ESAIM: COCV 29 (2023) 41
https://doi.org/10.1051/cocv/2023028

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

SPACETIME FINITE ELEMENT METHODS FOR CONTROL PROBLEMS SUBJECT TO THE WAVE EQUATION

Erik Burman[1,*], Ali Feizmohammadi[2], Arnaud Münch[3] and Lauri Oksanen[4]

Abstract. We consider the null controllability problem for the wave equation, and analyse a stabilized finite element method formulated on a global, unstructured spacetime mesh. We prove error estimates for the approximate control given by the computational method. The proofs are based on the regularity properties of the control given by the Hilbert Uniqueness Method, together with the stability properties of the numerical scheme. Numerical experiments illustrate the results.

Mathematics Subject Classification. 35L05, 93-08, 65M12, 65M60.

Received October 18, 2021. Accepted April 11, 2023.

1. Introduction

We consider the classical null controllability problem for the wave equation, both with distributed and boundary control. Let $T > 0$, and let $\Omega \subset \mathbb{R}^n$, with $n \geq 2$, be a connected bounded open set with smooth boundary. We write

$$M = (0,T) \times \Omega, \quad \Gamma = (0,T) \times \partial\Omega.$$

For a fixed initial state $(u_0, u_1) \in H_0^1(\Omega) \times L^2(\Omega)$, the distributed null control problem on M reads: find $\phi \in L^2(M)$ such that the solution u of

$$\begin{cases} \Box u = \chi\phi, \\ u|_\Gamma = 0, \\ u|_{t=0} = u_0, \ \partial_t u|_{t=0} = u_1, \end{cases} \tag{1.1}$$

Keywords and phrases: Wave equation, control, finite element method, space time, stabilisation.

[1] Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK.
[2] Department of Mathematics, University of Toronto, 3359 Mississauga Road Deerfield Hall, Mississauga, Ontario, Canada L5L 1C6.
[3] Laboratoire de Mathématiques Blaise Pascal, Université Clermont Auvergne, CNRS, LMBP, F-63000 Clermont-Ferrand, France.
[4] Department of Mathematics and Statistics, University of Helsinki, P.O 68, 00014 University of Helsinki, Finland.

[*] Corresponding author: e.burman@ucl.ac.uk

satisfies

$$u|_{t=T} = 0, \quad \partial_t u|_{t=T} = 0. \tag{1.2}$$

Here $\Box = \partial_t^2 - \Delta$ is the wave operator, and χ is a cutoff function that localizes the control in a subset of M. More precisely, we consider a cutoff of the form

$$\chi(t,x) = \chi_0(t)\chi_1^2(x)$$

where $\chi_0 \in C_0^\infty([0,T])$ and $\chi_1 \in C^\infty(\Omega)$ take values in $[0,1]$.

Our main assumption is that

(A) $\chi = 1$ on $(a,b) \times \omega \subset M$ satisfying the geometric control condition.

Here $0 < a < b < T$ and $\omega \subset \Omega$ is open. The geometric control condition means that every compressed generalized bicharacteristic intersects the set $(a,b) \times \omega$, when projected to M. We refer to [4] for the rather technical definition of a compressed generalized bicharacteristic. Roughly speaking, all continuous paths on M, consisting of lightlike line segments in its interior and reflected on Γ according to Snell's law, must intersect $(a,b) \times \omega$. However, projections of compressed generalized bicharacteristics may also glide along Γ under suitable convexity.

For a simple example, suppose ω is a neighbourhood of the boundary $\partial\Omega$, then the geometric control condition holds when $b - a$ is larger than the diameter of Ω. Observe that also T needs to be larger than the diameter.

For our main result we assume, furthermore, that $(u_0, u_1) \in H^{k+1}(\Omega) \times H^k(\Omega)$ for some $k = 2, 3, \dots,$ and that the following compatibility conditions of order k are satisfied

$$(\Delta^j u_0)|_{\partial\Omega} = 0 \quad \text{for} \quad j = 0, 1, \dots, \left\lfloor \frac{k}{2} \right\rfloor, \tag{C1}$$

$$(\Delta^j u_1)|_{\partial\Omega} = 0 \quad \text{for} \quad j = 0, 1, \dots, \left\lfloor \frac{k-1}{2} \right\rfloor. \tag{C2}$$

Here $\lfloor \cdot \rfloor$ is the floor function that gives the greatest integer less than or equal to its argument. We recall that the compatibility conditions guarantee that, for smooth enough ϕ, the solution u of (1.1) is in $H^{k+1}(M)$, see *e.g.* Theorem 6, p. 412 of [22].

Under these assumptions, we show that the stabilized finite element method introduced below gives such an approximation ϕ_h of a certain minimum norm solution ϕ to the control problem that

$$\|\chi(\phi_h - \phi)\|_{L^2(M)} \lesssim h^q, \tag{1.3}$$

where $h > 0$ is the mesh size and $q \leq k - 1$ is the polynomial order of the finite element space. The implicit constant in the above inequality is independent of h and the functions ϕ_h and ϕ. This notation is used in the paper when confusion is not likely to arise. See Theorem 2.4 for the precise formulation. In this result both u and ϕ are assumed to be at least $H^2(M)$-smooth, corresponding to the above constraint $k \geq 2$. Moreover, we show that the approximation ϕ_h can be used to compute an approximation $\tilde{u}_h$ of u solving (1.1) and (1.2) such that

$$|||u - \tilde{u}_h|||_E \lesssim h^q,$$

where the norm is the natural energy norm

$$|||u|||_E = \max_{0 \leq t \leq T} (\|u(t)\|_{H^1(\Omega)} + \|\partial_t u(t)\|_{L^2(\Omega)}), \tag{1.4}$$

see Corollary 2.7. We prove also a weak convergence result for our method assuming only the smoothness $u_1 \in L^2(\Omega)$ in the case that $u_0 = 0$, see Theorem 3.1. The case with general rough data is left for future work.

Let us now sketch our result in the case of the boundary null control problem of the following form: given an initial state $(u_0, u_1) \in H_0^1(\Omega) \times L^2(\Omega)$, find $\psi \in L^2(\Gamma)$ such that the solution u of

$$\begin{cases} \Box u = 0, \\ u|_\Gamma = \chi\psi, \\ u|_{t=0} = u_0, \ \partial_t u|_{t=0} = u_1, \end{cases} \tag{1.5}$$

satisfies the final time condition (1.2). Under a geometric control condition analogous to (A), and the above regularity assumptions on the data (u_0, u_1), we introduce a finite element method that converges as

$$\|\chi(\psi_h - \psi)\|_{L^2(\Gamma)} \lesssim h^{q-\frac{1}{2}}, \tag{1.6}$$

with notations analogous to (1.3) and $q \leq k+1$. See Theorem 4.4 for the precise formulation. Observe that, although there is a loss of order $1/2$ in (1.6) in comparison to (1.3), if the highest possible polynomial orders are used, the order in (1.6) becomes $k + 1/2$ versus $k - 1$ in (1.3). We can also rescale the method in the distributed case to get the order $q - 1/2$ for $q \leq k$, leading to $k - 1/2$ for the highest possible order.

1.1. Literature

This work is a contribution to the finite dimensional approximation of null controls for the linear wave equation. The seminal work is due to Glowinski and Lions in [25] where the search of the control of minimal L^2 norm is reduced (using the Fenchel–Rockafellar duality theory) to the unconstrained minimization of the corresponding conjugate functional involving the homogeneous adjoint problem. Minimization of the discrete functional, associated with centered finite difference approximation in time and P^1 finite element method in space is discussed at length in [25] and exhibits a lack of convergence of the approximation with respect to the discretization parameter h. This is due to spurious high frequencies discrete modes which are not exactly controllable uniformly in h.

This pathology can easily be avoided in practice by adding to the conjugate functional a regularized Tikhonov parameter, but this leads to so called approximate controls, solving the control problem only up to a small remainder term. Several cures aiming to filter out the high frequencies have been proposed and analyzed, mainly for simple geometries (1d interval, unit square in 2d, etc) with finite differences schemes. The simplest, but artificial, approach is to eliminate the highest eigenmodes of a discrete approximation of the initial condition as discussed in one space dimension in [39], and extended in [37]. We mention spectral methods initially developed [5] then used in [35]. We also mention so called bi-grid method (based on the projection of the discrete gradient on a coarse grid) proposed in [25] and analyzed in [30, 38] leading to convergence results. One may also design more elaborated discrete schemes avoiding spurious modes: we mention [24] based on a mixed reformulation of the wave equation analyzed later with finite difference schemes in [3, 11, 12] at the semi-discrete level and then extended in [43] to a full space-time discrete setting, leading to strong convergent results.

The above previous works, notably reviewed in [21, 50], fall within an approach that can be called "discretize then control" as they aim to control exactly to zero a finite dimensional approximation of the wave equation. A relaxed controllability approach is analyzed in [10] using a stabilized finite element method in space and leading for smooth two and three dimensional geometries to a strong convergent approximation. To our knowledge, [10] is the only previous work that fully proves a convergence rate for a numerical method solving the null controllability problem assuming only the geometric control condition. There linear rate is obtained when piecewise affine finite elements are used in spatial discretization and simple forward and backward finite differences are used in temporal discretization. The main novelty of the present paper is that we use higher order finite elements in spacetime and show that this leads to convergence rates of higher order.

Let us also mention [45] based on the Russel principle, extended in [2, 14, 26] for least-squares based method. On the other hand, one may also employ a "control then discretize" procedure, where the optimality system (for instance associated with the control of minimal L^2 norm) mixing the boundary condition in time and space and involving the primal and adjoint state is discretized within *a priori* a space-time approximation. The well-posedness of such system is achieved by using so called global or generalized observability inequalities. Such approach avoids the numerical pathologies mentioned above and is notably well-suited for mesh adaptivity. On the other hand, the numerical analysis, within a conformal approximation is delicate since it requires to prove inf-sup stablity that is uniform with respect to h. We mention [15] where this approach has been introduced within a conformal approximation leading to convergent numerical results for the control of minimal L^2 norm. It has been extended in [42] where the wave equation is reformulated as a first order system, solved in the one dimensional case with a stabilized formulation allowing to bypass the inf-sup property issue. We also mention [13] in the 1d case where the optimality system associated to cost involving both the control and the state is reformulated as a space-time elliptic problem of order four, leading to strong convergent result with respect to the discretization parameter. The present paper falls into this category and aims, in the spirit of [8] devoted to the dual data assimilation problem, to provide some convergent results, including rate of convergence, with respect to the discrete parameter. We mention a growing interest for space-time (finite element) methods of approximation for the wave equation, initially advocated in [23, 29, 31] and more recently in [1, 16, 17, 32, 47].

1.2. Outline of the paper

In Section 2 we introduce a weak formulation for the null controllability problem in the distributed control case, furthermore, we discretize and regularize the weak formulation and show that the control function obtained by solving the resulting finite dimensional linear system converges to a solution of the null controllability problem. The main theorem of Section 2, Theorem 2.4, gives a convergence rate for this numerical method.

In Section 3 we show that the method converges weakly even when u_1 in (1.1) has limited regularity in the sense that it is only in $L^2(\Omega)$. The result is incomplete, though, as we make the assumption $u_0 = 0$, see Theorem 3.1, the main theorem of the section. We leave the case of a non-vanishing $u_0 \in H_0^1(\Omega)$ as a topic of future research.

In Section 4, the case of boundary control is discussed. We focus on explaining the differences between this and the distributed control case. The main theorem, Theorem 4.4, is analogous to Theorem 2.4 in Section 2.

In Section 5 we present numerical experiments illustrating the theory developed in Sections 2 and 4, and brief conclusions are given in Section 6.

Appendix A gathers results on the continuum theory. We also rectify some inaccuracies in a related previous work [9] by three of the authors of the present paper, see also Remark 2.8.

Finally, Appendix B gives two proofs concerning the particular way that we deal with the domain Ω having smooth, non-polygonal boundary.

1.3. Notations

We write $(t, x) = (x^0, x^1, \dots, x^n)$ for the coordinates on $\mathbb{R}^{1+n}$. Let g stand for the Minkowski metric on $\mathbb{R}^{1+n}$, and denote by $g(\cdot,\cdot)$ the scalar product with respect to g. The wave operator can be written as $\Box = -\operatorname{div}\operatorname{grad} u$, where the divergence and gradient are defined with respect to g. Let $K \subset M$ be an open set with piecewise smooth boundary, and let $N = (N_0, \dots, N_n)$ be the outward pointing unit normal vector field on ∂K, defined with respect to the Euclidean metric on $\mathbb{R}^{1+n}$. We write

$$\partial_\nu u = N \cdot \operatorname{grad} v = -N_0 \partial_{x^0} v + N_1 \partial_{x^1} v + \dots + N_n \partial_{x^n} v.$$

Note that div coincides with the Euclidean divergence, and we can apply the Euclidean divergence theorem to obtain

$$\int_K u\Box v \ \mathrm{d}x = \int_K g(du, dv) \ \mathrm{d}x - \int_{\partial K} u\partial_\nu v \ \mathrm{d}s, \tag{1.7}$$

where $\mathrm{d}s$ is the Euclidean surface measure on ∂K, and du is the spacetime differential of u, that is, the covector with the components $\partial_{x^j} u$, $j = 0, \dots, n$.

2. Distributed control

2.1. Weak formulation of the control problem

We recall that, assuming (A), the distributed control problem can be solved by finding u and ϕ such that

$$\begin{cases} \Box u = \chi\phi, \\ u|_{x\in\partial\Omega} = 0, \\ u|_{t=0} = u_0, \ \partial_t u|_{t=0} = u_1, \\ u|_{t=T} = 0, \ \partial_t u|_{t=T} = 0, \end{cases} \qquad \begin{cases} \Box\phi = 0, \\ \phi|_{x\in\partial\Omega} = 0. \end{cases} \tag{2.1}$$

Moreover, if $(u_0, u_1) \in H^{k+1}(\Omega) \times H^k(\Omega)$ satisfies the compatibility conditions of order k, then the unique solution (u, ϕ) to (2.1) satisfies

$$\phi|_{t=T} \in H^k(\Omega), \quad \partial_t\phi|_{t=T} \in H^{k-1}(\Omega), \tag{2.2}$$

and this initial data for ϕ satisfies the compatibility conditions of order $k-1$, see Theorem 5.1 of [20]. It follows that $\phi \in H^k(M)$, and this again implies that $u \in H^{k+1}(M)$. The convergence proof for our finite element method is based on the fact that the solution of (2.1) has this regularity.

The control given by ϕ can be characterized also as the control with the minimum norm on M with respect to the weighted measure $\chi \mathrm{d}t\mathrm{d}x$. The fact that (2.1) has a unique solution follows from this characterization. Uniqueness of the solution can also be proven directly, and we have included a short proof in Appendix A, see Lemma A.8 there.

For purposes of the numerical analysis below, it turns out to be convenient to rescale (2.1) in a "semiclassical" manner, that is, we multiply the equations in (2.1) by suitable powers of a small parameter $h > 0$, so that each derivative is multiplied by h. Then (2.1) can be formulated weakly as

$$a(u, \psi) = h^2 c(\phi, \psi) + L(\psi), \quad a(v, \phi) = 0, \tag{2.3}$$

for all $v, \psi \in C^\infty(M)$ vanishing on Γ, where

$$\begin{aligned} a(u, \psi) &= \int_M g(hdu, hd\psi) \ \mathrm{d}x - h(u, h\partial_\nu\psi)_{L^2(\partial M\setminus\Gamma)} \\ L(\psi) &= h(hu_1, \psi|_{t=0})_{L^2(\Omega)} - h(u_0, h\partial_t\psi|_{t=0})_{L^2(\Omega)}, \end{aligned} \tag{2.4}$$

and $c(\phi, \psi) = (\chi\phi, \psi)_{L^2(M)}$. Indeed, it follows from (1.7) that if smooth (u, ϕ) solves (2.1) then (2.3) holds for all smooth (v, ψ) vanishing on Γ.

Let us remark that the converse holds as well. That is, if (2.3) holds for a smooth pair of functions (u, ϕ) vanishing on Γ and for all smooth (v, ψ) vanishing on Γ, then (u, ϕ) solves (2.1). Indeed, to see that $\Box\phi = 0$ it suffices to let v vary in $C_0^\infty(M)$ and integrate by parts. Analogously $\Box u = \chi\phi$. To see that u satisfies the initial and final conditions, we let $\psi \in C^\infty(M)$ satisfy $\psi|_\Gamma = 0$ and integrate by parts

$$0 = a(u, \psi) - h^2 c(\phi, \psi) - L(\psi)$$

$$= h(\psi, h\partial_\nu u)_{L^2(\partial M \setminus \Gamma)} - h(u, h\partial_\nu \psi)_{L^2(\partial M \setminus \Gamma)} - L(\psi).$$

This reduces to

$$\begin{aligned} 0 = (u_0 - u|_{t=0}, \partial_t \psi|_{t=0})_{L^2(\Omega)} + (\partial_t u|_{t=0} - u_1, \psi|_{t=0})_{L^2(\Omega)} \\ - (\partial_t u|_{t=T}, \psi|_{t=T})_{L^2(\Omega)} + (u|_{t=T}, \partial_t \psi|_{t=T})_{L^2(\Omega)}. \end{aligned}$$

Taking $\psi(t,x) = \psi_0(t)\psi_1(x)$, with $\psi_0 \in C^\infty(\mathbb{R})$ satisfying

$$\psi_0(t) = \begin{cases} t & \text{for } t \text{ near } 0, \\ 0 & \text{for } t \text{ near } T, \end{cases}$$

and varying ψ_1 in $C_0^\infty(\Omega)$, we see that $u|_{t=0} = u_0$. The rest of the initial and final conditions follow similarly.

2.2. Discretization

We will formulate a discretized and regularized version of the weak formulation (2.3). This discretized version is given in the form of a finite dimensional linear system, see (2.13) below, and in Proposition 2.1 we will show that the system has a unique solution. Our computational method to solve the distributed control problem boils down to solving (2.13).

A special feature of the regularization used here is that it vanishes when applied to a smooth solution to (2.1). This leads to Galerkin orthogonality, see (2.16), and ultimately allows us to prove a good convergence rate for the method. Similar regularizations have been used for numerical analysis of unique continuation type problems, the closest previous result being that in [8].

Consider a family $\mathcal{T} = \{\mathcal{T}_h : h > 0\}$ where $\mathcal{T}_h$ is a set of $1+n$-dimensional simplices forming a simplicial complex. To keep the discussion as simple as possible, we assume in this section that

$$\bigcup_{K \in \mathcal{T}_h} = M \tag{2.5}$$

for all $h > 0$. This is a restrictive assumption since we also assumed that the spatial boundary $\partial\Omega$ is smooth. We will explain later, see Remark 4.6, how this issue can be avoided by allowing the simplices adjacent to the boundary to have curved faces, fitting Ω. This fitting technique is also described in detail in the context of the boundary control problem below.

If the set ω in assumption (A) is a neighbourhood of the boundary $\partial\Omega$ then the distributed observability estimate in Theorem A.4 holds in the case of piecewise smooth $\partial\Omega$ and large enough $T > 0$. In particular, we can consider polyhedral Ω and then $\bigcup_{K \in \mathcal{T}_h} = M$ is straightforward to arrange. The multiplier method can also be used to derive the distributed observability estimate for polyhedral Ω and more general observation regions ω, however, this method can not reproduce the sharp geometric control condition in the case of smooth boundary [41].

