Numerical null controllability of semi-linear 1-D heat equations: fixed point, least squares and Newton methods

ENRIQUE FERNÁNDEZ-CARA* and ARNAUD MÜNCH†

May 22, 2011

Abstract

This paper deals with the numerical computation of distributed null controls for semi-linear 1D heat equations, in the sublinear and slightly superlinear cases. Under sharp growth assumptions, the existence of controls has been obtained in *[Fernandez-Cara & Zuazua, Null and approximate controllability for weakly blowing up semi-linear heat equation, 2000]* via a fixed point reformulation; see also *[Barbu, Exact controllability of the superlinear heat equation, 2000].* More precisely, Carleman estimates and Kakutani's Theorem together ensure the existence of solutions to fixed points for an equivalent fixed point reformulated problem. A nontrivial difficulty appears when we want to extract from the associated Picard iterates a convergent (sub)sequence. In this paper, we introduce and analyze a least squares reformulation of the problem; we show that this strategy leads to an effective and constructive way to compute fixed points. We also formulate and apply a Newton-Raphson algorithm in this context. Several numerical experiments that make it possible to test and compare these methods are performed.

Keywords: One-dimensional semi-linear heat equation, null controllability, blow up, numerical solution, least squares method.

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

1 Introduction. The null controllability problem

We are concerned in this paper with the null controllability problem for a semi-linear one-dimensional heat equation. The state equation is the following:

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

*Dpto. EDAN, University of Sevilla, Aptdo. 1160, 41080 Sevilla, Spain. E-mail: `cara@us.es`. Partially supported by grants MTM2006-07932 and MTM2010-15592 (DGI-MICINN, 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).

Here, $\omega \subset\subset (0,1)$ is a (small) non-empty open interval, 1_ω is the associated characteristic function, $T > 0$, $a \in C^1([0,1])$ with $a(x) \geq a_0 > 0$, $y_0 \in L^2(0,1)$, $v \in L^\infty(\omega \times (0,T))$ is the *control* and y is the associated *state*. We will assume that the function $f : \mathbb{R} \mapsto \mathbb{R}$ is, at least, locally Lipschitz-continuous.

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. The symbols C, C_0, C_1, etc. and K will be used to denote positive constants.

Following [14], we will always assume for simplicity that f satisfies

$$|f'(s)| \leq C(1 + |s|^m) \quad \text{a.e., with } 1 \leq m \leq 5. \tag{2}$$

Under this condition, (1) possesses exactly one local in time solution. Moreover, in accordance with the results in [5], under the growth condition

$$|f(s)| \leq C(1 + |s| \log(1 + |s|)) \quad \forall s \in \mathbb{R}, \tag{3}$$

the solutions to (1) are globally defined in $[0,T]$ and one has

$$y \in C^0([0,T]; L^2(0,1)) \cap L^2(0,T; H^1_0(0,1)). \tag{4}$$

The main goal of this paper is to analyze numerically the null controllability properties of (1), in particular when blow-up occurs. Recall that, without a growth condition of the kind (3), the solutions to (1) can blow up before $t = T$; in general, the blow-up time depends on the sizes of $\|y_0\|_{L^2(0,1)}$ and $\|a\|_{L^\infty}$.

The system (1) is said to be *null-controllable* at time T if, for any $y_0 \in L^2(0,1)$ and any globally defined bounded trajectory $y^\star \in C^0([0,T]; L^2(0,1))$ (corresponding to the data $y^\star_0 \in L^2(0,1)$ and $v^\star \in L^\infty(q_T)$), there exist controls $v \in L^\infty(q_T)$ and associated states y that are again globally defined in $[0,T]$ and satisfy (4) and

$$y(x,T) = y^\star(x,T), \quad x \in (0,1). \tag{5}$$

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 [24, 25, 22, 21, 7]. For semi-linear systems of this kind, the first contributions have been given in [26, 20, 9, 18, 15]. From the practical viewpoint, null controllability results are crucial since, roughly speaking, they make it possible to drive the state to an equilibrium exactly and, consequently, are associated to finite time work.

Let us recall briefly the main results of [14]; see also [1].

The first one states that, if f is "too super-linear" at infinity, then the control cannot compensate the blow-up phenomena occurring in $(0,1)\backslash\overline{\omega}$:

THEOREM 1.1 *There exist locally Lipschitz-continuous functions f with $f(0) = 0$ and*

$$|f(s)| \sim |s| \log^p(1 + |s|) \quad as \quad |s| \to \infty, \quad p > 2,$$

such that (1) *fails to be null-controllable for all $T > 0$.*

The second result provides conditions under which (1) is null-controllable:

THEOREM 1.2 *Let $T > 0$ be given. Assume that* (1) *admits at least one solution $y^\star$, globally defined in $[0,T]$ and bounded in Q_T. Assume that $f : \mathbb{R} \mapsto \mathbb{R}$ is locally Lipschitz-continuous and satisfies* (2) *and*

$$\frac{f(s)}{|s| \log^{3/2}(1+|s|)} \to 0 \quad as \quad |s| \to \infty.$$

Then (1) *is null-controllable at time T.*

In accordance with these results, if $|f(s)|$ grows like $|s| \log^p(1+|s|)$ as $|s| \to +\infty$ for some $p > 2$, (1) is not null-controllable in general at time T; contrarily, if $|f(s)|$ does not grow at infinity faster than $|s| \log^p(1+|s|)$ for some $p < 3/2$, then (1) is null-controllable. These results leave open the case where $|f(s)|$ behaves like $|s| \log^p(1+|s|)$ with $p \in [3/2, 2]$.

Let us also mention [8] in the context of Theorem 1.1 where a positive boundary controllability result is proved for a specific class of initial and final data and T large enough.

The goal of this paper is to design, analyze and implement numerical methods allowing to solve the previous null controllability problem, with the aim to clarify the role of the exponent p in the function f. In a similar linear situation, this has already been the objective of some important work, where it has been shown that the numerical computation of controls is difficult; see for instance [4, 16, 23, 19, 3].

This paper can be viewed as a continuation of our previous works [12] and [13], devoted to the numerical null controllability of the linear 1D heat equation. There, the starting point is a variational approach introduced and systematically employed by Fursikov and Imanuvilov in [15] in order to obtain controllability results for parabolic systems.

In the sequel, we will assume that $f(0) = 0$. We will be concerned only with the search for state-control pairs (y,v) satisfying (4) and (5) for $y^* \equiv 0$, that is:

$$y(x,T) = 0, \quad x \in (0,1). \tag{6}$$

Let us consider the non-empty set

$$\mathcal{C}(y_0,T) = \{\, (y,v) : v \in L^2(q_T), \quad y \text{ solves (1) and satisfies (6)}\,\}.$$

Following [12, 13], we would like to solve the following extremal problem:

$$\begin{cases} \text{Minimize} \ \ J(y,v) = \dfrac{1}{2} \displaystyle\iint_{Q_T} \rho^2 |y|^2 \, dx \, dt + \dfrac{1}{2} \displaystyle\iint_{q_T} \rho_0^2 |v|^2 \, dx \, dt \\ \text{Subject to} \ \ (y,v) \in \mathcal{C}(y_0,T) \end{cases} \tag{7}$$

Here, we assume (at least) that the weights ρ and ρ_0, that can blow up as $t \to T^-$, satisfy:

$$\begin{cases} \rho = \rho(x,t), \ \rho_0 = \rho_0(x,t) \text{ are continuous and } \geq \rho_* > 0 \text{ in } Q_T \\ \rho, \rho_0 \in L^\infty(Q_{T-\delta}) \quad \forall \delta > 0 \end{cases} \tag{8}$$

This would furnish (very appropriate) state-control pairs that solve the null controllability problem.

However, we will see later that all what we are able to obtain is just pairs $(y,v) \in \mathcal{C}(y_0,T)$ satisfying $J(y,v) < +\infty$. This will be sufficient for our purposes.

The plan of the paper is the following.

In Section 2, we review some results from [12] and [13] concerning the numerical solution of the null controllability problem for linear 1D parabolic equations. Section 3 is devoted to provide several strategies concerning the general (semi-linear) case: first, we present an approach that relies on a standard fixed point formulation of the null controllability problem; then, we introduce a least squares reformulation and we apply appropriate gradient techniques; finally, a Newton-Raphson strategy is also considered. In Section 4, we perform several numerical experiments that make it possible to test and compare these methods. In particular, we exhibit some data for which least squares algorithms converge, contrarily to what happens to the fixed point and Newton-Raphson approaches. Finally, some conclusions and perspectives are indicated in Section 5.

Let us end this Introduction with two comments:

- First, to our knowledge, this is the first work addressing the numerical approximation of null controls for (general) semi-linear heat equations, at least in the slightly super-linear case.
- Secondly, it can be observed that all the ideas and arguments below, with the possible exception of the theoretical results in Section 3.2, can be extended to the N-dimensional setting, i.e. to semi-linear heat equations of the form

$$y_t - \nabla \cdot (A(x)\nabla y) + f(y) = v 1_\omega, \quad (x,t) \in \Omega \times (0,T),$$

  where $\Omega \subset \mathbf{R}^N$ is a regular, bounded, connected open set and $\omega \subset\subset \Omega$ is a non-empty subdomain.

2 A linear control problem

In this Section, we deal with the controllability properties of the following linear system

$$\begin{cases} y_t - (a(x)y_x)_x + A(x,t)y = v\,1_\omega + B(x,t), & (x,t) \in Q_T \\ y(x,t) = 0, & (x,t) \in \Sigma_T \\ y(x,0) = y_0(x), & x \in (0,1) \end{cases} \tag{9}$$

that arises naturally after linearization of (1).

We consider the extremal problem

$$\begin{cases} \text{Minimize} \quad J(y,v) = \dfrac{1}{2} \displaystyle\iint_{Q_T} \rho^2 |y|^2 \,dx\,dt + \dfrac{1}{2} \iint_{q_T} \rho_0^2 |v|^2 \,dx\,dt \\ \text{Subject to} \quad (y,v) \in \mathcal{C}_{lin}(y_0,T) \end{cases} \tag{10}$$

where $\mathcal{C}_{lin}(y_0,T)$ is the linear manifold

$$\mathcal{C}_{lin}(y_0,T) = \{\, (y,v) : v \in L^2(q_T),\ y \text{ solves (9) and satisfies (6)} \,\}$$

and we recall some related numerical methods. We assume that $A \in L^\infty(Q_T)$ and $B \in L^2(Q_T)$ and, also, that B vanishes at $t = T$ in an appropriate sense (see the condition (13) below).

In order to solve this problem, we can follow the ideas in [15]. Accordingly, we can adapt the results in [12] and [13], where primal and dual methods are respectively considered in the simpler case $A \equiv 0$, $B \equiv 0$.

Since we are looking for controls such that the associated states satisfy (5), it is a good idea to choose weights ρ and ρ_0 blowing up to $+\infty$ as $t \to T^-$. Indeed, this can be viewed as a reinforcement of the constraint (5); in fact, by prescribing ρ and ρ_0 we are even telling how fast $y(\cdot\,,t)$ and $v(\cdot\,,t)$ must decay to zero as $t \to T^-$.

There are "good" weight functions ρ and ρ_0 that blow up at $t = T$ and provide a very suitable solution to the null controllability problem. They were determined and systematically used by Fursikov and Imanuvilov [15] and are given as follows:

$$\left\{ \begin{array}{l} \rho(x,t) = \exp\left(\dfrac{\beta(x)}{T-t}\right), \ \rho_0(x,t) = (T-t)^{3/2}\rho(x,t), \ \beta(x) = K_1\left(e^{K_2} - e^{\beta_0(x)}\right) \\ \text{the } K_i \text{ are large positive constants (depending on } T,\ a_0,\ \|a\|_{C^1} \text{ and } \|A\|_\infty) \\ \text{and } \beta_0 \in C^\infty([0,1]),\ \beta_0 > 0 \text{ in } (0,1),\ \beta_0(0) = \beta_0(1) = 0,\ |\beta_0'| > 0 \text{ outside } \omega. \end{array} \right. \tag{11}$$

Together with ρ and ρ_0, we will need later another weights, ρ_1 and ρ_2, given by

$$\rho_1(x,t) \equiv (T-t)^{1/2}\rho(x,t) \ \text{ and } \ \rho_2(x,t) \equiv (T-t)^{-1/2}\rho(x,t). \tag{12}$$

It can be shown that, if

$$\iint_{Q_T} \rho_0^2 |B(x,t)|^2 \, dx \, dt < +\infty, \tag{13}$$

the extremal problem (7) is well-posed: for any $T > 0$ and any $y_0 \in L^2(0,1)$, there exists exactly one minimizer of J in $\mathcal{C}_{lin}(y_0,T)$.

