Explicit Runge-Kutta schemes with hybrid high-order methods for the wave equation in first-order form

Alexandre Ern, Rekha Khot

Explicit Runge-Kutta schemes with hybrid high-order methods for the wave equation in first-order form*

Alexandre Ern† Rekha Khot‡

November 2, 2024

Abstract

We analyze the approximation of the acoustic wave equation in its first-order Friedrichs formulation by explicit Runge-Kutta (ERK) schemes in time combined with hybrid high-order (HHO) methods in space. We propose two general assumptions (I1)-(I2) for an interpolation operator to evaluate the consistency error, and establish energy-error estimates in the time-continuous setting. We give several examples of interpolation operators: the classical one in the HHO literature based on L^2-orthogonal projections and others from, or inspired from, the hybridizable discontinuous Galerkin (HDG) literature giving improved convergence rates on simplices. In the fully discrete analysis, the key observation is that it becomes crucial to bound the consistency error in space by means of the stabilization seminorm only. We formulate three abstract properties (A1)-(A3) to lead the analysis and prove that, under suitable CFL conditions for second- and third-order ERK schemes, the energy error converges optimally in time and quasi-optimally in space, with optimal rates recovered on simplicial meshes. The abstract foundations of our analysis should facilitate its application to other nonconforming hybrid methods such as HDG and weak Galerkin (WG) methods.

Mathematics Subject Classification: 35L05, 65M15, 65M60.

Keywords: Acoustic wave equation, hybrid high-order, hybridizable discontinuous-Galerkin, Friedrichs formulation, explicit Runge-Kutta, energy-error estimates, CFL conditions, interpolation operators.

1 Introduction

Wave propagation is encountered in a variety of physical phenomena, e.g., earthquakes and other seismic activities, ultrasound wave imaging, etc. In this work, we focus on the acoustic wave equation. Many numerical methods exist to approximate this equation, either in its original second-order formulation in time or as a first-order system in time. Here, we are concerned with the (well-known) first-order setting in time based on a Friedrichs-type formulation involving a skew-symmetric differential operator in space (namely, gradient and divergence operators). The resulting formulation involves two unknowns: a vector-valued dual variable and a scalar-valued primal variable.

There are many available schemes in the literature to discretize in space the acoustic wave equation in its Friedrichs formulation. The first possibility is to use conforming finite element methods (FEM) with fluctuation-based stabilization, e.g., subgrid viscosity, local projection stabilization or gradient-jump penalty (also called continuous interior penalty). Implicit Euler schemes combined with stabilized

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034255.

†CERMICS, Ecole nationale des ponts et chaussées, IP Paris, 77455 Marne-la-Vallee Cedex 2, France; SERENA, Centre Inria de Paris, 48 rue Barrault, 75647, France, Email: alexandre.ern@enpc.fr.

‡SERENA, Centre Inria de Paris, 48 rue Barrault, 75647, France; CERMICS, Ecole nationale des ponts et chaussées, IP Paris, 77455 Marne-la-Vallee Cedex 2, France, Email: rekha.khot@inria.fr.

conforming FEM are reviewed in [18, Chapter 77], following ideas originally introduced in [22] for subgrid stabilization and linear monotone operators. Moreover, explicit Runge-Kutta (ERK) schemes combined with stabilized conforming FEM for time-dependent Friedrichs systems are analyzed in [5] (see also [18, Chapter 78]). The second possibility for space discretization is to use virtual element methods (VEM), which can be viewed as a generalization of conforming FEM to polyhedral grids. A VEM discretization for the first-order form of the acoustic wave equation, based on a dG approximation for the primal variable and a VEM approximation for the dual variable along with a θ-scheme in time, is studied in [12]. The third possibility for space discretization is to use discontinuous Galerkin (dG) methods. ERK schemes combined with dG methods for system of conservation laws were developed in [10]. Moreover, space-time dG methods for time-dependent Friedrichs systems are investigated in [20, 11], and more specifically in [2] for the acoustic wave equation with point singularities. Notice also that stabilized conforming FEM and dG are amenable to a unified analysis in the context of ERK schemes, as already shown in [5].

The fourth possibility for space discretization is to use nonconforming hybrid methods, such as hybridizable discontinuous Galerkin (HDG), weak Galerkin (WG), or hybrid high-order (HHO) methods. These methods are closely interlinked, as highlighted in [7] (see also [6, Chapters 1&3]). Let us briefly comment on these methods in the context of elliptic problems. Both WG and HHO methods are formulated in terms of cell- and face-based polynomials which approximate the primal variable in the mesh cells and on its faces, respectively. The equal-order setting considers cell- and face-based polynomials having the same degree, whereas the cell-based polynomials are one order higher than the face-based polynomials in the mixed-order setting. Both HHO and WG methods reconstruct locally a discrete gradient from the local unknowns (called weak gradient in WG). The devising viewpoint for HDG methods is somewhat different, as additional cell unknowns are introduced to approximate the dual variable, and the stabilization is introduced via the numerical flux trace. However, as shown in [7], the HDG dual variable is nothing but the discrete gradient reconstructed from the local cell and face variables. Moreover, the numerical flux trace in HHO and WG methods can be explicitly identified in terms of the reconstructed gradient and the stabilization. Furthermore, we notice that nonconforming VEM is also closely related to HDG, WG, and HHO. For instance, the computable projection of the gradient of virtual functions can be expressed using the gradient reconstruction operator.

Nonconforming hybrid methods have already been used to approximate the acoustic wave equation in its Friedrichs formulation. For both HDG and HHO methods, previous work approximated the dual variable in a dG fashion (by using piecewise vector-valued polynomials) and the primal variable using cell and face unknowns. Implicit and explicit RK schemes for HDG methods are proposed in [24] and [25], respectively, whereas the convergence analysis is performed in [9] in the time-continuous case and in [21] for implicit schemes using continuous finite elements in time. Implicit and explicit RK schemes for HHO methods are proposed in [3], and the convergence analysis is performed in [4] in the time-continuous case. In the WG setting, we mention [28], where a dG method is used for the primal variable and a WG method for the dual variable, together with an implicit Euler scheme in time. Furthermore, the space semi-discrete analysis of the WG approximation of time-dependent abstract Friedrichs systems is discussed in [27]. Nonconforming hybrid methods offer some advantages with respect to dG methods. For implicit time-schemes, one can use static condensation to reduce substantially the number of unknowns. Moreover, interior penalty dG methods come with a lower bound on the stabilization weight, which is not the case for nonconforming hybrid methods. A third advantage in the context of nonlinear problems is that the integration of nonlinear behavior laws is only required at the quadrature nodes in the cells, but not on the faces [1].

The goal of the present work is to derive error estimates for the acoustic wave equation in a *fully discrete* setting, using second- and third-order ERK schemes (referred as ERK2 and ERK3, respectively) in time and HHO methods in space (ERK-HHO schemes in short). Owing to the above discussion, our results also apply when other nonconforming hybrid methods, such as HDG and WG, are employed for space discretization. To our knowledge, this is the first fully discrete analysis for nonconforming hybrid methods with explicit time-stepping schemes applied to the wave equation in its Friedrichs formulation. One of the main challenges in the analysis is to deal with the static coupling

between face and cell unknowns. This has already been done in [19], but only for the second-order formulation in time.

Following the ideas in [5], our fully discrete analysis hinges on energy estimates. ERK schemes are antidissipative in nature, that is, they produce energy at each time step, as originally emphasized in [23], and this energy production needs to be compensated by the dissipation of the stabilization in space. The main flavour of our results on ERK-HHO schemes is similar to those achieved for ERK-dG methods in [26, 5], namely energy-error estimates invoke a $\frac{4}{3}$-CFL condition for ERK2 and the usual CFL condition for ERK3. There is, however, an important difference between ERK-HHO and ERK-dG schemes, since we stabilize only the primal variable whereas both primal and dual variables are stabilized in ERK-dG. This in itself makes the present method more appealing as it reduces the amount of energy dissipation. In addition, we achieve higher convergence rates in space, $\mathcal{O}(h^{k+1})$, on simplices and the same convergence rates as ERK-dG, $\mathcal{O}(h^{k+\frac{1}{2}})$, on polyhedra.

Concerning the analysis, we mention three main contributions.

- To address the static coupling between face and cell unknowns, we formulate three abstract properties **(A1)**-**(A3)** (see Section 5.1) to lead the analysis, thereby facilitating the application of our results to other nonconforming hybrid methods.
- We highlight a nontrivial novelty in the analysis when treating the fully discrete case instead of the time-continuous setting (as in previous works). Indeed, in the fully discrete setting, it becomes crucial to bound the consistency error in space produced at each time step by means of the stabilization seminorm only, and not the full HHO norm. This subtle aspect of the analysis is perhaps not that well-known since previous works essentially focused on the time-continuous setting.
- As usual, the consistency error is defined through an interpolation operator. Here, we formulate two abstract assumptions **(I1)**-**(I2)** (see Section 4.1) on the interpolation operator to lead the analysis, again with the aim to facilitate the application to other nonconforming hybrid methods. Interestingly, we consider altogether five interpolation operators depending on the discrete setting (equal- vs. mixed-order, simplicial vs. polyhedral meshes), and some of these operators come from the HDG literature (see Table 1.1 below and Table 7.1 for more details; these operators are commonly known as HDG projections, but we call them HDG interpolation operators for uniformity). This illustrates the mutual benefits promoted by building bridges among methods.

HDG	HDG interpolation on simplices	[8]
HHO	HHO interpolation on polyhedra	[14]
HDG-HHO	HDG interpolation modified with HHO stabilization on simplices	Present
HDG$^+$	HDG interpolation modified for mixed-order on simplices	[16]
H$^+$	H$^+$ interpolation for dual variable combined with HHO interpolation for primal variable on polyhedra	[17]

Table 1.1: Names of the interpolation operators (left), their main usage (middle), and references (right)

The paper is organized as follows. Section 2 introduces the acoustic wave equation as a first-order PDE system. Section 3 provides basic tools such as mesh assumptions, discrete spaces, and discrete operators required to formulate the HHO method. Section 4 is devoted to the time-continuous error analysis and introduces the main tools to estimate the consistency error in space. Section 5 states our main assumptions and results on the time-discrete error analysis. Section 6 contains the proofs of the results from Section 5. In Sections 4 through 6, we focus on the equal-order setting with plain

least-squares (LS) stabilization. Finally, the mixed-order setting with other choices of stabilization is discussed in Section 7.

We use standard notation for the Lebesgue and Sobolev spaces, as well as for the Bochner–Sobolev spaces in the space-time setting. Boldface notation is used for vectors and vector-valued fields. For a weight function $\phi \in L^\infty(\Omega)$ taking positive values uniformly bounded from below away from zero, we introduce the shorthand notation $\|w\|_{L^2(\phi;\Omega)} := \|\phi^{\frac{1}{2}} w\|_{L^2(\Omega)}$ for all $w \in L^2(\Omega)$, together with a similar notation for vector-valued fields in $\boldsymbol{L}^2(\Omega)$.

2 Model problem

Let Ω be an open bounded polyhedral domain in $\mathbb{R}^d$ for $d \in \{2,3\}$ with Lipschitz boundary Γ. The first-order Friedrichs-type formulation of the acoustic wave equation defined on the space domain Ω and the time domain $J := (0, T_f)$ with a final time $T_f > 0$ consists of the coupled PDEs in $J \times \Omega$,

$$\rho \partial_t \boldsymbol{\sigma} - \nabla v = \mathbf{0}, \tag{2.1a}$$

$$\frac{1}{\kappa} \partial_t v - \nabla \cdot \boldsymbol{\sigma} = f, \tag{2.1b}$$

involving as unknowns the dual variable (or flux) $\boldsymbol{\sigma} : J \times \Omega \to \mathbb{R}^d$ and the primal variable (or velocity) $v : J \times \Omega \to \mathbb{R}$. The parameters are the source term, $f : J \times \Omega \to \mathbb{R}$, the bulk modulus, $\kappa > 0$, and the density, $\rho > 0$, both assumed (for simplicity) to be piecewise constant on a polyhedral partition of Ω. The wave speed is defined as $\mathfrak{c} := \sqrt{\frac{\kappa}{\rho}}$. The initial conditions are

$$\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0, \quad v(0) = v_0 \quad \text{in } \Omega, \tag{2.2}$$

with given data $\boldsymbol{\sigma}_0 : \Omega \to \mathbb{R}^d$ and $v_0 : \Omega \to \mathbb{R}$, and the boundary condition is (for simplicity)

$$v = 0 \quad \text{on } J \times \Gamma. \tag{2.3}$$

We set $\boldsymbol{\Sigma} := \boldsymbol{H}(\mathrm{div}, \Omega)$, $V_0 := H_0^1(\Omega)$, and $\underline{V}_0 := \boldsymbol{\Sigma} \times V_0 \subset \underline{L} := \boldsymbol{L}^2(\Omega) \times L^2(\Omega)$. Our convention is to underline pairs of functions composed of one dual variable and one primal variable. We assume that $f \in L^2(J; L^2(\Omega))$, $\underline{v} := (\boldsymbol{\sigma}, v) \in H^1(J; \underline{L}) \cap L^2(J; \underline{V}_0)$, and $\underline{v}_0 := (\boldsymbol{\sigma}_0, v_0) \in \underline{L}$. The space $\underline{L}$ is equipped with the usual L^2-inner product. We introduce the linear operators $\underline{\mathcal{M}} : \underline{L} \to \underline{L}$ and $\underline{\mathcal{B}} : \underline{V}_0 \to \underline{L}$ defined by

$$\underline{\mathcal{M}}(\underline{w}) := (\rho \boldsymbol{\tau}, \tfrac{1}{\kappa} w) \quad \forall \underline{w} \in \underline{L} \quad \text{and} \quad \underline{\mathcal{B}}(\underline{w}) := -(\nabla w, \nabla \cdot \boldsymbol{\tau}) \qquad \forall \underline{w} := (\boldsymbol{\tau}, w) \in \underline{V}_0. \tag{2.4}$$

The norms $\|\cdot\|_{\underline{L}}, \|\cdot\|_{\rho,\frac{1}{\kappa}}$ and $\|\cdot\|_{\frac{1}{\rho},\kappa}$ in $\underline{L}$ are induced by the inner products $\langle \cdot, \cdot \rangle_{\underline{L}}, \langle \cdot, \cdot \rangle_{\rho,\frac{1}{\kappa}} := \langle \underline{\mathcal{M}}(\cdot), \cdot \rangle_{\underline{L}}$, and $\langle \cdot, \cdot \rangle_{\frac{1}{\rho},\kappa} := \langle \underline{\mathcal{M}}^{-1}(\cdot), \cdot \rangle_{\underline{L}}$. With the above notation, the model problem (2.1) can be rewritten as to seek, for all $t \in \overline{J}$, $\underline{v}(t) := (\boldsymbol{\sigma}(t), v(t)) \in \underline{V}_0$ satisfying in $L^2(J)$,

$$\underline{\mathcal{M}}(\partial_t \underline{v}(t)) + \underline{\mathcal{B}}(\underline{v}(t)) = \underline{f}(t), \tag{2.5}$$

with $\underline{f}(t) := (\mathbf{0}, f(t))$, and the initial condition $\underline{v}(0) = \underline{v}_0$. It is well-known that the energy $E(t) := \frac{1}{2}\|\underline{v}(t)\|^2_{\rho,\frac{1}{\kappa}}$ satisfies the following balance law a.e. in J:

$$E(t) = E(0) + \int_0^t \langle \underline{f}(s), \underline{v}(s) \rangle_{\underline{L}} \, \mathrm{d}s. \tag{2.6}$$

If the source term vanishes, (2.6) implies energy conservation, i.e., $E(t) = E(0)$ a.e. in J.

3 Space discretization

This section presents the mesh assumptions, the local and global discrete spaces, and the discrete operators in the HHO setting. The HHO method was originally introduced in [13] for linear elasticity problems and in [14] for diffusion problems.

3.1 Mesh

Let $\{\mathcal{T}_h\}_{h>0}$ be a sequence of shape-regular polyhedral meshes of Ω [6, 15]. The set $\mathcal{F}_h$ contains the mesh faces, which is divided into the set of mesh interfaces $\mathcal{F}_h^i$ and the set of boundary faces $\mathcal{F}_h^\partial$. A generic cell is denoted by $T \in \mathcal{T}_h$ with diameter h_T, unit outward normal $\boldsymbol{n}_T$, and the set $\mathcal{F}_{\partial T}$ collects the mesh faces located at the boundary of T. We assume that each mesh $\mathcal{T}_h$ fits the partition of Ω into polyhedral subdomains associated with the coefficients ρ and κ. Thus, both coefficients are piecewise constant on $\mathcal{T}_h$. Their restriction to every $T \in \mathcal{T}_h$ is denoted by ρ_T and κ_T, and we set $\mathfrak{c}_T := \sqrt{\frac{\kappa_T}{\rho_T}}$.

In what follows, the inequality $a \le Cb$ for positive numbers a and b is often abbreviated as $a \lesssim b$. The value of C can be different at each occurence provided it is independent of the parameters $\rho, \kappa, \ell_\Omega := \mathrm{diam}(\Omega), T_f$, the mesh-size h, and (from Section 5 onwards) the time step τ; the value can depend on the mesh shape-regularity, the polynomial degree, and the space dimension.

3.2 Discrete spaces

To approximate $\boldsymbol{\sigma}$ locally, we consider the vector-valued d-variate polynomial space $\boldsymbol{\Sigma}_T^k := \mathbb{P}^k(T;\mathbb{R}^d)$ of degree $k \ge 0$ for all $T \in \mathcal{T}_h$, and define the dG space on the mesh $\mathcal{T}_h$ as

$$\boldsymbol{\Sigma}_{\mathcal{T}}^k := \bigtimes_{T\in\mathcal{T}_h} \boldsymbol{\Sigma}_T^k. \tag{3.1}$$

To approximate v locally, we consider discrete unknowns attached to the mesh cells and to the mesh faces. Let $k \ge 0$ (resp. $k' \in \{k, k+1\}$) be the polynomial degree associated with the face (resp. cell) unknowns. The HHO space $\hat{V}_h^k$ is defined as

$$\hat{V}_h^k := V_{\mathcal{T}}^{k'} \times V_{\mathcal{F}}^k, \qquad V_{\mathcal{T}}^{k'} := \bigtimes_{T\in\mathcal{T}_h} V_T^{k'}, \qquad V_{\mathcal{F}}^k := \bigtimes_{F\in\mathcal{F}_h} V_F^k, \tag{3.2}$$

where $V_T^{k'} := \mathbb{P}^{k'}(T;\mathbb{R})$ (resp. $V_F^k := \mathbb{P}^k(F;\mathbb{R})$) is composed of the scalar-valued d-variate (resp. $(d-1)$-variate) polynomials of degree at most k' (resp. k) restricted to the cell T (resp. face F). Combining everything together, we define the triple

$$\underline{\hat{V}}_h^k := \overbrace{\boldsymbol{\Sigma}_{\mathcal{T}}^k \times \underbrace{V_{\mathcal{T}}^{k'} \times V_{\mathcal{F}}^k}_{=:\hat{V}_h^k}}^{\underline{V}_{\mathcal{T}}^k :=}. \tag{3.3}$$

To impose the zero Dirichlet boundary condition, we define $V_{\mathcal{F}0}^k := \{v_F \in V_{\mathcal{F}}^k : v_F = 0 \quad \forall F \in \mathcal{F}_h^\partial\}$ and set $\underline{\hat{V}}_{h0}^k := \underline{V}_{\mathcal{T}}^k \times V_{\mathcal{F}0}^k = \boldsymbol{\Sigma}_{\mathcal{T}}^k \times \hat{V}_{h0}^k$ with the HHO subspace $\hat{V}_{h0}^k := V_{\mathcal{T}}^{k'} \times V_{\mathcal{F}0}^k$.

Consistently with the above notation, we underline pairs of discrete functions defined cellwise (composed of one dual variable and one primal variable), and we use a hat for HHO pairs of primal functions composed of one function defined cellwise and one function defined facewise. A generic element in $\underline{\hat{V}}_h^k$ is denoted by

$$\underline{\hat{w}}_h := (\overbrace{\boldsymbol{\tau}_{\mathcal{T}}, \underbrace{w_{\mathcal{T}}}_{=:\hat{w}_h}}^{\underline{w}_{\mathcal{T}} :=}, w_{\mathcal{F}}). \tag{3.4}$$

The local components of $\underline{\hat{w}}_h \in \underline{\hat{V}}_h^k$ attached to the cell $T \in \mathcal{T}_h$ and its faces $F \in \mathcal{F}_{\partial T}$ are denoted by

$$\underline{\hat{w}}_T := (\overbrace{\boldsymbol{\tau}_T, \underbrace{w_T}_{=:\hat{w}_T}}^{\underline{w}_T :=}, w_{\partial T}) \in \underline{\hat{V}}_T^k := \overbrace{\boldsymbol{\Sigma}_T^k \times \underbrace{V_T^{k'} \times V_{\partial T}^k}_{=:\hat{V}_T^k}}^{\underline{V}_T^k :=}, \tag{3.5}$$

where $w_{\partial T} := (w_F)_{F\in\mathcal{F}_{\partial T}}$ and $V_{\partial T}^k := \bigtimes_{F\in\mathcal{F}_{\partial T}} V_F^k$. We fix $k' = k$ until the end of Section 6, and discuss the mixed-order case $k' = k+1$ in Section 7.

Let Π_T^k (resp. $\Pi_{\partial T}^k$) be the $L^2(T)$-orthogonal (resp. $L^2(\partial T)$-orthogonal) projection onto V_T^k (resp. $V_{\partial T}^k$). Let $\Pi_{\mathcal{T}}^k$ (resp. $\Pi_{\mathcal{F}}^k$) be the piecewise L^2-orthogonal projection onto $V_{\mathcal{T}}^k$ (resp. $V_{\mathcal{F}}^k$) and let $\underline{\Pi}_{\mathcal{T}}^k := (\boldsymbol{\Pi}_{\mathcal{T}}^k, \Pi_{\mathcal{T}}^k)$. For later use, we recall that, for all $T \in \mathcal{T}_h$ and all $(\boldsymbol{\tau}, w) \in \boldsymbol{H}^{\ell_{\boldsymbol{\tau}}}(\Omega) \times H^{\ell_w}(\Omega)$ with $\ell_{\boldsymbol{\tau}} \in [0, k+1]$ and $\ell_w \in [0, k+1]$, the following holds:

$$\|\boldsymbol{\Pi}_T^k(\boldsymbol{\tau}) - \boldsymbol{\tau}\|_{\boldsymbol{L}^2(T)} \lesssim h_T^{\ell_{\boldsymbol{\tau}}} |\boldsymbol{\tau}|_{\boldsymbol{H}^{\ell_{\boldsymbol{\tau}}}(T)}, \qquad \|\Pi_T^k(w) - w\|_{L^2(T)} \lesssim h_T^{\ell_w} |w|_{H^{\ell_w}(T)}. \tag{3.6}$$

3.3 Discrete operators

The HHO formulation is defined locally through a gradient reconstruction operator and a stabilization operator. The local gradient reconstruction operator $\boldsymbol{G}_T : \hat{V}_T^k \to \boldsymbol{\Sigma}_T^k$ is defined such that, for all $\hat{w}_T \in \hat{V}_T^k$ and all $T \in \mathcal{T}_h$,

$$(\boldsymbol{G}_T(\hat{w}_T), \boldsymbol{q})_{\boldsymbol{L}^2(T)} = (\nabla w_T, \boldsymbol{q})_{\boldsymbol{L}^2(T)} + (w_{\partial T} - w_T|_{\partial T}, \boldsymbol{q}\cdot\boldsymbol{n}_T)_{L^2(\partial T)} \quad \forall \boldsymbol{q} \in \boldsymbol{\Sigma}_T^k. \tag{3.7}$$

The global gradient reconstruction operator $\boldsymbol{G}_{\mathcal{T}} : \hat{V}_h^k \to \boldsymbol{\Sigma}_{\mathcal{T}}^k$ is defined as $(\boldsymbol{G}_{\mathcal{T}}(\hat{w}_h))|_T := \boldsymbol{G}_T(\hat{w}_T)$ for all $T \in \mathcal{T}_h$ and all $\hat{w}_h \in \hat{V}_h^k$.

