \otimes \mathbb{P}_{1,t})(K) \ \ \forall K \in \mathcal{Q}_h, \varphi_h(0,t) = \varphi_h(1,t) = 0 \ \ \forall t \in (0,T) \}.
$$

Here $\mathbb{P}_{m,\xi}$ denotes the space of polynomial functions of order m in the variable ξ. Accordingly, the functions in X_{0h} reduce on each quadrangle $K \in \mathcal{Q}_h$ to a linear polynomial in both x and t. The space X_{0h}, conformal approximation of $L^2(Q_T)$ is a finite-dimensional subspace of $L^2(0, T; H_0^1(0,1))$. Moreover, the functions $\varphi_h \in X_{0h}$ are uniquely determined by their values at the nodes (x_i, t_j) of $\mathcal{Q}_h$ such that $0 < x_i < 1$.

Let us now introduce other finite dimensional spaces. First, we set

$$
\Phi_{\Delta x} = \{ z \in C^0([0,1]) : z|_k \in \mathbb{P}_{1,x}(k) \ \ \forall k \in \mathcal{P}_{\Delta x} \}.
$$

Then, $\Phi_{\Delta x}$ is a finite dimensional subspace of $L^2(0,1)$ and the functions in $\Phi_{\Delta x}$ are uniquely determined by their values at the nodes of $\mathcal{P}_{\Delta x}$.

Secondly, since the variable $\lambda(2s-1) \in L^\infty(Q_T)$ appears in the right hand side of the forward problem in y, it is natural to approximate $\lambda(2s-1) \in L^\infty(Q_T)$ by a piecewise constant function. Thus, let M_h be the space defined by

$$M_h = \{\, \mu_h \in L^\infty(Q_T) : \mu_h|_K \in (\mathbb{P}_{0,x} \otimes \mathbb{P}_{0,t})(K) \;\; \forall K \in \mathcal{Q}_h \,\}.$$

M_h is a finite dimensional subspace of $L^\infty(Q_T)$ and the functions in M_h are uniquely determined by their (constant) values on the quadrangles $K \in \mathcal{Q}_h$.

Therefore, for any $s_h \in M_h$ and any $\lambda_h \in \mathbb{R}$, the approximation $y_h \in X_{0h}$ of the solution of (RT_α) is given as follows :

(i) Consider the times $t_j = j\Delta t$ and set $y_h|_{t=0} = \pi_{\Delta x}(y_0) \in \Phi_{\Delta x}$.

(ii) Then, $y_h|_{t=t_1}$ is the solution to the in $\Psi \in \Phi_{\Delta x}$

$$\left\{ \begin{aligned}
& \int_0^1 \frac{1}{\Delta t}(\Psi - y_h|_{t=0})z\,dx + \frac{1}{2}\int_0^1 (\Psi_x z_x + \pi_{\Delta x} A(x,t_1)\Psi z)\,dx \\
& \quad + \frac{1}{2}\int_0^1 ((y_h|_{t=0})_x z_x + \pi_{\Delta x} A(x,t_{N_t})y_h|_{t=0}z)\,dx \\
& \quad = \frac{1}{2}\lambda_h \int_0^1 ((2s_h(x,t_1)-1) + (2s_h(x,t_0)-1))\,z(x)\,dx \quad \forall z \in \Phi_{\Delta x}.
\end{aligned} \right.$$

(iii) Finally, for given $n = 1,\ldots,N-1$, $\Psi^\star = y_h|_{t=t_{n-1}}$ and $\overline{\Psi} = y_h|_{t=t_n}$, $y_h|_{t=t_{n+1}}$ is the solution to the linear problem

$$\left\{ \begin{aligned}
& \int_0^1 \frac{1}{2\Delta t}(3\Psi - 4\overline{\Psi} + \Psi^\star)z\,dx + \int_0^1 (\Psi_x z_x + \pi_{\Delta x}(A(x,t_{n-1}))\Psi z)\,dx \\
& \quad = \int_0^1 \mu_h(x,t_{n-1})z(x)\,dx \quad \forall z \in \Phi_{\Delta x}.
\end{aligned} \right.$$

Here, $\pi_{\Delta x}$ denotes the projection over $\Phi_{\Delta x}$.

At the finite dimensional level, algorithm (23) reads as follows :

$$\left\{ \begin{aligned}
& (\lambda_h^0, s_h^0) \in \mathbb{R} \times M_h, \\
& \lambda_h^{n+1} = \lambda_h^n - a_{nh}\left(\lambda_h^n - \int_{q_T} p_h^n(2s_h^n - 1)\,dx\,dt\right), \quad n \geq 0, \\
& s_h^{n+1} = P_{[0,1]}(s_h^n + b_{nh}\lambda_h^n p_h^n), \quad n \geq 0
\end{aligned} \right.$$

where p_h^n is an approximation of the backward problem (8), obtained by using P_1 finite element in space and the Gear scheme for the time integration, as described above.

3.2 Experiments in 1D - Oscillatory behavior of the bang-bang control

3.2.1 Distributed case

We take $\Omega = (0,1)$ and first consider the data $\omega = (0.25, 0.75)$, $a \equiv 0$ and $y_0(x) = \sin(2\pi x)$. In order to have a better control of the diffusion, we also replace the operator $-\Delta$ by $-c\Delta$ with c lower than one.

The algorithm is initialized with $\lambda = 1$ and $s = 1/2$. We take $\sigma := 10^{-3}$ as stopping criterion parameter.

We take a uniform partition $\mathcal{P}_{\Delta x}$ for Ω : $x_{i+1} - x_i = 1/N_x$ with $N_x = 400$. On the other hand, in order to describe correctly the oscillations of the density near T, we take a non uniform partition $\mathcal{P}_{\Delta t}$ of the time interval $(0, T)$: precisely we define

$$t_1 = 0; \quad t_{j+1} - t_j = \frac{T}{e^{pT} - 1}(e^{\frac{pT}{N_t}} - 1)e^{\frac{p(N_t+1-j)}{N_t}T}, \quad j = 1, ..., N_t,$$

where $N_t \in \mathbb{N}$ is the number of sub-interval of the partition $\mathcal{P}_{\Delta t}$ and any $p \in \mathbb{N}$. The points t_j, distributed along $(0, T)$, are thus *exponentially* concentrated near $t = T$. Here $p := 6$ and $N_t = 400$.

Table 1 collects the value of λ_h and $\|y_h(\cdot, T)\|_{L^2(0,1)}$ with respect to the penalty parameter α. We take $c := 1/10$. We check that the L^∞-norm λ_h of the control increases as α goes to zero : in other words, the amount of work needed to get closer to the zero target at time T is more important. However, as a consequence of the null controllability of (1), we check that λ_h is uniformly bounded by above w.r.t. α.

