Uniform null-controllability properties for space/time-discretized parabolic equations

Franck Boyer, Florence Hubert, Jérôme Le Rousseau

UNIFORM NULL-CONTROLLABILITY PROPERTIES FOR SPACE/TIME-DISCRETIZED PARABOLIC EQUATIONS*

FRANCK BOYER[†§], FLORENCE HUBERT[‡§], AND JÉRÔME LE ROUSSEAU[¶]

Abstract. This article is concerned with the analysis of semi-discrete-in-space and fully-discrete approximations of the null controllability (and controllability to the trajectories) for parabolic equations. We propose an abstract setting for space discretizations that potentially encompasses various numerical methods and we study how the controllability problems depend on the discretization parameters. For time discretization we use θ-schemes with $\theta \in [\frac{1}{2}, 1]$.

For the proofs of controllability we rely on the strategy introduced in 1995 by G. Lebeau and L. Robbiano for the null-controllability of the heat equation, which is based on a spectral inequality. We obtain relaxed uniform observability estimates in both the semi-discrete and fully-discrete frameworks, and associated uniform controllability properties.

For the practical computation of the control functions we follow J.-L. Lions' Hilbert Uniqueness Method strategy. Algorithms for the computation of the controls are proposed and analysed in the semi-discrete and fully-discrete cases. Additionally, we prove error estimates with respect to the time step for the control functions obtained in these two cases. The theoretical results are illustrated through numerical experimentations.

Key words. Discrete Lebeau-Robbiano spectral inequality, parabolic equation, semi-discrete scheme, fully-discrete scheme, uniform controllability / observability.

AMS subject classifications. 35K05 - 65M06 - 93B05 - 93B07 - 93B40

1. Introduction.

The null-controllability of parabolic equations was proven in the 90's in two seminal works, [LR95] and [FI96]. Let Ω, ω be connected non-empty bounded open subsets of $\mathbb{R}^n$ with $\omega \Subset \Omega$ and consider the following parabolic distributed control problem in $(0,T)\times\Omega$, with $T>0$,

$$\partial_t y - \nabla_x \cdot (\gamma \nabla_x y) = \mathbf{1}_\omega v \text{ in } (0,T)\times\Omega, \quad y|_{\partial\Omega} = 0, \quad \text{and} \quad y|_{t=0} = y_0. \tag{1.1}$$

The null controllability states that for all $y_0 \in L^2(\Omega)$, there exists $v \in L^2((0,T)\times\Omega)$, such that $y(T)=0$ and $\|v\|_{L^2((0,T)\times\Omega)} \leq C|y_0|_{L^2(\Omega)}$, where $C>0$ only depends on Ω, ω, γ and T.

If we consider discretized version of this parabolic problem we hope to retain some of the controllability result features. It is known however that controllability and discretization do not "commute" well. This question has been quite extensively studied in the context of hyperbolic equations and yet very little for parabolic equations.

Regarding space discretization, let us mention the work of [LZ98b], where the null controllability of the heat equation with a constant diffusion coefficient γ is proven for a finite-difference scheme in one dimension on a uniform mesh. In higher dimension, a counter-example for finite differences due to O. Kavian (see e.g. [Zua06]) shows

Date: November 1, 2009

*The three authors were partially supported by l'Agence Nationale de la Recherche under grant ANR-07-JCJC-0139-01. The CNRS Pticrem project facilitated the writting of this article (`http://pticrem.math.cnrs.fr/`).

[†]Université Paul Cézanne (`fboyer@latp.univ-mrs.fr`).

[‡]Université de Provence (`fhubert@latp.univ-mrs.fr`).

[§]Laboratoire d'Analyse Topologie Probabilités (LATP), CNRS UMR 6632, Universités d'Aix-Marseille, 39 rue F. Joliot-Curie, 13453 Marseille cedex 13, France.

[¶]Université d'Orléans, Laboratoire Mathématiques et Applications, Physique Mathématique d'Orléans, CNRS UMR 6628, Fédération Denis Poisson, FR CNRS 2964, B.P. 6759 - 45067 Orléans cedex 2, France (`jlr@univ-orleans.fr`).

that localized eigenfunctions for the discrete Laplace operator are an obstruction to null controllability with an arbitrary control region ω. These problematic discrete eigenfunctions correspond to the high end of the discrete spectrum. The authors of the present article lately proved a result of null-controllability for a constant portion of the lower part of the spectrum with an arbitrary control region ω (see [BHL09a] for the one-dimensional case and [BHL09b] for higher dimensions). The L^2 norm of the control function v_h is estimated by $C\|y_0\|_{L^2}$, with C uniform with respect to the spatial discretization step h. Moreover, the resulting final state, $y_h(T)$ decays super-algebraically to zero as h goes to zero. This also yields a relaxed observation inequality which form resembles the case studied in [LT06] for general controlled semi-discrete scheme.

The results of [BHL09a, BHL09b] are based on a discrete extension of a spectral inequality due to G. Lebeau and L. Robbiano [LR95] (see also [LZ98a, JL99]). This extension to discrete elliptic operator is only partial and holds for a constant portion of the lower part of the spectrum, hence the form of the null-controllability result. The proof of this partial spectral inequality is based on semi-discrete Carleman estimates for elliptic operators. The null-controllability is obtained via the Lebeau-Robbiano strategy that takes advantage of parabolic dissipation.

There is also some work on the time discretization of controlled parabolic systems. In [Zhe08], the author studies the time-discretized Lebeau-Robbiano strategy. A filtering of the high frequencies is required (in the spectral representation of the *continuous* Laplace operator) and the convergence results obtained are far from optimal. More interesting is the result of [EV09], where the authors prove that any controllable parabolic equation, be it discrete or continuous in space, is null controllable after time discretization upon the application of an appropriate filtering of the high-frequencies (in the spectral representation of the *continuous* or *discrete* Laplace operator). In [EV09] there is however no study of the convergence of the control function as the time step goes to zero.

Here, we consider fully-discrete schemes and we avoid high-frequency filtering. For the time discretization we shall use a θ-scheme with $\frac{1}{2} \leq \theta \leq 1$, *i.e.*, ranging from the Crank-Nicolson to the implicit Euler scheme,

$$\begin{cases} y^0 = y_0, \\ \mathcal{M}_h \dfrac{y^{n+1} - y^n}{\delta t} + \mathcal{A}_h(\theta y^{n+1} + (1-\theta) y^n) = \mathcal{B}_h v^{n+1}, \ \forall n \in [\![0, M-1]\!]. \end{cases} \tag{1.2}$$

The matrix $\mathcal{M}_h$ should be understood as a mass matrix, $\mathcal{A}_h$ as the discrete version of the elliptic operator and $\mathcal{B}_h$ as the control operator. In space, the discretization we use is assumed to yield a partial spectral inequality like that proven in [BHL09a, BHL09b]. This inequality, with the control strategy of Lebeau and Robbiano, yields the controllability of the fully discrete system (1.2) up to a small remainder that decays exponentially as the spatial step size goes to zero. This result thus compares to that obtain in the semi-discrete case in [BHL09a, BHL09b].

We then deduce an observability inequality of the form

$$\left\| q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \right\|_h \leq C_{\text{obs}} \left(\sum_{n=1}^{M} \delta t \left\| \mathcal{B}_h^\star q^n \right\|_h^2 \right)^{\frac{1}{2}} + C_1 e^{-C_2/h^\gamma} \left\| q_F \right\|_h ,$$

for some $\gamma > 0$, where $q = (q^n)_{1\leq n\leq M+1}$ is solution of the

$$\begin{cases} \mathcal{M}_h \dfrac{q^M - q^{M+1}}{\delta t} + \theta \mathcal{A}_h q^M = 0, \\ \mathcal{M}_h \dfrac{q^n - q^{n+1}}{\delta t} + \mathcal{A}_h(\theta q^n + (1-\theta) q^{n+1}) = 0, \quad \forall n \in [\![1, M-1]\!]. \end{cases} \tag{1.3}$$

This system is the proper adjoint system for (1.2) (compare with [Zhe08, Section 7.2]). Following the so-called Hilbert Uniqueness Method (HUM) (see for instance [GL94]), minimizing a certain quadratic functional then allows us to build a control function by solving the coupled system of primal and adjoint parabolic equations.

The following results are obtained:

- *The bound of the L^2 norm of the fully-discrete control is independent of the discretization parameters δt and h.*
- *The distance to the target of the final state of the controlled solution is small with respect to h, uniformly in δt.*

Computing error estimates for the forward problem (1.2) and backward problem (1.3) we then prove the following convergence results for the fully-discrete control $v_{\delta t} = (v^n)_n$, $n \in [\![1, M]\!]$.

The control of the fully-discrete parabolic equation converges to that of the semi-discrete equation with:

- *a* first-*order rate for the θ-scheme, $\frac{1}{2} < \theta \leq 1$;*
- *a* second-*order rate for the Crank-Nicolson scheme, $\theta = \frac{1}{2}$.*

Note however that these convergence results are not uniform with respect to the spatial discretization.

Finally, let us mention [FCM] where the control problem is addressed through the minimization of a weighted functional. Numerical experiments then indicate a better behaved control function in time. The weight in the functional is introduced in connection to the global parabolic Carleman estimates of [FI96].

1.1. Outline. In Section 2 we introduce the setting of the article, in particular the discrete setting. We also present some examples of applications. In Section 3 we show how the partial spectral inequality can be used to prove the null controllability of the lower part of the spectrum and the resulting observability in the semi-discrete case. The fully-discrete problem is first considered in Section 4. This section is devoted to (i) the derivation of a proper adjoint problem and the duality between controllability and observability and (ii) the extension of the Lebeau-Robbiano control strategy to the fully-discrete case. A large part of the proof lays in an effort to recover an exponential decay for the remainder at final time. The control result yields a relaxed observability inequality. This inequality is exploited in Section 5 for the actual computation of a fully-discrete control to the trajectories. In Section 6 we derive error estimates for the forward problem and the backward problem. The Crank-Nicolson scheme is studied with special care. These estimates then yield error estimates between the semi-discrete and fully-discrete control functions. In Section 7, numerical results are presented illustrating the theoretical results we have obtained.

2. Notation and examples.

2.1. Notation. For any $h > 0$ (which is supposed to represent the space discretization parameter), let us consider:

- A Euclidean space E_h, whose inner product and its associated norm are denoted by $(\cdot,\cdot)_h$ and $|\cdot|_h$ respectively. The dimension of E_h is denoted by N_h. In practice, $N_h \to \infty$ when $h \to 0$ but this is not a necessary assumption.
- Two linear operators $\mathcal{M}_h, \mathcal{A}_h$ on E_h which are supposed to be symmetric and definite positive for the scalar product $(\cdot,\cdot)_h$.
- The scalar product and the norm induced by M_h, defined as follows

$$\forall x, y \in E_h, \quad \langle x, y\rangle_h = (\mathcal{M}_h x, y)_h\,, \quad \|x\|_h = \langle x, x\rangle_h^{\frac{1}{2}} = \left|\mathcal{M}_h^{\frac{1}{2}} x\right|_h .$$

- The previous scalar product and the norm can be generalized by introducing arbitrary powers of $\mathcal{M}_h^{-1}\mathcal{A}_h$ as follows

$$\forall x, y \in E_h, \quad \langle x, y\rangle_{s,h} = \left\langle (\mathcal{M}_h^{-1}\mathcal{A}_h)^s x, y\right\rangle_h ,$$

$$\|x\|_{s,h} = \langle x, x\rangle_{s,h}^{\frac{1}{2}} = \left\|(\mathcal{M}_h^{-1}\mathcal{A}_h)^{\frac{s}{2}} x\right\|_h ,$$

  for any $s \in \mathbb{R}$. Observe that $\mathcal{M}_h^{-1}\mathcal{A}_h$ is symmetric for $\langle .,.\rangle_h$.
- The spectral radius of $\mathcal{M}_h^{-1}\mathcal{A}_h$ is denoted by ρ_h. We assume that there exists $\rho_0 > 0$ such that

$$\rho_h \geq \rho_0,\ \forall h > 0. \tag{2.1}$$

  Notice that, this assumption is reasonable since in all practical cases, we have $\rho_h \to +\infty$ when $h \to 0$.
- A second Euclidean space U_h, whose inner product and its associated norm are denoted by $[\cdot,\cdot]_h$ and $[\![\cdot]\!]_h$.
- A linear operator $\mathcal{B}_h : U_h \to E_h$. We denote by $\mathcal{B}_h^\star$ its adjoint with respect to the initial scalar product on E_h and the scalar product on U_h, that is,

$$\forall u \in U_h, \forall x \in E_h, \quad (\mathcal{B}_h u, x)_h = [\mathcal{B}_h^\star x, u]_h .$$

- For a linear map $\mathcal{F} : E_h \to E_h$ we denote by $\|\mathcal{F}\|_h$ its application norm with respect to the norm $\|\cdot\|_h$ on E_h.
- For a linear map $\mathcal{G} : U_h \to E_h$ we denote by $|\!|\!|\mathcal{G}|\!|\!|_h$ its application norm with respect to the norm $\|\cdot\|_h$ on E_h and the norm $[\![\cdot]\!]_h$ on U_h.
- We shall assume that the following condition is ensured

$$\sup_{h>0} |\!|\!|\mathcal{M}_h^{-1}\mathcal{B}_h|\!|\!|_h < +\infty. \tag{2.2}$$

  This assumption is equivalent to the fact that there is $C > 0$ such that

$$[\![\mathcal{B}_h^\star x]\!]_h \leq C\,\|x\|_h\,, \quad \forall h > 0, \forall x \in E_h, \tag{2.3}$$

  that is to say that $\mathcal{B}_h^\star$ is a family of uniformly bounded operators.

With the above properties, there exists an orthonormal basis of $(E_h, \langle\cdot,\cdot\rangle_h)$ whose elements are eigenvectors of $\mathcal{M}_h^{-1}\mathcal{A}_h$. Such a basis will be denoted by $(\psi_i)_{1\leq i\leq N_h}$ and the corresponding eigenvalues by $(\mu_i)_{1\leq i\leq N_h}$. These eigenvalues are positive and we assume that they are sorted in a non-decreasing order.

We are interested here in the following controllability problem: Given $y_0 \in E_h$, find $v : [0,T] \to U_h$, such that the solution $y : [0,T] \to E_h$ of the following problem

$$\begin{cases} \mathcal{M}_h \partial_t y + \mathcal{A}_h y = \mathcal{B}_h v, \\ y(0) = y_0, \end{cases} \tag{2.4}$$

satisfies $y(T) = 0$ and such that $\|v\|_{L^2(0,T,U_h)} \leq C \|y_0\|_h$, for a constant $C > 0$ which is independent of h.

In fact, in many practical situations, this problem may not have a solution. However, with suitable assumptions on the operators, that we shall state below, there exists a control function v, uniformly bounded in h, such that $\|y(T)\|_h$ is exponentially small with respect to h.

We conclude this introductory section with some elementary inequalities that will be useful in the sequel.

LEMMA 2.1.

1. *The following inequality holds*

$$\|x\|_{s+\alpha,h} \leq \rho_h^{\frac{\alpha}{2}} \|x\|_{s,h}, \quad \forall x \in E_h, \forall s \in \mathbb{R}, \forall \alpha \geq 0. \tag{2.5}$$

2. *The following interpolation inequalities hold*

$$|\langle x, y \rangle_{s,h}| \leq \|x\|_{s-\alpha,h} \|y\|_{s+\alpha,h}, \quad \forall x, y \in E_h, \forall s, \alpha \in \mathbb{R}, \tag{2.6}$$

$$\|x\|_{\alpha s+(1-\alpha)t,h} \leq \|x\|_{s,h}^{\alpha} \|x\|_{t,h}^{1-\alpha}, \quad \forall x \in E_h, \forall s, t \in \mathbb{R}, \forall \alpha \in [0,1]. \tag{2.7}$$

The proof of the interpolation inequality can be readily obtained by using a basis of orthonormal eigenvectors for $\mathcal{M}_h^{-1}\mathcal{A}_h$ and the Hölder inequality.

2.2. Examples. The main examples we have in mind when introducing the above framework are those of some space discretization techniques. More precisely, we consider the following parabolic control problem in a bounded polygonal domain $\Omega \subset \mathbb{R}^d$

$$\begin{cases} \partial_t y - \nabla_x(\gamma(x)\nabla_x y) = 1_\omega v, \text{ in } (0,T) \times \Omega, \\ y = 0, \text{ on } (0,T) \times \partial\Omega, \\ y|_{t=0} = y_0, \quad \text{in } \Omega, \end{cases}$$

where $\omega \subset \Omega$ is a distributed control domain, $x \mapsto \gamma(x)$ is the diffusion coefficient and v is the control function that we wish to characterize. The map γ may be matrix-valued but, for simplicity, we assume here that γ is a scalar coefficient. We consider now semi-discretization in space for such a problem.

2.2.1. Finite differences schemes. In the case when Ω has a Cartesian geometry, for instance $\Omega = (0,1)^2$ in dimension 2, then one may be interested in using the elementary finite-difference method. This method will lead to a semi-discrete problem of the form (2.4), with

- $E_h = \mathbb{R}^N$, $N = n_1 \times n_2$ being the total number of discretization cells in the domain equipped with the inner product $(x,y)_h = \sum_{i,j} h_i h_j x_{i,j} y_{i,j}$, where (h_i, h_j) is the size of the cell labelled (i,j).
- $U_h = \mathbb{R}^k$, k being the number of discretization cells which intersect the control domain ω equipped with the same inner product as E_h.
- $\mathcal{A}_h \in M_N(\mathbb{R})$ is the classical 5-diagonal matrix given by

$$(\mathcal{A}_h y)_{i,j} = -\frac{\gamma_{i+\frac{1}{2},j}\dfrac{y_{i+1,j}-y_{i,j}}{h_{i+\frac{1}{2}}} - \gamma_{i-\frac{1}{2},j}\dfrac{y_{i,j}-y_{i-1,j}}{h_{i-\frac{1}{2}}}}{h_i} - \frac{\gamma_{i,j+\frac{1}{2}}\dfrac{y_{i,j+1}-y_{i,j}}{h_{j+\frac{1}{2}}} - \gamma_{i,j-\frac{1}{2}}\dfrac{y_{i,j}-y_{i,j-1}}{h_{j-\frac{1}{2}}}}{h_j},$$

where $h_{i+\frac{1}{2}}$ is the distance between the centers of the cells (i,j) and $(i+1,j)$ and $h_{j+\frac{1}{2}}$ is the distance between the cells (i,j) and $(i,j+1)$. The boundary conditions are taken into account in those formulas by imposing that $y_{0,j} = y_{n_1+1,j} = 0$, $1 \leq j \leq n_2$ and $y_{i,0} = y_{i,n_2+1} = 0$, $1 \leq i \leq n_1$.

- $\mathcal{M}_h \in M_N(\mathbb{R})$ is the identity matrix.
- $\mathcal{B}_h \in M_{N,k}(\mathbb{R})$ is the rectangle matrix corresponding the natural embedding of ω in Ω.

Notice that condition (2.2) is automatically satisfied in this case. Furthermore the spectral radius ρ_h behaves like h^{-2} in this case.

2.2.2. Galerkin methods. We consider a finite dimensional subspace X_h of the Sobolev space $H_0^1(\Omega)$ and a finite dimensional subspace Y_h of the space $L^2(\Omega)$ and we denote by $(\phi_i^h)_i \subset X_h$ and $(\psi_j^h)_j \subset Y_h$ two basis for these spaces, respectively. Such spaces and associated basis might be obtained through finite elements methods in connection to some mesh of Ω or through spectral methods if the geometry of Ω is simple enough. This situation enters the general framework proposed above by choosing:

- $E_h = \mathbb{R}^N$ where $N = \dim X_h$, the elements in E_h being the coordinates vectors of the elements of X_h in the basis $(\phi_i^h)_i$ and the inner product $(\cdot,\cdot)_h$ is the usual Euclidean one.
- $U_h = \mathbb{R}^k$ where $k = \dim Y_h$, the elements in U_h representing the coordinates of elements in Y_h in the basis $(\psi_j^h)_j$ and the inner product $[\cdot,\cdot]_h$ is the usual Euclidean one.
- The matrix $\mathcal{M}_h \in M_N(\mathbb{R})$ is the mass matrix associated to $(\phi_i^h)_i$ that is the matrix whose entries are $\int_\Omega \phi_i^h \phi_j^h \, dx$.
- The matrix $\mathcal{B}_h \in M_{N,k}(\mathbb{R})$ is the matrix whose entries are $\int_\omega \phi_i^h \psi_j^h \, dx$.
- The matrix $\mathcal{A}_h \in M_N(\mathbb{R})$ is the so-called rigidity matrix associated to the diffusion operator, whose entries are given by $\int_\Omega \gamma(x) \nabla \phi_i^h \cdot \nabla \phi_j^h \, dx$.

Notice that condition (2.2) is also automatically satisfied in this case since it corresponds to the fact that the multiplication by 1_ω is a bounded operator in $L^2(\Omega)$. In the standard situation where X_h is built upon a $\mathbb{P}^1$ finite-element approximation space associated to a triangulation of Ω, the spectral radius ρ_h also behaves like h^{-2}. Notice that the spaces X_h and Y_h can be chosen in an independent way.

This framework can also be slightly modified if one uses the so-called mass lumping technique, which consists in replacing the mass matrix $\mathcal{M}_h$ by a diagonal matrix whose entries are the sum of the entries of $\mathcal{M}_h$ in each line. In that case, assumption (2.2) is not necessarily trivial and depend on the choice made for the space X_h.

2.3. Additional notation. We shall denote by $\lfloor . \rfloor$ the floor function and use the following notation $[\![a,b]\!] = [a,b] \cap \mathbb{N}$.

In the sequel, C will denote a generic constant independent of h, whose value may change from line to line.

3. The semi-discrete situation. We shall assume that the following discrete spectral inequality holds. In the continuous case, for the Laplace-Beltrami operator, this inequality is originally due to G. Lebeau and L. Robbiano [LR95] (see also [LZ98a, JL99] and [LL09] for an introductory presentation).

ASSUMPTION 3.1. *There exists $h_0 > 0$, $\alpha \in [0,1)$, $\beta > 0$, and $\kappa, \ell > 0$ such that the following holds. For any $h < h_0$ and for any $(a_j)_j \in \mathbb{R}^{\mathbb{N}}$, we have*

$$\left\| \sum_{\mu_j \leq \mu} a_j \psi_j \right\|_h^2 = \sum_{\mu_j \leq \mu} |a_j|^2 \leq \kappa e^{\kappa \mu^\alpha} \left[\!\!\left[\mathcal{B}_h^\star \left(\sum_{\mu_j \leq \mu} a_j \psi_j \right) \right]\!\!\right]_h^2, \quad \forall \mu < \frac{\ell}{h^\beta}. \qquad (\mathcal{H}_{\alpha,\beta})$$

Without loss of generality, we shall always assume that $\kappa \geq 1$.

A spectral inequality of this type is proven in [BHL09a], in the case of a finite-difference discretization of the operator $\partial_x(\gamma(x)\partial_x)$ in one-space dimension. The higher dimensional cases, *i.e.*, for elliptic operators of the form $\nabla_x \cdot (\gamma(x)\nabla_x)$, again for finite-differences, are treated in [BHL09b]. To our knowledge, the proof of such a property in the finite elements framework is still an open problem up to now.

For $j \in \mathbb{N}$, we introduce the following subspace of E_h

$$E_{h,j} = \operatorname{span}\{\psi_j;\ \mu_j \leq 2^{\frac{j}{\alpha}}\}, \tag{3.1}$$

and denote by $\Pi_{h,j}$ the orthogonal projector onto $E_{h,j}$ in $(E_h, \langle \cdot, \cdot \rangle_h)$. Note that we have the following properties.

