

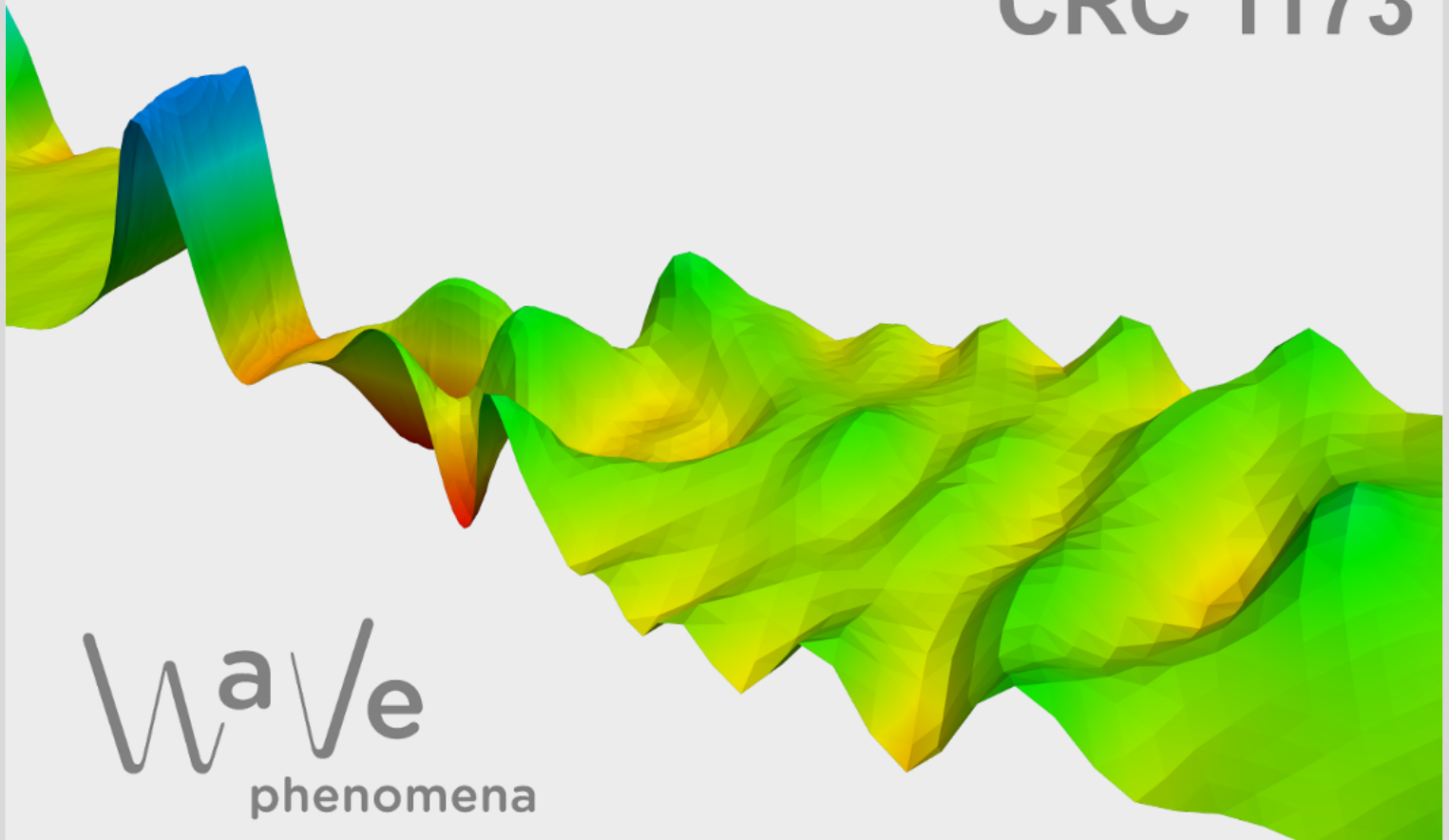
Conservative finite element methods for Kirchhoff-type equations

Benjamin Dörich, Felix Krumbiegel, Jemal Rogava,
Andreas Rupp, Zurab Vashakidze

CRC Preprint 2026/59, October 2026

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Participating universities



Funded by



Conservative Finite Element Methods for Kirchhoff-Type Equations

Benjamin Dörich 

Department of Mathematics, Saarland
University,
Campus E1.1, 66123 Saarbrücken, Germany
benjamin.doerich@uni-saarland.de

Jemal Rogava 

Ivane Javakhishvili Tbilisi State University, Ilia
Vekua Institute of Applied Mathematics,
11 University Street, 0186, Tbilisi, Georgia
jemal.rogava@tsu.ge

Zurab Vashakidze 

School of Science and Technology, The
University of Georgia (UG),
77 Merab Kostava Street, 0171, Tbilisi, Georgia;
Ilia Vekua Institute of Applied Mathematics,
11 University Street, 0186, Tbilisi, Georgia
z.vashakidze@ug.edu.ge |
zurab.vashakidze@tsu.ge

Felix Krumbiegel 

Department of Mathematics, Saarland
University,
Campus E1.1, 66123 Saarbrücken, Germany
felix.krumbiegel@uni-saarland.de

Andreas Rupp 

Department of Mathematics, Saarland
University,
Campus E1.1, 66123 Saarbrücken, Germany
rupp.ink | andreas.rupp@uni-saarland.de

Abstract

We construct and analyse a fully discrete scheme for a Kirchhoff-type nonlinear integro-differential equation in two spatial dimensions with time-dependent coefficients. The scheme combines a symmetric three-layer time discretization of Crank–Nicolson type with a conforming finite element method of arbitrary polynomial degree p and a Picard iteration for the nonlinearity. We prove that the Picard iteration is well defined and, that the iterates converge to the Ritz projection of the exact solution measured in fractional-power norms of the discrete Laplacian. Running the same energy argument in the unshifted norms, and under the same hypotheses, we then obtain the optimal order $\mathcal{O}(h^{p+1} + \tau^2)$ in L^2 and $\mathcal{O}(h^p + \tau^2)$ in H^1 for the full error. For the homogeneous problem with constant coefficients, the exact solution of the scheme moreover preserves a discrete energy identically. Numerical experiments with manufactured solutions reproduce the predicted second order in time and the full spatial order h^{p+1} in the L^2 -norm.

Keywords: Kirchhoff-type equation; nonlinear wave equation; finite element method; three-layer scheme; Picard iteration; optimal-order error estimate.

MSC 2020: 65M60; 65M12; 65M15; 35L70; 35L90.

1 Introduction

1.1 Problem Formulation

Kirchhoff-type equations constitute an important class of nonlocal partial differential equations arising from the classical model proposed by Kirchhoff as an extension of the linear theory of vibrating strings; see [17]. In this model the tension varies according to the total stretching induced by the displacement, so that the principal part of the equation depends on a global quantity involving the solution itself, typically its Dirichlet energy. In two spatial dimensions, the equation models the transverse vibrations of an elastic membrane whose effective tension depends on its total deformation energy, a setting that arises in contemporary work on nanomechanical oscillators [35]; allowing the material parameters to depend on time reflects the variation of those properties during vibration, for instance under thermal fluctuation or strong external driving [29, 32].

We consider the nonlinear Kirchhoff-type integro-differential equation

$$\frac{\partial^2 u}{\partial t^2} - \left(\alpha(t) + \beta(t) \iint_{\Omega} |\nabla u|^2 \, dx \, dy \right) \Delta u = f, \quad (1.1)$$

posed in $[0, T]$ and a domain $\Omega \subset \mathbb{R}^2$, together with the initial and boundary conditions

$$u(0) = u_0, \quad \partial_t u(0) = v_0, \quad u(t)|_{\partial\Omega} = 0 \quad \text{for } t \in [0, T]. \quad (1.2)$$

Here u is the transverse displacement, $|\cdot|$ the Euclidean norm and ∇ the gradient. Throughout both this paper and the companion paper [19], u and f are regarded as functions of t with values in a space of functions on Ω , so that the spatial argument is never displayed. $u(t)$ is the state at time t and u_h^n the fully discrete approximation at the n -th time level. The coefficients are scalar functions of t and do carry that argument, *i.e.*, $\alpha, \beta \in C^1[0, T]$ satisfy $\alpha_{\min} > 0$ and $\beta_{\min} > 0$, where throughout $\alpha_{\min} = \min_{[0, T]} \alpha(t)$, $\alpha_{\max} = \max_{[0, T]} \alpha(t)$ and $\beta_{\min}, \beta_{\max}$ are defined correspondingly. The source term is continuous in time with values in $L^2(\Omega)$, so that the grid values $f(n\tau)$, where τ is the time step size, are defined. Section 3 imposes more regularity on the right-hand side where it is needed. Throughout this paper Ω is a bounded convex *polytopal* domain, so that it is meshed exactly by the simplicial families used below. For a general bounded convex domain, the analysis applies verbatim to a polytopal approximation $\Omega_h \subset \Omega$, at the cost of an additional geometric consistency error, or with isoparametric elements of matching degree, and we do not pursue either variant here.

1.2 Related work and contribution

Numerical methods for Kirchhoff-type equations have a long history. Christie and Sanz-Serna [8] combined a finite element discretization in space with a Crank–Nicolson scheme in time for the one-dimensional problem with constant coefficients. In [24], Peradze proposed and analyzed a numerical scheme combining a sine-based Galerkin spatial discretization with a Crank–Nicolson-type time-stepping method for the nonlinear one-dimensional homogeneous Kirchhoff string problem with constant coefficients. Gudi [15] investigated a finite-element approximation of a second-order nonlinear elliptic boundary value problem that arises as the stationary counterpart of the generalized Kirchhoff equation. More recently, Mallik [22] developed and analyzed a hybrid high-order finite element method for a nonlocal nonlinear elliptic boundary value problem of Kirchhoff type, establishing optimal-order convergence in the discrete energy norm. Rogava and Tsiklauri [26, 25] analysed symmetric three-layer semi-discrete schemes in an abstract setting and established local convergence; the one-dimensional problem with *time-dependent* coefficients was treated in [28], and related schemes appear in [27, 34, 33]. Well-posedness of the continuous problem for constant coefficients goes back to Arosio and Panizzi [2].

Beyond discretizations developed specifically for Kirchhoff equations, there has also been extensive work on the numerical analysis of quasilinear wave problems. Classical finite-element and multistep analyses for nonlinear second-order hyperbolic equations were given by Bales and Dougalis [3] and Makridakis [21]. More recently, spatial and fully discrete error analyses for general classes of quasilinear wave-type equations were developed by Hochbruck and Maier [16] and Maier [20], while strong-norm estimates under weak CFL-type conditions were obtained by Dörich [12]. Related finite-element

analyses for nonlinear acoustic wave models include Nikolić and Wohlmuth [23] and Dörich and Nikolić [13]. These works concern different, predominantly local or damped quasilinear structures and different time integrators. In contrast, the coefficient considered here is a time-dependent spatially nonlocal scalar depending on the total Dirichlet energy, and the symmetric three-layer discretization is designed to preserve an exact discrete energy in the homogeneous constant-coefficient case.

The companion paper [19] analyses the temporal discretization of (1.1)–(1.2) in two space dimensions: it establishes energy conservation for two symmetric three-layer schemes, uniform *a priori* bounds, local second-order convergence in time. The present paper takes the semi-discrete scheme analysed there as its starting point and supplies the spatial discretization. The contributions are the following.

We formulate a fully discrete scheme combining the symmetric three-layer time stepping with a conforming finite element method of arbitrary polynomial degree p and a Picard iteration for the nonlinearity. The discretization of the nonlinear coefficient leaves one free parameter, and our first result is that for the symmetric weighting in time and time-independent coefficients, each choice of that parameter has its own discrete energy, which the scheme conserves exactly on the homogeneous problem (Proposition 3.3). Further, we prove that at each time level the nonlinear scheme admits a unique solution, to which the Picard iteration converges (Lemma 4.1). We then prove convergence of the error relative to the Ritz projection of the exact solution, measured in fractional-power norms of the discrete Laplacian, locally in time (Theorem 4.4) and then over all time levels (Corollary 4.6). A single argument covers a whole range of these norms. The stronger pair is what the existence theory needs and yields $\mathcal{O}(h^{p+\frac{1}{2}} + \tau^2)$, while the weaker pair yields the optimal order. To our knowledge this step has not been taken for Kirchhoff-type equations with time-dependent coefficients. From the latter we obtain an estimate of *optimal* order for the full error $u(t_n) - u_h^n$, namely $\mathcal{O}(h^{p+1} + \tau^2)$ in L^2 and $\mathcal{O}(h^p + \tau^2)$ in H^1 (Corollary 4.7). The half-order gap is a property of the norms in which the first estimate is measured, not of the scheme. This is the rate the numerical experiments exhibit.

Two works are the natural points of comparison. The work [8] treats the one-dimensional problem with constant coefficients by a Galerkin method in space and a Crank–Nicolson discretization in time. Closer to the present setting, Dautov and Ivanova [11] construct a conservative finite element scheme for the Kirchhoff equation: an implicit two-layer scheme for the first-order formulation, which conserves a discrete energy, is realised by Newton’s method, and is reported to achieve $\mathcal{O}(h^2 + \tau^2)$ accuracy in L^2 for both unknowns in numerical tests. The present work differs from these in the following respects. We treat the case of time-dependent coefficients and employ a symmetric three-layer time-discretization scheme for the second-order formulation, coupled with a finite-element spatial discretization of arbitrary polynomial degree p . For the resulting nonlinear system, we establish the existence and uniqueness of the solution and prove the convergence of the associated Picard iteration.

The paper is organized as follows: Section 2 fixes notation, states the standing hypothesis on the continuous problem and records what [19] proves about the semi-discrete scheme. Section 3 contains the fully discrete scheme with its energy conservation, and Section 4 its convergence analysis. Section 5 reports the numerical experiments and Section 6 concludes.

2 Setting and the Semi-Discrete Scheme

We write $\mathcal{L} = -\Delta$ for the Dirichlet Laplacian on Ω , a self-adjoint positive definite operator with domain $D(\mathcal{L}) = H^2(\Omega) \cap H_0^1(\Omega)$, the last equality holds because Ω is bounded and convex. Its fractional powers $\mathcal{L}^a$ are defined through the spectral resolution. Denote with $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ the norm and inner product of $L^2(\Omega)$, where $\langle \cdot, \cdot \rangle$ also carries the $H^{-1}(\Omega) - H_0^1(\Omega)$ duality where the first argument calls for it, as noted at its first occurrence. Interpolation between $L^2(\Omega)$ and $D(\mathcal{L})$ gives the embedding

$$D(\mathcal{L}^a) \hookrightarrow H^{2a}(\Omega), \quad 0 \leq a \leq 1, \quad (2.1)$$

which is the only direction used below: it lets results quoted from the literature in the Sobolev scale be applied under the operator-domain hypotheses stated here. The reverse inclusion is false for $a > \frac{1}{4}$, since membership of $D(\mathcal{L}^a)$ encodes the homogeneous Dirichlet condition. Temporal regularity

hypotheses are therefore stated in the operator domains, spatial approximation hypotheses in the Sobolev scale.

First, we comment on the genuinely nonlinear of the two three-layer schemes considered in [19] and the main result therein. It is the one whose fully discrete counterpart we analyse. On the uniform temporal grid $t_n = n\tau$, $\tau = \frac{T}{N}$ with the number of time steps $N \in \mathbb{N}$, the semi-discrete scheme

$$\frac{u^{n+1} - 2u^n + u^{n-1}}{\tau^2} + \frac{1}{2} \left(\alpha^n + \frac{\beta^n}{2} (\|\nabla u^{n+1}\|^2 + \|\nabla u^{n-1}\|^2) \right) \mathcal{L}(u^{n+1} + u^{n-1}) = f^n \quad (2.2)$$

is given for $n = 1, \dots, N-1$, with $\alpha^n = \alpha(t_n)$, $\beta^n = \beta(t_n)$ and $f^n = f(t_n)$, and with the starting values

$$u^0 = u_0, \quad u^1 = u_0 + \tau v_0 + \frac{\tau^2}{2} w_0, \quad w_0 = f^0 - q^0 \mathcal{L}u_0, \quad q^0 = \alpha^0 + \beta^0 \|\nabla u_0\|^2. \quad (2.3)$$

Proposition 2.1 ([19, Thm. 5.4, Cor. 5.5 and Cor. 5.6]). *Under the following regularity hypotheses $u_0 \in D(\mathcal{L}^2)$, $v_0 \in D(\mathcal{L})$, $f \in C([0, T]; D(\mathcal{L}^{\frac{1}{2}}))$ with $f^0 \in D(\mathcal{L})$, three time derivatives of u with $\partial_t^3 u$ Lipschitz in t , and the corresponding hypotheses on the second spatial derivatives, the time-discrete solution of (2.2) converges to the exact solution with order two in time, locally on an interval $[0, \bar{T}]$ whose length is determined by the data. The same order holds for the central-difference approximation of the first time derivative.*

Remark 2.2 (What is not imported from the companion paper). **Proposition 2.1** is recorded to place the present paper next to its companion, and is never invoked below. Neither is the existence and uniqueness of the time-discrete trajectory. Nothing below requires either. The hypotheses are stated with each result and not inherited, in particular the existence of the fully discrete solution is proved here and not assumed.

We close this section with an existence result for the PDE (1.1)–(1.2). The analysis below mainly uses the following assumptions, and only for convergence, it is necessary that higher regularity in time and space is assumed.

Assumption 2.3 (Existence of a solution). Let $u_0 \in D(\mathcal{L}^{\frac{3}{4}})$, $v_0 \in D(\mathcal{L}^{\frac{1}{4}})$ and $f \in C^0([0, T_{\max}]; D(\mathcal{L}^{\frac{1}{4}}))$ for some possibly infinite final time $T_{\max}$. Then there exists a unique solution u of (1.1)–(1.2) with

$$u \in C^0([0, T]; D(\mathcal{L}^{\frac{3}{4}})) \cap C^1([0, T]; D(\mathcal{L}^{\frac{1}{4}}))$$

for a final time $T < T^{**} \leq T_{\max}$.

For constant α and β a result of this type is proved in [2, Thm. 2.1], where the blow-up time T^{**} is also specified. For time-dependent α and β the coefficient $\alpha(t) + \beta(t)\|\nabla u(t)\|^2$ is nonautonomous and that theorem does not apply, so there we take the statement as a hypothesis on the continuous problem. It is a hypothesis about the exact solution only; it is not used to construct anything, so no circularity arises. As in [19], we do not claim that the hypotheses of that theorem — in particular whatever regularity it imposes on the domain — are met by every domain admitted here; where they are not, the statement above is to be read as a hypothesis in the constant-coefficient case as well.