2.1 Numerical solution via a primal method

We refer to [12] for the details related to this Section. Let us introduce the following notation :

$$L_A z = z_t - (a(x)z_x)_x + A(x,t)z\,, \quad L_A^* q = -q_t - (a(x)q_x)_x + A(x,t)q.$$

The roles of ρ and ρ_0 are clarified by the following arguments and results, which are mainly due to Fursikov and Imanuvilov [15].

First, let us set $P_0 = \{\, q \in C^2(\overline{Q}_T) : q = 0 \ \text{ on } \ \Sigma_T \,\}$. In this linear space, the bilinear form

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

is a scalar product. Let P be the completion of P_0 for this scalar product. Then P is a Hilbert space and the usual global Carleman estimates for the solutions to parabolic equations lead to the following result (see [12] and the references therein):

Lemma 2.1 *Let ρ and ρ_0 be given by (11). Then, for any $\delta > 0$, one has $P \hookrightarrow C^0([0,T-\delta]; H_0^1(0,1))$ and the embedding is continuous. In particular, there exists $C_0 > 0$, only depending on ω, T, a_0, $\|a\|_{C^1}$ and $\|A\|_\infty$, such that*

$$\|q(\cdot\,,0)\|^2_{H_0^1(0,1)} \le C_0 \left(\iint_{Q_T} \rho^{-2} |L_A^* q|^2 \, dx \, dt + \iint_{q_T} \rho_0^{-2} |q|^2 \, dx \, dt \right) \tag{14}$$

for all $q \in P$.

The primal methods in [12] rely on the following result.

PROPOSITION 2.1 *Assume that $a \in C^1([0,1])$ and let ρ and ρ_0 be given by (11). Let (y,v) be the corresponding optimal state-control pair. Then there exists $p \in P$ such that*

$$y = \rho^{-2} L_A^* p \equiv \rho^{-2}(-p_t - (a(x)p_x)_x + A(x,t)p), \quad v = -\rho_0^{-2} p|_{q_T}\,. \tag{15}$$

The function p is the unique solution to

$$\begin{cases} \displaystyle\iint_{Q_T} \rho^{-2} L_A^* p\, L_A^* q\,dx\,dt + \iint_{q_T} \rho_0^{-2} p\,q\,dx\,dt = \int_0^1 y_0(x)\,q(x,0)\,dx + \iint_{Q_T} B(x,t) q\,dx\,dt \\ \forall q \in P; \quad p \in P \end{cases} \tag{16}$$

The well-posedness of (16) is a consequence of Lax-Milgram Lemma and (14). This estimate justifies the specific choice of the weights that we have made.

Remark 1 In view of (15) and (16), the function p furnished by Proposition 2.1 solves in a weak (distributional) sense the following differential problem, that is second order in time and fourth order in space:

$$\begin{cases} L_A(\rho^{-2} L_A^* p) + \rho_0^{-2} p\,1_\omega = B, & (x,t) \in Q_T \\ p(x,t) = 0, \quad (\rho^{-2} L_A^* p)(x,t) = 0, & (x,t) \in \Sigma_T \\ (\rho^{-2} L_A^* p)(x,0) = y_0(x), \quad (\rho^{-2} L_A^* p)(x,T) = 0, & x \in (0,1). \end{cases}$$

Notice that the "boundary" conditions at $t = 0$ and $t = T$ appear in (16) as Neumann-like conditions. Also, notice that no information is obtained on $p(\cdot\,,T)$. □

In [12], in order to keep explicit the variable y, we rewrote (16) as an equivalent mixed variational problem: find $(y,p,\lambda) \in Z \times P \times Z$ such that

$$\begin{cases} \displaystyle\iint_{Q_T} \rho^2 y\,z\,dx\,dt + \iint_{q_T} \rho_0^{-2} p\,q\,dx\,dt + \iint_{Q_T} (L_A^* q - \rho^2 z)\,\lambda\,dx\,dt \\ \displaystyle\qquad = \int_0^1 y_0(x)\,q(x,0)\,dx + \iint_{Q_T} B\,q\,dx\,dt \quad \forall (z,q) \in Z \times P \\ \displaystyle\iint_{Q_T} (L_A^* p - \rho^2 y)\,\mu\,dx\,dt = 0 \quad \forall \mu \in Z \end{cases} \tag{17}$$

where

$$Z = L^2(\rho^2; Q_T) := \{\, z \in L^1_{\rm loc}(Q_T) : \iint_{Q_T} \rho^2 |z|^2\,dx\,dt < +\infty \,\}. \tag{18}$$

We have shown in [12] that this mixed formulation is well-posed.

Let us now present a second mixed formulation, equivalent to (17), that does not use unbounded weights. This will be particularly important at the numerical level.

The key idea is to perform the following change of variables:

$$\eta = \rho_0^{-1} p, \quad m = \rho y = \rho^{-1} L_A^\star(\rho_0 \eta).$$

Let us introduce the space $P_* := \rho_0^{-1}P$. Then, (17) is rewritten as follows: find $(m,\eta,\mu) \in L^2(Q_T)\times P_* \times L^2(Q_T)$ such that

$$\left\{\begin{aligned} &\iint_{Q_T} m\overline{m}dx\,dt + \iint_{q_T}\eta\,\overline{\eta}\,dx\,dt + \iint_{Q_T}\Big(\rho^{-1}L_A^\star(\rho_0\overline{\eta}) - \overline{m}\Big)\mu\,dx\,dt \\ &\qquad = \int_0^1 \rho_0(x,0)y_0(x)\overline{\eta}(x,0)dx + \iint_{Q_T}\rho_0 B\,\overline{\eta}\,dx\,dt \quad \forall(\overline{m},\overline{\eta})\in L^2(Q_T)\times P_* \\ &\iint_{Q_T}\Big(\rho^{-1}L_A^\star(\rho_0\eta) - m\Big)\overline{\mu}\,dx\,dt = 0, \quad \forall\overline{\mu}\in L^2(Q_T)\end{aligned}\right. \tag{19}$$

Proceeding as in [12], we get the following result:

PROPOSITION 2.2 *There exists a unique solution $(m,\eta,\mu)\in L^2(Q_T)\times P_*\times L^2(Q_T)$ to (19). Moreover, $y=\rho^{-1}m$ is, together with $v=\rho^{-1}\eta\,1_\omega$, the unique solution to (10).*

PROOF: Let us introduce the bilinear forms $a(\cdot\,,\cdot)$ and $b(\cdot\,,\cdot)$, given by

$$a((m,\eta),(\overline{m},\overline{\eta})) = \iint_{Q_T} m\overline{m}\,dx\,dt + \iint_{q_T}\eta\,\overline{\eta}\,dx\,dt \quad \forall(m,\eta),(\overline{m},\overline{\eta})\in L^2(Q_T)\times P_*$$

and

$$b((m,\eta),\mu) = \iint_{Q_T}\mu\Big(m - \rho^{-1}L_A^\star(\rho_0\eta)\Big)\,dx\,dt \quad \forall(m,\eta,\mu)\in L^2(Q_T)\times P_*\times L^2(Q_T)$$

and the linear form ℓ, with

$$\langle \ell,(m,\eta)\rangle = \int_0^1 \rho_0(x,0)y_0(x)\eta(x,0)dx + \iint_{Q_T}\rho_0 B\eta\,dx\,dt \quad \forall(m,\eta)\in L^2(Q_T)\times P_*.$$

Then, $a(\cdot,\cdot)$, $b(\cdot\,,\cdot)$ and ℓ are well-defined and continuous and (19) reads as follows: find $(m,\eta,\mu)\in L^2(Q_T)\times P_*\times L^2(Q_T)$ such that

$$\left\{\begin{aligned} a(\,(m,p),(\overline{m},\overline{\eta})\,) + b(\,(\overline{m},\overline{\eta}),\mu\,) &= \langle\ell,(\overline{m},\overline{\eta})\rangle & &\forall(\overline{m},\overline{\eta})\in L^2(Q_T)\times P_* \\ b(\,(m,\eta),\overline{\mu}\,) &= 0 & &\forall\overline{\mu}\in L^2(Q_T)\end{aligned}\right.$$

Finally, we introduce the space

$$V = \{\,(m,\eta)\in L^2(Q_T)\times P_* : b(\,(m,\eta),\mu\,) = 0 \ \ \forall\mu\in L^2(Q_T)\,\}.$$

We first notice that, for any $(m,\eta)\in V$, $m=\rho^{-1}L_A^\star\rho_0\eta$ and thus

$$\begin{aligned} a(\,(m,\eta),(m,\eta)\,) &= \frac{1}{2}\iint_{Q_T}|m|^2\,dx\,dt + \frac{1}{2}\iint_{Q_T}\rho^{-2}|L_A^\star(\rho_0\eta)|^2\,dx\,dt + \iint_{q_T}|\eta|^2\,dx\,dt \\ &\geq \frac{1}{2}\,\|(m,\eta)\|^2_{L^2(Q_T)\times P_*}.\end{aligned}$$

This shows that $a(\cdot\,,\cdot)$ is coercive on V.

On the other hand, for any $\mu\in L^2(Q_T)$, we have:

$$\sup_{(m,P)\in L^2(Q_T)\times P_*}\frac{b(\,(m,P),\mu\,)}{\|(m,P)\|_{L^2\times P}\|\mu\|_{L^2(Q_T)}} \geq \frac{b(\,(\mu,0),\mu\,)}{\|(\mu,0)\|_{L^2(Q_T)\times P_*}\|\mu\|_{L^2(Q_T)}} \geq 1.$$

Therefore, $b(\cdot\,,\cdot)$ satisfies the usual "inf-sup" condition with respect to $L^2(Q_T)\times P_*$ and $L^2(Q_T)$. This suffices to ensure the existence and uniqueness of a solution to (19). □

In the formulation (19), the weights ρ_0 and ρ appear in the last integral of the first equation. However, we have assumed that $\rho_0 B$ belongs to $L^2(Q_T)$, see (13). In practice, in the linear problems we are going to find below, the function B is not zero only in the Newton-Raphson approach, where it is of order $O(\rho^{-2})$, see (62).

More importantly, both weights ρ and ρ_0 appear in the expression $\rho^{-1}L_A^\star\rho_0\eta$, where they somehow compensate to each other. In fact, if we assume that $\mu_x\in L^2(Q_T)$ and $\mu_{|\Sigma_T}=0$ and we expand the associated term, after some integrations by parts we find that

$$\begin{aligned}&\iint_{Q_T}\Big(\rho^{-1}L_A^\star(\rho_0\overline{\eta})-\overline{m}\Big)\mu\,dx\,dt\\&\quad=\iint_{Q_T}\Big(B_1\mu\overline{\eta}+B_2\mu\overline{\eta}_x+B_3\mu_x\overline{\eta}+B_4\mu_x\overline{\eta}_x+B_5\mu\overline{\eta}_t-\mu\overline{m}\Big)\,dx\,dt\end{aligned}\tag{20}$$

where

$$\begin{aligned}&B_1=-\rho^{-1}(\rho_0)_t+a\rho_x^{-1}(\rho_0)_x+A\rho^{-1}\rho_0,\quad B_2=a\rho_x^{-1}\rho_0\\&B_3=a\rho^{-1}(\rho_0)_x,\quad B_4=a\rho^{-1}\rho_0,\quad B_5=-(T-t)^{3/2}\end{aligned}\tag{21}$$

In this formulation, there is no weight blowing up exponentially in time near T. The unique singularity appears in B_1 and is polynomial of order $(T-t)^{-1/2}$ and can be removed easily, by replacing the Lagrangian multiplier μ by $(T-t)^{1/2}\mu$.

The mixed formulation (19) is used below, in Section 4, to solve numerically (10).

