

ON THE PENALISED HUM APPROACH AND ITS APPLICATIONS TO THE NUMERICAL APPROXIMATION OF NULL-CONTROLS FOR PARABOLIC PROBLEMS

F. BOYER¹

Abstract. This article deals with the problem of computing numerical approximations of null-controls for parabolic equations or systems by using the Hilbert Uniqueness Method (HUM). We mainly review recent results on this subject but we also provide new elements to emphasize the main ideas underlying the penalised HUM approach which is at the heart of the methods used in practice. We give many numerical illustrations.

Résumé. Cet article est consacré à l'étude du problème du calcul approché de contrôles à zéro pour des équations ou des systèmes paraboliques par le biais de la méthode HUM (Hilbert Uniqueness Method). On donne un aperçu des résultats récents sur le sujet tout en mettant en lumière certains aspects fondamentaux de la méthode HUM pénalisée, qui se trouve au coeur des algorithmes proposés. De nombreuses illustrations numériques sont également données.

INTRODUCTION

The main goal of this article is to present methods and results concerning the numerical computation of null-controls for parabolic problems. Basically one can think of the following general strategy: first build a semi- or fully-discrete approximation of the PDE under study; then compute a null-control for this discrete system (which is an ODE or a finite-dimensional discrete-time dynamical system); finally try to prove that this discrete control is actually an approximation of a null-control for the initial continuous problem. More precisely, the strategy that we follow is the one initially introduced in the pioneering works [GL94, GLH08] and is based on a formulation of the control problem under the form of a suitable convex quadratic optimisation problem. This approach, which consists in characterizing and building the minimal L^2 -norm control if it exists, is often referred to as the Hilbert Uniqueness Method (HUM in what follows).

Unfortunately, this natural and naive approach may lead to some theoretical and/or numerical issues:

- (1) Due to the fact that parabolic equations enjoy smoothing properties along time, the original HUM dual functional (see formula (1.5) with $\varepsilon = 0$) is known to be coercive only on a very large abstract space. This leads to severe issues when the method is applied for numerical purposes (see for instance the discussion in [MZ10]).
- (2) Additionally, it may happen that a fully- or semi-discrete system associated with a null-controllable continuous problem is not itself controllable for some values of the discretisation parameters. The above strategy will thus fail.

¹ Aix Marseille Université, CNRS, Centrale Marseille, LATP, UMR 7353, 13453 Marseille, FRANCE

- (3) Finally, even if the discrete system is known to be controllable, the actual computation of a corresponding null-control can be very expensive on fine meshes, mainly because one tries to approximate a problem which is ill-posed in the natural energy space (see the first point above).

For all these reasons, in view of numerical computations, the HUM approach for parabolic problems is almost never used in its original version. Many authors, including [GL94, GLH08], make use of a suitable penalised version of HUM which aims to relax the problem. The discrete problem which is obtained thus depend not only on the discretisation parameters but also on the penalisation parameter ε .

We would like in this paper to contribute to the clarification of the relations between these parameters by reviewing the theoretical results known up to now on this problem. To this end, we start this article by giving a careful analysis of the penalised HUM approach at the continuous level for the approximate- and null-controllability problems (Section 1). In particular, this analysis allows us to recover some well-known results in PDE controllability theory, using constructive proofs based on HUM, instead of more standard functional analysis arguments. In addition, our approach allows us to analyse how the controllability properties of the system depend on the initial data. This is important in view of applications to numerical analysis since the initial data of discrete problems are, by definition, approximations of the initial data of the continuous problem (see for instance the discussion in Section 1.7). The main result of this first section is Theorem 1.7 which states that the controllability of the system can be completely characterised by means of the behavior of the penalised HUM functional when the penalisation parameter ε goes to 0. This result is then sharpened in Theorem 1.11 which shows, in some sense, that the controlled solution at final time always converges as $\varepsilon \rightarrow 0$. It is then sufficient to check whether or not this limit is 0 to determine if the system is approximately controllable or not.

In Section 2, we apply the ideas developed in the first section to a family of semi-discrete (in space) approximations of the original problem. We explain why uniform null-controllability (with respect to discretisation parameter) of these problems is not necessarily a satisfactory notion and we rather consider the notion of $\phi(h)$ -null-controllability in which the penalisation idea is used in conjunction with the discretisation process by choosing a penalisation parameter $\varepsilon = \phi(h)$ depending on the discretisation parameter. We then review some recent theoretical results on this problem in various situations.

In Section 3, we discuss the case of fully-discrete approximations of the parabolic problem under study. Most of the results of the semi-discrete case can be adapted, yet two additional issues appear: it is first needed to compute the correct adjoint problem in each case and then, suitable conditions between the time step and the mesh size are identified that ensure a good behavior of the overall algorithm.

We conclude this article by giving in Section 4 various numerical results aiming to illustrate the available theoretical results but we also investigate some more complex problems, in particular, for parabolic systems.

Before starting our analysis, we wish to mention that alternative approaches have been recently proposed for computing approximate controls for parabolic problems.

The first one, introduced and developed in [FCM12a, FCM12b, FCM13], is to formulate the control problem as an optimisation problem for a functional different from HUM which involves weighted L^2 norms of the control and of the solution of the parabolic equation itself. The (singular) weights are chosen similar to those that naturally appear in Carleman estimates for the corresponding problem if available (see [FI96]). The schemes derived from this strategy are shown to give satisfactory results. However, the price to pay is that the PDE problem that one has to solve at the end is much more complex than the initial one: typically a fourth order 2D elliptic problem has to be solved to address a 1D heat equation. This method therefore seems to be quite intricate to extend to more general problems (multi-D or systems), in particular because suitable Carleman weights are not necessarily easy to determine in some cases (in fact Carleman estimates are not always available for complex systems).

The second alternative strategy that we want to mention, the so-called “variational approach”, was introduced in [Ped10, MP11]. In short, it consists in viewing the partial differential equation itself as a constraint for an energy functional defined on the space of all the functions satisfying the boundary conditions and the initial/final conditions. From a numerical point of view, in the case of the heat equation for instance, each iteration of the conjugate gradient type algorithm requires the resolution of a corrector equation that is a space/time 2D elliptic

problem. Here also the dimension of the system is thus in some sense increased, and the development of the algorithm strongly depends on the problem under study.

Compared to those methods, the simple HUM approach that we study in the present work has the very important advantage that, at the end, the numerical scheme for solving the forward and the backward problem (see Section 4.1) can be almost chosen at the user's convenience. In particular, the two solvers can actually be used as black-boxes. Moreover, the strategy is almost problem independent (at least for linear parabolic problems) that is to say that the same overall algorithm can be used without modifications for N-dimensional problems and parabolic systems rather than scalar equations even with variable coefficients, as we shall see in the final section.

1. THE HILBERT UNIQUENESS METHOD AND ITS PENALISED VERSION

1.1. Introduction

Let $(E, \langle \cdot, \cdot \rangle)$ be an Hilbert space whose norm is denoted by $\|\cdot\|$. Let $(\mathcal{A}, D(\mathcal{A}))$ be an unbounded operator in E such that $-\mathcal{A}$ generates an analytic semi-group in E that we denote by $t \mapsto e^{-t\mathcal{A}}$. We denote by $(\mathcal{A}^*, D(\mathcal{A}^*))$ the adjoint of this operator and by $t \mapsto e^{-t\mathcal{A}^*}$ the corresponding semi-group.

Let $(U, [\cdot, \cdot])$ be another Hilbert space whose norm is denoted by $\|\cdot\|$ and \mathcal{B} be a bounded operator from U into $D(\mathcal{A}^*)'$ and $\mathcal{B}^* : D(\mathcal{A}^*) \rightarrow U$ its adjoint. We assume that \mathcal{B}^* satisfies the following regularity property

$$\left(t \mapsto \mathcal{B}^* e^{-t\mathcal{A}^*} \psi \right) \in L^2(0, T; U), \quad \forall \psi \in E,$$

and for some $C > 0$

$$\left\| \mathcal{B}^* e^{-\cdot\mathcal{A}^*} \psi \right\|_{L^2(0, T; U)} \leq C \|\psi\|, \quad \forall \psi \in E. \quad (1.1)$$

Let $T > 0$ be given. For any $y_0 \in E$ and $v \in L^2(0, T; U)$ we consider the non-homogeneous evolution problem

$$\begin{cases} \partial_t y + \mathcal{A}y = \mathcal{B}v, \\ y(0) = y_0. \end{cases} \quad (\mathcal{P})$$

Since \mathcal{B} is not necessarily bounded from U into E , solutions have to be understood in the following weak sense (see e.g. [Cor07]).

Theorem 1.1. *For any $y_0 \in E$, there exists a unique $y \in \mathcal{C}^0([0, T], E)$ such that for any $t \in [0, T]$ and any $\psi \in E$, we have*

$$\langle y(t), \psi \rangle - \langle y_0, e^{-t\mathcal{A}^*} \psi \rangle = \int_0^t \left[v(s), \mathcal{B}^* e^{-(t-s)\mathcal{A}^*} \psi \right] ds. \quad (1.2)$$

This solution is now referred to as $t \mapsto y_{v, y_0}(t) \in E$.

Moreover, the map

$$(y_0, v) \in E \times L^2(0, T; U) \mapsto y_{v, y_0} \in \mathcal{C}^0([0, T], E),$$

is continuous.

Notice that we have

$$y_{0, y_0}(t) = e^{-t\mathcal{A}} y_0.$$

For simplicity, the solution at time T , which is of particular interest in what follows, will be denoted by

$$\mathcal{L}_T(v|y_0) = y_{v, y_0}(T).$$

The linear operator $\mathcal{L}_T(\cdot|\cdot)$ is then continuous from $L^2(0, T; U) \times E$ into E .

Assumptions on \mathcal{B} are satisfied by bounded operators $\mathcal{B} : U \rightarrow E$ but more general operators are allowed. For instance, for the usual heat equation $E = U = L^2(\Omega)$, $\mathcal{A} = -\Delta$ with $D(\mathcal{A}) = H^2(\Omega) \cap H_0^1(\Omega)$, the operator \mathcal{B} is

allowed to contain first order differential operators, or to be the operator associated with the Neumann boundary control. However, (1.1) does not hold for the Dirichlet boundary control problem for the same equation. Indeed, in that case (1.1) is only valid for $\psi \in H^{\frac{3}{2}+\varepsilon}(\Omega) \subset D(\mathcal{A})$ and it is necessary to adapt all the framework to solutions defined in a weaker sense, namely in $\mathcal{C}^0([0, T], D(\mathcal{A})')$. The HUM presented below can be adapted to this context but we shall not present the details here, see e.g. [BDPDM93, GL94, CT06, GLH08].

For any $\delta \geq 0$ we define the (possibly empty) closed convex set

$$\text{Adm}(y_0, \delta) = \{v \in L^2(0, T; U), \text{ s.t. } \|\mathcal{L}_T(v|y_0)\| \leq \delta\}.$$

Definition 1.2. • We say that Problem (\mathcal{P}) is approximately null-controllable at time T , from the initial data y_0 , if

$$\text{Adm}(y_0, \delta) \neq \emptyset, \quad \forall \delta > 0.$$

If this holds for any $y_0 \in E$, we simply say that Problem (\mathcal{P}) is approximately null-controllable at time T .

• We say that Problem (\mathcal{P}) is null-controllable at time T from the initial data y_0 , if

$$\text{Adm}(y_0, 0) \neq \emptyset.$$

If this holds for any $y_0 \in E$, we simply say that the problem is null-controllable at time T .

Remark 1.3. (1) Since the control time T is fixed all along this paper, we shall not systematically recall that these definitions depend on T .

(2) Since we assumed that the semi-group is analytic, it can be observed that the range of $e^{-T\mathcal{A}}$ is dense in E . As a consequence, the approximate null-controllability as defined above is equivalent to the more general property

$$\forall y_0, y_T \in E, \forall \delta > 0, \exists v \in L^2(0, T; U), \text{ such that } \|\mathcal{L}_T(v|y_0) - y_T\| \leq \delta.$$

We shall thus use “approximate controllability” in place of “approximate null-controllability”.

(3) The problem being linear, the null-controllability property is well-known to be equivalent to the controllability to trajectories that is

$$\forall \tilde{y}_0, y_0 \in E, \exists v \in L^2(0, T; U), \text{ such that } \mathcal{L}_T(v|y_0) = \mathcal{L}_T(0|\tilde{y}_0) = e^{-T\mathcal{A}}\tilde{y}_0.$$

1.2. The HUM approach

In the framework of both controllability notions, if one control exists it is certainly not unique. This is the reason why for proving controllability results (but also for the actual computation of such controls, see Section 2), it is useful to specify one such control. For instance, the HUM approach consists in finding the control with the minimal $L^2(0, T; U)$ -norm. More precisely, for any $\delta \geq 0$ such that $\text{Adm}(y_0, \delta)$ is not empty we shall define $v^\delta \in \text{Adm}(y_0, \delta)$ to be the unique control satisfying

$$F(v^\delta) = \inf_{v \in \text{Adm}(y_0, \delta)} F(v), \tag{1.3}$$

where

$$F(v) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^T \|v(t)\|^2 dt = \frac{1}{2} \|v\|_{L^2(0, T; U)}^2, \quad \forall v \in L^2(0, T; U).$$

Definition 1.4. In the case where $\text{Adm}(y_0, 0) \neq \emptyset$, the control v^0 is called the HUM (null-)control associated with the initial data y_0 for the problem under study.

Even though (1.3) has a unique solution, it can be a difficult task to solve it directly because of the nature of the constraints that are involved. As usual, the duality theory can help in dealing with the constraints but will lead to functionals that are either non coercive on the usual dual state space E (for $\delta = 0$; in this case considering a much larger space is required) or that are non smooth (for $\delta > 0$). Both issues can lead to important difficulties when using this approach for numerical purposes.

1.3. The penalised HUM approach

Because of the issues described above, it is convenient to deal with a penalised version of the above mentioned optimisation problems. To this end, for any $\varepsilon > 0$, we define the following quadratic functional

$$F_\varepsilon(v) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^T \|v(t)\|^2 dt + \frac{1}{2\varepsilon} \|\mathcal{L}_T(v|y_0)\|^2, \quad \forall v \in L^2(0, T; U),$$

that we wish to minimise onto the whole space $L^2(0, T; U)$.

We shall show in this first part of the article that the careful study of the behavior of the solution of this penalised optimisation problem as a function of ε makes it possible to recover the most usual theoretical results in the controllability theory for abstract evolution problems.

Moreover, following [GL94, GLH08], this penalised method (applied to a family of finite-dimensional discretised problems) will be at the heart of the numerical methods that we shall analyze in Section 2 and 3.

Let us make some preliminary remarks.

- For any $\varepsilon > 0$, the functional F_ε has a unique minimiser on $L^2(0, T; U)$ that we denote by v_ε . This is due to the fact that F_ε is strictly convex, continuous and coercive.
- For a given $\varepsilon > 0$, it is very likely that $v_\varepsilon \notin \{v^\delta, \delta > 0\}$, where the notation v^δ is introduced in section 1.2.
- This minimiser is characterised by the following Euler-Lagrange equation

$$\int_0^T [v_\varepsilon(t), \tilde{v}(t)] dt + \frac{1}{\varepsilon} \langle \mathcal{L}_T(v_\varepsilon|y_0), \mathcal{L}_T(\tilde{v}|0) \rangle = 0, \quad \forall \tilde{v} \in L^2(0, T; U). \quad (1.4)$$

Applying general results of the Fenchel-Rockafellar theory (see e.g. [ET99]) we can build an associated dual problem as follows. For any $\varepsilon > 0$, we define the functional

$$J_\varepsilon(q^F) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^T \left[\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q^F \right]^2 dt + \frac{\varepsilon}{2} \|q^F\|^2 + \langle y_0, e^{-T\mathcal{A}^*} q^F \rangle, \quad \forall q^F \in E. \quad (1.5)$$

- For any $\varepsilon > 0$, the functional J_ε has a unique minimiser denoted by $q_\varepsilon^F \in E$. Here also, this is due to standard arguments of convex optimisation. Note that the coercivity of the functional in E comes from the term $\frac{\varepsilon}{2} \|q^F\|^2$ which corresponds, in the dual framework, to the penalisation term introduced in F_ε . Observe also that $\inf_E J_\varepsilon \leq J_\varepsilon(0) = 0$.
- This minimiser is characterised by the Euler-Lagrange equation

$$\int_0^T \left[\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F, \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} \tilde{q}^F \right] dt + \varepsilon \langle q_\varepsilon^F, \tilde{q}^F \rangle + \langle y_0, e^{-T\mathcal{A}^*} \tilde{q}^F \rangle = 0, \quad \forall \tilde{q}^F \in E. \quad (1.6)$$

The duality properties between these two functional are consequence of the general results mentioned above. It is however instructive to write an explicit argument here.

Proposition 1.5. *For any $\varepsilon > 0$, the minimisers v_ε and q_ε^F of the functionals F_ε and J_ε respectively, are related through the formulas*

$$v_\varepsilon(t) = \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F, \quad \text{for a.e. } t \in]0, T[,$$

and

$$\mathcal{L}_T(v_\varepsilon|y_0) = y_{v_\varepsilon, y_0}(T) = -\varepsilon q_\varepsilon^F. \quad (1.7)$$

As a consequence, we have

$$\inf_{L^2(0, T; U)} F_\varepsilon = F_\varepsilon(v_\varepsilon) = -J_\varepsilon(q_\varepsilon^F) = -\inf_E J_\varepsilon.$$

Finally, we have

$$\|\mathcal{L}_T(v_\varepsilon|y_0)\| \leq \|\mathcal{L}_T(0|y_0)\| = \|e^{-T\mathcal{A}}y_0\|. \quad (1.8)$$

Proof. We set $w_\varepsilon(t) = \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F$ and we want to show that $w_\varepsilon = v_\varepsilon$. Notice that $w_\varepsilon \in L^2(0, T; U)$, thanks to (1.1).

- We start from the Euler-Lagrange equation for J_ε that reads, for any $\tilde{q}^F \in E$,

$$\int_0^T \left[w_\varepsilon(t), \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} \tilde{q}^F \right] dt + \varepsilon \langle q^F, \tilde{q}^F \rangle + \langle y_0, e^{-T\mathcal{A}^*} \tilde{q}^F \rangle = 0.$$

By definition of y_{w_ε, y_0} , see (1.2), we have

$$\langle y_{w_\varepsilon, y_0}(T), \tilde{q}^F \rangle - \langle y_0, e^{-T\mathcal{A}^*} \tilde{q}^F \rangle = \int_0^T \left[w_\varepsilon(t), \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} \tilde{q}^F \right] dt.$$

Comparing the last two formulas, we obtain

$$\langle y_{w_\varepsilon, y_0}(T) + \varepsilon q_\varepsilon^F, \tilde{q}^F \rangle = 0. \quad (1.9)$$

As this equality holds for any $\tilde{q}^F \in E$, we find

$$\mathcal{L}_T(w_\varepsilon|y_0) = y_{w_\varepsilon, y_0}(T) = -\varepsilon q_\varepsilon^F. \quad (1.10)$$

- We consider now any $\tilde{v} \in L^2(0, T; U)$ and write the equation (1.2) satisfied by $y_{\tilde{v}, 0}$ using $\psi = q_\varepsilon^F$ as a test function. Since $y_{\tilde{v}, 0}(0) = 0$, we get

$$\langle y_{\tilde{v}, 0}(T), q_\varepsilon^F \rangle = \int_0^T \underbrace{\left[\tilde{v}, \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F \right]}_{[\tilde{v}, w_\varepsilon]} dt$$

Using (1.10), we have

$$\int_0^T [\tilde{v}, w_\varepsilon] dt + \frac{1}{\varepsilon} \langle y_{\tilde{v}, 0}(T), y_{w_\varepsilon, y_0}(T) \rangle = 0.$$

This formula holds for any $\tilde{v} \in L^2(0, T; U)$; we have thus shown that w_ε solves the Euler-Lagrange equation (1.4) implying that we have $w_\varepsilon = v_\varepsilon$.

- We can now compute, using (1.9),

$$\begin{aligned} F_\varepsilon(v_\varepsilon) &= F_\varepsilon(\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q_\varepsilon^F) \\ &= \frac{1}{2} \int_0^T \left[\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F \right]^2 dt + \frac{1}{2\varepsilon} \|\mathcal{L}_T(v_\varepsilon|y_0)\|^2 \\ &= \frac{1}{2} \int_0^T \left[\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F \right]^2 dt + \frac{\varepsilon}{2} \|q_\varepsilon^F\|^2. \end{aligned} \quad (1.11)$$

Moreover, taking $\tilde{q}^F = q_\varepsilon^F$ in (1.6) gives

$$\int_0^T \left[\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F \right]^2 dt + \varepsilon \|q_\varepsilon^F\|^2 = - \left\langle y_0, e^{-T\mathcal{A}^*} q_\varepsilon^F \right\rangle, \quad (1.12)$$

so that we have

$$J_\varepsilon(q_\varepsilon^F) = -\frac{1}{2} \int_0^T \left[\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q_\varepsilon^F \right]^2 dt - \frac{\varepsilon}{2} \|q_\varepsilon^F\|^2.$$

Comparing the values of $F_\varepsilon(v_\varepsilon)$ and $J_\varepsilon(q_\varepsilon^F)$ precisely gives the relation $F_\varepsilon(v_\varepsilon) = -J_\varepsilon(q_\varepsilon^F)$.

- By definition of v_ε , we have $F_\varepsilon(v_\varepsilon) \leq F_\varepsilon(0) = \frac{1}{2\varepsilon} \|\mathcal{L}_T(0|y_0)\|^2$. We deduce, in particular, that $\|\mathcal{L}_T(v_\varepsilon|y_0)\| \leq \|\mathcal{L}_T(0|y_0)\|$.

□

From (1.6), (1.7) and (1.8) we deduce the following result.