The value α also affects the shape of the bang-bang control. Figure 1 depicts the iso-values of the optimal density for $\alpha = 10^{-2}, 10^{-4}, 10^{-6}$ and $\alpha = 10^{-8}$. According to the symmetry of ω and of the initial datum y_0, we obtain a symmetric density over Q_T. For $\alpha = 10^{-2}$, the density is constant in time and related to the sign of y_0: precisely, for all t, $s_h(x, t) = 0$ if $y_0(x) > 0$ and $s_h(x, t) = 1$ if $y_0(x) < 0$. However, for α small enough, for instance here, $\alpha = 10^{-3}$, the optimal density exhibits some variations with respect to the variable t. These variations are mainly located at the end of the time interval. Moreover, as α decreases, the number of theses oscillations, that is, the number of theses changes of sign of v_h increases so that, at the null controllability limit ($\alpha = 0$) one may expected an oscillatory behavior of the bang-bang control, both in space and time, in an arbitrarily close neighborhood of $(0, 1) \times \{T\}$. This is in agreement with our observations in the L^2-case (see [20]). Of course, as in the L^2-case, this behavior may only be captured with an arbitrarily fine mesh. The increasing number of iterates needed to satisfy the criterion (25) as α decreases is also the consequence of these oscillations near T.

In Figure 1 we also observe that (except for $\alpha = 10^{-2}$) the optimal density s_h is not strictly a bi-valued function (as it should be almost everywhere) and takes some intermediate values: this is due to the numerical approximation and to the very low variation of the cost function with respect to the density near the minimum (a bi-valued density is obtained after a very large number of iterates). Figure 2 displays the sign of the adjoint solution p (see (8)) in Q_T and also clearly exhibits the variation of the bang-bang control: recall that from (11), s and p are related through the relation $2s - 1 = \text{sign}(p)$.

More interesting is the fact that, whatever be the initial guess $(\lambda_h^0, s_h^0) \in \mathbb{R}^+ \times L^\infty(Q_T, [0, 1])$ for the algorithm, we always get the same limit. This is of course in agreement with the fact that the control v of minimal L^∞-norm is unique, and so the couple (λ, s) defined by $v = \lambda(2s - 1)1_\omega$ is.

α	10^{-1}	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$\|y_h(\cdot, T)\|_{L^2(\Omega)}$	8.96×10^{-2}	5.34×10^{-2}	6.24×10^{-3}	4.37×10^{-4}	9.17×10^{-5}
λ_h	0.087	0.471	1.309	1.831	1.948
$\sharp$ iterates	11	213	561	1032	4501

Table 1: $N_x = N_t = 400$ - $y_0(x) = \sin(2\pi x)$ - $c = 0.1$

3.2.2   Boundary case

In the boundary situation, the density s is simply a time function. This allows to observe clearer the singular character of the control as α goes to zero when one wants to recover the null control of minimal L^∞-norm. The procedure is similar, except that we consider a non uniform partition $\mathcal{P}_{\Delta x}$ (concentrated on $x = 1$) in order to better describe the final state $p(\cdot, T)$ of the adjoint system (oscillating near $x = 1$):

$$x_1 = 0; \quad x_{i+1} - x_i = \frac{1}{e^p - 1}(e^{\frac{p}{N_x}} - 1)e^{\frac{p(N_x+1-i)}{N_x}}, \quad i = 1, ..., N_x,$$
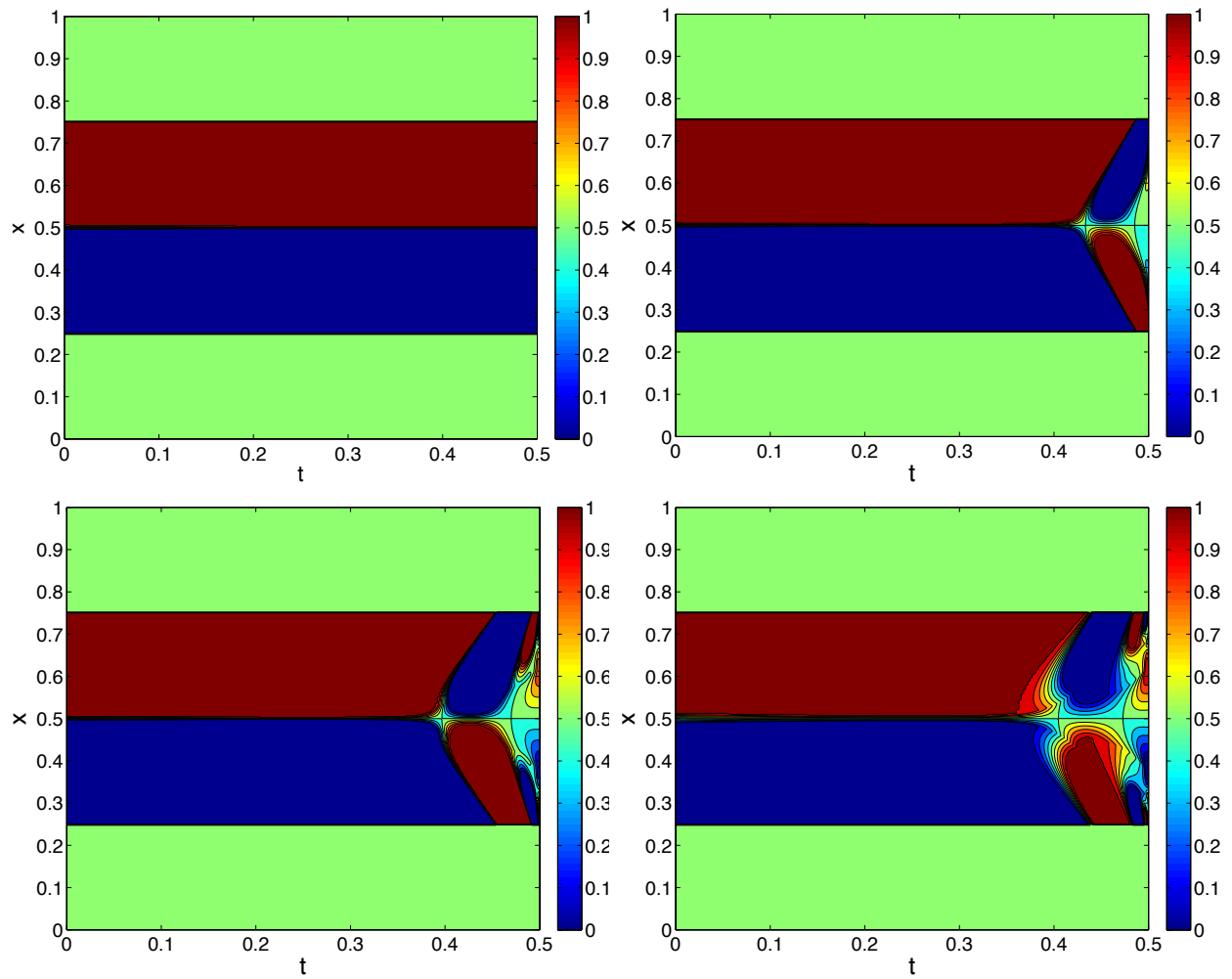
Figure 1: $N_x = N_t = 400$ - $y_0(x) = \sin(2\pi x)$ - $\omega = (0.25, 0.75)$ - $c = 0.1$ - From left to right and from top to bottom, iso-values of the density function s_h in Q_T for $\alpha = 10^{-2}, 10^{-4}, 10^{-6}$ and $\alpha = 10^{-8}$.
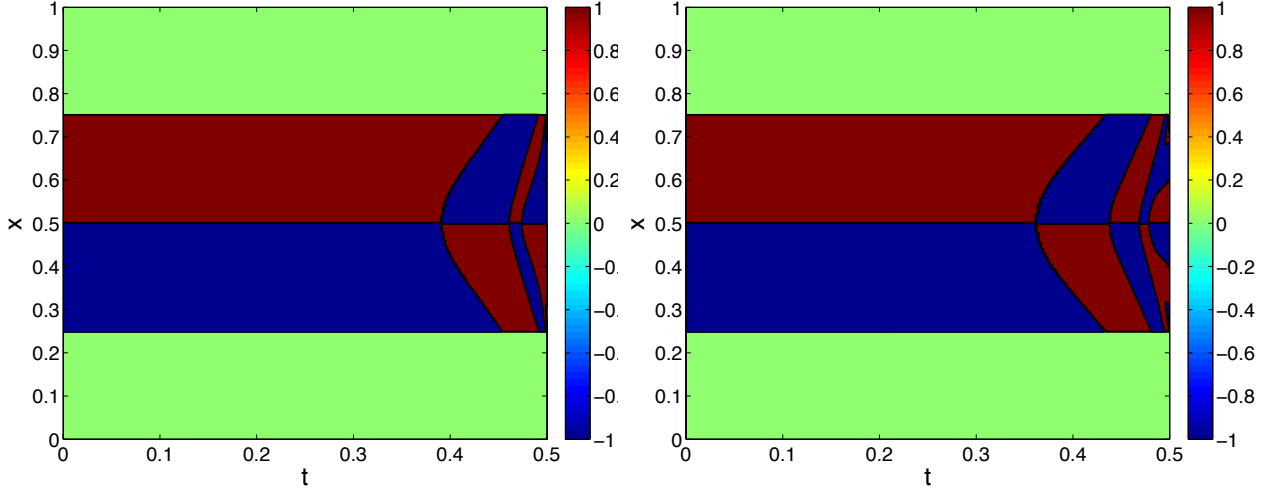
Figure 2: $N_x = N_t = 400$ - $y_0(x) = \sin(2\pi x)$ - $\omega = (0.25, 0.75)$ - $c = 0.1$ - Sign of the adjoint state p for $\alpha = 10^{-6}$ (left) and $\alpha = 10^{-8}$ (right).

