ESAIM: M2AN 59 (2025) 613–641
https://doi.org/10.1051/m2an/2024085

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

ASYMPTOTIC-PRESERVING HYBRIDIZABLE DISCONTINUOUS GALERKIN METHOD FOR THE WESTERVELT QUASILINEAR WAVE EQUATION

Sergio Gómez[1,*] and Mostafa Meliani[2,3]

Abstract. We discuss the asymptotic-preserving properties of a hybridizable discontinuous Galerkin method for the Westervelt model of ultrasound waves. More precisely, we show that the proposed method is robust with respect to small values of the sound diffusivity damping parameter δ by deriving low- and high-order energy stability estimates, and *a priori* error bounds that are independent of δ. Such bounds are then used to show that, when $\delta \to 0^+$, the method remains stable and the discrete acoustic velocity potential $\psi_h^{(\delta)}$ converges to $\psi_h^{(0)}$, where the latter is the singular vanishing dissipation limit. Moreover, we prove optimal convergence rates for the approximation of the acoustic particle velocity variable $\underline{\boldsymbol{v}} = \nabla\psi$. The established theoretical results are illustrated with some numerical experiments.

Mathematics Subject Classification. 65M60, 65M15, 35L70.

Received May 6, 2024. Accepted December 29, 2024.

1. Introduction

Let $Q_T = \Omega \times (0, T)$ be a space–time cylinder, where $\Omega \subset \mathbb{R}^d$ ($d \in \{2, 3\}$) is an open, bounded polytopic domain with Lipschitz boundary $\partial\Omega$, and $T > 0$ is the final time. We consider the following Westervelt equation of nonlinear acoustics [36]:

$$\begin{cases} (1 + 2k\partial_t\psi)\partial_{tt}\psi - c^2\Delta\psi - \delta\Delta(\partial_t\psi) = 0 & \text{in } Q_T, \\ \psi = 0 & \text{on } \partial\Omega \times (0, T), \\ \psi = \psi_0, \qquad \partial_t\psi = \psi_1 & \text{on } \Omega \times \{0\}, \end{cases} \tag{1.1}$$

where the unknown $\psi : Q_T \to \mathbb{R}$ is the acoustic velocity potential. In the IBVP (1.1), the constant $k \in \mathbb{R}$ is a medium-dependent nonlinearity parameter, $c > 0$ is the speed of sound, ψ_0 and ψ_1 are given initial data, and $\delta \geq 0$ is the sound diffusivity coefficient.

Keywords and phrases. Asymptotic-preserving method, nonlinear acoustics, Westervelt equation, hybridizable discontinuous Galerkin method.

[1] Department of Mathematics and Applications, University of Milano-Bicocca, 20125 Milan, Italy.

[2] Department of Mathematics, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

[3] Institute of Mathematics, Czech Academy of Sciences, Prague, Czech Republic.

[*]Corresponding author: sergio.gomezmacias@unimib.it

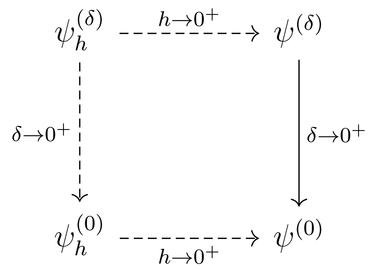

FIGURE 1. Asymptotic-preserving commutative diagram for the Westervelt equation. This diagram represents the connections between $\psi_h^{(\delta)}$ and $\psi^{(\delta)}$ as $h \to 0^+$ (even in the limit case $\delta = 0$) as well as between $\psi_h^{(\delta)}$ and $\psi_h^{(0)}$ as $\delta \to 0^+$. The superscript (δ) is used to emphasize the dependence on the parameter δ of the continuous solution and its numerical approximation.

Introducing the acoustic particle velocity variable $\underline{\boldsymbol{v}} : Q_T \to \mathbb{R}^d$, defined by $\underline{\boldsymbol{v}} := \nabla\psi$, the Westervelt equation in (1.1) can be rewritten in mixed form as

$$\begin{cases} (1+2k\partial_t\psi)\partial_{tt}\psi - c^2\nabla\cdot\underline{\boldsymbol{v}} - \delta\nabla\cdot(\partial_t\underline{\boldsymbol{v}}) = 0 & \text{in } Q_T, \\ \underline{\boldsymbol{v}} = \nabla\psi & \text{in } Q_T, \\ \psi = 0 & \text{on } \partial\Omega\times(0,T), \\ \psi = \psi_0, \qquad \partial_t\psi = \psi_1 & \text{on } \Omega\times\{0\}. \end{cases} \tag{1.2}$$

Since we study the limit as $\delta \to 0^+$, we make the purely technical assumption that $\delta \in [0, \overline{\delta})$ for some fixed $\overline{\delta} > 0$. Such an assumption is helpful in the limiting behavior analysis in Section 5, as it allows us to make the estimates depend on $\overline{\delta}$ but never on δ itself.

The Westervelt equation models the propagation of sound in a fluid medium, and it is a well-accepted model in nonlinear acoustics (see *e.g.*, [18], Sect. 5.3). Nonlinear sound propagation finds a multitude of technical and medical applications, such as ultrasound imaging, lithotripsy, welding, and sonochemistry; see [12, 22].

When the parameter δ is strictly positive, equation (1.1) is strongly damped, and its solution enjoys global existence properties for initial conditions satisfying some smallness and regularity assumptions as shown in [19, 30]. Conversely, when $\delta = 0$, the main mechanism preventing the formation of singularities is lost and no global existence results are known. The stark contrast between these two regimes gives rise to interesting issues, such as the continuous dependence of the solution on the damping parameter $\delta \to 0^+$, and the interplay of this limit and numerical discretizations. A numerical method for the Westervelt equation is said to be *asymptotic preserving* if it allows for interchanging the vanishing limits of the mesh size parameter h and the sound diffusivity parameter δ, *i.e.*, if it satisfies the commutative diagram in Figure 1. The main focus of this work is to show that the proposed method is asymptotic preserving.

In the literature, *a priori* error results for the approximation of the solution to the Westervelt equation initially relied on the assumption of strictly positive values of the damping parameter δ (see, *e.g.*, [2, 33]). Nevertheless, as the damping parameter is relatively small in practice and it can become negligible in certain applications, there have been recent efforts to devise numerical methods that are robust with respect to small values of the sound diffusivity parameter δ. In particular, estimates for the standard and mixed finite element discretizations of the Westervelt equation with $\delta = 0$ follow as particular cases of those in [16, 26, 29], whereas the asymptotic behaviour of such methods for $\delta \to 0^+$ has been recently studied in [14, 32]. The main challenge resides in the limited regularity offered by most standard finite element spaces, which hinders the extension of the arguments used to study the vanishing viscosity limit in the continuous setting (see, *e.g.*, [20]).

This work concerns the asymptotic analysis of a hybridizable discontinuous Galerkin (HDG) method for the Westervelt equation when $\delta \to 0^+$. HDG methods, originally introduced in [6] for an elliptic PDE, are a class of discontinuous Galerkin methods characterized by the possibility of performing a local *static condensation* procedure to reduce the number of unknowns of the linear system stemming from the discretization of a d-dimensional linear PDE. Such a procedure leads to linear systems involving only unknowns associated with degrees of freedom on $(d-1)$-dimensional mesh-facets. Although this hybridization property does not naturally extend to nonlinear PDEs, it can be used in combination with suitable nonlinear solvers (see, *e.g.*, Sect. 6.1 below). Moreover, provided that the exact solution is smooth enough, the Local Discontinuous Galerkin-hybridizible (LDG-H) method in [5,6] for the Poisson equation converges with optimal order $\mathcal{O}(h^{p+1})$ for the $L^2(\Omega)$-error of the flux variable when approximations of degree p are used, and allows for a local postprocessing that produces an approximation of degree $p+1$ of the primal variable that superconverges with order $\mathcal{O}(h^{p+2})$ in the $L^2(\Omega)$-norm.

To the best of our knowledge, there are four different versions of the HDG method for the linear acoustic wave equation ($c^{-2}\partial_{tt}u - \Delta u = f$):

(a) the dissipative HDG method introduced in [31] and analyzed in [4], which is based on the first-order system ($\partial_t \underline{\boldsymbol{q}} = \nabla v;\ c^{-2}\partial_t v - \nabla\cdot\underline{\boldsymbol{q}} = f$) with $v := \partial_t u$ and $\underline{\boldsymbol{q}} := \nabla u$;
(b) the conservative HDG method in [15] based on the same first-order system, whose energy conserving property is enforced by choosing the *numerical fluxes* of $\underline{\boldsymbol{q}}_h$ in dependence of $\partial_t v_h$, which in turn causes a theoretical loss of convergence of half an order;
(c) the HDG method in [34] for the Hamiltonian formulation ($\partial_t u = v;\ c^{-2}\partial_t v = f + \nabla\cdot\underline{\boldsymbol{q}}$); and
(d) the conservative HDG method in [11], which is based on the mixed formulation ($\underline{\boldsymbol{q}} = -\nabla u;\ c^{-2}\partial_{tt}u + \nabla\cdot\underline{q} = f$).

The theoretical results in (a), (c), and (d) predict optimal convergence for the approximation of all the variables involved, and superconvergence for some (locally computable) postprocessed approximations of the scalar variables.

In this work, we design an HDG method for the Westervelt model, which is based on the conservative HDG method in [11] for the linear second-order wave equation. This choice allows us to directly approximate the variables of interest $(\psi, \underline{\boldsymbol{v}})$, eliminate efficiently the discrete vector variable $\underline{\boldsymbol{v}}_h$ from the nonlinear ODE system, and obtain optimal convergence in the low- and high-order energy norms. Moreover, it facilitates the extension of the techniques used in [29] for the analysis of mixed FEM discretizations of the Westervelt equation.

Main contributions. The main theoretical results in this work are as follows: under some sensible assumptions on the smallness and regularity of the exact solution, we show that

(i) There exists a unique solution to the proposed HDG semidiscrete formulation.
(ii) Optimal convergence rates of order $\mathcal{O}(h^{p+1})$ are achieved for the error of the method in some energy norms. In particular, the higher accuracy obtained for the approximation of the acoustic particle velocity $\underline{\boldsymbol{v}}$ exceeds the one expected for standard DG discretizations; cf. [2]. An accurate numerical approximation of $\underline{\boldsymbol{v}}$ is relevant, *e.g.*, for enforcing absorbing conditions [35] or gradient-based shape optimization of focused ultrasound devices [21,28].
(iii) The method is asymptotic preserving (*i.e.*, the commutative diagram in Fig. 1 holds), which implies that the semidiscrete approximation does not degenerate when $\delta \to 0^+$.

These theoretical results are validated in Section 6 below by some numerical examples. In addition, we numerically observe superconvergence of the discrete approximation of ψ obtained by the local postprocessing technique in equation (2.2) of [11].

Outline of the paper. In Section 2, we introduce the discrete spaces and the HDG semidiscrete formulation for model (1.2). In Section 3, we study the well-posedness and derive *a priori* error estimates for an auxiliary linearized problem. By means of a fixed-point argument, such results are extended in Section 4 to the nonlinear Westervelt equation. Section 5 is devoted to establishing the convergence of the numerical scheme to its vanishing δ-limit. In Section 6, we describe a fully discrete scheme obtained by combining the proposed HDG method with a predictor-corrector Newmark time discretization, and illustrate our theoretical findings with some numerical experiments. We end this work with some concluding remarks in Section 7.

Notation. We denote the first, second, and third partial derivatives with respect to the time variable t of a function v by $\partial_t v$, $\partial_{tt} v$, and $\partial_{ttt} v$, respectively.

We shall use the notation $x \lesssim y$, which stands for $x \leq Cy$, where C is a generic constant that does not depend on the mesh size parameter h nor on the sound diffusivity parameter δ.

Standard notation for L^p, Sobolev, and Bochner spaces is employed throughout. For example, for a given bounded, Lipschitz domain $D \subset \mathbb{R}^d$ ($d \in \mathbb{N}$) and $s \in \mathbb{R}^+$, the Sobolev space $H^s(D)$ is endowed with the standard inner product $(\cdot,\cdot)_{s,D}$, the seminorm $|\cdot|_{H^s(D)}$, and the norm $\|\cdot\|_{H^s(D)}$. In particular, for $s = 0$, the space $H^0(D) := L^2(D)$ is the space of Lebesgue square integrable functions over D, and we simply denote its standard inner product by $(\cdot,\cdot)_D$.

Let $n \in \mathbb{N}$, $p \in [1,\infty]$, and X be a Banach space, and denote by ∂_t^i the ith partial derivative with respect to time. The Bochner space

$$W^{n,p}(0,T;X) := \{u \in L^p(0,T;X), \quad \partial_t^i u \in L^p(0,T;X) \quad \forall i \leq n\}$$

is endowed with the norm

$$\|u\|_{W^{n,p}(0,T;X)} := \sum_{i=0}^{n} \|\partial_t^i u\|_{L^p(0,T;X)} \qquad \text{for all } u \in W^{n,p}(0,T;X).$$

2. Semidiscrete HDG formulation

Let $\{\mathcal{T}_h\}_{h>0}$ be a family of conforming simplicial meshes for the domain Ω satisfying the standard shape-regularity and quasi-uniformity conditions. We denote by $\mathcal{F}_h = \mathcal{F}_h^{\mathcal{I}} \cup \mathcal{F}_h^{\mathcal{D}}$ the set of mesh facets of $\mathcal{T}_h$, where $\mathcal{F}_h^{\mathcal{I}}$ and $\mathcal{F}_h^{\mathcal{D}}$ are the sets of internal and Dirichlet boundary facets, respectively. For each element $K \in \mathcal{T}_h$, we denote by $(\partial K)^{\mathcal{I}}$ and $(\partial K)^{\mathcal{D}}$ the union of the facets of K that belong to $\mathcal{F}_h^{\mathcal{I}}$ and $\mathcal{F}_h^{\mathcal{D}}$, respectively. Denoting the diameter of each element K by h_K, we define the mesh size $h := \max_{K \in \mathcal{T}_h} h_K$.

Given $p \in \mathbb{N}$, we define the following piecewise polynomial spaces:

$$\mathcal{S}_h^p := \prod_{K \in \mathcal{T}_h} \mathbb{P}^p(K), \qquad \boldsymbol{\mathcal{Q}}_h^p := \prod_{K \in \mathcal{T}_h} \mathbb{P}^p(K)^d, \qquad \mathcal{M}_h^p := \prod_{F \in \mathcal{F}_h^{\mathcal{I}}} \mathbb{P}^p(F), \tag{2.1}$$

where $\mathbb{P}^p(K)$ and $\mathbb{P}^p(F)$ denote the spaces of polynomials of total degree at most p on K and F, respectively. We denote by $[\![\cdot]\!]_{\mathsf{N}}$ the normal jump operator, which is defined for all $w_h \in \mathcal{S}_h^p$ and $\underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p$ as

$$\begin{cases} [\![w_h]\!]_{\mathsf{N}} := w_{h|_{K_1}} \underline{\mathbf{n}}_{K_1} + w_{h|_{K_2}} \underline{\mathbf{n}}_{K_2} \\ [\![\underline{\boldsymbol{r}}_h]\!]_{\mathsf{N}} := \underline{\boldsymbol{r}}_{h|_{K_1}} \cdot \underline{\mathbf{n}}_{K_1} + \underline{\boldsymbol{r}}_{h|_{K_2}} \cdot \underline{\mathbf{n}}_{K_2} \end{cases} \quad \text{on } F = \partial K_1 \cap \partial K_2 \in \mathcal{F}_h^{\mathcal{I}}, \text{ for some } K_1, K_2 \in \mathcal{T}_h,$$

where $\underline{\mathbf{n}}_K$ denotes the outward-pointing unit normal vector on ∂K. For any positive real number s, we define the following broken Sobolev space:

$$H^s(\mathcal{T}_h) := \big\{ v \in L^2(\Omega) \ : \ v_{|_K} \in H^s(K) \ \forall K \in \mathcal{T}_h \big\}.$$

The proposed hybridizable discontinuous Galerkin semidiscrete formulation for the Westervelt equation in (1.2) is[1]: for all $t \in (0,T]$, find $(\psi_h(\cdot,t), \underline{\boldsymbol{v}}_h(\cdot,t), \lambda_h(\cdot,t)) \in \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p \times \mathcal{M}_h^p$ such that the following equations are satisfied for all $K \in \mathcal{T}_h$:

$$\int_K \underline{\boldsymbol{v}}_h \cdot \underline{\boldsymbol{r}}_h \, \mathrm{d}\boldsymbol{x} = \int_{\partial K} \widehat{\psi}_h \underline{\boldsymbol{r}}_h \cdot \underline{\mathbf{n}}_K \, \mathrm{d}S - \int_K \psi_h \nabla \cdot \underline{\boldsymbol{r}}_h \, \mathrm{d}\boldsymbol{x} \qquad \forall \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p, \tag{2.2a}$$

[1] In this work, the vector variable $\underline{\boldsymbol{v}}_h$ approximates $\nabla\psi$, whereas it typically approximates $-\nabla\psi$ in elliptic problems. As a consequence, there are some slight differences in the standard HDG tools used in the coming sections.

$$\begin{aligned}\int_K (1+2k\partial_t\psi_h)\partial_{tt}\psi_h w_h\,\mathrm{d}\boldsymbol{x} - \int_{\partial K} w_h(c^2\widehat{\underline{\boldsymbol{v}}}_h + \delta\partial_t\widehat{\underline{\boldsymbol{v}}}_h)\cdot\underline{\mathbf{n}}_K\,\mathrm{d}S \\ + \int_K (c^2\underline{\boldsymbol{v}}_h + \delta\partial_t\underline{\boldsymbol{v}}_h)\cdot\nabla w_h\,\mathrm{d}\boldsymbol{x} = 0 \qquad \forall w_h \in \mathcal{S}_h^p,\end{aligned} \tag{2.2b}$$

the following compatibility equation is satisfied for all $F \in \mathcal{F}_h^{\mathcal{I}}$:

$$\int_F \mu_h [\![\widehat{\underline{\boldsymbol{v}}}_h]\!]_{\mathsf{N}}\,\mathrm{d}S = 0 \qquad \forall \mu_h \in \mathcal{M}_h^p, \tag{2.2c}$$

and appropriate discrete initial conditions, which will be specified in Section 3.3, are prescribed.