2.2 Numerical solution via a dual approach

As noticed in [13], the extremal problem (10) can also be solved by dual methods. In a first step, we fix $\varepsilon>0$ and $R>0$ and we "relax" (10) by introducing the problem

$$\begin{cases}\text{Minimize}\ \ J_{R,\varepsilon}(y,v)=\dfrac12\displaystyle\iint_{Q_T}\rho_R^2|y|^2\,dx\,dt+\dfrac12\iint_{q_T}\rho_0^2|v|^2\,dx\,dt+\dfrac{1}{2\varepsilon}\|y(\cdot\,,T)\|_{L^2}^2\\ \text{Subject to}\ \ (y,v)\in\mathcal{A}_{lin}(y_0,T)\end{cases}\tag{22}$$

where $\rho_R=\min(\rho,R)$ and $\mathcal{A}_{lin}(y_0,T)=\{\,(y,v): v\in L^2(q_T),\ y \text{ solves } (9)\,\}$. Then, we consider the associated (well-posed) dual problem:

$$\begin{cases}\text{Minimize}\ \ J_{R,\varepsilon}^\star(\mu,\varphi_T)=\dfrac12\Big(\displaystyle\iint_{Q_T}\rho_R^{-2}|\mu|^2\,dx\,dt+\iint_{q_T}\rho_0^{-2}|\psi|^2\,dx\,dt\Big)\\ \qquad\qquad +\displaystyle\iint_{Q_T}B(x,t)\psi\,dx\,dt+\int_0^1 y_0(x)\,\varphi(x,0)\,dx+\frac{\varepsilon}{2}\|\varphi_T\|_{L^2}^2\\ \text{Subject to}\ \ (\mu,\varphi_T)\in L^2(Q_T)\times L^2(0,1)\end{cases}\tag{23}$$

where ψ solves

$$\begin{cases}-\psi_t-(a(x)\psi_x)_x+A(x,t)\psi=\mu, & (x,t)\in Q_T\\ \psi(x,t)=0, & (x,t)\in\Sigma_T\\ \psi(x,T)=\varphi_T(x), & x\in(0,1)\end{cases}\tag{24}$$

The following result from [13] justifies the introduction of $J_{R,\varepsilon}^\star$:

PROPOSITION 2.3 *The unconstrained extremal problem* (23) *is the dual problem to* (22) *in the sense of the Fenchel-Rockafellar theory. Furthermore,* (23) *is stable and possesses a unique solution. Finally, if we denote by* $(y_{R,\varepsilon}, v_{R,\varepsilon})$ *the unique solution to* (22) *and we denote by* $(\mu_{R,\varepsilon}, \varphi_{T,R,\varepsilon})$ *the unique solution to* (23)*, then the following relations hold:*

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

Here, $\psi_{R,\varepsilon}$ *solves (24) with* $\mu = \mu_{R,\varepsilon}$ *and* $\varphi_T = \varphi_{T,R,\varepsilon}$*. Moreover,* $(v_{R,\varepsilon}, y_{R,\varepsilon})$ *converges strongly in* $L^2(q_T) \times L^2(Q_T)$ *to* (y,v) *as* $\varepsilon \to 0^+$ *and* $R \to +\infty$*, where* (y,v) *is the unique minimizer of* J *in* $\mathcal{C}_{lin}(y_0,T)$*.*

3 The solution of the nonlinear problem

We will now present several numerical methods for the computation of the solution(s) to the null controllability problem for the nonlinear system (1).

For simplicity, we will assume in this Section that $y_0 \in L^\infty(0,1)$ and $f \in C^1(\mathbf{R})$ and is globally Lipschitz-continuous (and even more regular in Sections 3.2 and 3.3; see below). Let us introduce the function g, with

$$g(s) = \frac{f(s)}{s} \;\text{ if }\; s \neq 0, \quad g(0) = f'(0) \;\text{ otherwise.} \tag{25}$$

Then $g \in C^0_b(\mathbf{R})$ and $f(s) = g(s)\,s$ for all s (recall that $f(0) = 0$). We will set $G_0 = \|g\|_{L^\infty(\mathbf{R})}$.

For any $z \in L^1(Q_T)$, let us introduce the bilinear form

$$m(z;p,q) = \iint_{Q_T} \rho^{-2} L^*_{g(z)}p\, L^*_{g(z)}q\,dx\,dt + \iint_{q_T} \rho_0^{-2} p\,q\,dx\,dt \quad \forall p,q \in P_0. \tag{26}$$

Then $m(z;\cdot\,,\cdot)$ is a scalar product in P_0. As before, it can be used to construct a Hilbert space P that, in principle, may depend on z.

We will use the following result, which is a direct consequence of the Carleman estimates in [15]:

LEMMA 3.1 *Under the previous conditions, if the constants* K_i *in* (11) *are large enough (depending on* ω*,* T*,* a_0*,* $\|a\|_{C^1}$ *and* G_0*), there exist positive constants* C_1 *and* C_2 *such that*

$$C_1\, m(0;p,p) \le m(z;p,p) \le C_2\, m(0;p,p) \quad \forall p \in P_0 \tag{27}$$

for all $z \in L^1(Q_T)$*.*

In the sequel, it will be assumed that the constants K_i have been chosen as in Lemma 3.1. Accordingly, all the spaces P provided by the bilinear forms $m(z;\cdot\,,\cdot)$ are the same and, in fact, (27) holds for all $p \in P$:

$$C_1\, m(0;p,p) \le m(z;p,p) \le C_2\, m(0;p,p) \quad \forall p \in P. \tag{28}$$

We will fix the following norm in P:

$$\|p\|_P = m(0;p,p)^{1/2} \quad \forall p \in P. \tag{29}$$

Let us introduce the mapping $\Lambda_0 : L^2(Q_T) \mapsto L^2(Q_T)$ where, for any $z \in L^2(Q_T)$, $y_z = \Lambda_0(z)$ is, together with v_z, the unique solution to the linear extremal problem

$$\begin{cases} \text{Minimize } \; J(z;y,v) = \dfrac{1}{2}\displaystyle\iint_{Q_T} \rho^2|y|^2\,dx\,dt + \dfrac{1}{2}\iint_{q_T} \rho_0^2|v|^2\,dx\,dt \\ \text{Subject to } \; (y,v) \in \mathcal{C}_{lin}(z;y_0,T) \end{cases} \tag{30}$$

where

$$\mathcal{C}_{lin}(z; y_0, T) = \{\, (y, v) : v \in L^2(q_T),\ y \text{ solves (31) and satisfies (6)}\,\}.$$

Here, (31) is the following particular system of the kind (9):

$$\begin{cases} y_t - (a(x)y_x)_x + g(z)\, y = v\, 1_\omega & (x,t) \in Q_T \\ y(x,t) = 0, & (x,t) \in \Sigma_T \\ y(x,0) = y_0(x), & x \in (0,1) \end{cases} \tag{31}$$

In view of the results recalled in Section 2, $\Lambda_0 : L^2(Q_T) \mapsto L^2(Q_T)$ is well defined. Furthermore, applying Proposition 2.1 with $A = g(z)$ and $B = 0$, we obtain that y_z and v_z are characterized as follows :

$$y_z = \Lambda_0(z) = \rho^{-2} L^*_{g(z)} p_z, \quad v_z = -\rho_0^{-2} p_z|_{q_T}, \tag{32}$$

where p_z is the unique solution to the linear problem

$$m(z; p_z, q) = \int_0^1 y_0(x)\, q(x,0)\, dx \quad \forall q \in P; \quad p_z \in P. \tag{33}$$

3.1 A fixed point method

In order to solve the null controllability problem for (1), it suffices to find a solution to the fixed point equation

$$y = \Lambda_0(y), \quad y \in L^2(Q_T). \tag{34}$$

A natural algorithm is thus the following.

ALG 1 (fixed point):

(i) Choose $y^0 \in L^2(Q_T)$, $\varepsilon > 0$ and $N_\infty \in \mathbb{N}$.

(ii) Then, for any given $n \geq 0$ and $y^n \in L^2(Q_T)$, compute $y^{n+1} = \Lambda_0(y^n)$ (until $\|\Lambda_0(y^n) - y^n\|_{L^2(Q_T)} \leq \varepsilon$ or $n > N_\infty$) , i.e. find the unique solution (y^{n+1}, v^{n+1}) to the linear extremal problem

$$\text{Minimize } J(y^n; y^{n+1}, v^{n+1}) = \frac{1}{2} \iint_{Q_T} \rho^2 |y^{n+1}|^2 \, dx\, dt + \frac{1}{2} \iint_{q_T} \rho_0^2 |v^{n+1}|^2 \, dx\, dt \tag{35}$$

subject to $v^{n+1} \in L^2(q_T)$ and

$$\begin{cases} y_t^{n+1} - (a(x)y_x^{n+1})_x + g(y^n)\, y^{n+1} = v^{n+1}\, 1_\omega & (x,t) \in Q_T \\ y^{n+1}(x,t) = 0, & (x,t) \in \Sigma_T \\ y^{n+1}(x,0) = y_0(x), & x \in (0,1). \end{cases} \tag{36}$$

This is a standard fixed point method for the mapping Λ_0. At each step, we have to solve a null controllability problem for a linear parabolic system. Thus, it can be combined with the techniques in Sections 2.1 or 2.2 to produce sequences $\{y^n\}$ and $\{v^n\}$ such that, if v^n converges weakly in $L^2(q_T)$ to v, then v and its associated state solve the null controllability problem for (1).

This fixed point formulation has been used in [14] to prove Theorem 1.2. Precisely, it is shown that there exists $M > 0$ such that Λ_0 maps the closed ball $B(0; M) \subset L^2(Q_T)$ into itself. Then, Kakutani's Theorem provides the existence of at least one fixed point for Λ_0.

It is however important to note that this does not imply the convergence of the sequence $\{y^n\}$ defined by $y^{n+1} = \Lambda_0(y^n)$. This motivates the arguments and algorithms in the following Section.

It may also be noted that a couple (y, v) such that y is a fixed point of Λ_0 is not necessarily a solution to (7). It has not to be even a local minimizer of J.

3.2 Least squares and (conjugate) gradient methods

We now introduce the function $\zeta(t) \equiv (T-t)^{-1/2}$ and the space $Z := L^2(\zeta^2, Q_T)$, see (18). We will denote by Λ the restriction to Z of the mapping Λ_0. Obviously, $\Lambda(z) \in Z$ for all $z \in Z$.

Let us consider the following least squares reformulation of (34):

$$\begin{cases} \text{Minimize} \quad R(z) := \dfrac{1}{2}\|z - \Lambda(z)\|_Z^2 \\ \text{Subject to} \quad z \in Z \end{cases} \tag{37}$$

Any solution to (34) solves (37). Conversely, if y solves (37), we necessarily have $R(y) = 0$ (because (1) is null controllable with control-states (y, v) such that $J(z; y, v) < +\infty$); hence, y also solves (34). This shows that (34) and (37) are, in the present context, equivalent.

We will solve (37) by using gradient techniques. To this purpose, it is crucial to determine conditions under which R is differentiable and, also, to compute $R'(z)$.

For simplicity and clarity, in the following results it will be assumed that the function g given by (25) is C^1 and g' is uniformly bounded. This is the case if, for instance, f is C^1 and globally Lipschitz-continuous and C^2 in a neighborhood of 0.

PROPOSITION 3.1 *Let us assume that $g \in C^1_b(\mathbf{R})$. Then $R \in C^1(Z)$. Moreover, for any $z \in Z$, the derivative $R'(z)$ is given by*

$$R'(z) = \big(1 - \rho^{-2} g'(z) p_z\big)(z - y_z) + \zeta^{-2} g'(z)\big(y_z \lambda_z + p_z \mu_z\big), \tag{38}$$

where p_z is the unique solution to (33), $y_z = \rho^{-2} L^*_{g(z)} p_z$, λ_z is the unique solution to the linear (adjoint) problem

$$m(z; q, \lambda_z) = (z - y_z, \rho^{-2} L^*_{g(z)} q)_Z \quad \forall q \in P; \quad \lambda_z \in P \tag{39}$$

and, finally, $\mu_z = \rho^{-2} L^*_{g(z)} \lambda_z$.

For the proof, we will need the following Lemma:

LEMMA 3.2 *For any $q \in P$ one has $(\zeta\rho)^{-1} q \in L^\infty(Q_T)$. Furthermore, there exists $C > 0$, only depending on ω, T, a_0, $\|a\|_{C^1}$ and G_0, such that*

$$\|(\zeta\rho)^{-1} q\|^2_{L^\infty(Q_T)} \le C\, \|q\|_P^2 \quad \forall q \in P. \tag{40}$$

PROOF: We will use the following Carleman inequality from [15] (see also [11]):

There exists $C > 0$, only depending on ω, T, a_0, $\|a\|_{C^1}$ and G_0, such that

$$\begin{cases} \displaystyle\iint_{Q_T} \left[\rho_2^{-2}\left(|q_t|^2 + |q_{xx}|^2\right) + \rho_1^{-2}|q_x|^2 + \rho_0^{-2}|q|^2\right] dx\,dt \\ \qquad\qquad \displaystyle\le C\left(\iint_{Q_T} \rho^{-2}|L^*_{g(z)} q|^2\, dx\,dt + \iint_{q_T} \rho_0^{-2}|q|^2\, dx\,dt\right) \end{cases} \tag{41}$$

for all $q \in P$ and any $z \in L^1(Q_T)$.

