The unique continuation problem for the wave equation discretized with a high-order space-time nonconforming method

Erik Burman, Guillaume Delay, Alexandre Ern

The unique continuation problem for the wave equation discretized with a high-order space-time nonconforming method

Erik Burman[*] Guillaume Delay[†] Alexandre Ern[‡]

July 19, 2024

Abstract

We are interested in solving the unique continuation problem for the wave equation, i.e., we want to reconstruct the solution of the wave equation given its (noised) value in a subset of the computational domain. Homogeneous Dirichlet boundary conditions are imposed, whereas the initial datum is unknown. We discretize this problem using a space-time discontinuous Galerkin method (including hybrid variables in space and in time) and look for the solution corresponding to the saddle-point of a discrete Lagrangian. We establish discrete inf-sup stability and bound the consistency error, leading to a priori estimates on the residual. Our main result proves the convergence of the discrete solution to the exact solution in a shifted energy norm involving weaker Sobolev norms than the standard energy norm for the wave equation. The proof combines the above a priori bound with a conditional stability estimate at the continuous level. Finally, we run numerical simulations to assess the performance of the method in practice. A static condensation procedure is used to eliminate the cell unknowns and reduce the size of the linear system.

Keywords: unique continuation, data assimilation, wave equation, hybridized discontinuous Galerkin, error estimate, static condensation

1 Introduction

In the present work, we are interested in solving numerically a data assimilation problem subject to the wave equation. Homogeneous Dirichlet boundary conditions are imposed, whereas initial conditions are unknown. In order to compensate for the lack of initial datum, we use the knowledge of the solution in a subdomain. We also investigate the influence of noise on this additional datum. Specifically, we consider a bounded open domain $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, with a smooth boundary $\partial\Omega$, a subset $\varpi \subset \Omega$, and a time interval $J := (0, T_f)$ with final time $T_f > 0$. The assumption that $\partial\Omega$ is smooth has its limits in practice, owing to the hybrid nature of the discretization method; see Remark 2. We denote by $Q := J \times \Omega$ the computational domain.

[*]Department of Mathematics, University College London, London, UK–WC1E 6BT, UK

[†]Laboratoire Jacques-Louis Lions, Sorbonne Université, CNRS, Inria, 4 place Jussieu, Paris, F-75005, France

[‡]CERMICS, École des Ponts, 8 av. Blaise Pascal, Marne-la-Vallée cedex 2, 77455, France and Inria Paris, 48, rue Barrault, CS 61534, 75647 Paris Cedex, France

Our goal is to approximate the function $u : Q \to \mathbb{R}$ that satisfies

$$L(u) := \partial_{tt}^2 u - \Delta u = f \quad \text{in } Q, \tag{1a}$$

$$u = 0 \quad \text{on } J \times \partial\Omega, \tag{1b}$$

$$u = g \quad \text{in } J \times \varpi, \tag{1c}$$

where $f \in L^2(Q)$ is a given source term and $g \in H^1(J; L^2(\varpi)) \cap L^2(J; H^1(\varpi))$ is the restriction to $J \times \varpi$ of a solution to the wave equation in Q. Notice that f and g have to be chosen in a compatible way such that there exists a solution to the problem (1). In the sequel, we consider perturbed data $g_\delta := g + \delta$ to account for some noise $\delta \in L^2(J \times \varpi)$ in the measurements.

This work is carried out under the so-called assumption of geometric control condition.

Assumption 1 (Geometric control condition)**.** *Every compressed generalized bicharacteristic intersects the set $J \times \varpi$ when projected to Q.*

Roughly this means that every ray has to cross the domain where measurements are available (accounting for reflexions on the computational domain boundary). We refer to [1] for a technical definition of a compressed generalized bicharacteristic, see also [2] where a time-dependent measurement domain is considered. This condition is necessary and sufficient for the exact controllability of the wave equation by a boundary control, see [3]. Notice that the assumption that $\partial\Omega$ is smooth is used to fit the setting of [2].

We consider the following shifted energy norm:

$$\|v\|_{\mathrm{sft}} := \big\{ \|v\|^2_{L^\infty(J;L^2(\Omega))} + \|\nabla v\|^2_{H^{-1}(J;L^2(\Omega))} + \|\partial_t v\|^2_{L^2(J;H^{-1}(\Omega))} \big\}^{\frac{1}{2}}. \tag{2}$$

The terminology is motivated by the use of a weaker Sobolev norm than the standard energy norm for the wave equation, which is $\{\|\nabla v\|^2_{L^2(J;L^2(\Omega))} + \|\partial_t v\|^2_{L^2(J;L^2(\Omega))}\}^{\frac{1}{2}}$. Under the above geometric control condition, we have the following stability estimate.

Lemma 1 (Stability estimate)**.** *Let the geometric control condition (Assumption 1) be satisfied. There exists $C_{\mathrm{stb}} > 0$ such that for all $v \in H^1(J; H^{-1}(\Omega)) \cap H^{-1}(J; H^1_0(\Omega)) \cap L^\infty(J; L^2(\Omega))$, we have*

$$\|v\|_{\mathrm{sft}} \le C_{\mathrm{stb}} \big(\|v\|_{L^2(J\times\varpi)} + \|L(v)\|_{H^{-1}(Q)} \big). \tag{3}$$

The proof of this lemma is given in Section A. It is one of the building blocks to prove the convergence of our approximation method.

Several works already deal with the numerical resolution of data assimilation subject to a transient PDE. For instance, data assimilation for the heat and wave equations in 1d has already been considered in [4] using a quasi-reversibility method. The case of the wave equation in a multi-dimensional configuration is addressed in [5], where a mixed formulation is proposed and proven to be well-posed at the continuous level only. A discretization of this mixed problem using a C^1 time-space finite element method is proposed as well. The reader is also referred to [6] for the reconstruction of the wave celerity based on Carleman estimates. Also notice the work [7] where the authors develop a general framework for least-squares methods to solve ill-posed problems that are conditionally stable. The norm in the least-squares method and the regularization are determined by the components of the conditional stability estimate. The method is applied to data assimilation problems subject to the heat and wave equations.

It is well-known that data assimilation problems and control problems are closely linked. The discretization of the control problem for the wave equation has been studied in [8]. A mixed formulation is proposed and C^1-conforming time-space finite elements are used for the discretization. Another work on discretizing the control problem of the wave equation is [9], where a low-order finite element method is considered for space discretization and a low-order finite difference method for time discretization. This work has been extended in [10], where space-time arbitrary high-order finite elements were considered. More examples of control and data assimilation methods for the wave equation can be found in [11].

We are interested in a class of methods that consists in first discretizing the ill-posed continuous problem and then adding stabilization terms at the discrete level. The discrete system to be solved corresponds to the equations characterizing the saddle point of a Lagrangian functional minimizing the solution discrepancy with respect to the data under the constraint of the wave equation. Some stabilization terms are also considered to enforce the well-posedness of the discrete problem. This framework for data assimilation has been developed in [12, 13] for elliptic problems. Several other application problems have been considered since, such as the Stokes [14] and Helmholtz [15, 16] equations. Unstationary problems have also been considered in this framework; see for instance [17, 18, 19] for the heat equation. The wave equation has been dealt with in [20], where a low-order finite element method for the space discretization and a finite difference for time have been considered. This work has been continued in [21], with an arbitrary high-order time-space finite element method. Notice though that the above framework is computationally challenging since all the time-space unknowns are globally coupled, thereby adding one more dimension compared to a discretization of a well-posed problem by a time-marching scheme. Finally, in the recent work [22], a dG approach in time is combined with a standard FEM approach in space. The wave equation is written in a mixed formulation. Two preconditioning approaches are also proposed and studied numerically.

The main asset of the above framework is that a priori error bounds in a shifted energy norm can be derived. The method to prove these bounds follows the following two steps:

1. stability, consistency, and a priori error bounds in a residual norm are first established;
2. then, using the stability estimate of the continuous problem (see Lemma 1), a priori error bounds in the shifted energy norm are obtained.

In the present work, we follow these two steps. Stability, consistency and a priori error bounds in residual norm are respectively proved in Lemmas 3, 5 and 6. The a priori error bound in the shifted energy norm is proved in Theorems 9 and 10.

The discretization method we propose here is based on a time-space Galerkin discretization. Discontinuous polynomials are considered in time and in space. A distinctive feature to increase computational effectivity is that interfacial elements in space and in time are added to reduce the size of the linear system through a static condensation procedure that eliminates the cell unknowns. This is a significant difference with respect to [19], where the authors have proposed and analyzed a nonconforming discretization for the data assimilation problem subject to the heat equation without the use of time-interfacial unknowns, thereby precluding the possibility of performing static condensation in time. Actually, the numerical tests in [19] emphasized that the method was computationally expensive. In the present work, we benefit from the static condensation procedure to reduce the size of the linear problem to solve. Moreover, the present method can use arbitrary high-order polynomials, and we expect the static condensation procedure to save more degrees of freedom for higher-order polynomials.

The rest of this work is organized as follows. In Section 2, we present the discretization method. In Section 3, we perform the error analysis. Finally, in Section 4, we present some numerical results that corroborate our theoretical error estimates.

2 Problem discretization

In this section, we describe the time and space discretization, and we present the numerical scheme studied in the present work.

2.1 Time-space meshes

We discretize the time-space cylinder $Q = J \times \Omega$ by a tensor-product time-space mesh.

Let $\mathcal{T}_\tau^t$ be a quasi-uniform mesh of J. We denote by τ the diameter of its largest cell. We also define the outward time-cell normal n_{T^t} associated with a generic time cell $T^t \in \mathcal{T}_\tau^t$ which

is conventionally taken to be an open subset of $\mathbb{R}$. The mesh faces in time (which coincide with the discrete time nodes) are collected in the set $\mathcal{F}_\tau^t$, and every generic time face $F^t \in \mathcal{F}_\tau^t$ is oriented by the fixed unit normal vector $\boldsymbol{n}^t \in \mathbb{R}^1$ positively orienting the time line.

Let $\mathcal{T}_h^x$ be a shape-regular mesh of Ω, where h refers to the maximal diameter of the mesh cells. We assume that $\mathcal{T}_h^x$ covers Ω exactly and that it is fitted to the measurement subset ϖ (so that, in particular, at least one mesh cell is contained in ϖ). In principle, the mesh cells in $\mathcal{T}_h^x$ can be polyhedra with planar faces in $\mathbb{R}^d$, and hanging nodes are also possible. However, the analysis below requires the mesh to be such that the underlying discontinuous polynomial approximation space has a global H^1-conforming subspace with optimal approximation properties. For simplicity, we will therefore restrict the discussion to meshes composed of simplices (one can also readily consider meshes composed of cuboids). The mesh cells are conventionally taken to be open subsets of $\mathbb{R}^d$, and $\boldsymbol{n}_{T^x} \in \mathbb{R}^d$ denotes the unit outward normal of the generic mesh cell $T^x \in \mathcal{T}_h^x$. The mesh faces in space are collected in the set $\mathcal{F}_h^x$, and every generic face $F^x \in \mathcal{F}_h^x$ is oriented by a fixed unit normal vector $\boldsymbol{n}_{F^x}^x \in \mathbb{R}^d$. To avoid technicalities, we assume henceforth that the mesh $\mathcal{T}_h^x$ is quasi-uniform. Therefore, we will use h to measure the diameter of any cell in $\mathcal{T}_h^x$ or any face in $\mathcal{F}_h^x$.

The time-space mesh of Q is defined as $\mathcal{Q}_{h,\tau} := \{q := T^t \times T^x \mid T^t \in \mathcal{T}_\tau^t, T^x \in \mathcal{T}_h^x\}$. For the sake of simplicity, we assume that there exists a constant $C > 0$ such that $C^{-1}h \le \tau \le Ch$. We could lift this assumption, but we would have to include the ratio h/τ in several estimates in the error analysis. With the above assumption, we will use h as the mesh parameter of the time-space mesh instead of $\max(h, \tau)$. Thus, we simply write $\mathcal{Q}_h$ instead of $\mathcal{Q}_{h,\tau}$ for the time-space mesh. We notice that, for every time-space cell $q := T^t \times T^x \in \mathcal{Q}_h$, its boundary ∂q can be decomposed as $\partial q = \partial^t q \cup \partial^x q$ with

$$\partial^t q := \bigcup_{\substack{F^t \in \mathcal{F}_\tau^t \\ F^t \subset \partial T^t}} \{F^t \times T^x\}, \qquad \partial^x q := \bigcup_{\substack{F^x \in \mathcal{F}_h^x \\ F^x \subset \partial T^x}} \{T^t \times F^x\}. \tag{4}$$

We loosely say that $\partial^x q$ collects the space faces of q and that $\partial^t q$ collects the time faces of q. We let $\boldsymbol{n}_q$ be the unit outward normal to q. We also set $\boldsymbol{n}_{\partial^t q} := \boldsymbol{n}_q|_{\partial^t q} := \boldsymbol{n}^t$ and $\boldsymbol{n}_{\partial^x q} := \boldsymbol{n}_q|_{\partial^x q}$ with $\boldsymbol{n}_{\partial^x q}|_{T^t \times F^x} := \boldsymbol{n}_{F^x}^x$ for all $T^t \times F^x \in \partial^x q$.