The *numerical fluxes* $\widehat{\psi}_h$ and $\widehat{\underline{\boldsymbol{v}}}_h$ are approximations of the traces of ψ_h and $\underline{\boldsymbol{v}}_h$ on $\mathcal{F}_h$, and are defined as follows (see [7], Sect. 3.2):

$$\widehat{\psi}_h := \begin{cases}\lambda_h & \text{if } F \in \mathcal{F}_h^{\mathcal{I}},\\ 0 & \text{if } F \in \mathcal{F}_h^{\mathcal{D}},\end{cases} \qquad \widehat{\underline{\boldsymbol{v}}}_h := \begin{cases}\underline{\boldsymbol{v}}_h - \tau(\psi_h - \lambda_h)\underline{\mathbf{n}}_K & \text{if } F \in \mathcal{F}_h^{\mathcal{I}},\\ \underline{\boldsymbol{v}}_h - \tau\psi_h\underline{\mathbf{n}}_\Omega & \text{if } F \in \mathcal{F}_h^{\mathcal{D}},\end{cases} \tag{2.3}$$

for some piecewise constant function τ that is double valued on $\mathcal{F}_h^{\mathcal{I}}$ and single valued on $\mathcal{F}_h^{\mathcal{D}}$. In particular, we consider the single-facet choice introduced in equation (1.6) of [5], *i.e.*, given a strictly positive constant $\bar{\tau}$, we define τ on each element $K \in \mathcal{T}_h$ as

$$\tau_{|\partial K} := \begin{cases}0 & \text{on } \partial K \backslash F_K^\tau,\\ \bar{\tau} & \text{on } F_K^\tau,\end{cases} \tag{2.4}$$

for a fixed facet F_K^τ of K. The compatibility condition (2.2c) implies that the normal component of $\widehat{\underline{\boldsymbol{v}}}_h$ is single valued on the mesh skeleton, *i.e.*, $[\![\widehat{\underline{\boldsymbol{v}}}_h]\!]_{\mathsf{N}} = 0$ on $\mathcal{F}_h^{\mathcal{I}}$.

We define the following inner products:

$$(u,v)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(u,v)_K, \qquad (u,v)_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(u,v)_{\partial K}, \qquad (u,v)_{(\partial\mathcal{T}_h)^{\mathcal{I}}} := \sum_{K\in\mathcal{T}_h}(u,v)_{(\partial K)^{\mathcal{I}}}.$$

Given bases for the spaces in (2.1), let M, $\boldsymbol{M}$, B, S, E, F, and G be the matrix representations of the following bilinear forms[2]:

$$\begin{aligned}
m_h(\psi_h, w_h) &:= (\psi_h, w_h)_{\mathcal{T}_h} && \forall \psi_h, w_h \in \mathcal{S}_h^p,\\
\boldsymbol{m}_h(\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{r}}_h) &:= (\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{r}}_h)_{\mathcal{T}_h} && \forall \underline{\boldsymbol{v}}_h, \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p,\\
b_h(\psi_h, \underline{\boldsymbol{r}}_h) &:= (\psi_h, \nabla\cdot\underline{\boldsymbol{r}}_h)_{\mathcal{T}_h} && \forall (\psi_h, \underline{\boldsymbol{r}}_h) \in \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p,\\
s_h(\psi_h, w_h) &:= (\tau\psi_h, w_h)_{\partial\mathcal{T}_h} && \forall \psi_h, w_h \in \mathcal{S}_h^p,\\
e_h(\lambda_h, \underline{\boldsymbol{r}}_h) &:= -(\lambda_h, [\![\underline{\boldsymbol{r}}_h]\!]_{\mathsf{N}})_{\mathcal{F}_h^{\mathcal{I}}} && \forall (\lambda_h, \underline{\boldsymbol{r}}_h) \in \mathcal{M}_h^p \times \boldsymbol{\mathcal{Q}}_h^p,\\
f_h(\lambda_h, w_h) &:= -(\tau\lambda_h, w_h)_{(\partial\mathcal{T}_h)^{\mathcal{I}}} && \forall (\lambda_h, w_h) \in \mathcal{M}_h^p \times \mathcal{S}_h^p,\\
g_h(\lambda_h, \mu_h) &:= (\tau\lambda_h, \mu_h)_{(\partial\mathcal{T}_h)^{\mathcal{I}}} && \forall \lambda_h, \mu_h \in \mathcal{M}_h^p,
\end{aligned}$$

and $\mathcal{N}_h(\cdot,\cdot)$ be the vector representation of the nonlinear operator

$$n_h(\phi_h; \theta_h, w_h) := \sum_{K\in\mathcal{T}_h}\int_K (1+2k\phi_h)\theta_h w_h\,\mathrm{d}\boldsymbol{x} \qquad \forall \phi_h, \theta_h, w_h \in \mathcal{S}_h^p.$$

Then, after summing up over all the elements $K \in \mathcal{T}_h$, replacing the numerical fluxes by their definition in (2.3), and using the following notation:

$$\widetilde{\lambda}_h = \lambda_h + \frac{\delta}{c^2}\partial_t\lambda_h, \quad \widetilde{\psi}_h = \psi_h + \frac{\delta}{c^2}\partial_t\psi_h, \quad \text{and} \quad \widetilde{\underline{\boldsymbol{v}}}_h = \underline{\boldsymbol{v}}_h + \frac{\delta}{c^2}\partial_t\underline{\boldsymbol{v}}_h, \tag{2.5}$$

[2]These bilinear forms are also well defined for sufficiently regular functions.

the semidiscrete HDG formulation (2.2) can be written in operator form as follows: for all $t \in (0,T]$, find $(\psi_h(\cdot,t), \underline{\boldsymbol{v}}_h(\cdot,t), \lambda_h(\cdot,t)) \in \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p \times \mathcal{M}_h^p$ such that

$$\boldsymbol{m}_h(\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{r}}_h) + b_h(\psi_h, \underline{\boldsymbol{r}}_h) + e_h(\lambda_h, \underline{\boldsymbol{r}}_h) = 0 \qquad \forall \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p, \tag{2.6a}$$

$$n_h(\partial_t\psi_h, \partial_{tt}\psi_h, w_h) - c^2 b_h(w_h, \widetilde{\underline{\boldsymbol{v}}}_h) + c^2 s_h(\widetilde{\psi}_h, w_h) + c^2 f_h(\widetilde{\lambda}_h, w_h) = 0 \qquad \forall w_h \in \mathcal{S}_h^p, \tag{2.6b}$$

$$-e_h(\mu_h, \underline{\boldsymbol{v}}_h) + f_h(\mu_h, \psi_h) + g_h(\lambda_h, \mu_h) = 0 \qquad \forall \mu_h \in \mathcal{M}_h^p, \tag{2.6c}$$

which leads to the following system of nonlinear ordinary differential equations (ODEs):

$$\begin{aligned} \boldsymbol{M}\mathbf{V}_h + B\Psi_h + E\Lambda_h &= 0, \\ \mathcal{N}_h\left(\frac{\mathrm{d}}{\mathrm{d}t}\Psi_h, \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Psi_h\right) - c^2 B^T \widetilde{\mathbf{V}}_h + c^2 S\widetilde{\Psi}_h + c^2 F\widetilde{\Lambda}_h &= 0, \\ -E^T\mathbf{V}_h + F^T\Psi_h + G\Lambda_h &= 0. \end{aligned}$$

Remark 2.1 (Structure of $\mathcal{N}_h(\cdot,\cdot)$)**.** Since the nonlinear operator $\mathcal{N}_h(\cdot,\cdot)$ is linear with respect to its second argument, it can also be written as $\mathcal{N}_h(\frac{\mathrm{d}}{\mathrm{d}t}\Psi_h, \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Psi_h) = N_h(\frac{\mathrm{d}}{\mathrm{d}t}\Psi_h)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Psi_h$, for some block diagonal matrix $N_h = N_h(\frac{\mathrm{d}}{\mathrm{d}t}\Psi_h)$.

Remark 2.2 (Linear case)**.** Setting $\delta = 0$ and $k = 0$ in the semidiscrete formulation (2.6), the conservative HDG method in [11] for the linear acoustic wave equation is recovered.

3. Linearized semidiscrete HDG formulation

As an intermediate step for the asymptotic and convergence analysis of the semidiscrete HDG formulation (2.6) for the Westervelt equation, we analyze an auxiliary linearized problem with damping parameter $\delta \geq 0$ and a variable coefficient. We first make some assumptions on the data of the linearized problem. In Section 3.1, we show some low- and high-order energy stability estimates and discuss the existence of a unique semidiscrete solution. In Section 3.2, we show some *a priori* error bounds in the energy norms. The choice of the discrete initial conditions is discussed in Section 3.3. Optimal h-convergence rates for the error in the energy norms are proven in Section 3.4.

We consider the following auxiliary, potentially damped, perturbed linear wave equation:

$$\begin{cases} (1+2k\alpha)\partial_{tt}\psi - c^2\nabla\cdot\underline{\boldsymbol{v}} - \delta\nabla\cdot(\partial_t\underline{\boldsymbol{v}}) = \varphi & \text{in } Q_T, \\ \underline{\boldsymbol{v}} = \nabla\psi + \underline{\boldsymbol{\Upsilon}} & \text{in } Q_T, \\ \psi = 0 & \text{on } \partial\Omega\times(0,T), \\ \psi = \psi_0, \qquad \partial_t\psi = \psi_1 & \text{on } \Omega\times\{0\}, \end{cases} \tag{3.1}$$

for some given functions $\varphi : Q_T \to \mathbb{R}$, α: $Q_T \to \mathbb{R}$, and $\underline{\boldsymbol{\Upsilon}}$: $Q_T \to \mathbb{R}^d$. The force term φ will be used to represent the consistency error due to the approximation of $\partial_t\psi$ by α. The perturbation function $\underline{\boldsymbol{\Upsilon}}$ will be used in Theorem 3.2 to represent the error resulting from the low-order $L^2(\Omega)$-orthogonality properties of the HDG projection in (3.9) of $\underline{\boldsymbol{v}}$. This will be useful in proving the error bounds of Theorem 3.7; see the system of error equations (3.16). Such an error term also appears in the analysis of the HDG method for the linear acoustic wave equation in Lemma 3.1 of [4].

We consider the following semidiscrete HDG formulation for the auxiliary problem in (3.1): for all $t \in (0,T]$, find $(\psi_h(\cdot,t), \underline{\boldsymbol{v}}_h(\cdot,t), \lambda_h(\cdot,t)) \in \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p \times \mathcal{M}_h^p$ such that

$$\boldsymbol{m}_h(\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{r}}_h) + b_h(\psi_h, \underline{\boldsymbol{r}}_h) + e_h(\lambda_h, \underline{\boldsymbol{r}}_h) = (\underline{\boldsymbol{\Upsilon}}, \underline{\boldsymbol{r}}_h)_\Omega \qquad \forall \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p, \tag{3.2a}$$

$$m_h((1+2k\alpha_h)\partial_{tt}\psi_h, w_h) - c^2 b_h(w_h, \widetilde{\underline{\boldsymbol{v}}}_h) + c^2 s_h\left(\widetilde{\psi}_h, w_h\right)$$

$$+c^2 f_h\left(\widetilde{\lambda}_h, w_h\right) = (\varphi, w_h)_\Omega \qquad \forall w_h \in \mathcal{S}_h^p, \tag{3.2b}$$

$$-e_h(\mu_h, \underline{\boldsymbol{v}}_h) + f_h(\mu_h, \psi_h) + g_h(\lambda_h, \mu_h) = 0 \qquad \forall \mu_h \in \mathcal{M}_h^p, \tag{3.2c}$$

where α_h is a discrete approximation of α. To complete the system of differential equations (3.2), it is necessary to compute appropriate discrete initial conditions from the initial data of the continuous problem ψ_0, ψ_1. A suitable choice for these initial conditions is essential in the error analysis below. We discuss our choice for the discrete initial conditions in Section 3.3.

To show the well-posedness of the semidiscrete problem (3.2), we make the following assumptions on the semidiscrete coefficient α_h, the forcing function φ, and the perturbation function $\underline{\boldsymbol{\Upsilon}}$.

Assumption 1. *Let $T > 0$. We assume that $\varphi \in H^1(0,T;L^2(\Omega))$, $\underline{\boldsymbol{\Upsilon}} \in W^{3,1}(0,T;L^2(\Omega)^d)$, and the coefficient $\alpha_h \in H^1(0,T;\mathcal{S}_h^p)$ is non degenerate, i.e., there exist constants $\underline{\alpha}$, $\overline{\alpha} > 0$ independent of h and δ, such that*

$$0 < 1 - 2|k|\underline{\alpha} \le 1 + 2k\alpha_h(\boldsymbol{x}, t) \le 1 + 2|k|\overline{\alpha} \qquad \forall (\boldsymbol{x}, t) \in \Omega \times (0,T). \tag{3.3}$$

Furthermore, we assume that there exist constants $0 < \gamma_0 < \sigma_0 < 1$ independent of h and the damping parameter δ such that

$$\frac{|k|}{1 - 2|k|\underline{\alpha}} \|\partial_t \alpha_h\|_{L^1(0,T;L^\infty(\Omega))} + \frac{\gamma_0}{2} \le \frac{\sigma_0}{2}. \tag{3.4}$$

Remark 3.1 (Linearization argument)**.** It is fairly common in the (numerical) analysis of quasilinear wave equations to combine a linearized problem with nondegeneracy assumptions on the variable coefficient. Such assumptions are then shown to be verified by the solution to the nonlinear problem by using a fixed-point strategy; see Theorem 4.1 below. See also Theorem 3 of [2], Theorem 6.1 of [33], and Theorem 4.1 of [29] for similar arguments.

3.1. Well-posedness and energy estimates

In this section, we discuss the existence and uniqueness of the solution to the semidiscrete formulation (3.2), and derive some low- and high-order energy stability estimates.

We first write the semidiscrete formulation (3.2) in matrix form as

$$\boldsymbol{M}\mathbf{V}_h + B\Psi_h + E\Lambda_h = \mathbf{\Gamma}, \tag{3.5a}$$

$$N_h(\alpha_h)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Psi_h - c^2 B^T \widetilde{\mathbf{V}}_h + c^2 S\widetilde{\Psi}_h + c^2 F\widetilde{\Lambda}_h = \Phi, \tag{3.5b}$$

$$-E^T\mathbf{V}_h + F^T\Psi_h + G\Lambda_h = 0, \tag{3.5c}$$

where Φ and $\mathbf{\Gamma}$ are, respectively, the vector representations of the terms in (3.2a) and (3.2b) involving ϕ and $\underline{\boldsymbol{\Upsilon}}$. The matrix $N_h = N_h(\alpha_h)$, defined in Remark 2.1, is symmetric positive definite on account of the nondegeneracy assumption made in (3.3) on α_h. From (3.5a) and (3.5c), we deduce that

$$\begin{pmatrix} \boldsymbol{M} & E \\ -E^T & G \end{pmatrix} \begin{pmatrix} \widetilde{\mathbf{V}}_h \\ \widetilde{\Lambda}_h \end{pmatrix} = \begin{pmatrix} -B\widetilde{\Psi}_h + \widetilde{\mathbf{\Gamma}} \\ -F^T\widetilde{\Psi}_h \end{pmatrix}. \tag{3.6}$$

Since $\boldsymbol{M}$ and G are symmetric positive definite matrices, the block matrix on the left-hand side of (3.6) is nonsingular. Therefore, $\widetilde{\mathbf{V}}_h$ and $\widetilde{\Lambda}_h$ can be expressed in terms of $\widetilde{\Psi}_h$ and $\widetilde{\mathbf{\Gamma}}$ through (3.6). This implies that the ODE system (3.5) can be reduced to a second-order linear ODE system involving only Ψ_h by multiplying equation (3.5b) by the matrix $N_h(\alpha_h)^{-1}$. If Assumption 1 holds, classical ODE theory (see, *e.g.*, [1], Thm. 1.8) predicts the existence of a unique solution $\psi_h \in W^{3,1}(0,T;\mathcal{S}_h^p)$. Moreover, through (3.5a) and (3.5c), we obtain that $\underline{\boldsymbol{v}}_h \in W^{3,1}(0,T;\boldsymbol{\mathcal{Q}}_h^p)$ and $\lambda_h \in W^{3,1}(0,T;\mathcal{M}_h^p)$. In the analysis below, the embedding $W^{3,1}(0,T) \hookrightarrow C^2([0,T])$ is of utmost relevance.

We derive low- and high-order energy stability estimates for the semidiscrete formulation (3.2).

Theorem 3.2 (Energy estimates for the discrete linearized problem)**.** *Let $T > 0$, $c > 0$, and $\delta \geq 0$. Assume that the semidiscrete-in-space coefficient α_h, the forcing function φ, and the perturbation function $\underline{\boldsymbol{\Upsilon}}$ satisfy Assumption 1. Then, the solution to semidiscrete formulation (3.2) satisfies the following energy stability estimates:*

$$\sup_{t\in(0,T)} \mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) \leq (1-\sigma_0)^{-1}\Bigg(\mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](0) + \frac{T}{2\gamma_0(1-2|k|\underline{\alpha})}\|\varphi\|^2_{L^2(0,T;L^2(\Omega))} + \left(\frac{\delta}{4} + \frac{c^2 T}{2\sigma_0}\right)\|\partial_t \underline{\boldsymbol{\Upsilon}}\|^2_{L^2(0,T;L^2(\Omega)^d)}\Bigg), \tag{3.7a}$$

$$\sup_{t\in(0,T)} \mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) \leq (1-\sigma_0)^{-1}\Bigg(\mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](0) + \frac{T}{2\gamma_0(1-2|k|\underline{\alpha})}\|\partial_t\varphi\|^2_{L^2(0,T;L^2(\Omega))} + \left(\frac{\delta}{4} + \frac{c^2 T}{2\sigma_0}\right)\|\partial_{tt} \underline{\boldsymbol{\Upsilon}}\|^2_{L^2(0,T;L^2(\Omega)^d)}\Bigg), \tag{3.7b}$$

where σ_0 is the constant in the smallness assumption (3.4), and the discrete energy functionals $\mathcal{E}_h^{(0)}[\cdot,\cdot,\cdot](t)$ and $\mathcal{E}_h^{(1)}[\cdot,\cdot,\cdot](t)$ are given by

$$\mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) := \frac{1}{2}\left\|\sqrt{1+2k\alpha_h}\,\partial_t\psi_h\right\|^2_{L^2(\Omega)} + \frac{c^2}{2}\left(\|\underline{\boldsymbol{v}}_h\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}(\lambda_h - \psi_h)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\psi_h\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\right),$$

$$\mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) := \frac{1}{2}\left\|\sqrt{1+2k\alpha_h}\,\partial_{tt}\psi_h\right\|^2_{L^2(\Omega)} + \frac{c^2}{2}\left(\|\partial_t\underline{\boldsymbol{v}}_h\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}(\partial_t\lambda_h - \partial_t\psi_h)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\partial_t\psi_h\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\right).$$

Proof. The proofs of the energy estimates in (3.7a) and (3.7b) are postponed to Appendices A and B, respectively. ◻

Remark 3.3 (Regularity of $\underline{\boldsymbol{\Upsilon}}$)**.** As can be seen from estimates (3.7a) and (3.7b), it is sufficient to have $\underline{\boldsymbol{\Upsilon}} \in H^2(0,T;L^2(\Omega)^d)$. However, this would degrade the regularity to be expected from the solution to the semidiscrete problem (3.5). In particular, we would only get that $\underline{\boldsymbol{v}}_h \in H^2(0,T;\boldsymbol{\mathcal{Q}}_h^p)$ and $\lambda_h \in H^2(0,T;\mathcal{M}_h^p)$. Since $\underline{\boldsymbol{\Upsilon}}$ is only an auxiliary function used to represent the error introduced by the low-order $L^2(\Omega)$-orthogonality of the HDG projection used in the error analysis (see Thm. 3.7 below), we assume $\underline{\boldsymbol{\Upsilon}} \in W^{3,1}(0,T;L^2(\Omega)^d)$, thus retaining the expected regularity of $\underline{\boldsymbol{v}}_h$ and λ_h when $\underline{\boldsymbol{\Upsilon}} = 0$ as in the original problem (2.6).