LEMMA 3.2. *The operator $\mathcal{M}_h \Pi_{h,j}$ is symmetric in $(E_h, (\cdot,\cdot)_h)$ and the operators $\Pi_{h,j}$ and $\mathcal{M}_h^{-1}\mathcal{A}_h$ commute.*

Under Assumption 3.1, we define

$$\mu_{\max,h} = \frac{\ell}{h^\beta}, \qquad j_h = \max\left\{ j \in \mathbb{N};\ 2^{\frac{j}{\alpha}} \leq \mu_{\max,h} \right\}. \tag{3.2}$$

Let us now consider the adjoint problem

$$\begin{cases} -\mathcal{M}_h \partial_t q + \mathcal{A}_h q = 0, \quad t \in [0,T) \\ q(T) = q_F, \end{cases} \tag{3.3}$$

for which we can prove the following partial observability inequality

THEOREM 3.3. *Under Assumption 3.1, for any $T > 0$, $h < h_0$ and $j \leq j_h$, the solution of (3.3) satisfies the following inequality*

$$\|q(0)\|_h^2 \leq \frac{\kappa e^{\kappa 2^j}}{T} \int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \, dt, \tag{3.4}$$

provided that $q_F \in E_{h,j}$.

With this observability inequality we obtain the following partial controllability result:

THEOREM 3.4. *Under Assumption 3.1, there exists $C > 0$, such that for any $T > 0$, $h < h_0$, $j \leq j_h$, and any initial data $y_0 \in E_{h,j}$, there exists a control $v \in L^2(0,T,U_h)$ such that the solution y of*

$$\begin{cases} \mathcal{M}_h \partial_t y + \mathcal{A}_h y = \mathcal{M}_h \Pi_{h,j} \mathcal{M}_h^{-1} \mathcal{B}_h v, \\ y(0) = y_0, \end{cases} \tag{3.5}$$

satisfies $y(T) = 0$ and furthermore we have the estimate

$$\|v\|_{L^2(0,T,U_h)} \leq \kappa^{\frac{1}{2}} T^{-\frac{1}{2}} e^{\kappa 2^j} \|y_0\|_h .$$

In the sequel, such a control function v will be denoted by

$$V_j(T, y_0) \in L^2(0, T, U_h).$$

Proof. Notice first that, for an arbitrary v, upon applying $\mathcal{M}_h \Pi_{h,j} \mathcal{M}_h^{-1}$ to (3.5), we find

$$\mathcal{M}_h \partial_t(\Pi_{h,j} y) + \mathcal{A}_h(\Pi_{h,j} y) = \mathcal{M}_h \Pi_{h,j} \mathcal{M}_h^{-1} \mathcal{B}_h v. \tag{3.6}$$

In particular, $y - \Pi_{h,j} y$ satisfies a linear differential equation without source term. It follows that $y(t) \in E_{h,j}$ for $t \geq 0$ as $y_0 \in E_{h,j}$.

For any $q_F \in E_{h,j}$, we define

$$J(q_F) = \frac{1}{2} \int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \, dt + \langle q(0), y_0 \rangle_h ,$$

where $t \mapsto q(t)$ is the solution of the adjoint problem (3.3) with final condition $q_F \in E_{h,j}$. Note that $q(t) \in E_{h,j}$ for $t \in [0, T]$ with the same arguments as above.

The observability inequality (3.4) implies that J is quadratic and strictly coercive on $E_{h,j}$. The functional J thus admits a unique minimizer q_F. The Euler-Lagrange equation then reads

$$0 = \int_0^T [\mathcal{B}_h^\star q(t), \mathcal{B}_h^\star \tilde{q}(t)]_h \, dt + \langle \tilde{q}(0), y_0 \rangle_h , \tag{3.7}$$

for $\tilde{q}$ solution of the adjoint problem associated to an arbitrary final condition $\tilde{q}_F \in E_{h,j}$.

We choose $v(t) = \mathcal{B}_h^\star q(t)$ as a control function in (3.5).

For any $\tilde{q}_F \in E_{h,j}$, we form the inner product $(\cdot, \cdot)_h$ of (3.6) with $\tilde{q}(t)$ and we integrate in time. An integration by parts with respect to time yields

$$\langle \Pi_{h,j} y(T), \tilde{q}_F \rangle_h = \langle \Pi_{h,j} y_0, \tilde{q}(0) \rangle_h + \int_0^T \left\langle \Pi_{h,j} \mathcal{M}_h^{-1} \mathcal{B}_h \mathcal{B}_h^\star q(t), \tilde{q}(t) \right\rangle_h dt. \tag{3.8}$$

Since $y_0 \in E_{h,j}$, the first term in the right-hand side equals $\langle y_0, \tilde{q}(0) \rangle_h$, and since $\tilde{q}(t) \in E_{h,j}$, the second term equals

$$\int_0^T \left\langle \mathcal{M}_h^{-1} \mathcal{B}_h \mathcal{B}_h^\star q(t), \tilde{q}(t) \right\rangle_h dt = \int_0^T (\mathcal{B}_h \mathcal{B}_h^\star q(t), \tilde{q}(t))_h \, dt = \int_0^T [\mathcal{B}_h^\star q(t), \mathcal{B}_h^\star \tilde{q}(t)]_h \, dt.$$

Comparing (3.8) and (3.7) leads to

$$\langle \Pi_{h,j} y(T), \tilde{q}_F \rangle_h = 0, \quad \forall \tilde{q}_F \in E_{h,j},$$

which gives $\Pi_{h,j} y(T) = 0$ and then $y(T) = 0$ as $y(T) \in E_{h,j}$.

Choosing now $\tilde{q} = q$ in (3.7) and using the observability inequality (3.4) we obtain

$$\int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \, dt \leq \|q(0)\|_h \, \|y_0\|_h \leq \kappa^{\frac{1}{2}} T^{-\frac{1}{2}} e^{\kappa 2^j} \left(\int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \, dt \right)^{\frac{1}{2}} \|y_0\|_h ,$$

which finally leads to

$$\|v\|_{L^2(0,T,U_h)} = \left(\int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \, dt \right)^{\frac{1}{2}} \leq \kappa^{\frac{1}{2}} T^{-\frac{1}{2}} e^{\kappa 2^j} \, \|y_0\|_h \, .$$

□

The following estimate is of interest. Take any $y_0 \in E_h$, then take $v = V_j(T, \Pi_{h,j} y_0)$ as a control function in the problem (2.4).

If one applies $\mathcal{M}_h \Pi_{h,j} \mathcal{M}_h^{-1}$ to Problem (2.4), we find that $\Pi_{h,j} y$ solves (3.5) with initial data $\Pi_{h,j} y_0$ and $v = V_j(T, \Pi_{h,j} y_0)$. Hence, by definition of this control function, we deduce that $\Pi_{h,j} y(T) = 0$. Let us now estimate the norm of $y(T)$. To this end, we use the following energy inequality

$$\begin{aligned} \frac{1}{2} \partial_t \|y\|_h^2 + \|y\|_{1,h}^2 = (\mathcal{B}_h v, y)_h = \left\langle \mathcal{M}_h^{-1} \mathcal{B}_h v, y \right\rangle_h \\ \leq \left\| \mathcal{M}_h^{-1} \mathcal{B}_h v \right\|_h \|y\|_h \leq [\![\mathcal{M}_h^{-1} \mathcal{B}_h]\!]_h \, [\![v]\!]_h \, \|y\|_h \, , \end{aligned}$$

from which we deduce that

$$\begin{aligned} \|y(T)\|_h &\leq \|y_0\|_h + [\![\mathcal{M}_h^{-1} \mathcal{B}_h]\!]_h \int_0^T [\![v(t)]\!]_h \, dt \\ &\leq \|y_0\|_h + [\![\mathcal{M}_h^{-1} \mathcal{B}_h]\!]_h \sqrt{T} \left(\int_0^T [\![v(t)]\!]_h^2 \, dt \right)^{\frac{1}{2}} \\ &\leq \left(1 + \kappa^{\frac{1}{2}} [\![\mathcal{M}_h^{-1} \mathcal{B}_h]\!]_h e^{\kappa 2^j} \right) \|y_0\|_h \, . \end{aligned}$$

With Assumption (2.2), $[\![\mathcal{M}_h^{-1} \mathcal{B}_h]\!]_h$ is uniformly bounded with respect to h so that we finally obtain, since $\kappa \geq 1$,

$$\|y(T)\|_h \leq C e^{\kappa 2^j} \, \|y_0\|_h \, .$$

Using the above results, we can now prove the following result. The proof can be adapted from the proof of Theorem 1.4 in [BHL09a].

THEOREM 3.5. *Under assumption* (3.1)*, for any $T > 0$, there exist $h_0 > 0$, $C_T > 0$ and C_1, $C_2 > 0$, such that for any $h \leq h_0$, and all initial data $y_0 \in E_h$, there exists a control function $v \in L^2(0, T, U_h)$ such that the solution to*

$$\mathcal{M}_h \partial_t y + \mathcal{A}_h y = \mathcal{B}_h v(t), \quad y|_{t=0} = y_0. \tag{3.9}$$

satisfies $\Pi_{h,j_h} y(T) = 0$, and

$$\int_0^T [\![v(t)]\!]_h^2 \, dt \leq C_T^2 \, \|y_0\|_h^2 \, , \text{ and } \|y(T)\|_h \leq C_1 e^{-C_2/h^\beta} \, \|y_0\|_h \, .$$

Thanks to this result, we deduce the following h-uniform approximate observability inequality for the semi-discrete problem under study.

THEOREM 3.6. *Under assumption* (3.1)*, for any $T > 0$, there exist $h_0 > 0$, $C_{\text{obs}} > 0$ and $C_1, C_2 > 0$, such that: for any $h < h_0$, the semi-discrete solution q in $\mathscr{C}^\infty(0, T, E_h)$ to*

$$\begin{cases} -\mathcal{M}_h \partial_t q + \mathcal{A}_h q = 0 \\ q(T) = q_F \in E_h, \end{cases}$$

satisfies

$$\|q(0)\|_h \leq C_{\rm obs}\left(\int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \ dt\right)^{\frac{1}{2}} + C_1 e^{-C_2/h^\beta}\,\|q_F\|_h\,.$$

4. Controllability and observability of fully-discrete systems.

4.1. General framework. We consider in this section the problem of controlling fully-discrete approximations of system (2.4) uniformly with respect to the discretization parameters. More precisely, for $M > 0$ and $\delta t = T/M$, We shall consider two time-discretization schemes:

- The implicit Euler scheme:

$$\begin{cases} y^0 = y_0 \in E_h, \\ \mathcal{M}_h \dfrac{y^{n+1}-y^n}{\delta t} + \mathcal{A}_h y^{n+1} = \mathcal{B}_h v^{n+1}, \quad \forall n \in [\![0, M-1]\!], \end{cases} \tag{4.1}$$

- The θ-scheme, with $\theta \in [1/2, 1)$:

$$\begin{cases} y^0 = y_0 \in E_h, \\ \mathcal{M}_h \dfrac{y^{n+1}-y^n}{\delta t} + \mathcal{A}_h(\theta y^{n+1} + (1-\theta) y^n) = \mathcal{B}_h v^{n+1}, \ \forall n \in [\![0, M-1]\!], \end{cases} \tag{4.2}$$

where, in both cases, $(v^n)_{1\leq n\leq M} \in (U_h)^M$ is a fully-discrete control function whose cost, *i.e.* L^2 norm, is given by

$$\left(\sum_{n=1}^M \delta t\, [\![v^n]\!]_h^2\right)^{\frac{1}{2}}.$$

Naturally, the θ-scheme (4.2) coincides with the implicit Euler scheme (4.1) in the case $\theta = 1$. We present the two schemes separately, even though most of the following results are similar for both schemes. In fact, the only particular case we shall encounter is the Crank-Nicolson scheme, that is $\theta = 1/2$, which is a limiting case for the scheme stability.

Let us first state a relationship between a partial controllability result and a suitable observability inequality, for both the implicit Euler scheme and the θ-scheme.

THEOREM 4.1. *Let E be any subspace of E_h such that $\mathcal{M}_h^{-1}\mathcal{A}_h E \subset E$ (that is to say that E is spanned by a suitable subset of the eigenvectors $(\psi_i)_{1\leq i\leq N_h}$). Let $\theta \in [0,1]$, and $F = \ker(\mathcal{M}_h - \delta t(1-\theta)\mathcal{A}_h)$. For a given $C_{\rm obs} > 0$, the following statements are equivalent.*

1. For any $y_0 \in E_h$, there exists $v = (v^n)_{1\leq n\leq M} \in (U_h)^M$ such that

$$\sum_{n=1}^M \delta t\, [\![v^n]\!]_h^2 \leq C_{\rm obs}^2\, \|y_0\|_h^2\,, \tag{4.3}$$

and such that the solution to

$$\begin{cases} y^0 = y_0, \\ \mathcal{M}_h \dfrac{y^{n+1}-y^n}{\delta t} + \mathcal{A}_h(\theta y^{n+1} + (1-\theta) y^n) = \mathcal{B}_h v^{n+1}, \quad \forall n \in [\![0, M-1]\!], \end{cases}$$

satisfies $\Pi_{E\cap F^\perp} y^M = 0$, where $F^\perp$ is the orthogonal of F in $(E_h, \langle\cdot,\cdot\rangle_h)$ and $\Pi_{E\cap F^\perp}$ is the orthogonal projector onto $E\cap F^\perp$ in the same Euclidean space.

2. *Any solution* $q = (q^n)_{1\le n\le M+1}$ *of the following adjoint problem, with* $q^{M+1} \in E \cap F^\perp$*:*

$$\begin{cases} \mathcal{M}_h \dfrac{q^M - q^{M+1}}{\delta t} + \theta \mathcal{A}_h q^M = 0, \\ \mathcal{M}_h \dfrac{q^n - q^{n+1}}{\delta t} + \mathcal{A}_h(\theta q^n + (1-\theta) q^{n+1}) = 0, \quad \forall n \in [\![1, M-1]\!], \end{cases} \tag{4.4}$$

satisfies

$$\left\| q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \right\|_h^2 \le C_{\rm obs}^2 \sum_{n=1}^{M} \delta t \, [\![\mathcal{B}_h^\star q^n]\!]_h^2 \,. \tag{4.5}$$

Notice, when $\theta < 1$, the particular form of the first iterate of the adjoint problem and of the left-hand side of the observability inequality (4.5). In many cases, the space F is trivial (in particular for the implicit Euler scheme).

Proof. The proof is based on the observation that, for any solution $(q^n)_n$ of the adjoint problem with any $q^{M+1} \in E_h$, any solution $(y^n)_n$ of the forward problem with a control term $(v^n)_n$ we have:

$$\begin{aligned} \left\langle y^M, q^{M+1} \right\rangle_h - \left\langle y^0, q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \right\rangle_h &= \sum_{n=1}^{M} \delta t \, (\mathcal{B}_h v^n, q^n)_h \\ &= \sum_{n=1}^{M} \delta t \, [v^n, \mathcal{B}_h^\star q^n]_h \,. \end{aligned} \tag{4.6}$$

Let us first prove that $2 \Rightarrow 1$. Assume that the observability inequality (4.5) holds and pick any $y_0 \in E_h$. Let us introduce a quadratic convex functional J defined for any $q^{M+1} \in E \cap F^\perp$ as follows:

$$J(q^{M+1}) = \frac{1}{2} \sum_{n=1}^{M} \delta t \, [\![\mathcal{B}_h^\star q^n]\!]_h^2 + \left\langle y_0, q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \right\rangle_h .$$

We first prove that J is coercive on $E \cap F^\perp$. As E_h is finite dimensional, it suffices to prove that $\sum_{n=1}^{M} \delta t \, [\![\mathcal{B}_h^\star q^n]\!]_h^2 = 0$ implies $q^{M+1} = 0$. By the observability inequality (4.5) we have $q^1 - \delta t \mathcal{M}_h^{-1}\mathcal{A}_h(1-\theta)q^1 = 0$, that is $q^1 \in F$, then we observe that $q^n = 0$ for any $n \in [\![1, M-1]\!]$ and that $q^M \in F$. Indeed, if we assume that for a given $n \le M-1$ we have $q^n \in F$, then we can use the definition of the adjoint scheme

$$(I - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h) q^{n+1} = (I + \delta t \theta \mathcal{M}_h^{-1}\mathcal{A}_h) q^n, \tag{4.7}$$

and take its $\langle \cdot, \cdot \rangle_h$ inner product with q^n. Since we assumed $q^n \in F$, we obtain

$$\left\langle (I + \delta t \theta \mathcal{M}_h^{-1}\mathcal{A}_h) q^n, q^n \right\rangle_h = 0,$$

so that $q^n = 0$. From (4.7), we deduce that $q^{n+1} \in F$. The result follows by induction.

In particular, we have that $q^M \in F$ and then $q^{M+1} = (I + \delta t \theta \mathcal{M}_h^{-1}\mathcal{A}_h) q^M$ also belongs to F. Since we initially assumed that $q^{M+1} \in F^\perp$, it follows that $q^{M+1} = 0$ and the coercivity of J is proven.

From the above properties of J, we know that it admits a unique minimizer that we denote by q^{M+1} and we denote by $(q^n)_n$ the associated solution to the adjoint problem. We now prove that the control defined by $v^n = \mathcal{B}_h^\star q^n$, satisfies the required properties.

The optimality conditions for J reads

$$\sum_{n=1}^{M} \delta t \Big[\underbrace{\mathcal{B}_h^\star q^n}_{v^n}, \mathcal{B}_h^\star \tilde{q}^n \Big]_h + \big\langle y_0, \tilde{q}^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h \tilde{q}^1 \big\rangle_h = 0,$$
$$\forall\ \tilde{q}^{M+1} \in E \cap F^\perp. \quad (4.8)$$

In particular, using (4.6), we deduce that

$$\big\langle y^M, \tilde{q}^{M+1} \big\rangle_h = 0, \quad \forall \tilde{q}^{M+1} \in E \cap F^\perp,$$

which says exactly that $\Pi_{E\cap F^\perp} y^M = 0$. Taking now $\tilde{q}^{M+1} = q^{M+1}$ in (4.8) and using the observability inequality, we obtain

$$\begin{aligned}
\sum_{n=1}^{M} \delta t \, [\![v^n]\!]_h^2 &= - \big\langle y_0, q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \big\rangle_h \\
&\le \|y_0\|_h \big\| q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \big\|_h \\
&\le C_{\rm obs} \|y_0\|_h \Big(\sum_{n=1}^{M} \delta t \, [\![v^n]\!]_h^2 \Big)^{\frac{1}{2}},
\end{aligned}$$

which gives the claimed estimate of the norm of the control.

Let us now prove that $1 \Rightarrow 2$. We choose $q^{M+1} \in E \cap F^\perp$ and denote by $(q^n)_n$ the associated solution of the adjoint problem (4.4).

We set $y^0 = q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1$ as an initial data for the forward control problem. By assumption, there exists a control $(v^n)_n$ satisfying (4.3), such that the solution $(y^n)_n$ to the controlled problem satisfies $\Pi_{E\cap F^\perp} y^M = 0$. Using these facts in (4.6), we obtain

$$\big\langle y^0, q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \big\rangle_h = - \sum_{n=1}^{M} \delta t \, [v^n, \mathcal{B}_h^\star q^n]_h \,,$$

which gives, by our choice of y^0

$$\big\| q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \big\|_h^2 \le \Big(\sum_{n=1}^{M} \delta t \, [\![v^n]\!]_h^2 \Big)^{\frac{1}{2}} \Big(\sum_{n=1}^{M} \delta t \, [\![\mathcal{B}_h^\star q^n]\!]_h^2 \Big)^{\frac{1}{2}}.$$

We conclude by (4.3) □

4.2. Partial observability inequalities and uniform controllability results. We show here that Assumption 3.1 on the existence of a uniform discrete Lebeau-Robbiano spectral inequality is enough to prove that, under suitable assumptions, the above observability inequality is satisfied in the spaces E_j defined in (3.1), for any $j \le j_h$ (with j_h defined in (3.2)).

THEOREM 4.2. *Assume that Assumption 3.1 holds. Let $T > 0$, and θ be given in $[1/2, 1]$. There exists $C > 0$ (independent of T) such that the following observability inequality holds for any $h \le h_0$*

$$\big\| q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1 \big\|_h^2 \le C(1+T^2)\frac{e^{\kappa 2^j}}{T} \sum_{n=1}^{M} \delta t \, [\![\mathcal{B}_h^\star q^n]\!]_h^2 \,,$$

for any solution of the adjoint problem (4.4) with $E = E_j$, and for any $j \le j_h$.

REMARK 4.3. *Notice that this result holds without any restriction on the time step δt. Note also that the observability inequality holds for any final data in E_j and in particular for any final data in $E_j \cap F^\perp$. This will allow us to apply Theorem 4.1 in the sequel.*

Proof. Let $(q^n)_n$ be a solution to (4.4), and let us introduce

$$\psi^n = (I - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h)q^n.$$

From the definition of q^n, we find that

$$\psi^n = (I + \delta t\theta\mathcal{M}_h^{-1}\mathcal{A}_h)^{-1}(I - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h)\psi^{n+1}, \quad \forall n \in [\![1, M-1]\!].$$

The operator $\mathcal{M}_h^{-1}\mathcal{A}_h$ is symmetric definite positive for the scalar product $\langle\cdot,\cdot\rangle_h$. As $\theta \in [1/2, 1]$, we classically deduce the bound

$$\|\psi^n\|_h \leq \left\|\psi^{n+1}\right\|_h, \quad \forall n \in [\![1, M-1]\!].$$

In particular, we find

$$\left\|\psi^1\right\|_h^2 \leq \frac{1}{T}\sum_{n=1}^{M} \delta t \left\|\psi^n\right\|_h^2.$$

Moreover, since $\psi^n \in E_j$ for any n, we have

$$\|\psi^n\|_h^2 \leq (1 + \delta t(1-\theta)2^{\frac{j}{\alpha}})^2 \left\|q^n\right\|_h^2.$$

It follows that

$$\left\|\psi^1\right\|_h^2 \leq (1 + T2^{\frac{j}{\alpha}})^2 \frac{1}{T}\sum_{n=1}^{M} \delta t \left\|q^n\right\|_h^2.$$

Then, since each q^n lies in E_j, we may apply to each of them the partial Lebeau-Robbiano spectral inequality of Assumption 3.1 which reads

$$\|q^n\|_h^2 \leq \kappa e^{\kappa 2^j} \left[\!\left[\mathcal{B}_h^\star q^n\right]\!\right]_h^2.$$

It follows that we have

$$\left\|\psi^1\right\|_h^2 \leq \kappa(1+T^2)(1+2^{\frac{j}{\alpha}})^2 e^{\kappa 2^j} \frac{1}{T}\sum_{n=1}^{M} \delta t \left[\!\left[\mathcal{B}_h^\star q^n\right]\!\right]_h^2.$$

Since e^{C2^j} increases much more rapidly than $(1+2^{\frac{j}{\alpha}})^2$, we deduce that, for a constant $C_{\alpha,\kappa} > 0$, we have

$$\left\|\psi^1\right\|_h^2 \leq C_{\alpha,\kappa}(1+T^2) e^{\kappa 2^j} \frac{1}{T}\sum_{n=1}^{M} \delta t \left[\!\left[\mathcal{B}_h^\star q^n\right]\!\right]_h^2,$$

which is the claimed observability inequality. □

With this result at hand we now prove the following uniform controllability results for the schemes under study.