We assume that the family $\mathcal{T}$ is quasi uniform, see *e.g.* Definition 1.140 of [18], and indexed by

$$h = \max_{K \in \mathcal{T}_h} \operatorname{diam}(K). \tag{2.6}$$

Then we define for $p \in \mathbb{N}^+ = \{1, 2, \dots\}$ the $H^1(M)$-conformal approximation space of polynomial degree p,

$$V_h^p = \{u \in H^1(M) : u|_\Gamma = 0,\, u|_K \in \mathbb{P}_p(K) \text{ for all } K \in \mathcal{T}_h\}, \tag{2.7}$$

where $\mathbb{P}_p(K)$ denotes the set of polynomials of degree less than or equal to p on K. Occasionally we write also $V_h = \bigcup_{p\in\mathbb{N}^+} V_h^p$.

Recall that the bilinear form a, defined by (2.4), uses the "semiclassical" scaling. We choose h in (2.4) as the mesh parameter (2.6). Then there is $C > 0$ such that for all $h > 0$ and $u, v \in H^2(M) + V_h$ there holds

$$a(u,v) \le C\|u\|_{H^2(\mathcal{T}_h)}\|v\|_{H^2(\mathcal{T}_h)}, \tag{2.8}$$

where the broken semiclassical Sobolev norm is defined for any $k \in \mathbb{N}$ by

$$\|u\|_{H^k(\mathcal{T}_h)}^2 = \sum_{j=0}^{k}\sum_{K\in\mathcal{T}_h}\|(hD)^j u\|_{L^2(K)}^2.$$

Here $D^j u$ is the tensor of order j that gives the jth total derivative of u. The continuity (2.8) is consequence of the following trace inequality, see *e.g.* equation 10.3.9 of [6]: there is $C > 0$ such that for all $h > 0$, $K \in \mathcal{T}_h$ and $u \in H^1(K)$ there holds

$$h^{\frac12}\|u\|_{L^2(\partial K)} \le C(\|u\|_{L^2(K)} + \|hdu\|_{L^2(K)}). \tag{2.9}$$

For $u \in H^k(M)$ the broken semiclassical norm $\|u\|_{H^k(\mathcal{T}_h)}$ reduces to the usual semiclassical norm defined by

$$\|u\|_{H_h^k(M)}^2 = \sum_{j=0}^{k}\|(hD)^j u\|_{L^2(M)}^2.$$

Moreover, there is $C > 0$ such that for all $h > 0$ and $u \in V_h$ there holds

$$\|u\|_{H^k(\mathcal{T}_h)} \le C\|u\|_{L^2(M)}. \tag{2.10}$$

This is due to the discrete inverse inequality, see *e.g.* Lemma 1.138 of [18]: there is $C > 0$ such that for all $h > 0$, $K \in \mathcal{T}_h$, $p \in \mathbb{N}^+$ and $u \in \mathbb{P}_p(K)$ there holds

$$\|hdu\|_{L^2(K)} \le C\|u\|_{L^2(K)}. \tag{2.11}$$

We will systematically use a scaling so that all the bilinear forms in the paper satisfy the bound (2.8).

We write $\mathcal{F}_h$ for the internal faces of $\mathcal{T}_h$ and

$$[\![\partial_\nu u]\!] = N_1 \cdot \operatorname{grad} u|_{K_1} + N_2 \cdot \operatorname{grad} u|_{K_2}$$

for the jump of $\partial_\nu u$ on $F \in \mathcal{F}_h$. Here $K_1, K_2 \in \mathcal{T}_h$ are the two simplices such that $K_1 \cap K_2 = F$, and N_j is the outward pointing unit normal vector field on ∂K_j, $j = 1, 2$.

Our finite element method has the form: find the critical point of the Lagrangian

$$\mathcal{L}(u,\phi) : V_h^p \times V_h^q \to \mathbb{R}, \quad \mathcal{L}(u,\phi) = \frac12 h^2 c(\phi,\phi) + L(\phi) - \frac12\mathcal{R}(u,\phi) - a(u,\phi),$$

where, writing $U = (u, \partial_t u)$ and $U_0 = (u_0, u_1)$, the regularization is given by

$$\begin{aligned}\mathcal{R}(u,\phi) = {} & h^{-\kappa}S(u) - h^{\kappa}S(\phi) + h^{-\kappa}E(U|_{t=0} - U_0) + h^{-\kappa}E(U|_{t=T}) \\ & + h^{4-\kappa}\widetilde{C}(\phi) + 2h^{2-\kappa}\rho(u,\phi),\end{aligned} \tag{2.12}$$

$$\begin{aligned}
E(U_0) &= h\|u_0\|^2_{L^2(\Omega)} + h\|hu_1\|^2_{L^2(\Omega)},\\
S(u) &= \sum_{K\in\mathcal{T}_h} \|h^2\Box u\|^2_{L^2(K)} + \sum_{F\in\mathcal{F}_h} h\|[\![h\partial_\nu u]\!]\|^2_{L^2(F)},\\
\widetilde{C}(\phi) &= \|\chi\phi\|^2_{L^2(M)}, \quad \rho(u,\phi) = -\sum_{K\in\mathcal{T}_h} (h^2\Box u, \chi\phi)_{L^2(K)},
\end{aligned}$$

where $\kappa < 2$ is a fixed constant. It can be used to tune the convergence order of the method, see (2.15) and Remark 2.6 below. When studying the control problem with limited regularity in Section 3, we can prove convergence only in the case $\kappa = 0$.

We have $\mathcal{R}(u,\phi) = 0$ for a smooth solution (u,ϕ) to (2.1). Indeed,

$$S(u) + 2h^2\rho(u,\phi) + h^4\widetilde{C}(\phi) = \sum_{K\in\mathcal{T}_h} \|h^2(\Box u - \chi\phi)\|^2_{L^2(K)} + \sum_{F\in\mathcal{F}_h} h\|[\![h\partial_\nu u]\!]\|^2_{L^2(F)} = 0,$$

and also $S(\phi) = 0$ and $E(U|_{t=0} - U_0) = E(U|_{t=T}) = 0$.

The equation $d\mathcal{L}(u,\phi) = 0$ can be written as

$$A[(u,\phi),(v,\psi)] = h^{-\kappa}e(U_0, V|_{t=0}) + L(\psi) \quad \text{for all } (v,\psi) \in V_h^p \times V_h^q, \tag{2.13}$$

where the bilinear form A is given by

$$\begin{aligned}
A[(u,\phi),(v,\psi)] = h^{-\kappa}s(u,v) - h^{\kappa}s(\phi,\psi) - h^2c(\phi,\psi) + h^{-\kappa}\sum_{\tau=0,T} e(U|_{t=\tau}, V|_{t=\tau})\\
+ a(v,\phi) + a(u,\psi)\\
+ h^{4-\kappa}\tilde{c}(\phi,\psi) + h^{2-\kappa}\rho(v,\phi) + h^{2-\kappa}\rho(u,\psi).
\end{aligned}$$

Here s is the bilinear form associated to the quadratic form S, and this lowercase–uppercase convention is systematically used also for other quadratic and bilinear forms in the paper. Let us emphasize that all the bilinear forms s, c, e, $\tilde{c}$ and ρ satisfy the same bound (2.8) as a.

System (2.13) can be viewed as defining a computational method:

Proposition 2.1. *The linear system* (2.13) *has a unique solution for small enough $h > 0$.*

In order to prove Proposition 2.1, we define the residual norm by

$$|||(u,\phi)|||^2 = h^{-\kappa}S(u) + h^{\kappa}S(\phi) + h^2C(\phi) + h^{-\kappa}\sum_{\tau=0,T} E(U|_{t=\tau}). \tag{2.14}$$

Lemma 2.2. *Suppose that (A) holds. Then $|||\cdot|||$ is a norm on $V_h \times V_h$.*

Proof. Suppose $|||(u,\phi)||| = 0$. Then $\Box u = 0$ elementwise and $[\![\partial_\nu u]\!] = 0$ for all internal faces. It follows that $\Box u = 0$ in the weak sense. As $u|_\Gamma = 0$ and $E(U|_{t=0}) = 0$, it follows that $u = 0$. Similarly $\Box\phi = 0$ in the weak sense. As $\phi|_\Gamma = 0$ and $\widetilde{C}(\phi) = 0$, the distributed observability estimate, see Theorem A.4, implies that $\phi = 0$. $\Box$

Lemma 2.3. *For all sufficiently small h and all $u,\phi \in H^2(M) + V_h$ there holds*

$$|||(u,\phi)|||^2 \lesssim A[(u,\phi),(u,-\phi)].$$

Proof. By the definition of A, we have

$$A[(u,\phi),(u,-\phi)] = |||(u,\phi)|||^2 - h^{4-\kappa}\widetilde{C}(\phi).$$

As $\kappa < 2$ and $\chi \le 1$, $h^{4-\kappa}\widetilde{C}(\phi)$ can be absorbed by $h^2 C(\phi)$ for small $h > 0$. □

Proof of Proposition 2.1. Observe that (2.13) is a square system of linear equations and that the above two lemmas imply that $(u,\phi) = 0$ is the only solution when the right-hand side is zero. (The right-hand side being zero is equivalent with $U_0 = 0$.) □

2.3. Error estimates

We will now prove the main theorem concerning the distributed control problem:

Theorem 2.4. *Suppose that (A) holds. Let $\kappa < 2$, $p, q \in \mathbb{N}^+$ and let (u_h, ϕ_h) in $V_h^p \times V_h^q$ be the solution of (2.13). Let $u \in H^{p+1}(M)$ and $\phi \in H^{q+1}(M)$ solve (2.1). Then for small enough $h > 0$*

$$|||(u - u_h, \phi - \phi_h)||| \lesssim h^{p+1-\frac{\kappa}{2}} \|u\|_{H^{p+1}(M)} + h^{q+1+\frac{\kappa}{2}} \|\phi\|_{H^{q+1}(M)}.$$

In particular,

$$\|\chi(\phi - \phi_h)\|_{L^2(M)} \lesssim h^{p-\frac{\kappa}{2}} \|u\|_{H^{p+1}(M)} + h^{q+\frac{\kappa}{2}} \|\phi\|_{H^{q+1}(M)}. \tag{2.15}$$

Let us begin by establishing some facts needed in the proof. Equation (2.13) defines a finite element method that is consistent in the sense that if smooth enough u and ϕ satisfy (2.1), then (2.13) holds for (u,ϕ). This follows from the weak formulation (2.3) of (2.1) together with the regularization vanishing for (u,ϕ). In particular, if $(u_h, \phi_h) \in V_h^p \times V_h^q$ solves (2.13) then the following Galerkin orthogonality holds

$$A[(u - u_h, \phi - \phi_h), (v, \psi)] = 0 \quad \text{for all } (v,\psi) \in V_h^p \times V_h^q. \tag{2.16}$$

It is straightforward to see that for all $u, \phi, v, \psi \in H^2(M) + V_h$ there holds

$$A[(u,\phi),(v,\psi)] - (a(v,\phi) + a(u,\psi)) \lesssim |||(u,\phi)||| \, |||(v,\psi)|||. \tag{2.17}$$

We will need the following continuity estimates for a.

Lemma 2.5. *For all $u, \phi, v, \psi \in H^2(M) + V_h$ vanishing on Γ there holds*

$$\begin{aligned} a(v,\phi) &\lesssim S^{\frac{1}{2}}(\phi) \|v\|_{H^1(\mathcal{T}_h)}, \\ a(u,\psi) &\lesssim \left(S^{\frac{1}{2}}(u) + \sum_{\tau=0,T} E^{\frac{1}{2}}(U|_{t=\tau}) \right) \|\psi\|_{H^2(\mathcal{T}_h)}. \end{aligned}$$

Proof. Recalling (1.7) we see that

$$\begin{aligned} a(v,\phi) &= \int_M g(hdv, hd\phi) \, \mathrm{d}x - h(v, h\partial_\nu \phi)_{L^2(\partial M \setminus \Gamma)} \\ &= \sum_{K \in \mathcal{T}_h} \int_K v h^2 \Box \phi \, \mathrm{d}x + \sum_{F \in \mathcal{F}_h} h \int_F v [\![h\partial_\nu \phi]\!] \, \mathrm{d}s \end{aligned}$$

and the first claimed estimate follows from the Cauchy–Schwarz inequality and the trace inequality (2.9).

Let us now turn to the second estimate. We have

$$\begin{aligned} a(u,\psi) &= \int_M g(hdu, hd\psi)\ \mathrm{d}x - h(u, h\partial_\nu \psi)_{L^2(\partial M\setminus\Gamma)} \\ &= \sum_{K\in\mathcal{T}_h}\int_K h^2\Box u\psi\ \mathrm{d}x + \sum_{F\in\mathcal{F}_h} h\int_F [\![h\partial_\nu u]\!]\psi\ \mathrm{d}s \\ &\qquad + h(h\partial_\nu u,\psi)_{L^2(\partial M\setminus\Gamma)} - h(u,h\partial_\nu\psi)_{L^2(\partial M\setminus\Gamma)}, \end{aligned}$$

and the second estimate follows. □

We denote the Scott–Zhang interpolant by

$$i_h^p : \{u\in H^1(M) : u|_\Gamma = 0\} \to V_h^p, \tag{2.18}$$

and recall that, contrary to the more classical nodal interpolant, the Scott–Zhang interpolant is well-defined for functions in $H^1(M)$, while having the desirable property that it preserves vanishing boundary conditions, just like the nodal interpolant. This guarantees that the interpolant gives indeed a map between the spaces in (2.18). The interpolant is constructed *via* local averaging over simplices in $\mathcal{T}_h$, or over their faces when dealing with boundary nodes. We refer to [46] for the precise definition. There it is also shown that for all $p\in\mathbb{N}^+$ and $k=1,\dots,p+1$ there is $C>0$ such that for all $h>0$ and $u\in H^k(M)$

$$\|u - i_h^p u\|_{H^k(\mathcal{T}_h)} \le C\|(hD)^k u\|_{L^2(M)}. \tag{2.19}$$

Proof of Theorem 2.4. We write

$$w = u_h - u, \quad \eta = \phi_h - \phi, \quad w_h = u_h - i_h^p u, \quad \eta_h = \phi_h - i_h^q\phi. \tag{2.20}$$

By Lemma 2.3 and the Galerkin orthogonality (2.16),

$$|\!|\!|(w,\eta)|\!|\!|^2 \lesssim A[(w,\eta),(w,-\eta)] = A[(w,\eta),(w-w_h,\eta_h-\eta)].$$

We write

$$w_i = i_h^p u - u, \quad \eta_i = i_h^q\phi - \phi.$$

Observing that $w - w_h = w_i$ and $\eta_h - \eta = -\eta_i$, it follows from (2.17) and Lemma 2.5 that

$$|\!|\!|(w,\eta)|\!|\!| \lesssim |\!|\!|(w_i,\eta_i)|\!|\!| + h^{-\frac{\kappa}{2}}\|w_i\|_{H^1(\mathcal{T}_h)} + h^{\frac{\kappa}{2}}\|\eta_i\|_{H^2(\mathcal{T}_h)}.$$

Recalling the scaling in (2.14) and using the bound (2.8), with a replaced by s, c and e, we see that

$$|\!|\!|(w_i,\eta_i)|\!|\!| \lesssim h^{-\frac{\kappa}{2}}\|w_i\|_{H^2(\mathcal{T}_h)} + h^{\frac{\kappa}{2}}\|\eta_i\|_{H^2(\mathcal{T}_h)}.$$

Finally, using (2.19),

$$h^{-\frac{\kappa}{2}}\|w_i\|_{H^2(\mathcal{T}_h)} + h^{\frac{\kappa}{2}}\|\eta_i\|_{H^2(\mathcal{T}_h)} \lesssim h^{p+1-\frac{\kappa}{2}}\|u\|_{H^{p+1}(M)} + h^{q+1+\frac{\kappa}{2}}\|\phi\|_{H^{q+1}(M)}.$$

□

Remark 2.6. Recall that if $(u_0, u_1) \in H^{k+1}(\Omega) \times H^k(\Omega)$ satisfies the compatibility conditions of order k, then the unique solution (u, ϕ) to (2.1) is in $H^{k+1}(M) \times H^k(M)$. Hence we can take $p \leq k$ and $q \leq k-1$ in Theorem 2.4. Choosing $\kappa = 0$ and $p = q \leq k-1$ leads to convergence of order q as stated in the introduction, see (1.3). On the other hand, we can also use the maximal polynomial orders $p = k$ and $q = k-1$, take $\kappa = 1$, and obtain convergence of order $k - 1/2$.

Recall that the norm $|||\cdot|||_E$ is defined by (1.4). Theorem 2.4 has the following corollary.

Corollary 2.7. *Suppose that (A) holds. Let $\phi_h \in V_h^q$ be as in Theorem 2.4. Suppose that $p \geq 2$ and $\kappa = 0$. Let $\tilde{u}_h \in V_h^p$ be the approximation of the solution to*

$$\begin{cases} \Box \tilde{u} = f, \\ \tilde{u}|_{x \in \partial\Omega} = 0, \\ \tilde{u}|_{t=T} = 0, \ \partial_t \tilde{u}|_{t=T} = 0, \end{cases} \tag{2.21}$$

with $f = \chi\phi_h$, from the DG-CG spacetime method [48], using polynomial order of degree less than or equal top $p-1$ in space and in time, with the time step equal to the space step. Let $u \in H^{p+1}(M) \cap W^{p,1}[0,T; H^2(\Omega)] \cap W^{2,1}[0,T; H^p(\Omega)]$ and $\phi \in H^{q+1}(M)$ solve (2.1). Then for small enough $h > 0$

$$|||u - \tilde{u}_h|||_E \lesssim h^p + h^q.$$

Proof. The following stability estimate holds ([48], Thm. 4.5)

$$|||\tilde{u}_h|||_E \lesssim \|f\|_{L^2(M)}, \tag{2.22}$$

for the DG-CG approximation of the solution to (2.21). In addition the following optimal error estimate holds ([48], Thm. 5.4): if $\tilde{u}$ is a smooth enough solution to (2.21), then

$$|||\tilde{u} - \tilde{u}_h|||_E \lesssim h^p. \tag{2.23}$$

Note that u is the solution to (2.21) with $f = \chi\phi$. Let v_h be the DG-CG approximation of the solution to (2.21) with $f = \chi\phi$. Then $v_h - \tilde{u}_h$ is the DG-CG approximation of the solution to (2.21) with $f = \chi(\phi - \phi_h)$. Then by (2.22) and (2.23)

$$\begin{aligned} |||u - u_h|||_E &\leq |||u - v_h|||_E + |||v_h - u_h|||_E \\ &\lesssim h^p + \|\chi(\phi - \phi_h)\|_{L^2(M)}. \end{aligned}$$

The conclusion follows now from (2.15). $\square$

Remark 2.8. We believe that u_h in Theorem 2.4 satisfies no better estimate than

$$\begin{aligned} &\|u - u_h\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t(u - u_h)\|_{L^2(0,T;H^{-1}(\Omega))} \\ &\quad \lesssim h^p \|u\|_{H^{p+1}(M)} + h^q \|\phi\|_{H^{q+1}(M)}. \end{aligned} \tag{2.24}$$

In a previous work by three of the authors of the present paper, an estimate similar to (2.24) was given for the unique continuation problem, that is the dual problem of the control problem studied here, see Theorem 2.2 of [9]. There the space for the time derivate was incorrect: the space $L^\infty(0,T; H^{-1}(\Omega))$ was used while the estimate holds only in the space $L^2(0,T; H^{-1}(\Omega))$ as in (2.24). The error was repeated in Theorem 1.1 of [8]. The results in the above references are nevertheless correct without further modifications after the former space is replaced by the latter one. See also Theorem A.4, Remark A.5 and Proposition A.1 in the appendix below for more details.