The difference operator $\delta_{\partial T}$ is such that, for all $\hat{w}_T \in \hat{V}_T^k$ and all $T \in \mathcal{T}_h$,

$$\delta_{\partial T}(\hat{w}_T) := w_T|_{\partial T} - w_{\partial T}. \tag{3.8}$$

With this notation, we define the stabilization bilinear form $s_h : \hat{V}_h^k \times \hat{V}_h^k \to \mathbb{R}$ as

$$s_h(\hat{w}_h, \hat{z}_h) := \sum_{T\in\mathcal{T}_h} s_T(\hat{w}_T, \hat{z}_T), \qquad s_T(\hat{w}_T, \hat{z}_T) := \lambda_T^{-1}(\delta_{\partial T}(\hat{w}_T), \delta_{\partial T}(\hat{z}_T))_{L^2(\partial T)}, \tag{3.9}$$

with the stabilization parameter

$$\lambda_T := \mathfrak{c}_T \rho_T = \mathfrak{c}_T^{-1} \kappa_T. \tag{3.10}$$

Notice that the weight λ_T is h-independent; we speak of $\mathcal{O}(1)$-stabilization, as opposed to $\mathcal{O}(\frac{1}{h})$-stabilization when λ_T scales as h_T. Moreover, s_T bounds a LS penalty directly on $\delta_{\partial T}$; we speak of LS stabilization, as opposite to other stabilizations which penalize a higher-order operator acting on $\delta_{\partial T}$. The bilinear form s_h is symmetric positive-semidefinite on $\hat{V}_h^k \times \hat{V}_h^k$. We define the following seminorm on $\hat{V}_h^k$: For all $\hat{w}_h \in \hat{V}_{h0}^k$,

$$|\hat{w}_h|_{\mathrm{S}}^2 := \sum_{T\in\mathcal{T}_h} |\hat{w}_T|_{\mathrm{S},T}^2, \qquad |\hat{w}_T|_{\mathrm{S},T}^2 := s_T(\hat{w}_T, \hat{w}_T).$$

The standard HHO norm is defined as

$$\|\hat{w}_h\|_{\mathrm{HHO}} := \left\{ \sum_{T\in\mathcal{T}_h} \lambda_T^{-1} \big(\|\nabla w_T\|_{\boldsymbol{L}^2(T)}^2 + h_T^{-1} \|\delta_{\partial T}(\hat{w}_T)\|_{L^2(\partial T)}^2 \big) \right\}^{\frac{1}{2}}. \tag{3.11}$$

Proceeding as in [14] proves that, for all $\hat{w}_T \in \hat{V}_T^k$ and all $T \in \mathcal{T}_h$,

$$\|\nabla w_T\|_{\boldsymbol{L}^2(T)}^2 + h_T^{-1} \|\delta_{\partial T}(\hat{w}_T)\|_{L^2(\partial T)}^2 \approx \|\boldsymbol{G}_T(\hat{w}_T)\|_{\boldsymbol{L}^2(T)}^2 + h_T^{-1} \|\delta_{\partial T}(\hat{w}_T)\|_{L^2(\partial T)}^2. \tag{3.12}$$

Moreover, invoking the Cauchy-Schwarz inequality, a discrete trace inequality and an inverse inequality shows that, for all $\hat{w}_h \in \hat{V}_h^k$,

$$\|\boldsymbol{G}_{\mathcal{T}}(\hat{w}_h)\|_{\boldsymbol{L}^2(\Omega)}^2 \lesssim \sum_{T\in\mathcal{T}_h} \left\{ h_T^{-2}\|w_T\|_{L^2(T)}^2 + \lambda_T h_T^{-1}|\hat{w}_T|_{\mathrm{s},T}^2 \right\}. \tag{3.13}$$

3.4 Space semi-discrete problem

We define the bilinear form $a_h : \underline{\hat{V}}_h^k \times \underline{\hat{V}}_h^k \to \mathbb{R}$ as

$$a_h(\underline{\hat{w}}_h, \underline{\hat{z}}_h) := -(\boldsymbol{G}_{\mathcal{T}}(\hat{w}_h), \boldsymbol{\xi}_{\mathcal{T}})_{\boldsymbol{L}^2(\Omega)} + (\boldsymbol{\tau}_{\mathcal{T}}, \boldsymbol{G}_{\mathcal{T}}(\hat{z}_h))_{\boldsymbol{L}^2(\Omega)} + s_h(\hat{w}_h, \hat{z}_h), \tag{3.14}$$

for all $\underline{\hat{w}}_h := (\boldsymbol{\tau}_{\mathcal{T}}, w_{\mathcal{T}}, w_{\mathcal{F}}), \underline{\hat{z}}_h := (\boldsymbol{\xi}_{\mathcal{T}}, z_{\mathcal{T}}, z_{\mathcal{F}}) \in \underline{\hat{V}}_h^k$. The key dissipativity property of the HHO discretization in the present Friedrichs-like formulation is

$$a_h(\underline{\hat{w}}_h, \underline{\hat{w}}_h) = |\hat{w}_h|_{\mathrm{s}}^2 \qquad \forall \underline{\hat{w}}_h \in \underline{\hat{V}}_{h0}^k. \tag{3.15}$$

The space semi-discrete problem consists of finding $\underline{\hat{v}}_h \in C^1(\overline{J}; \underline{\hat{V}}_{h0}^k)$ such that, for all $t \in \overline{J}$,

$$\langle \partial_t \underline{v}_{\mathcal{T}}(t), \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} + a_h(\underline{\hat{v}}_h(t), \underline{\hat{z}}_h) = \langle \underline{f}(t), \underline{z}_{\mathcal{T}} \rangle_{\underline{L}} \qquad \forall \underline{\hat{z}}_h \in \underline{\hat{V}}_{h0}^k, \tag{3.16}$$

with an initial condition specified in the next section.

4 Time-continuous error analysis

This section is devoted to the time-continuous error analysis. In particular, we introduce the notion of consistency error in space, which hinges on an interpolation operator satisfying the abstract assumptions **(I1)**-**(I2)** stated below.

4.1 Interpolation and consistency

To proceed generally, we introduce two functional spaces to define and estimate the consistency error. The space $\underline{Y} := \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(\Omega) \times H^{\nu_v}(\Omega)$ for some $\nu_{\boldsymbol{\sigma}} \in [0,1]$ and $\nu_v \in (\frac{1}{2}, 1]$ is the domain of the interpolation operator, whereas the space $\underline{Z} := \boldsymbol{H}^{\mu_{\boldsymbol{\sigma}}}(\Omega) \times H_0^1(\Omega)$ for some $\mu_{\boldsymbol{\sigma}} \in [\nu_{\boldsymbol{\sigma}}, 1]$ is a subspace of $\underline{Y}$ to be used to estimate the consistency error.

We consider an interpolation operator of the form

$$\hat{\underline{I}}_h : \underline{Y} \to \underline{\hat{V}}_h^k, \qquad \hat{\underline{I}}_h := (\overbrace{\boldsymbol{\mathcal{I}}_{\mathcal{T}}, \underbrace{I_{\mathcal{T}}}_{=:\hat{I}_h}}^{\underline{I}_{\mathcal{T}}:=}, \Pi_{\mathcal{F}}^k). \tag{4.1}$$

Notice that we allow for some generality in the first two components of $\hat{\underline{I}}_h$, whereas the third component is always the L^2-orthogonal projection $\Pi_{\mathcal{F}}^k$. Then, assuming that the exact solution $\underline{v}(t) \in \underline{Y} \cap \underline{V}_0$ for all $t \in \overline{J}$, the consistency error is the linear form $\psi_h(\underline{v}(t); \cdot)$ on $\underline{\hat{V}}_{h0}^k$ defined as

$$\psi_h(\underline{v}(t); \underline{\hat{z}}_h) := a_h(\hat{\underline{I}}_h(\underline{v}(t)), \underline{\hat{z}}_h) - \langle \underline{\mathcal{B}}(\underline{v}(t)), \underline{z}_{\mathcal{T}} \rangle_{\underline{L}} \qquad \forall \underline{\hat{z}}_h \in \underline{\hat{V}}_{h0}^k. \tag{4.2}$$

We also introduce the notation

$$\widehat{\psi}_h(\underline{v}(t); \hat{z}_h) := \psi_h(\underline{v}(t); (\mathbf{0}, \hat{z}_h)) \qquad \forall \hat{z}_h \in \hat{V}_{h0}^k. \tag{4.3}$$

For a linear form $\phi_h \in (\hat{V}_{h0}^k)'$, we define the following quantities:

$$\|\phi_h\|_{(\mathrm{HHO})'} := \sup_{\hat{z}_h \in \hat{V}_{h0}^k} \frac{|\phi_h(\hat{z}_h)|}{\|\hat{z}_h\|_{\mathrm{HHO}}} \quad \text{and} \quad \|\phi_h\|_{\mathrm{s}'} := \sup_{\hat{z}_h \in \hat{V}_{h0}^k} \frac{|\phi_h(\hat{z}_h)|}{|\hat{z}_h|_{\mathrm{s}}}. \tag{4.4}$$

Note that $\|\cdot\|_{(\mathrm{HHO})'}$ is always bounded, whereas $\|\cdot\|_{\mathrm{S}'}$ may be unbounded.

Our analysis is based on the following key assumptions on the interpolation operator $\hat{\underline{I}}_h$:

(I1) $\underline{I}_{\mathcal{T}}$ is bounded, i.e., for all $\underline{w} \in \underline{Y}$,

$$\|\underline{I}_{\mathcal{T}}(\underline{w})\|_{\rho,\frac{1}{\kappa}} \lesssim \|\underline{w}\|_{\underline{H}^{\nu}(\rho,\frac{1}{\kappa};\mathcal{T}_h)} := \Big\{ \sum_{T\in\mathcal{T}_h} \big(\rho_T \|\boldsymbol{\tau}\|^2_{\boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T)} + \frac{1}{\kappa_T}\|w\|^2_{H^{\nu_v}(T)}\big)\Big\}^{\frac{1}{2}}, \tag{4.5}$$

with $\|w\|^2_{H^{\nu_v}(T)} := \|w\|^2_{L^2(T)} + h_T^{2\nu_v}|w|^2_{H^{\nu_v}(T)}$, and a similar notation for $\|\boldsymbol{\tau}\|_{\boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T)}$.

(I2) The consistency error $\psi_h(\underline{v}(t);\cdot)$ satisfies the following two properties: For all $t \in \overline{J}$,

- **(I2a)** $\psi_h(\underline{v}(t);(\boldsymbol{\xi}_{\mathcal{T}},0,0)) = 0$ for all $\boldsymbol{\xi}_{\mathcal{T}} \in \boldsymbol{\Sigma}^k_{\mathcal{T}}$. This means that $\psi_h(\underline{v}(t);\hat{\underline{z}}_h) = \widehat{\psi}_h(\underline{v}(t);\hat{z}_h)$ for all $\hat{\underline{z}}_h := (\boldsymbol{\xi}_{\mathcal{T}},\hat{z}_h) \in \hat{\underline{V}}^k_{h0}$,
- **(I2b)** Assuming that $\underline{v}(t) \in \underline{Z}$, $\|\widehat{\psi}_h(\underline{v}(t);\cdot)\|_{\mathrm{S}'}$ is bounded in terms of a suitable norm $\|\underline{v}(t)\|_{\underline{Z}}$.

We give two examples of interpolation operator $\hat{\underline{I}}_h$ satisfying **(I1)**-**(I2)** in Sections 4.3 and 4.4, but before that, we show how the assumptions **(I1)**-**(I2)** can be used to derive an error estimate in the time-continuous case.

4.2 Error analysis

We use the following shorthand notation:

$$\|\underline{w}\|^p_{L^p(J_t;**)} := \int_{J_t} \|\underline{w}(s)\|^p_{**}\,\mathrm{d}s \quad \text{for } p\in[1,\infty), \qquad \|\underline{w}\|_{C^0(\overline{J}_t;**)} := \sup_{s\in\overline{J}_t}\|\underline{w}(s)\|_{**}, \tag{4.6}$$

where $J_t := (0,t)$ for all $t \in J$, and the (seminorm) $\|\cdot\|_{**}$ depends on the context. We define the space semi-discrete error as

$$\hat{\underline{e}}_h(t) := \hat{\underline{I}}_h(\underline{v}(t)) - \hat{\underline{v}}_h(t) := (\boldsymbol{\mathcal{E}}_{\mathcal{T}}(t), \overbrace{\underbrace{e_{\mathcal{T}}(t), e_{\mathcal{F}}(t)}_{=:\hat{e}_h(t)}}^{\underline{e}_{\mathcal{T}}(t):=}). \tag{4.7}$$

Lemma 4.1 (Energy identity)**.** *Assume that the exact solution satisfies $\underline{v} \in C^0(\overline{J};\underline{V}_0)\cap C^1(\overline{J};\underline{Y})$, and let $\hat{\underline{v}}_h \in C^1(\overline{J};\hat{\underline{V}}^k_{h0})$ solve (3.16) with the initial condition $\underline{v}_{\mathcal{T}}(0) := \underline{I}_{\mathcal{T}}(\underline{v}_0)$. The following energy identity holds for all $t \in \overline{J}$:*

$$\frac{1}{2}\frac{d}{dt}\|\underline{e}_{\mathcal{T}}(t)\|^2_{\rho,\frac{1}{\kappa}} + |\hat{e}_h(t)|^2_{\mathrm{S}} = \langle \underline{I}_{\mathcal{T}}(\partial_t\underline{v}(t)) - \partial_t\underline{v}(t), \underline{e}_{\mathcal{T}}(t)\rangle_{\rho,\frac{1}{\kappa}} + \widehat{\psi}_h(\underline{v}(t);\hat{e}_h(t)). \tag{4.8}$$

Proof. The semi-discrete problem (3.16) and the model problem (2.5) show that, for all $\hat{\underline{z}}_h := (\boldsymbol{\xi}_{\mathcal{T}}, z_{\mathcal{T}}, z_{\mathcal{F}}) \in \hat{\underline{V}}^k_{h0}$ and all $t \in \overline{J}$,

$$\begin{aligned}
\langle \partial_t\underline{e}_{\mathcal{T}}(t), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + a_h(\hat{\underline{e}}_h(t),\hat{\underline{z}}_h) &= \langle \partial_t\underline{I}_{\mathcal{T}}(\underline{v}(t)), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + a_h(\hat{\underline{I}}_h(\underline{v}(t)),\hat{\underline{z}}_h) - \langle \underline{f}(t), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} \\
&= \langle \partial_t\underline{I}_{\mathcal{T}}(\underline{v}(t)), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + a_h(\hat{\underline{I}}_h(\underline{v}(t)),\hat{\underline{z}}_h) - \langle \partial_t\underline{v}(t), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} - \langle \underline{\mathcal{B}}(\underline{v}(t)), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} \\
&= \langle \underline{I}_{\mathcal{T}}(\partial_t\underline{v}(t)) - \partial_t\underline{v}(t), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \widehat{\psi}_h(\underline{v}(t);\hat{z}_h),
\end{aligned}$$

where we used $\partial_t\underline{I}_{\mathcal{T}} = \underline{I}_{\mathcal{T}}\partial_t$, the definition (4.2) of ψ_h, and **(I2a)**. Choosing $\hat{\underline{z}}_h := \hat{\underline{e}}_h(t)$ and using the dissipativity property (3.15) concludes the proof. □

Theorem 4.2 (Energy-error estimate)**.** *In addition to the assumptions of Lemma 4.1, we suppose that $\underline{v} \in L^2(J;\underline{Z})$. The following holds:*

$$\max\Big\{\frac{1}{4}\|\underline{e}_{\mathcal{T}}\|^2_{C^0(\overline{J};\rho,\frac{1}{\kappa})}, \frac{3}{8}\|\hat{e}_h\|^2_{L^2(J;\mathrm{S})}\Big\} \le \|\underline{I}_{\mathcal{T}}(\partial_t\underline{v}) - \partial_t\underline{v}\|^2_{L^1(J;\rho,\frac{1}{\kappa})} + \|\widehat{\psi}_h(\underline{v};\cdot)\|^2_{L^2(J;\mathrm{S}')}. \tag{4.9}$$

Proof. Let $t \in J$. The integration from 0 to t in (4.8) and the observation that $\underline{e}_{\mathcal{T}}(0) = \underline{0}$ owing to the choice of initial condition lead to

$$\frac{1}{2}\|\underline{e}_{\mathcal{T}}(t)\|^2_{\rho,\frac{1}{\kappa}} + \|\hat{e}_h\|^2_{L^2(0,t;\mathrm{S})} = \int_{J_t} \Big(\langle \underline{I}_{\mathcal{T}}(\partial_t \underline{v}(s)) - \partial_t \underline{v}(s), \underline{e}_{\mathcal{T}}(s) \rangle_{\rho,\frac{1}{\kappa}} + \widehat{\psi}_h(\underline{v}(s); \hat{e}_h(s)) \Big)\, ds. \tag{4.10}$$

Applying Hölder's inequality in time to the first term on the right hand-side, the Cauchy-Schwarz inequality in time to the second term, and Young's inequality shows that

$$\frac{1}{2}\|\underline{e}_{\mathcal{T}}(t)\|^2_{\rho,\frac{1}{\kappa}} + \frac{3}{4}\|\hat{e}_h\|^2_{L^2(J_t;\mathrm{S})} \le \|\underline{I}_{\mathcal{T}}(\partial_t \underline{v}) - \partial_t \underline{v}\|^2_{L^1(J_t;\rho,\frac{1}{\kappa})} + \frac{1}{4}\|\underline{e}_{\mathcal{T}}\|^2_{C^0(\overline{J}_t;\rho,\frac{1}{\kappa})} + \|\widehat{\psi}_h(\underline{v};\cdot)\|^2_{L^2(J_t;\mathrm{S}')}.$$

Since the left hand-side is bounded by the right hand-side for any $t' \in J_t$, we have

$$\frac{1}{4}\|\underline{e}_{\mathcal{T}}\|^2_{C^0(\overline{J}_t;\rho,\frac{1}{\kappa})} \le \|\underline{I}_{\mathcal{T}}(\partial_t \underline{v}) - \partial_t \underline{v}\|^2_{L^1(J_t;\rho,\frac{1}{\kappa})} + \|\widehat{\psi}_h(\underline{v};\cdot)\|^2_{L^2(J_t;\mathrm{S}')},$$

and, consequently,

$$\begin{aligned}\frac{3}{4}\|\hat{e}_h\|^2_{L^2(J_t;\mathrm{S})} &\le \|\underline{I}_{\mathcal{T}}(\partial_t \underline{v}) - \partial_t \underline{v}\|^2_{L^1(J_t;\rho,\frac{1}{\kappa})} + \frac{1}{4}\|\underline{e}_{\mathcal{T}}\|^2_{C^0(\overline{J}_t;\rho,\frac{1}{\kappa})} + \|\widehat{\psi}_h(\underline{v};\cdot)\|^2_{L^2(J_t;\mathrm{S}')} \\ &\le 2\big(\|\underline{I}_{\mathcal{T}}(\partial_t \underline{v}) - \partial_t \underline{v}\|^2_{L^1(J_t;\rho,\frac{1}{\kappa})} + \|\widehat{\psi}_h(\underline{v};\cdot)\|^2_{L^2(J_t;\mathrm{S}')}\big).\end{aligned}$$

This concludes the proof. □

4.3 Example 1: Classical HHO analysis

The standard choice of interpolation operator in the HHO error analysis is to set $\underline{Y}^{\mathrm{HHO}} := \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(\Omega) \times H^{\nu_v}(\Omega)$, with $\nu_{\boldsymbol{\sigma}} = 0, \nu_v \in (\frac{1}{2}, 1]$, and

$$\underline{\hat{I}}_h^{\mathrm{HHO}}(\underline{w}) := (\overbrace{\underbrace{\boldsymbol{\Pi}_{\mathcal{T}}^k(\boldsymbol{\tau}), \Pi_{\mathcal{T}}^k(w)}_{=:\hat{I}_h^{\mathrm{HHO}}(w)}, \Pi_{\mathcal{F}}^k(w|_{\mathcal{F}})}^{\underline{I}_{\mathcal{T}}^{\mathrm{HHO}}(\underline{w}):=}) \in \underline{\hat{V}}_h^k. \tag{4.11}$$

An optimal energy error estimate for the semi-discrete problem (3.16) is proven in [4] for $\mathcal{O}(\frac{1}{h})$-stabilization leading to $\mathcal{O}(h^{k+1})$ convergence rates. Here, we consider instead $\mathcal{O}(1)$-stabilization, see (3.9). The consistency error defined using $\underline{\hat{I}}_h^{\mathrm{HHO}}$ is denoted as

$$\psi_h^{\mathrm{HHO}}(\underline{v}(t); \underline{\hat{z}}_h) := a_h(\underline{\hat{I}}_h^{\mathrm{HHO}}(\underline{v}(t)), \underline{\hat{z}}_h) - \langle \underline{\mathcal{B}}(\underline{v}(t)), \underline{z}_{\mathcal{T}} \rangle_{\underline{L}} \qquad \forall \underline{\hat{z}}_h \in \underline{\hat{V}}_{h0}^k. \tag{4.12}$$

Since $\underline{Y}^{\mathrm{HHO}} \cap \underline{V}_0 = \underline{V}_0$, no extra regularity is needed on the exact solution to define the consistency error. Let $\underline{Z}^{\mathrm{HHO}} := \boldsymbol{H}^{\mu_{\boldsymbol{\sigma}}}(\Omega) \times H_0^1(\Omega)$, with $\mu_{\boldsymbol{\sigma}} \in (\frac{1}{2}, 1]$. For a pair $\underline{w} := (\boldsymbol{\tau}, w) \in \underline{Z}^{\mathrm{HHO}}$, we consider the seminorm

$$|\underline{w}|_{\underline{Z}}^{\mathrm{HHO}} := \left\{ \sum_{T \in \mathcal{T}_h} \big(\|\boldsymbol{\gamma}_T^{\mathrm{HHO}}(\boldsymbol{\tau}) \cdot \boldsymbol{n}_T\|^2_{L^2(\lambda_T;\partial T)} + h_T \|\nabla \gamma_T^{\mathrm{HHO}}(w)\|^2_{\boldsymbol{L}^2(\lambda_T^{-1};T)} \big) \right\}^{\frac{1}{2}}, \tag{4.13}$$

with $\boldsymbol{\gamma}_T^{\mathrm{HHO}}(\boldsymbol{\tau}) := \boldsymbol{\tau} - \boldsymbol{\Pi}_T^k(\boldsymbol{\tau})$ and $\gamma_T^{\mathrm{HHO}}(w) := w - \Pi_T^k(w)$.

Lemma 4.3 (Properties (I1)-(I2))**.** *The interpolation operator $\underline{I}_{\mathcal{T}}^{\mathrm{HHO}}$ satisfies (I1)-(I2) with $\underline{Z}^{\mathrm{HHO}}$ as defined above.*

Proof. (1) The stability of L^2-projections implies that $\|\underline{I}_{\mathcal{T}}^{\mathrm{HHO}}(\underline{w})\|_{\rho,\frac{1}{\kappa}} \le \|\underline{w}\|_{\rho,\frac{1}{\kappa}} \le \|\underline{w}\|_{H^\nu(\rho,\frac{1}{\kappa};\mathcal{T}_h)}$. This proves (I1).

