A mixed formulation for the direct approximation of L^2-weighted controls for the linear heat equation

Arnaud Munch, Souza Diego A.

A mixed formulation for the direct approximation of L^2-weighted controls for the linear heat equation

Arnaud Münch* Diego A. Souza†

June 3, 2014

Abstract

This paper deals with the numerical computation of null controls for the linear heat equation. The goal is to compute approximations of controls that drive the solution from a prescribed initial state to zero at a given positive time. In [Fernandez-Cara & Münch, Strong convergence approximations of null controls for the 1D heat equation, 2013], a so-called primal method is described leading to a strongly convergent approximation of distributed control: the controls minimize quadratic weighted functionals involving both the control and the state and are obtained by solving the corresponding optimality conditions. In this work, we adapt the method to approximate the control of minimal square integrable-weighted norm. The optimality conditions of the problem are reformulated as a mixed formulation involving both the state and its adjoint. We prove the well-posedeness of the mixed formulation (in particular the inf-sup condition) then discuss several numerical experiments. The approach covers both the boundary and the inner situation and is valid in any dimension.

Keywords: Linear heat equation; Null controllability; Finite element methods; Mixed formulation.

Mathematics Subject Classification (2010)- 35K35, 65M12, 93B40, 65K10

Contents

*Laboratoire de Mathématiques, Université Blaise Pascal (Clermont-Ferrand 2), UMR CNRS 6620, Campus de Cézeaux, 63177, Aubière, France. E-mails: arnaud.munch@math.univ-bpclermont.fr.

†Dpto. EDAN, University of Sevilla, 41080 Sevilla, Spain and Departamento de Matemática, Universidade Federal da Paraíba, 58051-900, João Pessoa–PB, Brazil. E-mail: desouza@us.es. Partially supported by CAPES (Brazil) and grant MTM2010-15592 (DGI-MICINN, Spain).

1 Introduction. The null controllability problem

Let $\Omega \subset \mathbb{R}^N$ be a bounded connected open set whose boundary $\partial\Omega$ is regular enough (for instance of class C^2). Let $\omega \subset \Omega$ be a (small) nonempty open subset and assume that $T > 0$. In the sequel, for any $\tau > 0$ we denote by Q_τ, q_τ and Σ_τ the sets $\Omega \times (0,\tau)$, $\omega \times (0,\tau)$ and $\partial\Omega \times (0,\tau)$, respectively.

This work is concerned with the null controllability problem for the heat equation

$$\begin{cases} y_t - \nabla \cdot (c(x)\nabla y) + d(x,t)y = v\,1_\omega, & \text{in} \quad Q_T, \\ y = 0, & \text{on} \quad \Sigma_T, \\ y(x,0) = y_0(x), & \text{in} \quad \Omega. \end{cases} \tag{1}$$

Here, we assume that $c := (c_{i,j}) \in C^1(\overline{\Omega}; \mathcal{M}_N(\mathbb{R}))$ with $(c(x)\xi,\xi) \geq c_0|\xi|^2$ in $\overline{\Omega}$ ($c_0 > 0$), $d \in L^\infty(Q_T)$ and $y_0 \in L^2(\Omega)$; $v = v(x,t)$ is the *control* (a function in $L^2(q_T)$) and $y = y(x,t)$ is the associated state. Moreover, 1_ω is the characteristic function associated to the set ω.

In the sequel, we shall use the following notation :

$$L\,y := y_t - \nabla \cdot (c(x)\nabla y) + d(x,t)y, \qquad L^\star \varphi := -\varphi_t - \nabla \cdot (c(x)\nabla \varphi) + d(x,t)\varphi.$$

For any $y_0 \in L^2(\Omega)$ and $v \in L^2(q_T)$, there exists exactly one solution y to (1), with the regularity $y \in C^0([0,T]; L^2(\Omega)) \cap L^2(0,T; H^1_0(\Omega))$ (see [30, 7]). Accordingly, for any final time $T > 0$, the associated null controllability problem at time T is the following : for each $y_0 \in L^2(\Omega)$, find $v \in L^2(q_T)$ such that the corresponding solution to (1) satisfies

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

The controllability of PDEs is an important area of research and has been the subject of many papers in recent years. Some relevant references are [27, 29, 35] and [12]. In particular, we refer to [21] and [28] where the null controllability of (1) is proved.

The numerical approximation is also a fundamental issue, since it is not in general possible to get explicit expression of controls. Due to the strong regularization property of the heat kernel, numerical approximation of controls is a rather delicate issue. The same holds in inverse problems theory when parabolic equations and systems are involved (see [15]). This have been exhibited numerically in [5] who made use of duality argument and focused on the control of minimal square integrable norm: the problem reads

$$\begin{cases} \text{Minimize} \quad J_1(y,v) := \dfrac{1}{2} \displaystyle\iint_{q_T} |v(x,t)|^2\,dx\,dt \\ \text{Subject to} \quad (y,v) \in \mathcal{C}(y_0,T) \end{cases} \tag{3}$$

where $\mathcal{C}(y_0;T)$ denotes the linear manifold

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

The earlier contribution is due to Glowinski and Lions in [23] (updated in [24]) and relies on duality arguments. Duality allows to replace the original constrained minimization problem by an unconstrained and *a priori* easier minimization (dual) problem. The dual problem associated with (3) is :

$$\min_{\varphi_T\in\mathcal{H}} J_1^\star(\varphi_T) := \frac{1}{2}\iint_{q_T} |\varphi(x,t)|^2\,dxdt + \int_\Omega y_0(x)\varphi(x,0)dx \tag{4}$$

where the variable φ solves the backward heat equation :

$$L^\star\varphi = 0 \quad \text{in} \quad Q_T, \qquad \varphi = 0 \quad \text{on} \quad \Sigma_T; \quad \varphi(\cdot,T) = \varphi_T \quad \text{in} \quad \Omega, \tag{5}$$

and the Hilbert space $\mathcal{H}$ is defined as the completion of $\mathcal{D}(\Omega)$ with respect to the norm $\|\varphi_T\|_{\mathcal{H}} := \|\varphi\|_{L^2(q_T)}$. In view of the unique continuation property to (5), the mapping $\varphi_T \mapsto \|\varphi_T\|_{\mathcal{H}}$ is a Hilbertian norm in $\mathcal{D}(\Omega)$. Hence, we can certainly consider the completion of $\mathcal{D}(\Omega)$ for this norm. The coercivity of the functional $J_1^\star$ in $\mathcal{H}$ is a consequence of the so-called *observability inequality*

$$\|\varphi(\cdot,0)\|^2_{L^2(\Omega)} \le C\iint_{q_T} |\varphi(x,t)|^2\,dx\,dt \quad \forall\varphi_T\in\mathcal{H}, \tag{6}$$

where φ solves (5). This inequality holds for some constant $C = C(\omega,T)$ and, in turn, is a consequence of some appropriate global Carleman inequalities; see [21]. The minimization of $J_1^\star$ is numerically ill-posed, essentially because of the hugeness of the completed space $\mathcal{H}$. The control of minimal square integrable norm highly oscillates near the final time T, property which is hard to capture numerically. We refer to [1, 25, 31, 34] where this phenomenon is highlighted under several perspectives.

Moreover, at the level of the approximation, the minimization of $J_1^\star$ requires to find a finite dimensional and conformal approximation of $\mathcal{H}$ such that the corresponding discrete adjoint solution satisfies (5), which is in general impossible for polynomial piecewise approximations. In practice, the trick initially described in [23], consists first to introduce a discrete and consistent approximation of (1) and then to minimize the corresponding discrete conjugate functional. However, this requires to get some uniform discrete observability inequalities which is a delicate issue, strongly depend on the approximations used (we refer to [3, 16, 36] and the references therein) and is still open in the general case of the heat equation with non constant coefficients. This fact and the hugeness of $\mathcal{H}$ has raised many authors to relax the controllability problem: precisely, the constraint (2). We mention the references [3, 5, 36] and notably [2, 19, 26] for some numerical realizations.

In [18] (see also [17] in a semi-linear case), a different - so-called primal approach - allowing more general results has been used and consists to solve directly optimality conditions : specifically, the following general extremal problem (initially introduced by Fursikov and Imanuvilov in [21]) is considered :

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

The weights $\rho = \rho(x,t)$ and $\rho_0 = \rho_0(x,t)$ are continuous, uniformly positive and are assumed to belong to $L^\infty(Q_{T-\delta})$ for any $\delta > 0$ (hence, they can blow up as $t \to T^-$). Under those conditions, the extremal problem (7) is well-posed (see [18]).

Moreover, the explicit occurrence of the term y in the functional allow to solve directly the optimality conditions associated with (7): defining the Hilbert space P as the completion of the linear space $P_0 = \{q \in C^\infty(\overline{Q_T}) : q = 0 \text{ on } \Sigma_T\}$ with respect to the scalar product

$$(p,q)_P := \iint_{Q_T} \rho^{-2} L^\star p \, L^\star q \, dx \, dt + \iint_{q_T} \rho_0^{-2} \, p \, q \, dx \, dt, \tag{8}$$

the optimal pair (y, v) for J is characterized as follows

$$y = \rho^{-2} L^\star p \quad \text{in} \quad Q_T, \qquad v = -\rho_0^{-2} p \, 1_\omega \quad \text{in} \quad Q_T \tag{9}$$

in term of an additional variable $p \in P$ unique solution to the following variational equality :

$$(p,q)_P = \int_\Omega y_0(x) \, q(x,0) \, dx, \quad \forall q \in P. \tag{10}$$

The well-posedeness of this formulation is ensured as soon as the weights ρ_0, ρ are of Carleman type (in particular ρ and ρ_0 blow up exponentially as $t \to T^-$); this specific behavior near T reinforces the null controllability requirement and prevents the control of any oscillations near the final time.

The search of a control v in the manifold $\mathcal{C}(y_0, T)$ is reduced to solve the (elliptic) variational formulation (10). In [18], the approximation of (10) is performed in the framework of the finite element theory through a discretization of the space-time domain Q_T. In practice, an approximation p_h of p is obtained in a direct way by inverting a symmetric positive definite matrix, in contrast with the iterative (and possibly divergent) methods used within dual methods. Moreover, a major advantage of this approach is that a conformal approximation, say P_h of P, leads to the strong convergence of p_h toward p in P, and consequently from (9), to a strong convergence in $L^2(q_T)$ of $v_h := -\rho_0^{-2} p_h 1_\omega$ toward v, a null control for (1). It is worth to mention that, for any $h > 0$, v_h is not *a priori* an exact control for any finite dimensional system (which is not necessary at all in practice) but an approximation for the L^2-norm of the control v.

The variational formulation (10) derived from the optimality conditions (9) is obtained assuming that the weights ρ and ρ_0 are both strictly positive in Q_T and q_T respectively. In particular, this approach does not apply for the control of minimal L^2-norm, for which simply $\rho := 0$ and $\rho_0 := 1$. The main reason of the present work is to adapt this approach to cover the case $\rho := 0$ and therefore obtain directly an approximation v_h of the control of some minimal weighted L^2-norm. To do so, we adapt the idea developed in [11] devoted to the wave equation. We also mention [33] where a different space-time variational approach (based on Least-squares principles) is used to approximate null controls for the heat equation.

The paper is organized as follows. In Section 2, we associate to the dual problem (4) an equivalent mixed formulation which relies on the optimality conditions associated to the problem (7) with $\rho = 0$. In Section 2.1, we first address the penalization case and write the constraint $L^\star \varphi = 0$ as an equality in $L^2(Q_T)$. We then show the well-posedness of this mixed formulation, in particular we check the inf-sup condition (Theorem 2.1). The mixed formulation allows to approximate simultaneously the dual variable and the primal one, controlled solution of (1). Interestingly, we also derive an equivalent extremal problem in the primal variable y only (see Prop 2.2, Section 2.1.2). In Section 2.2, we reproduce the analysis relaxing the condition $L^\star \varphi = 0$ in the weaker space $L^2(0, T, H^{-1}(\Omega))$. Then, in Section 2.3, by using the Global Carleman estimate (34), we show that a well-posed mixed formulation is also available for the limit and singular case for which $\varepsilon = 0$ leading to Theorem 2.3. Section 3 is devoted to the numerical approximation of the mixed formulation (16) in the case $\varepsilon > 0$ (Section 3.1) and of the mixed formulation (32) in the case $\varepsilon = 0$ (section 3.2). Conformal approximations based on space-time finite elements are employed. In Section 3.3, we numerically check that the approximations used lead to discrete inf-sup properties, uniformly w.r.t. the discretization parameter h. Then the remaining of Section 3 is devoted to some experiments which emphasize the remarkable robuteness of the method. Section 4 concludes with some perspectives.

2 Control of minimal weighted L^2-norm: mixed reformulations

In order to avoid the minimization of the conjugate functional $J^\star$ with respect to the final state φ_T by an iterative process, we now present a direct way to approximate the control of minimal square integrable norm in the spirit of the primal approach recalled in the introduction and developed in [18]. We adapt the case of the wave equation studied in [11].

2.1 The penalized case: Mixed formulation I

Let $\rho_\star \in \mathbb{R}_\star^+$ and let $\rho_0 \in \mathcal{R}$ defined by

$$\mathcal{R} := \{w : w \in C(Q_T); w \geq \rho_\star > 0 \text{ in } Q_T; w \in L^\infty(Q_{T-\delta})\ \forall \delta > 0\} \tag{11}$$

so that in particular, the weight ρ_0 may blow up as $t \to T^-$. We first consider the approximate controllability case. For any $\varepsilon > 0$, the problem reads as follows:

$$\left\{ \begin{array}{l} \text{Minimize } \ J_\varepsilon(y,v) := \dfrac{1}{2}\displaystyle\iint_{q_T} \rho_0^2 |v|^2\, dt + \frac{1}{2\varepsilon}\|y(\cdot,T)\|^2_{L^2(\Omega)} \\ \text{Subject to } \ (y,v) \in \mathcal{A}(y_0;T) \end{array} \right. \tag{12}$$

where $\mathcal{A}(y_0;T)$ denotes the linear manifold $\mathcal{A}(y_0;T) := \{\,(y,v) : v \in L^2(q_T),\ y \text{ solves } (1)\,\}$ and where ε denotes a penalty parameter (see [2, 5, 19]). The corresponding conjugate and well-posed problem is given by

$$\left\{ \begin{array}{l} \text{Minimize } \ J_\varepsilon^\star(\varphi_T) := \dfrac{1}{2}\displaystyle\iint_{q_T} \rho_0^{-2} |\varphi(x,t)|^2 dx\, dt + \frac{\varepsilon}{2}\|\varphi_T\|^2_{L^2(\Omega)} + (y_0, \varphi(\cdot,0))_{L^2(\Omega)} \\ \text{Subject to } \ \varphi_T \in L^2(\Omega) \end{array} \right. \tag{13}$$

where φ solves (5).

We recall that the penalized problem (12) is a consistent approximation of the original null controllability problem, in the sense that its unique solution converges to the solution of (7) with $\rho = 0$ as $\varepsilon \to 0$. We refer for instance to [19], Prop. 3.3 for a proof of the following result, consequence of the null controllability for the heat equation.

Proposition 2.1 *Let $(y_\varepsilon, v_\varepsilon)$ be the solution of Problem (12) and let (y,v) be the solution of Problem (7) with $\rho = 0$. Then, one has*

$$y_\varepsilon \to y \ \text{ strongly in } \ L^2(Q_T), \quad v_\varepsilon \to v \ \text{ strongly in } \ L^2(q_T)$$

as $\varepsilon \to 0^+$.

2.1.1 Mixed formulation

Since the variable φ, solution of (5), is completely and uniquely determined by the data φ_T, the main idea of the reformulation is to keep φ as main variable.

We introduce the linear space $\Phi^0 := \{\varphi \in C^2(\overline{Q_T}),\ \varphi = 0 \text{ on } \Sigma_T\}$. For any $\eta > 0$, we define the bilinear form

$$(\varphi, \overline{\varphi})_{\Phi^0} := \iint_{q_T} \rho_0^{-2} \varphi\, \overline{\varphi}\, dx\, dt + \varepsilon(\varphi(\cdot,T), \overline{\varphi}(\cdot,T))_{L^2(\Omega)} + \eta \iint_{Q_T} L^\star\varphi\, L^\star\overline{\varphi}\, dx\, dt, \quad \forall \varphi, \overline{\varphi} \in \Phi^0.$$

From the unique continuation property for the heat equation, this bilinear form defines for any $\varepsilon \geq 0$ a scalar product. For any $\varepsilon > 0$, let Φ_ε be the completion of Φ^0 for this scalar product. We denote the norm over Φ_ε by $\|\cdot\|_{\Phi_\varepsilon}$ such that

$$\|\varphi\|^2_{\Phi_\varepsilon} := \|\rho_0^{-1}\varphi\|^2_{L^2(q_T)} + \varepsilon\|\varphi(\cdot,T)\|^2_{L^2(\Omega)} + \eta\|L^\star\varphi\|^2_{L^2(Q_T)}, \quad \forall \varphi \in \Phi_\varepsilon. \tag{14}$$

Finally, we defined the closed subset W_ε of Φ_ε by

$$W_\varepsilon = \{\varphi \in \Phi_\varepsilon : L^\star\varphi = 0 \text{ in } L^2(Q_T)\}$$

and we endow W_ε with the same norm than Φ_ε.

Then, we define the following extremal problem :

$$\min_{\varphi\in W_\varepsilon} \hat{J}^\star_\varepsilon(\varphi) := \frac{1}{2}\iint_{q_T} \rho_0^{-2}|\varphi(x,t)|^2 dx\,dt + \frac{\varepsilon}{2}\|\varphi(\cdot,T)\|^2_{L^2(\Omega)} + (y_0, \varphi(\cdot,0))_{L^2(\Omega)}. \tag{15}$$

Standard energy estimates for the heat equation imply that, for any $\varphi \in W_\varepsilon$, $\varphi(\cdot,0) \in L^2(\Omega)$ so that the functional $\hat{J}^\star_\varepsilon$ is well-defined over W_ε. Moreover, since for any $\varphi \in W_\varepsilon$, $\varphi(\cdot,T)$ belongs to $L^2(\Omega)$, Problem (15) is equivalent to the minimization problem (13). As announced, the main variable is now φ submitted to the constraint equality (in $L^2(Q_T)$) $L^\star\varphi = 0$. This constraint equality is addressed by introducing a Lagrangian multiplier.

We consider the following mixed formulation : find $(\varphi_\varepsilon, \lambda_\varepsilon) \in \Phi_\varepsilon \times L^2(Q_T)$ solution of

$$\left\{\begin{array}{rcll} a_\varepsilon(\varphi_\varepsilon,\overline{\varphi}) + b(\overline{\varphi},\lambda_\varepsilon) &=& l(\overline{\varphi}), & \forall \overline{\varphi} \in \Phi_\varepsilon \\ b(\varphi_\varepsilon,\overline{\lambda}) &=& 0, & \forall \overline{\lambda} \in L^2(Q_T), \end{array}\right. \tag{16}$$

where

$$\begin{aligned} a_\varepsilon : \Phi_\varepsilon \times \Phi_\varepsilon \to \mathbb{R}, \quad & a_\varepsilon(\varphi,\overline{\varphi}) := \iint_{q_T} \rho_0^{-2}\varphi\,\overline{\varphi}\,dx\,dt + \varepsilon(\varphi(\cdot,T),\overline{\varphi}(\cdot,T))_{L^2(\Omega)} \\ b : \Phi_\varepsilon \times L^2(Q_T) \to \mathbb{R}, \quad & b(\varphi,\lambda) := -\iint_{Q_T} L^\star\varphi\,\lambda\,dx\,dt \\ l : \Phi_\varepsilon \to \mathbb{R}, \quad & l(\varphi) := -(y_0,\varphi(\cdot,0))_{L^2(\Omega)}. \end{aligned}$$

We have the following result :

THEOREM 2.1 *(i) The mixed formulation* (16) *is well-posed.*

(ii) The unique solution $(\varphi_\varepsilon,\lambda_\varepsilon) \in \Phi_\varepsilon \times L^2(Q_T)$ is the unique saddle-point of the Lagrangian $\mathcal{L}_\varepsilon : \Phi_\varepsilon \times L^2(Q_T) \to \mathbb{R}$ defined by

$$\mathcal{L}_\varepsilon(\varphi,\lambda) := \frac{1}{2}a_\varepsilon(\varphi,\varphi) + b(\varphi,\lambda) - l(\varphi). \tag{17}$$

(iii) The optimal function φ_ε is the minimizer of $\hat{J}^\star_\varepsilon$ over W_ε while the optimal multiplier $\lambda_\varepsilon \in L^2(Q_T)$ is the state of the heat equation (1) *in the weak sense.*

PROOF - We easily check that the bilinear form a_ε is continuous over $\Phi_\varepsilon \times \Phi_\varepsilon$, symmetric and positive and that the bilinear form b_ε is continuous over $\Phi_\varepsilon \times L^2(Q_T)$. Furthermore, for any fixed ε, the continuity of the linear form l over Φ_ε can be viewed from the energy estimate :

$$\|\varphi(\cdot,0)\|^2_{L^2(\Omega)} \leq C\left(\iint_{Q_T} |L^\star\varphi|^2 dx\,dt + \|\varphi(\cdot,T)\|^2_{L^2(\Omega)}\right), \quad \forall \varphi \in \Phi_\varepsilon,$$

for some $C > 0$, so that $\|\varphi(\cdot,0)\|^2_{L^2(\Omega)} \leq C\max(\eta^{-1},\varepsilon^{-1})\|\varphi\|^2_{\Phi_\varepsilon}$.

