Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

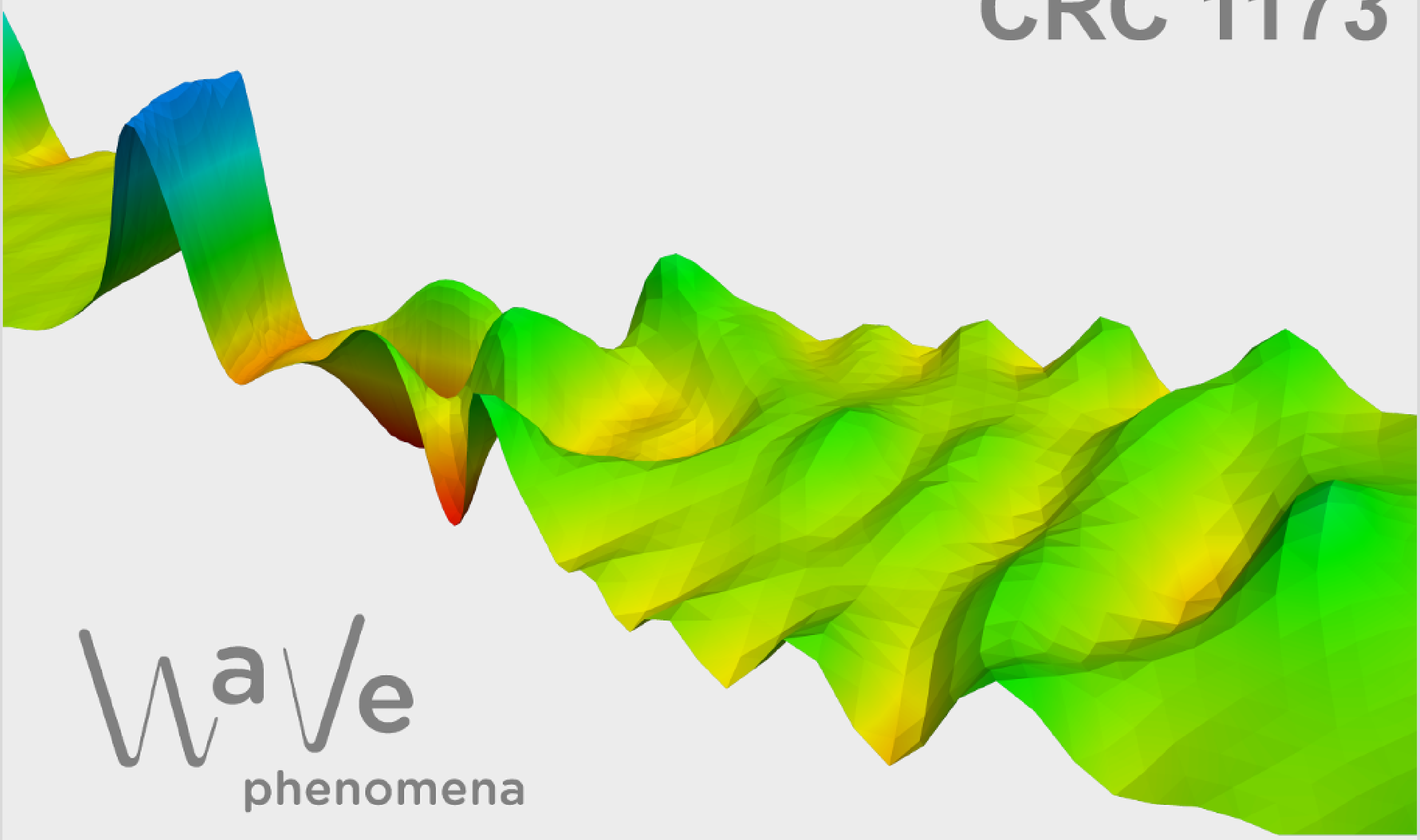
# Maximum norm error bounds for the full discretization of non-autonomous wave equations

Benjamin Dörich, Jan Leibold, Bernhard Maier

CRC Preprint 2021/47 (revised), September 2023

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave phenomena

**Participating universities**

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

TU WIEN

**Funded by**

DFG

ISSN 2365-662X

# Maximum norm error bounds for the full discretization of non-autonomous wave equations

BENJAMIN DÖRICH† , JAN LEIBOLD AND BERNHARD MAIER
*Institute for Applied and Numerical Mathematics, Karlsruhe Institute of Technology, Englerstr. 2, 76131 Karlsruhe, Germany*

In the present paper, we consider a specific class of non-autonomous wave equations on a smooth, bounded domain and their discretization in space by isoparametric finite elements and in time by the implicit Euler method. Building upon the work of Baker and Dougalis (1980), we prove optimal error bounds in the $W^{1,\infty} \times L^{\infty}$-norm for the semi discretization in space and the full discretization. The key tool is the gain of integrability coming from the inverse of the discretized differential operator. For this, we have to pay with (discrete) time derivatives on the error in the $H^1 \times L^2$-norm which are reduced to estimates of the differentiated initial errors. To confirm our theoretical findings, we also present numerical experiments.

*Keywords*: error analysis, full discretization, wave equation, maximum norm error bounds, nonconforming space discretization, isoparametric finite elements, a-priori error bounds

## 1. Introduction

In the present paper, we consider the non-autonomous wave equation

$$\partial_{tt} u(t,\mathbf{x}) = \lambda(t,\mathbf{x})^{-1} \Delta u(t,\mathbf{x}) + f(t,\mathbf{x}), \qquad t \in [0,T],\, \mathbf{x} \in \Omega, \tag{1.1}$$

on the domain $\Omega \subseteq \mathbb{R}^N$, $N = 2,3$. We assume it to be bounded and convex with a sufficiently regular boundary, and impose homogeneous Dirichlet boundary conditions and appropriate initial conditions. We discretize (1.1) with isoparametric finite elements in space and the implicit Euler scheme in time, and derive $W^{1,\infty} \times L^{\infty}$-norm error bounds both for the semi discretization in space and the full discretization.

A bound in the maximum norm allows us to control the numerical error at every point in the domain $\Omega$. Compared to the classical estimates in $L^2$, see, e.g., Bales *et al.* (1985); Bales & Dougalis (1989), which are implied (with non-optimal order) by our maximum norm error estimates, and in the energy space $H^1$, see, e.g., Maier (2022); Hochbruck & Maier (2021), they provide an additional insight in the approximation quality. For example, they become particularly interesting if one wants to approximate the quasilinear wave equation

$$\partial_{tt} u(t,\mathbf{x}) = \lambda(u(t,\mathbf{x}))^{-1} \Delta u(t,\mathbf{x}) + f(t,\mathbf{x},u(t,\mathbf{x}),\partial_t u(t,\mathbf{x})), \tag{1.2}$$

where for example in nonlinear acoustics $\lambda(u) = 1 - u^m$, $m = 2,3$. The reason is, that this equation is only well-posed as long as $\lambda(u)$ satisfies a pointwise lower bound away from zero. When discretizing (1.2) in space, it has to be ensured that the spatial discretization inherits this property. Since this requires

†Corresponding author. Email: benjamin.doerich@kit.edu

a pointwise bound of the numerical approximation, maximum norm estimates, as they are provided in this paper, are sufficient - once they are transferred to the nonlinear case - to guarantee such constraints. Indeed, for some spatial discretization $u_h$ the triangle inequality

$$\|u_h\|_{L^\infty} \leqslant \|u\|_{L^\infty} + \|u_h - u\|_{L^\infty}$$

allows to keep the maximum norm of the numerical solution arbitrarily close to the one of the exact solution once convergence is established. So far, to show convergence an inverse inequality has to be employed, which leads to an unsatisfactory CFL condition, even for methods which are known to be unconditionally stable, or a restriction to higher-order finite elements, see, e.g., Maier (2022); Antonietti *et al.* (2020); Makridakis (1993). Alternatively, $H^2$-conforming finite elements, as suggested in Zlámal (1968), can be employed. For those, Sobolev's embedding can be used to obtain maximum norm estimates, once the convergence in $H^2$ is established. However, in order to achieve this type of conformity, the number of degrees of freedom has to be increased significantly. Our hope is to show that these constraints are only of theoretical nature and can be removed. We are confident, that the analysis presented here for the linear problem (1.1) is an important step towards the quasilinear problem (1.2) using Lagrangian finite element methods without CFL conditions and also generalizes to higher-order methods in time.

In the articles Baker *et al.* (1979); Baker & Dougalis (1980), the space and time discretization of the linear autonomous wave-equation (i.e., $\lambda = 1, f = 0$ in (1.1)) by finite elements and one- or two-step methods, respectively, is analyzed. In our paper, we extend their analysis to the $W^{1,\infty} \times L^\infty$-norm, to the more general case of linear, non-autonomous wave equations and, also to nonconforming finite elements. We point out that the latter cannot be omitted due to the following reason: In the error analysis, we rely on elliptic regularity results only available on a smooth domain $\Omega$. Unfortunately, this prevents us from using these results on a computational domain $\Omega_h$ with a piecewise polynomial boundary. Our research is mainly inspired by Baker *et al.* (1979); Baker & Dougalis (1980) and we are not aware of further maximum norm estimates for wave equations discretized by finite elements, besides the one-dimensional case considered in Trautmann *et al.* (2018). For finite differences on a square combined with a fourth-order in time scheme, an error bound under a CFL condition is established in Liao & Sun (2011).

For the spatial semi discretization in Baker & Dougalis (1980), they trade integrability, coming from the inverse of the discretized differential operator $\Delta_h$, for time derivatives on the error in the $L^2$-norm. Those errors are controlled by the derivatives of the initial error which can be bounded using a properly preconditioned initial value. For our semi discretization, we use a similar approach to transfer and extend the results with additional technical effort to the non-autonomous case.

For the full discretization, the proofs in Baker *et al.* (1979); Baker & Dougalis (1980) rely on an expansion of the discrete error in the eigenbasis of $\Delta_h$. However, we are not aware of how to generalize this approach to the non-autonomous case. Hence, we pursue the strategy of the semi discretization. From the implicit Euler scheme we derive discrete derivatives and adapt the proofs to derive fully discrete error bounds. Let us also note that the bounds in Baker & Dougalis (1980) derived in the $L^\infty$-norm are of order $k+1$, but since our bounds involve the $W^{1,\infty}$-norm, we derive optimal error bounds of order $k$.

Further, we comment on maximum norm error bounds for finite element discretizations of elliptic problems as they are the fundamental tool for our error bounds in the time-dependent case. The first quasi-optimal error bounds in the maximum norm were given by Natterer (1975) and Scott (1976). Many extensions and refinements have been achieved in the following years, see, e.g., Nitsche (1975, 1977); Rannacher (1976); Rannacher & Scott (1982); Schatz & Wahlbin (1977, 1982, 1995); Wahlbin (1978).

More recently in the context of nonconforming space discretizations, maximum norm error bounds for linear finite elements applied to an inhomogeneous Neumann problem were derived in Kashiwabara & Kemmochi (2020). For (evolving) surface finite element methods, estimates on the finite element solution and the generalized Ritz map for isoparametric finite elements are considered in Demlow (2009); Kovács & Power Guerra (2018). In Dörich *et al.* (2023), the authors of the present paper extended the approach in Brenner & Scott (2008) to derive stability of a generalized Ritz map to higher-order isoparametric elements.

We also briefly comment on further work conducted in the context of maximum norm error estimates for parabolic problems. Here, we are aware of two strategies: In Bramble *et al.* (1977); Baker *et al.* (1977), a similar approach as for the wave equation is taken and integrability is gained for time derivatives. Alternatively, some kind of stability of the semigroup generated by $\Delta_h$ on $L^\infty$ is shown. This is done either directly using energy techniques, see, e.g., Schatz *et al.* (1980, 1998), or via resolvent estimates on $L^\infty$ and maximal parabolic regularity, see, e.g., Bakaev *et al.* (2003); Palencia (1996); Thomée & Wahlbin (2004); Chatzipantelidis *et al.* (2006); Li (2019). However, such stability estimates cannot be expected for hyperbolic problems in general, see (Arendt *et al.*, 2011, Exa. 8.4.9) and Littman (1963).

The paper is organized as follows: In Section 2, we present the analytical framework and the space discretization by isoparametric Lagrange finite elements. After providing some properties of the discretized objects, we state our main results on the error bounds for the semi discretization in space and the full discretization by the implicit Euler method.

The main parts of the proof of the semi-discrete error bound are given in Section 3. Here, we exchange the integrability in the error for time derivatives of the defect and trace those back to the initial values. We adapt the presented technique in Section 4 and transfer it from the continuous to the discrete derivatives in order to prove the theorem on the fully discrete error bound.

Section 5 is devoted to the final conclusion of our main results. We collect several approximation results and estimate the defects. Further, the (discrete) derivatives of the initial error as well as the errors in the first approximations of the fully discrete scheme are bounded.

In Appendix A, we collect some further results employed in the error analysis.

### Notation

In the rest of the paper we use the notation

$$a \lesssim b\,,$$

if there is a constant $C > 0$ independent of the spatial parameter $h$ and the time step-size $\tau$ such that $a \leqslant Cb$. For the sake of readability, we introduce the notation $t^n = n\tau$ and for an arbitrary time-dependent, continuous object $x(t)$ in some Banach space $X$ and a sequence $(x^m)$ in $X$, we define

$$\|x\|_{L^\infty(X)} := \max_{[0,T]} \|x(t)\|_X\,, \qquad \|x^n\|_{\ell^\infty(X)} := \max_{m=1,\dots,n} \|x^m\|_X\,.$$

If it is clear from the context, we write $L^p$ instead of $L^p(\Omega)$ or $L^p(\Omega_h)$.

## 2. General Setting

For a convex, bounded domain $\Omega \subset \mathbb{R}^N$, $N = 2,3$, with boundary $\partial\Omega \in C^{s,1}$, $s \in \mathbb{N}$, we study the non-autonomous wave equation (1.1) and the positive, self-adjoint operator $-\Delta$ on $L^2(\Omega)$ with homogeneous Dirichlet boundary conditions. Therefore, we introduce the spaces $H = L^2(\Omega)$ and $V =$

$\mathscr{D}\big((-\Delta)^{1/2}\big) = H_0^1(\Omega)$. The equation is further equipped with initial values

$$u(0) = u^0, \qquad\qquad \partial_t u(0) = v^0.$$

We expect that it is also possible to treat more general elliptic differential operators L with regular coefficients and Neumann boundary conditions, as long as one can establish the properties of the spatial discretization as stated in Section 2.1. Our analysis relies on the following regularity assumptions of $\lambda$.

ASSUMPTION 2.1 There are $\kappa, \ell_{\max} \in \mathbb{N}$ such that the following holds.

$(\lambda_1)$ There exist $C_\lambda \geqslant c_\lambda > 0$ such that the function $\lambda\colon [0,T] \times \Omega \to \mathbb{R}$ satisfies

$$c_\lambda \leqslant \lambda(t,\mathbf{x}) \leqslant C_\lambda, \qquad\qquad t \in [0,T],\, \mathbf{x} \in \Omega.$$

Moreover, we have $\lambda, \lambda^{-1} \in C^2\big([0,T], W^{\kappa,\infty}(\Omega)\big)$.

$(\lambda_2)$ For $0 \leqslant \ell \leqslant \ell_{\max}$ and $u \in \mathscr{D}\big((-\Delta)^{\ell/2}\big)$ it holds

$$\lambda u, \lambda^{-1} u \in \mathscr{D}\big((-\Delta)^{\ell/2}\big).$$

We note that assumption $(\lambda_2)$ guarantees that the multiplication with $\lambda$ preserves the boundary conditions incorporated in $\Delta$.

EXAMPLE 2.2 On a smooth domain $\Omega$ it holds

$$\mathscr{D}(\Delta) = \{u \in H^2(\Omega) \mid u|_{\partial\Omega} = 0\}, \quad \mathscr{D}(\Delta^2) = \{u \in H^4(\Omega) \mid u|_{\partial\Omega} = \Delta u|_{\partial\Omega} = 0\}.$$

In this case, we have the following sufficient conditions for $(\lambda_2)$.

(a) We always have $\ell_{\max} \geqslant 2$ and achieve $\ell_{\max} \geqslant 4$ by the product rule if

$$\nabla_x \lambda\big|_\Gamma = 0. \tag{2.1}$$

(b) Having the quasilinear case (1.2) in mind, and assuming $\lambda = \lambda(t,u)$ where $u$ is the solution in $H_0^1(\Omega)$, then a sufficient condition for (2.1) is given by $\partial_u \lambda(t,0) = 0$.

(c) If $\nabla_x \lambda$ has compact support in $\Omega$, $(\lambda_2)$ is satisfied for any $\ell_{\max} \in \mathbb{N}$.

Further, condition $(\lambda_1)$ directly yields the following lemma.

