A FINITE ELEMENT DATA ASSIMILATION METHOD FOR THE WAVE EQUATION

ERIK BURMAN, ALI FEIZMOHAMMADI, AND LAURI OKSANEN

Abstract. We design a primal-dual stabilized finite element method for the numerical approximation of a data assimilation problem subject to the acoustic wave equation. For the forward problem, piecewise affine, continuous, finite element functions are used for the approximation in space and backward differentiation is used in time. Stabilizing terms are added on the discrete level. The design of these terms is driven by numerical stability and the stability of the continuous problem, with the objective of minimizing the computational error. Error estimates are then derived that are optimal with respect to the approximation properties of the numerical scheme and the stability properties of the continuous problem. The effects of discretizing the (smooth) domain boundary and other perturbations in data are included in the analysis.

1. Introduction

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

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

The initial data u_0, u_1 are assumed to be a priori unknown functions, but it is assumed that we have the following additional data:

$$q = u|_{(0,T)\times\omega}, \tag{1.2}$$

where $\omega \subset \overline{\Omega}$ is open. The data assimilation problem then reads:

(DA) Find u_0 and u_1 given q.

Due to the finite speed of propagation, the length of time interval T needs to be large enough in order for (DA) to have unique solution. Assuming that

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

it follows from Holmgren's unique continuation theorem that (DA) is uniquely solvable. Here $\mathrm{dist}(x,\omega) = \min\{\mathrm{dist}(x,y) \,|\, y \in \overline{\omega}\}$ and $\mathrm{dist}(x,y)$ is the distance function in Ω, defined as the infimum over the lengths of continuous paths in Ω, joining x and y.

EB acknowledges funding by EPSRC grants EP/P01576X/1 and EP/P012434/1.
AF acknowledges funding by EPSRC grant EP/P01593X/1.
LO acknowledges funding by EPSRC grants EP/P01593X/1 and EP/R002207/1.

The problem (DA) can be exponentially ill-posed under the assumption (1.3). In order to avoid such severely ill-posed cases, we will suppose that the geometric control condition holds in the sense of [38]. This means roughly speaking that any billiard trajectory intersects ω before time T. A billiard trajectory leaving from a point in Ω consists of line segments that are joined together at points on $\partial\Omega$, with directions satisfying Snell's law of reflection. However, the exact formulation of the geometric control condition requires also a consideration of trajectories gliding along $\partial\Omega$. It is well-known that the geometric control condition characterizes the cases where the problem (DA) is stable, and in a slightly different context, the characterization originates from [5].

We will analyse the convergence of a finite element method that gives an approximate solution to (DA). Our method is based on piecewise affine elements in space and the use of backward finite differences in time. The main contribution of the paper is to show that, when complemented with a suitable stabilization, this standard, low order discretization yields a convergence that is optimal with respect to approximation and the stability of the continuous problem. Indeed, we are solving a linear problem and the truncation error of the time discretization used is of order $\mathcal{O}(\tau)$, where τ denotes the size of the time-steps. Consequently, the error estimates can not be better than $\mathcal{O}(\tau)$, which is what we obtain (see Theorem 4.6).

The stabilization terms that appear in our numerical scheme are carefully designed, balancing the numerical stability, the approximation properties of the scheme and the stability of the continuous problem. The analysis also considers the effect of discretizing the smooth domain, as well as other perturbations of the data. The resulting scheme is of the form of a time-space primal-dual system. The forward equation is independent of the dual. Therefore the gradient can be computed by a forward solve, followed by a dual backward solve, for steepest descent type iterative solving.

We hope that the present paper can act as a starting point for exploration of more applied, but also more advanced, stabilized finite element methods. Indeed although stabilization terms herein are tailored for the low order method, the approach is general and can be extended to other finite element methods. For instance, it might be desirable to use high order elements in space and a more sophisticated discretization in time in order to reduce the numerical dissipation.

1.1. **Previous literature.** There are two extensive traditions of research that are closely related to the problem (DA). A variation of (DA) arises as a mathematical model for the medical imaging technique called photoacoustic tomography (PAT), and works related to PAT form one of the two traditions. We refer to [28, 40, 53] for physical aspects of PAT, and to [32, 50] for mathematical reviews.

The problem (DA) models wave propagation in a cavity Ω, whereas the classical PAT problem is formulated in $\mathbb{R}^3$. However, the papers [1, 14, 35, 48] study the PAT problem in a cavity. All these papers consider methods based on using iterative time reversal for the continuum wave equation, an approach that originates from [47], and none of them consider the issues arising from discretization.

The second tradition draws from control theory, and it uses so-called Luenberger observers. The data assimilation problem (DA) arises as the dual problem of a control problem, and analysis of the latter is typically reduced to the analysis of (DA) by using the Hilbert uniqueness method originating from [41].

A Luenberger observers based algorithm was first analysed in a finite dimensional ODE context in [2]. An abstract version of the method, applicable to the problem (DA), was introduced in [45]. The two traditions have a significant overlap. For instance, as pointed out in [14], the result [35] on the PAT problem fits in the abstract setting of [45]. In particular, the methods in both the traditions can be formulated as Neumann series in infinite dimensional spaces.

The paper [23] studies a discretization of a Luenberger observers based algorithm. The error estimate in [23] depends linearly on the point of truncation of the Neumann series (see Theorem 1 there), and this ultimately leads to logarithmic convergence with respect to the mesh size. The issue with the truncation can be avoided if a stability estimate is available on a scale of discrete spaces. Such estimates were first derived in [27] and we refer the reader to the survey articles [54, 21], as well as the recent paper [20] for more details. Optimal-in-space discrete estimates can be derived from continuous estimates [43], however, spacetime optimal discrete estimates are known only for specific situations. We refer to the monograph [19], see in particular Chapter 5 on open problems, for a detailed discussion of the truncation issue in the context of the exact controllability problem, dual to (DA).

The closest work to the present paper is [15]. There, two finite element methods for (DA) are considered: one of them is stabilized while the other is not. The method without stabilization is shown to converge only under the further assumption that certain discrete inf-sup condition holds, see (42) there. On the other hand, the stabilized method is shown to converge to the exact solution only under a further regularity assumption on an auxiliary Lagrangian multiplier, see λ in Proposition 2 there. Under this assumption, it is then shown in the $1+1$-dimensional case, that the stabilized method converges with quadratic rate when the Bogner-Fox-Schmit C^1-elements, with third order polynomials, are used in spacetime rectangles.

The data assimilation problem (DA) can also be solved using the quasi-reversibility method. This method originates from [37], and it has been applied to data assimilation problems subject to the wave equation in [29, 30], and more recently to the PAT problem in [16]. Another interesting application is given in the recent preprint [6]. There the authors solve an obstacle detection problem by using a level set method together with the quasi-reversibility method applied to a variant of (DA).

The quasi-reversibility method introduces an auxiliary Tikhonov type regularization parameter. When deriving a rate of convergence for the method, this parameter needs to be chosen as a function of the mesh size h. In [16] the regularization parameter is called ε, and by balancing the estimates in Theorems 3.3, 4.6 and 5.3 there, we are lead to the choice $\varepsilon(h) = h^{2/3}$. This gives the convergence rate $h^{2/3}$ for the quasi-reversibility method [16].

Let us also mention that the method we propose can be seen as an instance of the well known 4DVAR algorithm for data assimilation [46, 51]. From this point of view the stabilization terms in the finite element method can be interpreted as a Tikhonov regularization of the discrete equations, where the analysis allows us to determine a regularization parameter that simultaneously balances both the errors from discretization and from perturbations. For the continuous equations the stabilization of the initial energy is indeed the natural regularization term in 4DVAR for the wave equation. The upshot here is that the analysis in the fully discrete framework prompts a bespoke stabilization that can not be derived from an analysis where only regularization on the continuous level is considered. We also show in a numerical example that, in general, the regularization of the initial energy is insufficient and that the added terms implied by the analysis improves the order of accuracy of the approximation.

To summarize, the linear convergence rate of our method is superior to that of the Neumann series based methods and the quasi-reversibility method. Contrary to [15], we also prove optimal convergence rate. The convergence proof is based on using the continuum estimates, and the only geometric assumption needed is the sharp geometric control condition. Finally, the method uses a very simple discretization of the spacetime, and it is likely that the ideas presented here can be adapted to various other discretizations.

As already mentioned above, the dual problem to (DA) is the exact controllability problem for the wave equation, and we refer to [22] for an excellent summary of early computational studies addressing the need to regularize the dual problem. The monograph [21] provides a thorough review of regularization via filtering of spurious high frequency modes, arising from discretization. We mention also our work on the exact controllability problem [10], that is a follow-up of the present work, and the recent similar studies [15, 44].

Related to the exact controllability problem there is a large body of literature on optimal control problems. A basic example of such a problem is a regularized version of the exact controllability problem, called the approximate controllability problem, where the regularization is viewed as describing the cost of control. We refer to the thesis [31] for an introduction to optimal control problems for the wave equation from the computation point of view. While most of the computational literature on optimal control problems focuses on elliptic and parabolic equations, see for example the monograph [24], we mention [36] where domain decomposition methods for optimal control problems for the wave equation are discussed. For recent numerical studies of problems close to the approximate controllability problem for the wave equation see [18, 52].

Let us also mention that the method in the present paper draws from our experience on stabilized finite element methods for the elliptic Cauchy problem [8, 9], and other types of data assimilation problems, see [12] for elliptic and [11, 13] for parabolic cases. In [12] we considered the Helmholtz equation. The convergence estimate there is explicit in the wave number, and exhibits a hyperbolic character in the sense that it relies on a convexity assumption that can viewed as a particular local version of the geometric control condition.

2. Continuum Estimates

The main aim of this section is to recall a continuum observability estimate for the wave operator under some geometric assumptions on the observable domain $\mathcal{O} = (0,T) \times \omega$. In order to state these geometric conditions we will need the following definition. We refer the reader to [38] for the definition of compressed generalized bicharacteristics.

Definition 2.1 (See [5],[38])**.** *We say that $\mathcal{O} \subset \mathcal{M}$ satisfies the geometric control condition in $\mathcal{M}$, if every compressed generalized bicharacteristic ${}^b\gamma(s) = (t(s), x(s), \tau(s), \xi(s))$ intersects the set $\mathcal{O}$ for some $s \in \mathbb{R}$.*

With this definition in mind, we can state the continuum estimate that is used to derive a convergence rate for our finite element method:

Theorem 2.2. *Suppose $\mathcal{M} = (0,T) \times \Omega$ where Ω is a domain with smooth boundary. Let $\omega \subset \overline{\Omega}$ and assume that $\mathcal{O} = (0,T) \times \omega$ satisfies the geometric control condition. If $u \in L^2(\mathcal{M})$ with $u(0,\cdot) \in L^2(\Omega)$, $\partial_t u(0,\cdot) \in H^{-1}(\Omega)$, $u|_{(0,T)\times\partial\Omega} \in L^2((0,T) \times \partial\Omega)$ and $\Box u \in H^{-1}(\mathcal{M})$, then*

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

and there exists a constant $\kappa > 0$ depending on the geometry, such that the following estimate holds:

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

Theorem 2.2 is a consequence of the following homogeneous version:

Theorem 2.3 (Observability estimate)**.** *Let $\mathcal{O}$ satisfy the geometric control condition. There exists a constant $C_{obs} > 0$ such that for any initial data $w|_{t=0} = g_1 \in L^2(\Omega)$ and $\partial_t w|_{t=0} = g_2 \in H^{-1}(\Omega)$,* *the corresponding unique weak solution w to $\Box w = 0$, $w|_{(0,T)\times\partial\Omega} = 0$ with*

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

satisfies:

$$\|g_1\|_{L^2(\Omega)} + \|g_2\|_{H^{-1}(\Omega)} \leqslant C_{obs}\|w\|_{L^2(\mathcal{O})}.$$

Theorem 2.3 is a classical result that yields an interior observability estimate under the geometric control condition. The proof of the theorem uses propagation of singularities for the wave equation and only works for smooth geometries. The geometric control condition is essentially a necessary and sufficient condition for obtaining the observability estimate and roughly states that all light rays in $\mathcal{M}$ must intersect $\mathcal{O}$ taking into account reflections at the boundary [5]. We refer the reader to [38, Proposition 1.2] for a proof of this theorem using a combination of the study of semiclassical defect measures and propagation of singularities. One can also look at [5, Theorem 3.3] for an alternative proof using propagation of singularites. The paper [5] deals with boundary observability but the proof can be applied to obtain interior observability as well. We omit rewriting these proofs here as they are well known in the literature. Let us remark at this point that there is a stronger geometric condition on the observable domain $\mathcal{O}$ known as the $\Gamma-$ condition which is much simpler to verify in general. We recall the $\Gamma-$condition defined as follows

Definition 2.4. *For each $x_0 \notin \Omega$, Let $\Gamma_{x_0} := \{x \in \partial\Omega \,|\, (x - x_0)\cdot\nu(x) \geqslant 0\}$. We say that $\mathcal{O} = (0,T)\times\omega$ satisfies the $\Gamma-$condition if*

$$\exists x_0 \notin \Omega, \quad \exists \delta > 0 \quad \text{such that} \quad \mathcal{N}_\delta(\Gamma_{x_0}) \cap \Omega \subset \omega,$$

$$T > 2\sup_{x\in\Omega} |x - x_0|,$$

where $\mathcal{N}_\delta(\Gamma_{x_0}) := \{y \in \mathbb{R}^n \,|\, |y - x| < \delta \quad \text{for some} \quad x \in \Gamma_{x_0}\}$.