Recall (11) and (12) for the definitions of the weights ρ and ρ_i. In view of (41), for any $q \in P$, one has

$$(\zeta\rho)^{-1}q \in L^2(0,T;H^2(0,1)) \text{ and } ((\zeta\rho)^{-1}q)_t \in L^2(0,T;L^2(0,1)).$$

Consequently, $(\zeta\rho)^{-1}q \in C^0([0,T];H^1(0,1))$ and, in particular, $(\zeta\rho)^{-1}q \in L^\infty(Q_T)$ and (40) holds. □

PROOF OF PROPOSITION 3.1: Let us first see that the mapping $z \mapsto p_z$ is C^1.

Thus, let us fix $z, m \in Z$ and $a \in \mathbf{R}_+$ and let us set $p = p_z$, $\tilde{z} = z + am$ and $\tilde{p} = p_{\tilde{z}}$. After some computations, we easily see that

$$\tilde{p} = p + ak + ak'_a\,,$$

where $k \in P$ and $k'_a \in P$ must respectively satisfy the following for all $q \in P$:

$$m(z;k,q) = -\iint_{Q_T} \rho^{-2} g'(z)\left(L^*_{g(z)}p\cdot q + p\cdot L^*_{g(z)}q\right) m\,dx\,dt$$

and

$$\begin{aligned}
m(z;k'_a,q) &= \iint_{Q_T} \rho^{-2} g'(z)\left(L^*_{g(z)}p\cdot q + p\cdot L^*_{g(z)}q\right) m\,dx\,dt \\
&\quad -\frac{1}{a}\left(m(\tilde{z};\tilde{p},q) - m(z;\tilde{p},q)\right) \\
&= \iint_{Q_T} \rho^{-2}\left(g'(z)m - \frac{g(\tilde{z})-g(z)}{a}\right)\left(L^*_{g(z)}p\cdot q + p\cdot L^*_{g(z)}q\right) dx\,dt \\
&\quad + a\iint_{Q_T} \rho^{-2}\left(\frac{g(\tilde{z})-g(z)}{a}\right)\left[L^*_{g(\tilde{z})}k\cdot q + k\cdot L^*_{g(\tilde{z})}q + \left(\frac{g(\tilde{z})-g(z)}{a}\right)pq\right] dx\,dt \\
&\quad + a\iint_{Q_T} \rho^{-2}\left(\frac{g(\tilde{z})-g(z)}{a}\right)\left(L^*_{g(\tilde{z})}k'_a\cdot q + k'_a\cdot L^*_{g(\tilde{z})}q\right) dx\,dt.
\end{aligned}$$

Taking $q = k'_a$ in these last identity, in account of (28), we find that

$$\begin{aligned}
\|k'_a\|_P^2 &\le C\iint_{Q_T} \rho^{-2}\left|g'(z)m - \frac{g(\tilde{z})-g(z)}{a}\right|\left|L^*_{g(z)}p\cdot k'_a + p\cdot L^*_{g(z)}k'_a\right| dx\,dt \\
&\quad + Ca\|k'_a\|_P + Ca\|k'_a\|_P^2.
\end{aligned}$$

On the other hand, in view of Lemma 3.2,

$$\begin{aligned}
&\iint_{Q_T} \rho^{-2}\left|g'(z)m - \frac{g(\tilde{z})-g(z)}{a}\right|\left|L^*_{g(z)}p\cdot k'_a + p\cdot L^*_{g(z)}k'_a\right| dx\,dt \\
&\quad \le \|\zeta[g'(z)m - \frac{1}{a}(g(\tilde{z})-g(z))]\|_{L^2(Q_T)}\left(\|\rho^{-1}L^*_{g(z)}p\|_{L^2(Q_T)}\|(\zeta\rho)^{-1}k'_a\|_\infty\right. \\
&\qquad \left.+ \|(\zeta\rho)^{-1}p\|_\infty\|\rho^{-1}L^*_{g(z)}k'_a\|_{L^2(Q_T)}\right) \\
&\quad \le C\|\zeta[g'(z)m - \frac{1}{a}(g(\tilde{z})-g(z))]\|_{L^2(Q_T)}\|p\|_P\|k'_a\|_P.
\end{aligned}$$

Consequently, we have the following for all small $a \in \mathbf{R}_+$:

$$\|k'_a\|_P^2 \le C\left(a^2 + \|\zeta[g'(z)m - \frac{1}{a}(g(\tilde{z})-g(z))]\|^2_{L^2(Q_T)}\right).$$

From Lebesgue's Theorem, we deduce that $k'_a \to 0$ in P as $a \to 0^+$. This shows that the mapping $z \mapsto p_z$ is differentiable in the direction m; it also shows that the derivative in that direction is given by $k_z = k$, that is, the unique solution to the problem

$$m(z;k_z,q) = -\iint_{Q_T} \rho^{-2} g'(z) \left(L^*_{g(z)} p\, q + p\, L^*_{g(z)} q \right) m \, dx\, dt \quad \forall q \in P; \quad k_z \in P. \tag{42}$$

An immediate consequence is that $\Lambda \in C^1(Z;Z)$, $R \in C^1(Z)$,

$$\Lambda'(z)\cdot m = \rho^{-2} g'(z) p_z m + \rho^{-2} L^*_{g(z)} k_z$$

and

$$\begin{aligned}(R'(z), m)_Z &= (z - \Lambda(z), m - \Lambda'(z)\cdot m)_Z \\ &= (z - y_z, (1 - \rho^{-2} g'(z) p_z)\, m)_Z - (z - y_z, \rho^{-2} L^*_{g(z)} k_z)_Z.\end{aligned}$$

Let λ_z be the unique solution to the well-posed problem (39). Then, taking into account the definitions of y_z and μ_z, one has:

$$\begin{aligned}(z - y_z, \rho^{-2} L^*_{g(z)} k_z)_Z &= m(z; k_z, \lambda_z) \\ &= -\iint_{Q_T} \rho^{-2} g'(z) \left(L^*_{g(z)} p_z\, \lambda_z + p_z\, L^*_{g(z)} \lambda_z \right) m\, dx\, dt \\ &= -(\zeta^{-2} g'(z) \big(y_z \lambda_z + p_z \mu_z\big), m)_Z.\end{aligned}$$

This proves (38). □

Proposition 3.2 *Let the assumptions in Proposition 3.1 be satisfied and let us introduce* $G_1 := \|g'\|_{L^\infty(\mathbf{R})}$. *There exists a constant* K *that depends on* ω, T, a_0, $\|a\|_{C^1}$ *and* G_0 *but is independent of* z *and* y_0, *such that the following holds for all* $z \in Z$:

$$\|R'(z)\|_Z \geq (1 - K\, G_1 \, \|y_0\|_{L^\infty})\, \|z - y_z\|_Z\,. \tag{43}$$

For the proof, the following technical result will be needed:

Lemma 3.3 *With the notation of Proposition 3.1, one has:*

$$\|p_z\|_P \leq C \|y_0\|_{L^\infty} \quad \forall z \in Z \tag{44}$$

and

$$\|\lambda_z\|_P \leq C \|\zeta \rho^{-1}\|_\infty \|z - y_z\|_Z \quad \forall z \in Z, \tag{45}$$

where C *depends on* ω, T, a_0, $\|a\|_{C^1}$ *and* G_0.

Proof: The estimate (44) is a consequence of (33), (28) and Lemma 2.1 with $A = g(0)$:

$$\|p_z\|_P^2 \leq \frac{1}{C_1} m(z; p_z, p_z) \leq \frac{1}{C_1} \|y_0\|_{L^\infty} \|p_z(\cdot\,, 0)\|_{L^2} \leq C \|y_0\|_{L^\infty} \|p_z\|_P.$$

On the other hand, (45) is a consequence of (39), the definition of Z and, again, Lemma 2.1:

$$\begin{aligned}\|\lambda_z\|_P^2 &\leq \frac{1}{C_1} m(z; \lambda_z, \lambda_z) \leq \frac{1}{C_1} \|\zeta \rho^{-1}\|_\infty \|\zeta (z - y_z)\|_2 \|\rho^{-1} L^*_{g(z)} \lambda_z\|_2 \\ &\leq \frac{C_2^{1/2}}{C_1} \|\zeta \rho^{-1}\|_\infty \|z - y_z\|_Z \|\lambda_z\|_P.\end{aligned}$$

□

PROOF OF PROPOSITION 3.2: Let $z \in Z$ be given and let us introduce $f := z - y_z$. In view of Proposition 3.1, one has

$$
\begin{aligned}
\|R'(z)\|_Z &\geq \frac{1}{\|f\|_Z}(R'(z), f) \\
&= \frac{1}{\|f\|_Z} \iint_{Q_T} \left(\zeta^2(1+\rho^{-2}g'(z)\,p_z)\,|f|^2 + g'(z)(y_z\lambda_z + p_z\mu_z)\,f\right)\,dx\,dt \\
&\geq \|f\|_Z - \frac{1}{\|f\|_Z}\iint_{Q_T} \zeta^2\rho^{-2}|g'(z)|\,|p_z|\,|f|^2\,dx\,dt \\
&\quad - \frac{1}{\|f\|_Z}\iint_{Q_T} |g'(z)|\,(|y_z||\lambda_z| + |p_z||\mu_z|)\,|f|\,dx\,dt.
\end{aligned}
$$

In view of Lemmas 3.2 and 3.3,

$$
\begin{aligned}
\iint_{Q_T} \zeta^2\rho^{-2}|g'(z)|\,|p_z|\,|f|^2\,dx\,dt &\leq \|\rho^{-2}g'(z)\,p_z\|_\infty \left(\iint_{Q_T}\zeta^2|f|^2\,dx\,dt\right) \\
&\leq \|\zeta\rho^{-1}\|_\infty\|g'(z)\|_\infty\,\|(\zeta\rho)^{-1}p_z\|_\infty\|f\|_Z^2 \\
&\leq C\,G_1\,\|p_z\|_P\|f\|_Z^2 \\
&\leq C\,G_1\,\|y_0\|_{L^\infty}\|f\|_Z^2.
\end{aligned}
$$

On the other hand, from Lemma 3.2 we also have

$$
\begin{aligned}
&\iint_{Q_T} |g'(z)|\,(|y_z||\lambda_z| + |p_z||\mu_z|)\,|f|\,dx\,dt \\
&= \iint_{Q_T} |g'(z)|\left(|\rho^{-2}L^*_{g(z)}p_z||\lambda_z| + |p_z||\rho^{-2}L^*_{g(z)}\lambda_z|\right)|f|\,dx\,dt \\
&\leq \|g'(z)\|_\infty\left[\left(\iint_{Q_T}\rho^{-2}|L^*_{g(z)}p_z|^2\,dx\,dt\right)^{1/2}\|(\zeta\rho)^{-1}\lambda_z\|_\infty\right. \\
&\quad \left. + \left(\iint_{Q_T}\rho^{-2}|L^*_{g(z)}\lambda_z|^2\,dx\,dt\right)^{1/2}\|(\zeta\rho)^{-1}p_z\|_\infty\right]\left(\iint_{Q_T}\zeta^2|f|^2\,dx\,dt\right)^{1/2} \\
&\leq C\,G_1\left[\|y_0\|_{L^\infty}\|\lambda_z\|_P + \|p_z\|_P\|\zeta\rho^{-1}\|_\infty\|f\|_Z\right]\|f\|_Z \\
&\leq C\,G_1\,\|y_0\|_{L^\infty}\|f\|_Z^2.
\end{aligned}
$$

Consequently,

$$
\|R'(z)\|_Z \geq \|f\|_Z - K\,G_1\,\|y_0\|_{L^\infty}\|f\|_Z
$$

and we get (43) for some K. □

A very relevant consequence of Proposition 3.2 is the following:

COROLLARY 3.1 *Let us the assumptions in Proposition 3.1 be satisfied and let K be the constant furnished by Proposition 3.2. If*

$$
K\,G_1\,\|y_0\|_{L^\infty} < 1,
$$

then the critical points of R are global minima and, consequently, solve (34).