LEMMA 2.3 Let Assumption 2.1 be satisfied for some $\kappa \in \mathbb{N}$. Then, we have for $t \in [0,T]$, $0 \leqslant \ell \leqslant \kappa$, $1 \leqslant p \leqslant \infty$, and $j = 0,1,2$ the bounds

$$\big\|\partial_t^j \lambda(t)\varphi\big\|_{W^{\ell,p}} \leqslant C\,\|\varphi\|_{W^{\ell,p}}, \qquad\qquad \big\|\partial_t^j \lambda(t)^{-1}\varphi\big\|_{W^{\ell,p}} \leqslant C\,\|\varphi\|_{W^{\ell,p}},$$

with a constant $C > 0$ depending on $\lambda$ and its derivatives.

Equivalently to (1.1), we consider the non-autonomous wave equation in first-order formulation

$$\partial_t y(t) = \Lambda(t)^{-1}\mathrm{A} y(t) + F(t), \qquad\qquad t \in [0,T], \tag{2.2}$$

with initial value $y(0) = y^0$ in the product space $X = V \times H$, with

$$y = \begin{pmatrix} u \\ \partial_t u \end{pmatrix}, \quad y^0 = \begin{pmatrix} u^0 \\ v^0 \end{pmatrix}, \quad \Lambda(t) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda(t) \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & \mathrm{Id} \\ \Delta & 0 \end{pmatrix}, \quad F(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

In particular, we emphasize that under Assumption 2.1 the operator $\Lambda$ generates the time-dependent inner product

$$(y \mid z)_{\Lambda(t)} = (\Lambda(t) y \mid z)_X\,, \qquad t \in [0,T],\, y,z \in X. \tag{2.3}$$

Since the multiplication with $\lambda, \lambda^{-1}$ is continuous on $L^2$, the corresponding norm is equivalent to the norm of $X$, i.e., we have

$$c_\Lambda \|z\|_X^2 \leqslant \|z\|_{\Lambda(t)}^2 \leqslant C_\Lambda \|z\|_X^2\,, \qquad t \in [0,T],\, z \in X, \tag{2.4}$$

with constants $c_\Lambda = \min\{1, c_\lambda\}$ and $C_\Lambda = \max\{1, C_\lambda\}$. Further, we conclude from $(\lambda_2)$ the continuity of the map

$$\Lambda(t)\colon \mathscr{D}(\mathrm{A}^\ell) \to \mathscr{D}(\mathrm{A}^\ell), \quad 0 \leqslant \ell \leqslant \ell_{\max},\, t \in [0,T]. \tag{2.5}$$

Our analysis relies on the solution operators of the Poisson equation in second- and first-order formulation, respectively. In particular, we introduce the second-order solution operator $\Delta^{-1}\colon H \to V$ given by

$$-\left(\Delta^{-1}\varphi \mid \psi\right)_V = (\varphi \mid \psi)_H\,, \qquad \varphi \in H,\, \psi \in V. \tag{2.6}$$

For the analysis, we heavily rely on the following elliptic regularity result (Grisvard, 1985, Thm. 2.4.2.5).

THEOREM 2.4 (Elliptic regularity) Let $\partial\Omega \in C^{1,1}$, then for all $1 < p < \infty$ there is a constant $C_p > 0$ such that for all $\varphi \in L^p(\Omega)$ it holds

$$\|\Delta^{-1}\varphi\|_{W^{2,p}} \leqslant C_p \|\varphi\|_{L^p}\,.$$

Furthermore, we define the first-order solution operator $\mathrm{T}\colon X \to \mathscr{D}(\mathrm{A})$ by

$$\mathrm{T} = \begin{pmatrix} 0 & \Delta^{-1} \\ \mathrm{Id} & 0 \end{pmatrix}.$$

In particular, this implies $\mathrm{TA} = \mathrm{Id}$ on $\mathscr{D}(\mathrm{A})$ and $\mathrm{AT} = \mathrm{Id}$ on $X$.

## 2.1 *Space discretization*

We study the nonconforming space discretization of (2.2) based on isoparametric finite elements. For further details on this approach, we refer to Elliott & Ranner (2021). In particular, we introduce a shape-regular and quasi-uniform mesh $\mathscr{T}_h$, consisting of isoparametric elements of degree $k \in \mathbb{N}$ and let $\partial\Omega \in C^{k+1,1}$. The computational domain $\Omega_h$ is given by

$$\Omega_h = \bigcup_{K \in \mathscr{T}_h} K \approx \Omega,$$

where the subscript $h$ denotes the maximal diameter of all elements $K \in \mathscr{T}_h$. In the following, we require that $h \leqslant h_0$ such that all cited results below hold true. We note that $h_0$ only depends on the geometry of the domain $\Omega$ and the polynomial degree $k$. Based on the transformations $F_K$ mapping the reference element $\widehat{K}$ to $K \in \mathscr{T}_h$, we introduce the finite element space of degree $k$

$$W_h = \{\varphi \in C_0(\overline{\Omega}_h) \mid \varphi|_K = \widehat{\varphi} \circ (F_K)^{-1} \text{ with } \widehat{\varphi} \in \mathscr{P}^k(\widehat{K}) \text{ for all } K \in \mathscr{T}_h\}.$$

Here, $\mathscr{P}^k(\widehat{K})$ consists of all polynomials on $\widehat{K}$ of degree at most $k$. The discrete approximation spaces are given by

$$H_h = \big(W_h, (\cdot \mid \cdot)_{L^2(\Omega_h)}\big), \quad V_h = \big(W_h, (\cdot \mid \cdot)_{H_0^1(\Omega_h)}\big),$$

and we set $X_h = V_h \times H_h$ Following the detailed construction in (Elliott & Ranner, 2021, Sec. 5), we introduce the lift operator $\mathscr{L}_h \colon H_h \to H$. In particular, for $p \in [1,\infty]$ there are constants $c_p, C_p > 0$ with

$$c_p \|\varphi_h\|_{L^p(\Omega_h)} \leqslant \|\mathscr{L}_h \varphi_h\|_{L^p(\Omega)} \leqslant C_p \|\varphi_h\|_{L^p(\Omega_h)}, \qquad \varphi_h \in L^p(\Omega_h), \tag{2.7a}$$

$$c_p \|\varphi_h\|_{W^{1,p}(\Omega_h)} \leqslant \|\mathscr{L}_h \varphi_h\|_{W^{1,p}(\Omega)} \leqslant C_p \|\varphi_h\|_{W^{1,p}(\Omega_h)}, \qquad \varphi_h \in W^{1,p}(\Omega_h), \tag{2.7b}$$

cf. (Elliott & Ranner, 2021, Prop. 5.8). Further by (Elliott & Ranner, 2013, Sec. 4), the lift preserves node values, i.e. in particular

$$I_h \mathscr{L}_h \varphi_h = \varphi_h, \quad \varphi_h \in V_h,$$

where we denote the nodal interpolation operator by $I_h \colon C(\Omega) \to V_h$. As shown in (Elliott & Ranner, 2021, Thm. 5.9), we have for $m = 0,1$, $1 \leqslant p \leqslant \infty$, and $1 \leqslant \ell \leqslant k$ the estimates

$$\|(\mathrm{Id} - \mathscr{L}_h I_h)\varphi\|_{W^{m,p}(\Omega)} \lesssim h^{\ell+1-m} \|\varphi\|_{W^{\ell+1,p}(\Omega)}, \qquad \varphi \in W^{\ell+1,p}(\Omega). \tag{2.8}$$

Further, $\ell = 0$ is allowed for $N < p \leqslant \infty$.

We define the adjoint lift operators $\mathscr{L}_h^{H*} \colon H \to H_h$ and $\mathscr{L}_h^{V*} \colon V \to V_h$ by

$$\big(\mathscr{L}_h^{H*}\varphi \mid \psi_h\big)_{H_h} = (\varphi \mid \mathscr{L}_h \psi_h)_H, \qquad \varphi \in H,\ \psi_h \in H_h, \tag{2.9a}$$

$$\big(\mathscr{L}_h^{V*}\varphi \mid \psi_h\big)_{V_h} = (\varphi \mid \mathscr{L}_h \psi_h)_V, \qquad \varphi \in V,\ \psi_h \in V_h, \tag{2.9b}$$

and note in the conforming case $\mathscr{L}_h^{H*}$ and $\mathscr{L}_h^{V*}$ coincide with the $L^2$- and the Ritz projection, respectively. From (Hipp *et al.*, 2019, Thm. 5.3) and (Elliott & Ranner, 2021, Lem. 8.24), we obtain for $1 \leqslant \ell \leqslant k$ the bounds

$$\|\mathscr{L}_h^{H*}\varphi\|_{H_h} \lesssim \|\varphi\|_{L^2(\Omega)}, \qquad \varphi \in L^2(\Omega). \tag{2.10a}$$

$$\big\|(I_h - \mathscr{L}_h^{H*})\varphi\big\|_{H_h} \lesssim h^{\ell+1} \|\varphi\|_{H^{\ell+1}(\Omega)}, \qquad \varphi \in H^{\ell+1}(\Omega), \tag{2.10b}$$

as well as for $1 \leqslant \ell \leqslant k$

$$\|\mathscr{L}_h^{V*}\varphi\|_{V_h} \lesssim \|\varphi\|_{H^1(\Omega)}, \qquad \varphi \in H^1(\Omega_h), \tag{2.11a}$$

$$\big\|(\mathrm{Id} - \mathscr{L}_h \mathscr{L}_h^{V*})\varphi\big\|_{H_h} \lesssim h^{\ell+1} \|\varphi\|_{H^{\ell+1}(\Omega)}, \qquad \varphi \in H^{\ell+1}(\Omega). \tag{2.11b}$$

In addition, we need the stability of $\mathscr{L}_h^{H*}$ in $H^1$

$$\big\|\mathscr{L}_h^{H*}\varphi\big\|_{V_h} \lesssim \|\varphi\|_{H^1(\Omega)}, \qquad \varphi \in H^1(\Omega), \tag{2.12}$$

Further, we define the (standard) $L^2$-projection $\pi_h$ onto $V_h$ for $\varphi \in L^2(\Omega_h)$ via

$$(\pi_h \varphi \mid \psi_h)_{L^2(\Omega_h)} = (\varphi \mid \psi_h)_{L^2(\Omega_h)}, \qquad \varphi \in L^2(\Omega_h), \psi_h \in V_h.$$

We note that it only differs from $\mathscr{L}_h^{H*}$ by geometric errors and thus, we obtain for $\varphi \in H^2(\Omega)$

$$\left\| \pi_h \mathscr{L}_h^{-1} \varphi - \mathscr{L}_h^{H*} \varphi \right\|_{L^2(\Omega_h)} \lesssim h^{k+1} \|\varphi\|_{H^2(\Omega)}. \tag{2.13}$$

The proofs for (2.12) and (2.13) are given in Appendix A. In addition, there hold the following stability estimates. Let $p \in [2,\infty]$ , then for any $\varphi \in L^p(\Omega_h)$ and $\psi \in H^1(\Omega_h)$

$$\|\pi_h \varphi\|_{L^p(\Omega_h)} \leqslant C \|\varphi\|_{L^p(\Omega_h)}, \qquad \|\pi_h \psi\|_{H^1(\Omega_h)} \leqslant C \|\psi\|_{H^1(\Omega_h)}, \tag{2.14}$$

with a constant $C > 0$ independent of $h$. We note that the case $p = 2$ is trivially satisfied by the definition of the projection. The case $p = \infty$ is covered by (Nitsche, 1975, Thm. 1), and an interpolation argument yields the first bound. The stability in $H^1$ is shown as for (2.12).

For the analysis in the following sections, we additionally rely on stability and approximation properties of $\mathscr{L}_h^{V*}$ in the maximum norm. These features are well known in the literature for conforming finite elements, see, e.g., the monograph (Brenner & Scott, 2008, Ch. 8). In the non-conforming case, the authors recently extended these results in (Dörich *et al.*, 2023, Thm. 2.5 & 2.6) to isoparametric finite elements. A special case is stated in the following proposition.

PROPOSITION 2.5 Let $\partial\Omega \in C^{k+1,1}$ and $h \leqslant h_0$. Then, the adjoint lift is stable in $W^{1,\infty}$ with

$$\|\mathscr{L}_h^{V*} \varphi\|_{W^{1,\infty}(\Omega_h)} \lesssim \|\varphi\|_{W^{1,\infty}(\Omega)}, \qquad \varphi \in W^{1,\infty}(\Omega). \tag{2.15}$$

For $0 \leqslant \ell \leqslant k$, it holds

$$\left\| (\mathrm{Id} - \mathscr{L}_h \mathscr{L}_h^{V*}) \varphi \right\|_{W^{1,\infty}(\Omega)} \lesssim h^\ell \|\varphi\|_{W^{\ell+1,\infty}(\Omega)}, \qquad \varphi \in W^{\ell+1,\infty}(\Omega). \tag{2.16}$$

Related estimates can be found in Kashiwabara & Kemmochi (2020) for the Neumann problem and linear elements, and in the context of evolving surfaces also in Kovács & Power Guerra (2018). We will also employ the inverse estimate, cf. (Brenner & Scott, 2008, Thm. 4.5.11) or (Maier, 2020, Lem. 5.6)

$$\|\varphi_h\|_{L^\infty(\Omega_h)} \leqslant Ch^{-N/p} \|\varphi_h\|_{L^p(\Omega_h)}, \qquad \|\varphi_h\|_{V_h} \leqslant Ch^{-1} \|\varphi_h\|_{L^2(\Omega_h)}, \qquad \varphi_h \in V_h. \tag{2.17}$$

We introduce the first-order lift operator $\boldsymbol{\mathscr{L}_h} \colon W^{\ell,p}(\Omega_h)^2 \to W^{\ell,p}(\Omega)^2$, for $\ell = 0,1$ and $1 \leqslant p \leqslant \infty$, and reference operator $J_h \colon V \times V \to X_h$ defined by

$$\boldsymbol{\mathscr{L}_h} = \begin{pmatrix} \mathscr{L}_h & 0 \\ 0 & \mathscr{L}_h \end{pmatrix}, \qquad J_h = \begin{pmatrix} \mathscr{L}_h^{V*} & 0 \\ 0 & \mathscr{L}_h^{V*} \end{pmatrix},$$

which are bounded uniformly in $h$ due to (2.7) and (2.11). In particular, we have by (2.16) the stability $J_h \in \mathcal{L}\big(\big(W^{1,\infty}(\Omega)\big)^2, \big(W^{1,\infty}(\Omega_h)\big)^2\big)$. For $t \in [0,T]$ we define the discrete operators $\lambda_h(t) \colon H_h \to H_h$, $\Lambda_h(t) \colon X_h \to X_h$, and the discrete right-hand side $F_h(t)$ by

$$\lambda_h(t)\varphi_h = \pi_h\big(I_h \lambda(t)\, \varphi_h\big), \qquad \Lambda_h(t) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda_h(t) \end{pmatrix}, \qquad F_h(t) = \begin{pmatrix} 0 \\ I_h f(t) \end{pmatrix}.$$

Correspondingly to Lemma 2.3, we collect important properties of $\lambda_h$ in the following lemma.

LEMMA 2.6 Let Assumption 2.1 be satisfied for some $\kappa \geqslant 2$ and $h \leqslant h_0$. Then, we have for $t \in [0,T]$, $1 \leqslant p \leqslant \infty$, and $j = 0,1,2$ the bounds

$$\left\| \partial_t^j \lambda_h(t) \varphi_h \right\|_{L^p} \leqslant C \left\| \varphi_h \right\|_{L^p}, \qquad \left\| \partial_t^j \lambda_h(t) \varphi_h \right\|_{V_h} \leqslant C \left\| \varphi_h \right\|_{V_h},$$

with a constant $C > 0$ depending only on $\lambda$ and its derivatives.