3 Extension to full discretization and energy conservation

3.1 Preliminaries

Following the assumption and the existence of a solution **Assumption 2.3** we may define

$$\begin{aligned} M_a &= 2\|\mathcal{L}^a u\|_{L^\infty(L^2)}, & 0 \leq a \leq \frac{3}{4}, \\ M'_a &= 2\|\mathcal{L}^a \partial_t u\|_{L^\infty(L^2)}, & C_a = \|\mathcal{L}^a f\|_{L^\infty(L^2)}, & 0 \leq a \leq \frac{1}{4} \end{aligned}$$

for the solution u and source term f .

The balls around the solution play a key role in the analysis, as we prove that the fully discrete solution lies inside these balls at all times, which is then used to show unique existence and convergence of the solution.

Let $\Omega \subset \mathbb{R}^2$ be an open, bounded, and convex domain (hence Lipschitz), as assumed in the introduction, and *polytopal*, that is, a convex polygon, and $\mathcal{T}_h$ be a regular decomposition of Ω into quasi-uniform triangles. Polytopality makes the decomposition into straight-sided simplices exact, so that $V_h \subset H_0^1(\Omega)$ holds without a variational crime. Convexity secures the H^2 -regularity behind the embedding (2.1). Further, let for any polynomial degree $p \geq 1$ the space of piecewise (*wrt.* the mesh) polynomials up to degree p with zero boundary conditions be denoted by

$$V_h := \{v \in C(\bar{\Omega}) \mid v|_E \in \mathbb{P}_p(E) \text{ for any element } E \in \mathcal{T}_h\} \cap H_0^1(\Omega).$$

Now, we define the discrete operator $\mathcal{L}_h: V_h \rightarrow V_h$ by

$$\langle \mathcal{L}_h v_h, \varphi_h \rangle = \langle \nabla v_h, \nabla \varphi_h \rangle \quad \text{for all } v_h, \varphi_h \in V_h. \quad (3.1)$$

The operator $\mathcal{L}_h$ plays the role of $\mathcal{L}$ for the discrete setting. This connection can be observed in the following. Let P_h denote the L^2 -orthogonal projection onto V_h , let R_h denote the Ritz projection onto V_h , given by

$$\langle P_h v, \varphi_h \rangle = \langle v, \varphi_h \rangle, \quad \langle \nabla(R_h v), \nabla \varphi_h \rangle = \langle \nabla v, \nabla \varphi_h \rangle, \quad \text{for all } \varphi_h \in V_h,$$

and let J_h denote a quasi-interpolation operator onto V_h : for the lowest order case $p = 1$ a Clément-type operator [9], and for the higher-order cases an Ern–Guermond-type operator [14] with the homogeneous boundary condition built in. Then, we have that the discrete operator $\mathcal{L}_h$ coincides with $\mathcal{L}$ weakly on V_h , *i.e.*, for any $v \in H_0^1(\Omega)$ we have that (with $\langle \mathcal{L}v, \varphi_h \rangle$ understood as the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$), which reduces to the L^2 -inner product when $v \in D(\mathcal{L})$)

$$\langle \mathcal{L}v, \varphi_h \rangle = \langle \nabla v, \nabla \varphi_h \rangle = \langle \nabla R_h v, \nabla \varphi_h \rangle = \langle \mathcal{L}_h R_h v, \varphi_h \rangle, \quad \text{for all } \varphi_h \in V_h, \quad (3.2)$$

and in particular, for $v \in D(\mathcal{L})$,

$$\mathcal{L}_h R_h v = P_h \mathcal{L}v. \quad (3.3)$$

We collect next the approximation and stability properties of these three operators that the analysis uses. Their proofs and the references for them are given in [Appendix A.2](#). Throughout, the constants implied by $\lesssim$ are independent of h , of the time step τ and of the time level n . For $v \in H^s(\Omega) \cap H_0^1(\Omega)$ with $1 \leq s \leq p+1$ and $\Pi \in \{P_h, R_h, J_h\}$,

$$\|v - \Pi v\| + h \|\nabla(v - \Pi v)\| \lesssim h^s |v|_{H^s(\Omega)}, \quad \text{and} \quad \|\mathcal{L}^{\frac{1}{4}}(v - \Pi v)\| \lesssim h^{s-\frac{1}{2}} |v|_{H^s(\Omega)}. \quad (3.4)$$

The analysis further compares discrete fractional norms of projected functions with continuous ones. For $v \in D(\mathcal{L}^a)$,

$$\|\mathcal{L}_h^a P_h v\| + \|\mathcal{L}_h^a J_h v\| \lesssim \|\mathcal{L}^a v\|, \quad 0 \leq a \leq \frac{1}{2}, \quad (3.5)$$

$$\|\mathcal{L}_h^a R_h v\| \leq \|\mathcal{L}^a v\|, \quad \frac{1}{2} \leq a \leq 1. \quad (3.6)$$

The Ritz projection is absent from (3.5) because it is defined only on $H_0^1(\Omega) = D(\mathcal{L}^{\frac{1}{2}})$, which does not contain $D(\mathcal{L}^a)$ for $a < \frac{1}{2}$. Where a discrete fractional norm of $R_h v$ below $a = \frac{1}{2}$ is needed, we use instead the uniform spectral bounds, valid for $\varphi_h \in V_h$, and $0 \leq a \leq b \leq 1$,

$$\|\mathcal{L}_h^a \varphi_h\| \leq \lambda_1^{a-b} \|\mathcal{L}_h^b \varphi_h\|, \quad (3.7)$$

$$\|\mathcal{L}_h^b \varphi_h\| \leq C_{\text{inv}}^{2(b-a)} h^{-2(b-a)} \|\mathcal{L}_h^a \varphi_h\|, \quad (3.8)$$

where $\lambda_1 > 0$ is the smallest eigenvalue of $\mathcal{L}$ and C_{inv} is the constant of the inverse inequality $\|\nabla \varphi_h\| \leq C_{\text{inv}} h^{-1} \|\varphi_h\|$, which depends only on the shape-regularity and quasi-uniformity of the mesh family and on p . From the estimate (3.7) with $0 \leq a \leq b = \frac{1}{2}$ follows using (3.6) for $v \in H_0^1(\Omega)$ that

$$\|\mathcal{L}_h^a R_h v\| \leq \lambda_1^{a-\frac{1}{2}} \|\mathcal{L}^{\frac{1}{2}} v\|, \quad (3.9)$$

which replaces (3.6) below $a = \frac{1}{2}$.

Finally, we present a lemma that will be used frequently in the analysis below. The proof is moved to [Appendix A.3](#) as it consists mostly of standard ideas. Let for any fully discrete sequence $v_h^{m+1}, v_h^m, v_h^{m-1} \in V_h$ and $\theta \in [\frac{1}{4}, \frac{1}{2}]$ the weighted average be defined as

$$v_h^{m+1;\theta} := \theta v_h^{m+1} + (1 - 2\theta) v_h^m + \theta v_h^{m-1}.$$

Lemma 3.1 (Coercivity of the displacement energy). *Let $\theta \in [\frac{1}{4}, \frac{1}{2}]$, $0 \leq a \leq \frac{1}{2}$, and $\tau > 0$ be some time step size. Define for any sequence $u_h^{m+1}, u_h^m, u_h^{m-1} \in V_h$ the functions $v_h^k := \mathcal{L}_h^{a+\frac{1}{2}} u_h^k$ for $k = m+1, m, m-1$. Then for $\mathcal{W}_h^m = \frac{1}{4} \|v_h^m + v_h^{m-1}\|^2 + (\theta - \frac{1}{4}) \|v_h^m - v_h^{m-1}\|^2$ we have that $\mathcal{W}_h^m \geq 0$, and it holds*

$$\langle v_h^{m+1;\theta}, v_h^{m+1} - v_h^{m-1} \rangle = \mathcal{W}_h^{m+1} - \mathcal{W}_h^m, \quad (3.10)$$

$$\mathcal{W}_h^m \leq \frac{1}{2} (\|v_h^m\|^2 + \|v_h^{m-1}\|^2), \quad (3.11)$$

$$\|v_h^m\|^2 \leq \frac{2}{4\theta-1} \mathcal{W}_h^m, \quad \theta > \frac{1}{4}, \quad (3.12)$$

$$\|v_h^m\|^2 \leq 2\mathcal{W}_h^m + \frac{C_{\text{inv}}^2}{2} \tau^2 h^{-2} \|\mathcal{L}_h^a \partial_\tau u_h^m\|^2, \quad \theta \geq \frac{1}{4}. \quad (3.13)$$

Two consequences of Lemma 3.1 are used so often that we record them separately. They are the two one-sided estimates the energy argument needs, and they are not an equivalence. At $\theta = \frac{1}{4}$ the displacement energy $\mathcal{W}_h^m$ alone does not control $\|v_h^m\|^2$, by the sawtooth of Remark A.1, and it is the combined displacement-plus-kinetic quantity $\mathcal{F}_h^m$ defined below that does.

Corollary 3.2 (Energy control). *In the setting of Lemma 3.1, put $\mathcal{F}_h^m := \|\mathcal{L}_h^a \partial_\tau u_h^m\|^2 + \mathcal{W}_h^m$. Then*

$$\frac{1}{4} \|\mathcal{L}_h^{a+\frac{1}{2}} (u_h^m + u_h^{m-1})\|^2 \leq \mathcal{W}_h^m, \quad \theta \geq \frac{1}{4}, \quad (3.14)$$

$$\|\mathcal{L}_h^{a+\frac{1}{2}} u_h^m\|^2 \leq C_\theta \mathcal{F}_h^m, \quad C_\theta = \begin{cases} \frac{2}{4\theta-1}, & \theta > \frac{1}{4}, \\ \max\{2, \frac{C_{\text{inv}}^2 \kappa^2}{2}\}, & \theta = \frac{1}{4}, \end{cases} \quad (3.15)$$

the second at $\theta = \frac{1}{4}$ under the mesh condition $\tau \leq \kappa h$.

Proof. (3.14) is the definition of $\mathcal{W}_h^m$ together with $\theta \geq \frac{1}{4}$, and (3.15) is (3.12) for $\theta > \frac{1}{4}$, respectively (3.13) with $\tau^2 h^{-2} \leq \kappa^2$ at $\theta = \frac{1}{4}$. $\blacksquare$

3.2 Fully discrete scheme and energy conservation

Now we may formulate the *fully discrete nonlinear scheme*: We rewrite the nonlinearity as follows

$$\alpha^n + \beta^n (\gamma \|\nabla v\|^2 + (1 - 2\gamma) \|\nabla u_h^n\|^2 + \gamma \|\nabla u_h^{n-1}\|^2) =: \Gamma_h^{n;\gamma}(v).$$

Given initial conditions $u_h^0, u_h^1 \in V_h$, we seek a sequence of functions $\{u_h^{n+1}\}_{n=1}^{N-1} \subset V_h$ that solves the following nonlinear equation

$$\partial_\tau^2 u_h^{n+1} + \Gamma_h^{n;\gamma}(u_h^{n+1}) \mathcal{L}_h u_h^{n+1;\theta} = f_h^{n+1;\eta}, \quad (3.16)$$

where the weighted average $u_h^{n+1;\theta}$ is given as above, and analogously $f_h^{n+1;\eta} := \eta f_h^{n+1} + (1 - 2\eta) f_h^n + \eta f_h^{n-1}$, with $\gamma, \eta \in [0, \frac{1}{2}]$ and $\theta \in [\frac{1}{4}, \frac{1}{2}]$, and $f_h^n = J_h f(n\tau)$ with an interpolation operator J_h . We may use for the lowest order case $p = 1$ a Cl  ment-type quasi-interpolation operator, and for the higher-order cases an Ern–Guermond-type quasi-interpolation operator. Note that for $\gamma = 0$, the scheme reduces to the locally linear scheme from [19]. In similar fashion, the scheme in [19] may also be extended by considering parameters γ, η , and θ as above.

Before turning to the convergence analysis we record that the scheme enjoys the following conservation property.

Proposition 3.3 (Fully discrete energy conservation). *Let $\theta = \frac{1}{2}$, α and β be constant, and $f \equiv 0$. Further, let $\{u_h^n\}_{n=0}^N \subset V_h$ solve the nonlinear scheme (3.16) exactly, and let V_h be any conforming subspace of $H_0^1(\Omega)$. Define the fully discrete energy*

$$\begin{aligned} \mathcal{E}_h^{n;\gamma} &:= \frac{1}{2} \|\partial_\tau u_h^n\|^2 + \frac{\alpha}{4} (\|\nabla u_h^{n-1}\|^2 + \|\nabla u_h^n\|^2) \\ &\quad + \frac{\beta}{4} [\|\nabla u_h^{n-1}\|^2 \|\nabla u_h^n\|^2 + \gamma (\|\nabla u_h^n\|^2 - \|\nabla u_h^{n-1}\|^2)^2]. \end{aligned} \quad (3.17)$$

Then $\mathcal{E}_h^{n+1;\gamma} = \mathcal{E}_h^{n;\gamma}$ for $n = 1, \dots, N-1$, and consequently $\mathcal{E}_h^{n;\gamma} = \mathcal{E}_h^{1;\gamma}$ for $n = 1, \dots, N$.

Proof. With $\theta = \frac{1}{2}$ we have $u_h^{n+1;\theta} = \frac{1}{2}(u_h^{n+1} + u_h^{n-1})$. Taking $u_h^{n+1} - u_h^{n-1}$ as a test function in (3.16) gives

$$\langle \partial_\tau^2 u_h^{n+1}, u_h^{n+1} - u_h^{n-1} \rangle + \frac{\Gamma_h^{n;\gamma}(u_h^{n+1})}{2} \langle \mathcal{L}_h(u_h^{n+1} + u_h^{n-1}), u_h^{n+1} - u_h^{n-1} \rangle = 0. \quad (3.18)$$

Throughout we abbreviate $\mathcal{V}_n =: \frac{1}{2} \|\nabla u_h^n\|^2$ and $\mathcal{K}_n =: \|\partial_\tau u_h^n\|^2$. For the first term in (3.18), write $u_h^{n+1} - 2u_h^n + u_h^{n-1} = \tau(\partial_\tau u_h^{n+1} - \partial_\tau u_h^n)$ and $u_h^{n+1} - u_h^{n-1} = \tau(\partial_\tau u_h^{n+1} + \partial_\tau u_h^n)$, so that it equals $\mathcal{K}_{n+1} - \mathcal{K}_n$. For the second, both arguments lie in V_h , so the definition (3.1) of $\mathcal{L}_h$ applies and turns it into $\langle \nabla(u_h^{n+1} + u_h^{n-1}), \nabla(u_h^{n+1} - u_h^{n-1}) \rangle = \|\nabla u_h^{n+1}\|^2 - \|\nabla u_h^{n-1}\|^2 = 2(\mathcal{V}_{n+1} - \mathcal{V}_{n-1})$. Hence (3.18) becomes

$$\mathcal{K}_{n+1} - \mathcal{K}_n + \Gamma_h^{n;\gamma}(u_h^{n+1})(\mathcal{V}_{n+1} - \mathcal{V}_{n-1}) = 0. \quad (3.19)$$

Inserting $\Gamma_h^{n;\gamma}(u_h^{n+1}) = \alpha + 2\beta(\gamma \mathcal{V}_{n+1} + (1 - 2\gamma) \mathcal{V}_n + \gamma \mathcal{V}_{n-1})$ into (3.19) therefore yields

$$(\mathcal{K}_{n+1} - \mathcal{K}_n) + \alpha(\mathcal{V}_{n+1} - \mathcal{V}_{n-1}) + 2\beta\gamma(\mathcal{V}_{n+1}^2 - \mathcal{V}_{n-1}^2) + 2\beta(1 - 2\gamma)\mathcal{V}_n(\mathcal{V}_{n+1} - \mathcal{V}_{n-1}) = 0.$$

Dividing by 2 and adding $\frac{\alpha}{2}\mathcal{V}_n$ and $\beta\gamma\mathcal{V}_n^2$ on both sides yields

$$\begin{aligned} \frac{1}{2}\mathcal{K}_{n+1} + \frac{\alpha}{2}(\mathcal{V}_{n+1} + \mathcal{V}_n) + \beta\gamma(\mathcal{V}_{n+1}^2 + \mathcal{V}_n^2) + \beta(1 - 2\gamma)\mathcal{V}_{n+1}\mathcal{V}_n \\ = \frac{1}{2}\mathcal{K}_n + \frac{\alpha}{2}(\mathcal{V}_n + \mathcal{V}_{n-1}) + \beta\gamma(\mathcal{V}_n^2 + \mathcal{V}_{n-1}^2) + \beta(1 - 2\gamma)\mathcal{V}_n\mathcal{V}_{n-1}. \end{aligned} \quad (3.20)$$

Applying the elementary identity $\gamma(a^2 + b^2) + (1 - 2\gamma)ab = ab + \gamma(a - b)^2$ for $a, b \in \mathbb{R}$ to the last two terms on both sides, then substituting back $\mathcal{V}$ yields the energy identity $\mathcal{E}_h^{n+1;\gamma} = \mathcal{E}_h^{n;\gamma}$. ■

Remark 3.4 (Scope of Proposition 3.3). The identity holds for the exact solution of the nonlinear scheme (3.16). Any approximation of it such as the Picard iterates defined below satisfies it only up to the nonlinear residual. So an implementation conserves $\mathcal{E}_h^n$ only to the tolerance at which the nonlinear system is solved and, in floating point, to the accuracy of the linear solves.

For $\beta \geq 0$ and any $\gamma \in [0, \frac{1}{2}]$ every term of (3.17) is nonnegative, and by (3.20),

$$\mathcal{E}_h^{n;\gamma} \geq \frac{1}{2} \|\partial_\tau u_h^n\|^2 + \frac{\alpha}{4} (\|\nabla u_h^{n-1}\|^2 + \|\nabla u_h^n\|^2) + \frac{\beta\gamma}{4} (\|\nabla u_h^{n-1}\|^4 + \|\nabla u_h^n\|^4),$$

so that Proposition 3.3 yields bounds on $\|\partial_\tau u_h^n\|$ and $\|\nabla u_h^n\|$ that are uniform in n and independent of τ and h . The quartic bound degenerates as $\gamma \rightarrow 0$, however, for $\alpha > 0$ this is immaterial, since the $\|\nabla u_h^n\|^2$ terms already provide the required control.