PROOF: It suffices to notice that, under the previous assumptions, if $R'(z) = 0$, we necessarily have $z - y_z = 0$, i.e. $\Lambda(z) = z$. □

Therefore, it is very appropriate to try to solve the null controllability problem for (1) by applying a gradient method to the extremal problem (37). An appropriate choice is the steepest descent algorithm, which is now described.

ALG 2 (Steepest descent gradient iterates for the least squares formulation):

(i) Set $\varepsilon > 0$ and $\mu_x, \mu_t > 0$ and choose $z^0 \in Z$ such that $z^0(\cdot\,, 0) = y_0$.

(ii) Then, for any given $n \geq 0$ and $z^n \in Z$, do the following until convergence, i.e. until $R(z^n) \leq \varepsilon$:

(a) Compute the solution p^n to the problem

$$m(z^n; p^n, q) = \int_0^1 y_0(x)\, q(x,0)\, dx \quad \forall q \in P; \quad p^n \in P,$$

set $y^n = \rho^{-2} L^*_{g(z^n)} p^n$, i.e. $y^n = \Lambda(z^n)$, compute the solution λ^n to the problem

$$m(z^n; q, \lambda^n) = (z^n - y^n, \rho^{-2} L^*_{g(z^n)} q)_Z \quad \forall q \in P; \quad \lambda^n \in P$$

and set $\mu^n = \rho^{-2} L^*_{g(z^n)} \lambda^n$ and

$$\tilde{g}^n = \big(1 - \rho^{-2} g'(z^n) p^n\big)(z^n - y^n) + \zeta^{-2} g'(z^n)\big(y^n \lambda^n + p^n \mu^n\big). \tag{46}$$

(b) Regularize the descent direction $\tilde{g}^n$ by solving the elliptic problem

$$\begin{cases} (1 - \mu_x \partial_x^2 - \mu_t \partial_t^2) g^n = \tilde{g}^n, & (x,t) \in Q_T \\ g^n = 0, & (x,t) \in \partial Q_T \end{cases}$$

(c) Compute the solution σ^n to the 1D extremal problem

$$\begin{cases} \text{Minimize} \;\; R^n(\sigma) := R(z^n - \sigma g^n) \\ \text{Subject to} \;\; \sigma > 0 \end{cases} \tag{47}$$

and set $z^{n+1} = z^n - \sigma^n g^n$.

Remark 2 We see from (46) that, strictly speaking, this algorithm only makes sense when $g \in C^1(\mathbf{R})$. It would be interesting to extend this method to the more general situation for which f is (only) globally Lipschitz-continuous. □

Remark 3

- Due to the term $\zeta^{-2} g'(z^n) p^n \mu^n$, the derivative $\tilde{g}^n$ vanishes everywhere on ∂Q_T except for $t = 0$. Therefore, Step (b) above ensures that the new iterate z^{n+1} defined by Step (c) satisfies $z^{n+1}(\cdot\,, 0) = y_0$ (in addition to play the role of a smoother of the descent direction).

- The scalar minimization problem (47) can be solved by computing a root of the first derivative of $R^n(\sigma)$. We have

$$T(\sigma) := \frac{dR^n(\sigma)}{d\sigma} = -\iint_{Q_T} R'(z^n - \sigma g^n) g^n \, dx \, dt.$$

For some small $\eta > 0$, the following Newton-like iterates are then computed:

$$\sigma_0^n = 0; \quad \sigma_{k+1}^n = \sigma_k^n - \frac{\eta T(\sigma^n)}{T(\sigma_k^n + \eta) - T(\sigma_k^n)}, \quad k \geq 1.$$

□

Notice that, under the assumptions of Corollary 3.1, **ALG 2** provides a sequence $\{z^n\}$, with $z^{n+1} = z^n - \sigma^n g^n$, that converges to a fixed point of the mapping Λ_0, in contrast with **ALG 1**.

Let us also describe a conjugate gradient algorithm.

In the sequel, the notation $\Gamma(g, h)$ is used to denote

$$\text{either } \frac{\|g\|_Z^2}{\|h\|_Z^2} \text{ or } \frac{(g, g-h)_Z}{\|h\|_Z^2}. \tag{48}$$

The first choice (resp. the second choice) leads to the Polak-Ribière (resp. Fletcher-Reeves) version of the conjugate gradient algorithm, which is as follows.

ALG 2' (Conjugate gradient iterates for the least squares formulation):

(i) Set $\varepsilon > 0$ and $\mu_x, \mu_t > 0$ and choose $z^0 \in Z$.

(ii) Do the following:

(a) Compute the solution p^0 to the problem

$$m(z^0; p^0, q) = \int_0^1 y_0(x)\, q(x, 0)\, dx \quad \forall q \in P; \quad p^0 \in P,$$

set $y^0 = -\rho^{-2} L^*_{g(z^0)} p^0$, i.e. $y^0 = \Lambda(z^0)$, compute the solution λ^0 to the problem

$$m(z^0; q, \lambda^0) = (z^0 - y^0, \rho^{-2} L^*_{g(z^0)} q)_Z \quad \forall q \in P; \quad \lambda^0 \in P$$

and set $\mu^0 = -\rho^{-2} L^*_{g(z^0)} \lambda^0$ and

$$\tilde{g}^0 = \big(1 + \rho^{-2} g'(z^0) p^0\big)(z^0 - y^0) + \zeta^{-2} g'(z^0)\big(y^0 \lambda^0 + p^0 \mu^0\big).$$

(b) Regularize the descent direction $\tilde{g}^0$ by solving the elliptic problem

$$\begin{cases} (1 - \mu_x \partial_x^2 - \mu_t \partial_t^2) g^0 = \tilde{g}^0, & (x,t) \in Q_T, \\ g^0 = 0, & (x,t) \in \partial Q_T. \end{cases}$$

(c) Set $d^0 = g^0$ and compute the solution σ^0 to the 1D extremal problem

$$\begin{cases} \text{Minimize } \; R^0(\sigma) := R(z^0 - \sigma d^0) \\ \text{Subject to } \; \sigma > 0 \end{cases}$$

and set $z^1 = z^0 - \sigma^0 d^0$.

(iii) Then, for any given $n \geq 1$, $z^n \in Z$, $g^{n-1} \in Z$ and $d^{n-1} \in Z$, do the following until convergence:

(a) Compute the solution p^n to the problem

$$m(z^n; p^n, q) = \int_0^1 y_0(x)\, q(x,0)\, dx \quad \forall q \in P; \quad p^n \in P,$$

set $y^n = -\rho^{-2} L^*_{g(z^n)} p^n$, i.e. $y^n = \Lambda(z^n)$, compute the solution λ^n to the problem

$$m(z^n; q, \lambda^n) = (z^n - y^n, \rho^{-2} L^*_{g(z^n)} q)_Z \quad \forall q \in P; \quad \lambda^n \in P$$

and set $\mu^n = -\rho^{-2} L^*_{g(z^n)} \lambda^n$ and

$$\tilde{g}^n = \big(1 + \rho^{-2} g'(z^n) p^n\big)(z^n - y^n) + \zeta^{-2} g'(z^n)\big(y^n \lambda^n + p^n \mu^n\big).$$

(b) Regularize the descent direction $\tilde{g}^n$ by solving the elliptic problem

$$\begin{cases} (1 - \varepsilon_x \partial_x^2 - \varepsilon_t \partial_t^2) g^n = \tilde{g}^n, & (x,t) \in Q_T, \\ g^n = 0, & (x,t) \in \partial Q_T, \end{cases}$$

compute

$$\gamma^n = \Gamma(g^n, g^{n-1})$$

and set

$$d^n = g^n + \gamma^n d^{n-1}.$$

(c) Compute the solution σ^n to the one-dimensional extremal problem

$$\begin{cases} \text{Minimize } \ R^n(\sigma) := R(z^n - \sigma d^0) \\ \text{Subject to } \ \sigma > 0 \end{cases}$$

and set $z^{n+1} = z^n - \sigma^n d^n$.

3.3 A Newton algorithm

The fixed point method of Section 3.1 relies on replacing at each step the nonlinear term $f(y^{n+1}) = y^{n+1} g(y^{n+1})$ by the linear term $y^{n+1} g(y^n)$. Let us now present a different way to linearize.

Let us introduce the spaces

$$Y := \{\, (y,v) : y \in L^2(\rho^2; Q_T),\ y_x \in L^2(\rho_1^2; Q_T),\ y_t - (a y_x)_x \in L^2(\rho_0^2; Q_T), \\ y(0,t) = y(1,t) = 0 \ \text{ a.e.},\ v \in L^2(\rho^2, Q_T) \,\}$$

and

$$W := L^2(\rho_0^2; Q_T) \times L^2(0,1)$$

and the mapping $F : Y \mapsto W$, with

$$F(y,v) = (y_t - (a y_x)_x + f(y) - v 1_\omega, y(\cdot\,,0) - y_0) \quad \forall (y,v) \in Y \tag{49}$$

(recall (11) and (12) for the definitions of ρ, ρ_0 and ρ_1).

Obviously, any solution to the nonlinear equation

$$F(y,v) = (0,0), \quad (y,v) \in Y \tag{50}$$

solves the null controllability problem for (1).

For their "natural" norms, Y and W are Hilbert spaces. On the other hand, since $f \in C_b^1(\mathbf{R})$, $F : Y \mapsto W$ is well defined and C^1. Indeed, for any $(y,v),(z,u) \in Y$ and any $a \in \mathbf{R}_+$, we have

$$\frac{1}{a}\left(F(y+az, v+au) - F(y,v)\right) = (z_t - (az_x)_x + \frac{1}{a}(f(y+az) - f(y)) - u1_\omega, y(\cdot\,,0))$$

and it is immediate from Lebesgue's Theorem that

$$\iint_{Q_T} \rho_0^2 \left|\frac{1}{a}(f(y+az) - f(y)) - f'(y)z\right|^2 dx\,dt \to 0 \quad \text{as} \quad a \to 0^+.$$

This argument also shows that

$$F'(y,v)\cdot(z,u) = (z_t - (az_x)_x + f'(y)z - u1_\omega, z(\cdot\,,0)) \quad \forall (y,v),(z,u) \in Y. \tag{51}$$

Notice that, for any $(y,v) \in Y$, the linear mapping $F'(y,v) \in \mathcal{L}(Y;W)$ is an isomorphism. Indeed, for any $(k,y_0) \in W$, the equation

$$F'(y,v)\cdot(z,u) = (k,z_0), \quad (z,u) \in Y$$

possesses exactly one solution, given as follows: $z = \rho^{-2}L^*_{f'(y)}p$ and $u = -\rho_0^{-2}p|_{q_T}$, where p is the unique solution to the linear problem

$$M(y;p,q) = \iint_{Q_T} kq\,dx\,dt + \int_0^1 y_0(x)\,q(x,0)\,dx \quad \forall q \in P; \quad p \in P$$

and $M(y;\cdot\,,\cdot)$ stands for the bilinear form

$$M(y;p,q) = \iint_{Q_T} \rho^{-2}L^*_{f'(y)}p\,L^*_{f'(y)}q\,dx\,dt + \iint_{Q_T} \rho_0^{-2}p\,q\,dx\,dt \quad \forall p,q \in P.$$

Therefore, in order to solve (50), it makes sense to apply Newton's method.

ALG 3 (Newton):

(i) Choose $(y^0,z^0) \in Y$.

(ii) Then, for any given $n \geq 0$ and $(y^n,z^n) \in Y$, solve the linear problem

$$F'(y^n,v^n)\cdot(z^n,u^n) = -F(y^n,v^n), \quad (z^n,u^n) \in Y,$$

i.e. find z^n and u^n such that $(z^n,u^n) \in Y$ and

$$\begin{cases} z^n_t - (a(x)z^n_x)_x + f'(y^n)\,z^n = k^n + u^n\,1_\omega, & (x,t) \in Q_T \\ z^n(x,t) = 0, & (x,t) \in \Sigma_T \\ z^n(x,0) = -y^n(x,0) + y_0(x), & x \in (0,1) \end{cases} \tag{52}$$

where $k^n = -y^n_t + (a(x)y^n_x)_x - f(y^n) + v^n\,1_\omega$ and set $y^{n+1} = y^n + z^n$ and $v^{n+1} = v^n + u^n$.

At each step of this algorithm, the task is thus to solve the linear problem

$$M(y^n; p^n, q) = \iint_{Q_T} k^n q \, dx \, dt + \int_0^1 (y_0(x) - y^n(x,0)) \, q(x,0) \, dx \quad \forall q \in P; \quad p^n \in P.$$

After this, we set $z^n = -\rho^{-2} L^*_{f'(y^n)} p^n$, $u^n = \rho_0^{-2} p^n|_{q_T}$, $y^{n+1} = y^n + z^n$ and $v^{n+1} = v^n + u^n$.