Estimates (3.7a) and (3.7b) show boundedness of the energy of the semidiscrete solution with respect to the initial energies, the forcing function φ, and the perturbation function $\underline{\boldsymbol{\Upsilon}}$. In order to show that these constitute indeed stability results, we need to show that the initial discrete energies, $\mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](0)$ and $\mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](0)$, remain bounded uniformly in h. We prove the stability result for the nonlinear problem in Lemma 4.3.

Remark 3.4 (Stabilization parameter)**.** In order to obtain the energy stability estimates in (3.7a) and (3.7b), we only require the stabilization parameter $\overline{\tau}$ in (2.4) to be strictly positive. Moreover, there are no polynomial inverse estimates involved in the proof of Theorem 3.2.

3.2. *A priori* error estimates

In this section, we carry out an *a priori* error analysis for the semidiscrete formulation (3.2). To do so, we first recall the properties of some special HDG projections. For all $\epsilon > 0$, let $\mathcal{P}_{\mathcal{M}} : H^{\frac{1}{2}+\epsilon}(\mathcal{T}_h) \to \mathcal{M}_h^p$ be the

L^2-orthogonal projection in $\mathcal{M}_h^p$, defined for all $u \in H^{\frac{1}{2}+\epsilon}(\mathcal{T}_h)$ as

$$(\mathcal{P}_{\mathcal{M}} u - u, \mu_h)_{(\partial \mathcal{T}_h)^{\mathcal{I}}} = 0 \qquad \forall \mu_h \in \mathcal{M}_h^p, \tag{3.8}$$

and let $\Pi_{\mathsf{HDG}} := (\Pi_{\mathcal{S}}, \Pi_{\boldsymbol{\mathcal{Q}}}) : H^{\frac{1}{2}+\epsilon}(\mathcal{T}_h) \times H^{\frac{1}{2}+\epsilon}(\mathcal{T}_h)^d \to \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p$ be the HDG projection in equation (2.1) of [8], defined for all $(\psi, \underline{\boldsymbol{v}}) \in H^{\frac{1}{2}+\epsilon}(\mathcal{T}_h) \times H^{\frac{1}{2}+\epsilon}(\mathcal{T}_h)^d$ and all $K \in \mathcal{T}_h$ as

$$(\Pi_{\boldsymbol{\mathcal{Q}}} \underline{\boldsymbol{v}} - \underline{\boldsymbol{v}}, \underline{\boldsymbol{r}}_h)_K = 0 \qquad \forall \underline{\boldsymbol{r}}_h \in \mathbb{P}^{p-1}(K)^d, \tag{3.9a}$$

$$(\Pi_{\mathcal{S}} \psi - \psi, w_h)_K = 0 \qquad \forall w_h \in \mathbb{P}^{p-1}(K), \tag{3.9b}$$

$$\left(\left(\widehat{\Pi_{\boldsymbol{\mathcal{Q}}} \underline{\boldsymbol{v}}} - \underline{\boldsymbol{v}}\right) \cdot \underline{\mathbf{n}}_K, \mu_h\right)_f = 0 \qquad \forall \text{ facets } F \subset \partial K, \ \forall \mu_h \in \mathbb{P}^p(F), \tag{3.9c}$$

where

$$\widehat{\Pi_{\boldsymbol{\mathcal{Q}}} \underline{\boldsymbol{v}}} \cdot \underline{\mathbf{n}}_K := \Pi_{\boldsymbol{\mathcal{Q}}} \underline{\boldsymbol{v}} \cdot \underline{\mathbf{n}}_K - \tau (\Pi_{\mathcal{S}} \psi - \mathcal{P}_{\mathcal{M}} \psi) \qquad \text{on } \partial K.$$

Let $(\psi, \underline{\boldsymbol{v}})$ be the solution to the continuous Westervelt equation in (1.2), and let $(\psi_h, \underline{\boldsymbol{v}}_h)$ be the solution to the semidiscrete formulation (3.2) for the linearized problem (3.1) with $\underline{\boldsymbol{\Upsilon}} = 0$ and $\varphi = 0$.

We define the following error functions:

$$\varepsilon_\psi := \psi - \psi_h, \qquad \boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}} := \underline{\boldsymbol{v}} - \underline{\boldsymbol{v}}_h, \qquad \varepsilon_\lambda := \psi - \lambda_h, \tag{3.10a}$$

$$\xi_\psi := \Pi_{\mathcal{S}} \psi - \psi, \qquad \boldsymbol{\xi}_{\underline{\boldsymbol{v}}} := \Pi_{\boldsymbol{\mathcal{Q}}} \underline{\boldsymbol{v}} - \underline{\boldsymbol{v}}, \qquad \xi_\lambda := \mathcal{P}_{\mathcal{M}} \psi - \psi, \tag{3.10b}$$

$$\eta_{\psi,h} := \Pi_{\mathcal{S}} \psi - \psi_h, \qquad \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h} := \Pi_{\boldsymbol{\mathcal{Q}}} \underline{\boldsymbol{v}} - \underline{\boldsymbol{v}}_h, \qquad \eta_{\lambda,h} := \mathcal{P}_{\mathcal{M}} \psi - \lambda_h, \tag{3.10c}$$

and recall the approximation properties of Π_{HDG} in Theorem 2.1 of [8].

Lemma 3.5 (Approximation properties of Π_{HDG}). *Suppose $p \geq 0$, $\tau_{|\partial K}$ is nonnegative, and $\tau_K^{\max} := \max \tau_{|\partial K} > 0$. Then, $\Pi_{\mathsf{HDG}}(\psi, \underline{\boldsymbol{v}}) = (\Pi_{\mathcal{S}} \psi, \Pi_{\boldsymbol{\mathcal{Q}}} \underline{\boldsymbol{v}})$ is well defined. Furthermore, there is a constant $C_\Pi > 0$ independent of K and τ such that*

$$\|\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}\|_{L^2(K)} \leq C_\Pi \left(h_K^{s_{\underline{\boldsymbol{v}}}+1} |\underline{\boldsymbol{v}}|_{H^{s_{\underline{\boldsymbol{v}}}+1}(K)^d} + h_K^{s_\psi+1} \tau_K^\star |\psi|_{H^{s_\psi+1}(K)} \right),$$

$$\|\xi_\psi\|_{L^2(K)} \leq C_\Pi \left(h_K^{s_\psi+1} |\psi|_{H^{s_\psi+1}(K)} + \frac{h_K^{s_{\underline{\boldsymbol{v}}}+1}}{\tau_K^{\max}} |\nabla \cdot \underline{\boldsymbol{v}}|_{H^{s_{\underline{\boldsymbol{v}}}}(K)} \right),$$

for $s_\psi, s_{\underline{\boldsymbol{v}}} \in [0, p]$ and $(\psi, \underline{\boldsymbol{v}}) \in H^{s_\psi+1}(K) \times H^{s_{\underline{\boldsymbol{v}}}+1}(K)^d$. Above, $\tau_K^\star := \max \tau_{|\partial K \setminus F^\star}$, where $F^\star$ is a facet of K at which $\tau_{|\partial K}$ is maximum.

For the single-facet choice in (2.4), we have that $\tau_K^\star = 0$ and $\tau_K^{\max} = \bar{\tau}$ for all $K \in \mathcal{T}_h$. In particular, the error bound for $\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}$ does not depend on the regularity of ψ.

The following lemma is crucial for the error analysis of HDG methods.

Lemma 3.6. *For all $(\psi, \underline{\boldsymbol{v}}) \in H^{\frac{1}{2}+\epsilon}(\mathcal{T}_h) \times H^1(\mathcal{T}_h)^d$, it holds*

$$b_h(w_h, \boldsymbol{\xi}_{\underline{\boldsymbol{v}}}) = s_h(w_h, \xi_\psi) \qquad \forall w_h \in \mathcal{S}_h^p. \tag{3.12}$$

Proof. This identity is an immediate consequence of the weak commutativity property in Proposition 2.1 of [8]. ☐

By the consistency of the proposed method and recalling the tilde ($\sim$) notation from (2.5), the following error equations are verified:

$$\boldsymbol{m}_h(\boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}, \underline{\boldsymbol{r}}_h) + b_h(\varepsilon_\psi, \underline{\boldsymbol{r}}_h) + e_h(\varepsilon_\lambda, \underline{\boldsymbol{r}}_h) = 0 \qquad \forall \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p,$$

$$
\begin{aligned}
m_h((1+2k\alpha_h)\partial_{tt}\varepsilon_\psi, w_h) - c^2 b_h(w_h, \widetilde{\boldsymbol{\varepsilon}}_{\underline{\boldsymbol{v}}}) \\
+c^2 s_h(\widetilde{\varepsilon}_\psi, w_h) + c^2 f_h(\widetilde{\varepsilon}_\lambda, w_h) &= -m_h(2k(\partial_t\psi - \alpha_h)\partial_{tt}\psi, w_h) && \forall w_h \in \mathcal{S}_h^p, \\
-e_h(\mu_h, \boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}) + f_h(\mu_h, \varepsilon_\psi) + g_h(\varepsilon_\lambda, \mu_h) &= 0 && \forall \mu_h \in \mathcal{M}_h^p.
\end{aligned}
$$

We are in a position to obtain *a priori* error bounds for the semidiscrete linearized formulation (3.2) with respect to the continuous solution to the Westervelt equation in (1.2).

Theorem 3.7 (Error bounds for the semidiscrete linearized formulation)**.** *Under the assumptions of Theorem 3.2, the following error bounds are satisfied:*

$$
\begin{aligned}
&\sup_{t\in(0,T)}\left(\frac{c^2}{2}\|\boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d} + \frac{1}{2}\|\sqrt{1+2k\alpha_h}\partial_t\varepsilon_\psi\|^2_{L^2(\Omega)}\right) \\
&\quad\le \sup_{t\in(0,T)}\left(c^2\|\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d} + \left\|\sqrt{1+2k\alpha_h}\partial_t\xi_\psi\right\|^2_{L^2(\Omega)}\right) + 2(1-\sigma_0)^{-1}\Bigg(\mathcal{E}_h^{(0)}[\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \eta_{\lambda,h}](0) \\
&\qquad + \frac{T}{2\gamma_0(1-2|k|\underline{\alpha})}\|\hat{\varphi}\|^2_{L^2(0,T;L^2(\Omega))} + \left(\frac{\delta}{4} + \frac{c^2T}{2\sigma_0}\right)\|\partial_t\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}\|^2_{L^2(0,T;L^2(\Omega)^d)}\Bigg),
\end{aligned}
\tag{3.14a}
$$

$$
\begin{aligned}
&\sup_{t\in(0,T)}\left(\frac{c^2}{2}\|\partial_t\boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d} + \frac{1}{2}\left\|\sqrt{1+2k\alpha_h}\partial_{tt}\varepsilon_\psi\right\|^2_{L^2(\Omega)}\right) \\
&\quad\le \sup_{t\in(0,T)}\left(c^2\|\partial_t\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d} + \left\|\sqrt{1+2k\alpha_h}\partial_{tt}\xi_\psi\right\|^2_{L^2(\Omega)}\right) + 2(1-\sigma_0)^{-1}\Bigg(\mathcal{E}_h^{(1)}\left[\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \eta_{\lambda,h}\right](0) \\
&\qquad + \frac{T}{2\gamma_0(1-2|k|\underline{\alpha})}\|\partial_t\hat{\varphi}\|^2_{L^2(0,T;L^2(\Omega))} + \left(\frac{\delta}{4} + \frac{c^2T}{2\sigma_0}\right)\left\|\partial_{tt}\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}\right\|^2_{L^2(0,T;L^2(\Omega)^d)}\Bigg),
\end{aligned}
\tag{3.14b}
$$

where $\hat{\varphi} \in H^1(0,T;\mathcal{S}_h^p)$ *is given by*

$$
\hat{\varphi} = \Pi_0[(1+2k\alpha_h)\partial_{tt}\xi_\psi + 2k(\partial_t\psi - \alpha_h)\partial_{tt}\psi], \tag{3.15}
$$

with Π_0 *denoting the* $L^2(\Omega)$*-orthogonal projection in* $\mathcal{S}_h^p$*.*

Proof. We only present the proof of the error bound in (3.14a), as the proof of (3.14b) is similar.

We split the error functions in (3.10a) as

$$
\varepsilon_\psi = \eta_{\psi,h} - \xi_\psi, \qquad \boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}} = \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h} - \boldsymbol{\xi}_{\underline{\boldsymbol{v}}}, \qquad \varepsilon_\lambda = \eta_{\lambda,h} - \xi_\lambda.
$$

The definition of the HDG projections in (3.8) and (3.9) implies that, for all $t \in (0,T]$, the discrete error functions $(\eta_{\psi,h}(\cdot,t), \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}(\cdot,), \eta_{\lambda,h}(\cdot,t)) \in \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p \times \mathcal{M}_h^p$ solve a semidiscrete linearized problem as in (3.2). More precisely, they satisfy the following equations for all $(w_h, \underline{\boldsymbol{r}}_h, \mu_h) \in \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p \times \mathcal{M}_h^p$:

$$
\boldsymbol{m}_h\big(\boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \underline{\boldsymbol{r}}_h\big) + b_h(\eta_{\psi,h}, \underline{\boldsymbol{r}}_h) + e_h(\eta_{\lambda,h}, \underline{\boldsymbol{r}}_h) = -\big(\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}, \underline{\boldsymbol{r}}_h\big)_\Omega \tag{3.16a}
$$

$$
m_h((1+2k\alpha_h)\partial_{tt}\eta_{\psi,h}, w_h) - c^2 b_h\big(w_h, \widetilde{\boldsymbol{\eta}}_{\underline{\boldsymbol{v}},h}\big) + c^2 s_h(\widetilde{\eta}_{\psi,h}, w_h) + c^2 f_h(\widetilde{\eta}_{\lambda,h}, w_h) = (\hat{\varphi}, w_h)_\Omega, \tag{3.16b}
$$

$$
-e_h\big(\mu_h, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}\big) + f_h(\mu_h, \eta_{\psi,h}) + g_h(\eta_{\lambda,h}, \mu_h) = 0, \tag{3.16c}
$$

where $\hat{\varphi} \in H^1(0,T;\mathcal{S}_h^p)$ is a lifting function defined by the following projection:

$$
\begin{aligned}
(\hat{\varphi}, w_h)_\Omega := {}& m_h((1+2k\alpha_h)\partial_{tt}\xi_\psi, w_h) + m_h(2k(\partial_t\psi - \alpha_h)\partial_{tt}\psi, w_h) \\
&- c^2 b_h(w_h, \widetilde{\boldsymbol{\xi}}_{\underline{\boldsymbol{v}}}) + c^2 s_h(\widetilde{\xi}_\psi, w_h) + c^2 f_h\Big(\widetilde{\xi}_\lambda, w_h\Big) \qquad \forall w_h \in \mathcal{S}_h^p.
\end{aligned}
$$

From the definition of $\mathcal{P}_{\mathcal{M}}$ in (3.8) and identity (3.12), we deduce that

$$f_h\left(\widetilde{\xi}_\lambda, w_h\right) = 0 \qquad \text{and} \qquad -b_h\left(w_h, \widetilde{\boldsymbol{\xi}}_{\underline{\boldsymbol{v}}}\right) + s_h\left(\widetilde{\xi}_\psi, w_h\right) = 0 \qquad \forall w_h \in \mathcal{S}_h^p,$$

which implies that $\hat{\varphi}$ satisfies (3.15).

The desired bound is then obtained from the triangle inequality and the energy estimate (3.7a) in Theorem 3.2. □

3.3. Choice of the discrete initial conditions

All the results presented so far are valid for any choice of the discrete initial conditions. However, in order to show optimal convergence rates for the error in the low- and high-order energy norms, we assume that $\psi_0, \psi_1 \in H^2(\Omega) \cap H_0^1(\Omega)$ and choose the discrete initial conditions $\psi_h^{(i)}$ $(i = 0, 1)$ as the solution to the following discrete HDG elliptic problem: find $(\psi_h^{(i)}, \underline{\boldsymbol{v}}_h^{(i)}, \lambda_h^{(i)}) \in \mathcal{S}_h^p \times \boldsymbol{\mathcal{Q}}_h^p \times \mathcal{M}_h^p$ such that

$$\boldsymbol{m}_h\left(\underline{\boldsymbol{v}}_h^{(i)}, \underline{\boldsymbol{r}}_h\right) + b_h\left(\psi_h^{(i)}, \underline{\boldsymbol{r}}_h\right) + e_h\left(\lambda_h^{(i)}, \underline{\boldsymbol{r}}_h\right) = 0 \qquad \forall \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p, \tag{3.17a}$$

$$-b_h\left(w_h, \underline{\boldsymbol{v}}_h^{(i)}\right) + s_h\left(\psi_h^{(i)}, w_h\right) + f_h\left(\lambda_h^{(i)}, w_h\right) = (-\Delta\psi_i, w_h)_{\mathcal{T}_h} \qquad \forall w_h \in \mathcal{S}_h^p, \tag{3.17b}$$

$$-e_h\left(\mu_h, \underline{\boldsymbol{v}}_h^{(i)}\right) + f_h\left(\mu_h, \psi_h^{(i)}\right) + g_h\left(\lambda_h^{(i)}, \mu_h\right) = 0 \qquad \forall \mu_h \in \mathcal{M}_h^p. \tag{3.17c}$$

This choice of the discrete initial conditions can be interpreted as an HDG variant of the well-known Ritz projection, which was used in the numerical analysis for the strongly damped Westervelt equation in [33].

The variational problem (3.17) corresponds to the HDG discretization of a Poisson problem with homogeneous Dirichlet boundary conditions and a source term given by $-\Delta\psi_i$. Therefore, the existence and uniquess of a solution to (3.17) follows from Theorem 2.3 of [5].

In next lemma, we provide bounds for the terms containing the discrete errors $(\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \eta_{\lambda,h})$ on the right-hand side of the *a priori* bounds (3.14a) and (3.14b).

Lemma 3.8 (Estimates at $t = 0$). *Assume that $\psi_0, \psi_1 \in H^2(\mathcal{T}_h) \cap H_0^1(\Omega)$, and the discrete initial conditions are chosen as in (3.17). Then, the following bounds hold:*

$$\mathcal{E}_h^{(0)}\left[\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \eta_{\lambda,h}\right](0) \leq \frac{(1 + 2|k|\overline{\alpha})}{2} \left\|\Pi_{\mathcal{S}}\psi_1 - \psi_h^{(1)}\right\|_{L^2(\Omega)}^2 + \frac{c^2}{2}\left\|\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}(\cdot, 0)\right\|_{L^2(\Omega)^d}^2, \tag{3.18a}$$

$$\begin{aligned}\mathcal{E}_h^{(1)}\left[\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \eta_{\lambda,h}\right](0) \leq{}& \frac{c^2}{2}\left\|\partial_t\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}(\cdot, 0)\right\|_{L^2(\Omega)^d}^2 + \frac{(1 + 2|k|\overline{\alpha})^2}{(1 - 2|k|\underline{\alpha})}\|\partial_{tt}\xi_\psi(\cdot, 0)\|_{L^2(\Omega)}^2 \\ &+ \frac{4k^2}{1 - 2|k|\underline{\alpha}}\|(\partial_t\psi - \alpha_h)(\cdot, 0)\partial_{tt}\psi(\cdot, 0)\|_{L^2(\Omega)}^2.\end{aligned} \tag{3.18b}$$

Moreover, if the domain Ω is such that

$$\varphi \in H_0^1(\Omega), \quad \Delta\varphi \in L^2(\Omega) \implies \varphi \in H^2(\Omega), \tag{3.19}$$

then, there exists a constant $C_* > 0$ independent of h and δ such that

$$\left\|\Pi_{\mathcal{S}}\psi_1 - \psi_h^{(1)}\right\|_{L^2(\Omega)} \leq C_* h\left\|\partial_t\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}(\cdot, 0)\right\|_{L^2(\Omega)^d}. \tag{3.20}$$

Proof. By using the nondegeneracy assumption in (3.3), the low-order bound in (3.18a) can be proven as in Lemma 3.6 of [11] for the linear wave equation, whereas estimate (3.20) follows from Theorem 4.1 of [8]. In contrast to Lemma 3.6 of [11], due to the choice of $\psi_h^{(1)}$ in (3.17), the term $\partial_t\eta_{\psi,h}(\cdot, 0) = \Pi_{\mathcal{S}}\psi_1 - \psi_h^{(1)}$ does not vanish.