The collection of all the faces of $\mathcal{Q}_h$ is denoted by $\mathcal{G}_h$, which is partitioned as $\mathcal{G}_h = \mathcal{G}_h^x \cup \mathcal{G}_h^t$, where $\mathcal{G}_h^x$ collects all the space faces and $\mathcal{G}_h^t$ collects all the time faces of the time-space mesh. Thus, $\partial^x q \subset \mathcal{G}_h^x$ and $\partial^t q \subset \mathcal{G}_h^t$ for all $q \in \mathcal{Q}_h$. Moreover, the set $\mathcal{G}_h^x$ is further partitioned as $\mathcal{G}_h^x = \mathcal{G}_h^{x,\mathrm{int}} \cup \mathcal{G}_h^{x,\mathrm{ext}}$, where $\mathcal{G}_h^{x,\mathrm{int}}$ collects all the space faces that are subsets of $J \times \Omega$ (called interior space faces), whereas $\mathcal{G}_h^{x,\mathrm{ext}}$ collects all the space faces that are subsets of $J \times \partial\Omega$ (called exterior space faces). Similarly, we consider the partition $\mathcal{G}_h^t = \mathcal{G}_h^{t,\mathrm{int}} \cup \mathcal{G}_h^{t,\mathrm{ext}}$ and observe that the time faces in $\mathcal{G}_h^{t,\mathrm{ext}}$ correspond either to the initial or to the final time.

Finally, we set $\mathcal{G}_h^{\mathrm{int}} := \mathcal{G}_h^{t,\mathrm{int}} \cup \mathcal{G}_h^{x,\mathrm{int}}$. For all $g \in \mathcal{G}_h^{\mathrm{int}}$, we define $\boldsymbol{n}_g := \boldsymbol{n}^t$ if $g \in \mathcal{G}_h^{t,\mathrm{int}}$ and $\boldsymbol{n}_g := \boldsymbol{n}_{F^x}^x$ if $g := T^t \times F^x \in \mathcal{G}_h^{x,\mathrm{int}}$, and $[\![v]\!]_g$ denotes the jump of the piecewise smooth, time-space function v across g in the direction of $\boldsymbol{n}_g$. The notation is summarized in Figure 1.

2.2 Time-space polynomial spaces

Let $k \ge 1$ be the polynomial degree of the hybridized dG method in space and let $\ell \ge 1$ be the polynomial degree of the hybridized dG method in time. We denote by $\mathbb{P}^k(S)$ the set of polynomials of total degree at most k on the subset $S \subseteq \Omega$. Moreover, for a linear space U composed of functions defined on S, we denote by $\mathbb{P}^\ell(I; U)$ the set of U-valued polynomials of degree at most ℓ on $I \subseteq \bar{J} = [0, T_f]$.

The discrete unknowns are tensor-product time-space polynomials attached to every time-space cell, $q \in \mathcal{Q}_h$, and to every space and time face, $g^x \in \mathcal{G}_h^x$ and $g^t \in \mathcal{G}_h^t$, respectively. For all $q := T^t \times T^x \in \mathcal{Q}_h$, we set

$$U_q := \mathbb{P}^\ell(T^t; \mathbb{P}^k(T^x)). \tag{5}$$

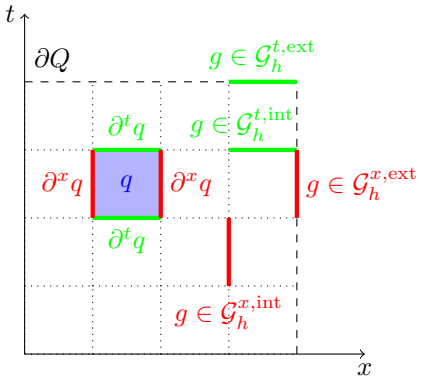

Figure 1: Mesh notation

For all $g^x := T^t \times F^x \in \mathcal{G}_h^x$ and all $g^t := F^t \times T^x \in \mathcal{G}_h^t$, we set

$$U_{g^x} := \mathbb{P}^\ell(T^t; \mathbb{P}^k(F^x)), \qquad U_{g^t} := \mathbb{P}^k(T^x). \tag{6}$$

(There is no time variation in U_{g^t} since F^t is just a discrete time node). We then set

$$U_{\partial q} := U_{\partial^x q} \times U_{\partial^t q}, \qquad U_{\partial^x q} := \prod_{g^x \in \partial^x q} U_{g^x}, \qquad U_{\partial^t q} := \prod_{g^t \in \partial^t q} U_{g^t}. \tag{7}$$

The discrete problem involves cell polynomials $u_\mathcal{Q} := (u_q)_{q\in\mathcal{Q}_h}$, and face polynomials $u_\mathcal{G} := (u_g)_{g\in\mathcal{G}_h} := \big((u_{g^t})_{g^t\in\mathcal{G}_h^t}, (u_{g^x})_{g^x\in\mathcal{G}_h^x}\big)$. Notice that we partitioned the face polynomials into those attached to the time faces and those attached to the space faces. We introduce the time-space discrete space

$$\hat{U}_\mathcal{Q} := U_\mathcal{Q} \times U_\mathcal{G}, \qquad U_\mathcal{G} := U_{\mathcal{G}^t} \times U_{\mathcal{G}^x}, \tag{8}$$

with

$$U_\mathcal{Q} := \{v_\mathcal{Q} \in L^2(Q) \mid v_q := v_\mathcal{Q}|_q \in U_q,\ \forall q \in \mathcal{Q}_h\}, \tag{9a}$$
$$U_{\mathcal{G}^t} := \{v_{\mathcal{G}^t} \in L^2(\mathcal{G}_h^t) \mid v_{g^t} := v_{\mathcal{G}^t}|_{g^t} \in U_{g^t},\ \forall g^t \in \mathcal{G}_h^t\}, \tag{9b}$$
$$U_{\mathcal{G}^x} := \{v_{\mathcal{G}^x} \in L^2(\mathcal{G}_h^x) \mid v_{g^x} := v_{\mathcal{G}^x}|_{g^x} \in U_{g^x},\ \forall g^x \in \mathcal{G}_h^{x,\mathrm{int}}, v_{g^x} = 0,\ \forall g^x \in \mathcal{G}_h^{x,\mathrm{ext}}\}, \tag{9c}$$

Notice that functions in $U_{\mathcal{G}^x}$ vanish identically on $J \times \partial\Omega$. For a generic function $\hat{v}_\mathcal{Q} \in \hat{U}_\mathcal{Q}$, we write $\hat{v}_\mathcal{Q} := (v_\mathcal{Q}, v_\mathcal{G})$ with $v_\mathcal{Q} \in U_\mathcal{Q}$ and $v_\mathcal{G} \in U_\mathcal{G}$. Finally, we denote by

$$\hat{U}_{\mathcal{Q},0} := U_\mathcal{Q} \times U_{\mathcal{G},0}, \qquad U_{\mathcal{G},0} := U_{\mathcal{G}^x} \times U_{\mathcal{G}^t,0}, \tag{10a}$$
$$U_{\mathcal{G}^t,0} := \{v_{\mathcal{G}^t} \in U_{\mathcal{G}^t} \mid v_{g^t} = 0,\ \forall g^t \in \mathcal{G}_h^{t,\mathrm{ext}}\}, \tag{10b}$$

the subspace of $\hat{U}_\mathcal{Q}$ composed of pairs whose time-face polynomials are null at the initial and at the final times. We will use the space $\hat{U}_\mathcal{Q}$ for the discrete primal unknowns and the space $\hat{U}_{\mathcal{Q},0}$ for the discrete dual unknowns.

For a time-space cell polynomial $v_\mathcal{Q} \in U_\mathcal{Q}$, we define the broken gradient and Laplacian operators, $\nabla_\mathcal{Q}$ and $\Delta_\mathcal{Q}$, such that $(\nabla_\mathcal{Q} v_\mathcal{Q})|_q := \nabla(v_\mathcal{Q}|_q)$ and $(\Delta_\mathcal{Q} v_\mathcal{Q})|_q := \Delta(v_\mathcal{Q}|_q)$ for all $q \in \mathcal{Q}_h$. A similar notation is adopted for the first-order and second-order time-derivatives, $\partial_{t,\mathcal{Q}}$ and $\partial^2_{tt,\mathcal{Q}}$. Finally, we set

$$L_\mathcal{Q}(v_\mathcal{Q}) := \partial^2_{tt,\mathcal{Q}} v_\mathcal{Q} - \Delta_\mathcal{Q} v_\mathcal{Q}, \qquad \forall v_\mathcal{Q} \in U_\mathcal{Q}. \tag{11}$$

Remark 1 (Cell polynomials)**.** *The time-space polynomial space defined in* (5) *has a time-space tensor-product form. Indeed, any polynomial $v_q \in U_q$ can be written as*

$$v_q = \sum_{\substack{\beta\in\{0:\ell\}\\ \alpha:=(\alpha_1,\dots,\alpha_d)\in\mathbb{N}^d\\ 0\le\alpha_1+\cdots+\alpha_d\le k}} \gamma_{\alpha\beta}\, t^\beta \boldsymbol{x}^\alpha,$$

with $\boldsymbol{x}^\alpha := x_1^{\alpha_1}\dots x_d^{\alpha_d}$.

2.3 Bilinear forms

We can now introduce the bilinear forms needed to formulate the discrete problem. Let $\hat{v}_\mathcal{Q}$, $\hat{w}_\mathcal{Q}$ be generic functions in $\hat{U}_\mathcal{Q}$ (primal variables) and let $\hat{\zeta}_\mathcal{Q}$, $\hat{\eta}_\mathcal{Q}$ be generic functions in $\hat{U}_{\mathcal{Q},0}$ (dual variables). The bilinear form associated with the discretization of the wave equation in a generic time-space cell $q \in \mathcal{Q}_h$ is

$$\begin{aligned} a_q(\hat{v}_q,\hat{\eta}_q) := {}& (\nabla v_q,\nabla\eta_q)_q - (\nabla v_q\cdot\boldsymbol{n}_{\partial^x q},\eta_q-\eta_{\partial^x q})_{\partial^x q} - (v_q - v_{\partial^x q},\nabla\eta_q\cdot\boldsymbol{n}_{\partial^x q})_{\partial^x q}\\ & - (\partial_t v_q,\partial_t\eta_q)_q + (\partial_t v_q\boldsymbol{n}_{\partial^t q},\eta_q-\eta_{\partial^t q})_{\partial^t q} + (v_q - v_{\partial^t q},\partial_t\eta_q\boldsymbol{n}_{\partial^t q})_{\partial^t q}. \end{aligned} \tag{12}$$

In order to allow for a more compact notation, we define

$$B := \begin{pmatrix} -1 & 0_{1,d}\\ 0_{d,1} & \mathbb{I}_{d,d}\end{pmatrix} \in \mathbb{R}^{d+1,d+1}, \tag{13}$$

together with the space-time gradient operator $\widetilde{\nabla} v := (\partial_t v,\nabla v)^{\mathrm{T}}$ and the space-time unit outward normal vector $\widetilde{\boldsymbol{n}}_q \in \mathbb{R}^{d+1}$ for all $q \in \mathcal{Q}_h$ such that $\widetilde{\boldsymbol{n}}_q|_{\partial^t q} := (\boldsymbol{n}_{\partial^t q},\boldsymbol{0})^{\mathrm{T}} \in \mathbb{R}^{d+1}$ and $\widetilde{\boldsymbol{n}}_q|_{\partial^x q} := (\boldsymbol{0},\boldsymbol{n}_{\partial^x q})^{\mathrm{T}} \in \mathbb{R}^{d+1}$. We also use a similar notation $\widetilde{\boldsymbol{n}}_g \in \mathbb{R}^{d+1}$ for all $g \in \mathcal{G}_h^{\mathrm{int}}$ and we consider the broken version of the above space-time gradient operator, $\widetilde{\nabla}_\mathcal{Q}$. We can then rewrite the local bilinear form a_q as follows:

$$a_q(\hat{v}_q,\hat{\eta}_q) = (B\widetilde{\nabla} v_q,\widetilde{\nabla}\eta_q)_q - (B\widetilde{\nabla} v_q\cdot\widetilde{\boldsymbol{n}}_q,\eta_q-\eta_{\partial q})_{\partial q} - (v_q - v_{\partial q},B\widetilde{\nabla}\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}.$$