3. Distributed control with limited regularity

In this section we will study the finite element method (2.13) in the case that the continuum solution (u, ϕ) to the control problem (2.1) is in the natural energy class $H^1(M) \times L^2(M)$. Our theory is not complete in this case, as we were able to show a convergence result only in the case that $u_0 = 0$ in (2.1). We make the standing assumption that (A) holds, so that (2.13) has a unique solution. We will prove the following result:

Theorem 3.1. *Suppose that $u_0 = 0$ and $u_1 \in L^2(\Omega)$. Let (u_h, ϕ_h) be the solution of (2.13) with $\kappa = 0$, and let (u, ϕ) be the solution of (2.1). Then there is a sequence $h_j \to 0$ such that (u_{h_j}, ϕ_{h_j}) converges weakly to (u, ϕ) in $L^2(M)$.*

Before proving the above theorem, we establish four lemmas used in the proof.

Lemma 3.2. *Let $(u_h, \phi_h) \in V_h^p \times V_h^q$ be the solution of (2.13) with $\kappa = 0$. Then*

$$h^{-1}\|h^2\Box\phi_h\|_{H^{-1}(M)} \lesssim h^{\frac{1}{2}}\|hu_1\|_{L^2(\Omega)} + |||(u_h, \phi_h)|||,$$
$$h^{-1}\|h^2\Box u_h\|_{H^{-1}(M)} \lesssim h^{\frac{1}{2}}\|u_0\|_{L^2(\Omega)} + |||(u_h, \phi_h)|||.$$

Proof. By duality there holds

$$\|\Box\phi_h\|_{H^{-1}(M)} = \sup_{v\in H_0^1(M)\setminus\{0\}} \frac{\langle\Box\phi_h, v\rangle_{H^{-1}\times H_0^1(M)}}{\|v\|_{H^1(M)}}.$$

Using also

$$\langle h^2\Box\phi_h, v\rangle_{H^{-1}\times H_0^1(M)} = \int_M g(hdv, hd\phi_h)\ \mathrm{d}x, \quad v \in H_0^1(M),$$

we see that to establish the first claimed inequality it is enough to show that for $v \in H_0^1(M)$ there holds

$$\int_M g(hdv, hd\phi_h)\ \mathrm{d}x \lesssim (h^{\frac{1}{2}}\|hu_1\|_{L^2(\Omega)} + |||(u_h, \phi_h)|||)\|hdv\|_{L^2(M)}. \tag{3.1}$$

We have

$$\int_M g(hdv, hd\phi_h)\ \mathrm{d}x = \sum_{K\in\mathcal{T}_h}\int_K vh^2\Box\phi_h\ \mathrm{d}x + \sum_{F\in\mathcal{F}_h} h\int_F v[\![h\partial_\nu\phi_h]\!]\ \mathrm{d}s.$$

Let $v_h \in V_h^p$ be the Scott–Zhang interpolant of v, and use the above equation with v replaced by $v - v_h$, together with the Cauchy–Schwarz inequality and (2.9), to get

$$\int_M g(hd(v - v_h), hd\phi_h)\ \mathrm{d}x \lesssim |||(0, \phi_h)|||\|v - v_h\|_{H^1(\mathcal{T}_h)}.$$

Now (2.19) implies

$$\int_M g(hd(v - v_h), hd\phi_h)\ \mathrm{d}x \lesssim |||(0, \phi_h)|||\|hdv\|_{L^2(M)}.$$

To establish (3.1) it remains to show that (3.1) holds with v replaced by v_h.

Using the assumption that (u_h, ϕ_h) satisfies (2.13), taking $\psi = 0$ there, we have

$$\begin{aligned} -\int_M g(hdv_h, hd\phi_h)\ \mathrm{d}x &= h^{-\kappa} s(u_h, v_h) + h^{-\kappa} \sum_{\tau=0,T} h(h\partial_t u_h|_{t=\tau}, h\partial_t v_h|_{t=\tau})_{L^2(\Omega)} \\ &\quad + h^{2-\kappa}\rho(v_h, \phi_h) - h^{-\kappa} h(hu_1, h\partial_t v_h|_{t=0})_{L^2(\Omega)}, \end{aligned} \tag{3.2}$$

where $\kappa = 0$. We recall that all the bilinear forms in the paper are scaled so that they are bounded by some semiclassical Sobolev norm. Moreover, all the semiclassical Sobolev norms are equivalent on piecewise polynomial functions, see (2.10). Hence, applying the Cauchy–Schwarz inequality to each term on the right-hand side of (3.2), and observing that the terms depend on v_h only *via* hdv_h, we obtain

$$\int_M g(hdv_h, hd\phi_h)\ \mathrm{d}x \lesssim (h^{\frac{1}{2}} \|hu_1\|_{L^2(\Omega)} + |||(u_h, \phi_h)|||) \|hdv_h\|_{L^2(M)}.$$

This finishes the proof of the first claimed inequality.

We turn to the second claimed inequality. Analogously to the above, it is enough to show that for $\psi \in H_0^1(M)$ there holds

$$\int_M g(hdu_h, hd\psi)\ \mathrm{d}x \lesssim (h^{\frac{1}{2}} \|u_0\|_{L^2(\Omega)} + |||(u_h, \phi_h)|||) h \|\psi\|_{H^1(M)}.$$

Let $\psi_h \in V_h^p$ be the Scott–Zhang interpolant of ψ. As above, we have

$$\int_M g(hdu_h, hd(\psi - \psi_h))\ \mathrm{d}x \lesssim |||(u_h, 0)||| \|hd\psi\|_{L^2(M)}.$$

Moreover, due to (2.13),

$$\begin{aligned} -\int_M g(hdu_h, hd\psi_h)\ \mathrm{d}x &= h(u_h, h\partial_\nu \psi_h)_{L^2(\partial M\setminus\Gamma)} - h^\kappa s(\phi_h, \psi_h) - h^2 c(\phi_h, \psi_h) \\ &\quad + h^{4-\kappa} \tilde{c}(\phi_h, \psi_h) + h^{2-\kappa}\rho(u_h, \psi_h) + h(u_0, h\partial_t \psi_h|_{t=0})_{L^2(\Omega)}. \end{aligned}$$

The second claimed inequality follows in a similar way as the first one. Here the right-hand side does not depend on ψ_h only *via* $hd\psi_h$, however, the terms depending on the whole ψ_h are scaled with positive powers of h. ☐

Lemma 3.3. *Let (u_h, ϕ_h) be the solution of (2.13) with $\kappa = 0$. Then*

$$|||(u_h, \phi_h)||| \lesssim h^{\frac{1}{2}} \|u_0\|_{L^2(\Omega)} + h\|u_0\|_{H_0^1(\Omega)} + h\|u_1\|_{L^2(\Omega)}. \tag{3.3}$$

Proof. By Lemma 2.3 there holds

$$\begin{aligned} |||(u_h, \phi_h)|||^2 &\lesssim A[(u_h, \phi_h), (u_h, -\phi_h)] = e(U_0, U_h|_{t=0}) - L(\phi_h) \\ &\le h^{\frac{1}{2}} (\|u_0\|_{L^2(\Omega)} + \|hu_1\|_{L^2(\Omega)}) E^{\frac{1}{2}}(U_h|_{t=0}) \\ &\quad + h^2 (\|u_0\|_{H_0^1(\Omega)} + \|u_1\|_{L^2(\Omega)}) (\|\phi_h|_{t=0}\|_{L^2(\Omega)} + \|\partial_t \phi_h|_{t=0}\|_{H^{-1}(\Omega)}). \end{aligned}$$

By the distributed observability estimate, see Theorem A.4 in Appendix A,

$$\|\phi_h|_{t=0}\|_{L^2(\Omega)} + \|\partial_t \phi_h|_{t=0}\|_{H^{-1}(\Omega)} \lesssim C^{\frac{1}{2}}(\phi_h) + \|\Box \phi_h\|_{H^{-1}(M)}.$$

Recalling that $hC^{\frac{1}{2}}(\phi_h) \lesssim |||(0, \phi_h)|||$, and using Lemma 3.2, we obtain

$$h(\|\phi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t \phi|_{t=0}\|_{H^{-1}(\Omega)}) \lesssim h^{\frac{1}{2}}\|hu_1\|_{L^2(\Omega)} + |||(u_h, \phi_h)|||.$$

As also $E^{\frac{1}{2}}(U_h|_{t=0}) \leq |||(u_h, \phi_h)|||$, we have

$$\begin{aligned}&|||(u_h, \phi_h)|||^2 \\ &\quad \lesssim (h^{\frac{1}{2}}\|u_0\|_{L^2(\Omega)} + h\|u_0\|_{H_0^1(\Omega)} + h\|u_1\|_{L^2(\Omega)})(h^{\frac{1}{2}}\|hu_1\|_{L^2(\Omega)} + |||(u_h, \phi_h)|||),\end{aligned}$$

leading to (3.3). □

Lemma 3.4. *Let (u_h, ϕ_h) be the solution of (2.13) with $\kappa = 0$ and $u_0 = 0$. Then*

$$\|u_h\|_{L^2(M)} + \|\phi_h\|_{L^2(M)} \lesssim \|u_1\|_{L^2(\Omega)}.$$

Proof. Lemmas 3.2 and 3.3 imply

$$\|\Box u_h\|_{H^{-1}(M)} + \|\Box \phi_h\|_{H^{-1}(M)} \lesssim \|u_1\|_{L^2(\Omega)}.$$

Moreover, it follows from (3.3) that

$$C^{\frac{1}{2}}(\phi_h) \lesssim \|u_1\|_{L^2(\Omega)}, \quad \|\partial_t^j u_h|_{t=0}\|_{L^2(\Omega)} \lesssim h^{\frac{1}{2}-j}\|u_1\|_{L^2(\Omega)}, \qquad j = 0, 1.$$

The bound $\|\phi_h\|_{L^2(M)} \leq C\|u_1\|$ follows from the distributed observability estimate, see Remark A.5. It remains to show the same bound for u_h. We face the complication that the above estimates do not allow us to conclude that $\partial_t u_h|_{t=0}$ is bounded.

To overcome this, we will employ $\tilde{u}_h \in V_h^p$ that coincides with u_h on ∂M and satisfies (3.4) and (3.5). We have

$$\|\Box \tilde{u}_h\|_{H^{-1}(M)} \lesssim \|u_1\|_{L^2(\Omega)}.$$

Indeed, for any $v \in H_0^1(M)$ there holds, using (2.11) and (3.5),

$$\begin{aligned}(h^2\Box \tilde{u}_h, v)_{L^2(M)} &= \int_M g(hd\tilde{u}_h, hdv)\ \mathrm{d}x \\ &\leq \int_M g(hdu_h, hdv)\ \mathrm{d}x + \|h\nabla(\tilde{u}_h - u_h)\|_{L^2(M)}\|h\nabla v\|_{L^2(M)} \\ &\lesssim \|h^2\Box u_h\|_{H^{-1}(M)}\|v\|_{H_0^1(M)} + h^{\frac{1}{2}}\|h(\partial_t u_h|_{t=0} - u_1)\|_{L^2(\Omega)}\|h\nabla v\|_{L^2(M)}. \\ &\lesssim h^2(\|u_1\|_{L^2(\Omega)} + h^{\frac{1}{2}}\|\partial_t u_h|_{t=0}\|_{L^2(\Omega)})\|h\nabla v\|_{L^2(M)} \leq h^2\|u_1\|_{L^2(\Omega)}\|h\nabla v\|_{L^2(M)}.\end{aligned}$$

Moreover, using (3.4),

$$\begin{aligned}\|\partial_t \tilde{u}_h|_{t=0}\|_{H^{-1}(\Omega)} &\leq \|\partial_t \tilde{u}_h|_{t=0} - u_1\|_{H^{-1}(\Omega)} + \|u_1\|_{H^{-1}(\Omega)} \\ &\lesssim h\|\partial_t u_h|_{t=0} - u_1\|_{L^2(\Omega)} + \|u_1\|_{H^{-1}(\Omega)} \leq h^{\frac{1}{2}}\|u_1\|_{L^2(\Omega)} + \|u_1\|_{H^{-1}(\Omega)}.\end{aligned}$$

Recalling that $\tilde{u}_h$ coincides with u_h on ∂M, we conclude that

$$\|\tilde{u}_h\|_{L^2(M)} \lesssim \|u_1\|_{L^2(\Omega)}$$

follows from an energy estimate, see Proposition A.1 in Appendix A. Finally, using (3.5),

$$\begin{aligned}\|u_h\|_{L^2(M)} &\le \|u_h - \tilde u_h\|_{L^2(M)} + \|\tilde u_h\|_{L^2(M)} \\ &\lesssim h^{\frac12}\|h(\partial_t u_h|_{t=0} - u_1)\|_{L^2(\Omega)} + \|u_1\|_{L^2(\Omega)} \lesssim \|u_1\|_{L^2(\Omega)}.\end{aligned}$$

□

Lemma 3.5. *Let $p \in \mathbb{N}^+$ and consider a family $u_h \in V_h^p$, $h > 0$. Let $u_1 \in L^2(\Omega)$. Then there is a family $\tilde u_h \in V_h^p$, $h > 0$, such that $\tilde u_h|_{\partial M} = u_h|_{\partial M}$ and*

$$h^{-1}\|\partial_t \tilde u_h|_{t=0} - u_1\|_{H^{-1}(\Omega)} \lesssim \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\Omega)}, \tag{3.4}$$

$$\|u_h - \tilde u_h\|_{L^2(M)} \lesssim h^{\frac12}\|h(\partial_t u_h|_{t=0} - u_1)\|_{L^2(\Omega)}. \tag{3.5}$$

Proof. Let us consider the trace mesh at $t = 0$,

$$\mathcal{F}_{h,0} = \{\partial K \cap \{t = 0\} : K \in \mathcal{T}_h\}.$$

We decompose $\mathcal{F}_{h,0}$ into a set of N_h disjoint patches $\mathcal{Q}_i$, $i = 1, \dots, N_h$, such that each patch contains several element faces but their area and diameter satisfy

$$h^n \lesssim |\mathcal{Q}_i| \lesssim h^n, \quad h \lesssim \operatorname{diam}(\mathcal{Q}_i) \lesssim h.$$

Then we define disjoint patches $\mathcal{P}_i$ consisting of elements of $\mathcal{T}_h$ so that

$$\mathcal{Q}_i = \mathcal{P}_i \cap \{t = 0\}$$

and that $h^{n+1} \lesssim |\mathcal{P}_i| \lesssim h^{n+1}$. Now we define the functions $p_i \in V_h^1$ such that $\operatorname{supp}(p_i) \subset \mathcal{P}_i$ and $p_i(x) = 1$ for every node x in the interior of $\mathcal{P}_i$. We require that the patches $\mathcal{Q}_i$ are large enough so that, writing

$$\alpha_i = \int_{\mathcal{Q}_i} \partial_t p_i|_{t=0}\, \mathrm{d}s, \quad \beta_i = \|\partial_t p_i|_{t=0}\|_{L^2(\mathcal{Q}_i)}, \quad \gamma_i = \|p_i\|_{L^2(\mathcal{P}_i)},$$

there holds $h^{n-1} \lesssim \alpha_i \lesssim h^{n-1}$, $h^{\frac n2 - 1} \lesssim \beta_i \lesssim h^{\frac n2 - 1}$ and $h^{\frac12(n+1)} \lesssim \gamma_i \lesssim h^{\frac12(n+1)}$.

We set

$$\tilde u_h = u_h + \sum_{i=1}^{N_h} w_i p_i, \quad w_i = -\alpha_i^{-1}\int_{\mathcal{Q}_i} (\partial_t u_h|_{t=0} - u_1)\, \mathrm{d}s.$$

Then

$$\int_{\mathcal{Q}_i} (\partial_t \tilde u_h|_{t=0} - u_1)\, \mathrm{d}s = 0. \tag{3.6}$$

To establish (3.4) we let $v \in H_0^1(\Omega)$ and show that

$$(\partial_t \tilde u_h|_{t=0} - u_1, v)_{L^2(\Omega)} \lesssim \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\Omega)}\|h\nabla v\|_{L^2(\Omega)}.$$

Let $\bar{v} \in L^2(\Omega)$ be equal to the average of v on each patch $\mathcal{Q}_i$, that is,

$$\bar{v}|_{\mathcal{Q}_i} = |\mathcal{Q}_i|^{-1} \int_{\mathcal{Q}_i} v \, \mathrm{d}s.$$

Now (3.6) implies $(\partial_t \tilde{u}_h|_{t=0} - u_1, \bar{v})_{L^2(\Omega)} = 0$, and

$$\begin{aligned}(\partial_t \tilde{u}_h|_{t=0} - u_1, v)_{L^2(\Omega)} &= (\partial_t \tilde{u}_h|_{t=0} - u_1, v - \bar{v})_{L^2(\Omega)} \\ &\le \|\partial_t \tilde{u}_h|_{t=0} - u_1\|_{L^2(\Omega)} \|v - \bar{v}\|_{L^2(\Omega)} \lesssim \|\partial_t \tilde{u}_h|_{t=0} - u_1\|_{L^2(\Omega)} \|h \nabla v\|_{L^2(\Omega)}.\end{aligned}$$

Here we used the Poincaré inequality as stated for example in [19]. To establish (3.4) it remains to show that

$$\|\partial_t \tilde{u}_h|_{t=0} - u_1\|_{L^2(\Omega)} \lesssim \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\Omega)}.$$

Using the fact that the patches $\mathcal{Q}_i$ are disjoint, we have

$$\|\partial_t \tilde{u}_h|_{t=0} - u_1\|_{L^2(\Omega)} \le \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\Omega)} + \sum_{i=1}^{N_h} |w_i| \|\partial_t p_i|_{t=0}\|_{L^2(\mathcal{Q}_i)}.$$

Recalling that α_i behaves like h^{n-1} and β_i like $h^{\frac{n}{2}-1}$, we obtain using the Cauchy–Schwarz inequality

$$\begin{aligned}|w_i| \|\partial_t p_i|_{t=0}\|_{L^2(\mathcal{Q}_i)} &= \alpha_i^{-1} \beta_i \left| \int_{\mathcal{Q}_i} (\partial_t u_h|_{t=0} - u_1) \, \mathrm{d}s \right| \\ &\lesssim h^{1-n} h^{\frac{n}{2}-1} h^{\frac{n}{2}} \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\mathcal{Q}_i)} = \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\mathcal{Q}_i)}.\end{aligned}$$

Let us now turn to (3.5). Note that

$$\|u_h - \tilde{u}_h\|_{L^2(M)}^2 = \sum_{i=1}^{N_h} |w_i|^2 \|p_i\|_{L^2(\mathcal{P}_i)}^2.$$

Recalling that γ_i^2 behaves like h^{n+1}, we obtain

$$|w_i|^2 \|p_i\|_{L^2(\mathcal{P}_i)}^2 = |w_i|^2 \gamma_i^2 \lesssim h^{n+1} h^{2(1-n)} h^n \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\mathcal{Q}_i)}^2,$$

leading to

$$\|u_h - \tilde{u}_h\|_{L^2(M)}^2 \lesssim h^3 \|\partial_t u_h|_{t=0} - u_1\|_{L^2(\Omega)}^2.$$

□

Proof of Theorem 3.1. By Lemma 3.4 both u_h and ϕ_h are bounded in $L^2(M)$. Thus there is a sequence $h_j \to 0$ such that (u_{h_j}, ϕ_{h_j}) converges weakly to a function (u_*, ϕ_*) in $L^2(M)$. By Lemma A.8 it is enough to show that (u_*, ϕ_*) satisfies (2.1).

