ESAIM: M2AN 55 (2021) S969–S991
https://doi.org/10.1051/m2an/2020062

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

SPACE TIME STABILIZED FINITE ELEMENT METHODS FOR A UNIQUE CONTINUATION PROBLEM SUBJECT TO THE WAVE EQUATION

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

Abstract. We consider a stabilized finite element method based on a spacetime formulation, where the equations are solved on a global (unstructured) spacetime mesh. A unique continuation problem for the wave equation is considered, where a noisy data is known in an interior subset of spacetime. For this problem, we consider a primal-dual discrete formulation of the continuum problem with the addition of stabilization terms that are designed with the goal of minimizing the numerical errors. We prove error estimates using the stability properties of the numerical scheme and a continuum observability estimate, based on the sharp geometric control condition by Bardos, Lebeau and Rauch. The order of convergence for our numerical scheme is optimal with respect to stability properties of the continuum problem and the approximation order of the finite element residual. Numerical examples are provided that illustrate the methodology.

Mathematics Subject Classification. 65M32, 35R30.

Received December 2, 2019. Accepted August 25, 2020.

1. Introduction

We consider a data assimilation problem for the acoustic wave equation, formulated as follows. Let $n \in \{1, 2, 3\}$, $T > 0$ and $\Omega \subset \mathbb{R}^n$ be an open, connected, bounded set with smooth boundary $\partial\Omega$. Let u be the solution of the initial boundary value problem

$$\begin{cases} \Box u = \partial_t^2 u - \Delta u = 0, & \text{on} \quad \mathcal{M} = (0,T) \times \Omega, \\ u = 0, & \text{on} \quad \Sigma = (0,T) \times \partial\Omega, \\ u|_{t=0} = u_0, \ \partial_t u|_{t=0} = u_1 & \text{on} \quad \Omega. \end{cases} \tag{1.1}$$

The initial data u_0, u_1 are assumed to be *a priori* unknown functions, but the measurements of u in some spacetime subset $\mathcal{O} = (0,T) \times \omega$, where $\omega \subset \overline{\Omega}$ is open, is assumed to be known:

$$u|_{\mathcal{O}} = u_{\mathcal{O}}. \tag{1.2}$$

Keywords and phrases. Unique continuation, data assimilation, wave equation, finite element method, geometric control condition, observability estimate.

[1] Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK.

[2] Laboratoire de Mathématiques Blaise Pascal, Université Clermont Auvergne, UMR CNRS 6620, Campus des Cézeaux, Aubière 63177, France.

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

The data assimilation problem then reads as follows:
Find u given $u_{\mathcal{O}}$. (DA)

The existence of a solution to the (DA) problem is always implicitly guaranteed in the sense that the measurements $u_{\mathcal{O}}$ correspond to a physical solution to the wave equation (1.1). On the other hand, assuming that

$$T > 2\max\{\operatorname{dist}(x,\omega)\,|\,x\in\overline{\Omega}\}, \tag{1.3}$$

with $\operatorname{dist}(x,\omega)$ defined as the infimum over the lengths of continuous paths in Ω, joining x and a point in ω, it follows from Holmgren's unique continuation theorem that the solution to (DA) is unique. Although uniquely solvable, (DA) might have poor stability properties if only (1.3) is assumed. We will require the (DA) problem to be Lipschitz stable, and for this reason we make the stronger assumption that the so-called *geometric control condition* holds. This condition originates from [3, 4] and we refer the reader to these works for the precise definition. Roughly speaking, the condition requires that all geometric optic rays in $\mathcal{M}$, taking into account their reflections at boundary, intersect the set $(0,T)\times\omega$.

We recall the following formulation of the observability estimates appearing in Theorem 3.3 of [4] and Proposition 1.2 of [35]. For the explicit derivation of this version of the estimate, we refer the reader to Theorem 2.2 of [14].

Theorem 1.1. *Let $\omega\subset\overline{\Omega}$, $T>0$ and suppose that $(0,T)\times\omega$ satisfies the geometric control condition. If $u(0,\cdot)\in L^2(\Omega)$, $\partial_t u(0,\cdot)\in H^{-1}(\Omega)$, $u|_{(0,T)\times\partial\Omega}\in L^2(\Sigma)$, and $\Box u\in H^{-1}(\mathcal{M})$, then*

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

and

$$\sup_{t\in[0,T]}\left(\|u(t,\cdot)\|_{L^2(\Omega)}+\|\partial_t u(t,\cdot)\|_{H^{-1}(\Omega)}\right)\leqslant C\left(\|u\|_{L^2(\mathcal{O})}+\|\Box u\|_{H^{-1}(\mathcal{M})}+\|u\|_{L^2(\Sigma)}\right),$$

where $C>0$ is a constant depending on $\mathcal{M}$ and ω.

Let us remark that the geometric control condition is sharp in the sense that Theorem 1.1 fails to hold if the geometric control condition does not hold on the set $(0,T)\times\omega$ [3].

The objective of the paper is to design a stabilized spacetime finite element method for the data assimilation problem (DA), which allows for higher order approximation spaces. The method will also allow for an error analysis exploiting the stability of Theorem 1.1 and the accuracy of the spaces in an optimal way. To the best of our knowledge this is the first complete numerical analysis of the data assimilation problem for the wave equation, using high order spaces. The statement of our main theoretical result appears as Theorem 4.4 in Section 4.

1.1. Previous literature

Spacetime methods for inverse problems subject to the wave equation were introduced in [17] with an application to the control problem in [18]. In those works however, the required H^2 regularity of the constraint equation was respected on the level of approximation leading to an approach using C^1-continuous approximation spaces in spacetime. Herein we instead use an approach where the approximating space is only H^1-conforming and we handle instabilities arising due to the lack of conformity of the space through the addition of stabilization terms. This approach to stabilization of ill-posed problems draws on the works [8, 9], for the elliptic Cauchy problem. In the context of time dependent problems, unique continuation for the heat equation was considered in [11,12], with piecewise affine finite elements in space and finite differences for the time discretization. Finally using a similar low order approach, with conventional, finite difference type time-discretization, the unique continuation and control problems for the wave equation were considered in [13,14] respectively. Another strategy requiring only H^1-regularity consists in reformulating the second order wave equation as a first order system; it is examined in [37] for the corresponding controllability problem.

Let us mention that the earlier works [13,14] studied the numerical implementation of data assimilation and control problems for the wave equation using the similar idea of implementing numerical stabilization terms in the discrete Lagrangian formulation. These earlier works are based on a first order finite element method in the space variables and a finite difference scheme in the time variable. Analogously to the theory here, the error estimates in these works are only based on the stability properties of the continuum problem along with the numerical approximation errors. The main aim of the authors in the works [13,14] was to show that even under a first order approximation scheme and with the simple time-stepping discretization in time, it is possible to obtain optimal rates of convergence for the numerical solutions to the (DA) and control problems. In contrast, the current work uses a mixed spacetime formulation, having the key advantage that it easily generalizes to higher order approximation spaces while still obtaining optimal error bounds, without resorting to C^1-type approximation spaces.

There are several works that approach the data assimilation problem (DA), or its close variants, by solving a sequence of classical initial–boundary value problems for the wave equation. Such methods have been proposed independently in [43], in the context of a particular application to Photoacoustic tomography, and in [41], based on the so-called Luenberger observers algorithm first introduced in an ODE context in [2]. An error estimate for a discretization of a Luenberger observers based algorithm was proven in [28], giving logarithmic convergence rate with respect to the mesh size. Better convergence rates can be proven if a stability estimate is available on a scale of discrete spaces. Such discrete estimates were first derived in [30,38] and we refer the reader to the survey articles [24,48] and the monograph [25], as well as the recent paper [26] for more details. Optimal-in-space discrete estimates can be derived from continuous estimates [36], however, spacetime optimal discrete estimates are known only for specific situations.

The data assimilation problem (DA) can also be solved using the quasi-reversibility method. This method originates from [34], and it has been applied to data assimilation problems subject to the wave equation in [31,32], and more recently to the Photoacoustic tomography problem in [19]. We are not aware of any works proving sharp convergence rates for the quasi-reversibility method with respect to mesh size.

The data assimilation problem (DA) arises in several applications. We mentioned already Photoacoustic tomography (PAT), and refer to [47] for physical aspects of PAT, to [33] for a mathematical review, and to [1,16,39,44] for the PAT problem in a cavity, the case closest to (DA). Another interesting application is given in [6] where an obstacle detection problem is solved by using a level set method together with the quasi-reversibility method applied to a variant of (DA).

1.2. Outline of the paper

The paper is organized as follows. In Section 2 we introduce a few notations used in the paper. In Section 3, we start by introducing the mixed spacetime mesh followed by the discrete representation of (DA) in terms of a primal dual Lagrangian formulation. The Euler–Lagrange equations are studied, showing in particular that there exists a unique solution to the discrete formulation of (DA). Section 4 is concerned with proving the convergence rates for the numerical error functions corresponding to the primal and dual variables. In Section 5, we provide two numerical examples that illustrate the theory while also making a comparison with the H^2-conformal finite element method introduced in [17]. Finally, in Section 6 we provide some concluding remarks.

2. Notations

We write $\nabla_{t,x}u = (\partial_t u, \nabla u)$, where $\nabla u \in \mathbb{R}^n$ is the usual gradient with respect to the space variables. The wave operator may be written as

$$\Box u = -\nabla_{t,x} \cdot (A\nabla_{t,x}u),$$

where A is the matrix associated to the Minkowski metric in $\mathbb{R}^{1+n}$, that is,

$$A = \begin{bmatrix} -1 & 0_{[1,n]} \\ 0_{[n,1]} & \mathbb{I}_{[n,n]} \end{bmatrix}$$

with $\mathbb{I}_{[n,n]}$ denoting the $n \times n$ identity matrix. We introduce the notations

$$(u,v)_{\mathcal{M}} = \int_0^T \int_\Omega u(t,x)v(t,x)\,\mathrm{d}x\,\mathrm{d}t, \quad \|u\|_{\mathcal{M}} = (u,u)_{\mathcal{M}}^{\frac{1}{2}}$$

and

$$(u,v)_{\mathcal{O}} = \int_0^T \int_\omega u(t,x)v(t,x)\,\mathrm{d}x\,\mathrm{d}t, \quad \|u\|_{\mathcal{O}} = (u,u)_{\mathcal{O}}^{\frac{1}{2}}$$

and use an analogous notation for inner products over other subsets of $\overline{\mathcal{M}}$, with the understanding that the natural volume measures are used in the case of subdomains. We also use the shorthand notation

$$a(u,z) = (A\nabla_{t,x}u, \nabla_{t,x}z)_{\mathcal{M}},$$

and note in passing that given any $(u_0,u_1) \in H_0^1(\Omega) \times L^2(\Omega)$, the solution u to equation (1.1) satisfies

$$a(u,v) = 0, \quad \forall v \in H_0^1(\mathcal{M}).$$

Finally, we fix an integer p, and make the standing assumption that the continuum solution u to (DA) satisfies

$$u \in H^{p+1}(\mathcal{M}) \quad \text{for some} \quad p \in \mathbb{N}.$$

The index p defined above will correspond to the highest order of the spacetime polynomial approximations that can be used for the discrete solution to (DA), while still getting optimal convergence for the numerical errors.

3. Discrete formulation of the data assimilation problem

Let us start this section by observing that the solution u to the data assimilation problem (DA) can be obtained by analyzing the saddle points for the continuum Lagrangian functional $\mathcal{J}(u,z)$ that is defined through

$$\mathcal{J}(u,z) = \frac{1}{2}\|u - u_{\mathcal{O}}\|_{\mathcal{O}}^2 + a(u,z),$$

for any

$$u \in H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega)) \quad \text{and} \quad z \in H_0^1(\mathcal{M}).$$

Here, the wave equation is imposed on the primal variable u by introducing a Lagrange multiplier z. It is easy to verify the the solution u to (DA) together with $z = 0$ is a saddle point for the Lagrangian functional.

