

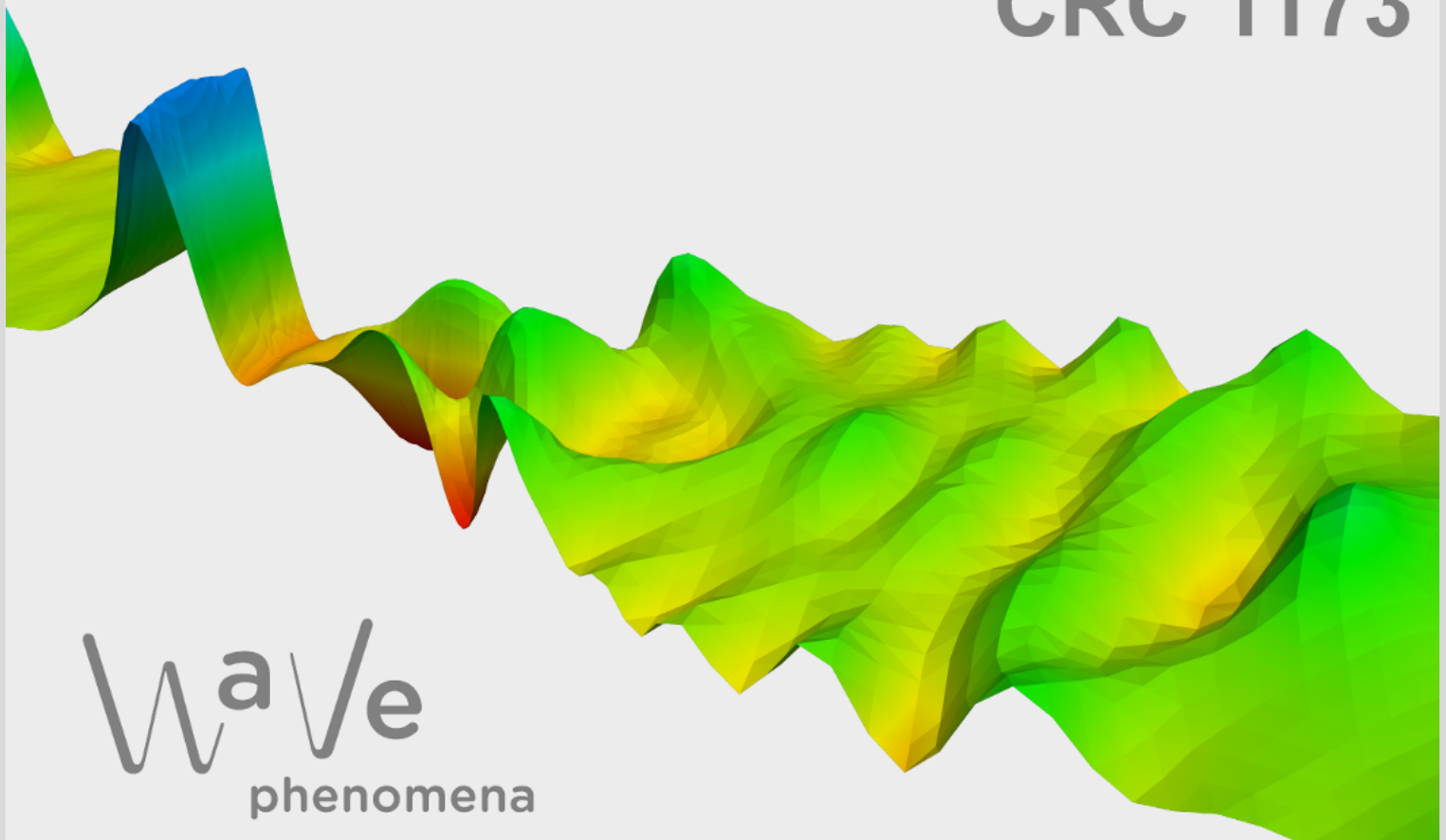
A finite element exponential integrator for rough solutions of semilinear wave equations. Part I: Dirichlet boundary conditions

Jiachuan Cao, Benjamin Dörich,
Marlis Hochbruck, Buyang Li

CRC Preprint 2026/29, July 2026

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave
phenomena

Participating universities



Funded by



A FINITE ELEMENT EXPONENTIAL INTEGRATOR FOR ROUGH SOLUTIONS OF SEMILINEAR WAVE EQUATIONS. PART I: DIRICHLET BOUNDARY CONDITIONS

JIACHUAN CAO, BENJAMIN DÖRICH, MARLIS HOCHBRUCK, AND BUYANG LI

ABSTRACT. We study a fully discrete scheme for semilinear wave equations on general bounded polygonal/polyhedral domains, combining an exponential Euler time integrator with a finite element spatial discretization, with initial data $(u^0, v^0) \in H^\gamma(\Omega) \times H^{\gamma-1}(\Omega)$ for $0 < \gamma \leq 1$. In contrast to existing low-regularity error analyses, which are mostly based on Fourier spectral discretizations, our approach applies to general bounded domains and finite element spatial discretizations. We prove rigorous error estimates for low-regularity solutions. The analysis is based on a frequency decomposition of the underlying elliptic operator, used solely as an analytical regularization device and not in the actual implementation, which allows low-regularity techniques to be extended beyond the Fourier spectral framework. The results also indicate that higher-order finite element methods remain advantageous in spatial approximation even for solutions of limited Sobolev regularity. Numerical experiments on different domains and with different polynomial degrees confirm the predicted convergence behavior.

1. Introduction

Semilinear wave equations of the form

$$\begin{cases} \partial_{tt}u - \Delta u = f(u), & \text{in } \Omega \times (0, T], \\ u = 0, & \text{on } \partial\Omega \times (0, T], \\ u|_{t=0} = u^0, \quad \partial_t u|_{t=0} = v^0, & \text{in } \Omega, \end{cases} \quad (1.1)$$

arise in a broad range of applications involving nonlinear wave propagation, including continuum mechanics, nonlinear optics, and materials science. Their accurate numerical approximation is therefore of considerable practical and theoretical interest.

For sufficiently smooth solutions, a wide variety of time discretization methods are available for semilinear wave equations, including trigonometric and exponential integrators, splitting methods, symplectic schemes, and classical finite difference or Runge–Kutta type methods; see, for example, [5, 13, 22, 23, 38]. Under appropriate regularity assumptions, these methods often achieve their expected convergence orders and exhibit good long-time behavior. However, in many situations of practical and mathematical interest, the solution possesses only limited Sobolev regularity. This may be caused by rough initial data, nonsmooth geometries, singular forcing, or boundary effects. In such regimes, standard error analyses are no longer applicable, and classical high-order schemes may suffer substantial order reduction or require assumptions incompatible with the actual solution regularity. Related regularity-dependent convergence results for time discretizations of nonsmooth problems have also been obtained in more abstract settings; see, for example, [8, 9, 31, 39].

In recent years, substantial progress has been made in low-regularity integrators for deterministic dispersive equations. Important advances have been reported for the Korteweg–de Vries

2010 *Mathematics Subject Classification.* 65M12, 65M15, 35Q55.

Key words and phrases. semilinear wave equation, low regularity, nontrivial boundary conditions, error estimates.

The first three authors acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 258734477 - SFB 1173. The work of B. Li was partially supported by the National Natural Science Foundation of China (project no. 12525111) and Hong Kong Research Grants Council (project no. 15306123).

(KdV) equation [25, 29] and the nonlinear Schrödinger equation [19, 32, 33], and other important dispersive equations [26, 34, 37]. For semilinear wave equations, second-order convergence in the $H^1 \times L^2$ norm has been proved under additional regularity assumptions [34]. Exponential-type low-regularity integrators based on integration-by-parts formulations have also shown efficient time stepping for finite-energy solutions [11]. Later, Cao, Li, Lin, and Yao [12] developed high-frequency-recovered exponential integrators based on a frequency-decomposition technique, and established convergence rates of 2γ and 1.5γ in one and two spatial dimensions, respectively, for initial data $(u^0, v^0) \in H^\gamma \times H^{\gamma-1}$ with $0 < \gamma \leq 1$. More recently, Ruff and Schnaubelt [35, 36] established convergence results for Lie- and Strang-splitting methods for semilinear wave equations with finite-energy solutions in three spatial dimensions. However, these analyses rely strongly on Fourier techniques, spectral discretizations, or periodic boundary conditions.

This reliance substantially limits the range of geometries and boundary conditions that can be treated. In particular, Fourier-based approaches are not directly suited to general bounded domains, variable coefficients, or nontrivial boundary conditions. From the perspective of practical computation, finite element methods are the natural framework in such situations. They offer geometric flexibility, compatibility with variational formulations, and a well-developed approximation theory on nonsmooth domains. Despite these advantages, a rigorous low-regularity error analysis for fully discrete finite element approximations of semilinear wave equations remains largely open.

In this paper, we study the semilinear wave equation (1.1) on a bounded polygonal/polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with low-regularity initial data

$$(u^0, v^0) \in H^\gamma(\Omega) \times H^{\gamma-1}(\Omega), \quad 0 < \gamma \leq 1. \quad (1.2)$$

We develop and analyze a fully discrete finite element method with an exponential Euler time discretization for this low-regularity regime. Our objective is to transfer low-regularity techniques, previously developed mainly in Fourier-based settings, to a finite element framework on more general domains.

To overcome the challenges posed by the low regularity of the solution, we combine an exponential-type time discretization with a frequency-localized analytical framework adapted to finite element spaces. A central ingredient is a frequency projection operator that decomposes the exact solution into low- and high-frequency parts. This device plays the role of an analytical smoothing mechanism, analogous in spirit to ideas used in Fourier-based low-regularity analysis, but it is employed here only as a tool in the proof. The numerical method itself remains a standard finite element discretization coupled with exponential time stepping and does not require the computation of eigenfunctions or the implementation of spectral truncations.

The main contributions of this paper can be summarized as follows.

- We develop a new analytical framework for deriving error estimates in the weak norm $L^2(\Omega) \times H^{-1}(\Omega)$ for the fully discrete exponential Euler finite element scheme on bounded polygonal/polyhedral domains with initial data satisfying (1.2). In abstract form, the scheme reads

$$U_h^{n+1} = e^{\tau L_h} U_h^n + \tau \varphi_1(\tau L_h) \mathcal{P}_h F(U_h^n).$$

A key step is a comparison estimate between $e^{tL} \Pi_r$ and $e^{tL_h} \mathcal{P}_h \Pi_r$, which links the frequency-localized auxiliary problem to the finite element discretization. Together with low-regularity consistency and stability estimates, this comparison yields a fully discrete convergence result under the low-regularity assumption (1.2). To the best of our knowledge, this provides the first fully discrete finite element convergence analysis for semilinear wave equations at this level of regularity.

- The resulting weak-norm estimate exhibits a degree-dependent spatial convergence rate. More precisely, for finite element spaces of polynomial degree k , the fully discrete error satisfies

$$\sup_{0 \leq n \leq T/\tau} \|U(t_n) - U_h^n\|_{L^2(\Omega) \times H^{-1}(\Omega)} \lesssim \tau h^{\frac{\ell_k}{\ell_k+1}(\gamma-1)} + h^{\frac{\ell_k}{\ell_k+1}\gamma},$$

where

$$\ell_k = \begin{cases} k+1, & k=1, \\ k+2, & k \geq 2. \end{cases}$$

In particular, the spatial contribution is of order $h^{\frac{\ell_k}{\ell_k+1}\gamma}$. Thus the convergence exponent improves from $2\gamma/3$ for linear elements to $4\gamma/5$ and $5\gamma/6$ for quadratic and cubic elements, respectively. This improvement is obtained under the same low-regularity assumption (1.2), without requiring any additional Sobolev regularity as the polynomial degree is increased. Thus polynomial enrichment improves the asymptotic spatial convergence rate itself, rather than merely the error constant, even for rough solutions of semilinear wave equations.

- Numerical experiments on different domains and for polynomial degrees $k = 1, 2, 3$ confirm the predicted degree-dependent rates. They also show a clear improvement in the error versus degrees-of-freedom comparison for higher-order finite elements in the low-regularity regime.

The remainder of this paper is organized as follows. In Section 2, we introduce the frequency-localized analytical framework for the abstract formulation of the semilinear wave equation and present the time semidiscrete problem, which serves as a bridge for the analysis of the fully discrete scheme. At the end of this section, we state the main convergence results. Section 3 collects several auxiliary results that support the proofs of the main theorems, including error estimates between the continuous and discrete solution operator semigroups. In Section 4, we combine these ingredients to establish the main theoretical results. Finally, Section 5 reports a series of numerical experiments for semilinear wave equations on different domains and under various regularity assumptions. These experiments verify the accuracy of the proposed method and confirm the theoretical findings.

Notations. For the sake of notational simplicity, we write $A \lesssim B$ (or equivalently $B \gtrsim A$) to denote that $A \leq CB$ for some constant $C > 0$. The constant C may vary from line to line, but it is always independent of the time step τ , the mesh size h , the frequency cut-off constant r and the bounds of the numerical solution. Similarly, we use $A \sim B$ to denote the equivalence $C^{-1}B \leq A \leq CB$ for some constant $C > 0$. In other words, $A \sim B$ is equivalent to $A \lesssim B \lesssim A$. We also use the notation $C(\cdot)$ to denote a positive constant whose value may change from line to line but depends only on the quantities specified in the parentheses.

2. Mathematical setting and the main result

In this section, we introduce the abstract functional framework used for the analysis and numerical approximation of semilinear wave equations on polygonal/polyhedral domains. To handle solutions with low regularity or strong singularities, we combine an exponential Euler integrator in time with a finite element discretization in space. Within this framework, we state the main convergence result under minimal regularity assumptions on the exact solution and introduce a frequency-localized auxiliary scheme that serves as a key analytical bridge in the error analysis.

2.1. Abstract formulation and functional framework

We begin with the weak formulation of a general semilinear wave equation, expressed as

$$\langle u_{tt}, v \rangle_H + a(u, v) = \langle f(u), v \rangle_H, \quad (2.1)$$

where $\langle \cdot, \cdot \rangle_H$ denotes a bilinear form defined on a Hilbert space H , and $a(\cdot, \cdot)$ is another bilinear operator on a Hilbert space V , with V densely embedded in H .

Remark 2.1. Throughout this work, we restrict our attention to problems with homogeneous Dirichlet boundary conditions. Within this setting, we have $H = L^2(\Omega)$, endowed with the

standard inner product

$$\langle u, v \rangle_H = \int_{\Omega} uv \, dx.$$

Moreover, $V = H_0^1(\Omega)$, the associated bilinear form is given by

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx.$$

Using the Riesz representation theorem for linear operators on Hilbert spaces, we reformulate the semilinear wave equation into an operator form involving an unbounded operator A :

$$A : D(A) \rightarrow H, \quad \langle Av, w \rangle_H = a(v, w) \quad \text{for all } v \in D(A), w \in V,$$

with

$$D(A) = \{v \in V \mid \exists C = C(v) > 0 \text{ such that } \forall w \in V : |a(v, w)| \leq C\|w\|_H\}.$$

The original semilinear wave equation is then reduced to the abstract form:

$$u_{tt} + Au = f(u).$$

We pose the following assumption on the operator A :

Assumption 2.2.

- (a) *The operator A is self-adjoint and strictly positive, satisfying the coercivity condition $\langle Av, v \rangle_H \geq C_A \langle v, v \rangle_H$, where $C_A > 0$ for all $v \in D(A)$.*
- (b) *A has a dense domain $D(A)$ in H and a purely point spectrum.*

We denote by $\{\lambda_{\alpha}, w_{\alpha}\}_{\alpha \in I}$ the eigenpairs of the operator $\Lambda = \sqrt{A}$, where I is a countable index set. The associated orthonormal eigenvectors $e_{\alpha} = w_{\alpha}/\|w_{\alpha}\|_H$ form a complete orthonormal basis of H .

We introduce a continuous scale of Hilbert spaces

$$\mathcal{Y}^{\ell} = D(\Lambda^{\ell}), \quad \ell \geq 0, \tag{2.2}$$

with $\mathcal{Y}^0 = H = L^2(\Omega)$. From the definition of A , it is straightforward to verify that $\mathcal{Y}^1 = H_0^1(\Omega)$. For negative indices ℓ , the spaces are defined by duality, $\mathcal{Y}^{\ell} = (\mathcal{Y}^{-\ell})^*$. For $\ell \geq 0$, the inner product on $\mathcal{Y}^{\ell}$ is given by

$$\langle u_1, u_2 \rangle_{\ell} = \langle u_1, u_2 \rangle_H + \langle \Lambda^{\ell} u_1, \Lambda^{\ell} u_2 \rangle_H. \tag{2.3}$$

Accordingly, each $u \in \mathcal{Y}^{\ell} \hookrightarrow H$ admits the expansion

$$u = \sum_{\alpha \in I} \widehat{u}_{\alpha} e_{\alpha}, \tag{2.4}$$

and its norm in $\mathcal{Y}^{\ell}$ satisfies

$$\|u\|_{\mathcal{Y}^{\ell}} = \sqrt{\langle u, u \rangle_{\ell}} \sim \left(\sum_{\alpha \in I} (1 + \lambda_{\alpha}^{2\ell}) |\widehat{u}_{\alpha}|^2 \right)^{\frac{1}{2}}. \tag{2.5}$$

By standard interpolation and duality arguments, one obtains

$$\mathcal{Y}^{\ell} = \begin{cases} H^{\ell}(\Omega), & 0 \leq \ell < \frac{1}{2}, \\ H_{00}^{1/2}(\Omega), & \ell = \frac{1}{2}, \\ H_0^{\ell}(\Omega), & \frac{1}{2} < \ell \leq 1, \end{cases} \tag{2.6}$$

with equivalence of norms. Consequently,

$$\mathcal{Y}^{\ell} \hookrightarrow H^{\ell}(\Omega) \quad \text{and} \quad H^{-\ell}(\Omega) \hookrightarrow \mathcal{Y}^{-\ell}, \quad 0 \leq \ell \leq 1. \tag{2.7}$$

2.2. Fully discrete scheme and main results

To derive a numerical approximation for the second-order semilinear wave equation, we first reformulate it as a first-order system:

$$\begin{cases} \partial_t U - LU = F(U) & \text{for } t \in (0, T], \\ U|_{t=0} = U^0, \end{cases} \quad (2.8)$$

where

$$U = \begin{pmatrix} u \\ v \end{pmatrix}, \quad U^0 = \begin{pmatrix} u^0 \\ v^0 \end{pmatrix}, \quad F(U) = \begin{pmatrix} 0 \\ f(u) \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 1 \\ -A & 0 \end{pmatrix}. \quad (2.9)$$

For each $U = (u, v)^\top$, we define

$$\|U\|_\ell := \|U\|_{\mathcal{Y}^\ell \times \mathcal{Y}^{\ell-1}} = (\|u\|_{\mathcal{Y}^\ell}^2 + \|v\|_{\mathcal{Y}^{\ell-1}}^2)^{\frac{1}{2}}.$$

As A is self-adjoint, the associated operator L is skew-adjoint, ensuring that all eigenvalues of L are purely imaginary. Consequently, the semigroup e^{tL} forms a C^0 -semigroup on H , satisfying the key stability property

$$\|e^{tL}\|_{\mathcal{L}(\mathcal{Y}^\ell \times \mathcal{Y}^{\ell-1})} \lesssim 1. \quad (2.10)$$

For the spatial discretization, let $(X_h^k)_h$ be a family of conforming Lagrange finite element spaces of degree $k \geq 1$, defined on a family of simplicial meshes $(\mathcal{T}_h)_h$, where $h > 0$ denotes the mesh size. We assume that the finite element spaces X_h^k satisfy the following approximation property for functions in $\mathcal{Y}^\ell$.