with $p = 3$ and $N_x = 400$.

For the Dirichlet boundary control considered in Section 2.2.1, from Theorem 2.4, we consider the following gradient algorithm :

$$\begin{cases} (\lambda_h^0, s_h^0) \in \mathbb{R} \times N_h, \\ \lambda_h^{n+1} = \lambda_h^n - a_{nh}\left(\lambda_h^n + \int_{\Sigma_0} \partial_\nu p_h^n (2s_h^n - 1)\, d\Sigma_0 \right), \quad n \geq 0, \\ s_h^{n+1} = P_{[0,1]}(s_h^n - b_{nh}\lambda_h^n \partial_\nu p_h^n(1, \cdot)), \quad n \geq 0 \end{cases}$$

where p_h^n is an approximation of the backward problem (8) and N_h the space defined by

$$N_h = \{\, \mu_h \in L^\infty([0, T]) : \mu_h|_k \in \mathbb{P}_{0,t}(K) \ \ \forall k \in \mathcal{P}_{\Delta t} \,\}.$$

As in the inner situation, the algorithm is controlled by the absolute value $|\lambda_h^n + \int_{\Sigma_0} \partial_\nu p_h^n (2s_h^n - 1)\, d\Sigma_0|$ of the residue. For the Neumann boundary control discussed in Section 2.2.2, the algorithm is

$$\begin{cases} (\lambda_h^0, s_h^0) \in \mathbb{R} \times N_h, \\ \lambda_h^{n+1} = \lambda_h^n - a_{nh}\left(\lambda_h^n + \int_{\Sigma_0} p_h^n (2s_h^n - 1)\, d\Sigma_0 \right), \quad n \geq 0, \\ s_h^{n+1} = P_{[0,1]}(s_h^n + b_{nh}\lambda_h^n p_h^n(1, \cdot)), \quad n \geq 0 \end{cases}$$

where p_h^n is an approximation of the solution p of (22).

Let us discuss the Neumann boundary case with $y_0(x) = \sin(\pi x)$, $T = 1/2$, $c = 1/10$, $a := 0$ and $\Sigma_0 = \{1\} \times (0, T)$. The algorithm is stopped as soon as $|\lambda_h^n + \int_0^1 p_h^n(1, t)\,(2s_h^n(t) - 1)\, dt|/|\lambda_h^n| \leq \sigma := 10^{-3}$. Table 2 reports the L^∞-norm λ_h and the $L^2(0, 1)$-norm of the $y_h(\cdot, T)$ for various values of α. The amplitude λ_h increases as $\alpha \to 0$, and is significantly bigger than for the inner case.

Figure 3 depicts the optimal density s_h with respect to the time variable. We observe here that the density is almost everywhere a characteristic function with an increasing number of change of sign as t goes to T^-. Figure 4 (left) depicts the trace in time of the approximate controlled solution y_h and Figure 4 (right) displays y_h along Q_T.

α	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$\|y_h(\cdot,T)\|_{L^2(\Omega)}$	2.96×10^{-1}	1.59×10^{-1}	5.56×10^{-2}	9.31×10^{-3}
λ_h	1.488	10.181	29.121	34.03
$\sharp$ iterates	33	512	6944	20122

Table 2: $(\lambda_h, y_h(T))$ with respect to α in the Neumann boundary case.

