Numerical exact controllability of the 1D heat equation: dual algorithms

Enrique Fernández-Cara* and Arnaud Münch†

Abstract

This work addresses the numerical approximation of distributed null controls (with small support) for the 1D heat equation, with Dirichlet boundary conditions: the goal is to compute a control that drives the solution from a prescribed initial state at $t = 0$ to zero at $t = T$. The earlier contribution by Carthel, Glowinski and Lions [3] considers boundary controls and exhibits numerical instability, which is closely related to the regularization effect of the heat equation. There, the controllability problem is solved through a dual reformulation. In this work, we deal with other related methods. We first introduce some constrained extremal problems (each of them corresponding to the minimization of a functional that involves weighted integrals of the state and the control) and we apply appropriate duality techniques that lead to the formulation of equivalent unconstrained extremal problems. Then, we introduce appropriate numerical approximations of the associated dual problems and we solve them by applying conjugate gradient techniques. We present several experiments and we discuss the robustness of this approach with respect to the control support. We also compare the results to those in the previous paper [7], where primal methods (directly connected to the original constrained problem) were considered.

Keywords: Heat equation, null controllability, numerical solution, duality.

Mathematics Subject Classification: 35L05, 49J05, 65K10.

1 Introduction

We are concerned in this work with the null controllability problem for the 1D heat equation.

The state equation is the following:

$$\begin{cases} y_t - (a(x)y_x)_x = v1_\omega\,, & (x,t) \in (0,1)\times(0,T) \\ y(x,t) = 0, & (x,t) \in \{0,1\}\times(0,T) \\ y(x,0) = y_0(x), & x \in (0,1). \end{cases} \tag{1}$$

Here, $\omega \subset\subset (0,1)$ is a (small) non-empty open interval, 1_ω is the associated characteristic function, $T > 0$, $a \in L^\infty(0,1)$ with $a(x) \geq a_0 > 0$ a.e., $y_0 \in L^2(0,1)$, $v \in L^2(\omega\times(0,T))$ is the

*Dpto. EDAN, University of Sevilla, Aptdo. 1160, 41080 Sevilla, Spain. E-mail: cara@us.es. Partially supported by grant MTM2006-07932 (Spain).

†Laboratoire de Mathématiques, Université Blaise Pascal (Clermont-Ferrand 2), UMR CNRS 6620, Campus des Cézeaux, 63177 Aubière, France. E-mail: arnaud.munch@math.univ-bpclermont.fr. Partially supported by grant ANR-07-JC-183284 (France).

control and y is the state. In the sequel, for any $\tau > 0$, we will denote by Q_τ, Σ_τ and q_τ the sets $(0,1)\times(0,\tau)$, $\{0,1\}\times(0,\tau)$ and $\omega\times(0,\tau)$, respectively.

It is well known that, for every $y_0 \in L^2(0,1)$, $T>0$ and $v \in L^2(q_T)$, there exists exactly one solution y to (1), with

$$y \in C^0([0,T];L^2(0,1)) \cap L^2(0,T;H_0^1(0,1)).$$

The null controllability problem for (1) at time T is the following:

For any $y_0 \in L^2(0,1)$, find $v \in L^2(q_T)$, such that the solution y to (1) satisfies

$$y(x,T) = 0, \quad x \in (0,1). \tag{2}$$

The controllability of differential systems is an important area of research and has been the subject of many papers in recent years. Some relevant references concerning the controllability of partial differential equations are [4, 13, 14] and [18]. From the practical viewpoint, null controllability results are crucial in control theory since, roughly speaking, they make it possible to drive the solution to rest and, consequently, are associated to finite time work.

In the particular case of the 1D heat equation in (1), the following result holds:

THEOREM 1.1 (ALESSANDRINI AND ESCAURIAZA [1]) *The linear system (1) is null controllable at any time $T>0$. In other words, for each $y_0 \in L^2(\Omega)$, there exists $v \in L^2(q_T)$ such that the associated solution of (1) satisfies (2).*

The goal of this paper is to design and analyze numerical methods allowing to solve the previous null controllability problem. This issue is challenging and has already been the objective of some important work; see for instance [2, 3, 7, 10, 11, 12, 16] and [19]. In particular, imposing some regularity properties to a and following [8], in [7] we have considered and solved numerically the extremal problem

$$\begin{cases} \text{Minimize} \quad J(y,v) = \dfrac{1}{2}\displaystyle\iint_{Q_T} \rho^2 |y|^2 \, dx\, dt + \frac{1}{2}\iint_{q_T} \rho_0^2 |v|^2 \, dx\, dt \\ \text{Subject to} \quad (y,v) \in \mathcal{C}(y_0,T) \end{cases} \tag{3}$$

where $\mathcal{C}(y_0,T)$ is the set

$$\mathcal{C}(y_0,T) = \{\, (y,v) : v \in L^2(q_T), \quad y \text{ solves (1) and satisfies (2)} \,\}$$

and the weights ρ and ρ_0 were chosen appropriately; see (7).

In fact, in order to find a solution to (3), we can use methods of two kinds:

- Primal methods, that provide an optimal couple (y,v) satisfying the constraint $(y,v) \in \mathcal{C}(y_0,T)$ and usually rely on the characterization of optimality. This was the class of methods considered in [7].
- Dual methods, that rely on appropriate reformulations and use new (dual) variables. The design and analysis of methods of this kind is the main objective of this paper.

An important previous contribution was made by Carthel, Glowinski and Lions in [3] (see also [9, 10]), dealing with approximate controllability issues for the $L^2(q_T)$-norm, that is, with $\rho \equiv 0$, $\rho_0 \equiv 1$ in (3) and $\mathcal{C}(y_0,T)$ replaced by

$$\mathcal{C}_\varepsilon(y_0,T) = \{(y,v) : v \in L^2(q_T), y \text{ solves (1) and } \|y(\cdot\,,T)\|_{L^2} \le \varepsilon\}$$

for small $\varepsilon > 0$. In view of the difficulty to construct pairs in $\mathcal{C}_\varepsilon(y_0, T)$ (and thus minimizing sequences for J), in [3] the authors use duality arguments and *Fenchel-Rockafellar* theory (see [6, 17]) in order to reformulate (3) in the form

$$\begin{cases} \text{Minimize} \quad J_\varepsilon^*(\varphi_T) = \dfrac{1}{2}\displaystyle\iint_{q_T} \varphi^2 \, dx\, dt + \int_0^1 \varphi(x,0)\, y_0(x)\, dx + \varepsilon \|\varphi_T\|_{L^2(0,1)} \\ \text{Subject to} \quad \varphi_T \in L^2(0,1) \end{cases} \tag{4}$$

where φ solves the homogeneous adjoint of (1) with "initial" condition φ_T at time T; see (21).

The difficulty to solve numerically (4) is related to the non-uniform coercivity of J_ε^* in $L^2(0,1)$ with respect to ε, which is in turn closely related to the regularizing property of the heat operator. At the limit $\varepsilon \to 0$, the coerciveness only holds in the space $\mathcal{H}$, defined as the completion of $\mathcal{D}(0,1)$ for the norm $\|\varphi_T\|_{\mathcal{H}} := \|\varphi\|_{L^2(q_T)}$. The space $\mathcal{H}$ is much larger than $L^2(0,1)$ and cannot be approximated easily by finite dimensional numerically reasonable spaces.

The numerical results in [3] and [10] show that, as $\varepsilon \to 0$, the minimizers $\varphi_{T,\varepsilon}$ become singular, which leads a highly oscillatory behavior of the control near $t = T$. Hence, the computation of approximations to the null control of minimal norm in $L^2(0,1)$ is severally ill-posed.

In [7], the weights ρ and ρ_0 are chosen as in related global Carleman inequalities, see [8]. In particular, they blow up exponentially as $t \to T^-$. This property, which can be seen as a reinforcement of the condition (2), ensures the well-posedness of the variational formulation associated to primal methods. Accordingly, it seems interesting to extend the techniques in [3] to more general (weighted) functionals and, then, to analyze the influence of the weights on the robustness of the corresponding dual problem. This is the main aim of this paper.

The paper is organized as follows. In Section 2, we briefly overview some previous results concerning the solution of (3) based on necessary optimality conditions. In Section 3, we apply results from the *Fenchel-Rockafellar* duality theory to (3). To this end, we first introduce some approximations to (3) that lead to well-posed dual problems (see Proposition 3.2). Then, we show that the solutions to the latter converge, in an appropriate sense, to the solution to the original problem (3) (see Propositions 3.1 and 3.3). In Section 4, we show how gradient methods can be applied in practice in this framework. More precisely, Subsection 4.1 is concerned with a conjugate gradient type algorithm, while Subsection 4.2 deals with the finite dimensional approximation of the control problems. In Section 5, we present several numerical experiments that show that the behavior of the considered algorithms is appropriate. Finally, some further comments and concluding remarks are given in Section 6.

2 Preliminary results

It will be assumed in the sequel that

$$a \in C^1([0,1]), \quad a(x) \geq a_0 > 0 \quad \forall x \in [0,1]. \tag{5}$$

Moreover, we will use the following notation

$$Lz = z_t - (a(x)z_x)_x\,, \quad L^*q = -q_t - (a(x)q_x)_x\,.$$

Very frequently, we will use the symbol C to design a generic positive constant, usually depending on ω, T, a_0, $\|a\|_{C^1([0,T])}$ and maybe other data.

Let ρ and ρ_0 be functions satisfying

$$\begin{cases} \rho = \rho(x,t),\ \rho_0 = \rho_0(x,t) \text{ are continuous and } \geq \rho_* > 0 \text{ in } Q_T \\ \rho, \rho_0 \in L^\infty(Q_{T-\delta}) \quad \forall \delta > 0 \end{cases} \tag{6}$$

Observe that, in principle, these functions can blow up as $t \to T^-$.

Let us consider the extremal problem (3). Then we have the following:

THEOREM 2.1 *For any $y_0 \in L^2(0,1)$ and any $T > 0$, there exists exactly one solution $(\hat{y}, \hat{v})$ to (3).*

The proof relies on the facts that $\mathcal{C}(y_0, T)$ is a non-empty closed convex set of $L^2(Q_T) \times L^2(q_T)$ and $(y, v) \mapsto J(y, v)$ is strictly convex, proper and lower semi-continuous on the space $L^2(Q_T) \times L^2(q_T)$ (the details can be found in [7]).

In [7], the following ρ and ρ_0 are considered:

$$\begin{cases} \rho(x,t) = \exp\left(\dfrac{\beta(x)}{T-t}\right), \quad \rho_0(x,t) = (T-t)^{3/2}\rho(x,t), \quad \beta(x) = K_1\left(e^{K_2} - e^{\beta_0(x)}\right) \\ \text{where the } K_i \text{ are sufficiently large positive constants (depending on } T,\ a_0 \text{ and } \|a\|_{C^1}) \\ \text{and } \beta_0 \in C^\infty([0,1]),\ \beta_0 > 0 \text{ in } (0,1),\ \beta_0(0) = \beta_0(1) = 0,\ |\beta_0'| > 0 \text{ in } (0,1) \setminus \omega. \end{cases} \tag{7}$$

These weight functions blow up at $t = T$ and provide a very suitable solution to the original null controllability problem. They were determined and systematically used by Fursikov and Imanuvilov in [8].

Let us introduce the linear space $P_0 = \{\, q \in C^2(\overline{Q}_T) : q = 0 \ \text{ on } \ \Sigma_T \,\}$. Then the bilinear form

$$(p,q)_P := \iint_{Q_T} \rho^{-2} L^* p\, L^* q \, dx\, dt + \iint_{q_T} \rho_0^{-2} p\, q \, dx\, dt$$

is a scalar product in P_0. Indeed, if we have $q \in P_0$, $L^*q = 0$ in Q_T and $q = 0$ in q_T then, by the well known *unique continuation property,* we must have $q \equiv 0$.

Let us denote by P be the completion of P_0 for this scalar product. Then P is a Hilbert space and the *primal* approach to the numerical solution of (3) that we have followed in [7] relies on the following proposition:

PROPOSITION 2.1 *Assume that a satisfies (5) and let ρ and ρ_0 be given by (7). Let (y, v) be the corresponding optimal pair. Then there exists $p \in P$ such that*

$$y = \rho^{-2} L^* p, \quad v = -\rho_0^{-2} p|_{q_T}\,. \tag{8}$$

The function p is the unique solution of

$$\begin{cases} \displaystyle\iint_{Q_T} \rho^{-2} L^* p\, L^* q \, dx\, dt + \iint_{q_T} \rho_0^{-2} p\, q \, dx\, dt = \int_0^1 y_0(x)\, q(x,0)\, dx \\ \forall q \in P; \quad p \in P. \end{cases} \tag{9}$$

The well-posedness of (9) in weighted L^2 spaces is a consequence of the following two lemmas, respectively proved in [8] and [7]:

LEMMA 2.1 *Assume that a satisfies (5) and let ρ and ρ_0 be given by (7). Let us also set*

$$\rho_1(x,t) = (T-t)^{1/2}\rho(x,t), \quad \rho_2(x,t) = (T-t)^{-1/2}\rho(x,t).$$

Then there exists $C > 0$, only depending on ω, T, a_0 and $\|a\|_{C^1}$, such that

$$\left\{ \begin{array}{l} \displaystyle\iint_{Q_T} \left[\rho_2^{-2}\left(|q_t|^2 + |q_{xx}|^2\right) + \rho_1^{-2}|q_x|^2 + \rho_0^{-2}|q|^2\right] dx\,dt \\ \qquad\qquad \displaystyle\leq C\left(\iint_{Q_T} \rho^{-2}|L^*q|^2\,dx\,dt + \iint_{q_T} \rho_0^{-2}|q|^2\,dx\,dt\right) \end{array} \right. \tag{10}$$

for all $q \in P$.

LEMMA 2.2 *Let the assumptions of Lemma 2.1 hold. Then, for any $\delta > 0$, one has*

$$P \hookrightarrow C^0([0, T-\delta]; H_0^1(0,1)),$$

where the embedding is continuous. In particular, there exists $C > 0$, only depending on ω, T, a_0 and $\|a\|_{C^1}$, such that

$$\|q(\cdot\,,0)\|^2_{H_0^1(0,1)} \leq C\left(\iint_{Q_T} \rho^{-2}|L^*q|^2\,dx\,dt + \iint_{q_T} \rho_0^{-2}|q|^2\,dx\,dt\right) \tag{11}$$

for all $q \in P$.

In the primal approach, the controllability problem is reduced to the numerical resolution of the variational equality (9) in P. This can be viewed as a boundary value problem for a PDE that is fourth order in space and second order in time.