Therefore, the well-posedness of the mixed formulation is a consequence of the following two properties (see [4]):

- a_ε is coercive on $\mathcal{N}(b)$, where $\mathcal{N}(b)$ denotes the kernel of b :

$$\mathcal{N}(b) := \{\varphi \in \Phi_\varepsilon \ : \ b(\varphi, \lambda) = 0 \text{ for every } \lambda \in L^2(Q_T)\};$$

- b satisfies the usual "inf-sup" condition over $\Phi_\varepsilon \times L^2(Q_T)$: there exists $\delta > 0$ such that

$$\inf_{\lambda \in L^2(Q_T)} \sup_{\varphi \in \Phi_\varepsilon} \frac{b(\varphi, \lambda)}{\|\varphi\|_{\Phi_\varepsilon} \|\lambda\|_{L^2(Q_T)}} \geq \delta. \tag{18}$$

From the definition of a_ε, the first point is clear : for all $\varphi \in \mathcal{N}(b) = W_\varepsilon$, $a_\varepsilon(\varphi, \varphi) = \|\varphi\|^2_{\Phi_\varepsilon}$. Let us check the inf-sup condition. For any fixed $\lambda^0 \in L^2(Q_T)$, we define the (unique) element φ^0 of

$$L^\star \varphi^0 = -\lambda^0 \quad \text{in} \quad Q_T, \qquad \varphi^0 = 0 \quad \text{on} \quad \Sigma_T; \quad \varphi^0(\cdot, T) = 0 \quad \text{in} \quad \Omega,$$

so that φ^0 solves the backward heat equation with source term $-\lambda^0 \in L^2(Q_T)$, null Dirichlet boundary condition and zero initial state. Since $-\lambda^0 \in L^2(Q_T)$, then using energy estimates, there exists a constant $C_{\Omega,T} > 0$ such that the solution φ^0 of the backward heat equation with source term λ^0 satisfies the inequality

$$\iint_{q_T} \rho_0^{-2} |\varphi^0|^2 dx\, dt \leq \rho_\star^{-2} \iint_{q_T} |\varphi^0|^2 dx\, dt \leq \rho_\star^2\, C_{\Omega,T}\, \|\lambda^0\|^2_{L^2(Q_T)}.$$

Consequently, $\varphi^0 \in \Phi_\varepsilon$. In particular, we have $b(\varphi^0, \lambda^0) = \|\lambda^0\|^2_{L^2(Q_T)}$ and

$$\sup_{\varphi \in \Phi_\varepsilon} \frac{b(\varphi, \lambda^0)}{\|\varphi\|_{\Phi_\varepsilon} \|\lambda^0\|_{L^2(Q_T)}} \geq \frac{b(\varphi^0, \lambda^0)}{\|\varphi^0\|_{\Phi_\varepsilon} \|\lambda^0\|_{L^2(Q_T)}} = \frac{\|\lambda^0\|^2_{L^2(Q_T)}}{\left(\|\rho_0^{-1}\varphi^0\|^2_{L^2(q_T)} + \eta \|\lambda_0\|^2_{L^2(Q_T)}\right)^{\frac{1}{2}} \|\lambda_0\|_{L^2(Q_T)}}.$$

Combining the above two inequalities, we obtain

$$\sup_{\varphi_0 \in \Phi_\varepsilon} \frac{b(\varphi_0, \lambda_0)}{\|\varphi_0\|_{\Phi_\varepsilon} \|\lambda_0\|_{L^2(Q_T)}} \geq \frac{1}{\sqrt{\rho_\star^2\, C_{\Omega,T} + \eta}}$$

and, hence, (18) holds with $\delta = \left(\rho_\star^2\, C_{\Omega,T} + \eta\right)^{-1/2}$.

The point (ii) is due to the symmetry and to the positivity of the bilinear form a_ε. (iii) Concerning the third point, the equality $b(\varphi_\varepsilon, \overline{\lambda}) = 0$ for all $\overline{\lambda} \in L^2(Q_T)$ implies that $L^\star \varphi_\varepsilon = 0$ as an $L^2(Q_T)$ function, so that if $(\varphi_\varepsilon, \lambda_\varepsilon) \in \Phi_\varepsilon \times L^2(Q_T)$ solves the mixed formulation, then $\varphi_\varepsilon \in W_\varepsilon$ and $\mathcal{L}_\varepsilon(\varphi_\varepsilon, \lambda_\varepsilon) = \hat{J}^\star_\varepsilon(\varphi_\varepsilon)$. Finally, the first equation of the mixed formulation reads as follows:

$$\iint_{q_T} \rho_0^{-2} \varphi_\varepsilon\, \overline{\varphi}\, dx\, dt + \varepsilon(\varphi_\varepsilon(\cdot, T), \overline{\varphi}(\cdot, T)) - \iint_{Q_T} L^\star \overline{\varphi}(x,t)\, \lambda_\varepsilon(x,t)\, dx\, dt = l(\overline{\varphi}), \quad \forall \overline{\varphi} \in \Phi_\varepsilon,$$

or equivalently, since the control is given by $v_\varepsilon := \rho_0^{-2} \varphi_\varepsilon\, 1_\omega$,

$$\iint_{q_T} v_\varepsilon\, \overline{\varphi}\, dx\, dt + (\varepsilon \varphi_\varepsilon(\cdot, T), \overline{\varphi}(\cdot, T)) - \iint_{Q_T} L^\star \overline{\varphi}(x,t)\, \lambda_\varepsilon(x,t)\, dx\, dt = l(\overline{\varphi}), \quad \forall \overline{\varphi} \in \Phi_\varepsilon.$$

But this means that $\lambda_\varepsilon \in L^2(Q_T)$ is solution of the heat equation in the transposition sense. Since $y_0 \in L^2(\Omega)$ and $v_\varepsilon \in L^2(q_T)$, λ_ε must coincide with the unique weak solution to (1) ($y_\varepsilon = \lambda_\varepsilon$) such that $\lambda_\varepsilon(\cdot, T) = -\varepsilon \varphi_\varepsilon(\cdot, T)$. As a conclusion, the optimal pair $(y_\varepsilon, v_\varepsilon)$ to (12) is characterized in term of the adjoint variable φ_ε solution of (16) by $v_\varepsilon = \rho_0^{-2} \varphi_\varepsilon\, 1_\omega$ and $y_\varepsilon(\cdot, T) = -\varepsilon \varphi_\varepsilon(\cdot, T)$. □

Theorem 2.1 reduces the search of the approximated control to the resolution of the mixed formulation (16), or equivalently the search of the saddle point for $\mathcal{L}_\varepsilon$. In general, it is convenient

to "augment" the Lagrangian (see [20]), and consider instead the Lagrangian $\mathcal{L}_{\varepsilon,r}$ defined for any $r > 0$ by

$$\begin{cases} \mathcal{L}_{\varepsilon,r}(\varphi,\lambda) := \dfrac{1}{2}a_{\varepsilon,r}(\varphi,\varphi) + b(\varphi,\lambda) - l(\varphi), \\ a_{\varepsilon,r}(\varphi,\varphi) := a_\varepsilon(\varphi,\varphi) + r\displaystyle\iint_{Q_T} |L^\star\varphi|^2\,dx\,dt. \end{cases}$$

Since $a_\varepsilon(\varphi,\varphi) = a_{\varepsilon,r}(\varphi,\varphi)$ on W_ε and since the function φ_ε such that $(\varphi_\varepsilon,\lambda_\varepsilon)$ is the saddle point of $\mathcal{L}_\varepsilon$ verifies $\varphi_\varepsilon \in W_\varepsilon$, the lagrangian $\mathcal{L}_\varepsilon$ and $\mathcal{L}_{\varepsilon,r}$ share the same saddle-point.

2.1.2 Dual problem of the extremal problem (15)

The mixed formulation allows to solve simultaneously the dual variable φ_ε, argument of the conjugate functional (15), and the Lagrange multiplier λ_ε. Since λ_ε turns out to be the (approximate) controlled state of (1), we may qualify λ_ε as the primal variable of the problem. We derive in this section the corresponding extremal problem involving only that variable λ_ε.

For any $r > 0$, let us define the linear operator $\mathcal{A}_{\varepsilon,r}$ from $L^2(Q_T)$ into $L^2(Q_T)$ by

$$\mathcal{A}_{\varepsilon,r}\lambda := L^\star\varphi, \quad \forall\lambda \in L^2(Q_T)$$

where $\varphi \in \Phi_\varepsilon$ is the unique solution to

$$a_{\varepsilon,r}(\varphi,\overline{\varphi}) = -b(\overline{\varphi},\lambda), \quad \forall\overline{\varphi} \in \Phi_\varepsilon. \tag{19}$$

Note that the assumption $r > 0$ is necessary here in order to guarantee the well-posedness of (19). Precisely, for any $r > 0$, the form $a_{\varepsilon,r}$ defines a norm equivalent to the norm on Φ_ε (see (14)).

We have the following crucial lemma :

LEMMA 2.1 *For any $r > 0$, the operator $\mathcal{A}_{\varepsilon,r}$ is a strongly elliptic, symmetric isomorphism from $L^2(Q_T)$ into $L^2(Q_T)$.*

PROOF- From the definition of $a_{\varepsilon,r}$, we easily get that $\|\mathcal{A}_{\varepsilon,r}\lambda\|_{L^2(Q_T)} \le r^{-1}\|\lambda\|_{L^2(Q_T)}$ and the continuity of $\mathcal{A}_{\varepsilon,r}$. Next, consider any $\lambda' \in L^2(Q_T)$ and denote by φ' the corresponding unique solution of (19) so that $\mathcal{A}_{\varepsilon,r}\lambda' := L^\star\varphi'$. Relation (19) with $\overline{\varphi} = \varphi'$ then implies that

$$\iint_{Q_T} (\mathcal{A}_{\varepsilon,r}\lambda')\lambda\,dx\,dt = a_{\varepsilon,r}(\varphi,\varphi') \tag{20}$$

and therefore the symmetry and positivity of $\mathcal{A}_{\varepsilon,r}$. The last relation with $\lambda' = \lambda$ implies that $\mathcal{A}_{\varepsilon,r}$ is also positive definite.

Finally, let us check the strong ellipticity of $\mathcal{A}_{\varepsilon,r}$, equivalently that the bilinear functional

$$(\lambda,\lambda') \mapsto \iint_{Q_T} (\mathcal{A}_{\varepsilon,r}\lambda)\lambda'\,dx\,dt$$

is $L^2(Q_T)$-elliptic. Thus we want to show that

$$\iint_{Q_T} (\mathcal{A}_{\varepsilon,r}\lambda)\lambda\,dx\,dt \ge C\|\lambda\|^2_{L^2(Q_T)}, \quad \forall\lambda \in L^2(Q_T) \tag{21}$$

for some positive constant C. Suppose that (21) does not hold; there exists then a sequence $\{\lambda_n\}_{n\ge0}$ of $L^2(Q_T)$ such that

$$\|\lambda_n\|_{L^2(Q_T)} = 1, \quad \forall n \ge 0, \qquad \lim_{n\to\infty}\iint_{Q_T} (\mathcal{A}_{\varepsilon,r}\lambda_n)\lambda_n\,dx\,dt = 0.$$

Let us denote by φ_n the solution of (19) corresponding to λ_n. From (20), we then obtain that

$$\lim_{n\to\infty} \|L^\star\varphi_n\|_{L^2(Q_T)} = 0, \qquad \lim_{n\to\infty} \|\rho_0^{-1}\varphi_n\|_{L^2(q_T)} = 0, \qquad \lim_{n\to\infty} \|\varphi_n(\cdot,T)\|_{L^2(\Omega)} = 0. \tag{22}$$

From (19) with $\lambda = \lambda_n$ and $\varphi = \varphi_n$, we have

$$\iint_{q_T} \rho_0^{-2}\varphi_n\,\overline{\varphi}\,dx\,dt + \varepsilon\int_0^1 \varphi_n(\cdot,T)\overline{\varphi}(\cdot,T)dx + \iint_{Q_T}(rL^\star\varphi_n - \lambda_n)L^\star\overline{\varphi}\,dx\,dt = 0, \quad \forall\overline{\varphi}\in\Phi_\varepsilon. \tag{23}$$

We define the sequence $\{\overline{\varphi}_n\}_{n\geq 0}$ as follows :

$$L^\star\overline{\varphi}_n = r\,L^\star\varphi_n - \lambda_n \quad \text{in} \quad Q_T, \qquad \overline{\varphi}_n = 0 \quad \text{on} \quad \Sigma_T; \quad \overline{\varphi}_n(\cdot,T) = 0 \quad \text{in} \quad \Omega,$$

so that, for all $n \geq 0$, $\overline{\varphi}_n$ is the solution of the backward heat equation with zero initial datum and source term $r\,L^\star\varphi_n - \lambda_n$ in $L^2(Q_T)$. Using again energy type estimates, we get

$$\|\rho_0^{-1}\overline{\varphi}_n\|_{L^2(q_T)} \leq \rho_\star^{-1}\|\overline{\varphi}_n\|_{L^2(q_T)} \leq \rho_\star^{-1}C_{\Omega,T}\|rL^\star\varphi_n - \lambda_n\|_{L^2(Q_T)},$$

so that $\overline{\varphi}_n \in \Phi_\varepsilon$. Then, using (23) with $\overline{\varphi} = \overline{\varphi}_n$, we get

$$\|rL^\star\varphi_n - \lambda_n\|_{L^2(Q_T)} \leq \rho_\star^{-1}C_{\Omega,T}\|\rho_0^{-1}\varphi_n\|_{L^2(q_T)}.$$

Then, from (22), we conclude that $\lim_{n\to+\infty}\|\lambda_n\|_{L^2(Q_T)} = 0$ leading to a contradiction and to the strong ellipticity of the operator $\mathcal{A}_{\varepsilon,r}$. □

The introduction of the operator $\mathcal{A}_{\varepsilon,r}$ is motivated by the following proposition:

PROPOSITION 2.2 *For any $r > 0$, let $\varphi^0 \in \Phi_\varepsilon$ be the unique solution of*

$$a_{\varepsilon,r}(\varphi^0,\overline{\varphi}) = l(\overline{\varphi}), \quad \forall\overline{\varphi}\in\Phi_\varepsilon$$

and let $J^{\star\star}_{\varepsilon,r} : L^2(Q_T) \to L^2(Q_T)$ be the functional defined by

$$J^{\star\star}_{\varepsilon,r}(\lambda) := \frac{1}{2}\iint_{Q_T}(\mathcal{A}_{\varepsilon,r}\lambda)\,\lambda\,dx\,dt - b(\varphi^0,\lambda).$$

The following equality holds :

$$\sup_{\lambda\in L^2(Q_T)}\inf_{\varphi\in\Phi_\varepsilon}\mathcal{L}_{\varepsilon,r}(\varphi,\lambda) = -\inf_{\lambda\in L^2(Q_T)}J^{\star\star}_{\varepsilon,r}(\lambda) \quad + \quad \mathcal{L}_{\varepsilon,r}(\varphi^0,0).$$

PROOF- For any $\lambda \in L^2(Q_T)$, let us denote by $\varphi_\lambda \in \Phi_\varepsilon$ the minimizer of $\varphi \mapsto \mathcal{L}_{\varepsilon,r}(\varphi,\lambda)$; φ_λ satisfies the equation

$$a_{\varepsilon,r}(\varphi_\lambda,\overline{\varphi}) + b(\overline{\varphi},\lambda) = l(\overline{\varphi}), \quad \forall\overline{\varphi}\in\Phi_\varepsilon$$

and can be decomposed as follows : $\varphi_\lambda = \psi_\lambda + \varphi^0$ where $\psi_\lambda \in \Phi_\varepsilon$ solves

$$a_{\varepsilon,r}(\psi_\lambda,\overline{\varphi}) + b(\overline{\varphi},\lambda) = 0, \quad \forall\overline{\varphi}\in\Phi_\varepsilon.$$

We then have

$$\begin{aligned}\inf_{\varphi\in\Phi_\varepsilon}\mathcal{L}_{\varepsilon,r}(\varphi,\lambda) &= \mathcal{L}_{\varepsilon,r}(\varphi_\lambda,\lambda) = \mathcal{L}_{\varepsilon,r}(\psi_\lambda+\varphi^0,\lambda)\\ &= \frac{1}{2}a_{\varepsilon,r}(\psi_\lambda+\varphi^0,\psi_\lambda+\varphi^0) + b(\psi_\lambda+\varphi^0,\lambda) - l(\psi_\lambda+\varphi^0)\\ &:= X_1 + X_2 + X_3\end{aligned}$$

with

$$\begin{cases} X_1 := \dfrac{1}{2}a_{\varepsilon,r}(\psi_\lambda,\psi_\lambda) + b(\psi_\lambda,\lambda) + b(\varphi^0,\lambda)\\ X_2 := a_{\varepsilon,r}(\varphi^0,\psi_\lambda) - l(\psi_\lambda), \quad X_3 := \dfrac{1}{2}a_{\varepsilon,r}(\varphi^0,\varphi^0) - l(\varphi^0).\end{cases}$$

From the definition of φ^0, $X_2 = 0$ while $X_3 = \mathcal{L}_{\varepsilon,r}(\varphi^0, 0)$. Eventually, from the definition of ψ_λ,

$$X_1 = \frac{1}{2} b(\psi_\lambda, \lambda) + b(\varphi^0, \lambda) = -\frac{1}{2} \iint_{Q_T} (\mathcal{A}_{\varepsilon,r}\lambda)\, \lambda \, dx\, dt + b(\varphi^0, \lambda) = -J^{\star\star}_{\varepsilon,r}(\lambda)$$

and the result follows. □

From the ellipticity of the operator $\mathcal{A}_{\varepsilon,r}$, the minimization of the functional $J^{\star\star}_{\varepsilon,r}$ over $L^2(Q_T)$ is well-posed. It is interesting to note that with this extremal problem involving only λ, we are coming to the primal variable, controlled solution of (1) (see Theorem 2.1, (iii)). This argument allows notably to avoid the direct minimization of J_ε (introduced in Problem (12)) with respect to the state y (ill-conditioned due to the term ε^{-1} for ε small). Here, any constraint equality is assigned to the variable λ.

2.2 The penalized case : Mixed formulation II (relaxing the condition $L^\star\varphi_\varepsilon = 0$ in $L^2(Q_T)$)

The previous mixed formulation amounts to find a backward solution φ_ε satisfying the condition $L^\star\varphi_\varepsilon = 0$ in $L^2(Q_T)$. For numerical purposes, it may be interesting to relax this condition, which typically leads to the use of C^1 type approximations in the space variable (see Section 3). In order to circumvent this difficulty, we introduce and analyze in this section a second penalized mixed formulation where the condition on φ_ε is relaxed, namely we impose the constraint $L^\star\varphi_\varepsilon = 0$ in $L^2(0,T;H^{-1}(\Omega))$.

Considering as before the full adjoint variable φ as the main variable, we associated to (13) the following extremal problem :

$$\min_{\varphi\in\widehat{W}_\varepsilon} \hat{J}^\star_\varepsilon(\varphi) = \frac{1}{2} \iint_{Q_T} \rho_0^{-2} |\varphi(x,t)|^2 dx\, dt + \frac{\varepsilon}{2} \|\varphi(\cdot,T)\|^2_{L^2(\Omega)} + \int_\Omega y_0(x)\varphi(x,0)dx, \tag{24}$$

over the space $\widehat{W}_\varepsilon = \Big\{\varphi \in \widehat{\Phi}_\varepsilon :\ L^\star\varphi = 0 \text{ in } L^2(0,T;H^{-1}(\Omega))\Big\}$. The space $\widehat{\Phi}_\varepsilon$ is again defined as the completion of Φ^0 with respect to the inner product

$$(\varphi,\overline{\varphi})_{\widehat{\Phi}_\varepsilon} := \iint_{q_T} \rho_0^{-2}\varphi\,\overline{\varphi}\, dx\, dt + \varepsilon(\varphi(\cdot,T),\overline{\varphi}(\cdot,T)) + \eta\left(\iint_{Q_T} \nabla\varphi\,\nabla\overline{\varphi}\, dx\, dt + \int_0^T (\varphi_t,\overline{\varphi}_t)_{H^{-1}} dt\right),$$

defined over Φ^0. We denote by $\|\cdot\|_{\widehat{\Phi}_\varepsilon}$ the associated norm such that

$$\|\varphi\|^2_{\widehat{\Phi}_\varepsilon} := \|\rho_0^{-1}\varphi\|^2_{L^2(q_T)} + \varepsilon\|\varphi(\cdot,T)\|^2_{L^2(\Omega)} + \eta(\|\nabla\varphi\|^2_{L^2(Q_T)} + \|\varphi_t\|^2_{L^2(0,T;H^{-1})}), \quad \forall\varphi\in\widehat{\Phi}_\varepsilon. \tag{25}$$

LEMMA 2.2 *The equality $\widehat{W}_\varepsilon = W_\varepsilon$ holds. Therefore, the minimization problem* (24) *is equivalent to the minimization* (15).

