

Combined DG–CG finite element method for the Westervelt equation

SERGIO GÓMEZ

Department of Mathematics and Applications, University of Milano-Bicocca, Via Cozzi 55,
20125 Milan, Italy
IMATI-CNR ‘E. Magenes’, Via Ferrata 5, 27100 Pavia, Italy

AND

VANJA NIKOLIĆ*

Department of Mathematics, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen,
The Netherlands

*Corresponding author: vanja.nikolic@ru.nl

[Received on 12 December 2024; revised on 18 June 2025]

We propose and analyze a space–time finite element method for Westervelt’s quasilinear model of ultrasound waves in its second-order formulation. The method combines conforming finite element spatial discretizations with a discontinuous–continuous Galerkin time stepping. Its analysis is challenged by the fact that standard Galerkin testing approaches for wave problems do not allow for bounding the discrete energy at all times. By means of redesigned energy arguments for a linearized problem combined with Banach’s fixed-point argument, we show the well-posedness of the scheme, *a priori* error estimates, and robustness with respect to the strong damping parameter δ . Moreover, the scheme preserves the asymptotic preserving property of the continuous problem; more precisely, we prove that the discrete solutions corresponding to $\delta > 0$ converge, in the singular vanishing dissipation limit, to the solution of the discrete inviscid problem. We use several numerical experiments in $(2 + 1)$ dimensions to validate our theoretical results.

Keywords: Westervelt’s equation; space–time method; discontinuous–continuous Galerkin time stepping; asymptotic-preserving method.

1. Introduction

Westervelt’s equation (Westervelt, 1963) is a classical model of nonlinear acoustic and ultrasonic wave propagation. It is especially attractive for simulating medical and industrial applications of ultrasound as it does not assume any directionality of the wave field. In this work we develop and analyze a space–time finite element approach for its discretization in the framework of the DG–CG scheme introduced in Walkington (2014) that combines continuous and discontinuous Galerkin time-stepping methodologies.

We consider a space–time cylinder $Q_T = \Omega \times (0, T)$, where $\Omega \subset \mathbb{R}^d$ ($d \in \{1, 2, 3\}$) is an open, bounded polytopic domain with Lipschitz boundary $\partial\Omega$, and $T > 0$ is the final propagation time. We define the surfaces $\Sigma_0 := \Omega \times \{0\}$ and $\Sigma_T := \Omega \times \{T\}$. Given a medium-dependent nonlinearity parameter $k \in \mathbb{R}$, a sound diffusivity coefficient $\delta \geq 0$, the speed of sound $c > 0$, a source term

$f : Q_T \rightarrow \mathbb{R}$, and initial data $u_0, u_1 : \Omega \rightarrow \mathbb{R}$, we consider the following Westervelt IBVP: find the acoustic pressure $u : Q_T \rightarrow \mathbb{R}$ such that

$$\begin{cases} \partial_t((1 + ku)\partial_t u) - c^2 \Delta u - \delta \Delta(\partial_t u) = f & \text{in } Q_T, \\ u = 0 & \text{on } \partial\Omega \times (0, T), \\ u = u_0 \quad \text{and} \quad \partial_t u = u_1 & \text{on } \Sigma_0. \end{cases} \quad (1.1)$$

It is known that the damping parameter plays a decisive role in the behavior of solutions to (1.1). When $\delta > 0$, Westervelt's equation has a parabolic character due to the strong damping $-\delta \Delta \partial_t u$, which unlocks the global well-posedness of (1.1) with an exponential decay of the energy of the system, as shown in [Kaltenbacher & Lasiecka \(2009\)](#). In the inviscid hyperbolic case where $\delta = 0$, on the other hand, one can expect only the local existence of solutions; see [Dörfler *et al.* \(2016\)](#). Given that the damping parameter is often small in practical scenarios, it becomes vital to understand how the solutions behave as δ approaches zero. The question is complex due to the quasilinear nature of (1.1). Such a singular limit analysis for (1.1) can be found in [Kaltenbacher & Nikolić \(2022\)](#). More recently, robustness with respect to the sound diffusivity has also become a prominent issue when considering numerical schemes for (1.1) and other nonlinear acoustic models; see [Dörich & Nikolić \(2024\)](#); [Meliani & Nikolić \(2024\)](#); [Gómez & Meliani \(2025\)](#) and [Nikolić \(2025\)](#). These works have built upon the insights from rigorous numerical investigations of the inviscid Westervelt equation in, e.g., [Hochbruck & Maier \(2022\)](#); [Maier \(2023\)](#) and [Dörich \(2025\)](#), and the strongly damped equation in, e.g., [Nikolić & Wohlmuth \(2019\)](#). The second major goal of this work is to shed further light on this problem by establishing the conditions under which the developed DG–CG method is robust with respect to δ and furthermore preserves the limiting behavior of the exact solutions in the vanishing dissipation limit.

We are particularly interested in fully discrete schemes based on Galerkin-type time discretizations, as they can be easily interpreted as space–time methods. Some of the appealing properties of space–time methods are that

- (i) they allow for simultaneous high-order convergence in space and time,
- (ii) they can be analyzed in a discrete variational setting closer to the continuous one and
- (iii) the error analysis does not rely on Taylor expansions, so they typically require weaker regularity assumptions than most standard time-stepping schemes.

Galerkin-type time stepping schemes for the wave equation can be classified in two groups: those based on a first-order-in-time system (i.e., introducing an auxiliary variable $v = \partial_t u$) and those defined directly for the second-order formulation. The first group includes the discontinuous Galerkin method by [Johnson \(1993\)](#) and the continuous Petrov–Galerkin method by [French & Peterson \(1996\)](#) (see also its more robust analysis in [Bales & Lasiecka, 1994, 1995](#); [Gómez, 2025b](#)). Both methods are unconditionally stable without requiring the presence of a damping term; however, at the expense of doubling the number of degrees of freedom. The DG time discretizations by [Hulbert & Hughes \(1990\)](#) and by [French \(1993\)](#) belong to the second group. In order to obtain well-posedness, the former uses a least-squares stability term, whereas the latter requires employing explicit exponential weights in the inner products. [Antonietti *et al.* \(2018\)](#) proposed and analyzed a DG time discretization for the second-order formulation relying on the presence of a first-order damping term. Recently, in [Shao \(2022\)](#), the DG method from [Antonietti *et al.* \(2018\)](#) was applied to a class of second-order nonlinear hyperbolic equations by transforming the

original problem in terms of a new variable $w = e^{-\gamma t}u$, so as to obtain the necessary damping term in the transformed problem.

Our method is based on the combined DG–CG scheme developed for the linear wave equation by Walkington (2014), which provides a discrete solution that is continuous in time, but with a possibly discontinuous first-order time derivative. The method in Walkington (2014) is applied directly to the second-order wave formulation and does not require the presence of a damping term to guarantee well-posedness. Moreover, unconditional stability and optimal convergence can be shown in $L^\infty(0, T; X)$ -type norms. Consequently, the scheme is endowed with good stability and approximation properties, and requires fewer degrees of freedom than the methods based on the Hamiltonian formulation (Johnson, 1993; French & Peterson, 1996). An additional important advantage for linear wave problems is that this method does not impose a Courant–Friedrichs–Lewy (CFL)-type restriction on the time step. Theorem 4.1 below shows that this advantage extends to damped linear problems with a variable leading coefficient.

1.1 Main contributions

Our first main contribution pertains to advancing the theoretical framework of the DG–CG finite element methodology from Walkington (2014) to accommodate quasilinear wave models in the form of (1.1). In addition, to the best of our knowledge, this work contains the first rigorous space–time finite element analysis of the Westervelt equation. Specifically, we determine the sufficient conditions for the discretization and exact solution u that ensure that the discrete solution $u_{h,\tau}$ satisfies the following *a priori* bounds:

$$\begin{aligned} \|\partial_t(u - u_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} &\lesssim \tau^m \|u\|_m + h^{\ell+1} \|u\|_\ell, \\ \|\nabla(u - u_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega)^d)} &\lesssim \tau^m \|u\|_{m,\tau} + h^\ell \|u\|_{\ell,h}, \end{aligned} \tag{1.2}$$

provided $h^{\ell+1-\frac{d}{2}} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$, $2 \leq m \leq q$, $1 \leq \ell \leq p$ and $\ell + 1 - d/2 \geq d/(2(m - 1))$, where p and q denote the degrees of approximation in space and time, respectively. The norms $\|\cdot\|_m$, $\|\cdot\|_\ell$, $\|\cdot\|_{m,\ell}$ and $\|\cdot\|_{\ell,h}$ on the right-hand side of (1.2) are norms in Bochner spaces, defined in (5.1) below. We refer to Theorem 5.4 for the details.

Secondly, our analysis guarantees that the constants in estimates (1.2) do not depend on the sound diffusivity δ . This fact paves the way for our second main theoretical contribution, which quantifies the δ -asymptotic behavior of discrete solutions of dissipative problems as $\delta \searrow 0$. More precisely, Theorem 6.1 reveals that, for a properly set up discretization, the discrete solutions preserve the asymptotic behavior of the exact ones in the vanishing dissipation limit, as they converge to the discrete solution to the inviscid problem at a linear rate.

1.2 Approach and challenges

We approach the analysis of the DG–CG method for (1.1) by first considering a linearized discrete problem with a variable coefficient. After deriving *a priori* bounds for the linearization, we combine them with a fixed-point argument set up on a ball centered at u with an (h, τ) -dependent radius, in the spirit of, e.g., Makridakis (1993); Ortner & Süli (2007) and Nikolić & Wohlmuth (2019). This approach yields the well-posedness of the discrete nonlinear problem and error bounds (1.2) simultaneously.

However, the reasoning used in the robustness analysis of the semidiscrete-in-space finite element formulation of the Westervelt equation (see, for example, Nikolić, 2025) cannot be applied to analyze the

fully discrete scheme in this case. The primary issue is that standard testing techniques would provide bounds for the energy norm of the discrete solution to the linearized problem only at discrete times $\{t_n\}_{n=1}^N$ (see also Proposition 3.5 below), which is not enough to bound the discrete energy at all times. Thus, a big challenge in the *a priori* analysis here lies in developing the right testing strategy for the linearized problem that leads to the bounds in the form of (1.2) and that is furthermore suitable for later closing the fixed-point argument.

The numerical analysis of the Westervelt equation must guarantee that the term $1 + ku_{h,\tau}$ does not degenerate, so that the discrete problem is still a valid wave model. This requirement translates to needing a uniformly small bound on $\|ku_{h,\tau}\|_{L^\infty(0,T;L^\infty(\Omega))}$. Another special feature of the Westervelt equation is that the second time derivative is involved in the nonlinearity. To prove the uniqueness of solutions to the discrete problem, we thus also need a suitable uniform bound on the discrete second time derivative, which is a nontrivial task in a time-discrete setting. The particular nonlinearities present in the model thus have significant implications for designing the testing strategy in the stability and error analysis, and for the assumptions made on the discretization parameters and polynomial degrees.

1.3 Notation

We use standard notation for L^p , Sobolev and Bochner spaces. For instance, given an open, bounded domain $\mathcal{D} \subset \mathbb{R}^d$ ($d \in \{1, 2, 3\}$) with Lipschitz boundary $\partial\mathcal{D}$, $p \in [1, \infty]$ and $m \in \mathbb{R}^+$, we denote the corresponding Sobolev space by $W_p^m(\mathcal{D})$, and its seminorm and norm, respectively, by $|\cdot|_{W_p^m(\mathcal{D})}$ and $\|\cdot\|_{W_p^m(\mathcal{D})}$. In particular, for $p = 2$, we use the notation $H^s(\mathcal{D}) := W_2^s(\mathcal{D})$, endowed with the seminorm $|\cdot|_{H^s(\mathcal{D})}$ and the norm $\|\cdot\|_{H^s(\mathcal{D})}$. For $s = 0$, $H^0(\mathcal{D}) := L^2(\mathcal{D})$ is the space of Lebesgue square integrable functions over \mathcal{D} with inner product $(\cdot, \cdot)_{\mathcal{D}}$, and $H_0^1(\mathcal{D})$ is the closure of $C_0^\infty(\mathcal{D})$ in the $H^1(\mathcal{D})$ norm. Given a time interval (a, b) , and a Banach space $(X, \|\cdot\|_X)$, the corresponding Bochner space is denoted by $W_p^s(a, b; X)$.

Given $q \in \mathbb{N}$, we denote by $\mathbb{P}^q(\mathcal{D})$ the space of polynomials defined on \mathcal{D} of degree at most q .

1.4 Structure of the remaining of the paper

In Section 2, we provide a precise description of the DG–CG method for the Westervelt equation. Section 3 contains the stability analysis of a linear problem with a variable coefficient, which is complemented in Section 4 by the *a priori* error analysis of that problem in the energy norm. The results for the linearized problem are then combined with Banach’s fixed-point theorem to arrive at the well-posedness and error bounds for the Westervelt equation in Section 5. In Section 6, we investigate the vanishing δ dynamics of the discrete problem and establish the convergence rates with respect to this parameter. Finally, our theoretical findings are illustrated in Section 7 with several numerical experiments, and we include some concluding remarks in Section 8.

2. Description of the method

Let $\{\mathcal{T}_h\}_{h>0}$ be a family of shape-regular conforming simplicial meshes for the spatial domain Ω . Let also $\{\mathcal{T}_\tau\}_\tau$ be a family of (not necessarily uniform) partitions of the time interval $(0, T)$, where τ denotes the maximum time step and each \mathcal{T}_τ is of the form $0 =: t_0 < t_1 < \dots < t_N := T$.

For each element $K \in \mathcal{T}_h$, we denote the diameter of K by h_K . Moreover, for $n = 1, \dots, N$, we define the time interval $I_n := (t_{n-1}, t_n) \in \mathcal{T}_\tau$, the time step $\tau_n := t_n - t_{n-1}$, the partial cylinder $Q_n := \Omega \times I_n$, and the surface $\Sigma_n := \Omega \times \{t_n\}$.

In addition, for $n = 1, \dots, N-1$ and any piecewise smooth function v , we define the time jump¹

$$\llbracket v \rrbracket_n(\mathbf{x}) := v(\mathbf{x}, t_n^-) - v(\mathbf{x}, t_n^+) \quad \forall \mathbf{x} \in \Omega, \quad (2.1)$$

where

$$v(\mathbf{x}, t_n^-) := \lim_{\varepsilon \rightarrow 0^+} v(\mathbf{x}, t_n - \varepsilon) \quad \text{and} \quad v(\mathbf{x}, t_n^+) := \lim_{\varepsilon \rightarrow 0^+} v(\mathbf{x}, t_n + \varepsilon).$$

Let $p, q \in \mathbb{Z}^+$ be some given degrees of approximation in space and time, respectively. We define the following piecewise polynomial spaces:

$$\begin{aligned} \mathcal{V}_h^p &:= \{v \in H_0^1(\Omega) : v|_K \in \mathbb{P}^p(K) \text{ for all } K \in \mathcal{T}_h\}, \\ \mathcal{V}_\tau^q &:= \{v \in C^0[0, T] : v|_{I_n} \in \mathbb{P}^q(I_n) \text{ for } n = 1, \dots, N\}, \\ \mathcal{V}_{h,\tau}^{p,q} &:= \{v \in C^0([0, T]; H_0^1(\Omega)) : v|_{Q_n} \in \mathbb{P}^q(I_n; \mathcal{V}_h^p) \text{ for all } I_n \in \mathcal{T}_\tau\}, \\ \mathcal{W}_{h,\tau}^{p,q-1} &:= \{w \in L^2(0, T; H_0^1(\Omega)) : w|_{Q_n} \in \mathbb{P}^{q-1}(I_n; \mathcal{V}_h^p) \text{ for all } I_n \in \mathcal{T}_\tau\}. \end{aligned} \quad (2.2)$$

Moreover, we denote by $\mathbb{P}^q(\mathcal{T}_\tau)$ the space of piecewise polynomials of degree at most q on each $I_n \in \mathcal{T}_\tau$.

In the spirit of the DG-CG time stepping scheme in [Walkington \(2014\)](#), the proposed space-time formulation reads: find $u_{h,\tau} \in \mathcal{V}_{h,\tau}^{p,q}$ such that

$$\mathcal{B}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}, z_h)) = \mathcal{L}(w_{h,\tau}, z_h) \quad \forall (w_{h,\tau}, z_h) \in \mathcal{W}_{h,\tau}^{p,q-1} \times \mathcal{V}_h^p, \quad (2.3)$$

where

$$\begin{aligned} \mathcal{B}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}, z_h)) &:= \sum_{n=1}^N (\partial_t((1 + ku_{h,\tau})\partial_t u_{h,\tau}), w_{h,\tau})_{Q_n} \\ &\quad - \sum_{n=1}^{N-1} ((1 + ku_{h,\tau}(\cdot, t_n)) \llbracket \partial_t u_{h,\tau} \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_\Omega + ((1 + ku_{h,\tau})\partial_t u_{h,\tau}, w_{h,\tau})_{\Sigma_0} \\ &\quad + c^2 (\nabla u_{h,\tau}, \nabla w_{h,\tau})_{Q_T} + \delta (\nabla \partial_t u_{h,\tau}, \nabla w_{h,\tau})_{Q_T} + (\nabla u_{h,\tau}(\cdot, 0), \nabla z_h)_\Omega, \\ \mathcal{L}(w_{h,\tau}, z_h) &:= (f, w_{h,\tau})_{Q_T} + ((1 + ku_0)u_1, w_{h,\tau}(\cdot, 0))_\Omega + (\nabla u_0, \nabla z_h)_\Omega. \end{aligned}$$

Since we are interested in the vanishing dynamics, we assume that $\delta \in [0, \bar{\delta}]$ for some fixed $\bar{\delta} > 0$.

¹ The time jumps are sometimes defined as $\llbracket v \rrbracket_n(\mathbf{x}) := v(\mathbf{x}, t_n^+) - v(\mathbf{x}, t_n^-)$. Instead, in the context of space-time methods, the definition in (2.1) corresponds to setting the normal vector on the so-called space-like facets pointing to the future.

2.1 Energy norm

We introduce the upwind-jump functional

$$|v|_{\mathbb{J}}^2 := \frac{1}{2} \left(\|v\|_{L^2(\Sigma_T)}^2 + \sum_{n=1}^{N-1} \left(\llbracket [v] \rrbracket_n \right\|_{L^2(\Omega)}^2 + \|v\|_{L^2(\Sigma_0)}^2 \right), \quad (2.4)$$

and the following energy norm in the space $\mathcal{V}_{h,\tau}^{p,q}$:

$$\begin{aligned} \|u_{h,\tau}\|_{\delta}^2 &:= \|\partial_t u_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))}^2 + c^2 \|\nabla u_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega)^d)}^2 + |\partial_t u_{h,\tau}|_{\mathbb{J}}^2 \\ &\quad + c^2 \|\nabla u_{h,\tau}\|_{L^2(\Sigma_T)^d}^2 + c^2 \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 + \delta \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_T)^d}^2. \end{aligned} \quad (2.5)$$

REMARK 2.1 (Linear case). By setting $\delta = 0$ and $k = 0$ in (2.3), the proposed method reduces to the DG–CG scheme in [Walkington \(2014\)](#) for the linear acoustic wave equation without damping, whereas, for $k = 0$ and $\delta > 0$, it corresponds to a space–time discretization of a strongly damped wave equation. ■

REMARK 2.2 (Discrete initial conditions). In the space–time variational formulation (2.3), the discrete initial condition for $u_{h,\tau}(\cdot, 0)$ is strongly enforced as the Ritz projection of u_0 , whereas the discrete initial condition for $\partial_t u_{h,\tau}(\cdot, 0)$ is imposed in a weak sense. ■

As announced, we approach the analysis of (2.3) by first considering a linearized problem with a variable coefficient, where we replace $1 + ku_{h,\tau}$ in the leading term by $1 + k\alpha_{h,\tau}$ for a suitably chosen $\alpha_{h,\tau}$. The estimates will later be combined with Banach’s fixed-point theorem to obtain the results for (2.3), uniformly in δ . Thus, $\alpha_{h,\tau}$ can be understood in this work as a placeholder for a fixed-point iterate.

3. Stability analysis of a linearized problem

In this section we analyze the stability of the following linearized space–time formulation: given a discrete approximation $\alpha_{h,\tau} \in \mathcal{V}_{h,\tau}^{p,q}$ of u , find $u_{h,\tau} \in \mathcal{V}_{h,\tau}^{p,q}$ such that

$$\tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}, z_h)) = \mathcal{L}_\delta(w_{h,\tau}, z_h) \quad \forall w_{h,\tau} \in \mathcal{W}_{h,\tau}^{p,q-1}, \forall z_h \in \mathcal{V}_h^p, \quad (3.1)$$

where

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}, z_h)) &:= \sum_{n=1}^N (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), w_{h,\tau})_{Q_n} \\ &\quad - \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n)) \llbracket \partial_t u_{h,\tau} \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_{\Omega} + ((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}, w_{h,\tau})_{\Sigma_0} \\ &\quad + c^2 (\nabla u_{h,\tau}, \nabla w_{h,\tau})_{Q_T} + \delta (\nabla \partial_t u_{h,\tau}, \nabla w_{h,\tau})_{Q_T} + (\nabla u_{h,\tau}(\cdot, 0), \nabla z_h)_{\Omega}, \end{aligned} \quad (3.2)$$

$$\begin{aligned} \mathcal{L}_\delta(w_{h,\tau}, z_h) &:= (f, w_{h,\tau})_{Q_T} + ((1 + ku_0)u_1, w_{h,\tau}(\cdot, 0))_\Omega \\ &\quad + \delta(\nabla\partial_t\mu, \nabla w_{h,\tau})_{Q_T} + \sum_{n=1}^N (\xi, \partial_t w_{h,\tau})_{Q_n} + (\nabla u_0, \nabla z_h)_\Omega, \end{aligned} \quad (3.3)$$

with $\mu \in H^1(0, T; H_0^1(\Omega)^d)$ being a perturbation function that will represent the interpolation error of the damping term, and $\xi \in L^2(Q_T)$ being a function representing the lack of orthogonality of the projection \mathcal{P}_τ used in the *a priori* error analysis in Section 4.3 below due to the presence of the discrete coefficient $\alpha_{h,\tau}$.

Henceforth, we denote by Id the identity operator, and by Π_{q-1}^t the $L^2(0, T)$ orthogonal projection onto $\mathbb{P}^q(\mathcal{T}_\tau)$. In Appendix A, we recall some standard results concerning stability and approximation properties of Π_{q-1}^t , polynomial inverse estimates, and standard trace inequalities that we heavily rely on in the analysis below.

Summary of results in this section We assume that the linearized problem does not degenerate in the following sense.

ASSUMPTION 1 (Nondegeneracy of the discrete coefficient). Assume that the discrete coefficient $\alpha_{h,\tau} \in \mathcal{V}_{h,\tau}^{p,q}$ does not degenerate, i.e., there exist positive constants $\underline{\alpha}$ and $\bar{\alpha}$ independent of h , τ and δ such that

$$0 < 1 - |k|\underline{\alpha} \leq 1 + k\alpha_{h,\tau}(\mathbf{x}, t) \leq 1 + |k|\bar{\alpha} \quad \forall(\mathbf{x}, t) \in Q_T. \quad (3.4)$$

Later on, in the fixed-point argument, we show that the nondegeneracy assumption is satisfied for $u_{h,\tau}$ as well.

As mentioned, the main difficulty in reproducing the arguments used for the semidiscrete-in-space formulation of the linearized Westervelt equation, e.g., [Nikolić \(2025\)](#) to analyze the linearization of the fully discrete scheme (2.3), is that standard techniques provide bounds for the energy norm of the discrete solution only at the discrete times $\{t_n\}_{n=1}^N$ (see also Proposition 3.5 below), which is not enough to bound the energy for all times. To address this issue, we first extend the ideas from [Walkington \(2014\)](#) to the present setting of the linear problem with a variable coefficient in order to obtain continuous dependence on the data controlling the energy at all times. Thus, the next theorem, which we prove in Section 3.3 below, is the main result of this section.