The global bilinear form is assembled by summing over the time-space cells:

$$a_\mathcal{Q}(\hat{v}_\mathcal{Q},\hat{\eta}_\mathcal{Q}) := \sum_{q\in\mathcal{Q}_h} a_q(\hat{v}_q,\hat{\eta}_q). \tag{14}$$

Moreover, the local stabilization bilinear forms read as follows:

$$s_q(\hat{v}_q,\hat{w}_q) := h^{-1}(v_q - v_{\partial q},w_q - w_{\partial q})_{\partial q}, \tag{15a}$$

$$\sigma_q(\hat{\zeta}_q,\hat{\eta}_q) := (\widetilde{\nabla}\zeta_q,\widetilde{\nabla}\eta_q)_q + s_q(\hat{\zeta}_q,\hat{\eta}_q), \tag{15b}$$

leading to the following global stabilization bilinear forms:

$$s_\mathcal{Q}(\hat{v}_\mathcal{Q},\hat{w}_\mathcal{Q}) := \sum_{q\in\mathcal{Q}_h} s_q(\hat{v}_q,\hat{w}_q), \tag{16a}$$

$$\sigma_\mathcal{Q}(\hat{\zeta}_\mathcal{Q},\hat{\eta}_\mathcal{Q}) := \sum_{q\in\mathcal{Q}_h} \sigma_q(\hat{\zeta}_q,\hat{\eta}_q) = (\widetilde{\nabla}_\mathcal{Q}\zeta_\mathcal{Q},\widetilde{\nabla}_\mathcal{Q}\eta_\mathcal{Q})_\mathcal{Q} + s_\mathcal{Q}(\hat{\zeta}_\mathcal{Q},\hat{\eta}_\mathcal{Q}). \tag{16b}$$

Notice that $s_\mathcal{Q}$ consists of least-squares penalties on the difference between the trace of the cell polynomials and the polynomials on the space and time faces of every time-space cell $q \in \mathcal{Q}_h$. For the dual variables, we enforce the same penalty together with a least-squares penalty on the broken time-space derivatives.

2.4 Lagrangian and discrete problem

We want to find the saddle-point of the Lagrangian defined for all $(\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$ by

$$\mathcal{L}_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) := \frac{1}{2}\|v_{\mathcal{Q}} - g_\delta\|^2_{J\times\varpi} + \frac{1}{2}s_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{v}_{\mathcal{Q}}) - \frac{1}{2}\sigma_{\mathcal{Q}}(\hat{\zeta}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) + a_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) - (f, \zeta_{\mathcal{Q}})_Q, \quad (17)$$

where $g_\delta := g + \delta$ denotes the available perturbed measurement of g. Notice that there is no initial nor final condition on the primal variable, whereas the dual variable is null at initial and final time.

The discrete problem is derived by seeking a critical point of the Lagrangian and reads as follows: Find $(\hat{u}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$ such that

$$(u_{\mathcal{Q}}, w_{\mathcal{Q}})_{J\times\varpi} + s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{w}_{\mathcal{Q}}) + a_{\mathcal{Q}}(\hat{w}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}) = (g_\delta, w_{\mathcal{Q}})_{J\times\varpi}, \quad (18a)$$

$$a_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) - \sigma_{\mathcal{Q}}(\hat{\xi}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) = (f, \eta_{\mathcal{Q}})_Q, \quad (18b)$$

where the first equation holds for all $\hat{w}_{\mathcal{Q}} \in \hat{U}_{\mathcal{Q}}$ and the second for all $\hat{\eta}_{\mathcal{Q}} \in \hat{U}_{\mathcal{Q},0}$. For all $(\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}})$ and $(\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}})$ in $\hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$, we define the bilinear form

$$\begin{aligned} A_{\mathcal{Q}}((\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}), (\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}})) := (v_{\mathcal{Q}}, w_{\mathcal{Q}})_{J\times\varpi} + s_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{w}_{\mathcal{Q}}) + a_{\mathcal{Q}}(\hat{w}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) \\ + a_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) - \sigma_{\mathcal{Q}}(\hat{\zeta}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}). \end{aligned} \quad (19)$$

The discrete problem (18) can be rewritten as follows: Find $(\hat{u}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$ such that, for all $(\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$,

$$A_{\mathcal{Q}}((\hat{u}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}})) = (g_\delta, w_{\mathcal{Q}})_{J\times\varpi} + (f, \eta_{\mathcal{Q}})_Q. \quad (20)$$

We will see in Corollary 1 that the discrete problem (18) (or equivalently, (20)) is well-posed.

The discrete problem (18) can be solved efficiently by eliminating locally all the cell degrees of freedom using a static condensation procedure. The global problem to be solved then involves only the face degrees of freedom. The stencil of the global problem couples unknowns attached to neighboring time and space faces (in the sense that two neighbors share a common time-space cell). To eliminate the cell degrees of freedom, one needs to solve the following linear system: Find $(u_{\mathcal{Q}}, \xi_{\mathcal{Q}}) \in U_{\mathcal{Q}} \times U_{\mathcal{Q}}$ such that, for all $(w_{\mathcal{Q}}, \eta_{\mathcal{Q}}) \in U_{\mathcal{Q}} \times U_{\mathcal{Q}}$,

$$\begin{aligned} A_{\mathcal{Q}}(((u_{\mathcal{Q}}, 0), (\xi_{\mathcal{Q}}, 0)), ((w_{\mathcal{Q}}, 0), (\eta_{\mathcal{Q}}, 0))) = (g_\delta, w_{\mathcal{Q}})_{J\times\varpi} + (f, \eta_{\mathcal{Q}})_Q \\ - A_{\mathcal{Q}}(((0, u_{\mathcal{G}}), (0, \xi_{\mathcal{G}})), ((w_{\mathcal{Q}}, 0), (\eta_{\mathcal{Q}}, 0))), \end{aligned} \quad (21)$$

where $u_{\mathcal{G}} \in U_{\mathcal{G}}$ and $\xi_{\mathcal{G}} \in U_{\mathcal{G},0}$ are given. We will see in Corollary 1 that the discrete problem (21) is well-posed. At this stage, we only record that the linear system (21) is easy to solve since it amounts to a block linear system on the cell unknowns. Indeed, (21) is equivalent to solving, for all $q := T^t \times T^x \in \mathcal{Q}_h$, all $u_{\partial q} \in U_{\partial q}$, and all $\xi_{\partial q} \in U_{\partial q}$, the following local system: Find $(u_q, \xi_q) \in U_q \times U_q$ such that, for all $(w_q, \eta_q) \in U_q \times U_q$,

$$(u_q, w_q)_{T^t\times(\varpi\cap T^x)} + s_q((u_q, 0), (w_q, 0)) + a_q((w_q, 0), (\xi_q, 0)) = \Phi^1_q(w_q), \quad (22a)$$

$$a_q((u_q, 0), (\eta_q, 0)) - \sigma_q((\xi_q, 0), (\eta_q, 0)) = \Phi^2_q(\eta_q), \quad (22b)$$

with $\Phi^1_q(w_q) := (g_\delta, w_q)_{T^t\times(\varpi\cap T^x)} - s_q((0, u_{\partial q}), (w_q, 0)) - a_q((w_q, 0), (0, \xi_{\partial q}))$ and $\Phi^2_q(\eta_q) := (f, \eta_q)_q - a_q((0, u_{\partial q}), (\eta_q, 0)) + \sigma_q((0, \xi_{\partial q}), (\eta_q, 0))$.

3 Analysis

In this section, we first study the stablity properties of the discrete operator $A_{\mathcal{Q}}$ in a suitable residual norm. Next, we introduce interpolation operators in space and in time, and we bound the consistency error. We combine these bounds with the abstract conditional estimate from Lemma 1 to derive error estimates in the shifted energy norm. Finally, we establish the convergence rates for the method.

In what follows, we use the convention $A \lesssim B$ to abbreviate the inequality $A \leq CB$ for positive real numbers A and B, where the constant $C > 0$ does not depend on h, τ, the solution of the continuous and discrete problems.

Remark 2 (Domain regularity)**.** *The analysis is conducted under the assumption that Ω has a smooth boundary, as stated in the introduction. The numerical tests are, however, performed on a square domain (in 2d). Indeed, treating a curved domain would be quite technical, and the modifications to the implementation would be non-trivial due to the hybrid nature of the discretization method; see for instance [23]. We conjecture that the analysis also holds true for a polygonal domain. In particular, a key issue is to establish Lemma 1.*

3.1 Residual stability

Recall the broken operator $L_{\mathcal{Q}}$ defined in (11) and the notation $\mathcal{G}_h^{\mathrm{int}} := \mathcal{G}_h^{t,\mathrm{int}} \cup \mathcal{G}_h^{x,\mathrm{int}}$ together with the unit normal vector $\widetilde{\boldsymbol{n}}_q \in \mathbb{R}^{d+1}$. For all $(\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$, we define the residual norm

$$\|\!|\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}|\!\|^2 := \|v_{\mathcal{Q}}\|_{\mathrm{R}}^2 + \|v_{\mathcal{Q}}\|_{J\times\varpi}^2 + s_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{v}_{\mathcal{Q}}) + \sigma_{\mathcal{Q}}(\hat{\zeta}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}), \tag{23}$$

with

$$\|v_{\mathcal{Q}}\|_{\mathrm{R}}^2 := \|hL_{\mathcal{Q}}(v_{\mathcal{Q}})\|_Q^2 + \|h^{\frac{1}{2}}[\![\widetilde{\nabla}_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_{\mathcal{G}^{\mathrm{int}}} \cdot \widetilde{\boldsymbol{n}}_{\mathcal{G}^{\mathrm{int}}}\|_{\mathcal{G}^{\mathrm{int}}}^2, \tag{24}$$

and $\|h^{\frac{1}{2}}[\![\widetilde{\nabla}_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_{\mathcal{G}^{\mathrm{int}}} \cdot \widetilde{\boldsymbol{n}}_{\mathcal{G}^{\mathrm{int}}}\|_{\mathcal{G}^{\mathrm{int}}}^2 := \sum_{g\in\mathcal{G}_h^{\mathrm{int}}} h\|[\![\widetilde{\nabla}_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_g \cdot \widetilde{\boldsymbol{n}}_g\|_g^2$.

Lemma 2 (Norm)**.** *$\|\!|\cdot|\!\|$ defines a norm on $\hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$.*

Proof. Let $(\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$ be such that $\|\!|\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}|\!\| = 0$.

(1) The dual component $\hat{\zeta}_{\mathcal{Q}} := (\zeta_{\mathcal{Q}}, \zeta_{\mathcal{G}}) \in \hat{U}_{\mathcal{Q},0}$ satisfies $\sigma_{\mathcal{Q}}(\hat{\zeta}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) = 0$. Hence, $\zeta_{\mathcal{Q}}$ is piecewise constant in Q and $\zeta_{\mathcal{G}}$ is equal to the trace of $\zeta_{\mathcal{Q}}$ on all faces. Hence, $\zeta_{\mathcal{Q}}$ is globally constant, and it vanishes since $\zeta_{\mathcal{G}}$ vanishes on $\mathcal{G}_h^{x,\mathrm{ext}}$ and $\mathcal{G}_h^{t,\mathrm{ext}}$. In conclusion, $\hat{\zeta}_{\mathcal{Q}} = (0,0)$.

(2) The primal component $\hat{v}_{\mathcal{Q}} := (v_{\mathcal{Q}}, v_{\mathcal{G}}) \in \hat{U}_{\mathcal{Q}}$ satisfies $\|v_{\mathcal{Q}}\|_{\mathrm{R}} = \|v_{\mathcal{Q}}\|_{J\times\varpi} = s_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{v}_{\mathcal{Q}}) = 0$. Since $L_{\mathcal{Q}}(v_{\mathcal{Q}}) = 0$, $v_{\mathcal{Q}}$ is a piecewise affine polynomial in space and in time. Moreover, $s_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{v}_{\mathcal{Q}}) = 0$ implies that $v_{\mathcal{Q}}$ is globally continuous and $v_{\mathcal{G}}$ is obtained from the traces of $v_{\mathcal{Q}}$ on the faces of $\mathcal{G}_h$. Combined with the fact that $\|h^{\frac{1}{2}}[\![\widetilde{\nabla}_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_{\mathcal{G}^{\mathrm{int}}} \cdot \widetilde{\boldsymbol{n}}_{\mathcal{G}^{\mathrm{int}}}\|_{\mathcal{G}^{\mathrm{int}}} = 0$, we infer that $v_{\mathcal{Q}}$ is globally affine in Q. Finally, invoking that ϖ contains at least one time-space cell or that the face components vanish on $\mathcal{G}_h^{x,\mathrm{ext}}$, we conclude that $\hat{v}_{\mathcal{Q}} = (0,0)$. □

Our key stability result is the following inf-sup condition.