It is known that the $\Gamma-$ condition implies the geometric control condition (see for example [42]). In essence, the $\Gamma-$condition roughly requires T and $\bar{\omega}\cap\overline{\partial\Omega}$ to be relatively large. Although not as sharp as the geometric control condition, the advantage of the $\Gamma-$ condition lies in its applicability in the presence of non-smooth geometries and the explicit derivation of the observability constant κ in Theorem 2.3. For an alternative proof of Theorem 2.3 in the case that $\mathcal{O}$ satisfies the $\Gamma-$condition, we refer the reader to [17, Theorem 2.2]. One can also use the Carleman estimate [4, Theorem 1.1] to derive this estimate although in this case one has to shift the Sobolev estimates.

A key ingredient in deriving the Lipschitz stability result in this paper is a corollary of the observability estimate for the wave equation as stated in Theorem 2.2. In the remainder of this section, we will show that Theorem 2.2 indeed follows from the observability estimate. To this end, we will need the following lemma concerning solutions to the mixed Dirichlet-Cauchy problem for the wave equation with weak Sobolev norms. We refer the reader to [39, Theorem 2.3] together with Remark 2.8 in that paper for the proof.

Lemma 2.5. *Let Ω be a bounded domain with smooth boundary. Suppose $(u_0, u_1, f, h) \in X$ where $X = L^2(\Omega) \times H^{-1}(\Omega) \times H^{-1}(\mathcal{M}) \times L^2((0,T)\times\partial\Omega)$ with the usual product topology. Then the equation (1.1) with a source term f on the right hand side and a lateral boundary Dirichlet data h, admits a unique solution*

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

Moreover, the linear mapping that maps (u_0, u_1, f, h) to u is continuous:

$$\|u\|_Y \leqslant C_e\,\|(u_0, u_1, f, h)\|_X,$$

where $C_e > 0$ is a constant depending only on the geometry $\mathcal{M}$.

We are now ready to show the derivation of Theorem 2.2 from Theorem 2.3.

Proof of Theorem 2.2. Let us consider the vector valued function $v := [v_1 \; v_2]^T$ with $v_i \in L^2(\mathcal{M})$ for $i \in \{1, 2\}$ defined as the solution to the following separable system of PDEs:

$$\begin{cases} \Box v = [\Box u \; 0]^T \\ v(t,x) = [u \; 0]^T \quad \forall x \in \partial\Omega, \forall t \in [0,T] \\ v(0,x) = [0 \; u(0,\cdot)]^T \quad \forall x \in \Omega \\ \partial_t v(0,x) = [0 \; \partial_t u(0,\cdot)]^T \quad \forall x \in \Omega. \end{cases}$$

Note that if $w := u - (v_1 + v_2)$, then $w \in L^2(\mathcal{M})$ and w satisfies the homogeneous wave equation

$$\begin{cases} \Box w = 0 \\ w(t,x) = 0 \quad \forall x \in \partial\Omega, \forall t \in [0,T] \\ w(0,x) = 0, \quad \forall x \in \Omega \\ \partial_t w(0,x) = 0 \quad \forall x \in \Omega. \end{cases}$$

By Lemma 2.5, we have $w = 0$, which implies that $u = v_1 + v_2$. Since $\mathcal{O}$ satisfies the geometric control condition, the observability estimate in Theorem 2.3 holds for the function v_2 and together with Lemma 2.5 we have that for all $t \in [0,T]$:

$$\|v_2(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t v_2(t,\cdot)\|_{H^{-1}(\Omega)} \leqslant C_{obs}\|v_2\|_{L^2(\mathcal{O})}.$$

Similarly, applying Lemma 2.5 to the function v_1 implies that:

$$\|v_1(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t v_1(t,\cdot)\|_{H^{-1}(\Omega)} \leqslant C_e \left(\|\Box u\|_{H^{-1}(\mathcal{M})} + \|u\|_{L^2((0,T)\times\partial\Omega)}\right).$$

Finally, combining the above estimates, we deduce that:

$$\begin{aligned} \|u(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t u(t,\cdot)\|_{H^{-1}(\Omega)} &\leqslant \|v_1(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t v_1(t,\cdot)\|_{H^{-1}(\Omega)} + \|v_2(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t v_2(t,\cdot)\|_{H^{-1}(\Omega)} \\ &\leqslant (C_e + C_{obs})\left(\|\Box u\|_{H^{-1}(\mathcal{M})} + \|u\|_{L^2((0,T)\times\partial\Omega)} + \|v_2\|_{L^2(\mathcal{O})}\right) \\ &\leqslant (C_e + C_{obs})\left(\|\Box u\|_{H^{-1}(\mathcal{M})} + \|u\|_{L^2((0,T)\times\partial\Omega)} + \|u - v_1\|_{L^2(\mathcal{O})}\right) \\ &\leqslant \kappa\left(\|\Box u\|_{H^{-1}(\mathcal{M})} + \|u\|_{L^2((0,T)\times\partial\Omega)} + \|u\|_{L^2(\mathcal{O})}\right). \end{aligned}$$

where $\kappa = (C_e + C_{obs})(1 + T^{\frac{1}{2}} C_e)$. ◻

3. Discretization

Let us begin with a brief discussion of the overall discretization approach employed in this paper. We consider the wave equation (1.1) and the preliminary Lagrangian functional

$$\mathcal{L}(u,z) = \frac{1}{2}\|u - q\|^2_{L^2((0,T)\times\omega)} + \int_{\mathcal{M}} (\partial_t^2 u)\, z + \nabla u \cdot \nabla z \, dxdt. \tag{3.1}$$

The Euler-Lagrange equations for $\mathcal{L}$ can be written as follows

$$\begin{aligned} \langle \partial_u \mathcal{L}(u,z), v\rangle &= \int_0^T \int_\omega (u - q)v \, dtdx + \int_{\mathcal{M}} (\partial_t^2 v)\, z + \nabla v \cdot \nabla z \, dtdx = 0, \\ \langle \partial_z \mathcal{L}(u,z), w\rangle &= \int_{\mathcal{M}} (\partial_t^2 u)\, w + \nabla u \cdot \nabla w \, dtdx = 0 \end{aligned}$$

for all v, w. It is clear that if u is equal to the unique solution to (1.1)-(1.2) and $z \equiv 0$, then these Euler-Lagrange equations are satisfied. This simple idea outlines the overall approach in this paper. We will employ a discrete Lagrangian functional whose critical points will converge to the unique solution to the continuum problem. However, as the term $\int_0^T \int_\omega (u-q)v \, dtdx$ does not seem to give enough stability for the discrete problem to converge, we will add certain regularization terms in the discrete setting. The design of these terms is driven by numerical stability and the stability of the continuous problem, with the objective of minimizing the computational error. In the final section of the paper we will briefly discuss the possibility of removing some of these regularization terms.

Let us now present the discretization of (1.1)-(1.2). We will first consider a family of polyhedral domains Ω_h approximating Ω and similarly let ω_h denote a family of domains approximating ω. Let $\mathcal{T}_h$ be a conforming triangulation of the polyhedral domain Ω_h. Let $h_K = diam(K)$ be the local mesh parameter and $h = max_{K\in\mathcal{T}_h} h_K$ the mesh size. We assume that the family of triangulations $\mathcal{T}_h$ is quasi uniform. Let V_h be the standard space of piecewise affine continuous finite elements satisfying the zero boundary condition,

$$V_h = \{v \in H_0^1(\Omega_h); v|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\}.$$

We assume that the approximate geometries Ω_h and ω_h are sufficiently close to Ω and ω in the following sense,

$$\mathrm{dist}(x, \partial\Xi) \leqslant c\, h^2 \quad \forall x \in \partial\Xi_h, \quad \Xi = \Omega \text{ or } \Xi = \omega, \tag{3.2}$$

where $c > 0$ is a constant that is independent of h. This is possible for domains Ω, ω with smooth boundary (see for example [7]). We have the following lemma:

Lemma 3.1. *(See [7, Lemma 2]) Let the condition (3.2) be satisfied. Then for all $v \in H^1(\Omega \cup \Omega_h)$ the following estimate holds:*

$$\int_{(\Omega\setminus\Omega_h)\cup(\Omega_h\setminus\Omega)} |v(x)|^2 \, dx \leqslant c\, h^2 \left(\int_{\partial\Omega} |v(x)|^2 \, ds + h^2 \int_\Omega |\nabla v(x)|^2 \, dx \right),$$

where $c > 0$ does not depend on h.

Remark 1. *Throughout the rest of the paper and for the sake of convenience, we use the uniform notation c to denote a generic constant that depends only on the geometry $\mathcal{M}$ and is independent of the mesh parameter h. This is useful in the proofs of lemmas and propositions where keeping track of uniform constants is of no particular interest.*

Following [13] we first discretize in space only. We may then write a semi-discrete finite element formulation of the problem as follows. Find $u \in \mathcal{C}^2(0, T; V_h)$, subject to (1.2), such that

$$(\partial_t^2 u, v)_h + a_h(u, v) = 0, \qquad \forall v \in V_h,$$

where

$$(u, v)_h = \int_{\Omega_h} uv \, dx, \qquad a_h(u, v) = \int_{\Omega_h} \nabla u \cdot \nabla v \, dx.$$

We also define

$$(u, v)_\Omega = \int_\Omega uv \, dx, \qquad a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

Let $N \in \mathbb{N}$ and $\tau > 0$ satisfy $N\tau = T$ and define $t_n = n\tau$. Furthermore, define for each discrete function $u = (u^n)_{n=0}^N \in V_h^{N+1}$,

$$\partial_\tau u^n = \frac{u^n - u^{n-1}}{\tau} \quad \text{for} \quad n = 1, \dots, N \qquad \partial_\tau^2 u^n = \frac{u^n - 2u^{n-1} + u^{n-2}}{\tau^2} \quad \text{for} \quad n = 2, \dots, N.$$

It is natural to assume that the two discretization scales τ and h should be comparable in size. We will therefore assume throughout the paper that $\tau = \mathcal{O}(h)$.

To take into account the mismatch between Ω_h and Ω, (and between ω_h and ω), we use the stable extension operator [49], $E : H^s(\mathcal{M}) \to H^s(\mathcal{M}_h)$, $s \geq 0$ with $\mathcal{M}_h := (0,T) \times (\Omega \cup \Omega_h)$ and set $q^e := Eq$. With some slight abuse of notation, we will write q to denote the extended function q^e, where no confusion is implied. Now, consider the Lagrangian functional

$$\mathcal{L} : V_h^{N+1} \times V_h^{N-1} \to \mathbb{R}$$

defined by:

$$\begin{aligned}
\mathcal{L}(u,z) &= \mathcal{L}_0(u,z) + \mathcal{L}_1(u),\\
\mathcal{L}_0(u,z) &= \frac{\tau}{2}\sum_{n=1}^{N} \|u^n - \tilde{q}^n\|_{\omega_h}^2 + G(u,z),\\
\mathcal{L}_1(u) &= \frac{1}{2}\left(\|h\nabla u^1\|_h^2 + \|h\partial_\tau u^1\|_h^2 + \|h\nabla\partial_\tau u^1\|_h^2 + \|h\nabla\partial_\tau u^N\|_h^2 + \tau\sum_{n=2}^{N}\|\tau\nabla\partial_\tau u^n\|_h^2 \right),\\
G(u,z) &= \tau\sum_{n=2}^{N}((\partial_\tau^2 u^n, z^n)_h + a_h(u^n, z^n)),
\end{aligned} \tag{3.3}$$

where

$$\tilde{q}^n = q(t_n) + \delta q(t_n) \quad \text{for} \quad n = 1, \ldots, N$$

and δq denotes some perturbation (or noise) in our observable data.