THEOREM 4.4 (Implicit Euler scheme and θ-scheme with $\theta > 1/2$). *Let $T > 0$ and $\theta \in (1/2, 1]$. Let $C_T > 0$ and $0 < \gamma \leq \beta$. There exist $C, C_{\text{obs}} > 0$ such that for any $h \leq h_0$, and any $M \in \mathbb{N}^*$ such that $\delta t = T/M \leq C_T h^\gamma$ we have:*

For any $y_0 \in E_h$, there exists a fully-discrete control $v = (v^n)_{1\leq n\leq M} \in U_h^M$ such that

- *The solution* $(y^n)_{0\le n\le M}$ *to* (4.2) *satisfies*

$$\Pi_{h,j_h} y^M = 0, \quad \textit{and} \quad \left\|y^M\right\|_h \le Ce^{-C/h^\gamma} \left\|y_0\right\|_h .$$

- *The control* v *satisfies*

$$\sum_{n=1}^{M} \delta t \left[\!\left[v^n\right]\!\right]_h^2 \le C_{\rm obs}^2 \left\|y_0\right\|_h^2 .$$

Proof. First of all, we note that if the partial spectral inequality of Assumption 3.1 is satisfied, then it is also satisfied for smaller values of the parameter β (changing the constants κ and ℓ if necessary). Once γ is chosen, we may therefore assume without any loss of generality that we have $\alpha\beta < \gamma \le \beta$. Let us also choose γ' such that $\alpha\beta < \gamma' < \gamma$.

Let $M \in \mathbb{N}^*$ be an integer such that $\delta t = \frac{T}{M} \le C_T h^\gamma$. This implies in particular that

$$\delta t(\mu_{\max,h})^{\frac{\gamma}{\beta}} \le \nu, \tag{4.9}$$

where ν depends on C_T, γ, β and ℓ. Observing that ν can be chosen as large as needed, we shall assume that $\nu \ge 1$ (this will be used in step 4 of the proof).

Step 1: a time-slicing procedure. Let $K > 0$ be such that

$$2K \sum_{j=0}^{+\infty} \frac{1}{(2^{\frac{j}{\alpha}})^{\frac{\gamma-\gamma'}{\beta}}} = \frac{T}{2}.$$

For any $0 \le j \le j_h$ we introduce the integers M_j and M_j' defined by

$$M_j = \left\lfloor \frac{K}{\delta t(2^{\frac{j}{\alpha}})^{\frac{\gamma-\gamma'}{\beta}}} \right\rfloor, \quad M_{j+1}' = M_j' + 2M_j, \quad M_0' = 0. \tag{4.10}$$

Notice that $(M_j)_j$ is a non increasing sequence and that

$$\begin{aligned} M_{j_h} = \left\lfloor \frac{K}{\delta t(2^{\frac{j_h}{\alpha}})^{\frac{\gamma-\gamma'}{\beta}}} \right\rfloor &\ge \left\lfloor \frac{K}{\delta t(\mu_{\max,h})^{\frac{\gamma-\gamma'}{\beta}}} \right\rfloor \\ &\ge \left\lfloor \frac{K(\mu_{\max,h})^{\frac{\gamma'}{\beta}}}{\nu} \right\rfloor = \left\lfloor \frac{K\ell^{\frac{\gamma'}{\beta}}}{\nu h^{\gamma'}} \right\rfloor . \end{aligned}$$

For h sufficiently small we thus have

$$\frac{K\ell^{\frac{\gamma'}{\beta}}}{\nu h^{\gamma'}} \ge 1 \tag{4.11}$$

and thus $\forall j \le j_h, M_j \ge M_{j_h} \ge 1$. Furthermore we have

$$M_{j_h+1}' = 2\sum_{j=0}^{j_h} M_j \le 2\sum_{j=0}^{\infty} \frac{K}{\delta t(2^{\frac{j}{\alpha}})^{\frac{\gamma-\gamma'}{\beta}}} \le \frac{T}{2\delta t} = \frac{M}{2}. \tag{4.12}$$

Hence, with this analysis, we may split the set of discrete times $\{0, \cdots, M_{j_h+1}'\delta t\}$ into $j_h + 1$ subsets, of size $2M_j$, $j \in \{0, \cdots, j_h\}$. We construct the restriction of the fully-discrete control $(v^n)_{M_j'+1\le n\le M_{j+1}'}$ in the jth sub-interval by induction on j as follows.

Step 2: active and passive control sequences. With Theorem 4.2, where M is replaced by M_j and T by $M_j\delta t$ and then applying Theorem 4.1, we obtain a control $(v^n)_{M_j'+1\leq n\leq M_j'+M_j}\in \big(U_h\big)^{M_j}$ such that

$$\sum_{n=M_j'+1}^{M_j'+M_j}\delta t\,[\![v^n]\!]_h^2\leq C\frac{e^{C2^j}}{M_j\delta t}\left\|y^{M_j'}\right\|_h^2, \tag{4.13}$$

and such that the corresponding controlled solution to the θ-scheme (4.2) satisfies

$$\Pi_{h,j}\Pi_{F^\perp}y^{M_j'+M_j}=0. \tag{4.14}$$

Note that we have (Duhamel principle)

$$y^{M_j'+M_j}=\mathcal{C}_h^{M_j}y^{M_j'}+\sum_{k=1}^{M_j}\delta t\mathcal{C}_h^{M_j-k+1}\mathcal{M}_h^{-1}\mathcal{B}_hv^{M_j'+k}, \tag{4.15}$$

with

$$\mathcal{C}_h=(\mathrm{Id}+\delta t\theta\mathcal{M}_h^{-1}\mathcal{A}_h)^{-1}(\mathrm{Id}-\delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h).$$

As $\|\mathcal{C}_h\|_h\leq 1$ (since $\theta\geq\frac12$), we deduce that

$$\begin{aligned}\left\|y^{M_j'+M_j}\right\|_h&\leq\left\|y^{M_j'}\right\|_h+|\!|\!|\mathcal{M}_h^{-1}\mathcal{B}_h|\!|\!|_h\sum_{k=1}^{M_j}\delta t\left[\!\!\left[v^{M_j'+k}\right]\!\!\right]_h\\&\leq\left\|y^{M_j'}\right\|_h+|\!|\!|\mathcal{M}_h^{-1}\mathcal{B}_h|\!|\!|_h\left(\sum_{k=1}^{M_j}\delta t\left[\!\!\left[v^{M_j'+k}\right]\!\!\right]_h^2\right)^{\frac12}\sqrt{M_j\delta t},\end{aligned}$$

and then by (4.13) and (2.2), for some $C_1>0$,

$$\left\|y^{M_j'+M_j}\right\|_h\leq e^{C_12^j}\left\|y^{M_j'}\right\|_h. \tag{4.16}$$

For $n\in[\![M_j'+M_j+1,M_{j+1}']\!]$, we choose $v^n=0$ so that the discrete solution y^n evolves free of any control for $n\in[\![M_j'+M_j+1,M_{j+1}']\!]$. We obtain

$$y^{M_{j+1}'}=\mathcal{C}_h^{M_j}y^{M_j'+M_j}. \tag{4.17}$$

Let us now study more precisely $y^{M_{j+1}'}$.

- For $\theta=1$, the space $F=\ker(\mathcal{M}_h)$ is trivial so that (4.14) gives immediately $\Pi_{h,j}y^{M_j'+M_j}=0$ and then, since $\mathcal{C}_h$ and $\Pi_{h,j}$ commute
$$\Pi_{h,j}y^{M_{j+1}'}=0. \tag{4.18}$$
- For $\frac12\leq\theta<1$, it may happen that $F=\ker(\mathcal{M}_h-\delta t(1-\theta)\mathcal{A}_h)$ is not trivial. In that case, using the definition of F and the fact that $\mathcal{C}_h$ and Π_F commute, we observe that
$$\begin{aligned}\Pi_Fy^{M_j'+M_j+1}&=\Pi_F\mathcal{C}_hy^{M_j'+M_j}\\&=(\mathrm{Id}+\delta t\theta\mathcal{M}_h^{-1}\mathcal{A}_h)^{-1}\underbrace{(\mathrm{Id}-\delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h)\Pi_F}_{=0}y^{M_j'+M_j}.\end{aligned}$$
We thus have $\Pi_Fy^{M_j'+M_j+1}=0$. Furthermore, with (4.14), we deduce that $\Pi_{h,j}y^{M_j'+M_j+1}=0$ and finally that (4.18) also holds in this case.

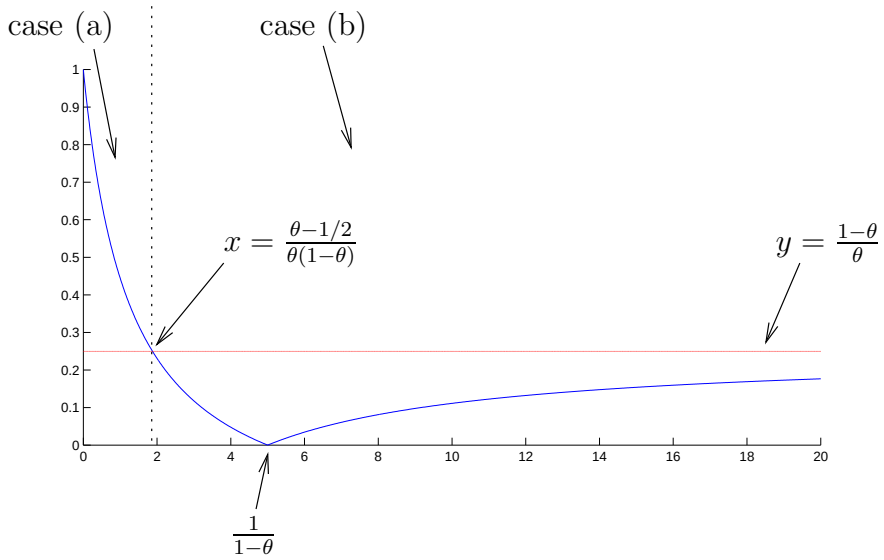


FIG. 4.1. *The map $x \mapsto |\Gamma_\theta(x)|$ and its limit at infinity*

Step 3: evolution of the L^2 norm of the solution. Let us now introduce the map

$$\Gamma_\theta : x \in [0, +\infty) \mapsto \Gamma_\theta(x) = \frac{1-(1-\theta)x}{1+\theta x},$$

which is such that the iteration matrix of the scheme is

$$\mathcal{C}_h = \Gamma_\theta(\delta t \mathcal{M}_h^{-1} \mathcal{A}_h). \tag{4.19}$$

For any $\theta \in [\frac{1}{2}, 1]$, we have $\Pi_{h,j} y^{M'_{j+1}} = 0$, we thus write

$$y^{M'_{j+1}} = (\mathrm{Id} - \Pi_{h,j}) y^{M'_{j+1}} = (\mathrm{Id} - \Pi_{h,j}) \mathcal{C}_h^{M_j} y^{M'_j + M_j} = \Big((\mathrm{Id} - \Pi_{h,j}) \mathcal{C}_h \Big)^{M_j} y^{M'_j + M_j},$$

and then

$$\left\| y^{M'_{j+1}} \right\|_h \leq \| (\mathrm{Id} - \Pi_{h,j}) \mathcal{C}_h \|_h^{M_j} \left\| y^{M'_j + M_j} \right\|_h. \tag{4.20}$$

Since $(\mathrm{Id} - \Pi_{h,j}) \mathcal{C}_h$ is symmetric for the scalar product $\langle \cdot, \cdot \rangle_h$, the norm $\| (\mathrm{Id} - \Pi_{h,j}) \mathcal{C}_h \|_h$ is actually equal to the spectral radius of $(\mathrm{Id} - \Pi_{h,j}) \mathcal{C}_h = \mathcal{C}_h (\mathrm{Id} - \Pi_{h,j})$.

By definition of the space $E_{h,j}$ and (4.19), we conclude that

$$\| (\mathrm{Id} - \Pi_{h,j}) \mathcal{C}_h \|_h \leq \sup_{x \in (\delta t 2^{\frac{j}{\alpha}}, +\infty)} |\Gamma_\theta(x)|. \tag{4.21}$$

The graph of the function $|\Gamma_\theta|$ is represented in Figure 4.1 and we observe that two cases are to be considered. We emphasize that we assume $\theta > \frac{1}{2}$ here.

Case (a). If j is such that $\delta t 2^{\frac{j}{\alpha}} \leq \frac{\theta - 1/2}{\theta(1-\theta)}$, we then have

$$\sup_{x \in (\delta t 2^{\frac{j}{\alpha}}, +\infty)} |\Gamma_\theta(x)| = |\Gamma_\theta(\delta t 2^{\frac{j}{\alpha}})| \leq (1 + \theta \delta t 2^{\frac{j}{\alpha}})^{-1}.$$

Estimates (4.16), (4.20) and (4.21) then lead to the following inequality

$$\left\|y^{M'_{j+1}}\right\|_h \leq e^{C_1 2^j - M_j \ln\left(1+\theta\delta t 2^{\frac{j}{\alpha}}\right)} \left\|y^{M'_j}\right\|_h. \tag{4.22}$$

For such a value of j, we define

$$\psi_j = C_1 2^j - M_j \ln\left(1+\theta\delta t 2^{\frac{j}{\alpha}}\right). \tag{4.23}$$

Case (b). If j is such that $\delta t 2^{\frac{j}{\alpha}} > \frac{\theta-1/2}{\theta(1-\theta)}$, then we have

$$\sup_{x\in(\delta t 2^{\frac{j}{\alpha}},+\infty)} |\Gamma_\theta(x)| = \lim_{+\infty} |\Gamma_\theta| = \frac{1-\theta}{\theta}.$$

We set $\xi = -\log\left(\frac{1-\theta}{\theta}\right)$. Note that this value of ξ only depends on θ and we have $\xi > 0$ as we assumed $\theta \in (1/2, 1]$.

In this case, estimates (4.16), (4.20) and (4.21) lead to

$$\left\|y^{M'_{j+1}}\right\|_h \leq e^{C_1 2^j - M_j \xi} \left\|y^{M'_j}\right\|_h. \tag{4.24}$$

For such a value of j, we define

$$\psi_j = C_1 2^j - M_j \xi. \tag{4.25}$$

Step 4: L^2 bounds for the control and the solution. Gathering the previous facts, we obtain

$$\Pi_{h,j_h} y^{M'_{j_h+1}} = 0, \qquad \left\|y^{M'_{j_h+1}}\right\|_h \leq e^{\sum_{j=0}^{j_h}\psi_j} \left\|y^0\right\|_h. \tag{4.26}$$

LEMMA 4.5. *There exists $C_2 > 0$ which does not depend on δt and h, such that for any $0 \leq J \leq j_h$ we have*

$$\sum_{j=0}^{J} \psi_j \leq \frac{1}{C_2} - C_2 2^{J\frac{\gamma'}{\alpha\beta}}.$$

Proof. We first estimate ψ_j. As seen above, two cases have to be considered.

Case (a). For $j \leq j_h$ such that $\delta t 2^{\frac{j}{\alpha}} \leq \frac{\theta-\frac{1}{2}}{\theta(1-\theta)}$, ψ_j is given by (4.23). Observing that $\ln(1+x) \geq \frac{x}{1+x}$ for any $x \geq 0$, to obtain

$$\psi_j = C_1 2^j - M_j \ln(1+\delta t 2^{\frac{j}{\alpha}}) \leq C_1 2^j - M_j \frac{\delta t 2^{\frac{j}{\alpha}}}{1+\delta t 2^{\frac{j}{\alpha}}}.$$

Then, the definition of M_j in (4.10) implies that

$$\psi_j \leq C_1 2^j - \left(\frac{K}{\delta t (2^{\frac{j}{\alpha}})^{\frac{\gamma-\gamma'}{\beta}}} - 1\right) \frac{\delta t 2^{\frac{j}{\alpha}}}{1+\delta t 2^{\frac{j}{\alpha}}} \leq C_1' 2^j - \frac{K 2^{\frac{j}{\alpha}(1-\frac{\gamma-\gamma'}{\beta})}}{1+\delta t 2^{\frac{j}{\alpha}}}.$$

Since we assumed $j \leq j_h$ and $\gamma \leq \beta$ we have

$$\delta t 2^{\frac{j}{\alpha}} \leq \delta t (2^{\frac{j_h}{\alpha}})^{\frac{\gamma}{\beta}} 2^{\frac{j}{\alpha}(1-\frac{\gamma}{\beta})} \leq \delta t (\mu_{\max,h})^{\frac{\gamma}{\beta}} 2^{\frac{j}{\alpha}(1-\frac{\gamma}{\beta})} \leq \nu 2^{\frac{j}{\alpha}(1-\frac{\gamma}{\beta})},$$

where ν is defined in (4.9). Recalling that we choose $\nu \geq 1$, we in fact write

$$1 + \delta t 2^{2j} \leq 2\nu 2^{\frac{j}{\alpha}(1-\frac{\gamma}{\beta})}.$$

It then follows that

$$\psi_j \leq C_1' 2^j - \frac{K}{6\nu} 2^{j\frac{\gamma'}{\alpha\beta}}. \tag{4.27}$$

Case (b). For $j \leq j_h$ such that $\delta t 2^{\frac{j}{\alpha}} > \frac{\theta-\frac{1}{2}}{\theta(1-\theta)}$, ψ_j is given by (4.25). From the definition of M_j, we have

$$\psi_j \leq C_1 2^j + \xi - \frac{K\xi}{\delta t (2^{\frac{j}{\alpha}})^{\frac{\gamma-\gamma'}{\beta}}}.$$

Note that we have

$$\delta t (2^{\frac{j}{\alpha}})^{\frac{\gamma-\gamma'}{\beta}} \leq \delta t (2^{\frac{j_h}{\alpha}})^{\frac{\gamma}{\beta}} 2^{-j\frac{\gamma'}{\alpha\beta}} \leq \delta t (\mu_{\max,h})^{\frac{\gamma}{\beta}} 2^{-j\frac{\gamma'}{\alpha\beta}} \leq \nu 2^{-j\frac{\gamma'}{\alpha\beta}}.$$

We thus find

$$\psi_j \leq C_1 2^j + \xi - \frac{K\xi}{\nu} 2^{j\frac{\gamma'}{\alpha\beta}} \leq (C_1 + \xi) 2^j - \frac{K\xi}{\nu} 2^{j\frac{\gamma'}{\alpha\beta}}. \tag{4.28}$$

Introducing $\tilde{C} = \max(C_1 + \xi, C_1', \frac{2\nu}{K}, \frac{\nu}{K\xi})$, estimates (4.27) and (4.28) give

$$\forall j \leq j_h, \qquad \psi_j \leq \tilde{C} 2^j - \frac{1}{\tilde{C}} 2^{j\frac{\gamma'}{\alpha\beta}}.$$

Using the Young inequality (with exponent $p = \frac{\gamma'+\alpha\beta}{2\alpha\beta} > 1$), we obtain

$$\tilde{C} 2^j = \tilde{C} 2^{-j\frac{\gamma'-\alpha\beta}{\gamma'+\alpha\beta}} 2^{j\frac{2\gamma'}{\gamma'+\alpha\beta}} \leq \tilde{C}' 2^{-j} + \frac{1}{2\tilde{C}} 2^{j\frac{\gamma'}{\alpha\beta}}.$$

It follows that

$$\psi_j \leq \tilde{C}' 2^{-j} - \frac{1}{2\tilde{C}} 2^{j\frac{\gamma'}{\alpha\beta}}.$$

Summing this inequality for $0 \leq j \leq J$, we obtain

$$\sum_{j=0}^{J} \psi_j \leq 2\tilde{C}' + \frac{1}{2\tilde{C}} \frac{1}{2^{\frac{\gamma'}{\alpha\beta}} - 1} - \frac{1}{2\tilde{C}} \frac{2^{\frac{\gamma'}{\alpha\beta}}}{2^{\frac{\gamma'}{\alpha\beta}} - 1} 2^{J\frac{\gamma'}{\alpha\beta}}.$$

The claim follows by choosing $C_2 > 0$ sufficiently small. □

Continuation of step 4 of the proof of Theorem 4.4. Using this lemma with (4.26), we obtain that the controlled solution $(y^n)_n$ we constructed satisfies

$$\left\| y^{M'_{j_h+1}} \right\|_h \leq e^{1/C_2} e^{-C_2 2^{j_h \frac{\gamma'}{\alpha\beta}}} \|y_0\|_h \leq C e^{-C'/h^{\gamma'}} \|y_0\|_h, \quad C > 0,\ C' > 0, \tag{4.29}$$

since, by the definition (3.2) of j_h, we have $2^{\frac{j_h}{\alpha}} \geq 2^{-\frac{1}{\alpha}} \frac{\ell}{h^\beta}$.

Furthermore, introducing

$$\Psi_0 = 0, \quad \Psi_j = \sum_{k=0}^{j-1} \psi_k, \quad 1 \le j \le j_h,$$

using (4.13), the total norm of the discrete control is bounded as follows

$$\begin{aligned}
\|v\|^2 &= \sum_{j=0}^{j_h} \sum_{n=M'_j+1}^{M'_j+M_j} \delta t \, [\![v^n]\!]_h^2 \le C \sum_{j=0}^{j_h} \frac{e^{C2^j}}{M_j \delta t} e^{2\Psi_j} \|y_0\|_h^2 \\
&\le C e^{2/C_2} \sum_{j=1}^{j_h} \frac{e^{C2^j}}{M_j \delta t} e^{-2C_2 2^{j\frac{\gamma'}{\alpha\beta}}} \|y_0\|_h^2 + C \frac{e^C}{M_0 \delta t} \|y_0\|_h^2 \\
&\le \frac{2Ce^{2/C_2}}{K} \sum_{j=1}^{j_h} 2^{j\frac{\gamma-\gamma'}{\alpha\beta}} e^{C2^j} e^{-2C_2 2^{j\frac{\gamma'}{\alpha\beta}}} \|y_0\|_h^2 + \frac{2Ce^C}{K} \|y_0\|_h^2 ,
\end{aligned}$$

as in fact

$$M_j \delta t \ge \frac{1}{2}(M_j + 1)\delta t \ge \frac{1}{2} \frac{K}{2^{j\frac{\gamma-\gamma'}{\alpha\beta}}},$$

by the definition of M_j in (4.10), and the fact that $M_j \ge 1$. We thus have

$$\|v\|^2 \le C_2' \|y_0\|_h^2 ,$$

as the series $\sum_{j=0}^{\infty} 2^{j\frac{\gamma-\gamma'}{\alpha\beta}} e^{C2^j} e^{-C_2 2^{j\frac{\gamma'}{\alpha\beta}}}$ converges, having assumed that $\alpha\beta < \gamma'$.

Step 5: L^2 estimate of the remainder. We now choose $v^n = 0$ for $M'_{j_h+1} < n \le M$. The above estimate on the cost of the control remains unchanged. We have

$$y^M = \mathcal{C}_h^{M-M'_{j_h+1}} y^{M'_{j_h+1}},$$

and as $\Pi_{h,j_h} y^{M'_{j_h+1}} = 0$ and by definition of j_h, we obtain

$$\left\| y^M \right\|_h \le \sup_{\mu \in (\frac{\ell 2^{-1/\alpha}}{h^\beta}, +\infty)} |\Gamma_\theta(\delta t \mu)|^{M-M'_{j_h+1}} \left\| y^{M'_{j_h+1}} \right\|_h .$$

Using (4.12) we see that $M - M'_{j_h+1} \ge \frac{M}{2}$ so that, with (4.29), we obtain

$$\left\| y^M \right\|_h \le C \sup_{\mu \in (\frac{\ell 2^{-1/\alpha}}{h^\beta}, +\infty)} |\Gamma_\theta(\delta t \mu)|^{M/2} e^{-\frac{C'}{h^{\gamma'}}} \|y_0\|_h . \tag{4.30}$$

We observe that there exists $\tilde{C} > 0$, such that, for h sufficiently small we have

$$\forall \mu \ge \frac{\ell 2^{-1/\alpha}}{h^\beta}, \quad |\Gamma_\theta(\delta t \mu)| \le e^{-\tilde{C}\frac{\delta t}{h^\gamma}}. \tag{4.31}$$

In fact, two cases have to be considered.