As for bound (3.18b), proceeding again as in Lemma 3.6 of [11], we get

$$\mathcal{E}_h^{(1)}[\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \eta_{\lambda,h}](0) \leq \frac{1}{2}\left\|\sqrt{1+2k\alpha_h(\cdot,0)}\partial_{tt}\eta_{\psi,h}(\cdot,0)\right\|_{L^2(\Omega)}^2 + \frac{c^2}{2}\left\|\partial_t \boldsymbol{\xi}_{\underline{\boldsymbol{v}}}(\cdot,0)\right\|_{L^2(\Omega)^d}. \tag{3.21}$$

Hence, it only remains to bound the first term on the right-hand side of (3.21). To do so, we choose $w_h = \partial_{tt}\eta_{\psi,h}(\cdot,0)$ in (3.16b) for $t=0$ (the explicit evaluation at $t=0$ is omitted in the subsequent steps), which leads to the following identity:

$$\begin{aligned}\left\|\sqrt{1+2k\alpha_h}\partial_{tt}\eta_{\psi,h}\right\|_{L^2(\Omega)}^2 &= c^2\big(b_h\big(\partial_{tt}\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}\big) - s_h(\eta_{\psi,h}, \partial_{tt}\eta_{\psi,h}) - f_h(\eta_{\lambda,h}, \partial_{tt}\eta_{\psi,h})\big)\\ &\quad + \delta\big(b_h\big(\partial_{tt}\eta_{\psi,h}, \partial_t\boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}\big) - s_h(\partial_t\eta_{\psi,h}, \partial_{tt}\eta_{\psi,h}) - f_h(\partial_t\eta_{\lambda,h}, \partial_{tt}\eta_{\psi,h})\big)\\ &\quad + (\hat{\varphi}, \partial_{tt}\eta_{\psi,h})_\Omega,\end{aligned}$$

where $\hat{\varphi} \in \mathcal{S}_h^p(\mathcal{T}_h)$ is defined in (3.15).

The choice of the discrete initial conditions $\psi_h^{(i)}$ $(i=0,1)$ in (3.17), the definition of $\mathcal{P}_{\mathcal{M}}$ in (3.8), and identity (3.12) imply that

$$\begin{aligned}b_h\big(\partial_{tt}\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}\big) - s_h(\eta_{\psi,h}, \partial_{tt}\eta_{\psi,h}) - f_h(\eta_{\lambda,h}, \partial_{tt}\eta_{\psi,h}) = 0,\\ b_h\big(\partial_{tt}\eta_{\psi,h}, \partial_t\boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}\big) - s_h(\partial_t\eta_{\psi,h}, \partial_{tt}\eta_{\psi,h}) - f_h(\partial_t\eta_{\lambda,h}, \partial_{tt}\eta_{\psi,h}) = 0.\end{aligned}$$

Therefore, using the Cauchy–Schwarz inequality and the stability of the $L^2(\Omega)$-orthogonal projection Π_0, we get

$$\begin{aligned}\frac{1}{2}\left\|\sqrt{1+2k\alpha_h}\partial_{tt}\eta_{\psi,h}\right\|_{L^2(\Omega)}^2 &\leq \frac{1}{2}\left\|(1+2k\alpha_h)^{-\frac{1}{2}}\hat{\varphi}\right\|_{L^2(\Omega)}^2\\ &\leq (1-2|k|\underline{\alpha})^{-1}\Big((1+2|k|\overline{\alpha})^2\|\partial_{tt}\xi_\psi(\cdot,0)\|_{L^2(\Omega)}^2 + 4k^2\|(\partial_t\psi-\alpha_h)(\cdot,0)\partial_{tt}\psi(\cdot,0)\|_{L^2(\Omega)}^2\Big),\end{aligned}$$

which, together with bound (3.21), completes the proof. □

3.4. *h*-convergence

In order to obtain optimal h-convergence rates in Theorem 3.9 below for the error in the low- and high-order energy norms, we will assume that the nonlinear Westervelt equation in (1.2) has a regular enough solution. We refer the reader to [20, 23] for δ-uniform analyses of the Westervelt equation. Higher-order regularity of the exact solution follows from Theorem 2.2 of [24] under stronger regularity and smallness assumptions on the initial conditions, and higher-order compatibility of the initial and boundary data.

Henceforth, we assume that $h<1$. We will also make the following assumption on how well the semidiscrete coefficient α_h approximates $\partial_t\psi$. This assumption will later be verified by means of a fixed-point argument.

Assumption 2. *For given $s_\psi, s_{\underline{\boldsymbol{v}}} \in [0,p]$, we assume that the semidiscrete coefficient α_h and its time derivative $\partial_t\alpha_h$ approximate $\partial_t\psi$ and $\partial_{tt}\psi$, respectively, up to the following accuracy:*

$$\begin{aligned}\|\partial_t\psi-\alpha_h\|_{L^\infty(0,t;L^2(\Omega))} &\leq C_*\Big(h^{s_\psi+1}\|\psi\|_{H^2(0,t;H^{s_\psi+1}(\Omega))} + h^{s_{\underline{\boldsymbol{v}}}+1}\|\underline{\boldsymbol{v}}\|_{H^2(0,t;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}\Big),\\ \|\partial_{tt}\psi-\partial_t\alpha_h\|_{L^2(0,t;L^2(\Omega))} &\leq C_*\Big(h^{s_\psi+1}\|\psi\|_{H^3(0,t;H^{s_\psi+1}(\Omega))} + h^{s_{\underline{\boldsymbol{v}}}+1}\|\underline{\boldsymbol{v}}\|_{H^3(0,t;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}\Big),\end{aligned}$$

for all $t\in[0,T]$, where the constant $C_*>0$ does not depend on h or δ.

To establish the higher-order-in-time error estimate in (3.23b) below, we make a uniform boundedness assumption on the time derivative of the linear coefficient α_h, namely, we require that

$$\|\partial_t\alpha_h\|_{L^2(0,T;L^\infty(\Omega))} \leq \breve{\alpha}, \tag{3.22}$$

for some positive constant $\breve{\alpha}$ independent of h and δ.

The smallness assumption in (3.22) matches the one made in Assumption W1 of [29] for the analysis of the mixed FEM approximation of the Westervelt equation.

Theorem 3.9 (Error estimate for the semidiscrete linearized problem)**.** *Let $h \in (0, \overline{h})$ and let the assumptions of Theorem 3.2 and Assumption 2 hold. Let additionally $\psi \in H^3(0,T; H_0^1(\Omega) \cap H^{s_\psi+1}(\Omega))$ for some $s_\psi \in [0,p]$ and $\underline{\boldsymbol{v}} \in H^3(0,T; H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)$ for some $s_{\underline{\boldsymbol{v}}} \in [0,p]$ be the solution to the IBVP for the Westervelt equation in (1.2). Let also Ω be such that the regularity condition in (3.19) holds, and the discrete initial condition be chosen as in Section 3.3. Then,*

$$\sup_{t\in(0,T)} \Big(\|\boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d} + \|\partial_t \varepsilon_\psi\|^2_{L^2(\Omega)} \Big) \lesssim \Big(h^{2s_\psi+2} \|\psi\|^2_{H^2(0,T;H^{s_\psi+1}(\Omega))} + h^{2s_{\underline{\boldsymbol{v}}}+2} \|\underline{\boldsymbol{v}}\|^2_{H^2(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)} \Big) \Big(1 + \|\partial_{tt}\psi\|^2_{L^2(0,T;L^\infty(\Omega))} \Big), \tag{3.23a}$$

and

$$\sup_{t\in(0,T)} \Big(\|\partial_t \boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d} + \|\partial_{tt} \varepsilon_\psi\|^2_{L^2(\Omega)} \Big) \lesssim \Big(h^{2s_\psi+2} \|\psi\|^2_{H^3(0,T;H^{s_\psi+1}(\Omega))} + h^{2s_{\underline{\boldsymbol{v}}}+2} \|\underline{\boldsymbol{v}}\|^2_{H^3(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)} \Big) \Big(1 + \|\partial_{tt}\psi\|^2_{L^\infty(0,T;L^\infty(\Omega))} + \|\partial_{ttt}\psi\|^2_{L^2(0,T;L^\infty(\Omega))} \Big), \tag{3.23b}$$

where the hidden constants are independent of h and δ.

Proof. We start from the estimates in Theorem 3.7. We then combine them with Lemma 3.8, the Hölder inequality, and the approximation properties in Lemma 3.5 of the HDG projection. Furthermore, the terms involving the forcing function $\hat{\varphi}$ in (3.15) are estimated using the Cauchy–Schwarz and the Hölder inequalities as follows:

$$\begin{aligned} \|\hat{\varphi}\|_{L^2(0,T;L^2(\Omega))} &\le \|1 + 2k\alpha_h\|_{L^\infty(0,T;L^\infty(\Omega))} \|\partial_{tt}\xi_\psi\|_{L^2(0,T;L^2(\Omega))} \\ &\quad + 2|k| \|\partial_t\psi - \alpha_h\|_{L^\infty(0,T;L^2(\Omega))} \|\partial_{tt}\psi\|_{L^2(0,T;L^\infty(\Omega))}, \\ \|\partial_t\hat{\varphi}\|_{L^2(0,T;L^2(\Omega))} &\le \|\partial_t\alpha_h\|_{L^2(0,T;L^\infty(\Omega))} \|\partial_{tt}\xi_\psi\|_{L^\infty(0,T;L^2(\Omega))} \\ &\quad + \|1 + 2k\alpha_h\|_{L^\infty(0,T;L^\infty(\Omega))} \|\partial_{ttt}\xi_\psi\|_{L^2(0,T;L^2(\Omega))}, \\ &\quad + \|\partial_{tt}\psi - \partial_t\psi_h\|_{L^2(0,T;L^2(\Omega))} \|\partial_{tt}\psi\|_{L^\infty(0,T;L^2(\Omega))} \\ &\quad + \|\partial_t\psi - \alpha_h\|_{L^\infty(0,T;L^2(\Omega))} \|\partial_{ttt}\psi\|_{L^2(0,T;L^\infty(\Omega))}. \end{aligned}$$

Finally, the terms involving the semidiscrete coefficient α_h can be bounded using Assumption 2.

The following estimates are then obtained:

$$\begin{aligned} &\sup_{t\in(0,T)} \Big(\|\boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d} + \|\partial_t\varepsilon_\psi\|^2_{L^2(\Omega)} \Big) \\ &\lesssim h^{2s_{\underline{\boldsymbol{v}}}+2} \bigg(|\underline{\boldsymbol{v}}(\cdot,0)|^2_{H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d} + h|\partial_t\underline{\boldsymbol{v}}(\cdot,0)|^2_{H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d} + |\nabla\cdot(\partial_{tt}\underline{\boldsymbol{v}})|^2_{L^2(0,T;H^{s_{\underline{\boldsymbol{v}}}}(\Omega))} \\ &\qquad + \sup_{t\in(0,T)} |\underline{\boldsymbol{v}}|^2_{H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d} + \sup_{t\in(0,T)} |\nabla\cdot(\partial_t\underline{\boldsymbol{v}})|^2_{H^{s_{\underline{\boldsymbol{v}}}}(\Omega)} + |\partial_t\underline{\boldsymbol{v}}|^2_{L^2(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)} \\ &\qquad + \|\underline{\boldsymbol{v}}\|^2_{H^2(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)} \Big(1 + \|\partial_{tt}\psi\|^2_{L^2(0,T;L^\infty(\Omega))} \Big) \bigg) \\ &\quad + h^{2s_\psi+2} \bigg(\sup_{t\in(0,T)} |\partial_t\psi|^2_{H^{s_\psi+1}(\Omega)} + |\partial_{tt}\psi|^2_{L^2(0,T;H^{s_\psi+1}(\Omega))} + \|\partial_{tt}\psi\|^2_{L^2(0,T;H^{s_\psi+1}(\Omega))} \\ &\qquad + \|\psi\|^2_{H^2(0,T;H^{s_\psi+1}(\Omega))} \|\partial_{tt}\psi\|^2_{L^2(0,T;L^\infty(\Omega))} \bigg), \end{aligned}$$

$$
\begin{aligned}
&\sup_{t\in(0,T)}\Big(\|\partial_t\boldsymbol{\varepsilon}_{\underline{\boldsymbol{v}}}\|^2_{L^2(\Omega)^d}+\|\partial_{tt}\varepsilon_\psi\|^2_{L^2(\Omega)}\Big)\\
&\quad\lesssim h^{2s_{\underline{\boldsymbol{v}}}+2}\Bigg(|\nabla\cdot(\partial_{tt}\underline{\boldsymbol{v}})(\cdot,0)|^2_{H^{s_{\underline{\boldsymbol{v}}}}(\Omega)}+|\partial_t\underline{\boldsymbol{v}}(\cdot,0)|^2_{H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d}+\sup_{t\in(0,T)}|\partial_t\underline{\boldsymbol{v}}|^2_{H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d}\\
&\qquad\qquad+\sup_{t\in(0,T)}|\nabla\cdot(\partial_{tt}\underline{\boldsymbol{v}})|^2_{H^{s_{\underline{\boldsymbol{v}}}}(\Omega)}+|\partial_{tt}\underline{\boldsymbol{v}}|^2_{L^2(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}+\|\nabla\cdot(\partial_{tt}\underline{\boldsymbol{v}})\|^2_{L^\infty(0,T;H^{s_{\underline{\boldsymbol{v}}}}(\Omega)^d)}\\
&\qquad\qquad+\|\nabla\cdot(\partial_{ttt}\underline{\boldsymbol{v}})\|^2_{L^2(0,T;H^{s_{\underline{\boldsymbol{v}}}}(\Omega)^d)}+\|\underline{\boldsymbol{v}}\|^2_{H^3(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}\|\partial_{ttt}\psi\|^2_{L^2(0,T;L^\infty(\Omega))}\Bigg)\\
&\qquad+h^{2s_\psi+2}\Bigg(|\partial_{tt}\psi(\cdot,0)|^2_{H^{s_\psi+1}(\Omega)}+\sup_{t\in(0,T)}|\partial_{tt}\psi|^2_{H^{s_\psi+1}(\Omega)}+\|\partial_{ttt}\psi\|^2_{L^2(0,T;H^{s_\psi+1}(\Omega))}\\
&\qquad\qquad+\|\partial_{tt}\psi\|^2_{L^\infty(0,T;H^{s_\psi+1}(\Omega))}+\|\partial_{ttt}\psi\|^2_{L^2(0,T;H^{s_\psi+1}(\Omega))}\\
&\qquad\qquad+\|\psi\|^2_{H^2(0,T;H^{s_\psi+1}(\Omega))}\|\partial_{tt}\psi\|^2_{L^\infty(0,T;L^\infty(\Omega))}\\
&\qquad\qquad+\|\psi\|^2_{H^3(0,T;H^{s_\psi+1}(\Omega))}\|\partial_{ttt}\psi\|^2_{L^2(0,T;L^\infty(\Omega))}\Bigg),
\end{aligned}
$$

where the hidden constants are independent of h and δ. Using the Sobolev embeddings $H^2(0,T)\hookrightarrow C^1([0,T])$ and $H^3(0,T)\hookrightarrow C^2([0,T])$, and the fact that $h\in(0,\overline{h})$, we get the desired result. □

4. Analysis of the semidiscrete HDG formulation for the Westervelt equation

We are now in a position to analyze the nonlinear semidiscrete formulation (2.6). The main idea consists of employing a Banach fixed-point argument applied to the mapping

$$\mathcal{F}:\mathcal{B}_{\text{F-P}}\ni(\psi_h^*,\underline{\boldsymbol{v}}_h^*)\mapsto(\psi_h,\underline{\boldsymbol{v}}_h),$$

$(\psi_h,\underline{\boldsymbol{v}}_h)$ being the two first components (*i.e.*, we omit the λ_h component, which is uniquely determined by $(\psi_h,\underline{\boldsymbol{v}}_h)$; see also Rem. 4.5 bellow) of the unique solution to linear problem (3.2) with discrete initial conditions as in Section 3.3, $\underline{\boldsymbol{\Upsilon}}=0$, $\varphi=0$, and

$$\alpha_h=\partial_t\psi_h^*$$

from

$$
\begin{aligned}
\mathcal{B}_{\text{F-P}}:=\Bigg\{&(\psi_h^*,\underline{\boldsymbol{v}}_h^*)\in W^{2,\infty}(0,T;\mathcal{S}_h^p)\times W^{1,\infty}(0,T;\boldsymbol{\mathcal{Q}}_h^p):(\psi_h^*,\partial_t\psi_h^*)_{|t=0}=\big(\psi_h^0,\psi_h^1\big),\\
&\sup_{t\in(0,T)}\Big(\|\underline{\boldsymbol{v}}-\underline{\boldsymbol{v}}_h^*\|^2_{L^2(\Omega)^d}+\|\partial_t\psi-\partial_t\psi_h^*\|^2_{L^2(\Omega)}\Big)\\
&\qquad\le C_0\Big(h^{2s_\psi+2}\|\psi\|^2_{H^2(0,T;H^{s_\psi+1}(\Omega))}+h^{2s_{\underline{\boldsymbol{v}}}+2}\|\underline{\boldsymbol{v}}\|^2_{H^2(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}\Big),\\
&\sup_{t\in(0,T)}\Big(\|\partial_t\underline{\boldsymbol{v}}-\partial_t\underline{\boldsymbol{v}}_h^*\|^2_{L^2(\Omega)^d}+\|\partial_{tt}\psi-\partial_{tt}\psi_h^*\|^2_{L^2(\Omega)}\Big)\\
&\qquad\le C_1\Big(h^{2s_\psi+2}\|\psi\|^2_{H^3(0,T;H^{s_\psi+1}(\Omega))}+h^{2s_{\underline{\boldsymbol{v}}}+2}\|\underline{\boldsymbol{v}}\|^2_{H^3(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}\Big)\Bigg\},
\end{aligned}
\tag{4.1}
$$

which is a ball centered at the exact solution $(\psi,\underline{\boldsymbol{v}})\in H^3(0,T;H_0^1(\Omega)\cap H^{s_\psi+1}(\Omega))\times H^3(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)$ for some s_ψ, $s_{\underline{\boldsymbol{v}}}\in(\frac{d}{2}-1,p]$. In the definition of $\mathcal{B}_{\text{F-P}}$, C_0 and C_1 are positive constants independent of h and δ that will be fixed in the proof of Theorem 4.1.