Let us insist again on that oscillatory behavior. For $\alpha = 10^{-6}$, Figure 5 depicts the optimal density $s_h(t)$ for $t \in [0.4, 0.5]$ and $t \in [0.48, 0.5]$ respectively. These figures indicate that the frequency of these sign changes increases as t gets closer to T. Moreover, the number of these sign changes increases as $\alpha \to 0$. At the limit in $\alpha \to 0$, we expect a bounded amplitude λ_h but an arbitrarily large number of oscillations near T, in agreement with our observations for the L^2 situation in [20]. The figures also confirm, in agreement with the optimality conditions for $\overline{J}_\alpha$, that the optimal density is almost everywhere a characteristic function, so that no relaxation occurs.
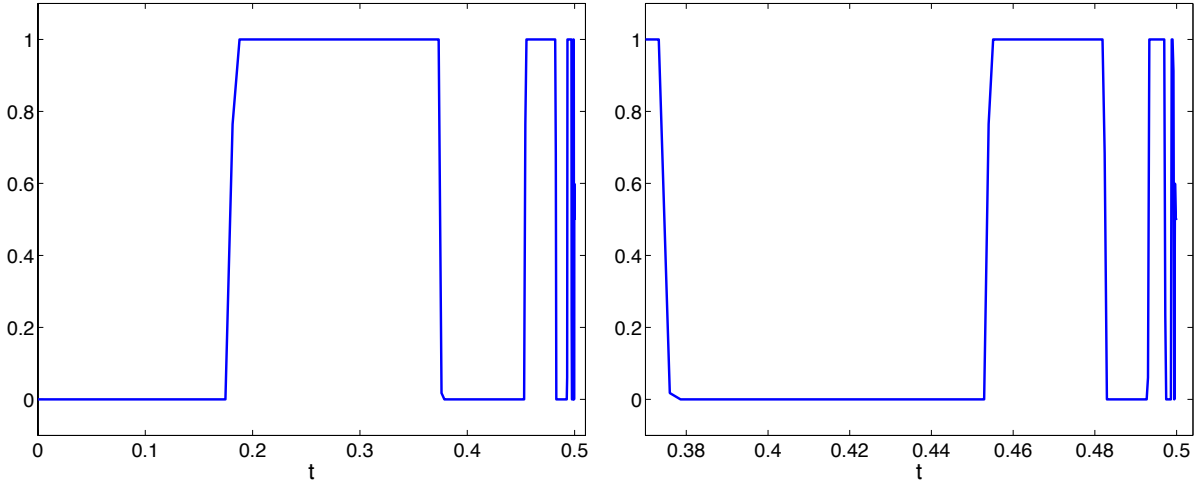
Figure 3: Neumann case - The optimal density $s_h(t)$ for $t \in [0, T]$ - $\alpha = 10^{-4}$.

For Dirichlet boundary control, the situation is similar: we report in Table 3 the optimal amplitude λ_h and $\|y_h(\cdot, T)\|_{L^2(0,1)}$ for several values of α.

α	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$\|y_h(\cdot,T)\|_{L^2(\Omega)}$	9.21×10^{-2}	3.28×10^{-3}	7.31×10^{-4}	2.34×10^{-5}
λ_h	0.98	5.12	7.30	9.02
$\sharp$ iterates	21	319	4912	9301

Table 3: $(\lambda_h, \|y_h(\cdot, T)\|_{L^2(0,1)})$ with respect to α in the Dirichlet boundary case.

3.3   Experiments in 2D

The algorithm developed in the previous section, which is valid for any N, may be modified by taking into account that the relaxed cost is quadratic with respect to λ. Precisely, we note $S := 2s - 1 \in L^\infty(q_T, [-1, 1])$ and remark that the solution y of (1) can be decomposed as $y = \overline{y} + \lambda z_S$, where $\overline{y}$ solves (1) with $v = 0$ and z_S solves
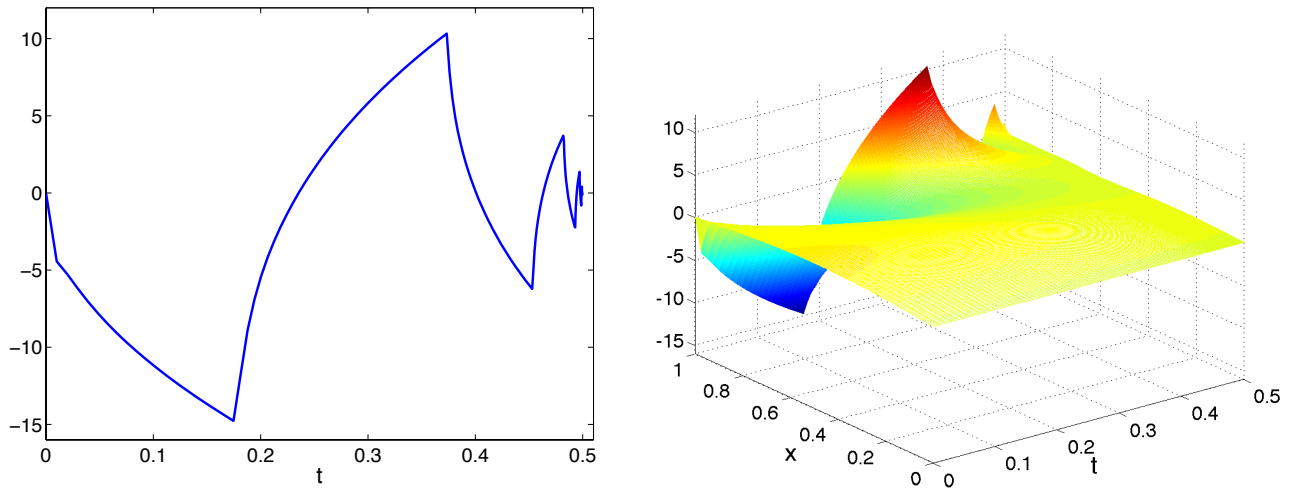
Figure 4: Neumann case - **Left** : $y_h(1,t)$ for $t \in (0,T)$; **Right**: Approximated controlled solution y_h - $\alpha = 10^{-4}$.
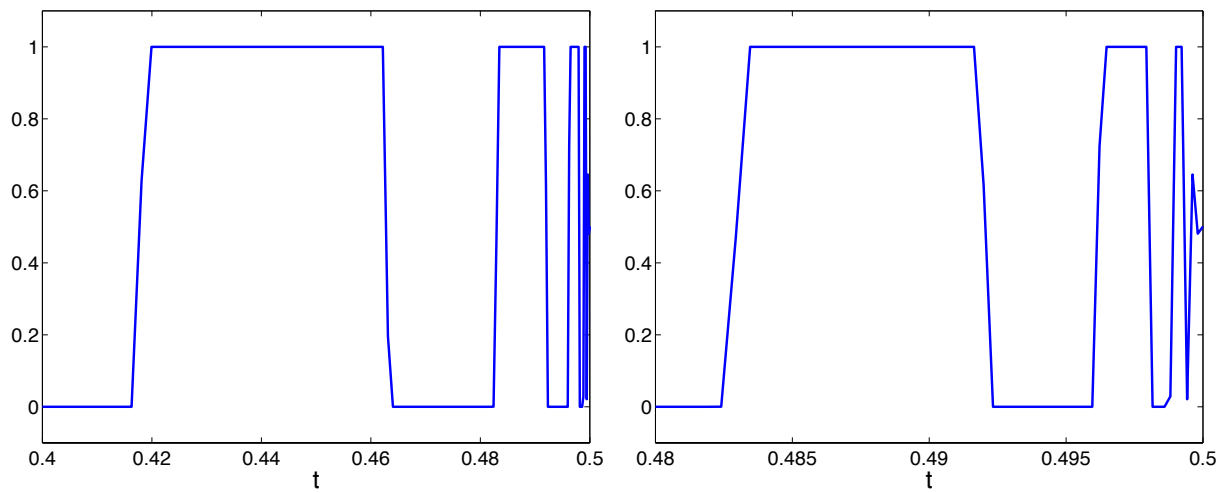