As the embedding $H^{-\epsilon}(M) \subset L^2(M)$ is compact for $\epsilon > 0$, by passing to a subsequence, we may assume that $(u_{h_j}, \phi_{h_j}) \to (u_*, \phi_*)$ in $H^{-\epsilon}(M)$. By Lemmas 3.2 and 3.3 we may further assume that $(\Box u_{h_j}, \Box \phi_{h_j}) \to (\Box u_*, \Box \phi_*)$ in $H^{-\epsilon-1}(M)$. For $\epsilon < 1/2$ it follows from Lemma A.7 that

$$0 = (u_{h_j}|_\Gamma, \phi_{h_j}|_\Gamma) \to (u_*|_\Gamma, \phi_*|_\Gamma).$$

Thus (u_*, ϕ_*) satisfies the homogeneous lateral boundary conditions in (2.1).

For any $\psi \in C^\infty(M)$ with $\psi|_\Gamma = 0$ and any $v \in C_0^\infty(M)$ there holds

$$h^{-2}(a(u_h, \psi) - h^2 c(\phi_h, \psi) - L(\psi)) \to 0, \tag{3.7}$$

$$h^{-2} a(v, \phi_h) \to 0, \tag{3.8}$$

as $h \to 0$. Before showing (3.7)–(3.8), let us show that they imply that (u_*, ϕ_*) satisfies (2.1). The equation $\Box \phi_* = 0$ follows immediately from (3.8). Observe that

$$h^{-2} a(u_h, \psi) = (u_h, \Box \psi)_{L^2(M)} \to (u_*, \Box \psi)_{L^2(M)}.$$

It follows from (3.7) that for any $\psi \in C^\infty(M)$ vanishing on Γ there holds

$$(u_*, \Box \psi)_{L^2(M)} = c(\phi_*, \psi) + (u_1, \psi|_{t=0})_{L^2(\Omega)} - (u_0, \partial_t \psi|_{t=0})_{L^2(\Omega)}. \tag{3.9}$$

In particular, taking $\psi \in C_0^\infty(M)$ we see that $\Box u_* = \chi \phi_*$.

To show that (u_*, ϕ_*) satisfies (2.1), it remains to verify the initial and final conditions for u_*. We have $u_* \in L^2(M)$ and

$$\partial_t^2 u_* = \Delta u_* + \chi \phi_* \in L^2(0, T; H^{-2}(\Omega)).$$

Now Theorem 3.1, page 19 of [36] gives

$$u_* \in C(0, T; H^{-\frac{1}{2}}(\Omega)), \quad \partial_t u_* \in C(0, T; H^{-\frac{3}{2}}(\Omega)).$$

Taking $\psi(t, x) = \psi_0(t)\psi_1(x)$, with $\psi_0 \in C^\infty(0, T)$ and $\psi_1 \in C_0^\infty(\Omega)$, we integrate by parts

$$\begin{aligned}
(u_*, \Box \psi)_{L^2(M)} &= \int_0^T \langle u_*, \psi_1 \rangle \partial_t^2 \psi_0 \, \mathrm{d}t - \int_0^T \langle u_*, \Delta \psi_1 \rangle \psi_0 \, \mathrm{d}t \\
&= \int_0^T (\chi \phi_*, \psi_1)_{L^2(\Omega)} \psi_0 \, \mathrm{d}t + \left[\langle u_*, \psi_1 \rangle \partial_t \psi_0 - \langle \partial_t u_*, \psi_1 \rangle \psi_0 \right]_{t=0}^{t=T},
\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is the pairing between distribution and test functions on Ω. Comparison with (3.9) shows that u_* satisfies the initial and final conditions in (2.1).

Let us now show (3.7). Denote by ψ_h the Scott–Zhang interpolant of ψ. By (2.13)

$$\begin{aligned}
&a(u_h, \psi) - h^2 c(\phi_h, \psi) - L(\psi) \\
&\quad = a(u_h, \phi - \psi_h) - h^2 c(\phi_h, \psi - \psi_h) - L(\psi - \psi_h) \\
&\qquad + s(\phi_h, \psi_h) - h^4 \tilde{c}(\phi_h, \psi_h) - h^2 \rho(u_h, \psi_h).
\end{aligned}$$

Using the continuity of a in Lemma 2.5, the interpolation estimate (2.19), and the bound (3.3) for the residual norm, we obtain

$$|a(u_h, \psi - \psi_h)| \lesssim |||(u_h, \phi_h)||| \|(hD)^2 \psi\|_{L^2(M)} \lesssim h \|u_1\|_{L^2(M)} \|(hD)^2 \psi\|_{L^2(M)}.$$

Recalling that $hC^{\frac12}(\phi_h) \lesssim |||(0,\phi_h)|||$, we use the continuity (2.8) for c and the interpolation estimate (2.19) to get

$$h^2|c(\phi_h, \psi - \psi_h)| \le h^2 C^{\frac12}(\phi_h) C^{\frac12}(\psi - \psi_h) \lesssim h|||(0,\phi_h)||| \|(hD)^2\psi\|_{L^2(M)}.$$

Using once again (2.19),

$$|L(\psi - \psi_h)| = h^2 |(u_1, (\psi - \psi_h)|_{t=0})_{L^2(\Omega)}| \lesssim h^{3/2} \|u_1\|_{L^2(\Omega)} \|(hD)^2\psi\|_{L^2(M)}.$$

Turning to the first term related to regularization, we have

$$|s(\phi_h, \psi_h)| \le S^{\frac12}(\phi_h) S^{\frac12}(\psi_h),$$

where the first factor is bounded by $|||(0,\phi_h)||| \lesssim h\|u_1\|_{L^2(\Omega)}$, and the second satisfies

$$\begin{aligned} S(\psi_h) &\lesssim \sum_{K\in\mathcal{T}_h} (\|h^2\Box(\psi_h - \psi)\|^2_{L^2(K)} + \|h^2\Box\psi\|^2_{L^2(K)}) + \sum_{F\in\mathcal{F}_h} h\|[\![h\partial_\nu(\psi_h - \psi)]\!]\|^2_{L^2(F)} \\ &\lesssim \|(hD)^2\psi\|^2_{L^2(M)}. \end{aligned}$$

Finally,

$$\begin{aligned} h^4|\tilde c(\phi_h, \psi_h)| &\lesssim h^3 |||(0,\phi_h)||| \|\psi\|_{L^2(M)}, \\ h^2|\rho(u_h, \psi_h)| &\lesssim h^2 |||(u_h, 0)||| \|\psi\|_{L^2(M)} \lesssim h^3 \|u_1\| \|\psi\|_{L^2(M)}, \end{aligned}$$

and (3.7) follows.

We turn to (3.8). Denote by v_h the Scott–Zhang interpolant of v. By (2.13)

$$a(v, \phi_h) = a(v - v_h, \phi_h) - s(u_h, v_h) + h^2 \rho(v_h, \phi_h).$$

Similarly to the bounds above, we have

$$\begin{aligned} |a(v - v_h, \phi_h)| + |s(u_h, v_h)| &\lesssim h\|u_1\|_{L^2(\Omega)} \|(hD)^2 v\|_{L^2(M)}, \\ h^2|\rho(v_h, \phi_h)| &\lesssim h|||(0,\phi_h)||| \|h^2\Box v\|_{L^2(M)}, \end{aligned}$$

and (3.8) follows. This finishes the proof that (u_*, ϕ_*) satisfies (2.1). □

4. Boundary control

Let us give a numerical method to solve the boundary control problem given by equations (1.5) and (1.2). We will follow the same strategy as in the case of distributed control. That is, we give a weak formulation analogous to (2.3), see (4.4) below, and discretize and regularize the weak formulation leading to the finite dimensional linear system (2.13) as before but with the bilinear form A modified to that given in (4.10). The only essential new feature in the proof is that we do not make the assumption (2.5), and deal with resulting technicalities that were avoided above.

Let us begin by formulating our assumptions on the cutoff function χ in (1.5). We consider a function of the form

$$\chi(t, x) = \chi_0(t)\chi_1^2(x),$$

where $\chi_0 \in C_0^\infty([0,T])$ and $\chi_1 \in C^\infty(\Gamma)$ take values in $[0,1]$, and suppose that

(A') $\chi = 1$ on $(a,b) \times \omega \subset \Gamma$ satisfying the geometric control condition.

Here $0 < a < b < T$ and $\omega \subset \partial\Omega$ is open in $\partial\Omega$. In the case of boundary control, the geometric control condition means that every compressed generalized bicharacteristic intersects the set $(a,b) \times \omega$, when projected to M. Moreover, the intersection must happen at a nondiffractive point and the lightlike lines must have finite order of contact with Γ. We refer again to [4] for the definitions.

4.1. Weak formulation of the control problem

We let $V \in C^\infty(\Omega)$ and consider the boundary control problem for the following operator

$$P = \Box + V. \tag{4.1}$$

Let $(u_0, u_1) \in L^2(\Omega) \times H^{-1}(\Omega)$. Then the distributed control problem for P can be solved by finding $(u,\phi) \in L^2(M) \times H^1(M)$ such that

$$\begin{cases} Pu = 0, \\ u|_\Gamma = \chi \partial_\nu \phi, \\ u|_{t=0} = u_0,\ \partial_t u|_{t=0} = u_1, \\ u|_{t=T} = 0,\ \partial_t u|_{t=T} = 0, \end{cases} \qquad \begin{cases} P\phi = 0, \\ \phi|_\Gamma = 0. \end{cases} \tag{4.2}$$

If $(u_0, u_1) \in H^{k+1}(\Omega) \times H^k(\Omega)$ satisfies the compatibility conditions of order k, then the unique solution (u,ϕ) to (4.2) satisfies

$$\phi|_{t=T} \in H^{k+2}(\Omega), \quad \partial_t \phi|_{t=T} \in H^{k+1}(\Omega), \tag{4.3}$$

and this initial data for ϕ satisfies the compatibility conditions of order $k+1$, see Thëorem 5.4 of [20]. It follows that $\phi \in H^{k+2}(M)$ and $u \in H^{k+1}(M)$, and the convergence proof for our finite element method is again based on this regularity.

The uniqueness of the solution (u,ϕ) to (4.2) is implictly contained in [20]. For the convenience of the reader, we give a short uniqueness proof in an appendix, see Lemma A.9.

For any $h > 0$, the control problem (4.2) can be formulated weakly as

$$a(u,\psi) = -c(h^{-1}\phi, \psi) + L(\psi), \quad a(v,\phi) = 0, \tag{4.4}$$

for all $v, \psi \in C^\infty(M)$, where

$$\begin{aligned} a(u,\psi) &= \int_M g(hdu, hd\psi)\, \mathrm{d}x + h^2(u, V\psi)_{L^2(M)} \\ &\qquad - h(u, h\partial_\nu\psi)_{L^2(\partial M)} - h(h\partial_\nu u, \psi)_{L^2(\Gamma)} \\ c(\phi,\psi) &= h(\chi h \partial_\nu \phi, h\partial_\nu \psi)_{L^2(\Gamma)}, \end{aligned} \tag{4.5}$$

and L is as in (2.4). Indeed, it follows from (1.7) that if smooth (u,ϕ) solves (4.2) then (4.4) holds for all smooth (v,ψ). We emphasize that a and c are chosen here so that they satisfy the continuity estimate (2.8).

4.2. Discretization

Let us consider a family $\hat{\mathcal{T}} = \{\hat{\mathcal{T}}_h : h > 0\}$ where $\hat{\mathcal{T}}_h$ is a set of $1+n$-dimensional simplices forming a simplicial complex. The family $\hat{\mathcal{T}}$ is parametrized by

$$h = \max_{K \in \hat{\mathcal{T}}_h} \operatorname{diam}(K).$$

Writing $M_h = \bigcup_{K \in \hat{\mathcal{T}}_h} K$, we assume that $M \subset M_h$. We define

$$\mathcal{T}_h = \{\hat{K} \cap M : \hat{K} \in \hat{\mathcal{T}}_h\}, \quad \mathcal{T} = \{\mathcal{T}_h : h > 0\},$$

and require that:

(T) There is $C > 0$ such that for all $h > 0$ and all $K \in \mathcal{T}_h$, letting $\hat{K} \in \hat{\mathcal{T}}_h$ satisfy $K = \hat{K} \cap M$, there are balls $B_1 \subset K$ and $B_2 \supset \hat{K}$ such that the radii r_j of B_j, $j = 1, 2$, satisfy

$$C^{-1} r_2 \le h \le C r_1, \tag{4.6}$$

and that

$$\nu(y) \cdot \rho(y) > C^{-1}, \quad \text{for all } y \in \partial K, \tag{4.7}$$

where ν is the outer unit normal vector of ∂K, and $\rho(y) = (y - x)/|y - x|$ with x is the centre of B_1.

Observe that M_h is a domain with polyhedral boundary, while M has a smooth part of boundary Γ. Passing from $\hat{\mathcal{T}}$ to $\mathcal{T}$ has the effect of cutting simplices near the boundary of M so that they fit perfectly to M, that is, $M = \bigcup_{K \in \mathcal{T}_h} K$. The assumption (T) is needed to guarantee that the family $\mathcal{T}_h$ admits a family interpolation operators with good properties, and that it satisfies the trace inequality, see Lemmas 4.1 and 4.2 below.

If M was polyhedral, then we could choose $\hat{\mathcal{T}}$ so that $M_h = M$ for all small enough $h > 0$. In this case (T) follows if $\hat{\mathcal{T}}$ is quasi-uniform, see Definition 1.140 of [18]. In the case of smooth Ω, we can construct $\hat{\mathcal{T}}$ so that (T) holds for all small enough $h > 0$ by choosing polyhedral sets $M_h \supset M$ that approximate M in the sense that the Hausdorff distance between ∂M_h and ∂M is of order $h^{1+\epsilon}$ for some $\epsilon > 0$, and meshing M_h in a quasi-uniform manner.

We define for $p \in \mathbb{N}^+ = \{1, 2, \dots\}$ the $H^1(M)$-conformal approximation space of polynomial degree p,

$$V_h^p = \{u \in H^1(M) : u|_K \in \mathbb{P}_p(K) \text{ for all } K \in \mathcal{T}_h\}, \tag{4.8}$$

where $\mathbb{P}_p(K)$ denotes the set of polynomials of degree less than or equal to p on K. We write also $V_h = \bigcup_{p \in \mathbb{N}^+} V_h^p$. Note that, contrary to (2.7) no boundary condition is imposed on Γ.

The following two lemmas are proven in Appendix B.

Lemma 4.1. *The trace inequality* (2.9) *holds for the family* $\mathcal{T}$.

Lemma 4.2. *There is a family of interpolation operators* $i_h^p : H^1(M) \to V_h^p$ *satisfying*

$$\|u - i_h^p u\|_{H^k(\mathcal{T}_h)} \lesssim h^k \|u\|_{H^k(M)}. \tag{4.9}$$

Our finite element method has the form: find the critical point of the Lagrangian

$$\mathcal{L}(u, \phi) : V_h^p \times V_h^q \to \mathbb{R}, \quad \mathcal{L}(u, \phi) = \frac{1}{2} c(\phi, \phi) + L(\phi) - \frac{1}{2} \mathcal{R}(u, \phi) + a(u, \phi),$$

where, writing $U = (u, \partial_t u)$ and $U_0 = (u_0, u_1)$, the regularization is given by

$$\begin{aligned}\mathcal{R}(u, \phi) &= h^{-\kappa} S(u) - h^{\kappa} S(\phi) + h^{-\kappa} E(U|_{t=0} - U_0) + h^{-\kappa} E(U|_{t=T}) \\ &\quad + \gamma h^{-\kappa} B(u) - h^{\kappa} B(\phi) + \gamma h^{-\kappa} \widetilde{C}(\phi) + 2\gamma h^{-\kappa} \rho(u, \phi), \\ B(u) &= h\|u\|^2_{L^2(\Gamma)}, \quad \widetilde{C}(\phi) = h\|\chi h \partial_\nu \phi\|^2_{L^2(\Gamma)}, \quad \rho(u, \phi) = -h(u, \chi h \partial_\nu \phi)_{L^2(\Gamma)},\end{aligned}$$

where $\kappa \le 0$ and $\gamma \in (0, 1)$ are fixed constants. Here E and S are as in (2.12) except that $\Box$ in S is replaced by P.

We have $\mathcal{R}(u, h^{-1}\phi) = 0$ for a smooth solution (u, ϕ) to (4.2). Indeed,

$$B(u) + 2\rho(u, h^{-1}\phi) + \widetilde{C}(h^{-1}\phi) = h\|u - \chi \partial_\nu \phi\|^2_{L^2(\Gamma)} = 0,$$

and also $S(u) = S(\phi) = B(\phi) = 0$ and $E(U|_{t=0} - U_0) = E(U|_{t=T}) = 0$. The equation $d\mathcal{L}(u, \phi) = 0$ can be written as (2.13) where the bilinear form A is now given by

$$\begin{aligned}&A[(u, \phi), (v, \psi)] \qquad (4.10)\\ &\quad = h^{-\kappa} s(u, v) - h^{\kappa} s(\phi, \psi) - h^2 c(\phi, \psi) + h^{-\kappa} \sum_{\tau=0,T} e(U|_{t=\tau}, V|_{t=\tau}) \\ &\qquad + \gamma b(u, v) - b(\phi, \psi) - a(v, \phi) - a(u, \psi) \\ &\qquad + \gamma h^{-\kappa} \tilde{c}(\phi, \psi) + 2\gamma h^{-\kappa} \rho(v, \phi) + 2\gamma h^{-\kappa} \rho(u, \psi).\end{aligned}$$

We define the residual norm by

$$|||(u, \phi)|||^2 = h^{-\kappa}(S(u) + B(u)) + h^{\kappa}(S(\phi) + B(\phi)) + C(\phi) + h^{-\kappa} \sum_{\tau=0,T} E(U|_{t=\tau}).$$

This is indeed a norm on $V_h \times V_h$ as can be seen by following the proof of Lemma 2.2. Observe that in this case the vanishing boundary conditions on Γ are not imposed in the spaces V_h but follow if $B(u) = B(\phi) = 0$.

Lemma 4.3. *For all $u, \phi \in H^2(M) + V_h$ there holds*

$$|||(u, \phi)|||^2 \lesssim A[(u, \phi), (u, -\phi)].$$

Proof. By the definition of A, we have

$$A[(u, \phi), (u, -\phi)] = |||(u, \phi)|||^2 - \gamma h^{-\kappa} \widetilde{C}(\phi).$$

As $\kappa \le 0$, $\gamma < 1$ and $\chi \le 1$, $\gamma h^{-\kappa} \widetilde{C}(\phi)$ can be absorbed by $C(\phi)$. $\Box$

The lemma implies that the finite dimensional linear system (2.13) has a unique solution, and thus defines a finite element method.

4.3. Error estimates

Our main theorem in the case of boundary control is

Theorem 4.4. *Suppose that (A') holds. Let $\kappa \le 0$, $p, q \in \mathbb{N}^+$ and let (u_h, ϕ_h) in $V_h^p \times V_h^q$ be the solution of (2.13). Let $u \in H^{p+1}(M)$ and $\phi \in H^{q+1}(M)$ solve (4.2). Then*

$$|||(u - u_h, h^{-1}\phi - \phi_h)||| \lesssim h^{p+1-\frac{\kappa}{2}} \|u\|_{H^{p+1}(M)} + h^{q+1+\frac{\kappa}{2}} \|h^{-1}\phi\|_{H^{q+1}(M)}.$$

In particular,

$$\|\chi\partial_\nu(\phi - h\phi_h)\|_{L^2(\Gamma)} \lesssim h^{p-\frac{\kappa}{2}+\frac{1}{2}}\|u\|_{H^{p+1}(M)} + h^{q+\frac{\kappa}{2}-\frac{1}{2}}\|\phi\|_{H^{q+1}(M)}. \tag{4.11}$$

Theorem 4.4 follows essentially by repeating the proof of Theorem 2.4. We recall the main steps.