THEOREM 3.1 (Continuous dependence on the data). Let Assumption 1 on the discrete coefficient $\alpha_{h,\tau}$ hold. There exists a positive constant C_α independent of h , τ and δ such that, if

$$|k|\|\partial_t\alpha_{h,\tau}\|_{L^1(0,T;L^\infty(\Omega))} \leq C_\alpha, \quad (3.5)$$

then the following stability estimate holds:

$$\begin{aligned} \|u_{h,\tau}\|_\delta^2 &\lesssim \|f\|_{L^1(0,T;L^2(\Omega))}^2 + \|(1 + ku_0)u_1\|_{L^2(\Omega)}^2 + c^2\|\nabla u_0\|_{L^2(\Omega)^d}^2 \\ &\quad + \bar{\delta}\|\nabla\partial_t\mu\|_{L^2(Q_T)^d}^2 + \|\bar{\tau}^{-1}\xi\|_{L^1(0,T;L^2(\Omega))}^2. \end{aligned} \quad (3.6)$$

REMARK 3.2 (Existence and uniqueness of a discrete solution to the linearized problem). From (3.6), we can deduce that the unique discrete solution to the linearized space-time formulation (3.1) with

zero data ($f = 0$, $u_0 = 0$, $u_1 = 0$, $\mu = 0$ and $\xi = 0$) is $u_{h,\tau} = 0$. Since (3.1) corresponds to a square system of linear equations, this implies the existence and uniqueness of a discrete solution to (3.1). \blacksquare

3.1 Auxiliary weight function

Following Walkington (2014) (see also Dong *et al.*, 2024 for a q -explicit analysis), we introduce a set of auxiliary linear functions that are used in the proof of Theorem 3.1 to derive stability estimates in the energy norm (2.5). Let θ be a strictly positive parameter that will be chosen later. For $n = 1, \dots, N$, we define the following linear function:

$$\varphi_n(t) := \theta - \lambda_n(t - t_{n-1}) \quad \text{with } \lambda_n := \frac{\zeta_q}{\tau_n} \quad \text{and } \zeta_q := \frac{1}{4(2q+1)} \quad \text{for all } t \in (t_{n-1}, t_n). \quad (3.7)$$

The constant ζ_q depends only on the degree of approximation q . From Lemma A.5, we deduce that

$$\frac{\zeta_q}{(1 + C_{\text{inv}})^2} \|w_\tau\|_{L^\infty(I_n; L^2(\Omega))}^2 \leq \lambda_n \|w_\tau\|_{L^2(Q_n)}^2 \quad \forall w_\tau \in \mathbb{P}^q(I_n; L^2(\Omega)),$$

which will be exploited in Section 3.3 to obtain L^∞ -bounds in time on the approximate solution.

We assume that $\theta > \zeta_q$. Then, the weight function φ_n satisfies the uniform bounds

$$0 < \theta - \zeta_q \leq \varphi_n(t) \leq \theta \quad \text{for all } t \in [t_{n-1}, t_n], \quad (3.8)$$

as well as

$$\varphi_n'(t) = -\lambda_n = -\frac{1}{4\tau_n(2q+1)} \quad \text{for all } t \in [t_{n-1}, t_n].$$

Furthermore, we have the following approximation result, which relies on the properties of Legendre polynomials.

LEMMA 3.3. For $n = 1, \dots, N$ and $K \in \mathcal{T}_h$, the following bounds hold:

$$\|(\text{Id} - \Pi_{q-1}^t)(\varphi_n w_{h,\tau})\|_{L^2(K_n)} \leq \zeta_q \|w_{h,\tau}\|_{L^2(K_n)} \quad \forall w_{h,\tau} \in \mathcal{W}_{h,\tau}^{p,q-1}, \quad (3.9)$$

$$\begin{aligned} \|(\text{Id} - \Pi_{q-1}^t)(\varphi_n w_{h,\tau})(\cdot, t_{n-1}^+)\|_{L^2(\Omega)} &\leq \frac{1}{4\sqrt{2q+1}\sqrt{\tau_n}} \|w_{h,\tau}\|_{L^2(Q_n)} \\ &= \frac{\sqrt{\lambda_n}}{2} \|w_{h,\tau}\|_{L^2(Q_n)} \quad \forall w_{h,\tau} \in \mathcal{W}_{h,\tau}^{p,q-1}. \end{aligned} \quad (3.10)$$

Proof. Let $L_q(t)$ denote the Legendre polynomial of degree q on the interval I_n (the explicit dependence on the interval I_n is omitted). Since $\varphi_n w_{h,\tau} \in \mathcal{W}_{h,\tau}^{p,q}$, the following identity holds for all $(\mathbf{x}, t) \in \mathcal{Q}_n$:

$$\begin{aligned} (\text{Id} - \Pi_{q-1}^t)(\varphi_n w_{h,\tau})(\mathbf{x}, t) &= \frac{(2q+1)L_q(t)}{\tau_n} \int_{I_n} L_q(t)\varphi_n(t)w_{h,\tau}(\mathbf{x}, t) dt \\ &= -\frac{(2q+1)\lambda_n L_q(t)}{\tau_n} \int_{I_n} L_q(t)(t-t_{n-1})w_{h,\tau}(\mathbf{x}, t) dt, \end{aligned} \quad (3.11)$$

where we have used the fact that $w_{h,\tau}(\mathbf{x}, \cdot) \in \mathbb{P}^{q-1}(I_n)$ for all $\mathbf{x} \in \Omega$. Therefore, using the Hölder and the Cauchy–Schwarz inequalities, we get

$$\begin{aligned} \|(\text{Id} - \Pi_{q-1}^t)(\varphi_n w_{h,\tau})\|_{L^2(K_n)} &\leq \frac{(2q+1)\lambda_n}{\tau_n} \left(\|L_q\|_{L^2(I_n)}^2 \int_K \left(\int_{I_n} L_q(t)(t-t_{n-1})w_{h,\tau}(\mathbf{x}, t) dt \right)^2 d\mathbf{x} \right)^{\frac{1}{2}} \\ &\leq \frac{(2q+1)\lambda_n}{\tau_n} \|t-t_{n-1}\|_{L^\infty(I_n)} \|L_q\|_{L^2(I_n)} \left(\int_K \|L_q\|_{L^2(I_n)}^2 \|w_{h,\tau}(\mathbf{x}, \cdot)\|_{L^2(I_n)}^2 d\mathbf{x} \right)^{\frac{1}{2}} \\ &\leq \frac{(2q+1)\lambda_n}{\tau_n} \|t-t_{n-1}\|_{L^\infty(I_n)} \|L_q\|_{L^2(I_n)} \|w_{h,\tau}\|_{L^2(K_n)} \\ &\leq \lambda_n \tau_n \|w_{h,\tau}\|_{L^2(K_n)} \\ &= \zeta_q \|w_{h,\tau}\|_{L^2(K_n)}. \end{aligned}$$

Similarly, using the fact that $L_q(t_{n-1}) = (-1)^q$, we have

$$\begin{aligned} \|(\text{Id} - \Pi_{q-1}^t)(\varphi_n w_{h,\tau})(\cdot, t_{n-1}^+)\|_{L^2(\Omega)} &\leq \frac{(2q+1)\lambda_n}{\tau_n} \left(\int_\Omega \left(\int_{I_n} L_q(t)(t-t_{n-1})w_{h,\tau}(\mathbf{x}, t) dt \right)^2 d\mathbf{x} \right)^{\frac{1}{2}} \\ &\leq \frac{(2q+1)\lambda_n}{\tau_n} \|t-t_{n-1}\|_{L^\infty(I_n)} \|L_q\|_{L^2(I_n)} \|w_{h,\tau}\|_{L^2(\mathcal{Q}_n)} \\ &\leq \frac{1}{4\sqrt{2q+1}\sqrt{\tau_n}} \|w_{h,\tau}\|_{L^2(\mathcal{Q}_n)} \\ &= \frac{\sqrt{\lambda_n}}{2} \|w_{h,\tau}\|_{L^2(\mathcal{Q}_n)}, \end{aligned}$$

which completes the proof. \square

In what follows, the projection operator Π_{q-1}^t is to be understood as applied pointwise in space. The constant C_S in Lemma A1 depends on the index s . In the analysis below, whenever Lemma A1 is used for different values of s , the constant C_S is to be understood as the maximum of such constants. Similar conventions are made for the constants C_{Π} in Lemma A2 and C_{inv} in (A.2c).

3.2 Weak partial bound on the discrete solution

In this section, we prove a partial bound on the discrete solution $u_{h,\tau}$ of the linearized space–time formulation (3.1) that will be used to build the ‘stronger’ result in Section 3.3 below. In the first step, we show the following lower bound.

PROPOSITION 3.4. Let Assumption 1 on the discrete coefficient $\alpha_{h,\tau}$ hold. For $n = 1, \dots, N$, the following bound holds:

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}^{*,n}, c^2 u_{h,\tau}(\cdot, 0))) &\geq \frac{(1-|k|\alpha)}{2} \left(\|\partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \sum_{m=1}^{n-1} \|\llbracket \partial_t u_{h,\tau} \rrbracket_m\|_{L^2(\Omega)}^2 + \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \right) \\ &\quad + \frac{c^2}{2} (\|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 + \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2) + \delta \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 \\ &\quad + \frac{k}{2} \int_0^{t_n} \int_{\Omega} \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 \, \mathbf{d}\mathbf{x} \, dt, \end{aligned} \quad (3.12)$$

where

$$w_{h,\tau}^{*,n}|_{Q_m} := \begin{cases} \partial_t u_{h,\tau}|_{Q_m} & \text{if } m \leq n, \\ 0 & \text{if } m > n. \end{cases} \quad (3.13)$$

Proof. We use the definition of $\tilde{\mathcal{B}}_{h,\tau}(\cdot, \cdot)$ given in (3.2) and the identity

$$\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau})\partial_t u_{h,\tau} = \frac{1}{2}\partial_t((1 + k\alpha_{h,\tau})(\partial_t u_{h,\tau})^2) + \frac{k}{2}\partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2,$$

to obtain

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}^{*,n}, c^2 u_{h,\tau}(\cdot, 0))) &= \sum_{m=1}^n (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \partial_t u_{h,\tau})_{Q_m} \\ &\quad - \sum_{m=1}^{n-1} ((1 + k\alpha_{h,\tau}(\cdot, t_m))\llbracket \partial_t u_{h,\tau} \rrbracket_m, \partial_t u_{h,\tau}(\cdot, t_m^+))_{\Omega} + ((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}, \partial_t u_{h,\tau})_{\Sigma_0} \\ &\quad + c^2 \int_0^{t_n} (\nabla u_{h,\tau}, \nabla \partial_t u_{h,\tau})_{\Omega} \, dt + \delta \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 + c^2 \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \\ &= \frac{1}{2} \sum_{m=1}^n \left(\int_{Q_m} \partial_t((1 + k\alpha_{h,\tau})(\partial_t u_{h,\tau})^2) \, dV + k \int_{Q_m} \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 \, dV \right) \\ &\quad - \sum_{m=1}^{n-1} ((1 + k\alpha_{h,\tau}(\cdot, t_m))\llbracket \partial_t u_{h,\tau} \rrbracket_m, \partial_t u_{h,\tau}(\cdot, t_m^+))_{\Omega} + ((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}, \partial_t u_{h,\tau})_{\Sigma_0} \\ &\quad + \frac{c^2}{2} \left(\|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 + \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \right) + \delta \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \|(1 + k\alpha_{h,\tau}(\cdot, t_n))^{\frac{1}{2}} \partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|(1 + k\alpha_{h,\tau})^{\frac{1}{2}} \partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \\
 &\quad + \sum_{m=1}^{n-1} \int_{\Omega} (1 + k\alpha_{h,\tau}(\cdot, t_m)) \left(\frac{1}{2} \|\llbracket \partial_t u_{h,\tau} \rrbracket_m\|_{L^2(\Omega)}^2 - \llbracket \partial_t u_{h,\tau} \rrbracket_m \partial_t u_{h,\tau}(\cdot, t_m^+) \right) \mathbf{d}\mathbf{x} \\
 &\quad + \frac{c^2}{2} \left(\|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 + \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \right) + \delta \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 \\
 &\quad + \frac{k}{2} \int_0^{t_n} \int_{\Omega} \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 \mathbf{d}\mathbf{x} \, dt,
 \end{aligned}$$

which, together with the flux-jump identity

$$\frac{1}{2} \llbracket w^2 \rrbracket_n - w(\cdot, t_n^+) \llbracket w \rrbracket_n = \frac{1}{2} \llbracket w \rrbracket_n^2, \quad n = 1, \dots, N-1,$$

leads to

$$\begin{aligned}
 \tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}^{*,n}, c^2 u_{h,\tau}(\cdot, 0))) &= \frac{1}{2} \|(1 + k\alpha_{h,\tau}(\cdot, t_n))^{\frac{1}{2}} \partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|(1 + k\alpha_{h,\tau})^{\frac{1}{2}} \partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \\
 &\quad + \sum_{m=1}^{n-1} \|(1 + k\alpha_{h,\tau})(\cdot, t_m)^{\frac{1}{2}} \llbracket \partial_t u_{h,\tau} \rrbracket_m\|_{L^2(\Omega)}^2 \\
 &\quad + \frac{c^2}{2} \left(\|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 + \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \right) + \delta \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 \\
 &\quad + \frac{k}{2} \int_0^{t_n} \int_{\Omega} \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 \mathbf{d}\mathbf{x} \, dt.
 \end{aligned}$$

The desired result then follows by using the nondegeneracy assumption (3.4) of $\alpha_{h,\tau}$. □

Next, we employ the previous result to derive a partial bound on $u_{h,\tau}$.

PROPOSITION 3.5 (Weak partial bound on the discrete solution). Let $\delta \in [0, \bar{\delta}]$ for some fixed $\bar{\delta} > 0$, and let Assumption 1 on the discrete coefficient $\alpha_{h,\tau}$ hold. For $n = 1, \dots, N$, the following bound holds:

$$\begin{aligned}
 &\frac{(1 - |k|\underline{\alpha})}{4} \left(\|\partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \sum_{m=1}^{n-1} \|\llbracket \partial_t u_{h,\tau} \rrbracket_m\|_{L^2(\Omega)}^2 + \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \right) \\
 &\quad + \frac{c^2}{4} \left(\|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 + \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \right) + \frac{\delta}{2} \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 \\
 &\leq \left(\|f\|_{L^1(0,t_n;L^2(\Omega))} + C_{\text{inv}} \|\bar{c}^{-1} \xi\|_{L^1(0,t_n;L^2(\Omega))} \right) \|\partial_t u_{h,\tau}\|_{L^\infty(0,t_n;L^2(\Omega))} \\
 &\quad + \frac{|k|}{2} \|\partial_t \alpha_{h,\tau}\|_{L^1(0,t_n;L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0,t_n;L^2(\Omega))}^2 \\
 &\quad + \frac{1}{1 - |k|\underline{\alpha}} \|(1 + k u_0) u_1\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_0\|_{L^2(\Omega)^d}^2 + \frac{\bar{\delta}}{2} \|\nabla \partial_t \mu\|_{L^2(0,t_n;L^2(\Omega)^d)}^2, \quad (3.14)
 \end{aligned}$$

where $\bar{\tau}$ is defined as

$$\bar{\tau}|_{I_m} := \tau_m \quad \text{for } m = 1, \dots, N.$$

Proof. Let $w_{h,\tau}^{*,n}$ be given by (3.13). Using the definition in (3.3) of the linear functional $\mathcal{L}_\delta(\cdot)$, the polynomial inverse estimate (A.2c), and the Hölder and the Cauchy–Schwarz inequalities, we obtain the following bound:

$$\begin{aligned} \mathcal{L}_\delta \left(w_{h,\tau}^{*,n}, c^2 u_{h,\tau}(\cdot, 0) \right) &= \sum_{m=1}^n (f, \partial_t u_{h,\tau})_{Q_m} + ((1 + ku_0)u_1, \partial_t u_{h,\tau}(\cdot, 0))_\Omega \\ &\quad + \delta \sum_{m=1}^n (\nabla \partial_t \mu, \nabla \partial_t u_{h,\tau})_{Q_m} + \sum_{m=1}^n (\xi, \partial_t u_{h,\tau})_{Q_m} + c^2 (\nabla u_0, \nabla u_{h,\tau}(\cdot, 0))_\Omega \\ &\leq \|f\|_{L^1(0,t_n;L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0,t_n;L^2(\Omega))} + \|(1 + ku_0)u_1\|_{L^2(\Omega)} \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)} \\ &\quad + \delta \|\nabla \partial_t \mu\|_{L^2(0,t_n;L^2(\Omega)^d)} \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)} \\ &\quad + C_{\text{inv}} \left(\sum_{m=1}^n \|\tau_m^{-1} \xi\|_{L^1(I_m;L^2(\Omega))} \right) \|\partial_t u_{h,\tau}\|_{L^\infty(0,t_n;L^2(\Omega))} \\ &\quad + c^2 \|\nabla u_0\|_{L^2(\Omega)^d} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}. \end{aligned}$$

Young’s inequality can then be used to deduce

$$\begin{aligned} \mathcal{L}_\delta(w_{h,\tau}^{*,n}, c^2 u_{h,\tau}(\cdot, 0)) &\leq \|f\|_{L^1(0,t_n;L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0,t_n;L^2(\Omega))} \\ &\quad + \frac{1}{1 - |k|\underline{\alpha}} \|(1 + ku_0)u_1\|_{L^2(\Omega)}^2 + \frac{(1 - |k|\underline{\alpha})}{4} \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \\ &\quad + \frac{\delta}{2} \|\nabla \partial_t \mu\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 + \frac{\delta}{2} \|\nabla \partial_t u_{h,\tau}\|_{L^2(0,t_n;L^2(\Omega)^d)}^2 \\ &\quad + C_{\text{inv}} \|\bar{\tau}^{-1} \xi\|_{L^1(0,t_n;L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0,t_n;L^2(\Omega))} \\ &\quad + c^2 \|\nabla u_0\|_{L^2(\Omega)^d}^2 + \frac{c^2}{4} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2. \end{aligned} \tag{3.15}$$

The desired result (3.14) then follows by using the definition of the linearized space–time formulation (3.1), and bounds (3.12) and (3.15). \square

REMARK 3.6 (Limitations of Proposition 3.5). The terms multiplying $\|\partial_t u_{h,\tau}\|_{L^\infty(0,t_n;L^2(\Omega))}$ on the right-hand side of (3.14) cannot be controlled, as the left-hand side contains terms only related to the energy at the discrete times $\{t_m\}_{m=1}^n$. For this reason, we refer to Proposition 3.5 as ‘weak partial bound’ on the discrete solution. We remedy this in the next section by using a nonstandard test function in combination with Proposition 3.5. \blacksquare

3.3 Strong continuous dependence on the data

Henceforth, we use the notation $a \lesssim b$ to indicate the existence of a positive constant C independent of the meshsize h , the maximum time step τ , and the damping parameter δ such that $a \leq Cb$. Moreover, we write $a \simeq b$ meaning that $a \lesssim b$ and $b \lesssim a$.

Proof of Theorem 3.1. We make use of the auxiliary weight function introduced in (3.7). For $n = 1, \dots, N$, we consider the following test functions:

$$w_{h,\tau}^{\bullet,n}|_{Q_m} := \begin{cases} \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}|_{Q_n}) & \text{if } m = n, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad z_h^{\bullet,n} := \begin{cases} c^2 \theta u_{h,\tau}(\cdot, 0) & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

For the sake of clarity, we distinguish two cases ($n = 1$) and ($n > 1$) in the proof. However, this is not a recursive argument.

(Case $n = 1$). We choose as test function $(w_{h,\tau}^{\bullet,1}, z_h^{\bullet,1})$ in (3.1), and obtain the identity

$$\tilde{\mathcal{B}}_{h,\tau} \left(u_{h,\tau}, \left(w_{h,\tau}^{\bullet,1}, z_h^{\bullet,1} \right) \right) = \mathcal{L}_\delta \left(w_{h,\tau}^{\bullet,1}, z_h^{\bullet,1} \right),$$

where

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau} \left(u_{h,\tau}, \left(w_{h,\tau}^{\bullet,1}, z_h^{\bullet,1} \right) \right) &= (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\ &\quad + ((1 + k\alpha_{h,\tau}(\cdot, 0))\partial_t u_{h,\tau}(\cdot, 0), \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0))_\Omega \\ &\quad + c^2(\nabla u_{h,\tau}, \nabla \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} + c^2 \theta \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \\ &\quad + \delta(\nabla \partial_t u_{h,\tau}, \nabla \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\ &=: L_1 + L_2 + L_3 + L_4 + L_5, \end{aligned} \tag{3.16}$$

and

$$\begin{aligned} \mathcal{L}_\delta \left(w_{h,\tau}^{\bullet,1}, z_h^{\bullet,1} \right) &= (f, \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} + ((1 + ku_0)u_1, \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0))_\Omega \\ &\quad + \delta(\nabla \partial_t \mu, \nabla \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} + (\xi, \partial_t \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\ &\quad + c^2 \theta (\nabla u_0, \nabla u_{h,\tau}(\cdot, 0))_\Omega \\ &=: R_1 + R_2 + R_3 + R_4 + R_5. \end{aligned} \tag{3.17}$$

We first estimate the terms $\{L_i\}_{i=1}^5$ on the right-hand side of (3.16).

Bound on $L_1 + L_2$. Splitting $\Pi_{q-1}^t(\varphi \partial_t u_{h,\tau})$ as $\varphi \partial_t u_{h,\tau} - (\text{Id} - \Pi_{q-1}^t)(\varphi \partial_t u_{h,\tau})$ leads to

$$\begin{aligned}
L_1 + L_2 &= (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\
&\quad + ((1 + k\alpha_{h,\tau}(\cdot, 0))\partial_t u_{h,\tau}(\cdot, 0), \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0))_{\Omega} \\
&= (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \varphi_1 \partial_t u_{h,\tau})_{Q_1} \\
&\quad + \theta((1 + k\alpha_{h,\tau}(\cdot, 0))\partial_t u_{h,\tau}(\cdot, 0), \partial_t u_{h,\tau}(\cdot, 0))_{\Omega} \\
&\quad - (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), (\text{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\
&\quad - ((1 + k\alpha_{h,\tau}(\cdot, 0))\partial_t u_{h,\tau}(\cdot, 0), (\text{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0))_{\Omega}, \tag{3.18}
\end{aligned}$$

where we have used the identity $\varphi_1(0) = \theta$.