Lemma 3 (Inf-sup condition)**.** *The following holds for all $(\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$,*

$$\|\!|\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}|\!\| \lesssim \sup_{(\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0} \setminus \{(0,0)\}} \frac{A_{\mathcal{Q}}((\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}), (\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}))}{\|\!|\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}|\!\|}. \tag{25}$$

Proof. Let us denote by S the right-hand side of (25).

(1) Stabilizations. We first use the test functions $\hat{w}_{\mathcal{Q}} := \hat{v}_{\mathcal{Q}}$ and $\hat{\eta}_{\mathcal{Q}} := -\hat{\zeta}_{\mathcal{Q}}$ to get

$$\|v_{\mathcal{Q}}\|_{J\times\varpi}^2 + s_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{v}_{\mathcal{Q}}) + \sigma_{\mathcal{Q}}(\hat{\zeta}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}) = A_{\mathcal{Q}}((\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}), (\hat{v}_{\mathcal{Q}}, -\hat{\zeta}_{\mathcal{Q}})) \leq S\|\!|\hat{v}_{\mathcal{Q}}, \hat{\zeta}_{\mathcal{Q}}|\!\|.$$

(2) Gradient jumps. Let $\hat\eta_{\mathcal{Q}} := (0,\eta_{\mathcal{G}}) \in \hat U_{\mathcal{Q},0}$ with $\eta_g := h[\![B\widetilde\nabla_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_g\cdot\widetilde{\boldsymbol{n}}_g$ for all $g\in\mathcal{G}_h^{\rm int}$ and $\eta_g := 0$ otherwise. Since $[\![0,\hat\eta_{\mathcal{Q}}]\!] = \sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})^{\frac12}$, integration by parts in time and in space gives

$$\begin{aligned}\|h^{\frac12}[\![\widetilde\nabla_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_{\mathcal{G}^{\rm int}}\cdot\widetilde{\boldsymbol{n}}_{\mathcal{G}^{\rm int}}\|_{\mathcal{G}^{\rm int}}^2 &= a_{\mathcal{Q}}(\hat v_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})\\ &= A_{\mathcal{Q}}((\hat v_{\mathcal{Q}},\hat\zeta_{\mathcal{Q}}),(0,\hat\eta_{\mathcal{Q}})) + \sigma_{\mathcal{Q}}(\hat\zeta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})\\ &\le S\sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})^{\frac12} + \sigma_{\mathcal{Q}}(\hat\zeta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})\\ &\le (S+\sigma_{\mathcal{Q}}(\hat\zeta_{\mathcal{Q}},\hat\zeta_{\mathcal{Q}})^{\frac12})\sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})^{\frac12}.\end{aligned}$$

Moreover, we also have

$$\sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}}) = \sum_{q\in\mathcal{Q}_h} h^{-1}\|\eta_{\partial q}\|_{\partial q}^2 \lesssim \|h^{\frac12}[\![\widetilde\nabla_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_{\mathcal{G}^{\rm int}}\cdot\widetilde{\boldsymbol{n}}_{\mathcal{G}^{\rm int}}\|_{\mathcal{G}^{\rm int}}^2.$$

This implies that

$$\|h^{\frac12}[\![\widetilde\nabla_{\mathcal{Q}} v_{\mathcal{Q}}]\!]_{\mathcal{G}^{\rm int}}\cdot\widetilde{\boldsymbol{n}}_{\mathcal{G}^{\rm int}}\|_{\mathcal{G}^{\rm int}}^2 \lesssim S^2 + \sigma_{\mathcal{Q}}(\hat\zeta_{\mathcal{Q}},\hat\zeta_{\mathcal{Q}}).$$

(3) Cell residuals. We now consider $\hat\eta_{\mathcal{Q}} := (\eta_{\mathcal{Q}},0)$ with $\eta_{\mathcal{Q}} := h^2 L_{\mathcal{Q}}(v_{\mathcal{Q}})$. Integration by parts in time and in space gives

$$\begin{aligned}a_{\mathcal{Q}}(\hat v_{\mathcal{Q}},\hat\eta_{\mathcal{Q}}) &= \sum_{q\in\mathcal{Q}_h}\big\{(B\widetilde\nabla v_q,\widetilde\nabla\eta_q)_q - (B\widetilde\nabla v_q\cdot\widetilde{\boldsymbol{n}}_q,\eta_q)_{\partial q} - (v_q - v_{\partial q}, B\widetilde\nabla\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}\big\}\\ &= (L_{\mathcal{Q}}(v_{\mathcal{Q}}),\eta_{\mathcal{Q}})_{\mathcal{Q}} - \sum_{q\in\mathcal{Q}_h}(v_q - v_{\partial q}, B\widetilde\nabla\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}.\end{aligned}$$

Using the definition of $\hat\eta_{\mathcal{Q}}$, we have

$$\begin{aligned}\|hL_{\mathcal{Q}}(v_{\mathcal{Q}})\|_{\mathcal{Q}}^2 &= a_{\mathcal{Q}}(\hat v_{\mathcal{Q}},\hat\eta_{\mathcal{Q}}) + \sum_{q\in\mathcal{Q}_h}(v_q - v_{\partial q}, B\widetilde\nabla\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}\\ &= A_{\mathcal{Q}}((\hat v_{\mathcal{Q}},\hat\zeta_{\mathcal{Q}}),(0,\hat\eta_{\mathcal{Q}})) + \sigma_{\mathcal{Q}}(\hat\zeta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}}) + \sum_{q\in\mathcal{Q}_h}(v_q - v_{\partial q}, B\widetilde\nabla\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}\\ &\lesssim S[\![0,\hat\eta_{\mathcal{Q}}]\!] + (\sigma_{\mathcal{Q}}(\hat\zeta_{\mathcal{Q}},\hat\zeta_{\mathcal{Q}}) + s_{\mathcal{Q}}(\hat v_{\mathcal{Q}},\hat v_{\mathcal{Q}}))^{\frac12}\sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})^{\frac12},\end{aligned}$$

since a discrete trace inverse inequality implies that

$$\sum_{q\in\mathcal{Q}_h} h\|\widetilde\nabla\eta_q\cdot\widetilde{\boldsymbol{n}}_q\|_{\partial q}^2 \lesssim \|\widetilde\nabla_{\mathcal{Q}}\eta_{\mathcal{Q}}\|_{\mathcal{Q}}^2 \le \sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}}).$$

Furthermore, invoking again inverse inequalities, we have

$$\sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}}) = \sum_{q\in\mathcal{Q}_h}\Big(\|\widetilde\nabla\eta_q\|_q^2 + h^{-1}\|\eta_q\|_{\partial q}^2\Big) \lesssim \|h^{-1}\eta_{\mathcal{Q}}\|_{\mathcal{Q}}^2 = \|hL_{\mathcal{Q}}(v_{\mathcal{Q}})\|_{\mathcal{Q}}^2.$$

Using $[\![0,\hat\eta_{\mathcal{Q}}]\!]^2 = \sigma_{\mathcal{Q}}(\hat\eta_{\mathcal{Q}},\hat\eta_{\mathcal{Q}})$ and Young's inequality, this implies that

$$\|hL_{\mathcal{Q}}(v_{\mathcal{Q}})\|_{\mathcal{Q}}^2 \lesssim S^2 + s_{\mathcal{Q}}(\hat v_{\mathcal{Q}},\hat v_{\mathcal{Q}}) + \sigma_{\mathcal{Q}}(\hat\zeta_{\mathcal{Q}},\hat\zeta_{\mathcal{Q}}).$$

Gathering the previous estimates leads to the expected inf-sup condition. □

Corollary 1 (Well-posedness)**.** *The discrete problems* (18) *and* (21) *are well-posed.*

Proof. Since both (18) and (21) amount to square linear systems, it suffices to prove that, for each system, the only solution with zero right-hand side is zero. For (18), this follows from the inf-sup condition and Lemma 2. For (21), one needs to adapt the above proofs. First, adapting the proof of Lemma 2, one observes that the norm

$$|||v_{\mathcal{Q}},\zeta_{\mathcal{Q}}|||^2 := \|hL_{\mathcal{Q}}(v_{\mathcal{Q}})\|_Q^2 + \|v_{\mathcal{Q}}\|_{J\times\varpi}^2 + s_{\mathcal{Q}}((v_{\mathcal{Q}},0),(v_{\mathcal{Q}},0)) + \sigma_{\mathcal{Q}}((\zeta_{\mathcal{Q}},0),(\zeta_{\mathcal{Q}},0)),$$

defines a norm on $U_{\mathcal{Q}}\times U_{\mathcal{Q}}$. Then, adapting the proof of Lemma 3, one observes that an inf-sup condition holds true for the above norm when considering the bilinear form $A_{\mathcal{Q}}(((u_{\mathcal{Q}},0),(\xi_{\mathcal{Q}},0)),((w_{\mathcal{Q}},0),(\eta_{\mathcal{Q}},0)))$. □

3.2 Interpolation operator

In this section, we define the space-time interpolation operator used in the error analysis. Let $q\in\mathcal{Q}_h$. For all $v\in L^2(q)$, we define $I_q(v)$ to be the L^2-orthogonal projection of v onto U_q. For all $v\in H^1(q)$, we define $I_{\partial q}(v)$ as the the L^2-orthogonal projection of $v|_{\partial q}$ onto $U_{\partial q}$. We also define $\hat{I}_q(v) := (I_q(v), I_{\partial q}(v))$. Similarly, for all $v\in H^1(Q)$, we define the interpolation on the whole discrete space by $\hat{I}_{\mathcal{Q}}(v) := (I_{\mathcal{Q}}(v), I_{\mathcal{G}}(v))$, where $I_{\mathcal{Q}}(v)|_q := I_q(v)$ for all $q\in\mathcal{Q}_h$ and $I_{\mathcal{G}}(v)|_g := I_{\partial q}(v)|_g$ for all $g\in\mathcal{G}_h$ with $g\subset\partial q$. Notice that $\hat{I}_{\mathcal{Q}}(v)\in\hat{U}_{\mathcal{Q}}$ if $v=0$ on $J\times\partial\Omega$, and $\hat{I}_{\mathcal{Q}}(v)\in\hat{U}_{\mathcal{Q},0}$ if additionally $v(0)=v(T_f)=0$.

The above interpolation operator enjoys the following convergence properties (the proof is omitted since it is standard).

Lemma 4 (Approximation)**.** *Let $k\geq 1$ and $\ell\geq 1$. The following holds for all $q := T^t\times T^x\in\mathcal{Q}_h$, $T^t\in\mathcal{T}_\tau^t$, $T^x\in\mathcal{T}_h^x$, all $v\in H^{\ell+1}(T^t;H^2(T^x))\cap H^2(T^t;H^{k+1}(T^x))$,*

$$\|I_q(v)-v\|_q \lesssim h^{\ell+1}\|v\|_{H^{\ell+1}(T^t;L^2(T^x))} + h^{k+1}\|v\|_{L^2(T^t;H^{k+1}(T^x))}, \tag{26a}$$

$$\|\widetilde{\nabla}(I_q(v)-v)\|_q \lesssim h^{\ell}\|v\|_{H^{\ell+1}(T^t;H^1(T^x))} + h^{k}\|v\|_{H^1(T^t;H^{k+1}(T^x))}, \tag{26b}$$

$$h^{\frac12}\|\widetilde{\nabla}(I_q(v)-v)\|_{\partial q} \lesssim h^{\ell}\|v\|_{H^{\ell+1}(T^t;H^2(T^x))} + h^{k}\|v\|_{H^2(T^t;H^{k+1}(T^x))}, \tag{26c}$$

$$h\|L(I_q(v)-v)\|_q \lesssim h^{\ell}\|v\|_{H^{\ell+1}(T^t;H^2(T^x))} + h^{k}\|v\|_{H^2(T^t;H^{k+1}(T^x))}. \tag{26d}$$

Moreover, we have

$$\begin{aligned} h^{-\frac12}\|I_q(v)-v\|_{\partial q} + h^{-\frac12}\|I_{\partial q}(v)-v\|_{\partial q} \\ \lesssim h^{\ell}\|v\|_{H^{\ell+1}(T^t;H^1(T^x))} + h^{k}\|v\|_{H^1(T^t;H^{k+1}(T^x))}. \end{aligned} \tag{27}$$

3.3 Error decomposition, consistency and a priori residual bound

Recall that u denotes the solution to the exact problem (1) and that $(\hat{u}_{\mathcal{Q}},\hat{\xi}_{\mathcal{Q}})$ denotes the solution to the discrete problem (18). The discrete and interpolation errors on the primal unknown are defined as

$$\hat{e}_{\mathcal{Q}} := \hat{u}_{\mathcal{Q}} - \hat{I}_{\mathcal{Q}}(u), \qquad \hat{\theta}_{\mathcal{Q}} := (\theta_{\mathcal{Q}},\theta_{\mathcal{G}}), \tag{28}$$

where $\theta_q := u|_q - I_q(u)$ for all $q\in\mathcal{Q}_h$ and $\theta_g := u|_g - I_{\mathcal{G}}(u)|_g$ for all $g\in\mathcal{G}_h$.