- If $\theta < 1$ and $\delta t \mu \ge \frac{\theta - \frac{1}{2}}{\theta(1-\theta)}$, we see on Figure 4.1 that

$$|\Gamma_\theta(\delta t \mu)| \le \frac{1-\theta}{\theta} = e^{-C_T \xi},$$

with $\xi = -\log((1-\theta)/\theta)/C_T > 0$. The result follows since $\delta t \le C_T h^\gamma$.

- If $\theta = 1$, or $\delta t\mu < \frac{\theta - \frac{1}{2}}{\theta(1-\theta)}$ for $\theta < 1$, we have

$$0 \leq \Gamma_\theta(\delta t\mu) \leq (1 + \theta\delta t\mu)^{-1}.$$

Since the function $x \mapsto \log(1+x)/x$ is non-increasing and $\delta t\mu \leq C_T \mu h^\gamma$, we have

$$\frac{\log(1+\theta\delta t\mu)}{\theta\delta t\mu} \geq \frac{\log(1+\theta C_T\mu h^\gamma)}{\theta C_T\mu h^\gamma},$$

so that

$$\begin{aligned}\log(1+\theta\delta t\mu) &\geq \frac{\delta t}{C_T h^\gamma}\log(1+\theta C_T\mu h^\gamma)\\ &\geq \frac{\delta t}{C_T h^\gamma}\log(1+\theta C_T \ell 2^{-1/\alpha} h^{\gamma-\beta}) \geq \frac{\delta t}{C_T h^\gamma}\log(1+\theta C_T\ell 2^{-1/\alpha}h_0^{\gamma-\beta}).\end{aligned}$$

Using (4.30) and (4.31), and since $\delta t M = T$, we deduce that

$$\left\|y^M\right\|_h \leq Ce^{-\frac{\tilde{C}_T}{2h^\gamma}} \left\|y_0\right\|_h.$$

This concludes the proof of Theorem 4.4. □

THEOREM 4.6 (Crank-Nicolson scheme). *Under assumption 3.1, for any $T > 0$, $0 < \gamma \leq \beta$, $C_T > 0$ and $\delta > 0$, there exist $C, C_{\text{obs}} > 0$ such that for any $h \leq h_0$, and any $M \in \mathbb{N}^*$ such that $\delta t = T/M \leq C_T h^\gamma$ and $\delta t\rho_h \leq \delta$, we have:*

For any $y_0 \in E_h$, there exists a fully-discrete control $v = (v^n)_{1\leq n\leq M} \in \left(U_h\right)^M$ such that

- *The solution $(y^n)_{0\leq n\leq M}$ to (4.2), with $\theta = \frac{1}{2}$, satisfies*

$$\Pi_{h,j_h} y^M = 0, \quad \text{and} \quad \left\|y^M\right\|_h \leq Ce^{-C/h^\gamma}\left\|y_0\right\|_h.$$

- *The control v satisfies*

$$\sum_{n=1}^{M} \delta t \left[\!\left[v^n\right]\!\right]_h^2 \leq C_{\text{obs}}^2 \left\|y_0\right\|_h^2.$$

We recall that ρ_h is the spectral radius of $\mathcal{M}_h^{-1}\mathcal{A}_h$.

Proof. The proof follows the same lines as the previous one. The main difference comes from the fact that high frequencies are not sufficiently damped by the Crank-Nicolson scheme. This phenomenon is well known and is related to the fact that the value $\theta = \frac{1}{2}$ is the limit of unconditional stability for the θ-scheme. For this reason we need to add the condition $\delta t\rho_h \leq \delta$ linking the time step and the space discretization for our result to hold. Without loss of generalities, we further assume that $\delta \geq 2$.

Let us only mention the points of the proof that require changes in this case (see also Figure 4.2):

- Formula (4.21) becomes

$$\left\|(\mathrm{Id} - \Pi_{h,j})\mathcal{C}_h\right\|_h \leq \sup_{x\in[\delta t 2^{\frac{j}{\alpha}},\delta]} |\Gamma_{\frac{1}{2}}(x)|.$$

- The two cases to be considered in the estimates are now

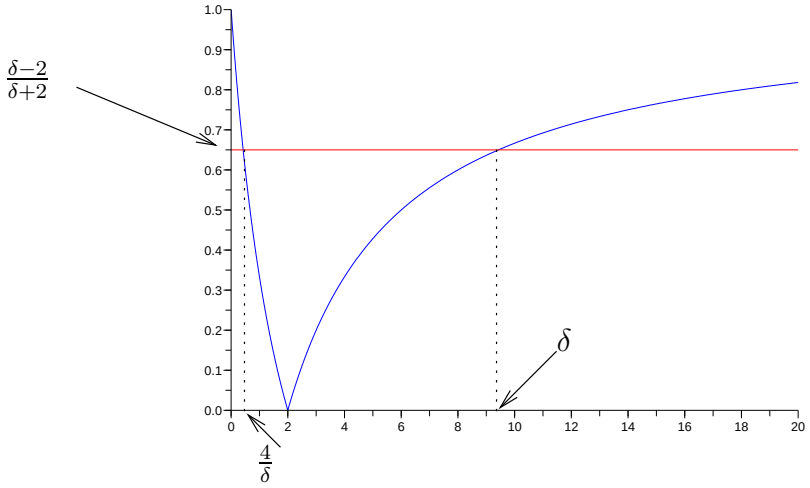


FIG. 4.2. *The map* $x \mapsto |\Gamma_{\frac{1}{2}}(x)|$

- *Case (a).* For j such that $\delta t 2^{\frac{j}{\alpha}} \leq \frac{4}{\delta}$.
- *Case (b).* For j such that $\delta t 2^{\frac{j}{\alpha}} > \frac{4}{\delta}$. We then have

$$\sup_{x \in [\delta t 2^{\frac{j}{\alpha}}, \delta]} |\Gamma_{\frac{1}{2}}(x)| = |\Gamma_{\frac{1}{2}}(\delta)| = \frac{\delta - 2}{\delta + 2},$$

and the same proof applies by choosing $\xi = -\log\left(\frac{\delta-2}{\delta+2}\right)$.

□

4.3. Global relaxed uniform observability inequalities. Using Theorems 4.4 and 4.6, we deduce the following global observability inequality, which improves that given in Theorem 4.2.

THEOREM 4.7. *Assume that Assumption 3.1 holds. Let* $T > 0$, *and* $\theta \in [1/2, 1]$. *Let* $C_T > 0$, $0 < \gamma \leq \beta$. *If* $\theta = \frac{1}{2}$, *we also suppose given some* $\delta > 0$.

There exists $h_0 > 0$, $C_{\text{obs}} > 0$ *and* $C_1, C_2 > 0$ *such that, for any* $h \leq h_0$ *and any* $M \in \mathbb{N}^*$ *such that* $\delta t = \frac{T}{M} \leq C_T h^\gamma$ *(and* $\delta t \rho_h \leq \delta$, *in the case* $\theta = \frac{1}{2}$*), we have*

$$\left\| q^1 - \delta t (1-\theta) \mathcal{M}_h^{-1} \mathcal{A}_h q^1 \right\|_h \leq C_{\text{obs}} \left(\sum_{n=1}^{M} \delta t \left\| \mathcal{B}_h^\star q^n \right\|_h^2 \right)^{\frac{1}{2}} + C_1 e^{-C_2/h^\gamma} \left\| q_F \right\|_h ,$$

for any $q_F \in E_h$, *and* $(q^n)_n \in E_h^M$ *the associated solution of the backward problem* (4.4).

5. Practical computation of semi-discrete and fully-discrete controls. We assume until the end of the article, that Assumption 3.1 is satisfied for some $h_0, \alpha, \beta, \kappa, \ell$, and that the final time $T > 0$ is fixed. We shall also assume that h_0 is sufficiently small such that Theorem 3.6 and Theorem 4.7 hold. We shall denote by C_{obs} a common value of the constant given by these two theorems.

Let us now consider a continuous function $h \mapsto \phi(h) \in \mathbb{R}^{+,*}$ satisfying

$$\lim_{h\to 0} \frac{e^{-C/h^\gamma}}{\phi(h)} = 0, \quad \forall C > 0, \gamma > 0. \tag{5.1}$$

This last assumption is certainly satisfied by any polynomial function of h, which is the typical case that we shall consider.

Choosing $\phi(h)$ constant, say $\phi(h) = \varepsilon > 0$, is also of interest. This situation corresponds to the case of an approximate control problem, tackled with a penalization technique, as proposed and studied in [GL94]. Our analysis thus includes such approximate control problems for parabolic equations.

Such a function ϕ is now fixed all along this section.

Here we address the control to the trajectories, which is known to be equivalent to null-controllability in the case of linear equations. However, here, a difficulty arises to obtain uniform estimates, as the semi-discrete and the fully-discrete free trajectories do not coincide. Let us present the framework we consider.

For any $h > 0$, we suppose given a target $\hat{y}_F \in E_h$ which is the final state of a free solution of the semi-discrete system from an initial datum $\hat{y}_0 \in E_h$, that is we assume that

$$\hat{y}_0, \hat{y}_F \in E_h, \quad \text{such that} \quad \hat{y}_F = e^{-T\mathcal{M}_h^{-1}\mathcal{A}_h}\hat{y}_0. \tag{5.2}$$

Notice that any element $\hat{y}_F$ in E_h can be written in this form. However, we are interested in the situation where some uniform bounds on $\hat{y}_0$ are available. This corresponds to attempting to control the system towards a semi-discrete approximation of an actual free trajectory of the original parabolic PDE for an initial data in $L^2(\Omega)$.

5.1. The semi-discrete case. We first deal with the semi-discrete situation.

THEOREM 5.1. *Let $\hat{y}_0, \hat{y}_F$ be given in* (5.2)*. For any $h \le h_0$, and any $y_0 \in E_h$, we consider the functional $q_F \in E_h \mapsto J^h(q_F)$ defined by*

$$J^h(q_F) = \frac{1}{2}\int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \, dt + \frac{\phi(h)}{2}\|q_F\|_h^2 - \langle \hat{y}_F, q_F\rangle_h + \langle y_0, q(0)\rangle_h,$$

where $t \mapsto q(t)$ is the solution to the adjoint problem $-\mathcal{M}_h\partial_t q(t) + \mathcal{A}_h q(t) = 0$ with final data $q(T) = q_F$.

This functional J^h has a unique minimiser denoted by $q_{opt}^F \in E_h$. This minimiser produces a solution q_{opt} of the adjoint problem such that, if we define the control function $v(t) = \mathcal{B}_h^\star q_{opt}(t)$:

- *The cost of the control is bounded as follows*

$$\int_0^T [\![v(t)]\!]_h^2 \, dt \le \left(C_{\text{obs}}^2 + \phi(h)\right)\|y_0 - \hat{y}_0\|_h^2. \tag{5.3}$$

- *The controlled solution y to* (3.9) *is such that*

$$\|y(T) - \hat{y}_F\|_h \le \sqrt{\phi(h)}\left(C_{\text{obs}} + \sqrt{\phi(h)}\right)\|y_0 - \hat{y}_0\|_h. \tag{5.4}$$

Finally, the optimal adjoint state $q_{opt,\delta t}^F$ satisfies

$$\sqrt{\phi(h)}\left\|q_{opt}^F\right\|_h \le \left(C_{\text{obs}} + \sqrt{\phi(h)}\right)\|y_0 - \hat{y}_0\|_h. \tag{5.5}$$

Proof. The functional J^h is smooth, strictly convex, and coercive on a finite dimensional space, thus it admits a unique minimizer. Furthermore, since $\hat{y}_F = e^{-T\mathcal{M}_h^{-1}\mathcal{A}_h}\hat{y}_0$ we have

$$\langle \hat{y}_F, q_F\rangle_h = \left\langle \hat{y}_0, e^{-T\mathcal{M}_h^{-1}\mathcal{A}_h} q_F\right\rangle_h = \langle \hat{y}_0, q(0)\rangle_h . \tag{5.6}$$

Then, J^h can be expressed as follows

$$J^h(q_F) = \frac{1}{2}\int_0^T [\![\mathcal{B}_h^\star q(t)]\!]_h^2 \, dt + \frac{\phi(h)}{2}\|q_F\|_h^2 + \langle y_0 - \hat{y}_0, q(0)\rangle_h .$$

The Euler-Lagrange equation associated to this minimization problem reads

$$\int_0^T [\mathcal{B}_h^\star q_{opt}(t), \mathcal{B}_h^\star q(t)]_h \, dt + \phi(h)\left\langle q_{opt}^F, q_F\right\rangle_h = -\langle y_0 - \hat{y}_0, q(0)\rangle_h , \tag{5.7}$$

for any $q_F \in E_h$, with the associated solution $t \mapsto q(t)$ of the adjoint problem. We consider now the solution y to the controlled problem $\mathcal{M}_h\partial_t y + \mathcal{A}_h y = \mathcal{B}_h\mathcal{B}_h^\star q_{opt}(t)$, with $y(0) = y_0$. By integration by parts, we deduce

$$\int_0^T [\mathcal{B}_h^\star q_{opt}(t), \mathcal{B}_h^\star q(t)]_h \, dt = \langle q_F, y(T)\rangle_h - \langle q(0), y_0\rangle_h .$$

Comparing with (5.7), and using (5.6) we obtain that

$$\forall q_F \in E_h, \langle q_F, y(T)\rangle_h = \langle \hat{y}_F, q_F\rangle_h - \phi(h)\left\langle q_{opt}^F, q_F\right\rangle_h ,$$

that is $y(T) = \hat{y}_F - \phi(h)q_{opt}^F$. Estimate (5.4) will thus be a consequence of (5.5).

Let us now prove (5.3) and (5.5). We move back to (5.7) in which we choose $q_F = q_{opt}^F$. It follows that

$$\int_0^T [\![\mathcal{B}_h^\star q_{opt}(t)]\!]_h^2 \, dt + \phi(h)\left\|q_{opt}^F\right\|_h^2 = \langle \hat{y}_0 - y_0, q_{opt}(0)\rangle_h \le \|y_0 - \hat{y}_0\|_h \|q_{opt}(0)\|_h . \tag{5.8}$$

By Assumption (5.1), we can choose h_0 sufficiently small such that $\phi(h) \ge C_1 e^{-C_2/h^\beta}$ for any $0 < h \le h_0$, where C_1 and C_2 are the constants introduced in Theorem 3.6. In particular, we deduce from this theorem that the following observability inequality holds

$$\|q_{opt}(0)\|_h \le C_{\rm obs}\left(\int_0^T [\![\mathcal{B}_h^\star q_{opt}(t)]\!]_h^2 \, dt\right)^{\frac{1}{2}} + \phi(h)\left\|q_{opt}^F\right\|_h ,$$

which leads to (5.8) by using the Young inequality. □

5.2. The fully-discrete case. We now state a result similar to that given above but in the fully-discrete case.

THEOREM 5.2. *Let $\hat{y}_0, \hat{y}_F$ be given in* (5.2). *Let $\theta \in [\frac{1}{2}, 1]$, $C_T > 0$ and $0 < \gamma \le \beta$. In the case $\theta = \frac{1}{2}$, we also let $\delta > 0$.*

For any $h \leq h_0$, and any $y_0 \in E_h$, we consider the functional $q_F \in E_h \mapsto J^{h,\delta t}(q_F)$ defined by

$$J^{h,\delta t}(q_F) = \frac{1}{2}\sum_{n=1}^{M} \delta t \left[\!\left[\mathcal{B}_h^\star q^n\right]\!\right]_h^2 + \frac{\phi(h)}{2}\left\|q_F\right\|_h^2 - \langle \hat{y}_F, q_F\rangle_h + \left\langle y_0, q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1\right\rangle_h,$$

where $(q^n)_n$ is the solution of the adjoint problem (4.4) with final data $q^{M+1} = q_F$.

This functional $J^{h,\delta t}$ has a unique minimiser $q^F_{opt,\delta t} \in E_h$. This minimiser produces a solution $q_{opt,\delta t} = (q^n_{opt,\delta t})_n$ to (4.4).

If one defines $v^n = \mathcal{B}_h^\star q^n_{opt,\delta t}$ for any $1 \leq n \leq M$ then, provided that $\delta t \leq C_T h^\gamma$ (and $\delta t \rho_h \leq \delta$ in the case $\theta = \frac{1}{2}$), we have

- *The cost of the control $v_{\delta t} = (v^n)_n \in U_h^M$ is bounded as follows*

$$\sum_{n=1}^{M} \delta t \left[\!\left[v^n\right]\!\right]_h^2 \leq (C_{\rm obs}^2 + \phi(h))\left(\left\|y_0 - \hat{y}_0\right\|_h + C_s \delta t^{\zeta_1 s}\left\|\hat{y}_0\right\|_{\zeta_2 s,h}\right)^2 + e^{-C/\delta t^{\zeta_3}}\left\|\hat{y}_0\right\|_h^2, \quad (5.9)$$

  where $\zeta_1 = \zeta_3 = 1$ and $\zeta_2 = 4$ for $\theta > \frac{1}{2}$ and $\zeta_1 = 2, \zeta_2 = 6$ and $\zeta_3 = 2/3$ for $\theta = \frac{1}{2}$.

- *The controlled solution $(y^n)_n$ associated to $v_{\delta t}$ and to the initial data $y^0 = y_0$ is such that*

$$\left\|y^M - \hat{y}_F\right\|_h \leq \sqrt{\phi(h)}(C_{\rm obs} + \sqrt{\phi(h)})\left(\left\|y_0 - \hat{y}_0\right\|_h + C\delta t^{\zeta_1 s}\left\|\hat{y}_0\right\|_{\zeta_2 s,h}\right) + e^{-C/\delta t^{\zeta_3}}\left\|\hat{y}_0\right\|_h. \quad (5.10)$$

Finally, the optimal adjoint state $q^F_{opt,\delta t}$ satisfies

$$\sqrt{\phi(h)}\left\|q^F_{opt,\delta t}\right\|_h \leq (C_{\rm obs} + \sqrt{\phi(h)})\left(\left\|y_0 - \hat{y}_0\right\|_h + C\delta t^{\zeta_1 s}\left\|\hat{y}_0\right\|_{\zeta_2 s,h}\right) + e^{-C/\delta t^{\zeta_3}}\left\|\hat{y}_0\right\|_h. \quad (5.11)$$

Remark 5.3.

1. *When δt goes to 0, the above estimates converge to their counterpart for the semi-discrete problem given in Theorem 5.1. Moreover, if we assume some regularity on the initial datum, that is if $\left\|\hat{y}_0\right\|_{r,h}$ is bounded w.r.t. h for some $r > 0$, then convergence is uniform with respect to h.*
2. *If we are interested in the null-controllability problem, then $\hat{y}_0 = \hat{y}_F = 0$ and the above estimates take simpler forms.*
3. *We have assumed $\delta t \leq C_T h^\gamma$. Together with Assumption (5.1) made on ϕ, and the fact that ϕ is bounded on $[0, h_0]$, we deduce (with $s = 0$) the simpler useful estimates*

$$\sum_{n=1}^{M} \delta t \left[\!\left[v^n\right]\!\right]_h^2 \leq C^2\left(\left\|y_0 - \hat{y}_0\right\|_h + \left\|\hat{y}_0\right\|_h\right)^2, \quad (5.12)$$

$$\left\|y^M - \hat{y}_F\right\|_h \leq C\sqrt{\phi(h)}\left(\left\|y_0 - \hat{y}_0\right\|_h + \left\|\hat{y}_0\right\|_h\right), \quad (5.13)$$

$$\sqrt{\phi(h)}\left\|q^F_{opt,\delta t}\right\|_h \leq C\left(\left\|y_0 - \hat{y}_0\right\|_h + \left\|\hat{y}_0\right\|_h\right). \quad (5.14)$$

Proof. Like in the semi-discrete situation, the functional $J^{h,\delta t}$ is smooth, strictly convex, and coercive which implies the existence and uniqueness of a minimizer. Using the same computations as in the proof of Theorem 4.1 we easily obtain, by a discrete integration by parts (see formula 4.6), that the controlled solution $(y^n)_n$ computed with the control function $v_{\delta t}$ satisfies $y^M = \hat{y}_F - \phi(h) q^F_{opt,\delta t}$. Hence, estimate (5.10) is a consequence of (5.11).

Let us prove (5.9) and (5.11). To this end, we need to take a special care of high frequencies. More precisely, let us introduce the following notation:

- For $\frac{1}{2} < \theta \leq 1$, we define
$$A_\theta = \frac{2\theta - 1}{\theta(1-\theta)}, \tag{5.15}$$
and we introduce the orthogonal projector $\Pi_{\theta,\delta t}$ onto the space spanned by the eigenvectors $\mathcal{M}_h^{-1}\mathcal{A}_h$ associated to the eigenvalues less than or equal to $A_\theta \delta t^{-1}$.
- For $\theta = \frac{1}{2}$, we define
$$A_{\frac{1}{2}} = T^{-1/3}. \tag{5.16}$$
ans similarly $\Pi_{\frac{1}{2},\delta t}$ denotes the orthogonal projector onto the space spanned by the eigenvectors $\mathcal{M}_h^{-1}\mathcal{A}_h$ associated to the eigenvalues less than or equal to $A_{\frac{1}{2}}\delta t^{-2/3}$, with

In each case, we denote by $\Pi^\perp_{\frac{1}{2},\delta t} = \mathrm{Id} - \Pi_{\frac{1}{2},\delta t}$. We now define $\tilde{y}_0^{\delta t}$ as the unique solution to the following system

$$\begin{cases} (\mathrm{Id} + \delta t \theta \mathcal{M}_h^{-1}\mathcal{A}_h)^{-M} (\mathrm{Id} - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h)^M \tilde{y}_0^{\delta t} = \Pi_{\theta,\delta t}\hat{y}_T, \\ \Pi_{\theta,\delta t}\tilde{y}_0^{\delta t} = \tilde{y}_0^{\delta t}. \end{cases} \tag{5.17}$$

Notice that this system is uniquely solvable since, by construction, the range of the projector $\Pi_{\theta,\delta t}$ intersect the kernel of $\mathrm{Id} - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h$ trivially even if this kernel is not reduced to $\{0\}$.