More precisely, the goal is to find a numerical approximation of the solution to the following linear problem:

$$\left\{ \begin{array}{ll} L(\rho^{-2}L^*p) + \rho_0^{-2}p\,1_\omega = 0, & (x,t) \in (0,1)\times(0,T) \\ p(x,t) = 0, \quad (-\rho^{-2}L^*p)(x,t) = 0 & (x,t) \in \{0,1\}\times(0,T) \\ (-\rho^{-2}L^*p)(x,0) = y_0(x), \quad (-\rho^{-2}L^*p)(x,T) = 0, & x \in (0,1). \end{array} \right. \tag{12}$$

Once p is known, the optimal pair (y, v) is found using (8).

In this paper, we will consider other methods, that rely on a different viewpoint.

Roughly speaking, in accordance with the previous contribution of Carthel, Glowinski and Lions [3], we will try to use the *Fenchel-Rockafellar* duality approach to convex optimization (see [6, 17]; see also [5]) in order to reduce the task to the solution of some *unconstrained* problems with good properties.

3 Approximation and duality

The main reason for using duality in the context of (3) is that it is difficult to construct minimizing sequences; in fact, it is already difficult to construct couples (y, v) in $\mathcal{C}(y_0, T)$.

However, it is not easy to apply the Fenchel-Rockafellar techniques to (3) directly, mainly because ρ blows up as $t \to T^-$; recall that the problem considered in [3] corresponds to $\rho \equiv 0$ and $\rho_0 \equiv 1$. Consequently, we will first work with well chosen approximations depending on appropriate parameters and, then, we will try to see what happens in the limit.

In some sense, the arguments we are going to present generalize those in [3] and [9]. It will be shown below that, from the viewpoint of numerical results, this generalized framework is more appropriate.

For each $R > 0$, we will first consider the following problem:

$$\begin{cases} \text{Minimize} \quad J_R(y,v) = \dfrac{1}{2}\displaystyle\iint_{Q_T} \rho_R^2 |y|^2 \, dx\, dt + \dfrac{1}{2}\iint_{q_T} \rho_0^2 |v|^2 \, dx\, dt \\ \text{Subject to} \quad (y,v) \in \mathcal{C}(y_0, T). \end{cases} \tag{13}$$

Here, we have used the notation $\rho_R = T_R(\rho) := \min(\rho, R)$. Notice that (13) is a new constrained extremal problem; again, it possesses exactly one solution (y_R, v_R).

PROPOSITION 3.1 *For any $R > 0$, let (y_R, v_R) be the unique minimizer of J_R in $\mathcal{C}(y_0, T)$. Then*

$$v_R \to \hat{v} \quad \textit{strongly in } L^2(q_T) \quad \textit{and} \quad y_R \to \hat{y} \quad \textit{strongly in } L^2(Q_T) \tag{14}$$

as $R \to +\infty$, where $(\hat{y}, \hat{v})$ is the unique solution to (3).

PROOF: First, notice that

$$J_R(\hat{y}, \hat{v}) = \frac{1}{2}\left(\iint_{Q_T} \rho_R^2 |\hat{y}|^2 \, dx\, dt + \iint_{q_T} \rho_0^2 |\hat{v}|^2 \, dx\, dt\right) \le J(\hat{y}, \hat{v})$$

for all $R > 0$. Consequently, the solutions to the problems (13) satisfy

$$J_R(y_R, v_R) = \frac{1}{2}\left(\iint_{Q_T} \rho_R^2 |y_R|^2 \, dx\, dt + \iint_{q_T} \rho_0^2 |v_R|^2 \, dx\, dt\right) \le J(\hat{y}, \hat{v}).$$

This shows that $\rho_R y_R$ is uniformly bounded in $L^2(Q_T)$ and $\rho_0 v_R$ is uniformly bounded in $L^2(q_T)$. Therefore, at least for some subsequence one has

$$\rho_0 v_R \to w \quad \text{weakly in } L^2(q_T) \quad \text{and} \quad \rho_R y_R \to z \quad \text{weakly in } L^2(Q_T). \tag{15}$$

Let us set $\tilde{y} = \rho^{-1} z$ and $\tilde{v} = \rho_0^{-1} w$. Then, it is clear from (15) that

$$v_R = \rho_0^{-1}(\rho_0 v_R) \to \tilde{v} \quad \text{weakly in } L^2(q_T) \quad \text{and} \quad y_R = \rho_R^{-1}(\rho_R y_R) \to \tilde{y} \quad \text{weakly in } L^2(Q_T).$$

In fact, $\tilde{y}$ is the state associated to $\tilde{v}$ and y_R converges strongly to $\tilde{y}$.

For every $(y', v') \in \mathcal{C}(y_0, T)$, one has

$$\begin{aligned} J(\tilde{y}, \tilde{v}) &\le \frac{1}{2}\liminf_{R\to+\infty}\left(\iint_{Q_T} \rho_R^2 |y_R|^2 \, dx\, dt + \iint_{q_T} \rho_0^2 |v_R|^2 \, dx\, dt\right) \\ &\le \frac{1}{2}\lim_{R\to+\infty}\left(\iint_{Q_T} \rho_R^2 |y'|^2 \, dx\, dt + \iint_{q_T} \rho_0^2 |v'|^2 \, dx\, dt\right) \\ &= J(y', v') \end{aligned} \tag{16}$$

Hence, $(\tilde{y}, \tilde{v}) = (\hat{y}, \hat{v})$. Finally, we also deduce from (15) that

$$\limsup_{R\to+\infty}\left(\iint_{Q_T} \rho_R^2 |y_R|^2 \, dx\, dt + \iint_{q_T} \rho_0^2 |v_R|^2 \, dx\, dt\right) \le J(\tilde{y}, \tilde{v}),$$

whence we see that (14) holds.

This ends the proof. □

Once again, it is difficult to construct a minimizing sequence for (13). On the other hand, as shown below, the constraint $y(\cdot\,,T)=0$ is related to the existence of *multipliers* in a (very) large space, difficult to handle in practice.

For these reasons, it is also convenient to consider for any $R>0$ and $\varepsilon>0$ the following *unconstrained* and *penalized* problem:

$$\begin{cases} \text{Minimize} \;\; J_{R,\varepsilon}(y,v)=\dfrac{1}{2}\displaystyle\iint_{Q_T}\rho_R^2|y|^2\,dx\,dt+\dfrac{1}{2}\iint_{q_T}\rho_0^2|v|^2\,dx\,dt+\dfrac{1}{2\varepsilon}\|y(\cdot\,,T)\|^2_{L^2} \\ \text{Subject to} \;\; (y,v)\in\mathcal{A}(y_0,T) \end{cases} \tag{17}$$

where ρ_R is as before and

$$\mathcal{A}(y_0,T)=\{\,(y,v): v\in L^2(q_T),\;\; y \text{ solves (1)}\,\}.$$

The introduction of the term in ε^{-1} may seem redundant with the use of the weight ρ_R, at least for large R and small ε. Indeed, in both cases the goal is to reinforce the null controllability condition (2). As we will see, this penalty term leads to a gain of regularity of the solution of the dual problem.

The analysis of problems (13) and (17) will provide useful indications on what to do in order to solve (3).

We will begin this analysis with (13).

To this end, let us denote by $\overline{y}$ the solution to (1) with $v=0$ and let us introduce the operators $M\in\mathcal{L}(L^2(q_T);L^2(Q_T))$ and $B\in\mathcal{L}(L^2(q_T);L^2(\Omega))$, with

$$Mv=z_v\,,\quad Bv=z_v(\cdot\,,T)$$

for all $v\in L^2(q_T)$, where z_v is the solution to

$$\begin{cases} z_t-(a(x)z_x)_x=v1_\omega\,, & (x,t)\in(0,1)\times(0,T) \\ z(x,t)=0, & (x,t)\in\{0,1\}\times(0,T) \\ z(x,0)=0, & x\in(0,1) \end{cases} \tag{18}$$

Accordingly, the solution y of (1) can be decomposed in the form

$$y=Mv+\overline{y}. \tag{19}$$

Clearly, M and B are linear bounded operators on $L^2(q_T)$. The *adjoints* M^* and B^* are given as follows:

- For each $\mu\in L^2(Q_T)$, $M^*\mu=\zeta|_{q_T}$, where ζ is the solution to the backwards system

$$\begin{cases} -\zeta_t-(a(x)\zeta_x)_x=\mu, & (x,t)\in(0,1)\times(0,T) \\ \zeta(x,t)=0, & (x,t)\in\{0,1\}\times(0,T) \\ \zeta(x,T)=0, & x\in(0,1) \end{cases} \tag{20}$$

- For each $\varphi_T\in L^2(0,1)$, $B^*\varphi_T=\varphi|_{q_T}$, where φ is the solution to

$$\begin{cases} -\varphi_t-(a(x)\varphi_x)_x=0, & (x,t)\in(0,1)\times(0,T) \\ \varphi(x,t)=0, & (x,t)\in\{0,1\}\times(0,T) \\ \varphi(x,T)=\varphi_T(x), & x\in(0,1) \end{cases} \tag{21}$$

We now associate to (13) the (dual) unconstrained problem

$$\left\{\begin{array}{l} \text{Minimize } J^*_R(\mu,\varphi_T)=\dfrac{1}{2}\displaystyle\iint_{q_T}\rho_0^{-2}|\psi|^2\,dx\,dt+\dfrac{1}{2}\iint_{Q_T}\rho_R^{-2}|\mu|^2\,dx\,dt+\int_0^1\psi(x,0)\,y_0(x)\,dx \\ \text{Subject to } (\mu,\varphi_T)\in L^2(Q_T)\times L^2(0,1) \end{array}\right. \tag{22}$$

where, for any $(\mu,\varphi_T)\in L^2(Q_T)\times L^2(0,1)$, we have set $\psi=M^*\mu+B^*\varphi_T$, i.e. ψ is the solution to

$$\left\{\begin{array}{ll} -\psi_t-(a(x)\psi_x)_x=\mu, & (x,t)\in(0,1)\times(0,T) \\ \psi(x,t)=0, & (x,t)\in\{0,1\}\times(0,T) \\ \psi(x,T)=\varphi_T(x), & x\in(0,1). \end{array}\right. \tag{23}$$

Problem (22) may have no solution. This is well known and is implied by the fact that J^*_R is coercive in the second variable φ_T with respect to a norm much weaker than $\|\cdot\|_{L^2}$, but not necessarily with respect to $\|\cdot\|_{L^2}$. Thus, any minimizing sequence is only bounded in a space strictly larger than $L^2(Q_T)\times L^2(0,1)$ and the "solution" to (22) does not belong to this space.

Nevertheless, we will be able to construct adequate minimizing sequences for (22) that converge in an appropriate sense to the unique solution to (13).

On the other hand, we can associate to (17) the following dual (well-posed) problem:

$$\left\{\begin{array}{l} \text{Minimize } \ J^*_{R,\varepsilon}(\mu,\varphi_T)=\dfrac{1}{2}\left(\displaystyle\iint_{Q_T}\rho_R^{-2}|\mu|^2\,dx\,dt+\iint_{q_T}\rho_0^{-2}|\psi|^2\,dx\,dt\right) \\ \qquad\qquad\qquad\qquad +\displaystyle\int_0^1\varphi(x,0)\,y_0(x)\,dx+\frac{\varepsilon}{2}\|\varphi_T\|^2_{L^2} \\ \text{Subject to } \ (\mu,\varphi_T)\in L^2(Q_T)\times L^2(0,1). \end{array}\right. \tag{24}$$

In the following result, we explain how (22) and (24) are respectively related to (13) and (17):

PROPOSITION 3.2 *The unconstrained extremal problems* (22) *and* (24) *are respectively dual problems to* (13) *and* (17) *in the sense of the Fenchel-Rockafellar theory. Furthermore,* (17) *and* (24) *are stable and possess unique solutions. Finally, if we denote by* $(y_{R,\varepsilon},v_{R,\varepsilon})$ *the unique solution to* (17)*, we denote by* $(\mu_{R,\varepsilon},\varphi_{T,R,\varepsilon})$ *the unique solution to* (24) *and we set* $\psi_{R,\varepsilon}=M^*\mu_{R,\varepsilon}+B^*\varphi_{T,R,\varepsilon}$*, then the following relations hold:*

$$v_{R,\varepsilon}=\rho_0^{-2}\psi_{R,\varepsilon}|_{q_T}\,,\quad y_{R,\varepsilon}=-\rho_R^{-2}\mu_{R,\varepsilon}\,,\quad y_{R,\varepsilon}(\cdot\,,T)=-\varepsilon\varphi_{T,R,\varepsilon}\,. \tag{25}$$

PROOF: Let us first check that (24) is the dual problem of (17).

In view of the decomposition (19), we can write that $J_{R,\varepsilon}(y,v)=F(Mv,Bv)+G(v)$ for any $(y,v)\in\mathcal{A}(y_0,T)$. Here, we have introduced the functions F and G, with

$$F(z,z_T)=\frac{1}{2}\iint_{Q_T}\rho_R^2|z+\overline{y}|^2\,dx\,dt+\frac{1}{2\varepsilon}\int_0^1|z_T(x)+\overline{y}(x,T)|^2\,dx$$

and

$$G(v)=\frac{1}{2}\iint_{Q_T}\rho_0^2|v|^2\,dx\,dt.$$

The functions $F:L^2(Q_T)\times L^2(q_T)\mapsto\mathbb{R}$ and $G:L^2(q_T)\mapsto\mathbb{R}$ are both convex and continuous and we can apply the duality Theorem of W. Fenchel and T.R. Rockafellar; see Theorem 4.2 p. 60 in [5]. We deduce that

$$\inf_{\mathcal{A}(y_0,T)}J_{R,\varepsilon}(y,v)=-\inf_{L^2(Q_T)\times L^2(0,1)}\Big\{G^*(M^*\mu+B^*\varphi_T)+F^*(-(\mu,\varphi_T))\Big\},$$

where F^* and G^* are the convex conjugate of F and G, respectively.