An equivalent and more useful formulation frome the practical viewpoint is the following:

ALG 3' (Newton):

(i) Choose $(y^0, z^0) \in Y$.

(ii) Then, given $n \geq 0$ and $(y^n, v^n) \in Y$, solve in $(y^{n+1}, v^{n+1}) \in Y$ the linear problem

$$F'(y^n, v^n) \cdot (y^{n+1} - y^n, v^{n+1} - v^n) = -F(y^n, v^n),$$

i.e. find y^{n+1} and v^{n+1} such that $(y^{n+1}, v^{n+1}) \in Y$ and

$$\begin{cases} y_t^{n+1} - (a(x) y_x^{n+1})_x + f'(y^n) \, y^{n+1} = v^{n+1} \, 1_\omega + f'(y^n) \, y^n - f(y^n), & (x,t) \in Q_T \\ y^{n+1}(x,t) = 0, & (x,t) \in \Sigma_T \\ y^{n+1}(x,0) = y_0(x), & x \in (0,1). \end{cases} \tag{53}$$

As we see, in Newton's method the nonlinear term $f(y^{n+1}) = y^{n+1} g(y^{n+1})$ is replaced by $f(y^n) + f'(y^n)(y^{n+1} - y^n)$.

4 Numerical approximation and numerical experiments

4.1 A finite element space-time approximation

In order to approximate the linear problems (9), we will use the second approach in Section 2.1; precisely, the mixed elliptic variational formulation (19), which requires a space-time finite element method.

For "large" integers N_x and N_t, we set $\Delta x = 1/N_x$, $\Delta t = T/N_t$ and $h = (\Delta x, \Delta t)$.

We introduce the associated uniform quadrangulations $\mathcal{Q}_h$, with $Q_T = \cup_{K \in \mathcal{Q}_h} K$. Then, we consider the following finite dimensional spaces:

$$M_h = \{ \, m_h \in C^{0,1}_{x,t}(\overline{Q}_T) : m_h|_K \in (\mathbb{P}_{1,x} \otimes \mathbb{P}_{3,t})(K) \ \ \forall K \in \mathcal{Q}_h, \quad m_h \in Z \, \},$$

$$Q_h = \{ \, q_h \in C^{0,1}_{x,t}(\overline{Q}_T) : q_h|_K \in (\mathbb{P}_{1,x} \otimes \mathbb{P}_{3,t})(K) \ \ \forall K \in \mathcal{Q}_h, \quad q_h|_{\Sigma_T} \equiv 0 \, \},$$

where $\mathbb{P}_{\ell,\xi}$ denotes the space of polynomial functions of order ℓ in the variable ξ.

The functions in M_h and Q_h coincide in each quadrangle $K \in \mathcal{Q}_h$ with a polynomial spanned by $\{1, t, t^2, t^3, x, xt, xt^2, xt^3\}$, involving 8 degrees of freedom. On each quadrangle, the coefficients must be such that the functions belong to $C^{0,1}_{x,t}(\overline{Q}_T)$, the space of the functions in $C^0(\overline{Q}_T)$ that are continuously differentiable with respect to the variable t in $\overline{Q_T}$. In fact, if the (x_i, t_j) denote the nodes of $\mathcal{Q}_h$, a function $m_h \in M_h$ (resp. $q_h \in Q_h$) is uniquely determined by the real numbers $m_h(x_i, t_j)$ and $(m_h)_t(x_i, t_j)$ (resp. $q_h(x_i, t_j)$ and $(q_h)_t(x_i, t_j)$). Notice that we have $M_h \subset L^2(Q_T)$ but $Q_h \not\subset P_*$, i.e. $\rho_0 Q_h \not\subset P$ (see [12]).

This approximation differs from the one used in [12] (where only continuous and piecewise linear functions in t are used) and allows much more accurate results. The main reasons are that both the state y and the control v may have important variations in time near $t = 0$ in the semi-linear case (see the experiments below) and, also, that the adjoint state p may be very oscillatory at time T; we refer to [12, 2, 23] for a discussion on this fact.

The variational problem (19) is then approximated as follows: find $(m_h, \eta_h, \mu_h) \in M_h \times Q_h \times M_h$ such that

$$\begin{cases} \displaystyle\iint_{Q_T} m_h \overline{m}_h dx\, dt + \iint_{q_T} \eta_h, \overline{\eta}_h\, dx\, dt + b_h((\overline{m}_h, \overline{\eta}_h), \mu_h) \\ \displaystyle\quad = \int_0^1 \rho_0(x,0) y_0(x) \overline{\eta}_h(x,0)\, dx + \iint_{Q_T} \rho_0 B(x,y)\, \overline{\eta}_h\, dx\, dt \quad \forall (\overline{m}_h, \overline{\eta}_h) \in M_h \times Q_h \\ b_h((m_h, \eta_h), \overline{\mu}_h) = 0 \quad \forall \overline{\mu}_h \in M_h \end{cases} \tag{54}$$

where we have introduced the notation

$$\begin{aligned} & b_h((m_h, \eta_h), \mu_h) \\ & \quad := \iint_{Q_T} \Big(B_1 \mu_h \eta_h + B_2 \mu_h (\eta_h)_x + B_3 (\mu_h)_x \eta_h + B_4 (\mu_h)_x (\eta_h)_x + B_5 \mu_h (\eta_h)_t - \mu_h m_h \Big)\, dx\, dt \\ & \forall (m_h, \eta_h, \mu_h) \in M_h \times Q_h \times M_h \end{aligned}$$

4.2 Experiments with the fixed point method

In this Section, we will assume that the function f is of the form :

$$f(s) = C_0 s\ \log^p(1 + |s|) \quad \forall s \in \mathbf{R} \tag{55}$$

with $C_0 \in \mathbf{R}$. Clearly, f satisfies $f(0) = 0$ and fulfills the hypotheses of Theorem 1.2 for $p < 3/2$.

Let us describe the numerical behavior of the fixed point method in Section 3.1 (**ALG 1**).

We consider the following data:

$$a(x) \equiv 1/10, \quad p = 1.4, \quad C_0 = -5, \quad T = 1/2. \tag{56}$$

We will first depict the influence of the L^∞-norm of the initial data on the convergence of **ALG 1**. To this purpose, we will take $y_0(x) \equiv \alpha \sin(\pi x)$ for several values of $\alpha > 0$.

In the uncontrolled situation, these data lead to the blow-up of the solution of (1) at times $t_c \approx 0.406, 0.367, 0.339, 0.318$ (before time $t = T$) for $\alpha = 10, 20, 40$ and 80, respectively. We first take $\omega = (0.2, 0.8)$ and initialize **ALG 1** with

$$y^0(x,t) \equiv y_0(x)(1 - t/T)^2. \tag{57}$$

The numerical results we present are obtained with $h = (1/60, 1/60)$.

Figure 1 shows the evolution of

$$e^n = \frac{\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}}{\|y_h^n\|_{L^2(Q_T)}} \tag{58}$$

with respect to n. The results are somehow unexpected, in the sense that, for $\alpha = 10$, $\alpha = 40$ and $\alpha = 80$, the fixed point method converges with a polynomial rate, while for $\alpha = 20$ the method produces a non convergent but bounded sequence $\{y_h^n\}$.

For $\alpha = 20$, we observe that the iterates y_h^n approach alternatively two different functions and, in particular, $y_h^{n+2} \approx y_h^n$ for n large enough (this means that we get two fixed points not for Λ_0 but for the application $\Lambda_0 \circ \Lambda_0$). This confirms that there is no guarantee to get convergence of the fixed point iterates even for $p < 3/2$, for which controllability is known to hold.

We stop the iterates as soon as the inequality

$$e^n = \frac{\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}}{\|y_h^n\|_{L^2(Q_T)}} \leq 10^{-3} \tag{59}$$

is satisfied. Table 1 gives some numerical values. In particular, we check there that the L^2-norm of the control increases with α.

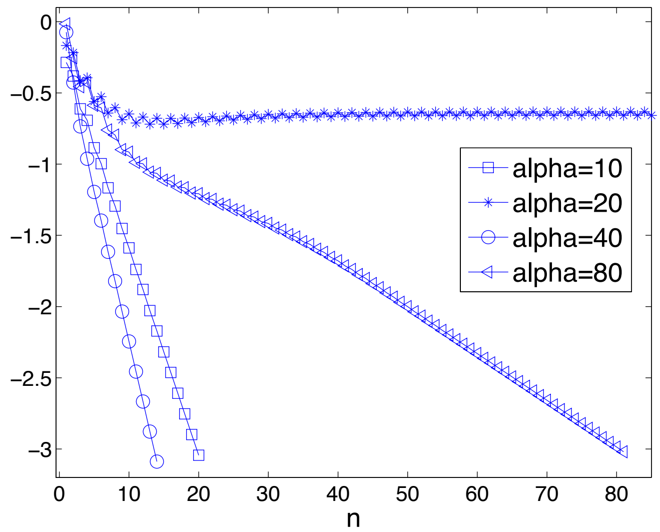

Figure 1: Fixed point method - $h = (1/60, 1/60)$ - $y_0(x) \equiv \alpha \sin(\pi x)$ - $p = 1.4$ - Evolution of $\log_{10}(\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}/\|y_h^n\|_{L^2(Q_T)})$ for $\alpha = 10, 20, 40$ and 80.

	$\|v_h\|_{L^2(q_T)}$	$\|v_h\|_{L^\infty(q_T)}$	$\|y_h\|_{L^2(Q_T)}$	♯ iterates
$\alpha = 10$	3.531×10^1	2.542×10^2	1.742	20
$\alpha = 40$	2.142×10^2	2.053×10^3	6.654	14
$\alpha = 80$	5.109×10^2	7.021×10^3	14.410	81

Table 1: Fixed point method - $h = (1/60, 1/60)$ - $y_0(x) \equiv \alpha \sin(\pi x)$ - $p = 1.4$ - Norms for $\alpha = 10, 40$ and $\alpha = 80$.

Let us now take $\alpha = 40$, i.e. $y_0(x) \equiv 40 \sin(\pi x)$. In Figure 4.2, we have displayed the control v_h and the corresponding controlled solution y_h obtained via **ALG 1**. The shape and values of v_h are similar to those reported in the linear case in [12, 13]. The main difference is the amplitude of the control near $t = 0$, that increases with α (and also with C_0 and p), in order to prevent the solution to (1) to blow up. For this reason, the norms of the v_h and y_h are much less sensible to the controllability time T than in the linear situation. Contrarily, y_h differs from what is found for the linear problem, since we observe some oscillations with respect to the spatial variable near $t = 0$ whose number increases with α.

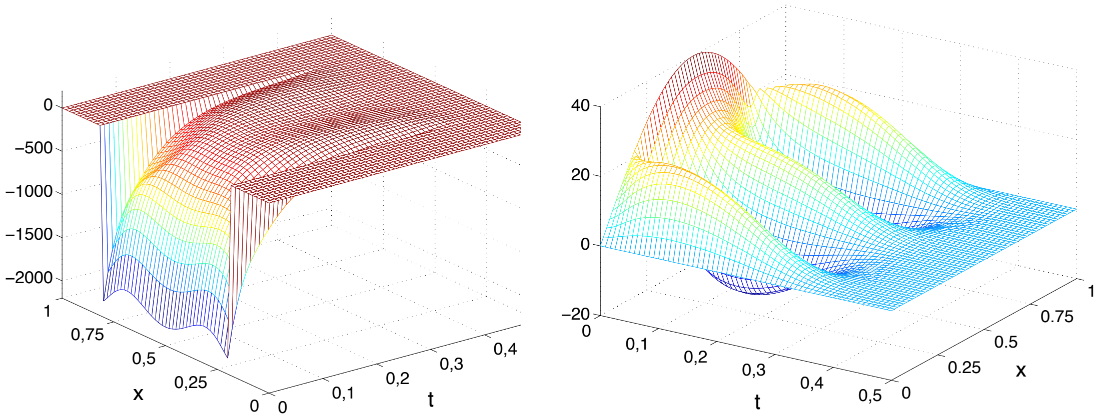

Figure 2: Fixed point method - $h = (1/60, 1/60)$ - $y_0(x) \equiv 40\sin(\pi x)$ - $p = 1.4$ - Control v_h (Left) and corresponding controlled solution y_h (Right) in Q_T.