(2a) The definition (4.12) of ψ_h^{HHO} and the definition (3.7) of $\boldsymbol{G}_T$ for all $T \in \mathcal{T}_h$ lead, for all $\boldsymbol{\xi}_{\mathcal{T}} \in \boldsymbol{\Sigma}_{\mathcal{T}}^k$, to

$$\begin{aligned}\psi_h^{\text{HHO}}(\underline{v}(t);(\boldsymbol{\xi}_{\mathcal{T}},0,0)) &= -(\boldsymbol{G}_{\mathcal{T}}(\hat{I}_h^{\text{HHO}}(v(t))),\boldsymbol{\xi}_{\mathcal{T}})_{\boldsymbol{L}^2(\Omega)} + (\nabla v(t),\boldsymbol{\xi}_{\mathcal{T}})_{\boldsymbol{L}^2(\Omega)}\\ &= \sum_{T\in\mathcal{T}_h}\left\{(\Pi_T^k(v(t)),\nabla\cdot\boldsymbol{\xi}_T)_{\boldsymbol{L}^2(T)} - (\Pi_{\partial T}^k(v(t)),\boldsymbol{\xi}_T\cdot\boldsymbol{n}_T)_{L^2(\partial T)} + (\nabla v(t),\boldsymbol{\xi}_T)_{\boldsymbol{L}^2(T)}\right\}\\ &= \sum_{T\in\mathcal{T}_h}\left\{(\Pi_T^k(v(t))-v(t),\nabla\cdot\boldsymbol{\xi}_T)_{\boldsymbol{L}^2(T)} + (v(t)-\Pi_{\partial T}^k(v(t)),\boldsymbol{\xi}_T\cdot\boldsymbol{n}_T)_{L^2(\partial T)}\right\}\\ &= 0,\end{aligned}$$

with an integration by parts and the L^2-orthogonalities of Π_T^k and $\Pi_{\partial T}^k$ in the last two steps. This proves **(I2a)**.
(2b) A direct calculation shows that, for all $\hat{z}_h := (z_{\mathcal{T}}, z_{\mathcal{F}}) \in \hat{V}_{h0}^k$,

$$\widehat{\psi}_h^{\text{HHO}}(\underline{v}(t);\hat{z}_h) = \sum_{T\in\mathcal{T}_h}\left((\boldsymbol{\sigma}(t)-\boldsymbol{\Pi}_{\mathcal{T}}^k(\boldsymbol{\sigma}(t)))\cdot\boldsymbol{n}_T, z_T - z_{\partial T}\right)_{L^2(\partial T)} + s_h(\hat{I}_h^{\text{HHO}}(v(t)),\hat{z}_h). \tag{4.14}$$

Using the definition of $\delta_{\partial T}$ (see (3.8)), we rewrite the above identity as

$$\begin{aligned}\widehat{\psi}_h^{\text{HHO}}(\underline{v}(t);\hat{z}_h) = \sum_{T\in\mathcal{T}_h}\left\{(\lambda_T^{\frac12}(\boldsymbol{\sigma}(t)\cdot\boldsymbol{n}_T - \boldsymbol{\Pi}_{\mathcal{T}}^k(\boldsymbol{\sigma}(t))\cdot\boldsymbol{n}_T), \lambda_T^{-\frac12}\delta_{\partial T}(\hat{z}_T))_{L^2(\partial T)}\right\}\\ + s_h(\hat{I}_h^{\text{HHO}}(v(t)),\hat{z}_h).\end{aligned}$$

This with the Cauchy-Schwarz inequality implies that

$$|\widehat{\psi}_h^{\text{HHO}}(\underline{v}(t);\hat{z}_h)| \lesssim \left\{\sum_{T\in\mathcal{T}_h}\left(\|\boldsymbol{\gamma}_T^{\text{HHO}}(\boldsymbol{\sigma}(t))\cdot\boldsymbol{n}_T\|_{L^2(\lambda_T;\partial T)}^2 + |\hat{I}_h^{\text{HHO}}(v(t))|_{\text{S},T}^2\right)\right\}^{\frac12}|\hat{z}_h|_{\text{S}}.$$

Finally, invoking the L^2-stability of $\Pi_{\partial T}^k$, a multiplicative trace inequality and a Poincaré inequality on T, we infer that

$$\begin{aligned}|\hat{I}_h^{\text{HHO}}(v(t))|_{\text{S},T}^2 &\le \lambda_T^{-1}\|v(t)-\Pi_T^k(v(t))\|_{L^2(\partial T)}^2\\ &\lesssim \lambda_T^{-1}h_T\|\nabla(v(t)-\Pi_T^k(v(t)))\|_{\boldsymbol{L}^2(T)}^2 \le h_T\|\nabla\gamma_T^{\text{HHO}}(v(t)))\|_{\boldsymbol{L}^2(\lambda_T^{-1};T)}^2.\end{aligned}$$

Hence, we obtain $|\widehat{\psi}_h^{\text{HHO}}(\underline{v}(t);\hat{z}_h)| \lesssim |\underline{v}(t)|_{\underline{Z}}^{\text{HHO}}|\hat{z}_h|_{\text{S}}$, and thus $\|\widehat{\psi}_h^{\text{HHO}}(\underline{v}(t);\cdot)\|_{\text{S}'} \lesssim |\underline{v}(t)|_{\underline{Z}}^{\text{HHO}}$. This concludes the proof of **(I2b)**. □

Following (4.7), we define the space semi-discrete error as

$$\hat{\underline{e}}_h^{\text{HHO}}(t) := \hat{\underline{I}}_h^{\text{HHO}}(\underline{v}(t)) - \underline{\hat{v}}_h(t) := (\overbrace{\boldsymbol{\mathcal{E}}_{\mathcal{T}}^{\text{HHO}}(t), \underbrace{e_{\mathcal{T}}^{\text{HHO}}(t)}_{=:\hat{e}_h^{\text{HHO}}(t)}}^{\underline{e}_{\mathcal{T}}^{\text{HHO}}(t):=}, e_{\mathcal{F}}^{\text{HHO}}(t)). \tag{4.15}$$

Corollary 4.4 (Energy identity and error estimate)**.** *Assume that the exact solution satisfies* $\underline{v} \in C^0(\overline{J};\underline{V}_0) \cap C^1(\overline{J};\underline{Y}^{\text{HHO}})$*, and let* $\underline{\hat{v}}_h \in C^1(\overline{J};\underline{\hat{V}}_{h0}^k)$ *solve* (3.16) *with the initial condition* $\underline{v}_{\mathcal{T}}(0) := \underline{I}_{\mathcal{T}}^{\text{HHO}}(\underline{v}_0)$*. The following identity holds for all* $t \in \overline{J}$*:*

$$\frac12\frac{d}{dt}\|\underline{e}_{\mathcal{T}}^{\text{HHO}}(t)\|_{\rho,\frac1\kappa}^2 + |\hat{e}_h^{\text{HHO}}(t)|_{\text{S}}^2 = \widehat{\psi}_h^{\text{HHO}}(\underline{v}(t);\hat{e}_h^{\text{HHO}}(t)). \tag{4.16}$$

In addition assuming that $\underline{v} \in L^2(J;\underline{Z}^{\text{HHO}})$*, we have*

$$\max\left\{\frac12\|\underline{e}_{\mathcal{T}}^{\text{HHO}}\|_{C^0(\overline{J};\rho,\frac1\kappa)}^2, \frac34\|\hat{e}_h^{\text{HHO}}\|_{L^2(J;\text{S})}^2\right\} \le \|\widehat{\psi}_h^{\text{HHO}}(\underline{v};\cdot)\|_{L^2(J;\text{S}')}^2. \tag{4.17}$$

Remark 4.5 (Convergence rates)**.** *In addition to the assumptions of Corollary 4.4, if there is* $\ell \in \{1,\dots,k+1\}$ *so that* $\underline{v} \in C^0(\overline{J};\boldsymbol{H}^\ell(\Omega)\times H^\ell(\Omega))$*, we obtain*

$$\|\underline{v}-\underline{v}_{\mathcal{T}}\|_{C^0(\overline{J};\rho,\frac{1}{\kappa})} \lesssim \mathcal{O}(h^{\ell-\frac{1}{2}}). \tag{4.18}$$

In the case where $\ell = k+1$*, this gives* $\mathcal{O}(h^{k+\frac{1}{2}})$ *convergence rate.*

4.4 Example 2: HDG interpolation on simplices

In this section, we evaluate the consistency error using the HDG interpolation operator from [8]. This improves the $\mathcal{O}(h^{k+\frac{1}{2}})$ convergence rate from Remark 4.5 to $\mathcal{O}(h^{k+1})$ on simplices.

For all $\underline{w} := (\boldsymbol{\tau}, w) \in \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T)\times H^{\nu_v}(T)$, with $\nu_{\boldsymbol{\sigma}} \in (\frac{1}{2},1]$, $\nu_v \in (\frac{1}{2},1]$, one defines $\underline{\Pi}_T^{\mathrm{H}}(\underline{w}) := (\boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{w}), \Pi_T^V(\underline{w})) \in \mathbb{P}^k(T;\mathbb{R}^d)\times\mathbb{P}^k(T;\mathbb{R})$ on any simplex $T\in\mathcal{T}_h$ by solving

$$(\Pi_T^V(\underline{w})-w, z)_{L^2(T)} = 0 \qquad \forall z\in\mathbb{P}^{k-1}(T;\mathbb{R}), \tag{4.19a}$$

$$(\boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{w})-\boldsymbol{\tau}, \boldsymbol{\xi})_{\boldsymbol{L}^2(T)} = 0 \qquad \forall \boldsymbol{\xi}\in\mathbb{P}^{k-1}(T;\mathbb{R}^d), \tag{4.19b}$$

$$((\boldsymbol{\tau}-\boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{w}))\cdot\boldsymbol{n}_T, \mu)_{L^2(\partial T)} = \lambda_T^{-1}(\delta_{\partial T}(\Pi_T^V(\underline{w}), \Pi_{\partial T}^k(w|_{\partial T})), \delta_{\partial T}(0,\mu))_{L^2(\partial T)} \quad \forall \mu\in V_{\partial T}^k. \tag{4.19c}$$

As shown in [8], the interpolation operator $\underline{\Pi}_T^{\mathrm{H}}$ is well-defined. Moreover, if $\underline{w} := (\boldsymbol{\tau}, w) \in \boldsymbol{H}^{\ell_{\boldsymbol{\tau}}}(T)\times H^{\ell_w}(T)$, with $\ell_{\boldsymbol{\tau}}\in[\nu_{\boldsymbol{\sigma}}, k+1]$, $\ell_w\in[\nu_v, k+1]$, the following approximation properties hold:

$$\|\Pi_T^V(\underline{w})-w\|_{L^2(T)} \lesssim h_T^{\ell_w}|w|_{H^{\ell_w}(T)} + \lambda_T h_T^{\ell_{\boldsymbol{\tau}}}|\boldsymbol{\tau}|_{\boldsymbol{H}^{\ell_{\boldsymbol{\tau}}}(T)}, \tag{4.20a}$$

$$\|\boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{w})-\boldsymbol{\tau}\|_{\boldsymbol{L}^2(T)} \lesssim h_T^{\ell_{\boldsymbol{\tau}}}|\boldsymbol{\tau}|_{\boldsymbol{H}^{\ell_{\boldsymbol{\tau}}}(T)} + \lambda_T^{-1} h_T^{\ell_w}|w|_{H^{\ell_w}(T)}. \tag{4.20b}$$

We define the global HDG interpolation operator $\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}$ on $\underline{Y}^{\mathrm{H}} := \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(\Omega)\times H^{\nu_v}(\Omega)$ by setting $(\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}(\underline{w}))|_T := \underline{\Pi}_T^{\mathrm{H}}(\underline{w}|_T)$ for all $T\in\mathcal{T}_h$, and we denote its two components by $\boldsymbol{\Pi}_{\mathcal{T}}^{\boldsymbol{\Sigma}}$ and $\Pi_{\mathcal{T}}^V$. Then we define the global interpolation operator such that, for all $\underline{w} := (\boldsymbol{\tau}, w) \in \underline{Y}^{\mathrm{H}}$,

$$\hat{\underline{I}}_h^{\mathrm{H}}(\underline{w}) := (\overbrace{\boldsymbol{\Pi}_{\mathcal{T}}^{\boldsymbol{\Sigma}}(\underline{w}), \underbrace{\Pi_{\mathcal{T}}^V(\underline{w})}, \Pi_{\mathcal{F}}^k(w|_{\mathcal{F}})}^{\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}(\underline{w}):=}) \in \hat{\underline{V}}_h^k.$$
$$=:\hat{I}_h^{\mathrm{H}}(\underline{w})$$

Assuming $\underline{v}(t)\in\underline{Y}^{\mathrm{H}}$ for all $t\in\overline{J}$, we define the consistency error as

$$\psi_h^{\mathrm{H}}(\underline{v}(t);\hat{\underline{z}}_h) := a_h(\hat{\underline{I}}_h^{\mathrm{H}}(\underline{v}(t)), \hat{\underline{z}}_h) - \langle \underline{\mathcal{B}}(\underline{v}(t)), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} \qquad \forall \hat{\underline{z}}_h\in\hat{\underline{V}}_{h0}^k. \tag{4.21}$$

Lemma 4.6 (Properties (I1)-(I2))**.** *The interpolation operator* $\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}$ *satisfies* (I1)-(I2) *with* $\underline{Z}^{\mathrm{H}} := \underline{Y}^{\mathrm{H}}$*.*

Proof. (1) The triangle inequality and the approximation estimate (4.20) show that

$$\begin{aligned}
\|\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}(\underline{w})\|_{\rho,\frac{1}{\kappa}} &\le \|\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}(\underline{w})-\underline{w}\|_{\rho,\frac{1}{\kappa}} + \|\underline{w}\|_{\rho,\frac{1}{\kappa}} \\
&\lesssim \Bigg\{ \sum_{T\in\mathcal{T}_h} \Big(\rho_T\big(h_T^{2\nu_{\boldsymbol{\sigma}}}|\boldsymbol{\tau}|^2_{\boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T)} + \lambda_T^{-2}h_T^{2\nu_v}|w|^2_{H^{\nu_v}(T)}\big) \\
&\quad + \frac{1}{\kappa_T}\big(h_T^{2\nu_v}|w|^2_{H^{\nu_v}(\Omega)} + \lambda_T^2 h_T^{2\nu_{\boldsymbol{\sigma}}}|\boldsymbol{\tau}|^2_{\boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T)}\big)\Big)\Bigg\}^{\frac{1}{2}} + \|\underline{w}\|_{\rho,\frac{1}{\kappa}} \\
&\le 2\|\underline{w}\|_{H^{\nu}(\rho,\frac{1}{\kappa};\mathcal{T}_h)},
\end{aligned}$$

using that $\rho_T\lambda_T^{-2} = \frac{1}{\kappa_T}$ for all $T\in\mathcal{T}_h$ in the last step. This proves (I1).
(2a) The definition (4.21) of ψ_h^{H} and the definition (3.7) of $\boldsymbol{G}_T$ for all $T\in\mathcal{T}_h$ lead, for all $\boldsymbol{\xi}_{\mathcal{T}}\in\boldsymbol{\Sigma}_{\mathcal{T}}^k$, to

$$\psi_h^{\mathrm{H}}(\underline{v}(t);(\boldsymbol{\xi}_{\mathcal{T}},0,0)) = -(\boldsymbol{G}_{\mathcal{T}}(\hat{I}_h^{\mathrm{H}}(v(t))), \boldsymbol{\xi}_{\mathcal{T}})_{\boldsymbol{L}^2(\Omega)} + (\nabla v(t), \boldsymbol{\xi}_{\mathcal{T}})_{\boldsymbol{L}^2(\Omega)}$$

$$= \sum_{T\in\mathcal{T}_h} \left\{ (\Pi_T^V(\underline{v}(t)) - v(t), \nabla\cdot\boldsymbol{\xi}_T)_{L^2(T)} + (v(t) - \Pi_{\partial T}^k(v(t)), \boldsymbol{\xi}_T\cdot\boldsymbol{n}_T)_{L^2(\partial T)} \right\}$$
$$= 0,$$

with an integration by parts, the definition (4.19a) of Π_T^V, and the L^2-orthogonality of $\Pi_{\partial T}^k$ in the last two steps. This proves **(I2a)**.
(2b) Similar arguments lead, for all $\hat{z}_h := (z_{\mathcal{T}}, z_{\mathcal{F}}) \in \hat{V}_{h0}^k$, to

$$\begin{aligned}
\widehat{\psi}_h^{\mathrm{H}}(\underline{v}(t); \hat{z}_h) &= (\boldsymbol{\Pi}_{\mathcal{T}}^{\boldsymbol{\Sigma}}(\underline{v}(t)), \boldsymbol{G}_{\mathcal{T}}(\hat{z}_h))_{\boldsymbol{L}^2(\Omega)} + s_h(\hat{I}_h^{\mathrm{H}}(\underline{v}(t)), \hat{z}_h) + (\nabla\cdot\boldsymbol{\sigma}(t), z_{\mathcal{T}})_{L^2(\Omega)} \\
&= \sum_{T\in\mathcal{T}_h} \Big\{ (\boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{v}(t)) - \boldsymbol{\sigma}(t), \nabla z_T)_{\boldsymbol{L}^2(T)} + ((\boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{v}(t)) - \boldsymbol{\sigma}(t))\cdot\boldsymbol{n}_T, z_{\partial T} - z_T)_{L^2(\partial T)} \\
&\qquad + s_T(\hat{I}_T^{\mathrm{H}}(\underline{v}(t)), \hat{z}_T) \Big\} \\
&= -\sum_{T\in\mathcal{T}_h} \Big\{ ((\boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{v}(t)) - \boldsymbol{\sigma}(t))\cdot\boldsymbol{n}_T, \delta_{\partial T}(\hat{z}_T))_{L^2(\partial T)} + \lambda_T^{-1}(\delta_{\partial T}(\hat{I}_T^{\mathrm{H}}(\underline{v}(t))), \delta_{\partial T}(\hat{z}_T))_{L^2(\partial T)} \Big\} \\
&= 0,
\end{aligned}$$

where we used (4.19b) and the definition (3.8) of $\delta_{\partial T}$ in the third step, and (4.19c) with $\mu := \delta_{\partial T}(\hat{z}_T)$ in the last step. This proves that $|\widehat{\psi}_h^{\mathrm{HHO}}(\underline{v}(t);\cdot)| = 0$ and hence **(I2b)**. □

Following (4.7), we define the space semi-discrete error as

$$\underline{\hat{e}}_h^{\mathrm{H}}(t) := \underline{\hat{I}}_h^{\mathrm{H}}(\underline{v}(t)) - \underline{\hat{v}}_h(t) := (\overbrace{\boldsymbol{\mathcal{E}}_{\mathcal{T}}^{\mathrm{H}}(t), \underbrace{e_{\mathcal{T}}^{\mathrm{H}}(t)}}^{\underline{e}_{\mathcal{T}}^{\mathrm{H}}(t):=}, \underbrace{e_{\mathcal{F}}^{\mathrm{HHO}}(t)}_{=:\hat{e}_h^{\mathrm{H}}(t)}). \tag{4.22}$$

Corollary 4.7 (Energy identity and error estimate)**.** *Assume that the exact solution satisfies* $\underline{v} \in C^0(\overline{J}; \underline{V}_0) \cap C^1(\overline{J}; \underline{Y}^{\mathrm{H}})$*, and let* $\underline{\hat{v}}_h \in C^1(\overline{J}; \underline{\hat{V}}_{h0}^k)$ *solve* (3.16) *with the initial condition* $\underline{v}_{\mathcal{T}}(0) := \underline{I}_{\mathcal{T}}^{\mathrm{H}}(\underline{v}_0)$*. The following identity holds for all* $t \in \overline{J}$*:*

$$\frac{1}{2}\frac{d}{dt}\|\underline{e}_{\mathcal{T}}^{\mathrm{H}}(t)\|_{\rho,\frac{1}{\kappa}}^2 + |\hat{e}_h^{\mathrm{H}}(t)|_{\mathrm{S}}^2 = \langle \underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}(\partial_t\underline{v}(t)) - \partial_t\underline{v}(t), \underline{e}_{\mathcal{T}}^{\mathrm{H}}\rangle_{\rho,\frac{1}{\kappa}}. \tag{4.23}$$

In addition, we have

$$\max\left\{\frac{1}{4}\|\underline{e}_{\mathcal{T}}^{\mathrm{H}}\|_{C^0(\overline{J};\rho,\frac{1}{\kappa})}^2, \frac{1}{2}\|\hat{e}_h^{\mathrm{H}}\|_{L^2(J;\mathrm{S})}^2\right\} \le \|\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}(\partial_t\underline{v}) - \partial_t\underline{v}\|_{\rho,\frac{1}{\kappa}}^2. \tag{4.24}$$

Remark 4.8 (Improved convergence rates on simplices)**.** *In addition to the assumptions of Corollary 4.7, if there is* $\ell \in \{1, \ldots, k+1\}$ *so that* $\underline{v} \in C^1(\overline{J}; \boldsymbol{H}^\ell(\Omega) \times H^\ell(\Omega))$*, we obtain*

$$\|\underline{v} - \underline{v}_{\mathcal{T}}\|_{C^0(\overline{J};\rho,\frac{1}{\kappa})} \lesssim \mathcal{O}(h^\ell). \tag{4.25}$$

In the case where $\ell = k+1$*, this gives* $\mathcal{O}(h^{k+1})$ *convergence rate.*

5 Main results on time-discrete error analysis

Here onwards, we assume that the mesh family $\{\mathcal{T}_h\}_{h>0}$ is quasi-uniform since we are going to invoke CFL conditions in the context of explicit time-stepping schemes. We could also formulate the CFL conditions using the minimum mesh-size. In this section, we state our main assumptions to lead the analysis, we present the ERK2 and ERK3 schemes, and state our main results. The proofs are postponed to Section 6.

5.1 CFL conditions

The time interval $\overline{J} = [0, T_f]$ is divided into N (open) sub-intervals J_n of equal size $\tau = \frac{T_f}{N}$ having time nodes $(t^n)_{n\in\{0,N\}}$ with $t^0 = 0$ and $t^N = T_f$. Let $\mathfrak{c}_\infty := \|\mathfrak{c}\|_{L^\infty(\Omega)}$. In what follows, the superscript $(\bullet)^n$ refers to a value computed at the discrete time t^n; for instance $\underline{f}_{\mathcal{T}}^n := \underline{f}_{\mathcal{T}}(t^n), \partial_t \underline{f}_{\mathcal{T}}^n := \partial_t \underline{f}_{\mathcal{T}}(t^n)$, and so on.

We always assume that the following usual CFL condition holds:

$$\tau \le \rho_1 \frac{h}{\mathfrak{c}_\infty}, \tag{5.1}$$

for some positive real number ρ_1. In the case of ERK2, we need the so-called $\frac{4}{3}$-CFL condition

$$\tau \le \rho_{\frac{4}{3}} T_f^{-\frac{1}{3}} \Big(\frac{h}{\mathfrak{c}_\infty}\Big)^{\frac{4}{3}}, \tag{5.2}$$

for some positive real number $\rho_{\frac{4}{3}}$. Notice that ρ_1 and $\rho_{\frac{4}{3}}$ are nondimensional numbers and that (5.2) implies (5.1) with $\rho_1 = \rho_{\frac{4}{3}} \Big(\frac{\ell_\Omega}{T_f \mathfrak{c}_\infty}\Big)^{\frac{1}{3}}$ (recall that $h \le \ell_\Omega$).

5.2 Main assumptions

We define the linear operators $\underline{\mathcal{A}}_{\mathcal{T}} : \underline{\hat{V}}_{h0}^k \to \underline{V}_{\mathcal{T}}^k$ and $\mathcal{A}_{\mathcal{F}} : \underline{\hat{V}}_{h0}^k \to V_{\mathcal{F}0}^k$ as follows:

$$\langle \underline{\mathcal{A}}_{\mathcal{T}}(\underline{\hat{w}}_h), \underline{z}_{\mathcal{T}} \rangle_{\underline{L}} := a_h(\underline{\hat{w}}_h, (\underline{z}_{\mathcal{T}}, 0)) \quad \text{and} \quad (\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h), z_{\mathcal{F}})_{L^2(\mathcal{F})} := a_h(\underline{\hat{w}}_h, (\underline{0}, z_{\mathcal{F}})), \tag{5.3}$$

for all $\underline{\hat{w}}_h := (\boldsymbol{\tau}_{\mathcal{T}}, w_{\mathcal{T}}, w_{\mathcal{F}}), \underline{\hat{z}}_h := (\boldsymbol{\xi}_{\mathcal{T}}, z_{\mathcal{T}}, z_{\mathcal{F}}) \in \underline{\hat{V}}_{h0}^k$. For all $w_{\mathcal{T}} \in V_{\mathcal{T}}^k$ and for all $w_{\mathcal{F}} \in V_{\mathcal{F}0}^k$, it is convenient to define

$$\|w_{\mathcal{T}}\|_{\mathrm{s}'} := \sup_{z_{\mathcal{T}} \in V_{\mathcal{T}}^k} \frac{|(w_{\mathcal{T}}, z_{\mathcal{T}})_{L^2(\Omega)}|}{|(z_{\mathcal{T}}, 0)|_{\mathrm{s}}}, \quad \|w_{\mathcal{F}}\|_{\mathrm{s}'} := \sup_{z_{\mathcal{F}} \in V_{\mathcal{F}0}^k} \frac{|(w_{\mathcal{F}}, z_{\mathcal{F}})_{L^2(\mathcal{F})}|}{|(0, z_{\mathcal{F}})|_{\mathrm{s}}}. \tag{5.4}$$

We can now state our main abstract assumptions to lead the analysis.

(A1) For all $w_{\mathcal{T}} \in V_{\mathcal{T}}^k$,

$$|(w_{\mathcal{T}}, 0)|_{\mathrm{s}} \lesssim \mathfrak{c}_\infty^{\frac{1}{2}} h^{-\frac{1}{2}} \|w_{\mathcal{T}}\|_{L^2(\frac{1}{\kappa};\Omega)}. \tag{5.5}$$

(A2) There exists a constant $C_{\mathcal{S}}$ such that, for all $\underline{\hat{w}}_h \in \underline{\hat{V}}_{h0}^k$ with $\|\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h)\|_{\mathrm{s}'}$ finite,

$$|\hat{w}_h|_{\mathrm{s}} \le C_{\mathcal{S}} (\mathfrak{c}_\infty^{\frac{1}{2}} h^{-\frac{1}{2}} \|\underline{w}_{\mathcal{T}}\|_{\rho,\frac{1}{\kappa}} + \|\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h)\|_{\mathrm{s}'}). \tag{5.6}$$

(A3) There exists a constant $C_{\mathcal{A}}$ such that, for all $\underline{\hat{w}}_h \in \underline{\hat{V}}_{h0}^k$ with $\|\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h)\|_{\mathrm{s}'}$ finite,

$$\|\underline{\mathcal{A}}_{\mathcal{T}}(\underline{\hat{w}}_h)\|_{\frac{1}{\rho},\kappa} \le C_{\mathcal{A}} (\mathfrak{c}_\infty h^{-1} \|\underline{w}_{\mathcal{T}}\|_{\rho,\frac{1}{\kappa}} + \mathfrak{c}_\infty^{\frac{1}{2}} h^{-\frac{1}{2}} \|\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h)\|_{\mathrm{s}'}). \tag{5.7}$$

The proof of the above properties for the HHO discretization is postponed to Section 6.1. We use a specific symbol $C_{\mathcal{S}}$ and $C_{\mathcal{A}}$ (instead of $\lesssim$) in **(A2)** and **(A3)**, respectively, because both constants are explicitly invoked in the CFL conditions for ERK3.