Figure 5: Neumann case -$\alpha = 10^{-6}$- Optimal density for $t \in [0.4, 0.5]$ (left) and for $t \in [0.48, 0.5]$ (right).

$$\begin{cases} z_t - \Delta z + a\,z = S\,1_\omega, & (x,t) \in Q_T \\ z(\sigma, t) = 0, & (\sigma, t) \in \Sigma_T, \qquad z(x, 0) = 0, \quad x \in \Omega. \end{cases}$$

The new expression of the relaxed cost $\overline{J}_\alpha$ is as follows

$$\overline{J}_\alpha(\lambda, S) = \frac{1}{2}\left(\lambda^2 + \alpha^{-1}\|\overline{y}(\cdot, T) + \lambda z_S(\cdot, T)\|^2_{L^2(\Omega)}\right).$$

Proceeding as above, the directional derivative at (λ, S) in the admissible direction $(\hat{\lambda}, \hat{S})$ is given by

$$\frac{\partial \overline{J}_\alpha(\lambda, S)}{\partial(\lambda, S)} \cdot (\hat{\lambda}, \hat{S}) = \hat{\lambda}\left(\lambda + \alpha^{-1}\int_\Omega (\overline{y}(\cdot, T) + \lambda z_S(\cdot, T))z_S(\cdot, T)\,dx\right) - \int_{q_T} p\hat{S}\,dx\,dt$$

where p solves $-p_t - \Delta p + a\,p = 0$ in Q_T, $p(\cdot, T) = -\lambda\alpha^{-1}(\overline{y}(\cdot, T) + \lambda z_S(\cdot, T))$ and $p = 0$ on Σ_T. A gradient algorithm to minimize the functional $\overline{J}_\alpha$ is then

$$\begin{cases} (\lambda^0, S^0) \quad \text{given in} \quad \mathbb{R} \times L^\infty(Q_T, [-1, 1]), \\ \lambda^{n+1} = \lambda^n - a_n\left(\lambda^n + \alpha^{-1}\int_\Omega (\overline{y}(\cdot, T) + \lambda^n z_{S^n}(\cdot, T))z_{S^n}(\cdot, T)\,dx\right), \quad n \geq 0, \\ S^{n+1/2} = S^n + b_n p^n, \quad S^{n+1} = P_{[-1,1]}(S^{n+1/2}), \quad n \geq 0 \end{cases} \tag{26}$$

where (a_n, b_n) denotes the minimizer of the functional $\overline{J}_\alpha(\lambda^{n+1}, S^{n+1/2})$ over $\mathbb{R}^+ \times \mathbb{R}^+$ easily obtained by a Newton type method. This algorithm is simpler than (23) which requires a line search technique to solve (24): it ensures that $\overline{J}_\alpha(\lambda^{n+1}, S^{n+1/2}) \leq \overline{J}_\alpha(\lambda^n, S^n)$. Moreover, we have checked in our numerical experiments that there is decreasing of the cost, that is, $\overline{J}_\alpha(\lambda^{n+1}, S^{n+1}) \leq \overline{J}_\alpha(\lambda^n, S^n)$ for all n. Eventually, from (23), the algorithm is stopped as soon as

$$\left|\lambda^n + \frac{\int_\Omega \overline{y}(\cdot, T)z_{S^n}(\cdot, T)dx}{\alpha + \|z_{S^n}(\cdot, T)\|^2_{L^2(\Omega)}}\right| \leq \sigma|\lambda^n| \tag{27}$$

for a given tolerance σ. Note that, at the miminal point, this leads to an expression of the optimal λ, that is the $L^\infty(q_T)$-norm of the bang-bang control in terms of the optimal distribution S: $\lambda = - <\overline{y}(\cdot, T), z_S(\cdot, T) >_{L^2(\Omega)} /(\alpha + < z_S(\cdot, T), z_S(\cdot, T) >_{L^2(\Omega)})$. The data we consider are the following:

$$\Omega = (0, 1)^2, \quad \omega = (0.2, 0.6)^2, \quad c = 0.1, \quad T = 0.5, \quad y_0(x) = 1_{(0.5, 0.8)^2}(x), \quad a \equiv 0.$$

Moreover, in order to reduce the natural dissipation of the heat from the boundary, we impose a free Neumann boundary condition, that is $\nabla y \cdot \nu = 0$ on Σ_T (instead of a null Dirichlet boundary condition). Up to that change, the relaxed formulation (RT_α) is unchanged. We use P_2 finite elements in space: the finite element discretization of $\overline{\Omega}$ composed of 3694 triangles is drawn in Figure 6. For the time integration, we still use the second order Gear scheme (see Section 3).

Finally, we do not take a too small value for α, precisely $\alpha = 10^{-4}$, in order to restrict the number of oscillations of the density near T, and therefore reduce the computational cost needed to capture them. The algorithm is initialized with $\lambda_h^0 = 1$ and $S_h^0 = 0$ in q_T.

Figure 7 displays the iso-values of the optimal density s_h in ω (defined from S_h by $s_h = (S_h + 1)/2$) for 8 values of $t \in [0, T]$: $t = kT/12$ with k in $\{0, 3, 5, 7, 9, 10, 11, 12\}$. As in the 1D-space situation, we recover that, for any $t \in [0, T)$, $x \to s_h(x, t)$ is a $(0, 1)$-function in ω and that the variations of s_h with respect to t increase abruptly when t becomes close enough to T. In particular, for $t = T$, the density we obtain has some important variations and takes some values in $(0, 1)$. Very likely, a very fine mesh together with a large number of iterates is needed to recover a bi-valued function. Table 4 collects some numerical values obtained after 3012 iterates of the algorithm (26) with $\sigma = 10^{-3}$ in (27). We observe in particular
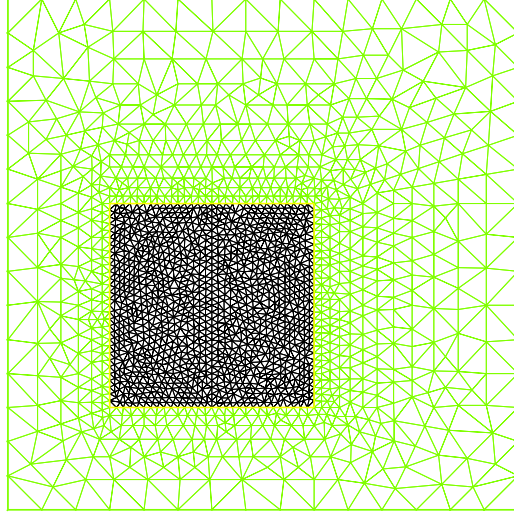
Figure 6: Mesh of $\overline{\Omega}$ - 3694 triangles - P_2 finite elements - 7449 degres of freedom.

that the final state satisfies $\|y_h(\cdot, T)\|_{L^2(\Omega)} \approx 5.06 \times 10^{-2}$, of the order of $\sqrt{\alpha}$. In the uncontrolled case, we obtain $\|y_h(\cdot, T)\|_{L^2(\Omega)} \approx 1.11 \times 10^{-1}$. The gap between the controlled and uncontrolled case seems rather small but is not: with Neumann free boundary condition, recall that, in the uncontrolled case, the heat equation is not dissipative and 0 is not a free trajectory (in the uncontrolled case, we get that $\lim_{T \to \infty} \|y(\cdot, T)\|_{L^2(\Omega)} \approx 9.98 \times 10^{-2} \neq 0$). The evolution of the norm $\|y_h(\cdot, t)\|_{L^2(\Omega)}$ with respect to the time variable is given in Figure 8. As is classical for the L^2-situation, we check that the bang-bang control v almost drives to rest the solution on q_T: we get $\|y_h(\cdot, T)\|_{L^2(\omega)} \approx 8.1 \times 10^{-4}$ of the order of α. The iso-values of $y_h(x, T)$ in Ω are given in Figure 9.