Since $\varphi_1(t_1) = \theta - \zeta_q$, using the identity

$$\begin{aligned}
\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}) \varphi \partial_t u_{h,\tau} &= \frac{1}{2} \partial_t \left((1 + k\alpha_{h,\tau}) \varphi (\partial_t u_{h,\tau})^2 \right) - \frac{1}{2} (1 + k\alpha_{h,\tau}) \varphi' (\partial_t u_{h,\tau})^2 \\
&\quad + \frac{k\varphi}{2} \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2,
\end{aligned}$$

the nondegeneracy bound on $(1 + k\alpha_{h,\tau})$ in (3.4), the uniform bound (3.8) for φ_n , the identity $\varphi'_n = -\lambda_n$, and the Hölder inequality the first two terms on the right-hand side of (3.18) can be bounded as follows:

$$\begin{aligned}
&(\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \varphi_1 \partial_t u_{h,\tau})_{Q_1} \\
&\quad + \theta((1 + k\alpha_{h,\tau}(\cdot, 0))\partial_t u_{h,\tau}(\cdot, 0), \partial_t u_{h,\tau}(\cdot, 0))_{\Omega} \\
&= \frac{1}{2} \int_{Q_1} \partial_t \left((1 + k\alpha_{h,\tau}) \varphi_1 (\partial_t u_{h,\tau})^2 \right) dV + \frac{\lambda_1}{2} \|(1 + k\alpha_{h,\tau})^{\frac{1}{2}} \partial_t u_{h,\tau}\|_{L^2(Q_1)}^2 \\
&\quad + \frac{k}{2} \int_{Q_1} \varphi_1 \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 dV + \theta \|(1 + k\alpha_{h,\tau})^{\frac{1}{2}} \partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \\
&= \frac{(\theta - \zeta_q)}{2} \|(1 + k\alpha_{h,\tau}(\cdot, t_1))\|^{\frac{1}{2}} \|\partial_t u_{h,\tau}(\cdot, t_1^-)\|_{L^2(\Omega)}^2 + \frac{\theta}{2} \|(1 + k\alpha_{h,\tau}(\cdot, 0))\|^{\frac{1}{2}} \|\partial_t u_{h,\tau}(\cdot, 0)\|_{L^2(\Omega)}^2 \\
&\quad + \frac{\lambda_1}{2} \int_{Q_1} (1 + k\alpha_{h,\tau}) (\partial_t u_{h,\tau})^2 dV + \frac{k}{2} \int_{Q_1} \varphi_1 \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 dV \\
&\geq \frac{(\theta - \zeta_q)(1 - |k|\underline{\alpha})}{2} \|\partial_t u_{h,\tau}(\cdot, t_1^-)\|_{L^2(\Omega)}^2 + \frac{\theta}{2} (1 - |k|\underline{\alpha}) \|\partial_t u_{h,\tau}(\cdot, 0)\|_{L^2(\Omega)}^2 \\
&\quad + \frac{(1 - |k|\underline{\alpha})\lambda_1}{2} \|\partial_t u_{h,\tau}\|_{L^2(Q_1)}^2 - \frac{|k|\theta}{2} \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}.
\end{aligned}$$

As for the third term on the right-hand side of (3.18), we use the orthogonality properties of Π_{q-1}^t , the Hölder inequality, estimate (3.9), the approximation properties of Π_0^t from Lemma A2, the inverse estimate (A.2c), Lemma A5, and the inequality $\|w\|_{L^2(I_n)} \leq \sqrt{\tau_n} \|w\|_{L^\infty(I_n)}$ to obtain the following bound:

$$\begin{aligned}
& -(\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), (\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\
&= -k(\partial_t \alpha_{h,\tau} \partial_t u_{h,\tau}, (\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} - k((\mathbf{Id} - \Pi_0^t)\alpha_{h,\tau} \partial_{tt} u_{h,\tau}, (\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\
&\geq -|k| \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \|(\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau})\|_{L^\infty(I_1; L^2(\Omega))} \\
&\quad - |k| \|(\mathbf{Id} - \Pi_0^t)\alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_{tt} u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \|(\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau})\|_{L^\infty(I_1; L^2(\Omega))} \\
&\geq -|k|(1 + C_{\text{inv}})\tau_1^{-\frac{1}{2}} \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \|(\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau})\|_{L^2(Q_1)} \\
&\quad - |k| C_\Pi C_{\text{inv}}(1 + C_{\text{inv}})\tau_1^{-\frac{1}{2}} \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_{tt} u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \|(\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau})\|_{L^2(Q_1)} \\
&\geq -|k|\zeta_q(1 + C_{\text{inv}})\tau_1^{-\frac{1}{2}} \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^2(Q_1)} \\
&\quad - |k|\zeta_q C_\Pi C_{\text{inv}}(1 + C_{\text{inv}})\tau_1^{-\frac{1}{2}} \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^2(Q_1)} \\
&\geq -\zeta_q |k|(1 + C_{\text{inv}})(1 + C_\Pi C_{\text{inv}}) \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}^2.
\end{aligned}$$

The last term on the right-hand side of (3.18) can be bounded using the Cauchy–Schwarz inequality, estimate (3.10), the nondegeneracy bound on $(1 + k\alpha_{h,\tau})$ in Assumption 1, the Young inequality, and the inequality $\|w\|_{L^2(I_n)} \leq \sqrt{\tau_n} \|w\|_{L^\infty(I_n)}$ as follows:

$$\begin{aligned}
& -((1 + k\alpha_{h,\tau}(\cdot, 0))\partial_t u_{h,\tau}(\cdot, 0), (\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0))_\Omega \\
&\geq -\frac{(1 - |k|\underline{\alpha})\theta}{4} \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 - \frac{1}{\theta} \left(\frac{1 + |k|\bar{\alpha}}{1 - |k|\underline{\alpha}} \right) \|(\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0)\|_{L^2(\Omega)}^2 \\
&\geq -\frac{(1 - |k|\underline{\alpha})\theta}{4} \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 - \frac{\lambda_1}{4\theta} \left(\frac{1 + |k|\bar{\alpha}}{1 - |k|\underline{\alpha}} \right) \|\partial_t u_{h,\tau}\|_{L^2(Q_1)}^2 \\
&\geq -\frac{(1 - |k|\underline{\alpha})\theta}{4} \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 - \frac{\zeta_q}{4\theta} \left(\frac{1 + |k|\bar{\alpha}}{1 - |k|\underline{\alpha}} \right) \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}^2.
\end{aligned}$$

Bound on $L_3 + L_4$. Adding and subtracting suitable terms, we have

$$\begin{aligned}
L_3 &= c^2(\nabla u_{h,\tau}, \varphi_1 \nabla \partial_t u_{h,\tau})_{Q_1} - c^2(\nabla u_{h,\tau}, (\mathbf{Id} - \Pi_{q-1}^t)(\varphi_1 \nabla \partial_t u_{h,\tau}))_{Q_1} \\
&=: L_3^{(a)} + L_3^{(b)}.
\end{aligned}$$

Using the Leibniz rule, the uniform bound (3.8) for φ_1 , and the fact that $\varphi_1'(t) = \lambda_1$, $\varphi_1(0) = \theta$, and $\varphi_1(t_1) = \theta - \zeta_q$, we get

$$\begin{aligned} L_3^{(a)} + L_4 &= \frac{c^2}{2} \int_{Q_1} \varphi_1 \partial_t (|\nabla u_{h,\tau}|^2) \, dV + c^2 \theta \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \\ &= \frac{c^2}{2} \left(\int_{Q_1} \partial_t (\varphi_1 |\nabla u_{h,\tau}|^2) - \int_{Q_1} \varphi_1' |\nabla u_{h,\tau}|^2 \, dV \right) + c^2 \theta \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \\ &= \frac{c^2}{2} \left(\int_{Q_1} \partial_t (\varphi_1 |\nabla u_{h,\tau}|^2) + \lambda_1 \|\nabla u_{h,\tau}\|_{L^2(Q_1)}^2 \right) + c^2 \theta \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 \\ &= \frac{c^2}{2} (\theta - \zeta_q) \|\nabla u_{h,\tau}\|_{L^2(\Sigma_1)^d}^2 + \frac{c^2 \theta}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2 + \frac{c^2 \lambda_1}{2} \|\nabla u_{h,\tau}\|_{L^2(Q_1)}^2. \end{aligned}$$

According to identity (3.11), there exists a function $\sigma_1(\mathbf{x})$ such that

$$u_{h,\tau} = (\text{Id} - \Pi_{q-1}^t) u_{h,\tau} + \Pi_{q-1}^t u_{h,\tau} = \sigma_1(\mathbf{x}) L_q(t) + \Pi_{q-1}^t u_{h,\tau}.$$

Moreover, we have

$$\partial_t u_{h,\tau} = \sigma_1(\mathbf{x}) L_q'(t) + \partial_t \Pi_{q-1}^t u_{h,\tau}.$$

Therefore, by using the orthogonality properties of Π_{q-1}^t and simple manipulations, we obtain

$$\begin{aligned} L_3^{(b)} &= -c^2 (\nabla u_{h,\tau}, (\text{Id} - \Pi_{q-1}^t) (\varphi_1 \nabla \partial_t u_{h,\tau}))_{Q_1} = -c^2 ((\text{Id} - \Pi_{q-1}^t) \nabla u_{h,\tau}, \varphi_1 \nabla \partial_t u_{h,\tau})_{Q_1} \\ &= -c^2 ((\text{Id} - \Pi_{q-1}^t) \nabla u_{h,\tau}, (\theta - \lambda_1 t) \nabla \sigma_1(\mathbf{x}) L_q'(t) + (\theta - \lambda_1 t) \partial_t \Pi_{q-1}^t \nabla u_{h,\tau})_{Q_1} \\ &= c^2 \lambda_1 ((\text{Id} - \Pi_{q-1}^t) \nabla u_{h,\tau}, \nabla \sigma_1(\mathbf{x}) t L_q'(t))_{Q_1} \\ &= c^2 \lambda_1 (\nabla \sigma_1(\mathbf{x}) L_q(t), \nabla \sigma_1(\mathbf{x}) t L_q'(t))_{Q_1} \\ &= c^2 \lambda_1 \|\nabla \sigma_1\|_{L^2(\Omega)^d}^2 \int_{I_1} t L_q(t) L_q'(t) \, dt \\ &= \frac{c^2 q \zeta_q}{2q+1} \|\nabla \sigma_1\|_{L^2(\Omega)^d}^2 \geq 0, \end{aligned} \tag{3.19}$$

where, in the last line, we have used the identity

$$\int_{I_1} t L_q(t) L_q'(t) \, dt = \frac{1}{2} \int_{I_1} t (L_q^2(t))' \, dt = \frac{\tau_1 q}{2q+1}.$$

Bound on L_5 . The commutativity of the orthogonal projection Π_{q-1}^t and the spatial gradient ∇ , the fact that $\nabla \partial_t u_{h,\tau}(\mathbf{x}, \cdot)$ belongs to $\mathbb{P}^{q-1}(\mathcal{T}_\tau)^d$ for all $\mathbf{x} \in \Omega$, and the uniform bound (3.8) for φ_1 lead to

$$\begin{aligned} L_5 &= \delta (\nabla \partial_t u_{h,\tau}, \nabla \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} = \delta (\nabla \partial_t u_{h,\tau}, \Pi_{q-1}^t \nabla(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\ &= \delta (\nabla \partial_t u_{h,\tau}, \varphi_1 \nabla \partial_t u_{h,\tau})_{Q_1} \\ &= \delta \|\sqrt{\varphi_1} \nabla \partial_t u_{h,\tau}\|_{L^2(Q_1)^d}^2 \\ &\geq \delta(\theta - \zeta_q) \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_1)^d}^2. \end{aligned}$$

We now focus on the terms $\{R_i\}_{i=1}^5$ on the right-hand side of (3.17).

Bound on R_1 . We use the orthogonality properties of Π_{q-1}^t , the stability bound in Lemma A1 for Π_{q-1}^t , the Hölder inequality, and the uniform bound (3.8) for φ_1 to obtain

$$\begin{aligned} R_1 &= (f, \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \leq \|f\|_{L^1(I_1; L^2(\Omega))} \|\Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})\|_{L^\infty(I_1; L^2(\Omega))} \\ &\leq C_S \theta \|f\|_{L^1(I_1; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}. \end{aligned}$$

Bound on R_2 . The second term on the right-hand side of (3.17) can be bounded using the Cauchy–Schwarz inequality, the polynomial trace inequality (A.1), the stability properties of Π_{q-1}^t , the uniform bound in (3.8) for φ_1 , and the inequality $\|w\|_{L^2(I_n)} \leq \sqrt{\tau_n} \|w\|_{L^\infty(I_n)}$ as follows:

$$\begin{aligned} R_2 &= ((1 + ku_0)u_1, \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0))_\Omega \\ &\leq \|(1 + ku_0)u_1\|_{L^2(\Omega)} \|\Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})(\cdot, 0)\|_{L^2(\Omega)} \\ &\leq C_{\text{tr}}^* \theta \tau_1^{-\frac{1}{2}} \|(1 + ku_0)u_1\|_{L^2(\Omega)} \|\partial_t u_{h,\tau}\|_{L^2(Q_1)} \\ &\leq C_{\text{tr}}^* \theta \|(1 + ku_0)u_1\|_{L^2(\Omega)} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \\ &\leq \frac{(C_{\text{tr}}^* \theta)^2}{2\gamma} \|(1 + ku_0)u_1\|_{L^2(\Omega)}^2 + \frac{\gamma}{2} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}^2 \end{aligned}$$

for all $\gamma > 0$.

Bound on R_3 . Using the commutativity of the spatial gradient operator and Π_{q-1}^t , the stability properties of the orthogonal projection Π_{q-1}^t , the uniform bound in (3.8) for φ_n , and the Young inequality, we get

$$\begin{aligned} R_3 &= \delta (\nabla \partial_t \mu, \nabla \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\ &\leq \delta \theta \|\nabla \partial_t \mu\|_{L^2(Q_1)^d} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_1)^d} \\ &\leq \frac{\bar{\delta} \theta^2}{2(\theta - \zeta_q)} \|\nabla \partial_t \mu\|_{L^2(Q_1)^d}^2 + \frac{\delta(\theta - \zeta_q)}{2} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_1)^d}^2. \end{aligned}$$

Bound on R_4 . We use the Hölder inequality, the polynomial inverse estimate (A.2c), the uniform bound in (3.8) for φ_n , and the inequality $\|w\|_{L^2(I_n)} \leq \sqrt{\bar{\tau}_n} \|w\|_{L^\infty(I_n)}$ to obtain

$$\begin{aligned} R_4 &= (\xi, \partial_t \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau}))_{Q_1} \\ &\leq \|\xi\|_{L^1(I_1; L^2(\Omega))} \|\partial_t \Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})\|_{L^\infty(I_1; L^2(\Omega))} \\ &\leq C_{\text{inv}} \|\tau_1^{-1} \xi\|_{L^1(I_1; L^2(\Omega))} \|\Pi_{q-1}^t(\varphi_1 \partial_t u_{h,\tau})\|_{L^\infty(I_1; L^2(\Omega))} \\ &\leq C_{\text{inv}} C_S \|\bar{\tau}^{-1} \xi\|_{L^1(I_1; L^2(\Omega))} \|\varphi_1 \partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \\ &\leq C_{\text{inv}} C_S \theta \|\bar{\tau}^{-1} \xi\|_{L^1(I_1; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}. \end{aligned}$$

Bound on R_5 . By using the Cauchy–Schwarz and the Young inequalities, we have

$$R_5 = c^2 \theta (\nabla u_0, \nabla u_{h,\tau}(\cdot, 0))_\Omega \leq \frac{c^2 \theta}{2} \|\nabla u_0\|_{L^2(\Omega)^d}^2 + \frac{c^2 \theta}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2.$$

Conclusion of case $n = 1$. From Lemma A5, we deduce that

$$\begin{aligned} \frac{\zeta_q}{(1 + C_{\text{inv}})^2} \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}^2 &\leq \lambda_n \|\partial_t u_{h,\tau}\|_{L^2(Q_n)}^2, \\ \frac{\zeta_q}{(1 + C_{\text{inv}})^2} \|\nabla u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega)^d)}^2 &\leq \lambda_n \|\nabla u_{h,\tau}\|_{L^2(Q_n)}^2, \end{aligned} \tag{3.20}$$

which, combined with all the above estimates, yields

$$\begin{aligned} &(\theta - \zeta_q) \left(\frac{(1 - |k|\underline{\alpha})}{2} \|\partial_t u_{h,\tau}(\cdot, t_1^-)\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_1)^d}^2 \right) + \frac{\theta(1 - |k|\underline{\alpha})}{4} \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \\ &+ \frac{\zeta_q}{2} \left(\frac{(1 - |k|\underline{\alpha})}{(1 + C_{\text{inv}})^2} - \frac{1}{2\theta} \left(\frac{1 + |k|\bar{\alpha}}{1 - |k|\underline{\alpha}} \right) - \frac{\gamma}{2} \right) \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}^2 + \frac{\zeta_q c^2}{2(1 + C_{\text{inv}})^2} \|\nabla u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega)^d)}^2 \\ &+ \frac{\delta(\theta - \zeta_q)}{2} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_1)^d}^2 \\ &\leq C_S \theta \|f\|_{L^1(I_1; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} + \frac{(C_{\text{ir}}^* \theta)^2}{2\gamma} \|(1 + k u_0) u_1\|_{L^2(\Omega)}^2 + \frac{c^2 \theta}{2} \|\nabla u_0\|_{L^2(\Omega)^d}^2 \\ &+ \frac{\bar{\delta} \theta^2}{2(\theta - \zeta_q)} \|\nabla \partial_t \mu\|_{L^2(Q_1)^d}^2 + C_{\text{inv}} C_S \theta \|\bar{\tau}^{-1} \xi\|_{L^1(I_1; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \\ &+ \left(\frac{\theta}{2} + \zeta_q (1 + C_{\text{inv}}) (1 + C_{\text{II}} C_{\text{inv}}) \right) |k| \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}. \end{aligned}$$

Taking θ large enough and γ small enough (independently of h , τ and δ) leads to

$$\begin{aligned} & \frac{1}{2} \|\partial_t u_{h,\tau}(\cdot, t_1^-)\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_1)^d}^2 + \|\partial_t u_{h,\tau}\|_{L^2(\Sigma_0)}^2 \\ & + \frac{1}{2} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega)^d)}^2 + \bar{\delta} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_1)}^2 \\ & \lesssim \left(\|f\|_{L^1(I_1; L^2(\Omega))} + \|\bar{\tau}^{-1} \xi\|_{L^1(I_1; L^2(\Omega))} \right) \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))} \\ & \quad + |k| \|\partial_t \alpha_{h,\tau}\|_{L^1(I_1; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}^2 \\ & \quad + \|(1 + ku_0)u_1\|_{L^2(\Omega)}^2 + c^2 \|\nabla u_0\|_{L^2(\Omega)^d}^2 + \bar{\delta} \|\nabla \partial_t \mu\|_{L^2(Q_1)^d}^2, \end{aligned}$$

where the hidden constant is independent of the time interval I_1 .

(Case $n > 1$). The rest of the proof consists in proving that, for $n = 2, \dots, N$, it holds

$$\begin{aligned} & \frac{1}{2} \|\partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 + \frac{1}{2} \|\llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}\|_{L^2(\Omega)}^2 \\ & + \frac{1}{2} \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega)^d)}^2 + \bar{\delta} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_n)}^2 \\ & \lesssim (\|f\|_{L^1(0, t_n; L^2(\Omega))} + \|\bar{\tau}^{-1} \xi\|_{L^1(0, t_n; L^2(\Omega))}) \|\partial_t u_{h,\tau}\|_{L^\infty(0, t_n; L^2(\Omega))} \\ & \quad + |k| \|\partial_t \alpha_{h,\tau}\|_{L^1(0, t_n; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0, t_n; L^2(\Omega))}^2 \\ & \quad + \|(1 + ku_0)u_1\|_{L^2(\Omega)}^2 + c^2 \|\nabla u_0\|_{L^2(\Omega)^d}^2 + \bar{\delta} \|\nabla \partial_t \mu\|_{L^2(0, t_n; L^2(\Omega)^d)}^2, \end{aligned} \tag{3.21}$$

with a hidden constant independent of the time interval I_n .

We choose in (3.1) as test function $(w_{h,\tau}^{\bullet,n}, z_h^{\bullet,n})$. Since $z_h^{\bullet,n} = 0$, with an abuse of notation, we omit it below. We obtain the identity

$$\tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, w_{h,\tau}^{\bullet,n}) = \mathcal{L}_\delta(w_{h,\tau}^{\bullet,n}),$$

where

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, w_{h,\tau}^{\bullet,n}) &= (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ & \quad - ((1 + k\alpha_{h,\tau}(\cdot, t_{n-1}))\llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}, \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau})(\cdot, t_{n-1}^+))_\Omega \\ & \quad + c^2 (\nabla u_{h,\tau}, \nabla \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ & \quad + \delta (\nabla \partial_t u_{h,\tau}, \nabla \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ & =: L_1^{(n)} + L_2^{(n)} + L_3^{(n)} + L_5^{(n)}, \end{aligned}$$

and

$$\begin{aligned}\mathcal{L}_\delta \left(w_{h,\tau}^{\bullet,n} \right) &= (f, \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}))_{Q_n} + \delta (\nabla \partial_t \mu, \nabla \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ &\quad + (\xi, \partial_t \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ &=: R_1^{(n)} + R_3^{(n)} + R_4^{(n)}.\end{aligned}$$

Most of the terms can be bounded as in the case $n = 1$, so we just focus on the slight differences.

Bound on $L_1^{(n)} + L_2^{(n)}$. By splitting $\Pi_{q-1}^t(\varphi \partial_t u_{h,\tau})$ as $\varphi \partial_t u_{h,\tau} - (\mathbf{Id} - \Pi_{q-1}^t)(\varphi \partial_t u_{h,\tau})$, we obtain

$$\begin{aligned}L_1^{(n)} + L_2^{(n)} &= (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \varphi_n \partial_t u_{h,\tau})_{Q_n} \\ &\quad - \theta((1 + k\alpha_{h,\tau}(\cdot, t_{n-1})) \llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}, \partial_t u_{h,\tau}(\cdot, t_{n-1}^+))_\Omega \\ &\quad - (\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), (\mathbf{Id} - \Pi_{q-1}^t)(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ &\quad + ((1 + k\alpha_{h,\tau}(\cdot, t_{n-1})) \llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}, (\mathbf{Id} - \Pi_{q-1}^t)(\varphi_n \partial_t u_{h,\tau})(\cdot, t_{n-1}^+))_\Omega.\end{aligned}\quad (3.22)$$

Since $\varphi_n(t_{n-1}) = \theta$ and $\varphi_n(t_n) = \theta - \zeta_q$, proceeding as for the case $n = 1$, and using the identity

$$\frac{1}{2}w(\cdot, t_{n-1}^+)^2 + \llbracket w \rrbracket_{n-1} w(\cdot, t_{n-1}^+) = -\frac{1}{2}\llbracket w \rrbracket_{n-1}^2 + \frac{1}{2}w(\cdot, t_{n-1}^-)^2,$$

we have

$$\begin{aligned}&(\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), \varphi_n \partial_t u_{h,\tau})_{Q_n} \\ &\quad - \theta((1 + k\alpha_{h,\tau}(\cdot, t_{n-1})) \llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}, \partial_t u_{h,\tau}(\cdot, t_{n-1}^+))_\Omega \\ &= \frac{1}{2} \int_{Q_n} \partial_t((1 + k\alpha_{h,\tau})\varphi_n(\partial_t u_{h,\tau})^2) \, dV + \frac{\lambda_n}{2} \|(1 + k\alpha_{h,\tau})^{\frac{1}{2}} \partial_t u_{h,\tau}\|_{L^2(Q_n)}^2 \\ &\quad + \frac{k}{2} \int_{Q_n} \varphi_n \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 \, dV - \theta((1 + k\alpha_{h,\tau}(\cdot, t_{n-1})) \llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}, \partial_t u_{h,\tau}(\cdot, t_{n-1}^+))_\Omega \\ &= \frac{(\theta - \zeta_q)}{2} \|(1 + k\alpha_{h,\tau}(\cdot, t_n))^{\frac{1}{2}} \partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 \\ &\quad - \theta \int_\Omega (1 + k\alpha_{h,\tau}(\cdot, t_{n-1})) \left(\frac{1}{2}(\partial_t u_{h,\tau}(\cdot, t_{n-1}^+))^2 + \llbracket \partial_t u_{h,\tau} \rrbracket_{n-1} \partial_t u_{h,\tau}(\cdot, t_{n-1}^+) \right) \, dx \\ &\quad + \frac{\lambda_n}{2} \int_{Q_n} (1 + k\alpha_{h,\tau}) (\partial_t u_{h,\tau})^2 \, dV + \frac{k}{2} \int_{Q_n} \varphi_n \partial_t \alpha_{h,\tau} (\partial_t u_{h,\tau})^2 \, dV \\ &\geq \frac{(\theta - \zeta_q)(1 - |k|\underline{\alpha})}{2} \|\partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \frac{(1 - |k|\underline{\alpha})\lambda_n}{2} \|\partial_t u_{h,\tau}\|_{L^2(Q_n)}^2 \\ &\quad + \frac{\theta(1 - |k|\underline{\alpha})}{2} \|\llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}\|_{L^2(\Omega)}^2 - \frac{\theta(1 + |k|\bar{\alpha})}{2} \|\partial_t u_{h,\tau}(\cdot, t_{n-1}^-)\|_{L^2(\Omega)}^2 \\ &\quad - \frac{|k|\theta}{2} \|\partial_t \alpha_{h,\tau}\|_{L^1(I_n; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}.\end{aligned}$$

The third term on the right-hand side of (3.22) can be bounded as for the case $n = 1$ as follows:

$$\begin{aligned} & -(\partial_t((1 + k\alpha_{h,\tau})\partial_t u_{h,\tau}), (\text{Id} - \Pi_{q-1}^t)(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ & \geq -\zeta_q |k| (1 + C_{\text{inv}}) \|\partial_t \alpha_{h,\tau}\|_{L^1(I_n; L^\infty(\Omega))} (1 + C_\Pi C_{\text{inv}}) \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}^2. \end{aligned}$$

Similarly, the following bound on the fourth term on the right-hand side of (3.22) can be obtained:

$$\begin{aligned} & ((1 + k\alpha_{h,\tau}(\cdot, t_{n-1})) \llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}, (\text{Id} - \Pi_{q-1}^t)(\varphi_n \partial_t u_{h,\tau})(\cdot, t_{n-1}^+))_{\Omega} \\ & \geq -\frac{(1 - |k|\underline{\alpha})\theta}{4} \|\llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}\|_{L^2(\Omega)}^2 - \frac{\zeta_q}{4\theta} \left(\frac{1 + |k|\bar{\alpha}}{1 - |k|\underline{\alpha}} \right) \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}^2. \end{aligned}$$

Bound on $L_3^{(n)}$. Adding and subtracting suitable terms, we have

$$\begin{aligned} L_3^{(n)} & = c^2 (\nabla u_{h,\tau}, \nabla \Pi_{q-1}^t(\varphi_n \partial_t u_{h,\tau}))_{Q_n} \\ & = c^2 (\nabla u_{h,\tau}, \varphi_n \nabla \partial_t u_{h,\tau})_{Q_n} - c^2 (\nabla u_{h,\tau}, (\text{Id} - \Pi_{q-1}^t)(\varphi_n \nabla \partial_t u_{h,\tau}))_{Q_n} \\ & \geq c^2 (\nabla u_{h,\tau}, \varphi_n \nabla \partial_t u_{h,\tau})_{Q_n}, \end{aligned}$$

as the last term in the second line is negative by the same reasoning used in (3.19) for $L_3^{(b)}$ above. Then, proceeding as for the case $n = 1$, we have

$$L_3^{(n)} \geq \frac{c^2}{2} (\theta - \zeta_q) \|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 - \frac{c^2 \theta}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_{n-1})^d}^2 + \frac{c^2 \lambda_n}{2} \|\nabla u_{h,\tau}\|_{L^2(Q_n)^d}^2.$$