Corollary 1.6. *For any $\varepsilon > 0$, the map $y_0 \in E \mapsto q_\varepsilon^F \in E$ is linear, continuous and we have*

$$\|q_\varepsilon^F\| \leq \frac{1}{\varepsilon} \|e^{-T\mathcal{A}} y_0\|.$$

As a consequence of the previous proposition, we observe that the penalised HUM approach always leads to a final state whose norm is lower than that of the free solution of the system (the one associated with the control $v = 0$). Note that this holds without any additional assumption on the system than the general ones given in the introduction.

We can now express the approximate- and null-controllability properties of the system, for a given initial data y_0 , in terms of the behavior of the penalised HUM approach described above.

Theorem 1.7. (1) *Problem (P) is approximately controllable from the initial data y_0 if and only if we have*

$$\mathcal{L}_T(v_\varepsilon|y_0) = y_{v_\varepsilon, y_0}(T) \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (1.13)$$

(2) *Problem (P) is null-controllable from the initial data y_0 if and only if we have*

$$M_{y_0}^2 \stackrel{\text{def}}{=} 2 \sup_{\varepsilon > 0} \left(\inf_{L^2(0, T; U)} F_\varepsilon \right) < +\infty. \quad (1.14)$$

In this case, we have

$$\|v_\varepsilon\|_{L^2(0, T; U)} \leq M_{y_0}, \quad (1.15)$$

$$\|\mathcal{L}_T(v_\varepsilon|y_0)\| \leq M_{y_0} \sqrt{\varepsilon}. \quad (1.16)$$

Moreover we have $\|v^0\|_{L^2(0, T; U)}^2 = M_{y_0}^2$ and

$$v_\varepsilon \xrightarrow{\varepsilon \rightarrow 0} v^0, \quad \text{strongly in } L^2(0, T; U), \quad (1.17)$$

where v^0 is the HUM control defined in Section 1.2.

Remark 1.8. Observe that for $\varepsilon_1 > \varepsilon_2 > 0$, and $v \in L^2(0, T; U)$ we have $F_{\varepsilon_2}(v) \geq F_{\varepsilon_1}(v)$. It follows that the supremum with respect to $\varepsilon > 0$ in (1.14) is actually equal to the limit when $\varepsilon \rightarrow 0$ of $\inf_{L^2(0, T; U)} F_\varepsilon$.

Proof. (1) It is clear that (1.13) implies the approximate controllability since it proves that for any $\delta > 0$, there exists $\varepsilon > 0$ such that $v_\varepsilon \in \text{Adm}(y_0, \delta)$.

Assume now that approximate controllability holds and that (1.13) does not hold. It then exists $\alpha > 0$ and a sequence $(\varepsilon_k)_k$ such that $\varepsilon_k \rightarrow 0$ as $k \rightarrow \infty$ and $\|\mathcal{L}_T(v_{\varepsilon_k}|y_0)\|^2 \geq \alpha$. By assumption, $\text{Adm}(y_0, \sqrt{\alpha/2})$ is not empty; let us choose

$$\hat{v} \in \text{Adm}(y_0, \sqrt{\alpha/2}). \quad (1.18)$$

By construction, we have

$$\begin{aligned} \frac{\alpha}{2\varepsilon_k} &\leq \frac{1}{2\varepsilon_k} \|\mathcal{L}_T(v_{\varepsilon_k}|y_0)\|^2 \\ &\leq F_{\varepsilon_k}(v_{\varepsilon_k}) \leq F_{\varepsilon_k}(\hat{v}) \\ &= \frac{1}{2} \|\hat{v}\|_{L^2(0,T;U)}^2 + \frac{1}{2\varepsilon_k} \underbrace{\|\mathcal{L}_T(\hat{v}|y_0)\|^2}_{\leq \alpha/2, \text{ by (1.18)}}, \end{aligned}$$

so that

$$\frac{\alpha}{4\varepsilon_k} \leq \frac{1}{2} \|\hat{v}\|_{L^2(0,T;U)}^2,$$

and we get a contradiction when $k \rightarrow \infty$.

- (2) If the problem is null-controllable, then $\text{Adm}(y_0, 0)$ is a non-empty closed and convex set and we have already seen that the HUM control v^0 is nothing but the orthogonal projection of 0 onto $\text{Adm}(y_0, 0)$ in $L^2(0, T; U)$.

By construction, we have $F_\varepsilon(v^0) = \frac{1}{2} \|\mathbb{[}v^0\mathbb{]}\|_{L^2(0,T;U)}^2$. This does not depend on ε so that we get

$$\sup_{\varepsilon > 0} \left(\inf_{L^2(0,T;U)} F_\varepsilon \right) \leq \frac{1}{2} \|\mathbb{[}v^0\mathbb{]}\|_{L^2(0,T;U)}^2.$$

This proves a first implication as well as the inequality

$$M_{y_0}^2 \leq \|\mathbb{[}v^0\mathbb{]}\|_{L^2(0,T;U)}^2. \quad (1.19)$$

Assume now that (1.14) holds. We immediately deduce (1.15) and (1.16). We can then find a subsequence $(v_{\varepsilon_k})_k$ which weakly converges towards some $v \in L^2(0, T; U)$ and such that $\mathcal{L}_T(v_{\varepsilon_k}|y_0) \xrightarrow[k \rightarrow \infty]{} 0$. The parabolic problem we consider being linear (and so is the operator $\mathcal{L}_T(\cdot|\cdot)$), we can easily pass to the limit and find that $\mathcal{L}_T(v_{\varepsilon_k}|y_0)$ weakly converges towards $\mathcal{L}_T(v|y_0)$ and in particular that $\mathcal{L}_T(v|y_0) = 0$. As a consequence, the weak limit we obtained is a null-control for our problem : $v \in \text{Adm}(y_0, 0)$.

By definition of the minimal-norm control we have $\|\mathbb{[}v^0\mathbb{]}\|_{L^2(0,T;U)} \leq \|v\|_{L^2(0,T;U)}$, so that with (1.19), we deduce $M_{y_0}^2 \leq \|\mathbb{[}v\mathbb{]}\|_{L^2(0,T;U)}^2$.

Moreover, taking the limit in (1.15) we have

$$\limsup_{k \rightarrow \infty} \|\mathbb{[}v_{\varepsilon_k}\mathbb{]}\|_{L^2(0,T;U)}^2 \leq M_{y_0}^2 \leq \|\mathbb{[}v\mathbb{]}\|_{L^2(0,T;U)}^2.$$

This proves that the convergence of v_{ε_k} towards v is actually strong and that $\|\mathbb{[}v\mathbb{]}\|_{L^2(0,T;U)}^2 = M_{y_0}^2$. By (1.19), and since v^0 is the unique minimal-norm control, we deduce that $v = v^0$, that $\|\mathbb{[}v^0\mathbb{]}\|_{L^2(0,T;U)}^2 = M_{y_0}^2$, and finally that the complete family $(v_\varepsilon)_\varepsilon$ strongly converges to v^0 . \square

Remark 1.9. We can slightly improve (1.16) by observing that

$$M_{y_0}^2/2 \geq F_\varepsilon(v_\varepsilon) \geq \frac{1}{2} \llbracket v_\varepsilon \rrbracket_{L^2(0,T;U)}^2 \xrightarrow{\varepsilon \rightarrow 0} M_{y_0}^2/2,$$

which implies that

$$\sqrt{\varepsilon} \|q_\varepsilon^F\| = \frac{1}{\sqrt{\varepsilon}} \|\mathcal{L}_T(v_\varepsilon|y_0)\| \xrightarrow{\varepsilon \rightarrow 0} 0.$$

1.4. General behavior of the penalised HUM

We define the space

$$Q_F = \{q^F \in E, \text{ s.t. } \mathcal{B}^* e^{-t\mathcal{A}^*} q^F = 0, \forall t \geq 0\}, \quad (1.20)$$

made of the non observable adjoint states and the space

$$Y_T = e^{-T\mathcal{A}^*} Q_F. \quad (1.21)$$

Notice that Q_F is closed and that

$$Q_F \text{ and } Y_T \text{ are stable through } e^{-s\mathcal{A}^*} \text{ for any } s \geq 0. \quad (1.22)$$

Remark 1.10. The semi-group generated by $-\mathcal{A}^*$ being analytic, we observe that we also have for any $T > 0$,

$$Q_F = \{q^F \in E, \text{ s.t. } \mathcal{B}^* e^{-t\mathcal{A}^*} q^F = 0, \forall t \in [0, T]\}.$$

The following result, which also holds without any additional assumption on the system, gives a general convergence result for the value of the controlled solution at the final time T given by the penalised HUM method.

Theorem 1.11. *For any $y_0 \in E$, the penalised-HUM sequence $(v_\varepsilon)_\varepsilon$ satisfies*

$$\mathcal{L}_T(v_\varepsilon|y_0) \xrightarrow{\varepsilon \rightarrow 0} \mathbb{P}_{Q_F}(e^{-T\mathcal{A}}y_0),$$

where \mathbb{P}_{Q_F} is the orthogonal projection onto Q_F in E .

Remark 1.12. In the finite dimensional case, it is easily seen that Q_F is the kernel of the adjoint of the Kalman matrix

$$K = [\mathcal{B}|\mathcal{A}\mathcal{B}|\dots|\mathcal{A}^{n-1}\mathcal{B}], \quad n = \dim(E).$$

Therefore, $Q_F = \{0\}$ if and only if K is full rank, and we recover the Kalman rank condition for the controllability of finite dimensional linear systems.

Proof of Theorem 1.11. Using assumption (1.1), we can define a bounded non-negative symmetric operator $\Lambda : E \rightarrow E$, usually referred to as the Gramian operator, as follows

$$\langle \Lambda\psi, \phi \rangle = \int_0^T [\mathcal{B}^* e^{-t\mathcal{A}^*} \psi, \mathcal{B}^* e^{-t\mathcal{A}^*} \phi] dt, \quad \forall \phi, \psi \in E.$$

Observe, using Remark 1.10, that $Q_F = \text{Ker } \Lambda$. With these notations, the Euler-Lagrange equation (1.6) can be written as

$$(\Lambda + \varepsilon \text{Id})q_\varepsilon^F = -e^{-T\mathcal{A}}y_0. \quad (1.23)$$

We set $z = -e^{-T\mathcal{A}}y_0$ and $r_\varepsilon = \varepsilon q_\varepsilon^F - \mathbb{P}_{Q_F}z$. From (1.7), the claim is equivalent to $\|r_\varepsilon\| \xrightarrow{\varepsilon \rightarrow 0} 0$.

To this end, we write the equation satisfied by r_ε

$$(\Lambda + \varepsilon \text{Id})r_\varepsilon = \varepsilon(z - \mathbb{P}_{Q_F}z).$$

Since $z - \mathbb{P}_{Q_F}z$ is orthogonal to $Q_F = \text{Ker } \Lambda$, we know that it belongs to $\overline{\text{Ran } \Lambda}$. Thus, there exists a sequence $(\psi_n)_n \subset E$ such that

$$\Lambda\psi_n \xrightarrow{n \rightarrow \infty} z - \mathbb{P}_{Q_F}z.$$

We have

$$(\Lambda + \varepsilon \text{Id})(r_\varepsilon - \varepsilon\psi_n) = \varepsilon(z - \mathbb{P}_{Q_F}z - \Lambda\psi_n - \varepsilon\psi_n).$$

Testing this equation against $r_\varepsilon - \varepsilon\psi_n$ we obtain

$$\langle (\Lambda + \varepsilon \text{Id})(r_\varepsilon - \varepsilon\psi_n), r_\varepsilon - \varepsilon\psi_n \rangle = \varepsilon \langle z - \mathbb{P}_{Q_F}z - \Lambda\psi_n - \varepsilon\psi_n, r_\varepsilon - \varepsilon\psi_n \rangle. \quad (1.24)$$

Using the non-negativity of Λ and the Cauchy-Schwarz inequality, we find

$$\varepsilon \|r_\varepsilon - \varepsilon\psi_n\|^2 \leq \varepsilon \|z - \mathbb{P}_{Q_F}z - \Lambda\psi_n - \varepsilon\psi_n\| \|r_\varepsilon - \varepsilon\psi_n\|,$$

so that

$$\|r_\varepsilon\| \leq 2\varepsilon \|\psi_n\| + \|z - \mathbb{P}_{Q_F}z - \Lambda\psi_n\|.$$

Taking the superior limit with respect to ε leads to

$$\limsup_{\varepsilon \rightarrow 0} \|r_\varepsilon\| \leq \|z - \mathbb{P}_{Q_F}z - \Lambda\psi_n\|.$$

Since $\Lambda\psi_n$ converges to $z - \mathbb{P}_{Q_F}z$, we finally get that $\|r_\varepsilon\| \xrightarrow{\varepsilon \rightarrow 0} 0$ and the claim is proved. \square

Remark 1.13. Observe that the previous theorem is a particular case of a general result saying that for any symmetric non-negative continuous operator in a Hilbert space, the family of operators $(\frac{1}{\varepsilon}\Lambda + \text{Id})^{-1}$ converges strongly towards the orthogonal projection onto the kernel of Λ as ε goes to 0.

In the case where \mathcal{A} is a symmetric operator we can obtain additional results.

Proposition 1.14. *Assume that \mathcal{A} is symmetric then we have*

$$e^{-s\mathcal{A}}\overline{Y_T} \subset Q_F, \quad \forall s \geq 0, \quad (1.25)$$

and

$$\mathbb{P}_{Q_F}(e^{-T\mathcal{A}}y_0) = e^{-T\mathcal{A}}\left(\mathbb{P}_{\overline{Y_T}}y_0\right), \quad \forall y_0 \in E. \quad (1.26)$$

Therefore, for any initial data $y_0 \in \overline{Y_T}$, the following properties hold

- (1) For any $\varepsilon > 0$, we have $v_\varepsilon = 0$.
- (2) The set $\text{Adm}(y_0, \delta)$ is not empty if and only if $\delta \geq \|\mathcal{L}_T(0|y_0)\|$ and in that case we have $v^\delta = 0$.

This means that, for such initial data, the value of the free solution y_{0,y_0} at time T is the closest to zero that one can achieve with any other control.

Proof. • Since $\mathcal{A}^* = \mathcal{A}$ and using (1.22) we conclude that, for any $s \geq 0$, we have

$$e^{-s\mathcal{A}}Y_T = e^{-(s+T)\mathcal{A}}Q_F \subset Q_F.$$

Property (1.25) immediately follows since Q_F is closed.

- Let us now prove (1.26). By definition of $Y_T = e^{-T\mathcal{A}}Q_F$ we get

$$e^{-T\mathcal{A}}\mathbb{P}_{Q_F} = \mathbb{P}_{\overline{Y_T}}e^{-T\mathcal{A}}\mathbb{P}_{Q_F}.$$

Taking the adjoint, we obtain

$$\mathbb{P}_{Q_F}e^{-T\mathcal{A}} = \mathbb{P}_{Q_F}e^{-T\mathcal{A}}\mathbb{P}_{\overline{Y_T}},$$

but the property (1.25) gives that

$$\mathbb{P}_{Q_F}e^{-T\mathcal{A}}\mathbb{P}_{\overline{Y_T}} = e^{-T\mathcal{A}}\mathbb{P}_{\overline{Y_T}},$$

and the claim is proved.

- Let $y_0 \in \overline{Y_T}$.
 - (1) Since $e^{-T\mathcal{A}}y_0 \in e^{-T\mathcal{A}}Y_T \subset Q_F = \text{Ker } \Lambda$, it is clear that the unique solution to (1.23) is given by

$$q_\varepsilon^F = -\frac{1}{\varepsilon}e^{-T\mathcal{A}}y_0 = -\frac{1}{\varepsilon}\mathcal{L}_T(0|y_0).$$

The control is then given by $v_\varepsilon(t) = \mathcal{B}^*e^{-(T-t)\mathcal{A}}q_\varepsilon^F$, that is $v_\varepsilon = 0$ by definition of Q_F .

- (2) Assume that there exists a $\tilde{v} \in L^2(0, T; U)$ such that $\|\mathcal{L}_T(\tilde{v}|y_0)\| < \|\mathcal{L}_T(0|y_0)\|$ then, for any $\varepsilon > 0$, we have

$$\frac{1}{2}\|\tilde{v}\|_{L^2(0, T; U)}^2 + \frac{1}{2\varepsilon}\|\mathcal{L}_T(\tilde{v}|y_0)\|^2 = F_\varepsilon(\tilde{v}) \geq \inf_{L^2(0, T; U)} F_\varepsilon = F_\varepsilon(0) = \frac{1}{2\varepsilon}\|\mathcal{L}_T(0|y_0)\|^2,$$

and it follows that

$$\|\tilde{v}\|_{L^2(0, T; U)}^2 \geq \frac{1}{\varepsilon}\left(\|\mathcal{L}_T(0|y_0)\|^2 - \|\mathcal{L}_T(\tilde{v}|y_0)\|^2\right) \xrightarrow{\varepsilon \rightarrow 0} +\infty,$$

which is a contradiction. □

Remark 1.15. The previous proposition shows, in particular, that proving a uniform bound on the approximate HUM controls $(v_\varepsilon)_\varepsilon$ is not sufficient in general to ensure that the problem is approximately controllable from y_0 (see for instance figure 13).

Remark 1.16. In the case where the operator Λ has a closed range (for instance in the finite dimensional setting), it can be proved that the sequence $(v_\varepsilon)_\varepsilon$ is bounded, that $\|r_\varepsilon\| \leq C\varepsilon$ and that, for any initial data, approximate controllability is equivalent to null-controllability.

Indeed with this additional assumption, and coming back to the proof of Theorem 1.11, we observe that there is a $\psi \in E$ such that $\Lambda\psi = z - \mathbb{P}_{Q_F}z$ so that, from (1.24), we have

$$\langle (\Lambda + \varepsilon\text{Id})(r_\varepsilon - \varepsilon\psi), (r_\varepsilon - \varepsilon\psi) \rangle = -\varepsilon^2 \langle \psi, r_\varepsilon - \varepsilon\psi \rangle,$$

yielding $\varepsilon\|r_\varepsilon - \varepsilon\psi\|^2 \leq \varepsilon^2\|\psi\|\|r_\varepsilon - \varepsilon\psi\|$, and finally $\|r_\varepsilon - \varepsilon\psi\| \leq \varepsilon\|\psi\|$, so that

$$\|r_\varepsilon\| \leq 2\varepsilon\|\psi\|.$$

The optimal control is then given by

$$v_\varepsilon(t) = \mathcal{B}^*e^{-(T-t)\mathcal{A}^*}q_\varepsilon^F = \frac{1}{\varepsilon}\mathcal{B}^*e^{-(T-t)\mathcal{A}^*}(r_\varepsilon + \mathbb{P}_{Q_F}z) = \frac{1}{\varepsilon}\mathcal{B}^*e^{-(T-t)\mathcal{A}^*}r_\varepsilon, \quad 0 \leq t \leq T,$$

recalling the definition of r_ε and that of Q_F . We conclude that

$$\|v_\varepsilon\|_{L^2(0,T;U)} \leq \frac{C}{\varepsilon} \|r_\varepsilon\| \leq 2C \|\psi\| = C_1,$$

which is independent of ε .

Assume now that Problem (\mathcal{P}) is approximately controllable from y_0 . From Theorems 1.7 and 1.11, this means that $\mathbb{P}_{Q_F} z = 0$ so that

$$\|\mathcal{L}_T(v_\varepsilon|y_0)\| = \|\varepsilon q_\varepsilon^F\| = \|r_\varepsilon\| \leq C_2\varepsilon,$$

and finally

$$\inf_{L^2(0,T;U)} F_\varepsilon = F_\varepsilon(v_\varepsilon) = \frac{1}{2} \|v_\varepsilon\|_{L^2(0,T;U)}^2 + \frac{1}{2\varepsilon} \|\mathcal{L}_T(v_\varepsilon|y_0)\|^2 \leq \frac{1}{2} C_1^2 + \frac{\varepsilon}{2} C_2^2,$$

so that

$$\sup_{\varepsilon > 0} \left(\inf_{L^2(0,T;U)} F_\varepsilon \right) < +\infty.$$

We conclude from Theorem 1.7 that the problem is indeed null-controllable from y_0 .

1.5. Approximate controllability and unique continuation

It is well known that the approximate controllability property of a linear system is related to the unique continuation property for the adjoint problem. Let us give a precise statement of this equivalence for a given fixed initial data y_0 as well as a proof based on the arguments developed in the previous section.

Proposition 1.17. *Let $y_0 \in E$ be given. The following properties are equivalent.*

- (1) *Problem (\mathcal{P}) is approximately controllable from y_0 .*
- (2) *The following weak unique continuation property holds*

$$\left[\mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q^F = 0, \quad \forall t \in [0, T] \right] \implies \left\langle y_0, e^{-T\mathcal{A}^*} q^F \right\rangle = 0.$$

- (3) *For any $\varepsilon > 0$ there exists a $C_{\varepsilon, y_0} > 0$ such that the following weak observability estimate holds*

$$\left| \left\langle y_0, e^{-T\mathcal{A}^*} q^F \right\rangle \right|^2 \leq C_{\varepsilon, y_0}^2 \left\| \left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^F \right] \right\|_{L^2(0,T;U)}^2 + \varepsilon \|q^F\|^2, \quad \forall q^F \in E.$$

Proof.

- (1) \Leftrightarrow (2):

Following Remark 1.10, the property (2) can be rewritten as $y_0 \in Y_T^\perp$ or also $e^{-T\mathcal{A}} y_0 \in Q_{\bar{F}}^\perp$, that is finally

$$(2) \iff \mathbb{P}_{Q_F} (e^{-T\mathcal{A}} y_0) = 0.$$

Therefore, the equivalence between (1) and (2) is just a consequence of Theorem 1.7 and Theorem 1.11.

- (3) \Rightarrow (2):

This is obvious by letting $\varepsilon \rightarrow 0$.