Remark 5.1 (Comparison with [5])**.** *The property* **(A1)** *evaluates HHO functions with zero face values in the stabilization seminorm, and is essentially the same as the property (2.21) in [5] in the dG setting. The second property* **(A2)** *is the novel property which bounds general HHO functions in the stabilization seminorm. We notice that the term $\|\mathcal{A}_{\mathcal{F}}(\cdot)\|_{\mathrm{s}'}$ accounts for the static coupling between face and cell unknowns. Finally, the property* **(A3)** *is similar to (2.29) in [5] with the additional face contribution coming from* **(A2)**.

5.3 ERK2 and ERK3 schemes

In the paper, we analyze ERK2 and ERK3 schemes of a particular form, but we show in Section 5.4 that they are equivalent, up to quadrature on f, to any ERK scheme specified using the more classical Butcher arrays.

ERK2 scheme. Given $\underline{v}_{\mathcal{T}}^{n,1} := \underline{v}_{\mathcal{T}}^{n} \in \underline{V}_{\mathcal{T}}^{k}$ from the previous time step or the initial condition, the ERK2 scheme consists of finding $\underline{v}_{\mathcal{T}}^{n+1} \in \underline{V}_{\mathcal{T}}^{k}$ in two stages as follows:

$$\langle \underline{v}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = \langle \underline{v}_{\mathcal{T}}^{n,1}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \tau a_h(\hat{\underline{v}}_h^{n,1}, \hat{\underline{z}}_h) + \tau \langle \underline{f}_{\mathcal{T}}^{n,1}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}}, \tag{5.8a}$$

$$\langle \underline{v}_{\mathcal{T}}^{n+1}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = \frac{1}{2}\langle \underline{v}_{\mathcal{T}}^{n,1} + \underline{v}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\tau a_h(\hat{\underline{v}}_h^{n,2}, \hat{\underline{z}}_h) + \frac{1}{2}\tau \langle \underline{g}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}}, \tag{5.8b}$$

for all $\hat{\underline{z}}_h \in \hat{\underline{V}}_{h0}^{k}$, with $\underline{f}_{\mathcal{T}}^{n,1} := \underline{f}_{\mathcal{T}}^{n}$, and, assuming $f \in C^2(\overline{J}; L^2(\Omega))$, $\underline{g}_{\mathcal{T}}^{n,2}$ satisfies

$$\|\underline{g}_{\mathcal{T}}^{n,2} - \underline{f}_{\mathcal{T}}^{n,2}\|_{\frac{1}{\rho},\kappa} \lesssim C_2^n(\underline{f})\tau^2, \tag{5.9}$$

with $\underline{f}_{\mathcal{T}}^{n,2} := \underline{f}_{\mathcal{T}}^{n} + \tau\partial_t \underline{f}_{\mathcal{T}}^{n}$ and $C_2^n(\underline{f}) := \|\partial_{tt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)}$.

ERK3 scheme. Given $\underline{v}_{\mathcal{T}}^{n,1} := \underline{v}_{\mathcal{T}}^{n} \in \underline{V}_{\mathcal{T}}^{k}$ from the previous time step or the initial condition, the ERK3 scheme consists of finding $\underline{v}_{\mathcal{T}}^{n+1} \in \underline{V}_{\mathcal{T}}^{k}$ in three stages as follows:

$$\langle \underline{v}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = \langle \underline{v}_{\mathcal{T}}^{n,1}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \tau a_h(\hat{\underline{v}}_h^{n,1}, \hat{\underline{z}}_h) + \tau \langle \underline{f}_{\mathcal{T}}^{n,1}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}}, \tag{5.10a}$$

$$\langle \underline{v}_{\mathcal{T}}^{n,3}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = \frac{1}{2}\langle \underline{v}_{\mathcal{T}}^{n,1} + \underline{v}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\tau a_h(\hat{\underline{v}}_h^{n,2}, \hat{\underline{z}}_h) + \frac{1}{2}\tau \langle \underline{f}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}}, \tag{5.10b}$$

$$\langle \underline{v}_{\mathcal{T}}^{n+1}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = \frac{1}{3}\langle \underline{v}_{\mathcal{T}}^{n,1} + \underline{v}_{\mathcal{T}}^{n,2} + \underline{v}_{\mathcal{T}}^{n,3}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{3}\tau a_h(\hat{\underline{v}}_h^{n,3}, \hat{\underline{z}}_h) + \frac{1}{3}\tau \langle \underline{g}_{\mathcal{T}}^{n,3}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}}, \tag{5.10c}$$

for all $\hat{\underline{z}}_h \in \hat{\underline{V}}_{h0}^{k}$, with $\underline{f}_{\mathcal{T}}^{n,1}, \underline{f}_{\mathcal{T}}^{n,2}$ defined as above, and, assuming $f \in C^3(\overline{J}; L^2(\Omega)) \cap C^2(\overline{J}; H^1(\Omega))$, $\underline{g}_{\mathcal{T}}^{n,3}$ satisfies

$$\|\underline{g}_{\mathcal{T}}^{n,3} - \underline{f}_{\mathcal{T}}^{n,3}\|_{\frac{1}{\rho},\kappa} \lesssim C_3^n(\underline{f})\tau^3, \tag{5.11}$$

with $\underline{f}_{\mathcal{T}}^{n,3} := \underline{f}_{\mathcal{T}}^{n} + \tau\partial_t \underline{f}_{\mathcal{T}}^{n} + \frac{1}{2}\tau^2\partial_{tt}\underline{f}_{\mathcal{T}}^{n}$ and $C_3^n(\underline{f}) := \|\partial_{ttt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)} + \mathfrak{c}_\infty\|\nabla\partial_{tt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)}$.

5.4 Statements of main results and comments

Let $\underline{I}_{\mathcal{T}} : \underline{Y} \to \underline{V}_{\mathcal{T}}^{k}$ be an interpolation operator satisfying the assumptions **(I1)** and **(I2)** stated in Section 4.1. Recall that $\hat{\underline{I}}_h = (\underline{I}_{\mathcal{T}}, \Pi_{\mathcal{F}}^{k})$. We define the following quantities:

$$\underline{v}^{n,1} := \underline{v}^{n}, \qquad \underline{v}^{n,2} := \underline{v}^{n} + \tau\partial_t\underline{v}^{n}, \qquad \underline{v}^{n,3} := \underline{v}^{n} + \tau\partial_t\underline{v}^{n} + \frac{1}{2}\tau^2\partial_{tt}\underline{v}^{n}, \tag{5.12a}$$

$$\widehat{\psi}_h^{n,1}(\cdot) := \widehat{\psi}_h(\underline{v}^{n,1};\cdot), \quad \widehat{\psi}_h^{n,2}(\cdot) := \widehat{\psi}_h(\underline{v}^{n,2};\cdot), \quad \widehat{\psi}_h^{n,3}(\cdot) := \widehat{\psi}_h(\underline{v}^{n,3};\cdot), \tag{5.12b}$$

where we recall that the consistency error $\widehat{\psi}_h$ is defined as a linear form acting on $\hat{V}_{h0}^{k}$ (see **(I2)**). We also recall that the space $\underline{Y}$ is the domain of the interpolation operator, whereas the space $\underline{Z}$ is used to estimate the consistency error. Let $\underline{\mathcal{D}} : \underline{Y} \to \underline{V}_{\mathcal{T}}^{k}$ be the linear operator such that $\underline{\mathcal{D}}(\cdot) := \underline{I}_{\mathcal{T}}(\cdot) - \underline{\Pi}_{\mathcal{T}}^{k}(\cdot)$ (recall that $\underline{\Pi}_{\mathcal{T}}^{k} = (\mathbf{\Pi}_{\mathcal{T}}^{k}, \Pi_{\mathcal{T}}^{k})$). We set

$$\partial_t\underline{v}^{n,1} := \partial_t\underline{v}^{n}, \qquad \partial_t\underline{v}^{n,2} := \partial_t\underline{v}^{n} + \tau\partial_{tt}\underline{v}^{n}, \quad \partial_t\underline{v}^{n,3} := \partial_t\underline{v}^{n} + \tau\partial_{tt}\underline{v}^{n} + \frac{1}{2}\tau^2\partial_{ttt}\underline{v}^{n}, \tag{5.13a}$$

$$\underline{\mathcal{D}}^{n,1} := \underline{\mathcal{D}}(\partial_t\underline{v}^{n,1}), \quad \underline{\mathcal{D}}^{n,2} := \underline{\mathcal{D}}(\partial_t\underline{v}^{n,2}), \quad \underline{\mathcal{D}}^{n,3} := \underline{\mathcal{D}}(\partial_t\underline{v}^{n,3}). \tag{5.13b}$$

The quantities $\underline{\mathcal{D}}^{n,1}, \underline{\mathcal{D}}^{n,2}$, and $\underline{\mathcal{D}}^{n,3}$ should be understood as space approximation errors on the time derivatives of the exact solution. We define the following discrete errors:

$$\hat{\underline{\zeta}}_h^{n,1} := \hat{\underline{v}}_h^{n,1} - \hat{\underline{I}}_h(\underline{v}^{n,1}), \qquad \hat{\underline{\zeta}}_h^{n,2} := \hat{\underline{v}}_h^{n,2} - \hat{\underline{I}}_h(\underline{v}^{n,2}), \qquad \hat{\underline{\zeta}}_h^{n,3} := \hat{\underline{v}}_h^{n,3} - \hat{\underline{I}}_h(\underline{v}^{n,3}). \tag{5.14}$$

Theorem 5.2 (Error estimate for ERK2)**.** *Assume $f \in C^2(\overline{J};L^2(\Omega))$ and $\underline{v} \in C^1(\overline{J};\underline{Z}\cap\underline{V}_0)\cap C^3(\overline{J};\underline{Y})$ and that the initial condition is $\underline{v}^0 := \underline{I}_{\mathcal{T}}(\underline{v}_0)$. Then, under the strengthened $\frac{4}{3}$-CFL condition (5.2), the following holds:*

$$\|\underline{\zeta}_{\mathcal{T}}^N\|_{\rho,\frac{1}{\kappa}}^2 + \sum_{n=0}^{N-1}\tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{S}}^2 + |\hat{\zeta}_h^{n,2}|_{\mathrm{S}}^2\big) \lesssim C_2(\underline{f},\underline{v})^2 T_f^2\tau^4 + \sum_{n=0}^{N-1}\tau\Big\{\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2 + \|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2 + T_f\Big\{\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \tau^2\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2\Big\}\Big\}, \tag{5.15}$$

where $C_2(\underline{f},\underline{v}) := \|\partial_{tt}\underline{f}\|_{C^0(\overline{J};\frac{1}{\rho},\kappa)} + \|\partial_{ttt}\underline{v}\|_{C^0(\overline{J};H^\nu(\rho,\frac{1}{\kappa};\mathcal{T}_h))}$.

Theorem 5.3 (Error estimate for ERK3)**.** *Assume $f \in C^3(\overline{J};L^2(\Omega))\cap C^2(\overline{J};H^1(\Omega))$ and $\underline{v} \in C^2(\overline{J};\underline{Z}\cap\underline{V}_0)\cap C^4(\overline{J};\underline{Y})$ and that the initial condition is $\underline{v}^0 := \underline{I}_{\mathcal{T}}(\underline{v}_0)$. Then, under the usual CFL condition (5.1) with*

$$\rho_1 \le \min\Big\{\frac{5}{308}C_{\mathcal{S}}^{-1}, \Big(\frac{3}{2}\Big)^{\frac{1}{2}}C_{\mathcal{A}}^{-1}\Big\}, \tag{5.16}$$

the following holds:

$$\|\underline{\zeta}_{\mathcal{T}}^N\|_{\rho,\frac{1}{\kappa}}^2 + \sum_{n=0}^{N-1}\tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{S}}^2 + |\hat{\zeta}_h^{n,2}|_{\mathrm{S}}^2 + |\hat{\zeta}_h^{n,3}|_{\mathrm{S}}^2\big) \lesssim C_3(\underline{f},\underline{v})^2 T_f^2\tau^6 + \sum_{n=0}^{N-1}\tau\Big\{\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2 + \|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2 + \|\widehat{\psi}_h^{n,3}\|_{\mathrm{S}'}^2 + T_f\Big\{\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \tau^2\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + \tau^4\|\underline{\mathcal{D}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2\Big\}\Big\}, \tag{5.17}$$

where $C_3(\underline{f},\underline{v}) := \|\partial_{ttt}\underline{f}\|_{C^0(\overline{J};\frac{1}{\rho},\kappa)} + \mathfrak{c}_\infty\|\nabla\partial_{tt}\underline{f}\|_{C^0(\overline{J};\frac{1}{\rho},\kappa)} + \|\partial_{tttt}\underline{v}\|_{C^0(\overline{J};H^\nu(\rho,\frac{1}{\kappa};\mathcal{T}_h))}$.

The proof of Theorems 5.2 and 5.3 is postponed to Sections 6.2 and 6.3, respectively.

Remark 5.4 (Convergence rates)**.** *In addition to the assumptions of Theorem 5.2 for ERK2 (Theorem 5.3 for ERK3), assume that $\underline{v} \in C^\ell(\overline{J};\boldsymbol{H}^{k+2-\ell}(\Omega)\times H^{k+2-\ell}(\Omega))$ with $\ell \in \{1,2\}$ for ERK2 ($\ell \in \{1,2,3\}$ for ERK3). The operator $\underline{\mathcal{D}} = \underline{I}_{\mathcal{T}} - \underline{\Pi}_{\mathcal{T}}^k$ with the choice $\hat{\underline{I}}_h := \hat{\underline{I}}_h^{\mathrm{HHO}}$ gives $\underline{\mathcal{D}}^{n,1} = \underline{\mathcal{D}}^{n,2} = \underline{\mathcal{D}}^{n,3} = 0$. This and Remark 4.5 lead for ERK2 to*

$$\|\underline{\zeta}_{\mathcal{T}}^N\|_{\rho,\frac{1}{\kappa}} + \Big(\sum_{n=0}^{N-1}\tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{S}}^2 + |\hat{\zeta}_h^{n,2}|_{\mathrm{S}}^2\big)\Big)^{\frac{1}{2}} \lesssim \mathcal{O}(\tau^2 + h^{k+\frac{1}{2}}), \tag{5.18}$$

and for ERK3 to

$$\|\underline{\zeta}_{\mathcal{T}}^N\|_{\rho,\frac{1}{\kappa}} + \Big(\sum_{n=0}^{N-1}\tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{S}}^2 + |\hat{\zeta}_h^{n,2}|_{\mathrm{S}}^2 + |\hat{\zeta}_h^{n,3}|_{\mathrm{S}}^2\big)\Big)^{\frac{1}{2}} \lesssim \mathcal{O}(\tau^3 + h^{k+\frac{1}{2}}). \tag{5.19}$$

Instead, the definition (5.13b) of $\underline{\mathcal{D}}^{n,1}$ with the choice $\hat{\underline{I}}_h := \hat{\underline{I}}_h^{\mathrm{H}}$ followed by a triangle inequality implies that

$$\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}} \le \|\partial_t\underline{v}^n - \underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}(\partial_t\underline{v}^n)\|_{\rho,\frac{1}{\kappa}} + \|\partial_t\underline{v}^n - \underline{\Pi}_{\mathcal{T}}^k(\partial_t\underline{v}^n)\|_{\rho,\frac{1}{\kappa}},$$

and similarly for $\underline{\mathcal{D}}^{n,2}$ and $\underline{\mathcal{D}}^{n,3}$. This followed by the approximation estimates (4.20) and (3.6) for $\underline{\Pi}_{\mathcal{T}}^{\mathrm{H}}$ and $\underline{\Pi}_{\mathcal{T}}^k$, and Remark 4.8 improves the convergence rate on simplices to $\mathcal{O}(\tau^2 + h^{k+1})$ for ERK2 and to $\mathcal{O}(\tau^3 + h^{k+1})$ for ERK3.

Remark 5.5 (Improved CFL for $k = 0$ in ERK2)**.** *For the lowest-order polynomial degree $k = 0$, we can utilize the estimate*

$$\|\underline{\mathcal{A}}_{\mathcal{T}}(\hat{\underline{w}}_h)\|_{\frac{1}{\rho},\kappa} \lesssim \mathfrak{c}_\infty\|\nabla_{\mathcal{T}}w_{\mathcal{T}}\|_{\boldsymbol{L}^2(\frac{1}{\kappa};\Omega)} + \mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}|\hat{w}_h|_{\mathrm{S}} \qquad \forall\hat{\underline{w}}_h \in \hat{\underline{V}}_{h0}^k.$$

Then we can prove Theorem 5.2 under the usual CFL condition using techniques similar to [5, Theorem 3.2]. Details are skipped for brevity.

5.5 Rewriting using operators

Recall the operators $\underline{\mathcal{A}}_{\mathcal{T}}$ and $\mathcal{A}_{\mathcal{F}}$ introduced in (5.3). Let us also define an operator $\phi_{\mathcal{F}} : \underline{V}_{\mathcal{T}}^k \to V_{\mathcal{F}0}^k$ through the following identity: Given $\underline{w}_{\mathcal{T}} \in \underline{V}_{\mathcal{T}}^k$, find $\phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}}) \in V_{\mathcal{F}0}^k$ such that

$$\mathcal{A}_{\mathcal{F}}(\underline{0}, \phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})) = -\mathcal{A}_{\mathcal{F}}(\underline{w}_{\mathcal{T}}, 0). \tag{5.20}$$

(Finding $\phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})$ amounts to solving the well-posed problem $a_h((\underline{0}, \phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})), (\underline{0}, z_{\mathcal{F}})) = -a_h((\underline{w}_{\mathcal{T}}, 0), (\underline{0}, z_{\mathcal{F}}))$ for all $z_{\mathcal{F}} \in V_{\mathcal{F}0}^k$. This is computationally inexpensive since the matrix associated with $a_h(\underline{0}, \cdot, (\underline{0}, \cdot))$ is block-diagonal.) Using $\phi_{\mathcal{F}}$, we define the operator $\tilde{\underline{\mathcal{A}}}_{\mathcal{T}} : \underline{V}_{\mathcal{T}}^k \to \underline{V}_{\mathcal{T}}^k$ such that

$$\tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{w}_{\mathcal{T}}) := \underline{\mathcal{A}}_{\mathcal{T}}(\underline{w}_{\mathcal{T}}, \phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})) \qquad \forall \underline{w}_{\mathcal{T}} \in \underline{V}_{\mathcal{T}}^k. \tag{5.21}$$

The operator $\tilde{\underline{\mathcal{A}}}_{\mathcal{T}}$ can be understood as a Schur complement on the cell unknowns after eliminating the face unknowns. This operator is useful to gain a more algebraic view on ERK schemes. Indeed, we can rewrite the semi-discrete scheme defined in (3.16) as follows for all $t \in \underline{J}$:

$$\underline{\mathcal{M}}(\partial_t \underline{v}_{\mathcal{T}}(t)) = -\tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}(t)) + \underline{f}_{\mathcal{T}}(t), \quad \underline{f}_{\mathcal{T}}(t) := \underline{\Pi}_{\mathcal{T}}^k(\underline{f}(t)). \tag{5.22a}$$

This provides a general template to rewrite the time-discrete schemes. Indeed, given $\underline{v}_{\mathcal{T}}^{n,1} := \underline{v}_{\mathcal{T}}^n \in \underline{V}_{\mathcal{T}}^k$, the ERK2 scheme (5.8a)-(5.8b) is equivalent to finding $(\underline{v}_{\mathcal{T}}^{n,2}, \underline{v}_{\mathcal{T}}^{n+1}) \in \underline{V}_{\mathcal{T}}^k \times \underline{V}_{\mathcal{T}}^k$ in the following two steps:

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,2}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau \tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) + \tau \underline{f}_{\mathcal{T}}^{n,1}, \tag{5.23a}$$

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1}) = \frac{1}{2}\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1} + \underline{v}_{\mathcal{T}}^{n,2}) - \frac{1}{2}\tau \tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,2}) + \frac{1}{2}\tau \underline{g}_{\mathcal{T}}^{n,2}. \tag{5.23b}$$

Similarly, given $\underline{v}_{\mathcal{T}}^{n,1} := \underline{v}_{\mathcal{T}}^n \in \underline{V}_{\mathcal{T}}^k$, the ERK3 scheme (5.10a)-(5.10c) is equivalent to finding $(\underline{v}_{\mathcal{T}}^{n,2}, \underline{v}_{\mathcal{T}}^{n,3}, \underline{v}_{\mathcal{T}}^{n+1}) \in \underline{V}_{\mathcal{T}}^k \times \underline{V}_{\mathcal{T}}^k \times \underline{V}_{\mathcal{T}}^k$ in the following three steps:

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,2}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau \tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) + \tau \underline{f}_{\mathcal{T}}^{n,1}, \tag{5.24a}$$

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,3}) = \frac{1}{2}\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1} + \underline{v}_{\mathcal{T}}^{n,2}) - \frac{1}{2}\tau \tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,2}) + \frac{1}{2}\tau \underline{f}_{\mathcal{T}}^{n,2}, \tag{5.24b}$$

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1}) = \frac{1}{3}\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1} + \underline{v}_{\mathcal{T}}^{n,2} + \underline{v}_{\mathcal{T}}^{n,3}) - \frac{1}{3}\tau \tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,3}) + \frac{1}{3}\tau \underline{g}_{\mathcal{T}}^{n,3}. \tag{5.24c}$$

Remark 5.6 (Implementation)**.** *In (5.23), one needs to compute $v_{\mathcal{F}}^{n,1} := \phi_{\mathcal{F}}(\underline{v}_{\mathcal{T}}^{n,1})$ and $v_{\mathcal{F}}^{n,2} := \phi_{\mathcal{F}}(\underline{v}_{\mathcal{T}}^{n,2})$. In other words, (5.23) can be decomposed into four substeps: compute $v_{\mathcal{F}}^{n,1}$, compute $\underline{v}_{\mathcal{T}}^{n,2}$ using (5.23a), compute $v_{\mathcal{F}}^{n,2}$, compute $\underline{v}_{\mathcal{T}}^{n+1}$ using (5.23b). A similar comment can be made for ERK3, which can be decomposed into six substeps.*

5.6 Equivalent ERK schemes using Butcher arrays

In general, s-stage ERK schemes for $s \geq 2$ are represented by their Butcher arrays $\{a_{ij}\}_{i,j\in\{1:s\}}$,$\{b_i\}_{i\in\{1:s\}}$, and $\{c_i\}_{i\in\{1:s\}}$. We define the intermediate time steps $t^{n,i} := t^n + c_i\tau$ and set $\widetilde{f}^{n,i} := f(t^{n,i})$ for all $i \in \{1{:}s\}$. For explicit schemes, the Butcher matrix a is always strictly lower triangular, i.e., $a_{ij} = 0$ for all $i \leq j$. Also we have $c_1 = 0$ under Butcher's simplifying assumptions (see, e.g., [18, Chapter 78]), and consequently, $t^{n,1} = t^n$.