Next theorem concerns the existence and uniqueness of the solution to the semidiscrete formulation (2.6). Moreover, it provides optimal *a priori* error estimates due to the definition of the ball $\mathcal{B}_{\text{F-P}}$. We denote by I_h the Lagrange interpolation operator in $\mathcal{S}_h^p$. In particular, we will use the approximation result in Theorem 4.4.20 of [3] and the inverse estimate in Theorem 4.5.11 of [3].

Theorem 4.1. *Let $\delta \in [0, \overline{\delta})$, $p > \frac{d}{2} - 1$, and s_ψ, $s_{\underline{\boldsymbol{v}}} \in (\frac{d}{2} - 1, p]$. Assume that $(\psi, \underline{\boldsymbol{v}}) \in H^3(0,T; H_0^1(\Omega) \cap H^{s_\psi+1}(\Omega)) \times H^3(0,T; H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)$ is the solution to the Westervelt equation in (1.2) for suitable initial conditions $(\psi, \psi_t)_{|t=0} = (\psi_0, \psi_1)$. Furthermore, let the discrete initial conditions $(\psi_h, \partial_t \psi_h)_{|t=0}$ be chosen as in Section 3.3. Then, there exist $T > 0$,*

$$\overline{h} = \overline{h}\Big(\|\psi\|_{H^3(0,T;H^{s_\psi+1}(\Omega))}, \|\underline{\boldsymbol{v}}\|_{H^3(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}\Big) < 1, \quad \text{and} \quad 0 < M = M(k,T),$$

such that, for $0 < h < \overline{h}$ and

$$\int_0^T \|\partial_{ttt}\psi(s)\|^2_{L^\infty(\Omega)}\, \mathrm{d}s + \sup_{t\in(0,T)} \|\partial_{tt}\psi(t)\|^2_{L^\infty(\Omega)} + \int_0^T \|\partial_{tt}\psi(s)\|^2_{L^\infty(\Omega)}\, \mathrm{d}s + \sup_{t\in(0,T)} \|\partial_t\psi(t)\|^2_{L^\infty(\Omega)} \leq M,$$

there is a unique solution $(\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h) \in \mathcal{B}_{\text{F-P}} \times W^{1,\infty}(0,T; \mathcal{M}_h^p)$ to the semidiscrete HDG formulation (2.6) for some constants C_0, $C_1 > 0$ in the definition of $\mathcal{B}_{\text{F-P}}$ that are independent of h and δ.

Proof. We proceed by using a Banach fixed-point argument. The ball $\mathcal{B}_{\text{F-P}}$ is nonempty as it contains the HDG projection of the exact solution thanks to the estimates given in Lemma 3.5.

We split the proof into three parts. The first two are intended to prove the existence and uniqueness of a fixed point. The third part discusses the reconstruction of λ_h.

Part I: Self-mapping. Let $(\psi_h^*, \underline{\boldsymbol{v}}_h^*) \in \mathcal{B}_{\text{F-P}}$ and set

$$(\psi_h, \underline{\boldsymbol{v}}_h) = \mathcal{F}(\psi_h^*, \underline{\boldsymbol{v}}_h^*).$$

To show the self-mapping property, we use the error estimates in Theorem 3.9. We first verify that its assumptions hold. We start by considering the nondegeneracy assumption in (3.3). Using the triangle inequality, the quasi-uniformity of the mesh, and the stability and inverse estimates in Theorems 4.4.20 and 4.5.11 from [3] for the Lagrange interpolation operator, we obtain

$$\begin{aligned}
\|\alpha_h\|_{L^\infty(0,T;L^\infty(\Omega))} &\leq \|\partial_t\psi_h^* - I_h\partial_t\psi\|_{L^\infty(0,T;L^\infty(\Omega))} + \|I_h\partial_t\psi\|_{L^\infty(0,T;L^\infty(\Omega))} \\
&\lesssim h^{-d/2}\|\partial_t\psi_h^* - I_h\partial_t\psi\|_{L^\infty(0,T;L^2(\Omega))} + \|I_h\partial_t\psi\|_{L^\infty(0,T;L^\infty(\Omega))} \\
&\lesssim h^{-d/2}\|\partial_t\psi_h^* - \partial_t\psi\|_{L^\infty(0,T;L^2(\Omega))} + h^{-d/2}\|\partial_t\psi - I_h\partial_t\psi\|_{L^\infty(0,T;L^2(\Omega))} \\
&\quad + \|I_h\partial_t\psi\|_{L^\infty(0,T;L^\infty(\Omega))}.
\end{aligned} \tag{4.2}$$

Thus, we can guarantee that the nondegeneracy condition in (3.3) holds with

$$\underline{\alpha} = \overline{\alpha} = \overline{C}\Big(\overline{h}^{s_\psi+1-d/2}\|\psi\|_{H^3(0,T;H^{s_\psi+1}(\Omega))} + \overline{h}^{s_{\underline{\boldsymbol{v}}}+1-d/2}\|\underline{\boldsymbol{v}}\|_{H^2(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)} + M^{1/2}\Big) \in \Big(0, \frac{1}{2|k|}\Big), \tag{4.3}$$

for sufficiently small M and $\overline{h}$, and some positive constant $\overline{C}$ depending on C_0 and C_1, but not on h or δ.

Similarly, the smallness assumptions in (3.4) and (3.22) can be shown to hold provided M, $\overline{h}$, and the final time T are sufficiently small. Assumption 2 is naturally verified since $(\psi_h^*, \underline{\boldsymbol{v}}_h^*) \in \mathcal{B}_{\text{F-P}}$. Therefore, Theorem 3.9 ensures the self-mapping property of $\mathcal{F}$ (*i.e.*, $\mathcal{F}(\mathcal{B}_{\text{F-P}}) \subseteq \mathcal{B}_{\text{F-P}}$) provided that C_0 and C_1 are large enough, and M is sufficiently small.

Part II: Strict contractivity. Contractivity of the mapping $\mathcal{F}$ follows similarly as in Theorem 5.1 from [29], where the δ-robustness of the mixed FEM for the Westervelt equation was proven. Indeed, one can obtain the contractivity of $\mathcal{F}$ with respect to the lower topology $\sup_{t\in(0,T)} \mathcal{E}_h^{(0)}[\cdot,\cdot,\cdot](t)$ by reducing M and $\overline{h}$. The arguments showing the closedness of $\mathcal{B}_{\text{F-P}}$ with respect to the lower topology are analogous to Theorem 1.4 of [20]. This shows that the fixed-point problem has a unique solution in $\mathcal{B}_{\text{F-P}}$, which solves the nonlinear problem (2.6).

Part III: Reconstructing λ_h. Parts I and II ensure the existence of a unique fixed point $(\psi_h, \underline{\boldsymbol{v}}_h) \in \mathcal{B}_{\text{F-P}}$ to the mapping $\mathcal{F}$. To finish constructing the solution to the semidiscrete HDG formulation (2.6) we reconstruct λ_h as a function of $(\psi_h, \underline{\boldsymbol{v}}_h)$ uniquely through (2.6c). The triplet $(\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h) \in \mathcal{B}_{\text{F-P}} \times W^{1,\infty}(0,T;\mathcal{M}_h^p)$ thus constructed is the unique solution to (2.6). ☐

Along the lines of the analysis performed in Section 4 of [32] for the conforming FEM, we state here a corollary of the previous existence and uniqueness theorem, which will be useful in Section 5 below for establishing the rate of convergence as $\delta \to 0^+$.

Corollary 4.2. *Under the assumptions of Theorem 4.1, the solution $(\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h)$ to (2.6) satisfies*

$$\|\partial_{tt}\psi_h\|_{L^\infty(0,T;L^\infty(\Omega))} \le C\Big(\|\psi\|_{H^3(0,T;H^{s_\psi+1}(\Omega))} + \|\underline{\boldsymbol{v}}\|_{H^3(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}\Big), \tag{4.4}$$

where $C > 0$ does not depend on h or δ. Furthermore, the following bound holds:

$$\|\partial_t\psi_h\|_{L^\infty(0,T;L^\infty(\Omega))} \le \overline{\alpha}. \tag{4.5}$$

Proof. The uniform-in-h-and-δ bounds follow from the use of inverse estimates as in (4.2). ☐

We end this section showing that the solution to the semidiscrete formulation (2.6) from Theorem 4.1 is energy stable. In the proof of the next result, we use the embedding $H^3(0,T;H_0^1(\Omega)\cap H^2(\Omega)) \hookrightarrow C^2([0,T];H_0^1(\Omega)\cap H^2(\Omega))$.

Lemma 4.3 (Energy stability)**.** *Let the assumptions of Theorem 4.1 hold. Moreover, assume that the solution $(\psi, \underline{\boldsymbol{v}})$ to the Westervelt equation in (1.2) belongs to $H^3(0,T;H_0^1(\Omega)\cap H^2(\Omega)) \times H^3(0,T;H^2(\Omega)^d)$. Then, there exists a constant $C_S > 0$ independent of $h \in (0,\overline{h})$ and $\delta \in [0,\overline{\delta})$ such that*

$$\sup_{t\in(0,T)} \mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) \le C_S\big(\|\psi_0\|_{H^2(\Omega)} + \|\psi_1\|_{H^2(\Omega)}\big), \tag{4.6a}$$

$$\sup_{t\in(0,T)} \mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) \le C_S\Big(\|\psi_1\|_{H^2(\Omega)} + \|\psi_{tt}(\cdot,0)\|_{H^2(\Omega)}\Big), \tag{4.6b}$$

with $\alpha_h = \partial_t\psi_h$ in the definition of $\mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t)$ and $\mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t)$.

Proof. The proof follows by considering the solution to the nonlinear semidiscrete problem in (2.6) as the solution to the linearized problem in (3.2) with $\alpha_h = \partial_t\psi_h$. We can then proceed similarly as in Section 3.1 to deduce that $(\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h) \in W^{3,1}(0,T;\mathcal{S}_h^p) \times W^{3,1}(0,T;\boldsymbol{\mathcal{Q}}_h^p) \times W^{3,1}(0,T;\mathcal{M}_h^p)$. By using similar arguments to those for the low- and high-order energy stability estimates in Theorem 3.2, we get

$$\sup_{t\in(0,T)} \mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) \le (1-\sigma_0)^{-1}\mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](0), \tag{4.7a}$$

$$\sup_{t\in(0,T)} \mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) \le (1-\sigma_0)^{-1}\mathcal{E}_h^{(1)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](0). \tag{4.7b}$$

Therefore, it only remains to bound the initial discrete energies.

The following estimate follows from the stability of the discrete HDG elliptic problem in (3.17):

$$\left\|\underline{\boldsymbol{v}}_h^{(i)}\right\|_{L^2(\Omega)^d}^2 + \left\|\tau^{\frac{1}{2}}\left(\lambda_h^{(i)} - \psi_h^{(i)}\right)\right\|_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})}^2 + \left\|\tau^{\frac{1}{2}}\psi_h^{(i)}\right\|_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}^2 \leq \frac{1}{2}\|\Delta\psi_i\|_{L^2(\mathcal{T}_h)}^2 + \frac{1}{2}\left\|\psi_h^{(i)}\right\|_{L^2(\Omega)}^2 \quad \text{for } i = 0, 1. \tag{4.8}$$

Using the triangle inequality and the error estimate in Corollary 2.7 of [5] for second-order elliptic problems, we obtain

$$\left\|\psi_h^{(i)}\right\|_{L^2(\Omega)} \leq \left\|\psi_h^{(i)} - \psi_i\right\|_{L^2(\Omega)} + \|\psi_i\|_{L^2(\Omega)} \leq \max\{1, Ch^2\}\|\psi_i\|_{H^2(\Omega)}, \tag{4.9}$$

for $i = 0, 1$. This shows that we can estimate the right-hand side of (4.8) independently of h. In particular, we can estimate

$$\begin{aligned}\left\|\sqrt{1 + 2k\partial_t\psi_h(\cdot, 0)}\psi_h^{(1)}\right\|_{L^2(\Omega)}^2 &\leq (1 + 2|k|\overline{\alpha})\left(\|\psi_1\|_{L^2(\Omega)} + \left\|\psi_h^{(1)} - \psi_1\right\|_{L^2(\Omega)}\right) \\ &\leq (1 + 2|k|\overline{\alpha})\max\{1, Ch^2\}\|\psi_1\|_{H^2(\mathcal{T}_h)},\end{aligned} \tag{4.10}$$

for some positive constant C independent of h. Bound (4.6a) then follows by combining (4.7a), bounds (4.8) and (4.9) for $i = 0$, and (4.10).

By the triangle inequality, we get

$$\begin{aligned}\left\|\sqrt{1 + 2k\partial_t\psi_h(\cdot, 0)}\partial_{tt}\psi_h(\cdot, 0)\right\|_{L^2(\Omega)} \leq{}& \left\|\sqrt{1 + 2k\partial_t\psi_h(\cdot, 0)}\partial_{tt}\psi(\cdot, 0)\right\|_{L^2(\Omega)} \\ &+ \left\|\sqrt{1 + 2k\partial_t\psi_h(\cdot, 0)}\partial_{tt}\xi_\psi(\cdot, 0)\right\|_{L^2(\Omega)} \\ &+ \left\|\sqrt{1 + 2k\partial_t\psi_h(\cdot, 0)}\partial_{tt}\eta_{\psi,h}(\cdot, 0)\right\|_{L^2(\Omega)}.\end{aligned} \tag{4.11}$$

The third term on the right-hand side of the above inequality satisfies

$$\left\|\sqrt{1 + 2k\partial_t\psi_h(\cdot, 0)}\partial_{tt}\eta_{\psi,h}(\cdot, 0)\right\|_{L^2(\Omega)}^2 \leq \mathcal{E}_h^{(1)}[\eta_{\psi,h}, \boldsymbol{\eta}_{\underline{\boldsymbol{v}},h}, \eta_{\lambda,h}](0),$$

which can be bounded using the approximation properties in Lemma 3.5 of the HDG projection Π_{HDG} due to (3.18b). Moreover, the following estimates hold:

$$\begin{aligned}\left\|\partial_t\boldsymbol{\xi}_{\underline{\boldsymbol{v}}}(\cdot, 0)\right\|_{L^2(\Omega)} &\lesssim h\|\psi_1\|_{H^2(\Omega)}, \\ \|\partial_{tt}\xi_\psi(\cdot, 0)\|_{L^2(\Omega)} &\lesssim h\Big(|\partial_{tt}\psi(\cdot, 0)|_{H^1(\Omega)} + \|\partial_{tt}\psi(\cdot, 0)\|_{H^2(\Omega)}\Big).\end{aligned}$$

Introducing these bounds into (4.11), combining it with bounds (4.8) and (4.9) for $i = 1$, and using the nondegeneracy of $\partial_t\psi_h$ complete the proof of bound (4.6b). ☐

Remark 4.4 (Minimum degree of approximation)**.** The condition $p > \frac{d}{2} - 1$ in the statement of Theorem 4.1, combined with the restriction $d \in \{2, 3\}$ on the spatial dimension, imposes that the degree of approximation must satisfy $p \geq 1$.

Nevertheless, in the case $d = 2$, we can ensure the nondegeneracy of $\partial_t\psi_h$ even for $p = \frac{d}{2} - 1 = 0$, by assuming smallness of the exact solution $\|\psi\|_{H^3(0,T;H^{s_\psi+1}(\Omega))} + \|\underline{\boldsymbol{v}}\|_{H^2(0,T;H^{s_{\underline{\boldsymbol{v}}}+1}(\Omega)^d)}$; see equation (4.3). This is relevant in practice, as is shown in the numerical experiments of Section 6.

Remark 4.5 (Omission of λ_h)**.** In the definition of the ball $\mathcal{B}_{\text{F-P}}$, we have omitted the component λ_h of the solution to the linearized semidiscrete problem in (3.2), as Theorem 3.9 does not provide an error control for this component. Nonetheless, given the fixed-point $(\psi_h, \underline{\boldsymbol{v}}_h)$ of the mapping $\mathcal{F}$, which solves the nonlinear semidiscrete formulation in (2.6), the component λ_h is uniquely determined by $(\psi_h, \underline{\boldsymbol{v}}_h)$ through (2.6c), as was used in the proof of Theorem 4.1. In fact, one can also define the mapping $\mathcal{F}$ in terms of the first component ψ_h only, as the nonlinearity solely depends on such a component.

5. Asymptotic behaviour at the vanishing viscosity limit

This section is dedicated to the proof of convergence of the numerical scheme as $\delta \to 0^+$. We denote in this section by $(\psi_h^{(\delta)}, \underline{\boldsymbol{v}}_h^{(\delta)}, \lambda_h^{(\delta)})$ the solution to the semidiscrete formulation (2.6), where we have stressed the dependence of the solution on the parameter δ. Then, we denote the difference

$$(\overline{\psi}_h, \overline{\underline{\boldsymbol{v}}}_h, \overline{\lambda}_h) = \left(\psi_h^{(\delta)} - \psi_h^{(0)}, \underline{\boldsymbol{v}}_h^{(\delta)} - \underline{\boldsymbol{v}}_h^{(0)}, \lambda_h^{(\delta)} - \lambda_h^{(0)}\right),$$

which satisfies the following system of equations:

$$\boldsymbol{m}_h(\overline{\underline{\boldsymbol{v}}}_h, \underline{\boldsymbol{r}}_h) + b_h(\overline{\psi}_h, \underline{\boldsymbol{r}}_h) + e_h(\overline{\lambda}_h, \underline{\boldsymbol{r}}_h) = 0 \qquad \forall \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p, \tag{5.1a}$$

$$\begin{aligned} m_h\Big(\Big(1 + 2k\partial_t\psi_h^{(\delta)}\Big)\partial_{tt}\overline{\psi}_h, w_h\Big) &+ m_h\Big(2k\partial_t\overline{\psi}_h\partial_{tt}\psi_h^{(0)}, w_h\Big) \\ -c^2 b_h(w_h, \overline{\underline{\boldsymbol{v}}}_h) + c^2 s_h(\overline{\psi}_h, w_h) &+ c^2 f_h(\overline{\lambda}_h, w_h) = \delta F(w_h) \qquad \forall w_h \in \mathcal{S}_h^p, \end{aligned} \tag{5.1b}$$

$$-e_h(\mu_h, \overline{\underline{\boldsymbol{v}}}_h) + f_h(\mu_h, \overline{\psi}_h) + g_h(\overline{\lambda}_h, \mu_h) = 0 \qquad \forall \mu_h \in \mathcal{M}_h^p, \tag{5.1c}$$

with zero initial conditions. Above, the forcing term $F(w_h)$ is given by

$$F(w_h) = b_h\Big(w_h, \partial_t\underline{\boldsymbol{v}}_h^{(\delta)}\Big) - s_h\Big(\partial_t\psi_h^{(\delta)}, w_h\Big) - f_h\Big(\partial_t\lambda_h^{(\delta)}, w_h\Big).$$

Theorem 5.1 (δ-convergence)**.** *Let the assumptions of Theorem 4.1 hold, and let $\overline{h}$ and T be fixed as in Theorem 4.1. Then, the family of solutions $\big\{(\psi_h^{(\delta)}, \underline{\boldsymbol{v}}_h^{(\delta)}, \lambda_h^{(\delta)})\big\}_{\delta\in[0,\bar{\delta})}$ converges to $(\psi_h^{(0)}, \underline{\boldsymbol{v}}_h^{(0)}, \lambda_h^{(0)})$ as $\delta \to 0^+$, and*

$$\sup_{t\in(0,T)} \mathcal{E}_h^{(0)}(t)\big[\overline{\psi}_h, \overline{\underline{\boldsymbol{v}}}_h, \overline{\lambda}_h\big] \le C(T)\delta, \tag{5.2a}$$

$$\sup_{t\in(0,T)} \big\|\overline{\psi}_h(t)\big\|_{L^2(\Omega)}^2 \le C(T)\delta^2, \tag{5.2b}$$

where $C(T)$ is a generic constant that depends exponentially on T.