PROOF - First, let us see that $W_\varepsilon \subset \widehat{W}_\varepsilon$. To do this, it is enough see that $\Phi_\varepsilon \subset \widehat{\Phi}_\varepsilon$. In fact, if $\varphi \in \Phi_\varepsilon$ then there exists a sequence $(\varphi^n)_{n=1}^\infty$ in Φ_0 such that $\varphi^n \to \varphi$ in Φ_ε. So, we can conclude that $\varphi^n \to \varphi$ in $L^2(0,T;H^1_0(\Omega))$ and $\varphi^n_t \to \varphi_t$ in $L^2(0,T;H^{-1}(\Omega))$. Hence, $\varphi^n \to \varphi$ in $\widehat{\Phi}_\varepsilon$.

Secondly, let us see that $\widehat{W}_\varepsilon \subset W_\varepsilon$. Indeed, if $\widehat{\varphi} \in \widehat{W}_\varepsilon$ then $\widehat{\varphi} \in \widehat{\Phi}_\varepsilon$ and $L^\star\widehat{\varphi} = 0$. Let us denote $\widehat{\varphi}_T := \widehat{\varphi}(\cdot,T)$, so there exists a sequence $(\varphi^n_T)_{n=1}^\infty$ in $C^\infty_0(\Omega)$ such that $\varphi^n_T \to \widehat{\varphi}_T$ in $L^2(\Omega)$. Now, if $(\varphi^n)_{n=1}^\infty$ is a sequence such that $L^\star\varphi^n = 0$, $\varphi^n = 0$ on Σ_T and $\varphi^n(\cdot,T) := \varphi^n_T$ then this sequence belongs to Φ^0. Hence, $\varphi^n \to \widehat{\varphi}$ in $\widehat{\Phi}_\varepsilon$ and $\varphi^n \to \widehat{\varphi}$ in Φ_ε. Therefore, $\widehat{\varphi}$ belongs to W_ε. □

The main variable is now φ submitted to the constraint equality $L^\star\varphi = 0 \in L^2(0,T;H^{-1})$. As before, this constraint is addressed by introducing a mixed formulation given as follows : find

$(\varphi_\varepsilon, \lambda_\varepsilon) \in \widehat{\Phi}_\varepsilon \times \widehat{\Lambda}_\varepsilon$ solution of

$$\left\{ \begin{array}{rcll} \hat{a}_\varepsilon(\varphi_\varepsilon, \overline{\varphi}) + \hat{b}(\overline{\varphi}, \hat{\lambda}_\varepsilon) & = & \hat{l}(\overline{\varphi}), & \forall \overline{\varphi} \in \widehat{\Phi}_\varepsilon \\ \hat{b}(\hat{\varphi}_\varepsilon, \overline{\lambda}) & = & 0, & \forall \overline{\lambda} \in \widehat{\Lambda}_\varepsilon, \end{array} \right. \tag{26}$$

where $\widehat{\Lambda}_\varepsilon := L^2(0,T; H_0^1(\Omega))$ and

$$\begin{aligned} &\hat{a}_\varepsilon : \widehat{\Phi}_\varepsilon \times \widehat{\Phi}_\varepsilon \to \mathbb{R}, \quad \hat{a}_\varepsilon(\varphi, \overline{\varphi}) := \iint_{q_T} \rho_0^{-2} \varphi\, \overline{\varphi}\, dx\, dt + \varepsilon(\varphi(\cdot,T), \overline{\varphi}(\cdot,T))_{L^2(\Omega)} \\ &\hat{b} : \widehat{\Phi}_\varepsilon \times \widehat{\Lambda}_\varepsilon \to \mathbb{R}, \\ &\hat{b}(\varphi, \lambda) := -\int_0^T < L^\star \varphi, \lambda >_{H^{-1}(\Omega), H_0^1(\Omega)} dt \\ &\qquad = \int_0^T \langle \varphi_t(t), \lambda(t) \rangle_{H^{-1}, H_0^1} dt - \iint_{Q_T} \Big((c(x) \nabla \varphi, \nabla \lambda) + d(x,t) \varphi \lambda \Big) dx\, dt \\ &\hat{l} : \widehat{\Phi}_\varepsilon \to \mathbb{R}, \quad \hat{l}(\varphi) := -\int_\Omega y_0(x) \varphi(x,0) dx. \end{aligned}$$

Similarly to Theorem 2.1, the following holds :

THEOREM 2.2 *(i) The mixed formulation* (26) *is well-posed.*

(ii) The unique solution $(\varphi_\varepsilon, \lambda_\varepsilon) \in \widehat{\Phi}_\varepsilon \times \widehat{\Lambda}_\varepsilon$ *is the unique saddle-point of the Lagrangian operator* $\widehat{\mathcal{L}}_\varepsilon : \widehat{\Phi}_\varepsilon \times \widehat{\Lambda}_\varepsilon \to \mathbb{R}$ *defined by*

$$\widehat{\mathcal{L}}_\varepsilon(\varphi, \lambda) := \frac{1}{2} \hat{a}_\varepsilon(\varphi, \varphi) + \hat{b}(\varphi, \lambda) - \hat{l}(\varphi). \tag{27}$$

(iii) The optimal function φ_ε *is the minimizer of* $\hat{J}_\varepsilon^\star$ *over* $\widehat{W}_\varepsilon$ *while the optimal multiplier* $\lambda_\varepsilon \in \hat{\Lambda}_\varepsilon$ *is the weak solution of the heat equation* (1).

PROOF - We easily check that the bilinear form $\hat{a}_\varepsilon$ is continuous over $\widehat{\Phi}_\varepsilon \times \widehat{\Phi}_\varepsilon$, symmetric and positive and that the bilinear form $\hat{b}$ is continuous over $\widehat{\Phi}_\varepsilon \times \widehat{\Lambda}_\varepsilon$. Furthermore, the continuity of the linear form $\hat{l}$ over $\widehat{\Phi}_\varepsilon$ is a direct by the continuous embedding $\widehat{\Phi}_\varepsilon \hookrightarrow C^0([0,T]; L^2(\Omega))$. Therefore, the well-posedness of the mixed formulation is a consequence of the following two properties (see [4]):

- $\hat{a}_\varepsilon$ is coercive on $\mathcal{N}(\hat{b})$, where $\mathcal{N}(\hat{b})$ denotes the kernel of $\hat{b}$:

$$\mathcal{N}(\hat{b}) = \Big\{ \varphi \in \widehat{\Phi}_\varepsilon \text{ such that } \hat{b}(\varphi, \lambda) = 0 \text{ for every } \lambda \in \widehat{\Lambda}_\varepsilon \Big\}.$$

- $\hat{b}$ satisfies the usual "inf-sup" condition over $\widehat{\Phi}_\varepsilon \times \widehat{\Lambda}_\varepsilon$: there exists $\delta > 0$ such that

$$\inf_{\lambda \in \widehat{\Lambda}_\varepsilon} \sup_{\varphi \in \widehat{\Phi}_\varepsilon} \frac{\hat{b}(\varphi, \lambda)}{\|\varphi\|_{\widehat{\Phi}_\varepsilon} \|\lambda\|_{\widehat{\Lambda}_\varepsilon}} \geq \delta. \tag{28}$$

From the definition of $\hat{a}_\varepsilon$, the first point is clear : for all $\varphi \in \mathcal{N}(\hat{b}_\varepsilon) = \widehat{W}_\varepsilon$, thanks to classical energy estimates, we have

$$\begin{aligned} \hat{a}_\varepsilon(\varphi, \varphi) &= \|\rho_0^{-1} \varphi\|^2_{L^2(q_T)} + \frac{\varepsilon}{2} \|\varphi(\cdot,T)\|^2_{L^2(\Omega)} + \frac{\varepsilon}{2} \|\varphi(\cdot,T)\|^2_{L^2(\Omega)} \\ &\geq \|\rho_0^{-1} \varphi\|^2_{L^2(q_T)} + \frac{\varepsilon}{2} \|\varphi(\cdot,T)\|^2_{L^2(\Omega)} + \varepsilon\, C(\|\nabla \varphi\|^2_{L^2(Q_T)} + \|\varphi_t\|^2_{L^2(0,T;H^{-1})}) \\ &\geq C_{\varepsilon,\eta} \|\varphi\|^2_{\widehat{\Phi}_\varepsilon}, \end{aligned}$$

where $C = C(T, c_0, \|d\|_\infty) > 0$ and $C_{\varepsilon,\eta} := \min(2^{-1}, C\,\varepsilon\,\eta^{-1})$.

Let us check the inf-sup condition. For any fixed $\lambda^0 \in \widehat{\Lambda}_\varepsilon$, we define the (unique) element φ^0 of

$$L^\star\varphi^0 = \Delta\lambda^0 \quad \text{in} \quad Q_T, \qquad \varphi^0 = 0 \quad \text{on} \quad \Sigma_T; \quad \varphi^0(\cdot,T) = 0 \quad \text{in} \quad \Omega,$$

so that φ^0 solves the backward heat equation with source term $\Delta\lambda^0$, null Dirichlet boundary condition and zero initial state. Since $\Delta\lambda^0 \in L^2(0,T;H^{-1})$, then $\varphi^0 \in \widehat{\Phi}_\varepsilon$: precisely, using energy estimates, there exists a constant $C > 0$ such that φ^0 satisfies the inequalities

$$\|\nabla\varphi^0\|^2_{L^2(Q_T)} \le C\|\nabla\lambda^0\|^2_{L^2(Q_T)}$$

and

$$\begin{aligned}\|\varphi^0\|^2_{\widehat{\Phi}_\varepsilon} &= \|\rho_0^{-1}\varphi^0\|^2_{L^2(q_T)} + \varepsilon\|\varphi^0(\cdot,T)\|^2_{L^2(\Omega)} + \eta(\|\nabla\varphi^0\|^2_{L^2(Q_T)} + \|\varphi^0_t\|^2_{L^2(0,T;H^{-1})}) \\ &= \|\rho_0^{-1}\varphi^0\|^2_{L^2(q_T)} + \eta(\|\nabla\varphi^0\|^2_{L^2(Q_T)} + \|\varphi^0_t\|^2_{L^2(0,T;H^{-1})}) \\ &\le C_\eta\|\nabla\lambda^0\|^2_{L^2(Q_T)}.\end{aligned}$$

where $C = C(T, \|c\|_\infty, \|d\|_\infty) > 0$ and $C_\eta := C(1+\eta)$.

Consequently, $\varphi^0 \in \widehat{\Phi}_\varepsilon$. In particular, we have $\hat{b}_\varepsilon(\varphi^0,\lambda^0) = \|\nabla\lambda^0\|^2_{L^2(Q_T)}$ and

$$\sup_{\varphi\in\widehat{\Phi}_\varepsilon} \frac{b_\varepsilon(\varphi,\lambda^0)}{\|\varphi\|_{\widehat{W}_\varepsilon}\|\lambda^0\|_{\widehat{\Lambda}_\varepsilon}} \ge \frac{b_\varepsilon(\varphi^0,\lambda^0)}{\|\varphi^0\|_{\widehat{W}_\varepsilon}\|\lambda^0\|_{\widehat{\Lambda}_\varepsilon}} = \frac{\|\nabla\lambda^0\|^2_{L^2(Q_T)}}{C_\eta^{1/2}\|\nabla\lambda_0\|_{L^2(Q_T)}\|\nabla\lambda_0\|_{L^2(Q_T)}}.$$

Combining the above two inequalities, we obtain

$$\sup_{\varphi_0\in\Phi} \frac{b(\varphi_0,\lambda_0)}{\|\varphi_0\|_\Phi\|\lambda_0\|_{L^2(Q_T)}} \ge \frac{1}{\sqrt{C_\eta}}$$

and, hence, (28) holds with $\delta = C_\eta^{-\frac{1}{2}}$.

The point (ii) is due to the symmetry and to the positivity of the bilinear form $\hat{a}_\varepsilon$.

Concerning the third assertion, the equality $b(\hat{\varphi}_\epsilon,\overline{\lambda}) = 0$ for all $\overline{\lambda} \in \widehat{\Lambda}_\varepsilon$ implies that $L^\star\varphi_\epsilon = 0$ as an $L^2(0,T;H^{-1})$ function, so that if $(\varphi_\varepsilon,\lambda_\varepsilon) \in \widehat{\Phi}_\varepsilon \times \widehat{\Lambda}_\varepsilon$ solves the mixed formulation (26), then $\varphi_\varepsilon \in \widehat{W}_\varepsilon$ and $\widehat{\mathcal{L}}_\varepsilon(\varphi_\varepsilon,\lambda_\varepsilon) = \hat{J}^\star_\varepsilon(\varphi_\varepsilon)$. This implies that φ_ε of the two mixed formulations coincide.

Assuming $y_0 \in L^2(\Omega)$ and $v \in L^2(q_T)$, it is said here that $y \in L^2(0,T;H^1_0(\Omega))$ is the (unique) solution by transposition of the heat equation (1) if and only if, for every $g \in L^2(0,T;H^{-1})$, we have

$$\int_0^T \langle g, y\rangle_{H^{-1},H^1_0} dt = \iint_{q_T} v\,\overline{\varphi}\,dx\,dt + (\overline{\varphi}(\cdot,0), y_0)_{L^2(\Omega)},$$

where $\overline{\varphi}$ solves

$$L^\star\overline{\varphi} = g \quad \text{in} \quad Q_T, \qquad \overline{\varphi} = 0 \quad \text{on} \quad \Sigma_T, \quad \overline{\varphi}(\cdot,T) = 0 \quad \text{in} \quad \Omega.$$

As $g \mapsto (v,\overline{\varphi})_{L^2(q_T)} + (\overline{\varphi}(\cdot,0), y_0)_{L^2(\Omega)}$ is linear and continuous on $L^2(0,T;H^{-1})$ the Riesz representation theorem guarantees that this definition makes sense.

Finally, the first equation of the mixed formulation (26) reads as follows:

$$\begin{aligned}&\iint_{q_T} \rho_0^{-2}\varphi_\varepsilon\,\overline{\varphi}\,dx\,dt + \varepsilon(\varphi_\varepsilon(\cdot,T),\overline{\varphi}(\cdot,T)) + \int_0^T \langle\overline{\varphi}_t,\lambda_\varepsilon\rangle_{H^{-1},H^1_0} \\ &- \iint_{Q_T} (c(x)\nabla\overline{\varphi},\nabla\lambda_\varepsilon) + d(x,t)\overline{\varphi}\lambda_\varepsilon\,dx\,dt = \hat{l}(\overline{\varphi}), \quad \forall\overline{\varphi} \in \widehat{\Phi}_\varepsilon,\end{aligned}$$

or equivalently, since the control is given by $v_\varepsilon = \rho_0^{-2}\varphi_\varepsilon$ (recall that the formulations (15) and (24) are equivalent),

$$\begin{aligned}&\iint_{q_T} v_\varepsilon\,\overline{\varphi}\,dx\,dt + (\varepsilon\varphi_\varepsilon(\cdot,T),\overline{\varphi}(\cdot,T)) + \int_0^T \langle\overline{\varphi}_t,\lambda_\varepsilon\rangle_{H^{-1},H^1_0}\,dt \\ &- \iint_{Q_T} (c(x)\nabla\overline{\varphi},\nabla\lambda_\varepsilon) + d(x,t)\overline{\varphi}\lambda_\varepsilon\,dx\,dt = \hat{l}(\overline{\varphi}), \quad \forall\overline{\varphi} \in \widehat{\Phi}_\varepsilon.\end{aligned}$$

But this means that $\lambda_\varepsilon \in \widehat{\Lambda}_\varepsilon$ is solution of the heat equation in the transposition sense. Since $y_0 \in L^2(\Omega)$ and $v_\varepsilon \in L^2(q_T)$, λ_ε must coincide with the unique weak solution to (1) ($y_\varepsilon = \lambda_\varepsilon$) and, in particular, we can conclude that $y_\varepsilon(\cdot, T) = -\varepsilon\varphi_\varepsilon(\cdot, T)$. So from the unique of the weak solution, the solution $(\varphi_\varepsilon, \lambda_\varepsilon)$ of the two mixed formulation coincides. □

The equivalence of the mixed formulation (26) with the mixed formulation (16) is related to the regularizing property of the heat kernel. At the numerical level, the advantage is that this formulation leads naturally to continuous spaces of approximation both in time and space.

2.3 Third mixed formulation of the controllability problem : the limit case $\varepsilon = 0$.

We consider in this section the limit case of Section 2.1 corresponding to $\varepsilon = 0$, i.e. to the null controllability. The conjugate functional $J^\star$ corresponding to this case is given in the introduction, see (4), with a weight ρ_0^{-2} (recall that $\rho_0 \in \mathcal{R}$ defined by (11)) in the first term, precisely

$$\min_{\varphi_T \in \mathcal{H}} J^\star(\varphi_T) := \frac{1}{2} \iint_{q_T} \rho_0^{-2}(x,t)|\varphi(x,t)|^2 \, dxdt + (y_0, \varphi(\cdot,0))_{L^2(\Omega)} \tag{29}$$

where the variable φ solves the backward heat equation (5) and $\mathcal{H}$ is again defined as the completion of the $L^2(\Omega)$ space with respect to the norm $\|\varphi_T\|_{\mathcal{H}} := \|\rho_0^{-1}\varphi\|_{L^2(q_T)}$. As explained in the introduction, the limit case is much more singular due to the hugeness of the space $\mathcal{H}$. At the limit $\varepsilon = 0$, the control of the terminal state $\varphi(\cdot, T)$ is lost in $L^2(\Omega)$.

Let $\rho \in \mathcal{R}$. Proceeding as before, we consider again the space $\widetilde{\Phi}_0 = \{\varphi \in C^2(\overline{Q_T}) : \varphi = 0 \,\text{on}\, \Sigma_T\}$ and then, for any $\eta > 0$, we define the bilinear form

$$(\varphi, \overline{\varphi})_{\widetilde{\Phi}_{\rho_0,\rho}} := \iint_{q_T} \rho_0^{-2} \varphi\, \overline{\varphi}\, dx\, dt + \eta \iint_{Q_T} \rho^{-2} L^\star\varphi\, L^\star\overline{\varphi}\, dx\, dt, \quad \forall \varphi, \overline{\varphi} \in \widetilde{\Phi}_0.$$

The introduction of the weight ρ, which does not appear in the original problem (29) will be motivated at the end of this Section. From the unique continuation property for the heat equation, this bilinear form defines for any $\eta > 0$ a scalar product. Let then $\widetilde{\Phi}_{\rho_0,\rho}$ be the completion of $\widetilde{\Phi}_0$ for this scalar product. We denote the norm over $\widetilde{\Phi}_{\rho_0,\rho}$ by $\|\cdot\|_{\widetilde{\Phi}_{\rho_0,\rho}}$ such that

$$\|\varphi\|^2_{\widetilde{\Phi}_{\rho_0,\rho}} := \|\rho_0^{-1}\varphi\|^2_{L^2(q_T)} + \eta\|\rho^{-1}L^\star\varphi\|^2_{L^2(Q_T)}, \quad \forall \varphi \in \widetilde{\Phi}_{\rho_0,\rho}. \tag{30}$$

Finally, we defined the closed subset $\widetilde{W}_{\rho_0,\rho}$ of $\widetilde{\Phi}_{\rho_0,\rho}$ by

$$\widetilde{W}_{\rho_0,\rho} = \{\varphi \in \widetilde{\Phi}_{\rho_0,\rho} : \rho^{-1}L^\star\varphi = 0 \text{ in } L^2(Q_T)\}$$

and we endow $\widetilde{W}_{\rho_0,\rho}$ with the same norm than $\widetilde{\Phi}_{\rho_0,\rho}$.