Assumption 2.3. *Let $\mathcal{T}_h$ be a simplicial partition of the domain Ω , and let X_h^k ($k \geq 1$) be a family of finite-dimensional subspaces satisfying $\dim(X_h^k) = O(h^{-d})$. Assume that for any $0 \leq \ell \leq k$ and any $w \in \mathcal{Y}^{\ell+1}$, the following approximation properties hold:*

$$\inf_{w_h \in X_h^k} \|w - w_h\|_{L^2(\Omega)} \leq Ch^{\ell+1} \|w\|_{\mathcal{Y}^{\ell+1}}, \quad (2.11)$$

$$\inf_{w_h \in X_h^k} \|w - w_h\|_{H^1(\Omega)} \leq Ch^\ell \|w\|_{\mathcal{Y}^{\ell+1}}, \quad (2.12)$$

where C is a constant independent of h , but possibly dependent on the domain Ω .

Remark 2.4. Typical situations in which Assumption 2.3 is satisfied are as follows.

- If $\Omega = [0, 1]^d$ with $d = 2, 3$, then by the definition of the space $\mathcal{Y}^\ell$ in (2.2), we have $\mathcal{Y}^\ell \hookrightarrow H^\ell(\Omega)$ for all $\ell \geq 0$. Therefore, Assumption 2.3 holds for all $k \geq 1$ on quasi-uniform, shape-regular simplicial partitions, by the standard approximation theory for finite element spaces.
- If $\Omega \subset \mathbb{R}^2$ is a polygonal domain, then the eigenfunctions of $\Delta = \sqrt{A}$ generally exhibit corner singularities and are not smooth near the vertices. A detailed singularity analysis, based on Kondrat'ev theory, yields

$$\mathcal{Y}^{\ell_\Omega} \hookrightarrow H^{\ell_\Omega}(\Omega), \quad \ell_\Omega = 1 + \frac{\pi}{\theta} - \varepsilon, \quad (2.13)$$

for arbitrarily small $\varepsilon > 0$, where $\theta = \max_{j=1, \dots, m} \theta_j$ is the largest interior angle of Ω ; see [7, Theorem 9.8 and Remark 9.12]. Consequently, Assumption 2.3 holds for quasi-uniform, shape-regular simplicial partitions $\mathcal{T}_h$, provided that the polynomial degree satisfies $1 \leq k \leq \ell_\Omega - 1$.

- More generally, if Ω is a polygonal domain, then Kondrat'ev theory shows that the eigenfunctions of $\Delta = \sqrt{A}$ belong to appropriate weighted Sobolev (Kondrat'ev) spaces that capture the corner singularities [15, 27]. In this setting, one can construct suitably graded simplicial meshes $\mathcal{T}_h$ such that Assumption 2.3 continues to hold; see, for example, [1, 28, 30] for graded-mesh constructions for functions in Kondrat'ev spaces.

We define the projection operator $\mathcal{P}_h : H \rightarrow X_h^k$ by

$$\langle \mathcal{P}_h u, v_h \rangle_H = \langle u, v_h \rangle_H \quad \text{for all } v_h \in X_h^k. \quad (2.14)$$

Furthermore, let $A_h : X_h^k \rightarrow X_h^k$ denote the finite element discretization of A , and define the Ritz projection $\mathcal{R}_h : V \rightarrow X_h^k$ through

$$a(u_h, v_h) = \langle A_h u_h, v_h \rangle_H \quad \text{and} \quad a(\mathcal{R}_h u, v_h) = a(u, v_h), \quad (2.15a)$$

It follows that for all $u \in D(A)$,

$$\langle A_h(\mathcal{R}_h u), v_h \rangle_H = \langle Au, v_h \rangle_H. \quad (2.15b)$$

For a time step size $\tau > 0$, let $t_n = n\tau$. By the variation-of-constants formula applied to (2.8), the exact solution satisfies

$$U(t_{n+1}) = e^{\tau L} U(t_n) + \int_0^\tau e^{(\tau-\sigma)L} F(U(t_n + \sigma)) d\sigma. \quad (2.16)$$

Applying the spatial discretization operators $\mathcal{P}_h$ and A_h , we obtain the spatially semidiscrete system

$$U_h(t_{n+1}) = e^{\tau L_h} U_h(t_n) + \int_0^\tau e^{(\tau-\sigma)L_h} \mathcal{P}_h F(U_h(t_n + \sigma)) d\sigma, \quad (2.17)$$

where, for notational simplicity, we use the same symbol $\mathcal{P}_h$ for its diagonal extension $\text{diag}(\mathcal{P}_h, \mathcal{P}_h)$, acting componentwise on vector-valued functions. The initial data are given by $U_h(0) = \mathcal{P}_h U(0)$, and the discrete operator L_h is defined by

$$L_h = \begin{pmatrix} 0 & 1 \\ -A_h & 0 \end{pmatrix}. \quad (2.18)$$

To construct a fully discrete scheme, we next discretize the integral term on the right-hand side of (2.17). Approximating the nonlinear term by

$$\mathcal{P}_h F(U_h(t_n + \sigma)) \approx \mathcal{P}_h F(U_h(t_n))$$

we arrive at the fully discrete exponential Euler scheme:

$$U_h^{n+1} = e^{\tau L_h} U_h^n + \tau \varphi_1(\tau L_h) \mathcal{P}_h F(U_h^n), \quad \text{with} \quad \varphi_1(z) = \frac{e^z - 1}{z}, \quad (2.19)$$

where $U_h^0 = \mathcal{P}_h U(0)$.

Throughout this article, we impose the following assumptions:

Assumption 2.5.

- (a) The projection operator $\mathcal{P}_h$ is stable in $\mathcal{Y}^\ell$ for $\ell = -1, 0, 1$, i.e.,

$$\|\mathcal{P}_h u\|_{\mathcal{Y}^\ell} \leq C_P \|u\|_{\mathcal{Y}^\ell}, \quad (2.20)$$

for some constant $C_P > 0$.

- (b) The linear operator L_h generates a C_0 -semigroup, and the following bound holds for $\ell = 0, 1$:

$$\|e^{tL_h} U\|_{\mathcal{Y}^\ell \times \mathcal{Y}^{\ell-1}} \leq e^{C_L t} \|U\|_{\mathcal{Y}^\ell \times \mathcal{Y}^{\ell-1}}, \quad (2.21)$$

for some constant $C_L > 0$ and for all $U \in \mathcal{Y}^\ell \times \mathcal{Y}^{\ell-1}$.

- (c) For $d = 2, 3$ and $\gamma \in (0, 1]$, the nonlinear function f is differentiable and satisfies the following growth conditions

$$|f(u)| \leq C_f (1 + |u|)^{\xi_\gamma}, \quad |f'(u)| \leq C_f (1 + |u|)^{\xi_\gamma - 1}, \quad (2.22)$$

with

$$\xi_\gamma < \begin{cases} \infty, & \text{for } d = 2, \gamma = 1, \\ 1 + \frac{2}{d-2\gamma}, & \text{for } d = 2, 0 < \gamma < 1, \text{ or } d = 3, 0 < \gamma \leq 1. \end{cases} \quad (2.23)$$

Assumption 2.5(a), in particular the boundedness in $\mathcal{Y}^1 = H_0^1(\Omega)$, is not restrictive, as it holds for quasi-uniform triangulations and certain adaptively refined meshes; see [4, 16]. By a standard duality argument, this immediately yields the corresponding boundedness on $\mathcal{Y}^{-1}$. Assumption 2.5(b) is standard for spatial discretizations for which A_h is symmetric and positive

definite. The growth condition in Assumption 2.5(c) is imposed in order to derive the stability estimates for the numerical scheme in the subsequent analysis.

Remark 2.6. By interpolation, Assumptions 2.5(a) and (b) imply that $\mathcal{P}_h$ is uniformly bounded on $\mathcal{Y}^\theta$ for all $\theta \in [-1, 1]$, and that e^{tL_h} is uniformly bounded on $\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$ for all $\gamma \in [0, 1]$.

Moreover, the nonlinear mapping F is locally Lipschitz continuous on $\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$ for all $\gamma \in (0, 1]$, under Assumptions 2.5(c). Indeed, for any $U_i = (u_i, v_i)^\top \in \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$, $i = 1, 2$, Lemma 3.5, in particular (3.10), yields

$$\|F(U_1) - F(U_2)\|_\gamma = \|f(u_1) - f(u_2)\|_{\mathcal{Y}^{\gamma-1}} \lesssim C(\|U_1\|_\gamma, \|U_2\|_\gamma)\|U_1 - U_2\|_\gamma.$$

Therefore, by the boundedness of e^{tL} and a standard fixed-point argument, problem (2.8) is locally well posed in $C([0, T]; \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1})$ for some $T > 0$.

We summarize the main convergence result as follows:

Theorem 2.7 (Main result). *Under Assumptions 2.2–2.5, suppose that the exact solution of (2.8) satisfies $U \in C([0, T]; \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1})$ for some $\gamma \in (0, 1]$. Then there exist $h_0 \in (0, 1)$ such that, for all $0 < h < h_0$ and $\tau > 0$ satisfying the step-size condition $\tau \lesssim h^{\frac{\ell_k}{\ell_k+1}(1-\gamma)+\mu}$ for an arbitrarily small but fixed $\mu > 0$, the fully discrete scheme (2.19) satisfies*

$$\sup_{0 \leq n \leq T/\tau} \|U(t_n) - U_h^n\|_0 \leq C(\tau h^{\frac{\ell_k}{\ell_k+1}(\gamma-1)} + h^{\frac{\ell_k}{\ell_k+1}\gamma}), \quad (2.24)$$

where

$$\ell_k := \begin{cases} k+1, & k=1, \\ k+2, & k \geq 2, \end{cases} \quad (2.25)$$

with $k \geq 1$ denoting the polynomial degree of the finite element space. The constant $C > 0$ depends only on T , $\sup_{t \in [0, T]} \|U(t)\|_\gamma$, and the constants in Assumptions 2.2–2.5, but is independent of h and τ .

Moreover, choosing $\tau = O(h^{\frac{\ell_k}{\ell_k+1}})$, which balances the two terms on the right-hand side of (2.24) and is independent of γ , yields the total error estimate

$$\sup_{0 \leq n \leq T/\tau} \|U(t_n) - U_h^n\|_0 \leq C\tau^\gamma, \quad (2.26)$$

for solutions $U \in C([0, T]; \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1})$.

Remark 2.8. The first term on the right-hand side of (2.24) represents the temporal discretization error, while the second term is the spatial discretization error. Owing to the limited regularity of the exact solution, the exponential integrator does not attain first-order convergence in time under the present assumptions. This loss is reflected in the additional factor $h^{\frac{\ell_k}{\ell_k+1}(\gamma-1)}$ appearing in the temporal error term. In the borderline case $\gamma = 1$, this factor reduces to 1, and the optimal first-order temporal convergence is recovered.

The step-size condition $\tau \lesssim h^{\frac{\ell_k}{\ell_k+1}(1-\gamma)+\mu}$ is imposed to guarantee the smallness requirement in the bootstrap argument of Lemma 4.3, in particular, to ensure that (4.31) holds. Moreover, under this condition, the error bound (2.24) implies

$$\sup_{0 \leq n \leq T/\tau} \|U(t_n) - U_h^n\|_0 \leq C(h^\mu + h^{\frac{\ell_k}{\ell_k+1}\gamma}) \rightarrow 0 \quad \text{as } h, \tau \rightarrow 0.$$

Finally, the choice $\tau = O(h^{\frac{\ell_k}{\ell_k+1}})$ balances the temporal and spatial contributions in (2.24), automatically satisfies the above step-size restriction, and yields the overall error estimate $O(\tau^\gamma)$.

Remark 2.9. We emphasize that the spatial convergence rates differ between the cases $k = 1$ and $k \geq 2$. This distinction is intrinsic to our weak-norm error analysis. More precisely, in the $\mathcal{Y}^{-1}$ norm the Ritz projection admits a higher-order estimate for $k \geq 2$; see (3.6) in Lemma 3.3 and also [10, Section 5.8]. By contrast, for $k = 1$ no analogous estimate is available in this form.

Consequently, the cases $k = 1$ and $k \geq 2$ must be treated separately in the error analysis; see Lemma 3.4.

Owing to the refined weak-norm estimate established in Lemma 3.4, the spatial convergence rate appearing on the right-hand side of (2.24), although somewhat involved in form, appears to be sharp. It is also worth noting that this rate is only of order $h^{\frac{\ell_k}{\ell_k+1}\gamma}$, due to the limited regularity of the exact solution, so that the optimal rate for k -th order finite elements cannot in general be expected. Nevertheless, our estimate shows that higher-order methods still yield a clear improvement in spatial convergence over lower-order ones, even in this low-regularity regime. These observations are confirmed by the numerical experiments in Section 5.

2.3. Frequency-localized auxiliary scheme

Since the classical error analysis of the fully discrete scheme (2.19) requires higher regularity of the exact solution, it is not directly applicable in our low-regularity setting. To overcome these difficulties, we introduce a frequency-localized auxiliary scheme, which serves as an analytical bridge in the error analysis.

For $t_n = n\tau$, the variation-of-constants formula gives

$$U(t_{n+1}) = e^{\tau L}U(t_n) + \int_0^\tau e^{(\tau-s)L}F(e^{sL}U(t_n))ds + R_1(t_n), \quad (2.27)$$

where

$$R_1(t_n) = \int_0^\tau e^{(\tau-s)L} \left(F(U(t_n + s)) - F(e^{sL}U(t_n)) \right) ds. \quad (2.28)$$

The remainder term $R_1(t_n)$ can be controlled by combining the semigroup property of e^{tL} in (2.10) with the variation-of-constants formula; see Lemma 4.1 below.

We next consider the approximation of the remaining integral term in (2.27). A standard first-order approximation would replace $F(e^{sL}U(t_n))$ by $F(U(t_n))$, as in the classical exponential integrator. However,

$$F(e^{sL}U(t_n)) - F(U(t_n)) = \int_0^s \frac{d}{d\sigma} F(e^{\sigma L}U(t_n)) d\sigma, \quad (2.29)$$

and the derivative on the right-hand side involves spatial derivatives of $U(t_n)$. This presents a difficulty under the low-regularity assumption $(u, v) \in \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$ with $\gamma \in (0, 1]$, since the term $\frac{d}{d\sigma} F(e^{\sigma L}U(t_n))$ requires one additional order of spatial differentiability beyond what is available.

Inspired by [12], we introduce a frequency decomposition operator Π_r , defined as the spectral projection of functions onto the low-frequency subspace $H_r \hookrightarrow H$, which is spanned by the basis functions

$$\mathcal{B}_r = \{e_\alpha \in H \mid \alpha \in I, |\lambda_\alpha| \leq r\}.$$

The projection operator Π_r satisfies

$$\Pi_r u \in H_r, \quad \langle \Pi_r u, v \rangle_H = \langle u, v \rangle_H, \quad \forall v \in H_r. \quad (2.30)$$

In our subsequent error analysis, we make essential use of the improved regularity of the low-frequency projected component to establish rigorous error bounds.

Thus we approximate the integral term on the right-hand side of (2.27) by

$$\begin{aligned} \int_0^\tau e^{(\tau-s)L} F(e^{sL}U(t_n)) ds &= \int_0^\tau e^{(\tau-s)L} F(e^{sL}\Pi_r U(t_n)) ds + R_2(t_n) \\ &= \tau \varphi_1(\tau L) F(\Pi_r U(t_n)) + R_2(t_n) + R_3(t_n) \\ &= \tau \varphi_1(\tau L) \Pi_r F(\Pi_r U(t_n)) + R_2(t_n) + R_3(t_n) + R_4(t_n), \end{aligned} \quad (2.31)$$

where φ_1 is given by (2.19). The residual terms $R_2(t_n)$, $R_3(t_n)$, and $R_4(t_n)$ are defined by:

$$R_2(t_n) = \int_0^\tau e^{(\tau-s)L} \left(F(e^{sL}U(t_n)) - F(e^{sL}\Pi_r U(t_n)) \right) ds, \quad (2.32a)$$

$$R_3(t_n) = \int_0^\tau e^{(\tau-s)L} \int_0^s \frac{d}{d\sigma} F(e^{\sigma L} \Pi_r U(t_n)) d\sigma ds, \quad (2.32b)$$

$$R_4(t_n) = \tau \varphi_1(\tau L) (F(\Pi_r U(t_n)) - \Pi_r F(\Pi_r U(t_n))). \quad (2.32c)$$

The approximation (2.31) together with (2.27) implies

$$U(t_{n+1}) = e^{\tau L} U(t_n) + \tau \varphi_1(\tau L) \Pi_r F(\Pi_r U(t_n)) + \sum_{i=1}^4 R_i(t_n). \quad (2.33)$$

And thus we arrive at the auxiliary scheme

$$U^{n+1} = e^{\tau L} U^n + \tau \varphi_1(\tau L) \Pi_r F(\Pi_r U^n), \quad n \geq 0. \quad (2.34)$$

Remark 2.10. The scheme (2.34) differs from the classical exponential integrator in that the frequency projection Π_r is applied both to the numerical solution and to the nonlinearity. This provides the additional regularity needed for the consistency analysis under low-regularity initial data. We emphasize, however, that (2.34) is introduced only as an analytical tool: in general, the spectral projector Π_r is not computationally practical on general domains, and our fully discrete method is therefore not obtained by directly discretizing (2.34).

3. Preliminary results

In this section, we present the key lemmas that will be used to prove the convergence result. These include the Bernstein estimates for the operator Π_r defined in (2.30) and the projection error estimates for the operator $\mathcal{P}_h$ defined in (2.14). Finally, we compare the semigroup e^{tL} of the continuous problem with the semigroup e^{tL_h} obtained after finite element spatial discretization. This comparison plays a crucial role in establishing the error estimates for the fully discrete scheme of the exponential integrator based on finite element discretization.

Lemma 3.1 (Bernstein-type inequality). *Let $u \in \mathcal{Y}^\ell$ for $\ell \in \mathbb{R}$, then there holds*

$$\|\Pi_r u\|_{\mathcal{Y}^{\ell+m}} \lesssim r^m \|u\|_{\mathcal{Y}^\ell} \quad \text{and} \quad \|\Pi_{>r} u\|_{\mathcal{Y}^{\ell-m}} \lesssim r^{-m} \|u\|_{\mathcal{Y}^\ell}, \quad (3.1)$$

for all $m \geq 0$.

Proof. The result follows directly from the definition of the norms in $\mathcal{Y}^\ell$ and the spectral localization properties of Π_r and $\Pi_{>r}$. $\square$

Next, we use a duality argument to derive an estimate for the L^2 -projection in $\mathcal{Y}^{-1}$.

Lemma 3.2 (Projection error estimate). *Let $\mathcal{P}_h$ be the projection operator defined in (2.14). Then, for every $\ell \in [0, 1]$,*

$$\|(\text{id} - \mathcal{P}_h)u\|_{\mathcal{Y}^0} \lesssim h^\ell \|u\|_{\mathcal{Y}^\ell}, \quad u \in \mathcal{Y}^\ell, \quad (3.2)$$

$$\|(\text{id} - \mathcal{P}_h)v\|_{\mathcal{Y}^{-1}} \lesssim h^\ell \|v\|_{\mathcal{Y}^{\ell-1}}, \quad v \in \mathcal{Y}^{\ell-1}. \quad (3.3)$$

Proof. We first prove (3.2). Since $\mathcal{Y}^0 = L^2(\Omega)$ and $\mathcal{Y}^1 = H_0^1(\Omega)$, it follows from the standard approximation property of the L^2 -projection and Assumption 2.3 that

$$\|(\text{id} - \mathcal{P}_h)u\|_{\mathcal{Y}^0} = \|(\text{id} - \mathcal{P}_h)u\|_{L^2(\Omega)} \lesssim h \|u\|_{H_0^1(\Omega)} = h \|u\|_{\mathcal{Y}^1}$$

for all $u \in \mathcal{Y}^1$. On the other hand, $\mathcal{P}_h$ is bounded on $\mathcal{Y}^0 = L^2(\Omega)$, and hence $\|(\text{id} - \mathcal{P}_h)u\|_{\mathcal{Y}^0} \lesssim \|u\|_{\mathcal{Y}^0}$ for all $u \in \mathcal{Y}^0$. Interpolating between the above two estimates, we obtain (3.2) for all $u \in \mathcal{Y}^\ell$, $\ell \in [0, 1]$.