Motivated by this example, we would like to present a discrete Lagrangian functional to numerically solve (DA). It is well-known that a naive discrete approximation of the continuum Lagrangian $\mathcal{J}(u,z)$ above will fail to work, due to the appearance of high frequency instabilities. This was first discovered by Glowinski *et al.* in a series of works in the early 1990s in the context of numerical controllability for the wave equation. We refer the reader to Sections 6.8 and 6.9 of [27] for a summary of these results. To remedy the issue of these spurious modes arising at high frequencies, we will use discrete stabilizer (also called regularizer) terms that guarantee the existence of a unique discrete saddle point. These terms will be designed with the goal of minimizing the numerical error functions for the primal and dual variables. The exact form of the discrete Lagrangian will be discussed later in Section 3.4.

We begin by introducing the spacetime mesh in Section 3.1. Then a discrete version of the bi-linear functional $a(\cdot,\cdot)$ is provided in Section 3.2, followed by the introduction of the stabilization terms in Section 3.3. Finally, we present the discrete formulation of (DA) in Section 3.4.

3.1. Spacetime discretization

In this section we will introduce the spacetime finite element method that we propose. The method is using an H^1-conforming piecewise polynomial space defined on a spacetime triangulation that can consist of simplices, or prisms. Herein for simplicity we restrict the discussion to the simplicial case. To be able to handle the case of curved boundaries without complicating the theory with estimations of the error in the approximation of the geometry we impose boundary conditions using a technique introduced by Nitsche [40]. See also Theorem 2.1 of [46] for a discussion of the application of the method to curved boundaries.

Consider a family $\mathcal{T} = \{\mathcal{T}_h;\ h > 0\}$ of quasi uniform triangulations of $\mathcal{M}$ consisting of simplices $\{K\}$ such that the intersection of any two distinct simplices is either a common vertex, a common edge or a common face. We let $h_K = \operatorname{diam}(K)$ and $h = \max_{K\in\mathcal{T}} h_K$, (see *e.g.* [22], Def. 1.140). By quasi-uniformity h/h_K is uniformly bounded, and therefore for simplicity the quasi-uniformity constant will be set to one below. Observe that we do not consider discretization of the smooth boundary $\partial\mathcal{M}$, but instead we allow triangles adjacent to the boundary to have curved faces, fitting $\mathcal{M}$. Finally, given any $k \in \mathbb{N}$, we let V_h^k be the $H^1(\mathcal{M})$-conformal approximation space of polynomial degree less than or equal to k, that is,

$$V_h^k = \{u \in C(\mathcal{M}) : u|_K \in \mathbb{P}_k(K), \quad \forall K \in \mathcal{T}_h\}, \tag{3.1}$$

where $\mathbb{P}_k(K)$ denotes the set of polynomials of degree less than or equal to $k \geq 1$ on K.

Next, we record two inequalities that will be used in the paper. The family $\mathcal{T}$ satisfies the following trace inequality, see *e.g.* equation 10.3.9 of [7],

$$\|u\|_{L^2(\partial K)} \lesssim h^{-\frac{1}{2}}\|u\|_{L^2(K)} + h^{\frac{1}{2}}\|\nabla u\|_{L^2(K)}, \quad u \in H^1(K). \tag{3.2}$$

The family $\mathcal{T}$ also satisfies the following discrete inverse inequality, see *e.g.* Lemma 1.138 of [22],

$$\|\nabla u\|_{L^2(K)} \lesssim h^{-1}\|u\|_{L^2(K)}, \quad u \in \mathbb{P}_p(K). \tag{3.3}$$

Remark 3.1. We will use the notations $A \lesssim B$ (resp., $A \gtrsim B$) to imply that there exists a constant $C > 0$ independent of the spacetime mesh parameter h, such that the inequalities $A \leqslant CB$ (resp., $A \geqslant CB$) hold.

Remark 3.2. It is also possible to consider the space $V_h^k \cap C^1(\mathcal{M})$ with $k \geqslant 2$, for the approximation of the primal variable. In this case the method coincides with that of [17] and the analysis shows that optimal error estimates are satisfied also in this case.

Remark 3.3. We will use the approximation space V_h^p for the primal variable u. We will also fix an integer $q \leqslant p$ and use the space V_h^q for the approximation of the dual variable z. As we will see, using our method, the approximation space of the dual variable can be quite coarse without sacrificing any rate of convergence for the discrete primal variable (*i.e.* we can take $q = 1$).

3.2. A discrete bi-linear formulation for the wave equation

Since no boundary conditions are imposed on the space V_h^p, the form $a(u, z)$ needs to be modified on the discrete level. For the formulation to remain consistent we propose the following modified bilinear form on $V_h^p \times V_h^q$,

$$a_h(u_h, z_h) = a(u_h, z_h) - (A\nabla_{t,x}u_h \cdot n_{\partial\mathcal{M}}, z_h)_{\partial\mathcal{M}} - (\nabla z_h \cdot n_{\partial\Omega}, u_h)_{\Sigma}.$$

Here $n_{\partial\mathcal{M}}$ and $n_{\partial\Omega}$ are the outward unit normal vectors on $\partial\mathcal{M}$ and $\partial\Omega$ respectively. The last term in the right hand side is added to make the weak form of the Laplace operator symmetric even in the case where no boundary conditions are imposed on the discrete spaces. Depending on how the stabilizing terms are chosen below, this term is not strictly necessary in this work, but becomes essential if the formulation must be consistent also for the adjoint equation.

Observe that, using integration by parts, there holds

$$a_h(u, z_h) = (\Box u, z_h)_{\mathcal{M}} - \underbrace{(\nabla z_h \cdot n_{\partial\Omega}, u)_\Sigma}_{=0} = (\Box u, z_h)_{\mathcal{M}} \tag{3.4}$$

for all $u \in H^2(\mathcal{M}) \cap L^2(0,T; H_0^1(\Omega))$.

3.3. Formulation of the discrete stabilization terms for primal and dual variables

We denote by $\mathcal{F}_h$ the set of internal faces of $\mathcal{T}_h$. For vector valued quantities u we define the jump across a face $F \in \mathcal{F}_h$ by

$$[\![n \cdot u]\!]_F = n_1 \cdot u|_{K_1} + n_2 \cdot u|_{K_2},$$

where $K_1, K_2 \in \mathcal{T}_h$ are the two simplices satisfying $K_1 \cap K_2 = F$ and n_j is the outward unit normal vector of K_j, $j = 1, 2$. Associating an arbitrary but fixed normal n_F to each face ($n_F = n_1$ or $n_F = n_2$), the jump of a scalar quantity u over a face F may be defined by

$$[\![u]\!]_F = u|_{K_1} n_F \cdot n_1 + u|_{K_2} n_F \cdot n_2.$$

The jump of a vector quantity may also be defined by applying the definition of the scalar jump componentwise, without modifications of the theory. Below we drop the normal to alleviate the notation. The norm over all the faces in $\mathcal{F}_h$ will be denoted by

$$\|v\|_{\mathcal{F}_h} = \left(\sum_{F \in \mathcal{F}_h} \|v\|_F^2 \right)^{\frac{1}{2}},$$

and the norm over all the simplices $\{K\}$ will be denoted by

$$\|v\|_{\mathcal{T}_h} = \left(\sum_{K \in \mathcal{T}_h} \|v\|_K^2 \right)^{\frac{1}{2}}.$$

For each $K \in \mathcal{T}_h$, we define the elementwise stabilizing form

$$s_K(u_h, u_h) = \|h \Box u_h\|_K^2 + \|h^{-\frac{1}{2}} u_h\|_{\partial K \cap \Sigma}^2 + \sum_{F \in \partial K \cap \mathcal{F}_h} \|h^{\frac{1}{2}} [\![A \nabla_{t,x} u_h]\!]\|_F^2. \tag{3.5}$$

A dual stabilizer is defined by

$$s_K^*(z_h, z_h) = \|\nabla_{t,x} z_h\|_K^2 + \|h^{-\frac{1}{2}} z_h\|_{\partial K \cap \partial \mathcal{M}}^2. \tag{3.6}$$

Subsequently, the global stabilizers are defined by summing over all the elements:

$$s = \sum_{K \in \mathcal{T}_h} s_K \quad \text{and} \quad s^* = \sum_{K \in \mathcal{T}_h} s_K^*$$

and we define the semi norms $|u|_s = s(u,u)^{\frac{1}{2}}$ and $|z|_{s^*} = s^*(z,z)^{\frac{1}{2}}$. Observe that the following stability estimate holds:

$$|w_h|_{s^*} \lesssim \|\nabla_{t,x} w_h\|_{\mathcal{M}} + \|h^{-\frac{1}{2}} w_h\|_{\partial \mathcal{M}} \quad \forall w_h \in V_h^q. \tag{3.7}$$

We also point out for future reference that for a solution to the data assimilation problem $u \in H^2(\mathcal{M})$ and all $u_h \in V_h^p$, there holds

$$s(u - u_h, u - u_h) = s(u_h, u_h). \tag{3.8}$$

Remark 3.4. There is some freedom in the choices of the discrete regularization terms that yield the same error estimates as in Theorem 4.4 below. The choice is more flexible for the dual variable z since the continuum analogue for this variable is zero. For instance we can define the following stabilization term for the dual variable:

$$s_K^*(z_h, z_h) = s_K(z_h, z_h) + \|h^{-\frac{1}{2}} z_h\|_{\partial K \cap (\partial \mathcal{M} \setminus \Sigma)} + \|h^{\frac{1}{2}} \partial_t z_h\|^2_{\partial K \cap (\partial \mathcal{M} \setminus \Sigma)}. \tag{3.9}$$

It is also possible to use a stabilization that is exclusively carried by the faces of the computational mesh provided $q \in \{p-2, p-1, p\}$. In this case we define

$$s_K(u_h, u_h) = \sum_{F \in \partial K \cap \mathcal{F}_h} \left(\|h^{\frac{1}{2}} [\![A \nabla_{t,x} u_h]\!]\|^2_F + \|h_F^{\frac{3}{2}} [\![\Box u_h]\!]\|^2_F \right) + \|h^{-\frac{1}{2}} u_h\|^2_{\partial K \cap \Sigma}. \tag{3.10}$$

In the second jump term we are allowed to split the operator in the time derivative and the Laplace operator (or second order derivatives in space) without sacrificing stability or consistency since

$$\|h_F^{\frac{3}{2}} [\![\Box u_h]\!]\|^2_F \leqslant \|h_F^{\frac{3}{2}} [\![\partial_t^2 u_h]\!]\|^2_F + \|h_F^{\frac{3}{2}} [\![\Delta u_h]\!]\|^2_F.$$

Weak consistency of the right order still holds since for a sufficiently smooth solution u

$$\|h_F^{\frac{3}{2}} [\![\partial_t^2 u]\!]\|^2_F + \|h_F^{\frac{3}{2}} [\![\Delta u]\!]\|^2_F = 0.$$

3.4. The discrete Lagrangian formulation for the data assimilation problem

Our finite element method is defined by the discrete Lagrangian functional

$$\mathcal{L} : V_h^p \times V_h^q \to \mathbb{R},$$

through

$$\mathcal{L}(u, z) = \frac{1}{2} \|u - \tilde{u}_{\mathcal{O}}\|^2_{\mathcal{O}} + \frac{\gamma}{2} s(u, u) - \frac{\gamma^*}{2} s^*(z, z) + a_h(u, z), \tag{3.11}$$

where $\gamma, \gamma^* > 0$ are fixed constants. Here,

$$\tilde{u}_{\mathcal{O}} = u_{\mathcal{O}} + \delta u_{\mathcal{O}},$$

with $u_{\mathcal{O}}$ denoting the restriction to the subset $\mathcal{O}$ of a continuum solution $u \in H^{p+1}(\mathcal{M})$ to (1.1) and $\delta u_{\mathcal{O}} \in L^2(\mathcal{O})$ denoting some experimental noise in our observable data.