Notice that

$$F^*(\mu,\varphi_T) = \sup_{L^2(Q_T)\times L^2(0,1)} \left\{ \iint_{Q_T} \mu\, z\, dx\, dt + \int_0^1 \varphi_T(x)\, z_T(x)\, dx - F(z,z_T) \right\}$$
$$= \frac{1}{2} \iint_{Q_T} \rho_R^{-2} |\mu|^2\, dx\, dt - \iint_{Q_T} \mu \overline{y}\, dx\, dt + \frac{\varepsilon}{2} \|\varphi_T\|_{L^2} - \int_0^1 \varphi_T(x)\, \overline{y}(x,T)\, dx$$

for all $(\mu,\varphi_T) \in L^2(Q_T)\times L^2(0,1)$. On the other hand,

$$G^*(w) = \frac{1}{2} \iint_{Q_T} \rho_0^{-2} |w|^2\, dx\, dt$$

for all $w \in L^2(Q_T)$. Therefore,

$$G^*(M^*\mu + B^*\varphi_T) + F^*(-(\mu,\varphi_T)) = \frac{1}{2} \iint_{Q_T} \rho_R^{-2} |\mu|^2\, dx\, dt + \frac{1}{2} \iint_{Q_T} \rho_0^{-2} |\psi|^2\, dx\, dt + \frac{\varepsilon}{2} \|\varphi_T\|_{L^2}^2$$
$$+ \iint_{Q_T} \mu\, \overline{y}\, dx\, dt + \int_0^1 \varphi_T(x)\, \overline{y}(x,T)\, dx$$

where we have used again the notation $\psi = M^*\mu + B^*\varphi_T$.

Finally, multiplying the state equation of (23) by $\overline{y}$ and integrating by parts, we obtain that

$$\iint_{Q_T} \mu\, \overline{y}\, dx\, dt + \int_0^1 \varphi_T(x)\, \overline{y}(x,T)\, dx = \int_0^1 \psi(x,0)\, y_0(x)\, dx,$$

whence

$$G^*(M^*\mu + B^*\varphi_T) + F^*(-(\mu,\varphi_T)) = \frac{1}{2} \left(\iint_{Q_T} \rho_R^{-2} |\mu|^2\, dx\, dt + \iint_{q_T} \rho_0^{-2} |\psi|^2\, dx\, dt \right)$$
$$+ \int_0^1 \varphi(x,0)\, y_0(x)\, dx + \frac{\varepsilon}{2} \|\varphi_T\|_{L^2}^2$$

This proves that (24) is the dual of (17).

The fact that (22) is the dual of (13) can be proved in a very similar way.

It is also easy to check that (17) and (24) are stable and possess unique solutions. Indeed, the hypotheses of Theorem 4.2 in [5] are satisfied for (17) (notice that this is not the case for (13), since the interior of the constraint set $\mathcal{C}(y_0,T)$ is empty).

Finally, let us deduce that the optimality conditions (25) hold.

Let us set $(y,v) = (y_{R,\varepsilon}, v_{R,\varepsilon})$ and $(\mu,\varphi_T) = (\mu_{R,\varepsilon}, \varphi_{T,R,\varepsilon})$. Then, since (24) and (17) are dual to each other, one has:

$$0 = \frac{1}{2} \iint_{Q_T} \rho_R^2 |y|^2\, dx\, dt + \frac{1}{2} \iint_{q_T} \rho_0^2 |v|^2\, dx\, dt + \frac{1}{2\varepsilon} \|y(\cdot\,,T)\|_{L^2}^2$$
$$+ \frac{1}{2} \iint_{Q_T} \rho_R^{-2} |\mu|^2\, dx\, dt + \frac{1}{2} \iint_{q_T} \rho_0^{-2} |\psi|^2\, dx\, dt + (\psi(\cdot\,,0), y_0) + \frac{\varepsilon}{2} \|\varphi_T\|_{L^2}^2$$
$$= \frac{1}{2} \iint_{Q_T} \rho_R^2 |y + \rho_R^{-2}\mu|^2\, dx\, dt + \frac{1}{2} \iint_{q_T} \rho_0^2 |v - \rho_0^{-2}\psi|^2\, dx\, dt + \frac{1}{2\varepsilon} \|y(\cdot\,,T) + \varepsilon\varphi_T\|_{L^2}^2$$
$$- \iint_{Q_T} \rho_R^2 y\mu\, dx\, dt + \iint_{q_T} \rho_0^2 v\psi\, dx\, dt - (y(\cdot\,,T), \varphi_T)_{L^2} + (\psi(\cdot\,,0), y_0)_{L^2} .$$

But the terms in the last line cancel, since $\psi = M^*\mu + B^*\varphi_T$. Consequently,

$$\iint_{Q_T} \rho_R^2 |y + \rho_R^{-2}\mu|^2 \, dx\, dt + \iint_{q_T} \rho_0^2 |v - \rho_0^{-2}\psi|^2 \, dx\, dt + \frac{1}{\varepsilon}\|y(\cdot\,,T) + \varepsilon\varphi_T\|_{L^2}^2 = 0$$

and we immediately get (25).

This ends the proof. □

We will now justify the introduction of the parameter ε by analyzing the behavior of the solutions to the problems (17) as $\varepsilon \to 0^+$.

PROPOSITION 3.3 *With the notation of Proposition 3.2, for each fixed $R > 0$ one has*

$$v_{R,\varepsilon} \to v_R \quad \textit{strongly in } L^2(q_T) \quad \textit{and} \quad y_{R,\varepsilon} \to y_R \quad \textit{strongly in } L^2(Q_T) \tag{26}$$

as $\varepsilon \to 0^+$.

PROOF : First, notice that, for each $R > 0$ and $\varepsilon > 0$, one has

$$\iint_{Q_T} \rho_R^2 |y_{R,\varepsilon}|^2 \, dx\, dt + \iint_{q_T} \rho_0^{-2} |\psi_{R,\varepsilon}|^2 \, dx\, dt + \frac{1}{\varepsilon}\|y_{R,\varepsilon}(\cdot\,,T)\|_{L^2}^2 = (\psi_{R,\varepsilon}(\cdot\,,0), y_0)_{L^2} \,. \tag{27}$$

Indeed, taking into account the equations satisfied by $y_{R,\varepsilon}$ and $\psi_{R,\varepsilon}$ and the identities (25), we find that the sum of the two integrals in the left hand side of (27) is equal to

$$\begin{aligned}
&\iint_{Q_T} (L^*\psi_{R,\varepsilon}\, y_{R,\varepsilon} - \psi_{R,\varepsilon}\, L y_{R,\varepsilon}) \, dx\, dt \\
&\qquad = -\,(\psi_{R,\varepsilon}(\cdot\,,t), y_{R,\varepsilon}(\cdot\,,t)_{L^2}|_{t=0}^{t=T} \\
&\qquad = (\psi_{R,\varepsilon}(\cdot\,,0), y_0)_{L^2} - \frac{1}{\varepsilon}\|y_{R,\varepsilon}(\cdot\,,T)\|_{L^2} \,.
\end{aligned}$$

Now, from Lemma 2.2 applied to $\psi_{R,\varepsilon}$, we deduce that the left hand side of (27) is uniformly bounded. Indeed, we have

$$\begin{aligned}
&\iint_{Q_T} \rho_R^2 |y_{R,\varepsilon}|^2 \, dx\, dt + \iint_{q_T} \rho_0^{-2} |\psi_{R,\varepsilon}|^2 \, dx\, dt + \frac{1}{\varepsilon}\|y_{R,\varepsilon}(\cdot\,,T)\|_{L^2}^2 \\
&\quad \le \|\psi_{R,\varepsilon}(\cdot\,,0)\|_{L^2}\, \|y_0\|_{L^2} \\
&\quad \le C\, \|y_0\|_{L^2} \left(\iint_{Q_T} \rho^{-2}\rho_R^4 |y_{R,\varepsilon}|^2 \, dx\, dt + \iint_{q_T} \rho_0^{-2} |\psi_{R,\varepsilon}|^2 \, dx\, dt \right)^{1/2} \\
&\quad \le C\, \|y_0\|_{L^2} \left(\iint_{Q_T} \rho_R^2 |y_{R,\varepsilon}|^2 \, dx\, dt + \iint_{q_T} \rho_0^{-2} |\psi_{R,\varepsilon}|^2 \, dx\, dt \right)^{1/2} .
\end{aligned}$$

Therefore, $\rho_R\, y_{R,\varepsilon}$ is uniformly bounded in $L^2(Q_T)$, $\rho_0\, v_{R,\varepsilon} = \rho_0^{-1}\psi_{R,\varepsilon}|_{q_T}$ is uniformly bounded in $L^2(q_T)$, $\|y_{R,\varepsilon}(\cdot\,,T)\|_{L^2} \le C\varepsilon^{1/2}$ and, at least for some subsequence, one has

$$\rho_R\, y_{R,\varepsilon} \to z_R = \rho_R\, \tilde{y}_R \quad \text{weakly in } L^2(Q_T) \quad \text{and} \quad \rho_0\, v_{R,\varepsilon} \to w_R = \rho_0\, \tilde{v}_R \quad \text{weakly in } L^2(q_T) \tag{28}$$

as $\varepsilon \to 0^+$.

Obviously, $\tilde{y}_R$ is the state associated to $\tilde{v}_R$ and $y_{R,\varepsilon}$ converges strongly to $\tilde{y}_R$ in $L^2(Q_T)$. Moreover, $\tilde{y}(\cdot\,,T) = 0$, that is, $(\tilde{y}_R, \tilde{v}_R) \in \mathcal{C}(y_0, T)$.

Now, arguing as in the proof of Proposition 3.1, it is not difficult to check that $(\tilde{y}_R, \tilde{v}_R)$ is the unique optimal pair of (13), i.e. $(\tilde{y}_R, \tilde{v}_R) = (y_R, v_R)$ and $v_{R,\varepsilon}$ also converges strongly.

This ends the proof. □

In view of these convergence results, it seems that an appropriate way to solve (3) is to first find the solution to (24) and then apply the relations (25) for various small ε and large R. This is confirmed by the experiments in Section 5.

We also observe that problems (13) and (17) are close for small ε. In fact, the experiments below will show that the parameter ε is in some sense useless, since the presence of the weighted integral of μ in the functional of (22) suffices to stabilize this extremal problem. The term in ε ensures that the second argument of the optimal pair $(\mu_{R,\varepsilon}, \varphi_{T,R,\varepsilon})$ belongs to $L^2(0,1)$ and therefore allows to apply duality techniques in a rigorous way without using "abstract" or "nonstandard" spaces.

Remark 1 Contrarily to the functional J^* in (4), as a consequence of the way we have decided to penalize the constraint (2), $J^*_{R,\varepsilon}$ is explicitly quadratic in $\|\varphi_T\|_{L^2(0,1)}$. In particular, this avoids the use of *operator-splitting methods* (see [10], Section 1.8.8). This does not affect the asymptotic limit in ε, since it can be shown that the state $y_{R,\varepsilon}$ associated to $v_{R,\varepsilon} = \rho_0^{-2}\psi_{R,\varepsilon}\,1_\omega$ satisfies

$$\|y_{R,\varepsilon}(\cdot\,,T)\|_{L^2(0,1)} \le C_R\sqrt{\varepsilon}\,\|y_0\|_{L^2(0,1)}$$

for some $C_R > 0$. □

Remark 2 It is reasonable to expect that $\|\varphi_{T,R,\varepsilon}\|_{L^2(0,1)}$ blows up as $\varepsilon \to 0^+$; this is confirmed by the numerical results in Section 5. A short explanation of this fact is the following: when we try to solve (3) or (13), the constraint (2) can be viewed as an equality in a "very small" space (due to the strong regularization effect of the heat operator); accordingly, the associated multipliers φ_T and $\varphi_{T,R}$ belong to a "large" dual space, much larger than L^2. We refer to [16] for more details. □

Remark 3 As briefly recalled in Section 2, we have followed other strategies in [7] to solve the same control problem (3). There, the key point was to find p in an appropriate space P such that

$$v = \rho_0^{-2}p|_{q_T}\,, \quad y = -\rho^{-2}L^*p;$$

recall Proposition 2.1 and compare with (25). The singular behavior of $\varphi_{T,R,\varepsilon}$ as $\varepsilon \to 0^+$ has a parallel in the computation of p. Indeed, the numerical results in [7] show that $p(\cdot\,,T)$ does not belong to $L^2(0,1)$. □

Remark 4 There are other ways to apply duality techniques to (3). For instance, we can use the fact that, if the first integral in (3) is finite, then necessarily (2) is satisfied. This leads to the extremal problem

$$\begin{cases} \text{Minimize } \dfrac{1}{2}\displaystyle\iint_{q_T}\rho_0^{-2}|\zeta|^2\,dx\,dt + \frac{1}{2}\iint_{Q_T}\rho^{-2}|\mu|^2\,dx\,dt + \int_0^1 \zeta(x,0)\,y_0(x)\,dx \\ \text{Subject to } \mu \in L^2(Q_T) \end{cases}$$

where, for each $\mu \in L^2(Q_T)$, we have set $\zeta = M^*\mu$; recall (20). However, similarly to (22), this formulation is formal since the unique minimizer μ may not belong to $L^2(Q_T)$. □

Remark 5 We can also apply duality techniques to problem (17) so as to get a dual functional expressed only in terms of the variable φ_T. Using once again that $y = Mv + \overline{y}$, we now decompose $J_{R,\varepsilon}$ as follows:

$$J_{R,\varepsilon}(y,v) = F_1(v) + F_2(Bv),$$

where

$$F_1(v) = \frac{1}{2}\iint_{Q_T} \rho_R^2 |Mv + \overline{y}|^2 \, dx\, dt + \frac{1}{2}\iint_{q_T} \rho_0^2 |v|^2 \, dx\, dt$$

and

$$F_2(Bv) = \frac{1}{2\varepsilon}\|Bv + \overline{y}(\cdot\,, T)\|^2,$$

so that

$$\inf_{L^2(q_T)} \{F_1(v) + F_2(Bv)\} = -\inf_{L^2(0,1)} \{F_1^*(B^*\varphi_T) + F_2^*(-\varphi_T)\}.$$

Introducing the operators $\mathcal{B}$ and $\mathcal{A}$, with $\mathcal{B}(\varphi_T) = B^*\varphi_T - M^*(\rho_R^2\overline{y})$ and $\mathcal{A} = M^*(\rho_R^2 M) + \rho_0^2\, 1_\omega$, it is not difficult to check that

$$F_1^*(B^*\varphi_T) = \frac{1}{2}\iint_{Q_T} (\mathcal{A}^{-1}\mathcal{B}(\varphi_T))\, \mathcal{B}(\varphi_T)\, dx\, dt - \frac{1}{2}\iint_{Q_T} \rho_R^2 |\overline{y}|^2 \, dx\, dt$$

and

$$F_2^*(-\varphi_T) = \frac{\varepsilon}{2}\|\varphi_T\|^2 + \int_0^1 \varphi_T(x)\, \overline{y}(x,T)\, dx$$

for all $\varphi_T \in L^2(0,1)$. Consequently, another extremal problem that can be put in duality with (17) is the following:

$$\left\{ \begin{array}{l} \text{Minimize } \dfrac{1}{2}\displaystyle\iint_{q_T} (\mathcal{A}^{-1}\mathcal{B}(\varphi_T))\, \mathcal{B}(\varphi_T)\, dx\, dt + \int_0^1 \varphi_T(x)\, \overline{y}(x,T)\, dx + \frac{\varepsilon}{2}\|\varphi_T\|^2 \\ \text{Subject to } \varphi_T \in L^2(0,1) \end{array} \right. \tag{29}$$

When we consider a problem similar to (17) but with weights $\rho \equiv 0$ and $\rho_0 \equiv 1$, we find again a dual problem of this kind, with $\mathcal{A}$ and $\mathcal{B}$ respectively replaced by 1_ω and B^*. This is just the formulation considered in [3].