We now bound the consistency error. To this purpose, we consider the norm

$$\begin{aligned} \|\hat{\theta}_{\mathcal{Q}}\|_{\#}^2 &:= \|\widetilde{\nabla}_{\mathcal{Q}}\theta_{\mathcal{Q}}\|_Q^2 + \|\theta_{\mathcal{Q}}\|_Q^2 \\ &\quad + \sum_{q\in\mathcal{Q}_h}\big\{\|hL(\theta_q)\|_q^2 + \|h^{\frac12}\widetilde{\nabla}\theta_q\|_{\partial q}^2 + \|h^{-\frac12}\theta_q\|_{\partial q}^2 + \|h^{-\frac12}\theta_{\partial q}\|_{\partial q}^2\big\}. \end{aligned} \tag{29}$$

Recall that the stability norm $|||\cdot|||$ is defined in (23)-(24). In what follows, we assume that

$$u\in H^2(J;L^2(\Omega))\cap L^2(J;H^2(\Omega)). \tag{30}$$

Lemma 5 (Consistency)**.** *Let $(\hat{u}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}})$ denote the solution to the discrete problem* (18)*. Let $\hat{e}_{\mathcal{Q}}$ and $\hat{\theta}_{\mathcal{Q}}$ be the discrete and interpolation errors on the primal unknown defined in* (28)*. Under the regularity assumption* (30)*, we have, for all $(\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$,*

$$|A_{\mathcal{Q}}((\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (\hat{w}_{\mathcal{Q}}, 0))| \lesssim \big(\|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}\big)|\!|\!|\hat{w}_{\mathcal{Q}}, 0|\!|\!|, \tag{31a}$$

$$|A_{\mathcal{Q}}((\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (0, \hat{\eta}_{\mathcal{Q}}))| \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#}|\!|\!|0, \hat{\eta}_{\mathcal{Q}}|\!|\!|. \tag{31b}$$

Proof. Proof of (31a). Since $(\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}) = (\hat{u}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}) - (\hat{I}_{\mathcal{Q}}(u), 0)$, the equation (18a) in the discrete problem, the definition (19) of $A_{\mathcal{Q}}$, and $g_\delta = g + \delta$ imply that, for all $\hat{w}_{\mathcal{Q}} \in \hat{U}_{\mathcal{Q}}$,

$$\begin{aligned} A_{\mathcal{Q}}((\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (\hat{w}_{\mathcal{Q}}, 0)) &= (g_\delta, w_{\mathcal{Q}})_{J\times\varpi} - (I_{\mathcal{Q}}(u), w_{\mathcal{Q}})_{J\times\varpi} - s_{\mathcal{Q}}(\hat{I}_{\mathcal{Q}}(u), \hat{w}_{\mathcal{Q}}) \\ &= (\delta, w_{\mathcal{Q}})_{J\times\varpi} + (\theta_{\mathcal{Q}}, w_{\mathcal{Q}})_{J\times\varpi} - s_{\mathcal{Q}}(\hat{I}_{\mathcal{Q}}(u), \hat{w}_{\mathcal{Q}}). \end{aligned}$$

The Cauchy–Schwarz inequality yields

$$\begin{aligned} |(\delta, w_{\mathcal{Q}})_{J\times\varpi}| + |(\theta_{\mathcal{Q}}, w_{\mathcal{Q}})_{J\times\varpi}| &\lesssim (\|\delta\|_{J\times\varpi} + \|\theta_{\mathcal{Q}}\|_{J\times\varpi})\|w_{\mathcal{Q}}\|_{J\times\varpi}, \\ |s_{\mathcal{Q}}(\hat{I}_{\mathcal{Q}}(u), \hat{w}_{\mathcal{Q}})| &\lesssim s_{\mathcal{Q}}(\hat{I}_{\mathcal{Q}}(u), \hat{I}_{\mathcal{Q}}(u))^{\frac{1}{2}} s_{\mathcal{Q}}(\hat{w}_{\mathcal{Q}}, \hat{w}_{\mathcal{Q}})^{\frac{1}{2}}. \end{aligned}$$

Moreover, we have

$$s_{\mathcal{Q}}(\hat{I}_{\mathcal{Q}}(u), \hat{I}_{\mathcal{Q}}(u)) = \sum_{q\in\mathcal{Q}_h} h^{-1}\|I_q(u) - I_{\partial q}(u)\|_{\partial q}^2 \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#}^2.$$

The bound (31a) follows by gathering the above estimates.

Proof of (31b). Proceeding as above and using now the equation (18b) and $f = L(u)$, we have, for all $\hat{\eta}_{\mathcal{Q}} \in \hat{U}_{\mathcal{Q},0}$,

$$A_{\mathcal{Q}}((\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (0, \hat{\eta}_{\mathcal{Q}})) = (L(u), \eta_{\mathcal{Q}})_Q - a_{\mathcal{Q}}(\hat{I}_{\mathcal{Q}}(u), \hat{\eta}_{\mathcal{Q}}).$$

Integrating by parts the first term on the right-hand side gives

$$(L(u), \eta_{\mathcal{Q}})_Q = \sum_{q\in\mathcal{Q}_h} \big\{(B\widetilde{\nabla}u, \widetilde{\nabla}\eta_q)_q - (B\widetilde{\nabla}u\cdot\widetilde{\boldsymbol{n}}_q, \eta_q - \eta_{\partial q})_{\partial q}\big\},$$

where we used the regularity assumption on u and the fact $\eta_{\partial q}$ vanishes on all the boundary faces in space and in time. Recalling the definition (12)-(14) of $a_{\mathcal{Q}}$, we infer that

$$\begin{aligned} A_{\mathcal{Q}}((\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (0, \hat{\eta}_{\mathcal{Q}})) = \sum_{q\in\mathcal{Q}_h} \big\{&(B\widetilde{\nabla}\theta_q, \widetilde{\nabla}\eta_q)_q - (B\widetilde{\nabla}\theta_q\cdot\widetilde{\boldsymbol{n}}_q, \eta_q - \eta_{\partial q})_{\partial q} \\ &- (\theta_q - \theta_{\partial q}, B\widetilde{\nabla}\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}\big\}. \end{aligned}$$

Using Cauchy–Schwarz and inverse inequalities, we readily obtain (31b). □

Lemma 6 (A priori residual bound)**.** *Under the regularity assumption* (30)*, we have*

$$|\!|\!|\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}|\!|\!| + s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}} \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}. \tag{32}$$

Proof. Using inf-sup stability (Lemma 3), we have

$$|\!|\!|\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}|\!|\!| \lesssim \sup_{(\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}})\in\hat{U}_{\mathcal{Q}}\times\hat{U}_{\mathcal{Q},0}\setminus\{(0,0)\}} \frac{A_{\mathcal{Q}}((\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}))}{|\!|\!|\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}|\!|\!|}.$$

Moreover, owing to consistency (Lemma 5), we have, for all $(\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) \in \hat{U}_{\mathcal{Q}} \times \hat{U}_{\mathcal{Q},0}$,

$$|A_{\mathcal{Q}}((\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}), (\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}))| \lesssim (\|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi})|\!|\!|\hat{w}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}|\!|\!|.$$

Combining the two bounds gives

$$|||\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}||| \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}.$$

Finally, observing that

$$s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}} \le s_{\mathcal{Q}}(\hat{I}_{\mathcal{Q}}(u), \hat{I}_{\mathcal{Q}}(u))^{\frac{1}{2}} + s_{\mathcal{Q}}(\hat{e}_{\mathcal{Q}}, \hat{e}_{\mathcal{Q}})^{\frac{1}{2}} \le \|\hat{\theta}_{\mathcal{Q}}\|_{\#} + |||\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}|||,$$

concludes the proof. □

3.4 Bound on dual error norm

In this section, we prove another important result bounding some dual norm of the error. Recall the operator $L := \partial_{tt}^2 - \Delta$ so that, for all $v \in H^1(Q)$ and all $\eta \in H_0^1(Q)$, we have $\langle L(v), \eta\rangle_{H^{-1},H_0^1} = (B\widetilde{\nabla}v, \widetilde{\nabla}\eta)_Q$, where $\langle\cdot,\cdot\rangle_{H^{-1},H_0^1}$ stands for the duality product between $H^{-1}(Q)$ and $H_0^1(Q)$. Observing that the broken operator $L_{\mathcal{Q}}$ acting on discrete functions $v_{\mathcal{Q}} \in U_{\mathcal{Q}}$ satisfies $\langle L_{\mathcal{Q}}(v_{\mathcal{Q}}), \eta\rangle_{H^{-1},H_0^1} = (B\widetilde{\nabla}_{\mathcal{Q}}v_{\mathcal{Q}}, \widetilde{\nabla}\eta)_Q$, we can consider the extended operator $L_{\mathcal{Q}} : H^1(Q)+U_{\mathcal{Q}} \to H^{-1}(Q)$ (we use the same notation for simplicity). For all $v \in H^1(Q)+U_{\mathcal{Q}}$, the corresponding dual norm is

$$\|L_{\mathcal{Q}}(v)\|_{H^{-1}(Q)} := \sup_{\substack{\eta\in H_0^1(Q)\\ \|\widetilde{\nabla}\eta\|_Q=1}} (B\widetilde{\nabla}_{\mathcal{Q}}v, \widetilde{\nabla}\eta)_Q.$$

Lemma 7 (Bound on dual error norm)**.** *Under the regularity assumption* (30)*, we have*

$$\|L_{\mathcal{Q}}(u_{\mathcal{Q}} - u)\|_{H^{-1}(Q)} \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}.$$

Proof. Let $\eta \in H_0^1(Q)$ with $\|\widetilde{\nabla}\eta\|_Q = 1$. We have

$$\langle L_{\mathcal{Q}}(u_{\mathcal{Q}} - u), \eta\rangle_{H^{-1},H_0^1} = (B\widetilde{\nabla}_{\mathcal{Q}}u_{\mathcal{Q}}, \widetilde{\nabla}\eta)_Q - (L(u), \eta)_Q.$$

Let $\hat{\eta}_{\mathcal{Q}} \in \hat{U}_{\mathcal{Q},0}$ be such that $\hat{\eta}_{\mathcal{Q}} := \hat{I}_{\mathcal{Q}}(\eta)$. Invoking equation (18b) gives

$$a_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) - \sigma_{\mathcal{Q}}(\hat{\xi}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) = (L(u), \eta_{\mathcal{Q}})_Q.$$

Subtracting this equation from the above expression, and re-arranging the terms gives

$$\langle L_{\mathcal{Q}}(u_{\mathcal{Q}} - u), \eta\rangle_{H^{-1},H_0^1} = A_1 + A_2 + A_3,$$

with

$$\begin{aligned}
A_1 &:= (L_{\mathcal{Q}}(u_{\mathcal{Q}} - u), \eta - \eta_{\mathcal{Q}})_Q,\\
A_2 &:= (B\widetilde{\nabla}_{\mathcal{Q}}u_{\mathcal{Q}}, \widetilde{\nabla}\eta)_Q - a_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}) - (L_{\mathcal{Q}}(u_{\mathcal{Q}}), \eta - \eta_{\mathcal{Q}})_Q,\\
A_3 &:= \sigma_{\mathcal{Q}}(\hat{\xi}_{\mathcal{Q}}, \hat{\eta}_{\mathcal{Q}}).
\end{aligned}$$

We have

$$\begin{aligned}
&\sum_{q\in\mathcal{Q}_h} \{(B\widetilde{\nabla}u_q, \widetilde{\nabla}\eta)_q - a_q(\hat{u}_q, \hat{\eta}_q)\}\\
&\quad= \sum_{q\in\mathcal{Q}_h} \big\{(B\widetilde{\nabla}u_q, \widetilde{\nabla}(\eta - \eta_q))_q + (B\widetilde{\nabla}u_q\cdot\widetilde{\boldsymbol{n}}_q, \eta_q - \eta_{\partial q})_{\partial q} + (u_q - u_{\partial q}, B\widetilde{\nabla}\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}\big\}\\
&\quad= \sum_{q\in\mathcal{Q}_h} \big\{(L(u_q), \eta - \eta_q)_q + (B\widetilde{\nabla}u_q\cdot\widetilde{\boldsymbol{n}}_q, \eta - \eta_{\partial q})_{\partial q} + (u_q - u_{\partial q}, B\widetilde{\nabla}\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}\big\}\\
&\quad= (L_{\mathcal{Q}}(u_{\mathcal{Q}}), \eta - \eta_{\mathcal{Q}})_Q + \sum_{q\in\mathcal{Q}_h} (u_q - u_{\partial q}, B\widetilde{\nabla}\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q},
\end{aligned}$$

where we used the definition (12) of the bilinear form a_q, the fact that $(B\widetilde{\nabla}u_q\cdot\widetilde{\boldsymbol{n}}_q)|_{\partial q}\in U_{\partial q}$ and the definition of $\eta_{\partial q}$ so that $(B\widetilde{\nabla}u_q\cdot\widetilde{\boldsymbol{n}}_q,\eta-\eta_{\partial q})_{\partial q}=0$. We can then rewrite A_2 as

$$A_2=\sum_{q\in\mathcal{Q}_h}(u_q-u_{\partial q},B\widetilde{\nabla}\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}.$$

It still remains to bound A_1, A_2, A_3.