We then define the following extremal problem :

$$\min_{\varphi \in \widetilde{W}_{\rho_0,\rho}} \hat{J}^\star(\varphi) = \frac{1}{2} \iint_{q_T} \rho_0^{-2}|\varphi(x,t)|^2 dx\, dt + (y_0,\, \varphi(\cdot,0))_{L^2(\Omega)}. \tag{31}$$

For any $\varphi \in \widetilde{W}_{\rho_0,\rho}$, $L^\star\varphi = 0$ a.e. in Q_T and $\|\varphi\|_{\widetilde{W}_{\rho_0,\rho}} = \|\rho_0^{-1}\varphi\|_{L^2(q_T)}$ so that $\varphi(\cdot, T)$ belongs by definition to the abstract space $\mathcal{H}$: consequently, extremal problems (31) and (29) are equivalent. In particular, from the regularizing property of the heat kernel, $\varphi(\cdot, 0)$ belongs to $L^2(\Omega)$ and the linear term in φ in $\hat{J}^\star$ is well defined.

Then, we consider the following mixed formulation : find $(\varphi, \lambda) \in \widetilde{\Phi}_{\rho_0,\rho} \times L^2(Q_T)$ solution of

$$\left\{ \begin{array}{rcll} \tilde{a}(\varphi, \overline{\varphi}) + \tilde{b}(\overline{\varphi}, \lambda) & = & \tilde{l}(\overline{\varphi}), & \forall \overline{\varphi} \in \widetilde{\Phi}_{\rho_0,\rho} \\ \tilde{b}(\varphi, \overline{\lambda}) & = & 0, & \forall \overline{\lambda} \in L^2(Q_T), \end{array} \right. \tag{32}$$

where

$$\begin{aligned}
&\tilde{a} : \widetilde{\Phi}_{\rho_0,\rho} \times \widetilde{\Phi}_{\rho_0,\rho} \to \mathbb{R}, \quad \tilde{a}(\varphi,\overline{\varphi}) = \iint_{q_T} \rho_0^{-2} \varphi\, \overline{\varphi}\, dx\, dt \\
&\tilde{b} : \widetilde{\Phi}_{\rho_0,\rho} \times L^2(Q_T) \to \mathbb{R}, \quad \tilde{b}(\varphi,\lambda) = -\iint_{Q_T} \rho^{-1} L^\star \varphi\, \lambda\, dx\, dt \\
&\tilde{l} : \widetilde{\Phi}_{\rho_0,\rho} \to \mathbb{R}, \quad \tilde{l}(\varphi) = -(y_0, \varphi(\cdot,0))_{L^2(\Omega)}.
\end{aligned}$$

Before to study this mixed formulation, let us do the following comment. The continuity of $\tilde{l}$ over the space $\widetilde{\Phi}_{\rho_0,\rho}$ holds true for a precise choice of the weights which appear in Carleman type estimates for parabolic equations (see [21]): we recall the following important result.

PROPOSITION 2.3 ([21]) *Let the weights $\rho^c, \rho_0^c \in \mathcal{R}$ (see (11)) be defined as follows :*

$$\begin{aligned}
&\rho^c(x,t) := \exp\left(\frac{\beta(x)}{T-t}\right), \quad \beta(x) := K_1\left(e^{K_2} - e^{\beta_0(x)}\right), \\
&\rho_0^c(x,t) := (T-t)^{3/2}\rho^c(x,t),
\end{aligned} \tag{33}$$

where the K_i are sufficiently large positive constants (depending on T, c_0 and $\|c\|_{C^1(\overline{\Omega})}$) such that

$$\beta_0 \in C^\infty(\overline{\Omega}),\ \beta > 0 \quad in \quad \Omega,\ \beta = 0 \quad on \quad \partial\Omega, \quad Supp(\nabla\beta) \subset \overline{\Omega} \setminus \omega.$$

Then, there exists a constant $C > 0$, depending only on ω, T, such that

$$\|\varphi(\cdot,0)\|_{L^2(\Omega)} \le C\|\varphi\|_{\widetilde{\Phi}_{\rho_0^c,\rho^c}}, \quad \forall \varphi \in \widetilde{\Phi}_{\rho_0^c,\rho^c}. \tag{34}$$

The estimate (34) is a consequence of the celebrated global Carleman inequality satisfied by the solution of (5), introduced and popularized in [21]. It allows to obtain the following existence and uniqueness result :

THEOREM 2.3 *Let $\rho_0 \in \mathcal{R}$ and $\rho \in \mathcal{R} \cap L^\infty(Q_T)$ and assume that there exists a positive constant K such that*

$$\rho_0 \le K\rho_0^c, \quad \rho \le K\rho^c \quad in \quad Q_T. \tag{35}$$

Then, we have :

(i) The mixed formulation (32) defined over $\widetilde{\Phi}_{\rho_0,\rho} \times L^2(Q_T)$ is well-posed.

(ii) The unique solution $(\varphi,\lambda) \in \widetilde{\Phi}_{\rho_0,\rho} \times L^2(Q_T)$ is the unique saddle-point of the Lagrangian $\widetilde{\mathcal{L}} : \widetilde{\Phi} \times L^2(Q_T) \to \mathbb{R}$ defined by

$$\widetilde{\mathcal{L}}(\varphi,\lambda) = \frac{1}{2}\tilde{a}(\varphi,\varphi) + \tilde{b}(\varphi,\lambda) - \tilde{l}(\varphi). \tag{36}$$

(iii) The optimal function φ is the minimizer of $\hat{J}^\star$ over $\widetilde{\Phi}_{\rho_0,\rho}$ while $\rho^{-1}\lambda \in L^2(Q_T)$ is the state of the heat equation (1) in the weak sense.

PROOF- The proof is similar to the proof of Theorem 2.1. From the definition, the bilinear form $\tilde{a}$ is continuous over $\widetilde{\Phi}_{\rho_0,\rho} \times \widetilde{\Phi}_{\rho_0,\rho}$, symmetric and positive and the bilinear form $\tilde{b}$ is continuous over $\widetilde{\Phi}_{\rho_0,\rho_1} \times L^2(Q_T)$. Furthermore, the continuity of the linear form $\tilde{l}$ over $\widetilde{\Phi}_{\rho_0,\rho}$ is the consequence of the estimate (34): precisely, from the assumptions (35), the inclusion $\widetilde{\Phi}_{\rho_0,\rho} \subset \widetilde{\Phi}_{\rho_0^c,\rho^c}$ hold true. Therefore, estimate (34) implies

$$\|\varphi(\cdot,0)\|_{L^2(\Omega)} \le C\|\varphi\|_{\widetilde{\Phi}_{\rho_0^c,\rho^c}} \le CK^{-1}\|\varphi\|_{\widetilde{\Phi}_{\rho_0,\rho}}, \quad \forall \varphi \in \widetilde{\Phi}_{\rho_0,\rho}. \tag{37}$$

Therefore, the well-posedeness of the formulation (32) is the consequence of two properties: first, the coercivity of the form $\tilde{a}$ on the kernel $\mathcal{N}(\tilde{b}) := \{\varphi \in \widetilde{\Phi}_{\rho_0,\rho} : \tilde{b}(\varphi,\lambda) = 0\, \forall \lambda \in L^2(Q_T)\}$: again, this holds true since the kernel coincides with the space $\widetilde{W}_{\rho_0,\rho}$. Second, the inf-sup property which reads as :

$$\inf_{\lambda\in L^2(Q_T)} \sup_{\varphi\in\widetilde{\Phi}_{\rho_0,\rho}} \frac{\tilde{b}(\varphi,\lambda)}{\|\varphi\|_{\widetilde{\Phi}_{\rho_0,\rho}}\|\lambda\|_{L^2(Q_T)}} \geq \delta \tag{38}$$

for some $\delta > 0$. For any fixed $\lambda^0 \in L^2(Q_T)$, we define the unique element φ^0 solution of

$$\rho^{-1}L^\star\varphi = -\lambda^0 \,\text{in}\, Q_T, \quad \varphi = 0 \ \text{on}\ \Sigma_T, \quad \varphi(\cdot,T) = 0 \ \text{in}\ \Omega.$$

Using energy estimates, we have

$$\|\rho_0^{-1}\varphi^0\|_{L^2(q_T)} \leq \rho_\star^{-1}\|\varphi^0\|_{L^2(Q_T)} \leq \rho_\star^{-1}\|\rho\lambda^0\|_{L^2(Q_T)} \leq \rho_\star^{-1}\|\rho\|_{L^\infty(Q_T)}\|\lambda^0\|_{L^2(Q_T)} \tag{39}$$

which proves that $\varphi^0 \in \widetilde{\Phi}_{\rho_0,\rho}$ and that

$$\sup_{\varphi\in\widetilde{\Phi}_{\rho_0,\rho}} \frac{\tilde{b}(\varphi,\lambda^0)}{\|\varphi\|_{\widetilde{\Phi}_{\rho_0,\rho}}\|\lambda^0\|_{L^2(Q_T)}} \geq \frac{\tilde{b}(\varphi^0,\lambda^0)}{\|\varphi^0\|_{\widetilde{\Phi}_{\rho_0,\rho}}\|\lambda^0\|_{L^2(Q_T)}} = \frac{\|\lambda^0\|_{L^2(Q_T)}}{\left(\|\rho_0^{-1}\varphi^0\|^2_{L^2(q_T)} + \eta\|\lambda_0\|^2_{L^2(Q_T)}\right)^{\frac{1}{2}}}.$$

Combining the above two inequalities, we obtain

$$\sup_{\varphi_0\in\widetilde{\Phi}_{\rho_0,\rho}} \frac{b(\varphi_0,\lambda_0)}{\|\varphi_0\|_{\widetilde{\Phi}_{\rho_0,\rho}}\|\lambda_0\|_{L^2(Q_T)}} \geq \frac{1}{\sqrt{\rho_\star^{-2}\|\rho\|^2_{L^\infty(Q_T)} + \eta}}$$

and, hence, (38) holds with $\delta = \left(\rho_\star^{-2}\|\rho\|^2_{L^\infty(Q_T)} + \eta\right)^{-1/2}$.

The point (ii) is again due to the positivity and symmetry of the form $\tilde{a}$.

Concerning the last point of the Theorem, the equality $\tilde{b}(\varphi,\overline{\lambda}) = 0$ for all $\overline{\lambda} \in L^2(Q_T)$ implies that $\rho^{-1}L^\star\varphi = 0$ as an $L^2(Q_T)$ function, so that if $(\varphi,\lambda) \in \widetilde{\Phi}_{\rho,\rho_0} \times L^2(Q_T)$ solves the mixed formulation (32), then $\varphi \in \widetilde{W}_{\rho,\rho_0}$ and $\widetilde{\mathcal{L}}(\varphi,\lambda) = \hat{J}^\star(\varphi)$. Finally, the first equation of the mixed formulation (32) reads as follows:

$$\iint_{q_T} \rho_0^{-2}\varphi\,\overline{\varphi}\,dx\,dt - \iint_{Q_T} \rho^{-1}L^\star\overline{\varphi}\lambda\,dx\,dt = \tilde{l}(\overline{\varphi}), \quad \forall\overline{\varphi} \in \widetilde{\Phi},$$

or equivalently, since the control is given by $v := \rho_0^{-2}\varphi\,1_\omega$,

$$\iint_{q_T} v\,\overline{\varphi}\,dx\,dt - \iint_{Q_T} L^\star\overline{\varphi}(\rho^{-1}\lambda)\,dx\,dt = \tilde{l}(\overline{\varphi}), \quad \forall\overline{\varphi} \in \widetilde{\Phi},$$

This means that $\rho^{-1}\lambda \in L^2(Q_T)$ is solution of the heat equation with source term $v\,1_\omega$ in the transposition sense and such that $(\rho^{-1}\lambda)(\cdot,T) = 0$. Since $y_0 \in L^2(\Omega)$ and $v \in L^2(q_T)$, $\rho^{-1}\lambda$ must coincide with the unique weak solution to (1) ($y = \rho^{-1}\lambda$) and, in particular, $y(\cdot,T) = 0$. □

Remark 1 *The well-posedness of the mixed formulation* (32)*, precisely the inf-sup property* (38)*, is open in the case where the weight ρ is simply in $\mathcal{R}$ (instead of $\mathcal{R} \cap L^\infty(Q_T)$): in this case, the weight ρ may blow up at $t \to T^-$. In order to get* (38)*, it suffices to prove that the function $\psi := \rho_0^{-1}\varphi$ solution of the boundary value problem*

$$\rho^{-1}L^\star(\rho_0\psi) = -\lambda^0 \ \textit{in}\ Q_T, \quad \psi = 0 \ \textit{on}\ \Sigma_T, \quad \psi(\cdot,T) = 0 \ \textit{in}\ \Omega$$

for any $\lambda_0 \in L^2(Q_T)$ satisfies the following estimate for some positive constant C

$$\|\psi\|_{L^2(q_T)} \leq C\|\rho^{-1}L^\star(\rho_0\psi)\|_{L^2(Q_T)}.$$

In the cases of interest for which the weights ρ_0 and ρ blow up at $t \to T^-$ (for instance given by ρ_0^c and ρ^c), this estimates is open and does not seem to be a consequence of the estimate (34)*.*

Let us now comment the introduction of the weight ρ. The solution φ of the mixed formulation (32) belongs to $\widetilde{W}_{\rho_0,\rho}$ and therefore does not depend on the weight ρ (recall that ρ is strictly positive); this is in agreement with the fact that ρ does not appear in the original formulation formulation (29). Therefore, this weight may be seen as a parameter to improve some specific properties of the mixed formulation, specifically at the numerical level. Precisely, in the limit case $\varepsilon = 0$, we recall that the trace $\varphi|_{t=T}$ of the solution does not belong to $L^2(\Omega)$ but to a much larger and singular space. Very likely, a similar behavior occurs for the function $L^\star\varphi$ near $\Omega\times\{T\}$ so that the constraint $L^\star\varphi = 0$ in $L^2(Q_T)$ introduced in Section 2.1 is too "strong" and must be replaced at the limit in ε by the relaxed one $\rho^{-1}L^\star\varphi = 0$ in $L^2(Q_T)$ with ρ^{-1} "small" near $\Omega\times\{T\}$. Remark that this is actually the effect and the role of the Carleman type weights ρ^c defined by (33) and initially introduced in [21].

As in Section 2.1, it is convenient to "augment" the Lagrangian and consider instead the Lagrangian $\mathcal{L}_r$ defined for any $r > 0$ by

$$\begin{cases} \mathcal{L}_r(\varphi,\lambda) := \dfrac{1}{2}\tilde{a}_r(\varphi,\varphi) + \tilde{b}(\varphi,\lambda) - \tilde{l}(\varphi), \\ \tilde{a}_r(\varphi,\varphi) := \tilde{a}(\varphi,\varphi) + r\displaystyle\iint_{Q_T} |\rho^{-1}L^\star\varphi|^2\,dx\,dt. \end{cases}$$

Finally, similarly to Lemma 2.1 and Proposition 2.2, we have the following result.

Let $\rho_0 \in \mathcal{R}$ and $\rho \in \mathcal{R}\cap L^\infty(Q_T)$

PROPOSITION 2.4 *For any $r > 0$, let $\rho_0 \in \mathcal{R}$ and $\rho \in \mathcal{R}\cap L^\infty(Q_T)$ verifying (35). Let us define the linear operator $\mathcal{A}_r$ from $L^2(Q_T)$ into $L^2(Q_T)$ by*

$$\mathcal{A}_r\lambda := \rho^{-1}L^\star\varphi, \quad \forall\lambda \in L^2(Q_T),$$

where $\varphi \in \widetilde{\Phi}_{\rho_0,\rho}$ is the unique solution to

$$a_r(\varphi,\overline{\varphi}) = -b(\overline{\varphi},\lambda), \quad \forall\overline{\varphi} \in \widetilde{\Phi}_{\rho_0,\rho}.$$

$\mathcal{A}_r$ is a strongly elliptic, symmetric isomorphism from $L^2(Q_T)$ into $L^2(Q_T)$. Let $\hat{J}_r^{\star\star} : L^2(Q_T) \to L^2(Q_T)$ be the functional defined by

$$\hat{J}_r^{\star\star}(\lambda) := \frac{1}{2}\iint_{Q_T} (\mathcal{A}_r\lambda)\,\lambda\,dx\,dt - \tilde{b}(\varphi^0,\lambda).$$

where $\varphi^0 \in \widetilde{\Phi}_{\rho_0,\rho}$ is the unique solution of

$$\tilde{a}_r(\varphi^0,\overline{\varphi}) = \tilde{l}(\overline{\varphi}), \quad \forall\overline{\varphi} \in \widetilde{\Phi}_{\rho_0,\rho}.$$

The following equality holds :

$$\sup_{\lambda\in L^2(Q_T)}\inf_{\varphi\in\widetilde{\Phi}_{\rho_0,\rho}} \mathcal{L}_r(\varphi,\lambda) = -\inf_{\lambda\in L^2(Q_T)} \hat{J}_r^{\star\star}(\lambda) \quad + \quad \mathcal{L}_r(\varphi^0,0).$$

3 Numerical approximation and experiments

3.1 Discretization of the mixed formulation (16)

We now turn to the discretization of the mixed formulation (16) assuming $r > 0$. Let then $\Phi_{\varepsilon,h}$ and $M_{\varepsilon,h}$ be two finite dimensional spaces parametrized by the variable h such that, for any $\varepsilon > 0$,

$$\Phi_{\varepsilon,h} \subset \Phi_\varepsilon, \quad M_{\varepsilon,h} \subset L^2(Q_T), \quad \forall h > 0.$$

Then, we can introduce the following approximated problems : find $(\varphi_h, \lambda_h) \in \Phi_{\varepsilon,h} \times M_{\varepsilon,h}$ solution of

$$\left\{ \begin{array}{rcll} a_{\varepsilon,r}(\varphi_h, \overline{\varphi}_h) + b(\overline{\varphi}_h, \lambda_h) & = & l(\overline{\varphi}_h), & \forall \overline{\varphi}_h \in \Phi_{\varepsilon,h} \\ b(\varphi_h, \overline{\lambda}_h) & = & 0, & \forall \overline{\lambda}_h \in M_{\varepsilon,h}. \end{array} \right. \tag{40}$$

The well-posedness of this mixed formulation is again a consequence of two properties : the coercivity of the bilinear form $a_{\varepsilon,r}$ on the subset $\mathcal{N}_h(b) = \{\varphi_h \in \Phi_{\varepsilon,h}; b(\varphi_h, \lambda_h) = 0 \quad \forall \lambda_h \in M_{\varepsilon,h}\}$. Actually, from the relation

$$a_{\varepsilon,r}(\varphi, \varphi) \geq C_{r,\eta} \|\varphi\|^2_{\Phi_\varepsilon}, \quad \forall \varphi \in \Phi_\varepsilon,$$

where $C_{r,\eta} = \min\{1, r/\eta\}$, the form $a_{\varepsilon,r}$ is coercive on the full space Φ_ε, and so *a fortiori* on $\mathcal{N}_h(b) \subset \Phi_{\varepsilon,h} \subset \Phi_\varepsilon$. The second property is a discrete inf-sup condition : there exists $\delta_h > 0$ such that

$$\inf_{\lambda_h \in M_{\varepsilon,h}} \sup_{\varphi_h \in \Phi_{\varepsilon,h}} \frac{b(\varphi_h, \lambda_h)}{\|\varphi_h\|_{\Phi_{\varepsilon,h}} \|\lambda_h\|_{M_{\varepsilon,h}}} \geq \delta_h. \tag{41}$$

For any fixed h, the spaces $M_{\varepsilon,h}$ and $\Phi_{\varepsilon,h}$ are of finite dimension so that the infimum and supremum in (41) are reached: moreover, from the property of the bilinear form $a_{\varepsilon,r}$, it is standard to prove that δ_h is strictly positive (see Section 3.3). Consequently, for any fixed $h > 0$, there exists a unique couple (φ_h, λ_h) solution of (40). On the other hand, the property $\inf_h \delta_h > 0$ is in general difficult to prove and depends strongly on the choice made for the approximated spaces $M_{\varepsilon,h}$ and $\Phi_{\varepsilon,h}$. We shall analyze numerically this property in Section 3.3.

Remark 2 *For $r = 0$, the discrete mixed formulation (40) is not well-posed over $\Phi_{\varepsilon,h} \times M_{\varepsilon,h}$ because the bilinear form $a_{\varepsilon,r=0}$ is not coercive over the discrete kernel of b: the equality $b(\lambda_h, \varphi_h) = 0$ for all $\lambda_h \in M_{\varepsilon,h}$ does not imply that $L^\star \varphi_h$ vanishes. Therefore, the term $r\|L^\star \varphi_h\|^2_{L^2(Q_T)}$ may be understood as a numerical stabilization term: for any $h > 0$, it ensures the uniform coercivity of the form $a_{\varepsilon,r}$ (and so the well-posedness) and vanishes at the limit in h. We also emphasize that this term is not a regularization term as it does not add any regularity to the solution φ_h.*

As in [10], the finite dimensional and conformal space $\Phi_{\varepsilon,h}$ must be chosen such that $L^\star \varphi_h$ belongs to $L^2(Q_T)$ for any $\varphi_h \in \Phi_{\varepsilon,h}$. This is guaranteed as soon as φ_h possesses second-order derivatives in $L^2_{loc}(Q_T)$. Any conformal approximation based on standard triangulation of Q_T achieves this sufficient property as soon as it is generated by spaces of functions continuously differentiable with respect to the variable x and spaces of continuous functions with respect to the variable t.