The remaining terms can be treated as in the case $n = 1$ as follows:

$$\begin{aligned} R_5^{(n)} & \geq \delta(\theta - \zeta_q) \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_n)^d}^2, \\ R_1^{(n)} & \leq C_S \theta \|f\|_{L^1(I_n; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}, \\ R_3^{(n)} & \leq \frac{\bar{\delta}\theta^2}{2(\theta - \zeta_q)} \|\nabla \partial_t \mu\|_{L^2(Q_n)^d}^2 + \frac{\delta(\theta - \zeta_q)}{2} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_n)^d}^2, \\ R_4^{(n)} & \leq C_{\text{inv}} C_S \theta \|\bar{\tau}^{-1} \xi\|_{L^1(I_n; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}. \end{aligned}$$

Conclusion of case $n > 1$. Collecting the estimates for the case $n > 1$ and using (3.20), we have

$$\begin{aligned}
& (\theta - \zeta_q) \left(\frac{(1 - |k|\underline{\alpha})}{2} \|\partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 \right) + \frac{\theta(1 - |k|\underline{\alpha})}{4} \|\llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}\|_{L^2(\Omega)}^2 \\
& + \frac{\zeta_q}{2} \left(\frac{(1 - |k|\underline{\alpha})}{(1 + C_{\text{inv}})^2} - \frac{1}{2\theta} \left(\frac{1 + |k|\bar{\alpha}}{1 - |k|\underline{\alpha}} \right) \right) \|\partial_t u_{h,\tau}\|_{L^\infty(I_1; L^2(\Omega))}^2 + \frac{\zeta_q c^2}{2(1 + C_{\text{inv}})^2} \|\nabla u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega)^d)}^2 \\
& + \frac{\delta(\theta - \zeta_q)}{2} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_n)^d}^2 \\
& \leq \frac{\theta(1 + |k|\bar{\alpha})}{2} \|\partial_t u_{h,\tau}(\cdot, t_{n-1}^-)\|_{L^2(\Omega)}^2 + \frac{c^2 \theta}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_{n-1})^d}^2 \\
& + C_S \theta \|f\|_{L^1(I_n; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0, t_n; L^2(\Omega))} \\
& + \frac{\bar{\delta} \theta^2}{2(\theta - \zeta_q)} \|\nabla \partial_t \mu\|_{L^2(Q_n)^d}^2 + C_{\text{inv}} C_S \theta \bar{\tau}^{-1} \xi \|L^1(I_n; L^2(\Omega))\| \|\partial_t u_{h,\tau}\|_{L^\infty(0, t_n; L^2(\Omega))} \\
& + \left(\frac{\theta}{2} + \zeta_q(1 + C_{\text{inv}})(1 + C_\Pi C_{\text{inv}}) \right) |k| \|\partial_t \alpha_{h,\tau}\|_{L^1(I_n; L^\infty(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0, t_n; L^2(\Omega))}.
\end{aligned}$$

The same parameter θ used for the case $n = 1$ leads to

$$\begin{aligned}
& \frac{1}{2} \|\partial_t u_{h,\tau}(\cdot, t_n^-)\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_n)^d}^2 + \frac{1}{2} \|\llbracket \partial_t u_{h,\tau} \rrbracket_{n-1}\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_{n-1})^d}^2 \\
& + \frac{1}{2} \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega)^d)}^2 + \bar{\delta} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_n)^d}^2 \\
& \lesssim \frac{1}{2} \|\partial_t u_{h,\tau}(\cdot, t_{n-1}^-)\|_{L^2(\Omega)}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^2(\Sigma_{n-1})^d}^2 \\
& + (\|f\|_{L^1(I_n; L^2(\Omega))} + \|\bar{\tau}^{-1} \xi\|_{L^1(I_n; L^2(\Omega))}) \|\partial_t u_{h,\tau}\|_{L^\infty(0, t_n; L^2(\Omega))} \\
& + \frac{|k|\theta}{2} \|\partial_t \alpha_{h,\tau}\|_{L^1(I_n; L^2(\Omega))} \|\partial_t u_{h,\tau}\|_{L^\infty(0, t_n; L^2(\Omega))} + \bar{\delta} \|\nabla \partial_t \mu\|_{L^2(Q_n)^d}^2, \tag{3.23}
\end{aligned}$$

with a hidden constant independent of the time interval I_n .

By combining (3.23) with the estimate in Proposition 3.5 to bound $\|\partial_t u_{h,\tau}(\cdot, t_{n-1}^-)\|_{L^2(\Omega)}^2$, and using the additive properties of the L^1 - and L^2 -norms, we obtain (3.21).

Conclusion of the proof. For each $n \in \{1, \dots, N\}$, summing (3.21) and the estimate in Proposition 3.5 for $n = N$, and neglecting some positive terms, we arrive at

$$\begin{aligned}
& \frac{1}{2} \|\partial_t u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega))}^2 + \frac{c^2}{2} \|\nabla u_{h,\tau}\|_{L^\infty(I_n; L^2(\Omega)^d)}^2 + |\partial_t u_{h,\tau}|_{\mathcal{J}}^2 \\
& + \frac{c^2}{2} (\|\nabla u_{h,\tau}\|_{L^2(\Sigma_T)^d}^2 + \|\nabla u_{h,\tau}\|_{L^2(\Sigma_0)^d}^2) + \bar{\delta} \|\nabla \partial_t u_{h,\tau}\|_{L^2(Q_n)^d}^2 \\
& \lesssim (\|f\|_{L^1(0, T; L^2(\Omega))} + \|\bar{\tau}^{-1} \xi\|_{L^1(0, T; L^2(\Omega))}) \|\partial_t u_{h,\tau}\|_{L^\infty(0, T; L^2(\Omega))} \\
& + C_\alpha \|\partial_t u_{h,\tau}\|_{L^\infty(0, T; L^2(\Omega))}^2 + \|(1 + k u_0) u_1\|_{L^2(\Omega)}^2 + c^2 \|\nabla u_0\|_{L^2(\Omega)^d}^2 + \bar{\delta} \|\nabla \partial_t \mu\|_{L^2(Q_T)^d}^2. \tag{3.24}
\end{aligned}$$

Taking n as the index where the left-hand side of (3.24) takes its maximum value and C_α small enough, we obtain the desired result (3.6). \square

4. Convergence analysis of the linearized problem

In this section, we derive *a priori* bounds for the discrete error $\Pi_{h\tau}u - u_{h,\tau}$ in the energy norm, where $\Pi_{h\tau}$ is the combined projection introduced in Section 4.1 below, and then estimate $u - u_{h,\tau}$ in the norms

$$\|\partial_t v\|_{L^\infty(0,T;L^2(\Omega))} \quad \text{and} \quad \|\nabla v\|_{L^\infty(0,T;L^2(\Omega)^d)}. \quad (4.1)$$

Henceforth, we further assume that the family of spatial meshes $\{\mathcal{T}_h\}_{h>0}$ is quasi-uniform, i.e., there exists a positive constant C_{qu} independent of h such that $\text{diam}(K) \leq C_{\text{qu}} \text{diam}(K')$ for all $K, K' \in \mathcal{T}_h$. This is required to obtain optimal convergence rates for the error of the Ritz projection \mathcal{R}_h in the $L^2(\Omega)$ norm (see Lemmas B3 and B6). In addition, we also need the following assumption on the spatial domain Ω .

ASSUMPTION 2 (Elliptic regularity for Ω). The spatial domain Ω is such that

$$\text{if } z \in H_0^1(\Omega) \text{ and } \Delta z \in L^2(\Omega), \text{ then } z \in H^2(\Omega).$$

Furthermore, we assume that the solution u to the Westervelt IBVP belongs to the following space for all $\delta \in [0, \bar{\delta}]$:

$$\begin{aligned} \mathcal{U}^{\ell,m} := & W_1^m(0, T; H^2(\Omega)) \cap W_\infty^{m+1}(0, T; H^1(\Omega)) \cap W_1^2(0, T; H^{\ell+1}(\Omega)) \\ & \cap C^1([0, T]; H^{\ell+1}(\Omega) \cap H_0^1(\Omega)) \cap C^2([0, T]; L^\infty(\Omega)), \end{aligned} \quad (4.2)$$

for some $2 \leq m \leq q$ and $1 \leq \ell \leq p$, where p and q are the degrees of approximation in space and time, respectively. In addition, we assume that the initial data have (at least) the following regularity:

$$u_0 \in H^{\ell+1}(\Omega) \cap H_0^1(\Omega) \quad \text{and} \quad u_1 \in H^{\ell+1}(\Omega) \cap H_0^1(\Omega). \quad (4.3)$$

A δ -uniform well-posedness result for the Westervelt IBVP can be found in Kaltenbacher & Nikolić (2022, Thm. 4.1). For sufficiently smooth and small initial data (and possibly small final time), it has been shown that the problem has a solution

$$u \in L^\infty(0, T; H_\diamond^3(\Omega)) \cap W_\infty^1(0, T; H^2(\Omega) \cap H_0^1(\Omega)) \cap H^2(0, T; H^1(\Omega)),$$

where $H_\diamond^3(\Omega) = \{u \in H^3(\Omega) : \text{tr}_{\partial\Omega} u = 0, \text{tr}_{\partial\Omega} \Delta u = 0\}$. The analysis in Kaltenbacher & Nikolić (2022) is geared toward minimal regularity assumptions on data; we expect, however, that the techniques in Kaltenbacher & Nikolić (2022) can be extended to allow showing uniform well-posedness in $\mathcal{U}^{\ell,m}$.

Next theorem is the main result in this section and is proven in Section 4.3 below.

THEOREM 4.1 (*A priori* estimates for the linearized problem). Assume that $2 \leq m \leq q$ and $1 \leq \ell \leq p$, and that the spatial domain Ω satisfies Assumption 2. Assume further that \mathcal{T}_h is quasi-uniform, $\delta \in [0, \bar{\delta}]$ for some fixed $\bar{\delta} > 0$, and that Assumption 1 holds. Let $u_{h,\tau}$ be the solution to the linearized problem

(3.1) with $\mu = 0$, $\xi = 0$, and $\alpha_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0$. If bound (3.5) holds with C_α such that estimate (3.6) holds, the solution u to the Westervelt IBVP belongs to $\mathcal{U}^{\ell,m}$, and u_0, u_1 satisfy (4.3), then the following estimates hold:

$$\begin{aligned} \|\partial_t e_u\|_{L^\infty(0,T;L^2(\Omega))} &\leq C_{\text{lin}}^{(1)} \left\{ |k| \|\partial_t(u - \alpha_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \right. \\ &\quad + |k| \|u - \alpha_{h,\tau}\|_{C^0([0,T];L^2(\Omega))} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} \\ &\quad + \tau^m \left(\|\Delta \partial_t^{(m)} u\|_{L^1(0,T;L^2(\Omega))} + \|\partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega))} \right. \\ &\quad \left. + |k| \|\nabla \partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega)^d)} + \bar{\delta} \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)} \right) \\ &\quad + h^{\ell+1} \left(|k| \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + (1 + |k|\bar{\alpha}) \|\partial_{tt} u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \right) \\ &\quad + |k| \|u_1\|_{L^\infty(\Omega)} \|u_0\|_{H^{\ell+1}(\Omega)} + (1 + |k| \|u_0\|_{H^2(\Omega)}) |u_1|_{H^{\ell+1}(\Omega)} \\ &\quad \left. + \|\partial_t u\|_{C^0([0,T];H^{\ell+1}(\Omega))} \right\}, \end{aligned} \quad (4.4)$$

$$\begin{aligned} \|\nabla e_u\|_{L^\infty(0,T;L^2(\Omega)^d)} &\leq C_{\text{lin}}^{(2)} \left\{ |k| \|\partial_t(u - \alpha_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \right. \\ &\quad + |k| \|u - \alpha_{h,\tau}\|_{C^0([0,T];L^2(\Omega))} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} \\ &\quad + \tau^m \left(\|\Delta \partial_t^{(m)} u\|_{L^1(0,T;L^2(\Omega))} + (|k| + \tau) \|\nabla \partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega)^d)} \right. \\ &\quad \left. + \bar{\delta} \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)} \right) \\ &\quad + h^\ell \left(\|u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} + |k|h \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \right) \\ &\quad + (1 + |k|\bar{\alpha}) h \|\partial_{tt} u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \\ &\quad \left. + |k|h \|u_1\|_{L^\infty(\Omega)} |u_0|_{H^{\ell+1}(\Omega)} + (1 + |k| \|u_0\|_{H^2(\Omega)}) h |u_1|_{H^{\ell+1}(\Omega)} \right\}. \end{aligned} \quad (4.5)$$

REMARK 4.2 (Error estimates for the linear damped wave equation with constant coefficients). Setting $k = 0$ and $\delta > 0$, and assuming that, in addition to (4.2), u belongs to $W_1^{m+1}(0, T; H^2(\Omega))$, the following error estimates follow from Theorem 4.1:

$$\begin{aligned} \|\partial_t e_u\|_{L^\infty(0,T;L^2(\Omega))} &\lesssim \tau^m \left(\|\Delta \partial_t^{(m)} u\|_{L^1(0,T;L^2(\Omega))} + \|\partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega))} + \delta \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)} \right) \\ &\quad + h^{\ell+1} \left(\|\partial_{tt} u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + \|\partial_t u\|_{C^0([0,T];H^{\ell+1}(\Omega))} \right), \\ \|\nabla e_u\|_{L^\infty(0,T;L^2(\Omega)^d)} &\lesssim \tau^m \left(\tau \|\Delta \partial_t^{(m+1)} u\|_{L^1(0,T;L^2(\Omega))} + \tau \|\nabla \partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega)^d)} \right. \\ &\quad \left. + \delta \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)} \right) \\ &\quad + h^\ell \left(\|u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} + h \|\partial_{tt} u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \right). \end{aligned}$$

The estimate of $\partial_t e_u$ is optimal, whereas the estimate of ∇e_u is optimal with respect to h , and it is suboptimal by one order with respect to τ unless the damping parameter is small ($\delta \lesssim \tau$), which is the case we are interested in this work. \blacksquare

4.1 Combined space-time projection

We first recall the definition of the auxiliary projection in [Walkington \(2014, Def. 5.1\)](#).

DEFINITION 4.3 (Projection \mathcal{P}_τ). Let $q \in \mathbb{N}$ with $q \geq 2$. Given a partition \mathcal{T}_τ of the time interval $(0, T)$, the projection operator $\mathcal{P}_\tau : C^1(0, T) + \mathcal{V}_\tau^q \rightarrow \mathcal{V}_\tau^q$ is defined for all $v \in C^1(0, T) + \mathcal{V}_\tau^q$ as follows:

$$\mathcal{P}_\tau v(0) = v(0), \quad (4.6a)$$

$$(\mathcal{P}_\tau v)'(t_n^-) = v'(t_n^-) \quad \text{for } n = 1, \dots, N, \quad (4.6b)$$

$$((\mathcal{P}_\tau v)' - v', p_{q-2})_{I_n} = 0 \quad \text{for } n = 1, \dots, N, \quad \forall p_{q-2} \in \mathbb{P}^{q-2}(I_n). \quad (4.6c)$$

Moreover, we denote by $\mathcal{R}_h : H_0^1(\Omega) \rightarrow \mathcal{V}_h^p$ the Ritz projection, defined for any $z \in H_0^1(\Omega)$ as the solution to the following variational problem:

$$(\nabla(\mathcal{R}_h z - z), \nabla v_h)_\Omega = 0 \quad \forall v_h \in \mathcal{V}_h^p.$$

In what follows, the projection operators \mathcal{P}_τ and \mathcal{R}_h are to be understood, respectively, as applied pointwise in space and pointwise in time. We also denote by $\Pi_{h\tau}$ the operator $\mathcal{P}_\tau \circ \mathcal{R}_h$. We next state its approximation properties that will be exploited in the error analysis.

LEMMA 4.4 (Estimates for $\Pi_{h\tau}$). Let v belong to the space

$$W_\infty^{m+1}(0, T; L^2(\Omega)) \cap C^1([0, T]; H^{\ell+1} \cap H_0^1(\Omega)) \cap L^\infty(0, T; H^{\ell+1}(\Omega)) \cap W_\infty^{m+1}(0, T; H^1(\Omega)),$$

for $2 \leq m \leq q$ and $1 \leq \ell \leq p$. Then, the following estimates hold:

$$\|\partial_t(\mathbf{Id} - \Pi_{h\tau})v\|_{L^\infty(0, T; L^2(\Omega))} \lesssim \tau^m \|\partial_t^{(m+1)} v\|_{L^\infty(0, T; L^2(\Omega))} + h^{\ell+1} \|\partial_t v\|_{C^0([0, T]; H^{\ell+1}(\Omega))}, \quad (4.7a)$$

$$\|\nabla(\mathbf{Id} - \Pi_{h\tau})v\|_{L^\infty(0, T; L^2(\Omega)^d)} \lesssim h^\ell \|v\|_{L^\infty(0, T; H^{\ell+1}(\Omega))} + \tau^{m+1} \|\nabla \partial_t^{(m+1)} v\|_{L^\infty(0, T; L^2(\Omega)^d)}, \quad (4.7b)$$

$$\begin{aligned} & \sum_{n=1}^N \left(\tau_n^{-1} \|\partial_t(\Pi_{h\tau} v - v)\|_{L^2(Q_n)}^2 + \tau_n \|\partial_H(\Pi_{h\tau} v - v)\|_{L^2(Q_n)}^2 \right)^{\frac{1}{2}} \\ & \lesssim \tau_{\min}^{-\frac{1}{2}} h^{\ell+1} \|v\|_{C^0([0, T]; H^{\ell+1}(\Omega))} + \tau^{m-\frac{1}{2}} \|\partial_t^{(m+1)} v\|_{L^2(0, T; L^2(\Omega))}. \end{aligned} \quad (4.7c)$$

Proof. We use the stability and approximation properties of the operators \mathcal{P}_τ and \mathcal{R}_h from Lemmas B1, B2 and B3 to obtain

$$\begin{aligned} \|\partial_t(\text{Id} - \Pi_{h\tau})v\|_{L^\infty(0,T;L^2(\Omega))} &\leq \|\partial_t(\text{Id} - \mathcal{P}_\tau)v\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t\mathcal{P}_\tau(\text{Id} - \mathcal{R}_h)v\|_{L^\infty(0,T;L^2(\Omega))} \\ &\lesssim \tau^m \|\partial_t^{(m+1)}v\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t(\text{Id} - \mathcal{R}_h)v\|_{C^0([0,T];L^2(\Omega))} \\ &\lesssim \tau^m \|\partial_t^{(m+1)}v\|_{L^\infty(0,T;L^2(\Omega))} + h^{\ell+1} \|\partial_tv\|_{C^0([0,T];H^{\ell+1}(\Omega))}, \\ \|\nabla(\text{Id} - \Pi_{h\tau})v\|_{L^\infty(0,T;L^2(\Omega)^d)} &\leq \|\nabla(\text{Id} - \mathcal{R}_h)v\|_{L^\infty(0,T;L^2(\Omega)^d)} + \|\nabla\mathcal{R}_h(\text{Id} - \mathcal{P}_\tau)v\|_{L^\infty(0,T;L^2(\Omega)^d)} \\ &\lesssim h^\ell \|v\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} + \tau^{m+1} \|\nabla\partial_t^{(m+1)}v\|_{L^\infty(0,T;L^2(\Omega)^d)}. \end{aligned}$$

The estimate (4.7c) follows in an analogous way. \square

4.2 Bounds on the discrete error in the linearized problem

Let u be the solution to the Westervelt IBVP (1.1) and $u_{h,\tau}$ be the solution of the linearized space–time formulation (3.1) with $\mu = 0$, $\xi = 0$ and $\alpha_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0$. We consider the following error functions:

$$e_u := u - u_{h,\tau} = e_\Pi - \Pi_{h\tau} e_u \quad \text{and} \quad e_\Pi = u - \Pi_{h\tau} u.$$

The next result establishes a bound on $\Pi_{h\tau} e_u$ in the $\|\cdot\|_\delta$ norm, which we recall is defined in (2.5).

PROPOSITION 4.5 (Estimates of the discrete error). Let \mathcal{T}_h be quasi-uniform, $\delta \in [0, \bar{\delta}]$ for some fixed $\bar{\delta} > 0$, $q \in \mathbb{N}$ with $q \geq 2$, and let also Assumption 1 hold. Let $u_{h,\tau}$ be the solution of (3.1) with $\mu = 0$, $\xi = 0$ and $\alpha_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0$. If bound (3.5) holds with C_α such that estimate (3.6) holds, the initial conditions $(u_0, u_1) \in (H_0^1(\Omega) \cap L^\infty(\Omega)) \times (H_0^1(\Omega) \cap L^\infty(\Omega))$, the source term $f \in L^1(0, T; L^2(\Omega))$, and the solution u to the Westervelt IBVP (1.1) belongs to $W_1^2(0, T; L^\infty(\Omega)) \cap W_1^1(0, T; L^2(\Omega)) \cap C^1(0, T; H^2(\Omega) \cap H_0^1(\Omega))$, then the following estimate holds:

$$\begin{aligned} \|\Pi_{h\tau} e_u\|_\delta &\lesssim |k| \|\partial_t(u - \alpha_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\ &\quad + |k| \|u - \alpha_{h,\tau}\|_{C^0([0,T];L^2(\Omega))} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\ &\quad + |k| \|(\text{Id} - \mathcal{R}_h)\partial_t u\|_{L^1(0,T;L^2(\Omega))} + (1 + |k|\bar{\alpha}) \|(\text{Id} - \mathcal{R}_h)\partial_{tt} u\|_{L^1(0,T;L^2(\Omega))} \\ &\quad + \|(\text{Id} - \mathcal{P}_\tau)\Delta u\|_{L^1(0,T;L^2(\Omega))} + |k| \|u_1\|_{L^\infty(\Omega)} \|(\text{Id} - \mathcal{R}_h)u_0\|_{L^2(\Omega)} \\ &\quad + (1 + |k| \|\mathcal{R}_h u_0\|_{L^\infty(\Omega)}) \|(\text{Id} - \mathcal{R}_h)u_1\|_{L^2(\Omega)} \\ &\quad + |k| \sum_{n=1}^N \tau_n^{-1} \|(\text{Id} - \Pi_0^n)\alpha_{h,\tau}\|_{L^1(I_n;L^\infty(\Omega))} \|\partial_t(\text{Id} - \mathcal{P}_\tau)\mathcal{R}_h u\|_{L^\infty(I_n;L^2(\Omega))} \\ &\quad + \bar{\delta} \|\partial_t(\text{Id} - \mathcal{P}_\tau)\nabla u\|_{L^2(Q_T)^d}, \end{aligned} \tag{4.8}$$

where the hidden constant is independent of h , τ and δ .

Proof. Since the space–time variational formulation (3.1) is consistent except for the linearized term, and $u_{h,\tau}(\cdot, 0) = \alpha_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0$, the following error equation holds:

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(e_u, (w_{h,\tau}, z_h)) &= - \sum_{n=1}^N k(\partial_t((u - \alpha_{h,\tau})\partial_t u), w_{h,\tau})_{Q_n} \\ &\quad - k((u_0 - \mathcal{R}_h u_0)u_1, w_{h,\tau}(\cdot, 0))_{\Omega} \quad \forall (w_{h,\tau}, z_h) \in \mathcal{V}_{h,\tau}^{p,q} \times \mathcal{V}_h^p, \end{aligned}$$

whence

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(\Pi_{h\tau} e_u, (w_{h,\tau}, z_h)) &= \tilde{\mathcal{B}}_{h,\tau}(e_{\Pi}, (w_{h,\tau}, z_h)) \\ &\quad + \sum_{n=1}^N k(\partial_t((u - \alpha_{h,\tau})\partial_t u), w_{h,\tau})_{Q_n} \\ &\quad + k((u_0 - \mathcal{R}_h u_0)u_1, w_{h,\tau}(\cdot, 0))_{\Omega} \quad \forall (w_{h,\tau}, z_h) \in \mathcal{V}_{h,\tau}^{p,q} \times \mathcal{V}_h^p. \end{aligned} \quad (4.9)$$

Due to the choice of the discrete initial condition $u_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0$ (see Remark 2.2), the above identity is independent of z_h . Therefore, with an abuse of notation, we omit the test function z_h in the subsequent steps.

We wish to represent $\Pi_{h\tau} e_u$ as a solution to a linearized space–time formulation of the form of (3.1) with a suitable choice of the functions f , μ and ξ . To this end, we further rewrite $\tilde{\mathcal{B}}_{h,\tau}(e_{\Pi}, w_{h,\tau})$ in (4.9). By the definition of the bilinear form $\tilde{\mathcal{B}}_{h,\tau}(\cdot, \cdot)$ in (3.2), we have

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(e_{\Pi}, w_{h,\tau}) &= \sum_{n=1}^N (\partial_t((1 + k\alpha_{h,\tau})\partial_t e_{\Pi}), w_{h,\tau})_{Q_n} \\ &\quad - \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n))\llbracket \partial_t e_{\Pi} \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_{\Omega} + ((1 + k\alpha_{h,\tau})\partial_t e_{\Pi}, w_{h,\tau})_{\Sigma_0} \\ &\quad + c^2(\nabla e_{\Pi}, \nabla w_{h,\tau})_{Q_T} + \delta(\nabla \partial_t e_{\Pi}, \nabla w_{h,\tau})_{Q_T} \\ &=: J_1 + J_2 + J_3 + J_4 + J_5. \end{aligned}$$

We now use the properties of the projection $\Pi_{h\tau}$ to simplify each term $\{J_i\}_{i=1}^6$.