The hypothesis $\theta = \frac{1}{2}$ enters only through $u_h^{n+1;\theta} = \frac{1}{2}(u_h^{n+1} + u_h^{n-1}) + (\theta - \frac{1}{2})\tau^2 \partial_\tau^2 u_h^{n+1}$. Testing (3.16) with $u_h^{n+1} - u_h^{n-1}$ as in the proof above, the second summand contributes $(\theta - \frac{1}{2})\tau^2 \Gamma_h^{n;\gamma}(u_h^{n+1})(D_{n+1} - D_n)$ with $D_m := \|\nabla \partial_\tau u_h^m\|^2$, so that

$$\mathcal{E}_h^{n+1;\gamma} - \mathcal{E}_h^{n;\gamma} = \frac{1}{2}(\frac{1}{2} - \theta)\tau^2 \Gamma_h^{n;\gamma}(u_h^{n+1})(D_{n+1} - D_n). \quad (3.21)$$

This additional term does not telescope, because the coefficient $\Gamma_h^{n;\gamma}$ depends on the solution and hence on n . Summation by parts leaves $\Gamma_h^{n;\gamma} D_{n+1} - \Gamma_h^{1;\gamma} D_1 - \sum_{m=2}^n (\Gamma_h^{m;\gamma} - \Gamma_h^{m-1;\gamma}) D_m$, whose last sum is $\mathcal{O}(\tau \cdot n) = \mathcal{O}(T)$ under the assumption by the increment bound $|\Gamma_h^{m;\gamma} - \Gamma_h^{m-1;\gamma}| \lesssim \tau$ of (4.8). Note that this uses bounds that are not available at this point. We assume that $\Gamma_h^{m;\gamma}(u_h^{m+1})$ and D_m are bounded uniformly in m , and that the increment bound above holds. The first and the third are of the kind established in Section 4 for the discrete solution, and the second is a half order stronger than the bounds used there, since $D_m = \|\mathcal{L}_h^{\frac{1}{2}} \partial_\tau u_h^m\|^2$. The estimate is therefore formal, and we do not use it anywhere below. Hence, for $\theta \in [\frac{1}{4}, \frac{1}{2}]$ the discrete energy is not conserved but satisfies

$$\max_{1 \leq n \leq N} |\mathcal{E}_h^{n;\gamma} - \mathcal{E}_h^{1;\gamma}| \lesssim (\frac{1}{2} - \theta)\tau^2, \quad (3.22)$$

that is, it is conserved to second order in τ with a constant growing linearly in T . This is confirmed numerically in Section 5.

Remark 3.5 (The cases $\gamma = 0$ and $\gamma = \frac{1}{2}$). The invariant $\mathcal{E}_h^{n;\gamma}$ is the fully discrete analogue of $\tilde{E}_{1,n}$, and $\tilde{E}_{2,n}$ in [19]. That is, for $\gamma \in \{0, \frac{1}{2}\}$ the equations coincide up to the spatial discretization. Thus,

we have that the energy $\mathcal{E}_h^{n;\gamma}$ does not conserve the continuous energy $E(t_n)$. For $\gamma = 0$ the coefficient $\Gamma_h^{n;0}(u_h^{n+1})$ does not involve u_h^{n+1} , so that (3.16) is linearly implicit and Picard iteration is not needed. Nonetheless, Proposition 3.3 shows that this energy is conserved. For general γ , comparing with the case $\gamma = \frac{1}{2}$ gives

$$\mathcal{E}_h^{n;\frac{1}{2}} - \mathcal{E}_h^{n;\gamma} = \frac{\beta(1-2\gamma)}{8} (\|\nabla u_h^n\|^2 - \|\nabla u_h^{n-1}\|^2)^2. \quad (3.23)$$

Since $\|\nabla u_h^n\|^2 - \|\nabla u_h^{n-1}\|^2 = \tau \langle \nabla(u_h^n + u_h^{n-1}), \nabla \partial_\tau u_h^n \rangle$, the difference (3.23) is of order τ^2 . Moreover, it is itself a function of the two levels $n-1, n$ only, whence

$$\mathcal{E}_h^{n;\frac{1}{2}} = \mathcal{E}_h^{1;\frac{1}{2}} + \frac{\beta(1-2\gamma)}{8} [(\|\nabla u_h^n\|^2 - \|\nabla u_h^{n-1}\|^2)^2 - (\|\nabla u_h^1\|^2 - \|\nabla u_h^0\|^2)^2], \quad n = 1, \dots, N.$$

Thus, for $\gamma \neq \frac{1}{2}$ the quantity $\mathcal{E}_h^{n;\frac{1}{2}}$ is not conserved, but its deviation is bounded by $\mathcal{O}(\tau^2)$ uniformly in n and does not accumulate over $[0, T]$.

4 Analysis of the fully discrete scheme

To solve the nonlinear scheme (3.16) at every time step, we apply a Picard iteration that leads to solving a linear equation for every iterate. That is, we seek a sequence of functions $\{u_{h;\ell+1}^{n+1}\}_{\ell \geq 0} \subset V_h$ that solves the *fully discrete linear iteration scheme*

$$u_{h;\ell+1}^{n+1} + \tau^2 \theta \Gamma_h^{n;\gamma}(u_{h;\ell}^{n+1}) \mathcal{L}_h u_{h;\ell+1}^{n+1} = \tau^2 f_h^{n+1;\eta} + 2u_h^n - u_h^{n-1} - \tau^2 \Gamma_h^{n;\gamma}(u_{h;\ell}^{n+1}) \mathcal{L}_h((1-2\theta)u_h^n + \theta u_h^{n-1}) \quad (4.1)$$

with the initial guess $u_{h;0}^{n+1} = u_h^n$. In this section we will prove that this scheme admits a unique solution trajectory over the whole time interval $[0, T]$ where the exact solution is bounded. This is based on a contraction principle (fixed point argument) of the Picard iteration with the following linear operator

$$\mathcal{T}_v = \mathcal{I} + \tau^2 \theta \Gamma_h^{n;\gamma}(v) \mathcal{L}_h.$$

Then, the Picard iteration seeks a fixed point of (note that $\theta \geq \frac{1}{4}$)

$$v = F(v) = \mathcal{T}_v^{-1} g_h^n + (2 - \frac{1}{\theta})u_h^n - u_h^{n-1}, \quad g_h^n := \tau^2 f_h^{n+1;\eta} + \frac{1}{\theta}u_h^n. \quad (4.2)$$

4.1 Local existence of a solution

As a first step we show that the nonlinear equation (3.16) has a unique solution at each time level, and that the Picard iteration converges to it. Both follow from a contraction principle applied to the fixed point map (4.2), whose fixed point is that solution. This is formulated in the following lemma.

Lemma 4.1 (Existence of a solution of the nonlinear discrete equation). *Let Assumption 2.3 hold. Further, let the fully discrete solutions be bounded by*

$$\begin{aligned} \|\mathcal{L}_h^{\frac{3}{4}} u_h^n\|_{L^2(\Omega)} &\leq M_{\frac{3}{4}}, & \|\mathcal{L}_h^{\frac{1}{4}} u_h^n\|_{L^2(\Omega)} &\leq M_{\frac{1}{4}}, & \|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^n\|_{L^2(\Omega)} &\leq M'_{\frac{1}{4}}, \\ \|\mathcal{L}_h^{\frac{3}{4}} u_h^{n-1}\|_{L^2(\Omega)} &\leq M_{\frac{3}{4}}, & \|\mathcal{L}_h^{\frac{1}{4}} u_h^{n-1}\|_{L^2(\Omega)} &\leq M_{\frac{1}{4}} \end{aligned} \quad (4.3)$$

for some $1 \leq n \leq N-1$. Then there exists $\tau_0 > 0$, depending only on $\alpha_{\min}$, $\beta_{\max}$, γ , θ , the final time T and the constants $M_{\frac{1}{4}}, M_{\frac{3}{4}}, C_{\frac{1}{4}}$ of Assumption 2.3, but not on the time step n , the time step size τ or the mesh size h , such that for every $0 < \tau \leq \tau_0$ the nonlinear equation (3.16) has a unique solution $u_h^{n+1} \in V_h$, to which the Picard iteration (4.1) converges. Further, for the following norms we have that

$$\|\mathcal{L}_h^{\frac{3}{4}} u_h^{n+1}\|_{L^2(\Omega)} \lesssim 1, \quad \|\mathcal{L}_h^{\frac{1}{4}} u_h^{n+1}\|_{L^2(\Omega)} \lesssim 1, \quad \|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^{n+1}\|_{L^2(\Omega)} \lesssim 1$$

hold independent of the time step n , the time step size τ , and the mesh size h .

Proof. We show this lemma using a contraction principle on the mapping F introduced in (4.2). We first show that F is a self-map on a set bounded in both the $\mathcal{L}_h^{\frac{3}{4}}$ - and the $\mathcal{L}_h^{\frac{1}{4}}$ -norm, with bounds independent of the preimage: the contraction is performed in the $\mathcal{L}_h^{\frac{1}{4}}$ -norm, and the $\mathcal{L}_h^{\frac{3}{4}}$ -bound is what the error estimate below needs. Let $a \in \{\frac{1}{4}, \frac{3}{4}\}$. Using $\frac{1}{\theta} \geq 2$, we have for any $\varphi_h \in V_h$ that

$$\|\mathcal{L}_h^a F(\varphi_h)\| = \|\mathcal{L}_h^a \mathcal{T}_{\varphi_h}^{-1} (\tau^2 f_h^{n+1;\eta} + \frac{1}{\theta} u_h^n) + \mathcal{L}_h^a ((2 - \frac{1}{\theta})u_h^n - u_h^{n-1})\|$$

$$\begin{aligned}
 &\leq \tau^{-2(\frac{1}{4}-a)} \|\mathcal{L}_h^{a-\frac{1}{4}} \mathcal{T}_{\varphi_h}^{-1}\| \|\tau^{2(\frac{5}{4}-a)} \mathcal{L}_h^{\frac{1}{4}} f_h^{n+1;\eta}\| + \|\mathcal{T}_{\varphi_h}^{-1}\| \|\frac{1}{\theta} \mathcal{L}_h^a u_h^n\| + |2 - \frac{1}{\theta}| \|\mathcal{L}_h^a u_h^n\| + \|\mathcal{L}_h^a u_h^{n-1}\| \\
 &\leq \tau^{-2(\frac{1}{4}-a)} \tau^{2(\frac{1}{4}-a)} c(a - \frac{1}{4}) \|\tau^{2(\frac{5}{4}-a)} \mathcal{L}_h^{\frac{1}{4}} f_h^{n+1;\eta}\| + (\frac{2}{\theta} - 2) \|\mathcal{L}_h^a u_h^n\| + \|\mathcal{L}_h^a u_h^{n-1}\| \\
 &\leq c(a - \frac{1}{4}) \tau^{2(\frac{5}{4}-a)} \|\mathcal{L}_h^{\frac{1}{4}} f_h^{n+1;\eta}\| + (\frac{2}{\theta} - 2) \|\mathcal{L}_h^a u_h^n\| + \|\mathcal{L}_h^a u_h^{n-1}\|,
 \end{aligned}$$

where for $0 \leq a \leq 1$ we have that $c(0) = 1, c(a) = (\frac{a}{\theta\alpha_{\min}})^a (1-a)^{1-a}, c(1) = \frac{1}{\theta\alpha_{\min}}$. Note that the source term mean needs separate handling as we do not invoke the $\mathcal{L}_h^{\frac{3}{4}}$ -norm. Thus, for $a = \frac{1}{4}$ and $a = \frac{3}{4}$, respectively, using (3.5) and (3.6) we obtain

$$\begin{aligned}
 \|\mathcal{L}_h^{\frac{1}{4}} F(\varphi_h)\| &\leq \tau^2 C_{\frac{1}{4}} + (\frac{2}{\theta} - 1) M_{\frac{1}{4}} \leq T^2 C_{\frac{1}{4}} + (\frac{2}{\theta} - 1) M_{\frac{1}{4}} =: K_{\frac{1}{4}}, \\
 \|\mathcal{L}_h^{\frac{3}{4}} F(\varphi_h)\| &\leq \frac{1}{2\sqrt{\theta\alpha_{\min}}} \tau C_{\frac{1}{4}} + (\frac{2}{\theta} - 1) M_{\frac{3}{4}} \leq \frac{1}{2\sqrt{\theta\alpha_{\min}}} T C_{\frac{1}{4}} + (\frac{2}{\theta} - 1) M_{\frac{3}{4}} =: K_{\frac{3}{4}}
 \end{aligned}$$

without any *a priori* bounds on φ_h , specifically, this means that the constants K_a for $a \in \{\frac{1}{4}, \frac{3}{4}\}$ do not depend on the preimage. We construct a complete subset of the finite element space V_h to be

$$K := \{\varphi_h \in V_h \mid \|\mathcal{L}_h^a \varphi_h\| \leq K_a \text{ for } a \in \{\frac{1}{4}, \frac{3}{4}\}\},$$

such that the fixed point map F is a self map on K . The contraction is performed in the $\mathcal{L}_h^{\frac{1}{4}}$ -norm. The reason is the structure of the nonlinearity: since $\|\nabla \cdot\| = \|\mathcal{L}_h^{\frac{1}{2}} \cdot\|$, its increment is proportional to the increment of its argument in the $\mathcal{L}_h^a$ -norm for any $0 \leq a \leq 1$,

$$\begin{aligned}
 |\Gamma_h^{n;\gamma}(u) - \Gamma_h^{n;\gamma}(v)| &= \beta^n \gamma \|\nabla u\|^2 - \|\nabla v\|^2 = \beta^n \gamma \langle \nabla(u-v), \nabla(u+v) \rangle \\
 &\leq \beta^n \gamma (\|\mathcal{L}_h^{1-a} u\| + \|\mathcal{L}_h^{1-a} v\|) \|\mathcal{L}_h^a(u-v)\|.
 \end{aligned} \tag{4.4}$$

At $a = 0$ this would require H^2 -regularity, through $\|\mathcal{L}_h u_h^n\|$. Following [2], the choice $a = \frac{1}{4}$ avoids it, and it is the one made here. Using the definition (4.2) and the fact that $\mathcal{T}_v^{-1}$ and $\mathcal{L}_h$ commute, we have for any $\varphi_h, \psi_h \in V_h$ that

$$\begin{aligned}
 \|\mathcal{L}_h^a(F(\varphi_h) - F(\psi_h))\| &= \|\mathcal{L}_h^a(\mathcal{T}_{\varphi_h}^{-1} - \mathcal{T}_{\psi_h}^{-1})g_h^n\| = \|\mathcal{T}_{\varphi_h}^{-1}(\mathcal{T}_{\psi_h} - \mathcal{T}_{\varphi_h})\mathcal{L}_h^{-1}\mathcal{L}_h^{1+a}\mathcal{T}_{\psi_h}^{-1}g_h^n\| \\
 &\leq \|\mathcal{T}_{\varphi_h}^{-1}(\mathcal{T}_{\psi_h} - \mathcal{T}_{\varphi_h})\mathcal{L}_h^{-1}\| \|\mathcal{L}_h^{1+a}\mathcal{T}_{\psi_h}^{-1}g_h^n\|.
 \end{aligned} \tag{4.5}$$

The norm of the first term can be estimated by writing out the difference, and using the proportionality argument (4.4). We have using also $\|\mathcal{T}_v^{-1}\| \leq 1$ that

$$\begin{aligned}
 \|\mathcal{T}_{\varphi_h}^{-1}(\mathcal{T}_{\psi_h} - \mathcal{T}_{\varphi_h})\mathcal{L}_h^{-1}\| &= |\tau^2 \theta (\Gamma_h^{n;\gamma}(\varphi_h) - \Gamma_h^{n;\gamma}(\psi_h))| \\
 &\leq \tau^2 \theta \beta^n \gamma (\|\mathcal{L}_h^{1-a}\varphi_h\| + \|\mathcal{L}_h^{1-a}\psi_h\|) \|\mathcal{L}_h^a(\varphi_h - \psi_h)\|.
 \end{aligned}$$

Here, on the right-hand side we observe the difference of functions φ_h, ψ_h in the $\mathcal{L}_h^a$ -norm, as on the left-hand side of (4.5). Thus, it is left to bound the last term on the right-hand side of (4.5). With similar arguments as above we have that

$$\begin{aligned}
 \|\mathcal{L}_h^{1+a}\mathcal{T}_{\psi_h}^{-1}g_h^n\| &= \|\tau^2 \mathcal{T}_{\psi_h}^{-1} \mathcal{L}_h \mathcal{L}_h^a f_h^{n+1;\eta} + \frac{1}{\theta} \mathcal{T}_{\psi_h}^{-1} \mathcal{L}_h^{1-2a} \mathcal{L}_h^{3a} u_h^n\| \\
 &\leq \tau \|\mathcal{T}_{\psi_h}^{-1} \mathcal{L}_h\| \|\tau \mathcal{L}_h^a f_h^{n+1;\eta}\| + \frac{1}{\theta} \|\mathcal{T}_{\psi_h}^{-1} \mathcal{L}_h^{1-2a}\| \|\mathcal{L}_h^{3a} u_h^n\| \\
 &\leq \tau^{-1} c(1) \|\tau \mathcal{L}_h^a f_h^{n+1;\eta}\| + \tau^{2(2a-1)} \frac{1}{\theta} c(1-2a) \|\mathcal{L}_h^{3a} u_h^n\|.
 \end{aligned}$$

Combining the two previous estimates into (4.5), setting $a = \frac{1}{4}$ and using the bounds in Assumption 2.3 and in (4.3) we have that

$$\begin{aligned}
 \|\mathcal{L}_h^{\frac{1}{4}}(F(\varphi_h) - F(\psi_h))\| &\leq \tau^2 \theta \beta^n \gamma (\|\mathcal{L}_h^{\frac{3}{4}} \varphi_h\| + \|\mathcal{L}_h^{\frac{3}{4}} \psi_h\|) (\tau^{-1} \frac{1}{\theta \alpha_{\min}} \|\tau \mathcal{L}_h^{\frac{1}{4}} f_h^{n+1;\eta}\| + \tau^{-1} \frac{1}{\theta} c(\frac{1}{2}) \|\mathcal{L}_h^{\frac{3}{4}} u_h^n\|) \|\mathcal{L}_h^{\frac{1}{4}} \varphi_h - \mathcal{L}_h^{\frac{1}{4}} \psi_h\| \\
 &\leq \tau \beta^n \gamma (\|\mathcal{L}_h^{\frac{3}{4}} \varphi_h\| + \|\mathcal{L}_h^{\frac{3}{4}} \psi_h\|) (\tau \frac{1}{\alpha_{\min}} C_{\frac{1}{4}} + \frac{1}{2\sqrt{\theta\alpha_{\min}}} M_{\frac{3}{4}}) \|\mathcal{L}_h^{\frac{1}{4}}(\varphi_h - \psi_h)\|.
 \end{aligned}$$