Contrarily to (24), problem (29) involves minimization only with respect to the variable $\varphi_T \in L^2(0,1)$. However, it requires the inversion of a nonlocal operator $\mathcal{A}$ and is therefore *a priori* much harder to solve. □

4 Conjugate gradient algorithm and numerical approximation

In this Section, we address the numerical solution of the minimization problem (24). Following [3], the method combines conjugate gradient algorithms with finite difference and finite element approximations.

For convenience, we introduce the Hilbert space $V = L^2(Q_T) \times L^2(0,1)$.

The problem we want to solve reads as follows: for given $\varepsilon, R > 0$, $y_0 \in L^2(0,1)$ and $T > 0$, minimize over the Hilbert space V the functional

$$J_{R,\varepsilon}^*(\mu, \varphi_T) = \frac{1}{2}\iint_{q_T} \rho_0^{-2} |\psi|^2 \, dx\, dt + \frac{1}{2}\iint_{Q_T} \rho_R^{-2} |\mu|^2 \, dx\, dt + \int_0^1 \psi(x,0)\, y_0(x)\, dx + \frac{\varepsilon}{2}\|\varphi_T\|_{L^2}^2,$$

where $\psi = M^*\mu + B^*\varphi_T$, that is, ψ is the solution to (23). By definition, it will be said that ψ is the *adjoint state* associated to μ and φ_T.

Notice that, in view of the optimality condition (25), the optimal $\mu_{R,\varepsilon}$ satisfies $\mu_{R,\varepsilon}+\rho_R^2 y_{R,\varepsilon} = 0$ and therefore must vanish on $\{0,1\}\times(0,T)$.

Our aim is to apply a conjugate gradient method. First of all, notice that the Fréchet derivative of $J^*_{R,\varepsilon}$ at (μ,φ_T) in the direction (μ',φ'_T) is given by

$$DJ^*_{R,\varepsilon}(\mu,\varphi_T)\cdot(\mu',\varphi'_T) = \iint_{Q_T}(z+\rho_R^{-2}\mu)\,\mu'\,dx\,dt + \int_0^1 (z(x,T)+\varepsilon\varphi_T(x))\,\varphi'_T(x)\,dx,$$

where z is the unique solution of the following system:

$$\begin{cases} z_t - (a(x)z_x)_x = \rho_0^{-2}\psi 1_\omega, & (x,t)\in(0,1)\times(0,T)\\ z(x,t) = 0, & (x,t)\in\{0,1\}\times(0,T)\\ z(x,0) = y_0(x), & x\in(0,1)\end{cases} \tag{30}$$

and $\psi = M^*\mu + B^*\varphi_T$ is the solution to (23).

Consequently, the gradient of $J^*_{R,\varepsilon}$ at (μ,φ_T) is $(z+\rho^{-2}\mu, z(\cdot\,,T)+\varepsilon\varphi_T)$, where z is the solution to (30). This suggests to apply, at least at a first step, a classical gradient (steepest descent) method:

$$\begin{cases} (\mu^0,\varphi_T^0) \text{ is given in } V,\\ (\mu^{n+1},\varphi_T^{n+1}) = (\mu^n,\varphi_T^n) - \eta^n(z^n+\rho_R^{-2}\mu^n, z^n(\cdot\,,T)+\varepsilon\varphi_T), \quad n\geq 0.\end{cases} \tag{31}$$

Here, η^n is the minimizer of the function $\eta\mapsto J^*_{R,\varepsilon}((\mu^n,\varphi_T^n)-\eta(z^n+\rho_R^{-2}\mu^n, z^n(\cdot\,T)+\varepsilon\varphi_T^n))$.

Unfortunately, due to the lack of uniform coercivity of $J^*_{R,\varepsilon}$ with respect to R and ε, this gradient method is not sufficiently robust and does not provide good results as soon as the discretization parameters are small enough.

4.1 The conjugate gradient algorithm

Let us introduce the following symmetric and continuous bilinear form on V:

$$\begin{aligned} a_{R,\varepsilon}((\mu,\varphi_T),(\mu',\varphi'_T)) = &\iint_{q_T}\rho_0^{-2}(M^*\mu+B^*\varphi_T)(M^*\mu'+B^*\varphi'_T)\,dx\,dt + \iint_{Q_T}\rho_R^{-2}\mu\mu'\,dx\,dt\\ &+\varepsilon\int_0^1\varphi_T(x)\,\varphi'_T(x)\,dx \qquad \forall(\mu,\varphi_T),(\mu',\varphi'_T)\in V.\end{aligned}$$

Then one has

$$J^*_{R,\varepsilon}(\mu,\varphi_T) = \frac{1}{2}a_{R,\varepsilon}((\mu,\varphi_T),(\mu,\varphi_T)) + \int_0^1 (M^*\mu+B^*\varphi_T)(x,0)\,y_0(x)\,dx$$

for all $(\mu,\varphi_T)\in V$.

For any ε and R, the bilinear form $a_{R,\varepsilon}(\cdot\,,\cdot)$ is coercive with respect to the norm of V. It is therefore appropriate to apply conjugate gradient methods to (24); see [10].

In particular, the *Polak-Ribière* version reads as follows:

STEP 0: INITIALIZATION

Let ϵ be a small and strictly positive real number.

We choose $(\mu^0, \varphi_T^0) \in V$ and we compute the gradient g^0 of $J^*_{R,\varepsilon}$ at (μ^0, φ_T^0). Notice that $g^0 = (g_1^0, g_2^0)$, where g_1^0 and g_2^0 are respectively given by

$$g_1^0 = z^0 + \rho_0^{-2}\mu^0, \quad g_2^0 = z^0(\cdot\,, T) + \varepsilon\varphi_T^0$$

and z^0 solves, together with ψ^0, the cascade system

$$\begin{cases} L^*\psi^0 = \mu^0 \quad \text{in} \quad Q_T, \qquad \psi^0 = 0 \quad \text{on} \quad \Sigma_T, \qquad \psi^0(\cdot\,, T) = \varphi_T^0 \\ Lz^0 = \rho_0^{-2}\psi^0\, 1_\omega \quad \text{in} \quad Q_T, \qquad z^0 = 0 \quad \text{on} \quad \Sigma_T, \qquad z^0(\cdot\, 0) = y_0. \end{cases}$$

If $\|g^0\|_V / \|(\mu^0, \phi_T^0)\|_V \leq \epsilon$, then we take $(\mu, \phi_T) = (\mu^0, \varphi_T^0)$ and we stop; otherwise, we set

$$w^0 = (w_1^0, w_2^0) = g^0.$$

Then, for $n \geq 0$, assuming that (μ^n, φ_T^n), g^n and w^n are given, with $g^n \neq 0$ and $w^n \neq 0$, we compute $(\mu^{n+1}, \varphi_T^{n+1})$, g^{n+1} and (if necessary) w^{n+1} performing the following steps.

Step 1: Steepest descent

We set

$$\eta^n = \frac{DJ^*_{R,\varepsilon}(\mu^n, \varphi_T^n) \cdot w^n}{a_{R,\varepsilon}(w^n, w^n)}$$

and we take

$$(\mu^{n+1}, \varphi_T^{n+1}) = (\mu^n, \varphi_T^n) - \eta^n w^n.$$

Then, we compute the gradient g^{n+1} of $J^*_{R,\varepsilon}$ at $(\mu^{n+1}, \phi_T^{n+1})$. Now, we have $g^{n+1} = (g_1^{n+1}, g_2^{n+1})$, where g_1^{n+1} and g_2^{n+1} are respectively given by

$$g_1^{n+1} = z^{n+1} + \rho_0^{-2}\mu^{n+1}, \quad g_2^{n+1} = z^{n+1}(\cdot\,, T) + \varepsilon\varphi_T^n$$

and z^{n+1} is, together with ψ^{n+1}, the solution to the cascade system

$$\begin{cases} L^*\psi^{n+1} = \mu^{n+1} \quad \text{in} \quad Q_T, \qquad \psi^{n+1} = 0 \quad \text{on} \quad \Sigma_T, \qquad \psi^{n+1}(\cdot\,, T) = \varphi_T^{n+1} \\ Lz^{n+1} = \rho_0^{-2}\psi^{n+1}\, 1_\omega \quad \text{in} \quad Q_T, \qquad z^{n+1} = 0 \quad \text{on} \quad \Sigma_T, \qquad z^{n+1}(\cdot\,, 0) = y_0. \end{cases}$$

Step 2: Convergence test and construction of the new direction

If $\|g^{n+1}\|_{L^2(Q_T)\times L^2(0,1)} / \|g^0\|_{L^2(Q_T)\times L^2(0,1)} \leq \epsilon$, then we take $(\mu, \varphi_T) = (\mu^{n+1}, \varphi_T^{n+1})$ and we stop; otherwise, we compute

$$\gamma_n = \frac{(g^{n+1} - g^n, g^{n+1})_V}{\|g^n\|_V^2}\,, \tag{32}$$

we take

$$w^{n+1} = g^{n+1} + \gamma_n w^n$$

and we return to Step 1 with n replaced by $n + 1$.

Remark 6 In the present quadratic-linear situation, by construction the gradients g^n are conjugate to each other, that is,

$$(g^m, g^n)_V = 0 \quad \forall m, n \geq 0, \quad m \neq n.$$

Consequently, the parameter γ_n given by (32) can also be written in the form

$$\gamma_n = \frac{\|g^{n+1}\|_V^2}{\|g^n\|_V^2}\,. \tag{33}$$

For non necessarily quadratic-linear extremal problems, the choices (32) and (33) are not equivalent; they respectively lead to the *Polak-Ribiere* and the *Fletcher-Reeves* conjugate gradient algorithms. In our case, due to the numerical approximation, the orthogonality of the g^n is lost and strongly accentuated for small values of ε and large values of R. In that stiff case, the *Polak-Ribiere* version, mainly used in nonlinear situations, appears much more robust. □

Remark 7 With ρ_R and ρ_0 respectively replaced by the constants 0 and 1, we obtain exactly the conjugate gradient algorithm considered in Section 1.8 in [10], designed for the computation of the control of minimal norm in $L^2(q_T)$. Notice that the present situation does not lead to a significative increase of the computation cost. □

4.2 Full discrete approximations

For "large" integers N_x and N_t , we set $\Delta x = 1/N_x$, $\Delta t = T/N_t$ and $h = (\Delta x, \Delta t)$. Let us denote by $\mathcal{P}_{\Delta x}$ the uniform partition of $[0,1]$ associated to Δx and let us denote by $\mathcal{Q}_h$ the uniform quadrangulation of Q_T associated to h. In particular,

$$\overline{Q}_T = \bigcup_{K \in \mathcal{Q}_h} K.$$

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

$$X_h = \{\, \varphi_h \in C^0([0,1]\times[0,T]) : \varphi_h|_K \in (\mathbb{P}_{1,x} \otimes \mathbb{P}_{1,t})(K) \;\; \forall K \in \mathcal{Q}_h \,\}.$$

Here, $\mathbb{P}_{m,\xi}$ denotes the space of polynomial functions of order m in the variable ξ. Accordingly, we see that the functions in X_h reduce on each quadrangle $K \in \mathcal{Q}_h$ to a polynomial of the form

$$A + Bx + Ct + Dxt$$

involving 4 degrees of freedom. Obviously, the space X_h is a conformal approximation of $L^2(Q_T)$. We will also consider the space

$$X_{0h} = \{\, \varphi_h \in X_h : \varphi_h(0,t) = \varphi_h(1,t) = 0 \;\; \forall t \in (0,T) \,\}.$$

Of course, X_{0h} is a finite-dimensional subspace of $L^2(0,T;H_0^1(0,1))$ and the functions $\varphi_h \in X_{0h}$ are uniquely determined by their values at the nodes (x_j, t_j) of $\mathcal{Q}_h$ such that $0 < x_j < 1$.

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 μ appears in the right hand side of the backward equation $L^*\psi = \mu$, it is natural to approximate $\mu \in L^2(Q_T)$ by a piecewise constant function. Thus, let M_h be the space defined by

$$M_h = \{\, \mu_h \in L^2(Q_T) : \mu_h|_K \in (\mathbb{P}_{0,x} \otimes \mathbb{P}_{0,t})(K) \ \ \forall K \in \mathcal{Q}_h \,\}.$$

Obviously, M_h is a finite dimensional subspace of $L^2(Q_T)$ and the functions in M_h are uniquely determined by their (constant) values on the quadrangles $K \in \mathcal{Q}_h$.

For any h, we therefore consider the following approximation of (24):

$$\left\{ \begin{array}{l} \text{Minimize} \ \ J^*_{R,\varepsilon,h}(\mu_h, \varphi_{\Delta x,T}) = \dfrac{1}{2}\left(\displaystyle\iint_{Q_T} \rho_R^{-2}|\mu_h|^2 \, dx\, dt + \iint_{q_T} \rho_0^{-2}|\psi_h|^2 \, dx\, dt \right) \\ \qquad\qquad\qquad\qquad + \displaystyle\int_0^1 \varphi_h(x,0)\, y_0(x)\, dx + \frac{\varepsilon}{2}\|\varphi_{\Delta x,T}\|^2_{L^2(0,1)} \\ \text{Subject to} \ \ (\mu_h, \varphi_{\Delta x,T}) \in M_h \times \Phi_{\Delta x} \,. \end{array} \right. \tag{34}$$

In (34), for every $\mu_h \in M_h$ and every $\varphi_{\Delta x,T} \in \Phi_{\Delta x}$, we have denoted by ψ_h the associated *discrete adjoint state.* By definition, $\psi_h \in X_{0h}$ is given as follows:

(i) Let us introduce the times $t_j = j\Delta t$. We have $T = t_{N_t}$ and we first set $\psi_h|_{t=T} = \varphi_{\Delta x,T}$.