By splitting e_{Π} as $(u - \mathcal{P}_{\tau})u + \mathcal{P}_{\tau}(u - \mathcal{R}_h u)$, we have

$$\begin{aligned} J_1 + J_2 + J_3 &= \sum_{n=1}^N (\partial_t((1 + k\alpha_{h,\tau})\partial_t e_{\Pi}), w_{h,\tau})_{Q_n} - \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n))\llbracket \partial_t e_{\Pi} \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_{\Omega} \\ &\quad + ((1 + k\alpha_{h,\tau})\partial_t e_{\Pi}, w_{h,\tau})_{\Sigma_0} \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=1}^N (\partial_t((1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, w_{h,\tau}))_{Q_n} \\
&\quad - \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n))\llbracket \partial_t(\mathbf{Id} - \mathcal{P}_\tau)u \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_\Omega \\
&\quad + ((1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, w_{h,\tau})_{\Sigma_0} \\
&\quad + \sum_{n=1}^N (\partial_t(1 + k\alpha_{h,\tau})\partial_t\mathcal{P}_\tau(\mathbf{Id} - \mathcal{R}_h)u, w_{h,\tau})_{Q_n} \\
&\quad - \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n))\llbracket \partial_t\mathcal{P}_\tau(\mathbf{Id} - \mathcal{R}_h)u \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_\Omega \\
&\quad + ((1 + k\alpha_{h,\tau})\partial_t\mathcal{P}_\tau(\mathbf{Id} - \mathcal{R}_h)u, w_{h,\tau})_{\Sigma_0} \\
&=: J_{1,\tau} + J_{2,\tau} + J_{3,\tau} + J_{1,h} + J_{2,h} + J_{3,h}.
\end{aligned}$$

Integrating by parts in time, and using the definition in (4.6) of the projection \mathcal{P}_τ and the identity

$$v(\cdot, t_n^+)\llbracket u \rrbracket_n + u(\cdot, t_n^-)\llbracket v \rrbracket_n = \llbracket uv \rrbracket_n \quad \text{for } n = 1, \dots, N-1,$$

we deduce

$$\begin{aligned}
J_{1,\tau} + J_{2,\tau} + J_{3,\tau} &= \sum_{n=1}^N (\partial_t(1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, w_{h,\tau})_{Q_n} \\
&\quad - \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n))\llbracket \partial_t(\mathbf{Id} - \mathcal{P}_\tau)u \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_\Omega \\
&\quad + ((1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, w_{h,\tau})_{\Sigma_0} \\
&= - \sum_{n=1}^N ((1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, \partial_t w_{h,\tau})_{Q_n} + ((1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, w_{h,\tau})_{\Sigma_T} \\
&\quad + \sum_{n=1}^{N-1} \int_\Omega (1 + k\alpha_{h,\tau}(\cdot, t_n))(\llbracket w_{h,\tau} \partial_t(\mathbf{Id} - \mathcal{P}_\tau)u \rrbracket_n - w_{h,\tau}(\cdot, t_n^+)\llbracket \partial_t u_{h,\tau} \rrbracket_n) \, \mathbf{d}\mathbf{x} \\
&= - \sum_{n=1}^N ((1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, \partial_t w_{h,\tau})_{Q_n} + ((1 + k\alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, w_{h,\tau})_{\Sigma_T} \\
&\quad + \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n))\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u(\cdot, t_n^-), \llbracket w_{h,\tau} \rrbracket_n)_\Omega \\
&= - \sum_{n=1}^N (k(\alpha_{h,\tau} - \Pi_0^t \alpha_{h,\tau})\partial_t(\mathbf{Id} - \mathcal{P}_\tau)u, \partial_t w_{h,\tau})_{Q_n}, \tag{4.10}
\end{aligned}$$

where, in the last step, we have used the fact that $(\Pi_0^I \alpha_{h,\tau} \partial_t w_{h,\tau})(\mathbf{x}, \cdot) \in \mathbb{P}^{q-2}(\mathcal{T}_\tau)$ for all $\mathbf{x} \in \Omega$.

Integrating by parts in time and using similar arguments as above, we get

$$\begin{aligned}
 J_{1,h} + J_{2,h} + J_{3,h} &= \sum_{n=1}^N (\partial_t (1 + k\alpha_{h,\tau}) \partial_t \mathcal{P}_\tau (\mathbf{Id} - \mathcal{R}_h) u, w_{h,\tau})_{Q_n} \\
 &\quad - \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n)) \llbracket \partial_t \mathcal{P}_\tau (\mathbf{Id} - \mathcal{R}_h) u \rrbracket_n, w_{h,\tau}(\cdot, t_n^+))_\Omega \\
 &\quad + ((1 + k\alpha_{h,\tau}) \partial_t \mathcal{P}_\tau (\mathbf{Id} - \mathcal{R}_h) u, w_{h,\tau})_{\Sigma_0} \\
 &= - \sum_{n=1}^N ((1 + k\alpha_{h,\tau}) \partial_t \mathcal{P}_\tau (\mathbf{Id} - \mathcal{R}_h) u, \partial_t w_{h,\tau})_{Q_n} \\
 &\quad + ((1 + k\alpha_{h,\tau}) \partial_t \mathcal{P}_\tau (\mathbf{Id} - \mathcal{R}_h) u, w_{h,\tau})_{\Sigma_T} \\
 &\quad + \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n)) \partial_t \mathcal{P}_\tau (\mathbf{Id} - \mathcal{R}_h) u(\cdot, t_n^-), \llbracket w_{h,\tau} \rrbracket_n)_\Omega.
 \end{aligned}$$

Moreover, using the definition of \mathcal{P}_τ and integrating by parts back in time, we obtain

$$\begin{aligned}
 J_{1,h} + J_{2,h} + J_{3,h} &= - \sum_{n=1}^N ((1 + k\alpha_{h,\tau}) \partial_t (\mathbf{Id} - \mathcal{R}_h) u, \partial_t w_{h,\tau})_{Q_n} + ((1 + k\alpha_{h,\tau}) \partial_t (\mathbf{Id} - \mathcal{R}_h) u, w_{h,\tau})_{\Sigma_T} \\
 &\quad + \sum_{n=1}^{N-1} ((1 + k\alpha_{h,\tau}(\cdot, t_n)) \partial_t (\mathbf{Id} - \mathcal{R}_h) u(\cdot, t_n^-), \llbracket w_{h,\tau} \rrbracket_n)_\Omega \\
 &\quad + \sum_{n=1}^N (k\alpha_{h,\tau} \partial_t (\mathbf{Id} - \mathcal{P}_\tau) (\mathbf{Id} - \mathcal{R}_h) u, \partial_t w_{h,\tau})_{Q_n} \\
 &= \sum_{n=1}^N \left[(k\partial_t \alpha_{h,\tau} (\mathbf{Id} - \mathcal{R}_h) \partial_t u, w_{h,\tau})_{Q_n} + ((1 + k\alpha_{h,\tau}) (\mathbf{Id} - \mathcal{R}_h) \partial_{tt} u, w_{h,\tau})_{Q_n} \right. \\
 &\quad \left. + (k\alpha_{h,\tau} \partial_t (\mathbf{Id} - \mathcal{P}_\tau) (\mathbf{Id} - \mathcal{R}_h) u, \partial_t w_{h,\tau})_{Q_n} \right] \\
 &\quad + ((1 + k\alpha_{h,\tau}(\cdot, 0)) (\mathbf{Id} - \mathcal{R}_h) \partial_t u(\cdot, 0), w_{h,\tau})_\Omega. \tag{4.11}
 \end{aligned}$$

Finally, since $\alpha_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0$ and $\partial_t u(\cdot, 0) = u_1$, the last term in the above equation is equal to

$$((1 + k\mathcal{R}_h u_0) (\mathbf{Id} - \mathcal{R}_h) u_1, w_{h,\tau}(\cdot, 0))_\Omega.$$

As for the term J_4 , we split e_Π as $(\mathbf{Id} - \mathcal{R}_h) u + \mathcal{R}_h (\mathbf{Id} - \mathcal{P}_\tau) u$, and use the orthogonality properties of \mathcal{R}_h , the commutativity of \mathcal{P}_τ and the spatial gradient operator ∇ , and integration by parts in space to

obtain

$$\begin{aligned} J_4 &= c^2(\nabla e_\Pi, \nabla w_{h,\tau})_{Q_T} = c^2(\nabla(\text{Id} - \mathcal{R}_h)u, \nabla w_{h,\tau})_{Q_T} + c^2(\nabla \mathcal{R}_h(\text{Id} - \mathcal{P}_\tau)u, \nabla w_{h,\tau})_{Q_T} \\ &= c^2((\text{Id} - \mathcal{P}_\tau)\nabla u, \nabla w_{h,\tau})_{Q_T} = -c^2((\text{Id} - \mathcal{P}_\tau)\Delta u, w_{h,\tau})_{Q_T}. \end{aligned} \quad (4.12)$$

Similarly, we also have

$$\begin{aligned} J_5 &= \delta(\nabla \partial_t e_\Pi, \nabla w_{h,\tau})_{Q_T} = \delta(\nabla(\text{Id} - \mathcal{R}_h)\partial_t u, \nabla w_{h,\tau})_{Q_T} + \delta(\nabla \mathcal{R}_h \partial_t(\text{Id} - \mathcal{P}_\tau)u, \nabla w_{h,\tau})_{Q_T} \\ &= \delta(\nabla \partial_t(\text{Id} - \mathcal{P}_\tau)u, \nabla w_{h,\tau})_{Q_T}. \end{aligned} \quad (4.13)$$

From (4.9) and the simplified expressions of the terms $\{J_i\}_{i=1}^5$ in (4.10), (4.11), (4.12) and (4.13), we deduce that $\Pi_{ht}e_u$ solves a linearized space–time formulation of the form of (3.1) with

$$\begin{aligned} f &= k\partial_t \alpha_{h,\tau}(\text{Id} - \mathcal{R}_h)\partial_t u + (1 + k\alpha_{h,\tau})(\text{Id} - \mathcal{R}_h)\partial_{tt}u - c^2(\text{Id} - \mathcal{P}_\tau)\Delta u, \\ \xi &= -k(\text{Id} - \Pi_0^t)\alpha_{h,\tau}\partial_t(\text{Id} - \mathcal{P}_\tau)u - k(\text{Id} - \Pi_0^t)\alpha_{h,\tau}\partial_t(\text{Id} - \mathcal{P}_\tau)\mathcal{R}_h u, \\ \mu &= (\text{Id} - \mathcal{P}_\tau)u. \end{aligned}$$

Therefore, the continuous dependence on the data in Theorem 3.1 and the Hölder inequality yield the desired result. \square

4.3 A priori error estimates for the linearized problem

We are now in a position to prove Theorem 4.1.

Proof of Theorem 4.1. We first further estimate the discrete error by bounding the right-hand side terms in (4.8). By using the approximation properties of Π_0^t , \mathcal{P}_τ and \mathcal{R}_h from Lemmas A2, B2, B3 and B6, we obtain

$$\begin{aligned} |k| \|(\text{Id} - \mathcal{R}_h)\partial_t u\|_{L^1(0,T;L^2(\Omega))} &\lesssim |k|h^{\ell+1} \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))}, \\ (1 + |k|\bar{\alpha}) \|(\text{Id} - \mathcal{R}_h)\partial_{tt}u\|_{L^1(0,T;L^2(\Omega))} &\lesssim (1 + |k|\bar{\alpha})h^{\ell+1} \|\partial_{tt}u\|_{L^1(0,T;H^{\ell+1}(\Omega))}, \\ \|(\text{Id} - \mathcal{P}_\tau)\Delta u\|_{L^1(0,T;L^2(\Omega))} &\lesssim \tau^m \|\Delta \partial_t^{(m)}u\|_{L^1(0,T;L^2(\Omega))}, \\ |k| \|u_1\|_{L^\infty(\Omega)} \|(\text{Id} - \mathcal{R}_h)u_0\|_{L^2(\Omega)} &\lesssim h^{\ell+1} |k| \|u_1\|_{L^\infty(\Omega)} |u_0|_{H^{\ell+1}(\Omega)}, \\ (1 + |k| \|\mathcal{R}_h u_0\|_{L^\infty(\Omega)}) \|(\text{Id} - \mathcal{R}_h)u_1\|_{L^2(\Omega)} &\lesssim (1 + |k| \|u_0\|_{H^2(\Omega)}) h^{\ell+1} |u_1|_{H^{\ell+1}(\Omega)}. \end{aligned}$$

Furthermore,

$$\begin{aligned}
 |k| \sum_{n=1}^N \tau_n^{-1} \|(\mathbf{Id} - \Pi_0^t) \alpha_{h,\tau}\|_{L^1(I_n; L^\infty(\Omega))} & \|\partial_t(\mathbf{Id} - \mathcal{P}_\tau) \mathcal{R}_h u\|_{L^\infty(I_n; L^2(\Omega))} \\
 & \leq |k| C_\Pi \|\partial_t(\mathbf{Id} - \mathcal{P}_\tau) \mathcal{R}_h u\|_{L^\infty(0,T;L^2(\Omega))} \sum_{n=1}^N \|\partial_t \alpha_{h,\tau}\|_{L^1(I_n; L^\infty(\Omega))} \\
 & \lesssim |k| \|\partial_t \alpha_{h,\tau}\|_{L^1(0,T;L^\infty(\Omega))} \|\partial_t(\mathbf{Id} - \mathcal{P}_\tau) \nabla u\|_{L^\infty(0,T;L^2(\Omega)^d)} \\
 & \lesssim \tau^m |k| \|\nabla \partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega)^d)}, \\
 \bar{\delta} \|\partial_t(\mathbf{Id} - \mathcal{P}_\tau) \nabla u\|_{L^2(Q_T)} & \lesssim \bar{\delta} \tau^m \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)}.
 \end{aligned}$$

Therefore, the following estimate of the projected error is obtained:

$$\begin{aligned}
 \|\Pi_{h\tau} e_u\|_\delta & \lesssim |k| \|\partial_t(u - \alpha_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\
 & + |k| \|u - \alpha_{h,\tau}\|_{C^0([0,T];L^2(\Omega))} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} \\
 & + \tau^m \left(\|\Delta \partial_t^{(m)} u\|_{L^1(0,T;L^2(\Omega))} + |k| \|\nabla \partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega)^d)} + \bar{\delta} \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)} \right) \\
 & + h^{\ell+1} \left(|k| \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + (1 + |k| \bar{\alpha}) \|\partial_{tt} u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \right) \\
 & + |k| \|u_1\|_{L^\infty(\Omega)} |u_0|_{H^{\ell+1}(\Omega)} + (1 + |k| \|u_0\|_{H^2(\Omega)}) h |u_1|_{H^{\ell+1}(\Omega)}. \tag{4.14}
 \end{aligned}$$

Estimates (4.4) and (4.5) then follow using the triangle inequality, and combining estimates (4.14) and Lemma 4.4. \square

Due to condition (4.6a) in the definition of \mathcal{P}_τ , the following bound can be used to derive an error estimate of e_u :

$$\|e_u\|_{C^0([0,T];L^2(\Omega))} \leq \|(\mathbf{Id} - \mathcal{R}_h) u_0\|_{L^2(\Omega)} + T \|\partial_t e_u\|_{L^\infty(0,T;L^2(\Omega))}.$$

We use this estimate in Section 5 below; cf. (5.4).

5. *A priori* analysis of the nonlinear problem

In this section we prove the well-posedness and *a priori* bounds for the nonlinear problem (2.3) by relying on the results of Sections 3 and 4. To streamline the presentation, we introduce the short-hand notation

for (semi)norms involved in Theorem 4.1:

$$\begin{aligned}
\|u\|_m &:= \|\Delta \partial_t^{(m)} u\|_{L^1(0,T;L^2(\Omega))} + \|\partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega))} + |k| \|\nabla \partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega)^d)} \\
&\quad + \bar{\delta} \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)}, \\
\|u\|_{m,\tau} &:= \|\Delta \partial_t^{(m)} u\|_{L^1(0,T;L^2(\Omega))} + (|k| + \tau) \|\nabla \partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega)^d)} + \bar{\delta} \|\nabla \partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega)^d)}, \\
\|u\|_\ell &:= |k| \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + (1 + |k|\bar{\alpha}) \|\partial_{tt} u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + |k| \|u_1\|_{L^\infty(\Omega)} \|u_0\|_{H^{\ell+1}(\Omega)} \\
&\quad + \|\partial_t u\|_{C^0([0,T];H^{\ell+1}(\Omega))} + (1 + |k| \|u_0\|_{H^2(\Omega)}) h |u_1|_{H^{\ell+1}(\Omega)}, \\
\|u\|_{\ell,h} &:= \|u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} + |k|h \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + (1 + |k|\bar{\alpha}) h \|\partial_{tt} u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \\
&\quad + |k|h \|u_1\|_{L^\infty(\Omega)} |u_0|_{H^{\ell+1}(\Omega)} + (1 + |k| \|u_0\|_{H^2(\Omega)}) h |u_1|_{H^{\ell+1}(\Omega)}. \tag{5.1}
\end{aligned}$$

Recall that the variable coefficient $\alpha_{h,\tau}$ in the previous sections can be seen as a place holder for a fixed-point iterate. We next set up this fixed-point mapping to connect the linear results from Section 4 to the nonlinear problem (2.3).

5.1 The fixed-point operator

We define the mapping $\mathcal{T} : \mathbb{B} \ni \alpha_{h,\tau} \mapsto u_{h,\tau}$, where $\alpha_{h,\tau}$ is taken from

$$\begin{aligned}
\mathbb{B} = \left\{ \alpha_{h,\tau} \in \mathcal{V}_{h,\tau}^{p,q} : \|\partial_t(u - \alpha_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} \leq C_{\mathbb{B}}^{(1)} \left(\tau^m \|u\|_m + h^{\ell+1} \|u\|_\ell \right), \right. \\
\|\nabla(u - \alpha_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega)^d)} \leq C_{\mathbb{B}}^{(2)} \left(\tau^m \|u\|_{m,\tau} + h^\ell \|u\|_{\ell,h} \right), \\
\left. \alpha_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0 \right\},
\end{aligned}$$

and $u_{h,\tau}$ solves the corresponding linearized problem

$$\tilde{\mathcal{B}}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}, z_h)) = \ell(w_{h,\tau}, z_h) \quad \forall (w_{h,\tau}, z_h) \in \mathcal{V}_{h,\tau}^{p,q-1} \times \mathcal{V}_h^p.$$

The constants $C_{\mathbb{B}}^{(1)}$ and $C_{\mathbb{B}}^{(2)}$ will be chosen large enough by the upcoming proofs, but independent of the discretization parameters and δ .

The analysis of the nonlinear problem will follow by Banach's fixed-point theorem through two steps, where we first prove that \mathcal{T} is a well-defined self-mapping and then that it is strictly contractive.

5.2 The self-mapping property

In the first step we focus on proving the self-mapping property.

PROPOSITION 5.1. Let the assumptions of Theorem 4.1 hold, and let $\tau \lesssim h^{\frac{d}{2m}}$, $2 \leq m \leq q$ and $1 \leq \ell \leq p$. There exist large enough $C_{\mathbb{B}}^{(1)} > 0$ and $C_{\mathbb{B}}^{(2)} > 0$ (independent of h , τ and δ) in the definition of \mathbb{B} and

small enough $C_{\text{wellp}}^{(1)}$ (relative to $C_{\mathbb{B}}^{(1)}$), where

$$|k|(\|u\|_m + \|u\|_\ell + \|u\|_{W_\infty^1(0,T;H^{\ell+1}(\Omega))} + \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} + \|\partial_H u\|_{L^1(0,T;L^\infty(\Omega))}) \leq C_{\text{wellp}}^{(1)}, \quad (5.2)$$

such that \mathbb{B} is non-empty, the mapping \mathcal{T} is well defined, and $\mathcal{T}(\mathbb{B}) \subset \mathbb{B}$.

Proof. The ball \mathbb{B} is non-empty since $\Pi_{h\tau} u$ belongs to it as long as

$$C_{\mathbb{B}}^{(1)} \geq C_{\Pi_{h\tau}}^{(1)}, \quad C_{\mathbb{B}}^{(2)} \geq C_{\Pi_{h\tau}}^{(2)},$$

where $C_{\Pi_{h\tau}}^{(1)}$ and $C_{\Pi_{h\tau}}^{(2)}$ are the hidden constants in (4.7a) and (4.7b), respectively.

Let $\alpha_{h,\tau} \in \mathbb{B}$. The rest of the statement will follow by Theorem 4.1 once we check that the assumptions made there on $\alpha_{h,\tau}$ are satisfied.

Nondegeneracy and boundedness of the variable coefficient. We first verify Assumption 1. It is sufficient to show that

$$|k| \|\alpha_{h,\tau}\|_{L^\infty(0,T;L^\infty(\Omega))} \leq |k| \bar{\alpha} < 1. \quad (5.3)$$

Then, $\underline{\alpha} = \bar{\alpha}$ in (3.4).

We use the triangle inequality, the stability and approximation properties of the interpolant \mathcal{I}_h in Lemmas B4 and B5, and the polynomial inverse estimate (A.2b) to obtain

$$\begin{aligned} \|\alpha_{h,\tau}\|_{L^\infty(0,T;L^\infty(\Omega))} &\leq \|\alpha_{h,\tau} - \mathcal{I}_h u\|_{L^\infty(0,T;L^\infty(\Omega))} + \|\mathcal{I}_h u\|_{L^\infty(0,T;L^\infty(\Omega))} \\ &\leq C_{\text{inv}} h^{-d/2} \|\alpha_{h,\tau} - \mathcal{I}_h u\|_{L^\infty(0,T;L^2(\Omega))} + C_{\text{stab}}^{\mathcal{I}_h} \|u\|_{L^\infty(0,T;L^\infty(\Omega))} \\ &\leq C_{\text{inv}} h^{-d/2} (\|\alpha_{h,\tau} - u\|_{L^\infty(0,T;L^2(\Omega))} + \|\mathcal{I}_h u - u\|_{L^\infty(0,T;L^2(\Omega))}) \\ &\quad + C_{\text{stab}}^{\mathcal{I}_h} \|u\|_{L^\infty(0,T;L^\infty(\Omega))} \\ &\leq C_{\text{inv}} h^{-d/2} (\|\alpha_{h,\tau} - u\|_{L^\infty(0,T;L^2(\Omega))} + C_{\text{approx}}^{\mathcal{I}_h} h^{\ell+1} \|u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))}) \\ &\quad + C_{\text{stab}}^{\mathcal{I}_h} \|u\|_{L^\infty(0,T;L^\infty(\Omega))}. \end{aligned}$$

By (B.1a) and the fact that $\alpha_{h,\tau} \in \mathbb{B}$, we have

$$\begin{aligned} &h^{-d/2} \|\alpha_{h,\tau} - u\|_{C^0([0,T];L^2(\Omega))} \\ &\leq h^{-d/2} (\|(\text{Id} - \mathcal{R}_h)u_0\|_{L^2(\Omega)} + T \|\partial_t(u - \alpha_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))}) \\ &\leq C_{\text{approx}}^{\mathcal{R}_h} h^{\ell+1-d/2} \|u_0\|_{H^{\ell+1}(\Omega)} + T h^{-d/2} C_{\mathbb{B}}^{(1)} \left(\tau^m \|u\|_m + h^{\ell+1} \|u\|_\ell \right). \end{aligned} \quad (5.4)$$

Due to the condition $\tau \lesssim h^{\frac{d}{2m}}$, and the fact that $m \geq 2$ and $\ell \geq 1$, we have uniform bounds $h^{\ell+1-d/2} \leq C$ and $\tau^m h^{-d/2} \leq C$. Since $L^\infty(0,T;H^{\ell+1}(\Omega)) \hookrightarrow L^\infty(0,T;L^\infty(\Omega))$, the smallness condition in (5.3)

can be ensured to hold for small enough (relative to $C_{\mathbb{B}}^{(1)}$)

$$|k| \left(\|u_0\|_{H^{\ell+1}(\Omega)} + \|u\|_m + \|u\|_{\ell} + \|u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} \right).$$

Next, we check that condition (3.5) also holds; that is, we derive the uniform bound:

$$|k| \|\partial_t \alpha_{h,\tau}\|_{L^1(0,T;L^\infty(\Omega))} \leq C_\alpha. \quad (5.5)$$

We first have

$$|k| \|\partial_t \alpha_{h,\tau}\|_{L^1(0,T;L^\infty(\Omega))} \leq |k| T \|\partial_t \alpha_{h,\tau}\|_{L^\infty(0,T;L^\infty(\Omega))}$$

and we can further bound the right-hand side as follows:

$$\begin{aligned} \|\partial_t \alpha_{h,\tau}\|_{L^\infty(0,T;L^\infty(\Omega))} &\leq \|\partial_t(\alpha_{h,\tau} - \mathcal{I}_h u)\|_{L^\infty(0,T;L^\infty(\Omega))} + \|\mathcal{I}_h \partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))} \\ &\leq C_{\text{inv}} h^{-d/2} (\|\partial_t(\alpha_{h,\tau} - u)\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t(u - \mathcal{I}_h u)\|_{L^\infty(0,T;L^2(\Omega))}) \\ &\quad + C_{\text{stab}}^{\mathcal{I}_h} \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))} \\ &\leq C_{\text{inv}} h^{-d/2} (\|\partial_t(\alpha_{h,\tau} - u)\|_{L^\infty(0,T;L^2(\Omega))} + C_{\text{approx}}^{\mathcal{I}_h} h^{\ell+1} \|\partial_t u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))}) \\ &\quad + C_{\text{stab}}^{\mathcal{I}_h} \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))}. \end{aligned}$$

Moreover, since $\alpha_{h,\tau} \in \mathbb{B}$, we have

$$h^{-d/2} \|\partial_t(\alpha_{h,\tau} - u)\|_{L^\infty(0,T;L^2(\Omega))} \leq C_{\mathbb{B}}^{(1)} h^{-d/2} (\tau^m \|u\|_m + h^{\ell+1} \|u\|_{\ell}).$$

Thus, we can guarantee that (5.5) holds as long as

$$|k| \left(\|u\|_m + \|u\|_{\ell} + \|\partial_t u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} \right)$$

is sufficiently small (in other words, $C_{\text{wellp}}^{(1)}$ is sufficiently small), relative to $C_{\mathbb{B}}^{(1)}$.