Equation (2.13) defines a finite element method that is consistent in the sense that if smooth enough u and ϕ satisfy (4.2), then (2.13) holds for $(u, h^{-1}\phi)$. This follows from the weak formulation (4.4) of (4.2) together with the regularization vanishing at $(u, h^{-1}\phi)$. This again leads to Galerkin orthogonality, that reads now

$$A[(u - u_h, h^{-1}\phi - \phi_h), (v, \psi)] = 0 \quad \text{for all } (v, \psi) \in V_h^p \times V_h^q,$$

for smooth enough u and ϕ satisfying (4.2) and $(u_h, \phi_h) \in V_h^p \times V_h^q$ solving (2.13).

It is straightforward to see that for all $u, \phi, v, \psi \in H^2(M) + V_h$ there holds

$$A[(u, \phi), (v, \psi)] + (a(v, \phi) + a(u, \psi)) \lesssim |||(u, \phi)|||\,|||(v, \psi)|||.$$

The analogue of Lemma 2.5 is

Lemma 4.5. *For all $u, \phi, v, \psi \in H^2(M) + V_h$ there holds*

$$a(v, \phi) \lesssim \left(S^{\frac{1}{2}}(\phi) + B(\phi)\right)\|v\|_{H^2(\mathcal{T}_h)},$$
$$a(u, \psi) \lesssim \left(S^{\frac{1}{2}}(u) + B(u) + \sum_{\tau=0,T} E^{\frac{1}{2}}(U|_{t=\tau})\right)\|\psi\|_{H^2(\mathcal{T}_h)}.$$

We omit the proof, this being a modification of the earlier proof. The only difference is that the boundary terms on Γ need to be kept track of. The proof of Theorem 4.4 uses Lemma 4.5 exactly as the proof of Theorem 2.4 uses Lemma 2.5. We will not repeat the details, however, let us point out that (4.9) replaces (2.19) in the proof.

As discussed above, if $(u_0, u_1) \in H^{k+1}(\Omega) \times H^k(\Omega)$ satisfies the compatibility conditions of order k, then the solution to (4.2) satisfies

$$(u, \phi) \in H^{k+1}(M) \times H^{k+2}(M).$$

Hence we can take $\kappa = 0$, $q \le k + 1$ and $p = q - 1$ in the above theorem, leading to the convergence rate (1.6) stated in the introduction.

Remark 4.6. A finite element method for the distributed control problem can be formulated using the spaces V_h^p defined by (4.8) and the bilinear form a in (4.5). With these choices replacing V_h^p and a in the Lagrangian in Section 2.2, and with $B(u) - B(\phi)$ added in the regularization $\mathcal{R}(u, \phi)$ there, we obtain a method satisfying the estimates in Theorem 2.4. This method works for smooth Ω whenever the geometric control condition (A) holds. We omit proving this, the proof being very similar with those above.

5. Numerical experiments

We discuss some numerical experiments performed with the Freefem++ package (see [27]).

We address the distributed and boundary case in the one dimensional case and emphasize the influence of the regularity of the initial condition on the rate of convergence of the finite element method with respect to the size of the discretization. We use uniform (unstructured) meshes and the functions $\chi_0 \in C_0^\infty([0, T])$ and

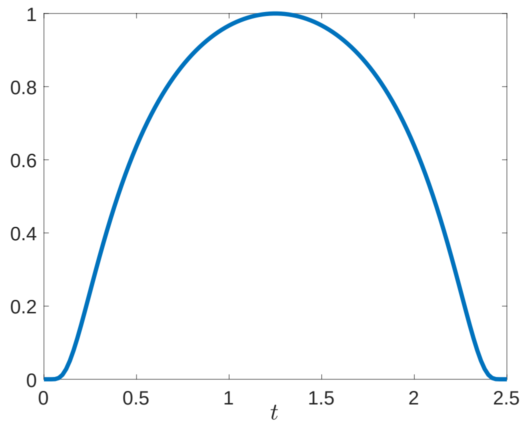

FIGURE 1. The $C_0^\infty([0,T])$ function $t \mapsto \chi_0(t)$, $t \in [0,T]$ with $T = 2.5$.

$\chi_1 \in C_0^\infty([0,1])$ such that

$$\chi_0(t) = \frac{e^{-\frac{1}{2t}} e^{-\frac{1}{2(T-t)}}}{e^{-\frac{1}{T}} e^{-\frac{1}{T}}} 1_{]0,T[}(t), \qquad \chi_1(x) = \frac{e^{-\frac{1}{5(x-a)}} e^{-\frac{1}{5(b-x)}}}{e^{-\frac{2}{5(b-a)}} e^{-\frac{2}{5(b-a)}}} 1_{]a,b[}(x) \tag{5.1}$$

for any $0 < a < b < 1$ and $T > 0$. In particular, $\chi_0(T/2) = 1$ and $\chi_1((a+b)/2) = 1$. Figure 1 depicts the function χ_0 for $T = 2.5$.

5.1. Distributed case: initial condition in $H^{k+1}(\Omega) \times H^k(\Omega)$ for all $k \in \mathbb{N}$

We consider the simplest situation for which

$$(u_0, u_1) = (\sin(\pi x), 0) \in H^{k+1}(\Omega) \times H^k(\Omega) \quad \forall k \in \mathbb{N}. \tag{Ex1}$$

Compatibility conditions (C1)–(C2) are satisfied for any j. Moreover, we use the cut-off functions $\chi_0 \in C_0^\infty([0,T])$ and $\chi_1 \in C_0^\infty([0,1])$ defined by (5.1) with $T = 2$, $a = 0.1$ and $b = 0.4$. The null controllability property (A) holds true for this set of data. Since explicit solutions are not available in the distributed case, we define as "exact" solution (u, ϕ) the one of (2.13) from a fine and structured mesh (composed of 409 000 triangles and 205 261 vertices) corresponding to $h \approx 4.41 \times 10^{-3}$ and $(u_h, \phi_h) \in V_h^p \times V_h^q$ with $(p,q) = (3,3)$.

Figure 2-left depicts the evolution of the relative error for the variable ϕ with respect to the L^2-norm

$$\mathrm{err}(\phi, \phi_h, \chi) := \|\chi(\phi - \phi_h)\|_{L^2(M)} / \|\chi\phi\|_{L^2(M)}$$

with respect to h for various pairs of (p,q). Table 1 collects the corresponding numerical values. We observe the convergence of the approximation w.r.t. h. Moreover, the figure exhibits the influence of the space V_h^q used for the variable ϕ_h while the choice of the space V_h^p for the variable u_h has no effect on the approximation. We observe rates close to 0.5, 2 and 3 for $(p,q) = (1,1)$, $(p,q) = (2,2)$ and $(p,q) = (3,3)$ respectively, in agreement with Theorem 2.4. For comparison, Figure 2-right depicts the evolution of the relative error $\mathrm{err}(\phi, \phi_h, \chi)$ for $\chi_0(t) = 1$ and $\chi_1(x) = 1_{(a,b)}(x)$, *i.e.* when no regularization of the control support is introduced. Table 2 collects the corresponding numerical values. The corresponding controlled pair (u, ϕ) is *a priori* only in $C([0,T]; H_0^1(\Omega)) \times C([0,T]; L^2(\Omega))$. Thus, if we still get the convergence with respect to the parameter h, we observe that the approximation is not improved beyond the value $q = 2$. As before, the choice of the approximation space V_h^p for u_h does not affect the result. The rate is also reduced: for $(p,q) = (2,2)$, the rate is close to 1.5. This highlights

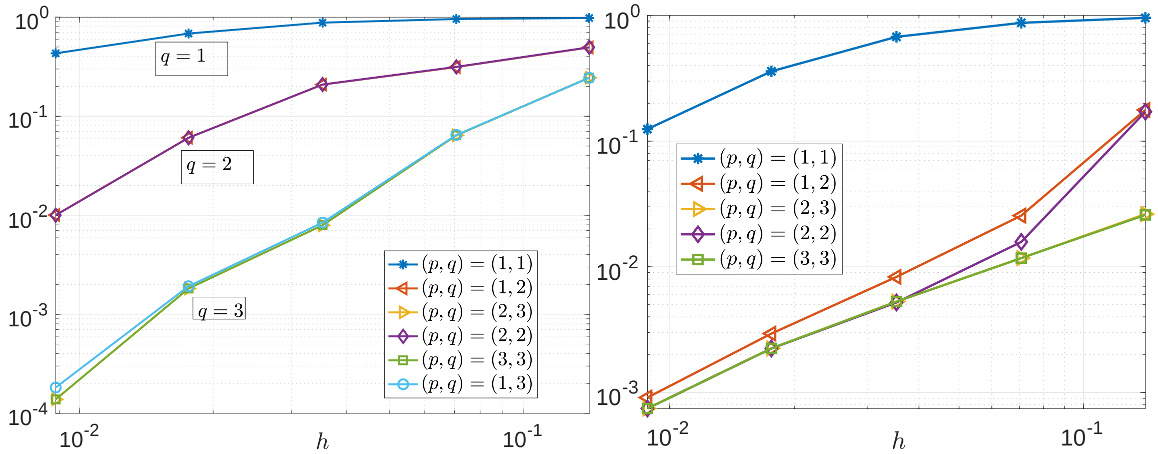

FIGURE 2. (**Ex1**); $\|\chi(\phi-\phi_h)\|_{L^2(M)}/\|\chi\phi\|_{L^2(M)}$ vs. h; *Left*: $\chi_0(t)$ and $\chi_1(x)$ given by (5.1); *Right*: $\chi_0(t)=1$ and $\chi_1(x)=1_{(a,b)}(x)$.

TABLE 1. (**Ex1**); $\|\chi(\phi_h-\phi)\|_{L^2(M)}/\|\chi\phi\|_{L^2(M)}$; χ from (5.1).

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$(p,q)=(1,1)$	9.81×10^{-1}	9.58×10^{-1}	8.81×10^{-1}	6.83×10^{-1}	4.31×10^{-1}
$(p,q)=(2,2)$	4.96×10^{-1}	3.15×10^{-1}	2.09×10^{-1}	6.05×10^{-2}	1.00×10^{-2}
$(p,q)=(3,3)$	2.45×10^{-1}	6.41×10^{-2}	7.93×10^{-3}	1.81×10^{-3}	1.38×10^{-4}

TABLE 2. (**Ex1**); $\|\chi(\phi_h-\phi)\|_{L^2(M)}/\|\chi\phi\|_{L^2(M)}$; $\chi_0(t)=1$; $\chi_1(x)=1_{(a,b)}(x)$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$(p,q)=(1,1)$	9.55×10^{-1}	8.71×10^{-1}	6.74×10^{-1}	3.58×10^{-1}	1.24×10^{-1}
$(p,q)=(2,2)$	1.71×10^{-1}	1.57×10^{-2}	5.20×10^{-3}	2.24×10^{-3}	7.47×10^{-4}
$(p,q)=(3,3)$	2.58×10^{-2}	1.17×10^{-3}	5.28×10^{-3}	2.24×10^{-3}	7.47×10^{-4}

TABLE 3. (**Ex1**); $(p,q)=(2,2)$ and χ from (5.1).

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi,\phi_h,1)$	6.22×10^{-1}	3.92×10^{-1}	2.33×10^{-1}	6.45×10^{-2}	1.04×10^{-2}
$\mathrm{err}(\partial_x\phi,\partial_x\phi_h,1)$	8.11×10^{-1}	6.84×10^{-1}	4.85×10^{-1}	1.64×10^{-1}	4.57×10^{-2}
$\mathrm{err}(u,u_h,1)$	4.29×10^{-1}	1.36×10^{-1}	5.00×10^{-2}	1.08×10^{-2}	1.10×10^{-3}
$\mathrm{err}(\phi,\phi_h,\chi)$	4.96×10^{-1}	3.15×10^{-1}	2.09×10^{-1}	6.05×10^{-2}	1.00×10^{-2}

the influence of the cut off functions, including for very smooth initial conditions. Table 3 collects some L^2 norms of u_h and ϕ_h with respect to h for the pair $(p,q)=(2,2)$: in particular, the relative error $err(u,u_h,1)$ associated with the controlled solution u is order of $h^{2.5}$ for h small enough.

5.2. Distributed case: initial condition in $H^1(\Omega)\times L^2(\Omega)$

We consider the initial condition

$$(u_0,u_1)=\big(4x1_{(0,1/2)}(x)+4(1-x)1_{[1/2,1)}(x),0\big)\in H^1(\Omega)\times H^0(\Omega) \qquad \textbf{(Ex2)}$$

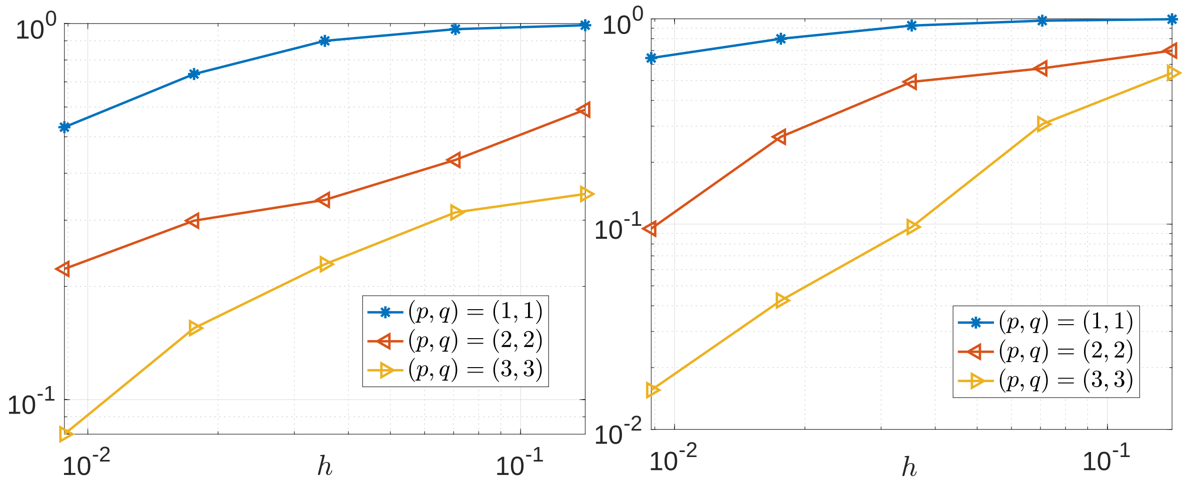

FIGURE 3. (**Ex2**) and (**Ex2b**); $\|\chi(\phi-\phi_h)\|_{L^2(M)}/\|\chi\phi\|_{L^2(M)}$ vs. h; χ from (5.1).

TABLE 4. (**Ex2**); $\|\chi(\phi-\phi_h)\|_{L^2(M)}/\|\chi\phi\|_{L^2(M)}$; χ from (5.1).

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$(p,q)=(1,1)$	9.89×10^{-1}	9.66×10^{-1}	8.99×10^{-1}	7.34×10^{-1}	5.30×10^{-1}
$(p,q)=(2,2)$	5.90×10^{-1}	4.31×10^{-1}	3.38×10^{-1}	2.98×10^{-1}	2.22×10^{-1}
$(p,q)=(3,3)$	3.51×10^{-1}	3.14×10^{-1}	2.28×10^{-1}	1.54×10^{-1}	8.08×10^{-2}

for which the compatibility conditions (C1)-(C2) are satisfied for any j. If the cut-off functions are introduced, the controlled pair (u,ϕ) belongs to $H^1(M)\times H^0(M)$. Theorem 2.4 does not provide a convergence rate in this case. The strong convergence is however observed: Figure 3 displays the relative error w.r.t. h for $(p,q)=(1,1)$, $(p,q)=(2,2)$ and $(p,q)=(3,3)$ with rates close to 1/2. Table 4 collects the corresponding numerical values.

A similar behavior is observed with the condition $u_0=0$ and $u_1=1_{(0.4,0.6)}(x)$ in $L^2(\Omega)$ in agreement with Theorem 3.1

5.3. Distributed case : initial condition in $H^2(\Omega)\times H^1(\Omega)$

We consider the initial condition

$$(u_0,u_1)=\left(\rho(x)\int_0^x u_0^{(2)}(t)dt,0\right)\in H^2(\Omega)\times H^1(\Omega) \tag{Ex2b}$$

where $u_0^{(2)}$ is the initial position defined in (**Ex2**) and $\rho\in C_0^\infty(\Omega)$ is introduced in order to preserve the compatibility conditions (C1)-(C2). Figure 3-right displays the convergence of the approximation for $(p,q)=(1,1)$, $(p,q)=(2,2)$ and $(p,q)=(3,3)$. Theorem 2.4 still does not provide a convergence rate in this case. However, with respect to the previous example, smaller relative error with rates close to 1 are observed for $(p,q)=(2,2)$ and $(p,q)=(3,3)$.

5.4. Boundary case: initial condition in $H^{k+1}(\Omega)\times H^k(\Omega)$ for all $k\in\mathbb{N}$

We consider again the simple situation given by the initial condition (**Ex1**). Compatibility conditions (C1)-(C2) are satisfied for any j. In contrast with the distributed case, explicit exact solutions are available in the boundary case when cut-off functions are not introduced. Precisely, the corresponding control of minimal $L^2(\Gamma)$

TABLE 5. (**Ex1**) - Boundary case - $(p,q)=(1,2)$ - $\chi\equiv 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi,\phi_h,1)$	7.17×10^{-1}	1.97×10^{-1}	4.54×10^{-2}	1.53×10^{-2}	4.50×10^{-3}
$\mathrm{err}(\partial_x\phi,\partial_x\phi_h,1)$	9.21×10^{-1}	3.05×10^{-1}	9.40×10^{-2}	5.41×10^{-2}	2.47×10^{-2}
$\mathrm{err}(u,u_h,1)$	1.88×10^{-1}	5.73×10^{-2}	1.92×10^{-2}	9.29×10^{-3}	4.09×10^{-3}
$\mathrm{err}(v,u_h)$	2.16×10^{-1}	5.83×10^{-2}	1.94×10^{-2}	7.21×10^{-3}	3.10×10^{-3}
$\mathrm{err}(v,h\chi\partial_x\phi_h)$	2.71×10^{-1}	8.41×10^{-2}	2.79×10^{-2}	1.22×10^{-2}	5.31×10^{-3}

TABLE 6. (**Ex1**) - Boundary case - $(p,q)=(2,3)$ - $\chi\equiv 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi,\phi_h,1)$	4.09×10^{-1}	1.15×10^{-1}	2.82×10^{-2}	9.29×10^{-3}	2.81×10^{-3}
$\mathrm{err}(\partial_x\phi,\partial_x\phi_h,1)$	6.73×10^{-1}	2.10×10^{-1}	6.13×10^{-2}	2.91×10^{-2}	1.24×10^{-2}
$\mathrm{err}(u,u_h,1)$	7.29×10^{-2}	2.81×10^{-2}	9.02×10^{-3}	3.45×10^{-3}	1.28×10^{-3}
$\mathrm{err}(v,u_h)$	1.00×10^{-1}	3.74×10^{-2}	1.13×10^{-2}	4.16×10^{-3}	1.43×10^{-3}
$\mathrm{err}(v,h\chi\partial_x\phi_h)$	1.05×10^{-1}	3.89×10^{-2}	1.23×10^{-2}	4.61×10^{-3}	1.65×10^{-3}

norm with $\Gamma=(0,T=2)\times\{1\}$ is given by $v(t)=\frac{1}{2}\sin(\pi(1-t))=\frac{1}{2}\sin(\pi t)$ leading to $\|v\|_{L^2(0,T)}=1/2$. The corresponding controlled solution is given by

$$u(t,x)=\begin{cases}\frac{1}{2}\big(u_0(x+t)+u_0(x-t)\big) & x+t\le 1,\\ \frac{1}{2}u_0(x-t), & x+t>1,\ x-t\ge -1,\\ 0, & x-t<-1,\end{cases}\tag{5.2}$$

leading to $\|u\|_{L^2(M)}=1/2$. The corresponding adjoint solution is given by $\phi(t,x)=-\frac{1}{2\pi}\sin(\pi t)\sin(\pi x)$ leading to $\|\phi\|_{L^2(M)}=\frac{1}{2\sqrt{2}\pi}$ and $\|\partial_x\phi\|_{L^2(M)}=\frac{1}{2\sqrt{2}}$.