Proof. We prove each estimate separately.

Proof of estimate (5.2a)**.** The proof follows by a similar energy argument to that used to establish the low-order stability bound in Appendix A. We differentiate (5.1a) in time once and then take the test functions $\underline{\boldsymbol{r}}_h = \overline{\underline{\boldsymbol{v}}}_h$, $w_h = \partial_t\overline{\psi}_h$, and $\mu_h = \partial_t\overline{\lambda}_h$. Multiplying the first and third equations by c^2, and summing the results, we get the identity

$$\begin{aligned} & m_h\Big(\Big(1 + 2k\partial_t\psi_h^{(\delta)}\Big)\partial_{tt}\overline{\psi}_h, \partial_t\overline{\psi}_h\Big) + m_h\Big(2k\partial_t\overline{\psi}_h\partial_{tt}\psi_h^{(0)}, \partial_t\overline{\psi}_h\Big) \\ & \quad + c^2\big(\boldsymbol{m}_h(\partial_t\overline{\underline{\boldsymbol{v}}}_h, \overline{\underline{\boldsymbol{v}}}_h) + s_h(\overline{\psi}_h, \partial_t\overline{\psi}_h) + f_h(\overline{\lambda}_h, \partial_t\overline{\psi}_h) + f_h(\partial_t\overline{\lambda}_h, \overline{\psi}_h) + g_h(\partial_t\lambda_h, \partial_t\overline{\lambda}_h)\big) \\ & \quad = \delta F(\partial_t\overline{\psi}_h). \end{aligned} \tag{5.3}$$

We consider the following identities, which follow from the definition of the discrete bilinear forms in Section 2:

$$m_h\Big(\Big(1 + 2k\partial_t\psi_h^{(\delta)}\Big)\partial_{tt}\overline{\psi}_h, \partial_t\overline{\psi}_h\Big) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big\|\sqrt{1 + 2k\partial_t\psi_h^{(\delta)}}\partial_t\overline{\psi}_h\Big\|_{L^2(\Omega)}^2 - m_h\Big(2k\partial_{tt}\psi_h^{(\delta)}\partial_t\overline{\psi}_h, \partial_t\overline{\psi}_h\Big),$$

$$\begin{aligned} & \boldsymbol{m}_h(\partial_t\overline{\underline{\boldsymbol{v}}}_h, \overline{\underline{\boldsymbol{v}}}_h) + s_h(\overline{\psi}_h, \partial_t\overline{\psi}_h) + f_h(\overline{\lambda}_h, \partial_t\overline{\psi}_h) + f_h(\partial_t\overline{\lambda}_h, \overline{\psi}_h) + g_h(\partial_t\lambda_h, \partial_t\overline{\lambda}_h) \\ & \quad = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\|\overline{\underline{\boldsymbol{v}}}_h\|_{L^2(\Omega)^d}^2 + \Big\|\tau^{\frac{1}{2}}(\overline{\lambda}_h - \overline{\psi}_h)\Big\|_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})}^2 + \Big\|\tau^{\frac{1}{2}}\overline{\psi}_h\Big\|_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}^2\Big). \end{aligned}$$

Using Corollary 4.2, we can ensure that

$$\int_0^t m_h\Big(\Big(1+2k\partial_t\psi_h^{(\delta)}\Big)\partial_{tt}\overline{\psi}_h,\partial_t\overline{\psi}_h\Big)\,\mathrm{d}s \geq \frac{1-2|k|\overline{\alpha}}{2}\big\|\partial_t\overline{\psi}_h\big\|^2_{L^2(\Omega)} - 2|k|\int_0^t\Big\|\partial_{tt}\psi_h^{(\delta)}\Big\|_{L^\infty(\Omega)}\big\|\partial_t\overline{\psi}_h\big\|^2_{L^2(\Omega)}\,\mathrm{d}s,$$

where the negative term on the right-hand side will be controlled using the Grönwall inequality.

It thus remains to control the term $\delta F(\partial_t\overline{\psi}_h)$. To this end, recall that the following equations hold:

$$\begin{aligned}
0 &= -\boldsymbol{m}_h(\partial_t\underline{\overline{\boldsymbol{v}}}_h,\underline{\boldsymbol{r}}_h) - b_h\big(\partial_t\overline{\psi}_h,\underline{\boldsymbol{r}}_h\big) - e_h\big(\partial_t\overline{\lambda}_h,\underline{\boldsymbol{r}}_h\big) && \forall\underline{\boldsymbol{r}}_h\in\boldsymbol{\mathcal{Q}}_h^p,\\
F(w_h) &= b_h\Big(w_h,\partial_t\underline{\boldsymbol{v}}_h^{(\delta)}\Big) - s_h\Big(\partial_t\psi_h^{(\delta)},w_h\Big) - f_h\Big(\partial_t\lambda_h^{(\delta)},w_h\Big) && \forall w_h\in\mathcal{S}_h^p,\\
0 &= e_h\Big(\mu_h,\partial_t\underline{\boldsymbol{v}}_h^{(\delta)}\Big) - f_h\Big(\mu_h,\partial_t\psi_h^{(\delta)}\Big) - g_h\Big(\partial_t\lambda_h^{(\delta)},\mu_h\Big) && \forall\mu_h\in\mathcal{M}_h^p.
\end{aligned}$$

Choosing $\underline{\boldsymbol{r}}_h = \partial_t\underline{\boldsymbol{v}}_h^{(\delta)}$, $w_h = \partial_t\overline{\psi}_h$, and $\mu_h = \partial_t\overline{\lambda}_h$ above, and summing up the results yield the following identity:

$$\begin{aligned}
F\big(\partial_t\overline{\psi}_h\big) = &-\boldsymbol{m}_h\Big(\partial_t\underline{\overline{\boldsymbol{v}}}_h,\partial_t\underline{\boldsymbol{v}}_h^{(\delta)}\Big) - s_h\Big(\partial_t\psi_h^{(\delta)},\partial_t\overline{\psi}_h\Big) - f_h\Big(\partial_t\lambda_h^{(\delta)},\partial_t\overline{\psi}_h\Big)\\
&- f_h\Big(\partial_t\overline{\lambda}_h,\partial_t\psi_h^{(\delta)}\Big) - g_h\Big(\partial_t\lambda_h^{(\delta)},\partial_t\overline{\lambda}_h\Big).
\end{aligned}$$

By the definition of $(\underline{\overline{\boldsymbol{v}}}_h,\overline{\psi}_h,\overline{\lambda}_h)$ and the Young inequality, we get

$$\begin{aligned}
F\big(\partial_t\overline{\psi}_h\big) = &-\Big\|\partial_t\underline{\boldsymbol{v}}_h^{(\delta)}\Big\|^2_{L^2(\Omega)^d} - \Big\|\tau^{\frac12}\Big(\partial_t\lambda_h^{(\delta)}-\partial_t\psi_h^{(\delta)}\Big)\Big\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} - \Big\|\tau^{\frac12}\partial_t\psi_h^{(\delta)}\Big\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\\
&+\Big(\partial_t\underline{\boldsymbol{v}}_h^{(\delta)},\partial_t\underline{\boldsymbol{v}}_h^{(0)}\Big)_\Omega + \Big(\tau\Big(\partial_t\psi_h^{(\delta)}-\partial_t\lambda_h^{(\delta)}\Big),\partial_t\psi_h^{(0)}-\partial_t\lambda_h^{(0)}\Big)_{(\partial\mathcal{T}_h)^{\mathcal{I}}} + \Big(\tau\partial_t\psi_h^{(\delta)},\partial_t\psi_h^{(0)}\Big)_{(\partial\mathcal{T}_h)^{\mathcal{D}}}\\
\leq\ & 3c^{-2}\mathcal{E}_h^{(1)}\Big[\psi_h^{(\delta)},\underline{\boldsymbol{v}}_h^{(\delta)},\lambda_h^{(\delta)}\Big](t) + c^{-2}\mathcal{E}_h^{(1)}\Big[\psi_h^{(0)},\underline{\boldsymbol{v}}_h^{(0)},\lambda_h^{(0)}\Big](t).
\end{aligned}$$

Moreover, for all $\tilde{t}\in(0,T)$,

$$\int_0^{\tilde{t}} F\big(\partial_t\overline{\psi}_h\big)\,\mathrm{d}t \leq 3c^{-2}\tilde{t}\Bigg(\sup_{t\in(0,\tilde{t})}\mathcal{E}_h^{(1)}\Big[\psi_h^{(\delta)},\underline{\boldsymbol{v}}_h^{(\delta)},\lambda_h^{(\delta)}\Big](t) + \sup_{t\in(0,\tilde{t})}\mathcal{E}_h^{(1)}\Big[\psi_h^{(0)},\underline{\boldsymbol{v}}_h^{(0)},\lambda_h^{(0)}\Big](t)\Bigg).$$

Thus, by the energy stability estimates in Lemma 4.3, the right-hand side is uniformly bounded with respect to both δ and h. Inserting the above estimates into (5.3) and using the Grönwall inequality yield estimate (5.2a).

Proof of estimate (5.2b). We follow the approach in Theorem 2 in Section 5.2 of [27] for establishing asymptotic behavior of wave equations in weak topologies. For simplicity of notation, we introduce the operator

$$\mathrm{I}_{t'}\,u(t) := \begin{cases}\int_t^{t'} u(s)\,\mathrm{d}s & \text{if } 0\leq t\leq t',\\ 0 & \text{if } t'\leq t\leq T.\end{cases}$$

We can then manipulate the system of equations in (5.1) to obtain

$$\boldsymbol{m}_h(\mathrm{I}_{t'}\,\underline{\overline{\boldsymbol{v}}}_h,\underline{\boldsymbol{r}}_h) + b_h\big(\mathrm{I}_{t'}\,\overline{\psi}_h,\underline{\boldsymbol{r}}_h\big) + e_h\big(\mathrm{I}_{t'}\,\overline{\lambda}_h,\underline{\boldsymbol{r}}_h\big) = 0 \qquad \forall\underline{\boldsymbol{r}}_h\in\boldsymbol{\mathcal{Q}}_h^p, \tag{5.5a}$$

$$\begin{aligned}
&m_h\Big(\partial_{tt}\overline{\psi}_h + k\partial_t\Big(\partial_t\overline{\psi}_h\partial_t\psi_h^{(\delta)} + \partial_t\overline{\psi}_h\partial_t\psi_h^{(0)}\Big),w_h\Big)\\
&\quad -c^2 b_h(w_h,\underline{\overline{\boldsymbol{v}}}_h) + c^2 s_h\big(\overline{\psi}_h,w_h\big) + c^2 f_h\big(\overline{\lambda}_h,w_h\big) = \delta F(w_h) \qquad \forall w_h\in\mathcal{S}_h^p,
\end{aligned} \tag{5.5b}$$

$$-e_h(\mu_h,\underline{\overline{\boldsymbol{v}}}_h) + f_h\big(\mu_h,\overline{\psi}_h\big) + g_h\big(\overline{\lambda}_h,\mu_h\big) = 0 \qquad \forall\mu_h\in\mathcal{M}_h^p, \tag{5.5c}$$

where, in the second equation, we have used the identity

$$\partial_t\Big(\partial_t\overline{\psi}_h\partial_t\psi_h^{(\delta)} + \partial_t\overline{\psi}_h\partial_t\psi_h^{(0)}\Big) = 2\partial_t\psi_h^{(\delta)}\partial_{tt}\psi_h^{(\delta)} - 2\partial_t\psi_h^{(0)}\partial_{tt}\psi_h^{(0)}.$$

We then choose the test functions $\underline{\boldsymbol{r}}_h = \overline{\underline{\boldsymbol{v}}}_h$, $w_h = \mathrm{I}_{t'}\,\overline{\psi}_h$, and $\mu_h = \mathrm{I}_{t'}\,\overline{\lambda}_h$. Multiplying the first and third equations in (5.5) by c^2 and summing the results, we get the identity

$$\begin{aligned} &m_h\Big(\partial_{tt}\overline{\psi}_h + k\partial_t\Big(\partial_t\overline{\psi}_h\partial_t\psi_h^{(\delta)} + \partial_t\overline{\psi}_h\partial_t\psi_h^{(0)}\Big), \mathrm{I}_{t'}\,\overline{\psi}_h\Big) \\ &\quad + c^2\big(\boldsymbol{m}_h(\mathrm{I}_{t'}\,\overline{\underline{\boldsymbol{v}}}_h, \overline{\underline{\boldsymbol{v}}}_h) + s_h\big(\overline{\psi}_h, \mathrm{I}_{t'}\,\overline{\psi}_h\big) + f_h\big(\overline{\lambda}_h, \mathrm{I}_{t'}\,\overline{\psi}_h\big) + f_h\big(\mathrm{I}_{t'}\,\overline{\lambda}_h, \overline{\psi}_h\big) + g_h\big(\overline{\lambda}_h, \mathrm{I}_{t'}\,\overline{\lambda}_h\big)\big) \\ &\quad = \delta F\big(\mathrm{I}_{t'}\,\overline{\psi}_h\big), \end{aligned}$$

which we integrate by parts in time on $(0, t')$ to obtain

$$\begin{aligned} &\int_0^{t'} m_h\Big(\Big[1 + k\Big(\partial_t\psi_h^{(\delta)} + \partial_t\psi_h^{(0)}\Big)\Big]\partial_t\overline{\psi}_h, \overline{\psi}_h\Big)\,\mathrm{d}s \\ &\qquad + \int_0^{t'} c^2\big[\boldsymbol{m}_h(\mathrm{I}_{t'}\,\overline{\underline{\boldsymbol{v}}}_h, \overline{\underline{\boldsymbol{v}}}_h) + s_h\big(\overline{\psi}_h, \mathrm{I}_{t'}\,\overline{\psi}_h\big) + 2f_h\big(\overline{\lambda}_h, \mathrm{I}_{t'}\,\overline{\psi}_h\big) + g_h\big(\overline{\lambda}_h, \mathrm{I}_{t'}\,\overline{\lambda}_h\big)\big]\,\mathrm{d}s = \delta F\big(\mathrm{I}_{t'}\,\overline{\psi}_h\big). \end{aligned} \tag{5.6}$$

For the first term on the left-hand side, we make use of the following identity:

$$\begin{aligned} m_h\Big(\Big[1 + k\Big(\partial_t\psi_h^{(\delta)} + \partial_t\psi_h^{(0)}\Big)\Big]\partial_t\overline{\psi}_h, \overline{\psi}_h\Big) &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\sqrt{1 + k\Big(\partial_t\psi_h^{(\delta)} + \partial_t\psi_h^{(0)}\Big)}\,\overline{\psi}_h\right\|^2_{L^2(\Omega)} \\ &\quad - m_h\Big(k\Big(\partial_{tt}\psi_h^{(\delta)} + \partial_{tt}\psi_h^{(0)}\Big)\overline{\psi}_h, \overline{\psi}_h\Big). \end{aligned}$$

The positivity of $1 + k(\partial_t\psi_h^{(\delta)} + \partial_t\psi_h^{(0)}) > 0$ follows from bound (4.5) in Corollary 4.2. Further, by using the Hölder inequality and bound (4.4), we obtain

$$\begin{aligned} m_h\Big(k\Big(\partial_{tt}\psi_h^{(\delta)} + \partial_{tt}\psi_h^{(0)}\Big)\overline{\psi}_h, \overline{\psi}_h\Big) &\le \Big\|\partial_{tt}\psi_h^{(\delta)} + \partial_{tt}\psi_h^{(0)}\Big\|_{L^\infty(0,T;L^\infty(\Omega))}\big\|\overline{\psi}_h\big\|^2_{L^2(\Omega)} \\ &\lesssim \big\|\overline{\psi}_h\big\|^2_{L^2(\Omega)}. \end{aligned}$$

Since $\partial_t\,\mathrm{I}_{t'}\,u(t) = -u(t)$, we can write

$$\begin{aligned} &\boldsymbol{m}_h(\mathrm{I}_{t'}\,\overline{\underline{\boldsymbol{v}}}_h, \overline{\underline{\boldsymbol{v}}}_h) + s_h\big(\overline{\psi}_h, \mathrm{I}_{t'}\,\overline{\psi}_h\big) + 2f_h\big(\overline{\lambda}_h, \mathrm{I}_{t'}\,\overline{\psi}_h\big) + g_h\big(\overline{\lambda}_h, \mathrm{I}_{t'}\,\overline{\lambda}_h\big) \\ &\qquad = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\|\mathrm{I}_{t'}\,\overline{\underline{\boldsymbol{v}}}_h\|^2_{L^2(\Omega)^d} + \Big\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\big(\overline{\lambda}_h - \overline{\psi}_h\big)\Big\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \Big\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\,\overline{\psi}_h\Big\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\Big). \end{aligned}$$

It only remains to treat the forcing term $F(\mathrm{I}_{t'}\,\overline{\psi}_h)$. To this end, we proceed similarly as in the proof of estimate (5.2a), and obtain

$$\begin{aligned} F\big(\mathrm{I}_{t'}\,\overline{\psi}_h\big) &= -\boldsymbol{m}_h\Big(\mathrm{I}_{t'}\,\overline{\underline{\boldsymbol{v}}}_h, \partial_t\underline{\boldsymbol{v}}_h^{(\delta)}\Big) - s_h\Big(\partial_t\psi_h^{(\delta)}, \mathrm{I}_{t'}\,\overline{\psi}_h\Big) \\ &\quad - f_h\Big(\partial_t\lambda_h^{(\delta)}, \mathrm{I}_{t'}\,\overline{\psi}_h\Big) - f_h\Big(\mathrm{I}_{t'}\,\overline{\lambda}_h, \partial_t\psi_h^{(\delta)}\Big) - g_h\Big(\partial_t\lambda_h^{(\delta)}, \mathrm{I}_{t'}\,\overline{\lambda}_h\Big) \\ &= -\Big(\mathrm{I}_{t'}\,\overline{\underline{\boldsymbol{v}}}_h, \partial_t\underline{\boldsymbol{v}}_h^{(\delta)}\Big)_\Omega - \Big(\tau\,\mathrm{I}_{t'}\big(\overline{\lambda}_h - \overline{\psi}_h\big), \partial_t\psi_h^{(\delta)} - \partial_t\lambda_h^{(\delta)}\Big)_{(\partial\mathcal{T}_h)^{\mathcal{I}}} - \Big(\tau\,\mathrm{I}_{t'}\,\overline{\psi}_h, \partial_t\psi_h^{(\delta)}\Big)_{(\partial\mathcal{T}_h)^{\mathcal{D}}} \\ &\lesssim \Big(\|\mathrm{I}_{t'}\,\overline{\underline{\boldsymbol{v}}}_h\|_{L^2(\Omega)^d} + \Big\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\big(\overline{\lambda}_h - \overline{\psi}_h\big)\Big\|_{L^2((\partial\mathcal{T}_h)^{o})} + \Big\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\,\overline{\psi}_h\Big\|_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\Big). \end{aligned}$$

In the last line, we have used the uniform-in-δ estimate of the high-order energy $\mathcal{E}_h^{(1)}[\psi_h^{(\delta)}, \underline{\boldsymbol{v}}_h^{(\delta)}, \lambda_h^{(\delta)}]$ from Lemma 4.3. Putting the above estimates into identity (5.6), we obtain

$$\begin{aligned}&\|\overline{\psi}_h(t')\|^2_{L^2(\Omega)} + \|\mathrm{I}_{t'}\,\underline{\overline{\boldsymbol{v}}}_h(0)\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\big(\overline{\lambda}_h - \overline{\psi}_h\big)(0)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\,\overline{\psi}_h(0)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\\ &\lesssim \int_0^{t'} \|\overline{\psi}_h\|^2_{L^2(\Omega)}\,\mathrm{d}s + \delta\int_0^{t'}\left(\|\mathrm{I}_{t'}\,\underline{\overline{\boldsymbol{v}}}_h\|_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\big(\overline{\lambda}_h - \overline{\psi}_h\big)\right\|_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\,\mathrm{I}_{t'}\,\overline{\psi}_h\right\|_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\right)\mathrm{d}s,\end{aligned} \tag{5.7}$$

where the hidden constant does not depend on δ or h.