We introduce a triangulation $\mathcal{T}_h$ such that $\overline{Q_T} = \cup_{K \in \mathcal{T}_h} K$ and we assume that $\{\mathcal{T}_h\}_{h>0}$ is a regular family. Then, we introduce the space $\Phi_{\varepsilon,h}$ as follows :

$$\Phi_{\varepsilon,h} = \{\varphi_h \in C^1(\overline{Q_T}) : \varphi_h|_K \in \mathbb{P}(K) \quad \forall K \in \mathcal{T}_h, \ \varphi_h = 0 \text{ on } \Sigma_T\}$$

where $\mathbb{P}(K)$ denotes an appropriate space of polynomial functions in x and t. In this work, we consider for $\mathbb{P}(K)$ the so-called *Bogner-Fox-Schmit* (BFS for short) C^1-element defined for rectangles.

In the one dimensional setting considered in the sequel, it involves 16 degrees of freedom, namely the values of $\varphi_h, \varphi_{h,x}, \varphi_{h,t}, \varphi_{h,xt}$ on the four vertices of each rectangle K. Therefore $\mathbb{P}(K) = \mathbb{P}_{3,x} \otimes \mathbb{P}_{3,t}$ where $\mathbb{P}_{r,\xi}$ is by definition the space of polynomial functions of order r in the variable ξ. We refer to [9] page 76.

We also define the finite dimensional space

$$M_{\varepsilon,h} = \{\lambda_h \in C^0(\overline{Q_T}) : \lambda_h|_K \in \mathbb{Q}(K) \quad \forall K \in \mathcal{T}_h\},$$

where $\mathbb{Q}(K)$ denotes the space of affine functions both in x and t on the element K.

Again, in the one dimensional setting, for rectangle, we simply have $\mathbb{Q}(K) = \mathbb{P}_{1,x} \otimes \mathbb{P}_{1,t}$.

We also mention that the approximation is conformal : for any $h > 0$, we have $\Phi_{\varepsilon,h} \subset \Phi_\varepsilon$ and $M_{\varepsilon,h} \subset L^2(Q_T)$.

Let $n_h = \dim \Phi_{\varepsilon,h}, m_h = \dim M_{\varepsilon,h}$ and let the real matrices $A_{\varepsilon,r,h} \in \mathbb{R}^{n_h,n_h}$, $B_h \in \mathbb{R}^{m_h,n_h}$, $J_h \in \mathbb{R}^{m_h,m_h}$ and $L_h \in \mathbb{R}^{n_h}$ be defined by

$$\begin{cases} a_{\varepsilon,r}(\varphi_h, \overline{\varphi_h}) =< A_{\varepsilon,r,h}\{\varphi_h\}, \{\overline{\varphi_h}\} >_{\mathbb{R}^{n_h},\mathbb{R}^{n_h}}, \quad \forall \varphi_h, \overline{\varphi_h} \in \Phi_{\varepsilon,h}, \\ b(\varphi_h, \lambda_h) =< B_h\{\varphi_h\}, \{\lambda_h\} >_{\mathbb{R}^{m_h},\mathbb{R}^{m_h}}, \quad \forall \varphi_h \in \Phi_{\varepsilon,h}, \forall \lambda_h \in M_{\varepsilon,h}, \\ \displaystyle\iint_{Q_T} \lambda_h \overline{\lambda_h}\, dx\, dt =< J_h\{\lambda_h\}, \{\overline{\lambda_h}\} >_{\mathbb{R}^{m_h},\mathbb{R}^{m_h}}, \quad \forall \lambda_h, \overline{\lambda_h} \in M_{\varepsilon,h}, \\ l(\varphi_h) =< L_h, \{\varphi_h\} >, \quad \forall \varphi_h \in \Phi_{\varepsilon,h} \end{cases}$$

where $\{\varphi_h\} \in \mathbb{R}^{n_h}$ denotes the vector associated to φ_h and $< \cdot, \cdot >_{\mathbb{R}^{n_h},\mathbb{R}^{n_h}}$ the usual scalar product over $\mathbb{R}^{n_h}$. With these notations, Problem (40) reads as follows : find $\{\varphi_h\} \in \mathbb{R}^{n_h}$ and $\{\lambda_h\} \in \mathbb{R}^{m_h}$ such that

$$\begin{pmatrix} A_{\varepsilon,r,h} & B_h^T \\ B_h & 0 \end{pmatrix}_{\mathbb{R}^{n_h+m_h,n_h+m_h}} \begin{pmatrix} \{\varphi_h\} \\ \{\lambda_h\} \end{pmatrix}_{\mathbb{R}^{n_h+m_h}} = \begin{pmatrix} L_h \\ 0 \end{pmatrix}_{\mathbb{R}^{n_h+m_h}}. \tag{42}$$

The matrix $A_{\varepsilon,r,h}$ as well as the mass matrix J_h are symmetric and positive definite for any $h > 0$ and any $r > 0$. On the other hand, the matrix of order $m_h + n_h$ in (42) is symmetric but not positive definite. We use exact integration methods developed in [14] for the evaluation of the coefficients of the matrices. The system (42) is solved using the direct LU decomposition method.

Let us also mention that for $r = 0$, although the formulation (16) is well-posed, numerically, the corresponding matrix $A_{\varepsilon,0,h}$ is not invertible in agreement with Remark 2. In the sequel, we shall consider strictly positive values for r.

Once an approximation φ_h is obtained, an approximation $v_{\varepsilon,h}$ of the control v_ε is given by $v_{\varepsilon,h} = \rho_0^{-2}\varphi_{\varepsilon,h}\, 1_\omega$. The corresponding controlled state $y_{\varepsilon,h}$ may be obtained by solving (1) with standard forward approximation (we refer to [10], Section 4 where this is detailed). Here, since the controlled state is directly given by the multiplier λ, we simply use λ_h as an approximation of y and we do not report here the computation of y_h.

In the sequel, we only report numerical experiments in the one dimensional setting. We use uniform rectangular meshes. Each element is a rectangle of lengths Δx and Δt; $\Delta x > 0$ and $\Delta t > 0$ denote as usual the discretization parameters in space and time, respectively. We note

$$h := \max\{\mathrm{diam}(K), K \in \mathcal{T}_h\}$$

where $\mathrm{diam}(K)$ denotes the diameter of K.

3.2 Normalization and discretization of the mixed formulation (32)

The same approximation may be used for the mixed formulation (32). In particular, we easily check that the finite dimensional spaces $M_{\varepsilon,h}$ and $\Phi_{\varepsilon,h}$ (which actually do not depend on ε) are conformal approximation of $L^2(Q_T)$ and $\widetilde{\Phi}_{\rho_0,\rho}$ respectively. However, in the limit case $\varepsilon = 0$, a normalization of the variable φ, which is singular and takes arbitrarily large amplitude in the neighborhood of $\Omega \times \{T\}$ is very convenient and suitable in practice. Following [18], we introduce the variable $\psi := \rho_0^{-1}\varphi \in \rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho}$ and replace the mixed formulation (32) by the equivalent one: find $(\psi, \lambda) \in \rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho} \times L^2(Q_T)$ solution of

$$\begin{cases} \hat{a}(\psi, \overline{\psi}) + \hat{b}(\overline{\psi}, \lambda) &= \hat{l}(\overline{\psi}), & \forall \overline{\psi} \in \rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho} \\ \hat{b}(\psi, \overline{\lambda}) &= 0, & \forall \overline{\lambda} \in L^2(Q_T), \end{cases} \tag{43}$$

where

$$\hat{a} : \rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho} \times \rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho} \to \mathbb{R}, \quad \hat{a}(\psi,\overline{\psi}) = \iint_{q_T} \psi\,\overline{\psi}\,dx\,dt$$
$$\hat{b} : \rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho} \times L^2(Q_T) \to \mathbb{R}, \quad \hat{b}(\psi,\lambda) = -\iint_{Q_T} \rho^{-1}L^\star(\rho_0\,\psi)\,\lambda\,dx\,dt$$
$$\hat{l} : \rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho} \to \mathbb{R}, \quad \hat{l}(\varphi) = -(y_0, \rho_0(\cdot,0)\psi(\cdot,0))_{L^2(\Omega)}.$$

Well-posedness of this formulation is the consequence of Theorem 2.3. Moreover, the optimal controlled state is still given by $\rho^{-1}\lambda$ while the optimal control is expressed in term of the variable ψ as $v = \rho_0^{-1}\psi\,1_\omega$.

The corresponding discretization approximation (augmented with the term $r\|\rho^{-1}L^\star(\rho_0\psi)\|_{L^2(Q_T)}$) reads as follows: find $(\psi_h,\lambda_h) \in \Phi_h \times M_h$ solution of

$$\left\{\begin{array}{rcll} \hat{a}_r(\psi_h,\overline{\psi}_h) + \hat{b}(\overline{\psi}_h,\lambda_h) &=& \hat{l}(\overline{\varphi}_h), & \forall \overline{\psi}_h \in \Phi_h \\ \hat{b}(\psi_h,\overline{\lambda}_h) &=& 0, & \forall \overline{\lambda}_h \in M_h. \end{array}\right. \tag{44}$$

with

$$\begin{aligned} a_r(\psi_h,\overline{\psi}_h) :=& a(\psi_h,\overline{\psi}_h) + r(\rho^{-1}L^\star(\rho_0\psi_h), \rho^{-1}L^\star(\rho_0\overline{\psi}_h))_{L^2(Q_T)} \\ =& (\psi_h,\overline{\psi}_h)_{L^2(q_T)} + r(\rho^{-1}L^\star(\rho_0\psi_h), \rho^{-1}L^\star(\rho_0\overline{\psi}_h))_{L^2(Q_T)}. \end{aligned}$$

for any $r > 0$.

Remark 3 *When the weights ρ_0 and ρ are chosen in such a way that they are compensated each other in the term $\rho^{-1}L^\star(\rho_0\psi)$, the change of variable has the effect to reduced the amplitude (with respect to the time variable) of the coefficients in the integrals of $\hat{a}_r$ and $\hat{b}$, and therefore, at the discrete level, to improve significantly the condition number of square matrix $\hat{A}_{r,h}$ so that $\hat{a}_r(\psi_h,\overline{\psi}_h) =< \hat{A}_{r,h}\{\psi_h\},\{\overline{\psi}_h\} >_{\mathbb{R}^{n_h},\mathbb{R}^{n_h}}$. In this respect, the change of variable, can be seen as a preconditioner for the mixed formulation (32).*

Similarly to (42), we note the matrix form of (44) as follows :

$$\left(\begin{array}{cc} \hat{A}_{r,h} & \hat{B}_h^T \\ \hat{B}_h & 0 \end{array}\right)_{\mathbb{R}^{n_h+m_h,n_h+m_h}} \left(\begin{array}{c} \{\psi_h\} \\ \{\lambda_h\} \end{array}\right)_{\mathbb{R}^{n_h+m_h}} = \left(\begin{array}{c} \hat{L}_h \\ 0 \end{array}\right)_{\mathbb{R}^{n_h+m_h}}, \tag{45}$$

where $\hat{B}_h$ is the matrix so that $\hat{b}(\psi_h,\lambda_h) =< \hat{B}_h\{\psi_h\},\{\lambda_h\} >_{\mathbb{R}^{m_h},\mathbb{R}^{m_h}}$ and $\hat{L}_h$ is the matrix so that $\hat{l}(\psi_h) =< \hat{L}_h,\{\psi_h\} >$.

3.3 The discrete inf-sup test

Before to give and discuss some numerical experiments, we first test numerically the discrete inf-sup condition (41). Taking $\eta = r > 0$ so that $a_{\varepsilon,r}(\varphi,\overline{\varphi}) = (\varphi,\overline{\varphi})_{\Phi_\varepsilon}$ exactly for all $\varphi,\overline{\varphi} \in \Phi_\varepsilon$, it is readily seen (see for instance [8]) that the discrete inf-sup constant satisfies

$$\delta_{\varepsilon,r,h} = \inf\left\{\sqrt{\delta} : B_h A_{\varepsilon,r,h}^{-1} B_h^T\{\lambda_h\} = \delta\, J_h\{\lambda_h\}, \quad \forall\,\{\lambda_h\} \in \mathbb{R}^{m_h} \setminus \{0\}\right\}. \tag{46}$$

The matrix $B_h A_{\varepsilon,r,h}^{-1} B_h^T$ enjoys the same properties than the matrix $A_{\varepsilon,r,h}$: it is symmetric and positive definite so that the scalar $\delta_{\varepsilon,h}$ defined in term of the (generalized) eigenvalue problem (46) is strictly positive. This eigenvalue problem is solved using the power iteration algorithm (assuming that the lowest eigenvalue is simple): for any $\{v_h^0\} \in \mathbb{R}^{n_h}$ such that $\|\{v_h^0\}\|_2 = 1$, compute for any $n \geq 0$, $\{\varphi_h^n\} \in \mathbb{R}^{n_h}$, $\{\lambda_h^n\} \in \mathbb{R}^{m_h}$ and $\{v_h^{n+1}\} \in \mathbb{R}^{m_h}$ iteratively as follows :

$$\left\{\begin{array}{l} A_{\varepsilon,r,h}\{\varphi_h^n\} + B_h^T\{\lambda_h^n\} = 0 \\ B_h\{\varphi_h^n\} = -J_h\{v_h^n\} \end{array}\right. , \quad \{v_h^{n+1}\} = \frac{\{\lambda_h^n\}}{\|\{\lambda_h^n\}\|_2}.$$

The scalar $\delta_{\varepsilon,r,h}$ defined by (46) is then given by : $\delta_{\varepsilon,r,h} = \lim_{n\to\infty}(\|\{\lambda_h^n\}\|_2)^{-1/2}$.

We now give some numerical values of $\delta_{\varepsilon,r,h}$ with respect to h for the C^1-finite element introduced in Section 3.1.

We consider the one dimensional case for which $\Omega = (0,1)$ and take for simplicity $c := 1/10$ and $d := 0$. Values of the diffusion c and of the potential d do not affect qualitatively the results.

In the spirit of the previous work [18], we consider the following choice for the weight $\rho_0 \in \mathcal{R}$:

$$\rho_0(x,t) := (T-t)^{3/2}\,\exp\left(\frac{K_1}{(T-t)}\right), \quad (x,t) \in Q_T, \qquad K_1 := \frac{3}{4} \tag{47}$$

so that ρ_0 blows exponentially as $t \to T^-$. This allows a smooth behavior of the corresponding control $v := \rho_0^{-2}\varphi\, 1_\omega$. Let us insist however that the mixed formulation is well-posed for any weight $\rho_0 \in \mathcal{R}$, in particular $\rho_0 := 1$ (leading to the control of minimal L^2-norm and for which we refer to [34]). ρ_0 is independent of the variable x for simplicity.

We consider the following data $\omega = (0.2, 0.5)$, $T = 1/2$, and $\Omega = (0,1)$. Tables 1, 2 and 3 provides the values of $\delta_{\varepsilon,r,h}$ with respect to h and ε for $r = 10^{-2}, 1$ and $r = 10^2$, respectively. In view of the definition, we check that $\delta_{\varepsilon,r,h}$ increases as $r \to 0$ and $\varepsilon \to 0$. We also observe, that for r large enough (see Tables 2 and 3), the value of the inf-sup constant is almost constant with respect to ε and behaves like

$$\delta_{\varepsilon,r,h} \approx C_{\varepsilon,r,h} \times r^{-1/2} \tag{48}$$

for some constant $C_{\varepsilon,r,h} \in (0,1)$. More importantly, we observe that for any r and ε, the value of $\delta_{\varepsilon,r,h}$ is bounded by below uniformly with respect to the discretization parameter h. The same behavior is observed for other discretizations such that $\Delta t \neq \Delta x$, other supports ω and other choices for the weight ρ_0 (in particular $\rho_0 := 1$).

Consequently, we may conclude that the finite approximation we have used do "pass" the discrete inf-sup test. It is interesting to note that this is in contrast with the situation for the wave equation for which the parameter r have to be adjusted carefully with respect to h; we refer to [11]. Moreover, as it is usual in mixed finite element theory, such a property together with the uniform coercivity of form $a_{\varepsilon,r}$ then implies the convergence of the approximation sequence (φ_h, λ_h) solution of (40).

h	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\varepsilon = 10^{-2}$	8.358	8.373	8.381	8.386
$\varepsilon = 10^{-4}$	9.183	9.213	9.229	9.237
$\varepsilon = 10^{-8}$	9.263	9.318	9.354	9.383

Table 1: $\delta_{\varepsilon,r,h}$ w.r.t. ε and h ; $r = 10^{-2}$; $\Omega = (0,1)$, $\omega = (0.2, 0.5)$, $T = 1/2$.

h	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\varepsilon = 10^{-2}$	9.933×10^{-1}	9.938×10^{-1}	9.940×10^{-1}	9.941×10^{-1}
$\varepsilon = 10^{-4}$	9.933×10^{-1}	9.938×10^{-1}	9.941×10^{-1}	9.942×10^{-1}
$\varepsilon = 10^{-8}$	9.933×10^{-1}	9.938×10^{-1}	9.941×10^{-1}	9.942×10^{-1}

Table 2: $\delta_{\varepsilon,r,h}$ w.r.t. ε and h ; $r = 1.$; $\Omega = (0,1)$, $\omega = (0.2, 0.5)$, $T = 1/2$.

Similarly, Table 4 displays the discrete inf-sup constant corresponding to the limit case of the

h	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\varepsilon = 10^{-2}$	9.933×10^{-2}	9.939×10^{-2}	9.940×10^{-2}	9.941×10^{-2}
$\varepsilon = 10^{-4}$	9.933×10^{-2}	9.939×10^{-2}	9.941×10^{-2}	9.942×10^{-2}
$\varepsilon = 10^{-8}$	9.933×10^{-2}	9.939×10^{-2}	9.941×10^{-2}	9.942×10^{-2}

Table 3: $\delta_{\varepsilon,r,h}$ w.r.t. ε and h ; $r = 10^2$; $\Omega = (0,1)$, $\omega = (0.2, 0.5)$, $T = 1/2$.

mixed formulation (43):

$$\inf_{\lambda_h \in \widehat{M}_h} \sup_{\psi_h \in \widehat{\Phi}_h} \frac{\hat{b}(\psi_h, \lambda_h)}{\|\lambda_h\|_{L^2(Q_T)} \|\psi_h\|_{\rho_0^{-1}\widetilde{\Phi}_{\rho_0,\rho}}}$$
$$= \inf_{\lambda_h \in \widehat{M}_h} \sup_{\psi_h \in \widehat{\Phi}_h} \frac{\displaystyle\iint_{q_T} \lambda_h \, \rho^{-1} L^\star(\rho_0 \psi_h) \, dx \, dt}{\|\lambda_h\|_{L^2(Q_T)} (\|\psi_h\|^2_{L^2(q_T)} + r \|\rho^{-1} L^\star(\rho_0 \, \psi_h)\|^2_{L^2(Q_T)})^{1/2}}.$$

We take here a weight ρ independent of the variable x given by

$$\rho(x,t) := \exp\left(\frac{K_1}{(T-t)}\right), \quad (x,t) \in Q_T, \quad K_1 := \frac{3}{4}. \tag{49}$$

Again, for the limit case, the value given in the Table suggest a similar behavior observed for $\varepsilon > 0$: the constant is uniformly bounded by below with respect to the parameter h and behaves like $r^{-1/2}$ for r large enough (up to 1). Remark that, due to the introduction of the weight $\rho \neq 1$, the inf-sup constants given by Table 4 are not the limit (as $\varepsilon \to 0$) of the previous Tables.

h	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}	4.41×10^{-3}
$r = 10^2$	6.9×10^{-2}	6.91×10^{-2}	7.06×10^{-2}	8.08×10^{-2}	9.52×10^{-2}
$r = 1$	6.89×10^{-1}	6.91×10^{-1}	6.96×10^{-1}	7.94×10^{-1}	8.66×10^{-1}
$r = 10^{-2}$	1.944	1.922	1.845	1.775	1.731

Table 4: $\varepsilon = 0$; $\delta_{r,h}$ w.r.t. r and h; $\Omega = (0,1)$, $\omega = (0.2, 0.5)$, $T = 1/2$.

3.4 Numerical experiments for the mixed formulation (16)

We report in this section experiments for the mixed formulation (16) and for simplicity we consider only the one dimensional case: $\Omega = (0,1)$ and $T = 1/2$.