*Proof.* Using the stability in (2.14), it is sufficient to show the assertion for the product $(I_h \partial_t^j \lambda)\, \varphi_h$. Combining the interpolation property (2.8) and Assumption 2.1, the Hölder inequality yields the desired bound. □

Finally, we introduce the operators $\Delta_h \colon V_h \to H_h$ and $\mathrm{A}_h \colon X_h \to X_h$ for $\varphi_h, \psi_h \in V_h$ given by

$$-\left(\Delta_h \varphi_h \mid \psi_h\right)_{H_h} = \left(\varphi_h \mid \psi_h\right)_{V_h}, \qquad \mathrm{A}_h = \begin{pmatrix} 0 & \mathrm{Id} \\ \Delta_h & 0 \end{pmatrix}.$$

Note that these operators are not uniformly bounded with respect to $h$. Correspondingly to (2.3) and (2.4) using the identity $(\lambda_h \varphi_h \mid \psi_h)_{L^2(\Omega_h)} = (I_h \lambda\, \varphi_h \mid \psi_h)_{L^2(\Omega_h)}$, the discrete operator $\Lambda_h$ generates the time-dependent inner product

$$\left(y_h \mid z_h\right)_{\Lambda_h(t)} = \left(\Lambda_h(t) y_h \mid z_h\right)_{X_h}, \qquad t \in [0,T],\ y_h, z_h \in X_h,$$

with the induced norm being equivalent to the norm of $X_h$, i.e., we have as in (2.4) by Lemma 2.6

$$c_{\Lambda_h} \|z_h\|_{X_h}^2 \leqslant \|z_h\|_{\Lambda_h(t)}^2 \leqslant C_{\Lambda_h} \|z_h\|_{X_h}^2, \qquad t \in [0,T],\ z_h \in X_h. \tag{2.18}$$

We define the discrete solution operator $\Delta_h^{-1} \colon H_h \to V_h$ by

$$-\left(\Delta_h^{-1} \varphi_h \mid \psi_h\right)_{V_h} = \left(\varphi_h \mid \psi_h\right)_{H_h}, \qquad \varphi_h, \psi_h \in V_h, \tag{2.19}$$

and further the corresponding first-order solution operator

$$\mathrm{T}_h = \begin{pmatrix} 0 & \Delta_h^{-1} \\ \mathrm{Id} & 0 \end{pmatrix},$$

which again satisfies $\mathrm{T}_h \mathrm{A}_h = \mathrm{Id}$ and $\mathrm{A}_h \mathrm{T}_h = \mathrm{Id}$ on $X_h$.

The spatially discrete non-autonomous wave equation in first-order formulation then reads

$$\partial_t y_h(t) = \Lambda_h(t)^{-1} \mathrm{A}_h y_h(t) + F_h(t), \qquad t \in [0,T], \tag{2.20}$$

with the initial value $y_h(0) = y_h^0$, where

$$y_h^0 = J_h y^0 = \left(\mathscr{L}_h^{V*} u^0, \mathscr{L}_h^{V*} v^0\right)^T. \tag{2.21}$$

This choice of the initial value guarantees the convergence of the expression $J_h \mathrm{A}^2 y^0 - \mathrm{A}_h^2 y_h^0$ with optimal order, see Lemma 5.4 below, and error bounds for the first approximations, see Section 5.3. We note that in practice, one would usually choose the interpolation $y_h(0) = \left(I_h u^0\ ,\ I_h v^0\right)$. However, already in Baker & Dougalis (1980) preconditioned initial values of the type $u_h^0 = (\Delta_h^{-1})^m \Delta^m u^0$, $m \geqslant 4$, had to be chosen for the full order of convergence. Note that (2.21) is equivalent to the case $m = 1$ in the conforming case, but we do not require knowledge on $\Delta u^0$. A discussion on the computation of high-order approximations to $\mathscr{L}_h^{V*}$ is given in (Dörich, 2022, Prop. 2.4).

In the spatially continuous case, the solution operator $\Delta^{-1}$ can be used to obtain regularity which is traded in for pointwise estimates via the bounded map

$$\Delta^{-1}\colon L^2 \to H^2 \hookrightarrow L^\infty.$$

However, since we use Lagrangian finite elements which are *not* $H^2$-conforming, this approach does not work with $\Delta_h^{-1}$. Hence, in the following we provide estimates of $\Delta_h^{-1}$ that directly give us integrability without a detour via higher-order Sobolev spaces. A weaker form of this result has already been proven in (Bramble *et al.*, 1977, Lem. 4.1) in the conforming case only, and was recently sharpened in Dörich (2022). We state a variant of this result and, for completeness, give the proof in Appendix A.

LEMMA 2.7 Let $\partial\Omega \in C^{1,1}$ and $h \leqslant h_0$. Then, the solution operator $\Delta_h^{-1}$ satisfies

$$\|\Delta_h^{-1}\varphi_h\|_{L^\infty} \lesssim \|\varphi_h\|_{L^2} \qquad \|\Delta_h^{-1}\varphi_h\|_{W^{1,\infty}} \lesssim \|\varphi_h\|_{L^4}$$

for $\varphi_h \in V_h$.

A direct consequence of the above lemma for $N = 2,3$, is the possibility to consider the maps

$$X_h \hookrightarrow L^4 \times L^2 \xrightarrow{\mathrm{T}_h} L^\infty \times L^4 \xrightarrow{\mathrm{T}_h} W^{1,\infty} \times L^\infty, \tag{2.22}$$

which allow us to bound the maximum norm $\|\cdot\|_{W^{1,\infty}\times L^\infty}$ in terms of the energy norm $\|\cdot\|_X$ if we apply the solution operator $\mathrm{T}_h$ twice. We explain in Section 3 how to employ this observation. Note that in fact, one could sharpen the result to show that $\Delta_h^{-1}\colon L^2 \to W^{1,p}$ is bounded uniform in $h$ as long as the embedding $H^2 \hookrightarrow W^{1,p}$ is valid. For $p > N$, this then also implies the first estimate of the lemma.

We can finally state our first main result on the semi discretization. The proof is given in Section 3. We let $\mathscr{H}_\Delta^\ell := \mathscr{D}\big((-\Delta)^{\ell/2}\big)$ and use the notation

$$k^* = \max\{k, 2\}$$

in order to treat linear and higher-order finite elements simultaneously.

THEOREM 2.8 Let $\partial\Omega \in C^{k+1,1}$, $h \leqslant h_0$, and let Assumption 2.1 hold for some $\ell_{\max} \in \mathbb{N}$ and $\kappa = k+1$. Further, assume that the right-hand side $f$ and the solution of (1.1) satisfy

$$\begin{aligned} u &\in C^4\big([0,T], H^{k^*}(\Omega)\big) \cap C^2\big([0,T], H^{k+1}(\Omega)\big) \cap C^1\big([0,T], W^{k+1,\infty}(\Omega)\big), \\ f &\in C^1\big([0,T], H^{k+1}(\Omega)\big) \cap C\big([0,T], \mathscr{H}_\Delta^1\big), \\ u^0 &\in H^{k^*+2}(\Omega) \cap \mathscr{H}_\Delta^3, \quad v^0 \in H^{k^*+1}(\Omega) \cap \mathscr{H}_\Delta^2, \end{aligned} \tag{2.23}$$

and the initial value $y_h(0)$ is chosen as in (2.21). Then we have the error bound

$$\|y(t) - \boldsymbol{\mathscr{L}_h} y_h(t)\|_{W^{1,\infty}\times L^\infty} \leqslant Ch^k,$$

where $C$ is independent of $h$.

### 2.2 *Full discretization*

We study the full discretization with the backward Euler scheme

$$\partial_\tau y_h^n = \Lambda_h^{-1}(t^n)\mathrm{A}_h y_h^n + F_h(t^n), \quad n \geqslant 1, \tag{2.24}$$

where $\tau > 0$ denotes the time step and the discrete approximation of the time derivative is for a sequence $(\varphi^n)$ given by

$$\partial_\tau \varphi^n = \frac{\varphi^n - \varphi^{n-1}}{\tau}. \tag{2.25}$$

For the fully discrete scheme, we use the same initial value and set $y_h^0 = J_h y^0 = \big(\mathscr{L}_h^{V*}u^0, \mathscr{L}_h^{V*}v^0\big)^T$ as in (2.21). Our second main result on the full discretization, which is proved in Section 4, then reads as follows.

THEOREM 2.9 Let $\partial\Omega \in C^{k+1,1}$, $h \leqslant h_0$, and let Assumption 2.1 hold for $\ell_{\max} \geqslant 2$ and $\kappa = k+1$. Further, assume that the right-hand side $f$ and the solution of (1.1) satisfy in addition to (2.23)

$$\begin{aligned} &u \in C^5\big([0,T], H^1(\Omega)\big) \cap C^4\big([0,T], \mathscr{H}_\Delta^1\big) \cap C^3\big([0,T], \mathscr{H}_\Delta^2\big) \cap C^2\big([0,T], \mathscr{H}_\Delta^3\big), \\ &f \in C^1\big([0,T], \mathscr{H}_\Delta^{k^*}\big) \cap C\big([0,T], \mathscr{H}_\Delta^{k^*+1}\big), \\ &u^0 \in \mathscr{H}_\Delta^{k^*+3}, \quad v^0 \in \cap \mathscr{H}_\Delta^{k^*+2}, \end{aligned}$$

and the initial value $y_h^0$ is chosen as in (2.21). Then, there is $\tau_0 > 0$ such that for $\tau \leqslant \tau_0$ we have the error bound

$$\|y(t^n) - \boldsymbol{\mathscr{L}_h} y_h^n\|_{W^{1,\infty}\times L^\infty} \leqslant C\tau + Ch^{\min\{k,\ell_{\max}\}}, \quad n \geqslant 2,$$

where $C$ is independent of $h$ and $\tau$, and $\tau_0$ is independent of $h$.

We refer to Remark 5.10 below in order to explain the minimum in the convergence rate. Further, we emphasize that the first two approximations do not enter the above error bound. However, we have the following convergence result for the first approximations.

THEOREM 2.10 Let the assumptions of Theorem 2.9 hold. Then, we have

$$\big\|y(t^\ell) - \boldsymbol{\mathscr{L}_h} y_h^\ell\big\|_{W^{1,\infty}\times L^\infty} \leqslant C\big(\tau + h^k\big), \quad \ell = 0,1.$$

The proof is given in Section 5.

REMARK 2.11 Considering the strategy of the proof, we see that the same ideas can be applied by only differentiating the error equations once and to exploit the relation

$$X_h = H^1 \times L^2 \xrightarrow{\mathrm{T}_h} L^\infty \times H^1 .$$

Following the lines of the presented proof below, we obtain the very same estimates for the $L^\infty \times H^1$ norm. The gain in using the "weaker" norm are the decreased regularity assumptions as well as the simpler choice of the initial value $y_h^0 = \big(\mathscr{L}_h^{V*}u^0, I_h v^0\big)^T$, where we save one computation of $\mathscr{L}_h^{V*}$.

## 3. Analysis of the space discretization

### 3.1 *Strategy of the proof*

We now prove Theorem 2.8, i.e., we derive an error bound for the spatially discrete approximation obtained by (2.20) in the maximum norm. To this end, we proceed as follows. We split the error in

$$y(t) - \mathscr{L}_{\boldsymbol{h}} y_h(t) = \big(\mathrm{Id} - \mathscr{L}_{\boldsymbol{h}} J_h\big) y(t) + \mathscr{L}_{\boldsymbol{h}}\big(J_h y(t) - y_h(t)\big) =: e_{J_h}(t) + \mathscr{L}_{\boldsymbol{h}} e_h(t) \tag{3.1}$$

and derive an equation for the discrete error $e_h$. With the solution operator T, we rewrite (2.2) as

$$\mathrm{T}\Lambda(t)\partial_t y = y + \mathrm{T}\Lambda(t)F(t), \qquad t \in [0,T],$$

with initial value $y(0) = y^0$. Correspondingly, we use the discrete solution operator $\mathrm{T}_h$ to obtain from (2.20) the semi-discrete equation

$$\mathrm{T}_h\Lambda_h(t)\partial_t y_h = y_h + \mathrm{T}_h\Lambda_h(t)F_h(t), \qquad t \in [0,T], \tag{3.2}$$

with initial value $y_h(0) = y_h^0$. Thus, we conclude that the discrete error $e_h$ solves the evolution equation

$$\mathrm{T}_h\Lambda_h(t)\partial_t e_h(t) = e_h(t) + \delta_{h,\mathrm{T}}(t), \qquad t \in [0,T], \tag{3.3}$$

with initial value $e_h(0) = e_h^0 = J_h y_0 - y_h^0$ and the defect

$$\delta_{h,\mathrm{T}}(t) = \big(\mathrm{T}_h\Lambda_h(t)J_h - J_h\mathrm{T}\Lambda(t)\big)\partial_t y(t) + J_h J_h \mathrm{T}\Lambda(t)F(t) - \mathrm{T}_h\Lambda_h(t)F_h(t). \tag{3.4}$$

As illustrated in Figure 1, the proof of Theorem 2.8 mainly consists of two steps. First, in Lemma 3.1 we exchange the maximum norm of $e_h(t)$ for bounds of time derivatives of $e_h(t)$ in $X_h$. To do so, we use (3.3) and Lemma 2.7, i.e., we rely on the property of the solution operator to gain integrability as sketched in (2.22). Note that we can view the error equation as an ordinary differential equation in a finite-dimensional space with right-hand side, say $g(t,e_h)$. Since the right-hand side $g$ is smooth by our assumptions, we obtain local existence. Since all norms are equivalent, the error bounds below guarantee existence up to the final time $T$.

Next, in Lemma 3.2 we trace back the time derivatives of $e_h(t)$ to time derivatives of the initial error $e_h(0)$, which can be bounded due to the choice (2.21) of the discrete initial value $y_h^0$. Here, we obtain from (2.2) and (2.20) for the discrete error $e_h$ the evolution equation

$$\Lambda_h(t)\partial_t e_h(t) = \mathrm{A}_h e_h(t) + \delta_{h,\mathrm{A}}(t), \qquad t \in [0,T], \tag{3.5}$$

with the defect

$$\delta_{h,\mathrm{A}}(t) = \big(\Lambda_h(t)J_h - J_h\Lambda(t)\big)\partial_t y(t) + \big(J_h\mathrm{A} - \mathrm{A}_h J_h\big)y(t) + J_h\Lambda(t)F(t) - \Lambda_h(t)F_h(t). \tag{3.6}$$

Note that we have the relation $\delta_{h,\mathrm{A}} = \mathrm{A}_h\delta_{h,\mathrm{T}}$. Moreover, we emphasize that a similar defect was already studied in the unified error analysis provided in Hipp *et al.* (2019). However, here we also have to bound time derivatives of $\delta_{h,\mathrm{T}}$ and $\delta_{h,\mathrm{A}}$ as well as a different choice of $J_h$. We postpone the derivation of these bounds as well as the estimates for the time derivatives of the initial error to Section 5.

$$
\begin{tikzcd}
\|e_h(t)\|_{W^{1,\infty}\times L^\infty} \arrow[r, "\text{Lemma 3.1}"] \arrow[d, bend left] & \|\partial_t^j e_h(t)\|_{X_h} \arrow[r, "\text{Lemma 3.2}"] \arrow[d, bend left] & \begin{array}{c}\|\partial_t^j e_h(0)\|_{X_h}\\ \text{Lemma 5.6}\end{array} \\
\begin{array}{c}\|\partial_t^j \delta_{h,\mathrm{T}}(t)\|_{W^{1,\infty}\times L^\infty}\\ \text{Lemma 5.2}\end{array} & \begin{array}{c}\|\partial_t^j \delta_{h,\mathrm{A}}(t)\|_{X_h}\\ \text{Lemma 5.5}\end{array} &
\end{tikzcd}
$$

FIG. 1: Strategy of the proof of Theorem 2.8. The forked arrows indicate how the corresponding term is estimated, and the lemmas below show where the term is bounded.

### 3.2 *Proof of* Theorem 2.8

Our first result shows how to bound the error in the maximum norm in terms of time derivatives of the error in the energy norm.

LEMMA 3.1 Let the assumptions of Theorem 2.8 hold. Then, it holds

$$\|e_h\|_{L^\infty(W^{1,\infty}\times L^\infty)} \lesssim \|\partial_t^1 e_h\|_{L^\infty(X_h)} + \|\partial_t^2 e_h\|_{L^\infty(X_h)} + \|\delta_{h,\mathrm{T}}\|_{L^\infty(W^{1,\infty}\times L^\infty)} + \|\partial_t \delta_{h,\mathrm{T}}\|_{L^\infty(L^\infty\times L^4)}.$$

*Proof.* From the representation in (3.3) we obtain with the properties of $\mathrm{T}_h$ in Lemma 2.7 and (2.22), and the bounds on $\Lambda_h$ in Lemma 2.6

$$\|e_h(t)\|_{W^{1,\infty}\times L^\infty} \lesssim \|\partial_t e_h(t)\|_{L^\infty\times L^4} + \|\delta_{h,\mathrm{T}}(t)\|_{W^{1,\infty}\times L^\infty}.$$

For the derivative we compute using the same estimates

$$\begin{aligned}
\|\partial_t e_h(t)\|_{L^\infty\times L^4} &\lesssim \|\mathrm{T}_h\Lambda_h(t)\partial_t^2 e_h(t)\|_{L^\infty\times L^4} + \|\mathrm{T}_h\partial_t\Lambda_h(t)e_h(t)\|_{L^\infty\times L^4} + \|\partial_t\delta_{h,\mathrm{T}}(t)\|_{L^\infty\times L^4} \\
&\lesssim \|\partial_t^2 e_h(t)\|_{X_h} + \|\partial_t\Lambda_h(t)e_h(t)\|_{X_h} + \|\partial_t\delta_{h,\mathrm{T}}(t)\|_{L^\infty\times L^4},
\end{aligned} \tag{3.7}$$

and, using Sobolev's embedding $\|\xi_h\|_{L^4\times L^2} \lesssim \|\xi_h\|_{X_h}$ for $\xi_h \in X_h$, concludes the proof. □

In the following lemma, we provide the bounds of the time derivatives appearing in Lemma 3.1 in the $X_h$ norm using the initial errors and certain defects.