(ii) Secondly, $\psi_h|_{t=t_{N_t-1}}$ is the solution of the linear problem

$$\left\{ \begin{array}{l} \displaystyle\int_0^1 \frac{1}{\Delta t}(\Psi - \varphi_{\Delta x,T})z\, dx + \frac{1}{2}\int_0^1 a(x)\Psi_x z_x \, dx = \frac{1}{2}\int_0^1 (\mu_h(x,t_{N_t-1}) + \mu_h(x,t_{N_t}))z(x)\, dx \\ \forall z \in \Phi_{\Delta x} \,, \quad \Psi \in \Phi_{\Delta x} \,. \end{array} \right.$$

(iii) Then, for given $n = N_t - 1, \ldots, 2$, $\Psi^* = \varphi_h|_{t=t_{n+1}}$ and $\overline{\Psi} = \varphi_h|_{t=t_n}$, $\varphi_h|_{t=t_{n-1}}$ is the solution of the linear problem

$$\left\{ \begin{array}{l} \displaystyle\int_0^1 \frac{1}{2\Delta t}(3\Psi - 4\overline{\Psi} + \Psi^*)z\, dx + \int_0^1 a(x)\Psi_x z_x \, dx = \int_0^1 \mu_h(x,t_{n-1})z(x)\, dx \\ \forall z \in \Phi_{\Delta x} \,, \quad \Psi \in \Phi_{\Delta x} \,. \end{array} \right.$$

We are thus using the two-step *implicit Gear* algorithm as a numerical tool to solve numerically the adjoint problem (22). As advocated in [3], where the influence of the time discretization is highlighted, it has been observed that this second order scheme ensures a better behavior of the underlying conjugate gradient algorithm than, for instance, the implicit Euler scheme.

For the computation of the gradient of $J^*_{R,\varepsilon,h}$, we also need to solve numerically systems of the form (30). This can be done in a similar way.

For any R and ε, the functional $J^*_{R,\varepsilon,h}$ enjoys the same properties than $J^*_{R,\varepsilon}$ when V is replaced by $V_h := X_h \times \Phi_{\Delta x}$. In particular $J^*_{R,\varepsilon,h}$ is coercive on V_h , uniformly with respect to h. Hence, (34) may be solved with the conjugate gradient algorithm stated in Section 4.1.

5 Numerical experiments

We will now present some numerical experiments and we will analyze the behavior of the computed controls with respect to the data and $h = (\Delta x, \Delta t)$. The numerical results will be compared to

those obtained in [7] by using some primal methods (that is, some methods devoted to solve (3) that "ignore" duality). Also, we will try to explain the role played by the weights ρ and ρ_0.

We emphasize that these weights appear in the conjugate functional $J^*_{R,\varepsilon}$ only through their inverse. Therefore, the fact that ρ and ρ_0 blow up as $t \to T^-$ does not lead *a priori* to any numerical pathology when we try to solve (24) numerically.

For any $s \in (0,1)$, we consider the following function $\beta_{0,s}$:

$$\beta_{0,s}(x) = \frac{x(1-x)e^{-(x-c_s)^2}}{s(1-s)e^{-(s-c_s)^2}}, \quad c_s = s - \frac{1-2s}{2s(1-s)}. \tag{35}$$

Then, if s belongs to ω, $\beta_{0,s}$ satisfies the conditions in (7). Indeed, notice that $\beta_{0,s}(0) = \beta_{0,s}(1) = 0$, $\beta_{0,s} > 0$ in $(0,1)$ and $|\beta'_{0,s}| > 0$ except at $x = s$.

In the following numerical experiments, we will take ρ and ρ_0 as in (7) with $\beta_0 = \beta_{0,s}$ for some $s \in \omega$. The function $\beta_{0,s}$ is plotted in Figure 1 for $s = 1/10, 1/4$ and $s = 1/2$. The weights ρ^{-2} and ρ_0^{-2} corresponding to $s = 1/2$ are displayed in Figure 2.

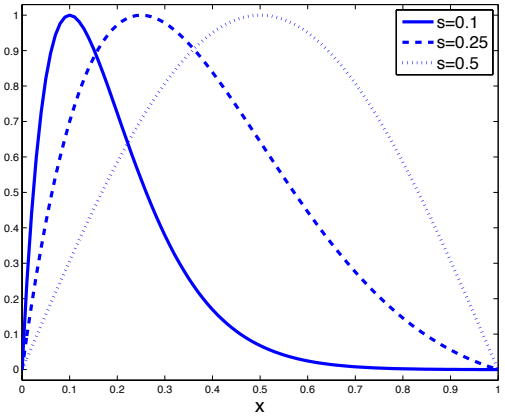

Figure 1: The function $\beta_{0,s}$ for several s.

For simplicity, we will assume that $\Delta x = \Delta t$. The following numerical values will be fixed:

$$T = 1/2, \qquad a(x) \equiv a_0 = 1/10, \quad x \in (0,1). \tag{36}$$

In particular, a is a positive constant (other situations have been considered and discussed in [7], in particular, discontinuous diffusion coefficients).

5.1 Experiment 1: Behavior of the numerical solutions $(v_{R,\varepsilon,h}, y_{R,\varepsilon,h})$ as $R \to +\infty$ and $\varepsilon \to 0$

In this first experiment, our main aim will be to illustrate the convergence of $\rho_R^{-2}\mu_{R,\varepsilon}$ and $\rho_0^{-2}\psi_{R,\varepsilon}\,1_\omega$ as $R \to +\infty$ and $\varepsilon \to 0$, in the sense stated in Propositions 3.1 and 3.3.

Once the unique minimizer $(\mu_{R,\varepsilon,h}, \varphi_{T,R,\varepsilon,h})$ of $J^*_{R,\varepsilon,h}$ is obtained through the conjugate gradient algorithm described in Section 4.1, we compute the associated discrete adjoint solution $\psi_{R,\varepsilon,h}$ using a $\mathbb{P}_1$-finite element in space and a second order implicit scheme in time, as discussed in Section 4.2. The control is then given by $v_{R,\varepsilon,h} = \rho_0^{-2}\psi_{R,\varepsilon,h}\,1_\omega$. Finally, the controlled solution $y_{R,\varepsilon,h}$ is given by $y_{R,\varepsilon,h} = -\rho_R^{-2}\mu_{R,\varepsilon,h}$, in accordance with the optimality relations (25).

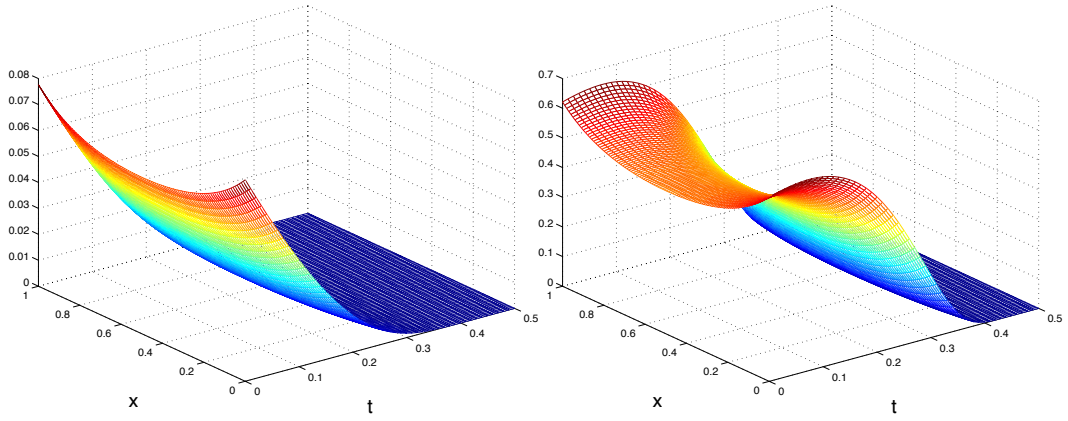

Figure 2: The weights ρ^{-2} and ρ_0^{-2} defined by (7), with $\beta_0 = \beta_{0,1/2}$ defined by (35), $K_1 = 1/10$ and $K_2 = 2\|\beta_0\|_{L^\infty(0,1)}$.

The function y_0 to be controlled is the eigenfunction of the Dirichlet-Laplace operator associated to the first eigenvalue, that is,

$$y_0(x) = \sin(\pi x) \quad \forall x \in (0,1). \tag{37}$$

With this initial condition and (36), the uncontrolled solution is given by

$$y(x,t) \equiv e^{-a_0\pi^2(T-t)} y_0(x),$$

so that $\|y(\cdot,T)\|_{L^2(0,1)} \approx 4.32 \times 10^{-1}$. Finally, we take $\omega = (0.3, 0.6)$ and $\beta_0 = \beta_{0,0.45}$ in (7), with $K_1 = 1/10$ and $K_2 = 2\|\beta_0\|_{L^\infty(0,1)} = 2$.

First, for $h = (1/100, 1/100)$, we show the behavior of the norms of $\mu_{R,\varepsilon,h}$ and $\varphi_{T,R,\varepsilon,h}$ with respect to ε and R. For each value of these parameters, we use the conjugate gradient algorithm with $\epsilon = 10^{-4}$. This is small enough to guarantee a good approximation of the control but, obviously, does not allow to fulfill exactly the optimality conditions (25). Notice however that the fact that ϵ (and h) is strictly positive allows to consider the limit cases $R = +\infty$ (for which $\rho_R^2 = \rho^2$) and $\varepsilon = 0$, as well.

The algorithm is initialized with $(\mu^0, \varphi_T^0) = (0,0)$ on $Q_T \times (0,1)$.

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-8}$	$\varepsilon = 0$
$R = 10^2$	1.3150	1.9209	2.5616	2.9020	2.9424
$R = 10^4$	1.8640	2.0898	2.6046	2.9043	2.9671
$R = 10^6$	n.c.	2.5886	2.6811	2.9263	2.9951
$R = 10^8$	n.c.	2.7376	2.7486	2.9374	2.9917
$R = +\infty$	n.c.	2.7405	2.7532	2.9368	2.9925

Table 1: $h = (10^{-2}, 10^{-2})$, $\omega = (0.3, 0.6)$, $y_0(x) \equiv \sin(\pi x)$. The $L^2(q_T)$-norm of $\rho_0^{-2}\psi_{R,\varepsilon,h}$ vs. ε and R (n.c. stands for non convergence).

Tables 1 and 2 report the norms $\|\rho_0^{-2}\psi_{R,\varepsilon,h}\|_{L^2(q_T)}$ and the norms $\|\rho_R^{-2}\mu_{R,\varepsilon,h}\|_{L^2(Q_T)}$ for $\varepsilon \in \{10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 0\}$ and $R \in \{10^2, 10^4, 10^6, 10^8, +\infty\}$. We check that these norms

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-8}$	$\varepsilon = 0$
$R = 10^2$	2.2383	1.9286	1.8168	1.7961	1.7988
$R = 10^4$	1.9241	1.8613	1.8016	1.7867	1.7899
$R = 10^6$	n.c.	1.7931	1.7919	1.7858	1.7885
$R = 10^8$	n.c.	1.7889	1.7878	1.7860	1.7879
$R = +\infty$	n.c.	1.7888	1.7876	1.7860	1.7879

Table 2: $h = (10^{-2}, 10^{-2})$, $\omega = (0.3, 0.6)$, $y_0(x) \equiv \sin(\pi x)$. The $L^2(Q_T)$-norm of $-\rho_R^{-2}\mu_{R,\varepsilon,h}$ ($\times 10^{-1}$) vs. R andε.

are uniformly bounded with respect to ε and R and both possess a limit as $\varepsilon \to 0$ and $R \to \infty$, in agreement with Propositions 3.1 and 3.3.

For small values of ε, we observe that the parameter R has a weak influence on the norm of $\rho_0^{-2}\psi_{R,\varepsilon,h}$; see columns $\varepsilon = 10^{-8}$ and $\varepsilon = 0$ in Table 1. Conversely, as soon as R is large enough (near $R = 10^6$), the norm of $\rho_R^{-2}\mu_{R,\varepsilon,h}$ is almost independent of ε; see Table 2. This is of course due to the choice of the weights ρ and ρ_0 .

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-8}$	$\varepsilon = 0$
$R = 10^2$	64	88	994	2 371	3 715
$R = 10^4$	434	338	1 686	6 194	12 073
$R = 10^6$	n.c.	1 772	1 710	2 467	4 261
$R = 10^8$	n.c.	2 737	2 524	4 519	6 196
$R = +\infty$	n.c.	2 759	2 892	4 451	6 278

Table 3: $h = (10^{-2}, 10^{-2})$, $\omega = (0.3, 0.6)$, $y_0(x) \equiv \sin(\pi x)$. The number of iterates to reach $\|g_h^n\|_V / \|g_h^0\|_V \leq \epsilon = 10^{-4}$ vs. R and ε.

Table 3 provides the number of iterates needed to achieve $\|g_h^n\|_V / \|g_h^0\|_V \leq \epsilon = 10^{-4}$, where g_h^n is the gradient of $J^*_{R,\varepsilon,h}$. In accordance with the results and conclusions in [3, 16], this number increases as $\varepsilon \to 0$ and/or $R \to +\infty$, which must be viewed as a numerical confirmation of the lack of uniform coercivity of $J^*_{R,\varepsilon}$ in V.

As soon as ϵ is small enough, depending on a_0 , T and the size of ω, the conjugate algorithm fails to converge. We also observe that the algorithm does not converge when ε and R are large, for instance when $\varepsilon = 10^{-2}$ and $R = 10^6$. This suggests that the ratio R versus ε^{-1}, both acting as penalty parameters for the null controllability constraint, should not be too large.

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-8}$	$\varepsilon = 0$
$R = 10^2$	1.03×10^1	7.456×10^0	7.530×10^0	7.802×10^0	7.816×10^0
$R = 10^4$	1.455×10^2	1.030×10^2	5.101×10^1	3.410×10^1	3.411×10^1
$R = 10^6$	n.c.	8.981×10^2	5.125×10^2	4.248×10^2	5.005×10^2
$R = 10^8$	n.c.	1.693×10^3	1.242×10^3	7.305×10^2	1.062×10^3
$R = +\infty$	n.c.	1.720×10^3	1.301×10^3	7.325×10^2	1.079×10^3

Table 4: $h = (10^{-2}, 10^{-2})$, $\omega = (0.3, 0.6)$, $y_0(x) \equiv \sin(\pi x)$. The $L^2(Q_T)$-norm of $\mu_{R,\varepsilon,h}$ vs. R and ε.