Altogether, we conclude that the assumptions of Proposition 4.5 are satisfied. By Proposition 4.5, we thus know that the mapping \mathcal{T} is well defined. Furthermore, by (4.4) and (4.5), and the fact that $\alpha_{h,\tau} \in \mathbb{B}$, we have

$$\begin{aligned} &\|\partial_t e_u\|_{L^\infty(0,T;L^2(\Omega))} \\ &\leq C_{\text{lin}}^{(1)} (\tau^m \|u\|_m + h^{\ell+1} \|u\|_{\ell}) (1 + |k| C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} + |k| C_{\mathbb{B}}^{(1)} \|\partial_H u\|_{L^1(0,T;L^\infty(\Omega))}), \end{aligned}$$

and

$$\begin{aligned} &\|\nabla e_u\|_{L^\infty(0,T;L^2(\Omega)^d)} \\ &\leq C_{\text{lin}}^{(2)} (\tau^m \|u\|_{1,\tau} + h^\ell \|u\|_{2,h}) (1 + |k| C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} + |k| C_{\mathbb{B}}^{(1)} \|\partial_H u\|_{L^1(0,T;L^\infty(\Omega))}), \end{aligned}$$

where we recall that $C_{\text{lin}}^{(1)}$ and $C_{\text{lin}}^{(2)}$ are the constants in Theorem 4.1.

By (5.2), we have

$$\|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} + \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} \leq C_{\text{wellp}}^{(1)},$$

uniformly in δ . Then, it follows that $u_{h,\tau} \in \mathbb{B}$ as long as

$$C_{\mathbb{B}}^{(1)} \geq C_{\text{lin}}^{(1)}(1 + |k|C_{\mathbb{B}}^{(1)}C_{\text{wellp}}^{(1)}), \quad C_{\mathbb{B}}^{(2)} \geq C_{\text{lin}}^{(2)}(1 + |k|C_{\mathbb{B}}^{(1)}C_{\text{wellp}}^{(1)}).$$

In other words, as long as $C_{\text{wellp}}^{(1)}$ is small enough so that

$$\max\{C_{\text{lin}}^{(1)}, C_{\text{lin}}^{(2)}\} |k|C_{\text{wellp}}^{(1)} < 1,$$

and $C_{\mathbb{B}}^{(1)}, C_{\mathbb{B}}^{(2)}$ are large enough so that

$$C_{\mathbb{B}}^{(1)} \geq \frac{C_{\text{lin}}^{(1)}}{1 - C_{\text{lin}}^{(1)}|k|C_{\text{wellp}}^{(1)}}, \quad C_{\mathbb{B}}^{(2)} \geq \frac{C_{\text{lin}}^{(2)}}{1 - C_{\text{lin}}^{(2)}|k|C_{\text{wellp}}^{(1)}}.$$

Thus, for properly calibrated constants, we can conclude that $\mathcal{T}(\mathbb{B}) \subset \mathbb{B}$. □

We observe that the condition $\tau \lesssim h^{\frac{d}{2m}}$ is invoked in the above proof by the need to use an inverse estimate in space to ensure the non-degeneracy of the problem and uniformly bound $\|\partial_t \alpha_{h,\tau}\|_{L^\infty(0,T;L^\infty(\Omega))}$.

REMARK 5.2 (On the smallness of u). We note that the smallness imposed on u via (5.3) (which can be enforced via the smallness of data) could be somewhat relaxed by complementing it with a smallness assumption on $|k|T(h^{\ell+1-d/2} + \tau^m h^{-d/2})$ and $|k|\|u\|_{L^\infty(0,T;L^\infty(\Omega))}$, especially for high-order approximations and sufficiently smooth, solutions. This might be relevant for ultrasound applications, where data is often smooth but not necessarily small in higher-order norms; see [Kaltenbacher & Nikolić \(2022, §4\)](#) for further details. We also point out that all the smallness conditions imposed in this work are mitigated by the fact that the parameter k is small in practice as it is proportional to $1/c^2$; see, e.g., [Kaltenbacher \(2007, Ch. 5\)](#). □

5.3 Strict contractivity of the operator

In the next step we prove contractivity of the mapping. Compared to Proposition 5.1, we strengthen here the assumption on the lower bound on ℓ to $\ell \geq d/2$.

PROPOSITION 5.3. Let the partitions \mathcal{T}_h and \mathcal{T}_τ be quasi-uniform, and the assumptions of Proposition 5.1 hold. Let also $h^{\ell+1-d/2} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$, $2 \leq m \leq q$, $1 \leq \ell \leq p$, and $\ell + 1 - d/2 \geq d/(2(m-1))$. There exists $C_{\text{wellp}}^{(2)} > 0$ such that, if

$$|k|T^{\frac{3}{2}}\|u\| \leq C_{\text{wellp}}^{(2)},$$

where

$$\begin{aligned} \|u\| &= \|u\|_m + \|u\|_\ell + \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + \|\partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega))} \\ &\quad + \|\partial_t u\|_{C^0([0,T];L^\infty(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} + \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\ &\quad + \|u\|_{C^0([0,T];H^{\ell+1}(\Omega))} + \|\partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega))}, \end{aligned} \quad (5.6)$$

then the mapping \mathcal{T} is strictly contractive.

Proof. Take $\alpha_{h,\tau}^{(1)}, \alpha_{h,\tau}^{(2)} \in \mathbb{B}$ and denote $\bar{\alpha}_{h,\tau} = \alpha_{h,\tau}^{(2)} - \alpha_{h,\tau}^{(1)}$. Let $u_{h,\tau}^{(2)} = \mathcal{T}\alpha_{h,\tau}^{(2)}$ and $u_{h,\tau}^{(1)} = \mathcal{T}\alpha_{h,\tau}^{(1)} \in \mathbb{B}$. The difference $\bar{u}_{h,\tau} = u_{h,\tau}^{(2)} - u_{h,\tau}^{(1)}$ then solves

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(\bar{u}_{h,\tau}, (w_{h,\tau}, 0)) &= - \sum_{n=1}^N (k\partial_t \bar{\alpha}_{h,\tau} \partial_t u_{h,\tau}^{(1)} + k\bar{\alpha}_{h,\tau} \partial_t^2 u_{h,\tau}^{(1)}, w_{h,\tau})_{Q_n} \\ &\quad + \sum_{n=1}^{N-1} (k\bar{\alpha}_{h,\tau}(\cdot, t_n) [\partial_t u_{h,\tau}^{(1)}]_n, w_{h,\tau}(\cdot, t_n^+))_\Omega \quad \forall w_{h,\tau} \in \mathcal{W}_{h,\tau}^{p,q-1}, \end{aligned} \quad (5.7)$$

where $\alpha_{h,\tau}$ in the definition of $\tilde{\mathcal{B}}_{h,\tau}(\cdot, \cdot)$ in (3.2) is replaced by $\alpha_{h,\tau}^{(2)}$ here.

By defining the following lifting operator $\mathfrak{L}_\alpha : \mathcal{V}_{h,\tau}^{p,q} \rightarrow \mathcal{W}_{h,\tau}^{p,q-1}$:

$$\int_{Q_T} \mathfrak{L}_\alpha u_{h,\tau} w_{h,\tau} \, dV = \sum_{n=1}^{N-1} (k\bar{\alpha}_{h,\tau}(\cdot, t_n) [\partial_t u_{h,\tau}]_n, w_{h,\tau}(\cdot, t_n^+))_\Omega \quad \forall w_{h,\tau} \in \mathcal{W}_{h,\tau}^{p,q-1}, \quad (5.8)$$

the variational problem (5.7) can be written as

$$\tilde{\mathcal{B}}_{h,\tau}(\bar{u}_{h,\tau}, (w_{h,\tau}, 0)) = - \sum_{n=1}^N (k\partial_t \bar{\alpha}_{h,\tau} \partial_t u_{h,\tau}^{(1)} + k\bar{\alpha}_{h,\tau} \partial_t^2 u_{h,\tau}^{(1)}, w_{h,\tau})_{Q_n}$$

We can then rely on an analogous stability estimate to (3.6) with zero initial and Dirichlet boundary data, $\mu = 0$, and $\xi = 0$; in other words,

$$\|\bar{u}_{h,\tau}\|_\delta \lesssim \sum_{n=1}^N \|f\|_{L^1(I_n; L^2(\Omega))}, \quad (5.9)$$

where $f = -k\bar{\alpha}_{h,\tau}\partial_t^2 u_{h,\tau}^{(1)} - k\partial_t\bar{\alpha}_{h,\tau}\partial_t u_{h,\tau}^{(1)} + \mathfrak{L}_\alpha u_{h,\tau}^{(1)}$ (recall the definition of $\|\cdot\|_\delta$ in (2.5)). Toward estimating f , we note that

$$\begin{aligned} |k| \sum_{n=1}^N \|\bar{\alpha}_{h,\tau} \partial_{tt} u_{h,\tau}^{(1)}\|_{L^1(0,T;L^2(\Omega))} &\leq |k| \sum_{n=1}^N \|\bar{\alpha}_{h,\tau}\|_{L^\infty(I_n;L^2(\Omega))} \|\partial_{tt} u_{h,\tau}^{(1)}\|_{L^1(I_n;L^\infty(\Omega))} \\ &\leq |k| \|\bar{\alpha}_{h,\tau}\|_{C^0([0,T];L^2(\Omega))} \sum_{n=1}^N \|\partial_{tt} u_{h,\tau}^{(1)}\|_{L^1(I_n;L^\infty(\Omega))}. \end{aligned}$$

Further, due to $\bar{\alpha}_{h,\tau}(\cdot, 0) = 0$, we have

$$\|\bar{\alpha}_{h,\tau}\|_{C([0,T];L^2(\Omega))} = \left\| \int_0^t \partial_t \bar{\alpha}_{h,\tau} \, ds \right\|_{L^\infty(0,T;L^2(\Omega))} \leq T \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))}.$$

Next, using the triangle inequality, the polynomial inverse estimates (A.2c) and (A.2b), and the stability and approximation properties of \mathcal{P}_τ and \mathcal{I}_h in Lemmas B1, B2, B4 and B5, we obtain

$$\begin{aligned} &\sum_{n=1}^N \|\partial_{tt} u_{h,\tau}^{(1)}\|_{L^1(I_n;L^\infty(\Omega))} \\ &\leq \sum_{n=1}^N \|\partial_{tt} u_{h,\tau}^{(1)} - \partial_t(\mathcal{P}_\tau \circ \mathcal{I}_h) \partial_t u\|_{L^1(I_n;L^\infty(\Omega))} + \|\partial_t(\mathcal{P}_\tau \circ \mathcal{I}_h) \partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\ &\leq C_{\text{inv}}^2 \sum_{n=1}^N \tau_n^{-1} h^{-d/2} \|\partial_t u_{h,\tau}^{(1)} - (\mathcal{P}_\tau \circ \mathcal{I}_h) \partial_t u\|_{L^1(I_n;L^2(\Omega))} \\ &\quad + C_{\text{stab}}^{\mathcal{P}_\tau} \|\mathcal{I}_h \partial_{tt} u\|_{C^0([0,T];L^\infty(\Omega))} \\ &\leq C_{\text{inv}}^2 \sum_{n=1}^N \tau_n^{-1} h^{-d/2} \left(\|\partial_t(u_{h,\tau}^{(1)} - u)\|_{L^1(I_n;L^2(\Omega))} + \|\partial_t u - \mathcal{I}_h \partial_t u\|_{L^1(I_n;L^2(\Omega))} \right. \\ &\quad \left. + \|\mathcal{I}_h(\partial_t u - \mathcal{P}_\tau \partial_t u)\|_{L^1(I_n;L^2(\Omega))} \right) + C_{\text{stab}}^{\mathcal{P}_\tau} C_{\text{stab}}^{\mathcal{I}_h} \|\partial_{tt} u\|_{C^0([0,T];L^\infty(\Omega))} \\ &\leq C_{\text{inv}}^2 \tau^{-1} h^{-d/2} T \|\partial_t(u_{h,\tau}^{(1)} - u)\|_{L^\infty(0,T;L^2(\Omega))} + C_{\text{inv}}^2 \tau^{-1} h^{\ell+1-d/2} \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \\ &\quad + C_{\text{inv}}^2 C_{\text{stab}}^{\mathcal{I}_h} \tau^{-1} h^{-d/2} \|\partial_t u - \mathcal{P}_\tau \partial_t u\|_{L^1(I_n;L^2(\Omega))} + C_{\text{stab}}^{\mathcal{P}_\tau} C_{\text{stab}}^{\mathcal{I}_h} \|\partial_{tt} u\|_{C^0([0,T];L^\infty(\Omega))} \\ &\leq C_{\text{inv}}^2 \tau^{-1} h^{-d/2} T C_{\mathbb{B}}^{(1)} (\tau^m \|u\|_m + h^{\ell+1} \|u\|_\ell) + C_{\text{inv}}^2 \tau^{-1} h^{\ell+1-d/2} \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} \\ &\quad + C_{\text{inv}}^2 C_{\text{stab}}^{\mathcal{I}_h} \tau^{m-1} h^{-d/2} \|\partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega))} + C_{\text{stab}}^{\mathcal{P}_\tau} C_{\text{stab}}^{\mathcal{I}_h} \|\partial_{tt} u\|_{C^0([0,T];L^\infty(\Omega))}, \end{aligned} \tag{5.10}$$

where we have used the quasi-uniformity of \mathcal{T}_τ .

The assumption $h^{\ell+1-\frac{d}{2}} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$ implies the existence of a positive constant C , independent of h and τ such that

$$\tau^{m-1} h^{-\frac{d}{2}} \leq C \quad \text{and} \quad \tau^{-1} h^{\ell+1-d/2} \leq C. \quad (5.11)$$

Therefore, we have

$$\begin{aligned} & |k| \sum_{n=1}^N \|\bar{\alpha}_{h,\tau} \partial_n u_{h,\tau}^{(1)}\|_{L^1(0,T;L^2(\Omega))} \\ & \lesssim |k| T (\|u\|_m + \|u\|_\ell + \|\partial_t u\|_{L^1(0,T;H^{\ell+1}(\Omega))} + \|\partial_t^{(m+1)} u\|_{L^\infty(0,T;L^2(\Omega))} \\ & \quad + \|\partial_{tt} u\|_{C^0([0,T];L^\infty(\Omega))}) \|\bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))}. \end{aligned}$$

The second term within f can be treated as follows:

$$|k| \|\partial_t \bar{\alpha}_{h,\tau} \partial_t u_{h,\tau}^{(1)}\|_{L^1(0,T;L^2(\Omega))} \leq |k| \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \|\partial_t u_{h,\tau}^{(1)}\|_{L^1(0,T;L^\infty(\Omega))}.$$

We can reason that $\|\partial_t u_{h,\tau}^{(1)}\|_{L^\infty(0,T;L^\infty(\Omega))}$ can be made small enough by reducing u in suitable norms similarly to above using the fact that $u_{h,\tau}^{(1)} \in \mathbb{B}$:

$$\begin{aligned} \|\partial_t u_{h,\tau}^{(1)}\|_{L^1(0,T;L^\infty(\Omega))} & \leq \|\partial_t u_{h,\tau}^{(1)} - \partial_t \mathcal{I}_h u\|_{L^1(0,T;L^\infty(\Omega))} + \|\mathcal{I}_h \partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\ & \leq C_{\text{inv}} h^{-\frac{d}{2}} T \|\partial_t u_{h,\tau}^{(1)} - \partial_t \mathcal{I}_h u\|_{L^\infty(0,T;L^2(\Omega))} + C_{\text{stab}}^{\mathcal{I}_h} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\ & \leq C_{\text{inv}} h^{-\frac{d}{2}} T (\|\partial_t u_{h,\tau}^{(1)} - \partial_t u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u - \mathcal{I}_h \partial_t u\|_{L^\infty(0,T;L^2(\Omega))}) \\ & \quad + C_{\text{stab}}^{\mathcal{I}_h} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} \\ & \leq C_{\text{inv}} h^{-\frac{d}{2}} T \left(C_{\mathbb{B}}^{(1)} (\tau^m \|u\|_m + h^{\ell+1} \|u\|_\ell) + h^{\ell+1} \|\partial_t u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} \right) \\ & \quad + C_{\text{stab}}^{\mathcal{I}_h} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))}. \end{aligned}$$

Thus,

$$\begin{aligned} & |k| \|\partial_t \bar{\alpha}_{h,\tau} \partial_t u_{h,\tau}^{(1)}\|_{L^1(0,T;L^2(\Omega))} \\ & \leq C |k| T (\|u\|_m + \|u\|_\ell + \|\partial_t u\|_{L^\infty(0,T;H^{\ell+1}(\Omega))} + \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))}) \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))}. \end{aligned}$$

It remains to estimate the \mathfrak{L}_α term:

$$\|\mathfrak{L}_\alpha u_{h,\tau}^{(1)}\|_{L^1(0,T;L^2(\Omega))} \leq \sqrt{T} \|\mathfrak{L}_\alpha u_{h,\tau}^{(1)}\|_{L^2(0,T;L^2(\Omega))}.$$

Taking $w_{h,\tau} = \mathfrak{L}_\alpha u_{h,\tau}^{(1)}$ in (5.8), and using the inverse estimates in Lemma A4, the trace inequalities in Lemma A3, and the continuity in time of $\partial_t u$, we obtain

$$\begin{aligned}
 \|\mathfrak{L}_\alpha u_{h,\tau}^{(1)}\|_{L^2(Q_T)}^2 &= \sum_{n=1}^{N-1} (k\bar{\alpha}_{h,\tau}(\cdot, t_n) \llbracket \partial_t u_{h,\tau}^{(1)} \rrbracket_n, \mathfrak{L}_\alpha u_{h,\tau}^{(1)}(\cdot, t_n^+))_\Omega \\
 &\leq |k| \sum_{n=1}^{N-1} \|\bar{\alpha}_{h,\tau}(\cdot, t_n) \llbracket \partial_t u_{h,\tau}^{(1)} \rrbracket_n\|_{L^2(\Omega)} \|\mathfrak{L}_\alpha u_{h,\tau}^{(1)}(\cdot, t_n^+)\|_{L^2(\Omega)} \\
 &\leq |k| T \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \sum_{n=1}^{N-1} \|\llbracket \partial_t u_{h,\tau}^{(1)} \rrbracket_n\|_{L^\infty(\Omega)} \|\mathfrak{L}_\alpha u_{h,\tau}(\cdot, t_n^+)\|_{L^2(\Omega)} \\
 &\leq C_{\text{inv}} |k| T \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \sum_{n=1}^{N-1} h^{-\frac{d}{2}} \|\llbracket \partial_t u_{h,\tau}^{(1)} \rrbracket_n\|_{L^2(\Omega)} \|\mathfrak{L}_\alpha u_{h,\tau}(\cdot, t_n^+)\|_{L^2(\Omega)} \\
 &\leq C_{\text{inv}} |k| T \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \sum_{n=1}^{N-1} h^{-\frac{d}{2}} \left(\|\llbracket \partial_t (u_{h,\tau}^{(1)} - \Pi_{h\tau} u) \rrbracket_n\|_{L^2(\Omega)} \right. \\
 &\quad \left. + \|\llbracket \partial_t (\Pi_{h\tau} u - u) \rrbracket_n\|_{L^2(\Omega)} \right) \|\mathfrak{L}_\alpha u_{h,\tau}^{(1)}(\cdot, t_n^+)\|_{L^2(\Omega)} \\
 &\leq C_{\text{tr}}^* C_{\text{inv}} |k| T \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \tau^{-\frac{1}{2}} h^{-\frac{d}{2}} \sum_{n=1}^{N-1} \left(\|\llbracket \partial_t (u_{h,\tau}^{(1)} - \Pi_{h\tau} u) \rrbracket_n\|_{L^2(\Omega)} \right. \\
 &\quad \left. + \|\llbracket \partial_t (\Pi_{h\tau} u - u) \rrbracket_n\|_{L^2(\Omega)} \right) \|\mathfrak{L}_\alpha u_{h,\tau}\|_{L^2(Q_{n+1})} \\
 &\leq \sqrt{2} C_{\text{tr}}^* C_{\text{inv}} |k| T \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \tau^{-\frac{1}{2}} h^{-\frac{d}{2}} \left(|\partial_t (u_{h,\tau}^{(1)} - \Pi_{h\tau} u)|_{\mathcal{J}} \right. \\
 &\quad \left. + C_{\text{tr}} \left(\sum_{n=1}^N \tau_n^{-1} \|\partial_t (\Pi_{h\tau} u - u)\|_{L^2(Q_n)}^2 + \tau_n \|\partial_{tt} (\Pi_{h\tau} u - u)\|_{L^2(Q_n)}^2 \right)^{\frac{1}{2}} \right) \|\mathfrak{L}_\alpha u_{h,\tau}\|_{L^2(Q_T)}.
 \end{aligned}$$

Thus,

$$\begin{aligned}
 \|\mathfrak{L}_\alpha u_{h,\tau}\|_{L^1(0,T;L^2(\Omega))} &\leq \sqrt{2} C_{\text{tr}}^* C_{\text{inv}} |k| T^{\frac{3}{2}} \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \tau^{-\frac{1}{2}} h^{-\frac{d}{2}} \left\{ |\partial_t (u_{h,\tau}^{(1)} - \Pi_{h\tau} u)|_{\mathcal{J}} \right. \\
 &\quad \left. + C_{\text{tr}} \left(\sum_{n=1}^N \tau_n^{-1} \|\partial_t (\Pi_{h\tau} u - u)\|_{L^2(Q_n)}^2 + \tau_n \|\partial_{tt} (\Pi_{h\tau} u - u)\|_{L^2(Q_n)}^2 \right)^{\frac{1}{2}} \right\}. \quad (5.12)
 \end{aligned}$$

Further, using (4.14), we have

$$\begin{aligned}
 |\partial_t (u_{h,\tau}^{(1)} - \Pi_{h\tau} u)|_{\mathcal{J}} &\lesssim \tau^m \|u\|_m (1 + |k| C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} + |k| C_{\mathbb{B}}^{(1)} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))}) \\
 &\quad + h^{\ell+1} \|u\|_\ell (1 + |k| C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))} + |k| C_{\mathbb{B}}^{(1)} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))}),
 \end{aligned}$$

which, combined with (4.7c), (5.12), and the assumption $h^{\ell+1-d/2} \lesssim \tau \lesssim h^{d/(2m)}$, gives

$$\begin{aligned} \|\mathcal{L}_\alpha u_{h,\tau}\|_{L^1(0,T;L^2(\Omega))} &\lesssim |k|T^{\frac{3}{2}} \left\{ \|u\|_m (1 + |k|C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))}) + |k|C_{\mathbb{B}}^{(1)} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} \right\} \\ &\quad + \|u\|_\ell (1 + |k|C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))}) + |k|C_{\mathbb{B}}^{(1)} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} \\ &\quad + \|u\|_{C^0([0,T];H^{\ell+1}(\Omega))} + \|\partial_t^{(m+1)} u\|_{L^2(0,T;L^2(\Omega))} \Big\} \|\partial_t \bar{\alpha}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))}. \end{aligned}$$

By combining the derived estimates with (5.9), we arrive at

$$\begin{aligned} \|\bar{u}_{h,\tau}\|_\delta &\lesssim |k|T^{\frac{3}{2}} \|u\| (1 + |k|C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))}) + |k|C_{\mathbb{B}}^{(1)} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} \|\bar{\alpha}_{h,\tau}\|_\delta \\ &\lesssim (1 + |k|C_{\mathbb{B}}^{(1)} \|\partial_t u\|_{L^1(0,T;L^\infty(\Omega))}) + |k|C_{\mathbb{B}}^{(1)} \|\partial_{tt} u\|_{L^1(0,T;L^\infty(\Omega))} C_{\text{wellp}}^{(2)} \|\bar{\alpha}_{h,\tau}\|_\delta. \end{aligned}$$

We can thus guarantee strict contractivity of the mapping by reducing $C_{\text{wellp}}^{(2)}$ (relative to $C_{\mathbb{B}}^{(1)}$). \square

The above proof of in Proposition 5.3 involves estimating $\sum_{n=1}^N \|\partial_{tt} u_{h,\tau}^{(1)}\|_{L^1(I_n;L^\infty(\Omega))}$ in (5.10). As this second-derivative term is not present in the discrete energy, we are forced to use inverse estimates in both time and space. This invokes the assumption $h^{\ell+1-d/2} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$ (strengthened from $\tau \lesssim h^{\frac{d}{2m}}$ needed for proving the self-mapping property). In the discretization approaches for nonlinear acoustic models that employ finite-element discretization in space with standard time-stepping schemes in Dörich & Nikolić (2024); Nikolić (2025) a similar issue is circumvented by considering time-differentiated (semi)discrete problems; this approach, however, does not seem easily transferrable to the present setting. We next prove the well-posedness and accuracy of the nonlinear discrete problem.