In order to rewrite (5.7) in a suitable form so as to be able to use the Grönwall inequality, we introduce the time-reversed operator $\widetilde{\mathrm{I}}_{t'}$, which we define for an integrable function u and $t \in (0,T)$ by

$$\widetilde{\mathrm{I}}_{t'}u(t) := \mathrm{I}_{t'}\,u(t'-t).$$

By the definition of $\widetilde{\mathrm{I}}_{t'}$, bound (5.7) can be conveniently rewritten as

$$\begin{aligned}&\|\overline{\psi}_h(t')\|^2_{L^2(\Omega)} + \left\|\widetilde{\mathrm{I}}_{t'}\underline{\overline{\boldsymbol{v}}}_h(t')\right\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}\widetilde{\mathrm{I}}_{t'}(\overline{\lambda}_h - \overline{\psi}_h)(t')\right\|^2_{L^2((\partial\mathcal{T}_h)^{o})} + \left\|\tau^{\frac{1}{2}}\widetilde{\mathrm{I}}_{t'}\overline{\psi}_h(t')\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\\ &\lesssim \delta^2 + \int_0^{t'}\left(\|\overline{\psi}_h\|^2_{L^2(\Omega)} + \left\|\widetilde{\mathrm{I}}_{t'}\underline{\overline{\boldsymbol{v}}}_h\right\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}\widetilde{\mathrm{I}}_{t'}(\overline{\lambda}_h - \overline{\psi}_h)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\widetilde{\mathrm{I}}_{t'}\overline{\psi}_h\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\right)\mathrm{d}s,\end{aligned}$$

where we have additionally used the Young inequality to get δ^2 on the right-hand side. The Grönwall inequality yields then the desired result. □

6. Numerical experiments

In this section, we assess the accuracy and robustness of the proposed method. In Section 6.1, we present some details for the implementation of the fully discrete scheme obtained by combining the semidiscrete formulation (2.6) with the Newmark time-marching scheme. The h- and δ-convergence of the proposed method are illustrated in Sections 6.2 and 6.3, respectively. In Section 6.4, we present an example of the effect of the nonlinearity parameter k on the solution.

Although our theory does not provide any superconvergence result, in the numerical experiments below, we consider the following local postprocessing technique (see [11], Sect. 2.2): given the numerical approximation $(\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h)$ of the solution to (1.2) at some time $t \geq 0$, we define $\psi_h^* \in \mathcal{S}_h^{p+1}(\mathcal{T}_h)$ such that, for all $K \in \mathcal{T}_h$, it satisfies

$$\int_K \nabla\psi_h^*\cdot\nabla q_{p+1}\,\mathrm{d}\boldsymbol{x} = \int_K \underline{\boldsymbol{v}}_h\cdot\nabla q_{p+1}\,\mathrm{d}\boldsymbol{x} \qquad \forall q_{p+1}\in\mathbb{P}^{p+1}(K), \tag{6.1a}$$

$$\int_K \psi_h^*\,\mathrm{d}\boldsymbol{x} = \int_K \psi_h\,\mathrm{d}\boldsymbol{x}. \tag{6.1b}$$

For the HDG discretization in [11] of the linear wave equation, the postprocessed variable ψ_h^* was shown to superconverge with order $\mathcal{O}(h^{p+2})$ if $p > 0$. Such a superconvergence is also numerically observed in Section 6.2 below for the nonlinear Westervelt equation.

An object-oriented MATLAB implementation of the fully discrete scheme described in the next section was developed to carry out the numerical experiments in two-dimensional domains.

6.1. Fully discrete scheme

We use the predictor-corrector Newmark scheme in Section 5.4.2 of [18] as time discretization. Let Δt be a fixed time step, $tol > 0$ be a given tolerance, $s_{\max}$ be a maximum number of linear iterations, and (γ, β) be the Newmark parameters with $\gamma \in [0,1]$ and $\beta \in [0, 1/2]$. In the numerical experiments below, we use $tol = 10^{-10}$ and $s_{\max} = 100$, and it will be useful to consider an inhomogeneous forcing term $\varphi : Q_T \to \mathbb{R}$. For convenience, we use the dot notation for discrete approximations of time derivatives.

In Algorithm 1, we describe an implementation of the proposed fully discrete scheme.

Algorithm 1: NEWMARK-HDG FULLY DISCRETE SCHEME.

1 **Set** a time step $\Delta t > 0$.
2 **Set** a tolerance *tol* > 0 and a maximum number of linear iterations $s_{\max} \in \mathbb{N}$.
3 **Compute** the coefficient $\mu = c^2(\Delta t)^2\beta + \delta\gamma\Delta t$ and the number of time steps $N_T = T/\Delta t$.
4 **Compute** the Schur complement matrices

$$\mathcal{S}_\psi = S + B^T \boldsymbol{M}^{-1} B, \quad \mathcal{A}_\lambda = G + E^T \boldsymbol{M}^{-1} E, \quad \text{and} \quad \mathcal{R}_\lambda = F + B^T \boldsymbol{M}^{-1} E.$$

5 **Solve** the matrix systems[a] for the pairs $(\boldsymbol{X}, Y)$ and $(\overline{\boldsymbol{X}}, \overline{Y})$

$$\begin{cases} (M + \mu S_\psi) Y = \mu \mathcal{R}_\lambda, \\ \boldsymbol{M}\boldsymbol{X} = E - BY. \end{cases} \qquad \begin{cases} \mathcal{S}_\psi \overline{Y} = \mathcal{R}_\lambda, \\ \boldsymbol{M}\overline{\boldsymbol{X}} = E - B\overline{Y}. \end{cases} \tag{6.2}$$

6 **Compute** the auxiliary matrices

$$\mathcal{S}_{\lambda,\mu} = G + E^T \boldsymbol{X} - F^T Y \quad \text{and} \quad \bar{\mathcal{S}}_\lambda = G + E^T \overline{\boldsymbol{X}} - F^T \overline{Y}.$$

7 **Compute** the discrete initial conditions $(\Psi_h^{(0)}, \mathbf{V}_h^{(0)}, \Lambda_h^{(0)})$ and $(\dot{\Psi}_h^{(0)}, \dot{\mathbf{V}}_h^{(0)}, \dot{\Lambda}_h^{(0)})$ by solving (3.17).
8 **Solve** the linear systems[b] for $(\ddot{\Psi}_h^{(0)}, \ddot{\Lambda}_h^{(0)})$

$$N_h\left(\dot{\Psi}_h^{(0)}\right)\ddot{\Psi}_h^{(0)} = -c^2\left(\mathcal{S}_\psi \widetilde{\Psi}_h^{(0)} + \mathcal{R}_\lambda \widetilde{\Lambda}_h^{(0)}\right) \quad \text{and} \quad \mathcal{A}_\lambda \ddot{\Lambda}_h^{(0)} = -\mathcal{R}_\lambda^T \ddot{\Psi}_h^{(0)}.$$

9 **for** $n = 0$ **to** N_T **do**
10 `%%% PREDICTOR STEP %%%`
11 **Compute** the approximations

$$\begin{aligned}
\widehat{\Psi}_h^{(n+1)} &= \Psi_h^{(n)} + \Delta t \dot{\Psi}_h^{(n)} + \frac{(\Delta t)^2}{2}(1-2\beta)\ddot{\Psi}_h^{(n)}, & \widehat{\dot{\Psi}}_h^{(n+1)} &= \dot{\Psi}^{(n)} + (1-\gamma)\Delta t \ddot{\Psi}_h^{(n)}, \\
\widehat{\Lambda}_h^{(n+1)} &= \Lambda_h^{(n)} + \Delta t \dot{\Lambda}_h^{(n)} + \frac{(\Delta t)^2}{2}(1-2\beta)\ddot{\Lambda}_h^{(n)}, & \widehat{\dot{\Lambda}}_h^{(n+1)} &= \dot{\Lambda}^{(n)} + (1-\gamma)\Delta t \ddot{\Lambda}_h^{(n)}, \\
\widetilde{\widehat{\Psi}}_h^{(n+1)} &= \widehat{\Psi}_h^{(n+1)} + \frac{\delta}{c^2}\widehat{\dot{\Psi}}_h^{(n+1)}, & \widetilde{\widehat{\Lambda}}_h^{(n+1)} &= \widehat{\Lambda}_h^{(n+1)} + \frac{\delta}{c^2}\widehat{\dot{\Lambda}}_h^{(n+1)}.
\end{aligned}$$

12 **Compute** the nth step vector[c] $\mathcal{L}^n = \Phi^{n+1} - c^2\big(\mathcal{S}_\psi \widetilde{\widehat{\Psi}}_h^{(n+1)} + \mathcal{R}_\lambda \widetilde{\widehat{\Lambda}}_h^{(n+1)}\big)$.
13 `%%% CORRECTOR STEP %%%`
14 **for** $s = 1$ **to** $s_{\max}$ **do**
15 **Compute** $R^{(n+1,s)} = \big(M - N_h(\dot{\Psi}_h^{(n+1,s)})\big)\ddot{\Psi}_h^{(n+1,s)} + \mathcal{L}^n$.
16 **Solve**[d] $(M + \mu\mathcal{S}_\psi) Z^{(n+1,s)} = R^{(n+1,s)}$.
17 **Solve**[e] $\mathcal{S}_{\lambda,\mu}\ddot{\Lambda}_h^{(n+1,s+1)} = -\mathcal{R}_\lambda^T Z^{(n+1,s)}$.
18 **Solve** $(M + \mu S_\psi)\ddot{\Psi}_h^{(n+1,s+1)} = R^{(n+1,s)} - \mu\mathcal{R}_\lambda \ddot{\Lambda}_h^{(n+1,s+1)}$.
19 **Compute** $\dot{\Psi}_h^{(n+1,s+1)} = \widehat{\dot{\Psi}}_h^{(n+1)} + \gamma\Delta t \ddot{\Psi}_h^{(n+1,s+1)}$.
20 `%%% STOPPING CRITERIA %%%`
21 **if** $\|\Psi_h^{n+1,s+1} - \Psi_h^{n+1,s}\| / \|\Psi_h^{n+1,s+1}\| <$ *tol* **then**
22 **stop**
23 **end**
24 **end**
25 **end**

[a]The matrix systems in (6.2) can be solved completely in parallel due to the block-diagonal structure of the matrices M, $\boldsymbol{M}$, and S_ψ.

[b]Here, $N_h(\cdot)$ is the block-diagonal matrix described in Remark 2.1.

[c]Here, Φ^{n+1} is the vector representation of the forcing term φ at $t = t_{n+1}$.

[d]The linear systems in lines **16** and **18** can be solved in parallel due to the block-diagonal structure of the matrix $M + \mu S_\psi$.

[e]The linear system in line **17** involves the solution of a statically condensed linear system, where the unknowns are associated with degrees of freedom on $(d-1)$-dimensional mesh facets only.

6.2. h-convergence

In order to assess the accuracy in space of the proposed method, we consider the Westervelt equation in (1.1) on the domain $Q_T = (0,1)^2 \times (0,T)$, with parameters $c = 100\,\mathrm{ms}^{-1}$, $\delta = 6 \times 10^{-9}\,\mathrm{ms}^{-1}$, and $k = 0.5\,\mathrm{s}^2\,\mathrm{m}^{-2}$. We add a forcing term $\varphi : Q_T \to \mathbb{R}$ and set the initial data such that the exact solution is given by

$$\psi(x,y,t) = A\sin(\omega t)\sin(\ell x)\sin(\ell y), \tag{6.3}$$

with $A = 10^{-2}\,\mathrm{m}^2\,\mathrm{s}^{-1}$, $\omega = 3.5\pi\,\mathrm{Hz}$, and $\ell = \pi\,\mathrm{m}^{-1}$; cf. Section 6 in [29].

We consider a set of structured simplicial meshes $\{\mathcal{T}_h\}_{h>0}$ for the spatial domain Ω, which we exemplify in Figure 2(left panel in the first row). We set the parameters $(\gamma,\,\beta) = (1/2,\,1/4)$ for the Newmark scheme, which guarantee second-order accuracy in time and unconditional stability in the linear setting (see, *e.g.*, [17], Sect. 9.1.2). The time step is chosen as $\Delta t = \mathcal{O}(h^{\frac{p+2}{2}})$, so as to balance the expected convergence rates of order $\mathcal{O}(h^{p+2})$ for the postprocessed approximation ψ_h^* with the second-order accuracy of the Newmark scheme.

In Figure 2, we show in *log–log* scale the following errors at the final time $T = 1\,\mathrm{s}$:

$$\left\|\psi(\cdot,T) - \psi_h^{(N_T)}\right\|_{L^2(\Omega)}, \quad \left\|\psi(\cdot,T) - \psi_h^{*(N_T)}\right\|_{L^2(\Omega)}, \quad \left\|\underline{\boldsymbol{v}}(\cdot,T) - \underline{\boldsymbol{v}}_h^{(N_T)}\right\|_{L^2(\Omega)^2}. \tag{6.4}$$

For $p = 0,1,2$, optimal convergence rates of order $\mathcal{O}(h^{p+1})$ are obtained for the $L^2(\Omega)$-errors of ψ_h and $\underline{\boldsymbol{v}}_h$, which is in agreement with the *a priori* error estimates derived in Section 4 for the semidiscrete HDG formulation. Moreover, when $p > 0$, superconvergence of order $\mathcal{O}(h^{p+2})$ is observed for the $L^2(\Omega)$-error of the postprocessed variable ψ_h^* defined in (6.1).

6.3. δ-convergence

We now validate the convergence of the method when the sound diffusivity parameter δ tends to zero. To do so, we consider the Westervelt equation in (1.1) on the domain $Q_T = (0,1)^2 \times (0,T)$, with parameters $c = 1\,\mathrm{ms}^{-1}$ and $k = 0.3\,\mathrm{s}^2\,\mathrm{m}^{-2}$. The initial data are given by

$$\psi_0(x,y) = 10^{-2}\sin(\pi x)\sin(\pi y), \qquad \psi_1(x,y) = \sin(\pi x)\sin(\pi y), \tag{6.5}$$

the spatial mesh is taken as in Figure 2(left panel in the first row), and the parameters (γ,β) and the time step is chosen as in the previous experiment; cf. Section 2.4.2 of [14]. We consider piecewise constant ($p = 0$) and piecewise linear ($p = 1$) approximations, and $\delta = 10^{-2i}$ with $i = 1,\ldots,5$.

In Figure 3, we show in *log–log* scale the following errors computed at the final time $T = 1\,\mathrm{s}$:

$$\left\|\psi_h^{(\delta)} - \psi_h^{(0)}\right\|_{L^2(\Omega)} \quad \text{and} \quad \left\|\underline{\boldsymbol{v}}_h^{(\delta)} - \underline{\boldsymbol{v}}_h^{(0)}\right\|_{L^2(\Omega)^2}. \tag{6.6}$$

Convergence rates of order $\mathcal{O}(\delta)$ are observed for both errors. For $p = 1$, these results are in agreement with estimate (5.2b), and suggest that estimate (5.2a) may be not sharp. In fact, in Theorem 2.2 from [14], convergence rates of order $\mathcal{O}(\delta)$ were established for the standard finite element method by exploiting the relation $\underline{\boldsymbol{v}}_h = \nabla\psi_h$. Moreover, it is likely that the exact solution is more regular than assumed in Theorem 4.1, in which case one could show full convergence rates of order $\mathcal{O}(\delta)$ in (5.2a), by deriving higher-order energy stability estimates.

6.4. Steepening of a wavefront

In this experiment, we illustrate the effect of the nonlinearity parameter k on the solution. We consider the Westervelt equation in (1.1) on the domain $Q_T = (0,1)^2 \times (0,T)$, with parameters $c = 1500\,\mathrm{ms}^{-1}$, $\delta = 6 \times 10^{-9}\,\mathrm{ms}^{-1}$, and $k = -10\,\mathrm{s}^2\,\mathrm{m}^{-2}$. We consider homogeneous initial conditions and the following forcing term:

$$\varphi(x,y,t) = \frac{a}{\sqrt{\sigma}}\exp(-\alpha t)\exp\left(-\frac{(x-0.5)^2 + (y-0.5)^2}{2\sigma^2}\right), \tag{6.7}$$

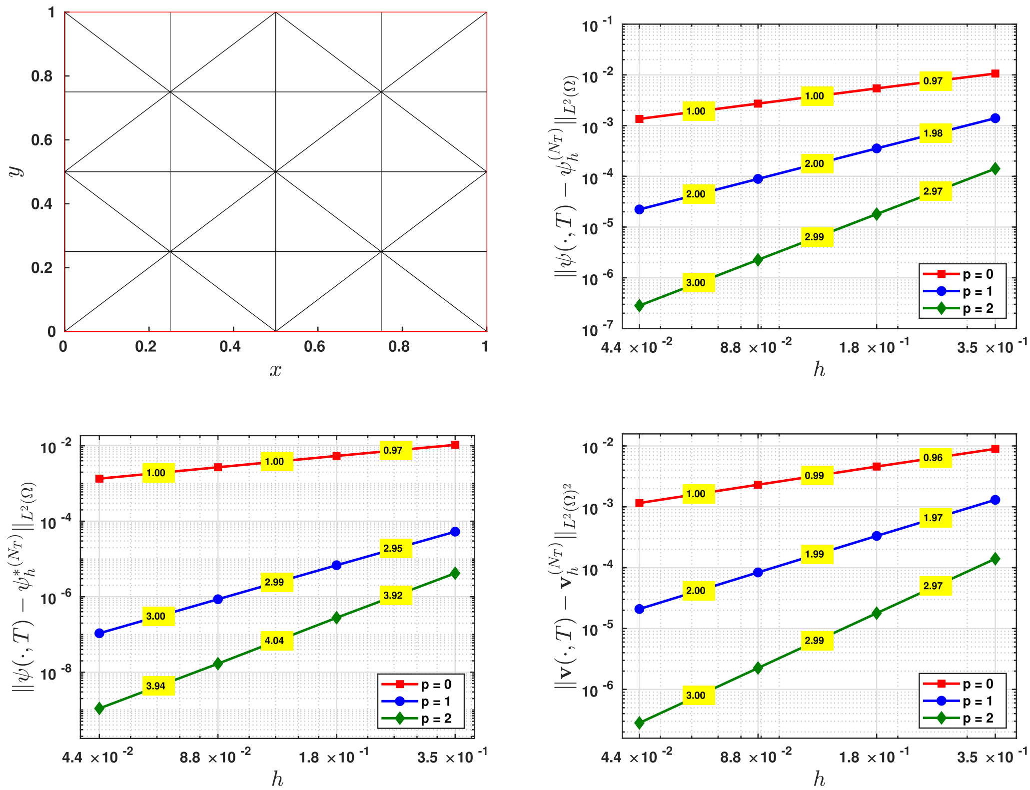

FIGURE 2. *First panel*: example of the simplicial meshes used in the numerical examples. *Remaining panels*: h-convergence of the errors in (6.4) at the final time $T = 1\,\mathrm{s}$ for the test case with exact solution (6.3). The numbers in the yellow rectangles denote the experimental rates of convergence.