Next we prove (3.3). By duality, for all $v \in \mathcal{Y}^0$,

$$\begin{aligned} \|(\text{id} - \mathcal{P}_h)v\|_{\mathcal{Y}^{-1}} &= \sup_{\|w\|_{\mathcal{Y}^1}=1} \langle (\text{id} - \mathcal{P}_h)v, w \rangle_H = \sup_{\|w\|_{\mathcal{Y}^1}=1} \langle v, (\text{id} - \mathcal{P}_h)w \rangle_H \\ &\leq \|v\|_{\mathcal{Y}^0} \sup_{\|w\|_{\mathcal{Y}^1}=1} \|(\text{id} - \mathcal{P}_h)w\|_{\mathcal{Y}^0} \lesssim h \|v\|_{\mathcal{Y}^0}, \end{aligned}$$

where we used (3.2) with $\ell = 1$ in the last step. Moreover, by Assumption 2.5(a), $\mathcal{P}_h$ is bounded on $\mathcal{Y}^{-1}$, and hence $\|(\text{id} - \mathcal{P}_h)v\|_{\mathcal{Y}^{-1}} \lesssim \|v\|_{\mathcal{Y}^{-1}}$ for all $v \in \mathcal{Y}^{-1}$. Interpolating between these two estimates, we obtain (3.3) for all $v \in \mathcal{Y}^{\ell-1}$, $\ell \in [0, 1]$. $\square$

Furthermore, under Assumption 2.2, by combining the duality argument with the Aubin–Nitsche technique, we obtain the following high-order approximation estimates for the projection operators $\mathcal{P}_h$ and $\mathcal{R}_h$.

Lemma 3.3 (Error estimates of projection operators $\mathcal{P}_h$ and $\mathcal{R}_h$). *Under Assumption 2.3, let $\mathcal{P}_h$ and $\mathcal{R}_h$ be the projection operators onto the finite element space X_h^k , defined by (2.14) and (2.15). Then for any $u \in \mathcal{Y}^{k+1}$, the following estimates hold for $k \geq 1$:*

$$\|u - \mathcal{P}_h u\|_{\mathcal{Y}^0} + \|u - \mathcal{R}_h u\|_{\mathcal{Y}^0} \lesssim h^{k+1} \|u\|_{\mathcal{Y}^{k+1}}, \quad (3.4)$$

and

$$\|u - \mathcal{P}_h u\|_{\mathcal{Y}^{-1}} \lesssim h^{k+2} \|u\|_{\mathcal{Y}^{k+1}}. \quad (3.5)$$

Moreover, if $k \geq 2$, there holds

$$\|u - \mathcal{R}_h u\|_{\mathcal{Y}^{-1}} \lesssim h^{k+2} \|u\|_{\mathcal{Y}^{k+1}}. \quad (3.6)$$

Proof. Since $\mathcal{Y}^0 = L^2(\Omega)$, the estimate (3.4) follows directly from the approximation property in Assumption 2.3, the best approximation property of $\mathcal{P}_h$ in $L^2(\Omega)$, and the best approximation property of $\mathcal{R}_h$ in $H_0^1(\Omega)$ together with the standard Aubin–Nitsche argument.

The estimate (3.5) is easily verified by duality and by using the error estimate (3.4):

$$\begin{aligned} \|u - \mathcal{P}_h u\|_{\mathcal{Y}^{-1}} &= \sup_{\|w\|_{\mathcal{Y}^1}=1} \langle u - \mathcal{P}_h u, w \rangle_H = \sup_{\|w\|_{\mathcal{Y}^1}=1} \langle u - \mathcal{P}_h u, w - \mathcal{P}_h w \rangle_H \\ &\leq \sup_{\|w\|_{\mathcal{Y}^1}=1} \|u - \mathcal{P}_h u\|_{\mathcal{Y}^0} \|w - \mathcal{P}_h w\|_{\mathcal{Y}^0} \lesssim h^{k+2} \|u\|_{\mathcal{Y}^{k+1}}. \end{aligned}$$

For (3.6), let $\phi = A^{-1}w$. Then

$$\begin{aligned} \|u - \mathcal{R}_h u\|_{\mathcal{Y}^{-1}} &= \sup_{\|w\|_{\mathcal{Y}^1}=1} \langle u - \mathcal{R}_h u, w \rangle_H = \sup_{\|w\|_{\mathcal{Y}^1}=1} \langle u - \mathcal{R}_h u, A\phi \rangle_H \\ &= \sup_{\|w\|_{\mathcal{Y}^1}=1} a(u - \mathcal{R}_h u, \phi) = \sup_{\|w\|_{\mathcal{Y}^1}=1} a(u - \mathcal{R}_h u, \phi - \mathcal{R}_h \phi), \end{aligned} \quad (3.7)$$

where we have used the orthogonality

$$a(u - \mathcal{R}_h u, \mathcal{R}_h \phi) = 0.$$

By using the elliptic estimate $\|\phi\|_{\mathcal{Y}^3} \lesssim \|w\|_{\mathcal{Y}^1}$, and the best approximation property of $\mathcal{R}_h$ for $u \in \mathcal{Y}^{k+1}$ and $\phi \in \mathcal{Y}^3$ in the $\mathcal{Y}^1$ norm, together with Assumption 2.3, we have

$$a(u - \mathcal{R}_h u, \phi - \mathcal{R}_h \phi) \lesssim \|u - \mathcal{R}_h u\|_{\mathcal{Y}^1} \|\phi - \mathcal{R}_h \phi\|_{\mathcal{Y}^1} \lesssim h^k \|u\|_{\mathcal{Y}^{k+1}} h^2 \|\phi\|_{\mathcal{Y}^3} \lesssim h^{k+2} \|u\|_{\mathcal{Y}^{k+1}} \|w\|_{\mathcal{Y}^1}.$$

By combining this with (3.7) we conclude (3.6). $\square$

We are now in a position to estimate the difference between the two semigroups.

Lemma 3.4. *Let $V(0) \in \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$, and let the operators L and L_h be as defined in (2.9) and (2.18), respectively. Then, if $r \leq h^{-1}$, under Assumptions 2.2, 2.3 and 2.5, the following estimate holds:*

$$\|e^{tL} \Pi_r V(0) - e^{tL_h} \mathcal{P}_h \Pi_r V(0)\|_0 \lesssim \begin{cases} h^2 r^{3-\gamma}, & \text{for } k = 1, \\ h^{k+2} r^{k+3-\gamma} + h^\gamma, & \text{for } k \geq 2, \end{cases} \quad (3.8)$$

where k is the polynomial degree of the finite element basis functions of X_h^k .

Proof (sketch). This result is equivalent to estimating the error of the finite element semidiscretization for the linear wave equation with initial data $\Pi_r V(0)$ in the weak norm $\mathcal{Y}^0 \times \mathcal{Y}^{-1}$. Owing to the presence of the frequency localization operator Π_r , the initial data are effectively regularized. In particular, by the Bernstein-type inequality in Lemma 3.1, arbitrary powers

of the differential operator A acting on the low-frequency part remain bounded, with bounds depending only on suitable positive powers of r .

This observation allows us to combine the regularizing property of Π_r with the approximation estimates in Lemma 3.3. We note, however, that the proof differs essentially between the cases $k \geq 2$ and $k = 1$. For $k \geq 2$, it relies on the sharp $\mathcal{Y}^{-1}$ -error estimate for the Ritz projection. For $k = 1$, since such an estimate is unavailable, we instead use the sharp $\mathcal{Y}^0$ -error estimate for the Ritz projection and a different energy-based argument.

The detailed proof is deferred to Appendix A. $\square$

We next collect several estimates for the nonlinear term in (1.1):

Lemma 3.5. *Suppose that the nonlinear function f satisfies the growth conditions (2.22). Then, the following estimates hold:*

$$\|f(u)\|_{\mathcal{Y}^{\gamma-1}} \leq C(\|u\|_{\mathcal{Y}^\gamma}), \quad (3.9)$$

$$\|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} \leq C(\|u\|_{\mathcal{Y}^\gamma}, \|v\|_{\mathcal{Y}^\gamma})\|u - v\|_{\mathcal{Y}^\gamma}, \quad (3.10)$$

$$\|f(u)\|_{\mathcal{Y}^{-1}} \leq C(\|u\|_{\mathcal{Y}^\gamma})(1 + \|u\|_{\mathcal{Y}^0}), \quad (3.11)$$

$$\|f(u) - f(v)\|_{\mathcal{Y}^{-1}} \leq C(\|u\|_{\mathcal{Y}^\gamma}, \|v\|_{\mathcal{Y}^\gamma})\|u - v\|_{\mathcal{Y}^0}, \quad (3.12)$$

$$\|f'(u)v\|_{\mathcal{Y}^{-1}} \leq C(\|u\|_{\mathcal{Y}^\gamma})\|v\|_{\mathcal{Y}^0}, \quad (3.13)$$

$$\|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} \leq C(\|u\|_{\mathcal{Y}^\gamma}, \|v\|_{\mathcal{Y}^\gamma})\|u - v\|_{\mathcal{Y}^0}^\alpha, \quad (3.14)$$

where $C(\|u\|_{\mathcal{Y}^\gamma})$ and $C(\|u\|_{\mathcal{Y}^\gamma}, \|v\|_{\mathcal{Y}^\gamma})$ denote constants that depend on $\|u\|_{\mathcal{Y}^\gamma}$ and both $\|u\|_{\mathcal{Y}^\gamma}$ and $\|v\|_{\mathcal{Y}^\gamma}$, respectively. Specifically, $\alpha \in (0, 1]$ can be chosen as

$$\alpha = \begin{cases} \frac{1}{2}, & \text{for } d = 2, \quad \gamma = 1, \\ \min \left\{ 1, \frac{d-2\gamma}{2\gamma} \left(\frac{d+2(1-\gamma)}{d-2\gamma} - \xi_\gamma \right) \right\}, & \text{for } d = 2, \quad 0 < \gamma < 1, \text{ or } d = 3, \quad 0 < \gamma \leq 1. \end{cases} \quad (3.15)$$

Proof (sketch). We note that, for $0 \leq \ell \leq 1$, the embeddings $\mathcal{Y}^\ell \hookrightarrow H^\ell(\Omega)$ and $H^{-\ell}(\Omega) \hookrightarrow \mathcal{Y}^{-\ell}$ hold. Moreover, since $0 < \gamma \leq 1$, the estimates follow from the growth assumption (2.22), the mean-value theorem, Hölder's inequality, and Sobolev embeddings. As the argument is standard but somewhat technical, we defer the proof to Appendix B. $\square$

4. Proof of the convergence result

In this section, we prove the main convergence theorem. We begin by estimating the local error of the frequency-localized auxiliary scheme (2.34). Using the semigroup difference estimates from Lemma 3.4 and the nonlinear estimates from Lemma 3.5, we then derive error bounds for the fully discrete scheme (2.19) in the $\mathcal{Y}^0 \times \mathcal{Y}^{-1}$ norm. We also establish uniform bounds for the numerical solution in $\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$, which allow us to control the nonlinear contributions. Combining these ingredients, we close the induction argument and obtain the desired error estimate.

4.1. Consistency error estimate for the auxiliary scheme

Recalling the definitions of the remainder terms in (2.28) and (2.32), we establish the following result under the stated assumptions.

Lemma 4.1. *Under the conditions of Theorem 2.7, the following estimates hold:*

$$\|R_1(t_n)\|_0 \lesssim \tau^2, \quad \|R_2(t_n)\|_0 + \|R_4(t_n)\|_0 \lesssim \tau r^{-\gamma}, \quad \text{and} \quad \|R_3(t_n)\|_0 \lesssim \tau^2 r^{1-\gamma}. \quad (4.1)$$

Proof. Using the Taylor expansion, we have

$$\begin{aligned} F(U(t_n + s)) - F(e^{sL}U(t_n)) &= F'(\theta U(t_n + s) + (1 - \theta)e^{sL}U(t_n)) (U(t_n + s) - e^{sL}U(t_n)) \\ &= F'(\theta U(t_n + s) + (1 - \theta)e^{sL}U(t_n)) \int_0^s e^{(s-\sigma)L} F(U(t_n + \sigma)) d\sigma, \end{aligned} \quad (4.2)$$

where $\theta \in [0, 1]$. Therefore,

$$\|R_1(t_n)\|_0 \lesssim \tau^2 \sup_{\theta, s, \sigma} \left\| F'(\theta U(t_n + s) + (1 - \theta)e^{sL}U(t_n)) e^{(s-\sigma)L} F(U(t_n + \sigma)) \right\|_0.$$

Let us denote $U_\theta = (U_{\theta,1}, U_{\theta,2})^\top := \theta U(t_n + s) + (1 - \theta)e^{sL}U(t_n)$, and

$$F'(U_\theta) = \begin{pmatrix} 0 & 0 \\ f'(U_{\theta,1}) & 0 \end{pmatrix}.$$

Using the nonlinear estimate (3.13), we find

$$\begin{aligned} &\left\| F'(\theta U(t_n + s) + (1 - \theta)e^{sL}U(t_n)) e^{(s-\sigma)L} F(U(t_n + \sigma)) \right\|_0 \\ &= \left\| f'(U_{\theta,1}) \left(e^{(s-\sigma)L} F(U(t_n + \sigma)) \right)_1 \right\|_{\mathcal{Y}^{-1}} \lesssim C(\|U_{\theta,1}\|_{\mathcal{Y}^\gamma}) \cdot \left\| \left(e^{(s-\sigma)L} F(U(t_n + \sigma)) \right)_1 \right\|_{\mathcal{Y}^0}. \end{aligned}$$

where $(e^{(s-\sigma)L} F(U(t_n + \sigma)))_1$ denotes the first component of the vector $e^{(s-\sigma)L} F(U(t_n + \sigma))$.

From the semigroup property (2.10) and the estimate (3.11), we further obtain

$$\left\| \left(e^{(s-\sigma)L} F(U(t_n + \sigma)) \right)_1 \right\|_{\mathcal{Y}^0} \lesssim \|f(u(t_n + \sigma))\|_{\mathcal{Y}^{-1}} \leq C(\|u(t_n + \sigma)\|_{\mathcal{Y}^\gamma}) \cdot (1 + \|u(t_n + \sigma)\|_{\mathcal{Y}^0}).$$

Moreover, from the definition of U_θ we have

$$\|U_{\theta,1}\|_{\mathcal{Y}^\gamma} \leq \|U_\theta\|_\gamma \lesssim \|U(t_n + s)\|_\gamma + \|U(t_n)\|_\gamma.$$

Combining the above estimates yields

$$\|R_1(t_n)\|_0 \lesssim \tau^2 C\left(\|U\|_{C([0,T]; \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1})}\right) \lesssim \tau^2, \quad (4.3)$$

provided $U \in C([0, T]; \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1})$.

For R_2 and R_4 , using the semigroup property (2.10) we note that

$$\|F(e^{sL}U(t_n)) - F(e^{sL}\Pi_r U(t_n))\|_0 = \|f(\tilde{u}(t_n + s)) - f(\Pi_r \tilde{u}(t_n + s))\|_{\mathcal{Y}^{-1}},$$

where $(\tilde{u}(t_n + s), \tilde{v}(t_n + s))^\top := e^{sL}(u(t_n), v(t_n))^\top$.

The boundedness of $e^{(\tau-s)L}$ and $\varphi_1(\tau L)$, together with Bernstein-type inequalities and the nonlinear estimates (3.12) and (3.11), gives

$$\begin{aligned} &\|R_2(t_n)\|_0 + \|R_4(t_n)\|_0 \\ &\lesssim \tau \sup_s \|f(\tilde{u}(t_n + s)) - f(\Pi_r \tilde{u}(t_n + s))\|_{\mathcal{Y}^{-1}} + \tau \sup_s \|f(\Pi_r \tilde{u}(t_n + s)) - \Pi_r f(\Pi_r \tilde{u}(t_n + s))\|_{\mathcal{Y}^{-1}} \\ &\lesssim \tau \sup_s C\left(\|\tilde{u}(t_n + s)\|_{\mathcal{Y}^\gamma}, \|\Pi_r \tilde{u}(t_n + s)\|_{\mathcal{Y}^\gamma}\right) \cdot \|\tilde{u}(t_n + s) - \Pi_r \tilde{u}(t_n + s)\|_{\mathcal{Y}^0} \\ &\quad + \tau r^{-\gamma} \sup_s \|f(\Pi_r \tilde{u}(t_n + s))\|_{\mathcal{Y}^0} \\ &\lesssim \tau r^{-\gamma} \sup_s C\left(\|\tilde{u}(t_n + s)\|_{\mathcal{Y}^\gamma}, \|\Pi_r \tilde{u}(t_n + s)\|_{\mathcal{Y}^\gamma}\right). \end{aligned}$$

For R_3 , we compute

$$\begin{aligned} \frac{d}{d\sigma} F(e^{\sigma L} \Pi_r U(t_n)) &= \begin{pmatrix} 0 & 0 \\ f'(\Pi_r \tilde{u}(t_n + \sigma)) & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -A & 0 \end{pmatrix} \begin{pmatrix} \Pi_r \tilde{u}(t_n + \sigma) \\ \Pi_r \tilde{v}(t_n + \sigma) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ f'(\Pi_r \tilde{u}(t_n + \sigma)) \cdot \Pi_r \tilde{v}(t_n + \sigma) \end{pmatrix}, \end{aligned} \quad (4.4)$$

where, as before, $(\tilde{u}(t_n + \sigma), \tilde{v}(t_n + \sigma))^\top = e^{\sigma L}(u(t_n), v(t_n))^\top$.

Applying (4.4), together with estimate (3.13) and Bernstein's inequality, we obtain

$$\begin{aligned} \left\| \frac{d}{d\sigma} F(e^{\sigma L} \Pi_r U(t_n)) \right\|_0 &\lesssim \|f'(\Pi_r \tilde{u}(t_n + \sigma)) \cdot \Pi_r \tilde{v}(t_n + \sigma)\|_{\mathcal{Y}^{-1}} \\ &\leq C \left(\|\Pi_r \tilde{u}(t_n + \sigma)\|_{\mathcal{Y}^\gamma} \right) \|\Pi_r v(t_n)\|_{\mathcal{Y}^0} \lesssim C \left(\|\Pi_r \tilde{u}(t_n + \sigma)\|_{\mathcal{Y}^\gamma} \right) \cdot r^{1-\gamma} \|U(t_n)\|_\gamma. \end{aligned}$$

Substituting this estimate into (2.32b) completes the proof of the bound for R_3 in (4.1). $\square$

4.2. Error estimate for the fully discrete scheme

Next, we employ the expansions of the exact and the numerical solution, and denote $\mathcal{L}^n = \sum_{i=1}^4 R_i(t_n)$. By iteratively applying the consistency error formula (2.33), we conclude that

$$U(t_{n+1}) = e^{t_{n+1}L} U(0) + \sum_{j=0}^n \tau e^{(t_{n+1}-t_{j+1})L} \varphi_1(\tau L) \Pi_r F(\Pi_r U(t_j)) + \sum_{j=0}^n e^{(t_{n+1}-t_{j+1})L} \mathcal{L}^j. \quad (4.5)$$

Similarly, for the fully discrete scheme (2.19), we have

$$U_h^{n+1} = e^{t_{n+1}L_h} \mathcal{P}_h U^0 + \sum_{j=0}^n \tau e^{(t_{n+1}-t_{j+1})L_h} \varphi_1(\tau L_h) \mathcal{P}_h F(U_h^j). \quad (4.6)$$

Subtracting (4.6) from (4.5), we obtain

$$\begin{aligned} U(t_{n+1}) - U_h^{n+1} &= e^{t_{n+1}L} U(0) - e^{t_{n+1}L_h} \mathcal{P}_h U^0 \\ &+ \sum_{j=0}^n \tau \left(e^{(t_{n+1}-t_{j+1})L} \varphi_1(\tau L) \Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h} \varphi_1(\tau L_h) \mathcal{P}_h F(U_h^j) \right) \\ &+ \sum_{j=0}^n e^{(t_{n+1}-t_{j+1})L} \mathcal{L}^j. \end{aligned} \quad (4.7)$$

In the next lemma, we estimate the three terms on the right-hand side of (4.7).