Let us first remark that in general explicit solutions $(\varphi_\varepsilon, \lambda_\varepsilon)$ of (16) are not available. However, when the coefficient c and d are constant, we may obtain a semi-explicit representation (using Fourier decomposition) of the minimizer $\varphi_{\varepsilon,T}$ of the conjugate functional $J^\star_\varepsilon$ (see (13)), and consequently of the corresponding adjoint variable φ_ε, the control of weighted minimal square integrable norm $v = \rho_0^{-2} \varphi_\varepsilon \, 1_\omega$ and finally the controlled state y_ε solution of (1). In practice, the obtention of the Fourier representation amounts to solve a symmetric linear system. We refer to the Appendix for the details.

Such representation allows to evaluate precisely the distance of the exact solution $(\varphi_\varepsilon, \lambda_\varepsilon)$ from the approximation (φ_h, λ_h) with respect to h and validate the convergence of the approximation with respect to h.

As for the initial data, we first simply take the first mode of the Laplacian, that is, $y_0(x) = \sin(\pi x)$, $x \in (0,1)$. In view of the regularization property of the heat equation, the regularity of the initial data has a very restricted effect on the optimal control and the robustness of the method.

We take $c(x) := 10^{-1}$, $d(x,t) := 0$ and recall that in the uncontrolled case ($\omega = \emptyset$), these data leads to $\|y(\cdot,T)\|_{L^2(0,1)} = \sqrt{1/2} e^{-\pi^2 cT} \approx 4.31 \times 10^{-1}$. Finally, we take $\omega = (0.2, 0.5)$.

For $r = 1$, Tables 5, 6 and 7 report some norms with respect to h for $\varepsilon = 10^{-2}$, 10^{-4} and $\varepsilon = 10^{-8}$ respectively. The cases $r = 10^2$ and $r = 10^{-2}$ are reported in the Appendix, in Tables 16, 17, 18 and 19, 20, 21 respectively. In the Tables, φ_ε and y_ε denotes the unique solution of (16) given by (51) and (53). In the Tables, $\kappa_{\varepsilon,h}$ denotes the condition number associated to (42), independent of the initial data y_0 [1].

We first check that the L^2-norm $\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$ of the final state is of the order of $\sqrt{\varepsilon}$ and that the condition number $\kappa_{\varepsilon,h}$ behave polynomially with respect to h; on the other hand, we observe a low variation of $\kappa_{\varepsilon,h}$ with respect to ε; $\kappa_{\varepsilon,h} \approx O(h^{5.9})$ for $\varepsilon = 10^{-2}$ and $\kappa_{\varepsilon,h} \approx O(h^{7.3})$ for $\varepsilon = 10^{-8}$.

Then, we check the convergence as h tends to zero of the approximations $(v_{\varepsilon,h}, \lambda_{\varepsilon,h})$ toward the optimal pair $(v_\varepsilon, y_\varepsilon)$ in $L^2(q_T) \times L^2(Q_T)$ for any values of ε and r.

More precisely, for large enough value of ε (here $\varepsilon = 10^{-2}$), we observe a quasi linear rate of convergence for both $\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$ and $\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$ with respect to h, independent of the value of the parameter r. We refer to Figure 1. For small values of ε, we observe a reduced convergence both for the control and the state (see Figure 2 for $\varepsilon = 10^{-4}$ and Figure 3 for $\varepsilon = 10^{-8}$). We recall that as ε tends to zero, the space Φ_ε degenerates into a much larger space and φ_ε highly oscillates near T. Remark also that for $\varepsilon = 10^{-8}$, the constraint $L^\star\varphi_\varepsilon = 0$ as an $L^2(Q_T)$ function is badly represented: this is due to the loss of regularity on the variable φ_ε (in the neighborhood of T) as $\varepsilon \to 0^+$. This does not prevent the convergence of the variable $\varphi_{\varepsilon,h}$ for the norm Φ_ε, in particular the control $v_{\varepsilon,h} = \rho^{-2}\varphi_{\varepsilon,h}\,1_\omega$, and of the variable $\lambda_{\varepsilon,h}$. We will come back to this situation in detail in the section devoted to the limit case $\varepsilon = 0$. Moreover, for small value of ε, the parameter r does have an influence; precisely, a low value of r (here $r = 10^{-2}$) leads to better relative errors : this is in agreement with the behavior of the inf-sup constant $\delta_{\varepsilon,r,h}$ which increases with $r^{-1/2}$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$m_h + n_h$	330	1 155	4 305	16 605	65 205
$\|L^\star\varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	1.32×10^{-1}	3.75×10^{-2}	9.66×10^{-3}	2.42×10^{-3}	7.82×10^{-4}
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	1.10×10^{-1}	6.21×10^{-2}	3.29×10^{-2}	1.68×10^{-2}	8.57×10^{-3}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	5.13×10^{-2}	2.84×10^{-2}	1.48×10^{-2}	7.60×10^{-3}	3.89×10^{-3}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	1.54×10^{-1}	1.61×10^{-1}	1.65×10^{-1}	1.67×10^{-1}	1.68×10^{-1}
$\kappa_{\varepsilon,h}$	1.52×10^{9}	1.10×10^{11}	6.80×10^{12}	3.83×10^{14}	1.96×10^{16}

Table 5: Mixed formulation (16) - $r = 1$ and $\varepsilon = 10^{-2}$ with $\omega = (0.2, 0.5)$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star\varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	1.383	1.471	9.05×10^{-1}	2.56×10^{-1}	6.54×10^{-2}
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	6.72×10^{-1}	3.22×10^{-1}	1.15×10^{-1}	5.49×10^{-2}	2.74×10^{-2}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	2.73×10^{-1}	1.86×10^{-1}	5.89×10^{-2}	2.51×10^{-2}	1.26×10^{-2}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	8.50×10^{-2}	5.74×10^{-2}	3.39×10^{-2}	3.11×10^{-2}	3.13×10^{-2}
$\kappa_{\varepsilon,h}$	3.02×10^{9}	3.91×10^{11}	3.86×10^{13}	3.25×10^{15}	2.46×10^{17}

Table 6: Mixed formulation (16) - $r = 1$ and $\varepsilon = 10^{-4}$ with $\omega = (0.2, 0.5)$.

[1] The condition number $\kappa(\mathcal{M}_h)$ of any square matrix $\mathcal{M}_h$ is defined by $\kappa(\mathcal{M}_h) = |||\mathcal{M}_h|||_2 |||\mathcal{M}_h^{-1}|||_2$ where the norm $|||\mathcal{M}_h|||_2$ stands for the largest singular value of $\mathcal{M}_h$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star \varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	1.48	2.03	2.50	2.52	2.61
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	1.44	1.01	7.92×10^{-1}	6.65×10^{-1}	4.89×10^{-1}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	8.42×10^{-1}	8.27×10^{-1}	5.73×10^{-1}	4.35×10^{-1}	2.89×10^{-1}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	8.63×10^{-2}	6.65×10^{-2}	2.39×10^{-2}	1.23×10^{-2}	4.43×10^{-3}
$\kappa_{\varepsilon,h}$	3.12×10^{9}	4.30×10^{11}	6.05×10^{13}	1.13×10^{16}	1.90×10^{18}

Table 7: Mixed formulation (16) - $r = 1$ and $\varepsilon = 10^{-8}$ with $\omega = (0.2, 0.5)$.

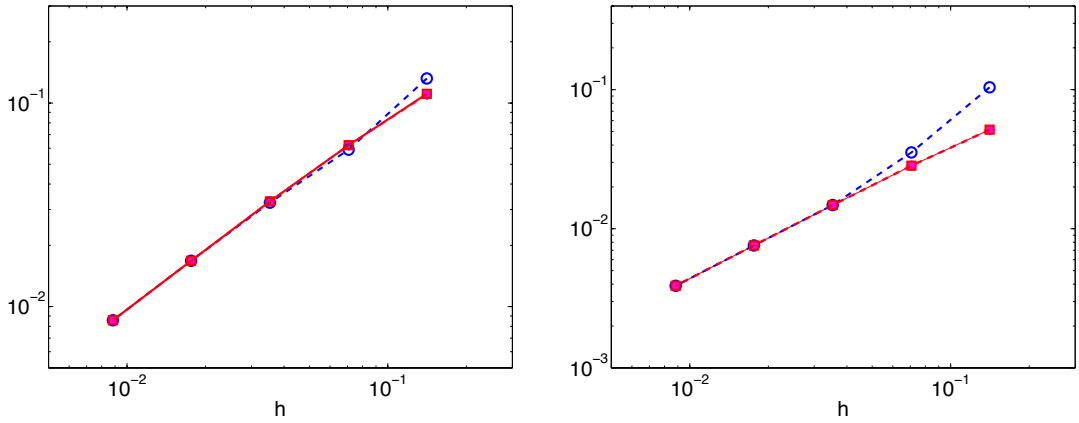

Figure 1: $\omega = (0.2, 0.5)$; $y_0(x) = \sin(\pi x)$: $\varepsilon = 10^{-2}$. ; $\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$ (**Left**) and $\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$ (**Right**) vs. h for $r = 10^2$ ($\circ$), $r = 1$. ($\star$) and $r = 10^{-2}$ ($\square$).

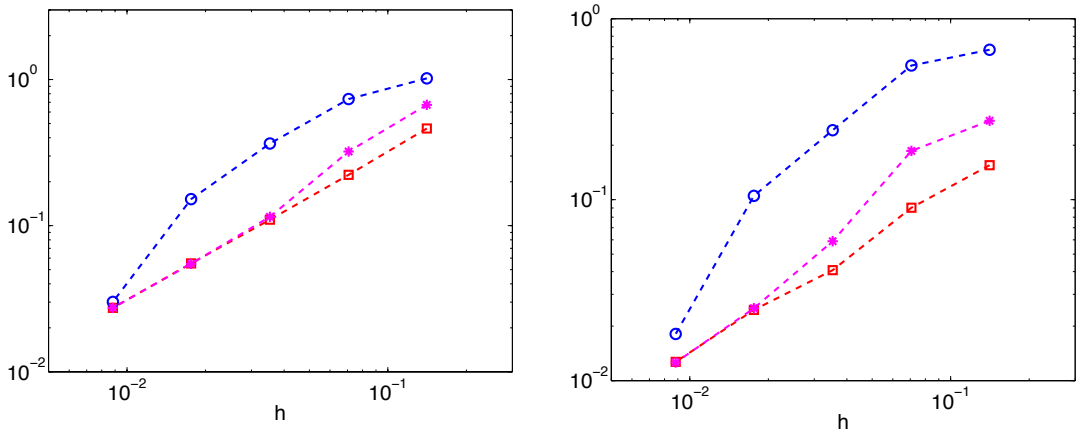

Figure 2: $\omega = (0.2, 0.5)$; $y_0(x) = \sin(\pi x)$: $\varepsilon = 10^{-4}$. ; $\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$ (**Left**) and $\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$ (**Right**) vs. h for $r = 10^2$ ($\circ$), $r = 1$. ($\star$) and $r = 10^{-2}$ ($\square$).

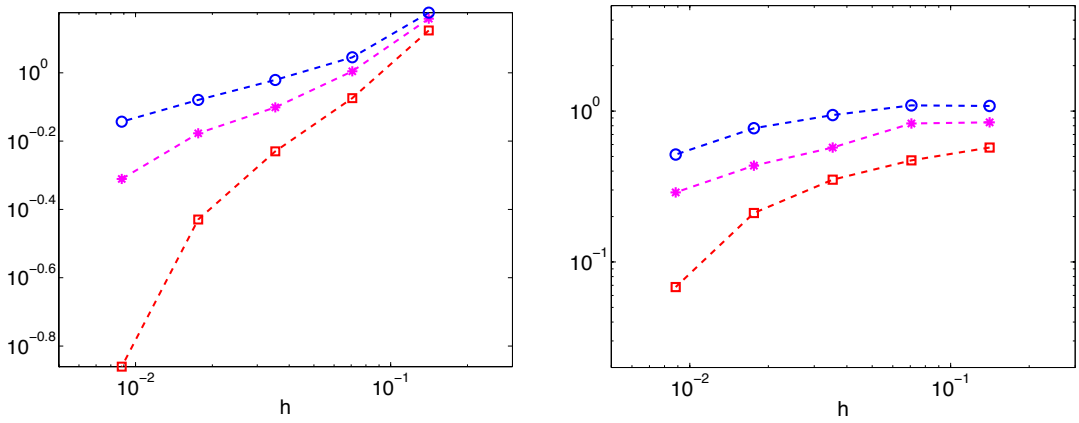

Figure 3: $\omega = (0.2, 0.5)$; $y_0(x) = \sin(\pi x)$: $\varepsilon = 10^{-8}$. ; $\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$ (**Left**) and $\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$ (**Right**) vs. h for $r = 10^2$ ($\circ$), $r = 1.$ ($\star$) and $r = 10^{-2}$ ($\square$).

Remarkably, we highlight that the variational approach developed here allows, for any ε, a direct and robust approximation of one control for the heat equation. As discussed at length in [19, 34], the minimization of the conjugate functional $J^\star_\varepsilon$ using conjugate gradient algorithm requires a number of iterates for small ε (typically $\varepsilon = 10^{-8}$ and $\omega = (0.2, 0.5)$) and diverge for small values of h.

3.5 Conjugate gradient for $J^{\star\star}_{\varepsilon,r}$

We illustrate here the Section 2.1.2 and minimize the functional $J^{\star\star}_{\varepsilon,r} : L^2(Q_T) \to \mathbb{R}$ defined in Proposition 2.2 with respect to the variable λ_ε. From the ellipticity of the operator $\mathcal{A}_{\varepsilon,r}$, we use a conjugate gradient method which in the present case reads as follows :

(i) *Initialization*
Let $\lambda^0_\varepsilon \in L^2(Q_T)$ be a given function;
Solve

$$\begin{cases} \overline{\varphi}^0_\varepsilon \in \Phi_\varepsilon \\ a_{\varepsilon,r}(\overline{\varphi}^0_\varepsilon, \overline{\varphi}) + b_\varepsilon(\overline{\varphi}, \lambda^0_\varepsilon) = l_\varepsilon(\overline{\varphi}), \quad \forall \overline{\varphi} \in \Phi_\varepsilon \end{cases}$$

and set $g^0_\varepsilon = L^\star \overline{\varphi}^0_\varepsilon$ and set $w^0_\varepsilon = g^0_\varepsilon$.

For $n \geq 0$, assuming that $\lambda^n_\varepsilon, g^n_\varepsilon$ and w^n_ε are known with $g^n_\varepsilon \neq 0$ and $w^n_\varepsilon \neq 0$, compute $\lambda^{n+1}_\varepsilon, g^{n+1}_\varepsilon$ and w^{n+1}_ε as follows

(ii) *Steepest descent*
Compute $\overline{\varphi}^n \in \Phi_\varepsilon$ solution to

$$a_{\varepsilon,r}(\overline{\varphi}^n_\varepsilon, \overline{\varphi}) = -b_\varepsilon(\overline{\varphi}, w^n_\varepsilon), \quad \forall \overline{\varphi} \in \Phi_\varepsilon$$

and $\overline{w}^n_\varepsilon = L^\star \overline{\varphi}^n_\varepsilon$ and then compute

$$\rho_n = \|g^n_\varepsilon\|^2_{L^2(Q_T)} / (\overline{w}^n_\varepsilon, w^n_\varepsilon)_{L^2(Q_T)}.$$

and set

$$\lambda^{n+1}_\varepsilon = \lambda^n_\varepsilon - \rho_n w^n_\varepsilon.$$

(iii) *Testing the convergence and construction of the new descent direction*
Update g_ε^n by

$$g_\varepsilon^{n+1} = g_\varepsilon^n - \rho^n \overline{w}_\varepsilon^n.$$

If $\|g_\varepsilon^{n+1}\|_{L^2(Q_T)}/\|g_\varepsilon^0\|_{L^2(Q_T)} \leq \gamma$, take $\lambda_\varepsilon = \lambda_\varepsilon^{n+1}$. Else, compute

$$\gamma_n = \|g_\varepsilon^{n+1}\|^2_{L^2(Q_T)}/\|g_\varepsilon^n\|^2_{L^2(Q_T)}$$

and update w_ε^n via

$$w_\varepsilon^{n+1} = g_\varepsilon^{n+1} + \gamma_n w_\varepsilon^n.$$

Do $n = n + 1$ and return to step (ii).

As mentioned in [22] where this approach is discussed at length for Stokes and Navier-Stokes systems, this algorithm can be viewed as a sophisticated version of Uzawa type algorithm to solve the mixed formulation (16). Concerning the speed of convergence of this algorithm, it follows, for instance, from [13] that

$$\|\lambda_\varepsilon^n - \lambda_\varepsilon\|_{L^2(Q_T)} \leq 2\sqrt{\nu(\mathcal{A}_{\varepsilon,r})}\left(\frac{\sqrt{\nu(\mathcal{A}_{\varepsilon,r})}-1}{\sqrt{\nu(\mathcal{A}_{\varepsilon,r})}+1}\right)^n \|\lambda_\varepsilon^0 - \lambda_\varepsilon\|_{L^2(Q_T)}, \quad \forall n \geq 1$$

where λ_ε minimizes $J^{\star\star}_{\varepsilon,r}$. $\nu(\mathcal{A}_{\varepsilon,r}) = \|\mathcal{A}_{\varepsilon,r}\|\|\mathcal{A}_{\varepsilon,r}^{-1}\|$ denotes the condition number of the operator $\mathcal{A}_{\varepsilon,r}$.

Eventually, once the above algorithm has converged we can compute $\varphi_\varepsilon \in \Phi_\varepsilon$ as solution of

$$a_{\varepsilon,r}(\varphi_\varepsilon, \overline{\varphi}) + b_\varepsilon(\overline{\varphi}, \lambda_\varepsilon) = l_\varepsilon(\overline{\varphi}), \quad \forall \overline{\varphi} \in \Phi_\varepsilon.$$

We use the same spaces $\Phi_{\varepsilon,h}$ and $M_{\varepsilon,h}$ as described in Section 3.1. In practice, each iteration amounts to solve a linear system involving the matrix $A_{\varepsilon,r,h}$ of size $n_h = 4m_h$ (see (42)) which is sparse, symmetric and positive definite. We use the Cholesky method.

From the previous estimate, the performances of the algorithm are related to the condition number of the operator $\mathcal{A}_{\varepsilon,r}$ restricted to $M_{\varepsilon,h} \subset L^2(Q_T)$, which coincides here (see [4]) with the condition number of the symmetric and positive definite matrix $B_h A^{-1}_{\varepsilon,r,h} B_h^T$ introduced in (46). Using again the power iteration algorithm, we obtain that, for any h, the largest eigenvalue of $B_h A^{-1}_{\varepsilon,r,h} B_h^T$ is very closed to r^{-1} (and bounded by r^{-1}). This is in agreement with the estimate $\|\mathcal{A}_{\varepsilon,r}\lambda\|_{L^2(Q_T)} \leq r^{-1}\|\lambda\|_{L^2(Q_T)}$ for all $\lambda \in L^2(Q_T)$. Consequently, the condition number is expressed in term of r and of the discrete inf-sup constant $\delta_{\varepsilon,h}$ as follows :

$$\nu(B_h A^{-1}_{\varepsilon,r,h} B_h^T) \approx r^{-1}\delta^{-2}_{\varepsilon,r,h}.$$

Since, from our observation in Section 3.3, the discrete inf-sup constant $\delta_{\varepsilon,r,h}$ is uniformly bounded by below with respect to h, we deduce that the condition number is uniformly bounded by above with respect to the discretization parameter. This implies that the convergence of the sequence $\{\lambda^n_{\varepsilon,h}\}_{(n>0)}$, minimizing for $J^{\star\star}_{\varepsilon,r}$ over $M_{\varepsilon,h}$ is independent of h. This is exactly what we observe from our numerical experiments. Morever, from (48), we get that the number $\nu(B_h A^{-1}_{\varepsilon,r,h} B_h^T) \approx C^{-2}_{\varepsilon,r,h}$ is very closed to one. We refer to Tables 8 and 9 for the values.

We consider the same data as in Section 3.4, that is, $\omega = (0.2, 0.5)$, $y_0(x) = \sin(\pi x)$ and $T = 1/2$. We take $\gamma = 10^{-10}$ as a stopping threshold for the algorithm (that is the algorithm is stopped as soon as the norm of the residue g^n at the iterate n satisfies $\|g^n_\varepsilon\|_{L^2(Q_T)} \leq 10^{-10}\|g^0_\varepsilon\|_{L^2(Q_T)}$). The algorithm is initiated with $\lambda^0_{\varepsilon,h} = 0$ in Q_T.