Let us make a few remarks about the discrete Lagrangian (3.3). The term $\mathcal{L}_0(u,z)$ denotes the discrete analogue of the continuum Lagrangian (3.1). The functional $\mathcal{L}_1(u)$ denotes the discrete regularization (stabilization) terms that formally converge to zero in the limit as $h \to 0$. These regularizing terms are designed with the goal of minimizing the errors that arise in the numerical approximation of (1.1)-(1.2). Stabilization in the initial energy seems natural to us since this is the piece of information that is missing when compared to a typical initial boundary value problem for the wave equation. This and the regularization in the cross derivatives was inspired by our previous work on a unique continuation problem for the heat equation [13]. The regularization in the cross derivative also appears in [3]. For a further discussion on the discrete regularization terms we refer the reader to Section 6.2.

The Euler-Lagrange equations for the Lagrangian functional $\mathcal{L}$ read as follows:

$$\langle D_u\mathcal{L}, v\rangle + \langle D_z\mathcal{L}, w\rangle = 0, \quad \forall (v,w) \in V_h^{N+1} \times V_h^{N-1}$$

where $D_u\mathcal{L}$ and $D_z\mathcal{L}$ denote the Fréchet derivatives of the discrete Lagrangian with respect to the variables u and z. The Euler-Lagrange equation, can be recast in the following form:

$$A_1(u,w) = 0 \quad \text{and} \quad A_2((u,z),v) = \tau\sum_{n=1}^{N}(\tilde{q}^n, v^n)_{\omega_h}, \tag{3.4}$$

where the bilinear forms A_1 and A_2 are defined as follows:

$$\begin{aligned} A_1(u,w) &= G(u,w), \\ A_2((u,z),v) &= \tau\sum_{n=1}^{N}(u^n,v^n)_{\omega_h} + G(v,z) + (h\nabla u^1, h\nabla v^1)_h + (h\partial_\tau u^1, h\partial_\tau v^1)_h \\ &\quad + (h\nabla\partial_\tau u^N, h\nabla\partial_\tau v^N)_h + (h\nabla\partial_\tau u^1, h\nabla\partial_\tau v^1)_h + \tau\sum_{n=2}^{N}(\tau\nabla\partial_\tau u^n, \tau\nabla\partial_\tau v^n)_h. \end{aligned} \tag{3.5}$$

The Euler Lagrange equations (3.4) define a system of equations with critical point(s) (u_h, z_h), if they exist. Let us observe that since no regularization is applied to the Lagrange multiplier z,u_h solves the discrete wave equation as can be seen from the first equation in (3.4), while z_h solves a wave equation with a small source term that formally approaches zero as $h \to 0$. Also, the discrete variables u_h and z_h are only weakly coupled, in the sense that A_1 does not depend on z, which allows for solution algorithms that use the classical forward-backward solving approach. In the remainder of this section, we aim to show that the discrete system of equations (3.4) indeed admits a unique solution (u_h, z_h).

3.1. **Inf-sup stability estimate.** We let $z^0 = z^1 = z^{N+1} = z^{N+2} = 0$ and define the following norms and semi-norms:

$$\begin{aligned} |||u|||_R^2 &= \tau\sum_{n=1}^{N}\|u^n\|_{\omega_h}^2 + \|h\nabla u^1\|_h^2 + \|h\partial_\tau u^1\|_h^2 \\ &\quad + \|h\nabla\partial_\tau u^1\|_h^2 + \|h\nabla\partial_\tau u^N\|_h^2 + \tau\sum_{n=2}^{N}\|\tau\nabla\partial_\tau u^n\|_h^2, \\ |||u|||_F^2 &= \tau\sum_{n=2}^{N}(\|\partial_\tau^2 u^n\|_h^2 + \|\partial_\tau u^n\|_h^2 + \|\nabla u^n\|_h^2) + \|\nabla u^N\|_h^2 + \|\partial_\tau u^N\|_h^2, \\ |||z|||_D^2 &= \frac{T}{2}\|z^N\|_h^2 + \frac{\tau}{4}\sum_{n=2}^{N}\|z^n\|_h^2 + \frac{\tau}{2(T+1)^2}\sum_{n=2}^{N}\|\nabla\tilde{z}^n\|_h^2 + \frac{1}{4(T+1)}\|\nabla\tilde{z}^N\|_h^2, \\ |||(u,z)|||_C^2 &= |||u|||_R^2 + \tau\sum_{n=2}^{N}\|z^n\|_h^2. \end{aligned} \tag{3.6}$$

Here $\tilde{z}^n := \tau\sum_{m=0}^{n}(1+m\tau)z^m$. Note that using the Poincaré inequality we have the following:

$$\|\nabla\tilde{z}^n\|_h \geqslant c\|\tilde{z}^n\|_h \qquad n = 1, 2, \ldots, N,$$

where $c > 0$ only depends on Ω. We also note that $|||(\cdot,\cdot)|||_C$ is a norm on V_h^{2N}. Heuristically, the norms signify the following principles. Firstly, $|||\cdot|||_R$ denotes the control that one naturally expects to obtain on the data fitting term and the regularization terms. The norms $|||u|||_F$ and $|||z|||_D$ are related to discrete energy estimates for the wave equation and in a sense measure the stability properties of the forward problem. There is a delicate counter balance on the control of the state variable u and the dual variable z as can be seen through the fact that $|||\cdot|||_F$ is analogous to $H^2(\mathcal{M})$ norm while $|||\cdot|||_D$ is somewhat reminiscent to the $H^{-1}(0,T;H^1(\Omega))$ norm. Finally, $|||\cdot|||_C$ is just a continuity norm that will be justified in Proposition 4.3.

Proposition 3.2. *For all $N \in \mathbb{N}$, $h > 0$, there exists a constant $c > 0$ only depending on $\mathcal{M}$ such that given any $(u, z) \in V_h^{2N}$ there exists $(v, w) \in V_h^{2N}$ satisfying:*

$$c\,(|||u|||_R^2 + h^2|||u|||_F^2 + |||z|||_D^2) \leqslant A_1(u, w) + A_2((u, z), v),$$

$$c\,|||(v, w)|||_C \leqslant |||u|||_R + h|||u|||_F + |||z|||_D.$$

We will start by proving two lemmas that are motivated by the following estimates for the wave operator at the continuum level:

$$\int_{\mathcal{M}} (2T - t)\partial_t w \,\Box w \,dt\,dx \geqslant c\left(\int_{\mathcal{M}} |w|^2\,dt\,dx + \int_{\mathcal{M}} |\nabla w|^2\,dt\,dx, \quad \forall\, w \in \mathcal{C}_c^\infty(\mathcal{M})\right),$$
$$\int_{\mathcal{M}} \mathcal{E}v \,\Box v\,dt\,dx \geqslant c\left(\int_{\mathcal{M}} |v|^2\,dt\,dx + \int_{\mathcal{M}} |\nabla \mathcal{E}v|^2\,dt\,dx, \quad \forall\, v \in \mathcal{C}_c^\infty(\mathcal{M})\right),$$

where $\mathcal{E}v(t,x) = \int_0^t (1+\tau)v(\tau,x)\,d\tau$ and $c > 0$ depends only on T. These estimates are both energy estimates for the wave equation with the latter being weaker in Sobolev scales compared to the former. This delicate counter-balance will be used to derive convergence in the Sobolev scale $\mathcal{C}(0,T;L^2(\Omega)) \cap \mathcal{C}^1(0,T;H^{-1}(\Omega))$ for the state variable u at the expense of the dual variable converging to zero in the weaker Sobolev scale $L^2(\mathcal{M})$.

Lemma 3.3. *Let $u \in V_h^{N+1}$. For $n = 2, 3, \ldots, N$ define $w^n := \partial_\tau^2 u^n + (2T - n\tau)\partial_\tau u^n$. Then:*

$$A_1(u, h^2 w) + |||u|||_R^2 \geqslant c h^2 |||u|||_F^2,$$

where $c > 0$ is independent of the parameter h.

Proof. Recall that $A_1(u, h^2 w) = h^2 G(u, w)$. Now, given the choice of the test function w we have

$$G(u, w) = S_1 + S_2 + S_3 + S_4,$$

where:

$$\begin{aligned}
S_1 &= \tau \sum_{n=2}^{N} \|\partial_\tau^2 u^n\|_h^2,\\
S_2 &= \tau \sum_{n=2}^{N} (\partial_\tau^2 u^n, (2T - n\tau)\partial_\tau u^n)_h\\
&= \tau \sum_{n=2}^{N} (2T - n\tau)(\partial_\tau v^n, v^n)_h = \sum_{n=2}^{N} (2T - n\tau)\frac{1}{2}(\|v^n\|_h^2 - \|v^{n-1}\|_h^2 + \|v^n - v^{n-1}\|_h^2)\\
&\geqslant \frac{T}{4}\|\partial_\tau u_N\|_h^2 + \frac{\tau}{2}\sum_{n=1}^{N} \|\partial_\tau u^n\|_h^2 - T\|\partial_\tau u^1\|_h^2 + \frac{T}{2}\tau^2 \sum_{n=2}^{N} \|\partial_\tau^2 u^n\|_h^2,
\end{aligned}$$

$$
\begin{aligned}
S_3 &= \tau \sum_{n=2}^{N} a_h(u^n, (2T - n\tau)\partial_\tau u^n) = \sum_{n=2}^{N} (2T - n\tau) a_h(u^n, u^n - u^{n-1}) \\
&= \sum_{n=2}^{N} (2T - n\tau) \frac{1}{2} (a_h(u^n, u^n) - a_h(u^{n-1}, u^{n-1}) + a_h(u^n - u^{n-1}, u^n - u^{n-1})) \\
&\geqslant \frac{T}{4} \|\nabla u^N\|_h^2 + \frac{\tau}{2} \sum_{2}^{N} \|\nabla u^n\|_h^2 - T\|\nabla u^1\|_h^2 + \frac{\tau^2}{2} \sum_{n=2}^{N} \|\tau \nabla \partial_\tau u^n\|_h^2, \\
S_4 &= \tau \sum_{n=2}^{N} a_h(u^n, \partial_\tau^2 u^n) \\
&= -\tau \sum_{n=2}^{N} (\nabla \partial_\tau u^{n-1}, \nabla \partial_\tau u^n)_h - (\partial_\tau \nabla u^1, \nabla u^1)_h + (\nabla u^N, \partial_\tau \nabla u^N)_h.
\end{aligned}
$$

Hence:

$$
|S_4| \leqslant \tau \sum_{n=1}^{N} \|\nabla \partial_\tau u^n\|_h^2 + \frac{1}{2} \|\partial_\tau \nabla u^1\|_h^2 + \frac{1}{2} \|\nabla u^1\|_h^2 + \frac{\delta}{2} \|\nabla u^N\|_h^2 + \frac{1}{2\delta} \|\partial_\tau \nabla u^N\|_h^2.
$$

One can see that by combining the above estimates the claim follows immediately for δ sufficiently small. □

Lemma 3.4. *Let $z \in V_h^{N-1}$. For $n = 0, 1, ..., N$ define $v^n = \tau \sum_{m=0}^{n} (1 + m\tau) z^m := \tilde{z}^n$. Then:*

$$
G(v, z) \geqslant |||z|||_D^2.
$$

Proof. Note that:

$$
\begin{aligned}
\tau \sum_{n=2}^{N} (\partial_\tau^2 v^n, z^n)_h &= \tau \sum_{n=2}^{N} (\partial_\tau((1 + n\tau) z^n), z^n)_h = \tau \sum_{n=2}^{N} (z^n, z^{n-1})_h + \sum_{n=2}^{N} (1 + n\tau)(z^n - z^{n-1}, z^n)_h \\
&= \tau \sum_{n=2}^{N} \|z^n\|_h^2 - \tau^2 \sum_{n=2}^{n} (z^n, \partial_\tau z^n)_h + \frac{1}{2} \sum_{n=2}^{N} (1 + n\tau)(\|z^n\|_h^2 - \|z^{n-1}\|_h^2 + \|z^n - z^{n-1}\|_h^2) \\
&\geqslant \frac{\tau}{2} \sum_{n=2}^{N} \|z^n\|_h^2 + \frac{T}{2} \|z^N\|_h^2 - \tau^2 \sum_{n=2}^{n} (z^n, \partial_\tau z^n)_h + \frac{\tau^2}{2} \sum_{n=2}^{N} \|\partial_\tau z^n\|_h^2 \\
&\geqslant \frac{\tau}{2} \sum_{n=2}^{N} \|z^n\|_h^2 + \frac{T}{2} \|z^N\|_h^2 - \tau^2 \sum_{n=2}^{n} \|z^n\|_h^2 - \frac{\tau^2}{4} \sum_{n=2}^{N} \|\partial_\tau z^n\|_h^2 + \frac{\tau^2}{2} \sum_{n=2}^{N} \|\partial_\tau z^n\|_h^2 \\
&\geqslant \frac{\tau}{4} \sum_{n=2}^{N} \|z^n\|_h^2 + \frac{T}{2} \|z^N\|_h^2 + \frac{\tau^2}{4} \sum_{n=2}^{N} \|\partial_\tau z^n\|_h^2.
\end{aligned}
$$