Given $\underline{v}_{\mathcal{T}}^{n,1} := \underline{v}_{\mathcal{T}}^n \in \underline{V}_{\mathcal{T}}^k$ from the previous time step or the initial condition, a general ERK scheme consists of finding $\underline{v}_{\mathcal{T}}^{n+1} := \underline{v}_{\mathcal{T}}^{n,s+1} \in \underline{V}_{\mathcal{T}}^k$ in the following s stages:

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,i}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau \sum_{j\in\{1:i-1\}} a_{ij}\big(\tilde{\underline{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,j}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,j}\big) \quad \forall i \in \{2{:}s\}, \tag{5.25a}$$

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau \sum_{j\in\{1:s\}} b_j\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,j}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,j}\big). \tag{5.25b}$$

For $s = 2$, this gives

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,2}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau a_{21}\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,1}\big), \tag{5.26a}$$

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,2}) - \tau\Big(b_1\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,1}\big) + b_2\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,2}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,2}\big)\Big). \tag{5.26b}$$

The second-order conditions are $c_1 = 0$, $c_2 = a_{21}$ from Butcher's simplifying assumptions, together with

$$b_1 + b_2 = 1, \quad b_2c_2 = \frac{1}{2}. \tag{5.27}$$

For $s = 3$, (5.25) gives

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,2}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau a_{21}\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,1}\big), \tag{5.28a}$$

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,3}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau\Big(a_{31}\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,1}\big) + a_{32}\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,2}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,2}\big)\Big), \tag{5.28b}$$

$$\begin{aligned}\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau\Big(&b_1\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,1}\big) + b_2\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,2}) - \underline{\widetilde{f}}_{\mathcal{T}}^{n,2}\big)\\ &+ b_3\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,3}) - \frac{1}{3}\tau\underline{\widetilde{f}}_{\mathcal{T}}^{n,3}\big)\Big).\end{aligned} \tag{5.28c}$$

The third-order conditions are $c_1 = 0$, $c_2 = a_{21}$, $c_3 = a_{31} + a_{32}$ from Butcher's simplifying assumptions, together with

$$b_1 + b_2 + b_3 = 1, \quad b_2c_2 + b_3c_3 = \frac{1}{2}, \quad b_2c_2^2 + b_3c_3^2 = \frac{1}{3}, \quad b_3a_{32}c_2 = \frac{1}{6}. \tag{5.29}$$

Lemma 5.7 (Comparison with ERK2 scheme)**.** *The sequence $(\underline{v}_{\mathcal{T}}^n)_{n\geq 0}$ generated by (5.26) with Butcher arrays satisfying the second-order condtions (5.27) is the same as the sequence $(\underline{v}_{\mathcal{T}}^n)_{n\geq 0}$ generated by the scheme (5.8) with*

$$\frac{1}{2}\underline{g}_{\mathcal{T}}^{n,2} := b_1\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_2\underline{\widetilde{f}}_{\mathcal{T}}^{n,2} - \frac{1}{2}\underline{f}_{\mathcal{T}}^{n}, \tag{5.30}$$

Consequently, we have

$$\|\underline{g}_{\mathcal{T}}^{n,2} - \underline{f}_{\mathcal{T}}^{n,2}\|_{\frac{1}{\rho},\kappa} \leq \tau^2\|\partial_{tt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)}. \tag{5.31}$$

Lemma 5.8 (Comparison with ERK3 scheme)**.** *The sequence $(\underline{v}_{\mathcal{T}}^n\}_{n\geq 0}$ generated by (5.28) with Butcher arrays satisfying the third-order conditions (5.29) is the same as the sequence $(\underline{v}_{\mathcal{T}}^n\}_{n\geq 0}$ generated by the scheme (5.10) with*

$$\begin{aligned}\frac{1}{3}\underline{g}_{\mathcal{T}}^{n,3} := {}& b_1\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_2\underline{\widetilde{f}}_{\mathcal{T}}^{n,2} + b_3\underline{\widetilde{f}}_{\mathcal{T}}^{n,3} - \frac{2}{3}\underline{f}_{\mathcal{T}}^{n} - \frac{1}{6}\tau\partial_t\underline{f}_{\mathcal{T}}^{n}\\ &- \tau\tilde{\mathcal{A}}_{\mathcal{T}}\underline{\mathcal{M}}^{-1}\big((b_2a_{21} + b_3a_{31})\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_3a_{32}\underline{\widetilde{f}}_{\mathcal{T}}^{n,2} - \tfrac{1}{2}\underline{f}_{\mathcal{T}}^{n} - \tfrac{1}{6}\tau\partial_t\underline{f}_{\mathcal{T}}^{n}\big).\end{aligned} \tag{5.32}$$

Consequently, we have

$$\|\underline{g}_{\mathcal{T}}^{n,3} - \underline{f}_{\mathcal{T}}^{n,3}\|_{\frac{1}{\rho},\kappa} \leq \tau^3\big(\chi_1\|\partial_{ttt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)} + \chi_2\mathfrak{c}_\infty\|\nabla\partial_{tt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)}\big), \tag{5.33}$$

where the constant χ_1 can depend on $\{b_j\}_{j\in\{2:3\}}$ if at least one b_j is negative and otherwise $\chi_1 = 1$, whereas the constant χ_2 is independent of the Butcher arrays.

The proof of Lemmas 5.7 and 5.8 is postponed to Section 6.4.

6 Proofs

In this section, we collect the proofs of the results stated in Section 5.

6.1 Verification of (A1)-(A3) for HHO methods

Proof of **(A1)**. For all $w_{\mathcal{T}} \in V_{\mathcal{T}}^k$, the definition of s_h using $\delta_{\partial T}(w_T, 0) = w_T|_{\partial T}$ for all $T \in \mathcal{T}_h$ shows that

$$|(w_{\mathcal{T}}, 0)|_{\mathrm{s}}^2 = s_h((w_{\mathcal{T}}, 0), (w_{\mathcal{T}}, 0)) = \sum_{T\in\mathcal{T}_h} \lambda_T^{-1}\|w_T\|_{L^2(\partial T)}^2$$
$$\lesssim \sum_{T\in\mathcal{T}_h} \lambda_T^{-1}h_T^{-1}\|w_T\|_{L^2(T)}^2 \lesssim \mathfrak{c}_\infty h^{-1}\|w_{\mathcal{T}}\|_{L^2(\frac{1}{\kappa};\Omega)}^2,$$

with a discrete trace inequality, $\lambda_T = \mathfrak{c}_T^{-1}\kappa_T$ from (3.10), and the quasi-uniformity assumption on the mesh in the last estimate. This concludes the proof of **(A1)**. □

Proof of **(A2)**. For all $\underline{\hat{w}}_h := (\boldsymbol{\tau}_{\mathcal{T}}, w_{\mathcal{T}}, w_{\mathcal{F}}) \in \underline{\hat{V}}_{h0}^k$ and all $z_{\mathcal{F}} \in V_{\mathcal{F}0}^k$, the definition (5.3) of $\mathcal{A}_{\mathcal{F}}$ gives

$$(\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h), z_{\mathcal{F}})_{L^2(\mathcal{F})} = (\boldsymbol{\tau}_{\mathcal{T}}, \boldsymbol{G}_{\mathcal{T}}(0, z_{\mathcal{F}}))_{\boldsymbol{L}^2(\Omega)} + s_h((w_{\mathcal{T}}, 0), (0, z_{\mathcal{F}})) + s_h((0, w_{\mathcal{F}}), (0, z_{\mathcal{F}})).$$

Choosing $z_{\mathcal{F}} := w_{\mathcal{F}}$, we obtain

$$|(0, w_{\mathcal{F}})|_{\mathrm{s}}^2 = -(\boldsymbol{\tau}_{\mathcal{T}}, \boldsymbol{G}_{\mathcal{T}}(0, w_{\mathcal{F}}))_{\boldsymbol{L}^2(\Omega)} - s_h((w_{\mathcal{T}}, 0), (0, w_{\mathcal{F}})) + (\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h), w_{\mathcal{F}})_{L^2(\mathcal{F})}$$
$$\le \|\boldsymbol{\tau}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{G}_{\mathcal{T}}(0, w_{\mathcal{F}})\|_{\boldsymbol{L}^2(\Omega)} + \big(|(w_{\mathcal{T}}, 0)|_{\mathrm{s}} + \|\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h)\|_{\mathrm{s}'}\big)|(0, w_{\mathcal{F}})|_{\mathrm{s}}. \tag{6.1}$$

Owing to the bound (3.13) on $\boldsymbol{G}_{\mathcal{T}}$, we infer that

$$\|\boldsymbol{G}_{\mathcal{T}}(0, w_{\mathcal{F}})\|_{\boldsymbol{L}^2(\Omega)}^2 \lesssim \sum_{T\in\mathcal{T}_h} \lambda_T h_T^{-1}|(0, w_{\mathcal{F}})|_{\mathrm{s},T}^2.$$

This estimate in (6.1) with $\lambda_T = \mathfrak{c}_T\rho_T$ from (3.10) proves that

$$|(0, w_{\mathcal{F}})|_{\mathrm{s}} \lesssim \mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}\|\boldsymbol{\tau}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\rho;\Omega)} + |(w_{\mathcal{T}}, 0)|_{\mathrm{s}} + \|\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h)\|_{\mathrm{s}'}. \tag{6.2}$$

The triangle inequality gives $|\hat{w}_h|_{\mathrm{s}} \le |(w_{\mathcal{T}}, 0)|_{\mathrm{s}} + |(0, w_{\mathcal{F}})|_{\mathrm{s}}$. This, followed by (6.2), **(A1)**, and the definition of $\|\underline{w}_{\mathcal{T}}\|_{\rho,\frac{1}{\kappa}}$, concludes the proof of **(A2)**. □

Proof of **(A3)**. For all $\underline{\hat{w}}_h := (\boldsymbol{\tau}_{\mathcal{T}}, w_{\mathcal{T}}, w_{\mathcal{F}}) \in \underline{\hat{V}}_{h0}^k$ and all $\underline{z}_{\mathcal{T}} := (\boldsymbol{\xi}_{\mathcal{T}}, z_{\mathcal{T}}) \in \underline{V}_{\mathcal{T}}^k$, the definition of $\underline{\mathcal{A}}_{\mathcal{T}}(\underline{\hat{w}}_h)$ leads to

$$\langle\underline{\mathcal{A}}_{\mathcal{T}}(\underline{\hat{w}}_h), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} = (\boldsymbol{\tau}_{\mathcal{T}}, \boldsymbol{G}_{\mathcal{T}}(z_{\mathcal{T}}, 0))_{\boldsymbol{L}^2(\Omega)} - (\boldsymbol{G}_{\mathcal{T}}(\hat{w}_h), \boldsymbol{\xi}_{\mathcal{T}})_{\boldsymbol{L}^2(\Omega)} + s_h(\hat{w}_h, (z_{\mathcal{T}}, 0))$$
$$\lesssim \|\boldsymbol{\tau}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{G}_{\mathcal{T}}(z_{\mathcal{T}}, 0)\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{G}_{\mathcal{T}}(\hat{w}_h)\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{\xi}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\Omega)} + \mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}|\hat{w}_h|_{\mathrm{s}}\|z_{\mathcal{T}}\|_{L^2(\frac{1}{\kappa};\Omega)},$$

with the Cauchy-Schwarz inequality and **(A1)** in the second step. Owing to the bound (3.13) on $\boldsymbol{G}_{\mathcal{T}}$, and the quasi-uniformity assumption on the mesh, we infer that

$$\langle\underline{\mathcal{A}}_{\mathcal{T}}(\underline{\hat{w}}_h), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}$$
$$\lesssim h^{-1}\|\boldsymbol{\tau}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\Omega)}\|z_{\mathcal{T}}\|_{L^2(\Omega)} + h^{-1}\|w_{\mathcal{T}}\|_{L^2(\Omega)}\|\boldsymbol{\xi}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\Omega)} + \mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}|\hat{w}_h|_s\big(\|\boldsymbol{\xi}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\rho;\Omega)} + \|z_{\mathcal{T}}\|_{L^2(\frac{1}{\kappa};\Omega)}\big)$$
$$\lesssim \big(\mathfrak{c}_\infty h^{-1}\|\underline{w}_{\mathcal{T}}\|_{\rho,\frac{1}{\kappa}} + \mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}|\hat{w}_h|_{\mathrm{s}}\big)\|\underline{z}_{\mathcal{T}}\|_{\rho,\frac{1}{\kappa}}, \tag{6.3}$$

where we used $\lambda_T = \mathfrak{c}_T\rho_T = \mathfrak{c}_T^{-1}\kappa_T$ from (3.10). Finally, invoking **(A2)** and the observation that $\|\underline{\mathcal{A}}_{\mathcal{T}}(\underline{\hat{w}}_h)\|_{\frac{1}{\rho},\kappa} = \sup_{\underline{z}_{\mathcal{T}}\in\underline{V}_{\mathcal{T}}^k} \frac{\langle\underline{\mathcal{A}}_{\mathcal{T}}(\underline{\hat{w}}_h), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}}{\|\underline{z}_{\mathcal{T}}\|_{\rho,\frac{1}{\kappa}}}$ concludes the proof of **(A3)**. □

6.2 Proof of Theorem 5.2

Lemma 6.1 (Error equation)**.** *Assume that* $\underline{v} \in C^1(\overline{J};\underline{Z}\cap\underline{V}_0)\cap C^3(\overline{J};\underline{Y})$*. The following holds: For all* $\hat{\underline{z}}_h \in \hat{\underline{V}}^k_{h0}$*,*

$$\langle \underline{\zeta}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} = \langle \underline{\zeta}_{\mathcal{T}}^{n,1}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} - \tau a_h(\hat{\underline{\zeta}}_h^{n,1}, \hat{\underline{z}}_h) - \tau\widehat{\psi}_h^{n,1}(\hat{z}_h) - \tau\langle \underline{\mathcal{D}}^{n,1}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}, \tag{6.4a}$$

$$\langle \underline{\zeta}_{\mathcal{T}}^{n+1}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} = \frac{1}{2}\langle \underline{\zeta}_{\mathcal{T}}^{n,1} + \underline{\zeta}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\tau a_h(\hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{z}}_h) - \frac{1}{2}\tau\widehat{\psi}_h^{n,2}(\hat{z}_h) - \frac{1}{2}\tau\langle \underline{\mathcal{D}}^{n,2}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\tau\langle \underline{\beta}_{\mathcal{T}}^{n}, \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}, \tag{6.4b}$$

where $\underline{\beta}_{\mathcal{T}}^{n} := \underline{g}_{\mathcal{T}}^{n,2} - \underline{f}_{\mathcal{T}}^{n,2} - \underline{\mathcal{M}}\,\underline{I}_{\mathcal{T}}(\underline{R}^{n,2})$ *with* $\underline{R}^{n,2} := \tau^{-1}\int_{J_n}(t_{n+1}-t)^2\partial_{ttt}\underline{v}(t)\,dt$*.*

Proof. (1) For all $\hat{\underline{z}}_h \in \hat{\underline{V}}^k_{h0}$, the L^2-orthogonality of $\underline{\Pi}^k_{\mathcal{T}}$ and the model problem (2.5) lead to

$$\begin{aligned}\langle \underline{\Pi}^k_{\mathcal{T}}(\partial_t\underline{v}(t)), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} &= \langle \partial_t\underline{v}(t), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} = \langle -\underline{\mathcal{B}}(\underline{v}(t)) + \underline{f}(t), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}\\ &= -a_h(\hat{\underline{I}}_h(\underline{v}(t)), \hat{\underline{z}}_h) + \widehat{\psi}_h(\underline{v}(t);\hat{z}_h) + \langle \underline{f}_{\mathcal{T}}(t), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}},\end{aligned} \tag{6.5}$$

with the definition (4.2) of ψ_h, **(I2a)**, and since $\underline{f}_{\mathcal{T}}(t) = \underline{\Pi}^k_{\mathcal{T}}(\underline{f}(t))$ in the last step. The definition of $\underline{v}^{n,1}$ and (6.5) evaluated at t^n yield, for all $\hat{\underline{z}}_h \in \hat{\underline{V}}^k_{h0}$,

$$\begin{aligned}\langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,2}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} &= \langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,1}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \tau\langle \underline{I}_{\mathcal{T}}(\partial_t\underline{v}^n), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}\\ &= \langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,1}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \tau\langle \underline{\mathcal{D}}^{n,1}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} - \tau a_h(\hat{\underline{I}}_h(\underline{v}^{n,1}), \hat{\underline{z}}_h) + \tau\widehat{\psi}_h^{n,1}(\hat{z}_h) + \tau\langle \underline{f}_{\mathcal{T}}^{n,1}, \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}.\end{aligned} \tag{6.6}$$

Subtracting the above equation from (5.8a) proves (6.4a).
(2) We start with the second-order Taylor expansion in time with exact remainder,

$$\underline{v}^{n+1} = \underline{v}^n + \tau\partial_t\underline{v}^n + \frac{1}{2}\tau^2\partial_{tt}\underline{v}^n + \frac{1}{2}\tau\underline{R}^{n,2} = \underline{v}^{n,2} + \frac{1}{2}\tau^2\partial_{tt}\underline{v}^n + \frac{1}{2}\tau\underline{R}^{n,2}.$$

An application of $\underline{I}_{\mathcal{T}}$ and (6.6) lead to

$$\begin{aligned}\langle \underline{I}_{\mathcal{T}}(\underline{v}^{n+1}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} &= \langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,2}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\tau^2\langle \underline{I}_{\mathcal{T}}(\partial_{tt}\underline{v}^n), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\tau\langle \underline{I}_{\mathcal{T}}(\underline{R}^{n,2}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}\\ &= \frac{1}{2}\langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,2} + \underline{v}^{n,1}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\tau a_h(\hat{\underline{I}}_h(\underline{v}^{n,1}), \hat{\underline{z}}_h) + \frac{1}{2}\tau\widehat{\psi}_h^{n,1}(\hat{z}_h) + \frac{1}{2}\tau\langle \underline{f}_{\mathcal{T}}^{n,1}, \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}\\ &\quad + \frac{1}{2}\tau^2\langle \underline{\Pi}^k_{\mathcal{T}}(\partial_{tt}\underline{v}^n), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\tau\langle \underline{\mathcal{D}}^{n,2}, \underline{z}^n_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\tau\langle \underline{I}_{\mathcal{T}}(\underline{R}^{n,2}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}},\end{aligned} \tag{6.7}$$

where we used that $\underline{\mathcal{D}}^{n,1} + \tau(\underline{I}_{\mathcal{T}}(\partial_{tt}\underline{v}^n) - \underline{\Pi}^k_{\mathcal{T}}(\partial_{tt}\underline{v}^n)) = \underline{\mathcal{D}}^{n,2}$. Differentiating (6.5) with respect to t, evaluating the result at t^n, and multiplying by τ gives

$$\begin{aligned}\tau\langle \underline{\Pi}^k_{\mathcal{T}}(\partial_{tt}\underline{v}^n), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} &= -\tau a_h(\hat{\underline{I}}_h(\partial_t\underline{v}^n), \hat{\underline{z}}_h) + \tau\widehat{\psi}_h(\partial_t\underline{v}^n;\hat{z}_h) + \tau\langle \partial_t\underline{f}^n_{\mathcal{T}}, \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}\\ &= -a_h(\hat{\underline{I}}_h(\underline{v}^{n,2} - \underline{v}^{n,1}), \hat{\underline{z}}_h) + \widehat{\psi}_h(\underline{v}^{n,2} - \underline{v}^{n,1};\hat{z}_h) + \tau\langle \partial_t\underline{f}^n_{\mathcal{T}}, \underline{z}_{\mathcal{T}}\rangle_{\underline{L}}.\end{aligned}$$

This simplifies (6.7) to

$$\begin{aligned}\langle \underline{I}_{\mathcal{T}}(\underline{v}^{n+1}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} &= \frac{1}{2}\langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,2} + \underline{v}^{n,1}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\tau a_h(\hat{\underline{I}}_h(\underline{v}^{n,2}), \hat{\underline{z}}_h) + \frac{1}{2}\tau\widehat{\psi}_h^{n,2}(\hat{z}_h)\\ &\quad + \frac{1}{2}\tau\langle \underline{f}_{\mathcal{T}}^{n,2}, \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} + \frac{1}{2}\tau\langle \underline{\mathcal{D}}^{n,2}, \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\tau\langle \underline{I}_{\mathcal{T}}(\underline{R}^{n,2}), \underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}.\end{aligned}$$

Subtracting this from (5.8b) gives (6.4b). □

Lemma 6.2 (Energy identity). *Under the assumptions of Lemma 6.1, the following identity holds:*

$$\|\underline{\zeta}_{\mathcal{T}}^{n+1}\|_{\rho,\frac{1}{\kappa}}^2 - \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + |\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2\big) = \|\underline{\zeta}_{\mathcal{T}}^{n+1} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \tau\widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,1}) - \tau\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,2})$$
$$-\tau\langle\underline{\mathcal{D}}^{n,1},\underline{\zeta}_{\mathcal{T}}^{n,1}\rangle_{\rho,\frac{1}{\kappa}} - \tau\langle\underline{\mathcal{D}}^{n,2},\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} + \tau\langle\underline{\beta}_{\mathcal{T}}^{n},\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\underline{L}}. \tag{6.8}$$

Proof. Choosing $\hat{\underline{z}}_h := \hat{\underline{\zeta}}_h^{n,1}$ in (6.4a) and $\hat{\underline{z}}_h := 2\hat{\underline{\zeta}}_h^{n,2}$ in (6.4b), elementary algebraic manipulations lead to

$$\|\underline{\zeta}_{\mathcal{T}}^{n+1}\|_{\rho,\frac{1}{\kappa}}^2 - \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 = \|\underline{\zeta}_{\mathcal{T}}^{n+1} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \tau a_h(\hat{\underline{\zeta}}_h^{n,1},\hat{\underline{\zeta}}_h^{n,1}) - \tau a_h(\hat{\underline{\zeta}}_h^{n,2},\hat{\underline{\zeta}}_h^{n,2}) - \tau\widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,1}) - \tau\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,2})$$
$$-\tau\langle\underline{\mathcal{D}}^{n,1},\underline{\zeta}_{\mathcal{T}}^{n,1}\rangle_{\rho,\frac{1}{\kappa}} - \tau\langle\underline{\mathcal{D}}^{n,2},\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} + \tau\langle\underline{\beta}_{\mathcal{T}}^{n},\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\underline{L}}.$$

Since $a_h(\hat{\underline{\zeta}}_h^{n,1},\hat{\underline{\zeta}}_h^{n,1}) = |\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2$ and $a_h(\hat{\underline{\zeta}}_h^{n,2},\hat{\underline{\zeta}}_h^{n,2}) = |\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2$ owing to (3.15), this concludes the proof. □

Lemma 6.3 (Preliminary stability estimate). *Under the assumptions of Lemma 6.1, we have*

$$\|\underline{\zeta}_{\mathcal{T}}^{n+1}\|_{\rho,\frac{1}{\kappa}}^2 - \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \frac{1}{2}\tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + |\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2\big) \le \|\underline{\zeta}_{\mathcal{T}}^{n+1} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2$$
$$+ C\tau\Big\{T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'}^2 + \|\widehat{\psi}_h^{n,2}\|_{\mathrm{s}'}^2 + T_f\big(C_2^n(\underline{f},\underline{v})^2\tau^4 + \|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2\big)\Big\}, \tag{6.9}$$

with $C_2^n(\underline{f},\underline{v}) := \|\partial_{tt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)} + \|\partial_{ttt}\underline{v}\|_{C^0(\overline{J}_n;H^\nu(\rho,\frac{1}{\kappa};\mathcal{T}_h))}$.