Bound on A_1. Since $\theta_{\mathcal{Q}}=u-I_{\mathcal{Q}}(u)$, we have

$$L_{\mathcal{Q}}(u_{\mathcal{Q}}-u)=L_{\mathcal{Q}}(u_{\mathcal{Q}}-I_{\mathcal{Q}}(u))-L_{\mathcal{Q}}(\theta_{\mathcal{Q}}).$$

Recalling the definition (24) of the $\|\cdot\|_{\mathrm{R}}$-norm, we have

$$|(L_{\mathcal{Q}}(u_{\mathcal{Q}}-I_{\mathcal{Q}}(u)),\eta-\eta_{\mathcal{Q}})_Q|\lesssim h^{-1}\|e_{\mathcal{Q}}\|_{\mathrm{R}}h\|\widetilde{\nabla}\eta\|_Q,$$

where we used that $\|\eta-\eta_{\mathcal{Q}}\|_Q\lesssim h\|\widetilde{\nabla}\eta\|_Q$. Owing to Lemma 6, we have $\|e_{\mathcal{Q}}\|_{\mathrm{R}}\lesssim\|\hat{\theta}_{\mathcal{Q}}\|_{\#}+\|\delta\|_{J\times\varpi}$. Moreover, the above bound on $\|\eta-\eta_{\mathcal{Q}}\|_Q$ also gives

$$|(L_{\mathcal{Q}}(\theta_{\mathcal{Q}}),\eta-\eta_{\mathcal{Q}})_Q|\lesssim h\|L_{\mathcal{Q}}(\theta_{\mathcal{Q}})\|_Q\|\widetilde{\nabla}\eta\|_Q\le\|\hat{\theta}_{\mathcal{Q}}\|_{\#}\|\widetilde{\nabla}\eta\|_Q.$$

Gathering the above estimates yields

$$|A_1|\lesssim\big(\|\hat{\theta}_{\mathcal{Q}}\|_{\#}+\|\delta\|_{J\times\varpi}\big)\|\widetilde{\nabla}\eta\|_Q.$$

Bound on A_2. Using Cauchy–Schwarz and inverse trace inequalities and interpolation stability, we have

$$\begin{aligned}|A_2|=\bigg|\sum_{q\in\mathcal{Q}_h}(u_q-u_{\partial q},B\widetilde{\nabla}\eta_q\cdot\widetilde{\boldsymbol{n}}_q)_{\partial q}\bigg|&\lesssim s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}},\hat{u}_{\mathcal{Q}})^{\frac12}\|\widetilde{\nabla}_{\mathcal{Q}}\eta_{\mathcal{Q}}\|_Q\\&\lesssim(\|\hat{\theta}_{\mathcal{Q}}\|_{\#}+\|\delta\|_{J\times\varpi})\|\widetilde{\nabla}\eta\|_Q,\end{aligned}$$

where we used Lemma 6 and $\|\widetilde{\nabla}_{\mathcal{Q}}\eta_{\mathcal{Q}}\|_Q\lesssim\|\widetilde{\nabla}\eta\|_Q$ in the last bound.

Bound on A_3. Recalling that $\sigma_{\mathcal{Q}}(\hat{\eta}_{\mathcal{Q}},\hat{\eta}_{\mathcal{Q}})=\|\widetilde{\nabla}_{\mathcal{Q}}\eta_{\mathcal{Q}}\|_Q^2+s_{\mathcal{Q}}(\hat{\eta}_{\mathcal{Q}},\hat{\eta}_{\mathcal{Q}})$, we infer that $\sigma_{\mathcal{Q}}(\hat{\eta}_{\mathcal{Q}},\hat{\eta}_{\mathcal{Q}})^{\frac12}\lesssim\|\widetilde{\nabla}_{\mathcal{Q}}\eta_{\mathcal{Q}}\|_Q\lesssim\|\widetilde{\nabla}\eta\|_Q$. This implies that

$$|\sigma_{\mathcal{Q}}(\hat{\xi}_{\mathcal{Q}},\hat{\eta}_{\mathcal{Q}})|\lesssim\sigma_{\mathcal{Q}}(\hat{\xi}_{\mathcal{Q}},\hat{\xi}_{\mathcal{Q}})^{\frac12}\sigma_{\mathcal{Q}}(\hat{\eta}_{\mathcal{Q}},\hat{\eta}_{\mathcal{Q}})^{\frac12}\lesssim\sigma_{\mathcal{Q}}(\hat{\xi}_{\mathcal{Q}},\hat{\xi}_{\mathcal{Q}})^{\frac12}\|\widetilde{\nabla}\eta\|_Q,$$

and $\sigma_{\mathcal{Q}}(\hat{\xi}_{\mathcal{Q}},\hat{\xi}_{\mathcal{Q}})^{\frac12}$ is bounded in Lemma 6. Hence, we have

$$|A_3|\lesssim\big(\|\hat{\theta}_{\mathcal{Q}}\|_{\#}+\|\delta\|_{J\times\varpi}\big)\|\widetilde{\nabla}\eta\|_Q.$$

Combining the above estimates and recalling that $\|\widetilde{\nabla}\eta\|_Q=1$ gives the expected bound. □

3.5 Main result: error estimate in shifted energy norm

We are now ready to derive our main error estimate. The idea of the proof is to combine the results of Sections 3.3 and 3.4 with the conditional stability estimate from Lemma 1. We use the shifted energy norm which we extend to $H^1(Q)+U_{\mathcal{Q}}$ by setting

$$\|v\|_{\mathrm{sft}}:=\|v\|_{L^\infty(J;L^2(\Omega))}+\|\partial_{t,\mathcal{Q}}v\|_{L^2(J;H^{-1}(\Omega))}+\|\nabla_{\mathcal{Q}}v\|_{H^{-1}(J;L^2(\Omega))}.\tag{33}$$

We shall invoke the following classical result on finite element interpolation, here written in a space-time setting using the space-time mesh regularity assumption (i.e. $h\lesssim\tau\lesssim h$). It is established first on pairs of the form $(v_{\mathcal{Q}},0)$ (see for instance [24, 25, 26, 27]) and then extended to pairs $\hat{v}_{\mathcal{Q}}\in\hat{U}_{\mathcal{Q}}$ by invoking a triangle inequality.

Lemma 8. *For all $\hat{v}_{\mathcal{Q}} \in \hat{U}_{\mathcal{Q}}$, there exists $\widetilde{v}_{\mathcal{Q}} \in U_{\mathcal{Q}} \cap H^1(Q)$ such that*

$$\|h^{-1}(\widetilde{v}_{\mathcal{Q}} - v_{\mathcal{Q}})\|_Q + \|\widetilde{\nabla}_{\mathcal{Q}}(\widetilde{v}_{\mathcal{Q}} - v_{\mathcal{Q}})\|_Q \lesssim s_{\mathcal{Q}}(\hat{v}_{\mathcal{Q}}, \hat{v}_{\mathcal{Q}})^{\frac{1}{2}}.$$

Recall the constant $C_{\mathrm{stb}} > 0$ introduced in Lemma 1. Recall that the $\|\cdot\|_{\#}$-norm is defined in (29). We have the following convergence result.

Theorem 9 (Error estimate in shifted energy norm)**.** *Under the regularity assumption* (30) *and Assumption 1, we have*

$$\|u - u_{\mathcal{Q}}\|_{\mathrm{sft}} \lesssim (1 + C_{\mathrm{stb}})\big(\|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}\big). \tag{34}$$

Proof. (1) H^1-conforming approximation. Let $\widetilde{u}_{\mathcal{Q}}$ be the H^1-conforming approximation to $\hat{u}_{\mathcal{Q}}$ given by Lemma 8. We have

$$\|h^{-1}(\widetilde{u}_{\mathcal{Q}} - u_{\mathcal{Q}})\|_Q + \|\widetilde{\nabla}_{\mathcal{Q}}(\widetilde{u}_{\mathcal{Q}} - u_{\mathcal{Q}})\|_Q \lesssim s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}}. \tag{35}$$

Invoking the triangle inequality yields

$$\|u - u_{\mathcal{Q}}\|_{\mathrm{sft}} \le \|u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}}\|_{\mathrm{sft}} + \|u - \widetilde{u}_{\mathcal{Q}}\|_{\mathrm{sft}},$$

and we are left with bounding the two terms on the right-hand side.

(2) Bound on $\|u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}}\|_{\mathrm{sft}}$. We observe that

$$\begin{aligned}\|\nabla_{\mathcal{Q}}(u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}})\|_{H^{-1}(J;L^2(\Omega))} + \|\partial_{t,\mathcal{Q}}(u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}})\|_{L^2(J;H^{-1}(\Omega))} &\\ \lesssim \|\widetilde{\nabla}_{\mathcal{Q}}(u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}})\|_{L^2(Q)} &\lesssim s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}},\\ \|u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}}\|_{L^\infty(J;L^2(\Omega))} \lesssim h^{-\frac{1}{2}}\|u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}}\|_{L^2(Q)} &\lesssim h^{\frac{1}{2}} s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}},\end{aligned}$$

where we used (35), and an inverse trace inequality for the last estimate. As a consequence, owing to Lemma 6 and since $h \lesssim 1$, we infer that

$$\|u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}}\|_{\mathrm{sft}} \lesssim s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}} \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}.$$

(3) Bound on $\|u - \widetilde{u}_{\mathcal{Q}}\|_{\mathrm{sft}}$. Since $u - \widetilde{u}_{\mathcal{Q}} \in H^1(Q)$, recalling Assumption 1, we can invoke the conditional stability estimate from Lemma 1. This gives

$$\|u - \widetilde{u}_{\mathcal{Q}}\|_{\mathrm{sft}} \lesssim C_{\mathrm{stb}}\Big(\|u - \widetilde{u}_{\mathcal{Q}}\|_{J\times\varpi} + \|L(u - \widetilde{u}_{\mathcal{Q}})\|_{H^{-1}(Q)}\Big).$$

It remains to bound $\|u - \widetilde{u}_{\mathcal{Q}}\|_{J\times\varpi}$ and $\|L(u - \widetilde{u}_{\mathcal{Q}})\|_{H^{-1}(Q)}$.