Similarly:

$$
\begin{aligned}
\tau \sum_{n=2}^{N} a_h(v^n, z^n) &= \tau \sum_{n=2}^{N} \frac{1}{(1 + n\tau)} a_h(v^n, \partial_\tau v^n) = \sum_{n=2}^{N} \frac{1}{(1 + n\tau)} a_h(v^n, v^n - v^{n-1}) \\
&= \frac{1}{2} \sum_{n=2}^{N} \frac{1}{1 + n\tau} (a_h(v^n, v^n) - a_h(v^{n-1}, v^{n-1}) + a_h(v^n - v^{n-1}, v^n - v^{n-1}))
\end{aligned}
$$

$$
\begin{aligned}
&\geqslant \frac{1}{2}\sum_{n=2}^{N}\frac{1}{1+n\tau}(a_h(v^n,v^n)-a_h(v^{n-1},v^{n-1}))\\
&\geqslant \frac{1}{2}\sum_{n=2}^{N}\frac{1}{1+n\tau}a_h(v^n,v^n)-\frac{1}{2}\sum_{n=2}^{N-1}\frac{1}{1+n\tau}a_h(v^n,v^n)\\
&+\frac{\tau}{2}\sum_{n=2}^{N}\frac{1}{(1+n\tau)(1+(n-1)\tau)}a_h(v^{n-1},v^{n-1})\\
&\geqslant \frac{1}{4(1+T)}a_h(v^N,v^N)+\frac{\tau}{2(1+T)^2}\sum_{n=2}^{N}a_h(v^n,v^n).
\end{aligned}
$$

Combining the above inequalities yields the claim. □

Proof of Proposition 3.2. Let α be a sufficiently small parameter independent of h and let us define

$$\hat{v}=u+\alpha v,\quad \text{and}\quad \hat{w}=-z+h^2\alpha w,$$

where w, v are chosen as in Lemma 3.3 and Lemma 3.4 respectively. We will show that the claim holds for this specific choice of $(\hat{v},\hat{w})\in V_h^{2N}$. We start by using the identity

$$A_1(u,-z)+A_2((u,z),u)=|||u|||_R^2,$$

together with Lemma 3.3 and Lemma 3.4 to write

$$
\begin{aligned}
A_1(u,\hat{w})+A_2((u,z),\hat{v}) &= |||u|||_R^2+\alpha A_1(u,h^2w)+\alpha A_2((u,z),v)\\
&\geqslant |||u|||_R^2+\alpha c\,h^2|||u|||_F^2-\alpha|||u|||_R^2+\alpha A_2((u,z),v)\\
&\geqslant \frac{1}{2}(|||u|||_R^2+\alpha c\,h^2|||u|||_F^2)+\alpha A_2((u,z),v).
\end{aligned}
\tag{3.7}
$$

Now recalling that $v^0=v^1=0$, we see that:

$$
\begin{aligned}
A_2((u,z),v) &= G(v,z)+\tau\sum_{n=1}^{N}(u^n,v^n)_{\omega_h}+(h\nabla\partial_\tau u^N,h\nabla\partial_\tau v^N)_h+\tau\sum_{n=2}^{N}(\tau\nabla\partial_\tau u^n,\tau\nabla\partial_\tau v^n)_h\\
&\geqslant |||z|||_D^2+\tau\sum_{n=1}^{N}(u^n,v^n)_{\omega_h}+(h\nabla\partial_\tau u^N,h\nabla\partial_\tau v^N)_h+\tau\sum_{n=2}^{N}(\tau\nabla\partial_\tau u^n,\tau\nabla\partial_\tau v^n)_h.
\end{aligned}
$$

We have:

$$\sum_{n=2}^{N}(\tau\nabla\partial_\tau u^n,\tau\nabla\partial_\tau v^n)_h\leqslant\sum_{n=2}^{N}(\frac{1}{2\delta}\|\tau\nabla\partial_\tau u^n\|_h^2+c\frac{\delta}{2}\|\nabla\tilde{z}^n\|_h^2).$$

Similarly using the Cauchy-Schwarz inequality we have:

$$|\sum_{n=1}^{N}(u^n,v^n)_{\omega_h}|\leqslant\sum_{n=1}^{N}(\frac{1}{2\delta}\|u^n\|_{\omega_h}^2+\frac{\delta}{2}\|v^n\|_h^2).$$

$$|(h\nabla\partial_\tau u^N,h\nabla\partial_\tau v^N)_h|\leqslant\frac{1}{2\delta}\|h\nabla\partial_\tau u^N\|_h^2+4T^2\frac{\delta}{2}\|z^N\|_h^2.$$

One can easily see that the first claimed estimate in the proposition holds for α small and δ sufficiently smaller than α. A similar argument can be used to prove the second claimed inequality in the statement of the proposition. Indeed we have:

$$\begin{aligned}|||(\hat{v},\hat{w})|||_C^2 \leqslant 2(\tau\sum_{n=1}^{N}(\|u^n\|_{\omega_h}^2+\alpha^2\|v^n\|_{\omega_h}^2)+\|h\nabla u^1\|_h^2+\|h\partial_\tau u^1\|_h^2\\ +\tau\sum_{n=2}^{N}(\|z^n\|_h^2+\alpha^2h^4\|w^n\|_h^2)+\tau\sum_{n=2}^{N}(\|\tau\nabla\partial_\tau u^n\|_h^2+\alpha^2\|\tau\nabla\partial_\tau v^n\|_h^2)\\ +\|h\nabla\partial_\tau u^1\|_h^2+\|h\nabla\partial_\tau u^N\|_h^2+\alpha^2\|h\nabla\partial_\tau v^N\|_h^2).\end{aligned}$$

with the following trivial estimates

$$\begin{aligned}\tau\sum_{n=1}^{N}\|v^n\|_{\omega_h}^2\leqslant c|||z|||_D^2, \qquad \tau\sum_{n=2}^{N}h^4\|w^n\|_h^2\leqslant c\,h^2|||u|||_F^2,\\ \tau\sum_{n=2}^{N}\|\tau\nabla\partial_\tau v^n\|_h^2\leqslant c|||z|||_D^2, \qquad \|h\nabla\partial_\tau v^N\|_h^2\leqslant c\|z^N\|_h^2.\end{aligned}$$

where $c>0$ only depends on $\mathcal{M}$. □

One can use Proposition 3.2 to show that the system of linear equations (3.4) has a unique solution. Indeed, denote by N_h the dimension of V_h. The equations (3.4) define a square linear system with $2N_h\times N$ unknowns. Setting $\tilde{q}^n=0$, the coercivity estimate in Proposition 3.2 implies that the kernel of this linear system is trivial and therefore there exists a unique solution for all choices of $\tilde{q}^n$. Henceforth, we will let (u_h,z_h) denote the unique solution to (3.4) subject to the measured noisy data $\tilde{q}^n$. Next section is concerned with proving the convergence of the discrete solution u_h to the continuum solution u of (1.1)-(1.2). The dual variable z_h is shown to converge to zero.

4. A Priori Error Estimates

An important feature of the error estimates below is that they include bounds of the perturbations from the discretization of the domain. To obtain such bounds we first prove some preliminary results.

Lemma 4.1. *For all $v_h\in V_h$ there holds*

$$\|v_h\|_{\partial\Omega}\leqslant ch\|\nabla v_h\|_{\Omega_h\setminus\Omega}.$$

Proof. First, note that using the trivial extension $v_h|_{\Omega\setminus\Omega_h}=0$ there holds $\|v_h\|_{\partial\Omega}=\|v_h\|_{\partial\Omega\cap\Omega_h}$. Now, for $x\in\partial\Omega\cap\Omega_h$, we write $v_h(x)=\int_{p(x)}^x\nabla v_h\cdot n\,ds$, with n the outward pointing unit normal of $\partial\Omega$, and where $p(x):=x+\zeta(x)n(x)$, with ζ is the (signed) distance from $\partial\Omega$ to $\partial\Omega_h$ in the n direction. By the assumption (3.2), $|\zeta|\leqslant ch^2$ and there holds

$$\int_{p(x)}^x\nabla v_h\cdot n\,ds\leqslant|\zeta(x)|^{\frac12}\left(\int_{p(x)}^x|\nabla v_h\cdot n|^2\,ds\right)^{\frac12}\leqslant c\,h\left(\int_{p(x)}^x|\nabla v_h\cdot n|^2\,ds\right)^{\frac12}. \qquad (4.1)$$

Using the above expression for $v_h|_{\partial\Omega}$ we have that

$$\|v_h\|_{\partial\Omega}^2\leqslant ch^2\int_{\partial\Omega}\int_{p(x)}^x|\nabla v_h\cdot n|^2\,ds\,dx\leqslant c\,h^2\|\nabla v_h\|_{\Omega_h\setminus\Omega}^2.$$

□

First we define an H^1-projection $\pi_h : H^1_0(\Omega) \to V_h(\Omega_h)$. Given $u \in H^1_0(\Omega)$, we let $\pi_h u \in V_h$ to be the unique solution of

$$a_h(\pi_h u, v_h) = a_h(Eu, v_h), \quad \forall v_h \in V_h \tag{4.2}$$

Lemma 4.2. *Let $u \in H^1_0(\Omega)$ and let $\pi_h u \in V_h$ be defined by (4.2). Then:*

$$\|u - \pi_h u\|_\Omega \leqslant c\,h \|u\|_{H^1(\Omega)} \tag{4.3}$$

and moreover

$$\|Eu - \pi_h u\|_{H^1(\Omega_h)} \leqslant c\,h \|u\|_{H^2(\Omega)} \text{ for } u \in H^1_0(\Omega) \cap H^2(\Omega), \tag{4.4}$$

for some constant $c > 0$ depending only on Ω.

Proof. First consider (4.4). Let $i_h u \in V_h$ denote the nodal interpolant of Eu. By the Poincaré's inequality there holds

$$\|i_h u - \pi_h u\|^2_{H^1(\Omega_h)} \leqslant c\, a_h(i_h u - \pi_h u, i_h u - \pi_h u).$$

Using the definition of $\pi_h u$, equation (4.2), we have

$$a_h(i_h u - \pi_h u, i_h u - \pi_h u) = a_h(i_h u - Eu, i_h u - \pi_h u) \leqslant \|i_h u - Eu\|_{H^1(\Omega_h)} \|i_h u - \pi_h u\|_{H^1(\Omega_h)}.$$

Combining the above estimate with

$$\|i_h u - Eu\|_{H^1(\Omega_h)} \leqslant ch\|u\|_{H^2(\Omega_h)} \leqslant ch\|u\|_{H^2(\Omega)}$$

Dividing with $\|i_h u - \pi_h u\|_{H^1(\Omega_h)}$ and using this estimate, it follows that

$$\|i_h u - \pi_h u\|_{H^1(\Omega_h)} \leqslant c\,h\|u\|_{H^2(\Omega)}.$$

The inequality (4.4) follows by the triangle inequality. For (4.3), first extend $\pi_h u$ to Ω by defining $\pi_h u = 0$ in $\Omega \setminus \Omega_h$. Then define the dual problem

$$\begin{aligned} -\Delta z &= u - \pi_h u \quad \text{in } \Omega \\ z &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

By the smoothness of Ω we know that $\|z\|_{H^2(\Omega)} \leqslant c\|u - \pi_h u\|_\Omega$. It follows that

$$\|u - \pi_h u\|^2_\Omega = (u - \pi_h u, -\Delta z)_\Omega = (\nabla(u - \pi_h u), \nabla z)_\Omega + (\pi_h u, \nabla z \cdot n)_{\partial\Omega \cap \Omega_h}.$$

For the first term in the right hand side we have

$$(\nabla(u - \pi_h u), \nabla z)_\Omega = (\nabla(u - \pi_h u), \nabla Ez)_{\Omega_h} - (\nabla(u - \pi_h u), \nabla Ez)_{\Omega_h \setminus \Omega} + (\nabla(u - \pi_h u), \nabla z)_{\Omega \setminus \Omega_h}.$$

Therefore, recalling that by trivial extension, $\pi_h u|_{\Omega \setminus \Omega_h} = 0$,

$$\begin{aligned} \|u - \pi_h u\|^2_\Omega &= (u - \pi_h u, -\Delta z)_\Omega = (\nabla(u - \pi_h u), \nabla(Ez - i_h Ez))_{\Omega_h} \\ &\quad - (\nabla(u - \pi_h u), \nabla Ez)_{\Omega_h \setminus \Omega} + (\nabla(u - \pi_h u), \nabla z)_{\Omega \setminus \Omega_h} + (\pi_h u, \nabla z \cdot n)_{\partial\Omega \cap \Omega_h} \\ &= (\nabla(u - \pi_h u), \nabla(z - i_h z))_{\Omega_h} - (\nabla(u - \pi_h u), \nabla Ez)_{\Omega_h \setminus \Omega} \\ &\quad + (\nabla u, \nabla z)_{\Omega \setminus \Omega_h} + (\pi_h u, \nabla z \cdot n)_{\partial\Omega \cap \Omega_h} = I + II + III + IV \end{aligned}$$