The corresponding Euler–Lagrange equations read as follows. Find $(u_h, z_h) \in V_h^p \times V_h^q$ such that for all $(v_h, w_h) \in V_h^p \times V_h^q$ there holds

$$a_h(u_h, w_h) - \gamma^* s^*(z_h, w_h) = 0, \tag{3.12}$$

$$(u_h, v_h)_{\mathcal{O}} + \gamma s(u_h, v_h) + a_h(v_h, z_h) = (\tilde{u}_{\mathcal{O}}, v_h)_{\mathcal{O}}. \tag{3.13}$$

To simplify the notation we introduce the bi-linear form

$$\mathcal{A}_h[(u_h, z_h), (v_h, w_h)] = (u_h, v_h)_{\mathcal{O}} + \gamma s(u_h, v_h) + a_h(v_h, z_h) + a_h(u_h, w_h) - \gamma^* s^*(z_h, w_h).$$

The discrete problem (3.12) and (3.13) can then be recast as follows. Find $u_h, z_h \in V_h^p \times V_h^q$ such that

$$\mathcal{A}_h[(u_h, z_h), (v_h, w_h)] = (\tilde{u}_{\mathcal{O}}, v_h)_{\mathcal{O}}, \quad \forall (v_h, w_h) \in V_h^p \times V_h^q. \tag{3.14}$$

Note that by definition,

$$\mathcal{A}_h[(u_h - u, z_h), (v_h, w_h)] = (u_h - u, v_h)_{\mathcal{O}} + \gamma s(u_h - u, v_h) + a_h(v_h, z_h) + a_h(u_h - u, w_h) - \gamma^* s^*(z_h, w_h).$$

By (3.4), $a_h(u, w_h) = 0$. Together with (3.8) we can simplify the above expression to obtain

$$\mathcal{A}_h[(u_h - u, z_h), (v_h, w_h)] = (u_h - u, v_h)_{\mathcal{O}} + \gamma s(u_h, v_h) + a_h(v_h, z_h) + a_h(u_h, w_h) - \gamma^* s^*(z_h, w_h).$$

Finally, applying (3.12) and (3.13) and the fact that $u|_{\mathcal{O}} = u_{\mathcal{O}}$, we conclude that

$$\mathcal{A}_h[(u_h - u, z_h), (v_h, w_h)] = (\delta u_{\mathcal{O}}, v_h)_{\mathcal{O}}. \tag{3.15}$$

Define the residual norm

$$|||(u, z)|||_S^2 = \|u\|_{\mathcal{O}}^2 + |u|_s^2 + |z|_{s^*}^2,$$

and a continuity norm

$$\|u\|_* = \|\nabla_{t,x} u\|_{\mathcal{M}} + \|h^{\frac{1}{2}} A \nabla_{t,x} u \cdot n\|_{\partial\mathcal{M}} + \|h^{-\frac{1}{2}} u\|_{\Sigma}.$$

For the purpose of our error analysis later, we also introduce a family of interpolants π_h^k, that are required to satisfy

Assumption 3.5. *$\pi_h^k : H^k(\mathcal{M}) \to V_h^k$ preserves Dirichlet boundary conditions and additionally satisfies*

$$\|u - \pi_h^k u\|_{H^m(\mathcal{M})} \lesssim h^{s-m} \|u\|_{H^s(\mathcal{M})} \text{ for all } u \in H^s(\mathcal{M}) \text{ with } s = 0, 1, \ldots, k+1 \text{ and } m = 0, 1, \ldots, s.$$

An example of such an interpolant is the Scott–Zhang interpolant [42]. For brevity, we will use the notations:

$$\pi_h = \pi_h^1 \quad \text{and} \quad \Pi_h = \pi_h^p.$$

We have the following lemma regarding the residual norm. We remind the reader that the notation $\lesssim$ is as defined in Remark 3.1.

Lemma 3.6. *Let $u \in H^{p+1}(\mathcal{M})$. There holds:*

$$|||(u - \Pi_h u, 0)|||_S \lesssim h^p \|u\|_{H^{p+1}(\mathcal{M})}.$$

Proof. Note that

$$|||(u - \Pi_h u, 0)|||_S = \|u - \Pi_h u\|_{\mathcal{O}} + |u - \Pi_h u|_s.$$

For the first term we immediately see that

$$\|u - \Pi_h u\|_{\mathcal{O}} \leqslant \|u - \Pi_h u\|_{\mathcal{M}} \lesssim h^{p+1} \|u\|_{H^{p+1}(\mathcal{M})}.$$

To bound the contribution from the stabilization term, recall by definition that

$$|u - \Pi_h u|_s^2 = \sum_{K \in \mathcal{T}_h} \left(\|h \Box (u - \Pi_h u)\|_K^2 + \|h^{-\frac{1}{2}} (u - \Pi_h u)\|_{\partial K \cap \Sigma}^2 + \sum_{F \in \partial K \cap \mathcal{F}_h} \|h^{\frac{1}{2}} [\![A \nabla_{t,x} (u - \Pi_h u)]\!]\|_F^2 \right).$$

We proceed to bound the three terms on the right hand side. For the first term, we note that

$$\sum_{K \in \mathcal{T}_h} \|h \Box (u - \Pi_h u)\|_K^2 \lesssim \sum_{K \in \mathcal{T}_h} \|\nabla_{t,x} (u - \Pi_h u)\|_K^2 \lesssim h^{2p} \|u\|_{H^{p+1}(\mathcal{M})}^2.$$

For the second term, we define $\Delta_F = \{K : \bar{K} \cap F \neq \emptyset\}$, $\tilde{\Delta}_F = \{K : \bar{K} \cap \bar{\Delta}_F \neq \emptyset\}$ and use (3.2) to write

$$\|h^{-\frac{1}{2}} (u - \Pi_h u)\|_F^2 \lesssim h^{-1} \|u - \Pi_h u\|_{\Delta_F}^2 + \|\nabla_{t,x} (u - \Pi_h u)\|_{\Delta_F}^2 \lesssim h^{2p} |u|_{H^{p+1}(\tilde{\Delta}_F)}^2.$$

By collecting the above local bounds and using the fact that $\tilde{\Delta}_F$ have finite overlaps, we conclude that

$$\sum_{K\in\mathcal{T}_h} \|h^{-\frac{1}{2}}(u-\Pi_h u)\|^2_{\partial K\cap\Sigma} \lesssim h^{2p}|u|^2_{H^{p+1}(\tilde{\Delta}_F)}.$$

For the last term, observe that using (3.2) again, we have:

$$\|h_F^{\frac{1}{2}}[\![A\nabla_{t,x}(u-\Pi_h u)]\!]\|_F^2 \lesssim \sum_{K\in\Delta_F}\left(\|\nabla_{t,x}(u-\Pi_h u)\|_K^2 + h^2\|D^2_{t,x}(u-\Pi_h u)\|_K^2\right) \lesssim h^{2p}|u|^2_{H^{p+1}(\tilde{\Delta}_F)},$$

where $D^2_{t,x}$ denotes the Hessian matrix consisting of second order derivatives in space and time variables. The claim follows by collecting the above local bounds analogously to the second term above. □

The next lemma is concerned with approximation properties of the continuity norm $\|\cdot\|_*$ defined earlier. The proof is analogous to the proof of the previous lemma and follows from the definition of the interpolant Π_h together with (3.2) and is therefore omitted.

Lemma 3.7. *Let $u\in H^{p+1}(\mathcal{M})$. There holds:*

$$\|u-\Pi_h u\|_* \lesssim h^p\|u\|_{H^{p+1}(\mathcal{M})},$$

where we recall that

$$\|u\|_* = \|\nabla_{t,x}u\|_{\mathcal{M}} + \|h^{\frac{1}{2}}A\nabla_{t,x}u\cdot n\|_{\partial\mathcal{M}} + \|h^{-\frac{1}{2}}u\|_\Sigma.$$

We end this section by proving that the solution to (3.14) exists and is unique.

Proposition 3.8. *The Euler–Lagrange equation (3.14) has a unique solution $(u_h,z_h)\in V_h^p\times V_h^q$.*

Proof. Since equation (3.14) defines a square system of linear equations, existence is equivalent to uniqueness and we only need to show that for $u_{\mathcal{O}}\equiv 0$, the solution $(u_h,z_h)=(0,0)$ is unique. Indeed, suppose that equation (3.14) with $u_{\mathcal{O}}\equiv 0$ holds for some $(u_h,z_h)\in V_h^p\times V_h^q$. First observe that

$$|||u_h,z_h|||_S^2 \lesssim \mathcal{A}_h((u_h,z_h),(u_h,-z_h)) = 0.$$

This means that $|z_h|_{s^*}=0$. Consequently, $z_h=0$ follows immediately by the Poincaré inequality. Next, considering u_h we immediately have that $u_h|_{\mathcal{O}}=0$ and $u_h|_\Sigma=0$. Using the definition of the stabilization (3.5) we see that by partial integration, followed by the Cauchy–Schwarz inequality and the trace inequality (3.2) there holds for all $w\in H^1_0(\mathcal{M})$

$$a(u_h,w) = (\Box u_h,w)_{\mathcal{T}_h} + \sum_{K\in\mathcal{T}_h}(A\nabla_{t,x}u_h\cdot n_{\partial K},w)_{\partial K} \lesssim h^{-1}(\underbrace{\|h\Box u_h\|_{\mathcal{T}_h}}_{=0} + \underbrace{\|h^{\frac{1}{2}}[\![A\nabla_{t,x}u_h]\!]\|_{\mathcal{F}_h}}_{=0})\|w\|_{H^1(\mathcal{M})}.$$

As a consequence

$$\|\Box u_h\|_{H^{-1}(\mathcal{M})} = \sup_{\substack{w\in H^1_0(\mathcal{M})\\ \|w\|_{H^1(\mathcal{M})}=1}} a(u_h,w) = 0.$$

By construction $u_h\in C^0(\overline{\mathcal{M}})\cap H^1(\mathcal{M})$ and in particular $u_h(0,\cdot)\in L^2(\Omega), \partial_t u_h(0,\cdot)\in L^2(\Omega), u_h|_{(0,T)\times\partial\Omega}\in L^2(\Sigma)$. We conclude that u_h vanishes thanks to Theorem 1.1. □

4. Error estimates

We will consider the derivation of error estimates in three steps. First, we will establish the continuity of $a_h(\,)$ with respect to $|||\cdot|||_S$ and $\|\cdot\|_*$ on the one hand (see Lem. 4.1) and then a continuity for the exact solution with respect to $|||\cdot|||_S$ and H^1 norms (see Lem. 4.2) on the other hand. Then, we will prove convergence of the error in the $|||\cdot|||_S$ norm. Finally, we will use these results to prove a posteriori and *a priori* error estimates based on the observability estimate of Theorem 1.1.

Lemma 4.1. *Let $v \in H^2(\mathcal{M}) + V_h^p$ and $w_h \in V_h^q$. Then there holds*

$$|a_h(v, w_h)| \lesssim \|v\|_* |w_h|_{s^*}.$$

Proof. First, observe that by (3.2) and (3.3) there holds

$$\|h^{\frac{1}{2}} A \nabla_{t,x} w_h \cdot n\|_\Sigma \lesssim \|\nabla_{t,x} w_h\|_{\mathcal{M}}. \tag{4.1}$$

Next, using the Cauchy–Schwarz inequality we write:

$$|a_h(v, w_h)| \leqslant \|\nabla_{t,x} v\|_{\mathcal{M}} \|\nabla_{t,x} w_h\|_{\mathcal{M}} + \|h^{\frac{1}{2}} \nabla_{t,x} v \cdot n\|_{\partial\mathcal{M}} \|h^{-\frac{1}{2}} w_h\|_{\partial\mathcal{M}} + \|h^{\frac{1}{2}} \nabla_{t,x} w_h \cdot n\|_\Sigma \|h^{-\frac{1}{2}} v\|_\Sigma.$$

The claim follows by combining the previous two bounds. □

Lemma 4.2. *Let $u \in H^{p+1}(\mathcal{M})$ be the exact solution of (1.1) satisfying (1.2) and let $(u_h, z_h) \in V_h^p \times V_h^q$ be the unique solution of the discrete Euler–Lagrange equation (3.14). There holds*

$$\|\Box(u - u_h)\|_{H^{-1}(\mathcal{M})} = \sup_{\substack{w \in H_0^1(\mathcal{M}) \\ \|w\|_{H^1(\mathcal{M})} = 1}} a(u_h, w) \lesssim |||(u_h, z_h)|||_S.$$

Proof. First observe that

$$\|\Box(u - u_h)\|_{H^{-1}(\mathcal{M})} = \sup_{\substack{w \in H_0^1(\mathcal{M}) \\ \|w\|_{H^1(\mathcal{M})} = 1}} a(u - u_h, w).$$

Since $\Box u = 0$, we have $a(u, w) = 0$ for all $w \in H_0^1(\mathcal{M})$ thus establishing the first equality in the claim. Using (3.12) we see that,

$$a(u_h, w) = a(u_h, w - \pi_h w) + a(u_h, \pi_h w) - a_h(u_h, \pi_h w) + \gamma^* s^*(z_h, \pi_h w). \tag{4.2}$$