(3.a) Bound on $\|u - \widetilde{u}_{\mathcal{Q}}\|_{J\times\varpi}$. We have

$$\|u - \widetilde{u}_{\mathcal{Q}}\|_{J\times\varpi} \le \|u - I_{\mathcal{Q}}(u)\|_{J\times\varpi} + \|I_{\mathcal{Q}}(u) - u_{\mathcal{Q}}\|_{J\times\varpi} + \|u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}}\|_{J\times\varpi},$$

where $I_{\mathcal{Q}}(u)$ is defined in Section 3.2. We have

$$\begin{aligned}\|u - I_{\mathcal{Q}}(u)\|_{J\times\varpi} &= \|\theta_{\mathcal{Q}}\|_{J\times\varpi} \le \|\theta_{\mathcal{Q}}\|_Q,\\ \|I_{\mathcal{Q}}(u) - u_{\mathcal{Q}}\|_{J\times\varpi} &= \|e_{\mathcal{Q}}\|_{J\times\varpi} \le |\!|\!|\hat{e}_{\mathcal{Q}}, \hat{\xi}_{\mathcal{Q}}|\!|\!|.\end{aligned}$$

Furthermore, the bound (35) and $h \lesssim 1$ imply that

$$\|u_{\mathcal{Q}} - \widetilde{u}_{\mathcal{Q}}\|_{J\times\varpi} \lesssim h s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}} \lesssim s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}}, \hat{u}_{\mathcal{Q}})^{\frac{1}{2}}.$$

Gathering the above estimates and using Lemma 6 gives

$$\|u - \widetilde{u}_{\mathcal{Q}}\|_{J\times\varpi} \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}.$$

(3.b) Bound on $\|L(u-\widetilde{u}_{\mathcal{Q}})\|_{H^{-1}(Q)}$. Lemma 7 gives

$$\|L_{\mathcal{Q}}(u-u_{\mathcal{Q}})\|_{H^{-1}(Q)} \lesssim \|\hat{\theta}_{\mathcal{Q}}\|_{\#} + \|\delta\|_{J\times\varpi}.$$

Moreover, using the Cauchy-Schwarz inequality and (35), we have

$$\begin{aligned}\|L_{\mathcal{Q}}(u_{\mathcal{Q}}-\widetilde{u}_{\mathcal{Q}})\|_{H^{-1}(Q)} &= \sup_{\substack{\eta\in H_0^1(Q)\\ \|\widetilde{\nabla}\eta\|_Q=1}} (B\widetilde{\nabla}_{\mathcal{Q}}(u_{\mathcal{Q}}-\widetilde{u}_{\mathcal{Q}}),\widetilde{\nabla}\eta)_Q \\ &\le \|B\widetilde{\nabla}_{\mathcal{Q}}(u_{\mathcal{Q}}-\widetilde{u}_{\mathcal{Q}})\|_Q \lesssim s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}},\hat{u}_{\mathcal{Q}})^{\frac{1}{2}}.\end{aligned}$$

(4) Combining the bounds from the above steps proves (34). □

3.6 Convergence

We consider the functional space

$$U_* := H^{\ell+1}(J;H^2(\Omega)) \cap H^2(J;H^{k+1}(\Omega)), \tag{36}$$

equipped with its natural norm. An important consequence of Lemma 4 is that, under the assumption $u \in U_*$, we have

$$\|\hat{\theta}_{\mathcal{Q}}\|_{\#} \lesssim h^{\ell}\|u\|_{H^{\ell+1}(J;H^2)} + h^k\|u\|_{H^2(J;H^{k+1})}, \tag{37}$$

where we used that $h \lesssim 1$ to simplify the expression.

Theorem 10 (Decay rates)**.** *Assume that $u \in U_*$. We have the following decay rates on the errors:*

$$\|\!|\hat{e}_{\mathcal{Q}},\hat{\xi}_{\mathcal{Q}}|\!\| + s_{\mathcal{Q}}(\hat{u}_{\mathcal{Q}},\hat{u}_{\mathcal{Q}})^{\frac{1}{2}} \lesssim h^{\ell}\|u\|_{H^{\ell+1}(J;H^2)} + h^k\|u\|_{H^2(J;H^{k+1})} + \|\delta\|_{J\times\varpi}, \tag{38a}$$

$$\|u-u_{\mathcal{Q}}\|_{\mathrm{sft}} \lesssim (1+C_{\mathrm{stb}})\big(h^{\ell}\|u\|_{H^{\ell+1}(J;H^2)} + h^k\|u\|_{H^2(J;H^{k+1})} + \|\delta\|_{J\times\varpi}\big). \tag{38b}$$

Proof. This is a consequence of plugging (37) into (32) and (34). □

4 Numerical results

In this section, we present numerical experiments to check the convergence rates established in Theorem 10. We also study the influence of noise in the measurements, and we illustrate the benefits of using a high-order discretization.

The space domain is $\Omega := (0,1)^d$ with $d \in \{1,2\}$. All the errors are computed as the difference between the numerical solution and the $L^2(Q)$-orthogonal projection of the exact solution onto the cell discretization space. These errors are measured in the $L^\infty(0,T_f;L^2(\Omega))$-norm. In practice, this norm is computed as the maximum value over all the Gauss points of each time interval.

Some noise can be added to the measurements in the following way: (i) the time-space domain is divided into 10^{d+1} subdomains; (ii) a random noise level $\delta := a * \mathrm{rand}()$ is assigned to each subdomain, with $a \ge 0$ the noise amplitude and rand() is a `C++` function returning a random number in $[-1,1]$. Thus, every subdomain has the same noise during the whole mesh-refinement process.

The time mesh is considered to be a uniform mesh made of N time cells.

All the tests are run with the `DiSk++` library [28], and all the linear systems are solved using the Pardiso solver from the MKL library. During the matrix assembly process, we compute the value of the coefficients for the first time interval I_1. Then, since the time step is constant, the coefficients associated with the following time intervals are the same and therefore do not need to be recomputed.

4.1 One-dimensional test cases

In this section, we consider the following one-dimensional setting:

$$\Omega := (0,1), \quad \varpi := \left(\frac{1}{4}, \frac{3}{4}\right), \quad T_f := 2, \quad u(t,x) := \cos(\pi t)\sin(\pi x).$$

We notice that this geometry fulfills Assumption 1, since every wave has to cross the set $J \times \varpi$ at some point.

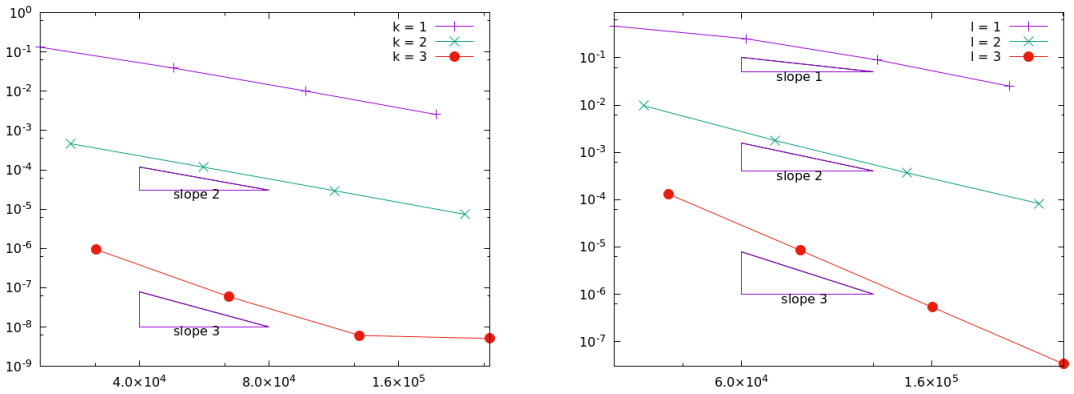

Figure 2: $L^\infty(L^2)$-norm errors for the 1d test case without noise (with respect to the number of degrees of freedom). Left: space convergence ($M \in \{16, 32, 64, 128\}$, $N = 128$, $\ell = 3$). Right: time convergence ($N \in \{10, 20, 40, 80\}$, $M = 256$, $k = 3$).

We use a uniform mesh in space and in time (N cells in time and M cells in space). Four levels of refinement are considered in space ($M \in \{16, 32, 64, 128\}$) and in time ($N \in \{10, 20, 40, 80\}$). We run two convergence studies. At first, for a good precision in time ($N = 128$, $\ell = 3$), we study the error for several successive space mesh sizes $M \in \{16, 32, 64, 128\}$ and several polynomial orders $k \in \{1, 2, 3\}$. Then, for a good precision in space ($M = 256$, $k = 3$), we study the error for several successive time mesh sizes $N \in \{10, 20, 40, 80\}$ and several polynomial orders $\ell \in \{0, 1, 2, 3\}$. The results without noise are reported in Figure 2. We observe optimal space convergence at rate k for $k \in \{2, 3\}$. Instead, we obtain superconvergence at rate 2 for $k = 1$. Moreover, we also observe the expected time convergence at rate ℓ, for $\ell \in \{1, 2, 3\}$. (Recall that the errors are computed as the difference between the numerical solution and a projection of the exact solution.) Another relevant observation is that a higher-order scheme is more efficient in terms of degrees of freedom.

Notice that the numerical error has also been computed using the $L^2(Q)$-norm, the $L^2(H^1)$-seminorm and the $H^1(\Omega)$-seminorm at initial time (using the associated face element). We observed for these errors the same convergence rate as the one for the $L^\infty(0, T_f; L^2(\Omega))$-norm used above. For the sake of clarity, these errors have not been drawn on Figure 2.

We report in Table 1 the number of degrees of freedom in the two situations of whether or not the static condensation procedure is used (see Section 2.4). We can observe a drastic reduction in the number of degrees of freedom offered by the static condensation procedure.

Without static condensation					With static condensation			
M	$k=1$	$k=2$	$k=3$		M	$k=1$	$k=2$	$k=3$
16	56320	76800	97280		16	23552	27648	31744
32	113664	154624	195584		32	48128	56320	64512
64	228352	310272	392192		64	97280	113664	130048
128	457728	621568	785408		128	195584	228352	261120

Table 1: Number of degrees of freedom ($1d$ test case) for space refinement with $\ell = 3$ and $N = 128$. Left: without static condensation. Right: with static condensation.

4.2 Two-dimensional test case

In this section, we consider the following two-dimensional setting:

$$\Omega := (0,1)^2, \quad \varpi := \Big((b,1)\times(0,1)\Big) \setminus \left(0,\frac{7}{8}\right) \times \left(\frac{1}{8},\frac{7}{8}\right), \quad T_f := 2,$$
$$u(t,x,y) := \cos(\sqrt{2}\pi\omega t)\sin(\pi\omega x)\sin(\pi\omega y),$$

where $b \in [0, \frac{7}{8})$ is a parameter that may vary and will be used to consider several geometric configurations, and $\omega := 3$.

One can verify that with $T_f := 2$ and $b := 0$, the geometry fulfills the geometric control condition (Assumption 1). However, this assumption is not fulfilled for $b > 0$. This latter configuration will enable us to evaluate the impact of this assumption on the approximation results.

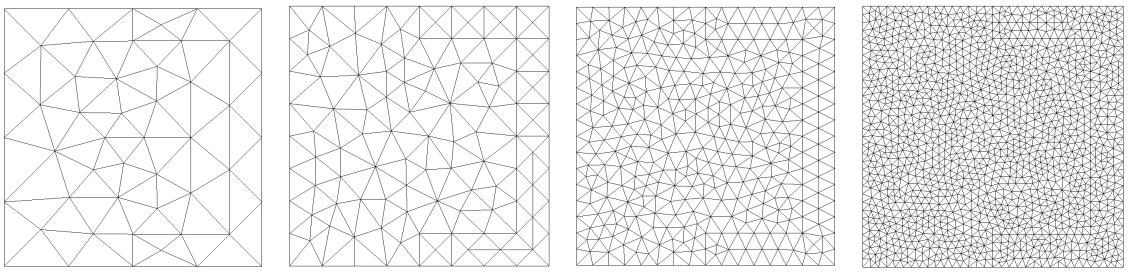

Figure 3: First, third, fifth, and seventh triangulations used for the 2d test case ($b = \frac{1}{2}$); all the triangulations are fitted to ϖ.

The errors in the $L^\infty(L^2)$-norm are presented in Figure 4 in the absence of noise on the measurements. On the left panel, we provide a convergence study with respect to the space discretization. Seven successive meshes are considered (see Figure 3) and the space polynomial degrees are $k \in \{1,2,3\}$. For this test, the time discretization remains constant ($\ell = 3$, $N = 40$). Keeping in mind that the x-axis of this figure is the number of degrees of freedom (that scales roughly as h^{-2} when we refine the space discretization), we retrieve the expected space convergence rates for the various values of k. We also observe that the high-order approach is much more efficient than the low-order one, which is a known behavior for the wave equation.

On the right panel, we provide a study of the time convergence for $\ell \in \{1,2,3\}$ and for $N \in \{5,10,15,20,30,40\}$. As stated before for the space convergence study, we observe here that the high-order approach is much more efficient. Surprisingly, the low-order approach ($k = 1$ or $\ell = 1$) does not seem to converge for the considered meshes.