α	$\overline{J}_{\alpha,h}(\lambda_h, S_h)$	λ_h	$\|y_h(\cdot, T)\|_{L^2(\Omega)}$	$\sharp$ iterate	CPU time
10^{-4}	18.177	3.265	5.06×10^{-2}	3012	$1729s$

Table 4: Direct method - Numerical results.

3.4 Comparison with the dual approach

We now compare in the 2D case our method with the one introduced and developed in [9]. We consider the inner controllability case which is not numerically discussed in [9]. The dual reformulation of the bang-bang controllability problem leads to the extremal problem

$$\min_{\varphi_{\alpha,T} \in L^2(\Omega)} \mathcal{J}_\alpha(\varphi_{\alpha,T}) = \frac{1}{2}\|\varphi\|^2_{L^1(q_T)} + \frac{\alpha}{2}\|\varphi_{\alpha,T}\|^2_{L^2(\Omega)} + \int_\Omega y_0(x)\, \varphi(x, 0)\, dx$$

where $(\varphi, \varphi_{\alpha,T})$ solves (4). The optimal bang-bang control is then given by (3). For any $\alpha > 0$, the minimizer of $\mathcal{J}_\alpha$ is approximated by the way of the conjugate gradient algorithm : contrarily to the L^2-situation (where the term $\|\varphi\|^2_{L^1(q_T)}$ is replaced by $\|\varphi\|^2_{L^2(q_T)}$, see [9, 20]), the optimal descent step at each iteration is not explicit but may be approximated numerically by the secant method.

The same discretization and data (as in the preceding section) are used and the algorithm is initialized with $\varphi^0_{T,h} = 0$ in Ω. Table 5 collects some numerical values obtained with the dual method obtained
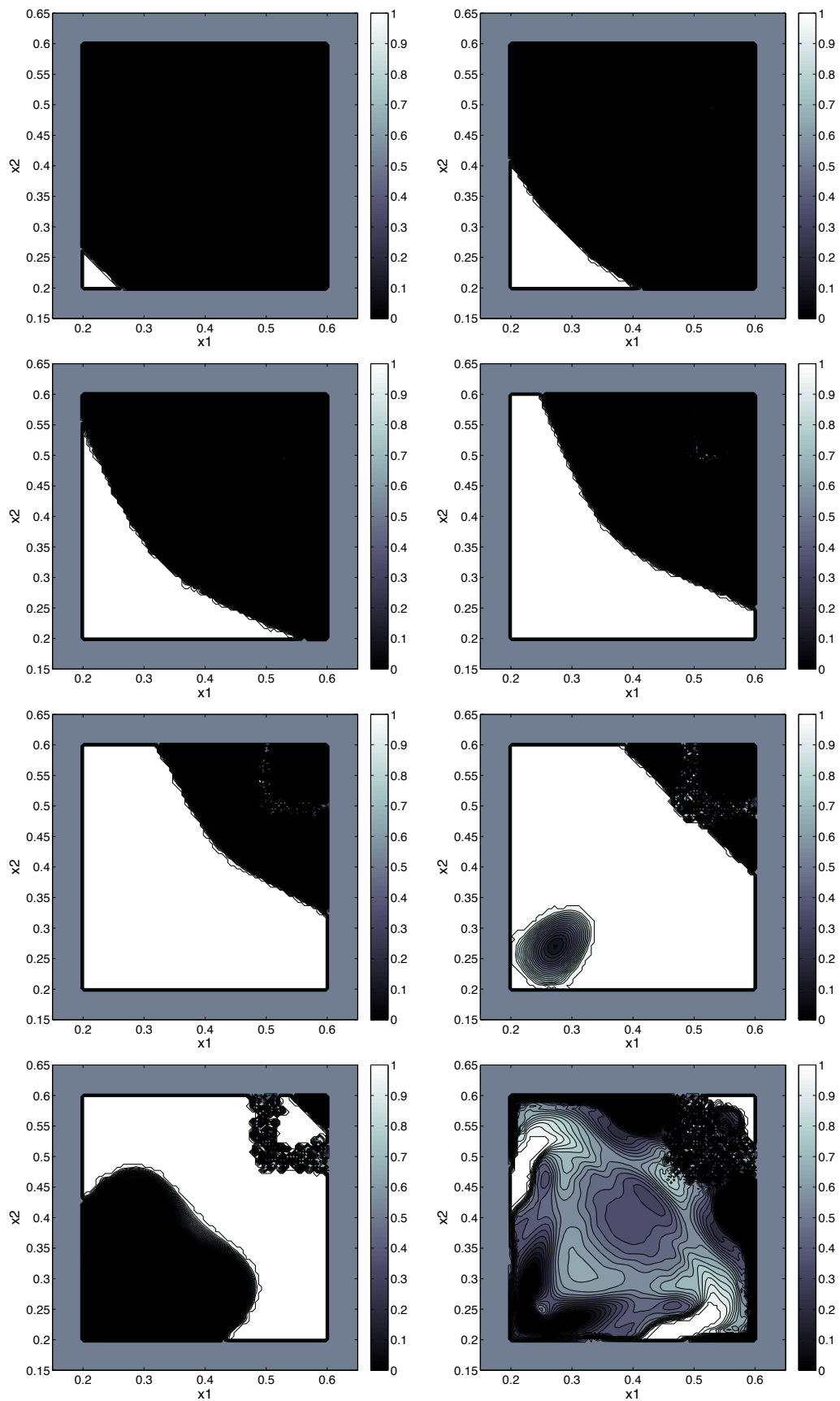
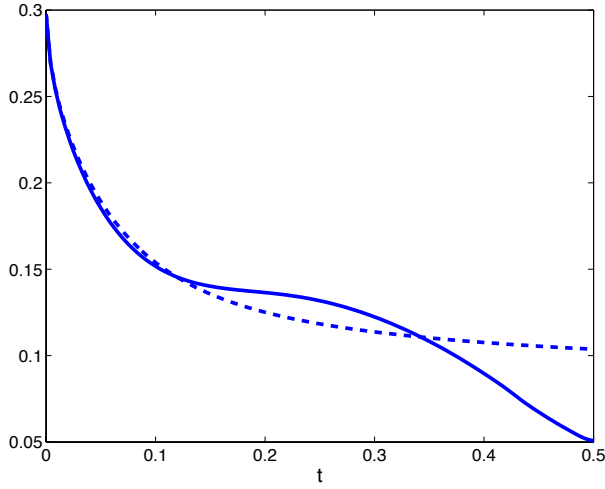Figure 7: Iso-values of the optimal density for $t = \frac{kT}{12}$ with $k = 0, 3, 5, 7, 9, 10, 11, 12$.