Using integration by parts in the first term of the right hand side we see that

$$\begin{aligned} a(u_h, w - \pi_h w) + a(u_h, \pi_h w) - a_h(u_h, \pi_h w) &= (\Box u_h, w - \pi_h w)_{\mathcal{T}_h} + \sum_{K \in \mathcal{T}_h} (A \nabla_{t,x} u_h \cdot n_{\partial K}, w - \pi_h w)_{\partial K} \\ &+ (A \nabla_{t,x} u_h \cdot n_{\partial\mathcal{M}}, \pi_h w)_{\partial\mathcal{M}} + (u_h, \nabla \pi_h w \cdot n)_\Sigma = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}. \end{aligned}$$

For the term I we have

$$|\mathrm{I}| = |(\Box u_h, w - \pi_h w)_{\mathcal{T}_h}| \leqslant \|h \Box u_h\|_{\mathcal{T}_h} (h^{-1} \|w - \pi_h w\|_{\mathcal{T}_h}) \lesssim \|h \Box u_h\|_{\mathcal{T}_h} \|w\|_{H^1(\mathcal{M})}.$$

Observe that the term III is absorbed by the same quantity with opposite sign in II, eliminating all terms $\pi_h w$ on the boundary. Since also $w|_{\partial\mathcal{M}} = 0$, we see that

$$|\mathrm{II} + \mathrm{III}| = |\sum_{K \in \mathcal{T}_h} (A \nabla_{t,x} u_h \cdot n_{\partial K}, w - \pi_h w)_{\partial K \setminus \partial\mathcal{M}}| \leqslant \|h^{\frac{1}{2}} [\![A \nabla_{t,x} u_h]\!]\|_{\mathcal{F}_h} \|w\|_{H^1(\mathcal{M})},$$

where we are using (3.2) and the approximation properties of the interpolant π_h. Finally, for the term IV we use the Cauchy–Schwarz inequality to get the bound

$$|\mathrm{IV}| = |(\nabla_{t,x}\pi_h w \cdot n, u_h)_\Sigma| \leqslant \|h^{-\frac{1}{2}}u_h\|_\Sigma \, \|h^{\frac{1}{2}}\nabla_{t,x}\pi_h w \cdot n\|_\Sigma \lesssim \|h^{-\frac{1}{2}}u_h\|_\Sigma \, \|w\|_{H^1(\mathcal{M})},$$

where we used the bound (4.1) in the last step.

Collecting the above bounds we have that

$$a(u_h, w - \pi_h w) + a(u_h, \pi_h w) - a_h(u_h, \pi_h w) \lesssim (\|h\Box u_h\|_{\mathcal{T}_h} + \|h^{\frac{1}{2}}[\![A\nabla_{t,x}u_h]\!]\|_{\mathcal{F}_h} + \|h^{-\frac{1}{2}}u_h\|_\Sigma)\|w\|_{H^1(\mathcal{M})}.$$

Using the definition of $|||(u_h, 0)|||_S$, we may rewrite this as

$$a(u_h, w - \pi_h w) + a(u_h, \pi_h w) - a_h(u_h, \pi_h w) \lesssim |||(u_h, 0)|||_S \|w\|_{H^1(\Omega)}.$$

For the remaining term in the right hand side of (4.2) we observe that by Cauchy–Schwarz inequality

$$s^*(z_h, \pi_h w) \leqslant |||(0, z_h)|||_S \, |\pi_h w|_{s^*}.$$

We can now use (3.7) to deduce

$$|\pi_h w|_{s^*} \lesssim \|\nabla_{t,x} w\|_{\mathcal{M}} + \|h^{-\frac{1}{2}}\pi_h w\|_{\partial\mathcal{M}} \lesssim \|w\|_{H^1(\mathcal{M})}.$$

□

We now prove convergence in the residual norm.

Proposition 4.3. *Let u be the solution to (1.1), satisfying (1.2). Let $(u_h, z_h) \in V_h^p \times V_h^q$ be the solution of (3.14). Then*

$$|||(u - u_h, z_h)|||_S \lesssim h^p \|u\|_{H^{p+1}(\mathcal{M})} + \|\delta u_{\mathcal{O}}\|_{L^2(\mathcal{O})}.$$

Proof. Let $u_h - u = \underbrace{u_h - \Pi_h u}_{e_h} + \underbrace{\Pi_h u - u}_{e_\Pi} = e_h + e_\Pi$. Using the triangle inequality we write

$$|||(u_h - u, z_h)|||_S \leqslant |||(e_\Pi, 0)|||_S + |||(e_h, z_h)|||_S.$$

Recalling Lemma 3.6 we only need an estimate for $|||(e_h, z_h)|||_S$. There holds

$$C|||(e_h, z_h)|||_S^2 \leqslant \mathcal{A}_h[(e_h, z_h), (e_h, -z_h)],$$

where $C > 0$ only depends on γ and γ^*. Using (3.15) we have

$$\mathcal{A}_h[(e_h, z_h), (e_h, -z_h)] = -\mathcal{A}_h[(e_\Pi, 0), (e_h, -z_h)] + (\delta u_{\mathcal{O}}, e_h)_{\mathcal{O}}.$$

Clearly,

$$|(\delta u_{\mathcal{O}}, e_h)_{\mathcal{O}}| \leqslant \|\delta u_{\mathcal{O}}\|_{\mathcal{O}} \|e_h\|_{\mathcal{O}} \leqslant \|\delta u_{\mathcal{O}}\|_{\mathcal{O}} |||(e_h, z_h)|||_S.$$

Also, by definition

$$\mathcal{A}_h[(e_\Pi, 0), (e_h, -z_h)] = (e_\Pi, e_h)_{\mathcal{O}} + \gamma s(e_\Pi, e_h) - a_h(e_\Pi, z_h).$$

As a consequence, applying the Cauchy–Schwarz inequality in the two first terms in the right hand side and the continuity of Lemma 4.1 in the last term we get the bound

$$\begin{aligned} |\mathcal{A}_h[(e_\Pi, 0), (e_h, -z_h)]| &\lesssim (\|e_\Pi\|_{\mathcal{O}} + |e_\Pi|_s + \|e_\Pi\|_*)(\|e_h\|_{\mathcal{O}} + |e_h|_s + |z_h|_{s^*}) \\ &\lesssim (\|e_\Pi\|_{\mathcal{O}} + |e_\Pi|_s + \|e_\Pi\|_*)|||(e_h, z_h)|||_S. \end{aligned}$$

Collecting the above bounds and applying Lemmas 3.6 and 3.7 we conclude that

$$|||(e_h, z_h)|||_S \lesssim \|e_\Pi\|_{\mathcal{O}} + |e_\Pi|_s + \|e_\Pi\|_* \lesssim h^p \|u\|_{H^{p+1}(\mathcal{M})} + \|\delta u_{\mathcal{O}}\|_{\mathcal{O}}.$$

□

We are now ready to state our main theorem as follows.

Theorem 4.4. *Let $p, q \in \mathbb{N}$, let $\mathcal{M} = (0,T) \times \Omega$ and $\omega \subset \Omega$ be such that the set $\mathcal{O} = (0,T) \times \omega$ satisfies the geometric control condition. Let $u \in H^{p+1}(\mathcal{M})$ solve the continuum equation (1.1) subject to some unknown initial data u_0, u_1 and assume that $u_\mathcal{O} = u|_\mathcal{O}$ is a priori known modulo some observable noise $\delta u_\mathcal{O} \in L^2(\mathcal{O})$. Let the discrete Lagrangian $\mathcal{L} : V_h^p \times V_h^q \to \mathbb{R}$ be defined by (3.11). Let $(u_h, z_h) \in V_h^p \times V_h^q$ be the unique solution to the Euler–Lagrange equations (3.14). Then we have the a posteriori error estimate*

$$\sup_{t\in[0,T]} \left(\|(u-u_h)(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t(u-u_h)(t,\cdot)\|_{H^{-1}(\Omega)} \right) \lesssim \left(\sum_{K\in\mathcal{T}} \eta_K^2 \right)^{\frac{1}{2}}$$

where

$$\eta_K^2 = \|u_h - u_\mathcal{O}\|_{\mathcal{O}\cap K}^2 + s_K(u_h,u_h) + s_K^*(z_h,z_h).$$

Moreover, the following a priori error estimate holds for the primal variable[3]

$$\sup_{t\in[0,T]} \left(\|(u-u_h)(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t(u-u_h)(t,\cdot)\|_{H^{-1}(\Omega)} \right) \lesssim h^p \|u\|_{H^{p+1}(\mathcal{M})} + \|\delta u_\mathcal{O}\|_\mathcal{O}.$$

Proof. Taking the square of the inequality of Theorem 1.1 we see that with $e = u - u_h$

$$\sup_{t\in[0,T]} \left(\|e(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t e(t,\cdot)\|_{H^{-1}(\Omega)} \right)^2 \lesssim \|e\|_\mathcal{O}^2 + \|\Box e\|_{H^{-1}(\mathcal{M})}^2 + \|e\|_\Sigma^2.$$

First we observe that

$$\|e\|_\mathcal{O}^2 = \sum_{K\in\mathcal{T}_h} \|u_h - u_\mathcal{O}\|_{\mathcal{O}\cap K}^2$$

and

$$\|e\|_\Sigma^2 = \sum_{K\in\mathcal{T}_h} \|u_h\|_{\Sigma\cap K}^2 \leqslant \sum_{K\in\mathcal{T}_h} \|h^{-\frac{1}{2}} u_h\|_{\Sigma\cap K}^2 \leqslant |u_h|_s^2.$$

Applying Lemma 4.2 we see that

$$\|\Box e\|_{H^{-1}(\mathcal{M})}^2 \lesssim \sum_{K\in\mathcal{T}_h} \eta_K^2$$

which proves the first claim.

For the *a priori* error estimate observe that by definition and by (3.8) we have

$$\left(\sum_{K\in\mathcal{T}_h} \eta_K^2 \right)^{\frac{1}{2}} \lesssim \|e\|_\mathcal{O} + |||(e,z_h)|||_S$$

and we conclude by applying the error bound of Proposition 4.3 to the right hand side. This concludes the proof. □

Remark 4.5. Tracking the influence of the stabilization parameters on the hidden constants in Theorem 4.4 leads to factors of the form $\gamma^{\frac{1}{2}}$, $\gamma^{-\frac{1}{2}}$, $(\gamma^*)^{\frac{1}{2}}$, $(\gamma^*)^{-\frac{1}{2}}$, showing that the present analysis does not allow either parameter to vanish or to become too large. The sensitivity to the parameter choice is explored in the numerical section.

[3]The convergence for the dual variable z_h is given by Proposition 4.3.

Remark 4.6. Theorem 4.4 can be used to make a number of observations. Firstly, the discrete algorithm is stable in the presence of the noise $\delta u_{\mathcal{O}}$. Secondly, when the size of the noise is comparable to h^p the method converges optimally with respect to the approximation order of the finite element spaces that are used. Finally, this latter observation suggests that in cases where some *a priori* knowledge of the sizes of the physical solution and the noise is known, the mesh size h should be taken to be of order $\|\delta u_{\mathcal{O}}\|_{\mathcal{O}}/\|u\|_{H^{p+1}(\mathcal{M})}$.

Remark 4.7. Observe that the preceding analysis shows that the order of the discretization space for the dual variable z_h, namely q, can be taken to be one without sacrificing any rate of convergence for the primal variable u_h. This is advantageous since then the system size only grows with increasing p.