Finally, for any $\varphi_h, \psi_h \in K$ we have that

$$\|\mathcal{L}_h^{\frac{1}{4}}(F(\varphi_h) - F(\psi_h))\| \leq \tau \beta_{\max} \gamma 2K_{\frac{3}{4}} (T \frac{1}{\alpha_{\min}} C_{\frac{1}{4}} + \frac{1}{2\sqrt{\theta\alpha_{\min}}} M_{\frac{3}{4}}) \|\mathcal{L}_h^{\frac{1}{4}}(\varphi_h - \psi_h)\|.$$

Thus, there exists $\tau_0 > 0$ that may depend on the coefficients α and β , the final time T , the parameters γ and θ , the norms of the exact solution u and the source term f , but does not depend on the time step n and the mesh size h such that, for $0 < \tau \leq \tau_0$, the mapping F is a contractive self-mapping on the complete metric space K . The bounds placing $F(\varphi_h)$ in K were obtained without any restriction on the preimage, so $F(V_h) \subset K$ and every fixed point of F lies in K , uniqueness in K is therefore uniqueness in V_h . By the contraction principle, there exists a unique fixed point $u_h^{n+1} \in V_h$ with the following bounds $\|\mathcal{L}_h^{\frac{3}{4}} u_h^{n+1}\| \leq K_{\frac{3}{4}}$, $\|\mathcal{L}_h^{\frac{1}{4}} u_h^{n+1}\| \leq K_{\frac{1}{4}}$. Lastly, $\|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^{n+1}\|$ is bounded by the discrete energy argument. Testing (3.16) with $\frac{1}{2} \mathcal{L}_h^{\frac{1}{2}} (\partial_\tau u_h^{n+1} + \partial_\tau u_h^n)$ and (3.10) gives

$$\begin{aligned} \|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^{n+1}\|^2 + \Gamma_h^{n;\gamma}(u_h^{n+1}) \mathcal{W}_h^{n+1} \\ \leq \|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^n\|^2 + \Gamma_h^{n;\gamma}(u_h^{n+1}) \mathcal{W}_h^n + \|\mathcal{L}_h^{\frac{1}{4}} f_h^{n+1;\eta}\| \|\mathcal{L}_h^{\frac{1}{4}} (u_h^{n+1} - u_h^{n-1})\|, \end{aligned}$$

where $\mathcal{W}_h^n$ as in Lemma 3.1. Since $\theta \geq \frac{1}{4}$ and $\Gamma_h^{n;\gamma}(u_h^{n+1}) \geq 0$, the left-hand side has the lower bound $\|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^{n+1}\|^2$. Further, by assumption (4.3) we have $\Gamma_h^{n;\gamma}(u_h^{n+1}) \leq \alpha_{\max} + \beta_{\max}(\gamma K_{\frac{1}{4}} K_{\frac{3}{4}} + (1 - \gamma) M_{\frac{1}{4}} M_{\frac{3}{4}}) =: C_\Gamma$, and using (3.11) we have that

$$\|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^{n+1}\|^2 \leq M_{\frac{1}{4}}'^2 + C_\Gamma M_{\frac{3}{4}}^2 + C_{\frac{1}{4}}(K_{\frac{1}{4}} + M_{\frac{1}{4}}),$$

which is independent of n , τ and h . ■

The previous lemma shows the existence of a solution u_h^{n+1} for some $1 \leq n \leq N - 1$, however, it has a crucial drawback. Since we are using a contraction argument, the norms of the solution are larger than the ones assumed at the previous time steps. A recursive application of this lemma would lead to a blow-up of the solution norms. The following sections provide an idea of how to avoid the blow-up and show that the solution exists for all time steps.

4.2 Convergence of the local solutions

We avoid the issue of a potential blow-up by showing that the solution u_h^{n+1} approximates the Ritz projection of the exact solution at time t_{n+1} , and, thus, lives in the same ball as assumed for the previous time steps. This leads to the fact that the norms of all discrete solutions are bounded and we may apply recursively the previous lemma. In the following we abbreviate $\hat{u}^n = u(n\tau)$ for the exact solution. Throughout this section, in addition to the assumptions in Lemma 4.1, we assume that a whole trajectory of fully discrete solutions exists and satisfies norm and error bounds. This is summarized in the following hypothesis.

Hypothesis 4.2 (Induction hypothesis). Let the fully discrete solution be uniformly bounded by

$$\begin{aligned} \|\mathcal{L}_h^{\frac{3}{4}} u_h^m\|_{L^2(\Omega)} &\leq M_{\frac{3}{4}}, & \|\mathcal{L}_h^{\frac{1}{4}} u_h^m\|_{L^2(\Omega)} &\leq M_{\frac{1}{4}}, & m &= 0, \dots, n, \\ \|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^m\|_{L^2(\Omega)} &\leq M_{\frac{1}{4}}', & & & m &= 1, \dots, n \end{aligned} \quad (4.6a)$$

for some $1 \leq n \leq N - 1$, and assume the following error estimates

$$\begin{aligned} \|\mathcal{L}_h^{\frac{1}{2}+a} (R_h u(t_m) - u_h^m)\|_{L^2(\Omega)} + \|\mathcal{L}_h^a (R_h u(t_m) - u_h^m)\|_{L^2(\Omega)} &\lesssim h^{p+1-2a} + \tau^2, & m &= 0, \dots, n, \\ \|\mathcal{L}_h^a \partial_\tau (R_h u(t_m) - u_h^m)\|_{L^2(\Omega)} &\lesssim h^{p+1-2a} + \tau^2, & m &= 1, \dots, n \end{aligned} \quad (4.6b)$$

hold for $a \in \{0, \frac{1}{4}\}$, where the constants are independent of the time step n , the time step size τ , and the mesh size h .

Note that Hypothesis 4.2 implies the hypotheses (4.3) of Lemma 4.1, so that a unique solution $u_h^{n+1} \in V_h$ of (3.16) exists. Regarding the choice of the parameter a , the hypothesis is stated for both $a = 0$ and $a = \frac{1}{4}$ for the reason that the existence of a trajectory requires the stronger norms with $a = \frac{1}{4}$, and standard error estimates are performed in the weaker norms with $a = 0$. To show the convergence consider the fully discrete scheme and insert the Ritz projection of the exact solution

$R_h \hat{u}^n$ into the fully discrete Kirchhoff PDE (3.16), then multiplying with a test function $\varphi_h \in V_h$ and integrating we obtain

$$\langle \partial_\tau^2 R_h \hat{u}^{n+1}, \varphi_h \rangle + \Gamma_h^{n;\gamma}(u_h^{n+1}) \langle \nabla R_h \hat{u}^{n+1;\theta}, \nabla \varphi_h \rangle = \langle f_h^{n+1;\eta}, \varphi_h \rangle + \langle \delta_{h;\frac{1}{4}}^n + \delta_{h;\frac{3}{4}}^n + S_h^n, \varphi_h \rangle. \quad (4.7)$$

Here, $\delta_{h;\frac{1}{4}}^n + \delta_{h;\frac{3}{4}}^n = [\delta_{h;f}^n + \delta_{h;1}^n] + [\delta_{h;2}^n + \delta_{h;3}^n + \delta_{h;4}^n]$ denotes a sum of defects, where the split is according to the analysis below, and S_h^n is a stabilization term that arises from the nonlinearity. Writing $q^m = \alpha^m + \beta^m \|\nabla \hat{u}^m\|^2$ for the coefficient of the continuous equation at t_m , applying P_h to that equation and using the commutation identity (3.2) in the form $\mathcal{L}_h R_h \hat{u}^m = P_h \mathcal{L} \hat{u}^m$, we have that

$$\begin{aligned} \delta_{h;f}^m &= P_h f^m - f_h^{m+1;\eta}, & \delta_{h;1}^m &= \partial_\tau^2 R_h \hat{u}^{m+1} - P_h \partial_t^2 \hat{u}^m, \\ \delta_{h;2}^m &= \Gamma_h^{m;\gamma}(u_h^{m+1}) \mathcal{L}_h R_h \theta(\hat{u}^{m+1} - 2\hat{u}^m + \hat{u}^{m-1}), \\ \delta_{h;3}^m &= \beta^m \gamma (\|\nabla \hat{u}^{m+1}\|^2 - 2\|\nabla \hat{u}^m\|^2 + \|\nabla \hat{u}^{m-1}\|^2) \mathcal{L}_h R_h \hat{u}^m, \end{aligned}$$

and, with the projection and error contributions to the coefficient $\rho^m = \|\nabla R_h \hat{u}^m\|^2 - \|\nabla \hat{u}^m\|^2$, $\sigma^m = \|\nabla u_h^m\|^2 - \|\nabla R_h \hat{u}^m\|^2$,

$$\begin{aligned} \delta_{h;4}^m &= \beta^m (\gamma \rho^{m+1} + (1 - 2\gamma) \rho^m + \gamma \rho^{m-1}) \mathcal{L}_h R_h \hat{u}^m, \\ S_h^m &= \beta^m (\gamma \sigma^{m+1} + (1 - 2\gamma) \sigma^m + \gamma \sigma^{m-1}) \mathcal{L}_h R_h \hat{u}^m. \end{aligned}$$

The split is exactly the one obtained by writing $\Gamma_h^{m;\gamma}(u_h^{m+1}) - q^m$ as a sum of a σ -part, a ρ -part and a purely temporal part, as carried out below. From the definitions of the defects, we readily observe that the split of the defect are according to if they carry a discrete Laplacian or not. The reason for it is that the terms are handled differently in the analysis. And only S_h^m is proportional to the coefficient *error*, which is why it alone requires the stabilization argument. Before we prove the theorem, we first provide the estimates for the defects and stabilization terms such that the proof is concise and readable. The proof is given in [Appendix A.1](#).

Lemma 4.3 (Consistency). *Let [Assumption 2.3](#) and [Hypothesis 4.2](#) hold and let $a \in \{0, \frac{1}{4}\}$. Further, assume the additional temporal regularity for the solution u to the PDE (1.1)*

$$\begin{aligned} \partial_t^j u &\in L^\infty(0, T; D(\mathcal{L}^{\frac{3}{4}})), & j &= 1, 2, 3, \\ \partial_t^4 u &\in L^\infty(0, T; D(\mathcal{L}^{\frac{1}{4}})), \end{aligned}$$

as well as the spatial regularity

$$\begin{aligned} \partial_t^j u &\in L^\infty(0, T; H^{p+1}(\Omega)), & j &= 0, 1, 2, \\ f &\in L^\infty(0, T; H^{p+1}(\Omega) \cap H_0^1(\Omega)), & \partial_t^2 f &\in L^\infty(0, T; D(\mathcal{L}^{\frac{1}{4}})). \end{aligned}$$

Then, we have that

$$|\Gamma_h^{m;\gamma}(u_h^{m+1}) - \Gamma_h^{m-1;\gamma}(u_h^m)| \lesssim \tau, \quad m = 2, \dots, n \quad (4.8)$$

$$\|\mathcal{L}_h^a \delta_{h;\frac{1}{4}}^m\| \lesssim h^{p+1-2a} + \tau^2, \quad m = 1, \dots, n \quad (4.9)$$

$$\|\mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^m\| \lesssim h^{2p} + \tau^2, \quad m \in \{1, n\} \quad (4.10)$$

$$\|\mathcal{L}_h^{a-\frac{1}{2}} \partial_\tau \delta_{h;\frac{3}{4}}^m\| \lesssim h^{2p} + \tau^2, \quad m = 2, \dots, n \quad (4.11)$$

$$|\langle \mathcal{L}_h^a S_h^m, \mathcal{L}_h^a \varphi_h \rangle| \lesssim \left(\sum_{k=m-1}^{m+1} \|\mathcal{L}_h^{\frac{1}{2}+a} e_h^k\| \right) \|\mathcal{L}_h^a \varphi_h\|, \quad m = 1, \dots, n, \quad (4.12)$$

where the constants are independent of the time steps n and m , the time step size τ , and the mesh size h .

Note that here, and only here, we assume the additional spatial regularity in the Sobolev scale, since these hypotheses feed the finite element approximation estimates. The requirement $f(t) \in H_0^1(\Omega)$ makes the approximation estimate (3.4) applicable to f . It is an additional hypothesis, not a consequence of [Assumption 2.3](#), whose $D(\mathcal{L}^{\frac{1}{4}}) = H_{00}^{\frac{1}{2}}(\Omega)$ does not embed in $H_0^1(\Omega)$.

Thus, we may give the main convergence theorem. The solution u_h^{n+1} of Lemma 4.1 approximates $R_h u$ with order $\mathcal{O}(h^{p+\frac{1}{2}} + \tau^2)$ in the $\mathcal{L}_h^{\frac{3}{4}}$ -norm, under stronger regularity in space and time and given boundedness of the discrete solution at all previous time steps. It is the induction step; the base case and the statement for all time levels are Lemma 4.5 and Corollary 4.6. All estimates below are for the error $R_h u(t_m) - u_h^m$ relative to the Ritz projection, not for the full error: the projection error $u(t_m) - R_h u(t_m)$ is not a finite element function and is therefore not measured in the discrete fractional norms at all. It is added back, in the L^2 - and H^1 -norms, in Corollary 4.7.

Theorem 4.4 (Convergence of the local solution). *Let the assumptions of Lemma 4.3 hold, let $\theta \in [\frac{1}{4}, \frac{1}{2}]$, and assume at the endpoint $\theta = \frac{1}{4}$ in addition the mesh condition $\tau \leq \kappa h$ for some fixed $\kappa > 0$. Then, the unique solution $u_h^{n+1} \in V_h$ of the nonlinear equation (3.16) supplied by Lemma 4.1 satisfies the same estimates at the level $n+1$ as the solutions at the previous steps. That is, we have that*

$$\begin{aligned} \|\mathcal{L}_h^{\frac{1}{2}+a}(R_h u(t_{n+1}) - u_h^{n+1})\|_{L^2(\Omega)} + \|\mathcal{L}_h^a \partial_\tau (R_h u(t_{n+1}) - u_h^{n+1})\|_{L^2(\Omega)} &\lesssim h^{p+1-2a} + \tau^2, \\ \|\mathcal{L}_h^a (R_h u(t_{n+1}) - u_h^{n+1})\|_{L^2(\Omega)} &\lesssim h^{p+1-2a} + \tau^2, \end{aligned}$$

where the constants are independent of the time step n , the time step size τ , and the mesh size h .

Proof. First, we derive the final error equation such that we may apply the discrete Grönwall inequality. Subtracting the fully discrete equation (3.16) from (4.7), we have that

$$\langle \partial_\tau^2 e_h^{n+1}, \varphi_h \rangle + \Gamma_h^{n;\gamma}(u_h^{n+1}) \langle \nabla e_h^{n+1;\theta}, \nabla \varphi_h \rangle = \langle \delta_{h;\frac{1}{4}}^n + \delta_{h;\frac{3}{4}}^n + S_h^n, \varphi_h \rangle. \quad (4.13)$$

Let $a \in \{0, \frac{1}{4}\}$, and define the test function $\varphi_h = \frac{1}{2} \mathcal{L}_h^{2a} (\partial_\tau e_h^{n+1} + \partial_\tau e_h^n) \in V_h$, and define the three layer displacement energy by $\mathcal{W}_h^m := \frac{1}{4} \|\mathcal{L}_h^{\frac{1}{2}+a}(e_h^m + e_h^{m-1})\|^2 + (\theta - \frac{1}{4}) \|\mathcal{L}_h^{\frac{1}{2}+a}(e_h^m - e_h^{m-1})\|^2$. Note that this definition coincides with the definition in Lemma 3.1. Using $\frac{1}{2}(\partial_\tau e_h^{n+1} + \partial_\tau e_h^n) = \frac{1}{2\tau}(e_h^{n+1} - e_h^{n-1})$ and (3.10) for the second term, the left-hand side becomes

$$\frac{1}{2\tau} (\|\mathcal{L}_h^a \partial_\tau e_h^{n+1}\|^2 - \|\mathcal{L}_h^a \partial_\tau e_h^n\|^2) + \frac{1}{2\tau} \Gamma_h^{n;\gamma}(u_h^{n+1})(\mathcal{W}_h^{n+1} - \mathcal{W}_h^n).$$

Writing, for $m \geq 2$,

$$\begin{aligned} \Gamma_h^{m;\gamma}(u_h^{m+1})(\mathcal{W}_h^{m+1} - \mathcal{W}_h^m) \\ = \Gamma_h^{m;\gamma}(u_h^{m+1})\mathcal{W}_h^{m+1} - \Gamma_h^{m-1;\gamma}(u_h^m)\mathcal{W}_h^m - (\Gamma_h^{m;\gamma}(u_h^{m+1}) - \Gamma_h^{m-1;\gamma}(u_h^m))\mathcal{W}_h^m, \end{aligned}$$

so that the first two terms on the right-hand side telescope, and keeping the term $m = 1$ as it stands. Then, similarly we write for $m \geq 2$,

$$\begin{aligned} \langle \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^m, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^{m+1} - e_h^{m-1}) \rangle &= \langle \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^m, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^{m+1} + e_h^m) \rangle - \langle \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^{m-1}, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^m + e_h^{m-1}) \rangle \\ &\quad - \tau \langle \mathcal{L}_h^{a-\frac{1}{2}} \partial_\tau \delta_{h;\frac{3}{4}}^m, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^m + e_h^{m-1}) \rangle, \end{aligned}$$

such that again the first two terms telescope. Inserting the previous three equations into (4.13), summing from 1 to n and multiplying by 2τ yields

$$\begin{aligned} \|\mathcal{L}_h^a \partial_\tau e_h^{n+1}\|^2 + \Gamma_h^{n;\gamma}(u_h^{n+1})\mathcal{W}_h^{n+1} &= \|\mathcal{L}_h^a \partial_\tau e_h^1\|^2 + \Gamma_h^{1;\gamma}(u_h^2)\mathcal{W}_h^1 \\ &\quad + \sum_{m=2}^n (\Gamma_h^{m;\gamma}(u_h^{m+1}) - \Gamma_h^{m-1;\gamma}(u_h^m))\mathcal{W}_h^m + \tau \sum_{m=1}^n \langle \mathcal{L}_h^a (\delta_{h;\frac{1}{4}}^m + S_h^m), \mathcal{L}_h^a (\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \rangle \\ &\quad + \langle \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^n, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^{n+1} + e_h^n) \rangle - \langle \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^1, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^1 + e_h^0) \rangle - \tau \sum_{m=2}^n \langle \mathcal{L}_h^{a-\frac{1}{2}} \partial_\tau \delta_{h;\frac{3}{4}}^m, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^m + e_h^{m-1}) \rangle. \end{aligned} \quad (4.14)$$