With the above notation, the functional $J^{h,\delta t}$ can be expressed as follows

$$\begin{aligned} J^{h,\delta t}(q_F) = \frac{1}{2}\sum_{n=1}^{M} \delta t \left[\!\left[\mathcal{B}_h^\star q^n\right]\!\right]_h^2 + \frac{\phi(h)}{2}\left\|q_F\right\|_h^2 \\ - \left\langle q_F, \Pi^\perp_{\theta,\delta t}\hat{y}_T\right\rangle_h + \left\langle y_0 - \tilde{y}_0^{\delta t}, q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1\right\rangle_h, \end{aligned}$$

and the associated Euler-Lagrange equation reads

$$\begin{aligned} \sum_{n=1}^{M} \delta t \left[\mathcal{B}_h^\star q^n_{opt,\delta t}, \mathcal{B}_h^\star q^n\right]_h + \phi(h)\left\langle q^F_{opt,\delta t}, q_F\right\rangle_h \\ - \left\langle q_F, \Pi^\perp_{\theta,\delta t}\hat{y}_T\right\rangle_h + \left\langle y_0 - \tilde{y}_0^{\delta t}, q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1\right\rangle_h = 0, \end{aligned} \tag{5.18}$$

for any $q_F \in E_h$, and $(q^n)_n$ its associated solution to the fully-discrete adjoint system. Choosing $q_F = q^F_{opt,\delta t}$ in (5.18) leads to

$$\begin{aligned} \sum_{n=1}^{M} \delta t \left[\!\left[\mathcal{B}_h^\star q^n_{opt,\delta t}\right]\!\right]_h^2 + \phi(h)\left\|q^F_{opt,\delta t}\right\|_h^2 \\ - \left\langle q^F_{opt,\delta t}, \Pi^\perp_{\theta,\delta t}\hat{y}_T\right\rangle_h + \left\langle y_0 - \tilde{y}_0^{\delta t}, q^1_{opt,\delta t} - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1_{opt,\delta t}\right\rangle_h = 0. \end{aligned} \tag{5.19}$$

We can now choose h_0 sufficiently small so that $\phi(h) \geq C_1 e^{-C_2/h^\gamma}$, for any $h \leq h_0$, where C_1 and C_2 are defined in Theorem 4.7. The following observability inequality then holds

$$\left\|q^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^1\right\|_h \leq C_{\text{obs}} \left(\sum_{n=1}^{M} \delta t \left\|\mathcal{B}_h^\star q^n\right\|_h^2\right)^{\frac{1}{2}} + \phi(h)\left\|q_F\right\|_h, \quad \forall q_F \in E_h.$$

Moving back to (5.19), we deduce by using Young inequalities that

$$\left\|q^F_{opt,\delta t}\right\|_h \leq \frac{1}{\phi(h)}\left\|\Pi^\perp_{\theta,\delta t}\hat{y}_T\right\|_h + (\phi(h) + C_{\text{obs}}\sqrt{\phi(h)}/2)\left\|y_0 - \tilde{y}_0^{\delta t}\right\|_h,$$

$$\sum_{n=1}^{M} \delta t \left[\!\left[\mathcal{B}_h^\star q^n_{opt,\delta t}\right]\!\right]_h^2 \leq \frac{1}{\phi(h)}\left\|\Pi^\perp_{\theta,\delta t}\hat{y}_T\right\|_h^2 + (C_{\text{obs}}^2 + \phi(h))\left\|y_0 - \tilde{y}_0^{\delta t}\right\|_h^2.$$

It thus remains to bound $\left\|\Pi^\perp_{\theta,\delta t}\hat{y}_T\right\|_h$ and $\left\|y_0 - \tilde{y}_0^{\delta t}\right\|_h$.

First by the definition of the projector $\Pi_{\theta,\delta t}$, and since $\hat{y}_F = e^{-T\mathcal{M}_h^{-1}\mathcal{A}_h}\hat{y}_0$, we have the estimate

$$\left\|\Pi^\perp_{\theta,\delta t}\hat{y}_T\right\|_h \leq \begin{cases} e^{-TA_\theta \delta t^{-1}}\left\|\hat{y}_0\right\|_h, & \text{for } \theta > \frac{1}{2}, \\ e^{-TA_{\frac{1}{2}}\delta t^{-\frac{2}{3}}}\left\|\hat{y}_0\right\|_h, & \text{for } \theta = \frac{1}{2}. \end{cases}$$

Since we assumed $\delta t \leq C_T h^\gamma$, and with Assumption (5.1) made on ϕ, we deduce that, if h_0 is sufficiently small, we have, for some $C > 0$

$$\frac{1}{\phi(h)}\left\|\Pi^\perp_{\theta,\delta t}\hat{y}_T\right\|_h^2 \leq \begin{cases} e^{-C\delta t^{-1}}\left\|\hat{y}_0\right\|_h^2, & \text{for } \theta > \frac{1}{2}, \\ e^{-C\delta t^{-2/3}}\left\|\hat{y}_0\right\|_h^2, & \text{for } \theta = \frac{1}{2}. \end{cases}$$

Second, we estimate $\tilde{y}_0^{\delta t} - y_0$, as follows, for any $0 \leq s \leq 1$:

$$\left\|\tilde{y}_0^{\delta t} - y_0\right\|_h \leq \begin{cases} \left\|\hat{y}_0 - y_0\right\|_h + C\delta t^s \left\|\hat{y}_0\right\|_{4s,h}, & \text{for } \theta > \frac{1}{2}, \\ \left\|\hat{y}_0 - y_0\right\|_h + C\delta t^{2s}\left\|\hat{y}_0\right\|_{6s,h}, & \text{for } \theta = \frac{1}{2}. \end{cases} \tag{5.20}$$

Indeed, it suffices to prove

$$\left\|\tilde{y}_0^{\delta t} - \hat{y}_0\right\|_h \leq \begin{cases} C\delta t^s \left\|\hat{y}_0\right\|_{4s,h}, & \text{for } \theta > \frac{1}{2}, \\ C\delta t^{2s}\left\|\hat{y}_0\right\|_{6s,h}, & \text{for } \theta = \frac{1}{2}. \end{cases}$$

We first write

$$\left\|\tilde{y}_0^{\delta t} - \hat{y}_0\right\|_h \leq \left\|\tilde{y}_0^{\delta t} - \Pi_{\theta,\delta t}\hat{y}_0\right\|_h + \left\|\Pi_{\theta,\delta t}\hat{y}_0 - \hat{y}_0\right\|_h.$$

By definition of the projector, the second term is bounded by $C\delta t^{2s}\left\|\hat{y}_0\right\|_{4s,h}$ for $\theta > \frac{1}{2}$ and by $C\delta t^{2s}\left\|\hat{y}_0\right\|_{6s,h}$ for $\theta = \frac{1}{2}$.

Then, by definition of the target $\hat{y}_T$, we can write

$$\begin{aligned} &\tilde{y}_0^{\delta t} - \Pi_{\theta,\delta t}\hat{y}_0 \\ &\quad = \left[(\text{Id} + \delta t\theta\mathcal{M}_h^{-1}\mathcal{A}_h)^M(\text{Id} - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h)^{-M}e^{-M\delta t\mathcal{M}_h^{-1}\mathcal{A}_h} - \text{Id}\right]\Pi_{\theta,\delta t}\hat{y}_0. \end{aligned} \tag{5.21}$$

In order to estimate this quantity, we develop $\Pi_{\theta,\delta t}\hat{y}_0$ in the basis of eigenvectors of $\mathcal{M}_h^{-1}\mathcal{A}_h$ and we use the inequality (5.22) given by Lemma 5.4 above. With this inequality and (5.21) we readily obtain the announced estimate (5.20). □

LEMMA 5.4. *For any M, δt such that $M\delta t = T$, for any $0 \le s \le 1$, there exists $C > 0$ only depending on θ and s such that*

$$\left|1 - \left(\frac{1+\delta t\theta\lambda}{1-\delta t(1-\theta)\lambda}\right)^M e^{-M\delta t\lambda}\right| \le \begin{cases} CT^s\delta t^s\lambda^{2s}, & \forall\lambda \ge 0, \text{ with } \lambda \le A_\theta\delta t^{-1}, \text{ for } \theta > \frac{1}{2}, \\ CT^s\delta t^{2s}\lambda^{3s}, & \forall\lambda \ge 0, \text{ with } \lambda \le A_{\frac{1}{2}}\delta t^{-2/3}, \text{ for } \theta = \frac{1}{2}. \end{cases} \tag{5.22}$$

We recall that A_θ and $A_{\frac{1}{2}}$ are defined in (5.15) *and* (5.16).

A proof of this technical lemma is given in Appendix A.

For instance if one chooses $\phi(h) = h^{2p}$ for any given p, with the above results, one can construct a uniformly bounded sequence of controls leading to a final state whose distance to the target $\hat{y}_F$ is no larger than $h^p C'_{\text{obs}}(\|y_0 - \hat{y}_0\|_h + \|\hat{y}_0\|_h)$, for h sufficiently small, where C'_{obs} is any given number larger than C_{obs}.

Notice in particular that the value of C'_{obs} does not depend on the particular choice of ϕ we use.

The practical computation of $q^F_{opt,\delta t}$ can be performed by a conjugate gradient solver as proposed in [GL94]. Each iteration of this solver consists in first solving for the solution to a fully-discrete adjoint problem, and second solving for the solution to a fully-discrete direct problem.

6. Error estimates. This section is devoted to the analysis of the convergence of $v_{\delta t}$ towards v when $\delta t \to 0$, where $v_{\delta t}$ and v are respectively the fully-discrete and the semi-discrete control functions obtained in Theorems 5.1 and 5.2.

We begin with simple lemmata on error estimates for the implicit Euler scheme for the simple ODE problem $\mathcal{M}_h\partial_t y + \mathcal{A}_h y = 0$.

LEMMA 6.1. *For any $\delta \ge 0$, there exists $C_\delta > 0$ such that*

$$\forall t > 0, \forall s \in \mathbb{R}, \forall y_0 \in E_h, \quad \|y(t)\|^2_{s,h} \le \frac{C}{t^\delta}\|y_0\|^2_{s-\delta,h}.$$

$$\forall T > 0, \forall s \in \mathbb{R}, \forall y_0 \in E_h, \quad \int_0^T t^\delta \|y(t)\|^2_{s,h}\, dt \le C\|y_0\|^2_{s-\delta-1,h}.$$

Proof. The solution y is explicitly given by $y(t) = e^{-t\mathcal{M}_h^{-1}\mathcal{A}_h}y_0$. The first property comes from the actual computation of the norm of $y(T)$ in an orthonormal basis of eigenvectors of $\mathcal{M}_h^{-1}\mathcal{A}_h$. Similarly, for the second property we write

$$\int_0^T t^\delta \|y(t)\|^2_{s,h}\, dt \le \int_0^{+\infty} t^\delta \left\langle (\mathcal{M}_h^{-1}\mathcal{A}_h)^s e^{-2t\mathcal{M}_h^{-1}\mathcal{A}_h}y_0, y_0\right\rangle_h dt.$$

We now write the decomposition of y_0 in the orthonormal basis $(\psi_i)_i$. The contribution of each component of y_0 in this decomposition in the above integral is then estimated through the following integral identity

$$\int_0^{+\infty} t^\delta \mu_i^s e^{-2t\mu_i}\, dt = \mu_i^{s-\delta-1}\int_0^{+\infty}(t\mu_i)^\delta e^{-2t\mu_i}\,\mu_i dt = C_\delta\mu_i^{s-\delta-1}.$$

The result then follows. □

6.1. The forward problem. Let us begin with the study of the case $\theta > \frac{1}{2}$. In that case, we know that the scheme is first order in time and we shall prove the following sharp estimates. These results are quite classical except that we want to obtain estimates which are uniform with respect to the approximation parameter h.

As E_h is finite dimensional the different norms $\|\cdot\|_{s,h}$, $s \in \mathbb{R}$, are equivalent. However, h is meant to go to zero (space discretization refinement), that is the dimension of E_h goes to infinity. For this reason, various norms $\|\cdot\|_{s,h}$, $s \in \mathbb{R}$, play different roles here.

PROPOSITION 6.2 (The θ-scheme for $\theta > \frac{1}{2}$). *Let $s \in \mathbb{R}$, $T > 0$ and $\frac{1}{2} < \theta \leq 1$. There exists $C > 0$ such that for any $y_0, y^0 \in E_h$, and any $M \in \mathbb{N}$, the solution $(y^n)_n \in E_h^M$ to the θ-scheme (with $\delta t = T/M$)*

$$\mathcal{M}_h \frac{y^{n+1} - y^n}{\delta t} + \mathcal{A}_h(\theta y^{n+1} + (1-\theta)y^n) = 0, \quad \forall n \in [\![0, M-1]\!],$$

and the solution $t \mapsto y(t)$ to $\mathcal{M}_h \partial_t y + \mathcal{A}_h y(t) = 0$ for the initial data y_0, satisfy the following error estimates

$$\sup_{0 \leq n \leq M} \|e^n\|_{s,h}^2 + \sum_{n=0}^{M-1} \delta t \left\|e^{n+\theta}\right\|_{s+1,h}^2 \leq C\delta t^2 \|y_0\|_{s+2,h}^2 + C \left\|y_0 - y^0\right\|_{s,h}^2, \tag{6.1}$$

$$\sum_{n=0}^{M-1} \delta t \left\|e^{n+1}\right\|_{s+1,h}^2 \leq C\delta t^2 \|y_0\|_{s+2,h}^2 + C \left\|y_0 - y^0\right\|_{s,h}^2 + C(1-\theta)^2 \delta t^2 \left[\delta t^2 \|y_0\|_{s+4,h}^2 + \left\|y_0 - y^0\right\|_{s+2,h}^2\right], \tag{6.2}$$

and

$$\sup_{0 \leq n \leq M} \|t^n e^n\|_{s,h}^2 \leq C\delta t^2 \|y_0\|_{s,h}^2 + C \left\|y_0 - y^0\right\|_{s-2,h}^2 + C\delta t^2 \left[\delta t^2 \|y_0\|_{s+2,h}^2 + \left\|y_0 - y^0\right\|_{s,h}^2\right], \tag{6.3}$$

with $e^n = y^n - y(t^n)$, $t^n = n\delta t$ and

$$e^{n+\theta} = \theta e^{n+1} + (1-\theta)e^n = (\theta - \frac{1}{2})(e^{n+1} - e^n) + \frac{1}{2}(e^{n+1} + e^n).$$

Proof.

1. Proof of (6.1). Observe that e^n solves

$$e^{n+1} - e^n + \delta t \mathcal{M}_h^{-1} \mathcal{A}_h e^{n+\theta} = \delta t R^{n+1}, \quad \forall n \in [\![0, M-1]\!], \tag{6.4}$$

where, the consistency residual is defined by

$$R^{n+1} = \delta t \int_0^1 (u + \theta - 1) y''(t^n + u\delta t)\, du. \tag{6.5}$$

As y is solution to $\mathcal{M}_h \partial_t y + \mathcal{A}_h y = 0$, we write

$$R^{n+1} = \delta t \int_0^1 (u + \theta - 1)(\mathcal{M}_h^{-1}\mathcal{A}_h)^2 y(t^n + u\delta t)\, du. \tag{6.6}$$

Forming the $\langle \cdot, \cdot \rangle_{s,h}$ inner product of (6.4) with $e^{n+\theta}$, we obtain

$$\frac{1}{2}\left\|e^{n+1}\right\|_{s,h}^2 - \frac{1}{2}\left\|e^{n}\right\|_{s,h}^2 + \left(\theta - \frac{1}{2}\right)\left\|e^{n+1} - e^{n}\right\|_{s,h}^2 + \delta t \left\|e^{n+\theta}\right\|_{s+1,h}^2 = \delta t \left\langle R^{n+1}, e^{n+\theta}\right\rangle_{s,h}.$$

Using (2.6) and Young's inequality yield

$$\left\|e^{n+1}\right\|_{s,h}^2 - \left\|e^{n}\right\|_{s,h}^2 + (2\theta - 1)\left\|e^{n+1} - e^{n}\right\|_{s,h}^2 + \delta t \left\|e^{n+\theta}\right\|_{s+1,h}^2 \leq \delta t \left\|R^{n+1}\right\|_{s-1,h}^2. \quad (6.7)$$

By (6.6), we have

$$\left\|R^{n+1}\right\|_{s-1,h}^2 \leq \delta t^2 \int_0^1 \|y(t^n + u\delta t)\|_{s+3,h}^2 \, du = \delta t \int_{t^n}^{t^{n+1}} \|y(t)\|_{s+3,h}^2 \, dt. \quad (6.8)$$

Summing (6.7) over n and using the above estimate for R^{n+1} gives

$$\sup_{0\leq n\leq M-1} \left\|e^{n+1}\right\|_{s,h}^2 + \sum_{n=0}^{M-1} \delta t \left\|e^{n+\theta}\right\|_{s+1,h}^2 \leq C\delta t^2 \int_0^T \|y(t)\|_{s+3,h}^2 \, dt + C\left\|e^0\right\|_{s,h}^2.$$

We conclude the proof of estimate (6.1) with Lemma 6.1.
Note that (6.7) also leads to the following useful estimate

$$\sum_{n=0}^{M-1} \left\|e^{n+1} - e^{n}\right\|_{s,h}^2 \leq C\delta t^2 \|y_0\|_{s+2,h}^2 + C\left\|y_0 - y^0\right\|_{s,h}^2. \quad (6.9)$$

2. Proof of (6.2). We write

$$e^{n+1} = e^{n+\theta} + (1-\theta)(e^{n+1} - e^{n}), \quad \forall n \in [\![0, M-1]\!],$$

so that

$$\left\|e^{n+1}\right\|_{s+1,h}^2 \leq 2\left\|e^{n+\theta}\right\|_{s+1,h}^2 + 2(1-\theta)^2\left\|e^{n+1} - e^{n}\right\|_{s+1,h}^2.$$

Using (6.4), we can then obtain

$$\left\|e^{n+1}\right\|_{s+1,h}^2 \leq 2\left\|e^{n+\theta}\right\|_{s+1,h}^2 + 4(1-\theta)^2\delta t^2 \left\|e^{n+\theta}\right\|_{s+3,h}^2 + 4(1-\theta)^2\delta t^2 \left\|R^{n+1}\right\|_{s+1,h}^2. \quad (6.10)$$

It is now possible to conclude by using (6.1) and (6.8).

3. Proof of (6.3). From (6.4) we deduce the following equation satisfied by $t^n e_n$:

$$t^{n+1}e^{n+1} - t^n e^n + \delta t \mathcal{M}_h^{-1}\mathcal{A}_h t^{n+1} e^{n+\theta} = \delta t\left(t^{n+1}R^{n+1}\right) + \delta t e^n. \quad (6.11)$$

We introduce $t^{n+\theta} = \theta t^{n+1} + (1-\theta)t^n$ and we observe that

$$t^{n+\theta}e^{n+\theta} = \theta t^{n+1}e^{n+1} + (1-\theta)t^n e^n - \delta t\theta(1-\theta)(e^{n+1} - e^n).$$

Forming the $\langle\cdot,\cdot\rangle_{s,h}$ inner product of (6.11) with $t^{n+\theta}e^{n+\theta}$, using the above formula and noting that $t^{n+\theta}\leq t^{n+1}$, we obtain

$$\frac{1}{2}\left\|t^{n+1}e^{n+1}\right\|_{s,h}^2-\frac{1}{2}\left\|t^ne^n\right\|_{s,h}^2+\left(\theta-\frac{1}{2}\right)\left\|t^{n+1}e^{n+1}-t^ne^n\right\|_{s,h}^2$$
$$+\delta t\left\|t^{n+\theta}e^{n+\theta}\right\|_{s+1,h}^2\leq\delta t\left\langle t^{n+1}R^{n+1},t^{n+\theta}e^{n+\theta}\right\rangle_{s,h}+\delta t\left\langle e^n,t^{n+\theta}e^{n+\theta}\right\rangle_{s,h}$$
$$+\delta t\theta(1-\theta)\left\langle t^{n+1}e^{n+1}-t^ne^n,e^{n+1}-e^n\right\rangle_{s,h}. \quad (6.12)$$

We recall that we assume $\theta>\frac{1}{2}$. By using (2.6) and Young's inequalities, we deduce

$$\left\|t^{n+1}e^{n+1}\right\|_{s,h}^2-\left\|t^ne^n\right\|_{s,h}^2\leq\delta t\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2+\delta t\left\|e^n\right\|_{s-1,h}^2$$
$$+C(1-\theta)^2\delta t^2\left\|e^{n+1}-e^n\right\|_{s,h}^2,\quad\forall n\in[\![1,M-1]\!],$$

and, by multiplying (6.7) by $\delta t^2=(t^1)^2$, we have

$$\left\|t^1e^1\right\|_{s,h}^2\leq\delta t\left\|t^1R^1\right\|_{s-1,h}^2+2\delta t^2\left\|e^0\right\|_{s,h}^2. \quad (6.13)$$

We now sum these inequalities to obtain

$$\sup_{1\leq n\leq M}\left\|t^ne^n\right\|_{s,h}^2\leq\sum_{n=0}^{M-1}\delta t\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2+\sum_{n=1}^{M-1}\delta t\left\|e^n\right\|_{s-1,h}^2$$
$$+C(1-\theta)^2\sum_{n=0}^{M-1}\delta t^2\left\|e^{n+1}-e^n\right\|_{s,h}^2+2\delta t^2\left\|y_0-y^0\right\|_{s,h}^2. \quad (6.14)$$

The contribution of the second term in the right-hand side was already estimated in (6.2) (with $s-2$ instead of s) and (6.9) gives a bound for the third term. It remains now to estimate the contribution of the first term in the right-hand side of (6.14) by writing

$$t^{n+1}R^{n+1}=\delta t\int_0^1(u+\theta-1)(t^n+u\delta t)(\mathcal{M}_h^{-1}\mathcal{A}_h)^2y(t^n+u\delta t)\,du$$
$$+\delta t^2\int_0^1(1-u)(u+\theta-1)(\mathcal{M}_h^{-1}\mathcal{A}_h)^2y(t^n+u\delta t)\,du. \quad (6.15)$$

It follows that

$$\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2\leq C\delta t^2\int_0^1(t^n+u\delta t)^2\left\|y(t^n+u\delta t)\right\|_{s+3,h}^2\,du$$
$$+C\delta t^4\int_0^1\left\|y(t^n+u\delta t)\right\|_{s+3,h}^2\,du.$$

Using finally Lemma 6.1, we obtain

$$\sum_{n=0}^{M-1}\delta t\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2$$
$$\leq C\delta t^2\int_0^T t^2\left\|y(t)\right\|_{s+3,h}^2\,dt+C\delta t^4\int_0^T\left\|y(t)\right\|_{s+3,h}^2\,dt$$
$$\leq C\delta t^2\left\|y_0\right\|_{s,h}^2+C\delta t^4\left\|y_0\right\|_{s+2,h}^2.$$

This concludes the proof. □

We now study the case of the Crank-Nicolson scheme. As expected, we find that the scheme is second order.

PROPOSITION 6.3 (The Crank-Nicolson scheme). *We consider the same notation as in Proposition 6.2, except that we assume now that* $\theta = \frac{1}{2}$.

For any $s \in \mathbb{R}$, *we have the following estimates*

$$\sup_{0\le n\le M} \|e^n\|_{s,h}^2 + \sum_{n=0}^{M-1} \delta t \left\|e^{n+\frac{1}{2}}\right\|_{s+1,h}^2 \le C\delta t^4 \|y_0\|_{s+4,h}^2 + C\left\|y_0 - y^0\right\|_{s,h}^2, \quad (6.16)$$

$$\sum_{n=0}^{M-1} \delta t \left\|e^{n+1}\right\|_{s+1,h}^2 \le C\delta t^4 \|y_0\|_{s+4,h}^2 + C\delta t^6 \|y_0\|_{s+6,h}^2 + C\left\|y_0 - y^0\right\|_{s,h}^2 + C\delta t^2 \left\|y_0 - y^0\right\|_{s+2,h}^2, \quad (6.17)$$

$$\sup_{0\le n\le M} \|t^n e^n\|_{s,h}^2 \le C(1+(\delta t\rho_h)^4)\left[\delta t^4 \|y_0\|_{s+2,h}^2 + \left\|y_0 - y^0\right\|_{s-2,h}^2\right], \quad (6.18)$$

$$\sup_{\frac{M}{2}\le n\le M} \left\|\left(t^n - \frac{T}{2}\right)e^n\right\|_{s,h}^2 \le C(1+(\delta t\rho_h)^8)\left[\delta t^4 \|y_0\|_{s,h}^2 + \left\|y_0 - y^0\right\|_{s-4,h}^2\right]. \quad (6.19)$$

Proof. The proof follows the same lines as that of Proposition 6.2, and we only mention here the points that need to be changed.