We check that the method provides, for the same value of r, ε and h, exactly the same approximation $\lambda_{\varepsilon,h}$ than the previous direct method (see Tables 5, etc). Table 10, 11 and 12, we simply give the number of iterates of the conjugate gradient algorithm for $r = 10^2$, $r = 1$ and

h	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\varepsilon = 10^{-2}$	1.431	1.426	1.423	1.423
$\varepsilon = 10^{-4}$	1.185	1.177	1.173	1.171
$\varepsilon = 10^{-8}$	1.165	1.151	1.142	1.135

Table 8: $r^{-1}\delta_{\varepsilon,r,h}^{-2}$ w.r.t. ε and h ; $r = 10^{-2}$; $\Omega = (0,1)$, $\omega = (0.2, 0.5)$, $T = 1/2$.

h	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\varepsilon = 10^{-2}$	1.013	1.012	1.012	1.011
$\varepsilon = 10^{-4}$	1.013	1.012	1.011	1.011
$\varepsilon = 10^{-8}$	1.013	1.012	1.011	1.011

Table 9: $r^{-1}\delta_{\varepsilon,r,h}^{-2}$ w.r.t. ε and h ; $r = 1$ and $r = 10^2$; $\Omega = (0,1)$, $\omega = (0.2, 0.5)$, $T = 1/2$.

$r = 10^{-2}$ with respect to h and ε respectively. For each case, the convergence is reached in very few iterates, independent of h. Once again, this is in contrast with the behavior of the conjugate gradient algorithm when this latter is used to minimize $J_\varepsilon^\star$ with respect to φ_T defined by (13). The number of convergence is also almost independent of ε and r. Since the gradient of $J_{\varepsilon,r}^{\star\star}$ is given by $\nabla J_{\varepsilon,r}^{\star\star}(\lambda') = \mathcal{A}_{\varepsilon,r}(\lambda') - L^\star\varphi_0$ for all $\lambda' \in L^2(Q_T)$, in particular $\nabla J_{\varepsilon,r}^{\star\star}(\lambda_\varepsilon) = L^\star\varphi_\varepsilon$, a larger value of the augmentation parameter r reduces (slightly here) the number of iterates.

According to this very low number of iterates, it seems more advantageous not only in term of memory resource but also in term of time execution to solve the extremal problem in the variable λ_ε than the (equivalent) mixed formulation (40). The matrix $A_{\varepsilon,r,h}$ of order n_h is very sparse, symmetric, definite positive, diagonal bloc (for which the Cholesky method is very efficient) while the matrix defined by (42), of order $m_h + n_h = 5/4n_h$ requires the use of for instance the Gauss decomposition method. Note however that the condition number of the matrix $A_{\varepsilon,r,h}$ is not independent of h but behaves polynomially (see Table 10 where the value is reported for $r = 1$.). On the other hand, the condition number slightly decreases with r (recall that the norm over Φ_ε contains the term $r\|L^\star\varphi\|_{L^2(Q_T)}$): consequently, for very stiff situation (typically ω very small), there may be a balance between large values of r leading to a better numerical robustness and low values of r leading to smaller relative errors on $v_{\varepsilon,h}$ and $\lambda_{\varepsilon,h}$.

For very small values of both h (leading to very fine meshes) of the order $h = 10^{-3}$ and ε, we observe some instabilities on the approximation $\lambda_{\varepsilon,h}$ (very likely due to the condition number of the matrix $A_{\varepsilon,r,h}$ which exceeds 10^{25} in this case). A preconditioning technique introduced in the next section is needed in these cases.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$m_h = card(\{\lambda_{\varepsilon,h}\})$	66	231	861	3 321	13 041
♯ iterates - $\varepsilon = 10^{-2}$	5	5	5	5	5
♯ iterates - $\varepsilon = 10^{-4}$	5	5	5	4	4
♯ iterates - $\varepsilon = 10^{-8}$	5	5	5	5	5
$\kappa(A_{\varepsilon,r,h})$ - $\varepsilon = 10^{-2}$	1.51×10^9	1.10×10^{11}	6.81×10^{12}	3.83×10^{14}	1.91×10^{16}

Table 10: Mixed formulation (16) - $r = 1$ - $\omega = (0.2, 0.5)$; Conjugate gradient algorithm.

We do not describe experiments for the mixed formulation introduced in 2.2, which require the use of continuous finite element approximation. We refer to [18] in a closed context.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
♯ iterates - $\varepsilon = 10^{-2}$	5	5	4	4	4
♯ iterates - $\varepsilon = 10^{-4}$	5	5	5	4	4
♯ iterates - $\varepsilon = 10^{-8}$	5	5	5	5	4

Table 11: Mixed formulation (16) - $r = 10^2$ - $\omega = (0.2, 0.5)$; Conjugate gradient algorithm.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
♯ iterates - $\varepsilon = 10^{-2}$	9	9	8	8	8
♯ iterates - $\varepsilon = 10^{-4}$	8	8	8	8	8
♯ iterates - $\varepsilon = 10^{-8}$	8	8	7	7	7

Table 12: Mixed formulation (16) - $r = 10^{-2}$ - $\omega = (0.2, 0.5)$; Conjugate gradient algorithm.

3.6 Numerical experiments for the mixed formulation (32) - limit case $\varepsilon = 0$.

We now report in this section some experiments corresponding to the limit case, that is $\varepsilon = 0$, of the mixed formulation (32). We consider again the first mode : $y_0(x) = \sin(\pi x)$, take $\omega = (0.2, 0.5)$, $T = 1/2$ and the exponential type weights ρ_0 and ρ given by (47) and (49) respectively.

This particular choice of the weights allows to rewrite the quantity $\rho^{-1} L^\star \varphi$ in term of the new variable ψ as follow

$$\begin{aligned} \rho^{-1} L^\star(\rho_0 \psi) &= \rho^{-1}\rho_0 L^\star \psi - \rho^{-1}\rho_{0t}\psi \\ &= (T-t)^{3/2} L^\star \psi + \left(-\frac{3}{2}(T-t)^{1/2} + K_1 (T-t)^{-1/2} \right) \psi \end{aligned} \tag{50}$$

and thus eliminate the exponential singularity near T^{-1}. Only a much weaker polynomial singularity, precisely $(T-t)^{-1/2}$ remains.

Moreover, we define as "exact" solution (y, v) the solution obtained with a very fine mesh corresponding to $h \approx 1.1 \times 10^{-3}$, a number of element equal to 819 200 and a number of degrees of freedom equal to $m_h + n_h = 3\ 284\ 484$. With these values, we get the following norms :

$$\|\rho^{-1}\lambda_{h=1.1\times 10^{-3}}\|_{L^2(Q_T)} \approx 3.592 \times 10^{-1}, \quad \|\rho_0 v_{h=1.1\times 10^{-3}}\|_{L^2(q_T)} \approx 18.6634.$$

We do not use the Fourier expansion approach described in the Appendix, since the optimality equation (52) is ill-posed for $\varepsilon = 0$ and leads to instability as the number of modes used in the sum increases. On the contrary, the minimization of $J_r^{\star\star}$- equivalent to the resolution of the mixed formulation (32) exhibits a remarkable robustness as $h \to 0$. Eventually, we mention that the mesh used is so fine that the corresponding result is (almost) independent of the parameter r.

Tables 13, 14 and 15 reports some norms with respect to h for $r = 10^{-2}$, $r = 1$ and $r = 10^2$, respectively. Let us first mention that we again obtain exactly the same approximations from the direct resolution of the system (45) and from the minimization of $J_r^{\star\star}$.

As in the case $\varepsilon > 0$, we observe the convergence of $\rho^{-1}\lambda_h$ and $\rho_0 v_h$ in $L^2(Q_T)$ and $L^2(q_T)$ respectively as $h \to 0^+$. For instance, for $r = 1$, we obtain

$$\frac{\|\rho_0(v - v_h)\|_{L^2(q_T)}}{\|\rho_0 v\|_{L^2(q_T)}} \approx e^{1.04} h^{0.429}, \quad \frac{\|y_- \rho^{-1}\lambda_h\|_{L^2(Q_T)}}{\|y\|_{L^2(Q_T)}} \approx e^{1.27} h^{0.704}.$$

Figure 4 depicts the evolution of these relatives errors with respect to h for $r = 10^{-2}, 1$ and $r = 10^2$. Again, in view of the values of the inf-sup constant of Table 4, we check that the lower value $r = 10^{-2}$ provides a faster convergence of the approximation. It is also interesting to remark

that low errors for the state $\rho^{-1}\lambda_h$ and the control v_h are obtained with a relatively large value of the norm $\|\rho^{-1}L^\star\varphi_h\|_{L^2(Q_T)}$. This suggests that the constraint equality $L^\star\varphi = 0$ in $L^2(Q_T)$ may be replaced by a weaker one as discussed in Section 2.2. We do not present experiments for the weaker formulation (26) and refer to Section 4 of [18] in a closed context. Tables 13, 14 and 15 also report some results from the minimization of the functional $J_r^{\star\star}$ using the conjugate gradient algorithm. For $r = 1$ and $r = 10^2$, the quantity $r^{-1}\delta_{r,h}^{-2}$ - bounded by above of the condition number of $\hat{B}_h\hat{A}_{r,h}^{-1}\hat{B}_h^T$ - slightly decreases with h; the convergence of the algorithm is reached in few iterations independent of h. The value $r = 10^{-2}$ requires about 50 iterations for all the discretization considered.

Remarkably, the change of variable performed in the limit case allows to reduce very significantly the condition number $\kappa(A_{r,h})$ of the matrix $A_{r,h}$ (almost independent of r): see Table 13. This allows to consider very small values of the parameter h without producing any instabilities.

This high robustness of the approximation is definitively in contrast classical dual methods discussed in [34] and the references therein: We recall that for $\varepsilon = 0$, the minimization of $J_{\varepsilon=0}^\star$ defined by (13) fails as soon as h is small enough.

Figure 5 and Figure 6 depict over Q_T the approximation $y_h := \rho^{-1}\lambda_h$ and $v_h := \rho_0^{-1}\psi_h\,1_{q_T}$ for $h = 8.83 \times 10^{-3}$. In particular, the smallness of both the diffusion coefficient and the size of the support ω leads to a large amplitude of the control at the initial time. This is in contrast with the boundary control situation where one acts directly on the state (or its first derivative).

Eventually, in order to validate one more time our computations, we have approximated by a standard time-marching algorithm the solution of (1) with $v = v_h$. Specifically, we have used a C^1-approximation with $\mathbb{P}_{3,x}(0,1)$ elements in space and the second-step implicit Gear scheme (of order two) for the time discretization. Tables 13, 14 and 15 report the L^2-norm of the state at the final time, i.e. $\|y_h(\cdot,T)\|_{L^2(0,1)}$. For each value of r, the L^2-norm decreases linearly to 0 with h. For any h, the non-zero value of $\|y_h(\cdot,T)\|_{L^2(0,1)}$ is, first due to the fact that v_h is not an exact null-control for any discrete system, and second to the consistency error of the approximation used.

h	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}	4.41×10^{-3}	2.2×10^{-3}
$\|\rho^{-1}L^\star(\rho_0\psi_h)\|_{L^2(Q_T)}$	29.76	24.86	21.12	17.92	15.42
$\frac{\|\rho_0(v-v_h)\|_{L^2(q_T)}}{\|\rho_0 v\|_{L^2(q_T)}}$	5.35×10^{-1}	3.34×10^{-1}	2.42×10^{-1}	1.63×10^{-1}	8.45×10^{-2}
$\|\rho_0 v_h\|_{L^2(q_T)}$	15.20	16.642	17.52	18.07	18.43
$\|\rho^{-1}\lambda_h\|_{L^2(Q_T)}$	3.15×10^{-1}	3.34×10^{-1}	3.46×10^{-1}	3.52×10^{-1}	3.56×10^{-1}
$\frac{\|y-\rho^{-1}\lambda_h\|_{L^2(Q_T)}}{\|y\|_{L^2(Q_T)}}$	1.96×10^{-1}	1.20×10^{-1}	6.97×10^{-2}	3.67×10^{-2}	1.49×10^{-2}
♯ CG iterates	52	55	56	56	55
$r^{-1}\delta_{r,h}^{-2}$	27.04	29.37	31.73	33.37	–
$\kappa(A_{r,h})$	9.5×10^4	1.4×10^7	3.03×10^9	1.1×10^{12}	–
n_h=size$(A_{r,h})$	3 444	13 284	52 264	206 724	823 044
$\|y_h(\cdot,T)\|_{L^2(0,1)}$	1.52×10^{-1}	6.109×10^{-2}	2.59×10^{-2}	1.162×10^{-2}	5.41×10^{-3}

Table 13: Mixed formulation (32) - $r = 10^{-2}$ and $\varepsilon = 0$ with $\omega = (0.2, 0.5)$.

The experiments reported here - in the limit case $\varepsilon = 0$ - are obtained for a specific choice of the weights ρ_0 and ρ. Precisely, the weight ρ_0 is such that the approximation $v_h := \rho_0^{-2}\varphi_h\,1_\omega$ vanishes exponentially as $t \to T^-$. This allows in particular to avoid the high oscillatory behavior of the control of minimal L^2-norm, that is when $\rho_0 := 1$ in q_T. The exponential behavior of the control implies a similar behavior of the corresponding controlled stated $\rho^{-1}\lambda$, so that the choice of the parameter ρ made here, is also natural. Remark that ρ is not bounded and therefore does not strictly satisfied the hypothesis of Theorem 2.3. Seemingly, this has no influence at the numerical level. This specific choice of the parameter ρ allows to perform a change of variable and therefore

h	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}	4.41×10^{-3}	2.2×10^{-3}
$\|\rho^{-1}L^\star(\rho_0\psi_h)\|_{L^2(Q_T)}$	3.659	3.276	2.808	2.377	2.002
$\frac{\|\rho_0(v-v_h)\|_{L^2(q_T)}}{\|\rho_0 v\|_{L^2(q_T)}}$	6.97×10^{-1}	4.82×10^{-1}	3.69×10^{-1}	2.81×10^{-1}	2.06×10^{-1}
$\|\rho_0 v_h\|_{L^2(q_T)}$	13.37	15.33	16.62	17.45	17.99
$\|\rho^{-1}\lambda_h\|_{L^2(Q_T)}$	3.35×10^{-1}	3.40×10^{-1}	3.41×10^{-1}	3.42×10^{-1}	3.52×10^{-1}
$\frac{\|y-\rho^{-1}\lambda_h\|_{L^2(Q_T)}}{\|y\|_{L^2(Q_T)}}$	3.28×10^{-1}	2.13×10^{-1}	1.33×10^{-1}	8.09×10^{-2}	4.63×10^{-2}
$\sharp$ CG iterates	12	11	10	9	9
$r^{-1}\delta_{r,h}^{-2}$	2.092	2.062	1.585	1.333	–
$\|y_h(\cdot,T)\|_{L^2(0,1)}$	1.19×10^{-1}	5.39×10^{-2}	2.42×10^{-2}	1.12×10^{-2}	5.29×10^{-3}

Table 14: Mixed formulation (32) - $r = 1.$ and $\varepsilon = 0$ with $\omega = (0.2, 0.5)$.

h	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}	4.41×10^{-3}	2.2×10^{-3}
$\|\rho^{-1}L^\star(\rho_0\psi_h)\|_{L^2(Q_T)}$	0.428	0.426	0.380	0.321	0.215
$\frac{\|\rho_0(v-v_h)\|_{L^2(q_T)}}{\|\rho_0 v\|_{L^2(q_T)}}$	8.83×10^{-1}	6.80×10^{-1}	5.24×10^{-1}	4.16×10^{-1}	3.25×10^{-1}
$\|\rho_0 v_h\|_{L^2(q_T)}$	9.880	12.706	14.82	16.256	17.338
$\|\rho^{-1}\lambda_h\|_{L^2(Q_T)}$	0.2546	0.2926	0.3189	0.3352	0.3477
$\frac{\|y-\rho^{-1}\lambda_h\|_{L^2(Q_T)}}{\|y\|_{L^2(Q_T)}}$	5.86×10^{-1}	4.04×10^{-1}	2.63×10^{-1}	1.66×10^{-1}	9.88×10^{-2}
$\sharp$ CG iterates	10	8	7	5	5
$r^{-1}\delta_{r,h}^{-2}$	2.092	2.007	1.53	1.103	–
$\|y_h(\cdot,T)\|_{L^2(0,1)}$	8.26×10^{-2}	4.24×10^{-2}	2.11×10^{-2}	1.03×10^{-2}	5.12×10^{-3}

Table 15: Mixed formulation (32) - $r = 10^2$ and $\varepsilon = 0$ with $\omega = (0.2, 0.5)$.

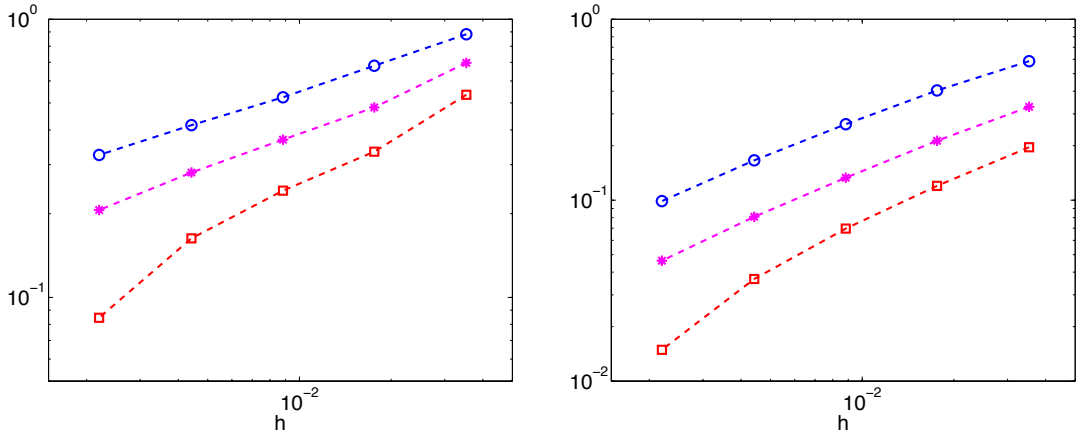

Figure 4: $\omega = (0.2, 0.5)$; $y_0(x) = \sin(\pi x)$: $\varepsilon = 0.$; $\frac{\|\rho_0(v-v_h)\|_{L^2(q_T)}}{\|\rho_0 v\|_{L^2(q_T)}}$ (**Left**) and $\frac{\|y-\rho^{-1}\lambda_h\|_{L^2(Q_T)}}{\|y\|_{L^2(Q_T)}}$ (**Right**) vs. h for $r = 10^2$ (∘), $r = 1.$ (⋆) and $r = 10^{-2}$ (□).

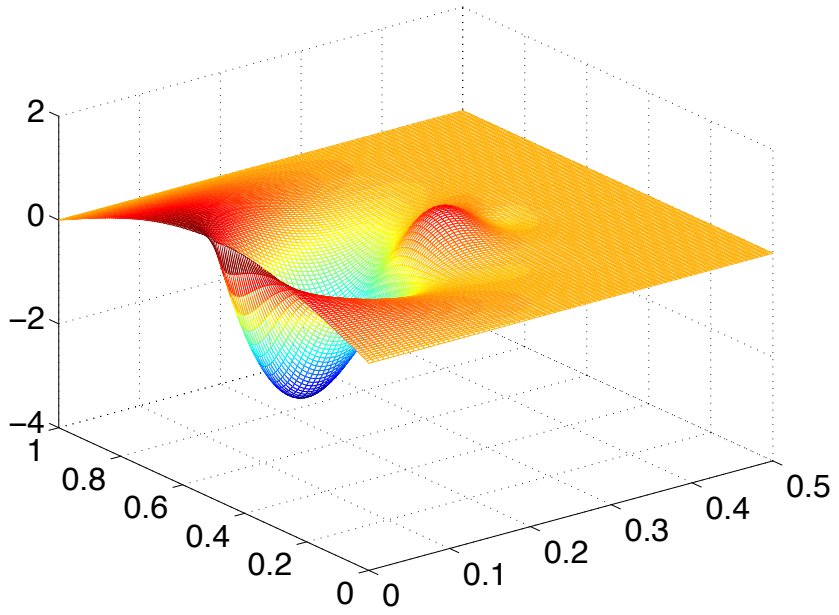

Figure 5: $\omega = (0.2, 0.5)$; Approximation $\rho^{-1}\lambda_h$ of the controlled state y over Q_T - $r = 1$ and $h = 8.83 \times 10^{-3}$.