LEMMA 3.2 Let the assumptions of Theorem 2.8 hold. Then, there is a constant $C > 0$ independent of $h$ such that for $j = 1,2$ we have

$$\|\partial_t^j e_h\|_{L^\infty(X_h)}^2 \lesssim \mathrm{e}^{CT}\sum_{\ell=1}^{j}\left(\|\partial_t^\ell e_h(0)\|_{X_h}^2 + \|\partial_t^\ell \delta_{h,\mathrm{A}}\|_{L^\infty(X_h)}^2\right).$$

*Proof.* In the following, we prove for $j = 1,2$ the estimate

$$\|\partial_t^j e_h\|_{L^\infty(X_h)}^2 \lesssim (1+T)\mathrm{e}^{CT}\left(\|\partial_t^j e_h(0)\|_{X_h}^2 + \|\partial_t^j \delta_{h,\mathrm{A}}\|_{L^\infty(X_h)}^2 + \sum_{\ell=1}^{j-1}\|\partial_t^\ell e_h\|_{L^\infty(X_h)}^2\right). \tag{3.8}$$

The result then follows from using this bound recursively. In the following, we often suppress the time arguments to increase the readability.

To prove (3.8), we first obtain by taking the derivative of (3.5) with respect to time

$$\sum_{\ell=0}^{j}\binom{j}{\ell}\partial_t^{j-\ell}\Lambda_h\partial_t^{\ell+1}e_h - \mathrm{A}_h\partial_t^j e_h = \partial_t^j\delta_{h,\mathrm{A}},$$

for $j = 0,1,2$. Taking the inner product with $\partial_t^j e_h$ gives

$$\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_{X_h} = \left(\mathrm{A}_h \partial_t^j e_h \mid \partial_t^j e_h\right)_{X_h} + \left(\partial_t^j \delta_{h,\mathrm{A}} \mid \partial_t^j e_h\right)_{X_h} - \sum_{\ell=0}^{j-1} \binom{j}{\ell} \left(\partial_t^{j-\ell} \Lambda_h \partial_t^{\ell+1} e_h \mid \partial_t^j e_h\right)_{X_h}.$$

Since $\mathrm{A}_h$ is skew-symmetric with respect to the inner product of $X_h$, we obtain with the triangle inequality and Young' inequality the bound

$$2\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_{X_h} \leqslant \left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_{X_h}^2 + 2^j \left\|\partial_t^j e_h\right\|_{X_h}^2 + \sum_{\ell=0}^{j-1} \binom{j}{\ell} \left\|\partial_t^{j-\ell} \Lambda_h \partial_t^{\ell+1} e_h\right\|_{X_h}^2.$$

In particular, due to the boundedness of $\Lambda_h$ by Lemma 2.6 and the corresponding time derivatives, we conclude

$$2\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_{X_h} \leqslant \left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_{X_h}^2 + C_j \left\|\partial_t^j e_h\right\|_{X_h}^2 + \widehat{C}_j \sum_{\ell=1}^{j-1} \left\|\partial_t^\ell e_h\right\|_{X_h}^2, \tag{3.9}$$

with the constants

$$C_j = 2^j + j\left\|\partial_t \Lambda_h\right\|_{L^\infty(\mathcal{L}(X_h))}^2, \qquad \widehat{C}_j = \max_{\ell=0,\dots,j-1} \binom{j}{\ell} \left\|\partial_t^{j-\ell} \Lambda_h\right\|_{L^\infty(\mathcal{L}(X_h))}^2.$$

Note that these constants are bounded independently of $j \leqslant 2$ by $C_2$ and $\widehat{C}_2$, respectively.

We rely on (3.9) to bound the first term on the right-hand side of

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\|\partial_t^j e_h\right\|_{\Lambda_h(t)}^2 = 2\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_{X_h} + \left(\partial_t \Lambda_h \partial_t^j e_h \mid \partial_t^j e_h\right)_{X_h}.$$

Moreover, integration in time, using the boundedness of $\partial_t \Lambda_h$ for the second term, and the norm equivalence (2.18) yields

$$\left\|\partial_t^j e_h\right\|_{\Lambda_h(t)}^2 \lesssim \left\|\partial_t^j e_h(0)\right\|_{\Lambda_h(0)}^2 + t\left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_{L^\infty(X_h)}^2 + t\sum_{\ell=1}^{j-1} \left\|\partial_t^\ell e_h\right\|_{L^\infty(X_h)}^2 + \int_0^t \left\|\partial_t^j e_h(s)\right\|_{\Lambda_h(s)}^2 \mathrm{d}s.$$

Finally, the Gronwall inequality implies for all $t \in [0,T]$

$$\left\|\partial_t^j e_h(t)\right\|_{\Lambda_h(t)}^2 \lesssim \mathrm{e}^{Ct}\left(\left\|\partial_t^j e_h(0)\right\|_{\Lambda_h(0)}^2 + t\left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_{L^\infty(X_h)}^2 + \sum_{\ell=1}^{j-1} \left\|\partial_t^\ell e_h\right\|_{L^\infty(X_h)}^2\right),$$

and (3.8) follows with (2.18). □

With these preparations we can prove our first main result.

*Proof of* Theorem 2.8. Using the decomposition (3.1) and the stability of the lift in (2.7), we estimate with the approximation property derived in (2.16)

$$\begin{aligned}\|y(t) - \mathscr{L}_h y_h(t)\|_{W^{1,\infty}\times L^\infty} &\leqslant \|e_{J_h}(t)\|_{W^{1,\infty}\times L^\infty} + C_{\mathscr{L}_h}\|e_h(t)\|_{W^{1,\infty}\times L^\infty} \\ &\leqslant Ch^k \|y(t)\|_{W^{k+1,\infty}\times W^{k+1,\infty}} + C_{\mathscr{L}_h}\|e_h(t)\|_{W^{1,\infty}\times L^\infty},\end{aligned}$$

and apply Lemmas 3.1 and 3.2. The remaining defects and errors in the initial values are bounded in Lemmas 5.2, 5.5, and 5.6. □

$$
\begin{tikzcd}
\|e_h^n\|_{W^{1,\infty}\times L^\infty} \arrow[rr, "\text{Lemma 4.4}"] \arrow[dr, bend left] & & \|\partial_\tau^j e_h^n\|_{X_h} \arrow[rr, "\text{Lemma 4.6}"] \arrow[dr, bend left] & & \begin{array}{c}\|\partial_\tau^j e_h^j\|_{X_h} \\ \text{Lemma 5.7}\end{array} \\
& \begin{array}{c}\|\partial_\tau^j \delta_{h,\mathrm{T}}^n\|_{W^{1,\infty}\times L^\infty} \\ \text{Lemma 5.2}\end{array} & & \begin{array}{c}\|\partial_\tau^j \delta_{h,\mathrm{A}}^n\|_{X_h} \\ \text{Lemma 5.5}\end{array} &
\end{tikzcd}
$$

FIG. 2: Strategy of the proof of Theorem 2.9. The forked arrows indicate how the corresponding term is estimated, and the lemmas below show where the term is bounded.

## 4. Analysis of the full discretization

In this section, we establish the proof of Theorem 2.9. The strategy is very similar to the one in Section 3, see Figure 2, where we replace the continuous by discrete derivatives. Hence, after introducing some useful calculus, we explain the adapted strategy.

### 4.1 *Calculus for discrete derivatives*

We first need some auxiliary results for the discrete derivatives defined in (2.25). A straightforward calculation yields the following.

LEMMA 4.1 It holds the discrete product rule

$$\partial_\tau(\varphi^n\psi^n) = (\partial_\tau\varphi^n)\psi^n + \varphi^{n-1}(\partial_\tau\psi^n)$$

and also the more general discrete Leibniz rule

$$\partial_\tau^j(\varphi^n\psi^n) = \sum_{\ell=0}^{j}\binom{j}{\ell}\big(\partial_\tau^{j-\ell}\varphi^{n-\ell}\big)\big(\partial_\tau^{\ell}\psi^n\big), \quad j \geqslant 0.$$

In order to mimic the strategy of the proof of Theorem 2.8, we state the well-known discrete version of the fundamental theorem of calculus and a direct consequence of a discrete Gronwall lemma.

LEMMA 4.2 Let $(\varphi^n)$ be a sequence in a Hilbert space with inner product $(\cdot \mid \cdot)$, and let $k_0 \in \mathbb{N}$.

(a) For any $M \geqslant k_0$, it holds

$$\frac{1}{2}\|\varphi^M\|^2 \leqslant \frac{1}{2}\|\varphi^{k_0-1}\|^2 + \tau\sum_{j=k_0}^{M}\big(\partial_\tau\varphi^j \mid \varphi^j\big)\,.$$

(b) If there are constants $\alpha, \beta_1, \beta_2 \geqslant 0$ such that

$$\big(\partial_\tau\varphi^j \mid \varphi^j\big) \leqslant \alpha^2 + \beta_1\big\|\varphi^{j-1}\big\|^2 + \beta_2\big\|\varphi^j\big\|^2, \quad j \geqslant k_0, \tag{4.1}$$

holds, then for $\tau \leqslant \frac{1}{4(\beta_1+\beta_2)}$ and $M \geqslant k_0$ we have

$$\|\varphi^M\| \leqslant \Big(\sqrt{1+2\tau\beta_1}\|\varphi^{k_0-1}\| + \sqrt{2t_N}\,\alpha\Big)\mathrm{e}^{2(\beta_1+\beta_2)t_N}.$$

*Proof.* Part $(a)$ is for example shown in (Hochbruck & Pažur, 2017, Lemma 4.2). Inserting (4.1) in $(a)$ yields

$$\begin{aligned}\|\varphi^N\|^2 &\leqslant \|\varphi^{k_0-1}\|^2 + 2\tau \sum_{j=k_0}^{N} \Big(\alpha^2 + \beta_1 \|\varphi^{j-1}\|^2 + \beta_2 \|\varphi^j\|^2\Big) \\ &\leqslant \big(1 + 2\tau\beta_1\big)\|\varphi^{k_0-1}\|^2 + 2t_N\alpha^2 + 2(\beta_1+\beta_2)\tau \sum_{j=k_0}^{N} \|\varphi^j\|^2\end{aligned}$$

and by a Gronwall argument, see, e.g., (Linz, 1969, Lemma 1), we obtain

$$\|\varphi^N\|^2 \leqslant \Big(\big(1+2\tau\beta_1\big)\|\varphi^{k_0-1}\|^2 + 2t_N\alpha^2\Big)\mathrm{e}^{4(\beta_1+\beta_2)t_N}.$$

Taking roots yields the assertion. □

We conclude with a useful bound which relates the discrete derivatives to their continuous limit.

LEMMA 4.3 Let $Z$ be some Banach space, $j \geqslant 1$ and $x\colon [0,T] \to Z$ be $j$-times differentiable with bounded derivatives, then

$$\|\partial_\tau^j x(t^n)\|_Z \leqslant \sup_{t\in[t^{n-j},t^n]} \|\partial_t^j x(t)\|_Z.$$

*Proof.* This simply follows from an iterative application of the fundamental theorem of calculus. □

### 4.2 *Proof of* Theorem 2.9

As in (3.1), we are interested in bounds on the discrete error

$$e_h^n = J_h y(t^n) - y_h^n,$$

and derive for the exact solution inserted in the numerical scheme similar to (3.2)

$$\mathrm{T}_h\Lambda_h(t^n)J_h\partial_\tau y(t^n) = J_h y(t^n) + \mathrm{T}_h\Lambda_h(t^n)F_h(t^n) + \delta_{h,\mathrm{T}}^n$$

with a defect of the form, using the representation of $\delta_{h,\mathrm{T}}$ in (3.4),

$$\delta_{h,\mathrm{T}}^n = \delta_{h,\mathrm{T}}(t^n) + \mathrm{T}_h\Lambda_h(t^n)J_h\big(\partial_\tau y(t^n) - \partial_t y(t^n)\big). \tag{4.2}$$

From this we obtain the fully discrete error equation

$$\mathrm{T}_h\Lambda_h(t^n)\partial_\tau e_h^n = e_h^n + \delta_{h,\mathrm{T}}^n. \tag{4.3}$$

For the estimates in the energy we need the equivalent formulation involving the operator $\mathrm{A}_h$. To this end, we insert $J_h y$ into (2.24) and obtain

$$\Lambda_h(t^n)J_h\partial_\tau y(t^n) = \mathrm{A}_h J_h y(t^n) + F_h(t^n) + \delta_{h,\mathrm{A}}^n$$

with the defect, using the representation of $\delta_{h,\mathrm{A}}$ in (3.6),

$$\delta_{h,\mathrm{A}}^n = \delta_{h,\mathrm{A}}(t^n) + \Lambda_h(t^n)J_h\big(\partial_\tau y(t^n) - \partial_t y(t^n)\big). \tag{4.4}$$

This gives us the second version of the error recursion

$$\Lambda_h(t^n)\partial_\tau e_h^n = \mathrm{A}_h e_h^n + \delta_{h,\mathrm{A}}^n. \tag{4.5}$$

Starting from (4.3), we obtain the following bound as a discrete counterpart to Lemma 3.1.

LEMMA 4.4 Let the assumptions of Theorem 2.9 hold. Then, there exists a constant $C > 0$ independent of $h$, $\tau$, and $n$ such that

$$\|e_h^n\|_{W^{1,\infty}\times L^\infty} \lesssim \|\partial_\tau^1 e_h^n\|_{X_h} + \|\partial_\tau^2 e_h^n\|_{X_h} + \|\delta_{h,\mathrm{T}}^n\|_{W^{1,\infty}\times L^\infty} + \|\partial_\tau^1 \delta_{h,\mathrm{T}}^n\|_{L^\infty\times L^4}$$

holds for $n \geqslant 2$.

REMARK 4.5 From this lemma it becomes clear that this technique does not provide bounds for $n = 0, 1$, since we can evaluate $\partial_\tau^j e_h^n$ only for $n \geqslant j$.

*Proof of* Lemma 4.4. As in (3.7), we obtain from (4.3) and Lemma 4.1

$$\partial_\tau^1 e_h^n = -\partial_\tau^1 \delta_{h,\mathrm{T}}^n + \mathrm{T}_h \Lambda_h(t^{n-1})\partial_\tau^2 e_h^n + \mathrm{T}_h \partial_\tau^1 \Lambda_h(t^n)\partial_\tau^1 e_h^n,$$

and the proof follows along the lines of Lemma 3.1. □

The next step is to establish the discrete analogue to Lemma 3.2 where the discrete derivatives are bounded in terms of discrete derivatives of the initial error and defects.

LEMMA 4.6 Let the assumptions of Theorem 2.9 hold. Then, there is $\tau_0 > 0$ such that for $\tau \leqslant \tau_0$

$$\|\partial_\tau^j e_h^n\|_{X_h}^2 \leqslant C(1+T)\mathrm{e}^{CT} \sum_{\ell=1}^{j} \Big( \|\partial_\tau^\ell e_h^\ell\|_{X_h}^2 + \|\partial_\tau^\ell \delta_{h,\mathrm{A}}^n\|_{\ell^\infty(X_h)}^2 \Big),$$

for $j = 1, 2$ and $n \geqslant j+1$.

*Proof.* As in the proof of Lemma 3.2, we provide the bound

$$\|\partial_\tau^j e_h^n\|_{X_h}^2 \leqslant C(1+T)\mathrm{e}^{CT} \Bigg( \|\partial_\tau^j e_h^j\|_{X_h}^2 + \|\partial_\tau^j \delta_{h,\mathrm{A}}^n\|_{\ell^\infty(X_h)}^2 + \sum_{\ell=1}^{j-1} \|\partial_\tau^\ell e_h^n\|_{X_h}^2 \Bigg), \tag{4.6}$$

cf. (3.8). Using this estimate recursively directly yields the assertion.