- When $\theta = \frac{1}{2}$, inequality (6.7) does not contain any numerical diffusion term like $\left\|e^{n+1} - e^n\right\|_{s,h}^2$. Yet, this term is useful in the sequel. Using (2.5) and (6.4), we can recover such a term by observing that

$$\begin{aligned}\left\|e^{n+1} - e^n\right\|_{s,h}^2 &\le \rho_h \left\|e^{n+1} - e^n\right\|_{s-1,h}^2 \\ &\le 2\rho_h\delta t^2 \left\|e^{n+\frac{1}{2}}\right\|_{s+1,h}^2 + 2\rho_h\delta t^2 \left\|R^{n+1}\right\|_{s-1,h}^2.\end{aligned}$$

  Combining this last inequality with (6.7), we obtain

$$\left\|e^{n+1}\right\|_{s,h}^2 - \|e^n\|_{s,h}^2 + \frac{1}{4\delta t\rho_h}\left\|e^{n+1} - e^n\right\|_{s,h}^2 + \frac{\delta t}{2}\left\|e^{n+\frac{1}{2}}\right\|_{s+1,h}^2 \le \frac{3}{2}\delta t \left\|R^{n+1}\right\|_{s-1,h}^2,$$

  which is the desired estimate. Note that, as expected, this estimate is only useful under the assumption that $\delta t\rho_h$ is bounded.
  Integrating by parts, we find that R^{n+1}, as given in (6.5), can be expressed as follows, in the case $\theta = \frac{1}{2}$,

$$R^{n+1} = \frac{\delta t^2}{2}\int_0^1 u(1-u)y'''(t^n + u\delta t)\,du,$$

  leading to the following estimate

$$\left\|R^{n+1}\right\|_{s-1,h}^2 \le C\delta t^3 \int_{t^n}^{t^{n+1}} \|y(t)\|_{s+5,h}^2\,dt.$$

Thus, with Lemma 6.1 we deduce

$$\sup_{0\le n\le M-1}\left\|e^{n+1}\right\|_{s,h}^2+\sum_{n=0}^{M-1}\delta t\left\|e^{n+\frac12}\right\|_{s+1,h}^2\le C\delta t^4\left\|y_0\right\|_{s+4,h}^2+C\left\|e^0\right\|_{s,h}^2,$$

as well as the estimate

$$\sum_{n=0}^{M-1}\left\|e^{n+1}-e^n\right\|_{s,h}^2\le C\delta t\rho_h\left(\delta t^4\left\|y_0\right\|_{s+4,h}^2+\left\|e^0\right\|_{s,h}^2\right). \tag{6.20}$$

- Estimate (6.10) still holds and gives

$$\left\|e^{n+1}\right\|_{s+1,h}^2\le 2\left\|e^{n+\frac12}\right\|_{s+1,h}^2+\delta t^2\left\|e^{n+\frac12}\right\|_{s+3,h}^2+\delta t^2\left\|R^{n+1}\right\|_{s+1,h}^2,$$

so that, using the previous bounds, we obtain (6.17).
- Estimate (6.12) now leads to

$$\begin{aligned}&\frac12\left\|t^{n+1}e^{n+1}\right\|_{s,h}^2-\frac12\left\|t^ne^n\right\|_{s,h}^2+\delta t\left\|t^{n+\frac12}e^{n+\frac12}\right\|_{s+1,h}^2\\&\le\delta t\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}\left\|t^{n+\frac12}e^{n+\frac12}\right\|_{s+1,h}+\delta t\left\|e^n\right\|_{s-1,h}\left\|t^{n+\frac12}e^{n+\frac12}\right\|_{s+1,h}\\&\qquad+\frac{\delta t}{4}\left\|t^{n+1}e^{n+1}-t^ne^n\right\|_{s,h}\left\|e^{n+1}-e^n\right\|_{s,h}.\end{aligned}$$

Applying Young's inequality we have

$$\begin{aligned}&\left\|t^{n+1}e^{n+1}\right\|_{s,h}^2-\left\|t^ne^n\right\|_{s,h}^2+\delta t\left\|t^{n+\frac12}e^{n+\frac12}\right\|_{s+1,h}^2\\&\qquad\le2\delta t\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2+2\delta t\left\|e^n\right\|_{s-1,h}^2\\&\qquad+\frac{1}{6\delta t\rho_h}\left\|t^{n+1}e^{n+1}-t^ne^n\right\|_{s,h}^2+\frac38\delta t^3\rho_h\left\|e^{n+1}-e^n\right\|_{s,h}^2.\end{aligned}$$

We now need to prove some bound for $\left\|t^{n+1}e^{n+1}-t^ne^n\right\|_{s,h}$ as this term is not present in the l.h.s. of the previous estimate, unlike the case $\theta>\frac12$ where such numerical diffusion is helpful. To this end, we proceed as follows by using (6.11)

$$\begin{aligned}\frac{1}{6\delta t\rho_h}\left\|t^{n+1}e^{n+1}-t^ne^n\right\|_{s,h}^2&\le\frac{1}{6\delta t}\left\|t^{n+1}e^{n+1}-t^ne^n\right\|_{s-1,h}^2\\&\le\frac{\delta t}{2}\left\|t^{n+1}e^{n+\frac12}\right\|_{s+1,h}^2+\frac{\delta t}{2}\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2+\frac{\delta t}{2}\left\|e^n\right\|_{s-1,h}^2.\end{aligned}$$

Adding the previous two inequalities, we obtain

$$\begin{aligned}&\left\|t^{n+1}e^{n+1}\right\|_{s,h}^2-\left\|t^ne^n\right\|_{s,h}^2+\frac{\delta t}{2}\left\|t^{n+\frac12}e^{n+\frac12}\right\|_{s+1,h}^2\\&\le\frac52\delta t\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2+\frac52\delta t\left\|e^n\right\|_{s-1,h}^2+\frac38\delta t^3\rho_h\left\|e^{n+1}-e^n\right\|_{s,h}^2.\end{aligned}\tag{6.21}$$

Furthermore, as $\theta = \frac{1}{2}$, we can integrate by parts in (6.15) to obtain

$$t^{n+1}R^{n+1} = \frac{\delta t^2}{2}\int_0^1 u(1-u)z'(t^n+u\delta t)\,du \\ + \delta t^2\int_0^1 (1-u)(u-\frac{1}{2})(\mathcal{M}_h^{-1}\mathcal{A}_h)^2 y(t^n+u\delta t)\,du,$$

where $z(t) = t(\mathcal{M}_h^{-1}\mathcal{A}_h)^2 y(t)$. It follows that

$$\sum_{n=0}^{M-1}\delta t\left\|t^{n+1}R^{n+1}\right\|_{s-1,h}^2 \\ \leq C\delta t^4\left(\int_0^T \|z'(t)\|_{s-1,h}^2\,dt + \int_0^T \|y(t)\|_{s+3,h}^2\,dt\right) \leq C\delta t^4\|y_0\|_{s+2,h}^2, \tag{6.22}$$

by using once more Lemma 6.1.
Finally, summing (6.21) for $n \in [\![1, M-1]\!]$ and (6.13), and using (6.17) (changing $s+1$ in $s-1$), (6.20) and (6.22), we obtain

$$\sup_{0\leq n\leq M}\|t^n e^n\|_{s,h}^2 \leq C\delta t^4\|y_0\|_{s+2,h}^2 + C(1+(\delta t\rho_h)^2)\delta t^6\|y_0\|_{s+4,h}^2 \\ + C\left\|e^0\right\|_{s-2,h}^2 + C(1+(\delta t\rho_h)^2)\delta t^2\left\|e^0\right\|_{s,h}^2.$$

Using (2.5), this proves the claimed estimate (6.18).

- It remains to prove (6.19). We define $M' = \lfloor M/2 \rfloor$. From (6.18), we obtain

$$\left\|e^{M'}\right\|_{s-2,h}^2 \leq \frac{C}{T^2}(1+(\delta t\rho_h)^4)\left[\delta t^4\|y_0\|_{s,h}^2 + \left\|e^0\right\|_{s-4,h}^2\right].$$

We now apply once more (6.18) for the same problem starting from $t = t^{M'}$ instead of $t = 0$ and with initial data $y(t^{M'})$ and $y^{M'}$ for the semi-discrete and fully-discrete problems respectively. We obtain

$$\left\|(t^M - \frac{T}{2})e^M\right\|_{s,h}^2 \leq \frac{C}{T^2}(1+(\delta t\rho_h)^4)\left[\delta t^4\left\|y(t^{M'})\right\|_{s+2,h}^2 + \left\|e^{M'}\right\|_{s-2,h}^2\right],$$

so that using Lemma 6.1 and the bound on $e^{M'}$ previously obtained we finally reach the claim.

□

6.2. The backward problem. For $X = E_h$ or $X = U_h$ and any family $x = (x^n)_{1\leq n\leq M} \in X^M$, we denote by $\mathcal{F}_0[x]$ the element of $L^2(0,T,X)$ defined by

$$\mathcal{F}_0[x](t) = \sum_{n=1}^M 1_{(t^{n-1},t^n)}(t)x^n. \tag{6.23}$$

Note that $\|\mathcal{F}_0[x]\|_{L^2(0,T,X)}^2 = \sum_{n=1}^M \delta t\|x^n\|_X^2$.

PROPOSITION 6.4 (The θ-scheme for $\theta > \frac{1}{2}$). *Let $q_F \in E_h$, $t \mapsto q(t) \in E_h$ be the solution to the adjoint problem $-\mathcal{M}_h \partial_t q(t) + \mathcal{A}_h q(t) = 0$ for the final data $q(T) = q_F$ and let $(q^n)_n \in E_h^M$ be the solution of the fully-discrete backward problem*

$$\begin{cases} \mathcal{M}_h \dfrac{q^M - q^{M+1}}{\delta t} + \theta \mathcal{A}_h q^M = 0, \\ \mathcal{M}_h \dfrac{q^n - q^{n+1}}{\delta t} + \mathcal{A}_h(\theta q^n + (1-\theta) q^{n+1}) = 0, \quad \forall n \in [\![1, M-1]\!], \end{cases} \tag{6.24}$$

for the same final data $q^{M+1} = q_F$. We have the following estimates

$$\sup_{1 \le n \le M+1} \|q^n\|_{s,h}^2 + \sum_{n=1}^{M} \delta t \left\|q^{n+1-\theta}\right\|_{s+1,h}^2 \le C \|q_F\|_{s,h}^2 + C(1-\theta)^2 \delta t \|q_F\|_{s+1,h}^2 , \tag{6.25}$$

$$\sum_{n=1}^{M} \delta t \|q^n\|_{s+1,h}^2 \le C \|q_F\|_{s,h}^2 + C(1-\theta)^4 \delta t^3 \|q_F\|_{s+3,h}^2 , \tag{6.26}$$

where $q^{n+1-\theta} = \theta q^n + (1-\theta) q^{n+1}$.

Moreover, if we introduce the error term

$$E^n = q^n - (1-\theta)\delta t \mathcal{M}_h^{-1} \mathcal{A}_h q^n - q(t^{n-1}), \qquad n \in [\![1, M+1]\!],$$

there exists $C > 0$ such that

$$\sup_{1 \le n \le M} \|E^n\|_{s,h}^2 + \sum_{n=1}^{M} \delta t \left\|E^{n+1-\theta}\right\|_{s+1,h}^2 \le C \delta t^2 \|q_F\|_{s+2,h}^2 , \tag{6.27}$$

$$\sum_{n=1}^{M} \delta t \|E^n\|_{s+1,h}^2 \le C \delta t^2 \|q_F\|_{s+2,h}^2 + C(1-\theta)^2 \delta t^4 \|q_F\|_{s+4,h}^2 , \tag{6.28}$$

$$\sup_{1 \le n \le M} \left\|(T - t^{n-1}) E^n\right\|_{s,h}^2 \le C \delta t^2 \|q_F\|_{s,h}^2 + C \delta t^4 \|q_F\|_{s+2,h}^2 . \tag{6.29}$$

Finally, we also have the estimate

$$\int_0^T \|q(t) - \mathcal{F}_0[(q^n)_n](t)\|_{s,h}^2 \, dt \le C \delta t^2 \|q_F\|_{s+1,h}^2 + C(1-\theta)^4 \delta t^5 \|q_F\|_{s+4,h}^2 . \tag{6.30}$$

Proof.

1. Arguing as in the proof of Proposition 6.2 we form the $\langle \cdot, \cdot \rangle_{s,h}$ inner product of (6.24) with $q^{n+1-\theta}$ and we sum over n to obtain

$$\begin{aligned} \sup_{1 \le n \le M+1} \|q^n\|_{s,h}^2 + (2\theta - 1) \sum_{n=1}^{M} \left\|q^{n+1} - q^n\right\|_{s,h}^2 + \sum_{n=1}^{M} \delta t \left\|q^{n+1-\theta}\right\|_{s+1,h}^2 \\ \le \|q_F\|_{s,h}^2 + 2\delta t (1-\theta) \left\langle \mathcal{M}_h^{-1} \mathcal{A}_h q_F, q^{M+1-\theta} \right\rangle_{s,h} . \end{aligned}$$

   By Young's inequality, we obtain

$$\sup_{1 \le n \le M+1} \|q^n\|_{s,h}^2 + \sum_{n=1}^{M} \delta t \left\|q^{n+1-\theta}\right\|_{s+1,h}^2 \le C \|q_F\|_{s,h}^2 + C\delta t (1-\theta)^2 \|q_F\|_{s+1,h}^2 .$$

Furthermore, by writing

$$q^n = q^{n+1-\theta} - (1-\theta)(q^{n+1} - q^n) = q^{n+1-\theta} - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^{n+1-\theta},$$

with $n \in [\![1, M-1]\!]$, and

$$\begin{aligned} q^M &= q^{M+1-\theta} - \delta t(1-\theta)\theta\mathcal{M}_h^{-1}\mathcal{A}_h q^M \\ &= q^{M+1-\theta} - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h q^{M+1-\theta} + \delta t(1-\theta)^2\mathcal{M}_h^{-1}\mathcal{A}_h q_F, \end{aligned}$$

and using (6.25) and (2.7), we obtain

$$\begin{aligned} \sum_{n=1}^{M} \delta t \left\|q^n\right\|_{s+1,h}^2 &\le C \sum_{n=1}^{M} \delta t \left\|q^{n+1-\theta}\right\|_{s+1,h}^2 + C(1-\theta)^4\delta t^2 \sum_{n=1}^{M} \delta t \left\|q^{n+1-\theta}\right\|_{s+3,h}^2 \\ &\quad + C\delta t^3(1-\theta)^4 \left\|q_F\right\|_{s+3,h}^2 \\ &\le C \left\|q_F\right\|_{s,h}^2 + C\delta t^3(1-\theta)^4 \left\|q_F\right\|_{s+3,h}^2. \end{aligned}$$

2. Let us introduce

$$z^{M+1} = q^{M+1} = q_F \quad \text{and} \quad z^n = (\mathrm{Id} - (1-\theta)\delta t\mathcal{M}_h^{-1}\mathcal{A}_h)q^n, \quad n \in [\![1, M]\!].$$

We observe that $(z^n)_{1\le n\le M+1}$ solves the usual backward θ-scheme (compare with (6.24))

$$\mathcal{M}_h \frac{z^n - z^{n+1}}{\delta t} + \mathcal{A}_h(\theta z^n + (1-\theta)z^{n+1}) = 0, \quad \forall n \in [\![1, M]\!].$$

Let us now define $y(t) = q(T-t)$ and $y^n = z^{M+1-n}$. We observe that $\mathcal{M}_h\partial_t y + \mathcal{A}_h y = 0$, $y(0) = q_F$ and that $(y^n)_n$ solves the forward θ-scheme with initial data $y^0 = q_F$. We also observe that

$$E^n = z^n - q(t^{n-1}) = y^{M+1-n} - y(T-t^{n-1}) = y^{M+1-n} - y(t^{M+1-n}) = e^{M+1-n},$$

in the notation of Proposition 6.2. Hence, (6.1), (6.2), (6.3) lead to (6.27), (6.28), (6.29) respectively.
It remains to prove (6.30). To this end we write

$$\begin{aligned} &\int_{t^{n-1}}^{t^n} \left\|q(t) - q^n\right\|_{s,h}^2 \, dt \\ &\quad = \int_{t^{n-1}}^{t^n} \left\|q(t) - q(t^{n-1}) - E^n - (1-\theta)\delta t\mathcal{M}_h^{-1}\mathcal{A}_h q^n)\right\|_{s,h}^2 \, dt \\ &\quad \le C \int_{t^{n-1}}^{t^n} \left\|q(t) - q(t^{n-1})\right\|_{s,h}^2 + C\delta t \left\|E^n\right\|_{s,h}^2 + C\delta t^3 \left\|q^n\right\|_{s+2,h}^2. \end{aligned}$$

The contributions of the sums over n of the last two terms can be estimated by using (6.26) and (6.28). It remains to consider the contribution of the first term. We proceed as follows

$$\begin{aligned} \int_{t^{n-1}}^{t^n} \left\|q(t) - q(t^{n-1})\right\|_{s,h}^2 \, dt &= \int_{t^{n-1}}^{t^n} \left\|\int_{t^{n-1}}^{t} q'(\tau)\, d\tau\right\|_{s,h}^2 \, dt \\ &\le \delta t \int_{t^{n-1}}^{t^n} \int_{t^{n-1}}^{t} \left\|q'(\tau)\right\|_{s,h}^2 \, d\tau \, dt \\ &\le \delta t^2 \int_{t^{n-1}}^{t^n} \left\|q(\tau)\right\|_{s+2,h}^2 \, d\tau, \end{aligned}$$

as $q(t)$ solves $\partial_t q(t) = \mathcal{M}_h^{-1}\mathcal{A}_h q(t)$. We obtain the claim by summing over n and using Lemma 6.1 (and the change of variable $y(t) = q(T-t)$).

□

We now proceed with the study of the backward problem for the Crank-Nicolson scheme.

To begin with, for any $x = (x^n)_{1\le n\le M} \in E_h^M$, we denote by $\mathcal{F}_1[x]$ the element of $L^2(0,T,E_h)$ defined as follows:

$$\mathcal{F}_1[x](t) = \sum_{n=1}^{M} 1_{]t^{n-1},t^n[}(t)(x^n + (t-t^{n-\frac12})\mathcal{M}_h^{-1}\mathcal{A}_h x^n). \tag{6.31}$$

PROPOSITION 6.5 (The Crank-Nicolson scheme). *If we take $\theta = \frac12$ in (6.24), then:*

- *Estimates (6.25) and (6.26) hold.*
- *With $E^n = q^n - \frac{\delta t}{2}\mathcal{M}_h^{-1}\mathcal{A}_h q^n - q(t^{n-1})$, $n \in [\![1, M+1]\!]$, Estimates (6.27)–(6.30) become*

$$\sup_{0\le n\le M} \|E^n\|_{s,h}^2 + \sum_{n=1}^{M} \delta t \left\|E^{n+\frac12}\right\|_{s+1,h}^2 \le C\delta t^4 \|q_F\|_{s+4,h}^2, \tag{6.32}$$

$$\sum_{n=1}^{M} \delta t \|E^n\|_{s+1,h}^2 \le C(1+(\delta t\rho_h)^2)\delta t^4 \|q_F\|_{s+4,h}^2, \tag{6.33}$$

$$\sup_{1\le n\le M/2} \left\|(T-t^{n-1})E^n\right\|_{s,h}^2 \le C\big(1+(\delta t\rho_h)^8\big)\delta t^4 \|q_F\|_{s,h}^2, \tag{6.34}$$

$$\int_0^T \|q(t) - \mathcal{F}_1[(q^n)_n](t)\|_{s,h}^2 \, dt \le C(1+(\delta t\rho_h)^2)\delta t^4 \|q_F\|_{s+3,h}^2. \tag{6.35}$$

Proof. The proof of (6.32) and (6.33) in Proposition 6.4 is not affected by the choice $\theta = \frac12$.

We proceed with the same change of variables as in the proof of Proposition 6.4 by defining $z^{M+1} = q^{M+1} = q_F$ and, for $n \in [\![1, M]\!]$, $z^n = (\mathrm{Id} - \frac{\delta t}{2}\mathcal{M}_h^{-1}\mathcal{A}_h)q^n$. We then set $y(t) = q(T-t)$ and $y^n = z^{M+1-n}$, $n \in [\![0, M]\!]$. Then (6.16), (6.17) and (6.19) imply (6.32), (6.33) and (6.34) respectively.

Let us now prove (6.35). We introduce $\tilde q = \mathcal{F}_1[(q^n)_n]$ and $\underline{q}(t)$ defined by

$$\underline{q}(t) = \sum_{n=1}^{M} 1_{]t^{n-1},t^n[}(t)\frac{1}{\delta t}\big(q(t^{n-1})(t^n - t) + q(t^n)(t - t^{n-1})\big).$$

Notice now that $\underline{q}$ and $\tilde q$ are continuous functions. In fact, for any $n \in [\![2, M]\!]$, we have

$$\tilde q(t_+^{n-1}) = q^n - \frac{\delta t}{2}\mathcal{M}_h^{-1}\mathcal{A}_h q^n, \quad \text{and} \quad \tilde q(t_-^{n-1}) = q^{n-1} + \frac{\delta t}{2}\mathcal{M}_h^{-1}\mathcal{A}_h q^{n-1},$$

and these two quantities are equal, as $(q^n)_n$ is solution of (6.24),

We shall now proceed in two steps. First, we estimate $\tilde q - \underline{q}$. We observe that this function is piecewise affine and continuous. We have

$$\tilde q(t^{n-1}) - \underline{q}(t^{n-1}) = q^n - \frac{\delta t}{2}\mathcal{M}_h^{-1}\mathcal{A}_h q^n - q(t^{n-1}) = E^n, \quad n \in [\![1, M]\!],$$

and

$$\tilde{q}(t^M) - \underline{q}(t^M) = q^M + \frac{\delta t}{2}\mathcal{M}_h^{-1}\mathcal{A}_h q^M - q(T) = q^{M+1} - q(T) = q_F - q_F = 0.$$

With a convexity argument we have

$$\int_{t^{n-1}}^{t^n} \left\|\tilde{q}(t) - \underline{q}(t)\right\|_{s,h}^2 \leq \frac{\delta t}{2}\big(\left\|E^{n+1}\right\|_{s,h}^2 + \|E^n\|_{s,h}^2 \big), \quad n \in [\![1, M]\!].$$

As a consequence, we find that

$$\int_0^T \left\|\tilde{q}(t) - \underline{q}(t)\right\|_{s,h}^2 dt \leq \sum_{n=1}^{M} \delta t \, \|E^n\|_{s,h}^2 .$$

This last term is bounded in (6.33).