Proof. The proof follows by estimating the last four terms on the right hand-side of the energy identity (6.8). The Cauchy-Schwarz inequality and Young's inequality lead to

$$\tau|\widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,1}) + \widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,2}) + \langle\underline{\mathcal{D}}^{n,1},\underline{\zeta}_{\mathcal{T}}^{n,1}\rangle_{\rho,\frac{1}{\kappa}} + \langle\underline{\mathcal{D}}^{n,2},\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} + \langle\underline{\beta}_{\mathcal{T}}^{n},\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\underline{L}}|$$
$$\le \frac{1}{2}\tau\Big\{\|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'}^2 + \|\widehat{\psi}_h^{n,2}\|_{\mathrm{s}'}^2 + |\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + |\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 + T_f\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + T_f\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2$$
$$+ 2T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + T_f\|\underline{\beta}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa}^2\Big\}. \tag{6.10}$$

Notice that this is the critical step where we bound the consistency errors in the $\|\cdot\|_{\mathrm{s}'}$-norm. It remains to bound $\|\underline{\beta}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa}$ and $\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}$. The definition of $\underline{\beta}_{\mathcal{T}}^{n}$, the triangle inequality, the assumption (5.9) on $\underline{g}_{\mathcal{T}}^{n,2} - \underline{f}_{\mathcal{T}}^{n,2}$, and **(I1)** imply that

$$\|\underline{\beta}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa} \le \|\underline{g}_{\mathcal{T}}^{n,2} - \underline{f}_{\mathcal{T}}^{n,2}\|_{\frac{1}{\rho},\kappa} + \|\underline{I}_{\mathcal{T}}(\underline{R}^{n,2})\|_{\rho,\frac{1}{\kappa}}$$
$$\lesssim C_2^n(\underline{f})\tau^2 + \|\underline{R}^{n,2}\|_{H^\nu(\rho,\frac{1}{\kappa};\mathcal{T}_h)} \lesssim C_2^n(\underline{f},\underline{v})\tau^2. \tag{6.11}$$

The definition of the $\|\cdot\|_{\mathrm{s}'}$-norm and the property **(A1)** imply that, for all $z_{\mathcal{T}} \in V_{\mathcal{T}}^k$,

$$|\widehat{\psi}_h^{n,1}(z_{\mathcal{T}},0)| \le \|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'}\|(z_{\mathcal{T}},0)\|_{\mathrm{s}} \lesssim \mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'}\|z_{\mathcal{T}}\|_{L^2(\frac{1}{\kappa};\Omega)}. \tag{6.12}$$

Using the error equation (6.4a) with $\hat{\underline{z}}_h = (\underline{\zeta}_{\mathcal{T}}^{n,2},0)$, and (6.12) leads to

$$\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}} \lesssim \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}} + \tau\|\underline{\mathcal{A}}_{\mathcal{T}}(\hat{\underline{\zeta}}_h^{n,1})\|_{\frac{1}{\rho},\kappa} + \tau\mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'} + \tau\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}.$$

Since $(\mathcal{A}_{\mathcal{F}}(\hat{\underline{\zeta}}_h^{n,1}),z_{\mathcal{F}})_{L^2(\mathcal{F})} = \widehat{\psi}_h^{n,1}(0,z_{\mathcal{F}})$ for all $z_{\mathcal{F}} \in V_{\mathcal{F}0}^k$ owing to (6.4a), we have $\|\mathcal{A}_{\mathcal{F}}(\hat{\underline{\zeta}}_h^{n,1})\|_{\mathrm{s}'} = \|\widehat{\psi}_h^{n,1}(0,\cdot)\|_{\mathrm{s}'} \le \|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'}$. Invoking **(A3)** then shows that

$$\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}} \lesssim \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}} + \tau\mathfrak{c}_\infty h^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}} + \tau\mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'} + \tau\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}. \tag{6.13}$$

Recalling the CFL condition (5.1) and since $\tau \le T_f$, we infer that

$$\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}} \lesssim \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}} + T_f^{\frac{1}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'} + T_f\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}. \tag{6.14}$$

This and (6.11) in (6.10) conclude the proof. □

Proof of Theorem 5.2. The first step is to control the anti-dissipative term $\|\underline{\zeta}_{\mathcal{T}}^{n+1}-\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2$ in the stability estimate from Lemma 6.3. The combination (6.4b) $-\frac{1}{2}\times$ (6.4a) gives, for all $\hat{\underline{z}}_h\in\hat{\underline{V}}_{h0}^k$,

$$\begin{aligned}\langle\underline{\zeta}_{\mathcal{T}}^{n+1}-\underline{\zeta}_{\mathcal{T}}^{n,2},\underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}&=\frac{1}{2}\tau a_h(\hat{\underline{\zeta}}_h^{n,1}-\hat{\underline{\zeta}}_h^{n,2},\hat{\underline{z}}_h)+\frac{1}{2}\tau\widehat{\psi}_h^{n,1}(\hat{\underline{z}}_h)-\frac{1}{2}\tau\widehat{\psi}_h^{n,2}(\hat{\underline{z}}_h)+\frac{1}{2}\tau\langle\underline{\mathcal{D}}^{n,1},\underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}\\&\quad-\frac{1}{2}\tau\langle\underline{\mathcal{D}}^{n,2},\underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}+\frac{1}{2}\tau\langle\underline{\beta}_{\mathcal{T}}^{n},\underline{z}_{\mathcal{T}}\rangle_{\underline{L}}.\end{aligned}\tag{6.15}$$

Analogously to Lemma 6.3, the identity (6.15) with $\hat{\underline{z}}_h:=(\underline{\zeta}_{\mathcal{T}}^{n+1}-\underline{\zeta}_{\mathcal{T}}^{n,2},0)$ followed by (6.12), the property **(A3)** for $\underline{\hat{w}}_h=\hat{\underline{\zeta}}_h^{n,1}-\hat{\underline{\zeta}}_h^{n,2}$, and $\mathcal{A}_{\mathcal{F}}(\hat{\underline{\zeta}}_h^{n,1}-\hat{\underline{\zeta}}_h^{n,2})=\widehat{\psi}_h^{n,1}(0,\cdot)-\widehat{\psi}_h^{n,2}(0,\cdot)$ show that

$$\|\underline{\zeta}_{\mathcal{T}}^{n+1}-\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}\lesssim\tau\Big(\mathfrak{c}_\infty h^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}-\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}+\mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}\|\widehat{\psi}_h^{n,1}-\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}+\|\underline{\mathcal{D}}^{n,1}-\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}+\|\underline{\beta}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa}\Big).\tag{6.16}$$

From (6.4a), we have, for all $\hat{\underline{z}}_h:=(\underline{z}_{\mathcal{T}},z_{\mathcal{F}})\in\hat{\underline{V}}_{h0}^k$,

$$\langle\underline{\zeta}_{\mathcal{T}}^{n,1}-\underline{\zeta}_{\mathcal{T}}^{n,2},\underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}}=\tau a_h(\hat{\underline{\zeta}}_h^{n,1},\hat{\underline{z}}_h)+\tau\widehat{\psi}_h^{n,1}(\hat{z}_h)+\tau\langle\underline{\mathcal{D}}^{n,1},\underline{z}_{\mathcal{T}}\rangle_{\rho,\frac{1}{\kappa}},\quad(\mathcal{A}_{\mathcal{F}}(\hat{\underline{\zeta}}_h^{n,1}),z_{\mathcal{F}})_{L^2(\mathcal{F})}=\widehat{\psi}_h^{n,1}(0,z_{\mathcal{F}}).$$

Similar arguments to (6.16) show that

$$\|\underline{\zeta}_{\mathcal{T}}^{n,1}-\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}\lesssim\tau\Big(\mathfrak{c}_\infty h^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+\mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}+\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}\Big).$$

Using this in (6.16), we arrive at

$$\begin{aligned}\|\underline{\zeta}_{\mathcal{T}}^{n+1}-\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}&\lesssim\tau^2\mathfrak{c}_\infty^2h^{-2}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+\tau^{\frac{1}{2}}(\tau\mathfrak{c}_\infty h^{-1})^{\frac{3}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}+\tau^{\frac{1}{2}}(\tau\mathfrak{c}_\infty h^{-1})^{\frac{1}{2}}\|\widehat{\psi}_h^{n,1}-\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}\\&\quad+\tau(\tau\mathfrak{c}_\infty h^{-1}+1)\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+\tau\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}+\tau\|\underline{\beta}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa}\\&\lesssim\tau^2\mathfrak{c}_\infty^2h^{-2}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+\tau^{\frac{1}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}+\tau^{\frac{1}{2}}\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}+\tau\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+\tau\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}\\&\quad+\tau\|\underline{\beta}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa},\end{aligned}\tag{6.17}$$

where we used the usual CFL condition (5.1) and the triangle inequality in the last step. Squaring both sides of (6.17), observing that $\tau^3\mathfrak{c}_\infty^4h^{-4}\le\rho_{\frac{4}{3}}^3T_f^{-1}$ owing to the strengthened $\frac{4}{3}$-CFL condition (5.2), and $\tau\le T_f$, we obtain

$$\begin{aligned}\|\underline{\zeta}_{\mathcal{T}}^{n+1}-\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2&\lesssim\tau\Big(T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2+\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2+T_f\big(\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\beta}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa}^2\big)\Big)\\&\lesssim\tau\Big(T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2+\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2+T_f\big(\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2+C^{n,2}(\underline{f},\underline{v})^2\tau^4\big)\Big),\end{aligned}$$

where we used (6.11) in the last step. This in (6.9) leads to

$$\begin{aligned}&\|\underline{\zeta}_{\mathcal{T}}^{n+1}\|_{\rho,\frac{1}{\kappa}}^2-\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\frac{1}{2}\tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{S}}^2+|\hat{\zeta}_h^{n,2}|_{\mathrm{S}}^2\big)\lesssim\tau T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2\\&\quad+\tau\Big(T_f(C^{n,2}(\underline{f},\underline{v})^2\tau^4+\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2)+\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2+\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2\Big).\end{aligned}$$

Using a discrete Gronwall's lemma (observing that $\tau N=T_f$) and invoking the initial condition $\underline{\zeta}_{\mathcal{T}}^0=\underline{0}$, we infer that

$$\begin{aligned}&\|\underline{\zeta}_{\mathcal{T}}^{N}\|_{\rho,\frac{1}{\kappa}}^2+\sum_{n=0}^{N-1}\frac{1}{2}\tau\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{S}}^2+|\hat{\zeta}_h^{n,2}|_{\mathrm{S}}^2\big)\\&\lesssim C_2(\underline{f},\underline{v})^2T_f^2\tau^4+\sum_{n=0}^{N-1}\tau\Big\{\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2+\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2+T_f\big(\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2\big)\Big\}.\end{aligned}$$

This completes the proof. □

6.3 Proof of Theorem 5.3

Lemma 6.4 (Error equation)**.** *Assume that* $\underline{v} \in C^2(\overline{J}; \underline{Z} \cap \underline{V}_0) \cap C^4(\overline{J}; \underline{Y})$*. The following holds: For all* $\hat{\underline{z}}_h \in \hat{\underline{V}}^k_{h0}$*,*

$$\langle \underline{\zeta}^{n,2}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = \langle \underline{\zeta}^{n,1}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \tau a_h(\hat{\underline{\zeta}}^{n,1}_h, \hat{\underline{z}}_h) - \tau \widehat{\psi}^{n,1}_h(\hat{z}_h) - \tau \langle \underline{\mathcal{D}}^{n,1}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}}, \tag{6.18a}$$

$$\langle \underline{\zeta}^{n,3}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = \frac{1}{2}\langle \underline{\zeta}^{n,1}_{\mathcal{T}} + \underline{\zeta}^{n,2}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\tau a_h(\hat{\underline{\zeta}}^{n,2}_h, \hat{\underline{z}}_h) - \frac{1}{2}\tau \widehat{\psi}^{n,2}_h(\hat{z}_h) - \frac{1}{2}\tau \langle \underline{\mathcal{D}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}}, \tag{6.18b}$$

$$\begin{aligned}\langle \underline{\zeta}^{n+1}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} &= \frac{1}{3}\langle \underline{\zeta}^{n,1}_{\mathcal{T}} + \underline{\zeta}^{n,2}_{\mathcal{T}} + \underline{\zeta}^{n,3}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{3} a_h(\hat{\underline{\zeta}}^{n,3}_h, \hat{\underline{z}}_h) - \frac{1}{3}\tau \widehat{\psi}^{n,3}_h(\hat{z}_h) - \frac{1}{3}\tau \langle \underline{\mathcal{D}}^{n,3}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} \\ &\quad + \frac{1}{3}\tau \langle \underline{\gamma}^n_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}},\end{aligned} \tag{6.18c}$$

with $\underline{\gamma}^n_{\mathcal{T}} := \underline{g}^{n,3}_{\mathcal{T}} - \underline{f}^{n,3}_{\mathcal{T}} - \underline{\mathcal{M}}\, \underline{I}_{\mathcal{T}}(\underline{R}^{n,3})$ *and* $\underline{R}^{n,3} := \tau^{-1} \int_{J_n} \frac{1}{2}(t_{n+1} - t)^3 \partial_{tttt} \underline{v}(t)\, dt$*.*

Proof. The error equations (6.18a)-(6.18b) can be obtained similarly to (6.4a)-(6.4b). For the third error equation, we start with the third-order Taylor expansion in time with exact remainder,

$$\underline{v}^{n+1} = \underline{v}^n + \tau \partial_t \underline{v}^n + \frac{1}{2}\tau^2 \partial_{tt} \underline{v}^n + \frac{1}{6}\tau^3 \partial_{ttt} \underline{v}^n + \frac{1}{3}\tau \underline{R}^{n,3} = \underline{v}^{n,3} + \frac{1}{6}\tau^3 \partial_{ttt} \underline{v}^n + \frac{1}{3}\tau \underline{R}^{n,3}.$$

An application of $\underline{I}_{\mathcal{T}}$ to the above identity, (6.7), an application of ∂_{tt} to (6.5), and elementary algebra lead to

$$\begin{aligned}&\langle \underline{I}_{\mathcal{T}}(\underline{v}^{n+1}), \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} \\ &= \langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,3}), \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{6}\tau^3 \langle \underline{\mathcal{D}}^{n,3}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{6}\tau^3 \langle \underline{\Pi}^k_{\mathcal{T}}(\partial_{ttt} \underline{v}^n, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{3}\tau \langle \underline{I}_{\mathcal{T}}(\underline{R}^{n,3}), \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} \\ &= \frac{1}{3}\langle \underline{I}_{\mathcal{T}}(\underline{v}^{n,1} + \underline{v}^{n,2} + \underline{v}^{n,3}), \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{3}\tau a_h(\hat{I}_h(\underline{v}^{n,3}), \hat{\underline{z}}_h) + \frac{1}{3}\tau \widehat{\psi}^{n,3}_h(\hat{z}_h) + \frac{1}{3}\tau \langle \underline{\mathcal{D}}^{n,3}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} \\ &\quad + \frac{1}{3}\tau \langle \underline{f}^{n,3}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}} + \frac{1}{3}\tau \langle \underline{I}_{\mathcal{T}}(\underline{R}^{n,3}), \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}}.\end{aligned}$$

Subtracting this from (5.10c) proves (6.18c). □

Lemma 6.5 (Energy identity)**.** *Under the assumptions of Lemma 6.4, the following identity holds:*

$$\begin{aligned}&\frac{1}{2}\|\underline{\zeta}^{n+1}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\|\underline{\zeta}^{n,1}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\tau|\hat{\zeta}^{n,1}_h|^2_{\mathrm{s}} + \frac{1}{6}\tau|\hat{\zeta}^{n,2}_h|^2_{\mathrm{s}} + \frac{1}{3}\tau|\hat{\zeta}^{n,3}_h|^2_{\mathrm{s}} + \frac{1}{6}\|\underline{\zeta}^{n,3}_{\mathcal{T}} - \underline{\zeta}^{n,2}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}} \\ &= \frac{1}{6}\tau|\hat{\zeta}^{n,1}_h - \hat{\zeta}^{n,2}_h|^2_{\mathrm{s}} + \frac{1}{2}\|\underline{\zeta}^{n+1}_{\mathcal{T}} - \underline{\zeta}^{n,3}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}} - \frac{1}{6}\tau \widehat{\psi}^{n,1}_h(2\hat{\zeta}^{n,1}_h + \hat{\zeta}^{n,2}_h) - \frac{1}{6}\tau \widehat{\psi}^{n,2}_h(\hat{\zeta}^{n,1}_h) - \frac{1}{3}\tau \widehat{\psi}^{n,3}_h(\hat{\zeta}^{n,3}_h) \\ &\quad - \frac{1}{6}\tau \langle \underline{\mathcal{D}}^{n,1}, 2\underline{\zeta}^{n,1}_{\mathcal{T}} + \underline{\zeta}^{n,2}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{6}\tau \langle \underline{\mathcal{D}}^{n,2}, \underline{\zeta}^{n,1}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{3}\tau \langle \underline{\mathcal{D}}^{n,3}, \underline{\zeta}^{n,3}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{3}\tau \langle \underline{\gamma}^n_{\mathcal{T}}, \underline{\zeta}^{n,3}_{\mathcal{T}} \rangle_{\underline{L}}.\end{aligned} \tag{6.19}$$

Proof. The algebraic manipulations (6.18b) $-\frac{1}{2} \times$ (6.18a) and (6.18c) $-\frac{2}{3} \times$ (6.18b) show that

$$\langle \underline{\zeta}^{n,3}_{\mathcal{T}} - \underline{\zeta}^{n,2}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} = -\frac{1}{2}\tau a_h(\hat{\underline{\zeta}}^{n,2}_h - \hat{\underline{\zeta}}^{n,1}_h, \hat{\underline{z}}_h) - \frac{1}{2}\tau(\widehat{\psi}^{n,2}_h(\hat{z}_h) - \widehat{\psi}^{n,1}_h(\hat{z}_h)) - \frac{1}{2}\tau \langle \underline{\mathcal{D}}^{n,2} - \underline{\mathcal{D}}^{n,1}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}}, \tag{6.20a}$$

$$\begin{aligned}\langle \underline{\zeta}^{n+1}_{\mathcal{T}} - \underline{\zeta}^{n,3}_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} &= -\frac{1}{3}\tau a_h(\hat{\underline{\zeta}}^{n,3}_h - \hat{\underline{\zeta}}^{n,2}_h, \hat{\underline{z}}_h) - \frac{1}{3}\tau(\widehat{\psi}^{n,3}_h(\hat{z}_h) - \widehat{\psi}^{n,2}_h(\hat{z}_h)) - \frac{1}{3}\tau \langle \underline{\mathcal{D}}^{n,3} - \underline{\mathcal{D}}^{n,2}, \underline{z}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} \\ &\quad + \frac{1}{3}\tau \langle \underline{\gamma}^n_{\mathcal{T}}, \underline{z}_{\mathcal{T}} \rangle_{\underline{L}}.\end{aligned} \tag{6.20b}$$

Set $A := \frac{1}{2}\|\underline{\zeta}^{n+1}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\|\underline{\zeta}^{n+1}_{\mathcal{T}} - \underline{\zeta}^{n,3}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\|\underline{\zeta}^{n,1}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}}$. Since $\frac{1}{2}a^2 - \frac{1}{2}(a-b)^2 = (a-b)b + \frac{b^2}{2}$ for any real numbers a and b, we obtain

$$A = \langle \underline{\zeta}^{n+1}_{\mathcal{T}} - \underline{\zeta}^{n,3}_{\mathcal{T}}, \underline{\zeta}^{n,3}_{\mathcal{T}} \rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{2}\|\underline{\zeta}^{n,3}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\|\underline{\zeta}^{n,1}_{\mathcal{T}}\|^2_{\rho,\frac{1}{\kappa}}.$$

Testing (6.20b) with $\hat{\underline{z}}_h := \hat{\underline{\zeta}}_h^{n,3}$ gives

$$A = -\frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{\zeta}}_h^{n,3}) + D_1 + \frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2 - \frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2,$$

$D_1 := -\frac{1}{3}\tau(\widehat{\psi}_h^{n,3}(\hat{\zeta}_h^{n,3}) - \widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,3})) - \frac{1}{3}\tau\langle\underline{\mathcal{D}}^{n,3} - \underline{\mathcal{D}}^{n,2}, \underline{\zeta}_{\mathcal{T}}^{n,3}\rangle_{\rho,\frac{1}{\kappa}} + \frac{1}{3}\tau\langle\underline{\gamma}_{\mathcal{T}}^{n}, \underline{\zeta}_{\mathcal{T}}^{n,3}\rangle_{\underline{L}}$. Choosing $\underline{z}_{\mathcal{T}} := \underline{\zeta}_{\mathcal{T}}^{n,1}$ (resp. $\underline{z}_{\mathcal{T}} := 2\underline{\zeta}_{\mathcal{T}}^{n,2}$) in (6.18a) (resp. (6.18b)), elementary algebraic manipulations lead to

$$\begin{aligned}\|\underline{\zeta}_{\mathcal{T}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2 - \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 = \|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 - \tau|\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 - \tau\widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,1}) - \tau\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,2})\\ - \tau\langle\underline{\mathcal{D}}^{n,1}, \underline{\zeta}_{\mathcal{T}}^{n,1}\rangle_{\rho,\frac{1}{\kappa}} - \tau\langle\underline{\mathcal{D}}^{n,2}, \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}}.\end{aligned}$$

Defining $D_2 := -\frac{1}{2}\tau\widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,1}) - \frac{1}{2}\tau\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,2}) - \frac{1}{2}\tau\langle\underline{\mathcal{D}}^{n,1}, \underline{\zeta}_{\mathcal{T}}^{n,1}\rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{2}\tau\langle\underline{\mathcal{D}}^{n,2}, \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} + D_1$, we arrive at

$$\begin{aligned}A &= -\frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{\zeta}}_h^{n,3}) + \frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \frac{1}{2}\tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 - \frac{1}{2}\tau|\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 + D_2\\ &= -\frac{1}{3}\tau|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2 + \frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{\zeta}}_h^{n,3}) + \frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \frac{1}{2}\tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 - \frac{1}{2}\tau|\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 + D_2\\ &= -\frac{1}{3}\tau|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2 + \frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}) - \frac{1}{6}\tau|\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 + \frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \frac{1}{2}\tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + D_2,\end{aligned}$$

since $a_h(\hat{\underline{\zeta}}_h^{n,3}, \hat{\underline{\zeta}}_h^{n,3}) = |\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2$ and $a_h(\hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{\zeta}}_h^{n,2}) = |\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2$ owing to (3.15). Let

$$\begin{aligned}B :&= \frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}) + \frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2\\ &= \frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,2}, \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}) - \frac{1}{6}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + \frac{2}{3}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2.\end{aligned}$$

Testing (6.20a) with $\hat{\underline{z}}_h := \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}$, we see that

$$\begin{aligned}\frac{2}{3}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 = -\frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,2} - \hat{\underline{\zeta}}_h^{n,1}, \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}) - \frac{1}{3}\tau(\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}) - \widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}))\\ - \frac{1}{3}\tau\langle\underline{\mathcal{D}}^{n,2} - \underline{\mathcal{D}}^{n,1}, \underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}}.\end{aligned}$$

This implies that

$$\begin{aligned}B &= \frac{1}{3}\tau a_h(\hat{\underline{\zeta}}_h^{n,1}, \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}) - \frac{1}{6}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \frac{1}{3}\tau(\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}) - \widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}))\\ &\quad - \frac{1}{3}\tau\langle\underline{\mathcal{D}}^{n,2} - \underline{\mathcal{D}}^{n,1}, \underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}}\\ &= \frac{1}{3}\langle\underline{\zeta}_{\mathcal{T}}^{n,1} - \underline{\zeta}_{\mathcal{T}}^{n,2}, \underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} - \frac{1}{6}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 - \frac{1}{3}\tau\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}) - \frac{1}{3}\tau\langle\underline{\mathcal{D}}^{n,2}, \underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}},\end{aligned}$$

where we used (6.18a) with $\hat{\underline{z}}_h := \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}$ in the last step. Hence, for $G := -\frac{1}{2}\tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 - \frac{1}{6}\tau|\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 - \frac{1}{3}\tau|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2 - \frac{1}{6}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2$ and $D_3 := -\frac{1}{3}\tau\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}) - \frac{1}{3}\tau\langle\underline{\mathcal{D}}^{n,2}, \underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} + D_2$, the term A simplifies to

$$\begin{aligned}A &= -\frac{1}{2}\tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 - \frac{1}{6}\tau|\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 - \frac{1}{3}\tau|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2 + B + D_2\\ &= G + \frac{1}{3}\langle\underline{\zeta}_{\mathcal{T}}^{n,1} - \underline{\zeta}_{\mathcal{T}}^{n,2}, \underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} + D_3\\ &= G - \frac{1}{6}\tau a_h(\hat{\underline{\zeta}}_h^{n,2} - \hat{\underline{\zeta}}_h^{n,1}, \hat{\underline{\zeta}}_h^{n,1} - \hat{\underline{\zeta}}_h^{n,2}) - \frac{1}{6}\tau(\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,1} - \hat{\zeta}_h^{n,2}) - \widehat{\psi}_h^{n,1}(\hat{\zeta}_h^{n,1} - \hat{\zeta}_h^{n,2}))\\ &\quad - \frac{1}{6}\tau\langle\underline{\mathcal{D}}^{n,2} - \underline{\mathcal{D}}^{n,1}, \underline{\zeta}_{\mathcal{T}}^{n,1} - \underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}} + D_3,\end{aligned}$$

using (6.20a) with $\hat{\underline{z}}_h := \hat{\underline{\zeta}}_h^{n,1} - \hat{\underline{\zeta}}_h^{n,2}$ in the third step. Rearranging the above terms concludes the proof. □

Lemma 6.6 (Stability). *Under the assumption* $\underline{v} \in C^0(\overline{J};\underline{V}_0)\cap C^4(\overline{J};\underline{Y})$ *and the CFL condition* (5.1) *with the choice* (5.16) *of* ρ_1*, the following holds*

$$\begin{aligned}
&\|\underline{\zeta}_{\mathcal{T}}^{n+1}\|_{\rho,\frac{1}{\kappa}}^2 - \|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \frac{1}{24}\tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + \frac{1}{6}\tau|\hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2 + \frac{1}{24}\tau|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2\\
&\lesssim \tau\Big\{T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \|\widehat{\psi}_h^{n,1}\|_{\mathrm{s}'} + \|\widehat{\psi}_h^{n,2}\|_{\mathrm{s}'} + \|\widehat{\psi}_h^{n,3}\|_{\mathrm{s}'}^2\\
&\quad + T_f\big(C_3^n(\underline{f},\underline{v})^2\tau^6 + \|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2 + \|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + \|\underline{\mathcal{D}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2\big)\Big\},
\end{aligned} \tag{6.21}$$

with $C_3^n(\underline{f},\underline{v}) = \|\partial_{ttt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)} + \mathfrak{c}_\infty\|\nabla\partial_{tt}\underline{f}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)} + \|\partial_{tttt}\underline{v}\|_{C^0(\overline{J}_n;H^\nu(\rho,\frac{1}{\kappa};\mathcal{T}_h))}$.

Proof. The aim is to bound each term on the right-hand side of the energy identity (6.19).