Tables 5 and 6 collects some relative errors w.r.t. h for $(p,q)=(1,2)$ and $(p,q)=(2,3)$ respectively including

$$\mathrm{err}(v,u_h):=\|v-u_h(\cdot,1)\|_{L^2(0,T)}/\|v\|_{L^2(0,T)},$$

and $\mathrm{err}(v,h\chi\partial_\nu\phi_h)$ while Figure 4-left depicts the relative error on the control w.r.t. h for several pairs of (p,q). Since compatibility conditions hold true here, the introduction of the cut off function $\chi\neq 1$ is *a priori* useless. However, we observe that the term $\partial_x\phi(\cdot,1)$ is not well approximated near $t=0$ and $t=T$. This somehow pollutes the approximation ϕ_h of ϕ inside the domain (precisely along the characteristics intersecting the points $(t,x)=(0,1)$ and $(t,x)=(T,1)$) and affects the optimal rate. We observe rate close to 0.75 for $(p,q)=(1,1)$ and close to 1.5 otherwise. Imposing in addition $\phi=0$ on the boundary $\partial\Omega$ slightly improves the approximation.

With $\chi\neq 1$, the explicit control of minimal $L^2(\chi^{1/2},(0,T))$ norm is not available anymore. As for the distributed controllability, we define as “exact” control the one obtained from a fine and uniform mesh (composed of 648 000 triangles and 325 261 vertices) corresponding to $h\approx 3.92\times10^{-3}$ and $(p,q)=(2,3)$. We take T large enough, precisely $T=2.5$, to ensure the null controllability property of the wave equation. Figure 4-right displays the evolution of $\|h\chi\partial_x\phi_h(\cdot,1)-v\|_{L^2(0,T)}/\|v\|_{L^2(0,T)}$ w.r.t. h. For $(p,q)=(1,1)$ and $(p,q)=(1,2)$, we observe rates close to 0.5 and 1.5 in agreement with (4.11) of Theorem 4.4 with $\kappa=0$. For $(p,q)=(2,3)$, we observe a rate close to 3, which is a bit better than the value 2.5 from (4.11). Those results also show that the boundary control can be approximated both from the quantity $h\partial_x\phi_h(\cdot,1)$ obtained from the adjoint dual variable and from the trace $u_h(\cdot,1)$ of the primal variable.

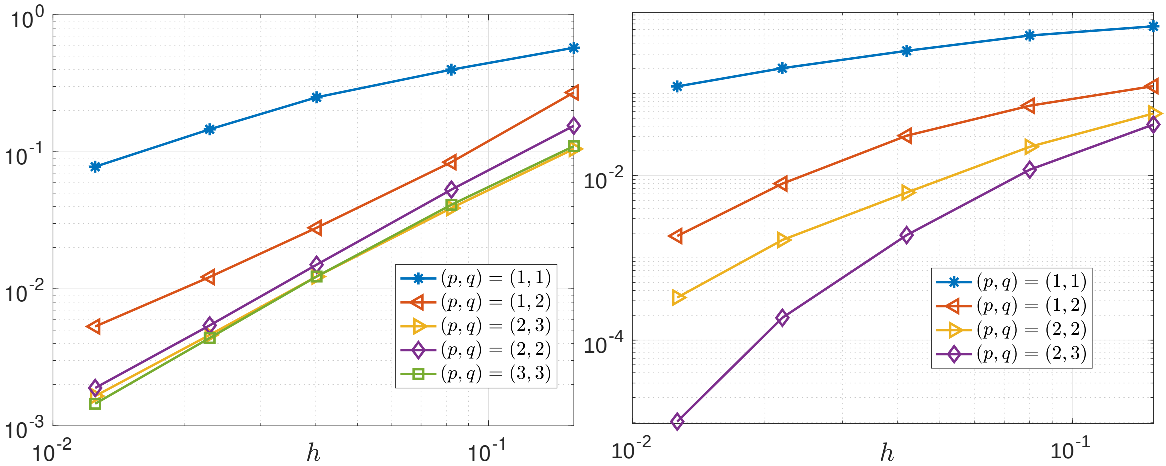

FIGURE 4. (**Ex1**) - Boundary case - $\|h\chi\partial_x\phi_h(1,\cdot)-v\|_{L^2(0,T)}/\|v\|_{L^2(0,T)}$ *vs.* h with $\chi=1$ (*left*) and $\chi=\chi_0$ from (5.1) (*right*).

5.5. Boundary case: initial condition in $H^1(\Omega)\times H^0(\Omega)$

We consider the initial data (**Ex2**). The corresponding control of minimal $L^2(\Gamma)$ norm with $\Gamma=(0,T=2)\times\{1\}$ (corresponding to $\chi\equiv 1$) is given by

$$v(t)=2t\,1_{(0,1/2)}(t)+2(1-t)1_{(1/2,3/2)}(t)+2(t-2)\,1_{(3/2,1)}(t).$$

The corresponding controlled solution is explicitly known as follows:

$$u(t,x)=\frac{1}{2}u_0(x-t)1_{(t-x\le 0)}+\frac{1}{2}u_0(x+t)1_{(t+x\le 1)}-\frac{1}{2}u_0(t-x)1_{(t-x\ge 0)}1_{(t-x\le 1)}. \tag{5.3}$$

The corresponding adjoint solution is given by $\phi(t,x)+\phi(t,1-x)$ where ϕ is defined in (5.5) with the following initial conditions

$$(\phi_0,\phi_1)=\big(0,-2x\,1_{(0,1/2)}(x)+2(x-1)\,1_{(1/2,1)}(x)\big)\in H^2(\Omega)\times H^1(\Omega).$$

Compatibility conditions are satisfied.

Figure 5-left displays the evolution of $\|\chi\partial_x\phi_h(1,\cdot)-v\|_{L^2(0,T)}/\|v\|_{L^2(0,T)}$ w.r.t. h for various pairs of (p,q) and $\chi\equiv 1$. We observe a rate close to 0.75 for $(p,q)=(1,1)$ and close to 1.5 otherwise. Remark that *a priori* $u\in H^1(M)$ and $\phi\in H^2(M)$ so that the choice $(p,q)=(2,3)$ does not lead to a better rate than the choice $(p,q)=(1,2)$. See also Table 7 and Table 8 for the corresponding values for (p,q) = (1,2) and (p,q) = (2,3) respectively. Moreover, as expected, the introduction of the cut off $\chi\neq 1$ does not improve here the rate of convergence: see Figure 5-right where similar rates are observed.

5.6. Boundary case: initial condition in $H^0(\Omega)\times H^{-1}(\Omega)$

We consider the following stiff situation given by

$$(u_0,u_1)=(4x1_{(0,1/2)}(x),0)\in H^0(\Omega)\times H^{-1}(\Omega) \tag{\textbf{Ex3}}$$

TABLE 7. (**Ex2**) - Boundary case - $(p, q) = (1, 2)$; $\chi \equiv 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi, \phi_h, 1)$	7.83×10^{-1}	2.53×10^{-1}	5.82×10^{-2}	1.97×10^{-2}	5.64×10^{-3}
$\mathrm{err}(\partial_x\phi, \partial_x\phi_h, 1)$	1.12×10^{0}	5.03×10^{-1}	2.02×10^{-1}	1.12×10^{-1}	5.04×10^{-2}
$\mathrm{err}(u, u_h, 1)$	1.91×10^{-1}	4.94×10^{-2}	2.42×10^{-2}	1.17×10^{-2}	5.04×10^{-3}
$\mathrm{err}(v, u_h)$	2.20×10^{-1}	6.48×10^{-2}	2.87×10^{-2}	1.16×10^{-2}	4.95×10^{-3}
$\mathrm{err}(v, h\chi\partial_x\phi_h)$	2.49×10^{-1}	6.41×10^{-2}	2.97×10^{-2}	1.40×10^{-2}	6.06×10^{-3}

TABLE 8. (**Ex2**) - Boundary case - $(p, q) = (2, 3)$; $\chi \equiv 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi, \phi_h, 1)$	4.90×10^{-1}	1.35×10^{-1}	3.28×10^{-2}	1.07×10^{-2}	3.41×10^{-3}
$\mathrm{err}(\partial_x\phi, \partial_x\phi_h, 1)$	9.79×10^{-1}	3.75×10^{-1}	1.35×10^{-1}	6.76×10^{-2}	3.21×10^{-2}
$\mathrm{err}(u, u_h, 1)$	6.70×10^{-2}	2.60×10^{-2}	9.34×10^{-3}	3.78×10^{-3}	1.44×10^{-3}
$\mathrm{err}(v, u_h)$	8.61×10^{-2}	3.25×10^{-2}	1.10×10^{-2}	4.19×10^{-3}	1.59×10^{-3}
$\mathrm{err}(v, h\chi\partial_x\phi_h)$	8.57×10^{-2}	3.23×10^{-2}	1.10×10^{-2}	4.45×10^{-3}	1.71×10^{-3}

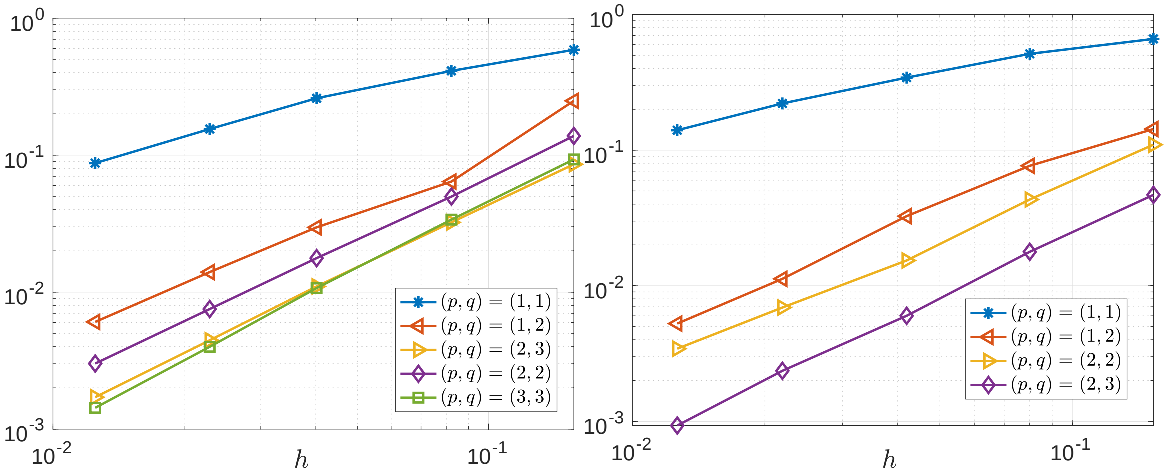

FIGURE 5. (**Ex2**) - Boundary case - $\|h\partial_x\chi\phi_h(1,\cdot) - v\|_{L^2(0,T)}/\|v\|_{L^2(0,T)}$ vs. h with $\chi = 1$ (*left*) and $\chi = \chi_0$ from (5.1) (*right*).

and extensively discussed in [15, 43] and $T = 2$. The corresponding control of minimal $L^2((0,T) \times \{1\})$ norm is given by $v(t) = 2(1-t)1_{(1/2,3/2)}(t)$ leading to $\|v\|_{L^2(0,T)} = 1/\sqrt{3}$. The corresponding controlled solution is explicitly known as follows:

$$u(t,x) = \left\{ \begin{array}{ccc} 4x & 0 \le x + t < \frac{1}{2}, & \\ 2(x-t) & -\frac{1}{2} < t - x < \frac{1}{2}, & x + t \ge \frac{1}{2}, \\ 0 & else, & \end{array} \right. \tag{5.4}$$

TABLE 9. (**Ex3**) - Boundary case - $(p, q) = (1, 1)$ - $\chi \equiv 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi, \phi_h, 1)$	6.38×10^{-1}	4.35×10^{-1}	2.85×10^{-1}	1.84×10^{-1}	1.05×10^{-1}
$\mathrm{err}(\partial_x\phi, \partial_x\phi_h, 1)$	8.38×10^{-1}	6.23×10^{-1}	4.85×10^{-1}	3.97×10^{-1}	3.17×10^{-1}
$\mathrm{err}(u, u_h, 1)$	5.78×10^{-1}	4.40×10^{-1}	3.59×10^{-1}	2.93×10^{-1}	2.27×10^{-1}
$\mathrm{err}(v, u_h)$	7.67×10^{-1}	5.69×10^{-1}	5.04×10^{-1}	4.05×10^{-1}	3.15×10^{-1}
$\mathrm{err}(v, h\chi\partial_x\phi_h)$	8.41×10^{-1}	6.47×10^{-1}	5.08×10^{-1}	4.09×10^{-1}	3.16×10^{-1}

TABLE 10. (**Ex3**) - Boundary case - $(p, q) = (1, 2)$ - $\chi \equiv 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi, \phi_h, 1)$	1.62×10^{0}	6.33×10^{-1}	2.72×10^{-1}	1.45×10^{-1}	7.36×10^{-2}
$\mathrm{err}(\partial_x\phi, \partial_x\phi_h, 1)$	2.33×10^{0}	1.52×10^{0}	1.22×10^{0}	1.08×10^{0}	1.05×10^{0}
$\mathrm{err}(u, u_h, 1)$	3.93×10^{-1}	3.00×10^{-1}	2.27×10^{-1}	1.74×10^{-1}	1.30×10^{-1}
$\mathrm{err}(v, u_h)$	5.03×10^{-1}	3.43×10^{-1}	2.41×10^{-1}	1.89×10^{-1}	1.48×10^{-1}
$\mathrm{err}(v, h\chi\partial_x\phi_h)$	4.73×10^{-1}	3.41×10^{-1}	2.64×10^{-1}	2.08×10^{-1}	1.60×10^{-1}

TABLE 11. (**Ex3**) - Boundary case - $(p, q) = (2, 3)$ - $\chi \equiv 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
$\mathrm{err}(\phi, \phi_h, 1)$	1.25×10^{0}	5.01×10^{-1}	2.02×10^{-1}	9.58×10^{-2}	4.52×10^{-2}
$\mathrm{err}(\partial_x\phi, \partial_x\phi_h, 1)$	3.44×10^{0}	2.34×10^{0}	1.88×10^{0}	1.57×10^{0}	1.45×10^{0}
$\mathrm{err}(u, u_h, 1)$	3.12×10^{-1}	2.22×10^{-1}	1.60×10^{-1}	1.24×10^{-1}	9.01×10^{-2}
$\mathrm{err}(v, u_h)$	3.52×10^{-1}	2.36×10^{-1}	1.71×10^{-1}	1.30×10^{-1}	9.64×10^{-2}
$\mathrm{err}(v, h\chi\partial_x\phi_h)$	3.40×10^{-1}	2.36×10^{-1}	1.66×10^{-1}	1.37×10^{-1}	9.56×10^{-2}

leading to $\|u\|_{L^2(M)} = 1/\sqrt{3}$. The corresponding initial conditions of the adjoint solution is $(\phi_0, \phi_1) = (0, -2x\, 1_{(0,1/2)}(x)) \in H^1(\Omega) \times H^0(\Omega)$ leading to

$$\phi(t,x) = \begin{cases} -2xt & 0 \le x+t < \frac{1}{2}, \quad x \ge 0, t \ge 0, \\ \frac{(x-t)^2}{2} - \frac{1}{8} & \frac{1}{2} \le x+t < \frac{3}{2}, \quad -\frac{1}{2} < x-t < \frac{1}{2}, \\ 2(x-1)(1-t) & \frac{3}{2} \le x+t, \quad -\frac{1}{2} < x-t, \\ -\frac{(x+t-2)^2}{2} + \frac{1}{8} & \frac{3}{2} < x+t < \frac{5}{2}, \quad -\frac{3}{2} < x-t \le -\frac{1}{2}, \\ 2x(2-t) & x-t \le -\frac{3}{2}, \end{cases} \tag{5.5}$$

leading to $\|\phi\|_{L^2(M)} \approx 9.86 \times 10^{2}$ and $\|\partial_x\phi\|_{L^2(M)} \approx 4.08 \times 10^{-1}$. In particular, we check that $\partial_x\phi(t,x)_{|x=1} = 2(1-t)\,1_{(1/2,3/2)}(t) = v(t)$. Both u and ϕ develop singularities (where u and $\nabla\phi$ are discontinuous).

Figure 6 depicts the evolution of $\|h\chi\partial_x\phi_h(\cdot,1) - v\|_{L^2(0,T)}/\|v\|_{L^2(0,T)}$ w.r.t. h with $\chi \equiv 1$ for several values of (p,q). See also Table 9,10 and 11 for $(p,q) = 1$, $(p,q) = 2$ and $(p,q) = (3,3)$ respectively. We observe a rate close to 0.5.

Let us also emphasize that the space-time discretization formulation is very well appropriated for mesh adaptivity. Using the $V_h^1 \times V_h^2$ approximation, Figure 7-left (resp. right) depicts the mesh obtained after seven adaptative refinements based on the local values of gradient of the variable ϕ_h (resp. u_h). Starting with a coarse mesh composed of 288 triangles and 166 vertices, the final mesh on the right is composed with 13068 triangles and 6700 vertices and leads to a relative error $\mathrm{err}(v, u_h)$ of the order of 10^{-3}. The final mesh follows the singularities of the controlled solution starting at the point $(0, 1/2)$ of discontinuity of u_0.

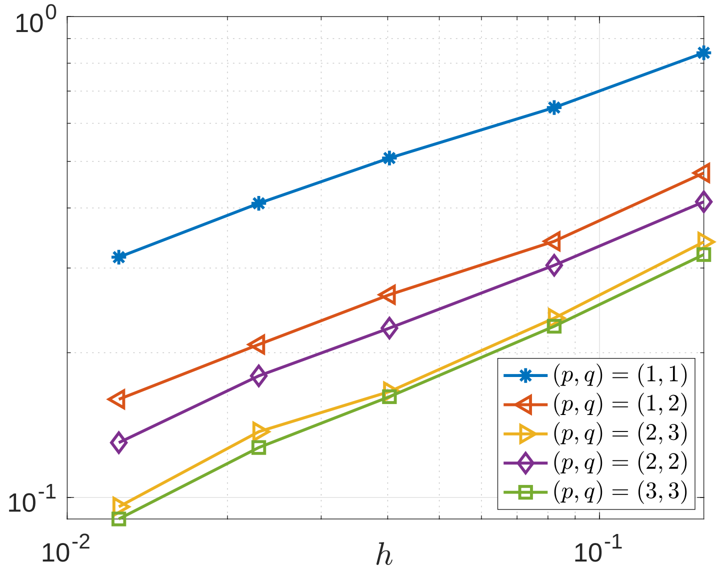

FIGURE 6. (**Ex3**); $\|h\chi\partial_x\phi_h(\cdot,1) - v\|_{L^2(0,T)}/\|v\|_{L^2(0,T)}$ w.r.t. h (rate ≈ 0.5).