In agreement with the lack of uniform coercivity in V, the results in Tables 4 and 5 seem to

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-8}$	$\varepsilon = 0$
$R = 10^2$	8.034×10^0	1.475×10^2	9.220×10^2	7.266×10^3	1.192×10^4
$R = 10^4$	2.092×10^0	1.122×10^2	8.784×10^2	5.732×10^3	1.164×10^4
$R = 10^6$	n.c.	1.726×10^1	7.339×10^2	5.393×10^3	1.102×10^4
$R = 10^8$	n.c.	5.254×10^0	5.294×10^2	5.171×10^3	8.324×10^3
$R = +\infty$	n.c.	5.133×10^0	5.154×10^2	5.154×10^3	8.404×10^3

Table 5: $h = (10^{-2}, 10^{-2})$, $\omega = (0.3, 0.6)$, $y_0(x) \equiv \sin(\pi x)$. The $L^2(0,1)$-norm of $\varphi_{T,R,\varepsilon,h}$ vs. R and ε.

indicate that $\mu_{R,\varepsilon,h}$ is not uniformly bounded in $L^2(Q_T)$ with respect to R and $\varphi_{T,R,\varepsilon,h}$ is not uniformly bounded in $L^2(0,1)$ with respect to ε. Contrarily, we observe that the norm of $\mu_{R,\varepsilon,h}$ is bounded with respect to ε and the norm of $\varphi_{T,R,\varepsilon,h}$ is bounded with respect to R (Tables 4 and 5). This highlights the influence of the weights on the multiplier $\varphi_{T,R,\varepsilon,h}$ and will be discussed more in depth in the next Section.

At the limit as $\varepsilon \to 0$, the L^2-norm of $\varphi_{T,R,\varepsilon}$, which can be viewed as the multiplier associated to the constraint $y(\cdot, T) = 0$, does not belong anymore to $L^2(0,1)$ (see Remark 2). We refer to [7] where a direct approach allows to satisfy the constraint $y_h(\cdot, T) = 0$ exactly (which corresponds to $\epsilon = 0$) and to observe arbitrarily large values of the L^2-norm of the multiplier.

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-8}$	$\varepsilon = 0$
$R = 10^2$	8.03×10^{-2}	1.47×10^{-2}	9.49×10^{-4}	4.61×10^{-5}	4.67×10^{-5}
$R = 10^4$	2.09×10^{-2}	1.12×10^{-2}	9.09×10^{-4}	9.06×10^{-5}	4.69×10^{-5}
$R = 10^6$	n.c.	1.71×10^{-3}	7.28×10^{-4}	8.76×10^{-5}	4.68×10^{-5}
$R = 10^8$	n.c.	5.26×10^{-4}	5.22×10^{-4}	8.25×10^{-5}	4.66×10^{-5}
$R = +\infty$	n.c.	5.14×10^{-4}	5.09×10^{-4}	8.33×10^{-5}	4.63×10^{-5}

Table 6: $h = (10^{-2}, 10^{-2})$, $\omega = (0.3, 0.6)$, $y_0(x) \equiv \sin(\pi x)$. The $L^2(0,1)$-norm of $y_h(\cdot, T)$ vs. R and ε.

Table 6 depicts the L^2-norm of the computed state at time T. Note that this solution satisfies $y_h(\cdot, 0) = y_{0h}$ and *a priori* differs from the function $-\rho_{R,h}^{-2}\mu_{R,\varepsilon,h}$. As expected, the weight ρ_R^2 reinforces slightly the null controllability constraint (2) as R increases.

These Tables suggest that it is not actually necessary to take $R = +\infty$ and $\varepsilon = 0$ to achieve a good approximation of the controls. Due to the weights, the norms of the computed controls and controlled solutions change only slightly with respect to these parameters. The singular case $R = +\infty$ and $\varepsilon = 0$ ensures a better approximation of the null controllability requirement, but leads to a significative increase of iterates, as the coercivity of $J^*_{R,\varepsilon}$ is lost.

Figure 3 depicts the computed controlled solution and control in Q_T for $R = 10^8$ and $\varepsilon = 10^{-8}$. In Figure 4, we have tried to illustrate the interplay between the adjoint solution $\varphi_{T,R,\epsilon,h}$ and the controlled solution $y_{R,\varepsilon,h}(\cdot, T)$ at time T for three pairs of (R, ε).

5.2 Experiment 2: The constant diffusion case for $R = +\infty$ and $\varepsilon = 0$

In order to understand the role played by the weights ρ and ρ_0, we will now take $R = +\infty$ and $\varepsilon = 0$. At the numerical level, this limit case still makes sense since, for any $h > 0$, the minimizer of $J^*_{+\infty,0,h}$ obtained with the conjugate gradient method depends on the stopping parameter ϵ and does not

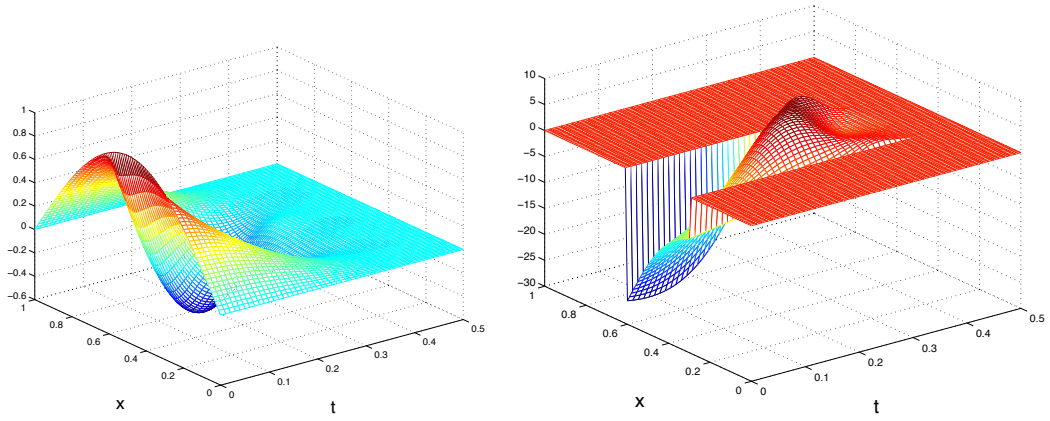

Figure 3: The state (**Left**) and the control (**Right**) with $\omega = (0.3, 0.6)$ for $R = 10^8$, $\varepsilon = 10^{-8}$ and $h = (10^{-2}, 10^{-2})$.

actually satisfy the constraint $y_h(\cdot\,, T) = 0$ exactly.

We will take a larger domain control, which allows to reach gradients closer to zero (i.e. to prescribe lower values of ϵ). More precisely, we will fix $\omega = (0.2, 0.8)$. This domain was also used in the numerical experiments in [7] and in [16] in a similar context, with $\rho \equiv 0$ and $\rho_0 \equiv 1$. On the other hand, we will fix again T and a as in (36) and y_0 as in 37. Finally, in (7) we take $\beta_0 = \beta_{0,1/2}$, $K_1 = 1/10$ and $K_2 = 2\|\beta_0\|_{L^\infty(0,1)} = 2$. The corresponding weights ρ^{-2} and ρ_0^{-2} are displayed in Figure 2.

In Table 7, we collect some relevant results obtained after 1 500 iterates of the conjugate gradient method initialized with $\mu^0 = 0$ and $\varphi_T^0 = 0$. The behavior of the method is shown for various meshsizes $h = (\Delta x, \Delta t)$. In particular, the convergence of the control v_h as $h \to 0$ becomes clear. The minimizer $(\mu_h, \varphi_{T,h})$ of $J^*_{+\infty,0,h}$ is plotted in Figure 5 for $h = (10^{-2}, 10^{-2})$. The associated v_h and y_h are plotted in Figure 6.

The approximations we obtain are very similar to those displayed in [7] with a primal method. Also, as a consequence of the behavior of the weights, we see in particular that the control v_h and the solution y_h are very close to zero near the controllability time T.

The evolution in $\log_{10}$ scale of the relative residue $r_h^n \equiv \|g_h^n\|_V / \|g_h^0\|_V$ is displayed in Figure 7. After 1 500 iterations, the relative residue is of the order $\mathcal{O}(10^{-5})$ and the corresponding $\|y_h(\cdot\,, T)\|_{L^2(0,1)}$ is of the order $\mathcal{O}(10^{-6})$, which is a good approximation of the null controllability requirement.

As is usual for ill-posed parabolic problems, the evolution of the residue is not linear. The slop reduces significantly after the first iterations. We check however that the residue remains decreasing, in contrast with what may be observed if we use a steepest descent algorithm (see 31). For instance, after 30 000 iterations, we get $r_h^{30000} \approx 4.17 \times 10^{-7}$ and $\|y_h(\cdot\,, T)\|_{L^2(0,1)} \approx 1.79 \times 10^{-7}$.

Let us now try to expain why the weights ρ and ρ_0 are crucial for a good numerical resolution of the controllability problem. To this end, let us consider the numerical results collected in Table 8.

We first observe that the norms $\|v_h\|_{L^2(q_T)}$ and $\|y_h\|_{L^2(Q_T)}$ are almost constant, contrarily to $\|y_h(\cdot\,, T)\|_{L^2(0,1)}$. The variations of the quantities $\|\mu_h\|_{L^2(q_T)}$ and $\|\varphi_{T,h}\|_{L^2(0,1)}$ are also small. This property, which partially explains the behavior of the algorithm, is a consequence of the fact that the weights ρ^{-2} and ρ_0^{-2} vanish exponentially as $t \to T^-$.

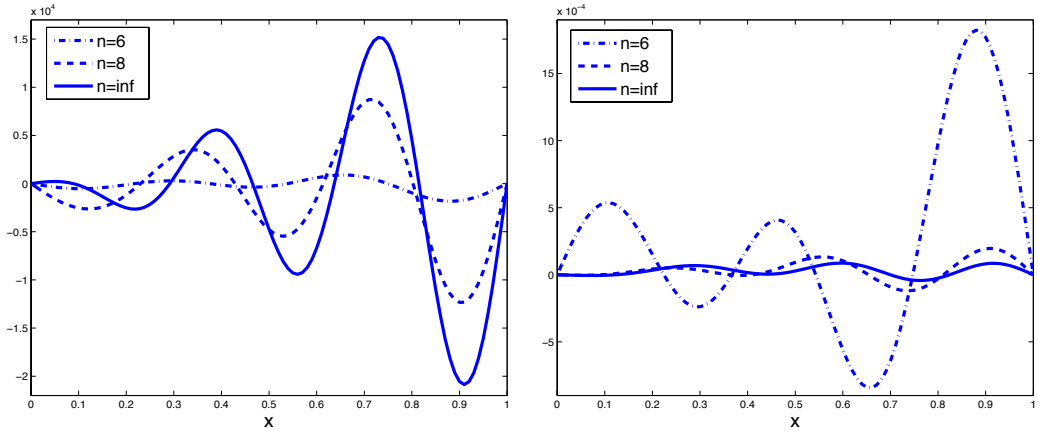

Figure 4: The functions $\varphi_{T,R,\varepsilon,h}$ (**Left**) and $y_{R,\varepsilon,h}(\cdot\,,T)$ (**Right**) with $\omega = (0.3, 0.6)$ for $R = 10^n$, $\varepsilon = 10^{-n}$ $(n = 6, 8, +\infty)$ and $h = (10^{-2}, 10^{-2})$.

$\Delta x = \Delta t$	1/25	1/50	1/100	1/200
$\|y_h(\cdot\,,T)\|_{L^2(0,1)}$	2.42×10^{-4}	8.46×10^{-6}	6.40×10^{-6}	8.64×10^{-6}
$\|\rho^{-2}\mu_h + y_h\|_{L^2(Q_T)}$	5.197×10^{-4}	4.229×10^{-4}	4.11×10^{-4}	3.139×10^{-4}
$\|v_h\|_{L^2(q_T)}$	1.283	1.273	1.292	1.302
$\|y_h\|_{L^2(Q_T)}$	1.905×10^{-1}	1.893×10^{-1}	1.886×10^{-1}	1.882×10^{-1}
$\|(\rho^{-2}\mu_h + y_h)(\cdot\,,0)\|_{L^2(0,1)}$	5.01×10^{-4}	5.35×10^{-5}	1.28×10^{-5}	2.59×10^{-6}
$\|\psi_h\|_{L^2(Q_T)}$	1.97×10^{1}	3.51×10^{1}	4.35×10^{1}	4.45×10^{1}
$\|\varphi_{T,h}\|_{L^2(0,1)}$	1.32×10^{2}	2.74×10^{2}	3.00×10^{2}	3.05×10^{2}
$\|\mu_h\|_{L^2(Q_T)}$	5.38×10^{1}	4.60×10^{1}	4.57×10^{1}	6.86×10^{1}

Table 7: $\omega = (0.2, 0.8)$, $y_0(x) \equiv \sin(\pi x)$. Numerical results obtained after 1 500 iterations of the conjugate gradient algorithm.

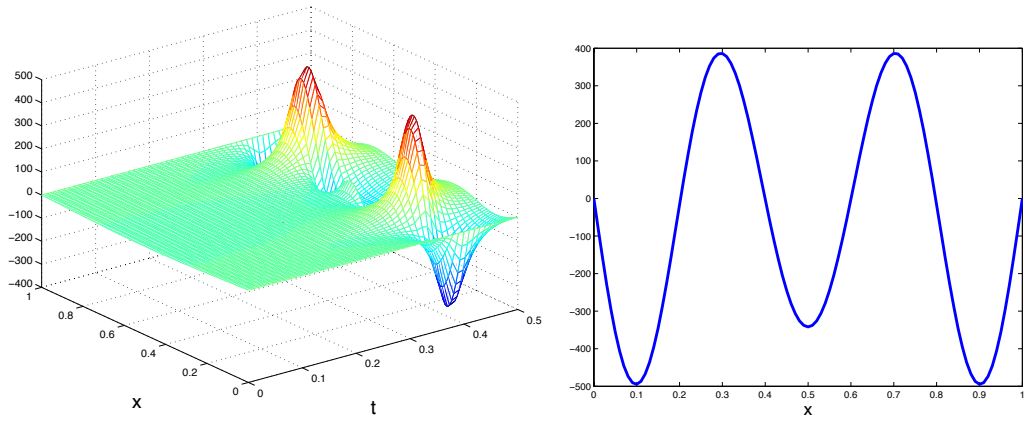

Figure 5: The minimizer μ_h (**Left**) and $\varphi_{T,h}$ (**Right**) of J_h^* after 1 500 iterations of the conjugate gradient algorithm with $\omega = (0.2, 0.8)$ for $h = (10^{-2}, 10^{-2})$.

When we look for the minimal L^2-norm null control, that is, when we try to solve a problem like (3) with $\rho \equiv 0$ and $\rho_0 \equiv 1$, the conjugate gradient algorithm for the associated dual problem, which is much more sensitive to the numerical approximation, behaves very differently. This was discussed in length in [16]. In Figure 8-Left, we have displayed the evolution of the residue for $h = (10^{-2}, 10^{-2})$ and $\epsilon = 10^{-8}$ in this case.