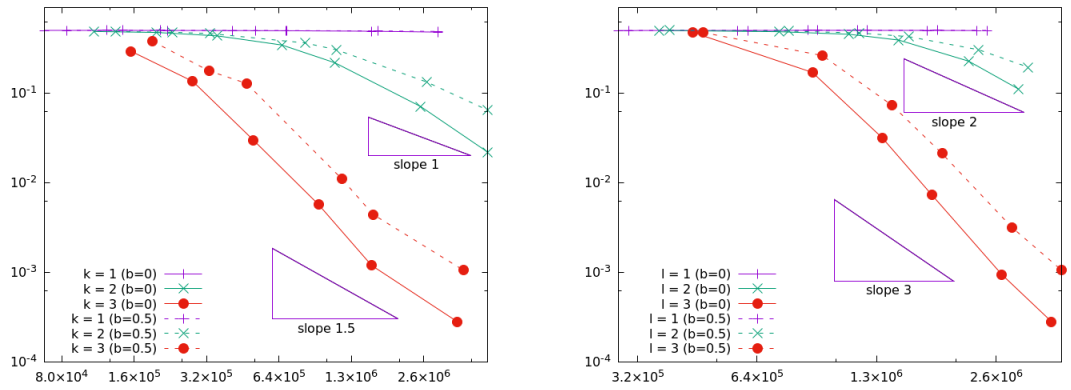

Figure 4: $L^\infty(L^2)$-norm errors for the 2d test case without noise (with respect to the number of degrees of freedom). Left: space convergence. Right: time convergence.

In both of these tests, we observe that the configuration $b = \frac{1}{2}$ is less favorable than the configuration with $b = 0$. However, convergence at optimal rate is still observed in the configuration $b = \frac{1}{2}$. The only visible drawback of that configuration is that the scaling of the errors, but not their decay rate, seems to be a bit worse. We also considered adding some noise to the right-hand side f, but this resulted only in a slight degradation of the solution. The results are not reported for the sake of conciseness.

The same setting is then considered with the presence of noise ($a = 10^{-2}$). The results are reported in Figure 5. We observe that the main consequence of noise is that convergence can be obstructed since the error stagnates at a certain level which depends on the geometry ($b = 0$ or $b = \frac{1}{2}$).

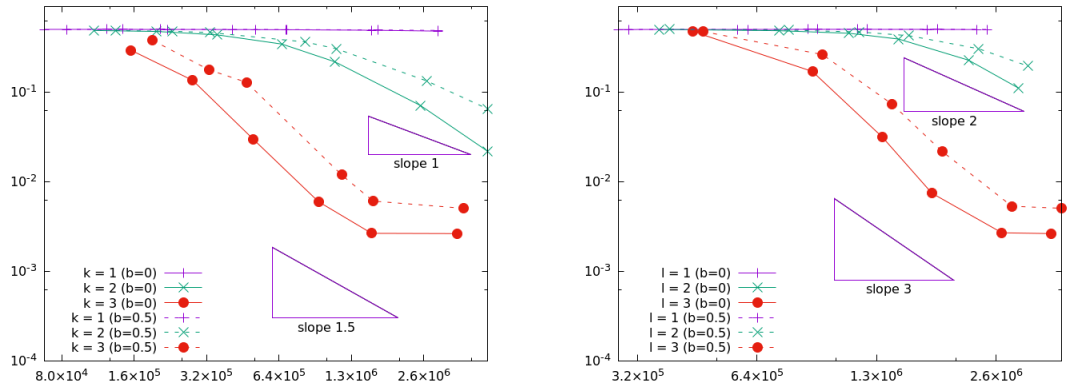

Figure 5: $L^\infty(L^2)$-norm errors for the 2d test case with noise $a = 10^{-2}$ (with respect to the number of degrees of freedom). Left: space convergence. Right: time convergence.

Computational time and memory usage for the configuration $b = \frac{1}{2}$ without noise are reported in Table 2. We observe that, since a direct solver is used, we attain the limits of the method in terms of memory usage. More refined meshes or 3d-space configurations need the development of an iterative solver.

Acknowledgement: GD was supported by the Emergence grant of Sorbonne University.

Computational time (in sec)					Memory usage (in GB)			
mesh	$k=1$	$k=2$	$k=3$		mesh	$k=1$	$k=2$	$k=3$
1	8.6	39	100		1	10^{-3}	4.1	8.8
3	38	170	440		3	4.3	11.6	23.7
5	610	2500	6300		5	20.9	54.5	107
7	13000	54000	–		7	135	345	–

Table 2: $2d$ test case for space refinement with $\ell = 3$ and $N = 40$ ($b = \frac{1}{2}$) and without noise. Left: computational time (in sec). Right: memory usage tip (in GB).

EB was supported by EPSRC grants EP/T033126/1 and EP/V050400/1.

Appendix A Technical proof

Here we prove Lemma 1. The proof is obtained by combining Proposition A.1 and Theorem A.4 from [10] where Ω is assumed to be smooth. Proposition A.1 yields that for $v = 0$ on $J \times \partial\Omega$ we have

$$\begin{aligned}\|v\|_{L^\infty(J;L^2(\Omega))} + \|\partial_t v\|_{L^2(J;H^{-1}(\Omega))} \\ \lesssim \|v|_{t=0}\|_{L^2(\Omega)} + \|\partial_t v|_{t=0}\|_{H^{-1}(\Omega)} + \|L(v)\|_{H^{-1}(J\times\Omega)}.\end{aligned}$$

Moreover, following the proof of Theorem A.4 from [10], we can show that under the geometric control condition (Assumption 1), we have

$$\|v|_{t=0}\|_{L^2(\Omega)} + \|\partial_t v|_{t=0}\|_{H^{-1}(\Omega)} \lesssim \|L(v)\|_{H^{-1}(J\times\Omega)} + \|v\|_{L^2(J\times\varpi)}.$$

Combining these two relations, we have

$$\|v\|_{L^\infty(J;L^2(\Omega))} + \|\partial_t v\|_{L^2(J;H^{-1}(\Omega))} \lesssim \|v\|_{L^2(J\times\varpi)} + \|L(v)\|_{H^{-1}(Q)}.$$

To estimate the missing term, we consider

$$\|\nabla v\|_{H^{-1}(J;L^2(\Omega))} \lesssim \|\Delta v\|_{H^{-1}(Q)} \leq \|L(v)\|_{H^{-1}(Q)} + \|\partial_{tt} v\|_{H^{-1}(Q)},$$

where, in the first bound, we use $\|\nabla\phi\|_\Omega \lesssim \|\Delta\phi\|_{H^{-1}(\Omega)}$ for all $\phi \in H_0^1(\Omega)$. Moreover, for $\phi \in H_0^1(Q)$ we have

$$\begin{aligned}< \partial_{tt} v, \phi >_{H^{-1}(Q),H_0^1(Q)} &= - < \partial_t v, \partial_t \phi >_{L^2(J;H^{-1}(\Omega)),L^2(J;H_0^1(\Omega))} \\ &\leq \|\partial_t v\|_{L^2(J;H^{-1}(\Omega))}\|\partial_t\phi\|_{L^2(J;H_0^1(\Omega))}.\end{aligned}$$

So that $\|\partial_{tt} v\|_{H^{-1}(Q)} \leq \|\partial_t v\|_{L^2(J;H^{-1}(\Omega))}$. This ends the proof.

References

[1] 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.*, vol. 30, no. 5, pp. 1024–1065, 1992.

[2] 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*, vol. 10, no. 4, pp. 983–1015, 2017.

[3] N. Burq and P. Gérard, "Condition nécessaire et suffisante pour la contrôlabilité exacte des ondes," *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 325, no. 7, pp. 749–752, 1997.

[4] E. Bécache, L. Bourgeois, L. Franceschini, and J. Dardé, "Application of mixed formulations of quasi-reversibility to solve ill-posed problems for heat and wave equations: the 1D case," *Inverse Probl. Imaging*, vol. 9, no. 4, pp. 971–1002, 2015.

[5] N. Cîndea and A. Münch, "Inverse problems for linear hyperbolic equations using mixed formulations," *Inverse Problems*, vol. 31, 2015.

[6] L. Baudouin, M. de Buhan, S. Ervedoza, and A. Osses, "Carleman-based reconstruction algorithm for waves," *SIAM J. Numer. Anal.*, vol. 59, no. 2, pp. 998–1039, 2021.

[7] W. Dahmen, H. Monsuur, and R. Stevenson, "Least squares solvers for ill-posed PDEs that are conditionally stable," *ESAIM Math. Model. Numer. Anal.*, vol. 57, no. 4, pp. 2227–2255, 2023.

[8] 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*, vol. 52, no. 3, pp. 245–288, 2015.

[9] E. Burman, A. Feizmohammadi, and L. Oksanen, "A fully discrete numerical control method for the wave equation," *SIAM J. Control Optim.*, vol. 58, no. 3, pp. 1519–1546, 2020.

[10] E. Burman, A. Feizmohammadi, A. Münch, and L. Oksanen, "Spacetime finite element methods for control problems subject to the wave equation," *ESAIM Control Optim. Calc. Var.*, vol. 29, pp. Paper No. 41, 40, 2023.

[11] S. Ervedoza and E. Zuazua, *Numerical approximation of exact controls for waves*. SpringerBriefs in Mathematics, Springer, New York, 2013.

[12] E. Burman, "Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: Elliptic equations," *SIAM J. Sci. Comput.*, vol. 35, no. 6, pp. A2752–A2780, 2013.

[13] E. Burman, "Error estimates for stabilized finite element methods applied to ill-posed problems," *C. R. Math. Acad. Sci. Paris*, vol. 352, no. 7-8, pp. 655–659, 2014.

[14] E. Burman and P. Hansbo, "Stabilized nonconforming finite element methods for data assimilation in incompressible flows," *Math. Comp.*, vol. 87, no. 311, pp. 1029–1050, 2018.

[15] E. Burman, M. Nechita, and L. Oksanen, "Unique continuation for the Helmholtz equation using stabilized finite element methods," *J. Math. Pures Appl. (9)*, vol. 129, pp. 1–22, 2019.

[16] E. Burman, G. Delay, and A. Ern, "A hybridized high-order method for unique continuation subject to the Helmholtz equation," *SIAM J. Numer. Anal.*, vol. 59, no. 5, pp. 2368–2392, 2021.

[17] E. Burman and L. Oksanen, "Data assimilation for the heat equation using stabilized finite element methods," *Numer. Math.*, vol. 139, no. 3, pp. 505–528, 2018.

[18] E. Burman, J. Ish-Horowicz, and L. Oksanen, "Fully discrete finite element data assimilation method for the heat equation," *ESAIM Math. Model. Numer. Anal.*, vol. 52, no. 5, pp. 2065–2082, 2018.

[19] E. Burman, G. Delay, and A. Ern, "The unique continuation problem for the heat equation discretized with a high-order space-time nonconforming method," *SIAM J. Numer. Anal.*, vol. 61, no. 5, pp. 2534–2557, 2023.

[20] E. Burman, A. Feizmohammadi, and L. Oksanen, "A finite element data assimilation method for the wave equation," *Math. Comp.*, vol. 89, no. 324, pp. 1681–1709, 2020.

[21] E. Burman, A. Feizmohammadi, A. Münch, and L. Oksanen, "Space time stabilized finite element methods for a unique continuation problem subject to the wave equation," *ESAIM Math. Model. Numer. Anal.*, vol. 55, no. suppl., pp. S969–S991, 2021.

[22] E. Burman and J. Preuss, "Unique continuation for the wave equation based on a discontinuous Galerkin time discretization," *preprint*, 2024. submitted.

[23] E. Burman, M. Cicuttin, G. Delay, and A. Ern, "An unfitted hybrid high-order method with cell agglomeration for elliptic interface problems," *SIAM J. Sci. Comput.*, vol. 43, no. 2, pp. A859–A882, 2021.

[24] O. A. Karakashian and F. Pascal, "A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems," *SIAM J. Numer. Anal.*, vol. 41, no. 6, pp. 2374–2399, 2003.

[25] E. Burman and A. Ern, "Continuous interior penalty *hp*-finite element methods for advection and advection-diffusion equations," *Math. Comp.*, vol. 76, no. 259, pp. 1119–1140, 2007.

[26] A. Ern and J.-L. Guermond, "Finite element quasi-interpolation and best approximation," *ESAIM Math. Model. Numer. Anal.*, vol. 51, no. 4, pp. 1367–1385, 2017.

[27] P. Houston, D. Schötzau, and T. P. Wihler, "Energy norm a posteriori error estimation of *hp*-adaptive discontinuous Galerkin methods for elliptic problems," *Math. Models Methods Appl. Sci.*, vol. 17, no. 1, pp. 33–62, 2007.

[28] M. Cicuttin, D. A. Di Pietro, and A. Ern, "Implementation of Discontinuous Skeletal methods on arbitrary-dimensional, polytopal meshes using generic programming," *J. Comput. Appl. Math.*, 2017.