5. Numerical experiments

We implement the stabilized finite element method introduced and analyzed in the previous sections with $p \geq q$ and assuming no noise in the observed data, that is to say $\delta u_{\mathcal{O}} = 0$. We also discuss the rate obtained according, notably, to the regularity of the initial condition to be reconstructed in the case $n = 1$. The results are compared with those obtained with the H^2-conformal finite element method introduced in [17] which reads as follows:

Find $(u, z) \in V \times L^2(0,T; H_0^1(\Omega))$, with $V = \{u \in L^2(\mathcal{M}), \Box u \in L^2(0,T; H^{-1}(\Omega)\}$, solution of

$$\begin{cases} (u,v)_{\mathcal{O}} + \gamma \displaystyle\int_0^T (\Box u, \Box v)_{H^{-1}(\Omega)} \mathrm{d}t + \int_0^T (z, \Box v)_{H_0^1(\Omega), H^{-1}(\Omega)} \mathrm{d}t = (u_{\mathcal{O}}, v)_{\mathcal{O}}, \quad \forall v \in V, \\ \displaystyle\int_0^T (w, \Box u)_{H_0^1(\Omega), H^{-1}(\Omega)} \mathrm{d}t = 0, \quad \forall w \in L^2(0,T; H_0^1(\Omega)), \end{cases} \tag{5.1}$$

where $(\cdot,\cdot)_{H_0^1(\Omega), H^{-1}(\Omega)}$ denotes the dual pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$ so that

$$(z, \Box u)_{H_0^1(\Omega), H^{-1}(\Omega)} = \big(\nabla z, \nabla(-\Delta^{-1}\Box u)\big)_{L^2(\Omega)}, \quad \forall z \in H_0^1(\Omega), u \in V.$$

For any $\gamma \geqslant 0$, this well-posed mixed formulation is associated to the Lagrangian

$$\widetilde{\mathcal{L}} : V \times L^2(0,T; H_0^1(\Omega)) \to \mathbb{R}$$

defined as follows

$$\widetilde{\mathcal{L}}(u,z) = \frac{1}{2}\|u - u_{\mathcal{O}}\|_{\mathcal{O}}^2 + \frac{\gamma}{2}\|\Box u\|_{L^2(H^{-1})}^2 - \int_0^T (z, \Box u)_{H_0^1(\Omega), H^{-1}(\Omega)} \mathrm{d}t. \tag{5.2}$$

At the finite dimensional level, the formulation reads: find $(u_h, z_h) \in V_h \times P_h$ solution of

$$\begin{cases} (u_h, v_h)_{\mathcal{O}} + \gamma h^2 (\Box u_h, \Box v_h)_{\mathcal{M}} + (z_h, \Box v_h)_{\mathcal{M}} = (u_{\mathcal{O}}, v_h)_{\mathcal{O}}, \quad \forall v_h \in V_h, \\ (w_h, \Box u_h)_{\mathcal{M}} = 0, \quad \forall w_h \in P_h, \end{cases} \tag{5.3}$$

where $V_h \subset V$ and $P_h \subset L^2(0,T; H_0^1(\Omega))$ for all $h > 0$. As in [17], we shall use a conformal approximation V_h based on the C^1 triangular reduced HCT element (see [5]). Concerning the approximation of the multiplier z, we consider $P_h = \{z \in C(\mathcal{M}) : u_{|K} \in \mathbb{P}^1(K), \forall K \in \mathcal{T}_h\}$. This method does not enter the above framework, however if a dual stabilizer as in (3.6) is added, the above theory may be applied and leads to error bounds also in this case.

The experiments are performed with the FreeFem++ package developed at the University Paris 6 (see [29]), very well-adapted to the spacetime formulation. In Table 1 we collect some data on the meshes that were used in the numerical experiments.

Remark 5.1. The cases where $p < q$ were included in the numerical study. Although the expected convergence rates were observed, such methods turned out to be very sensitive to the choice of stabilization parameters. This locking phenomenon lead to unsatisfactory results and the results are not reported here.

5.1. Example 1

For our first example, we take simply $\Omega = (0,1)$ and first consider an observation $u_{\mathcal{O}}$ based on the smooth initial condition $(u_0(x), u_1(x)) = (\sin(3\pi x), 0)$ completed with $T = 2$ and $\omega = (0.1, 0.3)$. Observe that the corresponding solution is simply the smooth function

$$(\mathbf{Ex1}) \quad u(t,x) = \sin(3\pi x)\cos(3\pi t).$$

In view of Theorem 4.4, we expect a rate equal to p for the primal variable u when approximated with elements in V_h^p. Tables 2–5 provide the norms

$$\|u-u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}, \quad \|(u-u_h)(0,\cdot)\|_{L^2(0,1)}/\|u(0,\cdot)\|_{L^2(0,1)}, \quad \|(u-u_h)_t(0,\cdot)\|_{H^{-1}(0,T)}$$

with respect to h for the primal variable and the norm $\|z_h\|_{L^2(0,T;H_0^1(0,1))}$ with respect to h for the dual variable. These are obtained from the formulation (5.3) based on a conformal approximation with $\gamma = 10^{-3}$ and from the formulation (3.14) based on a non conformal approximation with $p, q \in \{1,2,3\}$, $p \geq q$ (see (3.1)). For the latter, we use the dual stabilizer (3.6) with $\gamma = 10^{-3}$, $\gamma^\star = 1$. The linear system associated to the mixed formulation (3.12) and (3.13) is solved using the multi-frontal LU direct solver UMFPACK.

Concerning the primal variable u, we obtain the following behavior

$$\|u-u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})} \approx \beta \times h^\tau$$

with

$$\begin{aligned}
&\beta = e^{1.40}, \quad \tau = 1.66, \quad (u_h, z_h) \in V_h^1 \times V_h^1,\\
&\beta = e^{1.38}, \quad \tau = 3.06, \quad (u_h, z_h) \in V_h^2 \times V_h^1,\\
&\beta = e^{1.16}, \quad \tau = 4.06, \quad (u_h, z_h) \in V_h^3 \times V_h^1,\\
&\beta = e^{1.81}, \quad \tau = 3.35, \quad (u_h, z_h) \in V_h^2 \times V_h^2,\\
&\beta = e^{1.01}, \quad \tau = 4.01, \quad (u_h, z_h) \in V_h^3 \times V_h^2,\\
&\beta = e^{1.21}, \quad \tau = 4.08, \quad (u_h, z_h) \in V_h^3 \times V_h^3,\\
&\beta = e^{1.02}, \quad \tau = 3.67, \quad (u_h, z_h) \in \mathrm{HCT} \times V_h^1,
\end{aligned} \tag{5.4}$$

in agreement with Theorem 4.4. We observe actually a super convergence in all cases, very likely due to the regularity of the solution to be reconstructed. Figure 1 (left) depicts the evolution of the relative error $\|u-u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}$ with respect to h for the approximation $V_h^p \times V_h^1$, $p \in 1,2,3$ and for the HCT element based on a $H^2(Q_T)$ approximation. We notice that the case $p = 3$ produces similar results than the $H^2(Q_T)$ based approximation $\mathrm{HCT} \times V_h^1$. We also observe that the value of q does not affect the rate in agreement with Remark 4.7 (see Fig. 1 (right)). Table 6 collects the CPU time observed for various discretization and indicates that the use of the space $V_h^2 \times V_h^1$ seems very appropriate.

Since $u_{\mathcal{O}}$ is a well-prepared solution, we check from Table 2 that the approximation z_h of the dual variable (which has the meaning of a Lagrange multiplier for the weak formulation of the wave equation) goes to zero with h for the $L^2(0,T;H_0^1(0,1))$ norm.

With respect to the role of γ and $\gamma^\star$, here taken equal to 10^{-3} and 1 respectively, we have observed the following phenomenon: when p is strictly larger than q, *i.e.* when the primal variable is approximated in a richer space than the dual one, the value of $\gamma^\star$ has no influence on the quality of the result. In particular $\gamma^\star = 0$ still leads to a well-posed discrete formulation and provides the same results compared to for instance $\gamma^\star = 1$. This suggests, as least in the one dimensional setting for which $d = 1$, that the finite element spaces satisfy an inf-sup condition and that the stabilized term $s^\star(z_h, z_h)$ for the dual variable z is not necessary. Moreover, in that case, whatever be the value of $\gamma^\star$, γ must be small but strictly positive; the choice $\gamma = 0$ leads to a non invertible formulation. Still in the case $p \geqslant q$, we also observed that the rate of convergence with respect to h is independent of the value of the parameter γ. Small values strictly positive leads however to better constants

TABLE 1. Data of five triangular meshes associated to $\mathcal{M} = (0,1) \times (0,2)$.

Mesh	♯1	♯2	♯3	♯4	♯5
h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
card $(\mathcal{T}_h)$	442	1750	7164	29 182	116 300
♯vertices	252	936	3703	14 832	58 631
card (V_h) – $\mathbb{P}_2$	945	3621	14 569	58 845	233 561
card (V_h) – $\mathbb{P}_3$	2080	8056	32 599	132 040	524 791
card (V_h) – HCT	1449	5493	21 975	88 509	350 823

TABLE 2. (**Ex1**); $\|z_h\|_{L^2(0,T;H_0^1(0,1))}$ w.r.t h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times V_h^1$	9.88×10^{-4}	1.22×10^{-4}	1.83×10^{-5}	2.49×10^{-6}	3.52×10^{-7}
(3.14) – $V_h^1 \times V_h^1$	4.64×10^{-2}	4.79×10^{-2}	3.03×10^{-2}	1.51×10^{-2}	7.76×10^{-3}
(3.14) – $V_h^2 \times V_h^1$	6.14×10^{-3}	1.45×10^{-3}	4.07×10^{-4}	9.96×10^{-5}	2.64×10^{-5}
(3.14) – $V_h^3 \times V_h^1$	2.23×10^{-4}	3.02×10^{-5}	4.02×10^{-6}	4.99×10^{-7}	6.85×10^{-8}
(3.14) – $V_h^2 \times V_h^2$	4.10×10^{-2}	1.16×10^{-2}	2.68×10^{-3}	6.49×10^{-4}	1.61×10^{-4}
(3.14) – $V_h^3 \times V_h^2$	2.57×10^{-3}	3.98×10^{-4}	4.83×10^{-5}	6.43×10^{-6}	7.76×10^{-7}
(3.14) – $V_h^3 \times V_h^3$	9.06×10^{-3}	1.10×10^{-3}	1.25×10^{-4}	1.55×10^{-5}	1.92×10^{-6}

TABLE 3. (**Ex1**); $\|u - u_h\|_{L^2(\mathcal{M})} / \|u\|_{L^2(\mathcal{M})}$ w.r.t. h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times V_h^1$	1.32×10^{-2}	9.24×10^{-4}	8.72×10^{-5}	8.37×10^{-6}	1.20×10^{-6}
(3.14) – $V_h^1 \times V_h^1$	8.07×10^{-1}	4.94×10^{-1}	1.81×10^{-1}	4.90×10^{-2}	1.25×10^{-2}
(3.14) – $V_h^2 \times V_h^1$	1.00×10^{-1}	9.41×10^{-3}	1.23×10^{-3}	2.12×10^{-4}	4.03×10^{-5}
(3.14) – $V_h^3 \times V_h^1$	7.61×10^{-3}	5.16×10^{-4}	4.15×10^{-5}	2.64×10^{-6}	2.63×10^{-7}
(3.14) – $V_h^2 \times V_h^2$	1.58×10^{-1}	1.27×10^{-2}	1.21×10^{-3}	2.05×10^{-4}	3.01×10^{-5}
(3.14) – $V_h^3 \times V_h^2$	6.49×10^{-3}	3.97×10^{-4}	3.21×10^{-5}	2.29×10^{-6}	2.52×10^{-7}
(3.14) – $V_h^3 \times V_h^3$	9.07×10^{-3}	5.31×10^{-4}	3.92×10^{-5}	2.74×10^{-6}	3.01×10^{-7}

as reported in Figure 2. It should be noted that smaller value of γ leads to an increasing amount of CPU time needed to solve the formulation (3.14) through the UMFPACK solver (in view of the ill-posedeness for $\gamma = 0$).

On the contrary, when the same finite element space is used for primal and dual variables, *i.e.* when $p = q$, we observe that the stabilization of the dual variable, *i.e.* $\gamma^\star > 0$ is compulsory to achieve well-posedness. In that case, the choice $\gamma = 0$ provides excellent results (except for $p = q = 1$ and h not small enough). We remark that similar qualitative and quantitative conclusions are observed with structured meshes.