The chaotic evolution, firstly observed in [3], illustrates the lack of uniform coercivity of the dual functional with respect to h in $L^2(0,1)$. Thus, for any fixed small h, the minimizer $\varphi_{T,h}$, which is depicted in Figure 8-Right, exhibits an oscillatory behavior (specially in $\Omega \backslash \omega$) and possesses high frequency components.

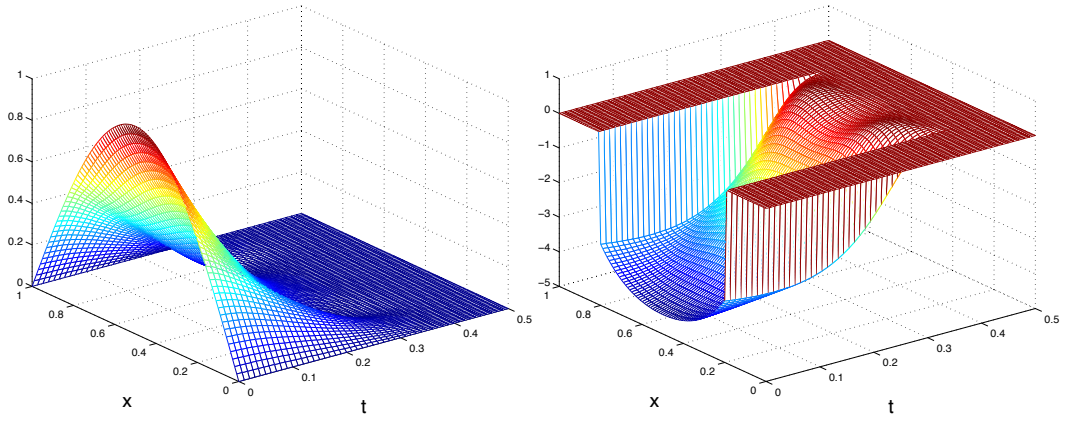

Figure 6: The associated controlled solution y_h (**Left**) and control v_h (**Right**).

For $\epsilon = 10^{-8}$, the algorithm converges after 485 iterations and leads to $\|y_h(\cdot\,, T)\|_{L^2(0,1)} \approx 1.69 \times 10^{-9}$, $\|\varphi_{T,h}\|_{L^2(0,1)} \approx 1.37 \times 10^{-6}$ and $\|v_h\|_{L^2(0,1)} \approx 1.049$.

When the discretization step h becomes too small, the conjugate algorithm fails to converge. Table 9 collects some values in function of the stopping parameter ϵ and allows to appreciate once again the balance between the norms $\|y_h(\cdot\,, T)\|_{L^2(0,1)}$ and $\|\varphi_{T,h}\|_{L^2(0,1)}$.

The algorithm behaves quite differently when it is applied to the dual formulation of (3) with $\rho \equiv 0$ and ρ_0 given by (7), which forces the control v_h, defined by $v_h = \rho_0^{-2}\varphi_h\, 1_\omega$, to vanish in a neighborhood of T. The residue and the minimizer $\varphi_{T,h}$ are displayed in Figure 9 for $\Delta x = \Delta t = 10^{-2}$.

The evolution is still very irregular. However, the number of iterates needed to reach convergence (here $\epsilon = 10^{-8}$) is lower (387 vs. 485) and, as discussed in [16] for a similar compactly supported function ρ_0^{-1}, this number does not blow up but decreases as $h \to 0$. For $\Delta x = \Delta t = 1/200$, the algorithm needs 172 iterations to achieve $r_h^n \leq \epsilon = 10^{-8}$ (to be compared with the 1 612 iterates required when $\rho \equiv 0$ and $\rho_0 \equiv 1$, see Figure 10).

From Figure 9-Right, we observe that the minimizer $\varphi_{T,h}$ is now smoother, and therefore easier to capture numerically. This suggests that the weight ρ_0 has the effect to filter out the high frequency components of the "initial" condition (at time $t = T$) of the backwards problem.

Table 10 collects some values with respect to the stopping parameter ϵ. As in Table 9, opposite (reciprocal) behaviors are observed for $\|y_h(\cdot\,, T)\|_{L^2(0,1)}$ and $\|\varphi_{T,h}\|_{L^2(0,1)}$. We also check that the L^2-norm of the control is greater (but still smaller than the norms obtained with positive weights ρ and ρ_0, see Table 7).

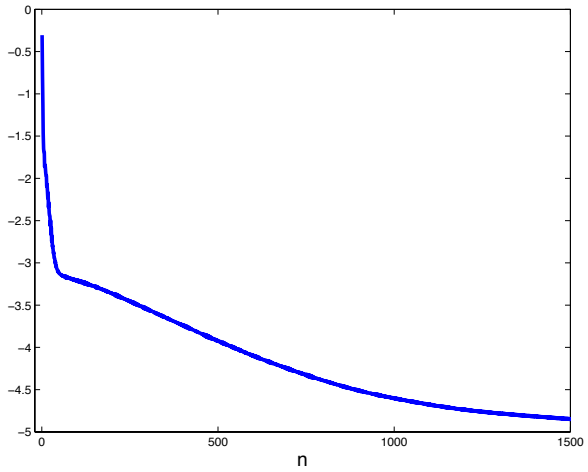

Figure 7: Evolution of $\log_{10}(r_h^n)$ with respect to the iteration number n with $\omega = (0.2, 0.8)$ and $y_0(x) \equiv \sin(\pi x)$ for $h = (10^{-2}, 10^{-2})$.

Let us come back to the case where ρ and ρ_0 are given by 7. Figure 5 suggests that the presence of ρ accentuates the regularization effect. Thus, for $\|y_h(\cdot\,, T)\|_{L^2(0,1)}$ of the order $\mathcal{O}(10^{-7})$, we get now $\|\varphi_{T,h}\|_{L^2(0,1)}$ of the order $\mathcal{O}(10^3)$ (see Table 8), while $\|\varphi_{T,h}\|_{L^2(0,1)} = \mathcal{O}(10^4)$ when we take $\rho \equiv 0$ and ρ_0 is given by (7).

However, this does not mean at all that the optimal minimizer $\varphi_{T,h}$ is uniformly bounded with respect to h in $L^2(0,1)$. Actually, the primal methods introduced in [7] allowed to achieve the null controllability requirement with a very good precision, typically $\|y_h(\cdot\,, T)\|_{L^2(0,1)} = \mathcal{O}(10^{-13})$, and showed nonzero high frequencies components for the minimizer $\varphi_{T,h}$. The key point here is that these high frequencies components have a very weak influence on the value of the minimum, because they are damped out in a small time interval, from T to, say, $T - \delta$, where the inverse ρ^{-1}, ρ_0^{-1} of the weights are almost zero. Thus, in Table 7, even if μ_h and ψ_h do not converge in $L^2(Q_T)$, the weighted functions $\rho_h^{-2}\mu_h$ and $\rho_{0,h}^{-2}\psi_h$ do. This weak influence with respect to the high frequencies explains the weak convergence of the algorithm, after the first iterates, once the lower frequency modes of $\varphi_{T,h}$ have been obtained.

♯ CG iterates	5×10^3	1×10^4	2×10^4	3×10^4
$\|y_h(\cdot\,, T)\|_{L^2(0,1)}$	3.05×10^{-6}	1.31×10^{-6}	3.16×10^{-7}	1.79×10^{-7}
$\|\rho^{-2}\mu_h + y_h\|_{L^2(Q_T)}$	1.19×10^{-4}	4.62×10^{-5}	2.04×10^{-5}	1.51×10^{-5}
$\|v_h\|_{L^2(q_T)}$	1.295	1.296	1.297	1.298
$\|y_h\|_{L^2(Q_T)}$	1.884×10^{-1}	1.883×10^{-1}	1.883×10^{-1}	1.883×10^{-1}
$\|(\rho^{-2}\mu_h + y_h)(\cdot\,, 0)\|_{L^2(0,1)}$	9.62×10^{-9}	3.70×10^{-10}	6.48×10^{-11}	2.54×10^{-13}
$\|\psi_h\|_{L^2(Q_T)}$	5.66×10^{1}	9.53×10^{1}	1.48×10^{2}	1.63×10^{2}
$\|\varphi_h(\cdot\,, T)\|_{L^2(0,1)}$	4.32×10^{2}	8.53×10^{2}	1.407×10^{3}	1.56×10^{3}
$\|\mu\|_{L^2(Q_T)}$	1.25×10^{2}	2.49×10^{2}	3.904×10^{2}	4.76×10^{2}
r_h^n	6.56×10^{-6}	2.81×10^{-6}	6.89×10^{-7}	4.17×10^{-7}

Table 8: $\omega = (0.2, 0.8)$, $y_0(x) \equiv \sin(\pi x)$. Numerical results obtained for $h = (10^{-2}, 10^{-2})$ for various numbers of conjugate gradient iterates.

♯ Stopping parameter ϵ	10^{-6}	10^{-7}	10^{-8}	10^{-9}
♯ CG iterates	277	453	485	1 062
$\|y_h(\cdot,T)\|_{L^2(0,1)}$	1.93×10^{-7}	3.68×10^{-8}	1.69×10^{-9}	4.11×10^{-10}
$\|v_h\|_{L^2(q_T)}$	1.026	1.049	1.051	1.072
$\|y_h\|_{L^2(Q_T)}$	2.48×10^{-1}	2.45×10^{-1}	2.45×10^{-1}	2.41×10^{-1}
$\|\varphi_{T,h}\|_{L^2(0,1)}$	6.70×10^{4}	1.37×10^{6}	1.77×10^{6}	3.87×10^{7}

Table 9: $h=(10^{-2},10^{-2})$, $\omega=(0.2,0.8)$, $y_0(x)\equiv\sin(\pi x)$. Numerical results for various stopping test parameters ϵ and $(\rho,\rho_0)\equiv(0,1)$.

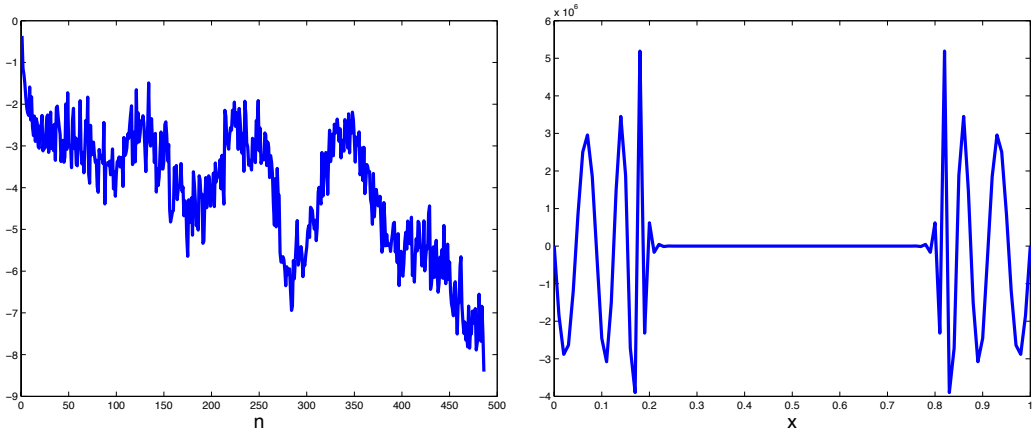

Figure 8: Evolution of the residue (in $\log_{10}$-scale) (**Left**) and $\varphi_{T,h}$ (**Right**) on $(0,1)$ with $\omega=(0.2,0.8)$ and $y_0(x)\equiv\sin(\pi x)$, for $(\rho,\rho_0)\equiv(0,1)$ and $h=(10^{-2},10^{-2})$.

5.3 Experiment 3: Controllability properties for different ω

This Section is devoted to discuss the sensitivity of the numerical controllability with respect to the support ω. We take $R=+\infty$, $\varepsilon=0$ and T and a given by (36). We will assume that the initial data to be controlled is now

$$y_0(x)=1,\quad x\in(0,1). \tag{38}$$

For any $L\in(0,1)$ and any $n\in\mathbb{N}$, we define the multiply connected set $\omega_n\subset(0,1)$ as follows:

$$\omega_n=\bigcup_{k=1}^{n}\left(d_n+(k-1)\left(\frac{L}{n}+d_n\right),k\left(\frac{L}{n}+d_n\right)\right),\quad d_n=\frac{1-L}{n+1}\,. \tag{39}$$

The Lebesgue measure of ω_n is equal to L for all n. Moreover, ω_n is composed of n disjoints components equi-distributed along $(0,1)$. For instance, with $L=0.4$, we have $\omega_2=\omega_{2,1}\cup\omega_{2,2}=(0.2,0.4)\cup(0.6,0.8)$ and $\mathrm{dist}(\{0\},\omega_{2,1})=\mathrm{dist}(\omega_{2,1},\omega_{2,2})=\mathrm{dist}(\{1\},\omega_{2,2})=0.2$.

We report in Table 11 the numerical results corresponding to various control supports ω_n for $\Delta x=\Delta t=1/250$ and $\epsilon=10^{-5}$.

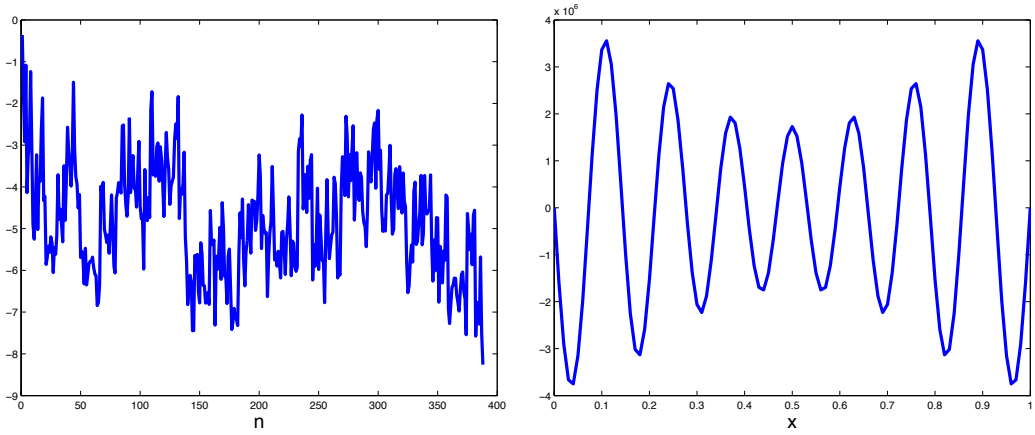

Figure 9: Evolution of the residue (in $\log_{10}$-scale) (**Left**) and $\varphi_{T,h}$ (**Right**) on $(0,1)$ with $\omega = (0.2, 0.8)$ and $y_0(x) \equiv \sin(\pi x)$, for $\rho \equiv 0$ and ρ_0 given by (7) and $h = (10^{-2}, 10^{-2})$.