Since $\beta^m \geq \beta_{\min} > 0$ and the weights $\gamma, 1 - 2\gamma, \gamma$ are non-negative, the coefficient obeys $\alpha_{\min} \leq \Gamma_h^{m;\gamma}(u_h^{m+1}) \leq C_\Gamma$, with C_Γ from the proof of Lemma 4.1. Let $\mathcal{F}_h^m = \|\mathcal{L}_h^a \partial_\tau e_h^m\|^2 + \mathcal{W}_h^m$ for $m =$

$1, \dots, n+1$, and $\mathcal{F}_h^0 := \|\mathcal{L}_h^{\frac{1}{2}+a} e_h^0\|^2$ be the quantities as defined in [Corollary 3.2](#) at $a \in \{0, \frac{1}{4}\}$ and u_h^m replaced by e_h^m . We now bound the seven terms on the right-hand side of (4.14) in turn. By the induction hypothesis (4.6) at the levels $m = 0, 1$, together with $\Gamma_h^{1;\gamma}(u_h^2) \leq C_\Gamma$, (4.10), Cauchy–Schwarz, and (3.11),

$$\|\mathcal{L}_h^a \partial_\tau e_h^1\|^2 + \Gamma_h^{1;\gamma}(u_h^2) \mathcal{W}_h^1 + |\langle \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^1, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^1 + e_h^0) \rangle| \lesssim (h^{p+1-2a} + \tau^2)^2, \quad (4.15a)$$

where we used $h^{2p} \leq h^{p+1-2a}$ for $p \geq 1$ and $h \leq 1$. By (4.8) and (3.11),

$$\sum_{m=2}^n |\Gamma_h^{m;\gamma}(u_h^{m+1}) - \Gamma_h^{m-1;\gamma}(u_h^m)| \mathcal{W}_h^m \lesssim \tau \sum_{m=2}^n \mathcal{W}_h^m \leq \tau \sum_{m=1}^n \mathcal{F}_h^m. \quad (4.15b)$$

By Cauchy–Schwarz, (4.9), the plain Young inequality and $\tau \sum_{m=1}^n 1 \leq T$,

$$\tau \sum_{m=1}^n |\langle \mathcal{L}_h^a \delta_{h;\frac{1}{4}}^m, \mathcal{L}_h^a (\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \rangle| \lesssim T(h^{p+1-2a} + \tau^2)^2 + \tau \sum_{m=1}^{n+1} \mathcal{F}_h^m. \quad (4.15c)$$

By (4.12), the plain Young inequality and (3.15),

$$\tau \sum_{m=1}^n |\langle \mathcal{L}_h^a S_h^m, \mathcal{L}_h^a (\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \rangle| \lesssim \tau \sum_{m=0}^{n+1} \mathcal{F}_h^m. \quad (4.15d)$$

This is the only term whose bound is quadratic in the error with a coefficient that is not small, and it is therefore the term that forces the use of a Grönwall argument below. The boundary term involving $\delta_{h;\frac{3}{4}}^n$ carries no factor τ and reaches the level $n+1$, so it is absorbed into the left-hand side. Using (3.14), $\Gamma_h^{n;\gamma}(u_h^{n+1}) \geq \alpha_{\min}$ and the weighted Young inequality with weight $\alpha_{\min}/8$ and (4.10), we have

$$\begin{aligned} |\langle \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^n, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^{n+1} + e_h^n) \rangle| &\leq \frac{2}{\alpha_{\min}} \|\mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;\frac{3}{4}}^n\|^2 + \frac{\alpha_{\min}}{8} \|\mathcal{L}_h^{\frac{1}{2}+a}(e_h^{n+1} + e_h^n)\|^2 \\ &\lesssim \frac{2}{\alpha_{\min}} (h^{p+1-2a} + \tau^2)^2 + \frac{1}{2} \Gamma_h^{n;\gamma}(u_h^{n+1}) \mathcal{W}_h^{n+1}. \end{aligned} \quad (4.15e)$$

Note that this absorption requires no restriction on τ . By (4.11), Young’s inequality, (3.15) and $\tau \sum_{m=2}^n 1 \leq T$,

$$\tau \sum_{m=2}^n |\langle \mathcal{L}_h^{a-\frac{1}{2}} \partial_\tau \delta_{h;\frac{3}{4}}^m, \mathcal{L}_h^{\frac{1}{2}+a}(e_h^m + e_h^{m-1}) \rangle| \lesssim T(h^{p+1-2a} + \tau^2)^2 + \tau \sum_{m=0}^n \mathcal{F}_h^m. \quad (4.15f)$$

It is here that the summation by parts pays off. Without it the same term would carry a factor τ^{-1} and the defects, being of order τ^2 , would not suffice.

Inserting (4.15) into (4.14), subtracting the second summand of (4.15e) from both sides and using $\Gamma_h^{n;\gamma}(u_h^{n+1}) \geq \alpha_{\min}$, so that the left-hand side of (4.14) is at least $\min\{1, \alpha_{\min}\} \mathcal{F}_h^{n+1}$, we obtain constants $c_1, c_2 > 0$, independent of n, τ and h , with

$$\frac{1}{2} \min\{1, \alpha_{\min}\} \mathcal{F}_h^{n+1} \leq c_1(1+T)(h^{p+1-2a} + \tau^2)^2 + c_2 \tau \sum_{m=0}^{n+1} \mathcal{F}_h^m.$$

The sum on the right still reaches the level $n+1$. Choosing $\tau_1 := \min\{1, \alpha_{\min}\}/(4c_2)$ and restricting to $\tau \leq \tau_1$, that term is absorbed as well, and we arrive at

$$\mathcal{F}_h^{n+1} \leq c_3(h^{p+1-2a} + \tau^2)^2 + c_4 \tau \sum_{m=0}^n \mathcal{F}_h^m, \quad c_3, c_4 \text{ independent of } n, \tau, h. \quad (4.16)$$

Every step leading to (4.16) used the hypotheses of the theorem only at the levels $m \leq n$, together with the bounds [Lemma 4.1](#) supplies at the level $n+1$. The derivation is therefore valid verbatim with n replaced by any $k \leq n$, so that (4.16) holds for $\mathcal{F}_h^{k+1}$, $k = 1, \dots, n$, while $\mathcal{F}_h^0 + \mathcal{F}_h^1 \lesssim (h^{p+1-2a} + \tau^2)^2$ by (4.6) at $m = 0, 1$. The discrete Grönwall inequality applied to this family yields

$$\max_{0 \leq m \leq n+1} \mathcal{F}_h^m \leq c_5 e^{c_4 T} (h^{p+1-2a} + \tau^2)^2,$$

with c_5 depending only on the data of the problem. Finally, (3.15) at $m = n + 1$ gives

$$\|\mathcal{L}_h^{\frac{1}{2}+a} e_h^{n+1}\|^2 + \|\mathcal{L}_h^a \partial_\tau e_h^{n+1}\|^2 \leq (1 + C_\theta) \mathcal{F}_h^{n+1} \lesssim (h^{p+1-2a} + \tau^2)^2,$$

which is the first asserted estimate, and since $\mathcal{L}_h$ is uniformly positive definite on V_h , (3.7) with $a \in \{0, \frac{1}{4}\}$, $b = a + \frac{1}{2}$ gives the second,

$$\|\mathcal{L}_h^a e_h^{n+1}\| \leq \lambda_1^{-\frac{1}{2}} \|\mathcal{L}_h^{a+\frac{1}{2}} e_h^{n+1}\| \lesssim h^{p+1-2a} + \tau^2. \quad \blacksquare$$

4.3 Initial data and closure of the induction

It remains to specify the discrete initial data and to close the induction. Mirroring the starting values (2.3) of the time-discrete scheme, we set

$$u_h^0 = R_h u_0, \quad u_h^1 = R_h(u_0 + \tau v_0 + \frac{\tau^2}{2}(f^0 - q^0 \mathcal{L} u_0)) \quad (4.17)$$

that is, the Ritz projection of the Taylor expansion of u about $t = 0$. This is what the scheme (3.16) at $n = 0$ produces once the ghost value is eliminated by the centred approximation $u_h^{-1} = u_h^1 - 2\tau v_0$ of $\partial_t u(0) = v_0$ and the stiffness term is evaluated explicitly, so that the level $n = 1$ is obtained from an explicit step; $q^0 = \alpha^0 + \beta^0 \|\nabla u_0\|^2$ denotes the coefficient of the continuous equation at $t = 0$, whence $f^0 - q^0 \mathcal{L} u_0 = \partial_t^2 u(0)$. Since Theorem 4.4 assumes both $\partial_t u$ and $\partial_t^2 u$ in $L^\infty(0, T; D(\mathcal{L}^{\frac{3}{4}}))$, we have $\partial_t u \in W^{1,\infty}(0, T; D(\mathcal{L}^{\frac{3}{4}})) \hookrightarrow C([0, T]; D(\mathcal{L}^{\frac{3}{4}}))$, so that $\partial_t u$ has a $D(\mathcal{L}^{\frac{3}{4}})$ -valued trace at $t = 0$ and $v_0 = \partial_t u(0) \in D(\mathcal{L}^{\frac{3}{4}})$. This is therefore not an additional hypothesis below.

Lemma 4.5 (Error bounds for the fully discrete initial data). *Let Assumption 2.3 hold, let $a \in \{0, \frac{1}{4}\}$, and let u_h^0, u_h^1 be given by (4.17). Then, for $m = 0, 1$*

$$\begin{aligned} \|\mathcal{L}_h^{\frac{1}{2}+a} (R_h u(t_m) - u_h^m)\| + \|\mathcal{L}_h^a (R_h u(t_m) - u_h^m)\| &\lesssim h^{p+1-2a} + \tau^3, \\ \|\mathcal{L}_h^a \partial_\tau (R_h u(t_1) - u_h^1)\| &\lesssim h^{p+1-2a} + \tau^2, \end{aligned}$$

and the uniform bounds (4.6) hold for $m = 0, 1$, for τ and h sufficiently small.

Proof. Write $e_h^m = R_h u(t_m) - u_h^m$. At $m = 0$ we have $u_h^0 = R_h u_0 = R_h u(t_0)$, so $e_h^0 = 0$ and the first estimate is trivial. At $m = 1$, since R_h is linear,

$$e_h^1 = R_h(u(t_1) - u_0 - \tau v_0 - \frac{\tau^2}{2}(f^0 - q^0 \mathcal{L} u_0)),$$

and the bracket is estimated by Taylor expansion about $t = 0$, which, together with $\partial_t^2 u(0) = f^0 - q^0 \mathcal{L} u_0$, is bounded by $\tau^3 \|\mathcal{L}^b \partial_t^3 u\|_{L^\infty(L^2)}$ in the $\mathcal{L}^b$ -norm for $b \in \{\frac{1}{2}, \frac{1}{2} + a\}$. Applying (3.9) at a and (3.6) at $\frac{1}{2} + a$ to pass from R_h to the continuous norm gives

$$\|\mathcal{L}_h^a e_h^1\| \lesssim \tau^3 \|\mathcal{L}^{\frac{1}{2}} \partial_t^3 u\|_{L^\infty(L^2)}, \quad \|\mathcal{L}_h^{\frac{1}{2}+a} e_h^1\| \lesssim \tau^3 \|\mathcal{L}^{\frac{1}{2}+a} \partial_t^3 u\|_{L^\infty(L^2)},$$

both $\lesssim \tau^3$ by the standing hypothesis $\partial_t^3 u \in L^\infty(0, T; D(\mathcal{L}^{\frac{3}{4}}))$, which also controls the $\mathcal{L}^{\frac{1}{2}}$ -norm because $\|\mathcal{L}^{\frac{1}{2}} w\| \leq \lambda_1^{-\frac{1}{4}} \|\mathcal{L}^{\frac{3}{4}} w\|$. Since $e_h^0 = 0$, we have $\partial_\tau e_h^1 = \tau^{-1} e_h^1$, whence $\|\mathcal{L}_h^a \partial_\tau e_h^1\| \lesssim \tau^2$. Note that from the proof, for this choice of starting values the error relative to $R_h u$ carries no spatial contribution at all. The asserted bounds are purely temporal ones, written in the form in which Corollaries 4.6 and 4.7 consume them.

The uniform bounds (4.6) follow by the triangle inequality from these estimates, the stability of R_h ((3.9) for a and (3.6) for $\frac{1}{2} + a$) and the factor 2 in the definition of M_a and M'_a . $\blacksquare$

Corollary 4.6 (Convergence of the fully discrete scheme). *Let the assumptions of Theorem 4.4 hold, in particular, the mesh condition $\tau \leq \kappa h$ at the endpoint case $\theta = \frac{1}{4}$, and let the discrete initial data be given by (4.17). Then there exist $\tau_0, h_0 > 0$, independent of n , such that for all $0 < \tau \leq \tau_0$ and $0 < h \leq h_0$ the fully discrete solution $\{u_h^n\}_{n=0}^N$ exists, is unique, and satisfies*

$$\max_{0 \leq n \leq N} \|\mathcal{L}_h^{\frac{1}{2}+a} (R_h u(t_n) - u_h^n)\| + \max_{1 \leq n \leq N} \|\mathcal{L}_h^a \partial_\tau (R_h u(t_n) - u_h^n)\| \lesssim h^{p+1-2a} + \tau^2.$$

Proof. Let n^* be the largest index for which the hypotheses of [Theorem 4.4](#), i.e., the uniform bounds (4.6) and the error estimate at all levels $m \leq n^*$ are satisfied. By [Lemma 4.5](#) they hold for $m = 0, 1$, so $n^* \geq 1$. If $n^* < N$, then [Lemma 4.1](#) provides $u_h^{n^*+1}$ and [Theorem 4.4](#) extends the error estimate to that level, the uniform bounds following from it by the triangle inequality together with the stability of R_h and the factor 2 in the definition of M_a and M'_a , contradicting the maximality of n^* . Hence $n^* = N$. The argument is admissible because the constants in [Theorem 4.4](#) are independent of n , τ and h , so a single pair (τ_0, h_0) works at every step. $\blacksquare$

Corollary 4.7 (Optimal-order error of the fully discrete scheme). *Under the assumptions of [Corollary 4.6](#) the full error of the fully discrete scheme satisfies*

$$\max_{0 \leq n \leq N} \|u(t_n) - u_h^n\| \lesssim h^{p+1} + \tau^2, \quad \max_{0 \leq n \leq N} \|\nabla(u(t_n) - u_h^n)\| \lesssim h^p + \tau^2,$$

both of which are of optimal order in h for the space V_h of continuous piecewise polynomials of degree p .

Proof. We split $u(t_n) - u_h^n = (\hat{u}^n - R_h \hat{u}^n) + e_h^n$. The approximation estimates for the Ritz projection with $s = p + 1$, admissible because $u \in L^\infty(0, T; H^{p+1}(\Omega))$ and $u(t) \in H_0^1(\Omega)$, give

$$\|\hat{u}^n - R_h \hat{u}^n\| + h \|\nabla(\hat{u}^n - R_h \hat{u}^n)\| \lesssim h^{p+1} |\hat{u}^n|_{H^{p+1}(\Omega)} \lesssim h^{p+1}.$$

For the remaining part, $e_h^n \in V_h$, so that (3.7) with $a = 0$ and $b = \frac{1}{2}$ together with [Corollary 4.6](#) yields $\|e_h^n\| \leq \lambda_1^{-\frac{1}{2}} \|\mathcal{L}_h^{\frac{1}{2}} e_h^n\| \lesssim h^{p+1} + \tau^2$, while $\|\nabla e_h^n\| = \|\mathcal{L}_h^{\frac{1}{2}} e_h^n\| \lesssim h^{p+1} + \tau^2$. The triangle inequality now gives both assertions, the second because $h^{p+1} \leq h^p$ for $h \leq 1$. $\blacksquare$

5 Numerical experiments

In this section we test the fully discrete scheme numerically on manufactured solutions. Throughout, $\Omega = (0, 1)^2$, the spatial discretization uses continuous Lagrange elements of degree $p = 1, 2, 3$, and the source term f is constructed from the prescribed exact solution and the coefficients given in the respective examples. The parameter $\lambda \in \mathbb{N}$ controls the spatial oscillation of the exact solution.

The scheme (3.16) is used with $\theta = \gamma = \frac{1}{2}$ and $\eta = \frac{1}{4}$, so that it reduces to the symmetric three-layer form (2.2), in particular $\theta > \frac{1}{4}$, so no mesh condition is active. The spatial convergence studies use the uniform mesh sequence $h = 2^{-2}, 2^{-3}, \dots, 2^{-6}$, and use a fixed number of time steps $N = 2^{12}$ such that $\tau = T2^{-12}$ is defined accordingly. This keeps the temporal contribution below the spatial error for most data points. The temporal convergence studies use the fixed mesh $h = 2^{-6}$ and $p = 3$ for time steps $\tau = T2^{-k}$, $k = 9, \dots, 12$.

The nonlinear system at each time level is solved by the Picard iteration (4.1), and the linear system of each Picard step by a conjugate-gradient (cg-) method with relative and absolute tolerance 10^{-12} and at most 1000 iterations. The Picard loop is started from the analysed initial guess $u_{h;0}^{n+1} = u_h^n$ and is stopped as soon as

$$\max \left\{ \frac{|\Gamma_h^{n;\gamma}(u_{h;\ell+1}^{n+1}) - \Gamma_h^{n;\gamma}(u_{h;\ell}^{n+1})|}{\max\{|\Gamma_h^{n;\gamma}(u_{h;\ell+1}^{n+1})|, 1\}}, \frac{\|u_{h;\ell+1}^{n+1} - u_{h;\ell}^{n+1}\|_\infty}{\max\{\|u_{h;\ell}^{n+1}\|_\infty, 1\}} \right\} < \varepsilon_{\text{NL}} := 10^{-10}, \quad (5.1)$$

with at most $L_{\text{max}} = 10$ iterations permitted. Note that here $\|\cdot\|_\infty$ denotes the maximum over the coefficient vectors with respect to the nodal basis of V_h . In every run reported below the criterion was met before L_{max} was reached, at every time level. Note that the tolerances for the nonlinearity and the cg-solve are tighter than required such that the errors are not polluted in any form.

The discrete initial data follow the Taylor start (4.17). The numerical experiments are implemented in Python using the FEniCSx software framework ([4, 31, 30, 1]). The versions of all software dependencies are specified in the enclosed `pyproject.toml` file in the repository. The scheme is realised in the auxiliary-variable form of [Remark 5.1](#). Errors are evaluated at the final time only, which reduces the cost per run. This is consistent with the theory, whose estimates are uniform over the time levels and therefore hold at the final time in particular. Further, it is the natural choice in practice, since the accumulation of the error means that the largest relative error is typically attained there.