Figure 8: Evolution of $\|y_h(\cdot, t)\|_{L^2(\Omega)}$ vs. $t \in (0, T)$ in the controlled (full line) and uncontrolled (dashed line) cases.
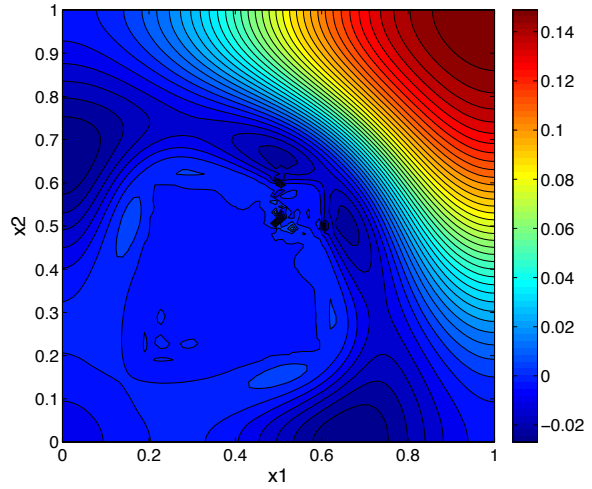
Figure 9: Iso-values of $y_h(\cdot, T)$ on Q_T.

after 2000 iterates. Comparing with Table 4, the dual method gives similar values[1]. At a first glance, the direct method seems to provide a slightly better result since the value of the relaxed function $\overline{J}_h$ is smaller than $-\mathcal{J}_h^\star$: actually, - still with $\alpha = 10^{-4}$ - when we push the algorithm (26) with a very fine mesh (9455 triangles and 31934 degrees of freedom) and $\sigma = 10^{-6}$ in (27), we obtain some intermediate values between Table 4 and Table 5: $\lambda_h \approx 3.298$ and $\overline{J}_h(\lambda_h, S_h) \approx 18.26$. Observe also that the methods provide the same value for the norm $\|y_h(\cdot, T)\|_{L^2(\Omega)}$ as the variation of the two costs, near the optimum, is very low with respect to S_h and $sign(p_h)$ respectively. The main qualitative difference is found in the behavior of the residue with respect to the iterates, reported in Figures 10. The dual method leads to a typical evolution when an ill-posed problem is considered (Figure 10-right): the algorithm needs few iterates to recover the low frequencies of the adjoint solution $\varphi_{T,\alpha}$ (leading to a fast decrease of the residue) then somehow stagnates - as the consequence of the ill-posedness - to recover the high frequencies (we refer to [4, 20] for a detailed discussion on that issue for the L^2-case). On the other hand, the evolution of the residue corresponding to the direct method (Figure 10-left) is much smoother, as the consequence of the relaxation procedure.

α	$J_h^\star(\varphi_{T,h})$	$\|\varphi_h\|_{L^1(q_T)}$	$\|y_h(\cdot, T)\|_{L^2(\Omega)}$	$\sharp$ iterate	CPU time
10^{-4}	-18.327	3.317	5.06×10^{-2}	2000	$1504s$

Table 5: Dual method - Numerical results.

Finally, it is important to emphasize that although the method introduced in this paper gives similar numerical results as the dual one, however the scope of both approaches is different. Indeed, if we look at the problem of computing bang-bang type controls with minimal amplitude (which is of practical interest since bang-bang controls are easier to implement than more general L^∞ controls) for other parabolic systems (where it is not known if the control of minimal L^∞-norm is of bang-bang type) then the relaxation method introduced in the present paper applies, but the one based of computing optimal controls of minimal L^∞-norm does not because a priori we don't know that optimal controls in L^∞-norm

---

[1]Recall that from the Legendre-Fenchel theorem, the following holds : $\inf_{v \in L^\infty(q_T)} J_\alpha(v_\alpha) = -\inf_{\varphi_T \in L^2(\Omega)} \mathcal{J}_\alpha^\star(\varphi_T)$. This explains the minus sign in Table 5.
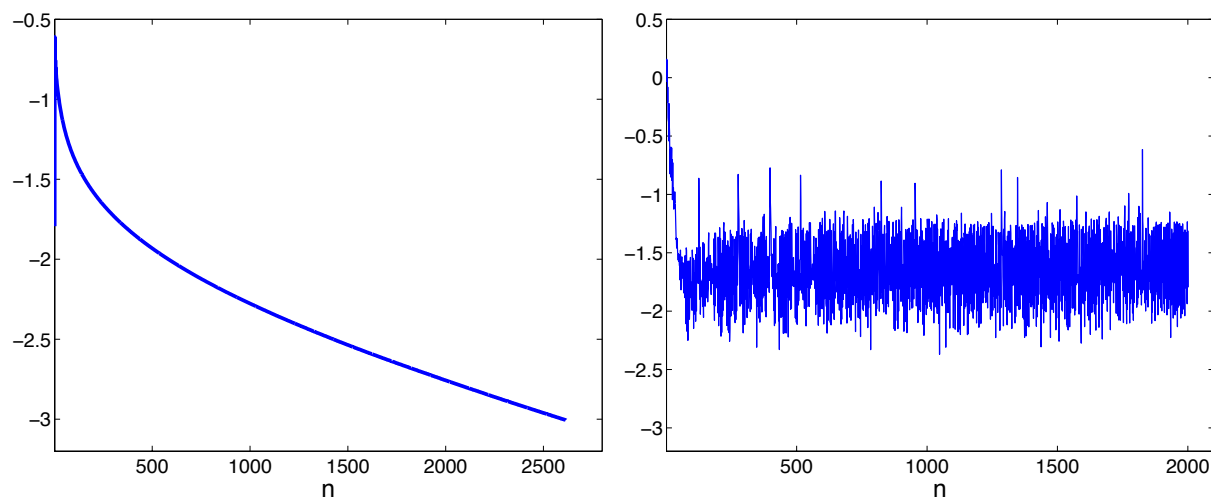
Figure 10: Evolution of the relative residue (in $\log_{10}$ scale) with respect to the iterates : Direct method (**Left**) and dual method (**Right**).

are of bang-bang type. For the contrary, if we are interested in computing controls with minimal L^∞-norm for a given parabolic system where it is unknown whether these controls are of bang-bang type, then the approach of this paper does not apply. As a conclusion, in these two scenarios there is not possibility for comparison because the methods introduced by Glowinski et al. and the one described here address different problems. Nevertheless, the above comparison with the dual method provides a useful information concerning the accuracy of the proposed method.