To do so, we apply the bounds from Lemma 4.2 for $\varphi^n = \Lambda_h^{1/2}(t^n)\partial_\tau^j e_h^n$. In particular, we first study the term

$$(\partial_\tau \varphi^n \mid \varphi^n)_{X_h} = \Big( \Lambda_h(t^n)\partial_\tau^{j+1} e_h^n \mid \partial_\tau^j e_h^n \Big)_{X_h} + \Big( \big(\partial_\tau \Lambda_h^{1/2}(t^n)\big)\partial_\tau^j e_h^{n-1} \mid \varphi^n \Big)_{X_h},$$

where we used Lemma 4.1 and the fact that $\Lambda_h^{1/2}$ is self-adjoint in $X_h$. Due to Assumption 2.1 and Young's inequality, this implies

$$(\partial_\tau \varphi^n \mid \varphi^n)_{X_h} \leqslant \Big( \Lambda_h(t^n)\partial_\tau^{j+1} e_h^n \mid \partial_\tau^j e_h^n \Big)_{X_h} + C\|\varphi^n\|_{X_h}^2 + C\,\|\varphi^{n-1}\|_{X_h}^2. \tag{4.7}$$

For the first term, we obtain with Lemma 4.1 applied to (4.5)

$$\Lambda_h(t^n)\partial_\tau^{j+1} e_h^n = \mathrm{A}_h \partial_\tau^j e_h^n + \partial_\tau^j \delta_{h,\mathrm{A}}^n - \sum_{\ell=0}^{j-1} \binom{j}{\ell} \big(\partial_\tau^{j-\ell}\Lambda_h(t^{n-\ell})\big)\big(\partial_\tau^{\ell+1} e_h^n\big).$$

Thus, taking the inner product in $X_h$ with $\partial_\tau^j e_h^n$ yields as in (3.9) the estimate

$$2\Big( \Lambda_h(t^n)\partial_\tau^{j+1} e_h^n \mid \partial_\tau^j e_h^n \Big)_{X_h} \leqslant \|\partial_\tau^j \delta_{h,\mathrm{A}}^n\|_{X_h}^2 + C\|\varphi^n\|_{X_h}^2 + C\sum_{\ell=1}^{j-1} \|\partial_\tau^\ell e_h^n\|_{X_h}^2.$$

Using this bound in (4.7) together with the discrete Gronwall lemma in Lemma 4.2, we obtain (4.6) for $\tau$ sufficiently small. □

Hence, we conclude our second main result.

*Proof of* Theorem 2.9. Using a decomposition analogous to (3.1), we estimate as in the proof of Theorem 2.8

$$\begin{aligned}\|y(t^n) - \mathscr{L}_h y_h^n\|_{W^{1,\infty}\times L^\infty} &\leqslant \|e_{J_h}(t^n)\|_{W^{1,\infty}\times L^\infty} + C_{\mathscr{L}_h}\|e_h^n\|_{W^{1,\infty}\times L^\infty} \\ &\leqslant Ch^k\|y(t^n)\|_{W^{k+1,\infty}\times W^{k+1,\infty}} + C_{\mathscr{L}_h}\|e_h^n\|_{W^{1,\infty}\times L^\infty},\end{aligned}$$

and apply Lemmas 4.4 and 4.6 for $n \geqslant 2$. The remaining defects and errors in the initial values are bounded in Lemmas 5.2, 5.5, and 5.7. □

## 5. Bounds on the defects and initial conditions

In this section, we provide all bounds missing in the proofs of Sections 3 and 4. Throughout this section, we mostly omit the time dependency for the sake of readability and assume $h \leqslant h_0$. Further, we take the assumptions of Section 2 as given, but will be precise about the regularity of the solution $u$ and the right-hand side $f$.

### 5.1 *Estimates of the defects*

We first provide certain approximation properties in the maximum norm which are used for the defects. We recall the notation $k^* = \max\{k, 2\}$.

LEMMA 5.1 Let $\xi = (\varphi, \psi) \in H^1(\Omega) \times H^{k+1}(\Omega)$ and $\lambda, f \in C^1([0,T], H^{k+1}(\Omega))$. Then, the discrete operators introduced in Section 2 satisfy the bounds

$$\big\|\big(\mathrm{T}_h J_h - J_h \mathrm{T}\big)\xi\big\|_{W^{1,\infty}\times L^\infty} \lesssim h^k\|\psi\|_{H^{k+1}}, \tag{5.1a}$$

$$\big\|\big(\mathrm{T}_h J_h - J_h \mathrm{T}\big)\xi\big\|_{L^\infty\times L^4} \lesssim h^k\|\psi\|_{H^{k^*}}, \tag{5.1b}$$

$$\big\|\big(\Lambda_h J_h - J_h \Lambda\big)\xi\big\|_{L^\infty\times L^4} \lesssim h^k\|\psi\|_{H^{k+1}}. \tag{5.1c}$$

For $j = 0, 1$ we further have

$$\big\|\partial_t^j\big(\Lambda_h J_h - J_h\Lambda\big)\xi\big\|_{H^1\times L^2} \lesssim h^k\|\psi\|_{H^{k^*}}, \tag{5.2a}$$

$$\big\|\partial_t^j\Lambda_h F_h - J_h\partial_t^j\Lambda F\big\|_{L^\infty\times L^4} \lesssim h^k\big\|\partial_t^j(\lambda f)\big\|_{H^{k+1}}. \tag{5.2b}$$

*Proof.* Using the identity $\mathscr{L}_h^{V*}\Delta^{-1}\psi = \Delta_h^{-1}\mathscr{L}_h^{H*}\psi$, we compute

$$\big(\mathrm{T}_h J_h - J_h\mathrm{T}\big)\xi = \begin{pmatrix}\Delta_h^{-1}\big(\mathscr{L}_h^{V*} - \mathscr{L}_h^{H*}\big)\psi \\ 0\end{pmatrix}.$$

Using Lemma 2.7, we obtain

$$\begin{aligned}\big\|\Delta_h^{-1}\big(\mathscr{L}_h^{V*} - \mathscr{L}_h^{H*}\big)\psi\big\|_{L^\infty} &\lesssim \big\|\big(\mathscr{L}_h^{V*} - \mathscr{L}_h^{H*}\big)\psi\big\|_{L^2}, \\ \big\|\Delta_h^{-1}\big(\mathscr{L}_h^{V*} - \mathscr{L}_h^{H*}\big)\psi\big\|_{W^{1,\infty}} &\lesssim \big\|\big(\mathscr{L}_h^{V*} - \mathscr{L}_h^{H*}\big)\psi\big\|_{V_h},\end{aligned}$$

and the bound (5.1a) and (5.1b) follow from (2.10) and (2.11). For the bound on (5.1c) we exploit the bounds in (2.8), (2.10), (2.11), (2.13), and (2.14) to obtain

$$\begin{aligned}\big\|\big(\Lambda_h J_h - J_h\Lambda\big)\xi\big\|_{L^\infty\times L^4} &\lesssim \big\|\lambda_h \mathscr{L}_h^{V*}\psi - \mathscr{L}_h^{V*}(\lambda\psi)\big\|_{H^1}\\ &\lesssim \big\|\pi_h\big(I_h\lambda\mathscr{L}_h^{V*}\psi - \mathscr{L}_h^{-1}(\lambda\psi)\big)\big\|_{H^1} + \big\|\big(\pi_h\mathscr{L}_h^{-1} - \mathscr{L}_h^{H*}\big)(\lambda\psi)\big\|_{H^1}\\ &\quad + \big\|\big(\mathscr{L}_h^{H*} - \mathscr{L}_h^{V*}\big)(\lambda\psi)\big\|_{H^1}\\ &\lesssim h^k\big(\|\psi\|_{H^{k+1}} + \|\lambda\psi\|_{H^{k+1}}\big),\end{aligned} \tag{5.3}$$

and in the same way for $j = 0,1$

$$\big\|\partial_t^j\big(\Lambda_h J_h - J_h\Lambda\big)\xi\big\|_{H^1\times L^2} \lesssim h^k\big(\|\psi\|_{H^{k^*}} + \|(\partial_t^j\lambda)\psi\|_{H^{k^*}}\big).$$

Similarly, (5.2b) follows from

$$\begin{aligned}\big\|\partial_t^j\Lambda_h F_h - J_h\partial_t^j\Lambda F\big\|_{L^\infty\times L^4} \leqslant \big\|\partial_t^j\pi_h\big(I_h\lambda I_h f - \mathscr{L}_h^{-1}(\lambda f)\big)\big\|_{L^4} + \big\|\partial_t^j\big(\pi_h\mathscr{L}_h^{-1} - \mathscr{L}_h^{H*}\big)(\lambda f)\big\|_{L^4}\\ + \big\|(\mathscr{L}_h^{H*} - \mathscr{L}_h^{V*})\partial_t^j(\lambda f)\big\|_{L^4}\end{aligned} \tag{5.4}$$

by (2.8), (2.10), (2.11), (2.13), and (2.14). □

From these approximation properties we conclude the bounds on the defects appearing in Lemmas 3.1 and 4.4.

LEMMA 5.2 Let the solution $u \in C^3([0,T],H^{k^*}(\Omega)\cap C^2([0,T],H^{k+1}(\Omega))$ and $\lambda, f \in C^1([0,T],H^{k+1}(\Omega))$.

(a) The defect $\delta_{h,\mathrm{T}}(t)$ introduced in (3.4) satisfies for $t \in [0,T]$

$$\big\|\delta_{h,\mathrm{T}}(t)\big\|_{W^{1,\infty}\times L^\infty} + \big\|\partial_t\delta_{h,\mathrm{T}}(t)\big\|_{L^\infty\times L^4} \lesssim h^k.$$

(b) If in addition $u \in C^2([0,T],W^{1,\infty}(\Omega))\cap C^4([0,T],H^1(\Omega))$, then the defect $\delta_{h,\mathrm{T}}^n$ introduced in (4.2) satisfies for $n \geqslant 2$

$$\big\|\delta_{h,\mathrm{T}}^n\big\|_{W^{1,\infty}\times L^\infty} + \big\|\partial_\tau\delta_{h,\mathrm{T}}^n\big\|_{L^\infty\times L^4} \lesssim \tau + h^k.$$

*Proof.* (a) We decompose the defect $\delta_{h,\mathrm{T}} = \delta_{h,\mathrm{T}}^1 + \delta_{h,\mathrm{T}}^2$ introduced in (3.4) with

$$\delta_{h,\mathrm{T}}^1 = \big(\mathrm{T}_h\Lambda_h J_h - J_h\mathrm{T}\Lambda\big)\partial_t y, \qquad \delta_{h,\mathrm{T}}^2 = J_h\mathrm{T}\Lambda F - \mathrm{T}_h\Lambda_h F_h,$$

and first consider $\delta_{h,\mathrm{T}}^1$. We write

$$\begin{aligned}\|\delta_{h,\mathrm{T}}^1\|_{W^{1,\infty}\times L^\infty} &\leqslant \|\mathrm{T}_h\big(\Lambda_h J_h - J_h\Lambda\big)\partial_t y\|_{W^{1,\infty}\times L^\infty} + \big\|\big(\mathrm{T}_h J_h - J_h\mathrm{T}\big)\Lambda\partial_t y\big\|_{W^{1,\infty}\times L^\infty}\\ &\lesssim \big\|\big(\Lambda_h J_h - J_h\Lambda\big)\partial_t y\big\|_{L^\infty\times L^4} + \big\|\lambda\partial_t^2 u\big\|_{H^{k+1}}\\ &\lesssim \big\|\partial_t^2 u\big\|_{H^{k^*}} + \big\|\lambda\partial_t^2 u\big\|_{H^{k+1}},\end{aligned}$$

where we used Lemma 2.7 together with (5.1a) and (5.1c). Differentiating and taking norms, gives with Lemma 2.3

$$\begin{aligned}\|\partial_t\delta_{h,\mathrm{T}}^1\|_{L^\infty\times L^4} &\leqslant \|\mathrm{T}_h\big(\Lambda_h J_h - J_h\Lambda\big)\partial_t^2 y\|_{L^\infty\times L^4} + \|\mathrm{T}_h\big(\partial_t\Lambda_h J_h - J_h\partial_t\Lambda\big)\partial_t y\|_{L^\infty\times L^4}\\ &\quad + \big\|\big(\mathrm{T}_h J_h - J_h\mathrm{T}\big)\partial_t(\Lambda\partial_t y)\big\|_{L^\infty\times L^4}\\ &\leqslant \big\|\big(\Lambda_h J_h - J_h\Lambda\big)\partial_t^2 y\big\|_{H^1\times L^2} + \big\|\big(\partial_t\Lambda_h J_h - J_h\partial_t\Lambda\big)\partial_t y\big\|_{H^1\times L^2}\\ &\quad + \big\|\big(\mathrm{T}_h J_h - J_h\mathrm{T}\big)\partial_t(\Lambda\partial_t y)\big\|_{L^\infty\times L^4}\end{aligned}$$

such that (5.1b) and (5.2a) together imply

$$\left\|\partial_t \delta^1_{h,\mathrm{T}}\right\|_{L^\infty \times L^4} \lesssim h^k \left(\left\|\partial_t^3 u\right\|_{H^{k^*}} + \left\|\partial_t^2 u\right\|_{H^{k^*}} + \left\|\partial_t(\lambda \partial_t u)\right\|_{H^{k^*}}\right).$$

With similar arguments, we obtain for $j = 0, 1$ with (5.1a), Lemma 2.7, and (5.2b)

$$\begin{aligned}\left\|\partial_t^j \delta^2_{h,\mathrm{T}}\right\|_{W^{1,\infty}\times L^\infty} &\leqslant \left\|\left(J_h\mathrm{T} - \mathrm{T}_h J_h\right)\partial_t^j\left(\Lambda(t)F(t)\right)\right\|_{W^{1,\infty}\times L^\infty} + \left\|\mathrm{T}_h\left(J_h\partial_t^j\left(\Lambda(t)F(t)\right) - \partial_t^j\left(\Lambda_h(t)F_h(t)\right)\right)\right\|_{W^{1,\infty}\times L^\infty} \\ &\lesssim h^k\left\|\partial_t^j(\lambda f)\right\|_{H^{k+1}},\end{aligned}$$

and thus the claim.

(b) Thanks to Lemma 4.3, the estimate from part (a) extends to this case, and we only have to provide the following bound for the additional defect

$$\begin{aligned}&\left\|\mathrm{T}_h\Lambda_h(t^n)J_h\left(\partial_\tau y(t^n) - \partial_t y(t^n)\right)\right\|_{W^{1,\infty}\times L^\infty} \lesssim \left\|\partial_\tau y(t^n) - \partial_t y(t^n)\right\|_{W^{1,\infty}\times H^1} \\ &\left\|\partial_\tau \mathrm{T}_h\Lambda_h(t^n)J_h\left(\partial_\tau y(t^n) - \partial_t y(t^n)\right)\right\|_{L^\infty\times L^4} \lesssim \left\|\partial_\tau y(t^n) - \partial_t y(t^n)\right\|_{H^1\times H^1} + \left\|\partial_\tau\left(\partial_\tau y(t^n) - \partial_t y(t^n)\right)\right\|_{H^1\times H^1}.\end{aligned}$$

We note the identity

$$\partial_\tau y(t^n) - \partial_t y(t^n) = \tau\int_0^1 (-s)\partial_t^2 y(t^{n-1} + \tau s)\,\mathrm{d}s, \tag{5.5}$$

and conclude the assertion. □

In the next step, we consider the time derivatives of the error which we estimate in the energy norm, and provide some more approximation properties. We first characterize the inverse of $\lambda_h$, which plays a crucial rule in the subsequent error analysis. We define the inner product $(\varphi_h \mid \psi_h)_{\lambda_h} := (I_h\lambda\,\varphi_h \mid \psi_h)_{L^2(\Omega_h)}$, and the corresponding $L^2$-projection $Q_h$ for $\psi \in L^2(\Omega_h)$ via

$$(Q_h\psi \mid \psi_h)_{\lambda_h} = (\psi \mid \psi_h)_{\lambda_h}, \quad \psi_h \in V_h.$$

This leads us to the following result, which is proved in Appendix A.

LEMMA 5.3 (a) For $\varphi \in H^{k+1}(\Omega)$ and $\psi \in H^1(\Omega_h)$, we have the bounds

$$\left\|Q_h\mathscr{L}_h^{-1}\varphi - \mathscr{L}_h^{H*}\varphi\right\|_{L^2(\Omega_h)} \lesssim h^{k+1}\left\|\varphi\right\|_{H^{k+1}(\Omega)}, \qquad \left\|Q_h\psi\right\|_{H^1(\Omega_h)} \lesssim \left\|\psi\right\|_{H^1(\Omega_h)}.$$

(b) The inverse of $\lambda_h$ is given by $\lambda_h{}^{-1}\varphi_h = Q_h\left((I_h\lambda)^{-1}\varphi_h\right)$ for $\varphi_h \in V_h$.

We further make use in the following of the operators

$$\widetilde{\Lambda}(t^n) := \begin{pmatrix}\lambda(t^n) & 0 \\ 0 & \mathrm{Id}\end{pmatrix} \tag{5.6}$$

and analogously defined, $\widetilde{\Lambda}_h(t^n)$, $\widetilde{\Lambda}_h(t^n)^{-1}$, and $\widetilde{\Lambda}(t^n)^{-1}$.