Remark 5.1 (Implementation). In practice, we try to avoid reassembling the right-hand side in every Picard iteration of (4.1) such that the assembly cost is only paid at each time step n . This can be achieved by introducing the auxiliary variable

$$v_h^{n+1} := u_h^{n+1;\theta}. \quad (5.2)$$

The idea is that this auxiliary variable cancels the terms that contain the nonlinearity on the right-hand side of (4.1). Multiplying (3.16) by $\tau^2\theta$ and plugging (5.2) into (3.16), the nonlinear scheme reads

$$v_h^{n+1} + \tau^2\theta\Gamma_h^{n;\gamma}(u_h^{n+1})\mathcal{L}_h v_h^{n+1} = \tau^2\theta f^{n+1;\eta} + u_h^n.$$

The weighted source term $f^{n+1;\eta} = \eta f((n+1)\tau) + (1-2\eta)f(n\tau) + \eta f((n-1)\tau)$ is realised by the standard quadrature used in FEniCSx of the L^2 scalar product of the right-hand side. This introduces only a quadrature error which is of greater order than an interpolation error and can thus be disregarded. Then, we solve this nonlinear equation for every time step with the following Picard iteration: We seek a sequence of functions $\{v_{h;\ell+1}^{n+1}\}_{\ell \geq 0} \subset V_h$ that solves the *fully discrete linear iteration scheme*

$$v_{h;\ell+1}^{n+1} + \tau^2\theta\Gamma_h^{n;\gamma}(u_{h;\ell}^{n+1})\mathcal{L}_h v_{h;\ell+1}^{n+1} = \tau^2\theta f^{n+1;\eta} + u_h^n$$

with the initial guess $u_{h;0}^{n+1} = u_h^n$ and update step

$$\theta u_{h;\ell+1}^{n+1} = v_{h;\ell+1}^{n+1} - (1-2\theta)u_h^n - \theta u_h^{n-1}.$$

Example 1: Optimal spatial and temporal convergence studies. We present a first very simple constructed example to show that the proposed method converges with the optimal orders. We set the time interval $[0, \frac{\pi}{2}]$ with $\alpha(t) = 1 + t^2$, and $\beta(t) = 2 - \sin(t)$. Then we solve the PDE with exact solution

$$u(x, y, t) = \sin(\lambda\pi x) \sin(\lambda\pi y) \sin(t).$$

Figure 5.1 reports the relative L^2 (top and bottom) and H_0^1 (middle) errors at the final time $T = \frac{\pi}{2}$ for $\lambda = 1$ (left) and $\lambda = 3$ (right), for polynomial degrees $p = 1, 2, 3$ under uniform mesh refinement. It confirms the optimal spatial orders $p+1$ in L^2 and p in H^1 predicted by Corollary 4.7. The time step size is fixed at $\tau = \frac{\pi}{2} \cdot 2^{-12}$. Note that the final point for $h = 2^{-6}$ and $p = 3$ does not attain the optimal order, because the error has reached the temporal plateau. The bottom plot shows that at this time step size the temporal contribution is of the same order of magnitude. That final point is therefore not a pure spatial error, and the degradation of the rate is the expected consequence rather than a property of the method. The bottom plot also shows the behaviour under refinement of the time step at a fixed mesh, for $p = 3$ and $h = 2^{-6}$. The error decreases at the second order predicted by Corollary 4.7 until it reaches the level of the spatial discretization error, after which further refinement in time brings no gain and the observed rate deteriorates. The observed orders in space and time are independent of the spatial oscillation parameter λ , although a larger λ increases the size of the error in Figure 5.1. Increasing λ has no significant effect on the Picard iteration either: the stopping criterion (5.1) was met at every time step, and $L_{\max}$ was never reached.

Example 2: Conservation of the discrete energy. In this example we observe the conservation of the fully discrete energy. We choose constant coefficients $\alpha = \beta = 1$ and a vanishing source f . Further, the initial conditions are given by $u_0 = 0$ and $v_0 = \sin(\pi x) \sin(\pi y)$. We choose $h = 2^{-5}$ and $p = 2$, with $T = 10^2 \cdot 2\pi$ and $\tau = 2\pi \cdot 2^{-13}$ for the conservation study, and $\tau = 2\pi \cdot 2^{-k}$ for $k = 11, \dots, 14$ for the convergence study.

In this setting the scheme conserves the fully discrete energy of Proposition 3.3 for the parameter $\theta = \frac{1}{2}$. The top plot of Figure 5.2 shows the relative drift of the energy, that is, the relative difference between the energy at time t and the energy at the initial time. We observe that over 100 periods the energy is conserved for the three parameters $\gamma \in \{0, \frac{1}{4}, \frac{1}{2}\}$. Since the identity holds for the exact solution of the nonlinear scheme only, the remaining error grows linearly in t . It stems from the nonlinear solve and is governed by the Picard tolerance, set to $\varepsilon_{\text{NL}} = 10^{-12}$ with $L_{\max} = 30$. The case

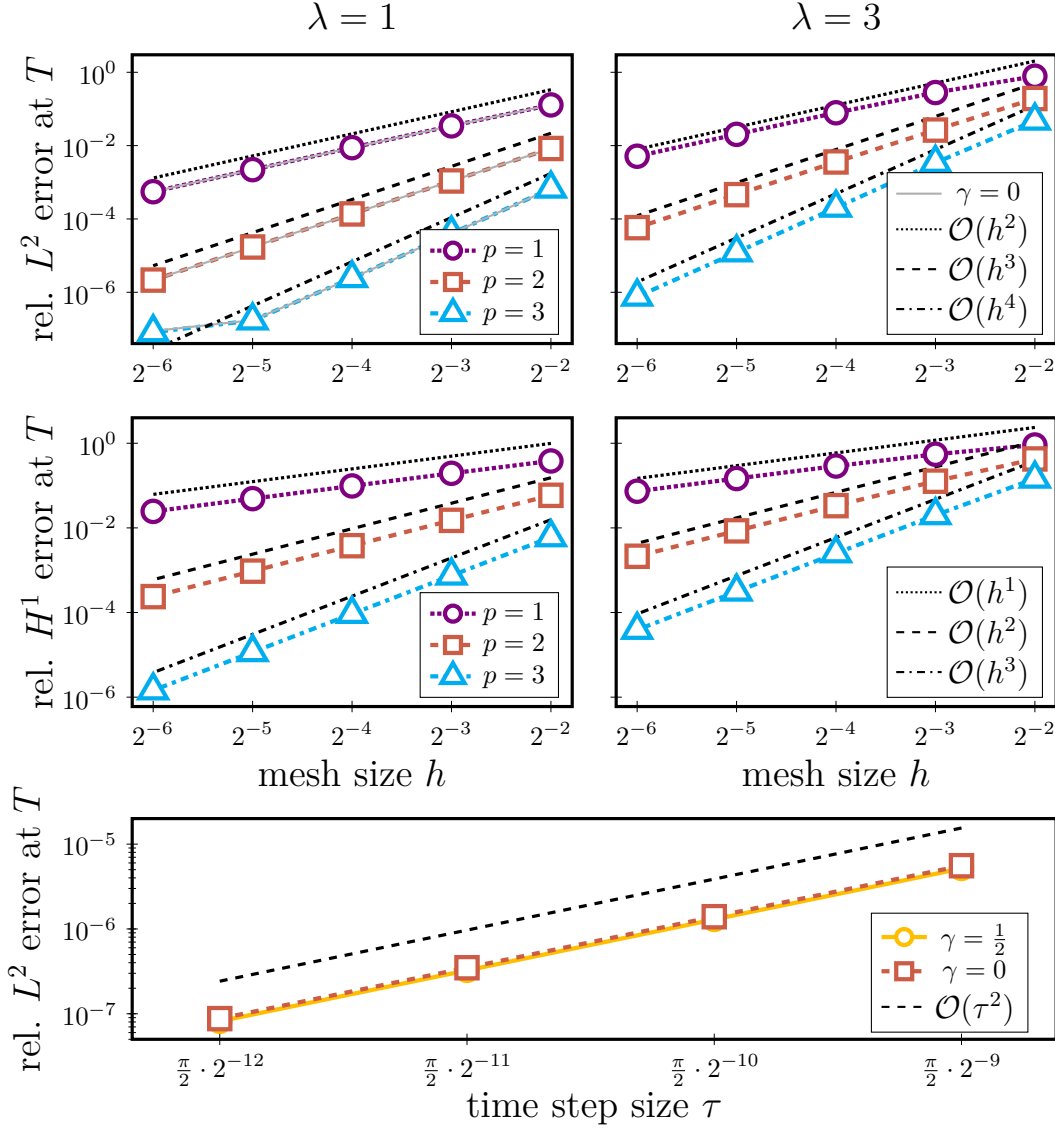


Figure 5.1: Example 1: relative L^2 errors (top and bottom) and H_0^1 errors (mid) at the last time step $T = \frac{\pi}{2}$ for $\lambda = 1$ (left) and $\lambda = 3$ (right). The order $p + 1$ for L^2 and p for H^1 in space are attained, with the exception of the last $p = 3$ point for $\lambda = 1$, where the fixed time step contributes at the same order. Further, the bottom plot shows the second order convergence in time predicted by the theory. Additionally, the plot on the top left and the plot on the bottom show the same convergence behaviour and similarly sized error for the locally linear scheme $\gamma = 0$.

$\gamma = 0$ confirms this reading. By Remark 3.5 the scheme is then linearly implicit, no Picard iteration is performed, and the energy is conserved all the same. Accordingly, its drift stays at the level of round-off over the whole interval and shows no growth in t , whereas the other two curves grow linearly at the rate set by ε_{NL} .

The bottom plot of Figure 5.2 shows the relative spread of the energy over the whole run, that is, the largest relative difference between any two of its values, plotted against the time step size τ . We observe that for $\gamma = \frac{1}{2}$ this spread is independent of τ when $\theta = \frac{1}{2}$, whereas for $\theta \neq \frac{1}{2}$ the energy is

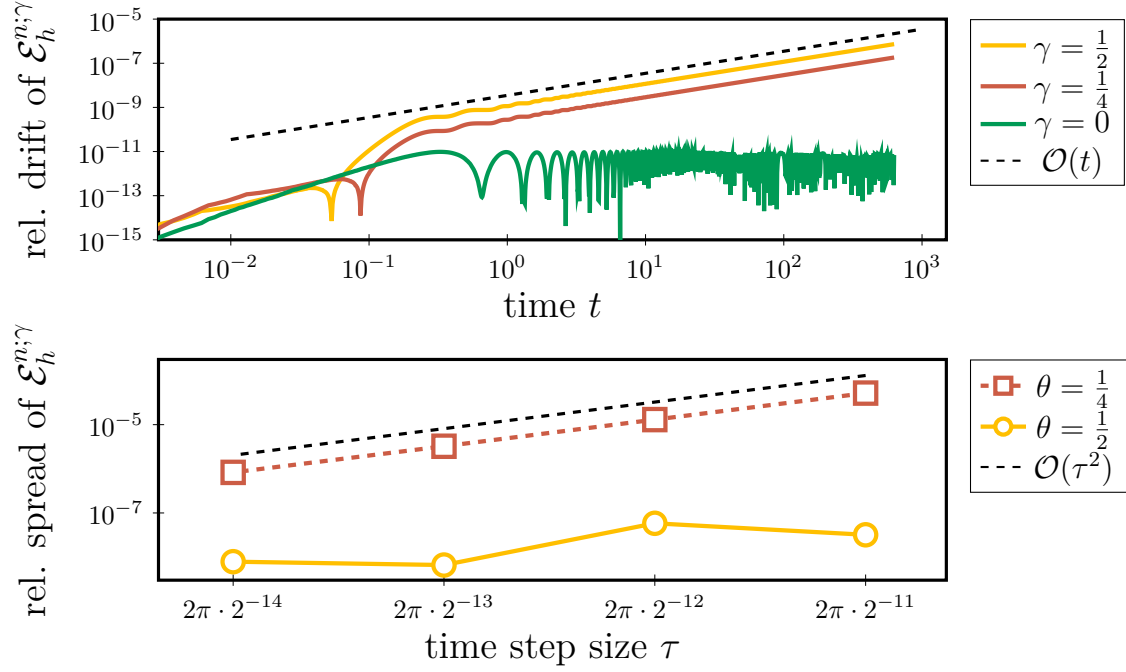


Figure 5.2: Example 2: relative drift (top) and relative spread (bottom) of the fully discrete energy. We observe that for $\theta = \frac{1}{2}$ the energy is conserved over 100 periods up to an error induced by the Picard tolerance, which grows linearly in time. Further, we observe for $\gamma = \frac{1}{2}$ that the spread of the energy does not depend on the time step size τ when $\theta = \frac{1}{2}$, whereas for $\theta = \frac{1}{4}$ the energy is conserved only to second order in τ .

not conserved but approximated to second order in τ , as predicted by [Remark 3.4](#).

6 Conclusion

We have combined the symmetric three-layer time discretization of [19] with a conforming finite element method of arbitrary polynomial degree p and a Picard iteration for the nonlinearity, and analysed the resulting fully discrete scheme for a two-dimensional Kirchhoff-type equation with time-dependent coefficients. The nonlinear equation solved at each time level was shown to have a unique solution, to which the Picard iteration converges (Lemma 4.1). For the homogeneous problem with constant coefficients and the parameter choice $\theta = \frac{1}{2}$, the exact solution of the scheme conserves the discrete energy $\mathcal{E}_h^{n;\gamma}$ identically (Proposition 3.3). This is the fully discrete counterpart of the temporal invariants of [19], see Remark 3.5, and it survives the spatial discretization because the method is conforming. We further proved convergence of optimal order $\mathcal{O}(h^{p+1} + \tau^2)$ in L^2 and $\mathcal{O}(h^p + \tau^2)$ in H^1 (Corollary 4.7) for the full error $u(t_n) - u_h^n$. The numerical experiments of Section 5 exhibit exactly these rates: second order in time up to the point at which the spatial error becomes dominant, and the full order h^{p+1} in the L^2 norm of the full error.

Several questions remain open. Both convergence results hold at every level of the grid on the interval $[0, T]$ of Assumption 2.3, itself local since $T < T^{**}$. Extending them up to the blow-up time would require a different argument. Both also rest on regularity going well beyond that assumption. Theorem 4.4 demands four time derivatives in the operator-domain scale and H^{p+1} spatial regularity of u , $\partial_t u$, $\partial_t^2 u$ and f , and whether the optimal rate survives under weaker hypotheses we do not know. At the endpoint $\theta = \frac{1}{4}$ the three-layer displacement energy is non-coercive in the displacement norm, whether it is necessary for the scheme or only for this method of proof is open. At $\theta = \frac{1}{2}$ the scheme conserves $\mathcal{E}_h^{n;\gamma}$ exactly (Proposition 3.3), but we have established no energy inequality for the whole range with time-dependent coefficients, and claim no unconditional stability there.

Verifying Proposition 3.3 for the finite element spaces used in the convergence studies, where the Picard and conjugate-gradient tolerances also enter, lies beyond the present scope, as does a systematic study of the sensitivity to those tolerances.

Declarations

Funding. The work of B. Dörich was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173. This work by A. Rupp was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 577175348. And in part supported by the Deutsche Forschungsgemeinschaft under Germany’s Excellence Strategy – EXC-2047/2 – 390685813, through F. Krumbiegel’s and A. Rupp’s stay at the Hausdorff Research Institute for Mathematics. Additionally, Z. Vashakidze was supported by the Shota Rustaveli National Science Foundation of Georgia (SRNSFG) under Grant No. FR-25-215. The authors declare no further funding for this work.

Competing interests. The authors declare that they have no competing interests.

Data and code availability. The raw error data underlying every figure, together with the sources of all figures, are contained in the `figures/` directory of the accompanying repository, <https://github.com/FelixKrumbiegel/Kirchhoff-conservative-FEM>. The `software/` directory of the same repository contains the implementation that produced the reported runs: a PYTHON package built on FENICSx/DOLFINx.

Declaration of generative AI and AI-assisted technologies. During the preparation of this work, the authors used Claude in order to improve formulation for parts of the manuscript and for parts of the code to create all-encompassing scripts for the various examples and to improve presentation of the software. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

- [1] M. S. Alnæs et al. “Unified Form Language: A domain-specific language for weak formulations of partial differential equations”. In: *ACM Transactions on Mathematical Software* 40.2 (2014), 9:1–9:37. URL: <https://doi.org/10.1145/2566630>.

- [2] A. Arosio and S. Panizzi. “On the well-posedness of the Kirchhoff string”. In: *Trans. Amer. Math. Soc.* 348.1 (1996), pp. 305–330. ISSN: 0002-9947,1088-6850. URL: <https://doi.org/10.1090/S0002-9947-96-01532-2>.
- [3] L. A. Bales and V. A. Dougalis. “Cosine methods for nonlinear second-order hyperbolic equations”. In: *Math. Comp.* 52.186 (1989), 299–319, S15–S33. ISSN: 0025-5718. URL: <https://doi.org/10.2307/2008469>.
- [4] I. A. Baratta et al. *DOLFINx: The next generation FEniCS problem solving environment*. 2023. URL: <https://doi.org/10.5281/zenodo.10447666>.
- [5] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976, pp. x+207.
- [6] J. H. Bramble, J. E. Pasciak, and O. Steinbach. “On the stability of the L^2 projection in $H^1(\Omega)$ ”. In: *Math. Comp.* 71.237 (2002), pp. 147–156. ISSN: 0025-5718,1088-6842. URL: <https://doi.org/10.1090/S0025-5718-01-01314-X>.
- [7] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*. Third. Vol. 15. Texts in Applied Mathematics. Springer, New York, 2008, pp. xviii+397. ISBN: 978-0-387-75933-3. URL: <https://doi.org/10.1007/978-0-387-75934-0>.
- [8] I. Christie and J. M. Sanz-Serna. “A Galerkin method for a nonlinear integro-differential wave system”. In: *Comput. Methods Appl. Mech. Engrg.* 44.2 (1984), pp. 229–237. ISSN: 0045-7825. URL: [https://doi.org/10.1016/0045-7825\(84\)90144-0](https://doi.org/10.1016/0045-7825(84)90144-0).
- [9] Ph. Clément. “Approximation by finite element functions using local regularization”. In: *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér.* 9.no. , no. R-2 (1975), pp. 77–84. ISSN: 0397-9342,2777-3493.
- [10] M. Crouzeix and V. Thomée. “The stability in L_p and W_p^1 of the L_2 -projection onto finite element function spaces”. In: *Math. Comp.* 48.178 (1987), pp. 521–532. ISSN: 0025-5718,1088-6842. URL: <https://doi.org/10.2307/2007825>.
- [11] R. Z. Dautov and M. V. Ivanova. “A conservative finite element scheme for the Kirchhoff equation”. In: *Uch. Zap. Kazan. Univ. Ser. Fiz.-Mat. Nauki* 165.2 (2023), pp. 115–131. ISSN: 2541-7746,2500-2198. URL: <https://doi.org/10.26907/2541-7746.2023.2.115-131>.
- [12] B. Dörich. “Strong norm error bounds for quasilinear wave equations under weak CFL-type conditions”. In: *Found. Comput. Math.* 25.1 (2025), pp. 303–350. ISSN: 1615-3375. URL: <https://doi.org/10.1007/s10208-024-09639-w>.
- [13] B. Dörich and V. Nikolić. “Robust fully discrete error bounds for the Kuznetsov equations in the inviscid limit”. In: *Numer. Math.* (2026). Online first. ISSN: 0029-599X. URL: <https://doi.org/10.1007/s00211-026-01574-y>.
- [14] A. Ern and J.-L. Guermond. “Finite element quasi-interpolation and best approximation”. In: *ESAIM Math. Model. Numer. Anal.* 51.4 (2017), pp. 1367–1385. ISSN: 2822-7840,2804-7214. URL: <https://doi.org/10.1051/m2an/2016066>.
- [15] T. Gudi. “Finite element method for a nonlocal problem of Kirchhoff type”. In: *SIAM J. Numer. Anal.* 50.2 (2012), pp. 657–668. ISSN: 0036-1429,1095-7170. URL: <https://doi.org/10.1137/110822931>.
- [16] M. Hochbruck and B. Maier. “Error analysis for space discretizations of quasilinear wave-type equations”. In: *IMA J. Numer. Anal.* 42.3 (2022), pp. 1963–1990. ISSN: 0272-4979,1464-3642. URL: <https://doi.org/10.1093/imanum/drab073>.
- [17] G. Kirchhoff. *Vorlesungen über mathematische Physik. Mechanik*. German. Vol. 1. B. G. Teubner, Leipzig, 1876, pp. ix+466.