THEOREM 5.4 (*A priori estimates for the discrete Westervelt equation*). Let the assumptions of Propositions 5.1 and 5.3 hold with $h^{\ell+1-d/2} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$, $2 \leq m \leq q$, $1 \leq \ell \leq p$ and $\ell+1-d/2 \geq d/(2(m-1))$. There exist $C_{\mathbb{B}}^{(1)} > 0$ and $C_{\mathbb{B}}^{(2)} > 0$, independent of h , τ and δ , and small enough C_{wellp} (relative to $C_{\mathbb{B}}^{(1)}$) such that, if the exact solution satisfies

$$|k| \|u\| \leq C_{\text{wellp}}$$

(where $\|\cdot\|$ is defined in (5.6)), then problem (2.3), that is,

$$\mathcal{B}_{h,\tau}(u_{h,\tau}, (w_{h,\tau}, z_h)) = \mathcal{L}(w_{h,\tau}, z_h) \quad \forall (w_{h,\tau}, z_h) \in \mathcal{W}_{h,\tau}^{p,q-1} \times \mathcal{V}_h^p,$$

has a unique solution $u_{h,\tau} \in \mathcal{V}_{h,\tau}^{p,q}$, which satisfies

$$\begin{aligned} \|\partial_t(u - u_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} &\leq C_{\mathbb{B}}^{(1)} \left(\tau^m \|u\|_m + h^{\ell+1} \|u\|_\ell \right), \\ \|\nabla(u - u_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega)^d)} &\leq C_{\mathbb{B}}^{(2)} \left(\tau^m \|u\|_{m,\tau} + h^\ell \|u\|_{\ell,h} \right). \end{aligned}$$

Proof. The statement follows by Propositions 5.1 and 5.3 and Banach's fixed-point theorem. \square

REMARK 5.5 (Assumption on τ). The assumption $h^{\ell+1-\frac{d}{2}} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$ determines just the scaling of τ , but it is not a CFL condition in the classical sense, as it is only necessary that the hidden constants are independent of h , τ and δ . For $d = 3$ and the minimal regularity $\ell = 2$ and $m = 2$, such an assumption becomes $\tau \simeq h^{\frac{3}{2}}$. Moreover, it relaxes for higher approximations in space and time and sufficiently smooth solutions. Note however, that the hidden constants in $h^{\ell+1-\frac{d}{2}} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$ influence the smallness needed from the exact data; see, e.g., (5.4) and (5.10).

We also note that, since the need for the condition $h^{\ell+1-d/2} \lesssim \tau$ arises in the proof of contractivity, we expect that it might still be possible to show existence of accurate solutions without this restriction by resorting to Schauder's fixed-point theorem instead of Banach's theorem; see, e.g., Nikolić (2025, Thm. 4.1) for similar arguments used in the analysis of a finite element semi-discretization of Westervelt's equation with nonlocal damping. ■

6. The vanishing dissipation parameter analysis

In this section, we discuss the behavior of discrete solutions as $\delta \searrow 0$. To emphasize the dependence on δ , we use here $u^{(\delta)}$ to denote the exact solution to the Westervelt IBVP for $\delta \in (0, \bar{\delta}]$ and $u^{(0)}$ to denote the exact solution to the inviscid problem. It is known that, as the dissipation vanishes, the exact dissipative solutions converge in the energy norm to the exact solution to the inviscid problem at a linear rate; see Kaltenbacher & Nikolić (2022, Thm. 5.1). In this section, we establish sufficient conditions under which this asymptotic behavior is preserved in the present discrete setting. Consider the family $\{u_{h,\tau}^{(\delta)}\}_{\delta \in (0, \bar{\delta}]}$ of solutions to

$$\mathcal{B}_{h,\tau}(u_{h,\tau}^{(\delta)}, (w_{h,\tau}, z_h)) = \mathcal{L}(w_{h,\tau}, z_h) \quad \forall (w_{h,\tau}, z_h) \in \mathcal{W}_{h,\tau}^{p,q-1} \times \mathcal{V}_h^p. \quad (6.1)$$

Denote the discrete solution corresponding to $\delta = 0$ (that is, of the inviscid discrete problem) by $u_{h,\tau}^{(0)}$. We have the following result that relates the two discrete problems.

THEOREM 6.1. Let the assumptions of Theorem 5.4 hold and let $\delta \in (0, \bar{\delta}]$. Let also $\tau \lesssim h^\vartheta$ with $\vartheta = \max\left\{\frac{d}{2(m-1)}, \frac{1}{m}\right\}$. Then the family of discrete solutions $\{u_{h,\tau}^{(\delta)}\}_{\delta \in (0, \bar{\delta}]}$ of (6.1) converges to the solution to the discrete inviscid problem at a linear rate as follows:

$$\begin{aligned} \|\partial_t(u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)})\|_{L^\infty(0,T;L^2(\Omega))} + \|\nabla(u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)})\|_{L^\infty(0,T;L^2(\Omega)^d)} + \left|\partial_t(u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)})\right|_{\mathcal{J}} \\ + \|\nabla(u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)})\|_{L^2(\Sigma_T)^d} \leq C\delta. \end{aligned}$$

Proof. Let $\bar{u}_{h,\tau} = u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)}$. Since $(\nabla \bar{u}_{h,\tau}(\cdot, 0), \nabla z_h)_{\Omega} = 0$ and $(k\bar{u}_{h,\tau} \partial_t u_{h,\tau}^{(0)}, w_{h,\tau})_{\Sigma_0} = 0$, this difference solves

$$\begin{aligned} \tilde{\mathcal{B}}_{h,\tau}(\bar{u}_{h,\tau}, (w_{h,\tau}, 0)) = \delta(\Delta_h \partial_t u_{h,\tau}^{(\delta)}, w_{h,\tau})_{Q_T} - (k\partial_t \bar{u}_{h,\tau} \partial_t u_{h,\tau}^{(0)} + k\bar{u}_{h,\tau} \partial_t^2 u_{h,\tau}^{(0)}, w_{h,\tau})_{Q_T} \\ + (\mathcal{L}_\alpha u_{h,\tau}^{(0)}, w_{h,\tau})_{Q_T} \quad \forall w_{h,\tau} \in \mathcal{W}_{h,\tau}^{p,q-1}, \quad (6.2) \end{aligned}$$

where $\Delta_h : H_0^1(\Omega) \rightarrow \mathcal{V}_h^p$ is the discrete Laplacian operator, defined for any $v \in H_0^1(\Omega)$ as follows:

$$(-\Delta_h v, z_h)_\Omega = (\nabla v, \nabla z_h)_\Omega \quad \forall z_h \in \mathcal{V}_h^p.$$

The last two terms on the right-hand side of (6.2) can be treated as in the proof of Proposition 5.3. Therefore, we obtain

$$\|\bar{u}_{h,\tau}\|_\delta \lesssim \delta \|\Delta_h u_{h,\tau}^{(\delta)}\|_{L^1(0,T;L^2(\Omega))},$$

where $C > 0$ does not depend on δ , h or τ . The right-hand side term above remains to be estimated. To treat it, we use the following inverse estimate:

$$\|\Delta_h v\|_{L^2(\Omega)} \lesssim h^{-1} \|\nabla v\|_{L^2(\Omega)^d} \quad \forall v \in H_0^1(\Omega),$$

together with the stability bound $\|\Delta_h v\|_{L^2(\Omega)} \leq \|\Delta v\|_{L^2(\Omega)}$ for all $v \in H_0^1(\Omega) \cap H^2(\Omega)$. Thus,

$$\begin{aligned} \delta \|\Delta_h u_{h,\tau}^{(\delta)}\|_{L^1(0,T;L^2(\Omega))} &\leq \delta T \|\Delta_h u_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} \\ &\lesssim \delta T (\|\Delta_h(u_{h,\tau}^{(\delta)} - u^{(\delta)})\|_{L^\infty(0,T;L^2(\Omega))} + \|\Delta_h u^{(\delta)}\|_{L^\infty(0,T;L^2(\Omega))}) \\ &\lesssim \delta T (h^{-1} \|\nabla(u_{h,\tau}^{(\delta)} - u^{(\delta)})\|_{L^\infty(0,T;L^2(\Omega)^d)} + \|\Delta u^{(\delta)}\|_{L^\infty(0,T;L^2(\Omega))}). \end{aligned}$$

Since $u_{h,\tau}^{(\delta)} \in \mathbb{B}$, we then have

$$\delta \|\Delta_h u_{h,\tau}^{(\delta)}\|_{L^1(0,T;L^2(\Omega))} \lesssim \delta T h^{-1} (\tau^m \|u^{(\delta)}\|_{m,\tau} + h^\ell \|u^{(\delta)}\|_{\ell,h} + \|u^{(\delta)}\|_{L^\infty(0,T;H^2(\Omega))}),$$

where the hidden constant does not depend on h , τ or δ . The statement then follows by using the assumption $\tau \lesssim h^{1/m}$ and noting that

$$\|\partial_t \bar{u}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega))} + \|\nabla \bar{u}_{h,\tau}\|_{L^\infty(0,T;L^2(\Omega)^d)} + |\partial_t \bar{u}_{h,\tau}|_{\mathcal{J}} + \|\nabla \bar{u}_{h,\tau}\|_{L^2(\Sigma_T)^d} \lesssim \|\bar{u}_{h,\tau}\|_\delta.$$

□

We note that $\tau \lesssim h^{1/m}$ is a more restrictive assumption than $\tau \lesssim h^{\frac{d}{2(m-1)}}$ only in a one-dimensional setting, that is, for $d = 1$. The present theory is, however, not optimized for that setting anyway, as there we can use, for example, the embedding $H^1(\Omega) \hookrightarrow L^\infty(\Omega)$ in place of inverse estimates in space in Section 5.

7. Numerical experiments

In this section we validate our theoretical results on the (h, τ) -convergence in Theorem 5.4, and the δ -convergence in Theorem 6.1. Moreover, we numerically assess the convergence of the (p, q) -version of the proposed method.

An object-oriented MATLAB implementation of the space–time method in (2.3) for (2 + 1)-dimensional problems was developed to carry out the numerical experiments.²

7.1 Implementation details

Since no continuity in time is required in the test space $\mathcal{W}_{h,\tau}^{p,q-1}$, the space–time formulation (2.3) reduces to solving the following sequence of nonlinear systems of equations on each time slab:

- For $n = 1$, $u_{h,\tau}(\cdot, 0) = \mathcal{R}_h u_0$ and

$$\begin{aligned} & (\partial_t((1 + ku_{h,\tau})\partial_t u_{h,\tau}), w_{h,\tau})_{Q_1} + ((1 + ku_{h,\tau})\partial_t u_{h,\tau}, w_{h,\tau})_{\Sigma_0} \\ & + c^2(\nabla u_{h,\tau}, \nabla w_{h,\tau})_{Q_1} + \delta(\nabla \partial_t u_{h,\tau}, \nabla w_{h,\tau})_{Q_1} \\ & = (f, w_{h,\tau})_{Q_1} + ((1 + ku_0)u_1, w_{h,\tau}(\cdot, 0))_{\Omega} \quad \forall w_{h,\tau} \in \mathbb{P}^{q-1}(I_1; \mathcal{V}_h^p). \end{aligned} \quad (7.1)$$

- For $n = 2, \dots, N$, $u_{h,\tau}(\cdot, t_{n-1}^+) = u_{h,\tau}(\cdot, t_{n-1}^-)$ and

$$\begin{aligned} & (\partial_t((1 + ku_{h,\tau})\partial_t u_{h,\tau}), w_{h,\tau})_{Q_n} + ((1 + ku_{h,\tau}(\cdot, t_{n-1}))\partial_t u_{h,\tau}(\cdot, t_{n-1}^+), w_{h,\tau}(\cdot, t_{n-1}^+))_{\Omega} \\ & + c^2(\nabla u_{h,\tau}, \nabla w_{h,\tau})_{Q_n} + \delta(\nabla \partial_t u_{h,\tau}, \nabla w_{h,\tau})_{Q_n} \\ & = (f, w_{h,\tau})_{Q_n} \\ & + ((1 + ku_{h,\tau}(\cdot, t_{n-1}))\partial_t u_{h,\tau}(\cdot, t_{n-1}^-), w_{h,\tau}(\cdot, t_{n-1}^+))_{\Omega} \quad \forall w_{h,\tau} \in \mathbb{P}^{q-1}(I_n; \mathcal{V}_h^p). \end{aligned} \quad (7.2)$$

Moreover, inspired by the linearized fixed-point iteration proposed in Kaltenbacher (2007, §5.4.2) (used there to solve the nonlinear problems stemming from the Newmark scheme applied to the Westervelt equation), we approximate the solution to the nonlinear systems of equations (7.2) using the following fixed-point iteration: for all $w_{h,\tau} \in \mathbb{P}^{q-1}(I_n; \mathcal{V}_h^p)$,

$$\begin{aligned} & (\partial_t u_{h,\tau}^{(s+1)}, w_{h,\tau})_{Q_n} + (\partial_t u_{h,\tau}^{(s+1)}(\cdot, t_{n-1}^+), w_{h,\tau}(\cdot, t_{n-1}^+))_{\Omega} \\ & + c^2(\nabla u_{h,\tau}^{(s+1)}, \nabla w_{h,\tau})_{Q_n} + \delta(\nabla \partial_t u_{h,\tau}^{(s+1)}, \nabla w_{h,\tau})_{Q_n} \\ & = (f, w_{h,\tau})_{Q_n} + ((1 + ku_{h,\tau}(\cdot, t_{n-1}))\partial_t u_{h,\tau}(\cdot, t_{n-1}^-), w_{h,\tau}(\cdot, t_{n-1}^+))_{\Omega} \\ & - k(\partial_t(u_{h,\tau}^{(s)}\partial_t u_{h,\tau}^{(s)}), w_{h,\tau})_{Q_n} - k(u_{h,\tau}(\cdot, t_{n-1})\partial_t u_{h,\tau}^{(s)}(\cdot, t_{n-1}^+), w_{h,\tau}(\cdot, t_{n-1}^+))_{\Omega}, \end{aligned} \quad (7.3)$$

whose associated linear system of equations has a unique solution due to Remark 3.2. The nonlinear system (7.1) is solved analogously.

In the numerical experiments below, for the fixed-point iteration (7.3), we use as initial guess $u_{h,\tau}^{(0)}$ the solution to problem (7.2) for $k = 0$. Moreover, we set a maximum number of linear iterations $s_{\max} = 15$ and a tolerance $tol = 10^{-12}$.

² The code used for the numerical tests are available in the GitHub repository Gómez (2025a).

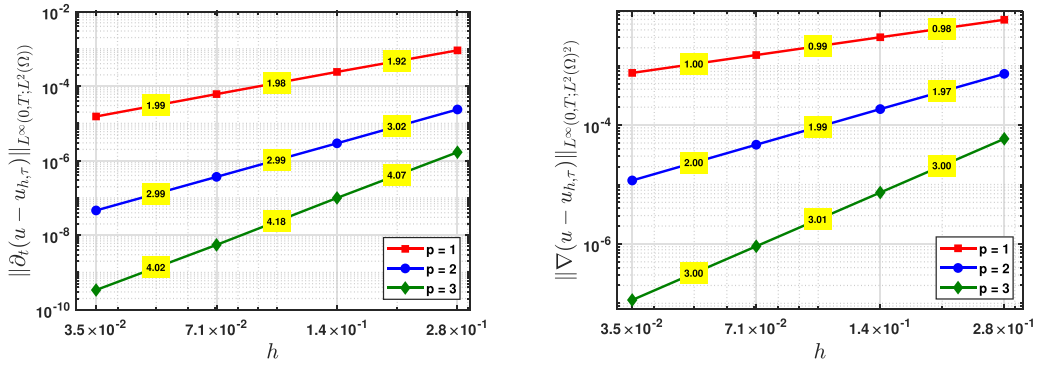


FIG. 1. h -convergence (in \log - \log scale) of the errors in (7.5) for the problem in Section 7.2 with exact solution u in (7.4) with $A = 10^{-2}$, $\omega = \pi/3$ and $\ell = \pi$. The numbers in the rectangles are the empirical convergence rates.

7.2 h -convergence

We consider a manufactured (2+1)-dimensional problem on the space–time domain $Q_T = (0, 1)^2 \times (0, 1)$ with $k = 0.5$, $c = 1$, $\delta = 6 \times 10^{-9}$, homogeneous Dirichlet boundary conditions, and initial data and source term chosen so that the exact solution to (1.1) is given by

$$u(x, y, t) = A \sin(\omega t) \sin(\ell x) \sin(\ell y), \tag{7.4}$$

where $A = 10^{-2}$, $\omega = \pi/3$ and $\ell = \pi$.

In order to numerically assess the h -convergence rates of the proposed scheme, we use approximations in time of degree $q = 5$ and a fixed partition of the time interval with uniform time step $\tau = 2 \times 10^{-1}$, so as to let the spatial error dominate. In Fig. 1, we show (in \log - \log scale) the results obtained for a set of structured simplicial meshes and approximations of degree $p = 1, 2, 3$. Optimal convergence rates of order $\mathcal{O}(h^p)$ and $\mathcal{O}(h^{p+1})$ are observed, respectively, for the following errors:

$$\|\partial_t(u - u_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega))} \quad \text{and} \quad \|\nabla(u - u_{h,\tau})\|_{L^\infty(0,T;L^2(\Omega)^2)}, \tag{7.5}$$

which are approximated by taking the maximum of the errors measured on uniformly distributed points on each time interval.

7.3 τ -convergence

We now assess the τ -convergence of the scheme. To do so, we consider the problem from the previous section with exact solution given by u in (7.4) with $A = 10^{-2}$, $\omega = 9\pi/2$ and $\ell = \pi$. For this value of ω , the solution u oscillates faster in time, thus increasing the influence of the temporal error.

We use approximations in space of degree $p = 6$, and a fixed coarse spatial mesh with $h \approx 2.82 \times 10^{-1}$. In Fig. 2, we show (in \log - \log scale) the results obtained for a set of uniform partitions of the time interval with $\tau = 0.5 \times 2^{-i}$, $i = 1, \dots, 4$, and approximations of degree $q = 2, 3, 4$. Optimal convergence rates of order $\mathcal{O}(\tau^q)$ and $\mathcal{O}(\tau^{q+1})$ are observed, respectively, for the errors in (7.5).

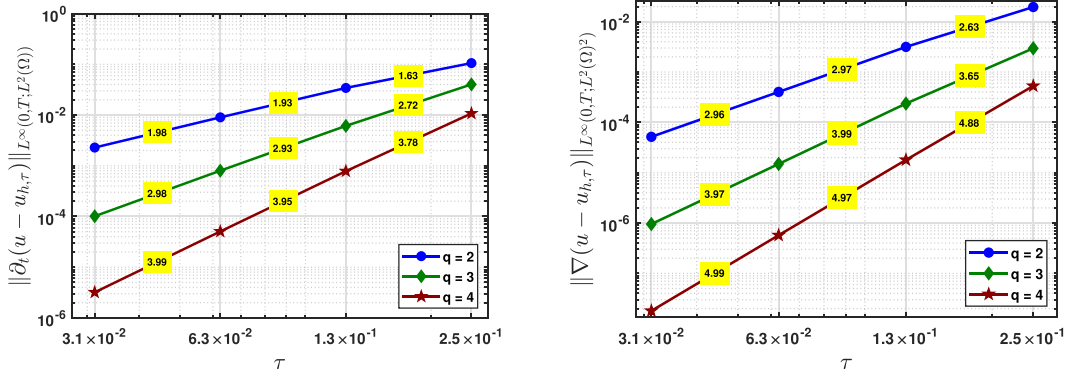


FIG. 2. τ -convergence (in \log - \log scale) of the errors in (7.5) for the problem in Section 7.3 with exact solution u in (7.4) with $A = 10^{-2}$, $\omega = 9\pi/2$ and $\ell = \pi$.

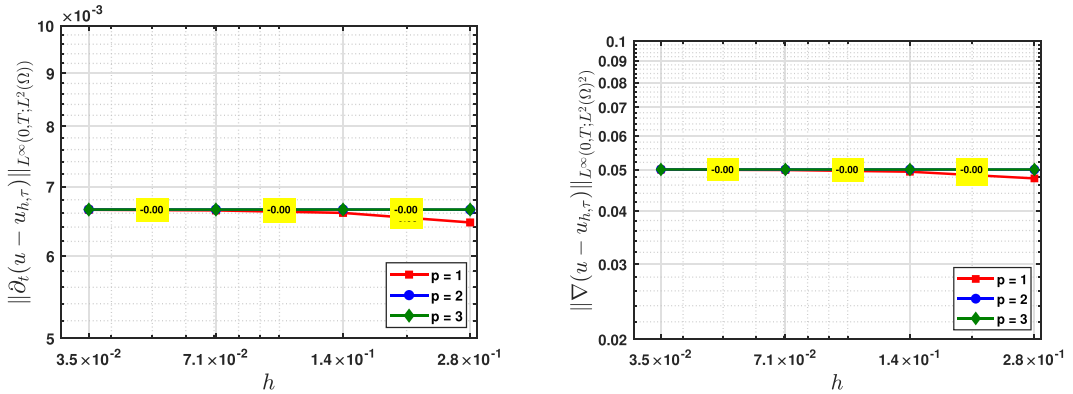


FIG. 3. Behavior (in \log - \log scale) of the errors in (7.5) for the problem in Section 7.2 with final time $T = 100$. The results correspond to $q = 2$ and $\tau = 20$.

7.4 Assessment of the constraints on space-time meshes

We now assess numerically the necessity of assumption $h^{\ell+1-\frac{d}{2}} \lesssim \tau \lesssim h^{\frac{d}{2(m-1)}}$ in Theorem 5.4, which, for $d = 2$, $\ell = p$ and $m = q$, simplifies to $h^p \lesssim \tau \lesssim h^{\frac{1}{q-1}}$. To this end, we study both constraints $h^p \lesssim \tau$ and $\tau \lesssim h^{\frac{1}{q-1}}$ separately.

We first consider the test case in Section 7.2, using a final time $T = 100$, time step $\tau = 20$, and degree of approximation in time $q = 2$. In Fig. 3, we present the results obtained for various spatial meshes and degrees of approximation in space $p = 1, 2, 3$, where we observe that the errors remain bounded even when the constraint $\tau \lesssim h$ is not satisfied.

Similarly, we consider the test case in Section 7.3, with final time $T = 1$, a coarse spatial mesh, and degree of approximation in space $p = 1$. In Fig. 4, we show the results obtained for a set of time partitions and degrees of approximation in time $q = 2, 3, 4$, which show that the errors remain bounded even when the constraint $h \lesssim \tau$ is not satisfied.