- (2) \Rightarrow (3):

We assume that, for some $\varepsilon > 0$, such a C_{ε, y_0} does not exist. Therefore, there exists a sequence $(q^{F,n})_n$ such as

$$1 = \left| \left\langle y_0, e^{-T\mathcal{A}^*} q^{F,n} \right\rangle \right| \geq n \left\| \left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^{F,n} \right] \right\|_{L^2(0,T;U)} + \sqrt{\varepsilon} \|q^{F,n}\|.$$

We deduce that the sequence $(q^{F,n})_n$ is bounded (here ε is fixed) and therefore there exists a subsequence (still referred to as $q^{F,n}$) that weakly converges towards some q^F in E . In particular $\langle y_0, e^{-T\mathcal{A}^*} q^{F,n} \rangle = 1$ converges towards $\langle y_0, e^{-T\mathcal{A}^*} q^F \rangle = 1$ and thus $q^F \neq 0$.

Moreover, we have by construction

$$\left[\left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^{F,n} \right] \right]_{L^2(0,T;U)} \xrightarrow{n \rightarrow \infty} 0,$$

so that $\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^F = 0$. This is a contradiction. \square

In particular, we recover here the usual statement saying that if $Q_F = \{0\}$ (that is the adjoint system satisfies the unique continuation property for any final data) then Problem (\mathcal{P}) is approximately controllable for any initial data.

1.6. Null-controllability and observability

We can now prove an observability-type inequality to characterise the null-controllability property of the system starting from a given initial data y_0 . We will see in the sequel that it is of some importance to be able to specify the influence of the initial data on the controllability of the system (see Sections 1.7 and 2). As a consequence of the following result we will of course recover the more usual statement of the equivalence between observability and controllability in Proposition 1.19.

Note that the following proof is based on the analysis of the penalised HUM approach and not on general functional analysis results.

Proposition 1.18. *Problem (\mathcal{P}) is null-controllable from the initial data y_0 if and only if, there exists $\widetilde{M}_{y_0} \geq 0$ such that*

$$\left| \left\langle y_0, e^{-T\mathcal{A}^*} q^F \right\rangle \right|^2 \leq \widetilde{M}_{y_0}^2 \left[\left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^F \right] \right]_{L^2(0,T;U)}^2, \quad \forall q^F \in E. \quad (1.27)$$

Moreover, in such case, the smallest value of \widetilde{M}_{y_0} satisfying (1.27) is equal to M_{y_0} as defined in (1.14).

Proof. • Assume that (1.27) holds. We get from (1.12)

$$\left[\left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q_\varepsilon^F \right] \right]_{L^2(0,T;U)}^2 + \varepsilon \|q_\varepsilon^F\|^2 \leq \widetilde{M}_{y_0} \left[\left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q_\varepsilon^F \right] \right]_{L^2(0,T;U)}.$$

It follows that we have $\varepsilon \|q_\varepsilon^F\|^2 \leq \widetilde{M}_{y_0}^2$ and

$$\|v_\varepsilon\|_{L^2(0,T;U)} = \left[\left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q_\varepsilon^F \right] \right]_{L^2(0,T;U)} \leq \widetilde{M}_{y_0}. \quad (1.28)$$

Therefore, by (1.11), $\sup_{\varepsilon > 0} \inf_{L^2(0,T;U)} F_\varepsilon$ is finite and the null-controllability follows by Theorem 1.7.

Moreover, with (1.28) we have

$$F_\varepsilon(v_\varepsilon) \leq \frac{1}{2} \widetilde{M}_{y_0}^2 + \frac{1}{2\varepsilon} \|\mathcal{L}_T(v_\varepsilon|y_0)\|^2 \xrightarrow{\varepsilon \rightarrow 0} \frac{1}{2} \widetilde{M}_{y_0}^2,$$

by using Remark 1.9. It follows that $M_{y_0} \leq \widetilde{M}_{y_0}$.

• Assume now that the null-controllability property holds. From (1.6) we get

$$\left\langle y_0, e^{-T\mathcal{A}^*} \tilde{q}^F \right\rangle = \left\langle \mathcal{L}_T(v_\varepsilon|y_0), \tilde{q}^F \right\rangle - \int_0^T \left[v_\varepsilon, \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} \tilde{q}^F \right] dt, \quad \forall \tilde{q}^F \in E.$$

By assumption and Theorem 1.7 we have $\mathcal{L}_T(v_\varepsilon|y_0) \xrightarrow{\varepsilon \rightarrow 0} 0$, therefore the first term in the right-hand side goes to 0 and we can use (1.17) to pass to the limit and get

$$\left\langle y_0, e^{-T\mathcal{A}^*} \tilde{q}^F \right\rangle = - \int_0^T \left[v^0, \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} \tilde{q}^F \right] dt.$$

By the Cauchy-Schwarz inequality, we finally get

$$\left| \left\langle y_0, e^{-T\mathcal{A}^*} \tilde{q}^F \right\rangle \right| \leq \llbracket v^0 \rrbracket_{L^2(0,T;U)} \left\| \mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} \tilde{q}^F \right\|_{L^2(0,T;U)},$$

so that (1.27) holds with $\widetilde{M}_{y_0} = \llbracket v^0 \rrbracket_{L^2(0,T;U)}$. By Theorem 1.7, we have $M_{y_0} = \llbracket v^0 \rrbracket_{L^2(0,T;U)}$. The proof is complete. \square

Proposition 1.19. *Problem (P) is null-controllable for any initial data if and only if there exists $C_{\text{obs}} > 0$ such that*

$$\left\| e^{-T\mathcal{A}^*} q^F \right\|^2 \leq C_{\text{obs}}^2 \left\| \mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^F \right\|_{L^2(0,T;U)}^2, \quad \forall q^F \in E. \quad (1.29)$$

Moreover, if C_{obs} is the smallest constant for which (1.29) holds, then it is also the smallest constant satisfying

$$M_{y_0} \leq C_{\text{obs}} \|y_0\|, \quad \forall y_0 \in E, \quad (1.30)$$

where M_{y_0} is the minimal cost of the control as defined in Theorem 1.7.

Proof. • It is clear that (1.29) implies (1.27) for any $y_0 \in E$, with $\widetilde{M}_{y_0} = C_{\text{obs}} \|y_0\|$, so that the null-controllability for any initial data follows by Proposition 1.18 as well as inequality (1.30).

- Assume now that the problem is null-controllable for any initial data. Then the maps $\Phi_\varepsilon : y_0 \mapsto v_\varepsilon$ are linear continuous and for any $y_0 \in E$, the family $(\Phi_\varepsilon y_0)_\varepsilon$ converges when ε goes to 0.

The Banach-Steinhaus theorem shows that there exists $K > 0$ such that

$$\llbracket \Phi_\varepsilon y_0 \rrbracket_{L^2(0,T;U)} \leq K \|y_0\|, \quad \forall y_0 \in E, \quad \forall \varepsilon > 0.$$

In particular, the HUM-null-control $v^0 = \lim_{\varepsilon \rightarrow 0} v_\varepsilon = \lim_{\varepsilon \rightarrow 0} \Phi_\varepsilon y_0$ associated with a given initial data satisfies

$$M_{y_0} = \llbracket v^0 \rrbracket_{L^2(0,T;U)} \leq K \|y_0\|.$$

From (1.27) and the fact that M_{y_0} is the optimal cost of the control, we get

$$\left| \left\langle y_0, e^{-T\mathcal{A}^*} q^F \right\rangle \right| \leq K \|y_0\| \left\| \mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^F \right\|_{L^2(0,T;U)}.$$

Since this is true for any y_0 , we deduce that (1.29) holds with $C_{\text{obs}} = K$. \square

1.7. The case of approximate initial data

In order to prepare the discussion of Section 2 concerning semi-discrete approximations of Problem (P), we introduce the following slightly modified problem.

For each value of $\varepsilon > 0$ we consider an initial data $y_{0,\varepsilon} \in E$ that we use in the definition of the functionals F_ε and J_ε instead of y_0 , the corresponding functionals are denoted by \tilde{F}_ε and \tilde{J}_ε

$$\tilde{F}_\varepsilon(v) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^T \llbracket v(t) \rrbracket^2 dt + \frac{1}{2\varepsilon} \left\| \mathcal{L}_T(v|y_{0,\varepsilon}) \right\|^2, \quad \forall v \in L^2(0,T;U),$$

$$\tilde{J}_\varepsilon(q^F) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^T \left\| \mathcal{B}^* e^{-(T-t)\mathcal{A}^*} q^F \right\|^2 dt + \frac{\varepsilon}{2} \|q^F\|^2 + \langle y_{0,\varepsilon}, e^{-T\mathcal{A}^*} q^F \rangle, \quad \forall q^F \in E.$$

We assume that $(y_{0,\varepsilon})_\varepsilon$ is a bounded sequence that satisfies, for some $y_0 \in E$,

$$e^{-T\mathcal{A}} y_{0,\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} e^{-T\mathcal{A}} y_0, \quad \text{that is } \mathcal{L}_T(0|y_{0,\varepsilon} - y_0) \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (1.31)$$

Notice that, this property implies the weak convergence of $y_{0,\varepsilon}$ towards y_0 as $\varepsilon \rightarrow 0$. We denote by \tilde{v}_ε the unique minimiser of \tilde{F}_ε .

The problem we want to deal with is that, even if Problem (\mathcal{P}) is approximately- or null-controllable from y_0 , it may not be approximately- or null-controllable from the approximate initial data $y_{0,\varepsilon}$. Therefore, the behavior of the corresponding penalised HUM approach as ε goes to 0 has to be made precise. We can obtain the following facts by slightly adapting the previous proofs.

- Problem (\mathcal{P}) is approximately controllable from y_0 if and only if

$$\mathcal{L}_T(\tilde{v}_\varepsilon|y_{0,\varepsilon}) \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (1.32)$$

- Assume that $\sup_{\varepsilon > 0} \inf_{L^2(0,T;U)} \tilde{F}_\varepsilon$ is finite, then Problem (\mathcal{P}) is null-controllable from y_0 .
- Conversely, assume that Problem (\mathcal{P}) is null-controllable from y_0 and that

$$\sup_{\varepsilon > 0} \frac{1}{\varepsilon} \left\| \mathcal{L}_T(0|y_0 - y_{0,\varepsilon}) \right\|^2 < +\infty, \quad (1.33)$$

then we have

$$\sup_{\varepsilon > 0} \inf_{L^2(0,T;U)} \tilde{F}_\varepsilon < +\infty.$$

If moreover we assume that

$$\frac{1}{\varepsilon} \left\| \mathcal{L}_T(0|y_0 - y_{0,\varepsilon}) \right\|^2 \xrightarrow{\varepsilon \rightarrow 0} 0, \quad (1.34)$$

then $(\tilde{v}_\varepsilon)_\varepsilon$ strongly converges towards the HUM control v^0 associated with the initial data y_0 .

- Notice that conditions (1.33),(1.34) are very important for the result above to be true. Let us consider a simple example.

Assume that \mathcal{A} is symmetric and that the space Y_T defined in (1.21) is not trivial (one can think for instance of a finite dimensional system such that the Kalman rank condition is not satisfied). We choose any $z \in Y_T \setminus \{0\}$ and set $y_0 = 0$ and $y_{0,\varepsilon} = \varepsilon^\alpha z$ with $\alpha < 1/2$. It is clear that (\mathcal{P}) is null-controllable from $y_0 = 0$, that it is not approximately controllable from $y_{0,\varepsilon}$ for any $\varepsilon > 0$, and that (1.31) holds.

However, we can simply compute

$$\tilde{F}_\varepsilon(v) = \varepsilon^{2\alpha} \left(\frac{1}{2} \|v/\varepsilon^\alpha\|_{L^2(0,T;U)}^2 + \frac{1}{2\varepsilon} \left\| \mathcal{L}_T(v/\varepsilon^\alpha|z) \right\|^2 \right), \quad \forall v \in L^2(0,T;U),$$

so that, using Proposition 1.14, we know that the parenthesis is minimal for $v = 0$. We deduce that

$$\inf_{L^2(0,T;U)} \tilde{F}_\varepsilon = \frac{\varepsilon^{2\alpha-1}}{2} \left\| \mathcal{L}_T(0|z) \right\|^2 \xrightarrow{\varepsilon \rightarrow 0} +\infty, \quad \text{since } \alpha < 1/2.$$

We can summarize this discussion as follows:

- As far as the approximate control property is concerned, the penalised HUM approach associated with a family of approximate initial data still provides a satisfactory criterion through (1.32), without any additional assumption.

- In order to be able to characterise null-controllability from the single observation of the behavior of the penalised HUM in the limit $\varepsilon \rightarrow 0$, it is necessary to assume that condition (1.33) is fulfilled. If moreover condition (1.34) is satisfied then we deduce the strong convergence of the family of controls \tilde{v}_ε towards the unique HUM control v^0 associated with the initial data y_0 .

In other words, it is necessary that $(y_{0,\varepsilon})_\varepsilon$ converges towards y_0 sufficiently rapidly.

- Moreover, considering a sequence of approximate initial data satisfying (1.33), we can state a relaxed observability inequality which says that Problem (\mathcal{P}) is null-controllable from y_0 if and only if there exists a $M > 0$ such that, for any $\varepsilon > 0$, we have

$$\left| \left\langle y_{0,\varepsilon}, e^{-T\mathcal{A}^*} q^F \right\rangle \right|^2 \leq M \left(\left\| \left[\mathcal{B}^* e^{-(T-\cdot)\mathcal{A}^*} q^F \right] \right\|_{L^2(0,T;U)}^2 + \varepsilon \|q^F\|^2 \right), \quad \forall q^F \in E. \quad (1.35)$$

This can be proved by adapting the proof of Proposition 1.18. Observe that this inequality aims to generalize (1.27) but is different from that given in Proposition 1.17, since the constant M here does not depend on ε .

Relaxed observability inequalities of the form (1.35) will be central in the analysis of semi-discrete and fully-discrete problems in the next sections. Indeed, we shall see that the space approximation of a null-controllable problem may lead to a non-controllable semi-discrete system, which typically corresponds to the situation above where (\mathcal{P}) is null-controllable from y_0 but is not null-controllable from the approximate data $y_{0,\varepsilon}$.

2. SEMI-DISCRETE APPROXIMATIONS

2.1. General framework

We consider now a semi-discrete approximation of our problem (\mathcal{P}) , that we write for the moment in the following abstract form. For any $h > 0$ (which is supposed to represent the space discretisation parameter), we are given

- A Euclidean space E_h , whose inner product and its associated norm are denoted by $\langle \cdot, \cdot \rangle_h$ and $\|\cdot\|_h$ respectively.
- A linear operator \mathcal{A}_h on E_h .
- A second Euclidean space U_h , whose inner product and its associated norm are denoted by $[\cdot, \cdot]_h$ and $\|\cdot\|_h$.
- A linear operator $\mathcal{B}_h : U_h \rightarrow E_h$. We denote by \mathcal{B}_h^* its adjoint, that is,

$$\forall u \in U_h, \forall x \in E_h, \quad \langle \mathcal{B}_h u, x \rangle_h = [\mathcal{B}_h^* x, u]_h.$$

We are now interested in the following semi-discrete control problem

$$\begin{cases} \partial_t y_h + \mathcal{A}_h y_h = \mathcal{B}_h v_h, \\ y_h(0) = y_{0,h}. \end{cases} \quad (\mathcal{P}_h)$$

The solution of this problem is referred to as $t \mapsto y_{v_h, y_{0,h}}(t) \in E_h$ and the value of this solution at time T is denoted by

$$\mathcal{L}_T^h(v_h | y_{0,h}) = y_{v_h, y_{0,h}}(T).$$

For each value of h , and each initial data $y_{0,h} \in E_h$, one can ask if this finite-dimensional problem is approximately- or null-controllable (notice that for a finite dimensional system, the two notions coincide, see for instance Remark 1.16). Moreover, if such a null-control v_h exists, can we prove uniform bounds for v_h ?

2.2. The main issues related to discretisation

Two main issues arise when applying the penalised HUM strategy to the family of semi-discrete problems (\mathcal{P}_h) .

- (1) It appears that, even if the initial problem (\mathcal{P}) is null-controllable (let say for any initial data for the moment), then the semi-discrete problem for a given $h > 0$ may not be null-controllable. This actually holds even for a very classical finite-difference approximation of the heat equation for instance, as observed by O. Kaviani (see [Zua06]). More precisely, there may exist an eigenfunction ψ_h of \mathcal{A}_h^* (we denote by μ_h the associated eigenvalue) such that $\mathcal{B}_h^* \psi_h = 0$. In that case, we see that ψ_h belongs to the set

$$Q_{F,h} = \{q_h^F \in E_h, \text{ s.t. } \mathcal{B}_h^* e^{-t\mathcal{A}_h^*} q_h^F = 0, \forall t \geq 0\}, \quad (2.1)$$

and therefore, from Theorem 1.11, we know that there exists an initial data $y_{0,h}$ which is not null-controllable for the semi-discrete system. This can also be checked easily here by taking the scalar product of (\mathcal{P}_h) with ψ_h . Since $\mathcal{B}_h^* \psi_h = 0$ and $\mathcal{A}_h^* \psi_h = \mu_h \psi_h$ we obtain

$$\frac{d}{dt} \langle y_h(t), \psi_h \rangle_h + \mu_h \langle y_h(t), \psi_h \rangle = 0,$$

so that, for any choice of the control v_h , we have

$$\langle \mathcal{L}_T^h(v_h | y_{0,h}), \psi_h \rangle_h = \langle y_h(T), \psi_h \rangle_h = e^{-\mu_h T} \langle y_{0,h}, \psi_h \rangle_h, \quad (2.2)$$

and this quantity (which does not depend on the control v_h) can not be zero as soon as the initial data $y_{0,h}$ satisfies $\langle y_{0,h}, \psi_h \rangle_h \neq 0$.

In the case of a Cartesian uniform mesh of the unit square in 2D, \mathcal{A}_h being the usual 5-point discrete Laplace operator in the finite-difference framework, we show in Figure 1 an example of a non-controllable eigenmode ψ_h . As we can see, the difficulty comes from the fact that ψ_h is supported exclusively on the diagonal cells of the mesh, that is to say that \mathcal{A}_h possesses eigenfunctions that are compactly supported. Observe that this pathology does not occur in the continuous setting and this is one of the reason why the analysis of semi-discrete problems contains new difficulties.

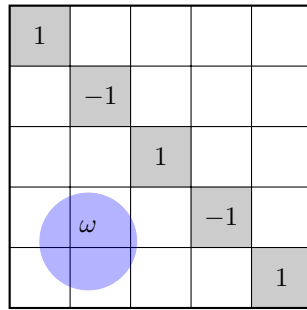


FIGURE 1. A non-controllable mode ψ_h for the five-point discrete Laplace operator on a 2D uniform mesh; the value of ψ_h is set to zero in the white cells

Actually, one can observe that, at least in this example, the eigenvalue associated with ψ_h is asymptotically very large ($\mu_h \sim C/h^2$). Therefore, even if this mode generates non-controllable initial data, the quantity $\langle \mathcal{L}_T^h(v_h | y_{0,h}), \psi_h \rangle_h$ is exponentially small (see (2.2)) and thus should not lead to numerical difficulties.

We shall see in Section 2.5 that this is, in some sense, the generic behavior. To the author's knowledge, the only case for which uniform null-controllability of a semi-discrete approximation of a parabolic

equation is known is the one given in [LZ98] where the authors study the 1D, constant coefficient, heat equation on a uniform mesh and with a boundary Dirichlet control. In this particular case, they managed to prove the result by an explicit computation of the control in Fourier variables. The same result, even in 1D, for variable coefficient operator and/or non uniform meshes is still an open problem. We should also mention the 2D result given in [Zua06] which essentially consists, in a very particular geometry, to reduce the problem to the above 1D case.

- (2) Even if we assume that the semi-discrete problem (\mathcal{P}_h) is null-controllable for any $h > 0$, we can wonder whether or not it is desirable to compute such a null-control since the dynamical system (\mathcal{P}_h) is itself an approximation of the original system (\mathcal{P}) . We shall see that, h being fixed, the computational cost of the penalised HUM control $v_{\varepsilon,h}$ may increase dramatically as $\varepsilon \rightarrow 0$.

Therefore we shall see that is reasonable to choose a penalization parameter ε which depends on h in such a way that the overall accuracy of the computation is of the same order as that of the numerical scheme under study while maintaining an acceptable computational cost (see Section 4.2).

2.3. Penalised HUM in E_h , $\phi(h)$ -null-controllability and relaxed observability

2.3.1. Notation and preliminary remarks

We are now led to the following problem. Given, for each $h > 0$, an initial data $y_{0,h} \in E_h$, we try to minimise the following two functionals for any value of the penalisation parameter $\varepsilon > 0$

$$F_{\varepsilon,h}(v_h) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^T \llbracket v_h(t) \rrbracket_h^2 dt + \frac{1}{2\varepsilon} \|\mathcal{L}_T^h(v_h|y_{0,h})\|_h^2, \quad \forall v_h \in L^2(0,T;U_h),$$

and

$$J_{\varepsilon,h}(q_h^F) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^T \left\| \mathcal{B}_h^* e^{-(T-t)A_h^*} q_h^F \right\|_h^2 dt + \frac{\varepsilon}{2} \|q_h^F\|_h^2 + \left\langle y_{0,h}, e^{-T A_h^*} q_h^F \right\rangle_h, \quad \forall q_h^F \in E_h.$$

For a fixed value of h , all the results of Section 1 directly apply.

Denoting by $v_{\varepsilon,h}$ the unique minimiser of $F_{\varepsilon,h}$, we have seen that, if $Q_{F,h}$ defined in (2.1) is not trivial, we may have

$$\lim_{\varepsilon \rightarrow 0} \|\mathcal{L}_T^h(v_{\varepsilon,h}|y_{0,h})\|_h \neq 0, \quad \forall h > 0,$$

but, however, since E_h is finite dimensional (and according to Remark 1.16) we have

$$\sup_{\varepsilon > 0} \llbracket v_{\varepsilon,h} \rrbracket_{L^2(0,T;U_h)} < +\infty, \quad \forall h > 0.$$

Of course, this bound on $(v_{\varepsilon,h})_\varepsilon$ may depend on h .