(1) Bound $\frac{1}{6}\tau|\hat{\zeta}_h^{n,2} - \hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2$. For appropriate choices of ϵ and $\hat{\epsilon}$ (to be chosen later), we observe that

$$\begin{aligned}
|\hat{\zeta}_h^{n,2} - \hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 &\le (1+\epsilon)|\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + (1+\epsilon^{-1})|\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2\\
&\le (1+\epsilon)(1+\hat{\epsilon})|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2 + (1+\epsilon)(1+\hat{\epsilon}^{-1})|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + (1+\epsilon^{-1})|\hat{\zeta}_h^{n,3} - \hat{\zeta}_h^{n,2}|_{\mathrm{s}}^2\\
&\le (1+\epsilon)(1+\hat{\epsilon})|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2 + (1+\epsilon)(1+\hat{\epsilon}^{-1})|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + (1+\epsilon^{-1})2C_{\mathcal{S}}\Big(\mathfrak{c}_\infty h^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2\\
&\quad + \|\widehat{\psi}_h^{n,3}(0,\cdot) - \widehat{\psi}_h^{n,2}(0,\cdot)\|_{\mathrm{s}'}^2\Big),
\end{aligned}$$

with **(A2)** and $\mathcal{A}_{\mathcal{F}}(\hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}) = \widehat{\psi}_h^{n,3}(0,\cdot) - \widehat{\psi}_h^{n,2}(0,\cdot)$ owing to (6.20b). Set $\epsilon = \frac{5}{72}$ and $\hat{\epsilon} = \frac{7}{11}$, and notice that

$$\frac{1}{6}(1+\epsilon)(1+\hat{\epsilon}) = \frac{7}{24}, \qquad \frac{1}{6}(1+\epsilon)(1+\hat{\epsilon}^{-1}) = \frac{11}{24}, \qquad \frac{1}{3}\tau(1+\epsilon^{-1})C_{\mathcal{S}}\mathfrak{c}_\infty h^{-1} \le \frac{1}{12},$$

owing to the choice (5.16) of ρ_1 in the last term. Hence,

$$\frac{1}{6}\tau|\hat{\zeta}_h^{n,2} - \hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 \le \frac{7}{24}\tau|\hat{\zeta}_h^{n,3}|_{\mathrm{s}}^2 + \frac{11}{24}\tau|\hat{\zeta}_h^{n,1}|_{\mathrm{s}}^2 + \frac{1}{12}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + C\tau\|\widehat{\psi}_h^{n,3} - \widehat{\psi}_h^{n,2}\|_{\mathrm{s}'}^2, \tag{6.22}$$

where we also used $\|\widehat{\psi}_h^{n,3}(0,\cdot) - \widehat{\psi}_h^{n,2}(0,\cdot)\|_{\mathrm{s}'} \le \|\widehat{\psi}_h^{n,3} - \widehat{\psi}_h^{n,2}\|_{\mathrm{s}'}^2$ for the last term on the right hand-side.
(2) Bound $\frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n+1} - \underline{\zeta}_{\mathcal{T}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2$. The identity (6.20b) with $\hat{\underline{z}}_h := (\underline{\zeta}_{\mathcal{T}}^{n+1} - \underline{\zeta}_{\mathcal{T}}^{n,3}, 0)$ followed by (6.12), the property **(A3)** for $\hat{\underline{w}}_h = \hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}$, and $\mathcal{A}_{\mathcal{F}}(\hat{\underline{\zeta}}_h^{n,3} - \hat{\underline{\zeta}}_h^{n,2}) = \widehat{\psi}_h^{n,3}(0,\cdot) - \widehat{\psi}_h^{n,2}(0,\cdot)$ show that

$$\begin{aligned}
&\frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n+1} - \underline{\zeta}_{\mathcal{T}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2\\
&\le \frac{1}{18}(\tau C_{\mathcal{A}}\mathfrak{c}_\infty h^{-1})^2\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + C\tau^2\Big\{\mathfrak{c}_\infty h^{-1}\|\widehat{\psi}_h^{n,3} - \widehat{\psi}_h^{n,2}\|_{\mathrm{s}'}^2 + \|\underline{\mathcal{D}}^{n,3} - \underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + \|\underline{\gamma}_{\mathcal{T}}^n\|_{\frac{1}{\rho},\kappa}^2\Big\}\\
&\le \frac{1}{12}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + C\Big\{\tau\|\widehat{\psi}_h^{n,3} - \widehat{\psi}_h^{n,2}\|_{\mathrm{s}'}^2 + \tau^2\big(\|\underline{\mathcal{D}}^{n,3} - \underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + \|\underline{\gamma}_{\mathcal{T}}^n\|_{\frac{1}{\rho},\kappa}^2\big)\Big\},
\end{aligned} \tag{6.23}$$

owing to the choice (5.16) of ρ_1 in the last step. The definition of $\underline{\gamma}_{\mathcal{T}}^n$, the triangle inequality, the assumption (5.11) on $\underline{g}_{\mathcal{T}}^{n,3} - \underline{f}_{\mathcal{T}}^{n,3}$, and **(I1)** lead to

$$\begin{aligned}
\|\underline{\gamma}_{\mathcal{T}}^n\|_{\frac{1}{\rho},\kappa} &\le \|\underline{g}_{\mathcal{T}}^{n,3} - \underline{f}_{\mathcal{T}}^{n,3}\|_{\frac{1}{\rho},\kappa} + \|\underline{I}_{\mathcal{T}}(\underline{R}^{n,3})\|_{\rho,\frac{1}{\kappa}}\\
&\lesssim C_3^n(\underline{f})\tau^3 + \|\underline{R}^{n,3}\|_{H^\nu(\rho,\frac{1}{\kappa};\mathcal{T}_h)} \lesssim C_3^n(\underline{f},\underline{v})\tau^3.
\end{aligned} \tag{6.24}$$

The bound (6.24) in (6.23) results in

$$\begin{aligned}
&\frac{1}{2}\|\underline{\zeta}_{\mathcal{T}}^{n+1} - \underline{\zeta}_{\mathcal{T}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2\\
&\le \frac{1}{12}\|\underline{\zeta}_{\mathcal{T}}^{n,3} - \underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + C\tau\Big\{\|\widehat{\psi}_h^{n,3} - \widehat{\psi}_h^{n,2}\|_{\mathrm{s}'}^2 + T_f\big(\|\underline{\mathcal{D}}^{n,3} - \underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2 + C_3^n(\underline{f},\underline{v})^2\tau^6\big)\Big\}.
\end{aligned} \tag{6.25}$$

(3) Bound on consistency errors. The Cauchy-Schwarz inequality and Young's inequality imply that

$$\begin{aligned}&\Big|-\frac{1}{6}\tau\widehat{\psi}_h^{n,1}(2\hat{\zeta}_h^{n,1}+\hat{\zeta}_h^{n,2})-\frac{1}{6}\tau\widehat{\psi}_h^{n,2}(\hat{\zeta}_h^{n,1})-\frac{1}{3}\tau\widehat{\psi}_h^{n,3}(\hat{\zeta}_h^{n,3})\Big|\\&\quad\le\frac{1}{12}\tau|\hat{\zeta}_h^{n,2}|_{\mathrm{S}}^2+\tau\frac{1}{48}\big(|\hat{\zeta}_h^{n,1}|_{\mathrm{S}}^2+|\hat{\zeta}_h^{n,3}|_{\mathrm{S}}^2\big)+C\tau\big(\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2+\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2+\|\widehat{\psi}_h^{n,3}\|_{\mathrm{S}'}^2\big).\end{aligned}$$

(4) Bound on remaining terms. The Cauchy-Schwarz inequality and Young's inequality show that

$$\begin{aligned}&\Big|-\frac{1}{6}\tau\langle\underline{\mathcal{D}}^{n,1},2\underline{\zeta}_{\mathcal{T}}^{n,1}+\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}}-\frac{1}{6}\tau\langle\underline{\mathcal{D}}^{n,2},\underline{\zeta}_{\mathcal{T}}^{n,1}\rangle_{\rho,\frac{1}{\kappa}}-\frac{1}{3}\tau\langle\underline{\mathcal{D}}^{n,3},\underline{\zeta}_{\mathcal{T}}^{n,3}\rangle_{\rho,\frac{1}{\kappa}}+\frac{1}{3}\tau\langle\underline{\gamma}_{\mathcal{T}}^{n},\underline{\zeta}_{\mathcal{T}}^{n,3}\rangle_{\underline{L}}\Big|\\&\lesssim\tau\Big\{T_f\big(\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\gamma}_{\mathcal{T}}^{n}\|_{\frac{1}{\rho},\kappa}^2\big)+T_f^{-1}\big(\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\zeta}_{\mathcal{T}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2\big)\Big\}.\end{aligned}\tag{6.26}$$

Recall from (6.14) that $\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}\lesssim\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+T_f^{\frac{1}{2}}\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}+T_f\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}$. The error equation (6.18b) with **(A1)**-**(A3)** results in

$$\begin{aligned}\|\underline{\zeta}_{\mathcal{T}}^{n,3}\|_{\rho,\frac{1}{\kappa}}&\le\frac{1}{2}\big(\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}\big)+\frac{1}{2}\tau\big(\mathfrak{c}_\infty h^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,2}\|_{\rho,\frac{1}{\kappa}}+\mathfrak{c}_\infty^{\frac{1}{2}}h^{-\frac{1}{2}}\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}\big)\\&\lesssim\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+T_f^{\frac{1}{2}}\big(\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}+\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}\big)+T_f\big(\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}\big),\end{aligned}$$

with the usual CFL condition and the observation that $\tau\le T_f$ in the last step. Using the above estimates in (6.26), we obtain

$$\begin{aligned}&\Big|-\frac{1}{6}\tau\langle\underline{\mathcal{D}}^{n,1},2\underline{\zeta}_{\mathcal{T}}^{n,1}+\underline{\zeta}_{\mathcal{T}}^{n,2}\rangle_{\rho,\frac{1}{\kappa}}-\frac{1}{6}\tau\langle\underline{\mathcal{D}}^{n,2},\underline{\zeta}_{\mathcal{T}}^{n,1}\rangle_{\rho,\frac{1}{\kappa}}-\frac{1}{3}\tau\langle\underline{\mathcal{D}}^{n,3},\underline{\zeta}_{\mathcal{T}}^{n,3}\rangle_{\rho,\frac{1}{\kappa}}+\frac{1}{3}\tau\langle\underline{\gamma}_{\mathcal{T}}^{n},\underline{\zeta}_{\mathcal{T}}^{n,3}\rangle_{\underline{L}}\Big|\\&\lesssim\tau\Big(T_f\big(C_3^n(\underline{f},\underline{v})\tau^6+\|\underline{\mathcal{D}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,2}\|_{\rho,\frac{1}{\kappa}}^2+\|\underline{\mathcal{D}}^{n,3}\|_{\rho,\frac{1}{\kappa}}^2\big)+\|\widehat{\psi}_h^{n,1}\|_{\mathrm{S}'}^2+\|\widehat{\psi}_h^{n,2}\|_{\mathrm{S}'}^2+T_f^{-1}\|\underline{\zeta}_{\mathcal{T}}^{n,1}\|_{\rho,\frac{1}{\kappa}}^2\Big).\end{aligned}\tag{6.27}$$

Combining all the estimates from Steps (1)-(4) in (6.19) concludes the proof. □

Proof of Theorem 5.2. This follows immediately from summing (6.21) from $n=0$ to N and applying a discrete Gronwall's lemma (observing that $\tau N=T_f$). □

6.4 Proofs of Lemmas 5.7 and 5.8

Proof of Lemma 5.7. Substituting (5.23a) in (5.23b), we obtain

$$\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1})=\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1})-\tau\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1})+\frac{1}{2}\tau^2\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}\underline{\mathcal{M}}^{-1}\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1})-\frac{1}{2}\tau^2\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}\underline{\mathcal{M}}^{-1}(\underline{f}_{\mathcal{T}}^{n})+\frac{1}{2}\tau\underline{f}_{\mathcal{T}}^{n}+\frac{1}{2}\tau\underline{g}_{\mathcal{T}}^{n,2}.\tag{6.28}$$

Similarly, (5.26a) simplifies (5.26b) to

$$\begin{aligned}\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1})&=\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1})-\tau(b_1+b_2)\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1})+\tau^2b_2a_{21}\big(\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}\underline{\mathcal{M}}^{-1}\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1})-\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}\underline{\mathcal{M}}^{-1}(\underline{\tilde{f}}_{\mathcal{T}}^{n,1})\big)\\&\quad+\tau(b_1\underline{\tilde{f}}_{\mathcal{T}}^{n,1}+b_2\underline{\tilde{f}}_{\mathcal{T}}^{n,2})\\&=\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1})-\tau\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1})+\frac{1}{2}\tau^2\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}\underline{\mathcal{M}}^{-1}\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1})-\frac{1}{2}\tau^2\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}\underline{\mathcal{M}}^{-1}(\underline{\tilde{f}}_{\mathcal{T}}^{n,1})+\tau(b_1\underline{\tilde{f}}_{\mathcal{T}}^{n,1}+b_2\underline{\tilde{f}}_{\mathcal{T}}^{n,2}),\end{aligned}\tag{6.29}$$

where we used the order conditions (5.27). Comparing (6.28) and (6.29), we obtain (5.30). Employing again the order conditions (5.27) (observe that b_2 is non-negative since $c_2\in[0,1]$) leads to

$$\frac{1}{2}(\underline{g}_{\mathcal{T}}^{n,2}-\underline{f}_{\mathcal{T}}^{n,2})=b_1\underline{\tilde{f}}_{\mathcal{T}}^{n,1}+b_2\underline{\tilde{f}}_{\mathcal{T}}^{n,2}-\underline{f}_{\mathcal{T}}^{n}-\frac{1}{2}\partial_t\tau\underline{f}_{\mathcal{T}}^{n}$$

$$
\begin{aligned}
&= b_1 \underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_2 \underline{\widetilde{f}}_{\mathcal{T}}^{n,2} - (b_1+b_2)\underline{f}_{\mathcal{T}}^{n} - b_2 c_2 \tau \partial_t \underline{f}_{\mathcal{T}}^{n} \\
&= b_2 (\underline{\widetilde{f}}_{\mathcal{T}}^{n,2} - \underline{\widetilde{f}}_{\mathcal{T}}^{n,1} - c_2 \tau \partial_t \underline{f}_{\mathcal{T}}^{n}) \\
&= b_2 \int_{t^{n,1}}^{t^{n,2}} (t^{n,2} - t) \partial_{tt} \underline{f}(t)\, dt,
\end{aligned}
$$

with $\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} = \underline{f}_{\mathcal{T}}^{n}$ and Taylor's expansion with integral remainder of order 2 and $t^{n,1} = t^n$ in the last step. Using $t^{n,2} - t^{n,1} = c_2\tau$ and (5.27) concludes the proof of (5.31). □

Proof of Lemma 5.8. Substituting (5.24a) and (5.24b) in (5.24c), we obtain

$$
\begin{aligned}
\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1}) = \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) &- \tau \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) + \frac{1}{2}\tau^2 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) - \frac{1}{6}\tau^3 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) \\
&+ \frac{2}{3}\tau \underline{f}_{\mathcal{T}}^{n} + \frac{1}{6}\tau^2 \partial_t \underline{f}_{\mathcal{T}}^{n} - \frac{1}{2}\tau^2 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{f}_{\mathcal{T}}^{n}) - \frac{1}{6}\tau^3 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\partial_t \underline{f}_{\mathcal{T}}^{n}) \\
&+ \frac{1}{6}\tau^3 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{f}_{\mathcal{T}}^{n}) + \frac{1}{3}\tau \underline{g}_{\mathcal{T}}^{n,3}.
\end{aligned} \tag{6.30}
$$

Similarly, (5.28a) and (5.28b) simplify (5.28c) to

$$
\begin{aligned}
\underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n+1}) &= \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau(b_1+b_2+b_3)\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) + (b_2a_{21} + b_3a_{31} + b_3a_{32})\tau^2 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) \\
&\quad - b_3a_{32}a_{21}\tau^3 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) + \tau\big(b_1 \underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_2 \underline{\widetilde{f}}_{\mathcal{T}}^{n,2} + b_3 \underline{\widetilde{f}}_{\mathcal{T}}^{n,3}\big) \\
&\quad - \tau^2\big((b_2a_{21} + b_3a_{31})\underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\widetilde{f}}_{\mathcal{T}}^{n,1}) + b_3a_{32}\underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\widetilde{f}}_{\mathcal{T}}^{n,2})\big) \\
&\quad + \tau^3 b_3a_{32}a_{21} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\widetilde{f}}_{\mathcal{T}}^{n,1}) \\
&= \underline{\mathcal{M}}(\underline{v}_{\mathcal{T}}^{n,1}) - \tau \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) + \frac{1}{2}\tau^2 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) - \frac{1}{6}\tau^3 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{v}_{\mathcal{T}}^{n,1}) \\
&\quad + \tau\big(b_1 \underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_2 \underline{\widetilde{f}}_{\mathcal{T}}^{n,2} + b_3 \underline{\widetilde{f}}_{\mathcal{T}}^{n,3}\big) - \tau^2\Big((b_2a_{21} + b_3a_{31})\underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\widetilde{f}}_{\mathcal{T}}^{n,1}) + b_3a_{32}\underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\widetilde{f}}_{\mathcal{T}}^{n,2})\Big) \\
&\quad + \frac{1}{6}\tau^3 \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1} \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\widetilde{f}}_{\mathcal{T}}^{n,1}),
\end{aligned} \tag{6.31}
$$

where we used the third-order conditions (5.29) in the last step. Comparing (6.30) and (6.31), we obtain (5.32). Consequently, we have

$$
\frac{1}{3}(\underline{g}_{\mathcal{T}}^{n,3} - \underline{f}_{\mathcal{T}}^{n,3}) = \underline{\delta}_{1,\mathcal{T}}^{n,3} - \tau \underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\delta}_{2,\mathcal{T}}^{n,3}),
$$

with

$$
\underline{\delta}_{1,\mathcal{T}}^{n,3} := b_1 \underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_2 \underline{\widetilde{f}}_{\mathcal{T}}^{n,2} + b_3 \underline{\widetilde{f}}_{\mathcal{T}}^{n,3} - \underline{f}_{\mathcal{T}}^{n} - \frac{1}{2}\tau \partial_t \underline{f}_{\mathcal{T}}^{n} - \frac{1}{6}\tau^2 \partial_{tt} \underline{f}_{\mathcal{T}}^{n} \tag{6.32a}
$$

$$
\underline{\delta}_{2,\mathcal{T}}^{n,3} := (b_2a_{21} + b_3a_{31})\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_3a_{32}\underline{\widetilde{f}}_{\mathcal{T}}^{n,2} - \frac{1}{2}\underline{f}_{\mathcal{T}}^{n} - \frac{1}{6}\tau \partial_t \underline{f}_{\mathcal{T}}^{n}. \tag{6.32b}
$$

The triangle inequality shows that

$$
\|\underline{g}_{\mathcal{T}}^{n,3} - \underline{f}_{\mathcal{T}}^{n,3}\|_{\frac{1}{\rho},\kappa} \le 3(\|\underline{\delta}_{1,\mathcal{T}}^{n,3}\|_{\frac{1}{\rho},\kappa} + \tau \|\underline{\tilde{\mathcal{A}}}_{\mathcal{T}} \underline{\mathcal{M}}^{-1}(\underline{\delta}_{2,\mathcal{T}}^{n,3})\|_{\frac{1}{\rho},\kappa}). \tag{6.33}
$$

(1) Bound on $\underline{\delta}_{1,\mathcal{T}}^{n,3}$. Employing again the order conditions (5.29) and $\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} = \underline{f}_{\mathcal{T}}^{n}$ leads to

$$
\begin{aligned}
\underline{\delta}_{1,\mathcal{T}}^{n,3} &= \sum_{j\in\{2:3\}} b_j (\underline{\widetilde{f}}_{\mathcal{T}}^{n,j} - \underline{f}_{\mathcal{T}}^{n,1} - c_j \tau \partial_t \underline{f}_{\mathcal{T}}^{n} - \frac{1}{2}c_j^2 \tau^2 \partial_{tt} \underline{f}_{\mathcal{T}}^{n}) \\
&= \sum_{j\in\{2:3\}} \frac{1}{2} b_j \int_{t^{n,1}}^{t^{n,j}} (t^{n,j} - t)^2 \partial_{ttt} \underline{f}_{\mathcal{T}}(t)\, dt,
\end{aligned} \tag{6.34}
$$

where we used Taylor's expansion with integral remainder of order 3 and $t^{n,1} = t^n$ in the last step. This proves that

$$\|\underline{\delta}_{1,\mathcal{T}}^{n,3}\|_{\frac{1}{\rho},\kappa} \leq \chi_1 \tau^3 \|\partial_{ttt} \underline{f}_{\mathcal{T}}\|_{C^0(\overline{J}_n;\frac{1}{\rho},\kappa)}, \tag{6.35}$$

where the constant χ_1 can depend on $\{b_j\}_{j\in\{2:3\}}$. However, for nonnegative b_j's, we can use $\sum_{j\in\{2:3\}} b_j c_j^3 \leq \sum_{j\in\{2:3\}} b_j c_j^2 = \frac{1}{3}$ since $\{c_j\}_{j\in\{2:3\}} \in [0,1]$, showing that χ_1 is independent of the b_j's.
(2) Bound for $\underline{\delta}_{2,\mathcal{T}}^{n,3}$. Using (5.29) and $\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} = \underline{f}_{\mathcal{T}}^n$, we arrive at

$$\begin{aligned}
\underline{\delta}_{2,\mathcal{T}}^{n,3} &= (b_2a_{21} + b_3a_{31})\underline{\widetilde{f}}_{\mathcal{T}}^{n,1} + b_3a_{32}\underline{\widetilde{f}}_{\mathcal{T}}^{n,2} - (b_2a_{21} + b_3a_{31} + b_3a_{32})\underline{f}_{\mathcal{T}}^n - b_3c_2a_{32}\tau\partial_t\underline{f}_{\mathcal{T}}^n \\
&= b_3a_{32}(\underline{\widetilde{f}}_{\mathcal{T}}^{n,2} - \underline{f}_{\mathcal{T}}^n - c_2\tau\partial_t\underline{f}_{\mathcal{T}}^n) \\
&= b_3a_{32}\int_{t^{n,1}}^{t^{n,2}} (t^{n,2} - t)\partial_{tt}\underline{f}_{\mathcal{T}}(t)\, dt,
\end{aligned} \tag{6.36}$$

where we used Taylor's expansion with integral remainder of order 2 and $t^{n,1} = t^n$ in the last step. Let

$$\underline{w}_{\mathcal{T}} := \underline{\mathcal{M}}^{-1}(\underline{\delta}_{2,\mathcal{T}}^{n,3}) = (\mathbf{0}, w_{\mathcal{T}}).$$

The definition of $\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}$ leads for $\hat{\underline{w}}_h := (\underline{w}_{\mathcal{T}}, \phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}}))$ and for all $\underline{z}_{\mathcal{T}} \in \underline{V}_{\mathcal{T}}^k$, to

$$\langle \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\hat{\underline{w}}_h), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} = -(\boldsymbol{G}_{\mathcal{T}}(\hat{w}_h), \boldsymbol{\xi}_{\mathcal{T}})_{\boldsymbol{L}^2(\Omega)} + s_h(\hat{w}_h, (z_{\mathcal{T}}, 0)).$$

The definition (3.7) of $\boldsymbol{G}_T$ and a discrete trace inequality show that

$$\|\boldsymbol{G}_{\mathcal{T}}(\hat{w}_h)\|_{\boldsymbol{L}^2(\Omega)}^2 \lesssim \sum_{T\in\mathcal{T}_h} \left\{ \|\nabla w_T\|_{L^2(T)}^2 + \lambda_T h_T^{-1} |\hat{w}_T|_{\mathrm{s},T}^2 \right\}.$$

This, **(A1)**, and $\lambda_T = \mathfrak{c}_T\rho_T = \mathfrak{c}_T^{-1}\kappa_T$ result in

$$\langle \underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\hat{\underline{w}}_h), \underline{z}_{\mathcal{T}}\rangle_{\underline{L}} \lesssim (\mathfrak{c}_\infty \|\nabla_{\mathcal{T}} w_{\mathcal{T}}\|_{L^2(\frac{1}{\kappa};\Omega)} + \mathfrak{c}_\infty^{-\frac{1}{2}} h^{-\frac{1}{2}} |\hat{w}_h|_{\mathrm{s}}) \|\underline{z}_{\mathcal{T}}\|_{\rho,\frac{1}{\kappa}}.$$

Consequently, we have

$$\|\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{w}_{\mathcal{T}})\|_{\frac{1}{\rho},\kappa} \lesssim \mathfrak{c}_\infty \|\nabla_{\mathcal{T}} w_{\mathcal{T}}\|_{\boldsymbol{L}^2(\frac{1}{\kappa};\Omega)} + \mathfrak{c}_\infty^{\frac{1}{2}} h^{-\frac{1}{2}} |\hat{w}_h|_{\mathrm{s}}. \tag{6.37}$$

Furthermore, the identity $a_h(\hat{\underline{w}}_h, (\mathbf{0}, 0, z_{\mathcal{F}})) = 0$ for all $z_{\mathcal{F}} \in V_{\mathcal{F}0}^k$ simplifies to

$$s_h((w_{\mathcal{T}}, \phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})), (0, z_{\mathcal{F}})) = 0 \quad \forall z_{\mathcal{F}} \in V_{\mathcal{F}0}^k.$$

This can be rewritten as $\sum_{F\in\mathcal{F}} \int_F z_{\mathcal{F}} \big(\sum_{T\in\mathcal{T}_F} \lambda_T^{-1}(w_T - \phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})|_F) \big)\, ds = 0$. Hence, we have

$$\phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})|_F = \frac{\sum_{T\in\mathcal{T}_F} \lambda_T^{-1} w_T}{\sum_{T\in\mathcal{T}_F} \lambda_T^{-1}},$$

where $\mathcal{T}_F := \{T \in \mathcal{T} : F \in \mathcal{F}_{\partial T}\}$ denotes the patch of F. Notice also that $\phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})|_F = w_T|_F$ for all $F \in \mathcal{F}_h^\partial$. An elementary algebraic manipulation shows that

$$\begin{aligned}
|\hat{w}_h|_{\mathrm{s}}^2 &= \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_{\partial T}\cap\mathcal{F}_h^i} \lambda_T^{-1} \|w_T|_F - \phi_{\mathcal{F}}(\underline{w}_{\mathcal{T}})|_F\|_{L^2(F)}^2 \\
&= \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_{\partial T}\cap\mathcal{F}_h^i} \frac{\lambda_T^{-1}}{(\sum_{T\in\mathcal{T}_F}\lambda_T^{-1})^2} \|(\sum_{T\in\mathcal{T}_F}\lambda_T^{-1}) w_T|_F - \sum_{T\in\mathcal{T}_F} \lambda_T^{-1} w_T\|_{L^2(F)}^2 \\
&\leq \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_{\partial T}\cap\mathcal{F}_h^i} \lambda_T^{-1} \|[w_{\mathcal{T}}]_F\|_{L^2(F)}^2 = \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_{\partial T}\cap\mathcal{F}_h^i} \lambda_T^{-1} \|[w_{\mathcal{T}} - w]_F\|_{L^2(F)}^2,
\end{aligned}$$

with $\kappa^{-1}w := (b_2a_{21}+b_3a_{31})\underline{\widetilde{f}}^{n,1} + b_3a_{32}\underline{\widetilde{f}}^{n,2} - \frac{1}{2}f^n - \frac{1}{6}\tau\partial_t f^n \in H^1(\Omega)$ since $f \in C^2(\overline{J};H^1(\Omega))$, and the notation $[\cdot]_F$ to denote the jump across the interface F. The approximation properties of $\underline{I}_{\mathcal{T}}^k$ give

$$|\hat{w}_h|_{\mathrm{S}}^2 \lesssim \sum_{T\in\mathcal{T}_h} \lambda_T^{-1} h_T \|\nabla w\|_{L^2(T)}^2.$$

This with $\lambda_T = \mathfrak{c}_T^{-1}\kappa_T$ and the H^1-stability of $\Pi_{\mathcal{T}}^k$ in (6.37) prove that

$$\|\underline{\tilde{\mathcal{A}}}_{\mathcal{T}}(\underline{w}_{\mathcal{T}})\|_{\frac{1}{\rho},\kappa} \lesssim \mathfrak{c}_\infty\tau^2\|\nabla\underline{\delta}_{2,\mathcal{T}}^{n,3}\|_{\frac{1}{\rho},\kappa} \le \frac{1}{6}\mathfrak{c}_\infty\tau^2\|\nabla\partial_{tt}\underline{f}_{\mathcal{T}}\|_{\frac{1}{\rho},\kappa}, \tag{6.38}$$

where we used (6.36) with $t^{n,2}-t^{n,1}\le c_2\tau$ and (5.29). The combination of (6.35) and (6.38) in (6.33) followed by the L^2-stability of $\underline{\Pi}_{\mathcal{T}}^k$ concludes the proof of (5.33). □

7 Variants of HHO

This section discusses variations in the HHO method depending on the equal- and mixed-order setting, and on the stabilization parameters. We provide examples of suitable interpolation operators for all these variations.