Lemma 4.2. *Under the conditions of Theorem 2.7, and further assume that $1 \leq r \leq h^{-1}$, there exist constants $C, C_0 > 0$, such that*

$$\begin{aligned} \|U(t_{n+1}) - U_h^{n+1}\|_0 &\leq C_0(r^{-\gamma} + h^{\ell_k} r^{\ell_k+1-\gamma}) + \sum_{j=0}^n \tau \cdot C \left(\|U_h^j\|_\gamma \right) \|U(t_j) - U_h^j\|_0 \\ &+ \sum_{j=0}^n C_0 \tau \cdot (\tau r^{1-\gamma} + r^{-\gamma} + h^{\ell_k} r^{\ell_k+1-\gamma}) + \sum_{j=0}^n C \left(\|U_h^j\|_\gamma \right) \tau \cdot r^{-\gamma}, \end{aligned} \quad (4.8)$$

where ℓ_k is defined in (2.25). The constant C_0 is independent of τ, h, r , and the bound on the numerical solution, whereas C depends only on that bound.

Proof. We begin by estimating the last term on the right-hand side of (4.7). By the semigroup property (2.10) and the semidiscrete consistency estimate in Lemma 4.1, we obtain

$$\left\| e^{(t_{n+1}-t_{j+1})L} \mathcal{L}^j \right\|_0 \lesssim \tau^2 + \tau r^{-\gamma} + \tau^2 r^{1-\gamma} \lesssim \tau (r^{-\gamma} + \tau r^{1-\gamma}). \quad (4.9)$$

For the first term on the right-hand side of (4.7), we decompose

$$\begin{aligned} &e^{t_{n+1}L} U(0) - e^{t_{n+1}L_h} \mathcal{P}_h U^0 \\ &= \underbrace{(e^{t_{n+1}L} U(0) - e^{t_{n+1}L} \Pi_r U(0))}_{(I)} + \underbrace{(e^{t_{n+1}L} \Pi_r U(0) - e^{t_{n+1}L_h} \mathcal{P}_h \Pi_r U(0))}_{(II)} \\ &+ \underbrace{(e^{t_{n+1}L_h} \mathcal{P}_h \Pi_r U(0) - e^{t_{n+1}L_h} \mathcal{P}_h U(0))}_{(III)}. \end{aligned} \quad (4.10)$$

For term (I), by the definition of Π_r and (2.10),

$$\|e^{t_{n+1}L}U(0) - e^{t_{n+1}L}\Pi_r U(0)\|_0 \lesssim \|(\text{id} - \Pi_r)U(0)\|_0 \lesssim r^{-\gamma}\|U(0)\|_\gamma. \quad (4.11)$$

For term (III), using assumption (2.21) and the stability of $\mathcal{P}_h$ (Assumption 2.5 (a)),

$$\begin{aligned} \|e^{t_{n+1}L_h}\mathcal{P}_h\Pi_r U(0) - e^{t_{n+1}L_h}\mathcal{P}_h U(0)\|_0 &\lesssim \|\mathcal{P}_h\Pi_r U(0) - \mathcal{P}_h U(0)\|_0 \\ &\lesssim \|\mathcal{P}_h(\Pi_r u(0) - u(0))\|_{\mathcal{Y}^0} + \|\mathcal{P}_h(\Pi_r v(0) - v(0))\|_{\mathcal{Y}^{-1}} \\ &\lesssim \|\Pi_r U(0) - U(0)\|_0 \lesssim r^{-\gamma}\|U(0)\|_\gamma. \end{aligned} \quad (4.12)$$

For term (II), by Lemma 3.4,

$$\|e^{t_{n+1}L}\Pi_r U(0) - e^{t_{n+1}L_h}\mathcal{P}_h\Pi_r U(0)\|_0 \lesssim (h^\gamma + h^{\ell_k}r^{\ell_k+1-\gamma})\|U(0)\|_\gamma. \quad (4.13)$$

Combining (4.11)–(4.13), and noting that $r \leq h^{-1}$, gives

$$\|e^{t_{n+1}L}U(0) - e^{t_{n+1}L_h}U_h^0\|_0 \lesssim (r^{-\gamma} + h^{\ell_k}r^{\ell_k+1-\gamma})\|U(0)\|_\gamma. \quad (4.14)$$

Similarly, for the nonlinear contribution in (4.7), we write

$$\begin{aligned} &e^{(t_{n+1}-t_{j+1})L}\varphi_1(\tau L)\Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h F(U_h^j) \\ &= \underbrace{\left(e^{(t_{n+1}-t_{j+1})L}\varphi_1(\tau L)\Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h\Pi_r F(\Pi_r U(t_j))\right)}_{(i)} \\ &\quad + \underbrace{\left(e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h\Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h F(\Pi_r U(t_j))\right)}_{(ii)} \\ &\quad + \underbrace{\left(e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h F(U_h^j)\right)}_{(iii)}. \end{aligned} \quad (4.15)$$

For term (ii), by assumption (2.21), the Bernstein inequality (3.1), and the stability of $\mathcal{P}_h$, together with Lemma 3.5,

$$\begin{aligned} &\left\|e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h\Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h F(\Pi_r U(t_j))\right\|_0 \\ &\lesssim \left\|\mathcal{P}_h\Pi_r F(\Pi_r U(t_j)) - \mathcal{P}_h F(\Pi_r U(t_j))\right\|_0 \\ &\lesssim \left\|\Pi_r F(\Pi_r U(t_j)) - F(\Pi_r U(t_j))\right\|_0 = \left\|\Pi_r f(\Pi_r u(t_j)) - f(\Pi_r u(t_j))\right\|_{\mathcal{Y}^{-1}} \\ &\lesssim r^{-\gamma}\|f(\Pi_r u(t_j))\|_{\mathcal{Y}^\gamma} \leq r^{-\gamma}C\left(\|u(t_j)\|_{\mathcal{Y}^\gamma}\right). \end{aligned} \quad (4.16)$$

For term (iii),

$$\begin{aligned} &\left\|e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h F(U_h^j)\right\|_0 \\ &\lesssim \left\|F(\Pi_r U(t_j)) - F(U_h^j)\right\|_0 = \|f(\Pi_r u(t_j)) - f(u_h^j)\|_{\mathcal{Y}^{-1}} \\ &\lesssim C\left(\|u(t_j)\|_{\mathcal{Y}^\gamma}, \|u_h^j\|_{\mathcal{Y}^\gamma}\right)\|\Pi_r u(t_j) - u_h^j\|_{\mathcal{Y}^0} \\ &\leq C\left(\|u(t_j)\|_{\mathcal{Y}^\gamma}, \|u_h^j\|_{\mathcal{Y}^\gamma}\right) \cdot \left(\|\Pi_{>r}u(t_j)\|_{\mathcal{Y}^0} + \|u(t_j) - u_h^j\|_{\mathcal{Y}^0}\right) \\ &\lesssim C\left(\|u(t_j)\|_{\mathcal{Y}^\gamma}, \|u_h^j\|_{\mathcal{Y}^\gamma}\right) \cdot \left(r^{-\gamma} + \|U(t_j) - U_h^j\|_0\right). \end{aligned} \quad (4.17)$$

For term (i), using the integral representation of φ_1 ,

$$\begin{aligned} &e^{(t_{n+1}-t_{j+1})L}\varphi_1(\tau L)\Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h}\varphi_1(\tau L_h)\mathcal{P}_h\Pi_r F(\Pi_r U(t_j)) \\ &= \frac{1}{\tau} \int_0^\tau \left(e^{(t_{n+1}-t_j-s)L}\Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_j-s)L_h}\mathcal{P}_h\Pi_r F(\Pi_r U(t_j))\right) ds, \end{aligned}$$

and applying Lemma 3.4 with $V(0) = F(\Pi_r U(t_j))$, together with Lemma 3.5, yields

$$\begin{aligned} & \left\| e^{(t_{n+1}-t_{j+1})L} \varphi_1(\tau L) \Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h} \varphi_1(\tau L_h) \mathcal{P}_h \Pi_r F(\Pi_r U(t_j)) \right\|_0 \\ & \lesssim (h^\gamma + h^{\ell_k} r^{\ell_k+1-\gamma}) \|F(\Pi_r U(t_j))\|_\gamma \\ & = (h^\gamma + h^{\ell_k} r^{\ell_k+1-\gamma}) \|f(\Pi_r u(t_j))\|_{\mathcal{Y}^{\gamma-1}} \lesssim (r^{-\gamma} + h^{\ell_k} r^{\ell_k+1-\gamma}) \cdot C \left(\|u(t_j)\|_{\mathcal{Y}^\gamma} \right), \end{aligned} \quad (4.18)$$

where, in the last inequality, we have used $r \leq h^{-1}$.

Collecting (4.16)–(4.18), we obtain

$$\begin{aligned} & \left\| e^{(t_{n+1}-t_{j+1})L} \varphi_1(\tau L) \Pi_r F(\Pi_r U(t_j)) - e^{(t_{n+1}-t_{j+1})L_h} \varphi_1(\tau L_h) \mathcal{P}_h F(U_h^j) \right\|_0 \\ & \leq C \left(\|u(t_j)\|_{\mathcal{Y}^\gamma}, \|u_h^j\|_{\mathcal{Y}^\gamma} \right) \|U(t_j) - U_h^j\|_0 + C \left(\|u(t_j)\|_{\mathcal{Y}^\gamma}, \|u_h^j\|_{\mathcal{Y}^\gamma} \right) r^{-\gamma} \\ & \quad + (r^{-\gamma} + h^{\ell_k} r^{\ell_k+1-\gamma}) C \left(\|u_h^j\|_{\mathcal{Y}^\gamma} \right). \end{aligned} \quad (4.19)$$

Substituting (4.14), (4.19), and the local consistency bound (4.9) into (4.7), and using the regularity $U \in C([0, T]; \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1})$, we conclude that there exists a constant $C_0 > 0$ such that (4.8) holds. $\square$

4.3. Estimate of the numerical solution in $\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$

At this point, a direct application of Gronwall's inequality is not yet possible, since the constant on the right-hand side of (4.8) may depend on $\|U_h^j\|_\gamma$, so such an argument requires a uniform bound for the numerical solution in $\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$. Moreover, the estimate in $\mathcal{Y}^0 \times \mathcal{Y}^{-1}$ at the $(n+1)$ -th time step depends on the boundedness of the numerical solution at all previous time steps, that is, on $\|U_h^j\|_\gamma$ for $j = 0, \dots, n$. We therefore estimate the error and the boundedness of the numerical solution simultaneously by induction. To this end, we take the $\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}$ -norm on both sides of (2.19), which yields

$$\begin{aligned} \|U_h^{n+1}\|_\gamma & \leq \|e^{t_{n+1}L_h} U_h^0\|_\gamma + \sum_{j=0}^n \tau \left\| e^{(t_{n+1}-t_{j+1})L_h} \varphi_1(\tau L_h) \mathcal{P}_h F(U_h^j) \right\|_\gamma \\ & \lesssim \|U_h^0\|_\gamma + \sum_{j=0}^n \tau \left\| F(U_h^j) \right\|_\gamma, \end{aligned} \quad (4.20)$$

where we used the representation $\varphi_1(\tau L_h) = \frac{1}{\tau} \int_0^\tau e^{sL_h} ds$, together with the boundedness of e^{tL_h} and $\mathcal{P}_h$ from Assumption 2.5(a),(b) and Remark 2.6. For the second term on the right-hand side of (4.20), we insert the term $F(U(t_j))$. By applying the estimates (3.9) and (3.14) in Lemma 3.5, we have

$$\begin{aligned} \left\| F(U_h^j) \right\|_\gamma & \leq \|f(u(t_j))\|_{\mathcal{Y}^{\gamma-1}} + \left\| f(u(t_j)) - f(u_h^j) \right\|_{\mathcal{Y}^{\gamma-1}} \\ & \leq C \left(\|u(t_j)\|_{\mathcal{Y}^\gamma} \right) + C \left(\|u(t_j)\|_{\mathcal{Y}^\gamma}, \|u_h^j\|_{\mathcal{Y}^\gamma} \right) \cdot \|u(t_j) - u_h^j\|_{\mathcal{Y}^0}^\alpha. \end{aligned} \quad (4.21)$$

with $\alpha > 0$ given by (3.15).

Substituting this estimate into (4.20) and noting that $U \in C([0, T]; \mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1})$ we deduce

$$\|U_h^{n+1}\|_\gamma \leq C_0 + \sum_{j=0}^n \tau C \left(\|U_h^j\|_\gamma \right) \cdot \|u(t_j) - u_h^j\|_{\mathcal{Y}^0}^\alpha, \quad (4.22)$$

where C_0 denotes a constant independent of τ , h , r , and the numerical solution bound.

4.4. The proof of Theorem 2.7

In (4.8) and (4.22), we have obtained estimates for $\|U(t_{n+1}) - U_h^{n+1}\|_0$ and $\|U_h^{n+1}\|_\gamma$. However, these bounds rely on the corresponding boundedness assumptions at the previous time steps.

To close the argument, we therefore employ a bootstrap procedure, for which we first state the following technical lemma.

Lemma 4.3 (Discrete bootstrap lemma). *Let $(a^n)_{n \geq 0}$ and $(b^n)_{n \geq 0}$ be nonnegative sequences such that*

$$a^{n+1} \leq A + \sum_{j=0}^n \tau C(b^j)(\delta + a^j), \quad (4.23)$$

$$b^{n+1} \leq B + \sum_{j=0}^n \tau C(b^j)(\delta + a^j)^\alpha, \quad (4.24)$$

$$a^0 \leq A \quad b^0 \leq B, \quad (4.25)$$

for $n \geq 0$, where $A, B, \delta \geq 0$, $\alpha > 0$, $\tau > 0$, and $C : [0, \infty) \rightarrow [0, \infty)$ is nondecreasing. Given $T > 0$, set

$$\tilde{C} := \sup_{0 \leq s \leq B+1} C(s).$$

If

$$\tilde{C}T \left(\delta + (A + \tilde{C}T\delta)e^{T\tilde{C}} \right)^\alpha \leq 1, \quad (4.26)$$

then for all integers n with $0 \leq n \leq T/\tau$,

$$b^n \leq B + 1, \quad a^n \leq (A + \tilde{C}T\delta)e^{n\tau\tilde{C}}. \quad (4.27)$$

Proof. We use induction on n . The claim is immediate for $n = 0$. Assume that

$$b^j \leq B + 1, \quad a^j \leq (A + \tilde{C}T\delta)e^{j\tau\tilde{C}}, \quad 0 \leq j \leq n.$$

Then $C(b^j) \leq \tilde{C}$ for $0 \leq j \leq n$, and hence

$$b^{n+1} \leq B + \tilde{C} \sum_{j=0}^n \tau (\delta + a^j)^\alpha \leq B + \tilde{C}T \left(\delta + (A + \tilde{C}T\delta)e^{T\tilde{C}} \right)^\alpha \leq B + 1,$$

by (4.26). Moreover,

$$a^{n+1} \leq A + \tilde{C} \sum_{j=0}^n \tau (\delta + a^j) \leq A + \tilde{C}T\delta + \tilde{C} \sum_{j=0}^n \tau a^j.$$

By the discrete Gronwall inequality,

$$a^{n+1} \leq (A + \tilde{C}T\delta)e^{(n+1)\tau\tilde{C}}.$$

This closes the induction and proves (4.27). $\square$

Going back to the estimate (4.8), to balance the high-frequency cut-off error and the finite element approximation error, we choose $r^{-\gamma} = h^{\ell_k} r^{\ell_k+1-\gamma}$, which leads to

$$r = h^{-\frac{\ell_k}{\ell_k+1}}, \quad \text{so that} \quad r^{-\gamma} = h^{\ell_k} r^{\ell_k+1-\gamma} = h^{\frac{\ell_k}{\ell_k+1}\gamma}.$$

Set

$$a^n = \|U(t_n) - U_h^n\|_0, \quad \text{and} \quad b^n = \|U_h^n\|_\gamma.$$

Then, the error estimate (4.8) together with the bound (4.22) implies that

$$a^{n+1} \leq C_0(\tau h^{\frac{\ell_k}{\ell_k+1}(\gamma-1)} + h^{\frac{\ell_k}{\ell_k+1}\gamma}) + \sum_{j=0}^n \tau \cdot C(b^j) \cdot \left(h^{\frac{\ell_k}{\ell_k+1}\gamma} + a^j \right), \quad (4.28)$$

$$b^{n+1} \leq C_0 + \sum_{j=0}^n \tau \cdot C(b^j) \cdot (a^j)^\alpha, \quad (4.29)$$

Moreover, by the projection error estimate in Lemma 3.2 and the boundedness of $\mathcal{P}_h$ from Assumption 2.5 and Remark 2.6, we have

$$a^0 = \|U(0) - \mathcal{P}_h U(0)\|_0 \leq C_0 h^\gamma \quad \text{and} \quad b^0 = \|\mathcal{P}_h U(0)\|_\gamma \leq C_0. \quad (4.30)$$

Here and below, C_0 denotes a generic constant independent of τ , h , r , and the a priori bound for the numerical solution. Thus a^n and b^n satisfy the inequalities (4.23)–(4.25) with

$$A = C_0(\tau h^{\frac{\ell_k}{\ell_{k+1}}(\gamma-1)} + h^{\frac{\ell_k}{\ell_{k+1}}\gamma}), \quad B = C_0, \quad \delta = h^{\frac{\ell_k}{\ell_{k+1}}\gamma}.$$

where the initial condition (4.25) is satisfied because $h^\gamma \leq h^{\frac{\ell_k}{\ell_{k+1}}\gamma}$ for $h \in (0, 1)$. Using the step-size condition from Theorem 2.7, namely, $\tau h^{\frac{\ell_k}{\ell_{k+1}}(\gamma-1)} \lesssim h^\mu$ for an arbitrarily small but fixed $\mu > 0$, we obtain

$$A \leq C_0(h^\mu + h^{\frac{\ell_k}{\ell_{k+1}}\gamma}).$$

Since $\mu > 0$ and $\frac{\ell_k}{\ell_{k+1}}\gamma > 0$, there exists $h_0 > 0$ such that

$$\tilde{C}T \left[h_0^{\frac{\ell_k}{\ell_{k+1}}\gamma} + \left(C_0(h_0^\mu + h_0^{\frac{\ell_k}{\ell_{k+1}}\gamma}) + \tilde{C}Th_0^{\frac{\ell_k}{\ell_{k+1}}\gamma} \right) e^{T\tilde{C}} \right]^\alpha \leq 1. \quad (4.31)$$

Hence, for $0 < h < h_0$ and $\tau \lesssim h^{\frac{\ell_k}{\ell_{k+1}}(1-\gamma)+\mu}$, the smallness condition (4.26) holds. By applying Lemma 4.3 and substituting the definitions of A and δ , we conclude that

$$\begin{aligned} \sup_{0 \leq n \leq T/\tau} \|U(t_n) - U_h^n\|_0 &\leq \left(C_0(\tau h^{\frac{\ell_k}{\ell_{k+1}}(\gamma-1)} + h^{\frac{\ell_k}{\ell_{k+1}}\gamma}) + \tilde{C}Th^{\frac{\ell_k}{\ell_{k+1}}\gamma} \right) e^{T\tilde{C}} \\ &= C_0 e^{T\tilde{C}} \cdot \tau h^{\frac{\ell_k}{\ell_{k+1}}(\gamma-1)} + (C_0 + \tilde{C}T) e^{T\tilde{C}} \cdot h^{\frac{\ell_k}{\ell_{k+1}}\gamma}. \end{aligned}$$

This completes the proof of Theorem 2.7.