LEMMA 5.4 Let $\xi = (\varphi, \psi) \in H^{k+1}(\Omega) \times H^{k^*}(\Omega)$ and $\lambda, f \in C^1([0,T], H^{k^*}(\Omega))$. Then, the discrete operators introduced in Section 2 satisfy for $j = 0, 1$ the bounds

$$\left\|\left(\widetilde{\Lambda}_h J_h - J_h\widetilde{\Lambda}\right)\xi\right\|_{X_h} \lesssim h^k\left\|\varphi\right\|_{H^{k+1}}, \tag{5.7a}$$

$$\left\|\left(\widetilde{\Lambda}_h^{-1} J_h - J_h\widetilde{\Lambda}^{-1}\right)\xi\right\|_{X_h} \lesssim h^k\left\|\varphi\right\|_{H^{k+1}}, \tag{5.7b}$$

$$\left\|\partial_t^j\Lambda_h F_h - J_h\partial_t^j\Lambda F\right\|_{X_h} \lesssim h^k\left\|\partial_t^j(\lambda f)\right\|_{H^{k^*}}. \tag{5.7c}$$

If $\xi = (\varphi, \psi) \in (H^{k+3}(\Omega) \cap \mathscr{D}(\Delta^2)) \times (H^{k^*+2}(\Omega) \cap \mathscr{D}(\Delta))$ and $f \in C([0,T], H^{k+1}(\Omega) \cap \mathscr{D}(\Delta))$, then

$$\left\| \mathrm{A}_h F_h - J_h \mathrm{A} F \right\|_{X_h} \lesssim h^k \left\| f \right\|_{H^{k+1}} , \tag{5.8a}$$

$$\left\| \big(\mathrm{A}_h J_h - J_h \mathrm{A}\big) \xi \right\|_{X_h} \lesssim h^k \left\| \Delta \varphi \right\|_{H^{k^*}} , \tag{5.8b}$$

$$\left\| \big(\mathrm{A}_h^2 J_h - J_h \mathrm{A}^2\big) \xi \right\|_{X_h} \lesssim h^k \left\| \xi \right\|_{H^{k+3} \times H^{k^*+2}} . \tag{5.8c}$$

*Proof.* The same computation as in (5.3) yields (5.7a) and with Lemma 5.3 also (5.7b), and the representation in (5.4) implies (5.7c). We obtain

$$\left\| \mathrm{A}_h F_h - J_h \mathrm{A} F \right\|_{X_h} = \left\| I_h f - \mathscr{L}_h^{V*} f \right\|_{V_h} ,$$

and thus with (2.11) and (2.8) the claim. We use the identity $\Delta_h \mathscr{L}_h^{V*} = \mathscr{L}_h^{H*} \Delta$ to derive

$$\big(\mathrm{A}_h J_h - J_h \mathrm{A}\big) \xi = \begin{pmatrix} 0 \\ (\mathscr{L}_h^{H*} - \mathscr{L}_h^{V*}) \Delta \varphi \end{pmatrix} , \qquad \big(\mathrm{A}_h^2 J_h - J_h \mathrm{A}^2\big) \xi = \begin{pmatrix} (\mathscr{L}_h^{H*} - \mathscr{L}_h^{V*}) \Delta \varphi \\ (\mathscr{L}_h^{H*} - \mathscr{L}_h^{V*}) \Delta \psi \end{pmatrix} .$$

We employ (2.10) and (2.11) to obtain (5.8b). Together with the inverse estimate (2.17), this further implies (5.8c). □

With these approximations at hand, we can finally provide bounds for the defects from Lemmas 3.2 and 4.6.

LEMMA 5.5 Let the solution $u \in C^4\big([0,T], H^{k^*}(\Omega)\big) \cap C^2\big([0,T], \mathscr{D}(\Delta)\big)$ and $\lambda, f \in C^2([0,T], H^{k^*}(\Omega))$.

(a) The defect defined in (3.6) satisfies for $j = 1,2$ and $t \in [0,T]$

$$\left\| \partial_t^j \delta_{h,\mathrm{A}}(t) \right\|_{X_h} \lesssim h^k .$$

(b) If in addition $u \in C^5([0,T], H^1(\Omega))$, then the defect defined in (4.4) satisfies for $j = 1,2$ and $n \geqslant j$

$$\left\| \partial_\tau^j \delta_{h,\mathrm{A}}^n \right\|_{X_h} \lesssim \tau + h^k .$$

*Proof.* (a) We estimate the defect with the help of Lemma 5.4. We employ Lemma 2.6 and the bounds (5.2a), (5.7c) and (5.8b) to obtain

$$\left\| \partial_t^j \delta_{h,\mathrm{A}}(t) \right\|_{X_h} \leqslant C h^k \Big( \sum_{\ell=2}^{4} \left\| \partial_t^\ell u \right\|_{H^{k^*}} + \left\| \partial_t^j \Delta u \right\|_{H^{k^*}} + \left\| \partial_t^j (\lambda f) \right\|_{H^{k^*}} \Big).$$

(b) Thanks to Lemma 4.3, it remains to show for $j = 1,2$

$$\left\| \partial_\tau^j \Lambda_h(t^n) J_h \big( \partial_\tau y(t^n) - \partial_t y(t^n) \big) \right\|_{X_h} \lesssim \tau \left\| \partial_t^4 y \right\|_{H^1 \times H^1} .$$

Using the continuity of $J_h$ in $H^1 \times H^1$ and the identity (5.5) immediately yields the claim. □

## 5.2 *Errors of the differentiated initial values*

The last part to prove Theorem 2.8 and Theorem 2.9 is to bound the initial error in Lemma 3.2 and Lemma 4.4. We recall the preconditioned initial values defined in (2.21) where $y_h^0 = y_h(0) = J_h y^0$. The aim of this section is to prove the following bounds.

LEMMA 5.6 Under the assumptions of Theorem 2.8 it holds for $\ell = 1,2$

$$\left\|\partial_t^\ell e_h(0)\right\|_{X_h} \leqslant Ch^k.$$

LEMMA 5.7 Under the assumptions of Theorem 2.9 it holds for $\ell = 1,2$

$$\left\|\partial_\tau^\ell e_h^\ell\right\|_{X_h} \leqslant C\left(\tau + h^k\right).$$

In order to conclude the desired assertion, we proceed in a series of lemmas and introduce the notation $\Lambda^n = \Lambda(t^n)$, $F(t^n) = F(t^n)$, $\Lambda_h^n = \Lambda_h(t^n)$, $F_h(t^n) = F_h(t^n)$ and also

$$\mathscr{R}^n = \left(\Lambda^n - \tau \mathrm{A}\right)^{-1}, \qquad \mathscr{R}_h^n = \left(\Lambda_h^n - \tau \mathrm{A}_h\right)^{-1}. \tag{5.9}$$

We provide a detailed proof of Lemma 5.7 first and explain afterwards how to conclude the assertion of Lemma 5.6. In order to keep the notation simple, we assume without loss of generality in the following that the spatial order satisfies

$$k^* \leqslant \ell_{\max}, \tag{5.10}$$

with $\ell_{\max}$ defined in Assumption 2.1. This induces the restriction $\ell_{\max} \geqslant 2$. In a first step, we find a suitable representation for the discrete and exact solutions.

LEMMA 5.8 The numerical solution can be expanded via

$$\begin{aligned}
\partial_\tau y_h^n &= \mathscr{R}_h^n \mathrm{A}_h y_h^{n-1} + G_h^n, \qquad G_h^n = \mathscr{R}_h^n \Lambda_h^n F_h^n, \\
\partial_\tau^2 y_h^n &= \mathscr{R}_h^n \mathrm{A}_h \mathscr{R}_h^{n-1} \mathrm{A}_h y_h^{n-2} + \left(\partial_\tau \mathscr{R}_h^n\right) \mathrm{A}_h y_h^{n-2} + \mathscr{R}_h^n \mathrm{A}_h G_h^{n-1} + \partial_\tau G_h^n.
\end{aligned}$$

The same holds for $\partial_\tau^\ell y(t^n)$, $\ell = 1,2,3$, with $h$ formally set to zero and

$$G_h^n \to G^n = \mathscr{R}^n \Lambda^n \left(F^n + \delta^n\right), \qquad \delta^n = \partial_\tau y(t^n) - \partial_t y(t^n). \tag{5.11}$$

*Proof.* Starting from (2.24), we multiply by $\Lambda_h^n$ and reorder the terms to obtain

$$\left(\Lambda_h^n - \tau \mathrm{A}_h\right) y_h^n = \Lambda_h^n y_h^{n-1} + \tau \Lambda_h^n F_h^n.$$

Using the resolvent, this further gives

$$y_h^n = y_h^{n-1} + \tau \mathscr{R}_h^n \mathrm{A}_h y_h^{n-1} + \tau \mathscr{R}_h^n \Lambda_h^n F_h^n,$$

which implies the representation for $\partial_\tau y_h^n$. The remaining identity is deduced from the product rule in Lemma 4.1. The results for the exact solution can be derived from the representation

$$\Lambda^n \partial_\tau y(t^n) = \mathrm{A} y(t^n) + \Lambda^n F^n + \Lambda^n \delta^n$$

and the same computations as above. □

In order to bound the expressions in Lemma 5.7, we subtract the representations for the exact and the numerical solution. Let us for example consider the first term in $\partial_\tau^2 e_h^2 = J_h \partial_\tau^2 y(t^2) - \partial_\tau^2 y_h^2$ given by

$$\left(\partial_\tau^2 e_h^2\right)_1 = \left(J_h \mathscr{R}^2 \mathrm{A} \mathscr{R}^1 \mathrm{A} - \mathscr{R}_h^2 \mathrm{A}_h \mathscr{R}_h^1 \mathrm{A}_h J_h\right) y^0,$$

where we used the initial value (2.21) with $y_h^0 = J_h y^0$. In order to bound the difference, we proceed in two steps. First, we move the operators A and $\mathrm{A}_h$ to the right. Therefore, we employ the identities

$$\mathrm{A}\mathscr{R}^n\Lambda^n = \Lambda^n\mathscr{R}^n\mathrm{A}, \qquad \mathrm{A}\Lambda^n = \widetilde{\Lambda}^n\mathrm{A}, \tag{5.12}$$

with $\widetilde{\Lambda}$ defined in (5.6). The corresponding equalities are also valid for the discrete objects and the inverse $\mathrm{A}(\Lambda^n)^{-1} = (\widetilde{\Lambda}^n)^{-1}\mathrm{A}$. Hence, we can write

$$J_h\mathscr{R}^2\mathrm{A}\mathscr{R}^1\mathrm{A}y^0 = J_h\mathscr{R}^2\Lambda^1\mathscr{R}^1(\widetilde{\Lambda}^1)^{-1}\mathrm{A}^2y^0$$

and similarly for the discrete counterpart. A reformulation of Lemma 5.8 according to the above strategy is given in Lemma A.1. The differences of A and $\mathrm{A}_h$ as well as $F$ and $F_h$ are bounded by (5.8). The remaining differences in front of them are treated by the following abstract estimate. As a shorthand notation, we set

$$\prod_{j=1}^m B^j := B^m \dots B^1,\ m \geqslant 1, \qquad \text{and} \qquad \prod_{j=1}^m B^j := \mathrm{Id},\ m < 1.$$

LEMMA 5.9 Let $\mathscr{Y} \subseteq X$ be a Hilbert space, $m \in \mathbb{N}$, and consider operators $B_h^j \in \mathcal{L}(X_h)$ and $B^j \in \mathcal{L}(X)$, $j = 1, \dots, m$, with the following properties:

$$\begin{aligned} \left\|(J_hB^j - B_h^jJ_h)x\right\|_{X_h} &\lesssim \left(\tau + h^k\right)\|x\|_{\mathscr{Y}}, \\ \left\|B^jx\right\|_{\mathscr{Y}} &\lesssim \|x\|_{\mathscr{Y}}. \end{aligned} \tag{5.13}$$

Then, the product is bounded by

$$\left\|\left(J_h\Big(\prod_{j=1}^m B^j\Big) - \Big(\prod_{j=1}^m B_h^j\Big)J_h\right)x\right\|_{X_h} \lesssim \left(\tau + h^k\right)\|x\|_{\mathscr{Y}}.$$

*Proof.* Using the telescopic sum

$$\left(J_h\Big(\prod_{j=1}^m B^j\Big) - \Big(\prod_{j=1}^m B_h^j\Big)J_h\right)x = \sum_{\ell=1}^m \Big(\prod_{j=\ell+1}^m B_h^j\Big)\left(J_hB^\ell - B_h^\ell J_h\right)\Big(\prod_{j=1}^{\ell-1} B^j\Big)x,$$

we immediately conclude the assertion. □

With these preparations, we finally conclude the bounds for the initial values.

*Proof of* Lemma 5.7. We estimate the differences of the continuous and discretized operators in Lemma A.1. We split the proof in two parts. First, we compare the products of bounded operators involving

$$B^j \in \{\Lambda^n, \partial_\tau\Lambda^n, (\widetilde{\Lambda}^n)^{-1}, \mathscr{R}^n\} \tag{5.14}$$

and the discrete counterparts using Lemma 5.9 with $\mathscr{Y} = \mathscr{D}(A^{k^*})$. In the second step, we deal with the powers of A and $\mathrm{A}_h$ applied to the initial value, the right-hand side and the defects. Because of $\partial\Omega \in C^{k+1,1}$, we several times use the embedding $\mathscr{D}(\mathrm{A}^k) \hookrightarrow H^{k+1}(\Omega) \times H^k(\Omega)$, see, e.g., (Grisvard, 1985, Rem. 2.5.1.2).

(i) We first consider the operators involving $\Lambda$. Under Assumption 2.1 for $\kappa = k+1$ and with Lemmas 2.3, 2.6, and 5.3 and the bounds (5.7a) and (5.7b), the properties (5.13) are satisfied. For the resolvent $\mathscr{R}^n$ , we directly employ Lemma A.2 to compute with (5.2a) and (5.8b)

$$\begin{aligned}\big\|\big(J_h\mathscr{R}-\mathscr{R}_hJ_h\big)y\big\|_{X_h} &\leqslant \tau\big\|\mathscr{R}_h\big(J_h\mathrm{A}-\mathrm{A}_hJ_h\big)\mathscr{R}y\big\|_{X_h}+\big\|\mathscr{R}_h\big(J_h\Lambda-\Lambda_hJ_h\big)\mathscr{R}y\big\|_{X_h}\\ &\leqslant Ch^k\tau\big\|\mathrm{A}^{k^*+1}\mathscr{R}y\big\|_{X_h}+Ch^k\big\|\mathrm{A}^{k^*}\mathscr{R}y\big\|_{X_h}\\ &\leqslant Ch^k\big\|\mathrm{A}^{k^*}y\big\|_X,\end{aligned}$$

and hence Lemma 5.9 is applicable.

(ii) We employ Lemma A.1 and denote any product of operators from part (i) by $\Pi$ and the discrete counterpart $\Pi_h$. Then, we have to compare expressions of the form

$$J_h\Pi(x+\delta)-\Pi_hx_h=(J_h\Pi-\Pi_hJ_h)x+\Pi_h(J_hx-x_h)+J_h\Pi\delta$$

with

$$x\in\{\mathrm{A}y^0,\mathrm{A}^2y^0,F^1,\partial_\tau F^2,\mathrm{A}F^1\},\qquad x_h\in\{\mathrm{A}_hJ_hy^0,\mathrm{A}_h^2J_hy^0,F_h^1,\partial_\tau F_h^2,\mathrm{A}_hF_h^1\},\tag{5.15}$$

and $\delta\in\{\delta^1,\partial_\tau\delta^2,\mathrm{A}\delta^1\}$ with $\delta^\ell$ given in (5.11). Then, the first part is covered by part (i), provided that $\|\mathrm{A}^{k^*}x\|_X\leqslant C$ holds for all $x$ in (5.15), which follows from the assumptions of the theorem. The second part is bounded due to (5.7c), (5.8a), (5.8b), and (5.8c). Lastly, the estimate

$$\|J_h\Pi\delta\|_{X_h}\lesssim\|\mathrm{A}\delta\|_X$$

together with (5.5) yields

$$\|J_h\Pi\delta\|_{X_h}\lesssim\tau\big(\big\|\mathrm{A}^2\partial_t^2y\big\|_{L^\infty(X)}+\big\|\mathrm{A}\partial_t^3y\big\|_{L^\infty(X)}\big),$$

and hence the desired bound. □

REMARK 5.10 The restriction $\ell_{\max}\geqslant k^*$ in (5.10) is necessary in order to guarantee (2.5) for $\ell=k^*$ which is needed in the estimates above to transport the regularity past $\Lambda$.

We now prove Lemma 5.6 and observe that formally setting $\tau=0$ in (2.24), we obtain the spatially discretized equation (2.20). This observation drives the following proof.