Still using that E_h is finite dimensional, we know from Remark 1.16 that there is a $C_h > 0$ such that

$$\|\mathcal{L}_T^h(v_{\varepsilon,h}|y_{0,h}) - \mathbb{P}_{Q_{T,h}}(e^{-T A_h} y_{0,h})\|_h \leq C_h \varepsilon, \quad \forall \varepsilon > 0.$$

Since C_h depends on h in a non explicit way, this estimate is not very useful from the practical point of view, when one wants to analyse the limit $(\varepsilon, h) \rightarrow (0, 0)$.

2.3.2. Coupling penalisation and discretisation

To circumvent the issues described in Section 2.2 and yet to achieve a satisfactory approximation of a null-control of the original problem, we propose to consider the penalised HUM method for the family of semi-discrete problems in the case where the penalisation parameter ε is chosen as a function of the discretisation parameter h . To this end, we suppose given a non-decreasing function $h \in]0, +\infty[\rightarrow \phi(h) \in]0, +\infty[$ such that $\lim_{h \rightarrow 0} \phi(h) = 0$.

Moreover, since we are interested in a family of problems indexed by the parameter h , we naturally need to introduce the space of families of initial data

$$E_{\text{init}} = \prod_{h>0} E_h.$$

An element $Y_0 \in E_{\text{init}}$ is a family $Y_0 = (y_{0,h})_h$ with $y_{0,h} \in E_h$. We say that Y_0 is bounded if $\sup_{h>0} \|y_{0,h}\|_h < +\infty$.

As we observed in Section 1.7, in a slightly different framework, the penalisation parameter ε should also be related in some sense to the approximation properties of the initial data (see in particular (1.33) and (1.34)) if one wants to produce sequences of approximate controls that converge towards a control of the initial problem. This is an additional motivation for introducing the following definition.

Definition 2.1. Let $Y_0 \in E_{\text{init}}$. We say that the semi-discrete problems (\mathcal{P}_h) are $\phi(h)$ -null-controllable from Y_0 if there exists a $h_0 > 0$ such that we have

$$M_{Y_0}^2 \stackrel{\text{def}}{=} 2 \sup_{0 < h < h_0} \left(\inf_{L^2(0,T;U_h)} F_{\phi(h),h} \right) < +\infty,$$

where the functional $F_{\phi(h),h}$ is built with the corresponding initial data given in the family Y_0 .

Observe that, in this semi-discrete setting, the map $h \mapsto \inf_{L^2(0,T;U_h)} F_{\phi(h),h}$ is not necessarily monotone.

Adapting the proofs given in Section 1, we can easily obtain the following result.

Proposition 2.2. For a given $Y_0 \in E_{\text{init}}$, the problems (\mathcal{P}_h) are $\phi(h)$ -null-controllable from Y_0 if and only if there exists $h_0 > 0$ and $\widetilde{M}_{Y_0} > 0$, independent of h , such that the following relaxed observability inequality holds

$$\left| \left\langle y_{0,h}, e^{-T\mathcal{A}_h^*} q_h^F \right\rangle_h \right|^2 \leq \widetilde{M}_{Y_0}^2 \left(\left[\mathcal{B}_h^* e^{-(T-\cdot)\mathcal{A}_h^*} q_h^F \right]_{L^2(0,T;U_h)}^2 + \phi(h) \|q_h^F\|_h^2 \right), \quad \forall q_h^F \in E_h, \forall 0 < h < h_0. \quad (2.3)$$

In such case, the smallest constant \widetilde{M}_{Y_0} satisfying (2.3) is equal to M_{Y_0} and we have

$$\|v_{\phi(h),h}\|_{L^2(0,T;U_h)} \leq M_{Y_0}, \quad \forall 0 < h < h_0,$$

$$\|\mathcal{L}_T^h(v_{\phi(h),h}|y_{0,h})\|_h \leq M_{Y_0} \sqrt{\phi(h)} \quad \forall 0 < h < h_0.$$

Remark 2.3. Note that (2.3) is automatically satisfied in the case where Y_0 is such that $\|y_{0,h}\|_h$ tends to zero sufficiently rapidly (using that all norms in the finite dimensional space E_h are equivalent).

Of course, in practice, this case is not interesting. The family Y_0 is rather meant to approximate a non trivial continuous initial data $y_0 \in E$, with for example the convergence of $(\|y_{0,h}\|_h)_h$ to $\|y_0\| \neq 0$ as $h \rightarrow 0$ (see for instance Theorem 2.5).

Proposition 2.4. Assume that, for some $C_{\text{obs}} > 0$, the following relaxed observability inequality holds

$$\|e^{-T\mathcal{A}_h^*} q_h^F\|_h^2 \leq C_{\text{obs}}^2 \left(\left[\mathcal{B}_h^* e^{-(T-\cdot)\mathcal{A}_h^*} q_h^F \right]_{L^2(0,T;U_h)}^2 + \phi(h) \|q_h^F\|_h^2 \right), \quad \forall q_h^F \in E_h, \forall 0 < h < h_0, \quad (2.4)$$

then for any bounded family of initial data $Y_0 \in E_{\text{init}}$, the problems (\mathcal{P}_h) are $\phi(h)$ -null-controllable from Y_0 and we have

$$M_{Y_0} \leq C_{\text{obs}} \left(\sup_{0 < h < h_0} \|y_{0,h}\|_h \right).$$

Notice that the value of C_{obs} depends in general on the function ϕ that we have chosen at the beginning.

In the sequel of the section, we review some recent results concerning the proof of (uniform with respect to h) relaxed observability inequalities similar to (2.4) for two slightly different discrete frameworks.

2.4. First example : an abstract Galerkin framework

In [LT06], following the framework introduced in [LT00], the authors study the above-mentioned problem in a quite general abstract setting that we summarize here.

Even though these authors are able to cope with some unbounded control operators (namely those for which there exists $\gamma \in [0, 1/2[$ such that $B : U \mapsto (D(\mathcal{A}^*)^\gamma)'$) we only state here the result for a bounded control operator (i.e. for $\gamma = 0$).

- We suppose given linear mappings $\tilde{P}_h : E_h \rightarrow D((\mathcal{A}^*)^{\frac{1}{2}})$ and $\tilde{Q}_h : U_h \rightarrow U$ such that

$$\|y_h\|_h = \left\| \tilde{P}_h y_h \right\|, \forall y_h \in E_h, \quad \text{and} \quad \llbracket u_h \rrbracket_h = \left\| \tilde{Q}_h u_h \right\|.$$

- We set $P_h = (\tilde{P}_h)^* : D((\mathcal{A}^*)^{\frac{1}{2}})' \rightarrow E_h$ and $Q_h = (\tilde{Q}_h)^* : U \rightarrow U_h$ and we assume that

$$P_h \tilde{P}_h = \text{Id}_{E_h}, \quad \text{and} \quad Q_h \tilde{Q}_h = \text{Id}_{U_h},$$

where we have implicitly identified E to its dual space and therefore we have the continuous embeddings $D((\mathcal{A}^*)^{\frac{1}{2}}) \subset E \subset D((\mathcal{A}^*)^{\frac{1}{2}})'$.

- We define now \mathcal{A}_h and \mathcal{B}_h through their adjoints by the formulas

$$\mathcal{A}_h^* = P_h \mathcal{A}^* \tilde{P}_h, \quad \mathcal{B}_h^* = Q_h \mathcal{B}^* \tilde{P}_h.$$

We further assume standard approximation properties (see the details in [LT06]) which require that, for some $s > 0$

- the following consistency estimates hold

$$\left\| \psi - \tilde{P}_h P_h \psi \right\| + \left\| \mathcal{B}^* \psi - \tilde{Q}_h Q_h \mathcal{B}^* \psi \right\| \leq Ch^s \|\mathcal{A}^* \psi\|, \quad \forall \psi \in D(\mathcal{A}^*),$$

$$\tilde{Q}_h Q_h u \xrightarrow{h \rightarrow 0} u, \quad \forall u \in U.$$

- An error estimate for the underlying elliptic problem holds that is

$$\left\| P_h (\mathcal{A}^*)^{-1} f - (\mathcal{A}_h^*)^{-1} P_h f \right\|_h \leq Ch^s \|f\|, \quad \forall f \in E.$$

- the semi-discrete semi-groups $t \mapsto e^{-t\mathcal{A}_h}$ are analytic in a uniform way with respect to h , that is to say that some uniform resolvent estimates hold.
- Finally, we assume that the numerical scheme is convergent in the following sense: for any $y_0 \in E$ and any $v \in L^2(0, T; U)$ we have

$$\left\| \mathcal{L}_T^h(Q_h v | P_h y_0) - P_h \mathcal{L}_T(v | y_0) \right\|_h = \left\| \tilde{P}_h \mathcal{L}_T^h(Q_h v | P_h y_0) - \tilde{P}_h P_h \mathcal{L}_T(v | y_0) \right\| \xrightarrow{h \rightarrow 0} 0. \quad (2.5)$$

As an example, it can be shown that the standard piecewise-affine finite element approximation for the heat equation on a quasi-uniform mesh of a polygonal domain enters this general framework.

Theorem 2.5 (see [LT06, Theorem 3.1]). *Assume that (\mathcal{P}) is null-controllable at time T . Under the previous assumptions on the semi-discrete systems, there exists a $\beta > 0$, such that the relaxed-observability inequality (2.4) holds as soon the function ϕ is chosen in such a way that*

$$\liminf_{h \rightarrow 0} \frac{\phi(h)}{h^\beta} > 0. \quad (2.6)$$

In that case, for any $y_0 \in E$, we define $y_{0,h} = P_h y_0$ and we consider the associated penalised HUM discrete controls $v_{\phi(h),h}$. For simplicity the corresponding solution of (\mathcal{P}_h) is denoted by y_h . Then, there exists a null-control $v \in \text{Adm}(y_0, 0)$ such that, up to a subsequence, we have

$$\begin{aligned} \tilde{Q}_h v_{\phi(h),h} &\xrightarrow{h \rightarrow 0} v, & \text{in } L^2(0, T; U), \\ \tilde{P}_h y_h &\xrightarrow{h \rightarrow 0} y_{v, y_0}, & \text{in } L^2(0, T; E), \\ \tilde{P}_h \mathcal{L}_T^h(v_{\phi(h),h} | y_{0,h}) &= \tilde{P}_h y_h(T) \xrightarrow{h \rightarrow 0} 0 = \mathcal{L}_T(v | y_0), & \text{in } E. \end{aligned}$$

Remark 2.6. (1) The main line followed during the proof of this result consists in deducing a relaxed-observability inequality for the semi-discrete adjoint system by using the corresponding observability inequality of the continuous problem. In particular, it is mandatory here to assume the null-controllability of (\mathcal{P}) .

Notice that this strategy is very different from the one described in Section 2.5.

- (2) In this approach, due to the strategy of proof we just described above, the number $\beta > 0$ actually depends explicitly on the various convergence rates that are assumed above. When the value of β is computed for specific situations, it appears that it can be quite small. The example of the piecewise-affine finite element approximation of the 1D heat equation with a Neumann boundary control is described in [LT06], and it is shown that $\beta = 0.45$ in that case.

Choosing for instance $\phi(h) = h^\beta$ (this is the best we can do according to the theorem above) leads to a convergence rate of $\tilde{P}_h y_h(T)$ towards 0 in $E = L^2(\Omega)$ which is given by $\sqrt{\phi(h)} = h^{\frac{\beta}{2}} = h^{0.225}$. This convergence rate is very small with respect to the expected accuracy of the numerical method, which is second order in the L^2 -norm, in that case.

We want to emphasize here that this theorem does not claim, in general, that the limit control v is indeed the unique HUM control v^0 of the continuous problem. This is the reason why we only have a weak convergence result up to a subsequence.

In general, even if we assume that the limit v is the HUM control v^0 , we are not able to show that the convergence of the approximate control is strong. Indeed, if we would like to mimick the proof of the strong convergence of penalised HUM controls given in Theorem 1.7, we see that we would need to take (in some sense) the limit control v as a test function in the minimisation problem associated with $F_{\phi(h),h}$. Since v does not belong to the semi-discrete space $L^2(0, T; U_h)$, it is clear that additional arguments are needed¹. Let us give an example.

Proposition 2.7. *Consider the same assumptions as in Theorem 2.5. Given an initial data $y_0 \in E$ and the associated HUM control v^0 , we set*

$$\Psi(h) \stackrel{\text{def}}{=} \left\| \mathcal{L}_T^h(Q_h v^0 | P_h y_0) \right\|_h = \left\| \tilde{P}_h \mathcal{L}_T^h(Q_h v^0 | P_h y_0) \right\|, \quad \forall h > 0.$$

By the convergence assumption (2.5), and using that $\mathcal{L}_T(v^0 | y_0) = 0$, we know that $\lim_{h \rightarrow 0} \Psi(h) = 0$.

We suppose that the penalisation function ϕ is chosen in such a way that

$$\lim_{h \rightarrow 0} \frac{\Psi(h)^2}{\phi(h)} = 0. \quad (2.7)$$

Then, the limit v in Theorem 2.5 is equal to the HUM control v^0 and the whole family $(v_{\phi(h),h}, y_h)_h$ satisfy the convergences given in that theorem. Moreover, the convergence of $\tilde{Q}_h v_{\phi(h),h}$ towards v^0 is strong in $L^2(0, T; U)$.

¹In the very particular case of the Dirichlet boundary control for a 1D equation, we have $U = U_h = \mathbb{R}$, so that strong convergence can be proved this way, as noticed in [LZ98].

Remark 2.8. Condition (2.7) says, in addition to (2.6), that $h \rightarrow \phi(h)$ should not converge to 0 too fast if one wants to ensure the strong convergence of the approximate controls. However, in practice, the function Ψ is unknown except if one postulates sufficient *a priori* regularity properties for the HUM control and the initial data so that $\Psi(h)$ can be estimated by some power of h , depending on the accuracy of the numerical method. Unfortunately the HUM control is not expected to be smooth enough in general.

Actually, in all the numerical experiments presented here, we never observed a behavior of the approximate controls that can indicate a possible defect of strong convergence even though it is very likely that condition (2.7) is not satisfied.

Proof. We need to compute an upper bound of the $L^2(0, T; U)$ -norm of the interpolation of the approximate control. To this end we use that this approximate control is the minimiser of $F_{\phi(h), h}$ and the definitions of the norms in E_h and U_h as follows

$$\begin{aligned} \frac{1}{2} \left[\left[\tilde{Q}_h v_{\phi(h), h} \right]_{L^2(0, T; U)} \right]^2 &= \frac{1}{2} \left[v_{\phi(h), h} \right]_{L^2(0, T; U_h)}^2 \\ &\leq F_{\phi(h), h}(v_{\phi(h), h}) = \inf_{L^2(0, T; U_h)} F_{\phi(h), h} \\ &\leq F_{\phi(h), h}(Q_h v^0) \\ &= \frac{1}{2} \left[Q_h v^0 \right]_{L^2(0, T; U_h)}^2 + \frac{1}{2\phi(h)} \left\| \mathcal{L}_T^h(Q_h v^0 | P_h y_0) \right\|_h^2 \\ &= \frac{1}{2} \left[\tilde{Q}_h Q_h v^0 \right]_{L^2(0, T; U)}^2 + \frac{1}{2\phi(h)} \left\| \tilde{P}_h \mathcal{L}_T^h(Q_h v^0 | P_h y_0) \right\|^2. \end{aligned}$$

Using (2.7), and the approximation properties of Q_h , we deduce that

$$\limsup_{h \rightarrow 0} \left[\left[\tilde{Q}_h v_{\phi(h), h} \right]_{L^2(0, T; U)} \right] \leq \left[v^0 \right]_{L^2(0, T; U)}.$$

Since v^0 is, by definition, the null-control with minimal $L^2(0, T; U)$ -norm, we have $\left[v^0 \right]_{L^2(0, T; U)} \leq \left[v \right]_{L^2(0, T; U)}$ and thus we deduce that the convergence of the subsequence of $(\tilde{Q}_h v_{\phi(h), h})_h$ towards v is strong and that $\left[v \right]_{L^2(0, T; U)} = \left[v^0 \right]_{L^2(0, T; U)}$. Therefore, we have $v = v^0$ and the whole sequence strongly converges towards v^0 . \square

Remark 2.9. Another sufficient condition for the strong convergence of the approximate control towards the HUM control is also given in [LT06, Proposition 3.2]. This condition is different from the one given in the previous Proposition since it consists in assuming that the sequence $\left\| q_h^F \right\|_h$ of the norms of the optimal adjoint states is bounded. We will see in the numerical results given in Section 4 that this boundedness property has almost no chance to be satisfied in general. We rather expect that $\left\| q_h^F \right\|_h \sim_0 C\phi(h)^{-1/2}$ or equivalently $\left\| \mathcal{L}_T^h(v_{\phi(h), h} | y_{0, h}) \right\|_h \sim_0 C\phi(h)^{\frac{1}{2}}$, excepted possibly in the case of a non-localised control as discussed in Section 4.2 and illustrated in Figure 6.

2.5. Second example : finite-difference schemes for the heat equation

In a series of articles [BHL10a, BHL10b, BHL11, BL13], a different strategy was developed consisting in directly mimicking at the discrete level various techniques coming from the analysis of PDE control problems. This led to more precise results (in some sense almost optimal) as we shall see below. The price to pay is that, for the moment, those results do not exist in a general setting but only in the framework of finite-difference schemes for scalar equations and distributed controls. However, numerical results seem to prove that the same results should hold in more general settings (see Section 4).

2.5.1. The discrete Lebeau-Robbiano strategy

In this spirit, it was first proposed to deal with a discrete Lebeau-Robbiano strategy following the original proof of null-controllability of the heat equation given in [LR95].

The starting point is to assume that a suitable spectral inequality for the underlying discrete elliptic operator holds uniformly with respect to h . Proving such a property will be discussed just after.

We assume that \mathcal{A}_h is symmetric positive definite in $(E_h, \langle \cdot, \cdot \rangle_h)$ and we denote by $(\psi_{j,h})_j$ an orthonormal basis of E_h made of eigenfunctions of \mathcal{A}_h . The corresponding eigenvalues will be denoted by $\mu_{j,h}$.

We also assume that \mathcal{B}_h^* is a family of uniformly bounded operators, that is, for some $C > 0$,

$$\|[\mathcal{B}_h^* x_h]\|_h \leq C \|x_h\|_h, \quad \forall x \in E_h, \forall h > 0. \quad (2.8)$$

This assumption prevents us in particular to deal with boundary control problems with this strategy.

Our main assumption is the following discrete Lebeau-Robbiano spectral inequality: there exists $h_0 > 0$, $\alpha \in [0, 1)$, $\beta > 0$, and $\kappa, \ell > 0$ such that, for any $h < h_0$ and for any $(a_j)_j \in \mathbb{R}^{\mathbb{N}}$, we have

$$\left\| \sum_{\mu_{j,h} \leq \mu} a_j \psi_{j,h} \right\|_h^2 = \sum_{\mu_{j,h} \leq \mu} |a_j|^2 \leq \kappa e^{\kappa \mu^\alpha} \left\| \mathcal{B}_h^* \left(\sum_{\mu_{j,h} \leq \mu} a_j \psi_{j,h} \right) \right\|_h^2, \quad \forall \mu < \frac{\ell}{h^\beta}. \quad (\mathcal{H}_{\alpha,\beta})$$

Remark 2.10. In general, the dimension of U_h is less than that of E_h (typically for a distributed control acting on some strict subdomain of the computational domain). Therefore we cannot expect $(\mathcal{H}_{\alpha,\beta})$ to be true for all values of μ . Actually, we easily see that $(\mathcal{H}_{\alpha,\beta})$ cannot hold for the values of μ such that $\#\{j, \mu_{j,h} \leq \mu\}$ is greater than the dimension of U_h .

Moreover, the existence of a particular eigenmode ψ_h satisfying $\mathcal{B}_h^* \psi_h = 0$ as described at the beginning of Section 2.2 (see Figure 1), also proves that $(\mathcal{H}_{\alpha,\beta})$ can not hold for any value of μ . Notice that the eigenvalue μ_h associated with ψ_h is asymptotically of order C/h^2 . Thus, assuming a critical value of μ of the form $\frac{\ell}{h^\beta}$ in $(\mathcal{H}_{\alpha,\beta})$ seems to be quite reasonable and actually essentially optimal.

With this assumption at hand, the following result can be proved. Notice that this theorem does not require any assumption on the underlying continuous problem (\mathcal{P}) . Actually it is an abstract result and the existence of an underlying continuous problem needs not be considered even.

Theorem 2.11. *Assume that assumption $(\mathcal{H}_{\alpha,\beta})$ holds, then there exists $h_0 > 0$, $C > 0$ such that, the relaxed observability inequality (2.4) holds as soon as the function ϕ is chosen in such a way that*

$$\liminf_{h \rightarrow 0} \frac{\phi(h)}{e^{-C/h^\beta}} > 0. \quad (2.9)$$

In such a case, for any family of initial data $Y_0 \in E_{\text{init}}$, the associated approximate controls and solutions satisfy

$$\begin{aligned} \|v_{\phi(h),h}\|_{L^2(0,T;U_h)} &\leq C_{\text{obs}} \|y_{0,h}\|_h, \quad \forall h > 0, \\ \|\mathcal{L}_T^h(v_{\phi(h),h}|y_{0,h})\|_h &\leq C_{\text{obs}} \|y_{0,h}\|_h \sqrt{\phi(h)}, \quad \forall h > 0. \end{aligned}$$

We emphasize the fact that condition (2.9) only requires that ϕ does not tend to 0 faster than some exponential term, which is far less restrictive than condition (2.6). In practice, we may thus choose for instance $\phi(h) = h^{2p}$ where p is the expected order of accuracy of the numerical scheme (compare with the discussion in Remark 2.6).