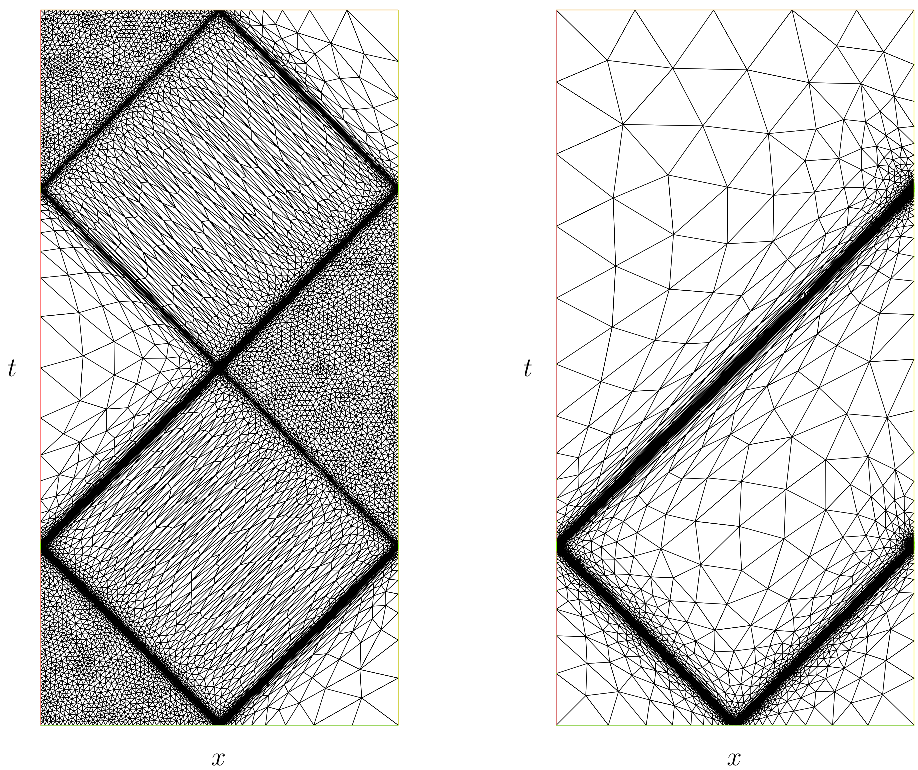

FIGURE 7. (**Ex3**); Locally refine space-time meshes with respect to ϕ_h (left) and u_h (right). $(p,q) = (1,2)$.

TABLE 12. (**Ex1**) - Boundary case - $(p,q) = (1,2)$ - χ from (5.1); $\|h\chi\partial_x\phi_h(\cdot,1) - v\|_{L^2(0,T)}/\|v\|_{L^2(0,T)}$ w.r.t. h and $V \in \{-10,-20,-30,-40\}$.

h	1.6×10^{-1}	8×10^{-2}	4×10^{-2}	2×10^{-2}	1×10^{-2}
$V=-10$	7.11×10^{-1}	2.43×10^{-1}	6.91×10^{-2}	2.48×10^{-2}	7.97×10^{-3}
$V=-20$	7.06×10^{-1}	2.81×10^{-1}	1.57×10^{-1}	6.13×10^{-2}	1.65×10^{-2}
$V=-30$	9.54×10^{-1}	6.91×10^{-1}	2.07×10^{-1}	7.21×10^{-2}	2.27×10^{-2}
$V=-40$	1.01×10^{-1}	9.44×10^{-1}	5.05×10^{-1}	1.13×10^{-1}	3.18×10^{-2}

5.7. Boundary case: the wave equation with a potential

To end these numerical illustrations, we report some results for the wave equation with non vanishing potential V, see (4.1). Non zero potentials notably appear from linearization of nonlinear wave equations of the form $\Box u + f(u) = \chi v$ (see [44]). Actually, we want to emphasize that this spacetime approach, based on the resolution of the optimal condition associated with the control of minimal L^2 norm is very relevant for potential with the "bad" sign for which $V(t,x)u(t,x) < 0$. Indeed, in this case, the usual "à la Glowinski" strategy developed in [25] is numerically inefficient and requires adaptations, since the uncontrolled solution (used to initialize the conjugate algorithm) grows exponentially in time, leading to numerical instabilities and overflow. Recall that the observability constant behaves like $e^{C(T,\omega)\|A\|^2_{L^\infty(0,T;L^n(\Omega))}}$ (see [49]) and appears notably in the constant in the *a priori* estimate (4.11). We consider the initial condition (**Ex1**), $T = 2.5$ and constant negative potentials $V(t,x) = V < 0$. Table 12 collects the relative error on the approximation of the boundary control with respect to h for several negatives values of V. In particular, for $V = -40$, the L^∞ norm of the corresponding uncontrolled solution is of order 10^5. We approximate u and ϕ in V_h^1 and V_h^2 respectively and observe a rate close to 1.5. The value of V only affects the constant. We refer to [40] for a semi-discrete (in space) approximation of exact boundary controls for a semi-discretized wave equation with potential, including experiments for small and potentials with good sign.

6. CONCLUSION

We have introduced and analyzed a spacetime finite element approximation of a controllability problem for the wave equation. Based on a non conformal H^1-approximation, the analysis yields error estimates for the control in the natural L^2-norm of order h^q (resp. $h^{q-\frac{1}{2}}$) where q is the degree of the polynomials used to describe the adjoint variable in the distributed (resp. boundary) case. The numerical experiments performed for initial data with various regularity exhibits the efficiency of the method. The convergence is also observed for initial data with minimal regularity.

We emphasize that spacetime formulations are easier to implement than time-marching methods, since in particular, there is no kind of CFL condition between the time and space discretization parameters. Moreover, as shown in the numerical section, they are well-suited for mesh adaptivity (as initially discussed in [29]).

Similarly to the formulation proposed in [13, 15], the present formulation follows the "control then discretize" approach. However, contrary to [13, 15], the H^1-formulation of the present work does not require the introduction of sophisticated finite element spaces. On the other hand, the formulation requires additional stabilized terms which are function of the jump of the gradient across the boundary of each element. The analysis is then inspired from [10] and also from [8] where an analogous spacetime formulation for a data assimilation is considered.

The implementation of the stabilized terms is not straightforward, in particular, in higher dimension, and is usually not available in finite element softwares. A possible way to circumvent the introduction of the gradient jump terms is to consider non-conforming approximation of the Crouzeix-Raviart type as in [7]. A penalty is then needed on the solution jump instead to control the H^1-conformity error. Another possible way, following [42] devoted to the boundary case, could be to consider the controllability problem associated to a first order reformulation of the wave equation:

$$\begin{cases} v_t - div\,\mathbf{p} = 0, \\ \mathbf{p}_t - \nabla v = 0. \end{cases} \tag{6.1}$$

with $v := u_t$ and $\mathbf{p} := \nabla u$. A H^1 conformal stabilized approximation is employed in [42] leading to promising numerical experiments in the one dimensional case. A rigorous numerical analysis however remains to be done.

Appendix A. Continuum estimates

Proposition A.1 (Energy estimate)**.** *There holds*

$$\begin{aligned} &\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{-1}(\Omega))} + \|\partial_\nu u\|_{H^{-1}((0,T)\times\partial\Omega)} \\ &\quad\lesssim \|u|_{t=0}\|_{L^2(\Omega)} + \|\partial_t u|_{t=0}\|_{H^{-1}(\Omega)} + \|u\|_{L^2((0,T)\times\partial\Omega)} + \|\Box u\|_{H^{-1}((0,T)\times\Omega)}. \end{aligned} \tag{A.1}$$

Proof. The estimate

$$\begin{aligned} &\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;H^{-1}(\Omega))} + \|\partial_\nu u\|_{H^{-1}((0,T)\times\partial\Omega)} \\ &\quad\lesssim \|u|_{t=0}\|_{L^2(\Omega)} + \|\partial_t u|_{t=0}\|_{H^{-1}(\Omega)} + \|u\|_{L^2((0,T)\times\partial\Omega)} + \|\Box u\|_{L^1(0,T;H^{-1}(\Omega))} \end{aligned} \tag{A.2}$$

follows from Theorem 2.3 of [33], see also Remark 2.2 there. Thus it is enough to consider the equation

$$\begin{cases} \Box u = f, \\ u|_{x\in\partial\Omega} = 0, \\ u|_{t<0} = 0, \end{cases} \tag{A.3}$$

and show that its solution satisfies

$$\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{-1}(\Omega))} + \|\partial_\nu u\|_{H^{-1}((0,T)\times\partial\Omega)} \lesssim \|f\|_{H^{-1}((0,T)\times\Omega)}.$$

Let us use the shorthand notations $M = (0,T)\times\Omega$ and $(t,x) = (x^0,\dots,x^n)$. We recall that for any $f \in H^{-1}(M)$ there are $f_j \in L^2(M)$, $j = -1, 0, \dots, n$, such that

$$f = f_{-1} + \sum_{j=0}^{n} \partial_{x^j} f_j, \tag{A.4}$$

and that

$$\|f\|^2_{H^{-1}(M)} = \inf \sum_{j=-1}^{n} \|f_j\|^2_{L^2(M)},$$

where the infimum is taken over all $f_j \in L^2(M)$ satisfying (A.4), see *e.g.* Theorem 1, page 299 of [22]. Thus it is enough to show that

$$\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u\|_{L^2(0,T;L^2(\Omega))} + \|\partial_\nu u\|_{H^{-1}((0,T)\times\partial\Omega)} \lesssim \|f_j\|_{L^2((0,T)\times\Omega)},$$

where u satisfies (A.3) with f replaced by $\partial_{x^j} f_j$ when $j \ge 0$ and by f_{-1} when $j = -1$. The cases $j = -1$ and $j > 0$ are contained in (A.2).

Let us consider the case $j = 0$. We denote by v the solution of (A.3) with $f = f_0$. Then $u = \partial_t v$ and it follows from Theorem 2.1 of [33] that

$$\begin{aligned}\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_\nu u\|_{H^{-1}((0,T)\times\partial\Omega)} &\lesssim \|\partial_t v\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_\nu v\|_{L^2((0,T)\times\partial\Omega)} \\ &\lesssim \|f_0\|_{L^2((0,T)\times\Omega)}.\end{aligned}$$

Moreover,

$$\|\partial_t u\|_{L^2(0,T;H^{-1}(\Omega))} = \|\partial_t^2 v\|_{L^2(0,T;H^{-1}(\Omega))} \le \|\Delta v\|_{L^2(0,T;H^{-1}(\Omega))} + \|f_0\|_{L^2(0,T;H^{-1}(\Omega))},$$

and using Theorem 2.1 of [33] again,

$$\|\Delta v\|_{L^2(0,T;H^{-1}(\Omega))} \lesssim \|v\|_{L^2(0,T;H^1(\Omega))} \lesssim \|f_0\|_{L^2((0,T)\times\Omega)}.$$

□

Remark A.2. It is not possible to improve the $L^2(0,T;H^{-1}(\Omega))$ norm of $\partial_t u$ on the left-hand side of (A.1) to its $L^\infty(0,T;H^{-1}(\Omega))$ norm.

Proof. To get a contradiction, we suppose that

$$\|\partial_t u\|_{L^\infty(0,T;H^{-1}(\Omega))} \lesssim \|\Box u\|_{H^{-1}((0,T)\times\Omega)}$$

for solutions u of (A.3). Let $f \in L^2(0,T)$, and denote by u and v the solutions of (A.3) with sources $\partial_t f$ and f, respectively. Then $u = \partial_t v$ and

$$f = \partial_t^2 v - \Delta v = \partial_t u - \Delta v \in L^\infty(0,T;H^{-1}(\Omega)).$$

But this implies that $f \in L^\infty(0,T)$. As $f \in L^2(0,T)$ was arbitrary, we get the contradiction $L^2(0,T) \subset L^\infty(0,T)$. □

If $\Box u = 0$ and $u|_{x\in\partial\Omega} = 0$ then the norm on the left-hand side of (A.1) controls the $L^\infty(0,T;H^{-1}(\Omega))$ norm of $\partial_t u$. In fact, we have:

Lemma A.3. *Suppose that $u \in L^2((0,T)\times\Omega)$ satisfies $\Box u = 0$ and $u|_{x\in\partial\Omega} = 0$. Then*

$$\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;H^{-1}(\Omega))} + \|\partial_\nu u\|_{H^{-1}((0,T)\times\partial\Omega)} \lesssim \|u\|_{L^2((0,T)\times\Omega)}.$$

Moreover, $u \in C(0,T;L^2(\Omega)) \cap C^1(0,T;H^{-1}(\Omega))$.

Proof. Let $\chi \in C^\infty(\mathbb{R})$ satisfy $\chi(t) = 0$ near $t = T$ and $\chi(t) = 1$ for $t \in (0,T/2)$. Applying Proposition A.1 to χu backwards in time, we obtain

$$\|u\|_{L^\infty(0,T/2;L^2(\Omega))} + \|\partial_t u\|_{L^2(0,T/2;H^{-1}(\Omega))} \lesssim \|[\Box,\chi]u\|_{H^1((0,T)\times\Omega)} \lesssim \|u\|_{L^2((0,T)\times\Omega)}.$$

Applying (A.2) backwards in time on the interval $(0,s)$ where $s < T/2$, we get

$$\|\partial_t u|_{t=0}\|^2_{H^{-1}(\Omega)} \lesssim \|u|_{t=s}\|^2_{L^2(\Omega)} + \|\partial_t u|_{t=s}\|^2_{H^{-1}(\Omega)}.$$

Integration in s gives

$$\|\partial_t u|_{t=0}\|^2_{H^{-1}(\Omega)} \lesssim \|u\|^2_{L^2(0,T/2;L^2(\Omega))} + \|\partial_t u\|^2_{L^2(0,T/2;H^{-1}(\Omega))}.$$

We conclude that

$$\|u|_{t=0}\|_{L^2(\Omega)} + \|\partial_t u|_{t=0}\|_{H^{-1}(\Omega)} \lesssim \|u\|_{L^2((0,T)\times\Omega)}.$$

The claimed estimate follows from (A.2).

Let us now turn the claimed continuity. Let $\epsilon > 0$ and u_j be a mollification in time that satisfies $u_j \to u$ in $L^2((\epsilon, T-\epsilon)\times\Omega)$. Then u_j converges in

$$C(\epsilon, T-\epsilon; L^2(\Omega)) \cap C^1(\epsilon, T-\epsilon; H^{-1}(\Omega))$$

and thus u is in this space. In particular, $u|_{t=T/2}$ and $\partial_t u|_{t=T/2}$ are well-defined. Solving the initial value problem starting from these gives the desired conclusion. □

Theorem A.4 (Distributed observability estimate)**.** *Let $T > 0$ and let an open set $\omega \subset \Omega$ satisfy the geometric control condition. Then*

$$\|\phi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\phi|_{t=0}\|_{H^{-1}(\Omega)} \lesssim \|\phi\|_{L^2((0,T)\times\omega)} + \|\phi\|_{L^2((0,T)\times\partial\Omega)} + \|\Box\phi\|_{H^{-1}((0,T)\times\Omega)}.$$

Proof. It is classical that

$$\|\psi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\psi|_{t=0}\|_{H^{-1}(\Omega)} \lesssim \|\psi\|_{L^2((0,T)\times\omega)} \tag{A.5}$$

when

$$\begin{cases} \Box\psi = 0, \\ \psi|_{x\in\partial\Omega} = 0, \end{cases}$$

see *e.g.* [34]. Let u solve

$$\begin{cases} \Box u = \Box\phi, \\ u|_{x\in\partial\Omega} = \phi|_{x\in\partial\Omega}, \\ u|_{t=0} = 0,\ \partial_t u|_{t=0} = 0, \end{cases}$$

and define $\psi = \phi - u$. Then using (A.5) and (A.1),

$$\begin{aligned} \|\phi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\phi|_{t=0}\|_{H^{-1}(\Omega)} &= \|\psi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\psi|_{t=0}\|_{H^{-1}(\Omega)} \\ &\lesssim \|\psi\|_{L^2((0,T)\times\omega)} \le \|\phi\|_{L^2((0,T)\times\omega)} + \|u\|_{L^2((0,T)\times\omega)} \\ &\lesssim \|\phi\|_{L^2((0,T)\times\omega)} + \|\phi\|_{L^2((0,T)\times\partial\Omega)} + \|\Box\phi\|_{H^{-1}((0,T)\times\Omega)}. \end{aligned}$$

□

Remark A.5. By applying Theorem A.4 to the function $(t,x) \mapsto \phi(T-t,x)$ we obtain the following variant

$$\|\phi|_{t=T}\|_{L^2(\Omega)} + \|\partial_t\phi|_{t=T}\|_{H^{-1}(\Omega)} \lesssim \|\phi\|_{L^2((0,T)\times\omega)} + \|\phi\|_{L^2((0,T)\times\partial\Omega)} + \|\Box\phi\|_{H^{-1}((0,T)\times\Omega)},$$

and by combining Proposition A.1 and Theorem A.4, we get

$$\begin{aligned} &\|\phi\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t\phi\|_{L^2(0,T;H^{-1}(\Omega))} \\ &\quad\lesssim \|\phi\|_{L^2((0,T)\times\omega)} + \|\phi\|_{L^2((0,T)\times\partial\Omega)} + \|\Box\phi\|_{H^{-1}((0,T)\times\Omega)}, \end{aligned}$$

assuming that $T > 0$ and $\omega \subset \Omega$ satisfy the geometric control condition.

Theorem A.6 (Boundary observability estimate)**.** *Let $T > 0$ and let an open set $\omega \subset \partial\Omega$ satisfy the geometric control condition. Let $V \in C^\infty(\Omega)$ and define P by (4.1). Then*

$$\begin{aligned}&\|\phi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\phi|_{t=0}\|_{H^{-1}(\Omega)} \\ &\quad\lesssim \|\partial_\nu\phi\|_{H^{-1}((0,T)\times\omega)} + \|\phi\|_{L^2((0,T)\times\partial\Omega)} + \|P\phi\|_{H^{-1}((0,T)\times\Omega)}.\end{aligned} \tag{A.6}$$

Proof. It is well-known that

$$\|\psi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\psi|_{t=0}\|_{H^{-1}(\Omega)} \lesssim \|\partial_\nu\psi\|_{H^{-1}((0,T)\times\omega)} \tag{A.7}$$

for solutions ψ of

$$\begin{cases} P\psi = 0, \\ \psi|_{x\in\partial\Omega} = 0. \end{cases} \tag{A.8}$$

However, we did not find this exact formulation in the literature, and give a short proof for the convenience of the reader. It follows from (3.11) of the classical paper [4] that

$$\|\psi\|_{L^2((0,T)\times\Omega)} \lesssim \|\partial_\nu\psi\|_{H^{-1}((0,T)\times\omega)}, \tag{A.9}$$

since the space of invisible solutions is empty in our case due to unique continuation. Let $\tau \in C^\infty(\mathbb{R})$ satisfy $\tau(t) = 1$ near $t = 0$ and $\tau(t) = 0$ near $t = T$. Writing $u = \tau\psi$ and $f = \partial_t\tau\,\partial_t\psi + \partial_t^2\tau\,\psi$, there holds

$$\begin{cases} Pu = f, \\ u|_{x\in\partial\Omega} = 0, \\ u|_{t>T} = 0. \end{cases}$$

As $\psi \in L^2((0,T)\times\Omega)$, we have using $P\psi = 0$,

$$\|\partial_t^2\psi\|_{L^2(0,T;H^{-2}(\Omega))} = \|\Delta\psi\|_{L^2(0,T;H^{-2}(\Omega))} + \|V\psi\|_{L^2(0,T;H^{-2}(\Omega))} \lesssim \|\psi\|_{L^2((0,T)\times\Omega)}.$$

Now interpolation, see *e.g.* Theorems 2.3 and 12.2 of [36], gives

$$\|\partial_t\psi\|_{L^2(0,T;H^{-1}(\Omega))} \lesssim \|\psi\|_{L^2(0,T;L^2(\Omega))}.$$

Hence also

$$\|f\|_{L^2(0,T;H^{-1}(\Omega))} \lesssim \|\psi\|_{L^2(0,T;L^2(\Omega))},$$

and (A.2), or rather its analogue backward in time, gives

$$\|u|_{t=0}\|_{L^2(\Omega)} + \|\partial_t u|_{t=0}\|_{H^{-1}(\Omega)} \lesssim \|\psi\|_{L^2(0,T;L^2(\Omega))}.$$

But the state of u at $t = 0$ coincides with that of ψ, and (A.7) follows from the above estimate and (A.9).