We also emphasize that the spacetime discretization introduced in the previous sections is very well-suited for mesh adaptivity. Using the $V_h^2 \times V_h^1$ approximation, Figure 4 (left) depicts the mesh obtained after seven adaptive refinements based on the local values of gradient of the primal variable u_h. Starting with a coarse mesh composed of 288 triangles and 166 vertices, the final mesh is composed by 13 068 triangles and 6700 vertices. We obtain the following values: $\|z_h\|_{L^2(H_0^1)} = 4.32 \times 10^{-4}$; $\|u - u_h\|_{L^2(\mathcal{M})} / \|u\|_{L^2(\mathcal{M})} = 8.42 \times 10^{-3}$; $\|(u - u_h)(0)\|_{L^2(0,1)} / \|u(0)\|_{L^2(0,1)} = 8.30 \times 10^{-3}$; $\|(u - u_h)_t(0)\|_{H^{-1}(0,1)} = 1.47 \times 10^{-3}$ for a CPU time equal to 3.01.

TABLE 4. (**Ex1**); $\|(u-u_h)(\cdot,0)\|_{L^2(0,1)}/\|u(\cdot,0)\|_{L^2(0,1)}$ w.r.t. h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times V_h^1$	6.64×10^{-3}	6.23×10^{-4}	6.55×10^{-5}	6.25×10^{-6}	1.20×10^{-6}
(3.14) – $V_h^1\times V_h^1$	8.02×10^{-1}	4.94×10^{-1}	1.81×10^{-1}	4.89×10^{-2}	1.25×10^{-2}
(3.14) – $V_h^2\times V_h^1$	1.04×10^{-1}	8.45×10^{-3}	9.30×10^{-4}	1.57×10^{-4}	2.32×10^{-5}
(3.14) – $V_h^3\times V_h^1$	4.81×10^{-3}	3.48×10^{-4}	3.68×10^{-5}	2.46×10^{-6}	1.85×10^{-7}
(3.14) – $V_h^2\times V_h^2$	1.55×10^{-1}	9.29×10^{-2}	1.03×10^{-3}	1.85×10^{-4}	2.00×10^{-5}
(3.14) – $V_h^3\times V_h^2$	4.22×10^{-3}	3.26×10^{-4}	2.23×10^{-5}	1.93×10^{-6}	1.65×10^{-7}
(3.14) – $V_h^3\times V_h^3$	5.23×10^{-3}	3.52×10^{-4}	2.87×10^{-5}	2.50×10^{-6}	1.99×10^{-7}

TABLE 5. (**Ex1**); $\|(u-u_h)_t(\cdot,0)\|_{H^{-1}(0,1)}$ w.r.t. h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times V_h^1$	1.74×10^{-1}	9.29×10^{-2}	3.81×10^{-2}	1.85×10^{-2}	8.96×10^{-3}
(3.14) – $V_h^1\times V_h^1$	2.85×10^{-2}	4.57×10^{-2}	2.68×10^{-2}	1.48×10^{-2}	7.09×10^{-3}
(3.14) – $V_h^2\times V_h^1$	3.87×10^{-2}	8.56×10^{-3}	2.03×10^{-3}	4.95×10^{-4}	1.19×10^{-4}
(3.14) – $V_h^3\times V_h^1$	6.10×10^{-3}	6.71×10^{-4}	6.32×10^{-5}	6.89×10^{-6}	8.32×10^{-7}
(3.14) – $V_h^2\times V_h^2$	3.18×10^{-2}	8.20×10^{-3}	2.01×10^{-3}	4.82×10^{-4}	1.14×10^{-4}
(3.14) – $V_h^3\times V_h^2$	6.82×10^{-3}	6.50×10^{-4}	6.24×10^{-5}	6.94×10^{-6}	8.28×10^{-7}
(3.14) – $V_h^3\times V_h^3$	7.36×10^{-3}	6.69×10^{-4}	6.26×10^{-5}	6.87×10^{-6}	8.26×10^{-7}

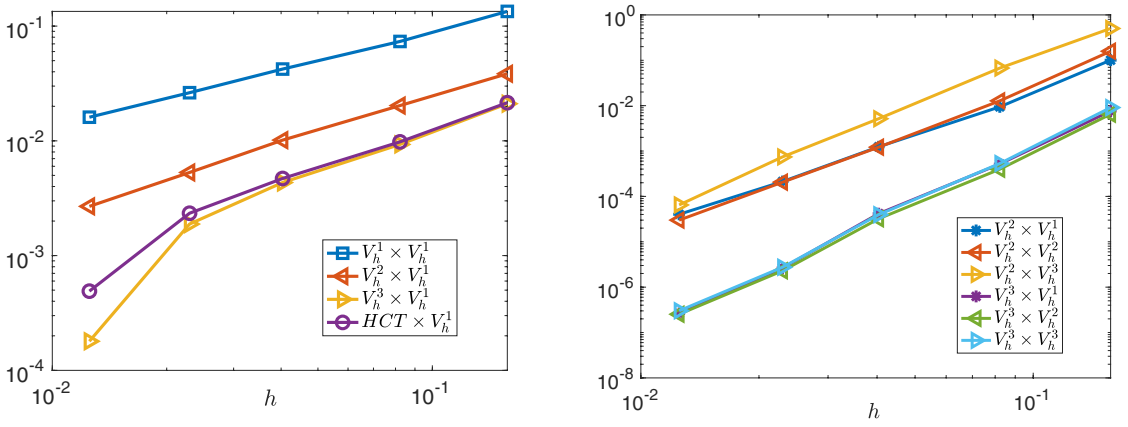

FIGURE 1. (**Ex1**) – Relative error $\|u-u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}$ with respect to h for various approximation spaces (see Tab. 3); $\gamma=10^{-3},\gamma^\star=1$.

TABLE 6. (**Ex1**); CPU time (in second) to solve (3.12) and (3.13) with the mesh ♯4.

$V_h^1\times V_h^1$	$V_h^2\times V_h^1$	$V_h^2\times V_h^2$	$V_h^3\times V_h^1$	$V_h^3\times V_h^2$	$V_h^3\times V_h^3$	HCT $\times V_h^1$
1.74	4.78	8.46	16.19	19.27	28.59	10.76

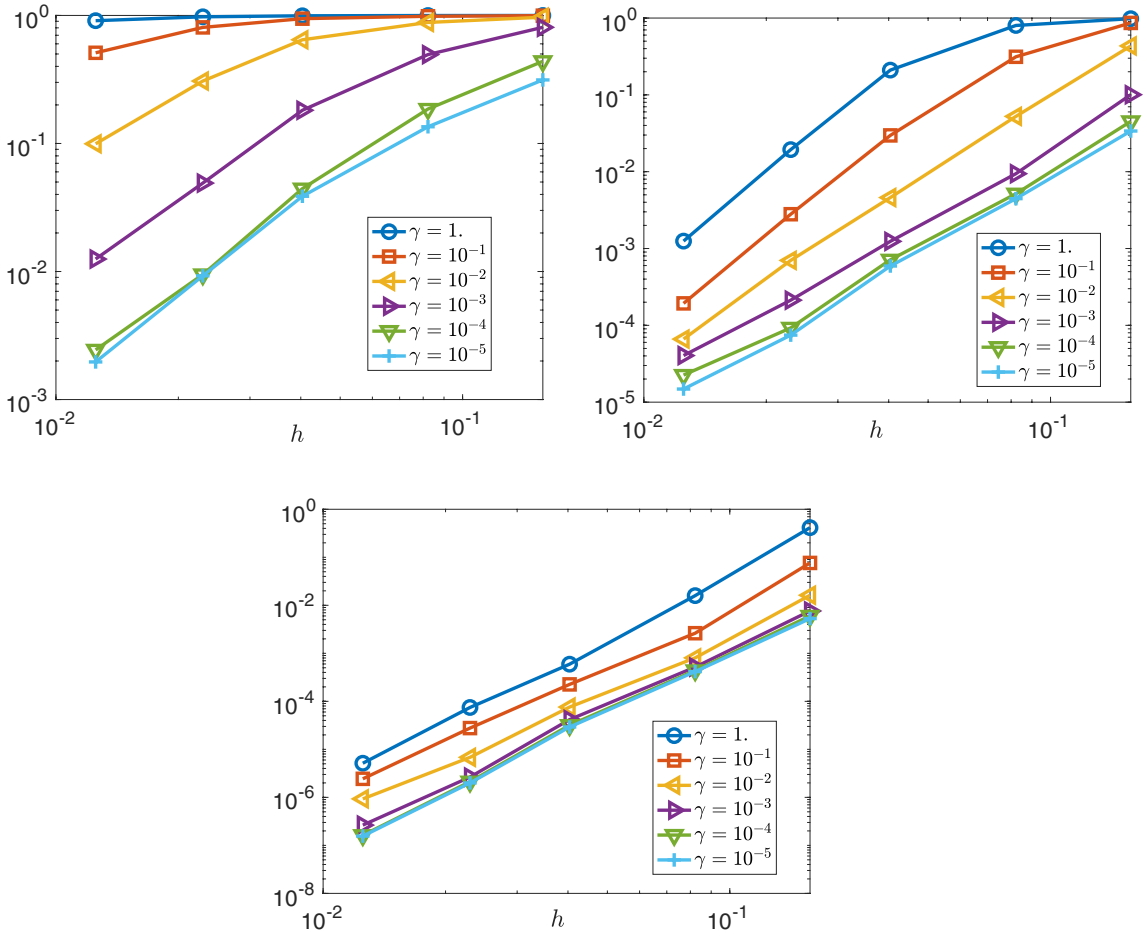

FIGURE 2. (**Ex1**) – Relative error $\|u - u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}$ with respect to h and γ for the approximation spaces $V_h^1 \times V_h^1$ (*left*), $V_h^2 \times V_h^1$ (*right*) and $V_h^3 \times V_h^1$ (*bottom*); $\gamma^\star = 1$.

5.1.1. Iterative solution using a conjugate gradient algorithm

Finally, we recall that the mixed formulation (3.12) and (3.13) associated to the Lagrangian $\mathcal{L}$ defined in (3.11) may be reformulated into an equivalent extremal problem involving only the dual variable z_h: precisely, if $(u_h, z_h) \in V_h^p \times V_h^q$ is the saddle-point of $\mathcal{L}$, then $z_h \in V_h^q$ is the minimizer of the functional $\mathcal{J}^\star : V_h^q \to \mathbb{R}$ defined by

$$\mathcal{J}^\star(z) := \frac{1}{2}(u_z, u_z)_{\mathcal{O}} + \frac{\gamma}{2} s(u_z, u_z) + \frac{\gamma^\star}{2} s^\star(z, z) - a_h(u_{z_0}, z)$$

where, for any $z \in V_h^q$, $u_z \in V_h^p$ solves, for all $v_h \in V_h^p$

$$(u_z, v_h)_{\mathcal{O}} + \gamma s(u_z, v_h) + a_h(v_h, z) = 0 \tag{5.5}$$

and where $u_{z_0} \in V_h^p$ solves, for all $v_h \in V_h^p$

$$(u_{z_0}, v_h)_{\mathcal{O}} + \gamma s(u_{z_0}, v_h) = (\tilde{u}_{\mathcal{O}}, v_h)_{\mathcal{O}}.$$

TABLE 7. (**Ex1**); Number of iterations for the CG algorithm w.r.t. h and γ; $\gamma^\star = 1$.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(3.14) – $V_h^1 \times V_h^1$ – $\gamma = 10^{-3}$	98	160	275	496	922
(3.14) – $V_h^2 \times V_h^1$ – $\gamma = 10^{-3}$	69	124	261	561	1257
(3.14) – $V_h^3 \times V_h^1$ – $\gamma = 10^{-3}$	49	101	216	474	997
(3.14) – $V_h^1 \times V_h^1$ – $\gamma = 10^{-2}$	37	54	92	168	314
(3.14) – $V_h^2 \times V_h^1$ – $\gamma = 10^{-2}$	50	89	169	338	699
(3.14) – $V_h^3 \times V_h^1$ – $\gamma = 10^{-2}$	43	82	174	357	682
(3.14) – $V_h^2 \times V_h^1$ – $\gamma = 10^{-1}$	24	41	74	140	280
(3.14) – $V_h^3 \times V_h^1$ – $\gamma = 10^{-1}$	25	46	85	160	292
(3.14) – $V_h^3 \times V_h^1$ – $\gamma = 1.$	11	18	31	58	102