We present in Figure 3 the evolution of (58) with respect to n for $p = 1.4, 1.5, 1.6, 1.7$, initializing in each case with $y^0(x,t) \equiv 40\sin(\pi x)(1-t/T)^2$. For the first three values of p, we observe convergence, with a linear rate decreasing with p. The numbers of iterates to get (59) are 14, 21 and 39, respectively. The case $p = 1.7$ provides a non-convergent sequence $\{y_h^n\}$. The same phenomenon is observed with this p if we initialize **ALG 1** with the controlled state obtained for $p = 1.6$. Values of p greater than 1.7 also provide non-convergent sequences.

We also observe that the fixed point method converges efficiently for these data for which the solution to (1) is globally defined: for instance, for $(\alpha, p) = (1000, 1)$ or $(\alpha, p) = (1, 30)$, convergence holds.

	$\|v_h\|_{L^2(q_T)}$	$\|v_h\|_{L^\infty(q_T)}$	$\|y_h\|_{L^2(Q_T)}$	$\sharp$ iterates
$p = 1.4$	2.142×10^2	2.053×10^3	6.654	14
$p = 1.5$	2.301×10^2	2.501×10^3	7.139	21
$p = 1.6$	2.481×10^2	3.110×10^3	7.423	39

Table 2: Fixed point method - $h = (1/60, 1/60)$ - $y_0(x) \equiv 40\sin(\pi x)$ - Norms for $p = 1.4, 1.5$ and $p = 1.6$.

We have compared the results obtained with uniform quadrangles and $P_1 \times P_3$ finite elements, see Section 4.1, to others on finer non-uniform meshes and $P_2 \times P_2$ finite elements, obtained with the FREE-FEM software developed by F. Hecht, O. Pironneau and collaborators at University Paris 6; see [17]. Qualitatively, we have found the same results and, in particular, lack of convergence for $\alpha = 20$.

4.3 Experiments with the least squares approach

We now present some numerical experiments obtained with the least squares approach introduced in Section 3.2.

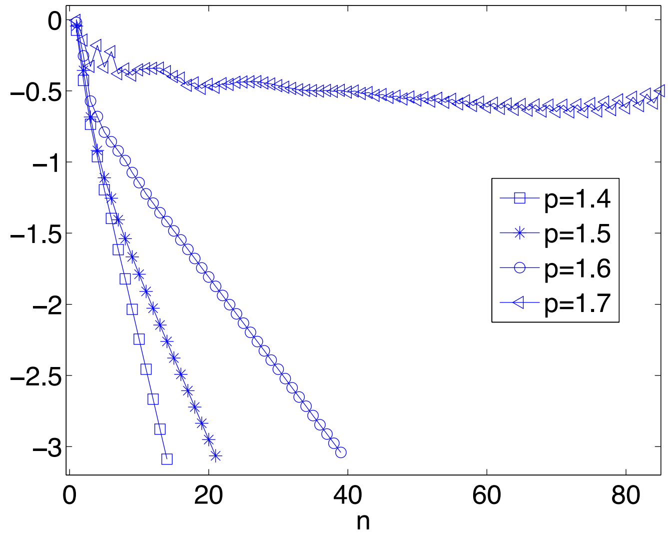

Figure 3: Fixed point method - $h = (1/60, 1/60)$ - $y_0(x) \equiv 40\sin(\pi x)$ - $p = 1.4$ - Evolution of $\log_{10}(\|\Lambda_0(y_h^n) - y_h^n\|_{L^2(Q_T)}/\|y_h^n\|_{L^2(Q_T)})$ for $p = 1.4, 1.5, 1.6$ and $p = 1.7$.

The function f is changed. Now we use f_η, given as follows:

$$f_\eta(s) = C_0 s \ \log^p(1 + |s|_\eta) \quad \forall s \in \mathbf{R}, \quad |s|_\eta := \sqrt{s^2 + \eta^2} - \eta, \tag{60}$$

so that, for any $\eta > 0$, the function $g_\eta(s) \equiv C_0 \log^p(1 + |s|_\eta)$ belongs to $C_b^1(\mathbf{R})$. Moreover, we have $f_\eta(0) = 0$ and Theorem 1.2 applies for f_η, since f_η and f are equivalent at infinity. We will take $\eta = 10^{-1}$ and, again, $p = 1.4$.

Once again, we initialize the iterates with $z^0(x, t) \equiv y_0(x)(1 - t/T)^2$.

We first fix a, p, C_0 and T as in Section and take $\alpha = 10, 20, 40, 80$; see (56). Table 3 reports the numerical results obtained after 100 iterates of the Polak-Ribière version of the conjugate gradient algorithm **ALG 2'**; see (48).

The evolution of the ratio e^n is depicted in Figure 4.3-Left. In contrast with the fixed point approach, we observe here that the sequence $\{z^n\}$ converges toward a fixed point of Λ also for $\alpha = 20$. We note however that the convergence rate is quite low and in fact decreases as α grows; in particular, the case $\alpha = 80$ does not give satisfactory results. The Fletcher-Reeves version of the conjugate gradient algorithm does not give better results.

In fact, for $\alpha = 80$, as reported in Figure 4.3-Right, we observe a better behavior of the steepest descent gradient iterates, i.e. **ALG 2** (recall that the cost R is not quadratic, so that the convergence of the conjugate gradient algorithm is not guaranteed).

By comparing the results in Tables 3 and 1 for $\alpha = 10$ and $\alpha = 40$, we observe similar norms for the controls and the associated states. We also see that, for $\alpha = 10$, $\alpha = 40$ and $\alpha = 80$, these results can be improved with **ALG 1** starting from the approximate solution $z_h^{n=100}$, obtained after 100 iterates. Indeed, this leads to convergence in a few additional steps.

However, once again, for $\alpha = 20$ the corresponding sequence $\{y_h^n\}$ does not converge and this strategy fails. This suggests that the behavior of the fixed point algorithm **ALG 1** does not rely only on its initialization, but also on the structure of the fixed points for a given set of data.

	$\|v_h\|_{L^2(q_T)}$	$\|v_h\|_{L^\infty(q_T)}$	$\|z_h\|_{L^2(Q_T)}$	$\|R'(z_h)\|_{L^2(Q_T)}$	e^n
$\alpha = 10$	3.507×10^1	2.532×10^2	1.753	1.27×10^{-3}	1.43×10^{-3}
$\alpha = 20$	8.781×10^1	7.323×10^2	3.180	1.44×10^{-3}	1.54×10^{-3}
$\alpha = 40$	2.137×10^2	2.048×10^3	6.651	5.42×10^{-3}	3.39×10^{-3}
$\alpha = 80$	2.526×10^2	3.299×10^3	14.73	2.23×10^{-1}	7.89×10^{-1}

Table 3: Least squares method approach after 100 iterates - $h = (1/60, 1/60)$ - $y_0(x) \equiv \alpha \sin(\pi x)$ - $p = 1.4$ - Norms for $\alpha = 10, 20, 40, 80$. Here, $e^n = \|\Lambda(z_h^n) - z_h^n\|_{L^2(Q_T)} / \|z_h^n\|_{L^2(Q_T)}$.

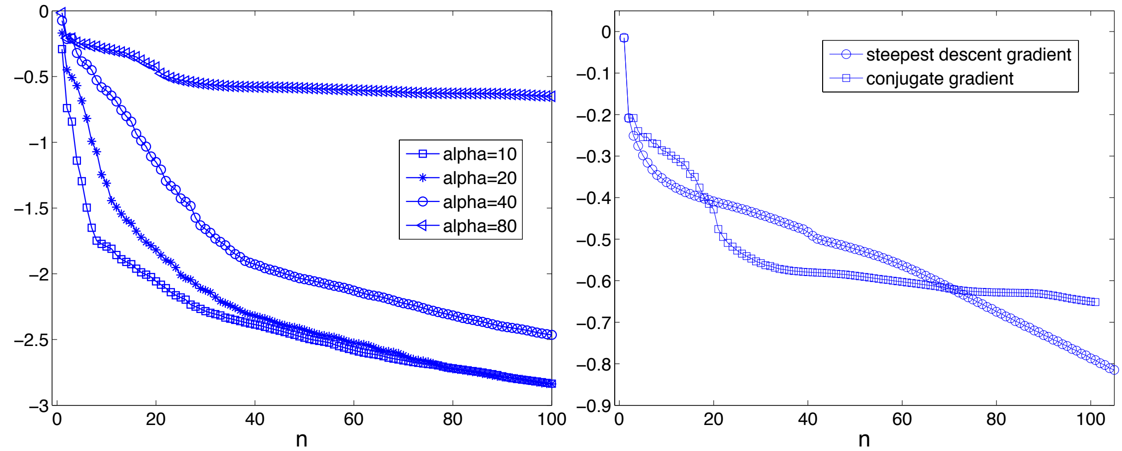

Figure 4: Least squares method - $h = (1/60, 1/60)$ - $y_0(x) \equiv 40 \sin(\pi x)$ - Evolution of $\log_{10}(\|\Lambda(z_h^n) - z_h^n\|_{L^2(Q_T)} / \|z_h^n\|_{L^2(Q_T)})$ - Left: $\alpha = 10, 20, 40, 80$ and algorithm **ALG 2'** - Right: $\alpha = 80$ and algorithms **ALG 2** and **ALG 2'**.

4.4 Experiments with the Newton-Raphson method

We now consider the Newton-Raphson algorithm introduced in Section 3.3.

With f defined by (55), we have

$$f'(y^n) y^{n+1} = C_0 \left[\log^p(1 + |y^n|) + p \frac{|y^n| \log^{p-1}(1 + |y^n|)}{1 + |y^n|} \right] y^{n+1} := k(y^n(x, t))\, y^{n+1}(x, t) \tag{61}$$

and

$$f'(y^n) y^n - f(y^n) = C_0 p y^n |y^n| \frac{\log^{p-1}(1 + |y^n|)}{1 + |y^n|} := \phi(y^n(x, t)). \tag{62}$$

Thus, we consider the mixed formulation (19) and its numerical approximation (54) with $A(x, t) = k(y^n(x, t))$ and $B(x, t) = \phi(y^n(x, t))$ to compute the pair (y^{n+1}, v^{n+1}).

We consider the same data as in Section 4.2, except that we initialize **ALG 3** with $y^0(x, t) \equiv y_0(x)(1 - t/T)^2 \rho^{-1}(x, t)$ instead of (57); this ensures that, at the initial step,

$$\iint_{Q_T} \rho_0^2 |B(x, t)|^2 \, dx \, dt = \iint_{Q_T} \rho_0^2 |\phi(y^0(x, t))|^2 \, dx \, dt < +\infty.$$

Figure 5 depicts the evolution of e^n, given by (58) for $\alpha = 10, 20, 40, 80$.

As fin the case of the fixed point approach **ALG 1**, we observe convergence except for $\alpha = 20$. In that case, we notice that, after some iterates, $\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}$ becomes approximately constant. We also observe that $y_h^{n+2} \approx y_h^n$ and y_h^n and y_h^{n+1} are very approximately symmetric to each other with respect to the line $\{1/2\} \times (0,T)$, so that $\|y_h^n\|_{L^2(Q_T)} \approx \|y_h^{n+1}\|_{L^2(Q_T)}$.

Moreover, we notice from Table 4 and Figure 4.4 that the Newton approach leads to a different state-control pair (y_h, v_h), with higher L^∞-norm for v_h (the L^∞ norm is reached at $t = 0$) and lower L^2 norm for y_h. This control is more active at the beginning (near $t = 0$), so that the dissipation of the state y is faster. This may be related to the fact that, with the Newton linearization of $f(y^{n+1})$, the potential term in the linearized state equation for y is no more $g(y^n)$ but $f'(y^n)$, which has a larger L^∞ norm, see (61).

We also remark that, for the considered data, **ALG 3** converges faster for $\alpha = 40$ than for $\alpha = 10$, as was the case for **ALG 1**. A bit more surprising is the fact the fixed point approach converges faster than the Newton-Raphson method. Thus, for $\alpha = 80$, **ALG 3** converges but very slowly; after 500 iterates, we only have

$$\frac{\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}}{\|y_h^n\|_{L^2(Q_T)}} \approx 1.34 \times 10^{-2}.$$

Figure 4.4 depicts the control v_h and the corresponding controlled solution y_h for $\alpha = 40$ (to be compared to Figure 4.2).