7.1 Main results

We consider the following variations in the HHO method: (i) Equal- or mixed-order settings, where the cell unknowns are of order $k'=k$ or $k'=k+1$, respectively. (ii) Two choices of the stabilization weight in (3.9), namely $\lambda_T^{-1}\tilde{h}_T^{-2\alpha}$ for $\alpha=\{0,\frac{1}{2}\}$ with $\tilde{h}_T = \frac{h_T}{\ell_\Omega}$. The analysis carried out so far is based on $\mathcal{O}(1)$-stabilization (i.e., $\alpha=0$), which is a natural choice for Friedrichs systems. However, $\mathcal{O}(\frac{1}{h})$-stabilization (i.e., $\alpha=\frac{1}{2}$) is also of interest since this choice can lead in some situations to $\mathcal{O}(h^{k+1})$ convergence rate in space (see Table 7.1 below). Since the LS $\mathcal{O}(\frac{1}{h})$-stabilization can only give $\mathcal{O}(h^k)$ convergence rate in space (for both equal- and mixed-order settings), we now consider instead the higher-order stabilization operators used in HHO methods for elliptic PDEs. Recalling the definition (3.8) of the difference operator $\delta_{\partial T}$ for all $T\in\mathcal{T}_h$, we define, for all $\hat{w}_T\in\hat{V}_T^k$,

$$S_{\partial T}^{\mathrm{HHO}}(\hat{w}_T) := \begin{cases} \Pi_{\partial T}^k\big(\delta_{\partial T}(\hat{w}_T) + (1-\Pi_T^k)R_T^{k+1}(0,\delta_{\partial T}(\hat{w}_T))|_{\partial T}\big) & \text{for } k'=k, \quad (7.1a)\\ \Pi_{\partial T}^k\big(\delta_{\partial T}(\hat{w}_T)\big) & \text{for } k'=k+1, \quad (7.1b)\end{cases}$$

with the high-order potential reconstruction operator $R_T^{k+1}:\hat{V}_T^k\to\mathbb{P}^{k+1}(T;\mathbb{R})$ such that

$$(\nabla R_T^{k+1}(\hat{w}_T),\nabla q)_{\boldsymbol{L}^2(T)} = (\nabla w_T,\nabla q)_{\boldsymbol{L}^2(T)} + (w_{\partial T}-w_T,\nabla q\cdot\boldsymbol{n}_T)_{L^2(\partial T)} \qquad \forall q\in\mathbb{P}_*^{k+1}(T;\mathbb{R}), \tag{7.2}$$

with $\mathbb{P}_*^{k+1}(T;\mathbb{R}) := \{q\in\mathbb{P}^{k+1}(T;\mathbb{R}) : (q,1)_{L^2(T)}=0\}$, and $(R_T^{k+1}(\hat{w}_T)-w_T,1)_{L^2(T)}=0$. Thus, we now consider the stabilization bilinear form $s_h:\hat{V}_h^k\times\hat{V}_h^k\to\mathbb{R}$ such that

$$s_h(\hat{w}_h,\hat{z}_h) := \sum_{T\in\mathcal{T}_h} s_T(\hat{w}_T,\hat{z}_T), \qquad s_T(\hat{w}_T,\hat{z}_T) := \tau_{\partial T}(S_{\partial T}^{\mathrm{HHO}}(\hat{w}_T),S_{\partial T}^{\mathrm{HHO}}(\hat{z}_T))_{L^2(\partial T)}, \tag{7.3}$$

with the stabilization parameter

$$\tau_{\partial T} := \lambda_T^{-1}\tilde{h}_T^{-2\alpha} \qquad \text{with } \lambda_T = \mathfrak{c}_T\rho_T = \mathfrak{c}_T^{-1}\kappa_T,\ \tilde{h}_T := \ell_\Omega^{-1}h_T,\ \alpha\in\left\{0,\frac{1}{2}\right\}. \tag{7.4}$$

This leads altogether to four variations: equal- vs. mixed-order setting and $\mathcal{O}(1)$- vs. $\mathcal{O}(\frac{1}{h})$-stabilization with $S_{\partial T}^{\mathrm{HHO}}$ defined in (7.1).

The first important point is that the stabilization (7.1) generally calls for implicit schemes, for both equal- and mixed-order settings. Indeed, in the equal-order case, the invertibility of the face-face stabilization matrix is not guaranteed (this property is required to handle the static coupling between cell and face unknowns in explicit schemes). Moreover, it is not clear how to prove **(A2)**. In the mixed-order case, the invertibility of the face-face stabilization matrix is guaranteed, but again **(A2)** fails. It is still possible to prove that

$$|\hat{w}_h|_{\mathrm{s}} \lesssim \mathfrak{c}_\infty^{\frac{1}{2}} \tilde{h}^{\alpha} h^{-\frac{1}{2}} \|\boldsymbol{\tau}_{\mathcal{T}}\|_{\boldsymbol{L}^2(\rho;\Omega)} + \mathfrak{c}_\infty^{\frac{1}{2}} \tilde{h}^{-\alpha} h^{-\frac{1}{2}} \|w_{\mathcal{T}}\|_{L^2(\frac{1}{\kappa};\Omega)} + \|\mathcal{A}_{\mathcal{F}}(\underline{\hat{w}}_h)\|_{\mathrm{s}'} \qquad \forall \underline{\hat{w}}_h \in \underline{\hat{V}}_{h0}^k. \tag{7.5}$$

Thus, for $\alpha = \frac{1}{2}$, we loose a factor $h^{-\frac{1}{2}}$ and this eventually leads to a 2-CFL condition for explicit schemes. For $\alpha = 0$ instead, we can repeat the analysis in Sections 5 and 6, but this leads to the same convergence rates as in the equal-order case with LS stabilization. In any case, if one is ready to run an implicit scheme with either equal- and mixed-order or to run an explicit scheme with mixed-order using a 2-CFL condition, one obtains $\mathcal{O}(\tilde{h}^{\alpha} h^{k+\frac{1}{2}})$ convergence rate on polyhedra, which is optimal for $\alpha = \frac{1}{2}$.

To summarize, Table 7.1 displays the convergence rates and CFL conditions for all the combinations of HHO discretizations discussed in the paper. The two rows labelled A correspond to the analysis of Sections 5-6, and the rows labelled B and C to the variations discussed in Section 7.

	Stabilization	HHO	Mesh	Convergence	Interpolation operator		
					Implicit	Explicit	CFL
A	$\mathcal{O}(1), \delta_{\partial T}$	$k' = k$	Simplices	$\mathcal{O}(h^{k+1})$	HDG	HDG	$\frac{4}{3}/1$
			Polyhedra	$\mathcal{O}(h^{k+\frac{1}{2}})$	HHO	HHO	$\frac{4}{3}/1$
B	$\mathcal{O}(1), S_{\partial T}^{\mathrm{HHO}}$	$k' = k$	Simplices	$\mathcal{O}(h^{k+1})$	HDG-HHO	-	-
			Polyhedra	$\mathcal{O}(h^{k+\frac{1}{2}})$	HHO		
		$k' = k+1$	Simplices	$\mathcal{O}(h^{k+1})$	HDG$^+$	HDG$^+$	$\frac{4}{3}/1$
			Polyhedra	$\mathcal{O}(h^{k+\frac{1}{2}})$	HHO or H$^+$	H$^+$	$\frac{4}{3}/1$
C	$\mathcal{O}(\frac{1}{h}), S_{\partial T}^{\mathrm{HHO}}$	$k' = k$	Polyhedra	$\mathcal{O}(h^{k+1})$	HHO	-	-
		$k' = k+1$	Polyhedra	$\mathcal{O}(h^{k+1})$	HHO or H$^+$	H$^+$	2/2

Table 7.1: Choices of stabilization parameters and order setting: convergence rates and suitable interpolation operators defined in Section 7.2

7.2 Interpolation operators

We discuss possible choices of the interpolation operator satisfying **(I1)** and **(I2)** to prove the convergence rates stated in Table 7.1. We focus on **(I2)** which is the more delicate property.

For equal-order on simplices, one can replace the stabilization weight λ_T^{-1} by $\tau_{\partial T}$ and the difference operator $\delta_{\partial T}$ by $S_{\partial T}^{\mathrm{HHO}}$ in the third condition (4.19c) defining the HDG interpolation operator $\underline{\Pi}_T^{\mathrm{H}}$ (we call this slightly modified HDG interpolation operator with HHO stabilization the HDG-HHO interpolation operator) and prove that the consistency error is still zero. Namely, (4.19c) is replaced by

$$((\boldsymbol{\tau} - \boldsymbol{\Pi}_T^{\boldsymbol{\Sigma}}(\underline{w}))\cdot\boldsymbol{n}_T, \mu)_{L^2(\partial T)} = \tau_{\partial T}(S_{\partial T}^{\mathrm{HHO}}(\Pi_T^V(\underline{w}), \Pi_{\partial T}^k(w|_{\partial T})), S_{\partial T}^{\mathrm{HHO}}(0,\mu))_{L^2(\partial T)} \qquad \forall \mu \in V_{\partial T}^k. \tag{7.6}$$

However, on general polyhedra, adapting the arguments from [4], an error analysis can be carried out using the classical HHO interpolation operator for $\alpha \in \{0, \frac{1}{2}\}$.

For mixed-order on polyhedra, the first term on the right hand-side of (4.14) can no longer be bounded using only the stabilization seminorm from (7.1b). To circumvent this problem, we propose to use the H$^+$ interpolation operator defined in [17] (where it is called HDG$^+$ interpolation operator). Specifically, the H$^+$ interpolation operator $\boldsymbol{\Pi}_T^{\mathrm{H}+} : \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T) \to \mathbb{P}^k(T;\mathbb{R}^d)$, $\nu_{\boldsymbol{\sigma}} \in (\frac{1}{2}, 1]$, is defined for

all $T \in \mathcal{T}_h$ as follows. We consider the L^2-orthogonal decomposition $\mathbb{P}^k(T;\mathbb{R}^d) = \nabla\mathbb{P}^{k+1}_*(T;\mathbb{R}) \oplus \mathbb{Z}^k(T;\mathbb{R}^d)$, where $\mathbb{P}^{k+1}_*(T;\mathbb{R}) := \{q \in \mathbb{P}^{k+1}(T;\mathbb{R}) : (q,1)_{L^2(T)} = 0\}$ and $\mathbb{Z}^k(T;\mathbb{R}^d) := \nabla\mathbb{P}^{k+1}_*(T;\mathbb{R})^\perp \cap \mathbb{P}^k(T;\mathbb{R}^d)$. Then, for all $\boldsymbol{\tau} \in \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T)$, we define $\boldsymbol{\Pi}^{\mathrm{H}^+}_T(\boldsymbol{\tau}) \in \boldsymbol{\Sigma}^k_T$ from the following conditions:

$$(\boldsymbol{\Pi}^{\mathrm{H}^+}_T(\boldsymbol{\tau}) - \boldsymbol{\tau}, \boldsymbol{\xi})_{\boldsymbol{L}^2(T)} = 0 \qquad \forall \boldsymbol{\xi} \in \mathbb{Z}^k(T;\mathbb{R}^d), \tag{7.7a}$$

$$(\boldsymbol{\Pi}^{\mathrm{H}^+}_T(\boldsymbol{\tau}) - \boldsymbol{\tau}, \nabla z)_{\boldsymbol{L}^2(T)} = (\Pi^k_{\partial T}(\boldsymbol{\tau}\cdot\boldsymbol{n}_T) - \boldsymbol{\tau}\cdot\boldsymbol{n}_T, z)_{L^2(\partial T)} \qquad \forall z \in \mathbb{P}^{k+1}_*(T;\mathbb{R}). \tag{7.7b}$$

The global interpolation operator $\boldsymbol{\Pi}^{\mathrm{H}^+}_{\mathcal{T}} : \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(\Omega) \to \boldsymbol{\Sigma}^k_{\mathcal{T}}$ is defined as $(\boldsymbol{\Pi}^{\mathrm{H}^+}_{\mathcal{T}}(\boldsymbol{\tau}))|_T := \boldsymbol{\Pi}^{\mathrm{H}^+}_T(\boldsymbol{\tau}|_T)$ for all $\boldsymbol{\tau} \in \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(\Omega)$. We then define the H^+ interpolation operator $\hat{\underline{I}}^{\mathrm{H}^+}_h : \underline{Y}^{\mathrm{H}^+} \to \hat{\underline{V}}^k_{h0}$ with $\underline{Y}^{\mathrm{H}^+} := \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(\Omega) \times H^{\nu_v}(\Omega)$, with $\nu_{\boldsymbol{\sigma}} \in (\frac{1}{2}, 1]$, $\nu_v \in (\frac{1}{2}, 1]$, such that, for all $\underline{w} := (\boldsymbol{\tau}, w) \in \underline{Y}^{\mathrm{H}^+}$,

$$\hat{\underline{I}}^{\mathrm{H}^+}_h(\underline{w}) := (\overbrace{\boldsymbol{\Pi}^{\mathrm{H}^+}_{\mathcal{T}}(\boldsymbol{\tau}), \underbrace{\Pi^{k'}_{\mathcal{T}}(w), \Pi^k_{\mathcal{F}}(w|_{\mathcal{F}})}_{=:\hat{I}^{\mathrm{HHO}}_h(w)}}^{\underline{I}^{\mathrm{H}^+}_{\mathcal{T}}(\underline{w}) :=}) \in \hat{\underline{V}}^k_h. \tag{7.8}$$

Thus, $\hat{\underline{I}}^{\mathrm{H}^+}_h$ approximates the dual variable $\boldsymbol{\tau}$ using the H^+ interpolation operator and the primal variable w using the usual L^2-based HHO interpolation operator. One can verify that the H^+ interpolation operator leads to $\mathcal{O}(\tilde{h}^\alpha h^{k+\frac{1}{2}})$ convergence rates on polyhedra. Finally, using similar arguments to those in Section 4.4 in the equal-order case, this rate can be improved to $\mathcal{O}(h^{k+1})$ on simplices for $\alpha = 0$ using the HDG^+ interpolation operator from [16, Section 4.1]. Specifically, for all $\underline{w} := (\boldsymbol{\tau}, w) \in \boldsymbol{H}^{\nu_{\boldsymbol{\sigma}}}(T) \times H^{\nu_v}(T)$, with $\nu_{\boldsymbol{\sigma}} \in (\frac{1}{2}, 1]$, $\nu_v \in (\frac{1}{2}, 1]$, one defines $\underline{\Pi}^{\mathrm{H}}_T(\underline{w}) := (\boldsymbol{\Pi}^{\boldsymbol{\Sigma}}_T(\underline{w}), \Pi^V_T(\underline{w})) \in \mathbb{P}^k(T;\mathbb{R}^d) \times \mathbb{P}^{k'}(T;\mathbb{R})$ on any simplex $T \in \mathcal{T}_h$ by using (4.19a)-(4.19b), replacing (4.19c) by

$$((\boldsymbol{\tau} - \boldsymbol{\Pi}^{\boldsymbol{\Sigma}}_T(\underline{w}))\cdot\boldsymbol{n}_T, \mu)_{L^2(\partial T)} = \tau_{\partial T}(\delta_{\partial T}(\Pi^V_T(\underline{w}), \Pi^k_{\partial T}(w|_{\partial T})), \delta_{\partial T}(0,\mu))_{L^2(\partial T)} \quad \forall \mu \in V^k_{\partial T}, \tag{7.9}$$

and adding that

$$(\nabla\cdot(\Pi^V_T(\underline{w}) - w), \chi)_{L^2(T)} = \tau_{\partial T}(S^{\mathrm{HHO}}_{\partial T}(\Pi^V_T(\underline{w}), \Pi^k_{\partial T}(w|_{\partial T})), S^{\mathrm{HHO}}_{\partial T}(\chi, 0))_{L^2(\partial T)} \qquad \forall \chi \in \widetilde{\mathbb{P}}^{k'}(T;\mathbb{R}), \tag{4.19d}$$

with $\widetilde{\mathbb{P}}^{k'}(T;\mathbb{R})$ is composed of homogeneous polynomials of degree $k' = k+1$. Notice that the two interpolation operators H^+ and HDG^+ are considered in the HDG^+ literature with $\mathcal{O}(\frac{1}{h})$-stabilization, but they are also well-defined for $\mathcal{O}(1)$-stabilization.

References

[1] M. Abbas, A. Ern, and N. Pignet, *A hybrid high-order method for incremental associative plasticity with small deformations*, Comput. Methods Appl. Mech. Eng., 346 (2019), pp. 891–912.

[2] P. Bansal, A. Moiola, I. Perugia, and C. Schwab, *Space–time discontinuous Galerkin approximation of acoustic waves with point singularities*, IMA J. Numer. Anal., 41 (2021), pp. 2056–2109.

[3] E. Burman, O. Duran, and A. Ern, *Hybrid high-order methods for the acoustic wave equation in the time domain*, Comm. App. Math. Comp. Sci., 4 (2022), pp. 597–633.

[4] E. Burman, O. Duran, A. Ern, and M. Steins, *Convergence analysis of hybrid high-order methods for the wave equation*, J. Sci. Comput., 87 (2021), p. 91.

[5] E. Burman, A. Ern, and M. A. Fernández, *Explicit Runge–Kutta schemes and finite elements with symmetric stabilization for first-order linear PDE systems*, SIAM J. Numer. Anal., 48 (2010), pp. 2019–2042.

[6] M. Cicuttin, A. Ern, and N. Pignet, *Hybrid high-order methods: a primer with applications to solid mechanics*, Springer, 2021.

[7] B. Cockburn, D. Di Pietro, and A. Ern, *Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods*, ESAIM: Math. Model. Numer. Anal., 50 (2016), pp. 635–650.

[8] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Math. Comp., 79 (2010), pp. 1351–1367.

[9] B. Cockburn and V. Quenneville-Bélair, *Uniform-in-time superconvergence of the HDG methods for the acoustic wave equation*, Math. Comp., 83 (2014), pp. 65–85.

[10] B. Cockburn and C.-W. Shu, *Tvb Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework*, Math. Comp., 52 (1989), pp. 411–435.

[11] D. Corallo, W. Dörfler, and C. Wieners, *Space-time discontinuous Galerkin methods for weak solutions of hyperbolic linear symmetric Friedrichs systems*, J. Sci. Comput., 94 (2023), p. 27.

[12] F. Dassi, A. Fumagalli, I. Mazzieri, and G. Vacca, *Mixed virtual element approximation of linear acoustic wave equation*, IMA J. Numer. Anal., (2023), p. drad078.

[13] D. Di Pietro and A. Ern, *A hybrid high-order locking-free method for linear elasticity on general meshes*, Comput. Meth. Appl. Mech. Engrg., 283 (2015), pp. 1–21.

[14] D. Di Pietro, A. Ern, and S. Lemaire, *An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators*, J. Comput. Methods. Appl. Math., 14 (2014), pp. 461–472.

[15] D. A. Di Pietro and J. Droniou, *The hybrid high-order method for polytopal meshes*, Number 19 in Modeling, Simulation and Application, (2020).

[16] S. Du and F.-J. Sayas, *An Invitation to the Theory of the Hybridizable Discontinuous Galerkin Method: Projections, Estimates, Tools*, Springer Nature, 2019.

[17] S. Du and F.-J. Sayas, *A note on devising HDG+ projections on polyhedral elements*, Math. Comp., 90 (2021), pp. 65–79.

[18] A. Ern and J.-L. Guermond, *Finite elements III: first-order and time-dependent PDEs*, vol. 74, Springer Nature, 2021.

[19] A. Ern and M. Steins, *Convergence analysis for the wave equation discretized with hybrid methods in space (HHO, HDG and WG) and the leapfrog scheme in time*, J. Sci. Comput., 101 (2024), p. 7.

[20] R. S. Falk and G. R. Richter, *Explicit finite element methods for symmetric hyperbolic equations*, SIAM J. Numer. Anal., 36 (1999), pp. 935–952.

[21] R. Griesmaier and P. Monk, *Discretization of the wave equation using continuous elements in time and a hybridizable discontinuous Galerkin method in space*, J. Sci. Comput., 58 (2014), pp. 472–498.

[22] J.-L. Guermond, *Subgrid stabilization of Galerkin approximations of linear monotone operators*, IMA J. Numer. Anal., 21 (2001), pp. 165–197.

[23] D. Levy and E. Tadmor, *From semidiscrete to fully discrete: Stability of Runge–Kutta schemes by the energy method*, SIAM review, 40 (1998), pp. 40–73.

[24] N. C. Nguyen, J. Peraire, and B. Cockburn, *High-order implicit hybridizable discontinuous Galerkin methods for acoustics and elastodynamics*, J. Comput. Phys., 230 (2011), pp. 3695–3718.

[25] M. Stanglmeier, N. C. Nguyen, J. Peraire, and B. Cockburn, *An explicit hybridizable discontinuous Galerkin method for the acoustic wave equation*, Comput. Meth. Appl. Mech. Engrg., 300 (2016), pp. 748–769.

[26] Q. Zhang and C.-W. Shu, *Error estimates to smooth solutions of Runge–Kutta discontinuous Galerkin methods for scalar conservation laws*, SIAM J. Numer. Anal., 42 (2004), pp. 641–666.

[27] T. Zhang and S. Zhang, *An explicit weak Galerkin method for solving the first order hyperbolic systems*, J. Comput. Appl. Math., 412 (2022), p. 114311.

[28] X. Zhang and M. Feng, *A weak Galerkin mixed finite element method for acoustic wave equation*, Adv. Appl. Math. Mech., 14 (2022), pp. 936–959.