Will now show (A.6). Let v be the solution of

$$\begin{cases} Pv = P\phi \\ v|_{x\in\partial\Omega} = \phi|_{x\in\partial\Omega} \\ v|_{t<0} = 0. \end{cases}$$

Then $\psi = \phi - v$ solves (A.8), and (A.7) and (A.1) imply

$$\begin{aligned} \|\phi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\phi|_{t=0}\|_{H^{-1}(\Omega)} &= \|\psi|_{t=0}\|_{L^2(\Omega)} + \|\partial_t\psi|_{t=0}\|_{H^{-1}(\Omega)} \\ &\lesssim \|\partial_\nu\phi\|_{H^{-1}((0,T)\times\omega)} + \|\partial_\nu v\|_{H^{-1}((0,T)\times\omega)} \\ &\lesssim \|\partial_\nu\phi\|_{H^{-1}((0,T)\times\omega)} + \|\phi\|_{L^2((0,T)\times\partial\Omega)} + \|P\phi\|_{H^{-1}((0,T)\times\Omega)}. \end{aligned}$$

□

The analogue of Remark A.5 holds also in the case of boundary observations. We need also the following classical result.

Lemma A.7 (Partial hypoellipticity)**.** *Let $I \subset (0,T)$ be a compact interval. Suppose that $s \in \mathbb{R}$ and $j \in \mathbb{N}$ satisfy $s + 1/2 > j$. Then*

$$\|\partial_\nu^j u\|_{H^{s-j-\frac{3}{2}}(I\times\partial\Omega)} \lesssim \|u\|_{H^s(M)} + \|\Box u\|_{H^{s-1}(M)}.$$

Proof. We define a norm $\|u\|_X$ by the right-hand side of the claimed inequality, and set $X = \{u \in H^s(M) : \|u\|_X < \infty\}$. It follows from the closed graph theorem that X is a Banach space. In normal coordinates of $\mathbb{R}\times\partial\Omega$, there holds

$$\Box = \partial_\nu^2 + A,$$

where A is a differential operator in the tangential directions to $\mathbb{R}\times\partial\Omega$, with coefficients depending on all the variables, see *e.g.* Corollary C.5.3 of [28].

We will use the spaces $\bar{H}_{(m,s)}(I\times\Omega)$, defined on p. 478 of [28], in the boundary normal coordinates, and use the shorthand notation $H_{(m,s)}$ for them. Here m measures Sobolev smoothness in all the variables and s additional smoothness in the tangential variables. However, s can be also negative, corresponding to a loss of smoothness in tangential directions.

Let $u \in X$. It follows from Theorem B.2.9 of [28] that $u \in H_{(m,r)}$ when $m + r \le s - 1$ and $m \le s + 1$. In particular, $u \in H_{(s+1,-2)}$ and the closed graph theorem implies

$$\|u\|_{H_{(s+1,-2)}} \lesssim \|u\|_X.$$

Moreover, using the assumption $s + 1/2 > j$, Theorem B.2.7 of [28] implies

$$\|\partial_\nu^j u\|_{H^{s-j-\frac{3}{2}}(I\times\partial\Omega)} \lesssim \|u\|_{H_{(s+1,-2)}}.$$

□

Lemma A.8. *Suppose that (A) holds. Let $u, \phi \in L^2(M)$ solve (2.1) with $u_0 = u_1 = 0$. Then $u = \phi = 0$.*

Proof. The lateral boundary traces on $(0,T) \times \partial\Omega$ are well-defined due to partial hypoellipticity, Lemma A.7. Typical energy estimates, see *e.g.* [33], give

$$u \in C(0,T; H_0^1(\Omega)) \cap C^1(0,T; L^2(\Omega)),$$

and Lemma A.3 implies that

$$\phi \in C(0,T; L^2(\Omega)) \cap C^1(0,T; H^{-1}(\Omega)).$$

In particular, we may parametrize ϕ by $\phi|_{t=0}$ and $\partial_t \phi|_{t=0}$. Let $\phi_j^0, \phi_j^1 \in C_0^\infty(\Omega)$ satisfy $\phi_j^0 \to \phi|_{t=0}$ in $L^2(\Omega)$ and $\phi_j^1 \to \partial_t \phi|_{t=0}$ in $H^{-1}(\Omega)$. Write (u_j, ϕ_j) for the solution of

$$\begin{cases} \Box u = \chi\phi, \\ u|_{x\in\partial\Omega} = 0, \\ u|_{t=0} = 0, \ \partial_t u|_{t=0} = 0, \end{cases} \qquad \begin{cases} \Box \phi = 0, \\ \phi|_{x\in\partial\Omega} = 0. \\ \phi|_{t=0} = \phi_j^0, \ \partial_t \phi|_{t=0} = \phi_j^1. \end{cases}$$

Then

$$\begin{aligned} (\chi\phi_j, \phi_j)_{L^2(M)} &= (\Box u_j, \phi_j)_{L^2(M)} - (u_j, \Box\phi_j)_{L^2(M)} \\ &= (\partial_t u_j|_{t=T}, \phi_j|_{t=T})_{L^2(\Omega)} - (u_j|_{t=T}, \partial_t \phi_j|_{t=T})_{H_0^1 \times H^{-1}(\Omega)}. \end{aligned}$$

Taking the limit $j \to \infty$ shows that $\phi = 0$ in $\mathrm{supp}(\chi)$. The distributed observability estimate, see Theorem A.4, implies that $\phi = 0$ in M. It follows that also $u = 0$ in M. □

Lemma A.9. *Suppose that (A') holds. Let $(u,\phi) \in L^2(M) \times H^1(M)$ solve (4.2) with $u_0 = u_1 = 0$. Then $u = \phi = 0$.*

Proof. For the convenience of the reader, we show first that

$$\phi \in C(0,T; H_0^1(\Omega)) \cap C^1(0,T; L^2(\Omega)), \quad \partial_\nu \phi|_\Gamma \in L^2(\Gamma). \tag{A.10}$$

The proof of this fact is very similar to the proof of Lemma A.3. The standard energy estimate implies that for all $s \in (0,T)$,

$$\|\phi|_{t=0}\|_{H^1(\Omega)} + \|\partial_t \phi|_{t=0}\|_{L^2(\Omega)} \lesssim \|\phi|_{t=s}\|_{H^1(\Omega)} + \|\partial_t \phi|_{t=s}\|_{L^2(\Omega)}.$$

Integration in s gives

$$\|\phi|_{t=0}\|_{H^1(\Omega)} + \|\partial_t \phi|_{t=0}\|_{L^2(\Omega)} \lesssim \|\phi\|_{H^1(M)},$$

and the regularity (A.10) follows now from [33]. It also follows from [33] that

$$u \in C(0,T; L^2(\Omega)) \cap C^1(0,T; H^{-1}(\Omega)).$$

In the case that u and ϕ are smooth

$$0 = (Pu, \phi)_{L^2(M)} - (u, P\phi)_{L^2(M)} = -(\chi \partial_\nu \phi, \partial_\nu \phi)_{L^2(\Gamma)},$$

and for $(u, \phi) \in L^2(M) \times H^1(M)$ this can be justified by approximating u and ϕ with smooth functions as in the proof of Lemma A.8. It follows from the boundary observability estimate, see Theorem A.6, that $\phi = 0$ identically, and hence also $u = 0$ identically. $\square$

Appendix B. Estimates for meshes fitted to the boundary

Proof of Lemma 4.1. Let $u \in C^\infty(K)$. Let $h > 0$, $K \in \mathcal{T}_h$ and consider spherical coordinates $(r, \theta) \in (0, \infty) \times S^n$ centered at x where x is as in (T). It follows from (4.7) that K is star-shaped with respect to x. In particular, there is $R : S^n \to (0, \infty)$ such that

$$K = \{(r, \theta) : 0 \le r \le R(\theta),\ \theta \in S^n\}.$$

As ∂K is piecewise smooth, it follows from (4.7) that R is piecewise smooth. Applying Theorem 6, p. 713 of [22] in a piecewise manner, we see that

$$\int_{\partial K} u^2\, \nu \cdot \rho \, \mathrm{d}s = \int_{S^n} u^2(R(\theta), \theta) R(\theta)^n \, \mathrm{d}\theta,$$

where $\mathrm{d}\theta$ is the canonical volume measure on the unit sphere S^n. It follows from (4.6) that $h \lesssim R(\theta)$. Hence, using (4.7) we have

$$\begin{aligned} h \int_{\partial K} u^2 \, \mathrm{d}s &\lesssim \int_{S^n} u^2(R(\theta), \theta) R(\theta)^{1+n} \, \mathrm{d}\theta = \int_{S^n} \int_0^{R(\theta)} \partial_r (u^2(r, \theta) r^{1+n}) \, \mathrm{d}r \, \mathrm{d}\theta \\ &\lesssim \int_{S^n} \int_0^{R(\theta)} (u^2 + |u||r\partial_r u|) r^n \, \mathrm{d}r \, \mathrm{d}\theta \lesssim \|u\|_K^2 + \|h \nabla u\|_K^2. \end{aligned}$$

$\square$

Proof of Lemma 4.2. We choose an extension $\hat{u}$ of u so that $\|\hat{u}\|_{H^k(\mathbb{R}^{1+n})} \lesssim \|u\|_{H^k(M)}$. Let $u_h \in \hat{V}_h^p$ be the Scott–Zhang interpolation of $\hat{u}$ where

$$\hat{V}_h^p = \{u \in H^1(M_h) : u|_K \in \mathbb{P}_p(K) \text{ for all } K \in \hat{\mathcal{T}}_h\}.$$

Clearly $u_h|_M \in V_h^p$, and the now classical result [46] says that the analogue

$$\|\hat{u} - u_h\|_{H^k(\hat{\mathcal{T}}_h)} \lesssim h^k \|\hat{u}\|_{H^k(\mathbb{R}^{1+n})}$$

of (4.9) holds. This implies (4.9) since $\mathcal{T}_h$ is obtained from $\hat{\mathcal{T}}_h$ *via* a restriction. $\square$

Acknowledgements. EB acknowledges funding by EPSRC grants EP/P01576X/1 and EP/V050400/1. AF acknowledges funding by EPSRC grant EP/P01593X/1 and support from the Fields institute for research in mathematical sciences. AM acknowledges funding by the French government research program "Investissements d'Avenir" through the IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25). LO acknowledges funding by EPSRC grants EP/P01593X/1 and EP/R002207/1.

References

[1] P.F. Antonietti, I. Mazzieri and F. Migliorini, A space-time discontinuous Galerkin method for the elastic wave equation. *J. Comput. Phys.* **419** (2020) 109685.

[2] E. Aranda and P. Pedregal, A variational method for the numerical simulation of a boundary controllability problem for the linear and semilinear 1D wave equations. *J. Franklin Inst.* **351** (2014) 3865–3882.

[3] M. Asch and A. Münch, Uniformly controllable schemes for the wave equation on the unit square. *J. Optim. Theory Appl.* **143** (2009) 417–438.
[4] C. Bardos, G. Lebeau and J. Rauch, Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary. *SIAM J. Control Optim.* **30** (1992) 1024–1065.
[5] F. Bourquin, Approximation theory for the problem of exact controllability of the wave equation with boundary control. In *Second International Conference on Mathematical and Numerical Aspects of Wave Propagation (Newark, DE, 1993)*. SIAM, Philadelphia, PA (1993), pages 103–112.
[6] S. Brenner and L. Scott, The mathematical theory of finite element methods. Springer-Verlag, third edition (2008).
[7] E. Burman, A stabilized nonconforming finite element method for the elliptic Cauchy problem. *Math. Comp.* **86** (2017) 75–96.
[8] E. Burman, A. Feizmohammadi, A. Münch and L. Oksanen, Space time stabilized finite element methods for a unique continuation problem subject to the wave equation. *ESAIM: Math. Model. Numer. Anal.* **55** (2021) S969–S991.
[9] E. Burman, A. Feizmohammadi and L. Oksanen, A finite element data assimilation method for the wave equation. *Math. Comp.* **89** (2020) 1681–1709.
[10] E. Burman, A. Feizmohammadi and L. Oksanen, A fully discrete numerical control method for the wave equation. *SIAM J. Control Optim.* **58** (2020) 1519–1546.
[11] C. Castro and S. Micu, Boundary controllability of a linear semi-discrete 1-D wave equation derived from a mixed finite element method. *Numer. Math.* **102** (2006) 413–462.
[12] C. Castro, S. Micu and A. Münch, Numerical approximation of the boundary control for the wave equation with mixed finite elements in a square. *IMA J. Numer. Anal.* **28** (2008) 186–214.
[13] N. Cîndea, E. Fernández-Cara and A. Münch, Numerical controllability of the wave equation through primal methods and Carleman estimates. *ESAIM: COCV* **19** (2013) 1076–1108.
[14] N. Cîndea, S. Micu and M. Tucsnak, An approximation method for exact controls of vibrating systems. *SIAM J. Control Optim.* **49** (2011) 1283–1305.
[15] N. Cîndea and A. Münch, A mixed formulation for the direct approximation of the control of minimal L^2-norm for linear type wave equations. *Calcolo* **52** (2015) 245–288.
[16] W. Dörfler, S. Findeisen and C. Wieners, Space-time discontinuous galerkin discretizations for linear first-order hyperbolic evolution systems. *Comput. Methods Appl. Math.* **16** (2016) 409–428.
[17] S. Dumont, F. Jourdan and T. Madani, 4D remeshing using a space-time finite element method for elastodynamics problems. *Math. Comput. Appl.* **23** (2018) Paper No. 29, 18.
[18] A. Ern and J.-L. Guermond, Theory and practice of finite elements. Vol. 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York (2004).
[19] A. Ern and J.-L. Guermond, Finite element quasi-interpolation and best approximation. *ESAIM Math. Model. Numer. Anal.* **51** (2017) 1367–1385.
[20] S. Ervedoza and E. Zuazua, A systematic method for building smooth controls for smooth data. *Discrete Contin. Dyn. Syst. Ser. B* **14** (2010) 1375–1401.
[21] S. Ervedoza and E. Zuazua, *Numerical approximation of exact controls for waves*. SpringerBriefs in Mathematics. Springer, New York (2013).
[22] L.C. Evans, Partial differential equations. Vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition (2010).
[23] D.A. French, A space-time finite element method for the wave equation. *Comput. Methods Appl. Mech. Eng.* **107** (1993) 145–157.
[24] R. Glowinski, W. Kinton and M.F. Wheeler, A mixed finite element formulation for the boundary controllability of the wave equation. *Internat. J. Numer. Methods Engrg.* **27** (1989) 623–635.
[25] R. Glowinski, C.H. Li and J.-L. Lions, A numerical approach to the exact boundary controllability of the wave equation. I. Dirichlet controls: description of the numerical methods. *Japan J. Appl. Math.* **7** (1990) 1–76.
[26] M. Gunzburger, L.S. Hou and L. Ju, A numerical method for exact boundary controllability problems for the wave equation. *Comput. Math. Appl.* **51** (2006) 721–750.
[27] F. Hecht, New development in Freefem++. *J. Numer. Math.* **20** (2012) 251–265.
[28] L. Hörmander, The analysis of linear partial differential operators. III. Classics in Mathematics. Springer, Berlin (2007).
[29] G.M. Hulbert and T.J.R. Hughes, Space-time finite element methods for second-order hyperbolic equations. *Comput. Methods Appl. Mech. Engrg.* **84** (1990) 327–348.
[30] L.I. Ignat and E. Zuazua, Convergence of a two-grid algorithm for the control of the wave equation. *J. Eur. Math. Soc. (JEMS)* **11** (2009) 351–391.
[31] C. Johnson, Discontinuous Galerkin finite element methods for second order hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.* **107** (1993) 117–129.
[32] U. Langer and O. Steinbach, editors, *Space-Time Methods*. De Gruyter (2019).
[33] I. Lasiecka, J.-L. Lions and R. Triggiani, Nonhomogeneous boundary value problems for second order hyperbolic operators. *J. Math. Pures Appl. (9)* **65** (1986) 149–192.
[34] J. Le Rousseau, G. Lebeau, P. Terpolilli and E. Trélat, Geometric control condition for the wave equation with a time-dependent observation domain. *Anal. PDE* **10** (2017) 983–1015.
[35] G. Lebeau and M. Nodet, Experimental study of the HUM control operator for linear waves. *Exp. Math.* **19** (2010) 93–120.

[36] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. I.* Springer-Verlag, New York-Heidelberg (1972). Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[37] P. Lissy and I. Rovenţa, Optimal filtration for the approximation of boundary controls for the one-dimensional wave equation using a finite-difference method. *Math. Comp.* **88** (2019) 273–291.

[38] P. Loreti and M. Mehrenberger, An Ingham type proof for a two-grid observability theorem. *ESAIM: COCV* **14** (2008) 604–631.

[39] S. Micu, Uniform boundary controllability of a semi-discrete 1-D wave equation. *Numer. Math.* **91** (2002) 723–768.

[40] S. Micu, I. Rovenţa and L.E. Temereancă, Approximation of the controls for the wave equation with a potential. *Numer. Math.* **144** (2020) 835–887.

[41] L. Miller, Escape function conditions for the observation, control, and stabilization of the wave equation. *SIAM J. Control Optim.* **41** (2002) 1554–1566.

[42] S. Montaner and A. Münch, Approximation of controls for linear wave equations: a first order mixed formulation. *Math. Control Relat. Fields* **9** (2019) 729–758.

[43] A. Münch, A uniformly controllable and implicit scheme for the 1-D wave equation. *M2AN Math. Model. Numer. Anal.* **39** (2005) 377–418.

[44] A. Münch and E. Trélat, Constructive exact control of semilinear 1-D wave equations by a least-squares approach (2020).

[45] P. Pedregal, F. Periago and J. Villena, A numerical method of local energy decay for the boundary controllability of time-reversible distributed parameter systems. *Stud. Appl. Math.* **121** (2008) 27–47.

[46] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.* **54** (1990) 483–493.

[47] O. Steinbach and M. Zank, A Stabilized Space–Time Finite Element Method for the Wave Equation. Springer International Publishing, Cham (2019), pp. 341–370.

[48] N.J. Walkington, Combined DG-CG time stepping for wave equations. *SIAM J. Numer. Anal.* **52** (2014) 1398–1417.

[49] X. Zhang, Explicit observability estimate for the wave equation with potential and its application. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **456** (2000) 1101–1115.

[50] E. Zuazua, Propagation, observation, and control of waves approximated by finite difference methods. *SIAM Rev.* **47** (2005) 197–243.