♯ Stopping parameter ϵ	10^{-6}	10^{-7}	10^{-8}	10^{-9}
♯ CG iterations	26	143	387	543
$\|y_h(\cdot, T)\|_{L^2(0,1)}$	3.87×10^{-7}	1.54×10^{-8}	2.93×10^{-9}	3.81×10^{-10}
$\|v_h\|_{L^2(q_T)}$	1.236	1.247	1.250	1.256
$\|y_h\|_{L^2(Q_T)}$	2.01×10^{-1}	2.00×10^{-1}	2.00×10^{-1}	1.99×10^{-1}
$\|\varphi_{T,h}\|_{L^2(0,1)}$	1.45×10^{4}	4.91×10^{6}	1.95×10^{6}	2.809×10^{7}

Table 10: $h = (10^{-2}, 10^{-2})$, $\omega = (0.2, 0.8)$, $y_0(x) \equiv \sin(\pi x)$. Numerical results for various stopping test parameters ϵ, $\rho \equiv 0$ and ρ_0 given by (7).

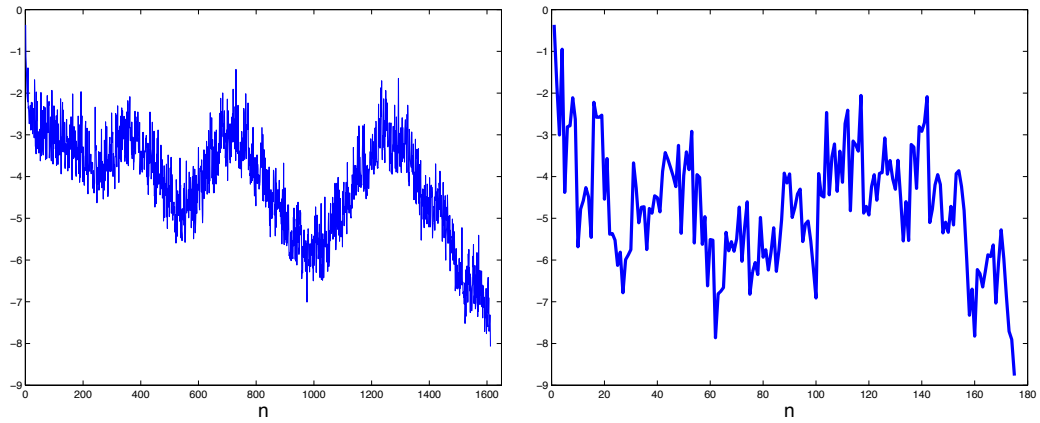

Figure 10: A comparison: evolution of the residues r_h^n (in $\log_{10}$-scale) vs. n for $(\rho, \rho_0) \equiv (0, 1)$ (**Left**) and for $\rho \equiv 0$ and ρ_0 done by (7) (**Right**). The data ω and y_0 are as in Figures 8 and 9 and $h = (1/200, 1/200)$.

We first observe that the L^2-norm of the control decreases with n, in agreement with the results in [15] where it is shown that, for y_0 as in (38), the support leading to the control of minimal L^2-norm is composed of an arbitrarily large number of disjoint components, uniformly distributed in $(0,1)$.

We also notice that not only the number of iterates needed to reach $\epsilon = 10^{-5}$ but also the norm in V of the corresponding minimizer $(\mu_h, \varphi_{T,h})$ are significantly reduced for large values of n.

We also report the numerical values corresponding to the control $\rho^{-2}\psi_h L\,1_{(0,1)}$, obtained when the characteristic function 1_{ω^n} is replaced by $L\,1_{(0,1)}$. This is the situation considered in [15] in the limit $n \to +\infty$, when the optimal distribution problem of ω over the spatial domain is relaxed.

Again, for any fixed n, $\varphi_{T,h}$ is not bounded in $L^2(Q_T)$ uniformly with respect to h. The uniform bound holds only at the limit in n, that is, after relaxation, when the control acts on the whole domain. Figure 11 shows the controlled solutions y_h corresponding to ω_2 and ω_{16}.

We have also checked that the conjugate algorithm allows to compute null controls for non-smooth initial conditions. In Figure 12, we present the controlled solutions for $y_0(x) = 1_D(x)$, where $D = (0.2, 0.4)$, with $\omega = (0.2, 0.4)$ and also with $\omega = (0.6, 0.8)$. As expected, the first case is more favorable, with $\|v_h\|_{L^2(q_T)} \approx 1.44$, than the second one, with $\|v_h\|_{L^2(q_T)} \approx 9.12$ (see Table 12).

	1_{ω_2}	1_{ω_4}	1_{ω_8}	$1_{\omega_{16}}$	$L1_\Omega$
♯ CG iterates	2 523	1 734	899	640	500
$\|y_h(\cdot,T)\|_{L^2(0,1)}$	4.65×10^{-6}	4.87×10^{-6}	3.68×10^{-6}	5.16×10^{-6}	1.32×10^{-7}
$\|\rho^{-2}\mu_h + y_h\|_{L^2(Q_T)}$	6.84×10^{-3}	6.67×10^{-3}	6.67×10^{-3}	6.64×10^{-3}	6.63×10^{-3}
$\|v_h\|_{L^2(q_T)}$	2.171	2.136	2.115	1.942	1.712
$\|y_h\|_{L^2(Q_T)}$	2.56×10^{-1}	2.63×10^{-1}	2.63×10^{-1}	2.64×10^{-1}	2.64×10^{-1}
$\|\psi_h\|_{L^2(Q_T)}$	9.23×10^{1}	1.52×10^{1}	1.21×10^{1}	1.12×10^{1}	1.10×10^{1}
$\|\varphi_{T,h}\|_{L^2(0,1)}$	6.34×10^{2}	8.54×10^{1}	4.31×10^{1}	2.56×10^{1}	1.78×10^{1}
$\|\mu_h\|_{L^2(Q_T)}$	1.42×10^{2}	3.24×10^{1}	1.95×10^{1}	1.67×10^{1}	1.58×10^{1}

Table 11: $y_0 \equiv 1$, $\epsilon = 10^{-5}$. Numerical results obtained with various $\omega_n \subset (0,1)$ for $h = (1/250, 1/250)$.

6 Concluding remarks

We have shown that appropriate duality techniques can be applied to solve (3).

The approach can be viewed as a nontrivial extension of the methods in [3] and [10], by introducing positive weights in the functional. These weights, that appear in related global Carleman inequalities, reinforce the null controllability condition and furnish controls that are smooth. They allow to decrease the number of iterates of the conjugate gradient algorithm needed to obtain controls for which $\|y_h(\cdot,T)\|_{L^2(0,1)} = \mathcal{O}(10^{-6})$, which is enough in practice.

In order to apply duality techniques to (3), we have to introduce new variables. One of them is the adjoint state of the heat equation at the final controllability time T. The resulting formulation of the problem is very useful to get theoretical controllability. However, in what concerns the heat equation, due to the regularization effect, the final adjoint state does not belong to $L^2(0,1)$, whatever be the weights. For this reason, the dual problem of (3) remains, at the level of finite dimension, severally ill-posed. This contrasts with some primal methods introduced and analyzed in [7], where the variables are y and ψ and the constraint (2) is satisfied exactly.

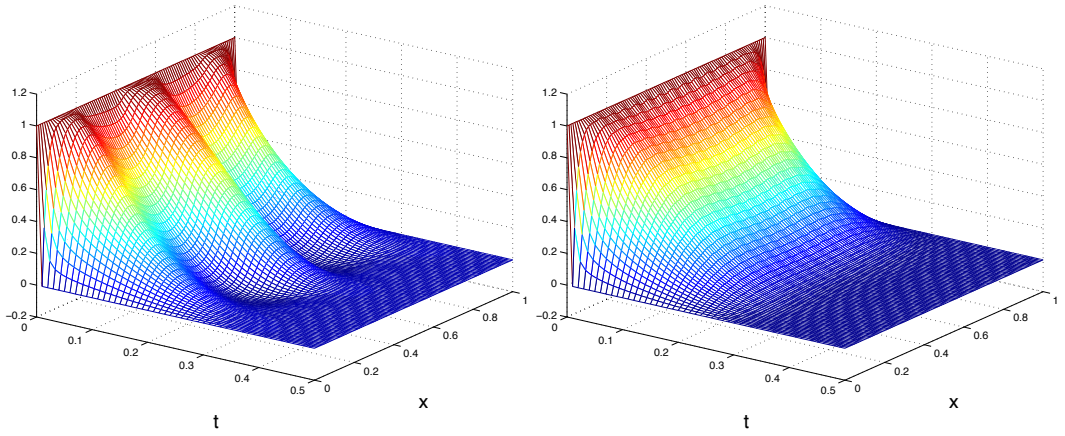

Figure 11: The controlled solutions corresponding to ω_2 (**Left**) and ω_{16} (**Right**), with $y_0 \equiv 1$, $\epsilon = 10^{-5}$ and $h = (10^{-2}, 10^{-2})$.

On the other hand, the dual approach allows to keep the time-dependent viewpoint and avoids the introduction of elliptic systems of dimension $N+1$ (N being the spatial dimension) and higher order, as in [7]. Therefore, since the related computational work is reasonable, it is expectable that dual methods can be adapted and extended to other more complex null controllability problems: N-dimensional heat equations, Stokes and Navier-Stokes systems, non-scalar diffusion systems, nonlinear system, etc.

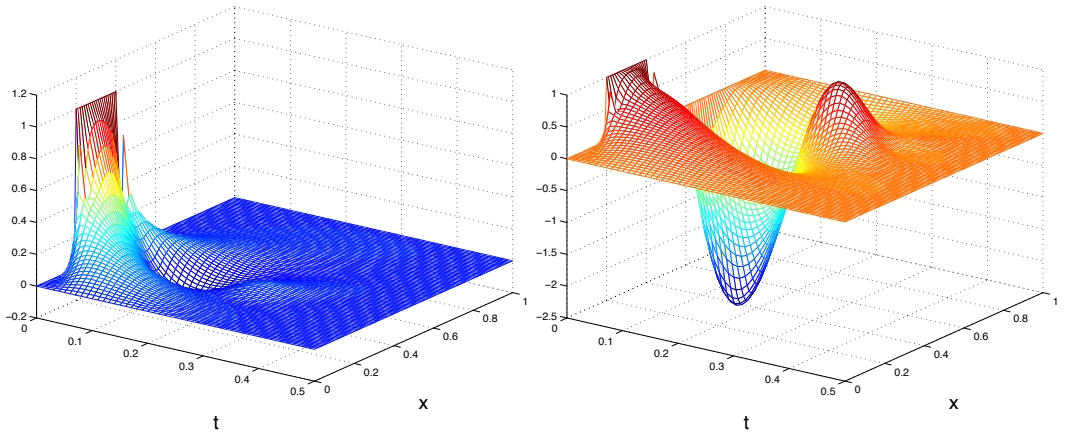

Figure 12: The controlled solutions corresponding to ω_2 (**Left**) and ω_{16} (**Right**), with $y_0(x) = 1_D(x), D = (0.2, 0.4)$, $\epsilon = 10^{-5}$ and $h = (10^{-2}, 10^{-2})$.

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.

	$\|y_h(\cdot,T)\|_{L^2(0,1)}$	$\|\varphi_{T,h}\|_{L^2(0,1)}$	$\|v_h\|_{L^2(q_T)}$	$\|\mu_h\|_{L^2(Q_T)}$	$\|y_h\|_{L^2(Q_T)}$
$\omega = (0.2, 0.4)$	3.29×10^{-5}	3.42×10^{3}	1.44	2.69×10^{2}	7.22×10^{-2}
$\omega = (0.6, 0.8)$	5.10×10^{-4}	4.99×10^{4}	9.12	4.25×10^{3}	4.22×10^{-1}

Table 12: $y_0(x) \equiv 1_D(x), D = (0.2, 0.4)$, $h = (10^{-2}, 10^{-2})$. Numerical results obtained after 10 000 iterates of the conjugate gradient algorithm.

[2] F. Boyer, F. Hubert and J. Le Rousseau, *Uniform null-controllability properties for space/time-discretized parabolic equations,* Preprint, 2009.

[3] 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.

[4] J-M. Coron, *Control and Nonlinearity*, Mathematical Surveys and Monographs, Volume 136, 2007.

[5] I. Ekeland, R. Temam, *Convex analysis and variational problems*, Classics in Applied Mathematics 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1999.

[6] W. Fenchel, *On conjugate convex functions*, Canad. J. Math., 1, 1949, p. 73–77.

[7] E. Fernandez-Cara and A. Münch, *Numerical null controllability of the 1D heat equation: primal algorithms,* Preprint (2009).

[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 and J.L. Lions, *Exact and approximate controllability for distributed parameter systems,* Acta Numerica (1996), 159–333.

[10] 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.

[11] S. Kindermann, *Convergence Rates of the Hilbert Uniqueness Method via Tikhonov regularization*, J. of Optimization Theory and Applications 103(3), (1999) 657-673.

[12] S. Labbé and E. Trélat, *Uniform controllability of semi-discrete approximations of parabolic control systems*, Systems and Control Letters 55 (2006) 597-609.

[13] G. Lebeau and L. Robbiano, *Contrôle exact de l'équation de la chaleur,* Comm. Partial Differential Equations 20 (1995), no. 1–2, 335–356.

[14] J-L. Lions, *Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués,* Recherches en Mathématiques Appliquées, Tomes 1 et 2, Masson, Paris 1988.

[15] A. Münch and F. Periago, *Optimal distribution of the internal null control for the 1D heat equation,* Preprint, 2009.

[16] A. Münch and E. Zuazua, *Numerical approximation of null controls for the heat equation through transmutation.* To appear in Inverse Problems (2010).

[17] E.T. Rockafellar, *Convex functions and duality in optimization problems and dynamics,* Lecture Notes Oper. Res. and Math. Ec., Vol. II, Springer, Berlin, 1969.

[18] D.L. Russell, *Controllability and stabilizability theory for linear partial differential equations. Recent progress and open questions,* SIAM Rev. 20, 1978, 639–739.

[19] E. Zuazua, *Control and numerical approximation of the wave and heat equations*, ICM2006, Madrid, Spain, Vol. III (2006) 1389–1417.