2.5.2. Proving discrete Lebeau-Robbiano inequalities

The original proof of the Lebeau-Robbiano spectral inequality in the continuous case was based on local Carleman elliptic estimates [LR95]. An alternative proof is available by using global Carleman estimates for an augmented elliptic operator [Le 07]. This last strategy was adapted to the semi-discrete augmented operator $-\partial_s^2 + \mathcal{A}_h$ by establishing uniform discrete global Carleman estimates in [BHL10a, BHL10b].

More precisely, it is proved that $(\mathcal{H}_{\alpha,\beta})$ holds with $\alpha = 1/2$ and $\beta = 2$ in the following framework:

- We suppose given a quasi-uniform Cartesian mesh of the unit cube $\Omega =]0, 1[^d$ (any domain for which a Cartesian mesh is available also works). The various norms are defined in the usual sense.
- The discrete operator \mathcal{A}_h is the usual $2d + 1$ -point finite difference approximation of the operator $-\operatorname{div}(\gamma \nabla \cdot)$, where $\gamma : \Omega \rightarrow \mathbb{R}$ is Lipschitz continuous and coercive.
- The discrete (distributed) control operator \mathcal{B}_h is simply $\mathbf{1}_\omega$ for some given $\omega \subset \Omega$. More precisely it means that \mathcal{B}_h only acts on the nodes of the discretisation which belongs to ω .

2.5.3. The discrete Fursikov-Imanuvilov strategy

A second strategy was proposed in [FI96] to prove controllability in the continuous case. It is based on the derivation of (global) Carleman estimates for the parabolic operator. This approach has been efficient to cope with more general parabolic problems (time dependent coefficients, non-linear, and so on).

It has been adapted to the semidiscrete setting in [BL13]. The general finite-difference framework that we use here is the same as in the previous section, excepted that the operator we consider is now time dependent and has the following form

$$\mathcal{A}_h = \mathcal{A}_h^0 + a_h,$$

where \mathcal{A}_h^0 is the $2d + 1$ -point discrete approximation of $-\operatorname{div}(\gamma \nabla \cdot)$ introduced above and a_h is any function in $L^\infty(0, T; E_h)$. Note that we use here the usual abuse of notation which consists in denoting also by $a_h(t)$ the operator from E_h into itself which is the point by point multiplication by the discrete function $a_h(t, \cdot)$.

Theorem 2.12 ([BL13]). *There exist $C > 0$ and $h_0 > 0$ such that for any ϕ satisfying*

$$\liminf_{h \rightarrow 0} \frac{\phi(h)}{e^{-C/h}} > 0, \quad (2.10)$$

any $A > 0$, and any $h \leq \min(h_0, A^{-2/3})$, the following $\phi(h)$ -relaxed observability inequality holds

$$\|q_h(0)\|_h^2 \leq C_A^2 \left(\|\mathcal{B}_h^* q_h\|_{L^2(0, T; U_h)}^2 + \phi(h) \|q_h^F\|_h^2 \right), \quad (2.11)$$

for any solution of the adjoint problem

$$\begin{cases} -\partial_t q_h + \mathcal{A}_h^0 q_h + a_h q_h = 0, \\ q_h(T) = q_h^F \in E_h, \end{cases}$$

where $a_h \in L^\infty(0, T; E_h)$ is any function satisfying $\|a_h\|_{L^\infty} \leq A$.

Notice that the form (2.11) of the inequality is slightly different from that in (2.4) since the adjoint operator we consider here has time-dependent coefficient so that the exponential notation is not suitable here.

Note that the dependency of C_A with respect to the L^∞ bound of a_h (that is to say with respect to A) can be precisely stated (see [BL13]). With this result at hand we deduced controllability results concerning semi-discrete semilinear problems can be deduced by adapting the strategy developed in [FCZ00].

Theorem 2.13 ([BL13]). *Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a Lipschitz continuous function and ϕ satisfying (2.10).*

- *If g is bounded, then there exists $C > 0$ such that for any $0 < h < h_0$, any $y_{0,h} \in E_h$ there exists a semi-discrete control $v_h \in L^2(0, T; U_h)$ for the problem*

$$\begin{cases} \partial_t y_h + \mathcal{A}_h^0 y_h + g(y_h) y_h = \mathcal{B}_h v_h, \\ y_h(0) = y_{0,h}, \end{cases} \quad (\mathcal{P}_h^g)$$

which satisfies

$$\begin{aligned} \llbracket v_h \rrbracket_{L^2(0,T;U_h)} &\leq C \|y_{0,h}\|_h, \\ \|y_h(T)\|_h &\leq C \sqrt{\phi(h)} \|y_{0,h}\|_h. \end{aligned}$$

- In dimension $d = 1$, the same result holds as soon as g satisfies, for some $M > 0$ and $0 < r < 3/2$

$$|g(s)| \leq M \log(1 + |s|)^r, \quad \forall s \in \mathbb{R}.$$

- In dimension $d > 1$, we can prove the existence of a semi-discrete control such that

$$\begin{aligned} \llbracket v_h \rrbracket_{L^2(0,T;U_h)} &\leq Ch^{-\alpha} \|y_{0,h}\|_h, \\ \|y_h(T)\|_h &\leq C \sqrt{\phi(h)} \|y_{0,h}\|_h, \end{aligned}$$

for some $\alpha > 0$. In that case, proving an uniform bound on v_h is still an open problem. Therefore, our result can be seen as a $\phi(h)$ -approximate controllability result for the nonlinear problem.

3. FULLY-DISCRETE APPROXIMATIONS

In the space discretisation framework introduced in Section 2.5.1, some fully-discrete numerical methods for (linear) parabolic problems have been studied in [BHL11].

For $M > 0$, we set $\delta t = T/M$ and we consider a θ -scheme, $\theta \in [1/2, 1]$, with respect to the time variable defined by

$$\begin{cases} \frac{y_h^{n+1} - y_h^n}{\delta t} + \mathcal{A}_h(\theta y_h^{n+1} + (1-\theta)y_h^n) = \mathcal{B}_h v_h^{n+1}, \quad \forall n \in \llbracket 0, M-1 \rrbracket, \\ y_h^0 = y_{0,h} \in E_h, \end{cases} \quad (\mathcal{P}_{h,\delta t})$$

where, $v_{h,\delta t} = (v_h^n)_{1 \leq n \leq M} \in (U_h)^M$ is a fully-discrete control function whose cost, that is the discrete $L^2_{\delta t}(0, T; U_h)$ -norm, is defined by

$$\llbracket v_{h,\delta t} \rrbracket_{L^2_{\delta t}(0,T;U_h)} \stackrel{\text{def}}{=} \left(\sum_{n=1}^M \delta t \llbracket v_h^n \rrbracket_h^2 \right)^{\frac{1}{2}}.$$

The final value of the controlled solution of $(\mathcal{P}_{h,\delta t})$ is denoted by $\mathcal{L}_T^{h,\delta t}(v_{h,\delta t} | y_{0,h}) \stackrel{\text{def}}{=} y_h^M$.

With the above notation, the primal fully-discrete penalised HUM functional reads as follows

$$v_{h,\delta t} \in U_h^M \mapsto F_{\varepsilon,h,\delta t}(v_{h,\delta t}) = \frac{1}{2} \llbracket v_{h,\delta t} \rrbracket_{L^2_{\delta t}(0,T;U_h)}^2 + \frac{1}{2\varepsilon} \|\mathcal{L}_T^{h,\delta t}(v_{h,\delta t} | y_{0,h})\|_h^2.$$

The first step in the analysis is to identify the correct dual functional.

Proposition 3.1. *We define the functional*

$$J_{\varepsilon,h,\delta t}(q_h^F) \stackrel{\text{def}}{=} \frac{1}{2} \llbracket \mathcal{B}_h^* \mathcal{L}_T^{*,h,\delta t}(q_h^F) \rrbracket_{L^2_{\delta t}(0,T;U_h)}^2 + \frac{\varepsilon}{2} \|q_h^F\|_h^2 - \langle y_{0,h}, q_h^1 - \delta t(1-\theta)\mathcal{A}_h q_h^1 \rangle_h, \quad \forall q_h^F \in E_h, \quad (3.1)$$

where $\mathcal{L}_T^{*,h,\delta t}(q_h^F) = (q_h^n)_{1 \leq n \leq M}$ is the solution of the following adjoint problem

$$\begin{cases} q_h^{M+1} = q_h^F, \\ \frac{q_h^M - q_h^{M+1}}{\delta t} + \theta \mathcal{A}_h q_h^M = 0, \\ \frac{q_h^n - q_h^{n+1}}{\delta t} + \mathcal{A}_h(\theta q_h^n + (1-\theta)q_h^{n+1}) = 0, \quad \forall n \in \llbracket 1, M-1 \rrbracket. \end{cases} \quad (\mathcal{P}_{h,\delta t}^*)$$

Notice that q_h^{M+1} is not taken into account in the definition of $\mathcal{L}_T^{*,h,\delta t}(q_h^F)$.

The functionals $F_{\varepsilon,h,\delta t}$ and $J_{\varepsilon,h,\delta t}$ are in duality, in the sense that their respective minimisers $v_{\varepsilon,h,\delta t} \in L^2(0,T;U_h)$ and $q_{\varepsilon,h,\delta t}^F \in E_h$ satisfy

$$\inf_{L^2_{\delta t}(0,T;U_h)} F_{\varepsilon,h,\delta t} = F_{\varepsilon,h,\delta t}(v_{\varepsilon,h,\delta t}) = -J_{\varepsilon,h,\delta t}(q_{\varepsilon,h,\delta t}^F) = -\inf_{E_h} J_{\varepsilon,h,\delta t},$$

and moreover

$$v_{\varepsilon,h} = \mathcal{B}_h^* \mathcal{L}_T^{*,h,\delta t}(q_{\varepsilon,h,\delta t}^F).$$

Notice that, for $\theta \in [1/2, 1[$, the first step of the adjoint problem ($\mathcal{P}_{h,\delta t}^*$) has a particular form which is not exactly that of the usual backward θ -scheme. The important point is that we cannot choose independently the time discretisation for the forward and the backward problem.

With this fully-discrete problem, we can follow the same lines as in the semi-discrete case and observe that for fixed h and δt , the computed target converges (as $\varepsilon \rightarrow 0$) to the projection of the free solution onto the set

$$Q_{F,h,\delta t} = \{q_h^F \in E_h, \text{ s.t. } \mathcal{B}_h^* \mathcal{L}_T^{*,h,\delta t}(q_h^F) = 0\}.$$

Here also, it make sense to choose ε depending on the discretisation parameters. More precisely, the following fully-discrete relaxed observability inequality is proved in [BHL11]. Observe that it is not necessary here to consider a time-step-dependent penalisation parameter.

Theorem 3.2 (Case $\theta \in]1/2, 1[$). *Assume that the uniform discrete Lebeau-Robbiano inequality ($\mathcal{H}_{\alpha,\beta}$) holds and let ϕ be such that*

$$\liminf_{h \rightarrow 0} \frac{\phi(h)}{e^{-C/h^\beta}} > 0.$$

Then, there exists $h_0 > 0, C_T > 0, C_{\text{obs}} > 0$ such that for any $0 < h < h_0$ and any $\delta t \leq C_T |\log \phi(h)|^{-1}$, the following relaxed observability inequality holds

$$\|q_h^1 - \delta t(1-\theta)\mathcal{A}_h q_h^1\|_h^2 \leq C_{\text{obs}}^2 \left(\|\mathcal{B}_h^* q_h^n\|_{L^2_{\delta t}(0,T;U_h)}^2 + \phi(h) \|q_h^F\|_h^2 \right).$$

Thus, for any such δt and h and any initial data $y_{0,h} \in E_h$, the full-discrete control $v_{\phi(h),h,\delta t}$, obtained by minimising $F_{\phi(h),h,\delta t}$ (or equivalently $J_{\phi(h),h,\delta t}$) is such that

$$\|v_{\phi(h),h,\delta t}\|_{L^2_{\delta t}(0,T;U_h)} \leq C_{\text{obs}} \|y_{0,h}\|_h, \text{ and } \|\mathcal{L}_T^{h,\delta t}(v_{\phi(h),h,\delta t} | y_{0,h})\|_h \leq C_{\text{obs}} \sqrt{\phi(h)} \|y_{0,h}\|_h.$$

Notice that the proof in [BHL11] is only given in the case $\phi(h) = e^{-C/h^\gamma}$ with $\gamma \leq \beta$ but it can be easily adapted to the more general case stated here. Moreover, the following result for the Crank-Nicolson scheme is also obtained in [BHL11].

Theorem 3.3 (Case $\theta = 1/2$). *Theorem 3.2 holds for $\theta = 1/2$ with the additional assumption $\delta t \rho(\mathcal{A}_h) \leq C_T$, where $\rho(\mathcal{A}_h)$ is the spectral radius of \mathcal{A}_h .*

Remark 3.4. Notice that the additional technical condition on the time step δt that is needed here is very strong (we typically expect $\rho(\mathcal{A}_h) \sim C/h^2$ for instance) and related to the bad dissipation properties for high frequencies of the Crank-Nicolson scheme.

We do not know whether this condition is really necessary or not, however some numerical tests seem to show that the Crank-Nicolson scheme may actually behave incorrectly if such a condition is not met.

An error estimate in time between semi-discrete and fully discrete controls obtained by the penalised HUM approach (as presented here) is also obtained in [BHL11]. More precisely, for $\theta \in]1/2, 1[$, the following estimate

holds

$$\left\| v_{\phi(h),h} - \sum_{n=1}^M \mathbf{1}_{(n-1)\delta t, n\delta t} v_{\phi(h),h}^n \right\|_{L^2(0,T;U_h)} \leq C_h \delta t \|y_{0,h}\|_h,$$

where the constant C_h “badly” depends on h . A second-order estimate in time is also obtained for $\theta = 1/2$.

4. NUMERICAL ILLUSTRATIONS

4.1. Computational method

We want to address here some questions related to the actual computation of the approximate (semi- or fully- discrete) controls for a given problem. Above, we saw that such controls are the minimisers of $F_{\phi(h),h}$ (resp. $F_{\phi(h),h,\delta t}$) but can also be computed by minimising the dual functionals $J_{\phi(h),h}$ (resp. $J_{\phi(h),h,\delta t}$). Since these dual functionals are defined on the finite dimensional space E_h (as opposed to the much larger spaces $L^2(0,T;U_h)$ or $L^2_{\delta t}(0,T;U_h)$ for the primal functionals), we decided (following [GL94, GLH08]) to apply an optimisation algorithm to the dual functionals.

Since these functionals are quadratic and coercive, the conjugate gradient algorithm is a natural and quite simple choice here even though finding an efficient preconditioner for such problems is still an important and mostly open problem. To apply this method we only need to compute, at each iteration, the gradient of $J_{\phi(h),h}$ (or of $J_{\phi(h),h,\delta t}$ in the fully discrete case). Gradients are understood here with respect to the inner product in E_h .

Corollary 4.1. *For any $h > 0, \delta t > 0, \varepsilon > 0$ and any $q_h^F \in E_h$, we have*

$$\nabla J_{\varepsilon,h}(q_h^F) = \underbrace{\mathcal{L}_T^h \left(\mathcal{B}_h^* e^{-(T-\cdot)\mathcal{A}_h^*} q_h^F \mid 0 \right)}_{\stackrel{\text{def}}{=} \Lambda^h q_h^F} + \varepsilon q_h^F + \mathcal{L}_T^h(0 \mid y_{0,h}), \quad (4.1)$$

$$\nabla J_{\varepsilon,h,\delta t}(q_h^F) = \underbrace{\mathcal{L}_T^{h,\delta t} \left(\mathcal{B}_h^* \mathcal{L}_T^{*,h,\delta t}(q_h^F) \mid 0 \right)}_{\stackrel{\text{def}}{=} \Lambda^{h,\delta t} q_h^F} + \varepsilon q_h^F + \mathcal{L}_T^{h,\delta t}(0 \mid y_{0,h}), \quad (4.2)$$

where $\mathcal{L}_T^{*,h,\delta t}(q_h^F)$ is the solution of the adjoint fully-discrete system $(\mathcal{P}_{h,\delta t}^*)$ associated with the final data q_h^F , as defined in Proposition 3.1.

Remark 4.2. The discrete Gramian operators Λ^h and $\Lambda^{h,\delta t}$ have to be understood as the suitable discrete version of the continuous Gramian Λ as defined at the beginning of the proof of Theorem 1.11.

Solving the optimisation problem for these functionals is thus equivalent to solve, in E_h , the equation

$$(\Lambda^h + \varepsilon \text{Id})q_h^F = -\mathcal{L}_T^h(0 \mid y_{0,h}), \quad \text{in the semi-discrete case,}$$

$$(\Lambda^{h,\delta t} + \varepsilon \text{Id})q_h^F = -\mathcal{L}_T^{h,\delta t}(0 \mid y_{0,h}), \quad \text{in the fully-discrete case.}$$

Let us comment on the actual computation of these gradients. The last term in (4.1) or (4.2) (which does not depend on q_h^F) is simply the solution at final time of the semi- or fully-discrete problem without any control and for the initial data $y_{0,h} \in E_h$. The second term is easy to compute. However the computation of the Gramian term for a given q_h^F requires some comments:

- We first need to solve the backward system associated with the final data q_h^F . In the semi-discrete case, this is $t \mapsto e^{-(T-t)\mathcal{A}_h^*} q_h^F$; in the fully-discrete case this is the term $\mathcal{L}_T^{*,h,\delta t}(q_h^F)$.
- Then, we apply the operator \mathcal{B}_h^* to these adjoint states. This gives us a control in $L^2(0,T;U_h)$ or $L^2_{\delta t}(0,T;U_h)$.
- Finally, we solve for the forward semi- or fully-discrete problem with this control in the right-hand side and the zero initial data.

In short, computing the gradient of these functionals requires first to solve an homogeneous backward parabolic problem, and second to solve a non-homogeneous forward parabolic problem. In the semi-discrete case, these computations can be made in an exact way provided that we are able to compute the eigenelements of \mathcal{A}_h and \mathcal{A}_h^* .

We can easily check that, since Λ^h is symmetric non-negative, the following estimates hold for $\varepsilon > 0$ small enough

$$\varepsilon \|q_h^F\|_h \leq \|(\Lambda^h + \varepsilon \text{Id})q_h^F\|_h \leq (C + \varepsilon) \|q_h^F\|_h.$$

We deduce that the condition number of the operator $\Lambda^h + \varepsilon \text{Id}$ behaves like ε^{-1} for ε small enough. Notice that Λ^h may not be invertible: for instance, in the setting introduced in Section 2.2 its kernel $Q_{F,h}$ contains the mode ψ_h . Similar properties hold in the fully-discrete case, independently of δt .

It follows from this analysis that the number of iterations expected to achieve a given accuracy with the conjugate gradient method, which is estimated by the square root of the condition number, should be of the order of $\varepsilon^{-1/2}$. Recall that our method consists in choosing $\varepsilon = \phi(h)$, which gives a number of iterations of the solver that should behave like $\phi(h)^{-1/2}$. We therefore see here that choosing a function ϕ which converges too rapidly towards 0 will lead to a much more complex linear problem to solve. Moreover we saw that $\sqrt{\phi(h)}$ is the expected size of the controlled target given by the computed controls, it seems that a reasonable practical rule is to systematically choose $\phi(h) \sim h^{2p}$ where p is the order of accuracy in space of the numerical method under study. This will be further discussed below.

Remark 4.3. (1) It is important to notice that, in the fully-discrete case, the size of the time-step has almost no influence on the condition number. Of course, choosing a small time step will increase the overall accuracy of the method but will also increase the global computational cost of the algorithm since we recall that two parabolic problems have to be solved (on the whole time interval $[0, T]$) at each iteration of the conjugate gradient solver (see Section 4.3.1).

(2) The behavior of the conjugate gradient solver is not directly related to the lack of uniform null-controllability of the systems under study. More precisely, even if we assume that the semi-discrete problems satisfy uniform discrete non-relaxed observability inequalities (that is (2.4) with $\phi(h) = 0$), then one may try to minimise $J_{0,h}$ and thus to compute an exact null-control from the semi-discrete problem. Nevertheless, we still have the following estimate

$$e^{-T\rho(\mathcal{A}_h^*)} \|q_h^F\|_h^2 \leq \|e^{-T\mathcal{A}_h^*} q_h^F\|_h^2 \leq C_{\text{obs}}^2 \langle \Lambda^h q_h^F, q_h^F \rangle_h,$$

which proves that the condition number of Λ^h in that case can be estimated by $Ce^{T\rho(\mathcal{A}_h^*)}$. In general, $\rho(\mathcal{A}_h^*)$ behaves like a negative power of h which gives a huge condition number for small values of h . Actually, the system is almost ill-posed and it appears that the penalisation term is absolutely necessary even in that case in order to regularise the problem.

This is somewhat related to the fact that, in the continuous setting, the Gramian operator Λ is in general not coercive in E but only in a suitably defined larger space (which the completion of E with respect to the norm defined through Λ , see [GLH08]).

4.2. Comments on the choice of the penalisation function

We want, in this section, to go further in the discussion on the choice of ε as a function of h .