5. Numerical experiments

In this section, we present two numerical experiments on two different domains, using quasi-uniform, shape-regular simplicial partitions. First, we consider the unit cube, for which Assumption 2.3 is satisfied for all $k \in \mathbb{N}$, see Remark 2.4. Second, we consider an equilateral triangle. In this case, $\theta = \frac{\pi}{3}$ in (2.13), and hence

$$\mathcal{Y}^{\ell_\Omega} \hookrightarrow H^{\ell_\Omega}(\Omega) \quad \text{with} \quad \ell_\Omega = 4 - \varepsilon,$$

for an arbitrarily small $\varepsilon > 0$. Therefore, Assumption 2.3 is fully justified for $k = 1, 2$, while for $k = 3$ it fails only by an arbitrarily small margin on quasi-uniform meshes.

We perform numerical experiments for polynomial degrees $k = 1, 2, 3$. In the triangular-domain case, the convergence rate for cubic elements is therefore expected to be reduced by an arbitrarily small amount compared with the ideal rate. Nevertheless, since this loss is negligible at the theoretical level, we still expect to observe essentially the predicted convergence rates in all experiments.

5.1. Implementation

We first discuss the implementation of the algorithm in Python. We first generate the meshes in `Gmsh` [20] and use `ADIOS4DOLFINx` [17] for writing and reading of the meshes and functions. We then use the `DOLFINx` environment [6] to assemble the mass and stiffness matrices $\mathbf{M}_h$ and $\mathbf{A}_h$ defined by

$$(\mathbf{M}_h)_{i,j} = \int_\Omega \varphi_i \varphi_j \, dx, \quad (\mathbf{A}_h)_{i,j} = \int_\Omega \nabla \varphi_i \cdot \nabla \varphi_j \, dx, \quad (5.1)$$

and store them as objects in the PETSc for Python library [3, 14], which allows for many options in the solution of the upcoming linear systems. We denote by $(\mathbf{u}_h^n, \mathbf{v}_h^n)$ the coefficient vectors

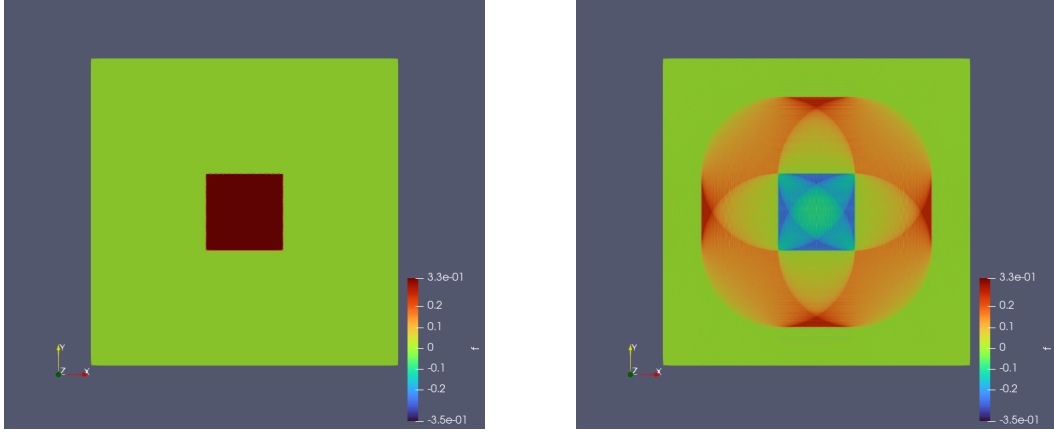


FIGURE 5.1. Left: Initial datum defined in (5.3). Right: Numerical approximation obtained from the reference solution u at end time $T = 0.25$. Both plots are obtained from the simulation with $k = 3$.

of $(u_h^n, v_h^n) \approx (u(t_n), u'(t_n))$, and note that the exponential integrator in (2.19) is equivalent to the exact solution $(\mathbf{u}_h, \mathbf{u}_h')$ at time $t_{n+1} = t_n + \tau$ of

$$\mathbf{M}_h \mathbf{u}_h''(t) = -\mathbf{A}_h \mathbf{u}_h(t) + \mathbf{f}_h^n, \quad t \in [t_n, t_n + \tau], \quad \mathbf{u}_h(t_n) = \mathbf{u}_h^n, \quad \mathbf{u}_h'(t_n) = \mathbf{v}_h^n, \quad (5.2)$$

where $\mathbf{f}_h^n$ denotes the load vector computed from $\mathcal{P}_h f(u_h(t_n))$, and the new approximations are defined as $\mathbf{u}_h^{n+1} = \mathbf{u}_h(t_{n+1})$ and $\mathbf{v}_h^{n+1} = \mathbf{u}_h'(t_{n+1})$. The solution of (5.2) is computed using a rational Krylov approximation as suggested in [21] and [24]. The code corresponding to the experiments in this section is made publicly available at

<https://github.com/BenjaminDoerich/ExpoFem-LowRegWave>.

5.2. Indicator function of a square in the unit square

In this subsection, we follow Section 7.2 in [12] and consider the semilinear wave equation defined on $\Omega = [0, 1]^2$, time interval $[0, 0.25]$, and $f(u) = 4 \sin(u)$ for the initial states $v^0 = 0$ and

$$u^0(x) = \begin{cases} 0.5, & x \in [0.375, 0.625]^2, \\ 0, & \text{else.} \end{cases} \quad (5.3)$$

We provide a plot of the initial state u^0 in Figure 5.1. Here, we have $u^0 \in H^{1/2-\varepsilon}(\Omega)$ for any $\varepsilon > 0$, and we are in the case of Theorem 2.7 with $\gamma < \frac{1}{2}$. For the experiments, we choose the scaling for ℓ_k defined in (2.25)

$$h \approx \frac{1}{T} C_{\text{scal}} \tau^{\frac{\ell_k+1}{\ell_k}} \quad (5.4)$$

with some factor C_{scal} to balance out the error contributions, cf. (2.26) and Remark 2.8, and to remain in a regime of a moderate number of degrees of freedom. This factor is the same for the numerical approximations as well as the reference solution, computed with some smaller time step size τ_{ref} and finer mesh parameter h_{ref} , see Table 5.1 for the precise data. The error $E(T)$ is always computed in the $L^2(\Omega) \times H^{-1}(\Omega)$ -norm at the final time $T = 0.25$ and scaled by the $L^2(\Omega) \times H^{-1}(\Omega)$ -norm of the reference solution.

We perform two different experiments for this domain. First, we study the convergence under the coupling (5.4), and we expect by Theorem 2.7 convergence of order $\tau^{1/2-\varepsilon}$ for any $\varepsilon > 0$. This is depicted in Figure 5.2 in the upper left plot, where the dashed lines indicate order $\frac{1}{2}$. The upper right plot shows the relation between the number of degrees of freedom and the error and intends to show the superior scaling for higher order polynomials, since fewer degrees of freedom are needed for $k = 2, 3$ to achieve the same error as linear elements at least for finer resolutions.

k	C_{scal}	τ_{ref}	h_{ref}	dofs_{ref}
1	10.8	$\sim 1 \cdot 10^{-3}$	$\sim 1 \cdot 10^{-3}$	986,035
2	7.2	$\sim 1 \cdot 10^{-3}$	$\sim 2 \cdot 10^{-3}$	1,116,329
3	7.2	$\sim 1 \cdot 10^{-3}$	$\sim 3 \cdot 10^{-3}$	1,526,890

TABLE 5.1. Choice of the parameters for the experiment corresponding to the initial datum defined in (5.3).

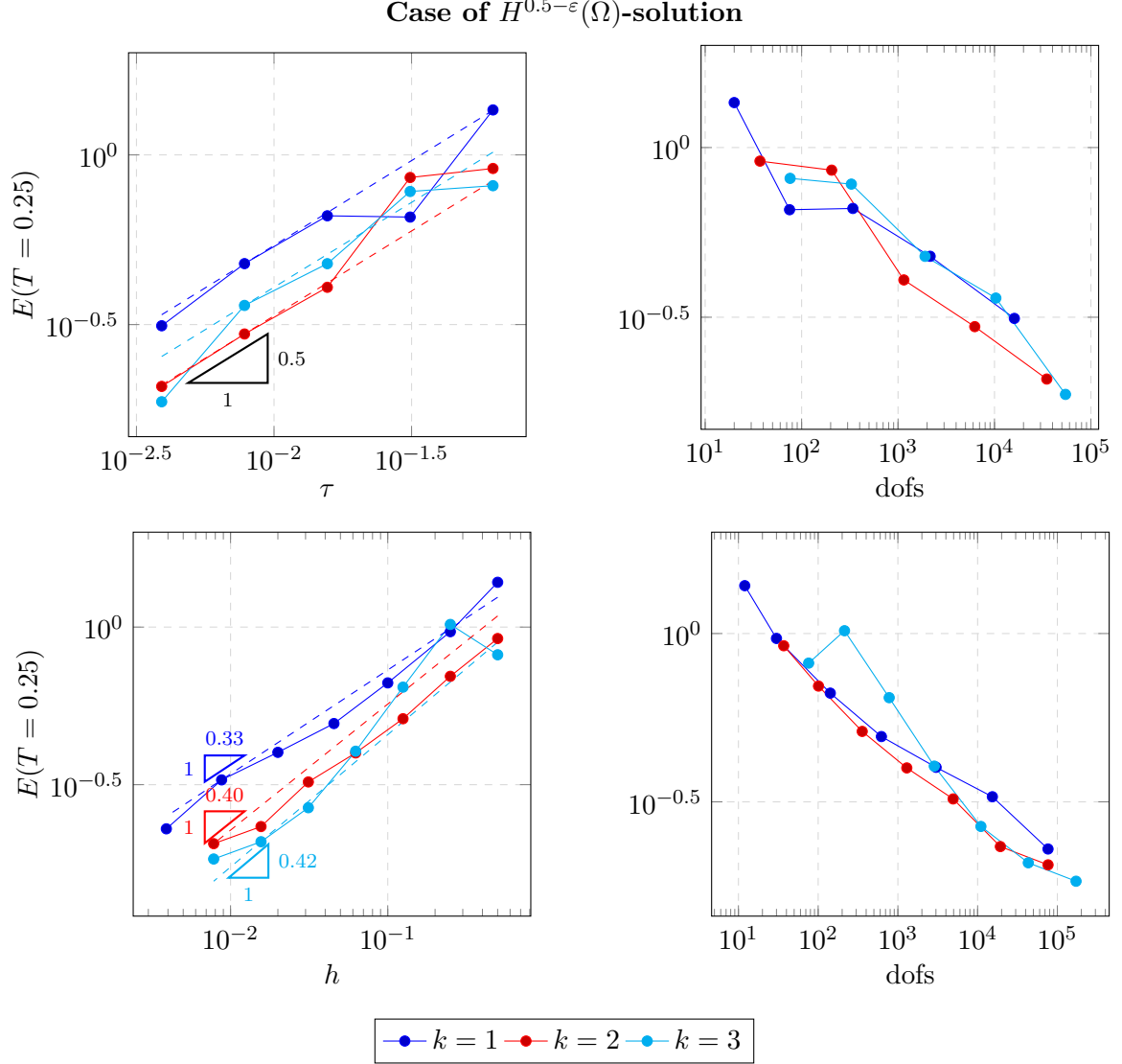


FIGURE 5.2. Error plots for the initial data (5.3). Top row: Equilibrated error $E(T = 0.25)$ via (5.4) for $k = 1, 2, 3$. In the left plot, the dashed lines indicate order $\tau^{\frac{1}{2}}$. Bottom row: Convergence in h for fixed τ . The dashed lines indicate the expected order of convergence $h^{\frac{\ell_k}{\ell_k+1} \frac{1}{2}}$ from Theorem 2.7.

Second, we compute solutions with a fixed $\tau_0 = 4\tau_{\text{ref}}$ and decreasing values of h in order to visualize the predicted order $h^{\frac{\ell_k}{\ell_k+1} \frac{1}{2}}$. This is shown in the lower left part of Figure 5.2. From our theory, we expect convergence with the orders $\frac{1}{3}, \frac{2}{5}, \frac{5}{12}$, which is again indicated with the dashed lines. As above also the scaling in the number of degrees of freedom shows a gain in the use of higher order polynomials.

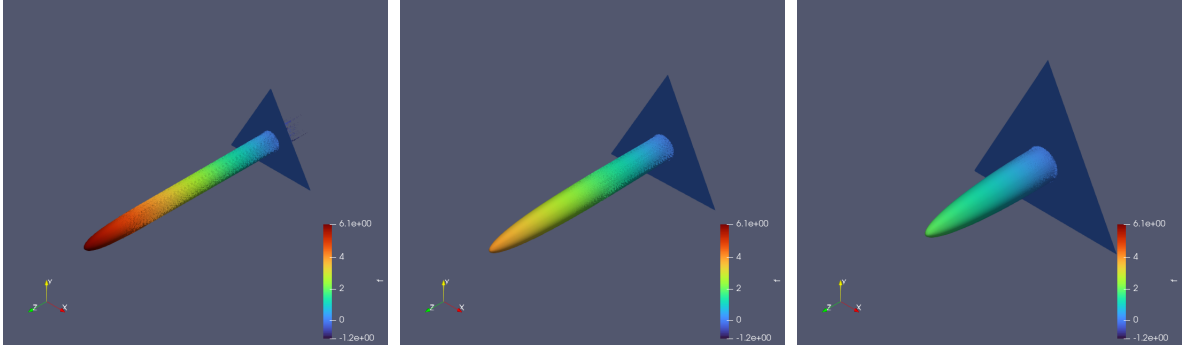


FIGURE 5.3. Initial datum defined in (5.5) for the parameters $\alpha = \frac{1}{10}, \frac{1}{4}, \frac{1}{2}$ (from left to right) computed with the L^2 -projection and cubic elements.

5.3. Scaled half sphere in an equilateral triangle

For the second experiment, we focus on a domain given by an equilateral triangle with center at the origin and corners lying on a circle of radius 1.2. We study the convergence on a scale of regularity parameters $\alpha \in (0, \frac{1}{2})$. i.e., we consider the family of initial values, cf. Figure 5.3,

$$u_\alpha^0(x) = \begin{cases} \mathcal{H}(r^2 - |x|^2)^\alpha, & \text{for } |x| \leq r, \\ 0, & \text{else where,} \end{cases} \quad (5.5)$$

for the radius $r = \frac{1}{4}$ and an amplitude $\mathcal{H} = 8$. Note that these initial values cover the range $u_\alpha^0 \in H^{1/2+\alpha-\varepsilon}(\Omega)$, see Corollary C.2 for a proof. As nonlinear term we use $f(u) = u^3$, and thus by Theorem 2.7 we expect convergence of order $\frac{1}{2} + \alpha - \varepsilon$. In the following, we present the results for $\alpha \in \{\frac{1}{10}, \frac{1}{4}, \frac{1}{2}\}$ on the time interval $[0, 0.3]$.

Remark 5.1. For $\alpha = \frac{1}{2}$, the initial data satisfy $(u^0, v^0) \in H^{1-\varepsilon}(\Omega) \times H^{-\varepsilon}(\Omega)$ for all $\varepsilon > 0$. In the two-dimensional case, the cubic nonlinearity $f(u) = u^3$ satisfies the growth condition (2.22). For sufficiently small initial data, one expects local solutions on a time interval $[0, T_{\max}]$, where $T_{\max} > 0$ depends on the size of the data, for instance on the amplitude parameter $\mathcal{H}$, and may become large when $\mathcal{H}$ is small.

Moreover, by finite speed of propagation, for points (x, t) such that $B(x, t) \subset \Omega$, the solution agrees with the corresponding full-space solution and hence admits the standard two-dimensional Duhamel representation formula; see, for example, [18]:

$$u(x, t) = \frac{1}{2} \oint_{B(x, t)} \frac{t u_\alpha^0(y) + t \nabla u_\alpha^0(y) \cdot (y - x)}{\sqrt{t^2 - |y - x|^2}} dy + \frac{1}{2} \int_0^t \oint_{B(x, t-s)} \frac{(t-s)^2 u(s, y)^3}{\sqrt{(t-s)^2 - |y - x|^2}} dy ds. \quad (5.6)$$

Here $\oint_{B(x, t)}$ denotes the average integral over the ball $B(x, t)$.

Using the expression of u_α^0 in (5.5), one formally finds that the linear contribution at the origin behaves like

$$\left| \oint_{B(0, t)} \frac{t \nabla u_\alpha^0(y) \cdot y}{\sqrt{t^2 - |y|^2}} dy \right| \sim \mathcal{H} |\ln |t - r||, \quad t \rightarrow r.$$

This indicates the formation of a logarithmic singularity in the linear part at $t = r$. Accordingly, one expects the full solution to lose boundedness in L^∞ near $t = r$, since the nonlinear term does not appear to provide a mechanism for cancelling the singular behaviour of the linear contribution. At the same time, the interaction with the nonlinear term may further amplify and complicate the singular profile. From a physical and analytical point of view, this reflects the combined effect of wave superposition (or energy focusing) and nonlinearity. This illustrates that a solution may remain in the low-regularity space $H^{1-\varepsilon}(\Omega) \times H^{-\varepsilon}(\Omega)$ while exhibiting singular pointwise behavior, which further motivates the use of low-regularity spaces in both the analysis and the numerical approximation.

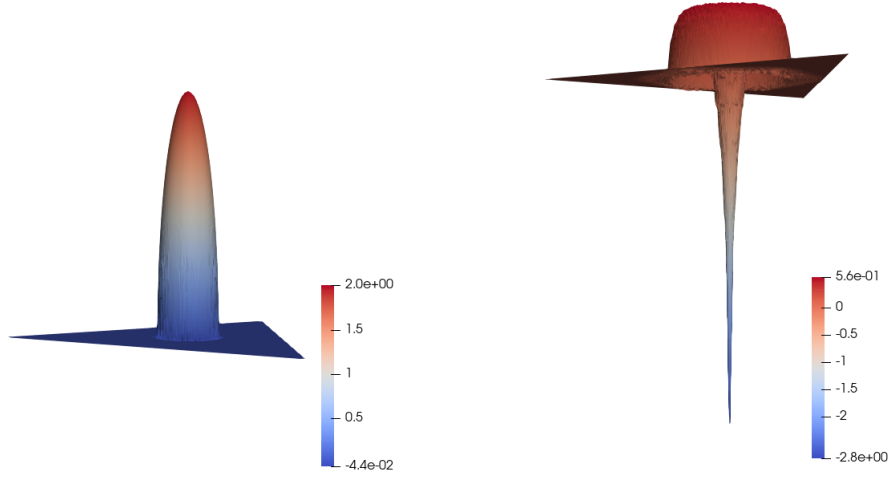


FIGURE 5.4. Plot of the initial datum (5.5) for $\alpha = \frac{1}{2}$ (left) as well as the numerical approximation at time $t \approx 0.25$ computed on a coarse mesh with quadratic elements.

This formal prediction is consistent with our numerical observations: in the right panel of Figure 5.4, as t approaches r , the solution develops a very sharp, high-amplitude spike near the focusing point. This numerically observed peak strongly suggests the onset of the singular behavior described above.