4   Concluding remarks

In this paper, a reformulation of the L^∞-controllability problem for the heat equation in terms of an optimal design one has been proposed. This new formulation, which needs of relaxation, leads to a simple descent algorithm for the numerical resolution of the original problem. In spite of the well-posedness character of the problem, optimal bang-bang controls highly oscillate near the final time. While the amplitude of approximate controls is bounded by above with respect to the penalty parameter α, the number of these oscillations is not. This feature, closely related to the heat kernel regularization property, renders severally ill-posed the numerical approximation of the null control of L^∞ norm. In this respect, the minimization of the relaxed cost is not easier than the minimization of the conjugate function (derived from duality arguments) commonly used. The subtle difference, strongly enhanced for small values of α, comes from the fact that the density s belongs to L^∞ while the dual variable degenerates in an abstract completed space, possibly larger than L^∞.

When the density is known to be almost everywhere a characteristic function, others methods may be used to solve the initial problem (5). For instance, in the boundary case of Section 2.2, one may write the density $s \in L^\infty([0, T]; \{-1, 1\})$ as the sum of characteristic functions and make use of algorithms from constrained optimization: such strategy is much more intricate when the density depends on several variables. In that case, one may use a level set approach (used in optimal shape design) which requires the resolution of a nonlinear Hamilton-Jacobi equation of first order (in our controllability context, we refer to [17]). As already mentioned in Remark 1, the approach we have used here, based on the introduction of a relaxed problem and justified mathematically, is easy to implement in any dimension.

This work may be extended in several ways. Let us first mention the semi-linear situation where the state equation of (1) is replaced by

$$y_t - \Delta y + f(y) = v \, 1_\omega \tag{28}$$

and for which positive controllability results exist (see [5]). A fixed point strategy reduces the search of nonlinear bang-bang controls for (28) to the construction of sequence of controls for (1). Consequently, our approach may be used here. We refer to [6] for a complete study in the L^2 case.

Also consider the following optimal control problem, closely related to $(\mathcal{P}_\alpha)$: for any fixed $\lambda > 0$

$$\min_{v \in L^\infty(q_T; [-\lambda, \lambda])} \|y(\cdot, T)\|_{L^2(\Omega)}$$

where $y = y(v)$ solves (1) . It turns out that the optimal control is of bang-bang form (we refer to [21], p.133) so that our method may be used as well.

Finally, let us mention the wave type equation where the situation is different: in that case, existence of bang-bang control does not hold in general, as it depends on the initial data (see [11]). This means that, by applying the same procedure, some data may exhibit relaxation, that is optimal densities taking values in (0,1). See also [10] where the controllability of the wave equation was studied by using penalization techniques.

References

[1] G. Alessandrini and L. Escauriaza, *Null-controllability of one-dimensional parabolic equations*, ESAIM Control Optim. Calc. Var. 14 (2008), no. 2, 284–293.

[2] S. Angenent, *The zero set of a solution of a parabolic equation*, J. reine angew. Math. 390 (1988), 79–96.

[3] F. Ben Belgacem and S.M. Kaber, *On the Dirichlet boundary controllability of the 1-D heat equation: semi-analytical calculations and ill-posedness degre*, Inverse Problems 27 (2011), no. 5.

[4] C. Carthel, R. Glowinski and J.-L. Lions, *On exact and approximate Boundary Controllabilities for the heat equation: A numerical approach*, J. Optimization, Theory and Applications 82 (3), (1994) 429–484.

[5] C. Fabre, J.-P. Puel and E. Zuazua, *Approximate controllability of the semilinear heat equation*, Proc. Roy. Soc. Edinburgh Sect. A 125 (1995), no. 1, 31–61.

[6] E. Fernandez-Cara and A. Münch, *Numerical null controllability of semi-linear 1D heat equations: fixed point, least squares and Newton methods*, Mathematical Control and Related Fields, 3 (2), 217-246 (2012).

[7] E. Fernandez-Cara and A. Münch, *Strong convergent approximations of null controls for the 1D heat equation*, SéMA journal, 61(1) (2013) 49-78.

[8] A.V. Fursikov and O. Yu. Imanuvilov, *Controllability of Evolution Equations,* Lecture Notes Series, number 34. Seoul National University, Korea,(1996) 1–163.

[9] R. Glowinski, J-L. Lions and J. He, *Exact and approximate controllability for distributed parameter systems. A numerical approach.* Encyclopedia of Mathematics and its Applications, 117. Cambridge University Press, Cambridge, 2008.

[10] M. Gugat, *Penalty techniques for state constrained optimal control problems with the wave equation*, SIAM J. Control Optim. 48 (5), (2009) 3026–3051.

[11] M. Gugat and G. Leugering, L^∞-*norm minimal control of the wave equation: on the weakness of the bang-bang principle*, ESAIM: COCV 14 (2008) 254-283.

[12] A. Henrot and M. Pierre, *Variation et optimisation de formes: une analyse géometrique*, in: Mathématiques et applications, 48, Springer, 2005.

[13] G. Lebeau and L. Robbiano, *Contrôle exact de l'équation de la chaleur,* Comm. Partial Differential Equations 20 (1995), no. 1–2,

[14] J. L. Lions, Contrôle optimal des systèmes gouvernés par des équations aux dérivés partielles, Dunod, collection E.M., 1968.

[15] J. L. Lions and E. Magenes, Problèmes aux limites non homogènes et applications. Vol. I and II, Dunod, 1968.

[16] S. Micu and E. Zuazua, *Regularity issues for the null-controllability of the linear 1-d heat equation.* C. R. Acad. Sci. Paris, Ser. I 349 (2011) 673-677.

[17] A. Münch, *Optimal design of the support of the control for the 2-D wave equation : numerical investigations*, International Journal of Numerical Analysis and Modeling, 5(2), 331-351 (2008).

[18] A. Münch, P. Pedregal and F. Periago, *Relaxation of an optimal design problem for the heat equation*, J. Math. Pures Appl. 89 (2008) 225-247.

[19] A. Münch and F. Periago, *Optimal distribution of the internal null control for the one-dimensional heat equation*, J. Diff. Eq. 250 (2011) 95-111.

[20] A. Münch and E. Zuazua, *Numerical approximation of null controls for the heat equation: Ill-posedness and remedies,* Inverse Problems 26 (8) 2010.

[21] F. Tröltzsch, Optimal Control of Partial Differential Equations. Theory, Methods and Applications, Graduate Studies in Mathematics, Vol. 112, AMS, 2010.

[22] G. Wang and E. Zuazua, *On the Equivalence of Minimal Time and Minimal Norm Controls for Internally Controlled Heat Equations*, SIAM J. Control Optim. 50 (5) (2012), 2938?2958.

[23] E. Zuazua, *Controllability and observability of partial differential equations: some results and open problems*, In Handbook of Differential Equations: Evolutionary Differential Equation, vol 3, C.M. Dafermos and E. Feireisl eds, Elsevier Science, 527–621 (2006).