	$\|v_h\|_{L^2(q_T)}$	$\|v_h\|_{L^\infty(q_T)}$	$\|y_h\|_{L^2(Q_T)}$	$\sharp$ iterates
$\alpha = 10$	3.489×10^1	3.12×10^2	1.467	58
$\alpha = 40$	2.110×10^2	2.587×10^3	5.248	18
$\alpha = 80$	5.033×10^2	8.589×10^3	10.976	> 500

Table 4: Newton-Raphson method - $h = (1/60, 1/60)$ - $y_0(x) \equiv \alpha \sin(\pi x)$ - $p = 1.4$ - Norms for $\alpha = 10, 40, 80$.

Let us finally consider a more difficult situation, with $\omega = (0.2, 0.5)$. We take $y_0(x) \equiv 10\sin(\pi x)$, the other data being unchanged, see (56).

In this situation, **ALG 1**, **ALG 2** and **ALG 2'** converge after a moderate amount of steps. As before, they provide similar numerical values (**ALG 1** and **ALG 3** are both initialized with $y_0(x)(1 - t/T)^2\rho^{-1}$).

As observed in [12], the L^2-norm of the control increases significantly when its support is reduced: we get $\|v_h\|_{L^2(q_T)} \approx 4.898 \times 10^2$ and $\|v_h\|_{L^\infty(q_T)} \approx 1.706 \times 10^4$. The variation in time of the corresponding controlled solution is also more intense; we get $\|y_h\|_{L^2(Q_T)} \approx 14.04$ and $\|y_h\|_{L^\infty(Q_T)} \approx 85.26$!

Figure 7 displays the computed pair (v_h, y_h) in Q_T. On the other hand, **ALG 3** does not converge. The evolution of $\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}/\|y_h^n\|_{L^2(Q_T)}$ is given in Figure **??**.

5 Conclusions

Although additional experiments are necessary to understand deeper the behavior of the previous algorithms, one may state the following partial conclusions:

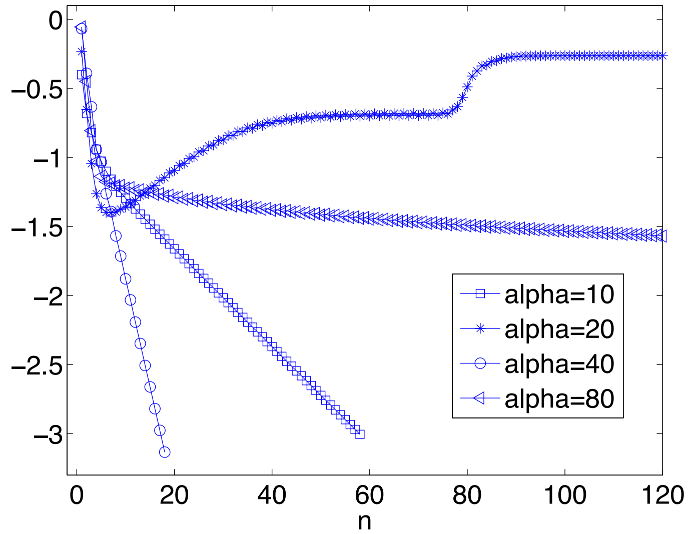

Figure 5: Newton-Raphson method - $h = (1/60, 1/60)$ - $y_0(x) \equiv \alpha \sin(\pi x)$ - $p = 1.4$ - Evolution of $\log_{10}(\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}/\|y_h^n\|_{L^2(Q_T)})$ for $\alpha = 10, 20, 40, 80$.

- The "natural" fixed point approach, obtained after first order linearization of the state equation, appears to be the most efficient in terms of convergence, at least when y_0 is not too large.
- On the contrary, the least squares formulation, for which the numerical implementation is more involved, seems to be more robust and provide convergence in more situations.
- The Newton-Raphson approach, obtained after a linearization of second order, is apparently more sensitive to the initialization and, in general, does not seem to improve convergence.

Since the numerical approximation of null controls for the linear heat equation still remains a challenge, *a fortiori,* in the considered nonlinear situations, it becomes an even more difficult task. We believe however that the least squares approach will make it possible to obtain good approximations of null controls in many nontrivial situations.

It is also remarkable that the previous approach and algorithms can be extended to higher spatial dimensions and can be used in the context of many other controllable systems for which appropriate Carleman estimates are available: wave-like equations, Navier-Stokes-like systems, etc.; see [7, 15, 20].

References

[1] V. Barbu, *Exact controllability of the superlinear heat equation*, Appl. Math. Optim. Optimization, Theory and Applications 42, (2000), no. 1, 73–89.

[2] 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.

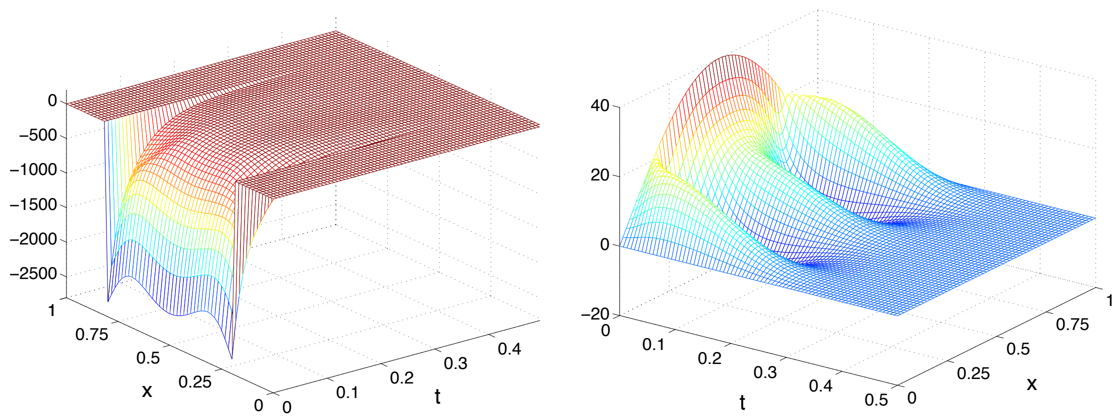

Figure 6: Newton-Raphson method - $h = (1/60, 1/60)$ - $y_0(x) \equiv 40\sin(\pi x)$ - $p = 1.4$ - Control v_h (Left) and corresponding controlled solution y_h (Right) in Q_T.

[3] F. Boyer, F. Hubert and J. Le Rousseau, *Uniform null-controllability properties for space/time-discretized parabolic equations.* To appear in Numerische Mathematik (2011).

[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, (1994), no. 3, 429–484.

[5] T. Cazenave and A. Haraux, *Introduction aux problèmes d'évolutions semi-linéaires*, Mathématiques et Applications, Ellipses, Paris 1989.

[6] I. Charpentier and Y. Maday, *Identifications numériques de contrôles distribués pour l'équation des ondes*, C.R. Acad. Sci. Paris Sér. I 322 (1996), no. 8, 779-784.

[7] J-M. Coron, *Control and Nonlinearity*, vol. 136 of Mathematical Surveys and Monographs, AMS, Providence, RI, 2007.

[8] J-M. Coron and E. Trélat, *Global steady-state controllability of one dimensional semilinear heat equations*, SIAM J. Control Optim. 43 (2004), no. 2, 549-569.

[9] 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.

[10] E. Fernández-Cara, *Null controllability of the semilinear heat equation,* ESAIM:COCV 87 (1997), no. 2, 87-103.

[11] E. Fernández-Cara and S. Guerrero, *Global Carleman inequalities for parabolic systems and applications to controllability,* SIAM J. Control Optim. 45 (2006), no. 4, 1399–1446.

[12] E. Fernández-Cara and A. Münch, *Numerical null controllability of the 1-d heat equation: primal algorithms.* Preprint (2010).

[13] E. Fernández-Cara and A. Münch, *Numerical null controllability of the 1-d heat equation: dual algorithms.* Preprint (2010).

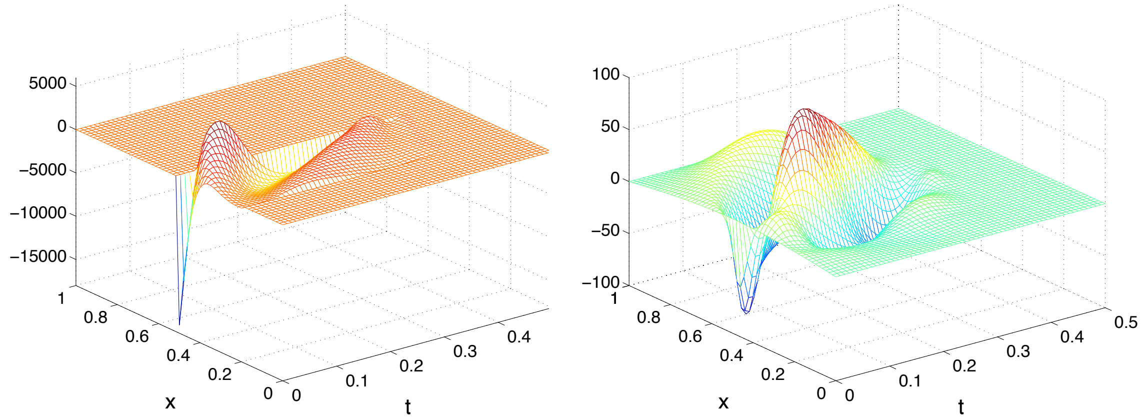

Figure 7: Fixed point method - $h = (1/60, 1/60)$ - $y_0(x) \equiv 10\sin(\pi x)$ - $p = 1.4$ - $\omega = (0.2, 0.5)$ - Control v_h (Left) and corresponding controlled solution y_h (Right) in Q_T.

[14] E. Fernández-Cara and E. Zuazua, *Null and approximate controllability for weakly blowing up semilinear,* Ann. Inst. Henri Poincaré, Analyse non linéaire. 17 (2000), no. 5, 583–616.

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

[16] 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 (2008).

[17] F. Hecht, A. Le Hyaric, J. Morice, K. Ohtsuka and O. Pironneau, *Free-Fem++: Third Edition, Version 3.12*, http://www.freefem.org/++.

[18] O. Yu. Imanuvilov, *Controllability of parabolic equations (Russian),* Mat. Sb. 186 (1995), no. 6, 109–132; translation in Sb. Math. 186 (1995), no. 6, 879–900.

[19] S. Labbé and E. Trélat, *Uniform controllability of semidiscrete approximations of parabolic control systems,* Systems Control Lett. 55 (2006), no. 7, 597–609.

[20] I. Lasiecka and R. Triggiani, *Exact controllability of semilinear abstract systems with applications to waves and plates boundary control,* Appl. Math. & Optim. 23 (1991), 109–154.

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

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

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

[24] D.L. Russell, *A unified boundary controllability theory for hyperbolic and parabolic partial differential equations,* Studies in Appl. Math. 52 (1973), 189–221.

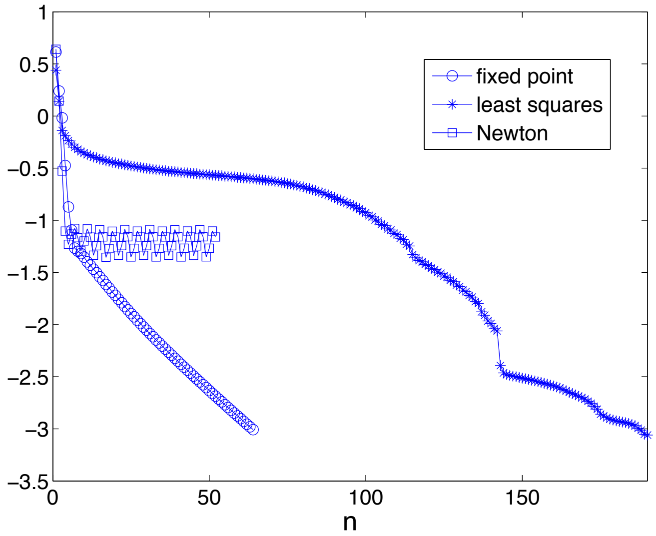

Figure 8: $h = (1/60, 1/60)$ - $y_0(x) \equiv 10\sin(\pi x)$ - $p = 1.4$ - $\omega = (0.2, 0.5)$ - Evolution of $\log_{10}(\|y_h^{n+1} - y_h^n\|_{L^2(Q_T)}/\|y_h^n\|_{L^2(Q_T)})$.

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

[26] E. Zuazua, *Exact boundary controllability for the semilinear wave equation,* In "Nonlinear Partial Differential Equations and their Applications", Vol. X (H. Brezis & J.-L. Lions eds.), Pitman, New York 1991.