For the first term of the right hand side we have

$$I \leqslant c\|\nabla(u - \pi_h u)\|_{\Omega_h} h \|z\|_{H^2(\Omega)} \leqslant c\|\nabla(u - \pi_h u)\|_{\Omega_h} h \|u - \pi_h u\|_\Omega \leqslant c\|\nabla u\|_\Omega h \|u - \pi_h u\|_\Omega,$$

where we used that by (4.2) and the stability of the extension operator there holds

$$\|\nabla \pi_h u\|_{\Omega_h} \leqslant \|\nabla Eu\|_{\Omega_h} \leqslant \|\nabla u\|_\Omega. \tag{4.5}$$

To bound the second term we recall that by Lemma 3.1, a trace inequality and the stability of the extension and of z there holds

$$\|\nabla Ez\|_{\Omega_h\setminus\Omega} \leqslant ch\|u-u_h\|_{\Omega}.$$

Hence, using once again the stability (4.5)

$$II \leqslant \|\nabla(u-\pi_h u)\|_{\Omega_h\setminus\Omega}\|\nabla Ez\|_{\Omega_h\setminus\Omega} \leqslant c\|\nabla u\|_{\Omega} h\|u-\pi_h u\|_{\Omega}.$$

Similarly we obtain for the third term

$$III \leqslant \|\nabla u\|_{\Omega\setminus\Omega_h}\|\nabla z\|_{\Omega\setminus\Omega_h} \leqslant c\|\nabla u\|_{\Omega} h\|u-\pi_h u\|_{\Omega}.$$

To estimate the fourth term, we use the Cauchy-Schwarz inequality, Lemma 4.1 and the trace inequality, followed by the stability estimate on z,

$$IV = (\pi_h u, \nabla z\cdot n)_{\partial\Omega\cap\Omega_h} \leqslant \|\pi_h u\|_{\partial\Omega\cap\Omega_h}\|\nabla z\|_{\partial\Omega\cap\Omega_h} \leqslant ch\|\nabla\pi_h u\|_{\Omega_h\setminus\Omega}\|z\|_{H^2(\Omega)} \leqslant ch\|\nabla u\|_{\Omega}\|u-\pi_h u\|_{\Omega}.$$

Collecting the bounds for terms I-IV we conclude. □

Proposition 4.3. *Suppose Ω_h, Ω are as before and that $u \in H^3(\mathcal{M})$. Let (u_h, z_h) be the unique solution to the Euler-Lagrange equations* (3.4) *with $q = u|_{(0,T)\times\omega}$. Then:*

$$|||u_h-\pi_h u|||_R + h|||u_h-\pi_h u|||_F + |||z_h|||_D \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))}), \tag{4.6}$$

where $\pi_h u$ is the orthogonal projection defined by equation (4.2) *and $c > 0$ only depends on $\mathcal{M}$.*

Proof. First we recall that by the stability estimate of Proposition 3.2, there is $(v, w) \in V_h^{2N}$ satisfying:

$$|||u_h-\pi_h u|||_R^2 + h^2|||u_h-\pi_h u|||_F^2 + |||z_h|||_D^2 \leqslant c(A_1(u_h-\pi_h u, w) + A_2((u_h-\pi_h u, z_h), v)) \tag{4.7}$$

and

$$|||(v,w)|||_C \leqslant c(|||u_h-\pi_h u|||_R + h|||u_h-\pi_h u|||_F + |||z_h|||_D). \tag{4.8}$$

We will now bound the two terms of the right hand side of (4.7). Note that if $u^n = u(t_n)$ then:

$$(\partial_t^2 u^n, \psi)_{\Omega} + a(u^n, \psi) = 0 \qquad \forall\psi \in H_0^1(\Omega).$$

Also note that for all $w \in V_h^{N-1}$ we have

$$A_1(u, w) = \tau\sum_{n=2}^{N}\left((\varsigma_E^n, w)_{\Omega_h\setminus\Omega} + (\partial_t^2 u^n, w)_h + a_h(u^n, w)\right),$$

where, with some abuse of notation we identify u^n with Eu^n outside Ω and $\varsigma_E^n := -\Box Eu^n$ denotes the geometry residual term. Together with equation (3.4) and (4.2), this implies that:

$$A_1(u_h-\pi_h u, w) = \tau\sum_{n=2}^{N}(\partial_t^2 u^n - \partial_\tau^2 u^n, w^n)_h + \tau\sum_{n=2}^{N}((1-\pi_h)\partial_\tau^2 u^n, w^n)_h + \tau\sum_{n=2}^{N}(\varsigma_E^n, w^n)_{\Omega_h\setminus\Omega}.$$

First we observe that by Lemma 3.1,

$$|(\varsigma_E^n, w^n)_{\Omega_h\setminus\Omega}| \leqslant \|\varsigma_E^n\|_{\Omega_h\setminus\Omega}\|w^n\|_{\Omega_h\setminus\Omega} \leqslant c(\|\partial_t^2 u^n\|_{\Omega} + \|u^n\|_{H^2(\Omega)})h^2\|\nabla w^n\|_h.$$

Let:

$$I_1 = \tau\sum_{n=2}^{N}\|(1-\pi_h)\partial_\tau^2 u^n\|_h^2,$$

$$I_2 = \tau\sum_{n=2}^{N}\|\partial_t^2 u^n - \partial_\tau^2 u^n\|_h^2,$$

$$I_3 = \tau \sum_{n=2}^{N} h^2 (\|\partial_t^2 u^n\|_\Omega^2 + \|u^n\|_{H^2(\Omega)}^2) \leqslant ch^2 (\|\partial_t^2 u\|_{H^1(0,T;L^2(\Omega))}^2 + \|u\|_{H^1(0,T;H^2(\Omega))}^2).$$

Then clearly we have:

$$A_1(u_h - \pi_h u, w) \leqslant c(I_1 + I_2 + I_3)^{\frac{1}{2}} |||(0,w)|||_C.$$

Here we used that since $\tau = O(h)$ and by a (discrete) Poincaré inequality

$$\tau \sum_{n=2}^{N} h^2 \|\nabla w^n\|_h^2 \leqslant c(\|h\nabla w^1\|_h^2 + \tau \sum_{n=2}^{N} \|\tau \nabla \partial_\tau w^n\|_h^2) \leqslant c|||(0,w)|||_C^2.$$

It remains to bound I_1 and I_2. To this end, observe that:

$$\partial_\tau^2 u^n = \frac{1}{\tau^2} (\int_{t_{n-2}}^{t_n} (t - t_{n-2}) \partial_t^2 u \, dt - 2 \int_{t_{n-1}}^{t_n} (t - t_{n-1}) \partial_t^2 u \, dt).$$

Hence:

$$\begin{aligned} I_1 &\leqslant \frac{1}{\tau^2} \sum_{n=2}^{N} (2 \int_{t_{n-2}}^{t_n} (t - t_{n-2})^2 \|(\pi_h - 1)\partial_t^2 u\|_h^2 \, dt + 8 \int_{t_{n-1}}^{t_n} (t - t_{n-1})^2 \|(\pi_h - 1)\partial_t^2 u\|_h^2 \, dt) \\ &\leqslant \sum_{n=2}^{N} (2 \int_{t_{n-2}}^{t_n} \|(\pi_h - 1)\partial_t^2 u\|_h^2 \, dt + 8 \int_{t_{n-1}}^{t_n} \|(\pi_h - 1)\partial_t^2 u\|_h^2 \, dt) \\ &\leqslant ch^2 \int_0^T \|\nabla \partial_t^2 u\|_h^2 \, dt. \end{aligned}$$

Similarly we have:

$$\partial_\tau^2 u^n - \partial_t^2 u^n = \frac{1}{2\tau^2} (- \int_{t_{n-2}}^{t_n} (t - t_{n-2})^2 \partial_t^3 u \, dt + 2 \int_{t_{n-1}}^{t_n} (t - t_{n-1})^2 \partial_t^3 u \, dt).$$

Using this identity we obtain:

$$\begin{aligned} I_2 &\leqslant \frac{1}{2\tau^3} (\sum_{n=2}^{N} (\int_{t_{n-2}}^{t_n} (t - t_{n-2})^4 \, dt)(\int_{t_{n-2}}^{t_n} \|\partial_t^3 u\|_h^2 \, dt) + 4 \sum_{n=2}^{N} (\int_{t_{n-1}}^{t_n} (t - t_{n-1})^4 \, dt)(\int_{t_{n-1}}^{t_n} \|\partial_t^3 u\|_h^2 \, dt)) \\ &\leqslant c\tau^2 \int_0^T \|\partial_t^3 u\|_h^2 dt. \end{aligned}$$

Considering now the contribution from A_2, note that by definition

$$\begin{aligned} A_2((u_h - \pi_h u, z_h), v) &= \tau \sum_{n=1}^{N} (\delta q^n, v^n)_{\omega_h} + \tau \sum_{n=1}^{N} (u^n - \pi_h^n u^n, v^n)_{\omega_h} - (h\partial_\tau \pi_h u^1, h\partial_\tau v^1) - (h\nabla \pi_h u^1, h\nabla v^1) \\ &- \tau \sum_{n=2}^{N} (\tau \nabla \partial_\tau \pi_h u^n, \tau \nabla \partial_\tau v^n) - (h\partial_\tau \nabla \pi_h u^N, h\partial_\tau \nabla u^N) - (h\partial_\tau \nabla \pi_h u^1, h\partial_\tau \nabla u^1). \end{aligned}$$

Hence:

$$A_2((u_h - \pi_h u, z_h), v) \leqslant c(I_3 + I_4 + I_5 + I_6 + I_7 + I_8 + I_9)^{\frac{1}{2}} |||(v,0)|||_C,$$

where:

$$I_3 = \tau \sum_{n=1}^{N} \|\delta q^n\|_{\omega_h} \leqslant \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))},$$

$$I_4 = \tau \sum_{n=1}^{N} \|u^n - \pi_h^n u^n\|_{\omega_h}^2 \leqslant ch^2 \|\nabla u\|_{L^2(0,T;H^1(\Omega_h))}^2,$$
$$I_5 = \|h\partial_\tau \pi_h u^1\|_h^2 \leqslant ch^2 \|u\|_{H^2(0,T;L^2(\Omega_h))}^2,$$
$$I_6 = \|h\nabla \pi_h u^1\|_h^2 \leqslant ch^2 \|\nabla u\|_{H^1(0,T;L^2(\Omega_h))}^2,$$
$$I_7 = \tau \sum_{n=2}^{N} \|\tau \nabla \partial_\tau \pi_h u^n\|_h^2 \leqslant c\tau^2 \int_0^T \|\nabla \partial_t u\|_h^2 \, dt,$$
$$I_8 = \|h\nabla \partial_\tau \pi_h u^N\|_h^2 \leqslant ch^2 \|\nabla u\|_{H^2(0,T;L^2(\Omega_h))}^2,$$
$$I_9 = \|h\nabla \partial_\tau \pi_h u^1\|_h^2 \leqslant ch^2 \|\nabla u\|_{H^2(0,T;L^2(\Omega_h))}^2.$$

By the stability of the extension, all the norms over Ω_h can now be bounded by norms of the same quantities over Ω. The claim follows by collecting the above bounds. □

Corollary 4.4. *Under the same assumptions as in Proposition 4.3 there holds*

$$|||u_h - u|||_R + h|||u_h - u|||_F + |||z_h|||_D \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))}),$$

and

$$h\|\nabla u_h^0\| + |||u|||_R + h|||u_h|||_F \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))}).$$

Proof. The first inequality is immediate by adding and subtracting $\pi_h u$, applying the triangle inequality followed by Proposition 4.3 and Lemma 4.2 and similar Taylor expansion arguments as in Proposition 4.3. In the second inequality we note that

$$\|\nabla u_h^1\|_h^2 + \|\nabla u_h^0\|_h^2 \leqslant c(\|\nabla u_h^1\|_h^2 + \tau^2 \|\nabla \partial_\tau u_h^1\|_h^2).$$

We then add and subtract u in the right hand side of the last inequality and in $|||u_h|||_F$ and proceed as before, using the first inequality of the result to control the $u - u_h$ part and a Taylor expansion argument for the second. □

Before presenting the main theorem of this section we need an additional definition and lemma as follows. For each $w \in H_0^1(\mathcal{M})$, let us introduce the time averaged function $(\bar{w}^n)_{n=1}^N$ through

$$\bar{w}^n = \tau^{-1} \int_{t^{n-1}}^{t^n} w \, dt,$$

and denote by $\bar{w}$ the piecewise constant function $\bar{w}|_{[t^{n-1},t^n]} = \bar{w}^n$.