We proceed as in the previous experiment and first consider the convergence under the coupling (5.4) and reference solution with data from Table 5.2.

k	C_{scal}	τ_{ref}	h_{ref}	dofs_{ref}
1	15	$\sim 1 \cdot 10^{-3}$	$\sim 2 \cdot 10^{-3}$	623,286
2	10	$\sim 1 \cdot 10^{-3}$	$\sim 4 \cdot 10^{-3}$	907,878
3	10	$\sim 1 \cdot 10^{-3}$	$\sim 5 \cdot 10^{-3}$	1,304,920

TABLE 5.2. Choice of the parameters for the experiment corresponding to the initial datum defined in (5.5).

We observe in the upper left plots of Figures 5.5, 5.6, and 5.7 that under the coupling condition (5.4) the convergence order roughly behaves as τ^γ with $\gamma = \frac{3}{5}, \frac{3}{4}, 1$, for $\alpha = \frac{1}{10}, \frac{1}{4}, \frac{1}{2}$, respectively (indicated by the dashed lines), as predicted. In the lower left plots we studied again the convergence in h and included dashed lines of order $h^{\frac{\ell_k}{\ell_k+1}\gamma}$.

Acknowledgments

We would like to thank Tim Buchholz for his valuable input on the DOLFINx and PETSc implementation of the finite element solver.

References

- [1] C. Bacuta, V. Nistor, and L. T. Zikatanov. Improving the rate of convergence of ‘high order finite elements’ on polygons and domains with cusps. *Numerische Mathematik*, 100(2):165–184, 2005.
- [2] G. A. Baker. Error estimates for finite element methods for second order hyperbolic equations. *SIAM journal on numerical analysis*, 13(4):564–576, 1976.

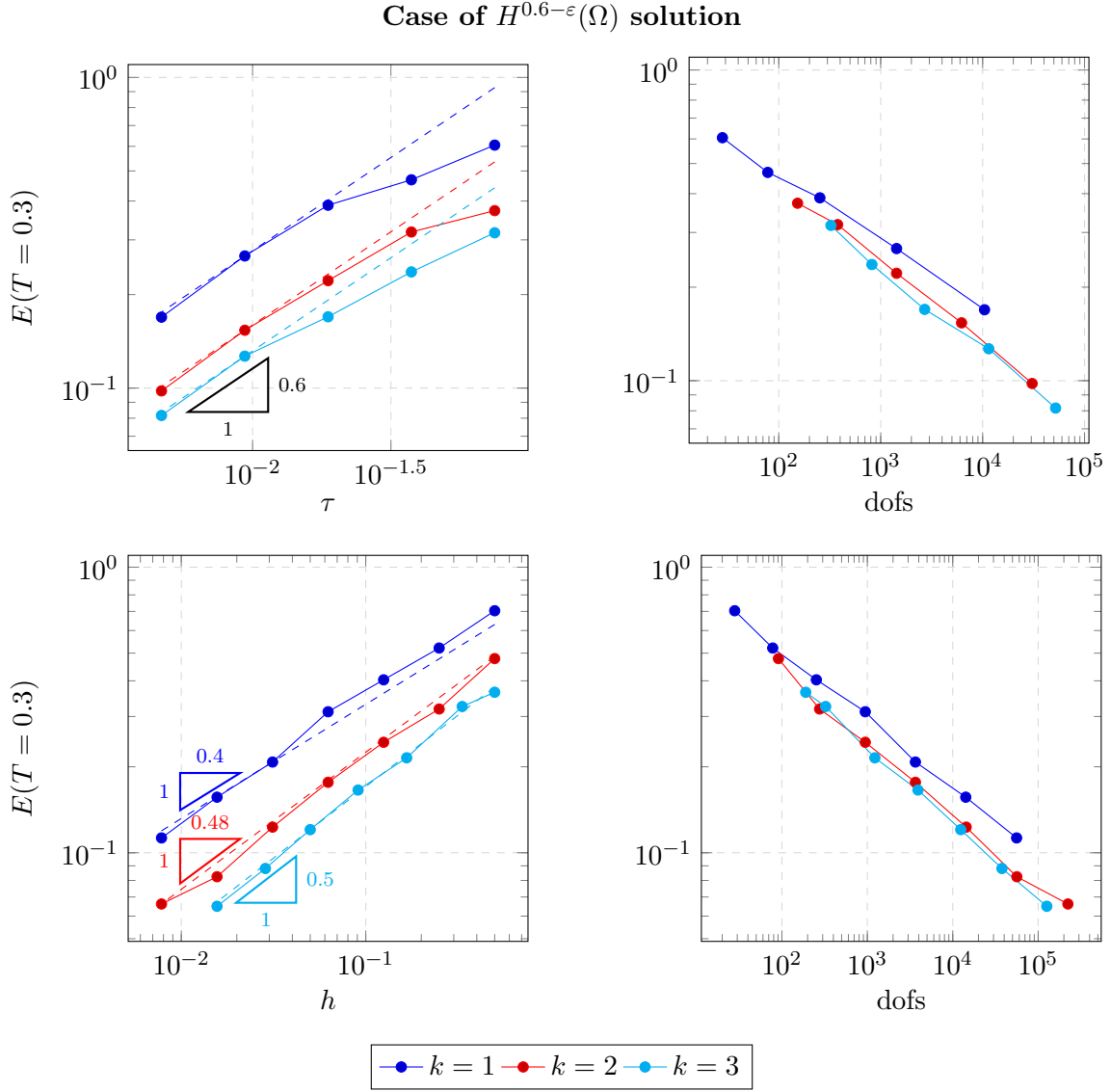


FIGURE 5.5. Error plots for the initial data (5.5) with $\alpha = 0.1$. Top row: Equilibrated error $E(T = 0.3)$ via (5.4) for $k = 1, 2, 3$. In the left plot, the dashed lines indicate order $\tau^{\frac{3}{5}}$. Bottom row: Convergence in h for fixed τ . The dashed lines indicate the expected order of convergence $h^{\frac{\ell_k}{\ell_k+1} \frac{3}{5}}$ from Theorem 2.7.

- [3] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, E. M. Constantinescu, L. Dalcin, S. Benson, A. Dener, et al. Petsc/tao users manual revision 3.24. Technical report, Argonne National Laboratory (ANL), Argonne, IL (United States), 09 2025.
- [4] R. E. Bank and H. Yserentant. On the h 1-stability of the l 2-projection onto finite element spaces. *Numerische Mathematik*, 126(2):361–381, 2014.
- [5] W. Bao, Y. Feng, and C. Su. Uniform error bounds of time-splitting spectral methods for the long-time dynamics of the nonlinear Klein-Gordon equation with weak nonlinearity. *Math. Comp.*, 91(334):811–842, 2022.
- [6] I. A. Baratta, J. P. Dean, J. S. Dokken, M. Habera, J. S. Hale, C. N. Richardson, M. E. Rognes, M. W. Scroggs, N. Sime, and G. N. Wells. Dolfinx: The next generation fenics problem solving environment, Dec. 2023.
- [7] M. Bourlard, M. Dauge, M.-S. Lubuma, and S. Nicaise. Coefficients of the singularities for elliptic boundary value problems on domains with conical points. iii: Finite element methods on polygonal domains. *SIAM journal on numerical analysis*, 29(1):136–155, 1992.
- [8] P. Brenner and V. Thomée. On rational approximations of semigroups. *SIAM Journal on Numerical Analysis*, 16(4):683–694, 1979.

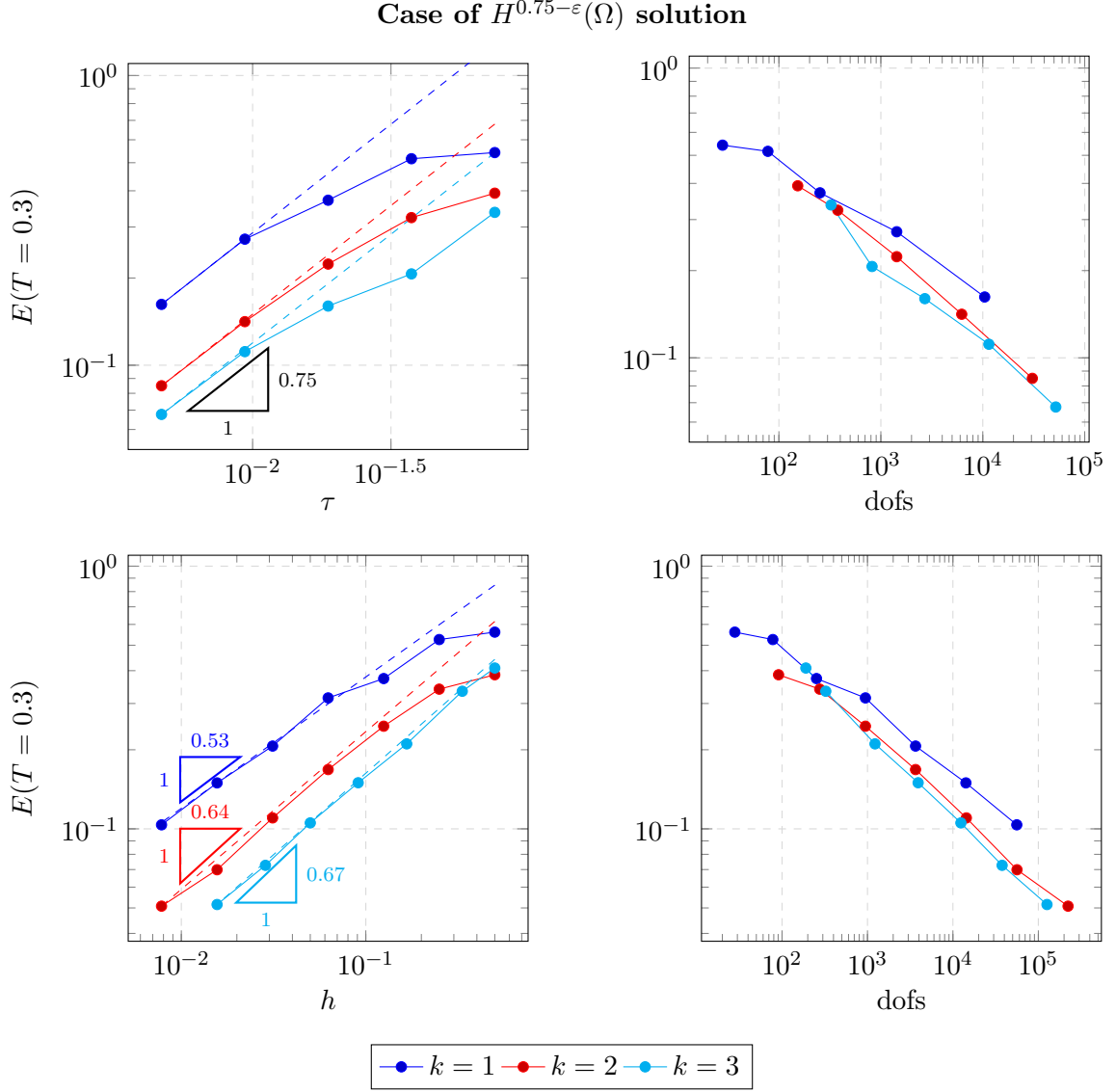


FIGURE 5.6. Error plots for the initial data (5.5) with $\alpha = 0.25$. Top row: Equilibrated error $E(T = 0.3)$ via (5.4) for $k = 1, 2, 3$. In the left plot, the dashed lines indicate order $\tau^{\frac{3}{4}}$. Bottom row: Convergence in h for fixed τ . The dashed lines indicate the expected order of convergence $h^{\frac{\ell_k}{\ell_k+1} \frac{3}{4}}$ from Theorem 2.7.

- [9] P. Brenner and V. Thomée. On rational approximations of groups of operators. *SIAM Journal on Numerical Analysis*, 17(1):119–125, 1980.
- [10] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*. Springer, 2008.
- [11] S. Buchholz, B. Dörich, and M. Hochbruck. On averaged exponential integrators for semilinear wave equations with solutions of low-regularity. *Partial Differ. Equ. Appl.*, 2(2):Paper No. 23, 27, 2021.
- [12] J. Cao, B. Li, Y. Lin, and F. Yao. Numerical approximation of discontinuous solutions of the semilinear wave equation. *SIAM J. Numer. Anal.*, 63(1):214–238, 2025.
- [13] C. Chen, J. Hong, C. Sim, and K. Sonwu. Energy and quadratic invariants preserving (EQUIP) multi-symplectic methods for Hamiltonian wave equations. *J. Comput. Phys.*, 418:109599, 18, 2020.
- [14] L. D. Dalcin, R. R. Paz, P. A. Kler, and A. Cosimo. Parallel distributed computing using python. *Advances in Water Resources*, 34(9):1124–1139, 2011. New Computational Methods and Software Tools.
- [15] M. Dauge. *Elliptic boundary value problems on corner domains: smoothness and asymptotics of solutions*. Springer, 2006.
- [16] L. Diening, J. Storn, and T. Tschierpel. On the sobolev and l^p -stability of the l^2 -projection. *SIAM Journal on Numerical Analysis*, 59(5):2571–2607, 2021.
- [17] J. S. Dokken. ADIOS4DOLFINx: A framework for checkpointing in FEniCS. *J. Open Source Softw.*, 9(96):6451, 2024.

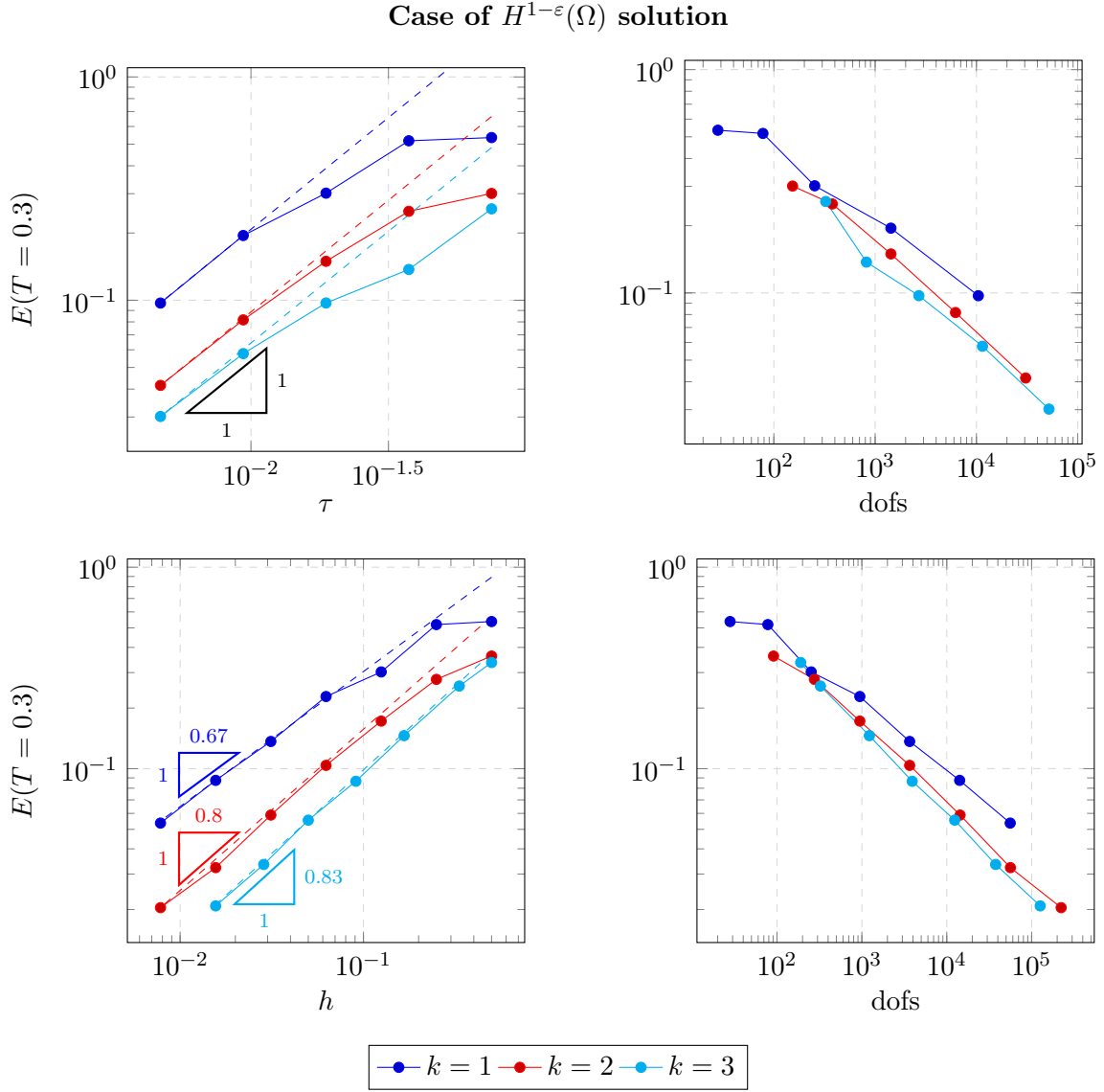


FIGURE 5.7. Error plots for the initial data (5.5) with $\alpha = 0.5$. Top row: Equilibrated error $E(T = 0.3)$ via (5.4) for $k = 1, 2, 3$. In the left plot, the dashed lines indicate order τ . Bottom row: Convergence in h for fixed τ . The dashed lines indicate the expected order of convergence $h^{\frac{\ell_k}{\ell_k+1}}$ from Theorem 2.7.

- [18] L. C. Evans. *Partial differential equations*, volume 19. American mathematical society, 2022.
- [19] Y. Feng, G. Maierhofer, and C. Wang. Explicit symmetric low-regularity integrators for the nonlinear schrödinger equation. *SIAM Journal on Scientific Computing*, 47(4):A2154–A2179, 2025.
- [20] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *Int. J. Numer. Meth. Engng.*, 79(11):1309–1331, 2009.
- [21] V. Grimm and M. Hochbruck. Rational approximation to trigonometric operators. *BIT*, 48(2):215–229, 2008.
- [22] M. Hochbruck and J. Leibold. An implicit-explicit time discretization scheme for second-order semilinear wave equations with application to dynamic boundary conditions. *Numer. Math.*, 147(4):869–899, 2021.
- [23] M. Hochbruck and C. Lubich. A Gautschi-type method for oscillatory second-order differential equations. *Numer. Math.*, 83(3):403–426, 1999.
- [24] M. Hochbruck, T. Pazur, A. Schulz, E. Thawinan, and C. Wieners. Efficient time integration for discontinuous Galerkin approximations of linear wave equations [Plenary lecture presented at the 83rd Annual GAMM Conference, Darmstadt, 26th–30th March, 2012]. *ZAMM Z. Angew. Math. Mech.*, 95(3):237–259, 2015.
- [25] M. Hofmanová and K. Schratz. An exponential-type integrator for the KdV equation. *Numer. Math.*, 136(4):1117–1137, 2017.
- [26] L. Ji, H. Li, A. Ostermann, and C. Su. Filtered lie-trotter splitting for the “good” boussinesq equation: Low regularity error estimates. *Mathematics of Computation*, 94(355):2345–2365, 2025.