TABLE 8. (**Ex2**); $\|z_h\|_{L^2(0,T;H_0^1(0,1))}$ w.r.t. h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times V_h^1$	4.70×10^{-4}	2.69×10^{-4}	1.48×10^{-4}	8.80×10^{-5}	2.44×10^{-5}
(3.14) – $V_h^1 \times V_h^1$	1.41×10^{-2}	9.02×10^{-3}	4.31×10^{-3}	2.37×10^{-3}	1.43×10^{-3}
(3.14) – $V_h^2 \times V_h^1$	9.47×10^{-4}	5.27×10^{-4}	3.13×10^{-4}	1.64×10^{-4}	9.55×10^{-5}
(3.14) – $V_h^3 \times V_h^1$	2.15×10^{-4}	1.23×10^{-4}	7.24×10^{-5}	4.70×10^{-5}	6.61×10^{-6}

TABLE 9. (**Ex2**); $\|(u - u_h)\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}$ w.r.t. h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times V_h^1$	2.15×10^{-2}	9.80×10^{-3}	4.70×10^{-3}	2.34×10^{-3}	4.90×10^{-4}
(3.14) – $V_h^1 \times V_h^1$	1.34×10^{-1}	7.34×10^{-2}	4.22×10^{-2}	2.62×10^{-2}	1.60×10^{-2}
(3.14) – $V_h^2 \times V_h^1$	3.84×10^{-2}	2.02×10^{-2}	1.01×10^{-1}	5.30×10^{-2}	2.68×10^{-3}
(3.14) – $V_h^3 \times V_h^1$	2.10×10^{-2}	9.30×10^{-3}	4.33×10^{-3}	1.88×10^{-3}	1.79×10^{-4}

The minimization of $\mathcal{J}^\star$ may be done iteratively through the conjugate gradient algorithm. Each iteration requires the resolution of the well-posed problem (5.5), simpler – notably in terms of computational ressource – than the direct resolution of (3.12) and (3.13). This reformulation is mainly of interest for multi-dimensional cases for which $n > 1$. We refer for instance to Section 2.2 of [17] where this is employed. The conjugate gradient is initialized with the zero function and stopped as soon as the gradient $g_k \in V_h^q$ at the iteration k satisfies $|g_k|_{s^\star}/|g_0|_{s^\star} < 10^{-4}$. Moreover, at each iteration, problem (5.5) is solved with the Cholesky solver. Once the minimizer z_h^{cg} of $\mathcal{J}^\star$ is obtained, the corresponding primal solution is defined as $u_h^{cg} = u_{z_h^{cg}} + u_{z_0}$. We checked that the above stopping criterion ensures a value of $\|u_h - u_h^{cg}\|_{L^2(\mathcal{M})} + \gamma|u_h - u_h^{cg}|_s + \gamma^\star|z_h - z_h^{cg}|_{s^\star}$ of the order 10^{-5}, 10^{-6} and 10^{-7} when the approximations $V_h^1 \times V_h^1$, $V_h^2 \times V_h^1$ and $V_h^3 \times V_h^1$ are employed respectively. Consequently, we simply report in Table 7 the number of iterations of the algorithm for $\gamma^\star = 1$ and $\gamma \in \{1, 10^{-1}, 10^{-2}, 10^{-3}\}$. As expected, a larger value of γ increases the coercivity property of the functional $\mathcal{J}^\star$. In view of Figure 2, the approximation $V_h^3 \times V_h^1$ combined with $\gamma = 1.$ leads to an appropriate choice.

5.2. Example 2

For our second numerical example, we consider the observation $u_\mathcal{O}$ based on the initial condition $u_0(x) = 1 - |2x - 1| \in H_0^1(\Omega)$, $u_1(x) = 1_{(1/3,2/3)}(x) \in L^2(\Omega)$ and $T = 2$, $\omega = (0.1, 0.3)$, considered in

TABLE 10. (**Ex2**); $\|(u-u_h)(\cdot,0)\|_{L^2(0,1)}/\|u(\cdot,0)\|_{L^2(0,1)}$ w.r.t. h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times$ V_h^1	1.84×10^{-2}	9.08×10^{-3}	4.20×10^{-3}	1.96×10^{-3}	1.33×10^{-3}
(3.14) – $V_h^1\times V_h^1$	1.30×10^{-1}	6.84×10^{-2}	4.10×10^{-2}	2.57×10^{-2}	1.57×10^{-2}
(3.14) – $V_h^2\times V_h^1$	3.70×10^{-2}	1.94×10^{-2}	9.69×10^{-3}	4.88×10^{-3}	2.36×10^{-3}
(3.14) – $V_h^3\times V_h^1$	1.72×10^{-2}	8.38×10^{-3}	3.65×10^{-3}	1.67×10^{-3}	1.73×10^{-3}

TABLE 11. (**Ex2**); $\|(u-u_h)_t(\cdot,0)\|_{H^{-1}(0,1)}/\|(u_t(\cdot,0)\|_{H^{-1}(0,1)}$ w.r.t. h.

h	1.57×10^{-1}	8.22×10^{-2}	4.03×10^{-2}	2.29×10^{-2}	1.25×10^{-2}
(5.3) – HCT $\times$ V_h^1	6.20×10^{-1}	4.28×10^{-1}	3.68×10^{-1}	3.26×10^{-1}	4.54×10^{-1}
(3.14) – $V_h^1\times V_h^1$	6.47×10^{-1}	4.35×10^{-1}	3.63×10^{-1}	3.21×10^{-1}	3.05×10^{-1}
(3.14) – $V_h^2\times V_h^1$	3.01×10^{-1}	2.85×10^{-1}	2.96×10^{-1}	2.90×10^{-1}	2.94×10^{-1}
(3.14) – $V_h^3\times V_h^1$	3.15×10^{-1}	2.87×10^{-1}	2.98×10^{-1}	2.91×10^{-1}	3.01×10^{-1}

Section 5.1 of [17]. The corresponding solution u belongs to $H^1(\mathcal{M})$ but not $H^2(\mathcal{M})$ and is given by

$$(\mathbf{Ex2})\qquad \begin{cases} u(t,x)=\displaystyle\sum_{k>0}\Big(a_k\cos(k\pi t)+\frac{b_k}{k\pi}\sin(k\pi t)\Big)\sqrt{2}\sin(k\pi t),\\ a_k=\dfrac{4\sqrt{2}}{\pi^2k^2}\sin(\pi k/2),\quad b_k=\dfrac{1}{\pi k}\big(\cos(\pi k/3)-\cos(2\pi k/3)\big),\quad k>0. \end{cases}$$

We define the observation $u_{\mathcal{O}}$ as the restriction over $(0.1,0.3)\times(0,2)$ of the first fifty terms in the previous sum. Tables 9–11 provide the norms

$$\|u-u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})},\quad \|(u-u_h)(0,\cdot)\|_{L^2(0,1)},\quad \|(u-u_h)_t(0,\cdot)\|_{H^{-1}(0,T)}$$

with respect to h for the primal variable, obtained from the formulation (5.3) and from the formulation (3.14) with $p\in\{1,2,3\}, q\in\{1,2\}$ and $q\leqslant p$ (see (3.1)). We use again the dual stabilizer (3.6) with $\gamma=10^{-3}$ and $\gamma^\star=1$. The convergence of the dual variable z_h is reported in Table 8.

Figure 3 (left) depicts the relative error $\|u-u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}\approx\beta\times h^\tau$ with respect to h. We get a convergence rate around one, independent of approximation order or smoothness of the approximation space. More precisely

$$\begin{aligned} &\beta=e^{-0.20},\quad \tau=0.83,\quad (u_h,z_h)\in V_h^1\times V_h^1,\\ &\beta=e^{-0.58},\quad \tau=1.02,\quad (u_h,z_h)\in V_h^2\times V_h^1,\\ &\beta=e^{-0.68},\quad \tau=1.23,\quad (u_h,z_h)\in V_h^3\times V_h^1,\\ &\beta=e^{-0.75},\quad \tau=1.17,\quad (u_h,z_h)\in \mathrm{HCT}\times V_h^1, \end{aligned}\tag{5.6}$$

and confirm, in agreement with Theorem 4.4, that the rate of convergence depends on the regularity of the solution u. Note that the observed rate also here is better than what is predicted by Theorem 4.4. We also check that increasing the order of the space for the dual variable does not improve the accuracy. Moreover, we observe the same property as the first example with respect to the choice of the parameter γ and $\gamma^\star$. In particular, as depicted in Figure 3 (left), a lower value of γ allows to improve the reconstruction of the solution.

We remark that, since the solution u to be reconstructed, develops singularities along characteristic lines starting from the point $x=1/2$ (due to the initial position u_0) and from the points $x=1/3,2/3$ (due to the initial velocity u_1), the adaptative refinement of the mesh mentioned in the previous subsection is of

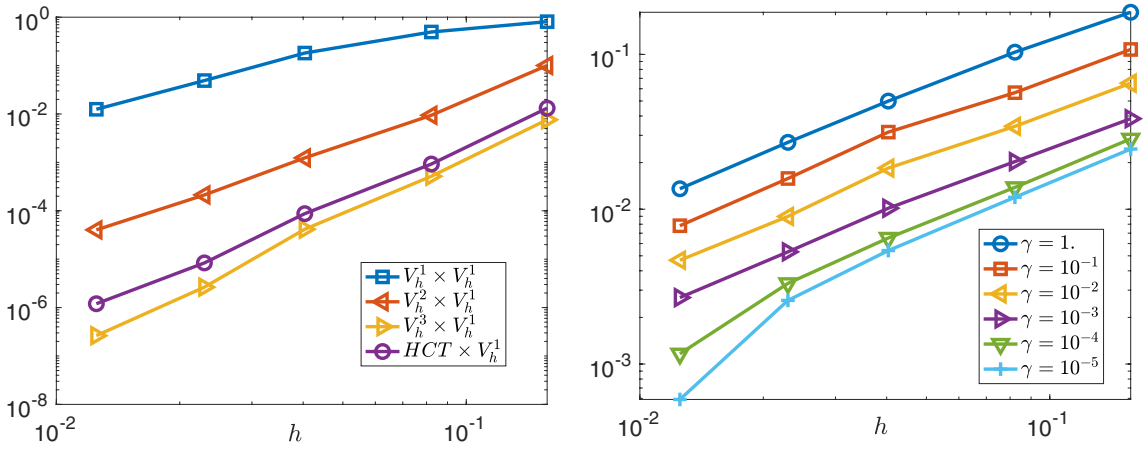

FIGURE 3. (**Ex2**) – *Left*: relative error $\|u - u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}$ with respect to h for various approximation spaces (see Tab. 9); *Right*: relative error $\|u - u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})}$ with respect to h and γ for the approximation $V_h^2 \times V_h^1$.

particular interest here. Using the $V_h^2 \times V_h^1$ approximation, Figure 4 (left) depicts the mesh obtained after ten adaptative refinements based on the local values of gradient of the primal variable u_h. Starting with a coarse mesh composed of 288 triangles and 166 vertex, the final mesh is composed with 12118 triangles and 6213 vertices. We obtain the following values: $\|z_h\|_{L^2(H_0^1)} = 5.15 \times 10^{-5}$; $\|u - u_h\|_{L^2(\mathcal{M})}/\|u\|_{L^2(\mathcal{M})} = 1.74 \times 10^{-3}$; $\|(u-u_h)(0,\cdot)\|_{L^2(0,1)}/\|u(0,\cdot)\|_{L^2(0,1)} = 7.63 \times 10^{-4}$; $\|(u-u_h)_t(0,\cdot)\|_{H^{-1}(0,1)}/\|u_t(0,\cdot)\|_{H^{-1}(0,1)} = 2.93 \times 10^{-1}$ for a CPU time equal to 3.13. The final mesh clearly exhibits the singularities generated by the initial data (u_0, u_1). On the contrary, the refinement strategy coupled with the HCT element does not permit to capture so clearly such singularities, in particular the weaker ones starting from the point $x = 1/3$ and $x = 2/3$ (see [17], Fig. 1).

6. Concluding remarks

We have introduced and analyzed a spacetime finite element approximation of a data assimilation problem for the wave equation. Based on an H^1-approximation that is nonconformal in H^2, the analysis yields error estimates for the natural norm

$$L^\infty(0,T;L^2(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$$

of order h^p, where p is the degree of the polynomials used to describe the primal variable u to be reconstructed. The numerical experiments performed for two initial data, the first one in $H^k \times H^{k-1}$ for all $k \in \mathbb{N}$, the second one in $H^1 \times L^2$, exhibit the efficiency of the method.

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.