Second, we give a bound for $q(t) - \underline{q}(t)$. Let $n \in [\![1, M+1]\!]$. By Taylor formulae, we find

$$\begin{aligned} q(t) - \frac{q(t^{n-1})(t^n - t) + q(t^n)(t - t^{n-1})}{\delta t} \\ = -\frac{(t - t^{n-1})(t^n - t)}{\delta t}\bigg((t - t^{n-1}) \int_0^1 (1-u) q''(t + u(t^{n-1} - t))\, du \\ + (t^n - t) \int_0^1 (1-u) q''(t + u(t^n - t))\, du \bigg), \end{aligned}$$

and we then obtain, with a Cauchy-Schwartz inequality,

$$\int_{t^{n-1}}^{t^n} \left\|q(t) - \underline{q}(t)\right\|_{s,h}^2 dt \leq \delta t^4 \int_{t^{n-1}}^{t^n} \|q''(t)\|_{s,h}^2 \, dt \leq \delta t^4 \int_{t^{n-1}}^{t^n} \|q(t)\|_{s+4,h}^2 \, dt.$$

By Lemma 6.1, we then obtain

$$\int_0^T \left\|q(t) - \underline{q}(t)\right\|_{s,h}^2 dt \leq C\delta t^4 \, \|q_F\|_{s+3,h}^2 .$$

Gathering the two estimates we obtain (6.35). □

6.3. Error estimate in time for the control problem. In this section, we give estimates of the error between the fully discrete control $v_{\delta t}$ and the semi-discrete one v corresponding to the same target $\hat{y}_F$, both defined in Theorems 5.1 and 5.2. Errors are estimated with respect to the time step, δt. As expected, the result will be different depending if $\theta > \frac{1}{2}$ or $\theta = \frac{1}{2}$.

6.3.1. The θ-scheme, $\theta > \frac{1}{2}$. We prove here a first-order in time error estimate.

THEOREM 6.6. *We consider the same assumptions as in Theorem 5.2, further assuming that $\theta > \frac{1}{2}$. Provided that the condition $\delta t \leq Ch^\gamma$ is fulfilled, the following first order error estimate holds*

$$\|\mathcal{F}_0[v_{\delta t}] - v\|_{L^2(0,T,U_h)} \leq C'\delta t \sqrt{\frac{\rho_h}{\phi(h)}}(1 + \delta t^{\frac{3}{2}} \rho_h^{\frac{3}{2}})(\|y_0\|_h + \|\hat{y}_0\|_h),$$

where C' is independent of δt and h.

We recall that we have chosen γ such that $0 < \gamma \leq \beta$. Notice that the constant in front of δt in the r.h.s. is not bounded with respect to h. The first-order convergence is thus not uniform with respect to h. The definition of $\mathcal{F}_0[.]$ can be found in (6.23).

Proof. We write the optimality conditions corresponding to the minimization of $J^{h,\delta t}$

$$0 = \sum_{n=1}^{M} \delta t \left[\mathcal{B}_h^\star q_{opt,\delta t}^n, \mathcal{B}_h^\star \tilde{q}^n\right]_h + \phi(h) \left\langle q_{opt,\delta t}^F, \tilde{q}^F \right\rangle_h - \left\langle \hat{y}_F, \tilde{q}^F \right\rangle_h + \left\langle y_0, \tilde{q}^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h \tilde{q}^1 \right\rangle_h, \quad (6.36)$$

for any $\tilde{q}_{\delta t} = (\tilde{q}^n)_n$ solution of the fully-discrete adjoint system (6.24) with $\tilde{q}^{M+1} = \tilde{q}^F \in E_h$. We also write the optimality conditions corresponding to the minimization of J^h

$$0 = \int_0^T \left[\mathcal{B}_h^\star q_{opt}(t), \mathcal{B}_h^\star \tilde{q}(t)\right]_h \, dt + \phi(h) \left\langle q_{opt}^F, \tilde{q}^F \right\rangle_h - \left\langle \hat{y}_F, \tilde{q}^F \right\rangle_h + \left\langle y_0, \tilde{q}(0) \right\rangle_h, \quad (6.37)$$

for any $t \mapsto \tilde{q}(t)$ solution of the adjoint problem (3.3) with $\tilde{q}(T) = \tilde{q}^F \in E_h$.

Let us now consider the first term in (6.36). For any $n \in [\![1, M]\!]$, we have

$$\delta t \left[\mathcal{B}_h^\star q_{opt,\delta t}^n, \mathcal{B}_h^\star \tilde{q}^n\right]_h = \int_{t^{n-1}}^{t^n} \left[\mathcal{F}_0[v_{\delta t}](t), \mathcal{B}_h^\star(\mathcal{F}_0[\tilde{q}_{\delta t}](t))\right]_h \, dt$$
$$= \int_{t^{n-1}}^{t^n} \left[\mathcal{F}_0[v_{\delta t}](t), \mathcal{B}_h^\star \tilde{q}(t)\right]_h \, dt + \int_{t^{n-1}}^{t^n} \left[\mathcal{F}_0[v_{\delta t}](t), \mathcal{B}_h^\star(\mathcal{F}_0[\tilde{q}_{\delta t}](t) - \tilde{q}(t))\right]_h \, dt. \quad (6.38)$$

Thus, (6.36) becomes

$$\int_0^T \left[\mathcal{F}_0[v_{\delta t}](t), \mathcal{B}_h^\star \tilde{q}(t)\right]_h \, dt + \phi(h) \left\langle q_{opt,\delta t}^F, \tilde{q}^F \right\rangle_h + \left\langle y_0, \tilde{q}(0) \right\rangle_h = -\int_0^T \left[\mathcal{F}_0[v_{\delta t}](t), \mathcal{B}_h^\star \tilde{e}(t)\right]_h \, dt - \left\langle y_0, E^1 \right\rangle_h, \quad (6.39)$$

where we recall that $E^1 = \tilde{q}^1 - \delta t(1-\theta)\mathcal{M}_h^{-1}\mathcal{A}_h\tilde{q}^1 - \tilde{q}(0)$ and we introduce

$$\tilde{e}(t) = \mathcal{F}_0[\tilde{q}_{\delta t}](t) - \tilde{q}(t), \ \forall t \in [0, T].$$

We now subtract (6.37) from (6.39) to obtain

$$\int_0^T \left[\mathcal{F}_0[v_{\delta t}](t) - v(t), \mathcal{B}_h^\star \tilde{q}(t)\right]_h \, dt + \phi(h) \left\langle q_{opt,\delta t}^F - q_{opt}^F, \tilde{q}^F \right\rangle_h = -\int_0^T \left[\mathcal{F}_0[v_{\delta t}](t), \mathcal{B}_h^\star \tilde{e}(t)\right]_h \, dt - \left\langle y_0, E^1 \right\rangle_h. \quad (6.40)$$

Let us now choose $\tilde{q}^F = q_{opt,\delta t}^F - q_{opt}^F$. The solution $\tilde{q}(t)$ of the semi-discrete backward problem associated to this data can be written

$$\tilde{q}(t) = \underline{q}_{\delta t}(t) - q_{opt}(t),$$

where $\underline{q}_{\delta t}$ denotes the solution of the Cauchy problem

$$-\mathcal{M}_h \partial_t \underline{q}_{\delta t} + \mathcal{A}_h \underline{q}_{\delta t} = 0, \qquad \underline{q}_{\delta t}(T) = q_{opt,\delta t}^F.$$

We split $\tilde{q}$ into two terms as follows

$$\tilde{q}(t) = \underline{q}_{\delta t}(t) - \mathcal{F}_0[q_{opt,\delta t}](t) + \mathcal{F}_0[q_{opt,\delta t}](t) - q_{opt}(t),$$

which gives

$$\mathcal{B}_h^\star \tilde{q}(t) = \mathcal{B}_h^\star \underbrace{(\underline{q}_{\delta t}(t) - \mathcal{F}_0[q_{opt,\delta t}](t))}_{=\underline{e}(t)} + \mathcal{F}_0[v_{\delta t}](t) - v(t).$$

Thus (6.40) leads to

$$\begin{aligned}
&\int_0^T [\![\mathcal{F}_0[v_{\delta t}](t) - v(t)]\!]_h^2 \, dt + \phi(h) \left\| q_{opt,\delta t}^F - q_{opt}^F \right\|_h^2 \\
&= -\int_0^T [\mathcal{F}_0[v_{\delta t}](t), \mathcal{B}_h^\star \tilde{e}(t)]_h \, dt - \left\langle y_0, E^1 \right\rangle_h - \int_0^T [\mathcal{F}_0[v_{\delta t}](t) - v(t), \mathcal{B}_h^\star \underline{e}(t)]_h \, dt.
\end{aligned} \tag{6.41}$$

It remains to estimate the three terms T_1, T_2 and T_3 in the r.h.s. of this inequality.

- With (2.3), (2.5), the bound (5.12) and to the error estimate (6.30), we find

$$\begin{aligned}
\left|T_1\right| &\le C \|\mathcal{F}_0[v_{\delta t}]\|_{L^2(0,T,U_h)} \|\tilde{e}\|_{L^2(0,T,E_h)} \\
&\le C(\|y_0\|_h + \|\hat{y}_0\|_h) \left(\delta t \left\| \tilde{q}^F \right\|_{1,h} + \delta t^{\frac{5}{2}} \left\| \tilde{q}^F \right\|_{4,h} \right) \\
&\le C(\|y_0\|_h + \|\hat{y}_0\|_h) \left\| \tilde{q}^F \right\|_h \left(\delta t \rho_h^{\frac{1}{2}} + \delta t^{\frac{5}{2}} \rho_h^2 \right).
\end{aligned}$$

- Using (6.29) and (2.5), the second term T_2 is bounded as follows

$$\begin{aligned}
\left|T_2\right| &\le \|y_0\|_h \left\| E^1 \right\|_h \le C \|y_0\|_h \left(\delta t \left\| \tilde{q}^F \right\|_h + \delta t^2 \left\| \tilde{q}^F \right\|_{2,h} \right) \\
&\le C \|y_0\|_h \left\| \tilde{q}^F \right\|_h (\delta t + \delta t^2 \rho_h).
\end{aligned}$$

- Using (2.3), we write for the last term T_3:

$$\left|T_3\right| \le C \|\mathcal{F}_0[v_{\delta t}] - v\|_{L^2(0,T,U_h)} \|\underline{e}\|_{L^2(0,T,E_h)},$$

  and $\underline{e}$ is estimated by (6.30) and (2.5) as follows

$$\begin{aligned}
\left|T_3\right| &\le C \|\mathcal{F}_0[v_{\delta t}] - v\|_{L^2(0,T,U_h)} (\delta t \left\| q_{opt,\delta t}^F \right\|_{1,h} + \delta t^{\frac{5}{2}} \left\| q_{opt,\delta t}^F \right\|_{4,h}) \\
&\le C \|\mathcal{F}_0[v_{\delta t}] - v\|_{L^2(0,T,U_h)} \left\| q_{opt,\delta t}^F \right\|_h (\delta t \rho_h^{\frac{1}{2}} + \delta t^{\frac{5}{2}} \rho_h^2).
\end{aligned}$$

We now collect the previous estimates in (6.41) and obtain

$$\begin{aligned}
&\|\mathcal{F}_0[v_{\delta t}] - v\|_{L^2(0,T,U_h)}^2 + \phi(h) \left\| \tilde{q}^F \right\|_h^2 \\
&\qquad \le C(\|y_0\|_h + \|\hat{y}_0\|_h) \left\| \tilde{q}^F \right\|_h (\delta t \rho_h^{\frac{1}{2}} + \delta t^{\frac{5}{2}} \rho_h^2 + \delta t + \delta t^2 \rho_h) \\
&\qquad\qquad + C \|\mathcal{F}_0[v_{\delta t}] - v\|_{L^2(0,T,U_h)} \left\| q_{opt,\delta t}^F \right\|_h (\delta t \rho_h^{\frac{1}{2}} + \delta t^{\frac{5}{2}} \rho_h^2).
\end{aligned} \tag{6.42}$$

Moreover, we have seen in (5.14) that $\sqrt{\phi(h)} \left\| q_{opt,\delta t}^F \right\|_h \le C(\|y_0\|_h + \|\hat{y}_0\|_h)$, so that, using Young's inequality and assumption (2.1), we finally obtain from (6.42) the expected error estimate

$$\|\mathcal{F}_0[v_{\delta t}] - v\|_{L^2(0,T,U_h)}^2 + \phi(h) \left\| \tilde{q}^F \right\|_h^2 \le C \frac{\rho_h}{\phi(h)} \delta t^2 (1 + \delta t^3 \rho_h^3)(\|y_0\|_h + \|\hat{y}_0\|_h)^2. \tag{6.43}$$

□

6.3.2. The Crank-Nicolson scheme. We prove now a second order in time error estimate.

THEOREM 6.7. *We consider the case $\theta = \frac{1}{2}$ and the same assumptions as in Theorem 5.2. Provided that the conditions $\delta t \leq Ch^\gamma$ and $\delta t \rho_h \leq \delta$ are fulfilled, the following second-order error estimate holds*

$$\|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}] - v\|_{L^2(0,T,U_h)} \leq C' \delta t^2 \frac{\rho_h}{\phi(h)^{\frac{1}{2}}} \left(\rho_h^{\frac{1}{2}} + \phi(h)^{-\frac{1}{2}}\right) (\|y_0\|_h + \|\hat{y}_0\|_h),$$

where C' is independent of δt and h.

The definition of $\mathcal{F}_1[.]$ can be found in (6.31). We recall that we have chosen γ such that $0 < \gamma \leq \beta$.

Proof. The proof follows the same lines as the previous one by replacing the operator $\mathcal{F}_0[\cdot]$ by $\mathcal{F}_1[\cdot]$. The main difference lies in the fact that, since $\mathcal{F}_1[\tilde{q}_{\delta t}]$ and $\mathcal{F}_1[q_{opt,\delta t}]$ are piecewise affine and not piecewise constant, we have to replace (6.38) by

$$\begin{aligned}
&\delta t \left[\mathcal{B}_h^\star q_{opt,\delta t}^n, \mathcal{B}_h^\star \tilde{q}^n\right]_h = \\
&\int_{t^{n-1}}^{t^n} [\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t), \mathcal{B}_h^\star \tilde{q}(t)]_h \, dt + \int_{t^{n-1}}^{t^n} [\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t), \mathcal{B}_h^\star (\mathcal{F}_1[\tilde{q}_{\delta t}](t) - \tilde{q}(t))]_h \, dt \\
&\qquad - \frac{\delta t^2}{12} \int_{t^{n-1}}^{t^n} \left[\mathcal{B}_h^\star \mathcal{F}_0\left[\mathcal{M}_h^{-1} \mathcal{A}_h q_{opt,\delta t}\right](t), \mathcal{B}_h^\star \mathcal{F}_0\left[\mathcal{M}_h^{-1} \mathcal{A}_h \tilde{q}_{\delta t}\right](t)\right]_h \, dt, \quad (6.44)
\end{aligned}$$

using that $\int_{t^{n-1}}^{t^n} (t - t^{n-\frac{1}{2}}) dt = 0$. Proceeding as in the proof of Theorem 6.6, using (2.3) in order to treat the new term (compare (6.44) with (6.38)), we obtain

$$\begin{aligned}
&\|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t) - v(t)\|_{L^2(0,T,U_h)}^2 + \phi(h) \left\|\tilde{q}^F\right\|_h^2 \\
&\qquad \leq \|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}]\|_{L^2(0,T,U_h)} \|\tilde{e}\|_{L^2(0,T,E_h)} + \|y_0\|_h \left\|E^1\right\|_h \\
&\qquad\quad + \|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t) - v(t)\|_{L^2(0,T,U_h)} \|\underline{e}\|_{L^2(0,T,E_h)} \\
&\qquad\quad + \frac{\delta t^2}{12} \Big(\sum_{n=1}^M \delta t \left\|q_{opt,\delta t}^n\right\|_{2,h}^2 \Big)^{\frac{1}{2}} \Big(\sum_{n=1}^M \delta t \left\|\tilde{q}^n\right\|_{2,h}^2 \Big)^{\frac{1}{2}}, \quad (6.45)
\end{aligned}$$

where

$$\tilde{e}(t) = \mathcal{F}_1[\tilde{q}_{\delta t}](t) - \tilde{q}(t), \text{and } \underline{e}(t) = \underline{q}_{\delta t}(t) - \mathcal{F}_1[q_{opt,\delta t}](t), \ \forall t \in [0,T].$$

By definition of the operator $\mathcal{F}_1[\cdot]$ we have

$$\begin{aligned}
\|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}]\|_{L^2(0,T,U_h)}^2 &= \|\mathcal{B}_h^\star \mathcal{F}_0[q_{opt,\delta t}]\|_{L^2(0,T,U_h)}^2 \\
&\quad + \frac{\delta t^2}{12} \|\mathcal{B}_h^\star \mathcal{F}_0\left[\mathcal{M}_h^{-1} \mathcal{A}_h q_{opt,\delta t}\right]\|_{L^2(0,T,U_h)}^2.
\end{aligned}$$

Using (5.12) to bound the first term, it follows

$$\|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}]\|_{L^2(0,T,U_h)} \leq C(\|y_0\|_h + \|\hat{y}_0\|_h) + C\delta t \Big(\sum_{n=1}^M \delta t \left\|q_{opt,\delta t}^n\right\|_{2,h}^2 \Big)^{\frac{1}{2}}.$$

Then, we use (2.3), (6.26) (notice that this inequality also holds for $\theta = \frac{1}{2}$, see Proposition 6.5) and (5.14) in order to bound the second term as follows

$$\begin{aligned}
\sum_{n=1}^M \delta t \left\|q_{opt,\delta t}^n\right\|_{2,h}^2 &\leq C \left\|q_{opt,\delta t}^F\right\|_{1,h}^2 + C\delta t^3 \left\|q_{opt,\delta t}^F\right\|_{4,h}^2 \\
&\leq C\rho_h \left(1 + \delta t^3 \rho_h^3\right) \left\|q_{opt,\delta t}^F\right\|_h^2 \leq C\rho_h \left\|q_{opt,\delta t}^F\right\|_h^2, \quad (6.46)
\end{aligned}$$

where we used the uniform bound $\delta t \rho_h \leq \delta$. We have thus proved that

$$\|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}]\|_{L^2(0,T,U_h)} \leq C'\left(1+\frac{\sqrt{\delta t}}{\sqrt{\phi(h)}}\right)(\|y_0\|_h + \|\hat{y}_0\|_h).$$

Similarly to (6.46) we have

$$\sum_{n=1}^{M} \delta t \, \|\tilde{q}^n\|_{2,h}^2 \leq C\rho_h \left\|\tilde{q}^F\right\|_h^2,$$

then, from (6.45) we deduce

$$\begin{aligned}\|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t) - v(t)\|_{L^2(0,T,U_h)}^2 + \phi(h)\left\|\tilde{q}^F\right\|_h^2 \\ \leq C(1+\delta t^{\frac{1}{2}}\phi(h)^{-\frac{1}{2}})(\|y_0\|_h + \|\hat{y}_0\|_h)(\|\tilde{e}\|_{L^2(0,T,E_h)} + \left\|E^1\right\|_h) \\ + C\|\underline{e}\|_{L^2(0,T,E_h)}^2 + C'\delta t^2 \rho_h \left\|q^F_{opt,\delta t}\right\|_h \left\|\tilde{q}^F\right\|_h.\end{aligned}$$

Using the Young inequality, estimates (5.14), (6.34) and (6.35) we finally obtain

$$\begin{aligned}\|\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t) - v(t)\|_{L^2(0,T,U_h)}^2 + \phi(h)\left\|\tilde{q}^F\right\|_h^2 \\ \leq C\frac{\delta t^4}{\phi(h)}\left(1+\rho_h^3+\frac{\delta t+\rho_h^2}{\phi(h)}\right)(\|y_0\|_h + \|\hat{y}_0\|_h)^2.\end{aligned}$$

The claim follows by using assumption (2.1) and the fact that $\delta t \leq T$. □

To conclude, we present here a second interpolation operator that also yields a second-order convergence result. For $X = E_h$ or $X = U_h$ and any $x = (x^n)_n \in X^n$ we define

$$\begin{aligned}\widetilde{\mathcal{F}}_1[x](t) = \sum_{n=1}^{M-1} 1_{]t^{n-1},t^n[}(t)\left(x^n + \frac{t-t^{n-\frac{1}{2}}}{\delta t}(x^{n+1}-x^n)\right) \\ + 1_{]t^{M-1},t^M[}(t)\left(x^M + \frac{t-t^{M-\frac{1}{2}}}{\delta t}(x^M - x^{M-1})\right). \quad (6.47)\end{aligned}$$

The interest of this new operator as compared to $\mathcal{F}_1[\cdot]$ is that it does not depend on the operators $\mathcal{M}_h$ and $\mathcal{A}_h$ and commutes with the operator $\mathcal{B}_h^\star$. Notice however that $\widetilde{\mathcal{F}}_1[x]$ is piecewise affine but not continuous on $[0,T]$.

Our result is then the following

THEOREM 6.8. *In the same conditions as in the previous theorem, we have*

$$\|\widetilde{\mathcal{F}}_1[v_{\delta t}] - v\|_{L^2(0,T,U_h)} \leq C\delta t^2 \frac{\rho_h}{\phi(h)^{\frac{1}{2}}}\left(\rho_h^{\frac{1}{2}} + \phi(h)^{-\frac{1}{2}}\right)(\|y_0\|_h + \|\hat{y}_0\|_h),$$

where $v_{\delta t} = \mathcal{B}_h^\star q_{opt,\delta t}$.

As we can see, the actual computation of $\widetilde{\mathcal{F}}_1[v_{\delta t}]$ only requires the knowledge of $(v^n)_n$ and not of the sequence $q_{opt,\delta t}$. In most practical cases, this can lead to saving a significant amount of memory for the storage of the control function.

Proof. With Theorem 6.7 it suffices to estimate the difference $\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}] - \widetilde{\mathcal{F}}_1[\mathcal{B}_h^\star q_{opt,\delta t}]$.

- For $n \in [\![1, M-1]\!]$, and $t^{n-1} < t < t^n$, we have (using the equations (6.24))

$$\begin{aligned}
&\left[\!\left[\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t) - \widetilde{\mathcal{F}}_1[\mathcal{B}_h^\star q_{opt,\delta t}](t)\right]\!\right]_h^2 \\
&\qquad \le |t - t^{n-\frac{1}{2}}|^2 \left[\!\left[\mathcal{B}_h^\star \left(\mathcal{M}_h^{-1}\mathcal{A}_h q^n - \frac{q^{n+1}-q^n}{\delta t}\right)\right]\!\right]_h^2 \\
&\qquad\qquad \le C\delta t^2 \left\|\mathcal{M}_h^{-1}\mathcal{A}_h q^n - \frac{q^{n+1}-q^n}{\delta t}\right\|_h^2 \\
&\qquad\qquad\qquad = \frac{C}{4}\delta t^2 \left\|\mathcal{M}_h^{-1}\mathcal{A}_h (q^{n+1}-q^n)\right\|_h^2 = \frac{C}{4}\delta t^4 \left\|q^{n+\frac{1}{2}}\right\|_{4,h}^2
\end{aligned}$$

  which by (6.25) (which is valid even for $\theta = \frac{1}{2}$, see Proposition 6.5), (2.5) and the bound $\delta t \rho_h \le \delta$, yields

$$\begin{aligned}
&\int_0^{T-\delta t} \left[\!\left[\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t) - \widetilde{\mathcal{F}}_1[\mathcal{B}_h^\star q_{opt,\delta t}](t)\right]\!\right]_h^2 dt \\
&\quad \le C\delta t^4 \left\|q^F_{opt,\delta t}\right\|_{3,h}^2 + C\delta t^5 \left\|q^F_{opt,\delta t}\right\|_{4,h}^2 \le C\delta t^4 \big(\rho_h^3 + \delta t \rho_h^4\big) \left\|q^F_{opt,\delta t}\right\|_h^2 \\
&\quad \le C\delta t^4 \frac{\rho_h^3}{\phi(h)} (\|y_0\|_h + \|\hat{y}_0\|_h)^2.
\end{aligned}$$

- Finally, the case $n = M$ can be bounded in a similar way as

$$\left[\!\left[\mathcal{B}_h^\star \mathcal{F}_1[q_{opt,\delta t}](t) - \widetilde{\mathcal{F}}_1[\mathcal{B}_h^\star q_{opt,\delta t}](t)\right]\!\right]_h^2 \le C\delta t^4 \left\|q^{M-\frac{1}{2}}\right\|_{4,h}^2.$$

□

7. Some numerical results. For all the tests we present below, we have chosen a one-dimensional domain $\Omega =]0,1[$, a distributed control domain $\omega =]0.3, 0.8[$, a final time $T = 1$, an initial data $y_0(x) = \sin(\pi x)^{10}$.