Lemma 4.5. *Suppose $u \in H^3(\mathcal{M})$ is the unique solution to the continuum problem* (1.1)-(1.2) *and let (u_h, z_h) denote the discrete solution to the Euler-Lagrange equations* (3.4). *The following estimate holds:*

$$\tau|(\nabla u_h^1, \nabla \bar{w}^1)_h| \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\Omega))})\|w\|_{H^1(\mathcal{M})},$$

where $w \in H_0^1(\mathcal{M})$ is arbitrary and $c > 0$ is independent of h and only depends on the geometry.

Proof. Note that

$$\tau(\nabla u_h^1, \nabla \bar{w}^1)_h = -\tau^2(\nabla \partial_\tau u_h^2, \nabla \bar{w}^1)_h + \tau(\nabla u_h^2, \nabla \bar{w}^1)_h. \tag{4.9}$$

For the first term on the right hand side of equation (4.9), observe that:

$$|\tau^2(\nabla \partial_\tau u_h^2, \nabla \bar{w}^1)_h| \leqslant \tau^2 \|\nabla \partial_\tau u_h^2\|_h \|\nabla \bar{w}^1\|_h \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\Omega))})\|w\|_{H^1(\mathcal{M})},$$

where we are using Corollary 4.4 to bound $\tau^{\frac{3}{2}}\|\nabla\partial_\tau u_h^2\|_h \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\Omega))})$ and the stability of $\bar{w}$ for the bound $\|\nabla\bar{w}^1\|_h \leqslant \tau^{-\frac{1}{2}}\|w\|_{H^1(\mathcal{M})}$. For the second term on the right hand side of equation (4.9) we have

$$\tau(\nabla u_h^2, \nabla\bar{w}^1)_h = \tau(\nabla u_h^2, \nabla\pi_h\bar{w}^1)_h = -\tau(\partial_\tau^2 u_h^2, (\pi_h - 1)\bar{w}^1)_h - \tau(\partial_\tau^2 u_h^2, \bar{w}^1)_h = I + II.$$

We note that Theorem 4.3 implies that $\tau^{\frac{3}{2}}\|\partial_\tau^2 u_h^2\|_h \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))})$. Finally,

$$|I| \leqslant c\tau\|\partial_\tau^2 u_h^2\|_h \tau^{\frac{1}{2}}\|w\|_{H^1(\mathcal{M})} \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))})\|w\|_{H^1(\mathcal{M})},$$

$$|II| \leqslant c\tau\|\partial_\tau^2 u_h^2\|_h\|\bar{w}^1\|_h \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))})\|w\|_{H^1(\mathcal{M})},$$

where in the last step we are using the standing assumption that $\tau \sim h$ and

$$\|\bar{w}^1\|_h \leqslant \tau^{-\frac{1}{2}}\left(\int_0^\tau \|w(t,\cdot)\|_h^2\,dt\right)^{\frac{1}{2}},$$

$$\int_0^\tau\int_\Omega\left|\int_0^t \partial_t w(t,\cdot)\,dt\right|^2 dx\,dt \leqslant \int_0^\tau\int_\Omega \tau\int_0^\tau |\partial_t w(t,\cdot)|^2\,dt\,dx\,dt = \tau^2\|\partial_t w\|^2_{L^2((0,t_1)\times\Omega)}.$$

□

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

Theorem 4.6. *Suppose $\mathcal{O} = (0,T)\times\omega$ satisfies the geometric control condition. Let $u \in H^3(\mathcal{M})$ denote the unique solution to the continuum problem (1.1)-(1.2). Let (u_h, z_h) denote the unique discrete solution to the Euler-Lagrange equations (3.4) subject to the noisy data $\tilde{q} = q + \delta q$ and $\delta q \in \mathcal{C}(0,T;L^2(\omega))$. Extend u_h to all of Ω by setting it equal to zero in $\Omega\setminus\Omega_h$. The following error estimate holds:*

$$\sup_{t\in[0,T]}\left(\|u(t,\cdot) - \tilde{u}_h(t,\cdot)\|_{L^2(\Omega)} + \|\partial_t u(t,\cdot) - \partial_t\tilde{u}_h(t,\cdot)\|_{H^{-1}(\Omega)}\right) \leqslant c\left(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))}\right),$$

where $c > 0$ is independent of h and only depends on the geometry and $\tilde{u}_h \in \mathcal{C}(\mathcal{M})$ denotes the linear interpolation

$$\tilde{u}_h = \frac{1}{\tau}((t - t_{n-1})u_h^n + (t_n - t)u_h^{n-1}) \quad \forall t \in [t_{n-1}, t_n].$$

Remark 2. *Theorem 4.6 can be used to make a number of observations. Firstly, it shows that the discretization method is stable in the presence of the noise δq, but stagnates when error reaches the level of the noise. This is the typical behaviour when solving a well-posed problem. Secondly, when the noise level is known and we have an apriori bound for u, Theorem 4.6 suggests that we should choose h to be of order $\|\delta q\|_{\mathcal{C}(0,T;L^2(\Omega))}/\|u\|_{H^3(\mathcal{M})}$. Finally, it should be noted that when $\|\delta q\|_{\mathcal{C}(0,T;L^2(\Omega))} = \mathcal{O}(h)$ then the method converges optimally corresponding to the approximation order of the lowest order finite difference method used for time discretization.*

Proof. Recall the standing assumption that $\tau = \mathcal{O}(h)$. Let $e = u - \tilde{u}_h$ and define the linear functional

$$\langle r, w\rangle = \int_0^T\int_\Omega(-\partial_t e\cdot\partial_t w + \nabla e\cdot\nabla w)\,dx\,dt \qquad \forall w \in H_0^1(\mathcal{M}). \tag{4.10}$$

Applying Theorem 2.2 we see that there holds

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

We will show that:

$$\|e\|_{L^2((0,T)\times\partial\Omega)} \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))}), \tag{4.11}$$

$$\|e\|_{L^2(\mathcal{O})} \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))}), \tag{4.12}$$

and

$$|\langle r, w\rangle| \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))})\|w\|_{H_0^1(\mathcal{M})}. \tag{4.13}$$

Estimate (4.11) and (4.12) will basically follow once we control the L^2 norm of the error function in $(0,T)\times\partial\Omega$ and $(0,T)\times\omega_h$, but (4.13) will be more delicate as there is no immediate relation that bounds $\|\Box e\|_{H^{-1}(\mathcal{M})}$ from above by $\|\Box e\|_{H^{-1}((0,T)\times\Omega_h)}$. Let us begin with (4.11). Since $u(t) \in H_0^1(\Omega)$

$$\|e\|^2_{L^2((0,T)\times\partial\Omega)} = \|\tilde{u}_h\|^2_{L^2((0,T)\times\partial\Omega)} \leqslant c\tau \sum_{n=0}^{N} \|u_h\|^2_{L^2((0,T)\times\partial\Omega)}.$$

Applying Lemma 4.1 followed by Corollary 4.4 we have

$$\begin{aligned}
\tau \sum_{n=0}^{N} \|u_h^n\|^2_{L^2((0,T)\times\partial\Omega)} &\leqslant c\tau \sum_{n=0}^{N} h^2\|\nabla u_h^n\|_h^2 \\
&\leqslant c(h^2\|\nabla u_h^0\|^2 + h^2\|\nabla u_h^1\|^2 + h^2|||u_h|||_F^2) \leqslant c(h^2\|u\|^2_{H^3(\mathcal{M})} + \|\delta q\|^2_{\mathcal{C}(0,T;L^2(\omega))}).
\end{aligned}$$

Now we consider the bounds (4.12) and (4.13). Define the time discrete projection operator π_0 as follows:

$$\pi_0 v := v(t^n) \quad \forall t \in (t_{n-1}, t_n], \quad n = 1, ..., N.$$

Then:

$$\|\pi_0 v - v\|_{L^2(0,T)} \leqslant \tau\|\partial_t v\|_{L^2(0,T)}.$$

We have:

$$\|e\|^2_{L^2((0,T)\times\omega_h)} \leqslant c(h^2+\tau^2)\|u\|^2_{H^1(\mathcal{M})} + \int_0^T \|\pi_0\pi_h u - \tilde{u}_h\|^2_{\omega_h}\,dt,$$

and

$$\begin{aligned}
\int_0^T \|\pi_0\pi_h u - \tilde{u}_h\|^2_{\omega_h}\,dt &\leqslant \int_0^T \|\pi_0\pi_h u - \pi_0\tilde{u}_h\|^2_{\omega_h}\,dt + \int_0^T \|\pi_0\tilde{u}_h - \tilde{u}_h\|^2_{\omega_h}\,dt \\
&= \tau\sum_{n=1}^{N} \|\pi_h u^n - u_h^n\|^2_{\omega_h} + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \|\pi_0\tilde{u}_h - \tilde{u}_h\|^2_{\omega_h}\,dt.
\end{aligned}$$

Here the first term is bounded by $|||u_h - \pi_h u|||_R$ and we use the identity

$$\tilde{u}_h(t) = u_h^n + (t-t_n)\partial_\tau u_h^n, \quad t \in (t_{n-1}, t_n]$$

to estimate the second one as follows:

$$\begin{aligned}
\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \|\pi_0\tilde{u}_h - \tilde{u}_h\|^2_{\omega_h}\,dt &= \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \|(t_n - t)\partial_\tau u_h^n\|^2_{\omega_h}\,dt \leqslant \tau\sum_{n=1}^{N}\|\tau\partial_\tau u_h^n\|_h^2 \\
&\leqslant \tau\sum_{n=1}^{N}\|\tau\partial_\tau\pi_h u^n\|_h^2 + \tau\sum_{n=1}^{N}\|\tau\partial_\tau(\pi_h u^n - u_h^n)\|_h^2.
\end{aligned}$$

The first term above is bounded by $\tau^2\|u\|^2_{H^3(\mathcal{M})}$ and as $\tau = \mathcal{O}(h)$, the second term is bounded by $h^2|||\pi_h u - u_h|||_F^2$. Hence, using Proposition 4.3 we deduce that

$$\|e\|^2_{L^2((0,T)\times\omega_h)} \leqslant c(h^2\|u\|^2_{H^3(\mathcal{M})} + \|\delta q\|^2_{\mathcal{C}(0,T;L^2(\omega))}).$$

Now, using Lemma 3.1 on the domains ω_h and ω, by choosing $v = u$ and noting that $v(t) \in H^1(\Omega)$ for a.e $t \in (0,T)$ we obtain that:

$$\begin{aligned}\|e\|^2_{L^2((0,T)\times(\omega\setminus\omega_h))} &\leqslant ch^2(\|e\|^2_{L^2((0,T)\times\partial\omega)} + h^2\|e\|^2_{L^2((0,T);H^1(\omega))}) \\ &\leqslant ch^2\|e\|^2_{L^2((0,T);H^1(\Omega))} \leqslant ch^2(\|u\|^2_{L^2((0,T);H^1(\Omega))} + \|\tilde{u}_h\|^2_{L^2((0,T);H^1(\Omega))}). \quad (4.14)\end{aligned}$$

By the Poincaré inequality and the definition of $\tilde{u}_h$

$$\|\tilde{u}_h\|^2_{L^2((0,T);H^1(\Omega))} \leqslant c\tau\sum_{n=1}^{N}(\|\nabla u_h^n\|_h^2 + \|\nabla u_h^{n-1}\|^2_{L^2(\Omega)}) \leqslant c\tau\sum_{n=0}^{N}\|\nabla u_h^n\|_h^2.$$

We now observe that, using the second inequality of Corollary 4.4

$$h^2\tau\sum_{n=0}^{N}\|\nabla u_h^n\|_h^2 = h^2\tau(\|\nabla u_h^0\|_h^2 + h^2\|\nabla u_h^1\|_h^2) + h^2|||u_h|||_F^2 \leqslant c(h^2\|u\|^2_{H^3(\mathcal{M})} + \|\delta q\|^2_{\mathcal{C}(0,T;L^2(\omega))}).$$

Finally, combining the preceding four inequalities yields the desired claim (4.12). We now prove (4.13). Using the definition of e and the equation $\Box u = 0$, we see that

$$\langle r, w\rangle = -\int_0^T\int_\Omega(-\partial_t\tilde{u}_h\cdot\partial_t w + \nabla\tilde{u}_h\cdot\nabla w)\,dx\,dt. \quad (4.15)$$