- [27] V. A. Kondrat'ev. Boundary value problems for elliptic equations in domains with conical or angular points. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 16:209–292, 1967.
- [28] B. Li. Maximum-norm stability of the finite element method for the neumann problem in nonconvex polygons with locally refined mesh. *Mathematics of Computation*, 91(336):1533–1585, 2022.
- [29] B. Li and Y. Wu. An unfiltered low-regularity integrator for the kdv equation with solutions below h 1. *Foundations of Computational Mathematics*, pages 1–60, 2025.
- [30] H. Li. The w_p^1 stability of the ritz projection on graded meshes. *Mathematics of Computation*, 86(303):49–74, 2017.
- [31] C. Lubich. On the multistep time discretization of linear\newline initial-boundary value problems and their boundary integral equations. *Numerische Mathematik*, 67(3):365–389, 1994.
- [32] A. Ostermann, F. Rousset, and K. Schratz. Fourier integrator for periodic NLS: low regularity estimates via discrete Bourgain spaces. *J. Eur. Math. Soc. (JEMS)*, 25(10):3913–3952, 2023.
- [33] A. Ostermann and K. Schratz. Low regularity exponential-type integrators for semilinear Schrödinger equations. *Found. Comput. Math.*, 18(3):731–755, 2018.
- [34] F. Rousset and K. Schratz. A general framework of low regularity integrators. *SIAM J. Numer. Anal.*, 59(3):1735–1768, 2021.
- [35] M. Ruff. Error analysis of the strang splitting for the 3d semilinear wave equation with finite-energy data. *arXiv preprint arXiv:2503.13126*, 2025.
- [36] M. Ruff and R. Schnaubelt. Error analysis of the Lie splitting for semilinear wave equations with finite-energy solutions. *Discrete Contin. Dyn. Syst.*, 45(9):2969–3008, sep 2025.
- [37] K. Schratz, Y. Wang, and X. Zhao. Low-regularity integrators for nonlinear dirac equations. *Mathematics of Computation*, 90(327):189–214, 2021.
- [38] B. Wang and X. Wu. The formulation and analysis of energy-preserving schemes for solving high-dimensional nonlinear Klein-Gordon equations. *IMA J. Numer. Anal.*, 39(4):2016–2044, 2019.
- [39] C. Wulff and C. Evans. Runge-Kutta time semidiscretizations of semilinear PDEs with non-smooth data. *Numer. Math.*, 134(2):413–440, 2016.

A. Proofs of Lemma 3.4

To prove the estimate of the difference between two semigroups e^{tL} and e^{tL_h} in the weak norm $\mathcal{Y}^0 \times \mathcal{Y}^{-1}$, we need first show the following equivalence norm result:

Lemma A.1 (Equivalence between the discrete and continuous $\mathcal{Y}^{-1}$ norms). *Under the Assumption 2.5(a), there exist constants $c, C > 0$, independent of h , such that*

$$c \|w_h\|_{\mathcal{Y}^{-1}} \leq \|A_h^{-1/2} w_h\|_{\mathcal{Y}^0} \leq C \|w_h\|_{\mathcal{Y}^{-1}} \quad \forall w_h \in X_h^k.$$

Proof. For $w_h \in X_h^k$, define

$$\|w_h\|_{-1,h} := \sup_{\substack{v_h \in X_h^k \\ \|v_h\|_{\mathcal{Y}^1}=1}} \langle w_h, v_h \rangle_H.$$

Since $X_h^k \subset H_0^1(\Omega) = \mathcal{Y}^1$, Poincaré's inequality implies that

$$\|w_h\|_{-1,h} \sim \sup_{0 \neq v_h \in X_h^k} \frac{\langle w_h, v_h \rangle_H}{\|\nabla v_h\|_{\mathcal{Y}^0}}.$$

Let $z_h := A_h^{-1} w_h$. Then, by the definition of A_h ,

$$\langle w_h, v_h \rangle_H = \langle A_h z_h, v_h \rangle_H = \langle \nabla z_h, \nabla v_h \rangle_H \quad \forall v_h \in X_h^k.$$

Therefore,

$$\|w_h\|_{-1,h} \sim \sup_{0 \neq v_h \in X_h^k} \frac{\langle \nabla z_h, \nabla v_h \rangle_H}{\|\nabla v_h\|_{\mathcal{Y}^0}} = \|\nabla z_h\|_{\mathcal{Y}^0},$$

where the last identity follows from the Cauchy–Schwarz inequality and the choice $v_h = z_h$ if $z_h \neq 0$.

On the other hand,

$$\|A_h^{-1/2} w_h\|_{\mathcal{Y}^0}^2 = \langle A_h^{-1} w_h, w_h \rangle_H = \langle z_h, A_h z_h \rangle_H = \|\nabla z_h\|_{\mathcal{Y}^0}^2.$$

Hence

$$\|A_h^{-1/2} w_h\|_{\mathcal{Y}^0} \sim \|w_h\|_{-1,h}. \tag{A.1}$$

Since $X_h^k \subset H_0^1(\Omega) = \mathcal{Y}^1$, we immediately have

$$\|w_h\|_{-1,h} \leq \|w_h\|_{\mathcal{Y}^{-1}}. \quad (\text{A.2})$$

Conversely, for any $v \in \mathcal{Y}^1$,

$$\langle w_h, v \rangle_H = \langle w_h, \mathcal{P}_h v \rangle_H \leq \|w_h\|_{-1,h} \|\mathcal{P}_h v\|_{\mathcal{Y}^1} \leq C \|w_h\|_{-1,h} \|v\|_{\mathcal{Y}^1},$$

where we used that $\mathcal{P}_h$ is the H -orthogonal projection onto X_h^k and is bounded in $\mathcal{Y}^1$ by Assumption 2.5(a).

Taking the supremum over all $v \in \mathcal{Y}^1$ with $\|v\|_{\mathcal{Y}^1} = 1$ yields

$$\|w_h\|_{\mathcal{Y}^{-1}} \leq C \|w_h\|_{-1,h}.$$

Combining this with (A.1) and (A.2) proves the result. $\square$

Proof of (3.8) for $k \geq 2$. We denote

$$\begin{aligned} e^{tL} \Pi_r V(0) &=: V_r(t) = (v_r(t), \partial_t v_r(t)), \\ e^{tL_h} \mathcal{P}_h \Pi_r V(0) &=: V_{h,r}(t) = (v_{h,r}(t), \partial_t v_{h,r}(t)). \end{aligned}$$

By the definitions of e^{tL} and e^{tL_h} , the functions $v_r(t)$ and $v_{h,r}(t)$ satisfy

$$\langle \partial_{tt} v_r(t), w_h \rangle_H + \langle A v_r(t), w_h \rangle_H = 0, \quad (\text{A.3})$$

$$\langle \partial_{tt} v_{h,r}(t), w_h \rangle_H + \langle A_h v_{h,r}(t), w_h \rangle_H = 0, \quad (\text{A.4})$$

for all $w_h \in X_h^k$.

We next decompose the error $v_r - v_{h,r}$ as

$$v_r(t) - v_{h,r}(t) = v_r(t) - \mathcal{R}_h v_r(t) + \theta_h(t), \quad \theta_h(t) := \mathcal{R}_h v_r(t) - v_{h,r}(t). \quad (\text{A.5})$$

Hence,

$$\begin{aligned} \|e^{tL} \Pi_r V(0) - e^{tL_h} \mathcal{P}_h \Pi_r V(0)\|_0 &= \|v_r(t) - v_{h,r}(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \partial_t v_{h,r}(t)\|_{\mathcal{Y}^{-1}} \\ &\leq \|\theta_h(t)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(t)\|_{\mathcal{Y}^{-1}} + \|v_r(t) - \mathcal{R}_h v_r(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \mathcal{R}_h \partial_t v_r(t)\|_{\mathcal{Y}^{-1}}. \end{aligned} \quad (\text{A.6})$$

Since

$$\langle \partial_{tt} v_r(t), w_h \rangle_H = \langle \mathcal{P}_h \partial_{tt} v_r(t), w_h \rangle_H, \quad \langle A v_r(t), w_h \rangle_H = \langle A_h \mathcal{R}_h v_r(t), w_h \rangle_H,$$

subtracting (A.4) from (A.3), and using the fact that ∂_{tt} commutes with $\mathcal{P}_h$, yields

$$\langle \partial_{tt} \theta_h(t), w_h \rangle_H + \langle A_h \theta_h(t), w_h \rangle_H = -\langle \partial_{tt} \rho_h(t), w_h \rangle_H, \quad \forall w_h \in X_h^k,$$

where

$$\rho_h(t) := \mathcal{P}_h v_r(t) - \mathcal{R}_h v_r(t). \quad (\text{A.7})$$

Since $\theta_h(t) \in X_h^k$ and A_h is invertible on X_h^k , we choose $w_h = A_h^{-1} \partial_t \theta_h(t)$. Using also that A_h commutes with ∂_t , we obtain

$$\begin{aligned} \frac{d}{dt} \left(\|A_h^{-1/2} \partial_t \theta_h(t)\|_{\mathcal{Y}^0}^2 + \|\theta_h(t)\|_{\mathcal{Y}^0}^2 \right) &\leq \|A_h^{-1/2} \partial_{tt} \rho_h(t)\|_{\mathcal{Y}^0} \cdot \|A_h^{-1/2} \partial_t \theta_h(t)\|_{\mathcal{Y}^0} \\ &\leq \frac{1}{2} \|A_h^{-1/2} \partial_{tt} \rho_h(t)\|_{\mathcal{Y}^0}^2 + \frac{1}{2} \|A_h^{-1/2} \partial_t \theta_h(t)\|_{\mathcal{Y}^0}^2. \end{aligned}$$

An application of Gronwall's inequality therefore gives

$$\|A_h^{-1/2} \partial_t \theta_h(t)\|_{\mathcal{Y}^0}^2 + \|\theta_h(t)\|_{\mathcal{Y}^0}^2 \lesssim \|A_h^{-1/2} \partial_t \theta_h(0)\|_{\mathcal{Y}^0}^2 + \|\theta_h(0)\|_{\mathcal{Y}^0}^2 + \sup_{s \in [0,t]} \|A_h^{-1/2} \partial_{tt} \rho_h(s)\|_{\mathcal{Y}^0}^2.$$

By the equivalence between the discrete and continuous $\mathcal{Y}^{-1}$ -norms, we further obtain

$$\|\partial_t \theta_h(t)\|_{\mathcal{Y}^{-1}}^2 + \|\theta_h(t)\|_{\mathcal{Y}^0}^2 \lesssim \|\partial_t \theta_h(0)\|_{\mathcal{Y}^{-1}}^2 + \|\theta_h(0)\|_{\mathcal{Y}^0}^2 + \sup_{s \in [0,t]} \|\partial_{tt} \rho_h(s)\|_{\mathcal{Y}^{-1}}^2. \quad (\text{A.8})$$

We next estimate the last term on the right-hand side of (A.8). By the definition of ρ_h in (A.7) and the estimates (3.5) and (3.6), we deduce, for $k \geq 2$,

$$\|\partial_{tt} \rho_h(s)\|_{\mathcal{Y}^{-1}} \leq \|\partial_{tt} v_r(s) - \mathcal{P}_h \partial_{tt} v_r(s)\|_{\mathcal{Y}^{-1}} + \|\partial_{tt} v_r(s) - \mathcal{R}_h \partial_{tt} v_r(s)\|_{\mathcal{Y}^{-1}}$$

$$\lesssim h^{k+2} \|\partial_{tt} v_r(s)\|_{\mathcal{Y}^{k+1}}. \quad (\text{A.9})$$

Moreover, since $\partial_{tt} v_r(s) = -A v_r(s)$, we have

$$\|\partial_{tt} v_r(s)\|_{\mathcal{Y}^{k+1}} \lesssim \|v_r(s)\|_{\mathcal{Y}^{k+3}}.$$

Combining this with (A.9), and using the boundedness of e^{tL} together with the spectral localization of Π_r , we deduce

$$\|\partial_{tt} \rho_h(s)\|_{H^{-1}(\Omega)} \lesssim h^{k+2} \|v_r(s)\|_{\mathcal{Y}^{k+3}} \lesssim h^{k+2} r^{k+3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}. \quad (\text{A.10})$$

We now estimate the first two terms on the right-hand side of (A.8). Since

$$\theta_h(0) = \mathcal{R}_h v_r(0) - \mathcal{P}_h v_r(0), \quad \partial_t \theta_h(0) = \mathcal{R}_h \partial_t v_r(0) - \mathcal{P}_h \partial_t v_r(0),$$

by (3.4), (3.5), and (3.6), we have

$$\begin{aligned} \|\theta_h(0)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(0)\|_{\mathcal{Y}^{-1}} &\leq \|v_r(0) - \mathcal{R}_h v_r(0)\|_{\mathcal{Y}^0} + \|v_r(0) - \mathcal{P}_h v_r(0)\|_{\mathcal{Y}^0} \\ &\quad + \|\partial_t v_r(0) - \mathcal{R}_h \partial_t v_r(0)\|_{\mathcal{Y}^{-1}} + \|\partial_t v_r(0) - \mathcal{P}_h \partial_t v_r(0)\|_{\mathcal{Y}^{-1}} \\ &\lesssim h^{k+1} \|v_r(0)\|_{\mathcal{Y}^{k+1}} + h^{k+2} \|\partial_t v_r(0)\|_{\mathcal{Y}^{k+1}}. \end{aligned} \quad (\text{A.11})$$

Noting that $(v_r(0), \partial_t v_r(0)) = \Pi_r V(0)$, we obtain from the Bernstein-type inequality in Lemma 3.1 that

$$\|v_r(0)\|_{\mathcal{Y}^{k+1}} \lesssim r^{k+1-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}, \quad \|\partial_t v_r(0)\|_{\mathcal{Y}^{k+1}} \lesssim r^{k+2-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}.$$

Therefore, combining this with (A.11), and using the constraint $r^{-1} \leq h$, we obtain

$$\begin{aligned} \|\theta_h(0)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(0)\|_{\mathcal{Y}^{-1}} &\lesssim \left(h^{k+1} r^{k+1-\gamma} + h^{k+2} r^{k+2-\gamma} \right) \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}} \\ &\lesssim h^\gamma \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}. \end{aligned} \quad (\text{A.12})$$

Combining (A.8), (A.10), and (A.12), we deduce

$$\|\partial_t \theta_h(t)\|_{\mathcal{Y}^{-1}} + \|\theta_h(t)\|_{\mathcal{Y}^0} \lesssim (h^{k+2} r^{k+3-\gamma} + h^\gamma) \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}. \quad (\text{A.13})$$

Recalling the decomposition of the error $v_r - v_{h,r}$ in (A.5), we have

$$\begin{aligned} \|e^{tL} \Pi_r V(0) - e^{tL_h} \mathcal{P}_h \Pi_r V(0)\|_0 &\leq \|v_r(t) - \mathcal{R}_h v_r(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \mathcal{R}_h \partial_t v_r(t)\|_{\mathcal{Y}^{-1}} \\ &\quad + \|\theta_h(t)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(t)\|_{\mathcal{Y}^{-1}}. \end{aligned} \quad (\text{A.14})$$

The first two terms can be estimated in the same way as above by using (3.4), (3.6), and the boundedness of e^{tL} :

$$\begin{aligned} &\|v_r(t) - \mathcal{R}_h v_r(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \mathcal{R}_h \partial_t v_r(t)\|_{\mathcal{Y}^{-1}} \\ &\lesssim h^{k+1} \|v_r(t)\|_{\mathcal{Y}^{k+1}} + h^{k+2} \|\partial_t v_r(t)\|_{\mathcal{Y}^{k+1}} \\ &\lesssim h^\gamma \|e^{tL} \Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}} \lesssim h^\gamma \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}. \end{aligned} \quad (\text{A.15})$$

Together with (A.13) and (A.14), this implies

$$\|e^{tL} \Pi_r V(0) - e^{tL_h} \mathcal{P}_h \Pi_r V(0)\|_0 \lesssim (h^{k+2} r^{k+3-\gamma} + h^\gamma) \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}},$$

which proves (3.8) for $k \geq 2$. $\square$

Proof of (3.8) for $k = 1$. When $k = 1$, the $\mathcal{Y}^{-1}$ -estimate for the Ritz projection is no longer optimal. Therefore, unlike the case $k \geq 2$, we need to reformulate the argument so as to rely on $\mathcal{Y}^0$ -estimates for the Ritz projection. To this end, we follow the argument in [2]. We continue to use the notation

$$\begin{aligned} e^{tL} \Pi_r V(0) &=: V_r(t) = (v_r(t), \partial_t v_r(t)), \\ e^{tL_h} \mathcal{P}_h \Pi_r V(0) &=: V_{h,r}(t) = (v_{h,r}(t), \partial_t v_{h,r}(t)), \end{aligned}$$

and the decomposition

$$v_r(t) - v_{h,r}(t) = v_r(t) - \mathcal{R}_h v_r(t) + \theta_h(t), \quad \theta_h(t) := \mathcal{R}_h v_r(t) - v_{h,r}(t). \quad (\text{A.16})$$

Then

$$\begin{aligned} \|e^{tL}\Pi_r V(0) - e^{tL_h}\mathcal{P}_h\Pi_r V(0)\|_0 &\leq \|\theta_h(t)\|_{\mathcal{Y}^0} + \|\partial_t\theta_h(t)\|_{\mathcal{Y}^{-1}} \\ &\quad + \|v_r(t) - \mathcal{R}_h v_r(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \mathcal{R}_h \partial_t v_r(t)\|_{\mathcal{Y}^{-1}}. \end{aligned} \quad (\text{A.17})$$

As in the case $k \geq 2$, the error term θ_h satisfies

$$\langle \partial_{tt}\theta_h(t), w_h \rangle_H + \langle A_h \theta_h(t), w_h \rangle_H = -\langle \partial_{tt}\rho_h(t), w_h \rangle_H, \quad \forall w_h \in X_h^1, \quad (\text{A.18})$$

where $\rho_h(t) := \mathcal{P}_h v_r(t) - \mathcal{R}_h v_r(t)$. Define

$$\Theta_h(t) := \int_0^t \theta_h(s) ds.$$

Integrating (A.18) from 0 to t , we obtain

$$\langle \partial_t \theta_h(t), w_h \rangle_H + \langle A_h \Theta_h(t), w_h \rangle_H = -\langle \partial_t \rho_h(t) - \partial_t \rho_h(0), w_h \rangle_H + \langle \partial_t \theta_h(0), w_h \rangle_H \quad (\text{A.19})$$

for all $w_h \in X_h^1$.

Choosing $w_h = \theta_h(t) = \partial_t \Theta_h(t)$, we obtain

$$\frac{1}{2} \frac{d}{dt} \|\theta_h(t)\|_{\mathcal{Y}^0}^2 + \frac{1}{2} \frac{d}{dt} a(\Theta_h(t), \Theta_h(t)) \leq \left(\|\partial_t \rho_h(t) - \partial_t \rho_h(0)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(0)\|_{\mathcal{Y}^0} \right) \|\theta_h(t)\|_{\mathcal{Y}^0}.$$

Hence, by Gronwall's inequality,

$$\begin{aligned} \|\theta_h(t)\|_{\mathcal{Y}^0} + \|\nabla \Theta_h(t)\|_{\mathcal{Y}^0} \\ \lesssim \|\partial_t \rho_h(0)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(0)\|_{\mathcal{Y}^0} + \|\theta_h(0)\|_{\mathcal{Y}^0} + \sup_{s \in [0, t]} \|\partial_t \rho_h(s)\|_{\mathcal{Y}^0}, \end{aligned} \quad (\text{A.20})$$

where we used that $\Theta_h(0) = 0$.