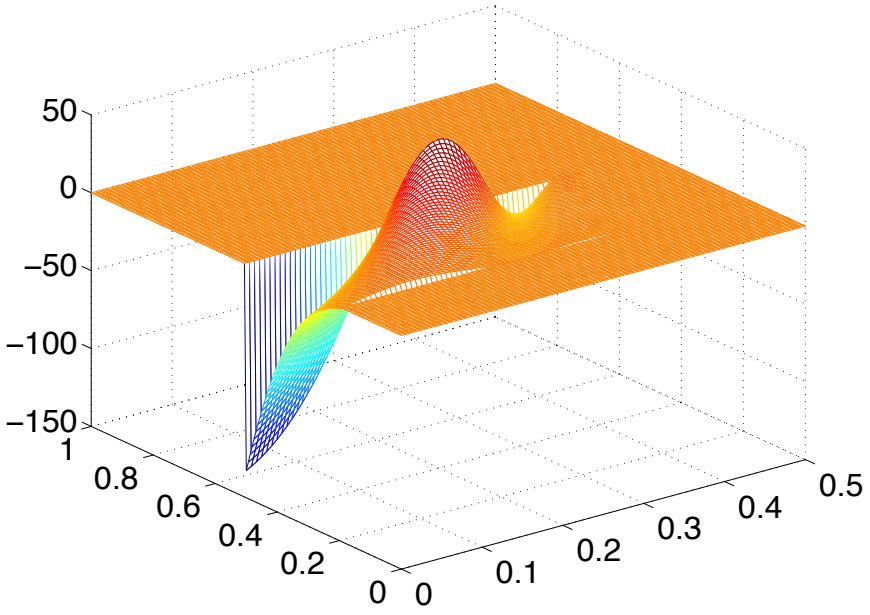

Figure 6: $\omega = (0.2, 0.5)$; Approximation $v_h = \rho_0^{-1}\psi_h$ of the null control v over Q_T - $r = 1$ and $h = 8.83 \times 10^{-3}$.

reduce significantly the condition number of the discrete problem. We also point out that, if the mixed formulation (32) is well-posed for any ρ, ρ_0 satisfying the hypothesis of Theorem 2.3, the constant of continuity of the linear form $\tilde{l}$ depends strongly - in view of the Carleman estimate (34) - of ρ and ρ_0. This affects the convergence and the robustness of the method. Thus, for ρ_0 as before and $\rho := 1$, the condition number is too large for small values of h (typically $h \approx 10^{-3}$) and leads to wrong results. Remark that for $\rho := 1$, the exponential decreases of ρ_0^{-1} cannot be compensated by ρ (see (34)) so that the change of variable is inefficient.

4 Concluding remarks and Perspectives

The mixed formulation we have introduced here in order to address the null controllability of the heat equation seems original and adapted the work [11] devoted to the wave equation. This formulation is nothing else than the Euler system associated to the conjugate functional and depends on both the dual adjoint variable and a Lagrange multiplier, which turns out to be the primal state of the heat equation to be controlled. The approach, recently used in a different way in [18], leads to a variational problem defined over time-space functional Hilbert spaces, without distinction between the time and the space variable. The main ingredients allowing to prove the well-posedness of the mixed formulation are an observability inequality and a direct inequality (usually deduced from energy estimates). For these reasons, the mixed reformulation may also be employed to any other controllable systems for which such inequalities are available. In particular, we may consider the Stokes system as in [32].

At the practical level, the discrete mixed time-space formulation is solved in a systematic way in the framework of the finite element theory: in contrast to the classical approach initially developed in [23], there is no need to take care of the time discretization nor of the stability of the resulting scheme, which is often a delicate issue. The resolution amounts to solve a sparse symmetric linear system : the corresponding matrix can be preconditioned if necessary, and may be computed once for all as it does not depend on the initial data to be controlled. Eventually, as discussed in [11], Section 4.3 (but not employed here), the space-time discretization of the domain allows an adaptation of the mesh so as to reduce the computational cost and capture the main features of the solutions. We also emphasize that the higher dimensional case is very similar as it requires C^1 approximation in space.

The numerical experiments reported in this work suggest a very good behavior of the approach: the strong convergence of the sequences $\{v_h\}_{h>0}$, approximation of the controls of minimal weighted square integrable norm, are clearly observed as the discretization parameter h tends to zero (as the consequence of the uniform inf-sup discrete property). It is worth to mention that, within this mixed formulation approach, the strong convergence of the approximations (as obtained within a closed but different approach in [18] assuming that the weights ρ_0 and ρ coincide with the Carleman weight) is still to be done. From the uniform coercivity of the bilinear form in the primal variable, a strong convergence is guaranteed by a uniform discrete inf-sup property. In view of the complicated and unusual constraint $L^\star\varphi = 0$ and of the C^1 nature of the approximation, the proof of such uniform property is probably very hard to get. However, it seems possible to bypass this property by adding to the Lagrangian the stabilization terms (for instance in the limit case $\varepsilon = 0$)

$$-\|L(\rho^{-1}\lambda_h) - \rho_0^{-2}\varphi_h\, 1_\omega\|^2_{L^2(Q_T)}, \quad -\|\lambda_h(\cdot,0) - y_0\|^2_{L^2(0,1)}$$

which vanish at the continuous level (writing $Ly = v\, 1_\omega$ with $y = \rho^{-1}\lambda$ and $v = \rho^{-2}\varphi 1_\omega$, see Theorem 2.3) and give coercivity property for the variable λ_h. This will be examined in a future work.

The approach may also be extended to the boundary case. We also mention that the variational approach developed here based on a space-time formulation is also very well-adapted to the case

where the support of the inner control evolves in time and takes the form

$$q_T := \{(x,t) \in Q_T; \quad a(t) < x < b(t) \quad t \in (0,T)\}$$

with any a, b in $C^0([0,T],]0,1[)$. We refer to [6] which examines this case for the wave equation.

Eventually, we also mention that this approach which consists in solving directly the optimality conditions of a controllability problem may be employed to solve inverse problems where, for instance, the solution of the heat equation has to be recovered from a partial observation, typically localized on a sub-domain q_T of the working domain: actually, the optimality conditions associated to a least-square type functional can be expressed as a mixed formulation very closed to (2.3). This issue will be analyzed in a future work.

A Appendix

A.1 Appendix : Fourier expansion of the control of minimal $L^2(\rho_0, q_T)$ norm.

We expand in term of Fourier series the control of minimal $L^2(\rho_0, q_T)$ norm v for the (1) and the corresponding controlled solution y. We use these expansions in Section 3.4 to evaluate with respect to h the error $\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}$ and $\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}$ where the sequence $(\varphi_{\varepsilon,h}, \lambda_{\varepsilon,h})$ solves the discrete mixed formulation (40). We use the characterization of the couple $(y_\varepsilon, v_\varepsilon)$ in term of the adjoint solution φ_ε (see (5)), unique minimizer in $L^2(\Omega)$ of $J^\star_\varepsilon$ defined by (13).

We first note $(a_{\varepsilon,p})_{(p>0)}$ the Fourier coefficients in $l^2(\mathbb{N})$ of the minimizer $\varphi_{T,\varepsilon} \in L^2(0,1)$ of (13) such that

$$\varphi_{\varepsilon,T}(x) = \sum_{p>0} a_{\varepsilon,p} \sin(p\pi x), \quad x \in (0,1). \tag{51}$$

The adjoint state takes the form $\varphi_\varepsilon(x,t) = \sum_{p\geq 1} a_{\varepsilon,p} e^{c\pi^2 p^2 (t-T)} \sin(p\pi x)$ in Q_T.

The optimality equation associated to the functional $J^\star_\varepsilon$ then reads,

$$DJ^\star_\varepsilon(\varphi_{\varepsilon,T}) \cdot \overline{\varphi_T} = \iint_{q_T} \rho_0^{-2} \varphi_\varepsilon \overline{\varphi}\, dx\, dt + \varepsilon \int_0^1 \varphi_{\varepsilon,T} \overline{\varphi_T} + (y_0, \overline{\varphi}(\cdot,0)) = 0, \forall \overline{\varphi_T} \in L^2(0,1)$$

and can be rewritten in terms of the $(a_{\varepsilon,p})_{p>0}$ as follows :

$$< \{\overline{a_p}\}_{p>0}, \mathcal{M}_{q_T,\varepsilon}\{a_p\}_{p>0} >=< \{\overline{a_p}\}_{p>0}, \mathcal{F}_{y_0} > \quad \forall \overline{a}_{\varepsilon,p} \in l^2(\mathbb{N}) \tag{52}$$

where $\mathcal{M}_{q_T,\varepsilon}$ denotes a symmetric positive definite matrix and $\mathcal{F}_{y_0}$ a vector obtained from the expansion (51). The resolution of the infinite dimensional system (reduced to a finite dimension one by truncation of the sums) allows an approximation of the minimizer $\varphi_{T,\varepsilon}$ of $J^\star_\varepsilon$.

Finally, we use that the control of minimal $L^2(\rho_0, q_T)$ norm is given by $v_\varepsilon = \rho_0^{-2} \varphi_\varepsilon\, 1_\omega$ and find that the corresponding controlled solution may be expanded as follows

$$y_\varepsilon(x,t) = \sum_{q>0} \left(e^{-c\pi^2 q^2 t} b^0_q + \sum_{p\geq 1} a_{\varepsilon,p} c_{q,p}(\omega) d_{q,p}(t)] \right) \sin(p\pi x), \quad (x,t) \in Q_T \tag{53}$$

with

$$c_{p,q}(\omega) := 2\int_\omega \sin(p\pi x)\sin(q\pi x)\, dx; \quad d_{p,q}(t) := \int_0^t \rho_0^{-2}(s) e^{c\pi^2 (p^2(s-T)+q^2(s-t))}\, ds, \quad t \in (0,T).$$

$(b^0_q)_{q>0}$ denotes the Fourier coefficients of the initial data $y_0 \in L^2(0,1)$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star \varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	3.84×10^{-2}	2.90×10^{-2}	9.27×10^{-3}	2.41×10^{-3}	7.78×10^{-4}
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	1.32×10^{-1}	5.90×10^{-2}	3.24×10^{-2}	1.68×10^{-2}	8.57×10^{-3}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	1.04×10^{-1}	3.54×10^{-2}	1.48×10^{-2}	7.59×10^{-3}	3.89×10^{-3}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	2.02×10^{-1}	1.68×10^{-1}	1.65×10^{-1}	1.67×10^{-1}	1.68×10^{-1}
κ_ε	4.44×10^{9}	4.20×10^{11}	3.84×10^{13}	3.25×10^{15}	5.72×10^{16}

Table 16: Mixed formulation (16) - $r = 10^2$ and $\varepsilon = 10^{-2}$ with $\omega = (0.2, 0.5)$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star \varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	6.19×10^{-2}	1.57×10^{-1}	1.56×10^{-1}	1.50×10^{-1}	6.21×10^{-2}
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	1.02	7.36×10^{-1}	3.65×10^{-1}	1.52×10^{-1}	3.01×10^{-2}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	6.74×10^{-1}	5.51×10^{-1}	2.42×10^{-1}	1.05×10^{-1}	1.81×10^{-2}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	2.23×10^{-1}	1.76×10^{-1}	7.86×10^{-2}	4.87×10^{-2}	3.28×10^{-2}
κ_ε	5.31×10^{9}	8.31×10^{11}	9.64×10^{13}	1.47×10^{16}	1.50×10^{18}

Table 17: Mixed formulation (16) - $r = 10^2$ and $\varepsilon = 10^{-4}$ with $\omega = (0.2, 0.5)$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star \varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	6.23×10^{-2}	1.63×10^{-1}	1.77×10^{-1}	2.66×10^{-1}	2.24×10^{-1}
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	1.50	1.11	9.53×10^{-1}	8.33×10^{-1}	7.19×10^{-1}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	1.08	1.09	9.4×10^{-1}	7.69×10^{-1}	5.15×10^{-1}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	2.24×10^{-1}	1.79×10^{-1}	8.10×10^{-2}	5.67×10^{-2}	1.71×10^{-2}
κ_ε	5.32×10^{9}	8.59×10^{11}	9.86×10^{13}	1.84×10^{16}	3.07×10^{18}

Table 18: Mixed formulation (16) - $r = 10^2$ and $\varepsilon = 10^{-8}$ with $\omega = (0.2, 0.5)$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star \varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	2.86×10^{-1}	7.15×10^{-2}	1.84×10^{-2}	4.86×10^{-3}	1.40×10^{-3}
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	1.11×10^{-1}	6.21×10^{-2}	3.29×10^{-2}	1.68×10^{-2}	8.57×10^{-3}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	5.16×10^{-2}	2.84×10^{-2}	1.48×10^{-2}	7.59×10^{-3}	3.89×10^{-3}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	1.53×10^{-1}	1.61×10^{-1}	1.65×10^{-1}	1.67×10^{-1}	1.68×10^{-1}
κ_ε	9.15×10^{8}	2.07×10^{10}	8.05×10^{11}	3.25×10^{13}	1.45×10^{15}

Table 19: Mixed formulation (16) - $r = 10^{-2}$ and $\varepsilon = 10^{-2}$ with $\omega = (0.2, 0.5)$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star \varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	10.77	3.821	1.018	2.59×10^{-1}	6.56×10^{-2}
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	4.63×10^{-1}	2.23×10^{-1}	1.10×10^{-1}	5.52×10^{-2}	2.74×10^{-2}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	1.55×10^{-1}	9.03×10^{-2}	4.08×10^{-2}	2.46×10^{-2}	1.27×10^{-2}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	3.22×10^{-2}	2.85×10^{-2}	2.99×10^{-2}	3.08×10^{-2}	3.12×10^{-2}
κ_ε	3.04×10^{9}	1.33×10^{11}	7.55×10^{12}	3.88×10^{14}	1.96×10^{16}

Table 20: Mixed formulation (16) - $r = 10^{-2}$ and $\varepsilon = 10^{-4}$ with $\omega = (0.2, 0.5)$.

h	1.41×10^{-1}	7.07×10^{-2}	3.53×10^{-2}	1.76×10^{-2}	8.83×10^{-3}
$\|L^\star \varphi_{\varepsilon,h}\|_{L^2(Q_T)}$	21.872	19.388	26.098	28.310	21.249
$\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}$	14.989	9.459	6.606	4.175	1.556
$\frac{\|\rho_0(v_\varepsilon - v_{\varepsilon,h})\|_{L^2(q_T)}}{\|\rho_0 v_\varepsilon\|_{L^2(q_T)}}$	1.33	8.43×10^{-1}	5.89×10^{-1}	3.72×10^{-1}	1.38×10^{-1}
$\frac{\|y_\varepsilon - \lambda_{\varepsilon,h}\|_{L^2(Q_T)}}{\|y_\varepsilon\|_{L^2(Q_T)}}$	5.73×10^{-1}	$4.71 \times 10^{-}$	3.51×10^{-1}	2.11×10^{-1}	6.82×10^{-2}
$\|\lambda_{\varepsilon,h}(\cdot,T)\|_{L^2(0,1)}$	3.31×10^{-2}	1.31×10^{-2}	5.99×10^{-3}	2.83×10^{-3}	8.26×10^{-4}
κ_ε	4.08×10^{9}	3.04×10^{11}	4.54×10^{13}	6.79×10^{15}	1.30×10^{18}

Table 21: Mixed formulation (16) - $r = 10^{-2}$ and $\varepsilon = 10^{-8}$ with $\omega = (0.2, 0.5)$.

A.2 Appendix: Tables

Acknowledgments: This work has been partially done while the second author was visiting the Blaise Pascal University (Clermont-Ferrand, France). He wishes to thank the members of the Laboratory of Mathematics for their kind hospitality.

References

[1] F. Ben Belgacem and S. M. Kaber, *On the Dirichlet boundary controllability of the one-dimensional heat equation: semi-analytical calculations and ill-posedness degree*, Inverse Problems, 27 (2011), pp. 055012, 19.

[2] F. Boyer, *On the penalised HUM approach and its applications to the numerical approximation of null-controls for parabolic problems*, in CANUM 2012, Super-Besse, ESAIM Proc., EDP Sci., Les Ulis, 2013.

[3] F. Boyer, F. Hubert, and J. Le Rousseau, *Uniform controllability properties for space/time-discretized parabolic equations*, Numer. Math., 118 (2011), pp. 601–661.

[4] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.

[5] C. Carthel, R. Glowinski, and J.-L. Lions, *On exact and approximate boundary controllabilities for the heat equation: a numerical approach*, J. Optim. Theory Appl., 82 (1994), pp. 429–484.

[6] C. Castro, N. Cîndea, and A. Münch, *Controllability of the linear wave equation with moving inner actions.*, (Preprint available at http://hal.archives-ouvertes.fr/hal-00927076).

[7] T. Cazenave and A. Haraux, *An introduction to semilinear evolution equations*, vol. 13 of Oxford Lecture Series in Mathematics and its Applications, The Clarendon Press Oxford University Press, New York, 1998. Translated from the 1990 French original by Yvan Martel and revised by the authors.

[8] D. Chapelle and K.-J. Bathe, *The inf-sup test*, Comput. & Structures, 47 (1993), pp. 537–545.

[9] P. G. Ciarlet, *The finite element method for elliptic problems*, vol. 40 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[10] N. Cîndea, E. Fernández-Cara, and A. Münch, *Numerical controllability of the wave equation through primal methods and carleman estimates*, ESAIM Control Optim. Calc. Var., 19 (2013), pp. 1076–1108.

[11] N. Cîndea and A. Münch, *A mixed formulation for the direct approximation of the control of minimal L^2-norm for linear type wave equations*, To appear in Calcolo (http://hal.archives-ouvertes.fr/hal-00853767).

[12] J.-M. Coron, *Control and nonlinearity*, vol. 136 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2007.

[13] J. W. Daniel, *The approximate minimization of functionals*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1971.

[14] D. A. Dunavant, *High degree efficient symmetrical Gaussian quadrature rules for the triangle*, Internat. J. Numer. Methods Engrg., 21 (1985), pp. 1129–1148.

[15] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1996.

[16] S. Ervedoza and J. Valein, *On the observability of abstract time-discrete linear parabolic equations*, Rev. Mat. Complut., 23 (2010), pp. 163–190.

[17] E. Fernández-Cara and A. Münch, *Numerical null controllability of semi-linear 1-D heat equations: fixed point, least squares and Newton methods*, Math. Control Relat. Fields, 2 (2012), pp. 217–246.

[18] E. Fernández-Cara and A. Münch, *Strong convergence approximations of null controls for the 1D heat equation*, SēMA J., 61 (2013), pp. 49–78.

[19] ——, *Numerical exact controllability of the 1d heat equation: Carleman weights and duality*, Journal of Optimization Theory and Applications, (2014).

[20] M. Fortin and R. Glowinski, *Augmented Lagrangian methods*, vol. 15 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, 1983. Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer.

[21] A. V. Fursikov and O. Y. Imanuvilov, *Controllability of evolution equations*, vol. 34 of Lecture Notes Series, Seoul National University Research Institute of Mathematics Global Analysis Research Center, Seoul, 1996.

[22] R. Glowinski, *Handbook of numerical analysis. Vol. IX*, Handbook of Numerical Analysis, IX, North-Holland, Amsterdam, 2003. Numerical methods for fluids. Part 3.

[23] R. Glowinski and J.-L. Lions, *Exact and approximate controllability for distributed parameter systems*, in Acta numerica, 1995, Acta Numer., Cambridge Univ. Press, Cambridge, 1995, pp. 159–333.

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

[25] S. Kindermann, *Convergence rates of the Hilbert uniqueness method via Tikhonov regularization*, J. Optim. Theory Appl., 103 (1999), pp. 657–673.

[26] S. Labbé and E. Trélat, *Uniform controllability of semidiscrete approximations of parabolic control systems*, Systems Control Lett., 55 (2006), pp. 597–609.

[27] I. Lasiecka and R. Triggiani, *Control theory for partial differential equations: continuous and approximation theories. I*, vol. 74 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2000. Abstract parabolic systems.

[28] G. Lebeau and L. Robbiano, *Contrôle exact de l'équation de la chaleur*, Comm. Partial Differential Equations, 20 (1995), pp. 335–356.

[29] J.-L. Lions, *Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. Tome 1*, vol. 8 of Recherches en Mathématiques Appliquées [Research in Applied Mathematics], Masson, Paris, 1988. Contrôlabilité exacte. [Exact controllability], With appendices by E. Zuazua, C. Bardos, G. Lebeau and J. Rauch.

[30] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. III*, Springer-Verlag, New York-Heidelberg, 1973. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 183.

[31] S. Micu and E. Zuazua, *On the regularity of null-controls of the linear 1-d heat equation*, C. R. Math. Acad. Sci. Paris, 349 (2011), pp. 673–677.

[32] A. Münch, *A least-squares formulation for the approximation of controls for the stokes system*, Math. Control Signals Systems, 27 (2014), pp. 1–27.

[33] A. Münch and P. Pedregal, *Numerical null controllability of the heat equation through a least squares and variational approach*, European J. Appl. Math., 25 (2014), pp. 277–306.

[34] A. Münch and E. Zuazua, *Numerical approximation of null controls for the heat equation: ill-posedness and remedies*, Inverse Problems, 26 (2010), pp. 085018, 39.

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

[36] E. Zuazua, *Control and numerical approximation of the wave and heat equations*, in International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 1389–1417.