Recalling that u_h has been extended by zero and that by extension $w|_{\Omega_h\setminus\Omega} = 0$, we have

$$\langle r, w\rangle = -\int_0^T\int_{\Omega_h}(-\partial_t\tilde{u}_h\cdot\partial_t w + \nabla\tilde{u}_h\cdot\nabla w)\,dx\,dt \quad (4.16)$$

Using integration by parts and recalling that $w(0,\cdot) = w(T,\cdot) = 0$ we have

$$\int_0^T\int_{\Omega_h}(-\partial_t\tilde{u}_h\cdot\partial_t w)\,dx\,dt = \tau\sum_{n=1}^{N-1}\int_{\Omega_h}\partial_\tau^2 u_h^{n+1}w(\cdot,t^n)\,dx$$

Now, recalling the definition of the time averaged function $\bar{w}$, and considering the right hand side of (4.16) we see that

$$\begin{aligned}\langle r, w\rangle = &-\underbrace{\tau\sum_{n=1}^{N-1}(\partial_\tau^2 u_h^{n+1}, w(\cdot,t^n) - \bar{w}^{n+1})_h}_{I} - \underbrace{\tau\sum_{n=2}^{N}[(\partial_\tau^2 u_h^n, \bar{w}^n)_h + (\nabla u_h^n, \nabla\bar{w}^n)_h]}_{II} \\ &- \underbrace{\tau(\nabla u_h^1, \nabla\bar{w}^1)_h}_{III} - \underbrace{\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(t-t_n)(\nabla\partial_\tau u_h^n, \nabla w)_h\,dt}_{IV}.\end{aligned}$$

We now proceed to bound the six terms *I*-*IV* of the right hand side. First, using that

$$\|w(\cdot,t) - \bar{w}^{n+1}\|_{L^2((t_n,t_{n+1});L^2(\Omega))} \leqslant c\tau\|\partial_t w\|_{L^2((t_n,t_{n+1});L^2(\Omega))},$$

we have for the term I:

$$I = \sum_{n=1}^{N-1} (\partial_\tau^2 u_h^{n+1}, \int_{t_n}^{t_{n+1}} (\int_{t_n}^{t} (\partial_s w(\cdot, s)\, ds + w(\cdot, t) - \bar{w}^{n+1})\, dt))_h$$
$$\leqslant c\tau \left(\sum_{n=2}^{N} \tau \|\partial_\tau^2 u_h^n\|_h^2 \right)^{\frac{1}{2}} \|w\|_{H^1(\mathcal{M})} \leqslant c\tau |||u_h|||_F \|w\|_{H^1(\mathcal{M})}.$$

For the term II, we use (3.4) and (4.2) to obtain

$$II = \tau \sum_{n=2}^{N} [-(\partial_\tau^2 u_h^n, \bar{w}^n - \pi_h \bar{w}^n)_h]$$
$$\leqslant c\tau |||u_h|||_F \|\nabla w\|_{L^2(\mathcal{M})} \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))})\|w\|_{H^1(\mathcal{M})}.$$

The estimate for III follows immediately from Lemma 4.5. Finally for the term IV, we use Cauchy-Schwarz inequality and Corollary 4.4 to write

$$IV \leqslant c \left(\tau \sum_{n=1}^{N} \|\tau \nabla \partial_\tau u_h^n\|_h^2 \right)^{\frac{1}{2}} \|w\|_{H^1(\mathcal{M})} \leqslant c(h\|u\|_{H^3(\mathcal{M})} + \|\delta q\|_{\mathcal{C}(0,T;L^2(\omega))})\|w\|_{H^1(\mathcal{M})}.$$

□

5. Computational examples

The Euler-Lagrange equations (3.4) form a non-singular, symmetric system of $2NN_h$ linear equations, where N_h is the dimension of V_h. In this section we will describe a computational implementation solving (3.4). Our aim is not to give a thorough computational study, but only to demonstrate the convergence rate in two model cases.

Our computational tests indicate that introducing a constant weight factor in the data fitting term in the Lagrangian (3.3) leads to improved performance. In the computations below we have rescaled the term

$$\frac{\tau}{2} \sum_{n=1}^{N} \|u^n - q^n\|_{\omega_h}$$

by the factor 10. Such a constant factor does not change the theoretical conclusions above. The other terms in (3.3) could be rescaled as well, however, we do not study this type of parameter tuning in the present paper.

We will consider only the case that Ω is the unit interval and take $T = 1$. The mesh $\mathcal{T}_h$ is chosen to be uniform and we will use the same number of degrees of freedom in space and in time, that is, we choose $N_h = N$. The LU factorization is employed to solve (3.4). We consider

$$u(t, x) = \cos(2\pi t) \sin(2\pi x)$$

solving (1.1). Figure 1 summarizes the convergence of the method in two cases $\omega = \omega_j$, $j = 1, 2$, where

$$\omega_1 = (0, 0.2) \cup (0.8, 1), \quad \omega_2 = (0, 0.2).$$

The errors reported are given by

$$\max_{n=0,\dots,N} \|u(t_n) - u_h^n\|_{L^2(\Omega)}, \tag{5.1}$$

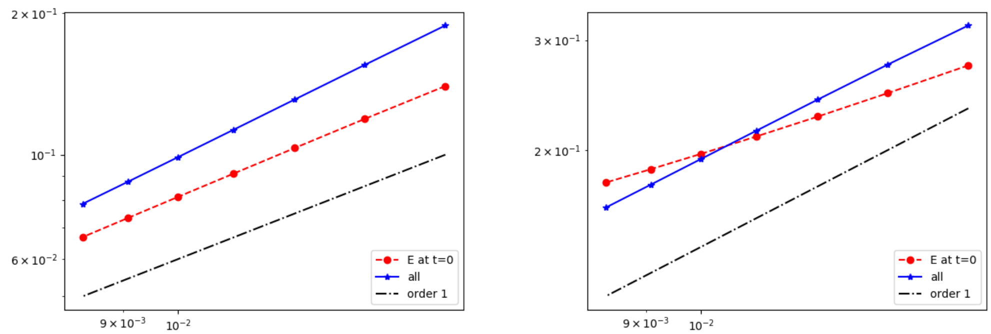

FIGURE 1. Convergence of the method compared with the predicted first order rate (log-log axes). The error on vertical axis is given by (5.1). Solid blue line gives the errors for the method analyzed in Theorem 4.6. The dashed red line gives the errors for the method where only the initial energy is used in regularization. The mesh sizes are $\tau = 1/60, 1/70, \ldots, 1/120$. *Left.* The case ω_1, convergence rates are 1.26 (blue) and 1.06 (red). *Right.* The case ω_2, rates are 0.97 (blue) and 0.62 (red).

where (u_h, z_h) is the solution of (3.4). We have also included computations using a method where the regularization $\mathcal{L}_1$ in (3.3) is replaced with the weaker regularization

$$\tilde{\mathcal{L}}_1(u) = \frac{1}{2}\left(\|h\nabla u^1\|_h^2 + \|h\partial_\tau u^1\|_h^2\right),$$

giving the initial energy on the discrete level.

Observe that the geometric control condition is satisfied in the case of ω_1 whereas it is not satisfied in the case of ω_2. However, as $\partial_t u$ vanishes identically at $t = 0$, the recovery of u is still stable, see e.g. [47]. In the case of ω_1, both the methods converge with a rate that is slightly better than that predicted by the theory. In the case of ω_2, the method with the weak regularization $\tilde{\mathcal{L}}_1$ fails to converge with a linear rate, whereas the method analyzed above converges with the linear rate. This is similar to our study of a data assimilation problem for the heat equation [13]. Also in that case it was necessary to use regularization terms involving the cross derivative $\nabla\partial_\tau$.

We conclude that the computational implementation gives convergence rates that are in line with the linear rate predicted by Theorem 4.6. In our numerical tests we also observed that simply discretizing the preliminary Lagrangian (3.1) leads to a method that fails to converge if no further regularization is present.

6. Further Remarks

We begin this section by making a remark about the choice of the time discretization employed here, since it does not conserve any notion of a discrete energy. The numerical dissipation introduced by the backward differentiation formula in time introduces a regularizing effect on the time derivatives of the approximate solution that we use to our advantage. This a priori control is used for the error estimate of Theorem 4.6.

It would be desirable to use an energy conservative discrete scheme, but the regularization of time derivatives of the forward solution must then be added to the stabilizing term. This shifts the energy balance of Proposition 3.2 and it is unclear if the stability result can be obtained without introducing regularization also for the dual variable z. Observe that regularization terms on the dual variable perturb energy conservation of the forward problem, so they must be avoided if the method is to be energy conservative. In contrast with the one-way coupling of the primal and dual variables u and z in (3.4), dual stabilization also introduces a two-way coupling between the systems for u and z. Therefore optimal estimates for energy conserving schemes is left as a topic for future research together with the extension to methods of higher accuracy.

6.1. **Polyhedral boundaries.** Recall that the proof of the key continuum estimate in Theorem 2.2 only works for smooth boundaries. Indeed the boundary smoothness assumption imposed in this paper is purely an artifact of the continuum estimate as the finite element method would be much simpler to apply for polyhedral boundaries and the discrete solution (u_h, z_h) would also exist and be unique. It is however possible to obtain a similar statement as in Theorem 4.6 for the case where Ω is a convex domain with a polyhedral boundary $\partial\Omega$. Here we present an admissibility condition that will be in some ways an alternative formulation of the geometric control condition or the $\Gamma-$condition for domains with polyhedral boundaries. Once this admissibility condition is satisfied for the observable domain $\mathcal{O}$, one can proceed to prove that Theorem 4.6 holds. To formulate this condition, we assume that there exists an auxiliary exhaustion of the polyhedral domain Ω by a sequence $\{\Omega_n\}_{n\in\mathbb{N}}$ such that the following properties are satisfied:

- $\forall n \in \mathbb{N} \quad \Omega_n \subset \Omega_{n+1}$,
- $\mu(\Omega \setminus \Omega_n) \leqslant \frac{1}{n}$, where μ denotes the Lebesgue measure,
- $\forall n \in \mathbb{N} \quad \partial\Omega_n \in C^\infty$.
- $(0,T) \times (\Omega_n \cap \omega)$ satisfies the geometric control condition in $(0,T) \times \Omega_n$ for all n.
- The constants C_n in the observability estimates corresponding to $(0,T) \times (\Omega_n \cap \omega)$ are uniformly bounded.

For polyhedral domains Ω, we call the sets $\mathcal{O} = (0,T)\times\omega$ with the above properties to be admissible. Note that the first three conditions will always be possible for any polyhedral domain Ω. It is merely the last two conditions which may not be true for an arbitrary domain $\mathcal{O}$. It is easy to check that if $\mathcal{O}$ satisfies the $\Gamma-$condition, then the admissibility condition above holds and therefore the implementation of the FEM in these cases works even for polyhedral boundaries Ω. It would be a very interesting question to study how this admissibility condition can more generally be written for Ω, ω, T without the use of the sequence Ω_n.

6.2. **Alternative choices of discrete regularization.** Let us now return to the explicit form of the Lagrangian functional $\mathcal{L}(u,z)$ in (3.3) and sketch some heuristic arguments regarding the discrete level regularization terms and the possibility of altering or removing them. The terms $\frac{\tau}{2}\sum_2^N \|u^n - q^n\|_\omega^2 + G(u,z)$ are absolutely necessary if we want the critical points of the Lagrangian functional to converge to the solution of (1.1). The two regularizer terms $\frac{1}{2}\|h\nabla u^1\|_h^2 + \frac{1}{2}\|h\partial_\tau u^1\|_h^2$ control the initial energy of the system and seem to be a natural term in the regularization. However, the additional terms $\frac{1}{2}\|h\nabla\partial_\tau u^1\|_h^2 + \frac{1}{2}\|h\nabla\partial_\tau u^N\|_h^2 + \frac{\tau}{2}\sum_2^N \|\tau\nabla\partial_\tau u^n\|_h^2$ control a particular choice of mixed derivatives of u. The advantage of using these additional regularizer terms is that it yields Lipschitz stability of the FEM with the optimal rate h and it avoids the use of any dual stabilizer terms for z in the Lagrangian. There is some freedom in the selection of these regularizers. For example, observe that the Lagrangian can be reduced by dropping the term $\|h\partial_\tau u^1\|$ without sacrificing stability since the contribution from this term is controlled by $\|h\partial_\tau\nabla u^1\|_h$ by the Poincaré inequality. This term however is kept due to its physical significance as the initial kinetic energy.

One may be able to remove the bulk regularizer term $\frac{\tau}{2}\sum_2^N \|\tau\nabla\partial_\tau u^n\|_h^2$ and replace it with only initial and final data regularizers such as $\frac{1}{2}\|h\partial_\tau^2 u^2\|_h^2 + \frac{1}{2}\|h\nabla\partial_\tau u^1\|_h^2 + \frac{1}{2}\|h\nabla\partial_\tau u^N\|_h^2$ and obtain the same error estimate. This will require an alternative energy estimate (see Lemma 3.3) and as such will require the smoothness class $u \in H^4(\mathcal{M})$.