Let us consider the following very simple example. We set $E = E_h = \mathbb{R}$, $\mathcal{A} = \lambda > 0$, $\mathcal{A}_h = (\lambda + \delta_h) \in \mathbb{R}$ with $\delta_h \xrightarrow{h \rightarrow 0} 0$, $\mathcal{B} = \mathcal{B}_h = 1$ and $y_0 = y_{0,h} = 1$. The exact and semi-discrete problems read respectively

$$(\tilde{\mathcal{P}}) \begin{cases} y' + \lambda y = v, \\ y(0) = 1, \end{cases} \quad \text{and} \quad (\tilde{\mathcal{P}}_h) \begin{cases} y_h' + (\lambda + \delta_h)y_h = v_h, \\ y_h(0) = 1. \end{cases}$$

The corresponding Gramian operators defined from \mathbb{R} into itself can be explicitly computed

$$\Lambda_h q^F = \frac{1 - e^{-2(\lambda+\delta_h)T}}{2(\lambda + \delta_h)} q^F, \quad \text{and} \quad \Lambda q^F = \frac{1 - e^{-2\lambda T}}{2\lambda} q^F, \quad \forall q^F \in \mathbb{R},$$

and we have

$$e^{-T\mathcal{A}_h} y_{0,h} = e^{-(\lambda+\delta_h)T}.$$

It follows that the optimal adjoint state which minimises $J_{\varepsilon,h}$ is given by

$$q_{\varepsilon,h}^F = -\frac{2(\lambda + \delta_h)e^{-(\lambda+\delta_h)T}}{1 - e^{-2(\lambda+\delta_h)T} + 2\varepsilon(\lambda + \delta_h)}.$$

The corresponding semi-discrete penalised HUM control is then given by

$$v_{\varepsilon,h}(t) = -e^{-(T-t)(\lambda+\delta_h)} \frac{2(\lambda + \delta_h)e^{-(\lambda+\delta_h)T}}{1 - e^{-2(\lambda+\delta_h)T} + 2\varepsilon(\lambda + \delta_h)},$$

while the exact HUM control is

$$v(t) = -e^{-(T-t)\lambda} \frac{2\lambda e^{-\lambda T}}{1 - e^{-2\lambda T}}.$$

It follows that the following (sharp) estimate holds

$$\|v - v_{\varepsilon,h}\|_{L^2(0,T;U)} \leq C(\lambda, T)(|\delta_h| + \varepsilon), \quad \text{for } \delta_h \text{ and } \varepsilon \text{ small.}$$

Moreover, since $U = U_h$ in this example, we can simply take the control $v_{\varepsilon,h}$ in the original problem ($\tilde{\mathcal{P}}$) to evaluate the accuracy of the method. A simple computation shows

$$\mathcal{L}_T(v_{\varepsilon,h}|1) = C_1(\lambda, T)\delta_h + C_2(\lambda, T)\varepsilon + O(\varepsilon^2 + \delta_h^2), \quad (4.3)$$

with

$$C_1(\lambda, T) = \frac{(-1 + 2\lambda T + e^{-2\lambda T})e^{-\lambda T}}{2\lambda(1 - e^{-2\lambda T})}, \quad C_2(\lambda, T) = \frac{2\lambda e^{-\lambda T}}{1 - e^{-2\lambda T}}.$$

Observe that the coefficients $C_1(\lambda, T)$ and $C_2(\lambda, T)$ are both positive. Thus, if δ_h is itself positive (recall that $\varepsilon > 0$), the two leading terms in (4.3) can not compensate each other and we deduce that taking ε of the same order as δ_h guarantees an optimal convergence rate.

Notice that, taking ε much smaller than $|\delta_h|$ (even $\varepsilon = 0$) is possible here but does not improve the accuracy of the approximation.

We want to emphasize that this very particular example is in fact quite relevant for understanding the case of general parabolic PDEs. Indeed, let us consider the finite difference approximation of a 1D heat equation on a uniform mesh (the mesh points are denoted by $(x_i)_{1 \leq i \leq N}$) with a control distributed on the whole domain (that is $\omega = \Omega =]0, 1[$). We can then reduce the system to the one of controlling each Fourier mode of the solution. The eigenfunctions of the 1D Laplace operator are $\phi_k(x) = \sin(k\pi x)$, $\lambda_k = k^2\pi^2$ but it is well known that the discrete Laplace operator in this case has the following eigenelements $\phi_{k,h} = (\phi_k(x_i))_i$, $\lambda_{k,h} = \frac{4 \sin^2(\frac{k\pi h}{2})}{h^2}$.

Therefore, for each value of k , we are exactly in the situation described above with

$$\delta_{k,h} = \lambda_{k,h} - \lambda_k \underset{h \rightarrow 0}{\sim} -\frac{k^4 \pi^4}{12} h^2.$$

The conclusion of this discussion is that the optimal choice should be $\varepsilon = \phi(h) \sim h^2$. This seems to be a contradiction with our general results (see Section 4.1) that rather lead to choose $\varepsilon = \phi(h) = h^4$ since the scheme is second order in space.

Actually, this very particular behavior is related to the choice of a non localised control domain $\omega = \Omega$. Indeed, in such a case, the controllability properties of the heat equation are much easier to prove and the following stronger observability inequality holds for the adjoint problem

$$\|q^F\|_{H^{-1}(\Omega)}^2 \leq C_{\text{obs}}^2 \int_0^T \|e^{(T-t)\Delta} q^F\|_{L^2(\Omega)}^2 dt, \quad \forall q^F \in L^2(\Omega).$$

This can be proved by using a Fourier decomposition of the solution. We conclude that, in such case, the HUM optimal adjoint state q_ε^F is bounded in $H^{-1}(\Omega)$ and by (1.7) we obtain the estimate

$$\|\mathcal{L}_T(v_\varepsilon|y_0)\|_{H^{-1}(\Omega)} \leq C\varepsilon.$$

It thus appears that the convergence rate towards zero of the target given by the HUM penalised approach, at least in the H^{-1} norm, is ε and not $\sqrt{\varepsilon}$ as in the general case. Such a behavior will be illustrated from a numerical point of view in Figure 6.

4.3. 1D scalar problems

We now present some results obtained in various situations with the methods described above. We use the standard finite-difference scheme on a uniform mesh of the domain $\Omega =]0, 1[$. We denote by N the number of points in the mesh. We either use the fully-discrete setting in the case $\theta = 1$ (that is the implicit Euler discretisation) and we denote by M the number of time intervals, or we use the semi-implicit discretisation, in which case we set $M = +\infty$. Note that results using the Crank-Nicolson scheme are given in [BHL11].

4.3.1. The heat equation

We consider here the following problem with a control time $T = 1$,

$$\begin{cases} \partial_t y - 0.1 \partial_x^2 y = \mathbf{1}_\omega v, \\ y(t, 0) = y(t, 1) = 0, \\ y_0(x) = \sin(\pi x)^{10}. \end{cases} \quad (4.4)$$

We start with the case of a localised control domain $\omega =]0.3, 0.8[$. In Figure 2, we show the time evolution of the uncontrolled solution as well as the one of the controlled solution. The control domain is represented as a darker zone on the plane (t, x) . Of course, we observe that the uncontrolled solution is damped with time but does not reach zero at time T , as opposed to the controlled solution.

(A) Case $\phi(h) = h^2$						(B) Case $\phi(h) = h^4$					
N	M					N	M				
	20	80	320	1280	$+\infty$		20	80	320	1280	$+\infty$
20	14	16	16	16	16	20	24	30	28	27	32
50	22	26	29	29	31	50	83	87	87	93	106
100	30	38	44	49	48	100	235	240	233	262	265
200	45	58	69	77	82	200	778	850	1098	1230	1374

TABLE 1. Conjugate gradient behavior for System (4.4); $\omega =]0.3, 0.8[$

In Table 1, we show the number of conjugate gradient iterations needed to achieve a (fixed) stopping criterion. Two cases are presented depending on the choice of the function $h \mapsto \phi(h)$ which serves as the penalisation term in the HUM functional. The results confirm the discussion of Section 4.1 that is:

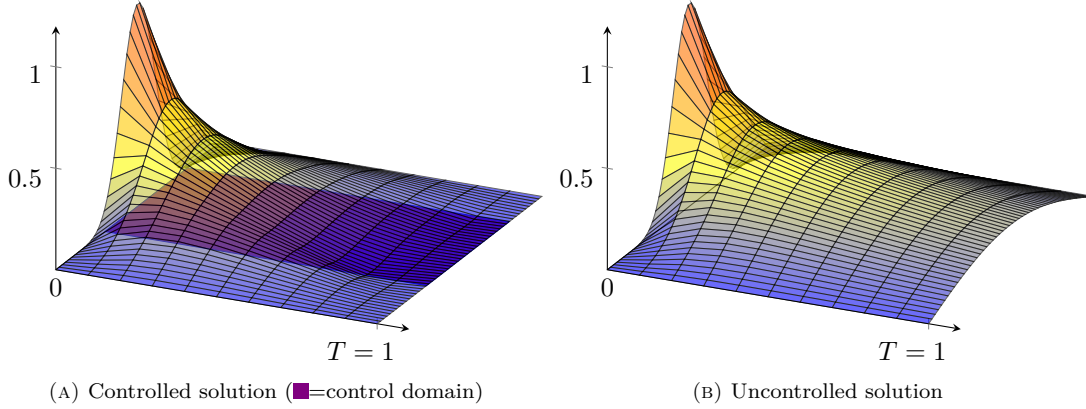


FIGURE 2. Time evolution of solutions of system (4.4); $\omega =]0.3, 0.8[$

- The converge speed of the conjugate gradient does not depend too much on the time step. Actually, the number of iterations required in the semi-discrete case seems to be an upper bound for the number of iterations of the fully-discrete case.
- For a given M , the number of iterations varies like $\phi(h)^{-1/2}$, with $h = 1/N$.

We observed the same kind of behavior in all the test cases below (except in Figure 6 which is very specific to the case $\omega = \Omega$).

In Table 2, we show the value (multiplied by 10^2) of the so-called optimal energy, that is the infimum of the penalised HUM functional $F_{\phi(h),h,\delta t}$. As predicted by theoretical results, we observe that this infimum is bounded as $(\delta t, h) \rightarrow (0, 0)$ which implies in particular that the $L^2_{\delta t}(0, T; U_h)$ -norm of the approximate control is bounded when refining the mesh.

More precisely, we can observe that for a given number of points N in the mesh, the value of the infimum decreases when M increases. This intuitively corresponds to the fact that, increasing M leads to a higher number of degrees of freedom for the approximate control, making the system easier to control.

(A) Case $\phi(h) = h^2$						(B) Case $\phi(h) = h^4$					
N	M					N	M				
	20	80	320	1280	$+\infty$		20	80	320	1280	$+\infty$
20	7.17	6.54	6.38	6.34	6.33	20	10.8	8.92	8.43	8.3	8.26
50	7.98	7.08	6.85	6.79	6.78	50	11.5	8.94	8.29	8.12	8.07
100	8.5	7.44	7.15	7.07	7.05	100	12.1	9.1	8.33	8.13	8.06
200	9.1	7.75	7.39	7.3	7.27	200	12.9	9.33	8.41	8.17	8.09

TABLE 2. Optimal energy (multiplied by 10^2) for System (4.4); $\omega =]0.3, 0.8[$

We begin now a discussion on the accuracy of the method. For a given mesh size h and time step δt , we are interested in

- The size of the computed target $\left\| \mathcal{L}_T^{h,\delta t}(v_{\phi(h),h,\delta t} | y_{0,h}) \right\|_h$,
- The size of the “exact solution” associated with the same computed control $v_{\phi(h),h,\delta t}$. To compute this quantity, we use a high order spectral method or a finite difference method with a much finer mesh to solve Problem (4.4) with $v = v_{\phi(h),h,\delta t}$ with a much higher accuracy.

In Figure 3 (resp. Figure 4) we plot those quantities as a function of h for various time steps in the case $\phi(h) = h^2$ (resp. $\phi(h) = h^4$).

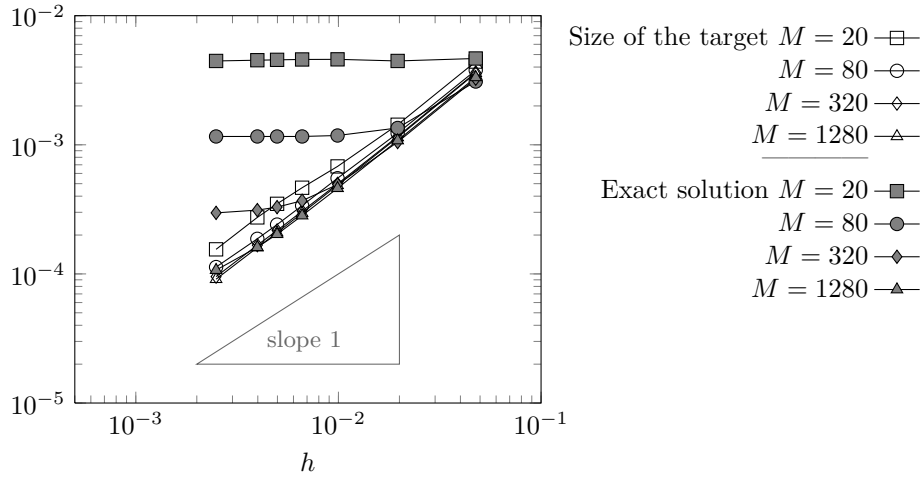


FIGURE 3. System (4.4) with $\phi(h) = h^2$; $\omega \in]0.3, 0.8[$

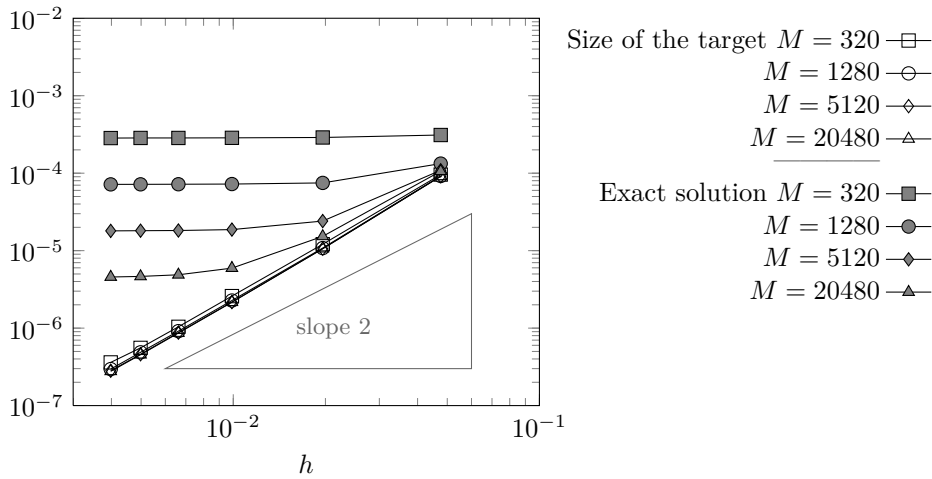


FIGURE 4. System (4.4) with $\phi(h) = h^4$; $\omega \in]0.3, 0.8[$

We observe that, as predicted by Theorem 3.2, the size of the computed target behaves like $\phi(h)^{-1/2}$ for any choice of the time step (the condition in Theorem 3.2 is satisfied here). Actually, all the plots with non-filled symbols are almost superimposed.

However, one has to be cautious with such a result. Indeed, for a large time step ($M = 20$), we can not hope to obtain a very precise approximation of the control since we are limited by the intrinsic accuracy of the finite-difference scheme itself. This is confirmed by the fact that, the size of the “exact solution” at time T (the plots with filled symbols), presents a plateau whose position corresponds to the first order accuracy in time of the approximation.

More importantly, we want to emphasize that choosing the penalisation parameter too small with respect to h and δt does not lead to more accurate results whereas it increases a lot the computational cost of the

overall method. Indeed, in Figure 5 we show the results obtained by choosing $\phi(h) = 1000h^6$ and changing the control domain to $\omega =]0.1, 0.9[$. As predicted by the theory, the size of the target which is computed by the numerical scheme behaves like $\phi(h)^{1/2} \sim Ch^3$. However, the size of the exact solution computed with the same approximate control does not tend to zero faster than h^2 which is the expected convergence rate of the scheme. We also observe the same plateau phenomenon as in the other cases, still due to the time accuracy of the scheme.

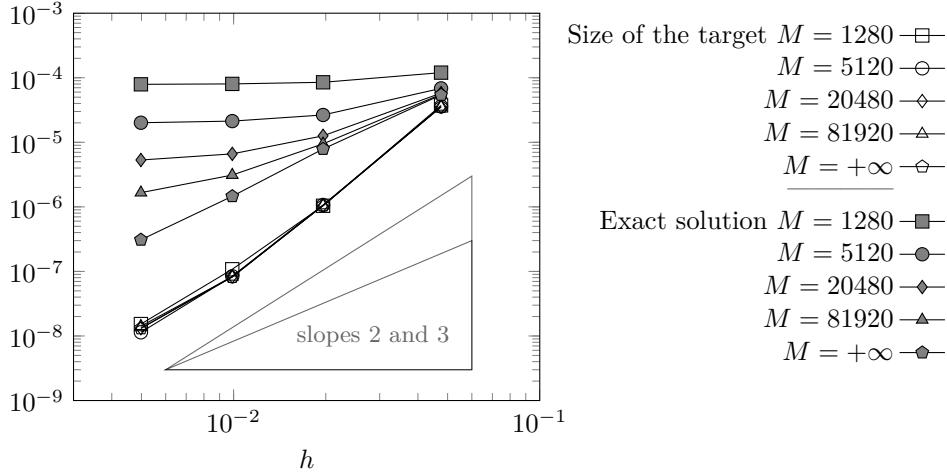


FIGURE 5. System (4.4) with $\phi(h) = 1000h^6$; $\omega =]0.3, 0.8[$

To conclude this section, we consider here the simpler case where $\omega = \Omega$ in order to illustrate the discussion of Section 4.2. We observe the predicted super-convergence phenomenon of the computed target whose size behaves like $\phi(h)$ instead of $\sqrt{\phi(h)}$. Observe that this appears to hold for the L^2 -norm even if the argument given in Section 4.2 only provides such an estimate for the H^{-1} -norm of the solution. However, we observe again that in any case the size of the exact solution at time T associated with the computed approximate control behaves like h^2 , which is the expected accuracy of the numerical scheme.

4.3.2. A semi-linear case

In this section, we are interested in the computation of an approximate null-control at time $T = 0.5$ for the following semi-linear problem (taken from [FCM12a])

$$\begin{cases} \partial_t y - 0.1\partial_x^2 y - 5y \log^{1.4}(1 + |y|) = \mathbf{1}_{]0.2, 0.8[} v, \\ y(t, 0) = y(t, 1) = 0, \\ y_0(x) = 20 \sin(\pi x). \end{cases} \tag{4.5}$$

Note that the initial data is chosen sufficiently large so as to observe a significant influence of the non-linear effects. Actually, with such a choice of y_0 , the uncontrolled solution blows up before the time $T = 0.5$. The controlled solution is shown on Figure 7a.

This solution is computed by a very simple fixed-point type algorithm with relaxation which consists in computing, at each iteration k , an approximate control for the linear problem

$$\begin{cases} \partial_t y^* - 0.1\partial_x^2 y^* - 5y^* \log^{1.4}(1 + |y^k|) = \mathbf{1}_{]0.2, 0.8[} v, \\ y^*(t, 0) = y^*(t, 1) = 0, \\ y^*(0, x) = y_0(x) = 20 \sin(\pi x), \end{cases} \tag{4.6}$$

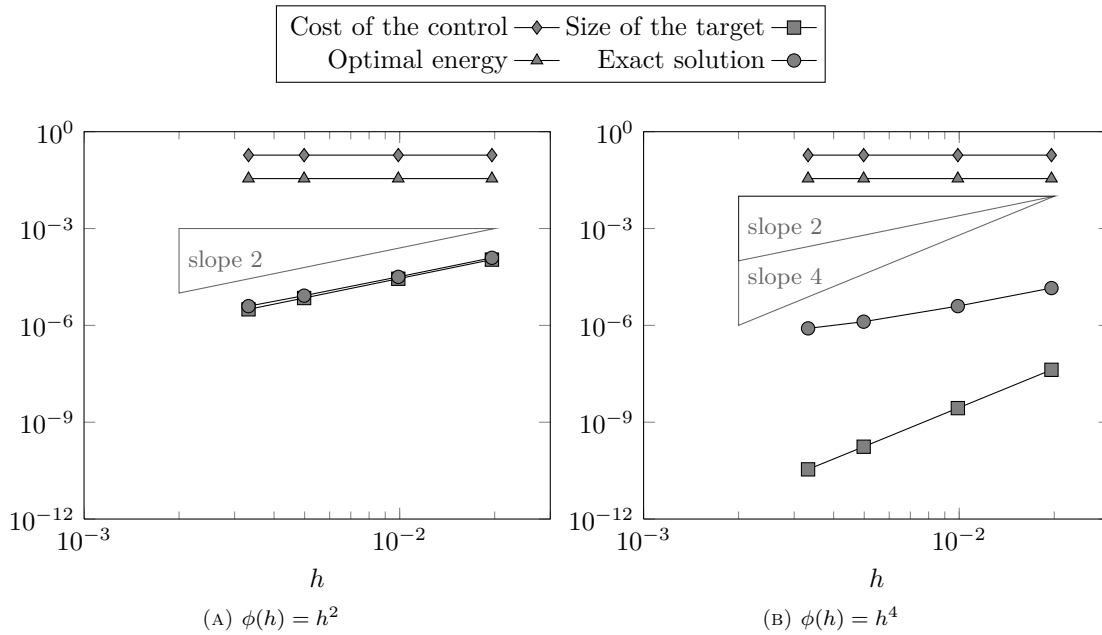


FIGURE 6. System (4.4) with a non-localised control $\omega = \Omega$; Semi-discrete scheme

and then, to set $y^{k+1} = \alpha y^* + (1 - \alpha)y^k$, for some relaxation parameter $\alpha \in]0, 1[$. The initial guess y^0 is for instance taken to be 0.

Note that we are not ensured that this fixed-point algorithm actually converges. Indeed, the theoretical argument for showing the existence of a null-control for this problem relies on a Schauder-type fixed-point argument (see [FCZ00]) and not on a contraction fixed-point argument. It appears that, in the test we performed below, we did not experienced any particular difficulty on this aspect, provided that a suitable relaxation parameter was choosed, namely $\alpha = 0.3$ here. Nevertheless, for other values of the parameters (for instance if one takes $\phi(h) = h^4$) we can experience convergence problems and then more subtle algorithms are needed (see for instance the algorithms discussed in [FCM12a]).