In comparison with the formulation introduced in [17], 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, see the definition of s_k in (3.5). These kinds of terms are known from non-conforming approximation of fourth order problems [21]. So the approach can be interpreted as a non-conforming, stabilized version of the method in [17]. The implementation of the stabilized terms is not straightforward, in particular, in higher dimension, and is usually not available in finite element softwares. In such cases one can apply the so-called orthogonal sub-scale stabilization [20], which can be shown to be equivalent, but requires the introduction of additional

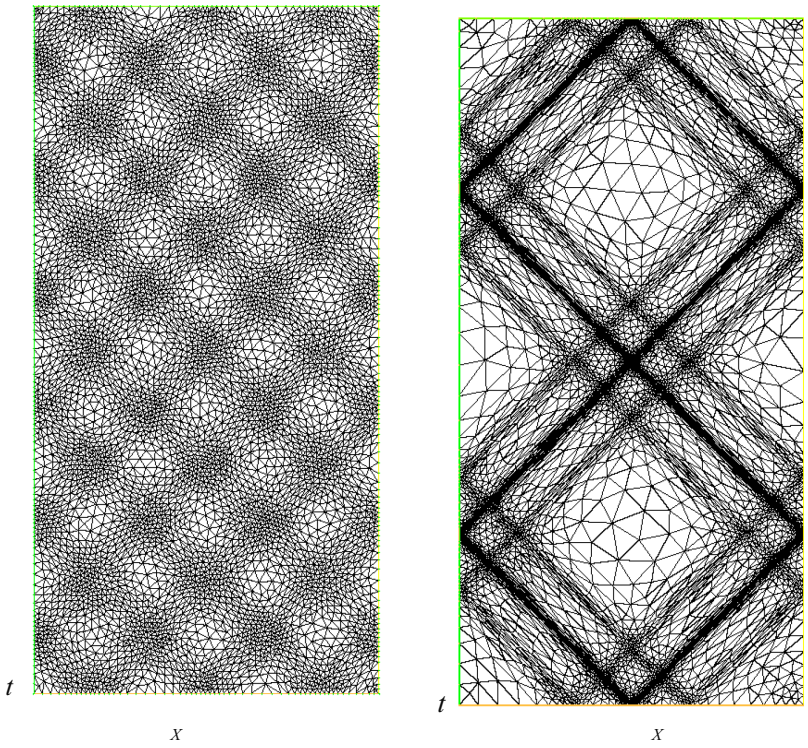

FIGURE 4. Locally refine spacetime meshes for the example 1 (*left*) and the example 2 (*right*).

degrees of freedom, one for each component of the spacetime gradient. Another possible way to circumvent the introduction of the gradient jump terms is to consider non-conforming approximation of the Crouzeix–Raviart type as in [10]. A penalty is then needed on the solution jump instead to control the H^1-conformity error. For the time discretization one could also explore the possibility of using discontinuous Petrov–Galerkin methods (see for instance [23, 45]).

The analysis performed can easily be extended to more general wave equations of the form

$$u_{tt} - \operatorname{div}(a(x)\nabla u) + p(x,t)u = 0$$

with $a \in L^\infty(\Omega, \mathbb{R}^+_\star)$ and $p \in L^\infty(\mathcal{M})$ allowing to consider, through appropriate linearization techniques, data assimilation problem for nonlinear wave equation of the form $u_{tt} - \Delta u + g(u) = 0$. From application viewpoint, it is also interesting to explore if a spacetime approach based on a non conformal H^1-approximation can be efficient to address data assimilation problem from boundary observation. We refer to [18] where a H^2 conformal approximation – similar to (5.3) – is discussed, assuming that the normal derivative $\partial_\nu u \in L^2(\Sigma)$ is available on a part large enough of $\Sigma = (0,T)\times\partial\Omega$. Eventually, the issue of the approximation of controllability problems for the wave equation through non conformal spacetime approach can very likely be addressed as well: remember that the control of minimal $L^2(\mathcal{O}\times(0,T))$ norm for the initial data $(z_0, z_1) \in H^1_0(\Omega)\times L^2(\Omega)$ is given by $v = u1_{\mathcal{O}}$ where u together with z is the saddle point of the following Lagrangian $\widehat{\mathcal{L}} : V \times L^2(0,T;H^1_0(\Omega)) \to \mathbb{R}$

$$\begin{aligned}\widehat{\mathcal{L}}(u,z) = \frac{1}{2}\|u\|^2_{\mathcal{O}} + \frac{\gamma}{2}\|\Box u\|^2_{L^2(0,T;H^{-1}(\Omega))} - \int_0^T (z,\Box u)_{H^1_0(\Omega),H^{-1}(\Omega)}\mathrm{d}t \\ + \langle u_t(\cdot,0), z_0\rangle_{H^{-1}(\Omega),H^1_0(\Omega)} - \langle u(\cdot,0), z_1\rangle_{L^2(\Omega)},\end{aligned}$$

very close to (5.2). We refer to [15].

Acknowledgements. EB acknowledges funding by EPSRC grant EP/P01576X/1. AF acknowledges funding by EPSRC grant EP/P01593X/1. LO acknowledges funding by EPSRC grants EP/P01593X/1 and EP/R002207/1.

References

[1] S. Acosta and C. Montalto, Multiwave imaging in an enclosure with variable wave speed. *Inverse Prob.* **31** (2015) 065009.

[2] D. Auroux and J. Blum, Back and forth nudging algorithm for data assimilation problems. *C. R. Math. Acad. Sci. Paris* **340** (2005) 873–878.

[3] C. Bardos, G. Lebeau and J. Rauch, Un exemple d'utilisation des notions de propagation pour le contrôle et la stabilisation de problèmes hyperboliques. *Rend. Sem. Mat. Univ. Politec. Torino* (Special Issue) (1988) 11–31.

[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] M. Bernadou and K. Hassan, Basis functions for general Hsieh–Clough–Tocher triangles, complete or reduced. *Int. J. Numer. Methods Eng.* **17** (1981) 784–789.

[6] L. Bourgeois, D. Ponomarev and J. Dardé, An inverse obstacle problem for the wave equation in a finite time domain. *Inverse Prob. Imaging* **13** (2019) 377–400.

[7] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods, Vol. 15 of *Texts in Applied Mathematics*, 3rd edition. Springer, New York (2008).

[8] E. Burman, Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: elliptic equations. *SIAM J. Sci. Comput.* **35** (2013) A2752–A2780.

[9] E. Burman, Error estimates for stabilized finite element methods applied to ill-posed problems. *C. R. Math. Acad. Sci. Paris* **352** (2014) 655–659.

[10] E. Burman, A stabilized nonconforming finite element method for the elliptic Cauchy problem. *Math. Comput.* **86** (2017) 75–96.

[11] E. Burman and L. Oksanen, Data assimilation for the heat equation using stabilized finite element methods. *Numer. Math.* **139** (2018) 505–528.

[12] E. Burman, J. Ish-Horowicz and L. Oksanen, Fully discrete finite element data assimilation method for the heat equation. *ESAIM: M2AN* **52** (2018) 2065–2082.

[13] E. Burman, A. Feizmohammadi and L. Oksanen, A fully discrete numerical control method for the wave equation. *SIAM J. Control Optim.* Preprint arXiv:1903.02320 (2019).

[14] E. Burman, A. Feizmohammadi and L. Oksanen, A finite element data assimilation method for the wave equation. *Math. Comput.* **89** (2020) 1681–1709.

[15] C. Castro, N. Cîndea and A. Münch, Controllability of the linear one-dimensional wave equation with inner moving forces. *SIAM J. Control Optim.* **52** (2014) 4027–4056.

[16] O. Chervova and L. Oksanen, Time reversal method with stabilizing boundary conditions for photoacoustic tomography. *Inverse Prob.* **32** (2016) 125004.

[17] N. Cîndea and A. Münch, Inverse problems for linear hyperbolic equations using mixed formulations. *Inverse Prob.* **31** (2015) 075001.

[18] 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.

[19] C. Clason and M.V. Klibanov, The quasi-reversibility method for thermoacoustic tomography in a heterogeneous medium. *SIAM J. Sci. Comput.* **30** (2007/2008) 1–23.

[20] R. Codina, Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods. *Comput. Methods Appl. Mech. Eng.* **190** (2000) 1579–1599.

[21] G. Engel, K. Garikipati, T.J.R. Hughes, M.G. Larson, L. Mazzei and R.L. Taylor, Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity. *Comput. Methods Appl. Mech. Eng.* **191** (2002) 3669–3750.

[22] A. Ern and J.-L. Guermond, Theory and Practice of Finite Elements. In: Vol. 159 of *Applied Mathematical Sciences*. Springer, New York (2004).

[23] J. Ernesti and C. Wieners, Space-time discontinuous Petrov–Galerkin methods for linear wave equations in heterogeneous media. *Comput. Methods Appl. Math.* **19** (2019) 465–481.

[24] S. Ervedoza and E. Zuazua, The wave equation: control and numerics. In: Vol. 2048 of *Lecture Notes in Mathematics. Control of Partial Differential Equations*. Springer, Berlin-Heidelberg (2012) 245–339.

[25] S. Ervedoza and E. Zuazua, Numerical Approximation of Exact Controls for Waves. *Springer Briefs in Mathematics*. Springer, New York (2013).

[26] S. Ervedoza, A. Marica and E. Zuazua, Numerical meshes ensuring uniform observability of one-dimensional waves: construction and analysis. *IMA J. Numer. Anal.* **36** (2016) 503–542.

[27] R. Glowinski and J.-L. Lions, Exact and Approximate Controllability for Distributed Parameter Systems. In: *Acta Numerica*. Cambridge University Press, Cambridge (1994) 269–378.

[28] G. Haine and K. Ramdani, Reconstructing initial data using observers: error analysis of the semi-discrete and fully discrete approximations. *Numer. Math.* **120** (2012) 307–343.

[29] F. Hecht, New development in Freefem++. *J. Numer. Math.* **20** (2012) 251–265.

[30] J.A. Infante and E. Zuazua, Boundary observability for the space semi-discretizations of the 1-D wave equation. *M2AN* **33** (1999) 407–438.

[31] M.V. Klibanov and J. Malinsky, Newton–Kantorovich method for three-dimensional potential inverse scattering problem and stability of the hyperbolic Cauchy problem with time-dependent data. *Inverse Prob.* **7** (1991) 577–596.

[32] M. Klibanov and Rakesh, Numerical solution of a time-like Cauchy problem for the wave equation. *Math. Methods Appl. Sci.* **15** (1992) 559–570.

[33] P. Kuchment and L. Kunyansky, Mathematics of thermoacoustic tomography. *Eur. J. Appl. Math.* **19** (2008) 191–224.

[34] R. Lattès and J.-L. Lions, Méthode de quasi-réversibilité et applications. In: Vol. 15 of *Travaux et Recherches Mathématiques*, Dunod, Paris (1967).

[35] 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.

[36] L. Miller, Resolvent conditions for the control of unitary groups and their approximations. *J. Spectr. Theory* **2** (2012) 1–55.

[37] 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.

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

[39] L.V. Nguyen and L.A. Kunyansky, A dissipative time reversal technique for photoacoustic tomography in a cavity. *SIAM J. Imaging Sci.* **9** (2016) 748–769.

[40] J. Nitsche, Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. [*Collection of articles dedicated to Lothar Collatz on his sixtieth birthday.*] *Abh. Math. Sem. Univ. Hamburg* **36** (1971) 9–15.

[41] K. Ramdani, M. Tucsnak and G. Weiss, Recovering and initial state of an infinite-dimensional system using observers. *Automatica J. IFAC* **46** (2010) 1616–1625.

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

[43] P. Stefanov and G. Uhlmann, Thermoacoustic tomography with variable sound speed. *Inverse Prob.* **25** (2009) 075011.

[44] P. Stefanov and Y. Yang, Multiwave tomography in a closed domain: averaged sharp time reversal. *Inverse Prob.* **31** (2015) 065007.

[45] O. Steinnach and M. Zank, A stabilized space-time finite element method for the wave equation. Technische Universität Graz Report 2018/5 (2018) 1–27.

[46] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems. In: Vol. 25 of *Springer Series in Computational Mathematics*. Springer, Berlin-Heidelberg (1997).

[47] L.V. Wang, Photoacoustic Imaging and Spectroscopy. CRC Press (2009).

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