One could also prove Theorem 4.6 using a less number of regularization terms but at the cost of a slower rate of decay. For example, using the Lagrangian functional

$$\hat{\mathcal{L}}(u,z) = \frac{\tau}{2}\sum_{n=2}^{N}\|u^n - q^n\|_\omega^2 + G(u,z) + \frac{1}{2}\|h\nabla u^1\|_h^2 + \frac{1}{2}\|h\partial_\tau u^1\|_h^2 + \frac{1}{2}\|h\partial_\tau\nabla u^1\|_h^2,$$

it is possible to prove Theorem 4.6 with a slower rate of decay of $\mathcal{O}(\sqrt{h})$ for the error function. It is also possible to obtain a linear convergence for the error function in weaker norms using the following 'minimal' Lagrangian functional:

$$\tilde{\mathcal{L}}(u,z) = \frac{\tau}{2}\sum_{n=2}^{N}\|u^n - q^n\|_\omega^2 + G(u,z) + \frac{1}{2}\|h\nabla u^1\|_h^2 + \frac{1}{2}\|h\partial_\tau u^1\|_h^2,$$

In this case, one can still prove Lemma 3.4 in the exact same manner. A similar estimate can be proved for u as well by choosing the test function w through:

$$w^n := (2T - n\tau)\partial_\tau u^n + \tau\sum_{m=0}^{n}(1+m\tau)u^m.$$

This will give positive control of $\|\tau\sum_{m=0}^n u^m\|_h^2$ together with $\tau\sum_{n=0}^N \|u^n\|_h^2$ and $\tau\sum_{n=1}^N \|\partial_\tau u^n\|_h^2$. Using these alternative estimates one can show that there exists a unique discrete solution (u_h, z_h) to

$$\partial_u\tilde{\mathcal{L}}(u_h,z_h) = 0,$$
$$\partial_z\tilde{\mathcal{L}}(u_h,z_h) = 0.$$

Now let $\mathcal{E}u^n = \tau\sum_{m=0}^n u^n$ and set

$$\tilde{u}_h := \frac{1}{\tau}((t - t_{n-1})\mathcal{E}{u_h}^n + (t_n - t)\mathcal{E}{u_h}^{n-1}) \quad \forall t \in [t_{n-1}, t_n].$$

One can then prove that if $e := \int_0^t u - \tilde{u}_h$, then the following weak stability estimate for the above FEM holds as well:

$$\|e\|_{L^2(\mathcal{M})} \leqslant ch\|u\|_{H^3(\mathcal{M})} \quad \text{and} \quad \|u_0 - u_h^0\|_{H^{-1}(\Omega)} \leqslant ch\|u\|_{H^3(\mathcal{M})}.$$

References

[1] S. Acosta, C. Montalto, Multiwave imaging in an enclosure with variable wave speed, Inverse Problems, 6, Vol. 31. (2015)

[2] D. Auroux and J. Blum, Back and forth nudging algorithm for data assimilation problems, C. R. Math. Acad. Sci. Paris, 12, 873–878, Vol. 340. (2005)

[3] L. Baudouin, S. Ervedoza, Convergence of an inverse problem for a 1-D discrete wave equation. SIAM J. Control Optim. 51 (2013), no. 1, 556–598.

[4] L. Baudouin , M. De Buhan, S. Ervedoza, Global Carleman estimates for waves and applications Commun. PDE 38 823-59, 2013.

[5] C. Bardos, G. Lebeau, J. Rauch, Un exemple dútilisation des notions de propagation pour le controle et la stabilisation de problémes hyperboliques. Rend. Sem. Mat. Univ. Politec. Torino, (Special Issue):11-31 (1989), 1988. Nonlinear hyperbolic equations in applied sciences.

[6] L. Bourgeois, D. Ponomarev, J. Dardé, An inverse obstacle problem for the wave equation in a finite time domain. Preprint hal-01818956.

[7] J. Bramble and J. T. King, A robust finite element method for nonhomogeneous Dirichlet problems in domains with curved boundaries. Math. Comp. 63 (1994), no. 207, 1–17.
[8] E. Burman, Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: Elliptic equations. SIAM J. Sci. Comput. 35, no. 6, (2013).
[9] E. Burman, Error estimates for stabilized finite element methods applied to ill-posed problems. C. R. Math. Acad. Sci. Paris 352, no. 7-8, 655–659, (2014).
[10] E. Burman, A. Feizmohammadi, L. Oksanen, A fully discrete numerical control method for the wave equation, 2019, arXiv preprint.
[11] E. Burman and L. Oksanen, Data assimilation for the heat equation using stabilized finite element methods, Numer. Math. (2018) 139: 505.
[12] E. Burman, M. Nechita, L. Oksanen, Unique continuation for the Helmholtz equation using stabilized finite element methods, J. Math. Pures Appl. (to appear), 2018..
[13] E. Burman, J. Ish-Horowicz and L. Oksanen, Fully discrete finite element data assimilation method for the heat equation, arXiv:1707.06908.
[14] O. Chervova, L. Oksanen, Time reversal method with stabilizing boundary conditions for photoacoustic tomography. Inverse Problems, 32(12):125004, 16, 2016.
[15] N. Cîndea, A. Münch, Inverse problems for linear hyperbolic equations using mixed formulations, Inverse Problems, 7, 075001, 38, Vol. 31 (2015).
[16] C. Clason and M. Klibanov, The quasi-reversibility method for thermoacoustic tomography in a heterogeneous medium. (English summary) SIAM J. Sci. Comput. 30 (2007/08), no. 1, 1–23.
[17] T. Duyckaerts, X. Zhang, and E. Zuazua, On the optimality of the observability inequalities for parabolic and hyperbolic systems with potentials, Ann. Inst. H. Poincaré Anal.
[18] S. Engel, P. Trautmann, B. Vexler, Optimal finite element error estimates for an optimal control problem governed by the wave equation with controls of bounded variation submitted, IMA Journal of Numerical Analysis, Preprint arXiv: 1907.11197, 2019.
[19] S. Ervedoza and E. Zuazua. Numerical approximation of exact controls for waves, Springer, New York, Springer Briefs in Mathematics (2013).
[20] S. Ervedoza, A. Marica, E. Zuazua, Numerical meshes ensuring uniform observability of one-dimensional waves: construction and analysis. (English summary) IMA J. Numer. Anal. 36 (2016), no. 2, 503–542.
[21] S. Ervedoza, E. Zuazua, The wave equation: control and numerics. Control of Partial Differential Equations (P. M. Cannarsa & J. M. Coron, eds). Lecture Notes in Mathematics, CIME Subseries. New York: Springer (2011).
[22] R. Glowinski, J.-L. Lions. Exact and approximate controllability for distributed parameter systems. Acta numerica, 159–333, 1995.
[23] G. Haine, K. Ramdani, Reconstructing initial data using observers: error analysis of the semi-discrete and fully discrete approximations. Numerische Mathematik, 120(2), 307–343. (2012)
[24] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich, Optimization with PDE Constraints, vol. 23 of Mathematical Modelling: Theory and Applications. Springer, 2009.
[25] L. Hörmander, The analysis of linear partial differential equations. Vol I-IV (Vol. 275) Berlin: Springer-Verlag (1985).
[26] O. Y. Imanuvilov, On Carleman estimates for hyperbolic equations, Asymptot. Anal., 32. (2002), pp. 185-220.
[27] J. Infante, E. Zuazua, Boundary observability for the space semi-discretizations of the 1-D wave equation. (English, French summary) M2AN Math. Model. Numer. Anal. 33 (1999), no. 2, 407–438.
[28] Y. Junjie, V. W. Lihong, J. Xia, Photoacoustic tomography: principles and advances, Electromagnetic waves (Cambridge, Mass), 147, 1-22 (2014).
[29] M. Klibanov, J. Malinsky, Newton-Kantorovich method for three-dimensional potential inverse scattering problem and stability of the hyperbolic Cauchy problem with time-dependent data. Inverse Problems 7 (1991), no. 4, 577–596.
[30] M. Klibanov, Rakesh, Numerical solution of a time-like Cauchy problem for the wave equation. (English summary) Math. Methods Appl. Sci. 15 (1992), no. 8, 559–570.
[31] A. Kröner, Numerical Methods for Control of Second Order Hyperbolic Equations, Ph.D. thesis, Fakultätfür Mathematik, Technische Universität München, 2011.
[32] P. Kuchment, L. Kunyansky, Mathematics of thermoacoustic tomography, European J. Appl. Math., 2,191-224,19 (2008).
[33] K. Kunisch, P. Trautmann, B. Vexler, Optimal control of the undamped linear wave equation with measure valued controls, SIAM J. Control Optim., 54 (2016), pp. 1212–1244.

[34] K. Kunisch, D. Wachsmuth, On time optimal control of the wave equation and its numerical realization as parametric optimization problem, SIAM J. Control Optim., 51 (2013), pp. 1232–1262.
[35] L. Kunyansky, L.V. Nguyen, A dissipative time reversal technique for photo-acoustic tomography in a cavity, To appear SIAM Journal on Imaging Sciences. Preprint (2015).
[36] J. Lagnese, G. Leugering, Domain decomposition methods in optimal control of partial differential equations, International Series of Numerical Mathematics, 148. Birkhäuser Verlag, Basel, 2004. xiv+443 pp.
[37] R. Lattès and J.-L. Lions, Méthode de quasi-réversibilité et applications. (French) Travaux et Recherches Mathématiques, No. 15 Dunod, Paris 1967 xii+368 pp.
[38] J. Le Rousseau, G. Lebeau, P. Terpolilli and E. Trélat, Geometric control condition for the wave equation with a time-dependent observation domain. Analysis PDE, MSP, 2017, 10 (4), pp.983-1015.
[39] I. Lasiecka, J.-L. Lions, R. Triggiani, Nonhomogeneous boundary value problems for second order hyperbolic operators. Journal de Mathematiques Pures et Appliques. Neuvime Srie, 65(2), 149192 (1986).
[40] V. W. Lihong, X. Minghua, Photoacoustic imaging in biomedicine, Review of Scientific Instruments, 4, Vol. 77 (2006).
[41] J.-L. Lions, Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. Tome 1, Masson (1988).
[42] L. Miller, Escape function conditions for the observation, control, and stabilization of the wave equation. SIAM Journal on Control and Optimization, 41(5), 1554–1566. (2002)
[43] L. Miller, Resolvent conditions for the control of unitary groups and their approximations. J. Spectr. Theory 2(1), 1–55 (2012).
[44] S. Montaner, A. Münch, Approximation of controls for linear wave equations: a first order mixed formulation. Preprint hal-01792949.
[45] K. Ramdani, M. Tucsnak and G. Weiss, Recovering and initial state of an infinite-dimensional system using observers, Automatica J. IFAC, 10, 1616,1625, Vol. 46 (2010).
[46] Y. Sasaki, Some basic formalisms in numerical variational analysis. Mon. Weather Rev. 98, (1970) pp. 875-883.
[47] P. Stefanov, G. Uhlmann, Thermoacoustic tomography with variable sound speed, Inverse Problems, 7, 075011, 16, 25 (2009).
[48] P. Stefanov and Y. Yang, Multiwave tomography in a closed domain: averaged sharp time reversal, Inverse Problems, 065007,23, Vol 31 (2015).
[49] E. M. Stein, Singular integrals and differentiability properties of functions. Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. 1970 xiv+290 pp.
[50] O. Scherzer, Handbook of mathematical methods in imaging, Springer Science & Business Media. (2010)
[51] O. Talagrand, On the mathematics of data assimilation. Tellus 33 (1981) 321–339.
[52] P. Trautmann, B. Vexler, A. Zlotnik, Finite Element Error Analysis For Measure-valued Optimal Control Problems Governed by a 1D Wave Equation with Variable Coefficients Mathematical Control and Related Fields, 8(2), pp. 411-449, 2018.
[53] L.V. Wang, Photoacoustic Imaging and Spectroscopy, CRC Press, 2009.
[54] E. Zuazua, Propagation, observation, and control of waves approximated by finite difference methods. SIAM Rev., 47, 197–243 (electronic) (2005).

Department of Mathematics, University College London, London, UK-WC1E 6BT, United Kingdom
Email address: e.burman@ucl.ac.uk

Department of Mathematics, University College London, London, UK-WC1E 6BT, United Kingdom
Email address: a.feizmohammadi@ucl.ac.uk

Department of Mathematics, University College London, London, UK-WC1E 6BT, United Kingdom
Email address: l.oksanen@ucl.ac.uk