In Table 3, we observe that the $L_{\delta t}^2(0, T; U_h)$ -norm of the approximate control appears to be bounded when one refines the time step and the mesh size, as in the linear situation. We also observe in Figure 7b that the size of the computed target $\left\| \mathcal{L}_T^{h, \delta t}(v_{\phi(h), h, \delta t} | y_{0, h}) \right\|_h$ behaves like $\phi(h)^{1/2}$.

N	M			
	50	100	150	200
50	97.5	95.2	94.4	94.1
100	96.7	94.3	93.5	93.2
150	96.4	94.1	93.3	92.9
200	96.3	93.9	93.2	92.8

TABLE 3. Cost of the control $\left\| v_{\phi(h), h, \delta t} \right\|_{L_{\delta t}^2(0, T; U_h)}$ for System (4.5) with $\phi(h) = h^2$

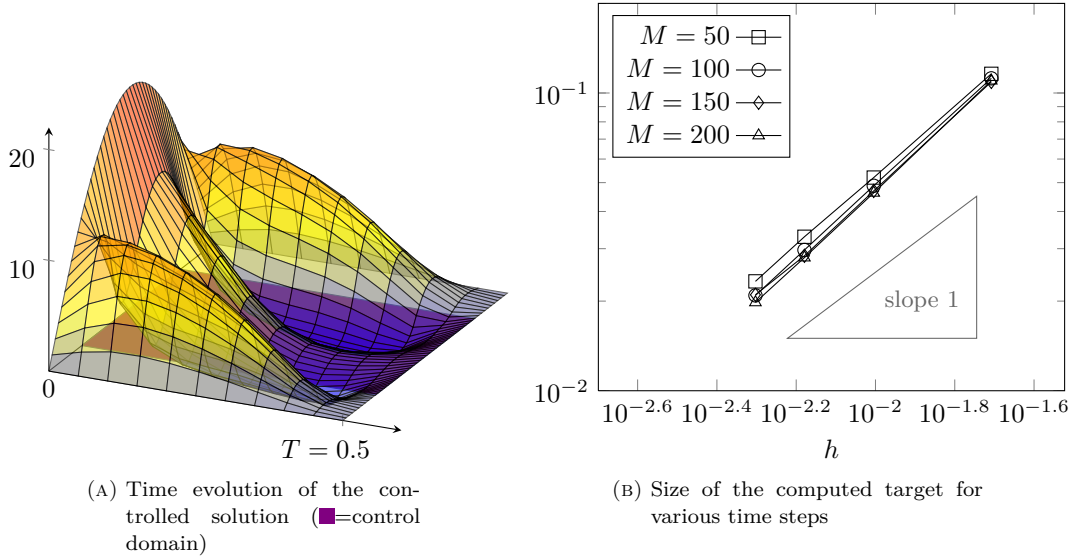


FIGURE 7. Computation of a null-control for Problem (4.5) with $\phi(h) = h^2$

4.4. 1D systems with one control force

In this section we illustrate the possibilities of the numerical methods discussed here in the case of more complex problems, in particular controlled parabolic systems characterized by a number of controls less than the number of equations in the system. This is a very important issue since the theoretical results available in this framework are far from being complete; we refer to the recent survey article [AKBGBdT11] and references therein. Numerical simulations may help to highlight expected or unexpected behaviors and then to go further in the understanding of controllability properties for such systems.

In this section, we systematically use the implicit Euler scheme with $M = 400$ time steps and the penalisation function is chosen to be $\phi(h) = h^4$.

4.4.1. A two equation cascade system

Let us consider the case of a two equation/one control cascade system

$$\partial_t y - 0.1 \partial_x^2 y + \underbrace{\begin{pmatrix} 0 & 0 \\ a_{21}(x) & 0 \end{pmatrix}}_{=A(x)} y = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{=B} \mathbf{1}_\omega v. \tag{4.7}$$

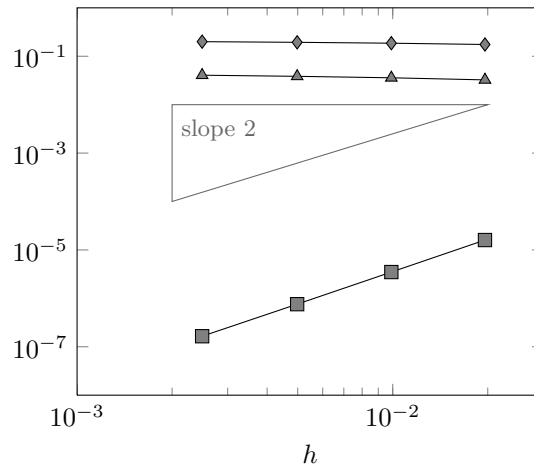
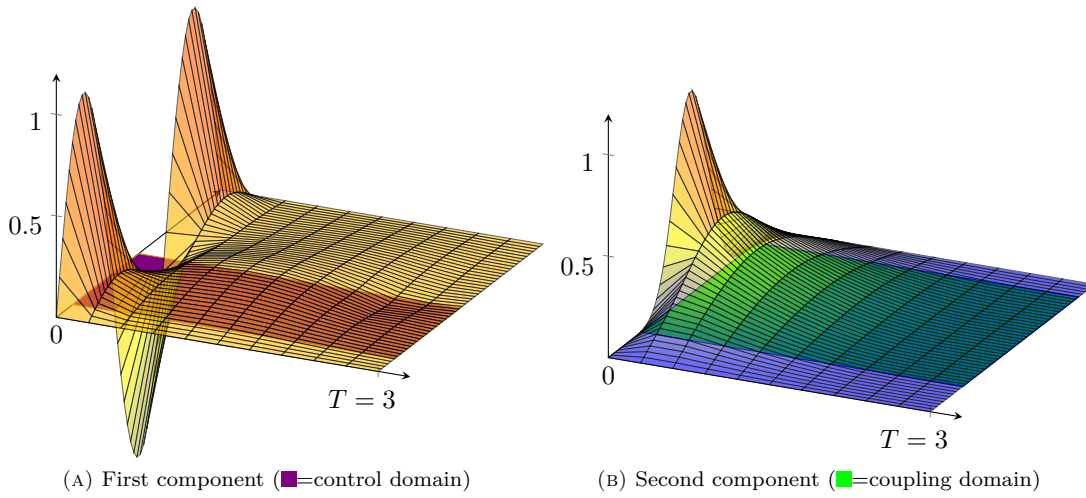
In such a system, the control v directly acts on the first component y_1 , whereas the second component y_2 is only controlled by the action of y_1 through the coupling matrix $A(x)$. In the case where $A(x)$ does not depend on x , it is shown in [AKBDGB09] that the null-controllability of this system is equivalent to the (algebraic) Kalman rank condition between A and B (in any dimension).

In the case of a variable coupling matrix, the situation is more intricate. Roughly speaking, if the support of the coupling term a_{21} intersects the control domain, then null-controllability can be proved (here also in any dimension) by means of Carleman estimates. If this geometric condition on the support is not fulfilled but if the sign of a_{21} is constant, then the same result can be achieved (in 1D) by means of the transmutation method (see [RdT11] and [ABL13] for some related results in any dimension). We recently gave in [BO13] a complete characterisation of the approximate controllability properties for (4.7) in the 1D case. As we will see

below, some unexpected behavior may occur when the sign of a_{21} is not constant. Let us describe the different simulations that we performed.

- Case 1 : $a_{21}(x) = \mathbf{1}_{\mathcal{O}}(x) \geq 0$ and $\mathcal{O} \cap \omega \neq \emptyset$ (see Figure 8)

$$\mathcal{O} =]0.2, 0.9[, \quad \omega =]0.1, 0.5[, \quad y_0(x) = (\sin(3\pi x), \sin(\pi x)^{10})^t, \quad T = 3.$$



(c) Same legend as in Figure 6

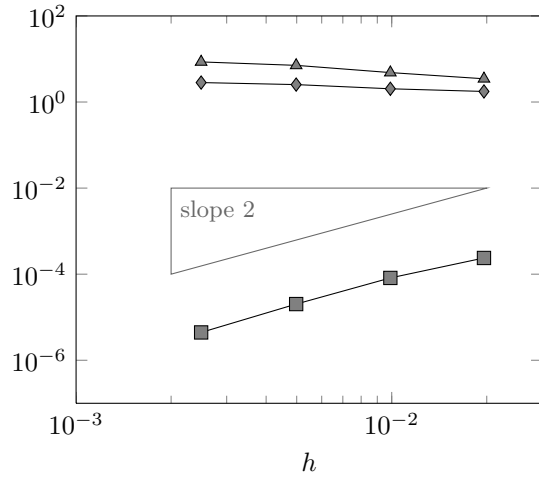
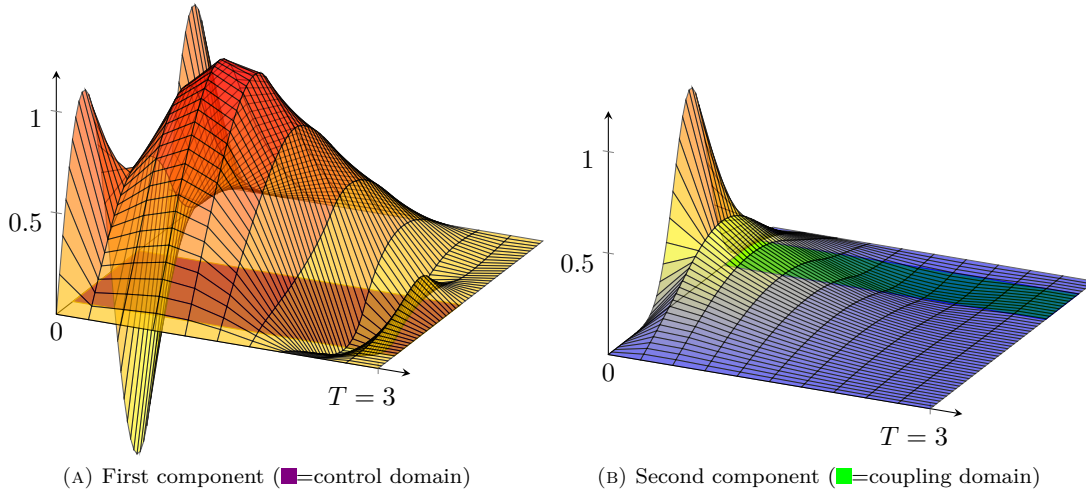
FIGURE 8. System (4.7). Case 1 : $\mathcal{O} \cap \omega \neq \emptyset$

- Case 2 : $a_{21}(x) = \mathbf{1}_{\mathcal{O}}(x) \geq 0$ and $\mathcal{O} \cap \omega = \emptyset$ (see Figure 9)

$$\mathcal{O} =]0.7, 0.9[, \quad \omega =]0.1, 0.5[, \quad y_0(x) = (\sin(3\pi x), \sin(\pi x)^{10})^t, \quad T = 3.$$

- Case 3 : $a_{21}(x) = (x - \alpha)\mathbf{1}_{\mathcal{O}}(x)$, with $\alpha \in \mathcal{O}$ and $\mathcal{O} \cap \omega = \emptyset$ (see Figure 10)

$$\mathcal{O} =]0, 0.5[, \quad \alpha \in \{1/4, 1/8\}, \quad \omega =]0.5, 1[, \quad y_0(x) = (\sin(2\pi x), 3 \sin(2\pi x))^t, \quad T = 2.$$

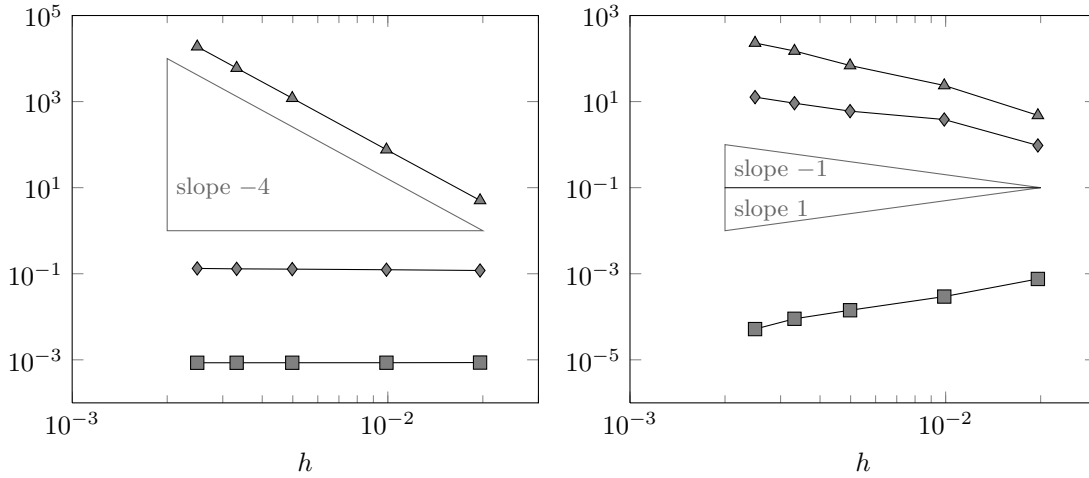


(c) Same legend as in Figure 6

FIGURE 9. System (4.7). Case 2 : $\mathcal{O} \cap \omega = \emptyset$

In cases 1 and 2, for which we have theoretical results proving that null-controllability holds, we observe that h^4 -null controllability of the discrete system seems to hold in both cases. Indeed, the cost of the control and the optimal energy are bounded as $h \rightarrow 0$ whereas the computed target goes to 0 as h^2 . In the case 1, the solution is damped quite rapidly since the coupling is strong (because it is supported on the control domain). By comparison, we observe that the evolution of y_1 is much more complex in case 2 since the coupling between the two components is much weaker.

In case 3, we illustrate the fact that the system is approximately controllable for $\alpha = 1/8$ but is not approximately controllable for $\alpha = 1/4$ (see [BO13]). Observe in Figure 10b that, when $\alpha = 1/8$, the size of the target actually goes to 0 but its convergence rate is only h and not h^2 . Similarly, the optimal energy and the control cost seems to behave like h^{-1} . We do not have any explanation yet for such a behavior, but two conjectures can possibly be made: either the continuous problem is approximately- but not null-controllable (this could even

(A) $\alpha = 1/4$. Same legend as in Figure 6(B) $\alpha = 1/8$. Same legend as in Figure 6FIGURE 10. System (4.7). Case 3 : $\mathcal{O} \cap \omega = \emptyset$ and a_{21} changes its sign

depend on the control time T) or the discrete problem is only $\phi(h)$ -null controllable for functions ϕ going to zero sufficiently slowly.

4.4.2. A three-equation cascade system

In this section, we consider the following cascade system

$$\partial_t y - 0.1 \partial_x^2 y + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ a_{21}(x) & 0 & 0 \\ 0 & a_{32}(x) & 0 \end{pmatrix}}_{\stackrel{\text{def}}{=} A(x)} y = \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\stackrel{\text{def}}{=} B} \mathbf{1}_\omega(x) v. \quad (4.8)$$

As in the previous system, null-controllability is known to hold as soon as the support of a_{21} and a_{32} do intersect the control domain ω with some additional sign conditions. We illustrate in Figure 11 the various behaviors described in [BO13]. More precisely, we choose $T = 1$, $y_0(x) = (-1, 0, \mathbf{1}_{]0,0.5[}(x))^t$ and we consider the following two cases

- Case 1 :

$$a_{21} = \mathbf{1}_{\mathcal{O}_2}, \quad \mathcal{O}_2 =]0, 0.5[, \quad a_{32} = 1, \quad \omega =]0.5, 1[.$$

Here the system is known to be approximately controllable and what we observe in Figure 11a is similar to the situation that we already encountered in Figure 10b: the target tends to zero like $h^{0.5}$ instead of $h^2 = \phi(h)^{1/2}$ that one could have expected. Here also the problem may not be null-controllable or the numerical approximation may require stronger conditions on the penalisation function ϕ .

- Case 2 :

$$a_{21} = \mathbf{1}_{\mathcal{O}_2}, \quad \mathcal{O}_2 =]0, 0.5[, \quad a_{32}(x) = (x - 0.5), \quad \omega =]0.5, 1[.$$

Here, it can be proved that the system is not approximately controllable and we actually observe that the size of the computed target does not tend to 0 when $h \rightarrow 0$. Observe that the support of a_{32} intersects the control domain and that its sign is constant in ω . The non-controllability is related here to the change of sign of this coefficient outside ω .

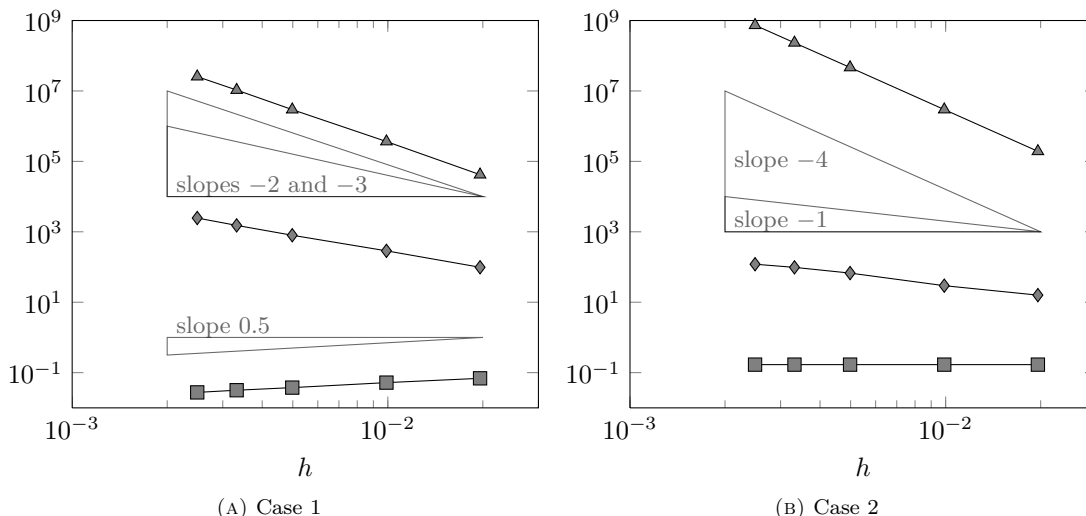


FIGURE 11. System (4.8). Same legend as in Figure 6

4.4.3. A three equation system with one single control

In this section, we consider the following system of three equations controlled by one single control

$$\partial_t y - 0.1 \partial_x^2 y + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ a_{21}(x) & 0 & 0 \\ a_{31}(x) & 0 & 0 \end{pmatrix}}_{\stackrel{\text{def}}{=} A(x)} y = \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\stackrel{\text{def}}{=} B} \mathbf{1}_\omega(x) v. \tag{4.9}$$

Observe that this is not a cascade system; in particular for any $x \in \Omega$, the Kalman rank condition between $A(x)$ and B is not satisfied. As a consequence, the results obtained for instance in [Mau13] do not apply to this system.

- In this system, the distributed control v only acts on the first component y_1 and this component serves as a simultaneous control for the second and third one through the coupling terms a_{21} and a_{31} .
- If one simply consider the first two equations (resp. equations 1 and 3), one faces a two-equation cascade system as in the previous section and theoretical results are known as discussed before. It follows that, in many cases, a control can be proved to exists that drive y_1 and y_2 (resp. y_1 and y_3) to zero at time T .

The more difficult problem that we want to deal with here is to find a single control v that drives simultaneously the three components to 0 at the same time T .

There is at least one case where the controllability for this system clearly fails. Indeed, assume that a_{21} and a_{31} are proportional, it is then clear that the difference $\alpha_2 y_2 - \alpha_3 y_3$ for a suitable choice of the constant coefficients α_2, α_3 simply satisfies a heat equation without any influence of v and y_1 and therefore we can not hope to drive this quantity to zero at time T .

We shall see below, in our numerical experiments, that the situation can be very different if a_{21} and a_{31} are linearly independent.

In all the following tests, we choose $T = 3$ and $y_0(x) = (0, \sin(\pi x)^{10}, 0)^t$.

Let us first consider the case where a_{21} and a_{31} are piecewise constant functions. More precisely, we choose $a_{21} = \mathbf{1}_{\mathcal{O}_2}$ and $a_{31} = \mathbf{1}_{\mathcal{O}_3}$ and we will investigate various geometric configurations for $\mathcal{O}_2, \mathcal{O}_3$ and ω .

For a first series of cases as shown in Figure 12, we proved in [BO13] that the system is approximately controllable. This is exactly what we observe in numerical results since the optimal energy and the control cost remain bounded and that the size of the target goes to zero as h^2 . Those results let us hope that null-controllability should also hold for such system but, to the author's knowledge, this is still an open problem.