*Proof of* Lemma 5.6. Several steps in the proof simplify, and we mainly explain why we can neglect the assumption on $\ell_{\max}$. First note, that compared to the proof above we replace due to $\tau=0$

$$\mathscr{R}^n\to(\Lambda^0)^{-1},\qquad\qquad \mathscr{R}_h^n\to(\Lambda_h^0)^{-1},$$

and the most delicate term is given by

$$\begin{aligned}\big(\partial_t^2e_h(0)\big)_1&=\big(J_h(\Lambda^0)^{-1}\mathrm{A}(\Lambda^0)^{-1}\mathrm{A}-(\Lambda_h^0)^{-1}\mathrm{A}_h(\Lambda_h^0)^{-1}\mathrm{A}_h\big)J_hy^0\\ &=\Big(J_h(\Lambda^0)^{-1}(\widetilde{\Lambda}^0)^{-1}\mathrm{A}^2-(\Lambda_h^0)^{-1}(\widetilde{\Lambda}_h^0)^{-1}\mathrm{A}_h^2J_h\Big)y^0,\end{aligned}$$

Further, the terms in (5.14) reduce to $B^j\in\{\Lambda^0,\partial_t\Lambda^0,(\widetilde{\Lambda}^0)^{-1}\}$. In contrast to above, we may employ Lemma 5.9 with $\mathscr{Y}=H^{k^*+1}(\Omega)\times H^{k^*}(\Omega)$, and thus obtain no restriction on $\ell_{\max}$. Then, we proceed along the lines of Lemma 5.7. □

### 5.3 *Bounds on the first approximations*

This section is devoted to the proof of Theorem 2.10, i.e., we provide bounds for the first two approximations of the fully discrete scheme (2.24), which are not covered by Theorem 2.9.

*Proof of* Theorem 2.10. (i) For $\ell = 0$, the assertion directly follow from Theorem 2.8.

(i) For $\ell = 1$, as before, we decompose the error using (3.1) and estimate by (2.7), (2.8) and (2.11)

$$\begin{aligned}\big\|y(t^1) - \boldsymbol{\mathscr{L}_h} y_h^1\big\|_{W^{1,\infty}\times L^\infty} &\leqslant \big\|\big(\mathrm{Id} - \boldsymbol{\mathscr{L}_h} J_h\big) y(t^1)\big\|_{W^{1,\infty}\times L^\infty} + C\big\|J_h y(t^1) - y_h^1\big\|_{W^{1,\infty}\times L^\infty}\\ &\leqslant Ch^k + \big\|J_h y(t^\ell) - y_h^\ell\big\|_{W^{1,\infty}\times L^\infty}.\end{aligned}$$

Hence, we only have to prove by Lemma 2.7 the bound

$$\big\|\mathrm{A}_h^2\big(J_h y(t^1) - y_h^1\big)\big\|_{X_h} \leqslant C\big(\tau + h^k\big). \tag{5.16}$$

We expand

$$\mathrm{A}_h^2\big(J_h y(t^1) - y_h^1\big) = \mathrm{A}_h^2\big(J_h y(t^0) - y_h^0\big) + \tau\big(\mathrm{A}_h^2 J_h \partial_\tau y(t^1) - \mathrm{A}_h^2 \partial_\tau y_h^1\big)$$

and note that the first summand vanishes by (2.21), and it holds with (2.10) and (2.12)

$$\big\|\mathrm{A}_h^2 J_h \partial_\tau y(t^1)\big\|_{X_h} = \left\|\begin{pmatrix}\mathscr{L}_h^{H*}\Delta u^0\\ \mathscr{L}_h^{H*}\Delta v^0\end{pmatrix}\right\|_{X_h} \lesssim \|\Delta u^0\|_{H^1} + \|\Delta v^0\|_{L^2}.$$

For the last term, Lemma 5.8 gives $\partial_\tau y_h^1 = \mathscr{R}_h^1 \mathrm{A}_h J_h y^0 + \mathscr{R}_h^1 \Lambda_h^1 F_h^1$. We treat the terms separately, using

$$\mathrm{A}_h^2 \mathscr{R}_h^1 \mathrm{A}_h J_h y^0 = \widetilde{\Lambda}_h^1 \Lambda_h^1 \mathscr{R}_h^1 (\widetilde{\Lambda}_h^1)^{-1} \mathrm{A}_h (\widetilde{\Lambda}_h^1)^{-1} \mathrm{A}_h^2 J_h y^0,$$

exploiting (5.12), we are by Lemma 2.6 left to bound

$$\big\|\mathrm{A}_h (\widetilde{\Lambda}_h^1)^{-1} \mathrm{A}_h^2 J_h y^0\big\|_{X_h} \leqslant \big\|\mathscr{L}_h^{H*}\Delta v^0\big\|_{V_h} + \big\|\Delta_h \lambda_h{}^{-1}(\mathscr{L}_h^{H*}\Delta u^0)\big\|_{L^2}.$$

The first term is bounded by (2.12) for $v^0 \in \mathscr{D}(\Delta^2)$. We first note that for any $\varphi \in \mathscr{D}(\Delta)$, we have

$$\|\Delta_h I_h \varphi\|_{L^2} \leqslant \big\|\Delta_h \mathscr{L}_h^{V*}\varphi\big\|_{L^2} + \big\|\Delta_h (I_h - \mathscr{L}_h^{V*})\varphi\big\|_{L^2} \lesssim \big\|\mathscr{L}_h^{H*}\Delta\varphi\big\|_{L^2} + h^{-1}\big\|(I_h - \mathscr{L}_h^{V*})\varphi\big\|_{V_h} \lesssim \|\Delta\varphi\|_{L^2},$$

where we used the inverse estimate (2.17) together with (2.11) and (2.8). Along the same lines we show $\big\|\Delta_h \mathscr{L}_h^{H*}\varphi\big\|_{L^2} \lesssim \|\Delta\varphi\|_{L^2}$. This implies with Lemma 5.3

$$\begin{aligned}\big\|\Delta_h \lambda_h{}^{-1}(\mathscr{L}_h^{H*}\varphi)\big\|_{L^2} &\leqslant \big\|\Delta_h \mathscr{L}_h^{H*}(\lambda^{-1}\varphi)\big\|_{L^2} + \big\|\Delta_h\big(\mathscr{L}_h^{H*} - Q_h \mathscr{L}_h^{-1}\big)(\lambda^{-1}\varphi)\big\|_{L^2}\\ &\qquad + \Big\|\Delta_h Q_h\Big(\mathscr{L}_h^{-1}(\lambda^{-1}\varphi) - \big((I_h\lambda)^{-1}\mathscr{L}_h^{H*}\varphi\big)\Big)\Big\|_{L^2}\\ &\lesssim \big\|\Delta(\lambda^{-1}\varphi)\big\|_{L^2},\end{aligned}$$

and thus the second term is bounded if $u^0 \in \mathscr{D}(\Delta^2) \cap H^4(\Omega)$. Similarly, we have

$$\mathrm{A}_h^2 \mathscr{R}_h^1 \Lambda_h^1 F_h^1 = \widetilde{\Lambda}_h^1 \Lambda_h^1 \mathscr{R}_h^1 (\widetilde{\Lambda}_h^1)^{-1} \mathrm{A}_h^2 F_h^1$$

and taking norms gives

$$\big\|\mathrm{A}_h^2 F_h\big\|_{X_h} = \|\Delta_h I_h f\|_{L^2} \lesssim \|\Delta f\|_{L^2},$$

and with this we established (5.16). □

## 6. Numerical experiments

To illustrate our theoretical findings, we present some numerical experiments. Let $\Omega = B_1(0) \subset \mathbb{R}^2$ be the two-dimensional unit sphere and consider equation (1.1) with data given by

$$u^0(\mathbf{x}) = \frac{1}{20}\sin(\pi r^2)^7, \qquad v^0(\mathbf{x}) = \frac{1}{20}\sin(\pi r^2)^7,$$
$$\lambda(t,\mathbf{x}) = 2 + (1-r^2)^2\big(\mathbf{x}_1 e^t + (1-r^2)\mathbf{x}_2\cos(t)\big),$$

where $r^2 = |\mathbf{x}|^2$. The right-hand side $f$ is chosen such that the exact solution is given by

$$u(t,\mathbf{x}) = \frac{1}{20}\mathrm{e}^t\sin(\pi r^2)^7.$$

A simple calculation shows, that the regularity assumptions of Theorems 2.8 and 2.9 are satisfied with $k = 3$ and $\ell_{\max} = 4$. The scaling by a factor 20 is used to approximately normalize the $W^{1,\infty}$-norm of the solution $u$.

### 6.1 *Discretization*

To discretize in space, we multiply (1.1) with $\lambda(t,\mathbf{x})$ and obtain the equivalent formulation

$$\lambda(t,\mathbf{x})\partial_{tt}u(t,\mathbf{x}) = \Delta u(t,\mathbf{x}) + \lambda(t,\mathbf{x})f(t,\mathbf{x}).$$

Using the mass and stiffness matrix defined by

$$\big(M_h(t)\big)_{i,j} := ((I_h\lambda(t))\varphi_i \mid \varphi_j)_{L^2(\Omega_h)}, \qquad \big(L_h\big)_{i,j} := (\nabla\varphi_i \mid \nabla\varphi_j)_{L^2(\Omega_h)},$$

where we denote by $(\varphi_i)_i$ the nodal basis of $V_h$, the discrete solution in (2.20) satisfies

$$M_h(t)\partial_{tt}u_h(t) = -L_h u_h(t) + M_h(t)I_h f(t),$$

by abusing the notation for the coefficient vectors and their corresponding function in $V_h$. The fully discrete method in (2.24) is then given for $n \geqslant 2$ by

$$M_h(t^n)\big(u_h^n - 2u_h^{n-1} + u_h^{n-2}\big) = -\tau^2 L_h u_h^n + \tau^2 M_h(t^n)I_h f(t^n),$$

and for $n = 1$, we have to solve

$$M_h(t^1)\big(u_h^1 - u_h^0 - \tau v_h^0\big) = -\tau^2 L_h u_h^1 + \tau^2 M_h(t^1)I_h f(t^1).$$

Further, we have $v_h^n = \partial_\tau u_h^n$ for $n \geqslant 1$. We implemented the numerical experiments in `C++` using the finite element library `deal.II` (version 9.4) Arndt *et al.* (2022); Bangerth *et al.* (2007). A precise description of the implementation can for example be found in (Leibold, 2021, Ch. 6.5.1). The codes written by Malik Scheifinger to reproduce the experiments are available at https://doi.org/10.5445/IR/1000157919.

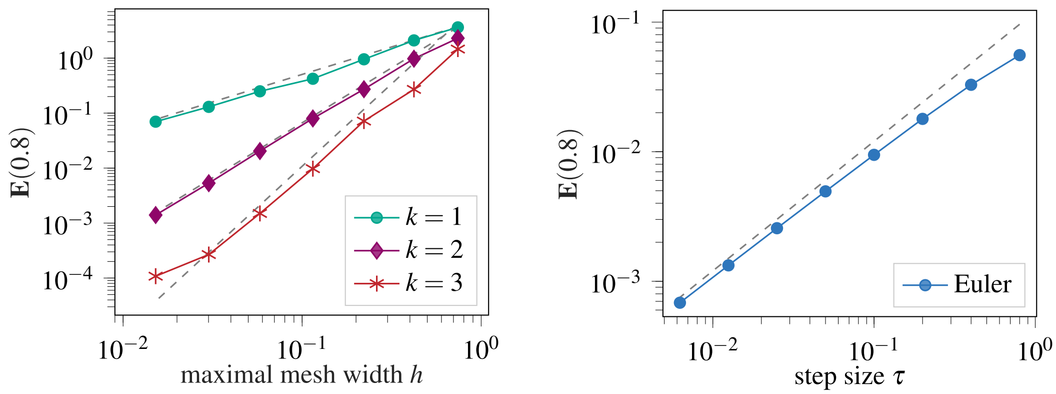



FIG. 3: Left: Error $\mathbf{E}(0.8)$ of the implicit Euler method (with time step size $\tau_{\mathrm{ref}} = 8 \cdot 10^{-5}$) combined with finite elements of order $k = 1,2,3$ plotted against the mesh width $h$. The dashed lines indicate order $h^k$ for $k = 1,2,3$. Right: Error $\mathbf{E}(0.8)$ of the implicit Euler method combined with finite elements of order $k = 3$ ($h_{\mathrm{ref}} = 1.52 \cdot 10^{-2}$) plotted against the time step size $\tau$. The dashed line indicates order 1.

### 6.2 *Numerical results*

For the problem described above, we performed experiments for the time and space discretization. To this end, we used finite elements of order $k = 1,2,3$. Since the computation of the lift of a finite element function is very laborious, we do not compute the full error in the form $\mathscr{L}_h u - u_h$. Instead, in our numerical examples we consider the error

$$\mathbf{E}(t) := \|u_h(t) - I_h u(t)\|_{W^{1,\infty}(\Omega_h)} + \|v_h(t) - I_h \partial_t u(t)\|_{L^\infty(\Omega_h)},$$

which is of the same order by the standard interpolation estimates. In the left part of Figure 3, the convergence of the error with respect to the spatial mesh width $h$ is shown when using the implicit Euler method with $\tau_{\mathrm{ref}} = 8 \cdot 10^{-5}$. We observe that for finite elements of order $k$ the error converges with order $k$ in space as predicted by Theorem 2.8 until the error for $k = 3$ is dominated by the error of the temporal approximation. In the right part of Figure 3, we consider the convergence of the error with respect to the time step size $\tau$. In space, we discretized with finite elements of order 3 and $h_{\mathrm{ref}} = 1.52 \cdot 10^{-2}$ such that the spatial error is negligible. Aligning to Theorem 2.9, we observe convergence of order 1 in time.

**Acknowledgment**

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173. The authors thank Roland Schnaubelt and Konstantin Zerulla for helpful discussions on elliptic regularity and in particular Konstantin Zerulla and Rebekka Zimmermann for their careful reading of this manuscript. Further, we like to thank Malik Scheifinger for providing us the codes for the numerical experiments.

## REFERENCES

ANTONIETTI, P. F., MAZZIERI, I., MUHR, M., NIKOLIĆ, V. & WOHLMUTH, B. (2020) A high-order discontinuous Galerkin method for nonlinear sound waves. *J. Comput. Phys.*, **415**, 109484, 27.

ARENDT, W., BATTY, C. J. K., HIEBER, M. & NEUBRANDER, F. (2011) *Vector-valued Laplace transforms and Cauchy problems*. Monographs in Mathematics, vol. 96, second edn. Birkhäuser/Springer Basel AG, Basel, pp. xii+539.

ARNDT, D., BANGERTH, W., FEDER, M., FEHLING, M., GASSMÖLLER, R., HEISTER, T., HELTAI, L., KRONBICHLER, M., MAIER, M., MUNCH, P., PELTERET, J.-P., STICKO, S., TURCKSIN, B. & WELLS, D. (2022) The `deal.II` library, version 9.4. *J. Numer. Math.*, **30**, 231–246.

BAKAEV, N. Y., THOMÉE, V. & WAHLBIN, L. B. (2003) Maximum-norm estimates for resolvents of elliptic finite element operators. *Math. Comp.*, **72**, 1597–1610.

BAKER, G. A., BRAMBLE, H. H. & THOMÉE, V. (1977) Single step Galerkin approximations for parabolic problems. *Math. Comp.*, **31**, 818–847.

BAKER, G. A., DOUGALIS, V. A. & SERBIN, S. M. (1979) High order accurate two-step approximations for hyperbolic equations. *RAIRO Anal. Numér.*, **13**, 201–226.

BAKER, G. A. & DOUGALIS, V. A. (1980) On the $L^\infty$-convergence of Galerkin approximations for second-order hyperbolic equations. *Math. Comp.*, **34**, 401–424.

BALES, L. A., DOUGALIS, V. A. & SERBIN, S. M. (1985) Cosine methods for second-order hyperbolic equations with time-dependent coefficients. *Math. Comp.*, **45**, 65–89.

BALES, L. A. & DOUGALIS, V. A. (1989) Cosine methods for nonlinear second-order hyperbolic equations. *Math. Comp.*, **52**, 299–319, S15–S33.

BANGERTH, W., HARTMANN, R. & KANSCHAT, G. (2007) deal.II—a general-purpose object-oriented finite element library. *ACM Trans. Math. Software*, **33**, Art. 24, 27.

BERNARDI, C. (1989) Optimal finite-element interpolation on curved domains. *SIAM J. Numer. Anal.*, **26**, 1212–1240.

BRAMBLE, J. H., SCHATZ, A. H., THOMÉE, V. & WAHLBIN, L. B. (1977) Some convergence estimates for semidiscrete Galerkin type approximations for parabolic equations. *SIAM J. Numer. Anal.*, **14**, 218–241.

BRENNER, S. C. & SCOTT, L. R. (2008) *The mathematical theory of finite element methods*. Texts in Applied Mathematics, vol. 15, third edn. Springer, New York, pp. xviii+397.

CHATZIPANTELIDIS, P., LAZAROV, R. D., THOMÉE, V. & WAHLBIN, L. B. (2006) Parabolic finite element equations in nonconvex polygonal domains. *BIT*, **46**, S113–S143.

DEMLOW, A. (2009) Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces. *SIAM J. Numer. Anal.*, **47**, 805–827.

DÖRICH, B. (2022) Strong norm error bounds for quasilinear wave equations under weak CFL-type conditions. *CRC 1173 Preprint* 2022/52. Karlsruhe Institute of Technology. `https://www.waves.kit.edu/downloads/CRC1173_Preprint_2022-52.pdf`.