We now estimate the terms on the right-hand side of (A.20). By the error estimate (3.4) in Lemma 3.3,

$$\|\partial_t \rho_h(t)\|_{\mathcal{Y}^0} \leq \|(\text{id} - \mathcal{P}_h) \partial_t v_r(t)\|_{\mathcal{Y}^0} + \|(\text{id} - \mathcal{R}_h) \partial_t v_r(t)\|_{\mathcal{Y}^0} \lesssim h^2 \|\partial_t v_r(t)\|_{\mathcal{Y}^2}. \quad (\text{A.21})$$

Similarly, since $\partial_t \theta_h(0) = \mathcal{R}_h \partial_t v_r(0) - \mathcal{P}_h \partial_t v_r(0)$, we have

$$\|\partial_t \theta_h(0)\|_{\mathcal{Y}^0} \leq \|(\text{id} - \mathcal{R}_h) \partial_t v_r(0)\|_{\mathcal{Y}^0} + \|(\text{id} - \mathcal{P}_h) \partial_t v_r(0)\|_{\mathcal{Y}^0} \lesssim h^2 \|\partial_t v_r(0)\|_{\mathcal{Y}^2}, \quad (\text{A.22})$$

and, since $\theta_h(0) = \mathcal{R}_h v_r(0) - \mathcal{P}_h v_r(0)$,

$$\|\theta_h(0)\|_{L^2(\Omega)} \lesssim h^2 \|v_r(0)\|_{\mathcal{Y}^2}. \quad (\text{A.23})$$

Using the boundedness of e^{tL} and the spectral localization of Π_r , we have

$$\|\partial_t v_r(t)\|_{\mathcal{Y}^2} \lesssim r^{3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}, \quad \|v_r(t)\|_{\mathcal{Y}^2} \lesssim r^{2-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}.$$

Substituting these bounds into (A.21), (A.22), and (A.23), and noting that $r \geq 1$, we arrive at

$$\sup_{s \in [0, t]} \|\partial_t \rho_h(s)\|_{\mathcal{Y}^0} + \|\partial_t \rho_h(0)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(0)\|_{\mathcal{Y}^0} + \|\theta_h(0)\|_{\mathcal{Y}^0} \lesssim h^2 r^{3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}.$$

Hence, by (A.20),

$$\|\theta_h(t)\|_{\mathcal{Y}^0} \lesssim h^2 r^{3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}. \quad (\text{A.24})$$

Next, we estimate $\partial_t \theta_h(t)$ in $\mathcal{Y}^{-1}$. By (A.19) and the norm equivalence in Lemma A.1,

$$\begin{aligned} \|\partial_t \theta_h(t)\|_{\mathcal{Y}^{-1}} &\lesssim \sup_{\substack{w_h \in X_h^1 \\ \|w_h\|_{\mathcal{Y}^1}=1}} \langle \partial_t \theta_h(t), w_h \rangle_H \\ &= \sup_{\substack{w_h \in X_h^1 \\ \|w_h\|_{\mathcal{Y}^1}=1}} \left(-\langle A_h \Theta_h(t), w_h \rangle_H - \langle \partial_t \rho_h(t) - \partial_t \rho_h(0), w_h \rangle_H + \langle \partial_t \theta_h(0), w_h \rangle_H \right) \\ &\lesssim \|\nabla \Theta_h(t)\|_{\mathcal{Y}^0} + \|\partial_t \rho_h(t)\|_{\mathcal{Y}^0} + \|\partial_t \rho_h(0)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(0)\|_{\mathcal{Y}^0}. \end{aligned} \quad (\text{A.25})$$

Combining (A.20), (A.21), (A.22), and (A.25), we infer

$$\|\theta_h(t)\|_{\mathcal{Y}^0} + \|\partial_t \theta_h(t)\|_{\mathcal{Y}^{-1}} \lesssim h^2 r^{3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}. \quad (\text{A.26})$$

Finally, applying the error estimate (3.4), we obtain

$$\begin{aligned} & \|v_r(t) - \mathcal{R}_h v_r(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \mathcal{R}_h \partial_t v_r(t)\|_{\mathcal{Y}^{-1}} \\ & \leq \|v_r(t) - \mathcal{R}_h v_r(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \mathcal{R}_h \partial_t v_r(t)\|_{\mathcal{Y}^0} \lesssim h^2 (\|v_r(t)\|_{\mathcal{Y}^2} + \|\partial_t v_r(t)\|_{\mathcal{Y}^2}). \end{aligned}$$

Using again the boundedness of e^{tL} and the spectral localization of Π_r , we obtain

$$\|v_r(t)\|_{\mathcal{Y}^2} \lesssim r^{2-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}, \quad \|\partial_t v_r(t)\|_{\mathcal{Y}^2} \lesssim r^{3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}.$$

Therefore,

$$\begin{aligned} \|v_r(t) - \mathcal{R}_h v_r(t)\|_{\mathcal{Y}^0} + \|\partial_t v_r(t) - \mathcal{R}_h \partial_t v_r(t)\|_{\mathcal{Y}^{-1}} & \lesssim (h^2 r^{2-\gamma} + h^2 r^{3-\gamma}) \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}} \\ & \lesssim h^2 r^{3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}}, \end{aligned} \quad (\text{A.27})$$

where we used $r \geq 1$.

Finally, combining (A.17), (A.26), and (A.27), we conclude that

$$\|e^{tL} \Pi_r V(0) - e^{tL_h} \mathcal{P}_h \Pi_r V(0)\|_0 \lesssim h^2 r^{3-\gamma} \|\Pi_r V(0)\|_{\mathcal{Y}^\gamma \times \mathcal{Y}^{\gamma-1}},$$

which proves (3.8) for $k = 1$. $\square$

B. Proofs of the estimates for the nonlinear terms

PROOF OF (3.9): Since $\gamma \in (0, 1]$, it follows from (2.7) and the Sobolev embedding $H_0^{1-\gamma} \hookrightarrow L^{\frac{2d}{d-2(1-\gamma)}}$ that $L^{\frac{2d}{d+2(1-\gamma)}} \hookrightarrow (H_0^{1-\gamma})^* = H^{\gamma-1}$. Therefore,

$$\|f(u)\|_{\mathcal{Y}^{\gamma-1}} \lesssim \|f(u)\|_{H^{\gamma-1}} \lesssim \|f(u)\|_{L^{\frac{2d}{d+2(1-\gamma)}}} \lesssim \|1 + |u|^{\xi_\gamma}\|_{L^{\frac{2d}{d+2(1-\gamma)}}} \lesssim 1 + \|u\|_{L^{\frac{2d\xi_\gamma}{d+2(1-\gamma)}}}^{\xi_\gamma}.$$

Moreover, by the Sobolev embedding $\mathcal{Y}^\gamma \hookrightarrow H^\gamma(\Omega) \hookrightarrow L^p(\Omega)$, where

$$p \begin{cases} < \infty, & \text{for } d = 2, \gamma = 1, \\ = \frac{2d}{d-2\gamma}, & \text{for } d = 2, 0 < \gamma < 1, \text{ or } d = 3, 0 < \gamma \leq 1, \end{cases} \quad (\text{B.1})$$

and by the assumption on ξ_γ in (2.23), it is straightforward to verify that

$$\frac{2d}{d-2\gamma} > \frac{2d\xi_\gamma}{d+2(1-\gamma)}.$$

Therefore,

$$\|f(u)\|_{\mathcal{Y}^{\gamma-1}} \lesssim 1 + \|u\|_{L^{\frac{2d\xi_\gamma}{d+2(1-\gamma)}}}^{\xi_\gamma} \lesssim 1 + \|u\|_{H^\gamma}^{\xi_\gamma} \lesssim 1 + \|u\|_{\mathcal{Y}^\gamma}^{\xi_\gamma},$$

which proves (3.9).

PROOF OF (3.10): We only prove the cases $d = 2, 0 < \gamma < 1$ and $d = 3, 0 < \gamma \leq 1$, since the case $d = 2, \gamma = 1$ is analogous. Note that

$$f(u) - f(v) = (u - v) \int_0^1 f'(\theta u + (1 - \theta)v) d\theta.$$

By (2.22),

$$|f(u) - f(v)| \lesssim \left(1 + |u|^{\xi_\gamma-1} + |v|^{\xi_\gamma-1}\right) |u - v|. \quad (\text{B.2})$$

As before, $L^{\frac{2d}{d+2(1-\gamma)}} \hookrightarrow (H_0^{1-\gamma})^* = H^{\gamma-1}$, and hence

$$\begin{aligned} \|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} & \lesssim \|f(u) - f(v)\|_{L^{\frac{2d}{d+2(1-\gamma)}}(\Omega)} \\ & \lesssim \left(1 + \| |u|^{\xi_\gamma-1} \|_{L^q(\Omega)} + \| |v|^{\xi_\gamma-1} \|_{L^q(\Omega)}\right) \|u - v\|_{L^p(\Omega)}, \end{aligned}$$

where $p = \frac{2d}{d-2\gamma}$ and

$$\frac{1}{q} = \frac{d+2(1-\gamma)}{2d} - \frac{1}{p} = \frac{1}{d},$$

so that $q = d$. By (2.23), $q(\xi_\gamma - 1) < p$. Therefore, by the Sobolev embedding $\mathcal{Y}^\gamma \hookrightarrow H^\gamma(\Omega) \hookrightarrow L^p(\Omega)$,

$$\begin{aligned} \|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} &\lesssim \left(1 + \|u\|_{L^p(\Omega)}^{\xi_\gamma-1} + \|v\|_{L^p(\Omega)}^{\xi_\gamma-1}\right) \|u - v\|_{L^p(\Omega)} \\ &\lesssim \left(1 + \|u\|_{\mathcal{Y}^\gamma}^{\xi_\gamma-1} + \|v\|_{\mathcal{Y}^\gamma}^{\xi_\gamma-1}\right) \|u - v\|_{\mathcal{Y}^\gamma}, \end{aligned}$$

which proves (3.10).

PROOF OF (3.11)–(3.13): Since the inequalities (3.11)–(3.13) can be proved in a similar way, we present only one representative case, namely (3.12).

Since $\mathcal{Y}^{-1} = H^{-1}(\Omega)$, the Sobolev embeddings yield $L^{p_1}(\Omega) \hookrightarrow H^{-1}(\Omega)$, where

$$p_1 = \begin{cases} 1+, & \text{for } d = 2, \\ \frac{2d}{d+2}, & \text{for } d = 3. \end{cases} \quad (\text{B.3})$$

Thus, by (B.2):

$$\begin{aligned} \|f(u) - f(v)\|_{\mathcal{Y}^{-1}} &\lesssim \|f(u) - f(v)\|_{L^{p_1}} \\ &\lesssim \left(1 + \| |u|^{\xi_\gamma-1} \|_{L^{q_1}(\Omega)} + \| |v|^{\xi_\gamma-1} \|_{L^{q_1}(\Omega)}\right) \cdot \|u - v\|_{L^2(\Omega)}, \end{aligned}$$

where Hölder's inequality is used with

$$\frac{1}{q_1} = \frac{1}{p_1} - \frac{1}{2}, \quad \text{that is } q_1 = \frac{2p_1}{2 - p_1}. \quad (\text{B.4})$$

Hence,

$$\|f(u) - f(v)\|_{\mathcal{Y}^{-1}} \lesssim \left(1 + \|u\|_{L^{q_1}(\xi_\gamma-1)}^{\xi_\gamma-1} + \|v\|_{L^{q_1}(\xi_\gamma-1)}^{\xi_\gamma-1}\right) \cdot \|u - v\|_{L^2}. \quad (\text{B.5})$$

When $d = 3$, combining (B.3) and (B.4) gives

$$q_1(\xi_\gamma - 1) = d(\xi_\gamma - 1) < \frac{2d}{d - 2\gamma}.$$

Recalling the Sobolev embedding $\mathcal{Y}^\gamma \hookrightarrow H^\gamma(\Omega) \hookrightarrow L^p(\Omega)$ with p given in (B.1), we obtain

$$\|f(u) - f(v)\|_{\mathcal{Y}^{-1}} \lesssim \left(1 + \|u\|_{H^\gamma}^{\xi_\gamma-1} + \|v\|_{H^\gamma}^{\xi_\gamma-1}\right) \cdot \|u - v\|_{L^2} \lesssim C\left(\|u\|_{\mathcal{Y}^\gamma}, \|v\|_{\mathcal{Y}^\gamma}\right) \|u - v\|_{\mathcal{Y}^0}.$$

This proves (3.12) for $d = 3$; the case $d = 2$ can be verified analogously.

PROOF OF (3.14): From the Sobolev embedding $H_0^{1-\gamma} \hookrightarrow L^{\frac{2d}{d-2(1-\gamma)}}$ and the boundedness estimate in (B.2), we have

$$\begin{aligned} \|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} &\lesssim \|f(u) - f(v)\|_{H^{\gamma-1}} \lesssim \|f(u) - f(v)\|_{L^{\frac{2d}{d+2(1-\gamma)}}} \\ &\lesssim \left\| \left(1 + |u|^{\xi_\gamma-1} + |v|^{\xi_\gamma-1}\right) \cdot |u - v| \right\|_{L^{\frac{2d}{d+2(1-\gamma)}}} \\ &\leq \left\| \left(1 + |u|^{\xi_\gamma-1} + |v|^{\xi_\gamma-1}\right) \cdot |u - v|^{1-\alpha} \right\|_{L^{\theta_1}} \| |u - v|^\alpha \|_{L^{\theta_2}}, \end{aligned} \quad (\text{B.6})$$

for $\alpha \in (0, 1]$, where Hölder's inequality is applied in the last step with exponents satisfying

$$\frac{1}{\theta_1} + \frac{1}{\theta_2} = \frac{d + 2(1 - \gamma)}{2d} \quad \text{and} \quad \theta_2 = \frac{2}{\alpha}. \quad (\text{B.7})$$

Since

$$\left(1 + |u|^{\xi_\gamma-1} + |v|^{\xi_\gamma-1}\right) \cdot |u - v|^{1-\alpha} \lesssim 1 + |u|^{\xi_\gamma-\alpha} + |v|^{\xi_\gamma-\alpha},$$

we deduce from (B.6) and (B.7) that

$$\|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} \lesssim \left(1 + \|u\|_{L^{(\xi_\gamma-\alpha)\theta_1}}^{\xi_\gamma-\alpha} + \|v\|_{L^{(\xi_\gamma-\alpha)\theta_1}}^{\xi_\gamma-\alpha}\right) \|u - v\|_{L^2}^\alpha. \quad (\text{B.8})$$

We first consider the case $(d, \gamma) \neq (2, 1)$. Set

$$\eta_\gamma := \frac{d - 2\gamma}{2\gamma} \left(\frac{d + 2(1 - \gamma)}{d - 2\gamma} - \xi_\gamma \right) > 0,$$

where the positivity follows from the growth condition on ξ_γ . If $0 < \eta_\gamma \leq 1$, we choose $\alpha = \eta_\gamma$. A direct calculation using (B.7) gives

$$(\xi_\gamma - \alpha)\theta_1 = \frac{2d}{d - 2\gamma}.$$

Hence, by the Sobolev embedding $\mathcal{Y}^\gamma \hookrightarrow H^\gamma(\Omega) \hookrightarrow L^{\frac{2d}{d-2\gamma}}(\Omega)$, we deduce from (B.8) that

$$\|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} \lesssim C(\|u\|_{\mathcal{Y}^\gamma}, \|v\|_{\mathcal{Y}^\gamma})\|u - v\|_{\mathcal{Y}^0}^\alpha.$$

It remains, still with $(d, \gamma) \neq (2, 1)$, to consider the case $\eta_\gamma > 1$. We choose $\alpha = 1$. Then $\theta_1 = \frac{2d}{2-2\gamma}$, with the usual interpretation $\theta_1 = \infty$ if $\gamma = 1$. Moreover, the condition $\eta_\gamma > 1$ implies

$$\xi_\gamma < \frac{d+2-4\gamma}{d-2\gamma}, \quad (\xi_\gamma - 1)\theta_1 < \frac{2d}{d-2\gamma}.$$

Thus the same argument based on (B.8) and the Sobolev embedding above yields

$$\|f(u) - f(v)\|_{\mathcal{Y}^{\gamma-1}} \lesssim C(\|u\|_{\mathcal{Y}^\gamma}, \|v\|_{\mathcal{Y}^\gamma})\|u - v\|_{\mathcal{Y}^0}.$$

Finally, let $d = 2$ and $\gamma = 1$. We choose $\alpha = \frac{1}{2}$. Then $\theta_1 < \infty$, and, since $\xi_\gamma < \infty$,

$$(\xi_\gamma - \alpha)\theta_1 < \infty.$$

Using the two-dimensional Sobolev embedding $H_0^1(\Omega) \hookrightarrow L^p(\Omega)$ for every $p < \infty$, and applying (B.8), we again obtain

$$\|f(u) - f(v)\|_{\mathcal{Y}^0} \lesssim C(\|u\|_{\mathcal{Y}^1}, \|v\|_{\mathcal{Y}^1})\|u - v\|_{\mathcal{Y}^0}^{1/2}.$$

Combining the above cases proves (3.14).

C. A regularity result

Lemma C.1. *Let $0 < a < b$ and consider the annulus around 0 denoted by $A_{a,b} = B_b(0) \setminus \overline{B_a(0)} \subset \mathbb{R}^2$. Further, let $f \in W^{1,p}(A_{a,b})$ be radially symmetric. Then, it holds*

$$\|f\|_{H^s(A_{a,b})} \lesssim \|f\|_{W^{1,p}(A_{a,b})} \quad (\text{C.1})$$

for $s < \frac{3}{2} - \frac{1}{p}$ and $p \in (1, 2)$.

Proof. We only prove this for smooth functions f . To each f , we associate a profile φ such that $f(x) = \varphi(|x|)$ for all $x \in A_{a,b}$. A straightforward calculation, then shows that there are constants only depending on a and b such that

$$c_0\|f\|_{L^p(A_{a,b})} \leq \|\varphi\|_{L^p(a,b)} \leq C_0\|f\|_{L^p(A_{a,b})} \quad (\text{C.2a})$$

$$c_1\|\nabla f\|_{L^p(A_{a,b})} \leq \|\varphi'\|_{L^p(a,b)} \leq C_1\|\nabla f\|_{L^p(A_{a,b})}. \quad (\text{C.2b})$$

An interpolation argument for $p = 2$ in the first estimates in (C.2a) and (C.2b) then shows that there is a constant which only depends on a and b such that for $s \in (0, 1)$

$$c_s\|f\|_{H^s(A_{a,b})} \leq \|\varphi\|_{H^s(a,b)}. \quad (\text{C.3})$$

Thus, the claim follows from the one dimensional Sobolev embedding $W^{1,p}(a,b) \hookrightarrow H^s(a,b)$ for $s < \frac{3}{2} - \frac{1}{p}$ and the second estimates in (C.2a) and (C.2b). $\square$

Corollary C.2. *The functions defined in (5.5) defined for $\alpha \in (0, \frac{1}{2}]$ is in $H^s(\Omega)$ for $s < \frac{1}{2} + \alpha$.*

Proof. We first consider a partition of unity of the circle by ψ_1 and ψ_2 with ψ_1 compactly supported on an annulus $A_{a,1}$, for $\frac{1}{4} < a$, and equal to one for $|x| \geq \frac{1}{4}$ and ψ_2 vanishes for $|x| \leq \frac{1}{3}$. Then, we write

$$u_\alpha^0 = \psi_1 u_\alpha^0 + \psi_2 u_\alpha^0 = w_1 + w_2, \quad (\text{C.4})$$

and directly see that w_2 is a smooth function. For w_1 one can easily compute that $w_2 \in W^{1,p}(\Omega)$ for $p < \frac{1}{1-\alpha}$. Further, since it is compactly supported inside the annulus, we conclude by the

above lemma that

$$\|w_2\|_{H^s(\Omega)} = \|w_2\|_{H^s(A_{a,1})} \lesssim \|w_2\|_{W^{1,p}(A_{a,1})} = \|w_2\|_{W^{1,p}(\Omega)}. \quad (\text{C.5})$$

Finally, as $s < \frac{3}{2} - \frac{1}{p} < \frac{1}{2} + \alpha$, we obtain the claim. $\square$

INSTITUTE FOR APPLIED AND NUMERICAL MATHEMATICS, KARLSRUHE INSTITUTE OF TECHNOLOGY, ENGLERSTR. 2, 76131 KARLSRUHE, GERMANY

Email address: {jiachuan.cao,benjamin.doerich,marlis.hochbruck}@kit.edu

DEPARTMENT OF APPLIED MATHEMATICS, THE HONG KONG POLYTECHNIC UNIVERSITY, HUNG HOM, HONG KONG

Email address: buyang.li@polyu.edu.hk