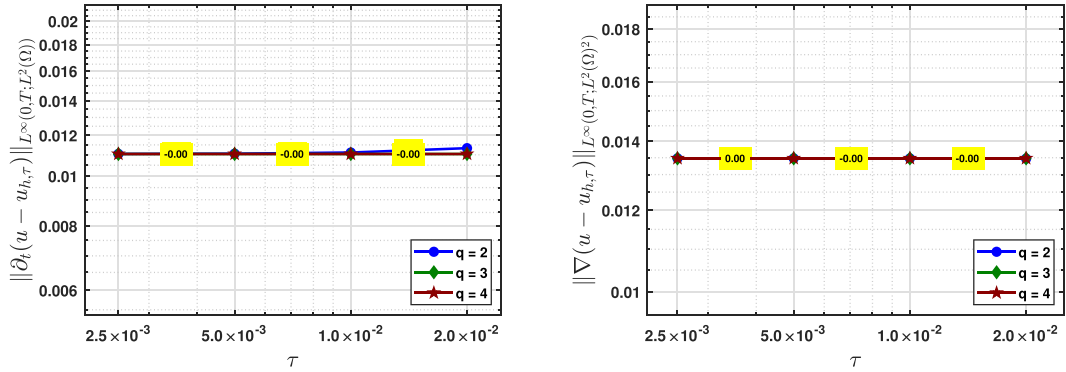


FIG. 4. Behavior (in \log - \log scale) of the errors in (7.5) for the problem in Section 7.3 with final time $T = 1$. The results correspond to $p = 1$ and the coarsest spatial mesh.

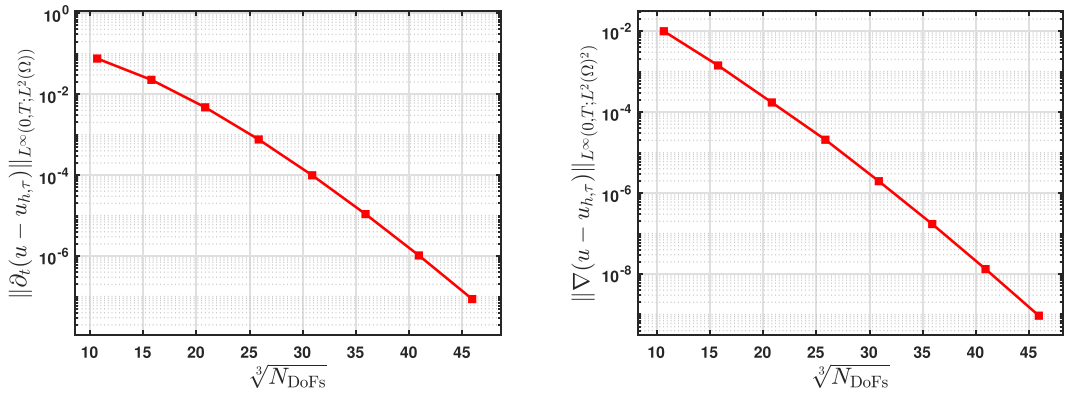


FIG. 5. p -convergence (in *semilogy* scale) of the errors in (7.5) for the problem in Section 7.5 with exact solution u in (7.4) with $A = 10^{-2}$, $\omega = 9\pi/2$ and $\ell = \pi$.

These numerical results suggest that the present theoretical constraint on the space–time meshes may be too restrictive, at least for solution fields of the type in (7.4) that do not exhibit nonlinear effects, such as steepening of the wavefront.

7.5 (p, q) -convergence

Taking advantage of the simultaneous high-order convergence in space and time provided by the proposed method, we numerically assess the (p, q) -version of the method. More precisely, given a fixed space–time mesh with $h = \sqrt{2}\tau \approx 2.82 \times 10^{-1}$, we increase the approximation degrees in space and time; for simplicity, we set $q = p$. In Fig. 5, we show (in *semilogy* scale) the errors obtained using approximations of degree $p = 2, \dots, 9$ for the problem in Section 7.3, and observe exponential convergence of order $\mathcal{O}(e^{-b\sqrt[3]{N_{\text{DoFs}}}})$ for the errors in (7.5), where N_{DoFs} denotes the total number of degrees of freedom. Such an exponential convergence was also observed in Dong *et al.* (2024, §4.1) for the DG–CG scheme applied to the linear wave equation, and it follows the expected decay for analytical solutions with p -FEM approximations (Schwab, 1998).

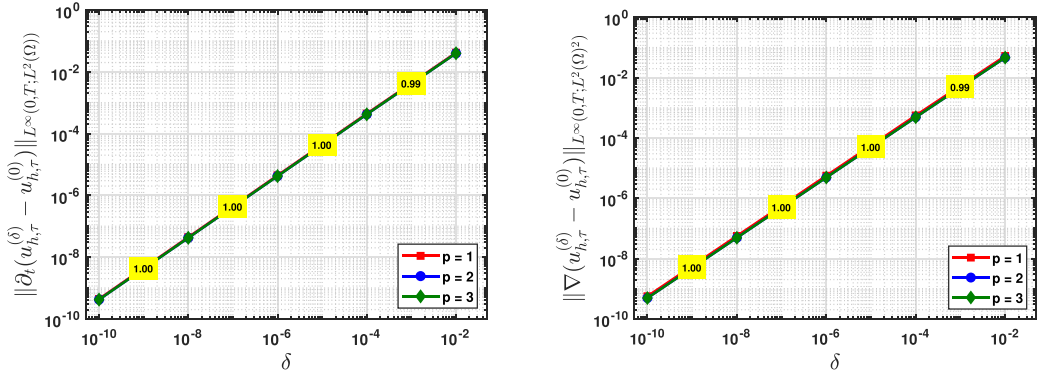


FIG. 6. δ -convergence (in \log - \log scale) for the problem in Section 7.6.1 with $f = 0$ and initial data in (7.6).

7.6 δ -convergence

We now validate the δ -convergence of the discrete solution to the inviscid discrete solution proven in Theorem 6.1. To do so, we just have to compute the difference $u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)}$, so it is not necessary to know the exact solution of the problem.

7.6.1 Homogeneous source term. We first consider the Westervelt equation (1.1) on the space–time domain $Q_T = (0, 1)^2 \times (0, 1)$ with $k = 0.3$, $c = 1$, homogeneous Dirichlet boundary conditions and source term, and initial data given by

$$u_0(x, y) = 10^{-2} \sin(\pi x) \sin(\pi y) \quad \text{and} \quad u_1(x, y) = \sin(\pi x) \sin(\pi y). \tag{7.6}$$

We set a space–time mesh with $h = \sqrt{2}\tau \approx 1.41 \times 10^{-1}$, and use approximations in time of degree $q = 4$. In Fig. 6, we show (in \log - \log scale) the decay of the following errors for $\delta = 10^{-2i}$, $i = 1, \dots, 5$ and $p = 1, 2, 3$:

$$\|\partial_t(u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)})\|_{L^\infty(0,T;L^2(\Omega))} \quad \text{and} \quad \|\nabla(u_{h,\tau}^{(\delta)} - u_{h,\tau}^{(0)})\|_{L^\infty(0,T;L^2(\Omega)^2)}, \tag{7.7}$$

where we observe the expected linear convergence in δ predicted by Theorem 6.1.

7.6.2 Nonhomogeneous source term. Finally, we consider the Westervelt equation (1.1) on the space–time domain $Q_T = (0, 1) \times (0, 2 \times 10^{-4})$, with $k = -10$, $c = 2000$, homogeneous Dirichlet boundary conditions and initial data, and a source term given by

$$f(x, y, t) = \frac{a}{\sqrt{\sigma}} \exp(-\alpha t) \exp\left(-\frac{(x - 1/2)^2 + (y - 1/2)^2}{2\sigma^2}\right), \tag{7.8}$$

with $a = 400$, $\alpha = 5 \times 10^4$ and $\sigma = 3 \times 10^{-2}$.

facet of the work is that the *a priori* error analysis guarantees that the estimates are robust with respect to the vanishing dissipation parameter δ . Furthermore, the asymptotic behavior of the exact problem is preserved in the singular limit.

This work can serve as a starting point for developing a space–time finite element methodology for different problems arising in the context of nonlinear acoustics. Extending the framework developed here to refined models of nonlinear acoustics, such as the Kuznetsov or the Blackstock equation, is of interest. Additionally, multiphysics interactions of nonlinear acoustic waves lie at the core of many developing applications of ultrasound, and their simulation could strongly benefit from the findings made here.

Acknowledgements

We thank the reviewers for their detailed comments, which have led to an improved version of the manuscript.

Funding

Italian Ministry of University and Research through the project PRIN2020 ‘Advanced polyhedral discretizations of heterogeneous PDEs for multiphysics problems’ and from the INdAM-GNCS through the project CUP E53C23001670001 to S.G.

REFERENCES

- ANTONIETTI, P. F., MAZZIERI, I., DAL SANTO, N. & QUARTERONI, A. (2018) A high-order discontinuous Galerkin approximation to ordinary differential equations with applications to elastodynamics. *IMA J. Numer. Anal.*, **38**, 1709–1734.
- BALES, L. & LASIECKA, I. (1994) Continuous finite elements in space and time for the nonhomogeneous wave equation. *Comput. Math. Appl.*, **27**, 91–102.
- BALES, L. & LASIECKA, I. (1995) Negative norm estimates for fully discrete finite element approximations to the wave equation with nonhomogeneous L_2 Dirichlet boundary data. *Math. Comp.*, **64**, 89–115.
- BRENNER, S. C. & SCOTT, L. R. (2008) *The Mathematical Theory of Finite Element Methods, vol. 15 of Texts in Applied Mathematics*, 3rd edn. New York: Springer.
- DI PIETRO, D. A. & ERN, A. (2012) *Mathematical Aspects of Discontinuous Galerkin Methods, vol. 69 of Mathématiques & Applications (Berlin)*. Berlin, Heidelberg: Springer.
- DONG, Z., MASCOTTO, L. & WANG, Z. (2024) *A priori* and *a posteriori* error estimates of a DG-CG method for the wave equation in second order formulation. arXiv:2411.03264.
- DÖRFLER, W., GERNER, H. & SCHNAUBELT, R. (2016) Local well-posedness of a quasilinear wave equation. *Appl. Anal.*, **95**, 2110–2123.
- DÖRICH, B. (2025) Strong norm error bounds for quasilinear wave equations under weak CFL-type conditions. *Found. Comput. Math.*, **25**, 303–350.
- DÖRICH, B. & NIKOLIĆ, V. (2024) Robust fully discrete error bounds for the Kuznetsov equation in the inviscid limit. arXiv preprint, arXiv:2401.06492.
- ERN, A. & GUERMOND, J.-L. (2021) *Finite Elements I—Approximation and Interpolation, vol. 72 of Texts in Applied Mathematics*. Cham: Springer.
- FRENCH, D. A. (1993) A space-time finite element method for the wave equation. *Comput. Methods Appl. Mech. Engrg.*, **107**, 145–157.
- FRENCH, D. A. & PETERSON, T. E. (1996) A continuous space-time finite element method for the wave equation. *Math. Comp.*, **65**, 491–506.
- GÓMEZ, S. (2025a) *DG-CG-Westervelt*. <https://github.com/SergioG5/DG-CG-Westervelt>.
- GÓMEZ, S. (2025b) A variational approach to the analysis of the continuous space–time FEM for the wave equation. *To appear in Math. Comp.* <https://doi.org/10.1090/mcom/4147>.

- GÓMEZ, S. & MELIANI, M. (2025) Asymptotic-preserving hybridizable discontinuous Galerkin method for the Westervelt quasilinear wave equation. *ESAIM Math. Model. Numer. Anal.*, **59**, 613–641.
- HOCHBRUCK, M. & MAIER, B. (2022) Error analysis for space discretizations of quasilinear wave-type equations. *IMA J. Numer. Anal.*, **42**, 1963–1990.
- HULBERT, G. M. & HUGHES, T. J. R. (1990) Space-time finite element methods for second-order hyperbolic equations. *Comput. Methods Appl. Mech. Engrg.*, **84**, 327–348.
- JOHNSON, C. (1993) Discontinuous Galerkin finite element methods for second order hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, **107**, 117–129.
- KALTENBACHER, B. & LASIECKA, I. (2009) Global existence and exponential decay rates for the Westervelt equation. *Discrete Contin. Dyn. Syst. Ser. S*, **2**, 503–523.
- KALTENBACHER, B. & NIKOLIĆ, V. (2022) Parabolic approximation of quasilinear wave equations with applications in nonlinear acoustics. *SIAM J. Math. Anal.*, **54**, 1593–1622.
- KALTENBACHER, M. (2007) *Numerical Simulation of Mechatronic Sensors and Actuators*, vol. 2. Berlin, Heidelberg: Springer.
- MAIER, B. (2023) Error analysis for full discretizations of quasilinear wave-type equations with two variants of the implicit midpoint rule. *IMA J. Numer. Anal.*, **43**, 1149–1180.
- MAKRIDAKIS, C. G. (1993) Finite element approximations of nonlinear elastic waves. *Math. Comp.*, **61**, 569–594.
- MELIANI, M. & NIKOLIĆ, V. (2024) Mixed approximation of nonlinear acoustic equations: well-posedness and *a priori* error analysis. *Appl. Numer. Math.*, **198**, 94–111.
- NIKOLIĆ, V. (2025) Asymptotic-preserving finite element analysis of Westervelt-type wave equations. *Anal. Appl.*, **23**, 577–605.
- NIKOLIĆ, V. & WOHLMUTH, B. (2019) *A priori* error estimates for the finite element approximation of Westervelt’s quasi-linear acoustic wave equation. *SIAM J. Numer. Anal.*, **57**, 1897–1918.
- ORTNER, C. & SÜLI, E. (2007) Discontinuous Galerkin finite element approximation of nonlinear second-order elliptic and hyperbolic systems. *SIAM J. Numer. Anal.*, **45**, 1370–1397.
- SCHWAB, C. (1998) *p- and Hp-Finite Element Methods. Numerical Mathematics and Scientific Computation*. New York: The Clarendon Press, Oxford University Press. Theory and applications in solid and fluid mechanics.
- SHAO, A. (2022) Discontinuous Galerkin discretization in time of systems of second-order nonlinear hyperbolic equations. *ESAIM Math. Model. Numer. Anal.*, **56**, 2255–2296.
- WALKINGTON, N. J. (2014) Combined DG-CG time stepping for wave equations. *SIAM J. Numer. Anal.*, **52**, 1398–1417.
- WESTERVELT, P. J. (1963) Parametric acoustic array. *J. Acoust. Soc. Am.*, **35**, 535–537.

Appendix A, Tools used in the stability analysis

A.1 Orthogonal projection and trace inequalities

LEMMA A1 (Stability of Π_r^t , see Ern & Guermond (2021, Thm. 18.16)). Let $r \in \mathbb{N}$ and $s \in [1, \infty]$. There exists a positive constant C_S independent of τ such that, for any $I_n \in \mathcal{T}_\tau$, it holds

$$\|\Pi_r^t \psi\|_{L^s(I_n)} \leq C_S \|\psi\|_{L^s(I_n)} \quad \forall \psi \in L^s(I_n).$$

LEMMA A2 (Estimates for Π_r^t , see Ern & Guermond (2021, Thm. 18.16)). Let $r \in \mathbb{N}$. There exists a positive constant C_Π independent of τ such that, for any $I_n \in \mathcal{T}_\tau$ and any sufficiently smooth function ψ defined on I_n , it holds

$$\|(\text{Id} - \Pi_r^t)\psi\|_{W_p^m(I_n)} \leq C_\Pi \tau_n^{s-m} |\psi|_{W_p^s(I_n)} \quad s, m \in \mathbb{N}, m \leq s \leq r + 1, p \in [1, \infty].$$

LEMMA A3 (Trace inequalities, see Di Pietro & Ern (2012, Lemma 1.46)). Let $I_n \in \mathcal{T}_\tau$. There exists a positive constant C_{tr} independent of τ such that

$$|\psi(t_{n-1})|^2 + |\psi(t_n)|^2 \leq C_{\text{tr}}(\tau_n^{-1} \|\psi\|_{L^2(I_n)}^2 + \tau_n \|\psi'\|_{L^2(I_n)}^2) \quad \forall \psi \in H^1(I_n).$$

Moreover, there exists a positive constant C_{tr}^* independent of τ such that

$$|p_q(t_{n-1})|^2 + |p_q(t_n)|^2 \leq C_{\text{tr}}^* \tau_n^{-1} \|p_q\|_{L^2(I_n)}^2 \quad \forall p_q \in \mathbb{P}^q(I_n). \quad (\text{A.1})$$

A.2 Polynomial inverse estimates

LEMMA A4 (Inverse estimates, see Brenner & Scott (2008, Lemma 4.5.3)). There exists a positive constant C_{inv} independent of h and τ such that

$$\|\nabla w_h\|_{L^2(Q_T)^d} \leq C_{\text{inv}} h_{\min}^{-1} \|w_h\|_{L^2(Q_T)} \quad \forall w_h \in L^2(0, T; \mathcal{V}_h^p), \quad (\text{A.2a})$$

$$\|w_h\|_{L^\infty(Q_T)^d} \leq C_{\text{inv}} h_{\min}^{-\frac{d}{2}} \|w_h\|_{L^2(Q_T)} \quad \forall w_h \in L^2(0, T; \mathcal{V}_h^p), \quad (\text{A.2b})$$

and, for $n = 1, \dots, N$,

$$\|\partial_t w_\tau\|_{L^r(Q_n)} \leq C_{\text{inv}} \tau_n^{-1} \|w_\tau\|_{L^r(Q_n)} \quad \forall w_\tau \in \mathbb{P}^q(I_n; L^r(\Omega)), \quad r \in [1, \infty], \quad (\text{A.2c})$$

$$\|\partial_t w_\tau\|_{L^2(Q_n)} \leq C_{\text{inv}} \tau_n^{-\frac{1}{2}} \|w_\tau\|_{L^\infty(I_n; L^2(\Omega))} \quad \forall w_\tau \in \mathbb{P}^q(I_n; L^\infty(\Omega)). \quad (\text{A.2d})$$

LEMMA A5. Let $I_n \in \mathcal{T}_\tau$. The following bound holds:

$$\|w_\tau\|_{L^\infty(I_n; L^2(\Omega))} \leq (1 + C_{\text{inv}}) \tau_n^{-\frac{1}{2}} \|w_\tau\|_{L^2(Q_n)} \quad \forall w_\tau \in \mathbb{P}^q(I_n; L^2(\Omega)),$$

where C_{inv} is the constant in the polynomial inverse estimate (A.2c).

Proof. The result follows by combining the inequality $\|w\|_{L^\infty(a,b)} \leq |b-a|^{-1/2} \|w\|_{L^2(a,b)} + |b-a|^{1/2} \|w'\|_{L^2(a,b)}$ (see Ern & Guermond (2021, Eq. (1.9))) with the polynomial inverse estimate (A.2c). \square

Appendix B, Tools used in the convergence analysis

B.1 Properties of the temporal projection

Next lemmas concern the stability and approximation properties of the projection \mathcal{P}_τ in Definition 4.3; see also Dong *et al.* (2024, Lemma 2.4 and Prop. 2.5), where q -explicit stability and approximation properties of \mathcal{P}_τ are derived.

LEMMA B1 (Stability of \mathcal{P}_τ , see Walkington (2014, Lemma 5.2)). Let $q \in \mathbb{N}$ with $q \geq 2$. For any $s \in [1, \infty)$, there exists a positive constant, depending only on q such that, for all $v \in C^1(0, T)$, the

following bounds hold:

$$\begin{aligned}\|\partial_t \mathcal{P}_\tau v\|_{L^s(0,T)} &\leq C_{\text{stab}}^{\mathcal{P}_\tau} T^{1/s} \|\partial_t v\|_{C^0([0,T])}, \\ \|\mathcal{P}_\tau v\|_{L^s(0,T)} &\leq C_{\text{stab}}^{\mathcal{P}_\tau} T^{1/s} (\|v\|_{C^0([0,T])} + \tau \|\partial_t v\|_{C^0([0,T])}).\end{aligned}$$

For $s = \infty$, the above bounds hold with $T^{1/s}$ replaced by 1.

LEMMA B2 (Estimates for \mathcal{P}_τ , see Walkington (2014, Lemma 5.2)). Let $q \in \mathbb{N}$ with $q \geq 2$. For any $s \in [1, \infty]$, the following estimates hold for all $v \in W_s^{m+1}(0, T)$ with $1 \leq m \leq q$: there exists a positive constant $C_{\text{approx}}^{\mathcal{P}_\tau}$ independent of τ such that

$$\begin{aligned}\|(\text{Id} - \mathcal{P}_\tau)v\|_{L^s(0,T)} &\leq C_{\text{approx}}^{\mathcal{P}_\tau} \tau^{m+1} |v|_{W_s^{m+1}(0,T)}, \\ \|\partial_t(\text{Id} - \mathcal{P}_\tau)v\|_{L^s(0,T)} &\leq C_{\text{approx}}^{\mathcal{P}_\tau} \tau^m |v|_{W_s^{m+1}(0,T)}.\end{aligned}$$

For any Banach space $(Z, \|\cdot\|_Z)$ with $Z \subseteq L^1(\Omega)$, the definition of \mathcal{P}_τ can be extended to functions in $C^1(0, T; Z)$ by requiring that (4.6) holds almost everywhere in Ω .

B.2 Properties of the Ritz projection and the Lagrange interpolant

We next recall certain approximation properties of the Ritz projection \mathcal{R}_h .

LEMMA B3 (Estimates for \mathcal{R}_h , see Brenner & Scott (2008, Thms 5.4.4 and 5.4.8)). Let $p \in \mathbb{N}$ with $p \geq 1$, and let Ω satisfy Assumption 2. Then, there exists a positive constant $C_{\text{approx}}^{\mathcal{R}_h}$ independent of h such that

$$\|(\text{Id} - \mathcal{R}_h)z\|_{L^2(\Omega)} \leq C_{\text{approx}}^{\mathcal{R}_h} h^{\ell+1} |z|_{H^{\ell+1}(\Omega)} \quad \forall z \in H^{\ell+1}(\Omega) \cap H_0^1(\Omega), \quad 0 \leq \ell \leq p, \quad (\text{B.1a})$$

$$\|\nabla(\text{Id} - \mathcal{R}_h)z\|_{L^2(\Omega)^d} \leq C_{\text{approx}}^{\mathcal{R}_h} h^\ell |z|_{H^{\ell+1}(\Omega)} \quad \forall z \in H^{\ell+1}(\Omega) \cap H_0^1(\Omega), \quad 0 \leq \ell \leq p. \quad (\text{B.1b})$$

We denote the Lagrange interpolant operator by $\mathcal{I}_h : C^0(\overline{\Omega}) \rightarrow \mathcal{V}_h^p$. The following stability and approximation properties follow from Brenner & Scott (2008, Thm. 4.4.20 in Ch. 4).

LEMMA B4 (Stability of \mathcal{I}_h). Let $p \in \mathbb{N}$ with $p \geq 1$. Then, there exists a positive constant $C_{\text{stab}}^{\mathcal{I}_h}$ independent of h such that

$$\|\mathcal{I}_h v\|_{L^\infty(\Omega)} \leq C_{\text{stab}}^{\mathcal{I}_h} \|v\|_{L^\infty(\Omega)} \quad \forall v \in C^0(\overline{\Omega}).$$

LEMMA B5 (Estimates for \mathcal{I}_h). Let $p \in \mathbb{N}$ with $p \geq 1$. Then, there exists a positive constant $C_{\text{approx}}^{\mathcal{I}_h}$ independent of h such that

$$\|\mathcal{I}_h v - v\|_{L^2(\Omega)} \leq C_{\text{approx}}^{\mathcal{I}_h} h^{\ell+1} |v|_{H^{\ell+1}(\Omega)} \quad \forall v \in H^{\ell+1}(\Omega), \quad (d-2)/2 < \ell \leq p.$$

We finish this section with some useful stability bounds for the Ritz projection.

LEMMA B6 (Stability of \mathcal{R}_h). Let $p \in \mathbb{N}$ with $p \geq 1$. There exists a positive constant $C_{\text{stab}}^{\mathcal{R}_h}$ independent of h such that

$$\|\mathcal{R}_h z\|_{L^2(\Omega)} \leq C_{\text{stab}}^{\mathcal{R}_h} \|\nabla z\|_{L^2(\Omega)^d} \quad \forall z \in H_0^1(\Omega), \quad (\text{B.2})$$

$$\|\mathcal{R}_h z\|_{L^\infty(\Omega)} \leq C_{\text{stab}}^{\mathcal{R}_h} \|z\|_{H^2(\Omega)} \quad \forall z \in H^2(\Omega) \cap H_0^1(\Omega). \quad (\text{B.3})$$

Proof. Bound (B.2) follows from the standard Poincaré inequality (see, e.g., Brenner & Scott, 2008, Prop. 5.3.2).

As for bound (B.3), we use the triangle inequality, the inverse estimate (A.2b), and Lemmas B4 and B5 to obtain

$$\begin{aligned} \|\mathcal{R}_h z\|_{L^\infty(\Omega)} &\leq \|\mathcal{R}_h z - \mathcal{I}_h z\|_{L^\infty(\Omega)} + \|\mathcal{I}_h z\|_{L^\infty(\Omega)} \\ &\lesssim h^{-d/2} \|\mathcal{R}_h z - \mathcal{I}_h z\|_{L^2(\Omega)} + \|z\|_{L^\infty(\Omega)} \\ &\lesssim h^{2-d/2} \|z\|_{H^2(\Omega)} + \|z\|_{L^\infty(\Omega)}, \end{aligned}$$

which, combined with the continuous Sobolev embedding $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$, completes the proof. \square