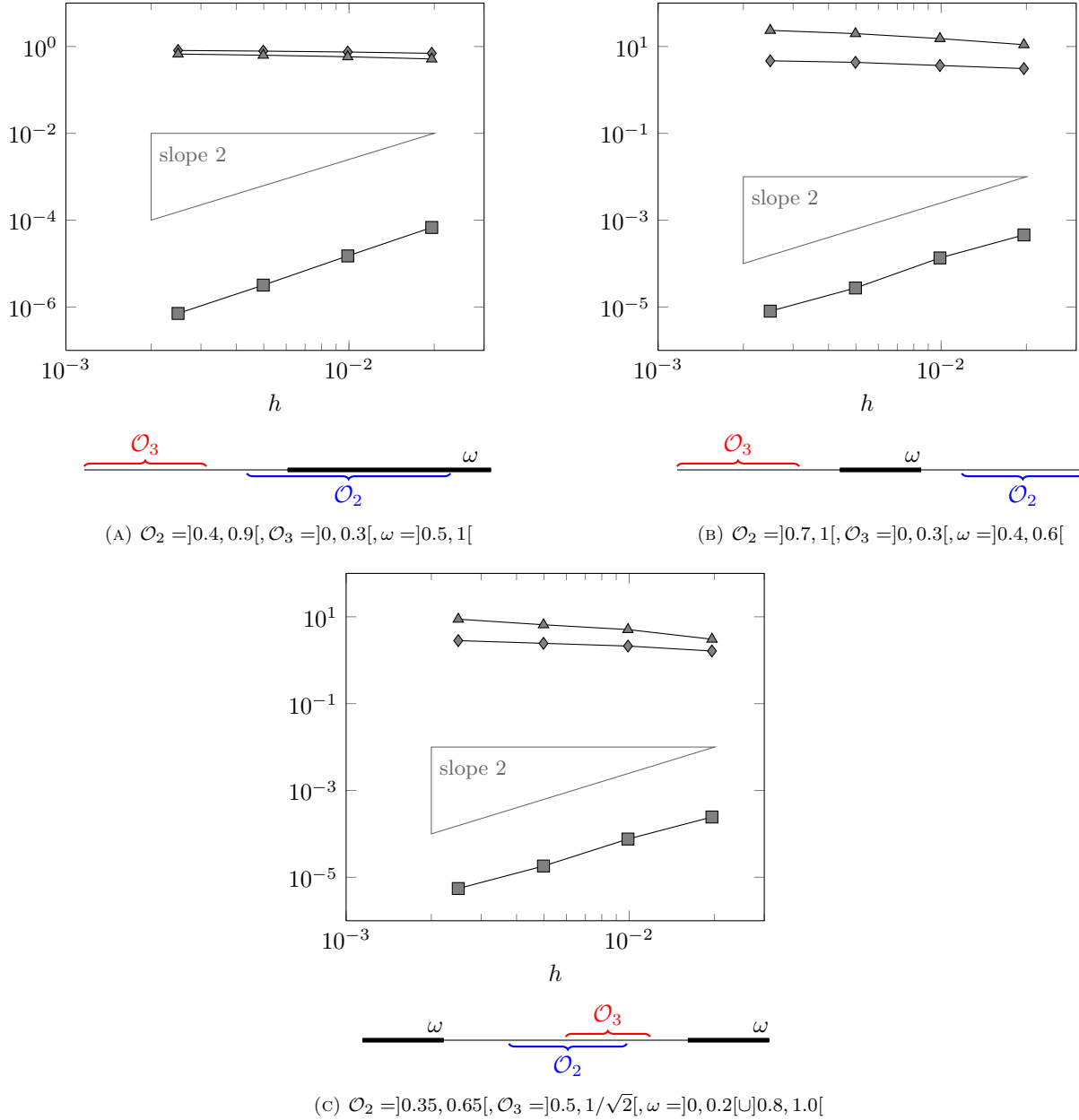
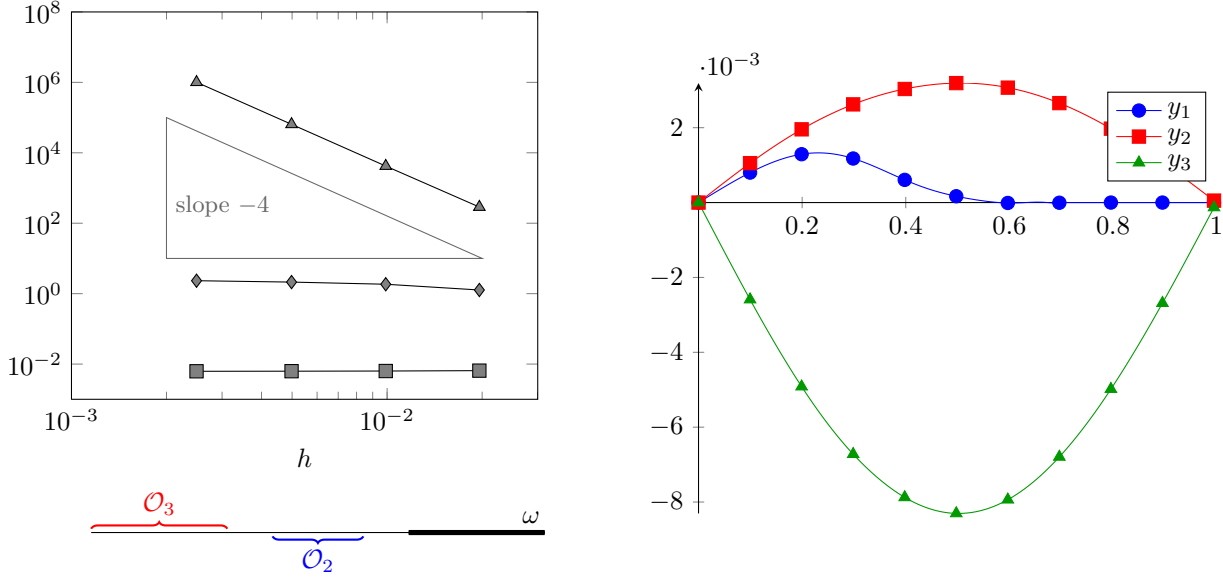


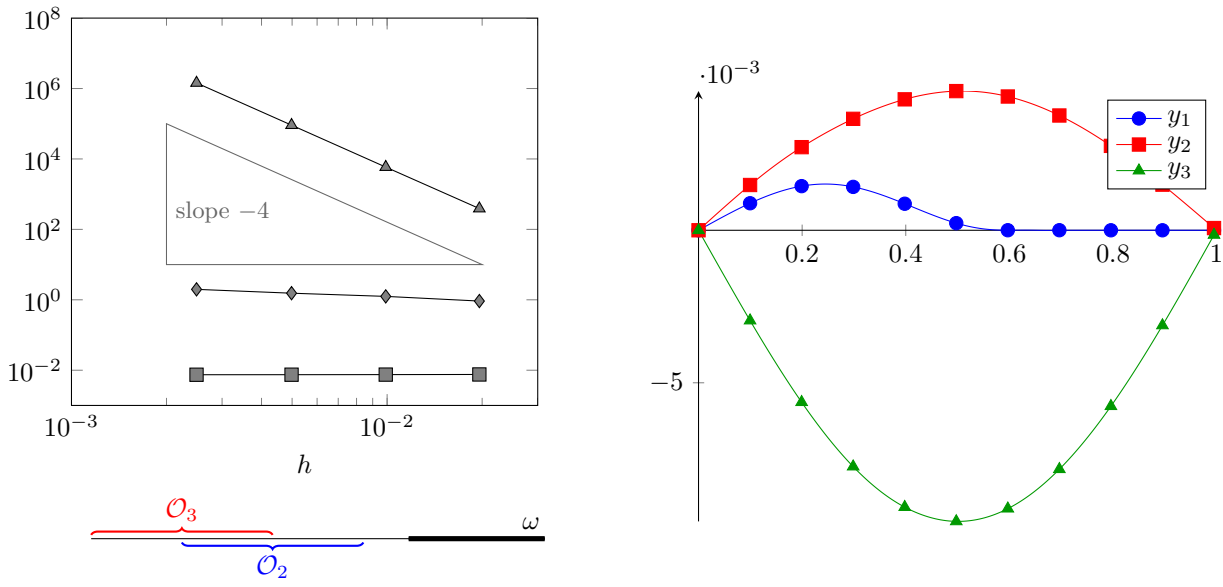
FIGURE 12. System (4.9). Controllable cases (same legend as in Figure 6)

Conversely, in the cases shown in Figure 13, we are able to prove that the system is not approximately controllable in general. Actually, we observe in numerical results that the optimal energy blows up like $h^{-4} =$

$\phi(h)^{-1}$, that the control cost remains bounded and that the size of the target does not tend to zero. We conclude that the controllability fails for this system in such cases. Note that, we observed that the computed target converges when $h \rightarrow 0$ towards some function, that we plot in the right-hand side of each figure. This function can be actually shown to be an element of the the space Q_F that we introduced in (1.20), see [BO13].



(A) $\mathcal{O}_2 =]0.4, 0.6[$, $\mathcal{O}_3 =]0, 0.3[$, $\omega =]0.7, 1[$



(B) $\mathcal{O}_2 =]0.2, 0.6[$, $\mathcal{O}_3 =]0, 0.4[$, $\omega =]0.7, 1[$

FIGURE 13. System (4.9). Non-controllable cases (same legend as in Figure 6). To be continued...

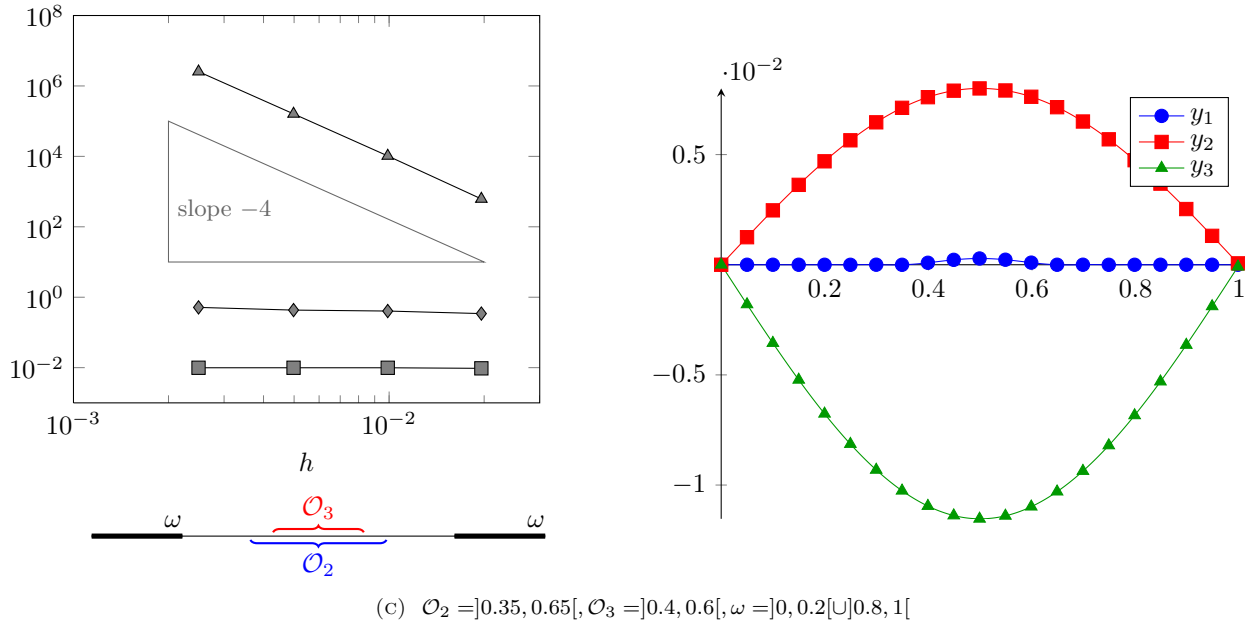


FIGURE 13. ... con't

We conclude this study by considering the case where a_{21} and a_{31} are not characteristic functions but, for instance, supported in the same coupling domain \mathcal{O} . More precisely, we choose $a_{21}(x) = x\mathbf{1}_{\mathcal{O}}(x)$ and $a_{31}(x) = (x+1)\mathbf{1}_{\mathcal{O}}(x)$. These examples are inspired from [BCGdT13] where the null-controllability of such systems is proved under some specific assumptions on the coefficients, and on the control domain.

In figure 14, we show the results obtained in the case where \mathcal{O} has a non empty intersection with the control domain. The behavior of the system seems to indicate that the system is controllable.

In Figure 15, we choose now a coupling domain \mathcal{O} which does not intersect the control domain. In this case, we observe that the system is not controllable and we show the computed target that we obtained on the right-hand side of the figure.

The author would like to warmly thank J. Le Rousseau and G. Olive for their useful comments on the material contained in this paper and the organizers of the *41ème Congrès d'Analyse Numérique* for their invitation to give a talk and to write the present article.

REFERENCES

- [ABL13] F. Alabau-Boussouira and M. Léautaud, *Indirect controllability of locally coupled wave-type systems and applications*, J. Math. Pures Appl. **to appear** (2013).
- [AKBDGB09] F. Ammar-Khodja, A. Benabdallah, C. Dupaix, and M. González-Burgos, *A Kalman rank condition for the localized distributed controllability of a class of linear parabolic systems*, J. Evol. Equ. **9** (2009), no. 2, 267–291.
- [AKBGBdT11] F. Ammar-Khodja, A. Benabdallah, M. González-Burgos, and L. de Teresa, *Recent results on the controllability of linear coupled parabolic problems: a survey*, Math. Control Relat. Fields **1** (2011), no. 3, 267–306.
- [BCGdT13] A. Benabdallah, M. Cristofol, P. Gaitan, and L. de Teresa, *Controllability to trajectories for some parabolic systems of three and two equations by one control force*, Mathematical Control and Related Fields **to appear** (2013).
- [BDPDM93] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and control of infinite-dimensional systems. Vol. II*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 1993.
- [BHL10a] F. Boyer, F. Hubert, and J. Le Rousseau, *Discrete Carleman estimates and uniform controllability of semi-discrete parabolic equations*, J. Math. Pures Appl. **93** (2010), 240–276.

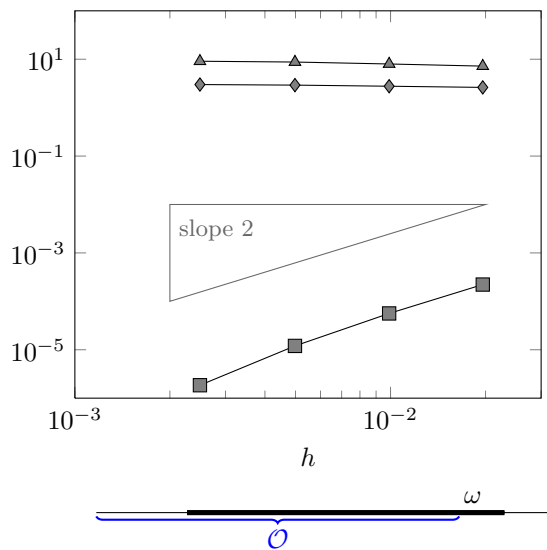


FIGURE 14. System (4.9). Variable coupling coefficients, $\mathcal{O} =]0, 0.8[$, $\omega =]0.2, 0.9[$
 $a_{21}(x) = x\mathbf{1}_{\mathcal{O}}(x)$, $a_{31}(x) = (x + 1)\mathbf{1}_{\mathcal{O}}(x)$.

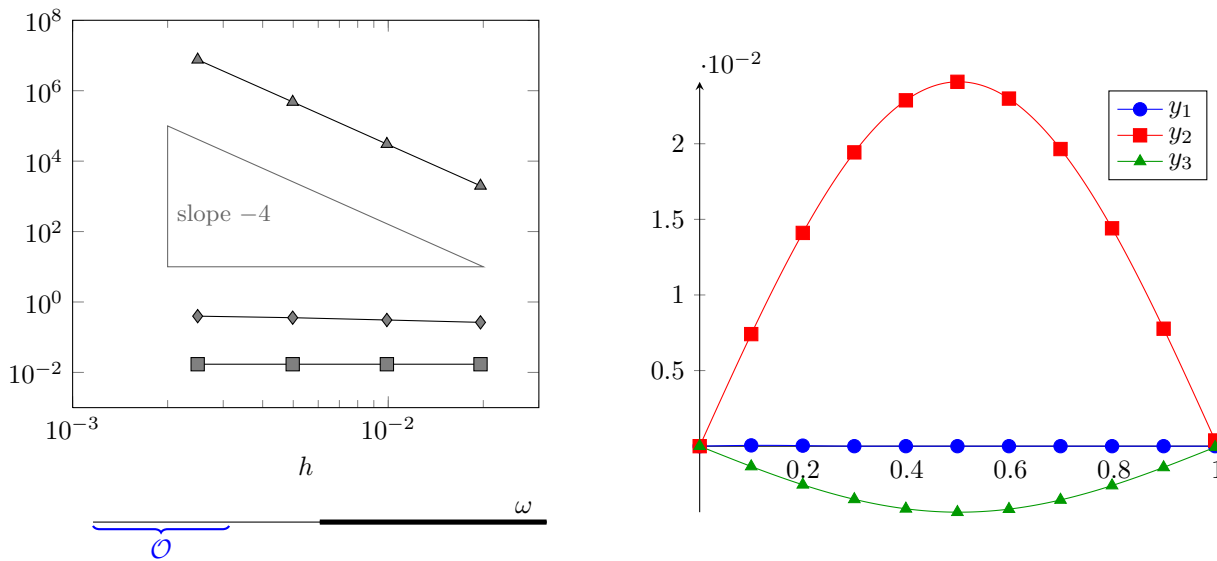


FIGURE 15. System (4.9). Variable coupling coefficients, $\mathcal{O} =]0, 0.2[$, $\omega =]0.5, 1[$
 $a_{21}(x) = x\mathbf{1}_{\mathcal{O}}(x)$, $a_{31}(x) = (x + 1)\mathbf{1}_{\mathcal{O}}(x)$.

[BHL10b] ———, *Discrete Carleman estimates for elliptic operators in arbitrary dimension and applications*, SIAM J. Control Optim. **48** (2010), 5357–5397.
 [BHL11] ———, *Uniform null-controllability properties for space/time-discretized parabolic equations*, Numer. Math, **118** (2011), 601–661.
 [BL13] F. Boyer and J. Le Rousseau, *Carleman estimates for semi-discrete parabolic operators and application to the controllability of semi-linear semi-discrete parabolic equations*, to appear in Annales de l’IHP - Analyse non Linéaire (2013), <http://dx.doi.org/10.1016/j.anihpc.2013.07.011>.

- [BO13] F. Boyer and G. Olive, *Approximate controllability conditions for some linear 1D parabolic systems with space-dependent coefficients*, preprint (2013), <http://hal.archives-ouvertes.fr/hal-00848709>.
- [Cor07] J.-M. Coron, *Control and nonlinearity*, Mathematical Surveys and Monographs, vol. 136, American Mathematical Society, Providence, RI, 2007.
- [CT06] Y. Chitour and E. Trélat, *Controllability of partial differential equations*, Advanced topics in control systems theory, Lecture Notes in Control and Inform. Sci., vol. 328, Springer, London, 2006, pp. 171–198.
- [ET99] I. Ekeland and R. Témam, *Convex analysis and variational problems*, english ed., Classics in Applied Mathematics, vol. 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999, Translated from the French.
- [FCM12a] E. Fernández-Cara and A. Münch, *Numerical null controllability of semi-linear 1D heat equations : Fixed point, Least squares and Newton methods*, Mathematical Control and Related Fields **2** (2012), no. 3, 217–246.
- [FCM12b] ———, *Numerical null controllability of the 1D heat equation: Carleman weights and duality*, preprint (2012).
- [FCM13] E. Fernández-Cara and A. Münch, *Strong convergent approximations of null controls for the 1D heat equation*, SĕMA J. **61** (2013), 49–78.
- [FCZ00] E. Fernández-Cara and E. Zuazua, *Null and approximate controllability for weakly blowing up semilinear heat equations*, Ann. Inst. H. Poincaré, Analyse non lin. **17** (2000), 583–616.
- [FI96] A. Fursikov and O. Yu. Imanuvilov, *Controllability of evolution equations*, vol. 34, Seoul National University, Korea, 1996, Lecture notes.
- [GL94] R. Glowinski and J.-L. Lions, *Exact and approximate controllability for distributed parameter systems*, Acta Numer. (1994), 269–378.
- [GLH08] R. Glowinski, J.-L. Lions, and J. He, *Exact and approximate controllability for distributed parameter systems*, Encyclopedia of Mathematics and its Applications, vol. 117, Cambridge University Press, Cambridge, 2008, A numerical approach.
- [Le 07] J. Le Rousseau, *Représentation Microlocale de Solutions de Systèmes Hyperboliques, Application à l’Imagerie, et Contributions au Contrôle et aux Problèmes Inverses pour des équations Paraboliques*, Mémoire d’habilitation à diriger des recherches, Universités d’Aix-Marseille, Université de Provence, 2007, <http://tel.archives-ouvertes.fr/tel-00201887/fr/>.
- [LR95] G. Lebeau and L. Robbiano, *Contrôle exact de l’équation de la chaleur*, Comm. Partial Differential Equations **20** (1995), 335–356.
- [LT00] I. Lasiecka and R. Triggiani, *Control theory for partial differential equations: continuous and approximation theories. I*, Encyclopedia of Mathematics and its Applications, vol. 74, Cambridge University Press, Cambridge, 2000, Abstract parabolic systems.
- [LT06] S. Labbé and E. Trélat, *Uniform controllability of semidiscrete approximations of parabolic control systems*, Systems Control Lett. **55** (2006), 597–609.
- [LZ98] A. Lopez and E. Zuazua, *Some new results to the null controllability of the 1-d heat equation*, Séminaire sur les Équations aux Dérivées Partielles, 1997–1998, Exp. No. VIII, 22 pp., École Polytech., Palaiseau (1998).
- [Mau13] K. Mauffrey, *On the null controllability of a parabolic system with non-constant coefficients by one or two control forces*, Journal de Mathématiques Pures et Appliquées **99** (2013), no. 2, 187 – 210.
- [MP11] A. Münch and P. Pedregal, *Numerical null controllability of the heat equation through a variational approach.*, preprint (2011).
- [MZ10] A. Münch and E. Zuazua, *Numerical approximation of null controls for the heat equation: ill-posedness and remedies*, Inverse Problems **26** (2010), no. 8, 085018, 39.
- [Ped10] P. Pedregal, *A variational perspective on controllability*, Inverse Problems **26** (2010), no. 1, 015004, 17.
- [RdT11] L. Rosier and L. de Teresa, *Exact controllability of a cascade system of conservative equations*, C. R. Math. Acad. Sci. Paris **349** (2011), no. 5-6, 291–296.
- [Zua06] E. Zuazua, *Control and numerical approximation of the wave and heat equations*, International Congress of Mathematicians, Madrid, Spain **III** (2006), 1389–1417.