- [18] P. Knabner and L. Angermann. *Numerical methods for elliptic and parabolic partial differential equations*. Vol. 44. Texts in Applied Mathematics. Extended edition; with contributions by Andreas Rupp. © 2021. Springer, Cham, 2021, pp. xix+802. ISBN: 978-3-030-79384-5. URL: <https://doi.org/10.1007/978-3-030-79385-2>. eBook ISBN: 978-3-030-79385-2.
- [19] F. Krumbiegel et al. *Conservative Three-Layer Schemes for Kirchhoff-Type Equations*. 2026. arXiv: [2609.27707](https://arxiv.org/abs/2609.27707) [math.NA]. URL: <https://arxiv.org/abs/2609.27707>.
- [20] B. Maier. “Error analysis for full discretizations of quasilinear wave-type equations with two variants of the implicit midpoint rule”. In: *IMA J. Numer. Anal.* 43.2 (2023), pp. 1149–1180. ISSN: 0272-4979. URL: <https://doi.org/10.1093/imanum/drac010>.
- [21] C. G. Makridakis. “Finite element approximations of nonlinear elastic waves”. In: *Math. Comp.* 61.204 (1993), pp. 569–594. ISSN: 0025-5718,1088-6842. URL: <https://doi.org/10.2307/2153241>.
- [22] G. Mallik. “A hybrid high-order finite element method for a nonlocal nonlinear problem of Kirchhoff type”. In: *Appl. Numer. Math.* 230 (2026), pp. 1–12. ISSN: 0168-9274,1873-5460. URL: <https://doi.org/10.1016/j.apnum.2026.07.005>.
- [23] V. Nikolić and B. Wohlmuth. “A priori error estimates for the finite element approximation of Westervelt’s quasi-linear acoustic wave equation”. In: *SIAM J. Numer. Anal.* 57.4 (2019), pp. 1897–1918. ISSN: 0036-1429,1095-7170. URL: <https://doi.org/10.1137/19M1240873>.
- [24] J. Peradze. “A numerical algorithm for the nonlinear Kirchhoff string equation”. In: *Numer. Math.* 102.2 (2005), pp. 311–342. ISSN: 0029-599X,0945-3245. URL: <https://doi.org/10.1007/s00211-005-0642-1>.
- [25] J. Rogava and M. Tsiklauri. “Convergence of a semi-discrete scheme for an abstract nonlinear second order evolution equation”. In: *Appl. Numer. Math.* 75 (2014), pp. 22–36. ISSN: 0168-9274. URL: <https://doi.org/10.1016/j.apnum.2012.09.004>.
- [26] J. Rogava and M. Tsiklauri. “On local convergence of a symmetric semi-discrete scheme for an abstract analogue of the Kirchhoff equation”. In: *J. Comput. Appl. Math.* 236.15 (2012), pp. 3654–3664. ISSN: 0377-0427. URL: <https://doi.org/10.1016/j.cam.2011.07.003>.
- [27] J. Rogava, M. Tsiklauri, and Z. Vashakidze. “On stability and convergence of a three-layer semi-discrete scheme for an abstract analogue of the Ball integro-differential equation”. In: *J. Math. Anal. Appl.* 518.1 (2023), Paper No. 126664, 25. ISSN: 0022-247X. URL: <https://doi.org/10.1016/j.jmaa.2022.126664>.
- [28] J. Rogava and Z. Vashakidze. “On convergence of a three-layer semi-discrete scheme for the non-linear dynamic string equation of Kirchhoff-type with time-dependent coefficients”. In: *ZAMM Z. Angew. Math. Mech.* 104.4 (2024), Paper No. e202300608, 36. ISSN: 0044-2267. URL: <https://doi.org/10.1002/zamm.202300608>.
- [29] C. Samanta et al. “Nonlinear nanomechanical resonators approaching the quantum ground state”. In: *Nat. Phys.* 19.9 (2023), pp. 1340–1344. URL: <https://doi.org/10.1038/s41567-023-02065-9>.
- [30] M. W. Scroggs et al. “Basix: a runtime finite element basis evaluation library”. In: *Journal of Open Source Software* 7.73 (2022), p. 3982. URL: <https://doi.org/10.21105/joss.03982>.
- [31] M. W. Scroggs et al. “Construction of arbitrary order finite element degree-of-freedom maps on polygonal and polyhedral cell meshes”. In: *ACM Transactions on Mathematical Software* 48.2 (2022), 18:1–18:23. URL: <https://doi.org/10.1145/3524456>.
- [32] C. P. Sun et al. “Quantum transducers: Integrating transmission lines and nanomechanical resonators via charge qubits”. In: *Phys. Rev. A* 73.2 (2006), p. 022318. URL: <https://doi.org/10.1103/PhysRevA.73.022318>.

- [33] Z. Vashakidze. “An application of the Legendre polynomials for the numerical solution of the nonlinear dynamical Kirchhoff string equation”. In: *Mem. Differ. Equ. Math. Phys.* 79 (2020), pp. 107–119. ISSN: 1512-0015.
- [34] Z. Vashakidze. “On the convergence of a three-layer semi-discrete scheme for the nonlinear dynamic Kirchhoff string equation”. In: *Georgian Math. J.* 29.4 (2022), pp. 615–627. ISSN: 1072-947X. URL: <https://doi.org/10.1515/gmj-2022-2149>.
- [35] X. Yao, D. Hoch, and M. Poot. “Relaxation and dynamics of stressed predisplaced string resonators”. In: *Phys. Rev. B* 106.17 (2022), p. 174109. URL: <https://doi.org/10.1103/PhysRevB.106.174109>.

Appendix

A.1 Proof of the full-order convergence

Proof of Lemma 4.3. In a first step, we estimate the difference in the nonlinearity in the first sum on the right-hand side of (4.14). Abbreviating $\Lambda^m := \gamma \|\mathcal{L}_h^{\frac{1}{2}} u_h^{m+1}\|^2 + (1-2\gamma) \|\mathcal{L}_h^{\frac{1}{2}} u_h^m\|^2 + \gamma \|\mathcal{L}_h^{\frac{1}{2}} u_h^{m-1}\|^2$, so that $\Gamma_h^{m;\gamma}(u_h^{m+1}) = \alpha^m + \beta^m \Lambda^m$. Adding and subtracting $\beta^m \Lambda^{m-1}$, we have that

$$|\Gamma_h^{m;\gamma}(u_h^{m+1}) - \Gamma_h^{m-1;\gamma}(u_h^m)| \leq |\alpha^m - \alpha^{m-1}| + |\beta^m - \beta^{m-1}| \Lambda^{m-1} + \beta_{\max} |\Lambda^m - \Lambda^{m-1}|. \quad (\text{A.1a})$$

The first two terms are $\lesssim \tau$ by the Taylor expansion of the coefficients $\alpha, \beta \in C^1[0, T]$ and the boundedness of the fully discrete solution at the previous time steps, using $\|\mathcal{L}_h^{\frac{1}{2}} \varphi_h\|^2 = \langle \mathcal{L}_h^{\frac{3}{4}} \varphi_h, \mathcal{L}_h^{\frac{1}{4}} \varphi_h \rangle$. For the third we apply the proportionality argument (4.4) to each of the three pairs of consecutive levels making up $\Lambda^m - \Lambda^{m-1}$,

$$\begin{aligned} \|\mathcal{L}_h^{\frac{1}{2}} u_h^{m+1}\|^2 - \|\mathcal{L}_h^{\frac{1}{2}} u_h^m\|^2 &= \langle \mathcal{L}_h^{\frac{3}{4}}(u_h^{m+1} + u_h^m), \mathcal{L}_h^{\frac{1}{4}}(u_h^{m+1} - u_h^m) \rangle \\ &\leq \tau (\|\mathcal{L}_h^{\frac{3}{4}} u_h^{m+1}\| + \|\mathcal{L}_h^{\frac{3}{4}} u_h^m\|) \|\mathcal{L}_h^{\frac{1}{4}} \partial_\tau u_h^{m+1}\| \lesssim \tau, \end{aligned} \quad (\text{A.1b})$$

and likewise at the levels $m, m-1$ and $m-1, m-2$, so that $|\Gamma_h^{m;\gamma}(u_h^{m+1}) - \Gamma_h^{m-1;\gamma}(u_h^m)| \lesssim \tau$. For $m < n$ these norms come from (4.6), whereas at the top index $m = n$, where Λ^n reaches u_h^{n+1} and $\partial_\tau u_h^{n+1}$, they come from the generic bounds Lemma 4.1 supplies for the level it constructs, uniformly in n, h and τ . Overall, (A.1) gives the bound (4.8). Note that this bound does not depend on the parameter a .

Next, we bound the defects $\delta_{h;\frac{1}{4}}^m$. The defect consists of the terms $\delta_{h;\frac{1}{4}}^m = \delta_{h;f}^m + \delta_{h;1}^m$. Then, for the first defect $\delta_{h;f}^m$ we have that

$$\begin{aligned} \langle \delta_{h;f}^m, \mathcal{L}_h^{2a}(\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \rangle &= \langle \mathcal{L}_h^a(P_h f^m - f_h^{m+1;\eta}), \mathcal{L}_h^a(\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \rangle \\ &\leq (\|\mathcal{L}_h^a(P_h f^m - J_h f^m)\| + \|\mathcal{L}_h^a(J_h f^m - f_h^{m+1;\eta})\|) \|\mathcal{L}_h^a(\partial_\tau e_h^{m+1} + \partial_\tau e_h^m)\|. \end{aligned} \quad (\text{A.2a})$$

For the first term in the sum we observe that $J_h f^m \in V_h$, so that $P_h f^m - J_h f^m = P_h(f^m - J_h f^m)$, and apply (3.5) at $a \in \{0, \frac{1}{4}\}$ followed by the approximation estimate (3.4), obtaining

$$\|\mathcal{L}_h^a(P_h f^m - J_h f^m)\| \lesssim \|\mathcal{L}_h^a(f^m - J_h f^m)\| \lesssim h^{p+1-2a}, \quad (\text{A.2b})$$

where the last step uses (3.4) with $s = p+1$, which is admissible by the assumed regularity $f \in L^\infty(0, T; H^{p+1}(\Omega) \cap H_0^1(\Omega))$. For the second term we use that $f_h^m = J_h f(m\tau)$ for every m , so that $J_h f^m - f_h^{m+1;\eta} = J_h(f^m - f^{m+1;\eta})$, and apply (3.5) at $a \in \{0, \frac{1}{4}\}$ directly to this discrete function, followed by a Taylor approximation of the source term:

$$\|\mathcal{L}_h^a(J_h f^m - f_h^{m+1;\eta})\| = \|\mathcal{L}_h^a J_h(f^m - f^{m+1;\eta})\| \lesssim \|\mathcal{L}_h^a(f^m - f^{m+1;\eta})\| \lesssim \tau^2 \eta \|\mathcal{L}_h^a \partial_t^2 f^m\| \lesssim \tau^2, \quad (\text{A.2c})$$

where the last step uses the assumed regularity $\partial_t^2 f \in L^\infty(0, T; D(\mathcal{L}_h^{\frac{1}{4}}))$. Similarly, we may treat the second defect $\delta_{h;1}^m$. We use a Taylor expansion of the solution, stability and approximation estimates to obtain

$$\begin{aligned} \frac{1}{2} \langle \delta_{h;1}^m, \mathcal{L}_h^{2a} (\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \rangle &= \frac{1}{2} \langle \mathcal{L}_h^a (\partial_\tau^2 R_h \hat{u}^{m+1} - P_h \partial_t^2 \hat{u}^m), \mathcal{L}_h^a (\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \rangle \\ &\leq \frac{1}{2} (\| \mathcal{L}_h^a (R_h - P_h) \partial_\tau^2 \hat{u}^{m+1} \| + \| \mathcal{L}_h^a P_h (\partial_\tau^2 \hat{u}^{m+1} - \partial_t^2 \hat{u}^m) \|) \| \mathcal{L}_h^a (\partial_\tau e_h^{m+1} + \partial_\tau e_h^m) \|, \end{aligned} \quad (\text{A.2d})$$

where we added and subtracted $P_h \partial_\tau^2 \hat{u}^{m+1}$ and used $\partial_\tau^2 R_h \hat{u}^{m+1} = R_h \partial_\tau^2 \hat{u}^{m+1}$. Note that R_h is applied only to $\partial_\tau^2 \hat{u}^{m+1}$, a linear combination of $\hat{u}^{m+1}, \hat{u}^m, \hat{u}^{m-1} \in D(\mathcal{L}^{\frac{3}{4}}) \subset H_0^1(\Omega)$, where it is defined. For the first term, $(R_h - P_h) \partial_\tau^2 \hat{u}^{m+1} \in V_h$, so (3.5) at $a \in \{0, \frac{1}{4}\}$ passes to the continuous norm and (3.4) with $s = p + 1$ applies to each summand,

$$\| \mathcal{L}_h^a (R_h - P_h) \partial_\tau^2 \hat{u}^{m+1} \| \lesssim \| \mathcal{L}_h^a (R_h \partial_\tau^2 \hat{u}^{m+1} - \partial_\tau^2 \hat{u}^{m+1}) \| + \| \mathcal{L}_h^a (\partial_\tau^2 \hat{u}^{m+1} - P_h \partial_\tau^2 \hat{u}^{m+1}) \| \lesssim h^{p+1-2a}, \quad (\text{A.2e})$$

because $\partial_\tau^2 \hat{u}^{m+1}$ is a weighted average of $\partial_t^2 u$ over $[t_{m-1}, t_{m+1}]$, whence $|\partial_\tau^2 \hat{u}^{m+1}|_{H^{p+1}(\Omega)} \leq \|\partial_t^2 u\|_{L^\infty(0,T;H^{p+1}(\Omega))}$, and, for the temporal part,

$$\| \mathcal{L}_h^a P_h (\partial_\tau^2 \hat{u}^{m+1} - \partial_t^2 \hat{u}^m) \| \lesssim \| \mathcal{L}_h^a (\partial_\tau^2 \hat{u}^{m+1} - \partial_t^2 \hat{u}^m) \| \lesssim \tau^2 \| \mathcal{L}_h^a \partial_t^4 \hat{u} \|_{L^\infty(L^2)} \lesssim \tau^2, \quad (\text{A.2f})$$

where the first step is (3.5) at $a \in \{0, \frac{1}{4}\}$ for P_h and the second is the standard consistency error of the central second difference, written in integral form so that only $\partial_t^4 u \in L^\infty(0, T; D(\mathcal{L}^{\frac{1}{4}}))$ is used. (A.2) yields exactly the bound (4.9) for $\delta_{h;\frac{1}{4}}^m$.

The next step consists of estimating the remaining defects $\delta_{h;\frac{3}{4}}^m = \delta_{h;2}^m + \delta_{h;3}^m + \delta_{h;4}^m$ for $m \in \{1, n\}$. Consider defect $\delta_{h;2}^m$ for $m \in \{1, n\}$, then, by Taylor expansion we have that

$$\| \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;2}^m \| = \Gamma_h^{m;\gamma}(u_h^{m+1}) \| \mathcal{L}_h^{\frac{1}{2}+a} R_h \theta(\hat{u}^{m+1} - 2\hat{u}^m + \hat{u}^{m-1}) \| \leq \tau^2 C_\Gamma \theta \| \mathcal{L}_h^{\frac{1}{2}+a} R_h \partial_t^2 \hat{u}^m \|. \quad (\text{A.3a})$$

For the defect $\delta_{h;3}^m$ for $m \in \{1, n\}$, we use Taylor expansion on the norm of the solution

$$\begin{aligned} \| \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;3}^m \| &= \beta^m |\gamma| \| \nabla \hat{u}^{m+1} \|^2 - 2\gamma \| \nabla \hat{u}^m \|^2 + \gamma \| \nabla \hat{u}^{m-1} \|^2 \| \mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m \| \\ &\leq \tau^2 \beta^m \gamma (\partial_t^2 \| \nabla \hat{u}^m \|^2) \| \mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m \|. \end{aligned} \quad (\text{A.3b})$$

Finally, the last consistency defect $\delta_{h;4}^m$ is the projection part of the coefficient. By Ritz orthogonality, $\langle \nabla(\hat{u}^m - R_h \hat{u}^m), \nabla \varphi_h \rangle = 0$ for all $\varphi_h \in V_h$, so that the quantities ρ^m satisfy

$$\rho^m = \| \nabla R_h \hat{u}^m \|^2 - \| \nabla \hat{u}^m \|^2 = - \| \nabla(\hat{u}^m - R_h \hat{u}^m) \|^2, \quad \text{hence} \quad |\rho^m| \lesssim h^{2p},$$

by the H^1 -approximation estimate for R_h with $s = p + 1$. Since the weights $\gamma, 1 - 2\gamma, \gamma$ are non-negative and sum to 1, the same bound holds for their combination, and therefore for $m \in \{1, n\}$ we have that

$$\begin{aligned} \| \mathcal{L}_h^{a-\frac{1}{2}} \delta_{h;4}^m \| &= \beta^m |\gamma \rho^{m+1} + (1 - 2\gamma) \rho^m + \gamma \rho^{m-1}| \| \mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m \| \\ &\leq \beta^m \max_{\mu \in \{m-1, m, m+1\}} |\rho^\mu| \| \mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m \| \lesssim \beta^m h^{2p} \| \mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m \|. \end{aligned} \quad (\text{A.3c})$$

All right-hand sides are controlled by the chain

$$\| \mathcal{L}_h^{\frac{1}{2}+a} R_h \partial_t^j \hat{u}^m \| \leq \lambda_1^{a-\frac{1}{4}} \| \mathcal{L}_h^{\frac{3}{4}} R_h \partial_t^j \hat{u}^m \| \leq \lambda_1^{a-\frac{1}{4}} \| \mathcal{L}_h^{\frac{3}{4}} \partial_t^j \hat{u}^m \| \lesssim 1, \quad j = 0, \dots, 3, \quad (\text{A.4})$$

valid for $0 \leq a \leq \frac{1}{4}$, whose first step is (3.7) and whose second is (3.6) at the exponent $\frac{3}{4}$. The final bound is Assumption 2.3 for $j = 0$, which gives $\| \mathcal{L}_h^{\frac{3}{4}} \hat{u}^m \| \leq \frac{1}{2} M_{\frac{3}{4}}$, and the assumed temporal regularity $\partial_t^j u \in L^\infty(0, T; D(\mathcal{L}^{\frac{3}{4}}))$ for $j = 1, 2, 3$. Applied with $j = 0, 2$, this turns (A.3) into (4.10).