DÖRICH, B., LEIBOLD, J. & MAIER, B. (2023) Optimal $W^{1,\infty}$-estimates for an isoparametric finite element discretization of elliptic boundary value problems. *Electron. Trans. Numer. Anal.*, **58**, 1–21.

ELLIOTT, C. M. & RANNER, T. (2013) Finite element analysis for a coupled bulk-surface partial differential equation. *IMA J. Numer. Anal.*, **33**, 377–402.

ELLIOTT, C. M. & RANNER, T. (2021) A unified theory for continuous-in-time evolving finite element space approximations to partial differential equations in evolving domains. *IMA J. Numer. Anal.*, **41**, 1696–1845.

GRISVARD, P. (1985) *Elliptic problems in nonsmooth domains*. Monographs and Studies in Mathematics, vol. 24. Pitman (Advanced Publishing Program), Boston, MA, pp. xiv+410.

HIPP, D., HOCHBRUCK, M. & STOHRER, C. (2019) Unified error analysis for nonconforming space discretizations of wave-type equations. *IMA J. Numer. Anal.*, **39**, 1206–1245.

HOCHBRUCK, M. & MAIER, B. (2021) Error analysis for space discretizations of quasilinear wave-type equations. *IMA J. Numer. Anal.*

HOCHBRUCK, M. & PAŽUR, T. (2017) Error analysis of implicit Euler methods for quasilinear hyperbolic evolution equations. *Numer. Math.*, **135**, 547–569.

KASHIWABARA, T. & KEMMOCHI, T. (2020) Pointwise error estimates of linear finite element method for Neumann boundary value problems in a smooth domain. *Numer. Math.*, **144**, 553–584.

KOVÁCS, B. & POWER GUERRA, C. A. (2018) Maximum norm stability and error estimates for the evolving surface finite element method. *Numer. Methods Partial Differential Equations*, **34**, 518–554.

LEIBOLD, J. (2021) A unified error analysis for the numerical solution of nonlinear wave-type equations with application to kinetic boundary conditions. *Ph.D. thesis*, Karlsruher Institut für Technologie (KIT). https://doi.org/10.5445/IR/1000130222.

LI, B. (2019) Analyticity, maximal regularity and maximum-norm stability of semi-discrete finite element solutions of parabolic equations in nonconvex polyhedra. *Math. Comp.*, **88**, 1–44.

LIAO, H. & SUN, Z. (2011) Maximum norm error estimates of efficient difference schemes for second-order wave equations. *J. Comput. Appl. Math.*, **235**, 2217–2233.

LINZ, P. (1969) Numerical methods for Volterra integral equations of the first kind. *Comput. J.*, **12**, 393–397.

LITTMAN, W. (1963) The wave operator and $L_p$ norms. *J. Math. Mech.*, **12**, 55–68.

MAIER, B. (2020) Error analysis for space and time discretizations of quasilinear wave-type equations. *Ph.D. thesis*, Karlsruhe Institute of Technology. https://doi.org/10.5445/IR/1000120935.

MAIER, B. (2022) Error analysis for full discretizations of quasilinear wave-type equations with two variants of the implicit midpoint rule. *IMA J. Numer. Anal.*

MAKRIDAKIS, C. G. (1993) Finite element approximations of nonlinear elastic waves. *Math. Comp.*, **61**, 569–594.

NATTERER, F. (1975) Über die punktweise Konvergenz finiter Elemente. *Numer. Math.*, **25**, 67–77.

NITSCHE, J. A. (1975) $L_\infty$-convergence of finite element approximation. *Journées "Éléments Finis" (Rennes, 1975)*. Univ. Rennes, Rennes, p. 18.

NITSCHE, J. A. (1977) $L_\infty$-convergence of finite element approximations. *Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*. *Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*., pp. 261–274. Lecture Notes in Math., Vol. 606.

PALENCIA, C. (1996) Maximum norm analysis of completely discrete finite element methods for parabolic problems. *SIAM J. Numer. Anal.*, **33**, 1654–1668.

RANNACHER, R. (1976) Zur $L^\infty$-Konvergenz linearer finiter Elemente beim Dirichlet-Problem. *Math. Z.*, **149**, 69–77.

RANNACHER, R. & SCOTT, R. (1982) Some optimal error estimates for piecewise linear finite element approximations. *Math. Comp.*, **38**, 437–445.

SCHATZ, A. H., THOMÉE, V. C. & WAHLBIN, L. B. (1980) Maximum norm stability and error estimates in parabolic finite element equations. *Comm. Pure Appl. Math.*, **33**, 265–304.

SCHATZ, A. H., THOMÉE, V. & WAHLBIN, L. B. (1998) Stability, analyticity, and almost best approximation in maximum norm for parabolic finite element equations. *Comm. Pure Appl. Math.*, **51**, 1349–1385.

SCHATZ, A. H. & WAHLBIN, L. B. (1977) Interior maximum norm estimates for finite element methods. *Math. Comp.*, **31**, 414–442.

SCHATZ, A. H. & WAHLBIN, L. B. (1982) On the quasi-optimality in $L_\infty$ of the $\dot{H}^1$-projection into finite element spaces. *Math. Comp.*, **38**, 1–22.

SCHATZ, A. H. & WAHLBIN, L. B. (1995) Interior maximum-norm estimates for finite element methods. II. *Math. Comp.*, **64**, 907–928.

SCOTT, R. (1976) Optimal $L^\infty$ estimates for the finite element method on irregular meshes. *Math. Comp.*, **30**, 681–697.

THOMÉE, V. & WAHLBIN, L. B. (2004) Maximum-norm estimates for finite-element methods for a strongly damped wave equation. *BIT*, **44**, 165–179.

TRAUTMANN, P., VEXLER, B. & ZLOTNIK, A. (2018) Finite element error analysis for measure-valued optimal control problems governed by a 1D wave equation with variable coefficients. *Math. Control Relat. Fields*, **8**, 411–449.

WAHLBIN, L. B. (1978) Maximum norm error estimates in the finite element method with isoparametric quadratic elements and numerical integration. *RAIRO Anal. Numér.*, **12**, 173–202, v.

ZLÁMAL, M. (1968) On the finite element method. *Numer. Math.*, **12**, 394–409.

## A. Appendix

In this appendix, we provide the proof of Lemma 2.7, properties (2.12) and (2.13), and the postponed calculations from Section 5. The following proof is adapted from the conforming case presented in (Bramble *et al.*, 1977, Lem. 4.1), and Dörich (2022).

*Proof of* Lemma 2.7. We first recall the solution operators $\Delta^{-1}$ and $\Delta_h^{-1}$ defined in (2.6) and (2.19), respectively. We further define the modified solution operator $\widetilde{S}_h = \mathscr{L}_h^{V*}\Delta^{-1}\mathscr{L}_h$, and use it to expand $\Delta_h^{-1} = \widetilde{S}_h + (\Delta_h^{-1} - \widetilde{S}_h)$. For the first term, we use Sobolev's embedding, the stability of the Ritz map in $W^{1,p}$, $2 \leqslant p \leqslant \infty$, (which is interpolated from (2.11) and (2.15)), as well as Theorem 2.4 to obtain

$$\|\widetilde{S}_h\varphi_h\|_{L^\infty(\Omega_h)} \lesssim \|\widetilde{S}_h\varphi_h\|_{W^{1,4}(\Omega_h)} \lesssim \|\Delta^{-1}\mathscr{L}_h\varphi_h\|_{W^{1,4}(\Omega)} \lesssim \|\Delta^{-1}\mathscr{L}_h\varphi_h\|_{H^2(\Omega)} \lesssim \|\varphi_h\|_{L^2(\Omega_h)},$$

and similarly using the same steps

$$\|\widetilde{S}_h\varphi_h\|_{W^{1,\infty}(\Omega_h)} \lesssim \|\Delta^{-1}\mathscr{L}_h\varphi_h\|_{W^{1,\infty}(\Omega)} \lesssim \|\Delta^{-1}\mathscr{L}_h\varphi_h\|_{W^{2,4}(\Omega)} \lesssim \|\mathscr{L}_h\varphi_h\|_{L^4(\Omega)} \lesssim \|\varphi_h\|_{L^4(\Omega_h)}.$$

It remains to bound the difference, stemming from the nonconformity, by the inverse estimate (2.17) via

$$\begin{aligned}
\|\widetilde{S}_h\varphi_h - \Delta_h^{-1}\varphi_h\|_{W^{1,\infty}(\Omega_h)} &\leqslant Ch^{-N/2}\|\widetilde{S}_h\varphi_h - \Delta_h^{-1}\varphi_h\|_{V_h} \\
&= Ch^{-N/2} \sup_{\|\psi_h\|_{V_h}=1} \left(\widetilde{S}_h\varphi_h - \Delta_h^{-1}\varphi_h \mid \psi_h\right)_{V_h} \\
&= Ch^{-N/2} \sup_{\|\psi_h\|_{V_h}=1} \big((\varphi_h \mid \psi_h)_{L^2(\Omega_h)} - (\mathscr{L}_h\varphi_h \mid \mathscr{L}_h\psi_h)_{L^2(\Omega)}\big),
\end{aligned}$$

where we used (2.6), (2.9), and (2.19) in the last step. Finally, we employ (Elliott & Ranner, 2021, Lem. 8.24) to obtain

$$\left|(\varphi_h \mid \psi_h)_{L^2(\Omega_h)} - (\mathscr{L}_h\varphi_h \mid \mathscr{L}_h\psi_h)_{L^2(\Omega)}\right| \lesssim h^{3/2}\|\varphi_h\|_{L^2}\|\psi_h\|_{V_h},$$

which yields the assertion. □

Next, we show the stability and convergence properties for $\mathscr{L}_h^{H*}$ and $\pi_h$.

*Proof of* (2.12) *and* (2.13). Let $\varphi_h \in V_h$, and estimate with (Elliott & Ranner, 2021, Lem. 8.24)

$$\|\mathscr{L}_h^{H*}\varphi - \varphi_h\|_{L^2} \lesssim h^k\|\varphi_h\|_{L^2} + \|\varphi - \mathscr{L}_h\varphi_h\|_{L^2}$$

and thus obtain with the inverse estimate (2.17)

$$\|\mathscr{L}_h^{H*}\varphi\|_{V_h} \lesssim \|\varphi_h\|_{V_h} + h^{-1}\|\mathscr{L}_h^{H*}\varphi - \varphi_h\|_{L^2} \lesssim \|\varphi_h\|_{V_h} + h^{-1}\|\varphi - \mathscr{L}_h\varphi_h\|_{L^2}.$$

We now use the quasi-interpolation from (Bernardi, 1989, Thm 5.1), which is stable in $H^1$ and converges linearly in $L^2$, and hence gives (2.12).

Employing again (Elliott & Ranner, 2021, Lem. 8.24), for $\varphi \in H^2(\Omega)$ we compute

$$\begin{aligned} \left| \left(\pi_h \mathscr{L}_h^{-1}\varphi - \mathscr{L}_h^{H*}\varphi \mid \psi_h\right)_{L^2(\Omega_h)} \right| &= \left| \left(\mathscr{L}_h^{-1}\varphi \mid \psi_h\right)_{L^2(\Omega_h)} - (\varphi \mid \mathscr{L}_h\psi_h)_{L^2(\Omega)} \right| \\ &\leqslant Ch^k \|\varphi\|_{L^2(U_h)} \|\mathscr{L}_h\psi_h\|_{L^2(\Omega)} \\ &\leqslant Ch^{k+1} \|\varphi\|_{H^2(\Omega)} \|\mathscr{L}_h\psi_h\|_{L^2(\Omega)}, \end{aligned}$$

where we used (Dörich, 2022, Appendix B) in the last step, and (2.13) is shown. □

Further, we verify the properties in Lemma 5.3 of the (weighted) $L^2$-projection $Q_h$ and the formula for the inverse of $\lambda_h$.

*Proof of* Lemma 5.3. (a) By the definition of $Q_h$, we have

$$\left(\mathscr{L}_h^{H*}v - Q_h\mathscr{L}_h^{-1}v \mid \mathscr{L}_h^{H*}v - Q_h\mathscr{L}_h^{-1}v\right)_{\lambda_h} = \left(\mathscr{L}_h^{H*}v - Q_h\mathscr{L}_h^{-1}v \mid \mathscr{L}_h^{H*}v - \mathscr{L}_h^{-1}v\right)_{\lambda_h},$$

and thus the equivalence of norms together with (2.8) and (2.10) gives the first claim. The stability in $H^1$ is shown as for (2.12).

(b) One easily verifies the identities for $\varphi_h, \psi_h \in V_h$

$$\left(\lambda_h({\lambda_h}^{-1}\varphi_h) \mid \psi_h\right)_{L^2(\Omega_h)} = (\varphi_h \mid \psi_h)_{L^2(\Omega_h)}, \quad \left({\lambda_h}^{-1}(\lambda_h\varphi_h) \mid \psi_h\right)_{\lambda_h} = (\varphi_h \mid \psi_h)_{\lambda_h},$$

which gives the last statement. □

Next, we give an extension of Lemma 5.8. We do not provide a proof here, since the expressions are derived by an iterative application of the identities (5.12).

LEMMA A.1 It holds the representation

$$\begin{aligned} \partial_\tau y_h^1 &= \mathscr{R}_h^1 \mathrm{A}_h J_h y^0 + \mathscr{R}_h^1 \Lambda_h(t^1) F_h^1, \\ \partial_\tau^2 y_h^2 &= \mathscr{R}_h^2 \Lambda_h^1 \mathscr{R}_h^1 (\widetilde{\Lambda}_h^1)^{-1} \mathrm{A}_h^2 J_h y^0 - \mathscr{R}_h^2 \partial_\tau \Lambda_h^2 \mathscr{R}_h^1 \mathrm{A}_h J_h y^0 + \mathscr{R}_h^2 \Lambda_h^1 \mathscr{R}_h^1 \mathrm{A}_h F_h^1 \\ &\quad + \mathscr{R}_h^2 \Lambda_h^2 \partial_\tau F_h^2 + \mathscr{R}_h^2 \partial_\tau \Lambda_h^2 F_h^1 - \mathscr{R}_h^2 \partial_\tau \Lambda_h^2 \mathscr{R}_h^1 \Lambda_h^1 F_h^1 . \end{aligned}$$

The same expansion holds for $\partial_\tau^\ell y(t^\ell)$, $\ell = 1,2$, with $h$ formally set to zero and

$$F_h^n \to F^n + \delta^n, \qquad \delta^n = \partial_\tau y(t^n) - \partial_t y(t^n).$$

Since, in the case $\Lambda \neq \mathrm{Id}$, the operator A does in general not commute with the resolvent, we provide the bounds which are still available under Assumption 2.1.

LEMMA A.2 Let $\mathscr{R}$ be the resolvent defined in (5.9) and $\ell_{\max}$ given in Assumption 2.1. Then, for $0 \leqslant \ell \leqslant \ell_{\max}$ there are constants $C_\ell$ such that

$$\begin{aligned} \left\|\mathrm{A}^\ell \mathscr{R} y\right\|_X &\leqslant C_\ell \left\|\mathrm{A}^\ell y\right\|_X \\ \tau\left\|\mathrm{A}^{\ell+1} \mathscr{R} y\right\|_X &\leqslant C_\ell \left\|\mathrm{A}^\ell y\right\|_X. \end{aligned}$$

for all $y \in \mathscr{D}(\mathrm{A}^{\ell_{\max}})$.

*Proof.* By the skew-adjointness of A and (2.4), we derive

$$\|\mathscr{R}y\|_X \leqslant C\,\|y\|_X,$$

which is the first equation for $\ell = 0$. For $0 \leqslant \ell \leqslant \ell_{\max} - 1$, we compute

$$\begin{aligned}\left\|\mathrm{A}^{\ell+1}\mathscr{R}y\right\|_X &= \left\|\mathrm{A}^{\ell}\Lambda\mathscr{R}(\widetilde{\Lambda})^{-1}\mathrm{A}y\right\|_X \\ &\leqslant \left\|\Lambda\right\|_{\mathscr{D}(\mathrm{A}^\ell)\leftarrow\mathscr{D}(\mathrm{A}^\ell)}\left\|\mathscr{R}\right\|_{\mathscr{D}(\mathrm{A}^\ell)\leftarrow\mathscr{D}(\mathrm{A}^\ell)}\left\|(\widetilde{\Lambda})^{-1}\right\|_{\mathscr{D}(\mathrm{A}^\ell)\leftarrow\mathscr{D}(\mathrm{A}^\ell)}\left\|\mathrm{A}^{\ell+1}y\right\|_X\end{aligned}$$

and the claim follows by induction. The second estimate is due to the resolvent identity

$$\tau\mathrm{A}^{\ell+1}\mathscr{R}y = \mathrm{A}^{\ell}\Lambda\mathscr{R}y - \mathrm{A}^{\ell}y$$

and an application of the above bound. □