We consider a finite-difference scheme in space on a mesh of Ω with N discretization points for solving the control problem (1.1) with a diffusion coefficient γ.

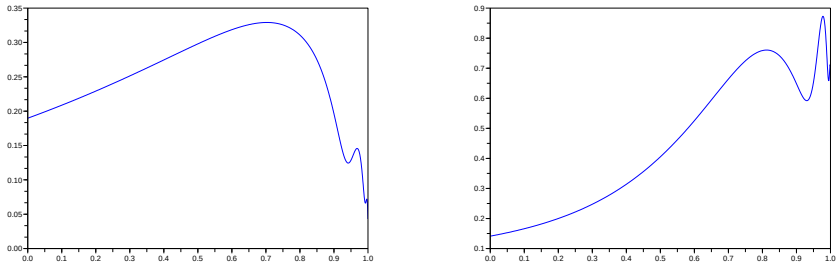
Tests #1 and #2. We choose here a constant diffusion coefficient $\gamma = 0.1$ and a uniform mesh of Ω so that $h = 1/N$. We are interested in the null-controllability problem, that is we choose a target $\hat{y}_F = 0$.

In Test #1, we consider the Implicit Euler time discretisation ($\theta = 1$) and in Test #2, we consider the Crank-Nicolson time discretisation ($\theta = \frac{1}{2}$).

The qualitative behavior of the control function obtained in each case is illustrated in Figure 7.1 where we represent the map $t \in [0,T] \mapsto \|v(t)\|_{L^2(\omega)}$.

We now illustrate the various convergence properties established in this article. In Figures 7.2 and 7.3 we plot the error $\|v - v_{\delta t}\|_{L^2(]0,T[\times\omega)}$ between the semi-discrete and the fully-discrete controls for various values of the time step and three different mesh size for the domain Ω for both Tests #1 and #2. For each test, two situations are presented depending on the choice of the penalization function $h \mapsto \phi(h)$.

As expected, we observe the first order convergence in δt for Test #1 and an asymptotic second order convergence for Test #2. Moreover, these convergences are not uniform with respect to h. In fact, for a given value of the time step, the error increases when N increases. Nevertheless, it appears that the dependences on h in

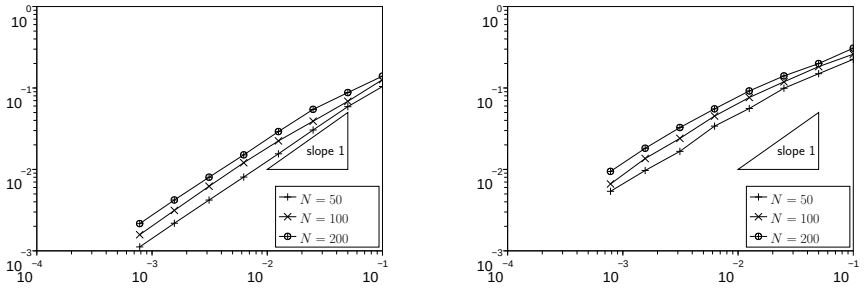


(a) Tests #1 and #2 (b) Test #3

FIG. 7.1. *The function* $t \in [0,T] \mapsto \|v(t)\|_{L^2(\omega)}$ *for Tests #1/#2 and Test #3.*

the error estimates we proved in Theorems 6.6, 6.7, and 6.8 are not optimal in this context.

We observe that, for large values of δt, the Crank-Nicholson scheme only behaves like a first order scheme, but yet produce smaller errors than the implicit Euler method.



(a) $\phi(h) = h^2$ (b) $\phi(h) = h^4$

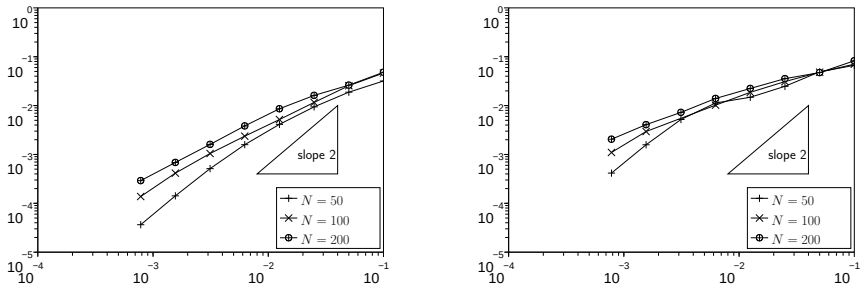
FIG. 7.2. *Error on the control* $\|v - \mathcal{F}_0[v_{\delta t}]\|_{L^2(]0,T[\times\omega)}$ *as a function of* δt *for Test #1.*

Table 7.1 (resp. Table 7.2) shows the size of the final state in the L^2 norm as well as the cost of the control for various values of δt and N and for $\phi(h) = h^2$ (resp. $\phi(h) = h^4$). We observe that, these values barely depend on δt; the cost of the control is uniformly bounded w.r.t. N. The size of the final state actually behaves like $\sqrt{\phi(h)}$, as proved in our results.

Test #3. In that test we consider random meshes of Ω with N points built in such a way that each cell in the mesh has a size $1/N \pm 40\%$. Furthermore, we choose a non constant diffusion coefficient whose formula is given by $\gamma(x) = 0.1 - 0.05 * \tanh((x - 0.5)/0.1)$ and we now consider the control to the trajectories problem by chosing a target $\hat{y}_F = 0.1 \sin(\pi x)$.

We choose here a penalization function $\phi(h) = h^2$ and we consider the Implicit Euler method in time.

We observe results qualitatively very similar to that of the previous tests (see Figure 7.1). In Figure 7.4, we only illustrate the convergence properties with respect



(a) $\phi(h) = h^2$ (b) $\phi(h) = h^4$

FIG. 7.3. *Error on the control* $\|v - \widetilde{\mathcal{F}}_1[v_{\delta t}]\|_{L^2(]0,T[\times\omega)}$ *as a function of* δt *for Test #2.*

	$N = 50$		$N = 100$		$N = 200$	
δt	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$
5.00E-02	1.36E-03	2.61E-01	6.55E-04	2.77E-01	3.44E-04	2.90E-01
1.25E-02	1.16E-03	2.47E-01	5.35E-04	2.61E-01	2.37E-04	2.71E-01
3.13E-03	1.09E-03	2.44E-01	4.83E-04	2.57E-01	2.14E-04	2.65E-01
7.81E-04	1.07E-03	2.43E-01	4.70E-04	2.55E-01	2.05E-04	2.64E-01
Semi-discrete	1.07E-03	2.43E-01	4.66E-04	2.55E-01	2.03E-04	2.63E-01

TABLE 7.1
Summary of the numerical results for Test #1 with $\phi(h) = h^2$.

	$N = 50$		$N = 100$		$N = 200$	
δt	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$
5.00E-02	2.58E-05	3.14E-01	4.70E-06	3.35E-01	1.70E-06	3.47E-01
1.25E-02	1.51E-05	2.81E-01	3.10E-06	2.92E-01	8.00E-07	2.99E-01
3.13E-03	1.14E-05	2.72E-01	2.50E-06	2.80E-01	5.00E-07	2.85E-01
7.81E-04	1.03E-05	2.69E-01	2.20E-06	2.77E-01	4.81E-07	2.81E-01
Semi-discrete	9.95E-06	2.68E-01	2.09E-06	2.76E-01	4.45E-07	2.80E-01

TABLE 7.2
Summary of the numerical results for Test #1 with $\phi(h) = h^4$.

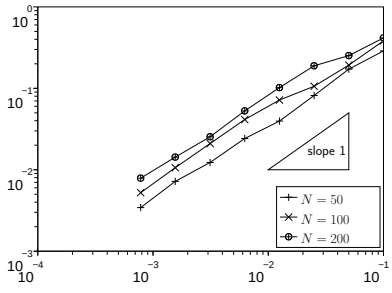
	$N = 50$		$N = 100$		$N = 200$	
δt	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$
5.00E-02	1.07E-03	2.43E-01	4.66E-04	2.51E-01	2.03E-04	2.64E-01
1.25E-02	1.07E-03	2.43E-01	4.66E-04	2.51E-01	2.03E-04	2.64E-01
3.13E-03	1.07E-03	2.43E-01	4.66E-04	2.51E-01	2.03E-04	2.64E-01
7.81E-04	1.07E-03	2.43E-01	4.66E-04	2.51E-01	2.03E-04	2.64E-01
Semi-discrete	1.07E-03	2.43E-01	4.66E-04	2.55E-01	2.03E-04	2.63E-01

TABLE 7.3
Summary of the numerical results for Test #2 with $\phi(h) = h^2$.

	$N = 50$		$N = 100$		$N = 200$	
δt	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$	$\left\|y^M\right\|_h$	$\|v_{\delta t}\|$
5.00E-02	9.85E-06	2.68E-01	2.06E-06	2.76E-01	5.07E-07	2.80E-01
1.25E-02	9.95E-06	2.68E-01	2.09E-06	2.76E-01	4.45E-07	2.80E-01
3.13E-03	9.95E-06	2.68E-01	2.09E-06	2.76E-01	4.45E-07	2.80E-01
7.81E-04	9.95E-06	2.68E-01	2.09E-06	2.76E-01	4.45E-07	2.80E-01
Semi-discrete	9.95E-06	2.68E-01	2.09E-06	2.76E-01	4.45E-07	2.80E-01

TABLE 7.4
Summary of the numerical results for Test #2 with $\phi(h) = h^4$.

to δt and the behavior of the distance of the final state to the target and of the cost of the control with respect to the mesh size (these results are those obtained with the smallest time step considered here, that is δt =7.81E-04).



(a) $\|v - \mathcal{F}_0[v_{\delta t}]\|_{L^2(]0,T[\times\omega)}$ as a function of δt

- $N = 50$
  $\left\|y^M - \hat{y}_F\right\|_h$ =3.44E-03
  $\|v_{\delta t}\|$ =4.50E-01
- $N = 100$
  $\left\|y^M - \hat{y}_F\right\|_h$ =1.38E-03
  $\|v_{\delta t}\|$ =4.69E-01
- $N = 200$
  $\left\|y^M - \hat{y}_F\right\|_h$ =8.79E-04
  $\|v_{\delta t}\|$ =5.14E-01

(b) distance to the target; norm of the control

FIG. 7.4. *Numerical results for Test #3.*

8. Concluding remarks. Most of the results of this article still hold if we assume that $\mathcal{B}_h^\star$ is a (formally) lower-order operator, that is, if we replace Assumption (2.3) by

$$[\![\mathcal{B}_h^\star x]\!]_h \leq C \|x\|_{\alpha,h}, \forall h > 0, \forall x \in E_h,$$

for some $\alpha \in]0, 1]$. Notice however that, for $\frac{1}{2} < \theta < 1$, we need to further assume that $\delta t \rho_h \leq \delta$ for our uniform controllability and observability results to hold.

In practice, this more general assumption should be useful to handle more complex control operators and possibly more complex discretization schemes. Note that boundary control problems do not enter this framework since Lebeau-Robbiano type spectral inequalities of the form (in the continuous case)

$$\sum_{\mu_k \leq \mu} |\alpha_k|^2 = \int_\Omega \left| \sum_{\mu_k \leq \mu} \alpha_k \phi_k(x) \right|^2 dx \leq C e^{C\sqrt{\mu}} \int_\omega \left| \sum_{\mu_k \leq \mu} \alpha_k \phi_k(x) \right|^2 dx.$$

are known to not hold for boundary observations (compare with the boundary observed inequality proven in [LR95]).

With Theorems 5.1 and 5.2 we note that the estimates of the optimal adjoint states q_{opt}^F and $q_{opt,\delta t}^F$ deteriorate as we take $\phi(h)$ to zero. This effect can be observed

numerically with the formation of the boundary layer at time $t = T$ for the adjoint system. It then leads to a poor conditioning of the numerical method. See for instance [MZ09] for numerical evidences. The idea of [FCM], introducing a weighted functional, may lead to a better treatment of this problem. The numerical analysis of this approach, in the spirit of the present work, yet needs to be carried out.

To apply the results we have obtained here, an important question that remains to be answered is the following: for what kind of linear parabolic problems (scalar with smooth or non-smooth coefficients, systems, etc ...) and for what kind of numerical methods and meshes (finite differences, finite elements, etc ...) do discrete Lebeau-Robbiano spectral inequalities ($\mathcal{H}_{\alpha,\beta}$) hold? The authors of the present article tackled this question for finite-difference discretizations of parabolic equations on smooth meshes and for smooth coefficients in [BHL09a] and [BHL09b]. Many other settings need to be studied.

Appendix A. Proof of Lemma 5.4.

We shall first consider the case $\theta > \frac{1}{2}$. Notice that the condition $\lambda \leq A_\theta \delta t^{-1}$ implies $1 - \delta t(1-\theta)\lambda > 0$.

Notice also that it suffices to prove (5.22) for $s = 0$ and $s = 1$, the general case being then deduced by interpolation. We denote by $I = \left[0, A_\theta \delta t^{-1}\right]$ the interval of the admissible values of λ in (5.22).

We set $\alpha(\lambda) = \left(\frac{1+\delta t\theta\lambda}{1-\delta t(1-\theta)\lambda}\right)^M e^{-M\delta t\lambda}$, and $f(\lambda) = 1 - \alpha(\lambda)$.

Case $s = 0$. We prove that $0 \leq f(\lambda) \leq 1$ for any $\lambda \in I$, which in turn implies (5.22). We see that $\alpha \geq 0$ on I. By computing f'

$$\begin{aligned} f'(\lambda) &= T\frac{\delta t(2\theta-1)\lambda - \delta t^2\theta(1-\theta)\lambda^2}{(1+\delta t\theta\lambda)(1-\delta t(1-\theta)\lambda)}\alpha(\lambda) \\ &= T\theta(1-\theta)\lambda\delta t\frac{A_\theta - \lambda\delta t}{(1+\delta t\theta\lambda)(1-\delta t(1-\theta)\lambda)}\alpha(\lambda), \end{aligned}$$

we deduce that f is non-decreasing on I. Hence

$$0 = f(0) \leq f(\lambda) \leq f\left(A_\theta\delta t^{-1}\right) = 1 - \alpha\left(A_\theta\delta t^{-1}\right) \leq 1. \tag{A.1}$$

Case $s = 1$. We set $g(\lambda) = f(\lambda)/\lambda^2$. As seen above, f is non-negative on I, so is g. Moreover, $f(0) = f'(0) = 0$. The function g can thus be extended to $\lambda = 0$ by setting $g(0) = \frac{1}{2}f''(0) = T\delta t(\theta - \frac{1}{2}) > 0$.

Let $\lambda_m \in I$ be such that $g(\lambda_m) = \sup_I g$. Three cases have to be considered.

1. If $\lambda_m = 0$ then $\sup_I g = g(0) = T\delta t(\theta - \frac{1}{2})$.
2. If $\lambda_m = A_\theta\delta t^{-1}$ then by (A.1),

$$\sup_I g = g\left(A_\theta\delta t^{-1}\right) = \frac{f\left(A_\theta\delta t^{-1}\right)}{\frac{(2\theta-1)^2}{\theta^2(1-\theta)^2}}\delta t^2 \leq \frac{\theta^2(1-\theta)^2}{(2\theta-1)^2}\delta t^2.$$

3. If $\lambda_m \in \overset{\circ}{I}$ then we have $g'(\lambda_m) = 0$. Computing g' as follows

$$g'(\lambda) = \frac{f'(\lambda)}{\lambda^2} - \frac{2f(\lambda)}{\lambda^3},$$

we deduce that λ_m is such that $\lambda_m f'(\lambda_m) = 2f(\lambda_m)$, which implies

$$\alpha(\lambda_m) = \frac{1}{1 + \lambda_m\frac{T}{2}\frac{\delta t(2\theta-1)\lambda_m - \delta t^2\theta(1-\theta)\lambda_m^2}{(1+\delta t\theta\lambda_m)(1-\delta t(1-\theta)\lambda_m)}},$$

and then

$$\begin{aligned}\sup_I g &= g(\lambda_m)\\ &= T\left(\theta-\frac{1}{2}\right)\delta t\frac{1-\frac{\delta t\lambda_m}{A_\theta}}{1+\delta t\lambda_m(1+\frac{T}{2}\lambda_m)\big[(2\theta-1)-\lambda_m\delta t\theta(1-\theta)\big]}\\ &\le T\left(\theta-\frac{1}{2}\right)\delta t.\end{aligned}$$

In each case estimate (5.22) follows.

We now consider the case $\theta=\frac{1}{2}$. The interval of admissible values of λ is now $I=[0,A_{\frac{1}{2}}\delta t^{-\frac{2}{3}}]$ and we also set $\alpha(\lambda)=\left(\frac{1+\delta t\lambda/2}{1-\delta t\lambda/2}\right)^M e^{-M\delta t\lambda}$ and $f(\lambda)=1-\alpha(\lambda)$.

Case s = 0. We observe that $f(0)=0$ and that

$$f'(\lambda)=-T\alpha(\lambda)\frac{\delta t^2\lambda^2}{4-\delta t^2\lambda^2}\le 0,\quad \forall\lambda\in I.$$

Hence we have $0\le|f(\lambda)|\le|f(A_{\frac{1}{2}}\delta t^{-\frac{2}{3}})|$. Furthermore we have

$$\left|f(A_{\frac{1}{2}}\delta t^{-\frac{2}{3}})\right|\le 1+\alpha(A_{\frac{1}{2}}\delta t^{-\frac{2}{3}})\le 1+3^{TA_{\frac{1}{2}}^3}.$$

The last inequality is obtained by using the two following facts

- The condition $\lambda\delta t^{\frac{2}{3}}\le A_{\frac{1}{2}}$ with the chosen value of $A_{\frac{1}{2}}$ implies that

$$\lambda\delta t\le A_{\frac{1}{2}}\delta t^{\frac{1}{3}}\le A_{\frac{1}{2}}T^{\frac{1}{3}}\le 1. \tag{A.2}$$

- The map $x\in[0,1]\mapsto x^{-3}(\log\left(\frac{1+x/2}{1-x/2}\right)-x)$ is positive, increasing on $[0,1]$, thus bounded by $\log(3)$, for instance, on this interval.

We conclude that

$$\alpha(\lambda)\le e^{\log(3)M(\delta t\lambda)^3}=3^{T\delta t^2\lambda^3}\le 3^{TA_{\frac{1}{2}}^3},$$

and the bound on f is proven.

Case s = 1. Here we set $g(\lambda)=|f(\lambda)|/\lambda^3$. Notice that g is non-negative and can be extended to $\lambda=0$ by letting $g(0)=T\delta t^2/12$.

We want to estimate $\sup_I g=g(\lambda_m)$, $\lambda_m\in I$. Here also three cases have to be considered:

1. If $\lambda_m=0$, then $\sup_I g=g(0)=T\delta t^2/12$ and the claim is proven.
2. If $\lambda_m=A_{\frac{1}{2}}\delta t^{-2/3}$, using the previously bound proven on α, we obtain

$$\sup_I g=g(A_{\frac{1}{2}}\delta t^{-2/3})\le CT\delta t^2\alpha(A_{\frac{1}{2}}\delta t^{-2/3})\le CT\delta t^2 e^{TA_{\frac{1}{2}}^3}.$$

3. If $\lambda_m\in\overset{\circ}{I}$ then $g'(\lambda_m)=0$ which is equivalent to $\lambda_m f'(\lambda_m)=3f(\lambda_m)$. Similar computations to those above yield

$$\alpha(\lambda_m)=\frac{1-\delta t^2\lambda_m^2/4}{1-\delta t^2\lambda_m^2/4-T\delta t^2\lambda_m^3/12},$$

and then

$$\sup_I g = g(\lambda_m) = \frac{T}{12}\delta t^2 \frac{1}{1-\delta t^2\lambda_m^2/4 - T\delta t^2\lambda_m^3/12} \leq \frac{T}{8}\delta t^2,$$

with (A.2) and since by (5.16) we have $TA^3_{\frac{1}{2}} = 1$.
This concludes the proof of Lemma 5.4.

REFERENCES

[BHL09a] F. Boyer, F. Hubert, and J. Le Rousseau, *Discrete Carleman estimates and uniform controllability of semi-discrete parabolic equations*, J. Math. Pures Appl., to appear, `http://hal.archives-ouvertes.fr/hal-00366496/fr` (2009).

[BHL09b] _____, *Discrete Carleman estimates for elliptic operators in arbitrary dimension and applications*, preprint (2009).

[EV09] S. Ervedoza and J. Valein, *On the observability of abstract time-discrete linear parabolic equations*, Rev. Mat. Complut., to appear (2009).

[FCM] E. Fernández-Cara and A. Münch, *Numerical exact controllability of the 1D heat equation: primal algorithms*, in prep.

[FI96] A. Fursikov and O. Yu. Imanuvilov, *Controllability of evolution equations*, vol. 34, Seoul National University, Korea, 1996, Lecture notes.

[GL94] R. Glowinski and J.-L. Lions, *Exact and approximate controllability for distributed parameter systems*, Acta Numer. (1994), 269–378.

[JL99] D. Jerison and G. Lebeau, *Harmonic analysis and partial differential equations (Chicago, IL, 1996)*, Chicago Lectures in Mathematics, ch. Nodal sets of sums of eigenfunctions, pp. 223–239, The University of Chicago Press, Chicago, 1999.

[LL09] J. Le Rousseau and G. Lebeau, *On Carleman estimates for elliptic and parabolic operators. Applications to unique continuation and control of parabolic equations*, Preprint (2009).

[LR95] G. Lebeau and L. Robbiano, *Contrôle exact de l'équation de la chaleur*, Comm. Partial Differential Equations **20** (1995), 335–356.

[LT06] S. Labbé and E. Trélat, *Uniform controllability of semidiscrete approximations of parabolic control systems*, Systems Control Lett. **55** (2006), 597–609.

[LZ98a] G. Lebeau and E. Zuazua, *Null-controllability of a system of linear thermoelasticity*, Arch. Rational Mech. Anal. **141** (1998), 297–329.

[LZ98b] A. Lopez and E. Zuazua, *Some new results to the null controllability of the 1-d heat equation*, Séminaire sur les Équations aux Dérivées Partielles, 1997–1998, Exp. No. VIII, 22 pp., École Polytech., Palaiseau (1998).

[MZ09] A. Münch and E. Zuazua, *Numerical approximation of trajectory controls for the heat equation through transmutation*, preprint (2009), 34 pages.

[Zhe08] C. Zheng, *Controllability of the time discrete heat equation*, Asymptotic Anal. **59** (2008), 139–177.

[Zua06] E. Zuazua, *Control and numerical approximation of the wave and heat equations*, International Congress of Mathematicians, Madrid, Spain **III** (2006), 1389–1417.