With the end points from the summation by parts estimated, we turn to the time derivatives of the last defects $\partial_\tau \delta_{h;\frac{3}{4}}^m = \partial_\tau \delta_{h;2}^m + \partial_\tau \delta_{h;3}^m + \partial_\tau \delta_{h;4}^m$ for $m = 2, \dots, n$. The ideas are completely analogous to (A.3) with the addition that due to the assumed temporal regularity we obtain an additional order τ which cancels the τ^{-1} in the definition of ∂_τ . For $\partial_\tau \delta_{h;2}^m$ for $2 \leq m \leq n$, by Taylor expansion, the estimate (A.1a), and $\Gamma_h^{n;\gamma}(u_h^{n+1}) \leq C_\Gamma$ by the bound established in the proof of Lemma 4.1, we have that

$$\begin{aligned} \| \mathcal{L}_h^{a-\frac{1}{2}} \partial_\tau \delta_{h;2}^m \| &\leq \tau^{-1} \theta (| \Gamma_h^{m;\gamma}(u_h^{m+1}) - \Gamma_h^{m-1;\gamma}(u_h^m) | \| \mathcal{L}_h^{\frac{1}{2}+a} R_h (\hat{u}^{m+1} - 2\hat{u}^m + \hat{u}^{m-1}) \| \\ &\quad + \Gamma_h^{m-1;\gamma}(u_h^m) \| \mathcal{L}_h^{\frac{1}{2}+a} R_h (\hat{u}^{m+1} - 3\hat{u}^m + 3\hat{u}^{m-1} - \hat{u}^{m-2}) \|) \end{aligned}$$

$$\lesssim \tau^2 \theta (\|\mathcal{L}_h^{\frac{1}{2}+a} R_h \partial_t^2 \hat{u}^m\| + C_\Gamma \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \partial_t^3 \hat{u}^m\|). \quad (\text{A.5a})$$

For the derivative of $\partial_\tau \delta_{h;3}^m$ for $2 \leq m \leq n$ using Taylor expansion of β for the first term, Taylor expansion of the norm of the solution for all three terms, and Taylor expansion of the solution for the last term, we have that

$$\begin{aligned} \|\mathcal{L}_h^{a-\frac{1}{2}} \partial_\tau \delta_{h;3}^m\| &\leq \tau^{-1} \left[|\beta^m - \beta^{m-1}| \|\gamma \|\nabla \hat{u}^{m+1}\|^2 - 2\gamma \|\nabla \hat{u}^m\|^2 + \gamma \|\nabla \hat{u}^{m-1}\|^2 \right] \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| \\ &\quad + \beta^{m-1} |\gamma \|\nabla \hat{u}^{m+1}\|^2 - 3\gamma \|\nabla \hat{u}^m\|^2 + 3\gamma \|\nabla \hat{u}^{m-1}\|^2 - \gamma \|\nabla \hat{u}^{m-2}\|^2| \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| \\ &\quad + \beta^{m-1} |\gamma \|\nabla \hat{u}^m\|^2 - 2\gamma \|\nabla \hat{u}^{m-1}\|^2 + \gamma \|\nabla \hat{u}^{m-2}\|^2| \|\mathcal{L}_h^{\frac{1}{2}+a} R_h (\hat{u}^m - \hat{u}^{m-1})\| \Big] \\ &\lesssim \tau^2 [(\partial_t^2 \|\nabla \hat{u}^m\|^2) \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| + (\partial_t^3 \|\nabla \hat{u}^m\|^2) \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| + (\partial_t^2 \|\nabla \hat{u}^{m-1}\|^2) \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \partial_t \hat{u}^m\|]. \end{aligned}$$

Finally, for the derivative of $\partial_\tau \delta_{h;4}^m$ for $2 \leq m \leq n$, we have that

$$|\rho'(t)| = 2 |\langle \nabla(I - R_h)u(t), \nabla(I - R_h)\partial_t u(t) \rangle| \leq 2 \|\nabla(I - R_h)u(t)\| \|\nabla(I - R_h)\partial_t u(t)\| \lesssim h^{2p},$$

which allows us to use Taylor expansion of ρ . This, Taylor expansion of β for the first term, and Taylor expansion of the solution for the last term we obtain

$$\begin{aligned} \|\mathcal{L}_h^{a-\frac{1}{2}} \partial_\tau \delta_{h;4}^m\| &\leq \tau^{-1} (|\beta^m - \beta^{m-1}| |\gamma \rho^{m+1} + (1 - 2\gamma)\rho^m + \gamma \rho^{m-1}| \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| \\ &\quad + \beta^{m-1} |\gamma(\rho^{m+1} - \rho^m) + (1 - 2\gamma)(\rho^m - \rho^{m-1}) + \gamma(\rho^{m-1} - \rho^{m-2})| \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| \\ &\quad + \beta^{m-1} |\gamma \rho^m + (1 - 2\gamma)\rho^{m-1} + \gamma \rho^{m-2}| \|\mathcal{L}_h^{\frac{1}{2}+a} R_h (\hat{u}^m - \hat{u}^{m-1})\|) \\ &\leq \beta_{\max} (h^{2p} \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| + h^{2p} \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \hat{u}^m\| + h^{2p} \|\mathcal{L}_h^{\frac{1}{2}+a} R_h \partial_t \hat{u}^m\|) \end{aligned} \quad (\text{A.5b})$$

(4.11) now follows from (A.5) and (A.4), applied with $j = 0, \dots, 3$.

It remains to treat the stabilization term S_h^m for $1 \leq m \leq n$. From the definition $e_h^m = R_h \hat{u}^m - u_h^m$ and the self-adjointness of $\mathcal{L}_h$, we obtain

$$\begin{aligned} |\sigma^m| &= |\|\nabla u_h^m\|^2 - \|\nabla R_h \hat{u}^m\|^2| = |\langle \mathcal{L}_h^{\frac{1}{2}}(u_h^m + R_h \hat{u}^m), \mathcal{L}_h^{\frac{1}{2}}(u_h^m - R_h \hat{u}^m) \rangle| \\ &\lesssim (\|\mathcal{L}_h^{\frac{1}{2}-a} u_h^m\| + \|\mathcal{L}_h^{\frac{1}{2}-a} R_h \hat{u}^m\|) \|\mathcal{L}_h^{\frac{1}{2}+a} e_h^m\| \lesssim \|\mathcal{L}_h^{\frac{1}{2}+a} e_h^m\|, \end{aligned}$$

where we used (3.7), the uniform bounds on the discrete solution, and (3.6) at $a = 0$, and the uniform bounds on the discrete solution together with the bound (3.9) at $a = \frac{1}{4}$, where the norm of the solution is controlled by Assumption 2.3. Moreover, by the commutation identity (3.2), we have $\mathcal{L}_h^{1+a} R_h \hat{u}^m = \mathcal{L}_h^a P_h \mathcal{L} \hat{u}^m$. By the continuous equation at the time step t_m , we obtain

$$\mathcal{L} \hat{u}^m = (q^m)^{-1} (f^m - \partial_t^2 \hat{u}^m).$$

Since $\alpha^m \geq \alpha_{\min} > 0$, the assumed regularity and the stability of P_h imply

$$\|\mathcal{L}_h^{1+a} R_h \hat{u}^m\| = \|\mathcal{L}_h^a P_h \mathcal{L} \hat{u}^m\| \lesssim \|\mathcal{L}^a \mathcal{L} \hat{u}^m\| \lesssim \|\mathcal{L}^a f^m\| + \|\mathcal{L}^a \partial_t^2 \hat{u}^m\| \lesssim 1.$$

Thus, for the stabilization term S_h^m , we have

$$|\langle \mathcal{L}_h^a S_h^m, \mathcal{L}_h^a \varphi_h \rangle| \lesssim (\|\mathcal{L}_h^{\frac{1}{2}+a} e_h^{m+1}\| + \|\mathcal{L}_h^{\frac{1}{2}+a} e_h^m\| + \|\mathcal{L}_h^{\frac{1}{2}+a} e_h^{m-1}\|) \|\mathcal{L}_h^a \varphi_h\|.$$

This is precisely the last estimate (4.12) we needed to establish. $\blacksquare$

A.2 Projection and interpolation operators

We prove here the estimates collected in Section 3 and record the references for the standard facts they rest on. The approximation estimates of (3.4) for P_h and R_h , and the L^2 -estimate for R_h obtained by the Aubin–Nitsche duality argument, are classical, see [7, Ch. 0, 4 & 5] or [18]; the quasi-interpolation operator J_h is the one of [9] for $p = 1$ and, for general p and with the homogeneous boundary condition built in, the one of [14], both of which are L^2 - and H_0^1 -stable and satisfy that estimate by construction. The stability we use is stated separately for each operator rather than as a blanket assertion: P_h

is L^2 -stable with constant 1, and its H_0^1 -stability holds on quasi-uniform families by [10] and, more generally, on shape-regular families under a local mesh condition by [6]; R_h is H_0^1 -stable with constant 1, and in weaker norms we use only (3.4), (3.6) and the commutation identity (3.3); J_h is L^2 - and H_0^1 -stable by construction.

The fractional estimate of (3.4) follows from the first by interpolation: for $w \in H_0^1(\Omega)$ we have $\|\mathcal{L}^{\frac{1}{4}}w\|^2 = \langle \mathcal{L}^{\frac{1}{2}}w, w \rangle \leq \|\nabla w\| \|w\|$, and applying this to $w = v - \Pi v$ and inserting the first estimate gives $\|\mathcal{L}^{\frac{1}{4}}(v - \Pi v)\| \lesssim h^{s-\frac{1}{2}}|v|_{H^s(\Omega)}$.

For (3.7), every eigenvalue of $\mathcal{L}_h$ is a Rayleigh quotient over $V_h \subset H_0^1(\Omega)$, hence bounded below by λ_1 independently of h , and the spectral calculus on V_h gives the first claim; combining it with (3.6) at $a = \frac{1}{2}$ gives (3.9). For (3.8), the inverse inequality states that the largest eigenvalue of $\mathcal{L}_h$ satisfies $\lambda_{\max,h} \leq C_{\text{inv}}^2 h^{-2}$, and the spectral calculus again gives the claim.

Proof of (3.5) and (3.6). Throughout we use that $\|\mathcal{L}_h^{\frac{1}{2}}\varphi_h\| = \|\nabla\varphi_h\|$ for $\varphi_h \in V_h$, and that $D(\mathcal{L}^a)$ and $D(\mathcal{L}_h^a)$ are the interpolation scales generated by the two self-adjoint positive definite operators $\mathcal{L}$ and $\mathcal{L}_h$, so that a linear map bounded at two exponents is bounded at every exponent in between, with the interpolated constant.

For (3.5) it suffices, by interpolation, to treat $a = 0$ and $a = \frac{1}{2}$. Both P_h and J_h are defined on $L^2(\Omega)$, so both endpoints are meaningful. At $a = 0$ the assertion is their L^2 -stability, and at $a = \frac{1}{2}$ it reads $\|\nabla P_h v\| + \|\nabla J_h v\| \lesssim \|\nabla v\|$, which is their H_0^1 -stability; both hold by the references recorded above, and the interpolation of the two fractional-power scales used here is the standard interpolation of Hilbert scales [5].

For (3.6) we again interpolate, now between $a = \frac{1}{2}$ and $a = 1$. At $a = \frac{1}{2}$,

$$\|\mathcal{L}_h^{\frac{1}{2}}R_h v\| = \|\nabla R_h v\| \leq \|\nabla v\| = \|\mathcal{L}^{\frac{1}{2}}v\|,$$

again because the Ritz projection is the H_0^1 -orthogonal projection onto V_h . At $a = 1$ we use the defining property of R_h : for every $\varphi_h \in V_h$ and $v \in D(\mathcal{L})$, $\langle \mathcal{L}_h R_h v, \varphi_h \rangle = \langle \nabla R_h v, \nabla \varphi_h \rangle = \langle \nabla v, \nabla \varphi_h \rangle = \langle \mathcal{L}v, \varphi_h \rangle = \langle P_h \mathcal{L}v, \varphi_h \rangle$, that is, $\mathcal{L}_h R_h v = P_h \mathcal{L}v$, which is (3.3), whence $\|\mathcal{L}_h R_h v\| = \|P_h \mathcal{L}v\| \leq \|\mathcal{L}v\|$. Both endpoints hold with constant 1, so does the interpolant. ■

For $\varphi_h \in V_h$ we have $P_h \varphi_h = \varphi_h$, so (3.5) reads $\|\mathcal{L}_h^a \varphi_h\| \lesssim \|\mathcal{L}^a \varphi_h\|$ for $0 \leq a \leq \frac{1}{2}$, which is the form in which it is used below.

A.3 The three-layer displacement energy

Proof of Lemma 3.1. Non-negativity of $\mathcal{W}_h^m$ holds by definition, both weights being non-negative for $\theta \geq \frac{1}{4}$. Expanding the squares in the definition of $\mathcal{W}_h^m$ gives the equivalent form

$$\mathcal{W}_h^m = \theta(\|v_h^m\|^2 + \|v_h^{m-1}\|^2) + (1 - 2\theta)\langle v_h^m, v_h^{m-1} \rangle. \quad (\text{A.6})$$

Since $\theta \leq \frac{1}{2}$ we have $1 - 2\theta \geq 0$, so applying $|\langle v_h^m, v_h^{m-1} \rangle| \leq \frac{1}{2}(\|v_h^m\|^2 + \|v_h^{m-1}\|^2)$ to (A.6) in both directions yields

$$\frac{4\theta-1}{2}(\|v_h^m\|^2 + \|v_h^{m-1}\|^2) \leq \mathcal{W}_h^m \leq \frac{1}{2}(\|v_h^m\|^2 + \|v_h^{m-1}\|^2),$$

which is (3.11), and, after discarding the term $\|v_h^{m-1}\|^2$ on the left, also (3.12) whenever $4\theta - 1 > 0$. For (3.13) split v_h^m into its half-sum and half-difference, $v_h^m = \frac{1}{2}(v_h^m + v_h^{m-1}) + \frac{1}{2}(v_h^m - v_h^{m-1})$, and using $\|x + y\|^2 \leq 2\|x\|^2 + 2\|y\|^2$,

$$\|v_h^m\|^2 \leq \frac{1}{2}\|v_h^m + v_h^{m-1}\|^2 + \frac{1}{2}\|v_h^m - v_h^{m-1}\|^2 \leq 2\mathcal{W}_h^m + \frac{1}{2}\|v_h^m - v_h^{m-1}\|^2.$$

For the remaining difference, we have that $v_h^m - v_h^{m-1} = \tau \mathcal{L}_h^{\frac{1}{2}}(\mathcal{L}_h^a \partial_\tau u_h^m)$ with $\mathcal{L}_h^a \partial_\tau u_h^m \in V_h$, so (3.8) with the pair $(0, \frac{1}{2})$ gives $\|v_h^m - v_h^{m-1}\| \leq C_{\text{inv}} \tau h^{-1} \|\mathcal{L}_h^a \partial_\tau u_h^m\|$, and (3.13) follows. Finally, we turn to (3.10). By algebraic manipulation, we have for the left-hand side

$$\begin{aligned} \langle v_h^{m+1;\theta}, v_h^{m+1} - v_h^{m-1} \rangle &= \frac{1}{4} \langle v_h^{m+1} + 2v_h^m + v_h^{m-1}, v_h^{m+1} + v_h^m - (v_h^m + v_h^{m-1}) \rangle \\ &\quad + (\theta - \frac{1}{4}) \langle v_h^{m+1} - 2v_h^m + v_h^{m-1}, v_h^{m+1} - v_h^m + (v_h^m - v_h^{m-1}) \rangle \\ &= \mathcal{W}_h^{m+1} - \mathcal{W}_h^m. \end{aligned} \quad \blacksquare$$

Remark A.1 (Sharpness at the endpoint). Estimate (3.12) degenerates as $\theta \downarrow \frac{1}{4}$ and fails at the endpoint, where $\mathcal{W}_h^m = \frac{1}{4}\|v_h^m + v_h^{m-1}\|^2$ vanishes for $v_h^m = -v_h^{m-1}$. The substitute is (3.13), which controls $\|v_h^m\|$ only together with the kinetic term and only up to the factor τh^{-1} . That factor is sharp. At $\theta = \frac{1}{4}$ the sawtooth $u_h^m = (-1)^m \varepsilon \psi_h$, with ψ_h an eigenfunction of $\mathcal{L}_h$ for the largest eigenvalue $\lambda_{\max,h}$, gives $\mathcal{W}_h^m = 0$ and quotient $\|v_h^m\|/\|\mathcal{L}_h^a \partial_\tau u_h^m\| = \frac{\tau}{2} \lambda_{\max,h}^{\frac{1}{2}} \sim \tau h^{-1}$. Every argument recovering $\|v_h^m\|$ from the three-layer energy there must therefore accept that factor or impose a mesh condition.