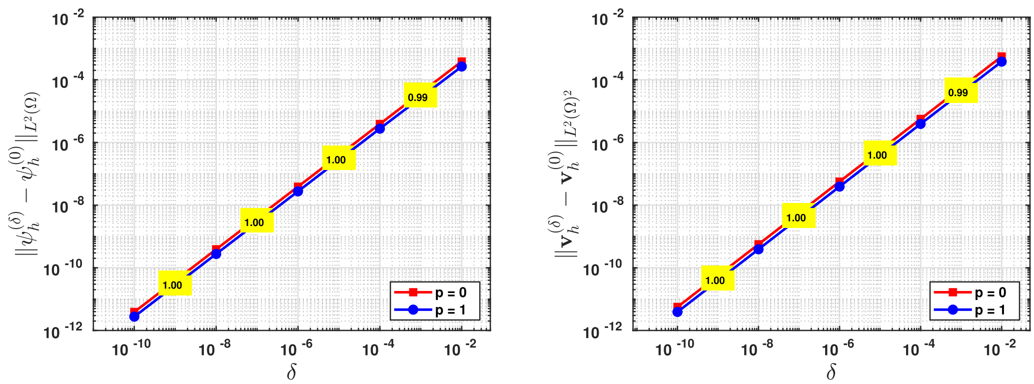

FIGURE 3. δ-convergence of the errors in (6.6) at the final time $T = 1\,\mathrm{s}$ for the test case in Section 6.3.

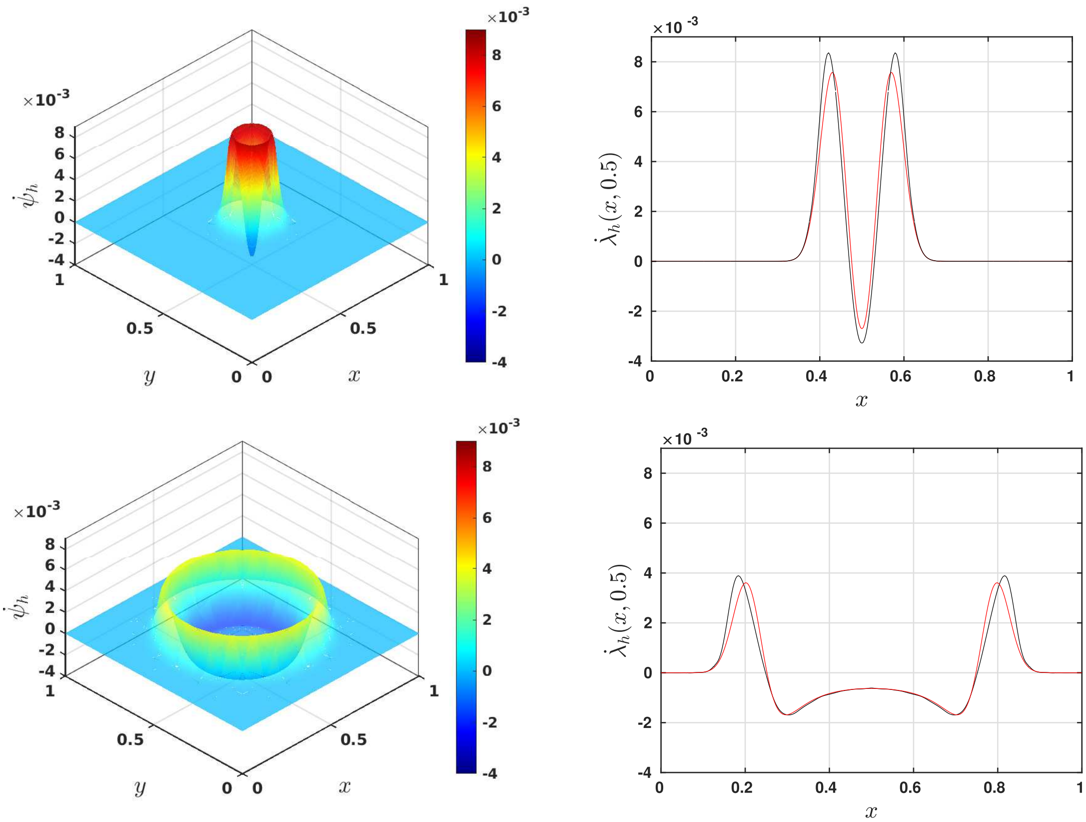

FIGURE 4. Results obtained at $t = 5 \times 10^{-5}\,$s (*first row*) and $t = 2 \times 10^{-4}\,$s (*second row*) for the test case in Section 6.4. *Left panels*: approximation of $\partial_t \psi$ obtained for $p = 5$ and $k = -10\,\mathrm{s}^2\,\mathrm{m}^{-2}$. *Right panels*: comparison of the approximations obtained for the Westervelt equation (**black** lines) and the linear damped wave equation (**red** lines) along the line $y = 0.5$.

where $a = 400$, $\alpha = 5 \times 10^4$, and $\sigma = 3 \times 10^{-2}$; cf. Section 6 in [29].

We employ a simplicial mesh $\mathcal{T}_h$ with $h \approx 8.83 \times 10^{-2}$, a fixed time step $\Delta t = 10^{-6}$, and $p = 5$. In order to deal with the steepening of the wave, the parameters for the Newmark scheme are chosen as $(\gamma, \beta) = (0.85, 0.45)$. In Figure 4(left panels), we show the approximation of $\partial_t \psi$ obtained at $t = 5 \times 10^{-5}\,$s and $t = 2 \times 10^{-4}\,$s. In Figure 4(right panels), we compare the approximation of $\partial_t \psi$ obtained for the nonlinear Westervelt equation ($k = -10$) and the damped linear wave equation ($k = 0$) along the line $y = 0.5$. A steepening at the wavefront of the solution is clearly observed for the nonlinear model.

Since the forcing term φ in (6.7) is independent of δ, the δ-convergence estimates in Theorem 5.1 are still valid. In Figure 5, we show the errors in (6.6) obtained at $t = 10^{-4}\,$s. Convergence rates of order $\mathcal{O}(\delta)$ are observed as in the numerical experiment of Section 6.3.

7. CONCLUSIONS

In this work, we have designed an asymptotic-preserving HDG method for the numerical simulation of the quasilinear Westervelt equation. We built up a well-posedness and approximation theory for this method, and illustrated our theoretical results with two-dimensional numerical experiments.

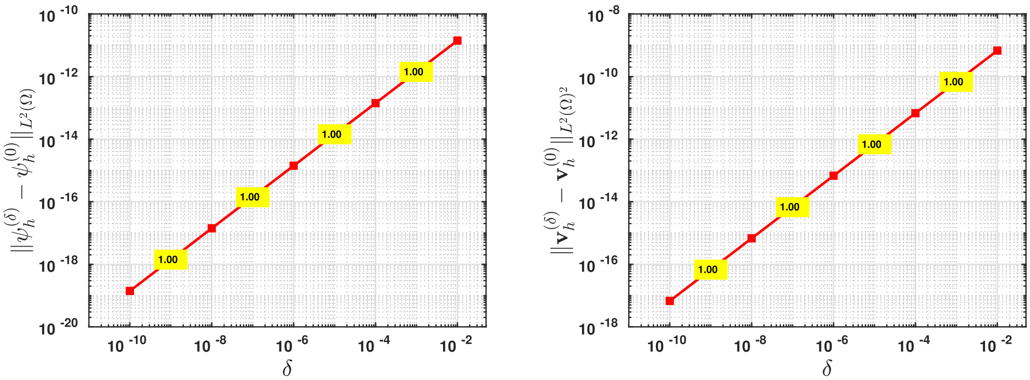

FIGURE 5. δ-convergence of the errors in (6.6) at $t = 10^{-4}$ s for the test case in Section 6.4 with degree of approximation $p = 5$.

Optimal h-convergence rates of order $\mathcal{O}(h^{p+1})$ are proven for the approximation of the acoustic particle velocity $\underline{\boldsymbol{v}} = \nabla\psi$, thus exceeding the expected convergence rates for most standard DG methods. Moreover, we have proven the convergence of the discrete approximation to the vanishing viscosity limit when the sound diffusivity parameter δ tends to zero. Such a result guarantees the robustness of the method for small values of δ.

Our analysis imposes a restriction on the degree of approximation of the method, namely $p \geq 1$. However, in the numerical experiments, we have obtained convergence of the method with respect to h and δ also for $p = 0$. This is most likely due to the fact that the numerical experiments were performed for two-dimensional domains. Indeed, the case $p = 0$ is critical for dimension $d = 2$, as commented in Remark 4.4.

The following are three interesting possible directions for the extension of our analysis:

– In view of the close relation between mixed FEM and HDG methods (see, *e.g.*, [6, 9]), we expect that the present analysis can be extended to a unified framework covering a large class of methods.
– More general polytopic meshes could be considered using the theory of M-decompositions [10], or hybrid high-order (HHO) methods [13]. In particular, the stabilization term used for HHO methods allows for a simpler analysis, in the context of polytopic meshes, that does not rely on specific HDG projections.
– The extension of the method to more general nonlinear sound propagation models such as the Kuznetsov equation [25].

In addition, the superconvergence of order $\mathcal{O}(h^{p+2})$ for the local postprocessed approximation ψ_h^* defined in (6.1), and the asymptotic-preserving properties of fully discrete schemes as in [14], with special attention to high-order time stepping schemes, are the subject of ongoing research.

Acknowledgments

We thank the reviewers for their careful reading of the manuscript and helpful remarks, which have led to marked improvements. We are greatful to Vanja Nikolić (Radboud University) for her very valuable suggestions to improve the presentation of this work. The first author acknowledges support from the Italian Ministry of University and Research through the project PRIN2020 "Advanced polyhedral discretizations of heterogeneous PDEs for multiphysics problems", and from the INdAM-GNCS through the project CUP_E53C23001670001.

References

[1] R.P. Agarwal, M. Meehan and D. O'regan, Fixed Point Theory and Applications. Vol. 141. Cambridge University Press (2001).

[2] P.-F. Antonietti, I. Mazzieri, M. Muhr, V. Nikolić and B. Wohlmuth, A high-order discontinuous Galerkin method for nonlinear sound waves. *J. Comput. Phys.* **415** (2020) 109484.

[3] S.C. Brenner and L.R. Scott, The mathematical theory of finite element methods, in Vol. 15 of Texts in Applied Mathematics, 3rd edition. Springer, New York (2008).

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

[5] B. Cockburn, B. Dong and J. Guzmán, A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. *Math. Comput.* **77** (2008) 1887–1916.

[6] B. Cockburn, J. Gopalakrishnan and R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.* **47** (2009) 1319–1365.

[7] B. Cockburn, J. Guzmán and H. Wang, Superconvergent discontinuous Galerkin methods for second-order elliptic problems. *Math. Comput.* **78** (2009) 1–24.

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

[9] B. Cockburn, W. Qiu and K. Shi, Conditions for superconvergence of HDG methods for second-order elliptic problems. *Math. Comput.* **81** (2012) 1327–1353.

[10] B. Cockburn, G. Fu and F.-J. Sayas, Superconvergence by M-decompositions. Part I: general theory for HDG methods for diffusion. *Math. Comput.* **86** (2017) 1609–1641.

[11] B. Cockburn, Z. Fu, A. Hungria, L. Ji, M.-A. Sánchez and F.-J. Sayas, Stormer–Numerov HDG methods for acoustic waves. *J. Sci. Comput.* **75** (2018) 597–624.

[12] L. Demi and M.D. Verweij, Nonlinear acoustics, in Comprehensive Biomedical Physics. Elsevier, Oxford (2014) 387–399.

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

[14] B. Dörich and V. Nikolić, Robust fully discrete error bounds for the Kuznetsov equation in the inviscid limit. Preprint: arXiv:2401.06492 (2024).

[15] 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) 472–498.

[16] M. Hochbruck and B. Maier, Error analysis for space discretizations of quasilinear wave-type equations. *IMA J. Numer. Anal.* **42** (2022) 1963–1990.

[17] T.J.R. Hughes, The Finite Element Method. Prentice Hall, Inc., Englewood Cliffs, NJ (1987).

[18] M. Kaltenbacher, Numerical Simulation of Mechatronic Sensors and Actuators. Vol. 2. Springer (2007).

[19] B. Kaltenbacher and I. Lasiecka, Global existence and exponential decay rates for the Westervelt equation. *Discrete Contin. Dyn. Syst. Ser. S* **2** (2009) 503–523.

[20] B. Kaltenbacher and V. Nikolić, Parabolic approximation of quasilinear wave equations with applications in nonlinear acoustics. *SIAM J. Math. Anal.* **54** (2022) 1593–1622.

[21] B. Kaltenbacher and G. Peichl, The shape derivative for an optimization problem in lithotripsy. *Evol. Equ. Control Theory* **5** (2016) 399–429.

[22] M. Kaltenbacher, H. Landes, J. Hoffelner and R. Simkovics, Use of modern simulation for industrial applications of high power ultrasonics, in 2002 IEEE Ultrasonics Symposium. Vol. 1. IEEE (2002) 673–678.

[23] B. Kaltenbacher, M. Meliani and V. Nikolić, Limiting behavior of quasilinear wave equations with fractional-type dissipation. *Adv. Nonlinear Stud.* **24** (2024) 748–774.

[24] S. Kawashima and Y. Shibata, Global existence and exponential stability of small solutions to nonlinear viscoelasticity. *Comm. Math. Phys.* **148** (1992) 189–208.

[25] V.P. Kuznetsov, Equations of nonlinear acoustics. *Sov. Phys. Acoustic* **16** (1970) 467–470.

[26] B. Maier, *Error analysis for space and time discretizations of quasilinear wave-type equations*. Ph.D. Thesis, Karlsruher Institut für Technologie (KIT) (2020).

[27] M. Meliani, A unified analysis framework for generalized fractional Moore–Gibson–Thompson equations: well-posedness and singular limits. *Fract. Calc. Appl. Anal.* **26** (2023) 2540–2579.

[28] M. Meliani and V. Nikolić, Analysis of general shape optimization problems in nonlinear acoustics. *Appl. Math. Optim.* **86** (2022) 39.

[29] M. Meliani and V. Nikolić, Mixed approximation of nonlinear acoustic equations: well-posedness and a priori error analysis. *Appl. Numer. Math.* **198** (2024) 94–111.

[30] S. Meyer and M. Wilke, Optimal regularity and long-time behavior of solutions for the Westervelt equation. *Appl. Math. Optim.* **64** (2011) 257–271.

[31] N.C. Nguyen, J. Peraire and B. Cockburn, Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwell's equations. *J. Comput. Phys.* **230** (2011) 7151–7175.

[32] V. Nikolić, Asymptotic-preserving finite element analysis of Westervelt-type wave equations. Preprint: arXiv:2303.10743 (2023).

[33] V. Nikolić and B. Wohlmuth, A priori error estimates for the finite element approximation of Westervelt's quasi-linear acoustic wave equation. *SIAM J. Numer. Anal.* **57** (2019) 1897–1918.

[34] M.A. Sánchez, C. Ciuca, N.C. Nguyen, J. Peraire and B. Cockburn, Symplectic Hamiltonian HDG methods for wave propagation phenomena. *J. Comput. Phys.* **350** (2017) 951–973.

[35] I. Shevchenko and B. Kaltenbacher, Absorbing boundary conditions for nonlinear acoustics: the Westervelt equation. *J. Comput. Phys.* **302** (2015) 200–221.

[36] P.J. Westervelt, Parametric acoustic array. *J. Acoust. Soc. Am.* **35** (1963) 535–537.

Appendix A. Proof of low-order energy stability estimate (3.7a)

The ideas for the proof of the stability estimates are inspired by the δ-robust analysis carried out in Section 5 from [29] for the mixed FEM approximation of the Westervelt equation.

Observe that (3.2a)–(3.2c) imply

$$c^2\boldsymbol{m}_h(\partial_t\underline{\boldsymbol{v}}_h,\underline{\boldsymbol{r}}_h)+c^2 b_h(\partial_t\psi_h,\underline{\boldsymbol{r}}_h)+c^2 e_h(\partial_t\lambda_h,\underline{\boldsymbol{r}}_h)=c^2(\partial_t\underline{\boldsymbol{\Upsilon}},\underline{\boldsymbol{r}}_h)_\Omega \qquad \forall \underline{\boldsymbol{r}}_h\in\boldsymbol{\mathcal{Q}}_h^p, \tag{A.1a}$$

$$\begin{aligned} m_h((1+2k\alpha_h)\partial_{tt}\psi_h,w_h)-c^2 b_h(w_h,\widetilde{\underline{\boldsymbol{v}}}_h)& \\ +c^2 s_h\left(\widetilde{\psi}_h,w_h\right)+c^2 f_h\left(\widetilde{\lambda}_h,w_h\right)&=(\varphi,w_h)_\Omega \qquad \forall w_h\in\mathcal{S}_h^p, \end{aligned}\tag{A.1b}$$

$$-c^2 e_h(\mu_h,\widetilde{\underline{\boldsymbol{v}}}_h)+c^2 f_h\left(\mu_h,\widetilde{\psi}_h\right)+c^2 g_h\left(\widetilde{\lambda}_h,\mu_h\right)=0 \qquad \forall\mu_h\in\mathcal{M}_h^p. \tag{A.1c}$$

Taking $\underline{\boldsymbol{r}}_h=\widetilde{\underline{\boldsymbol{v}}}_h$, $w_h=\partial_t\psi_h$, and $\mu_h=\partial_t\lambda_h$ in (A.1), and summing the results, we get

$$\begin{aligned} & m_h((1+2k\alpha_h)\partial_{tt}\psi_h,\partial_t\psi_h) \\ & \quad + c^2(\boldsymbol{m}_h(\partial_t\underline{\boldsymbol{v}}_h,\underline{\boldsymbol{v}}_h)+s_h(\psi_h,\partial_t\psi_h)+f_h(\lambda_h,\partial_t\psi_h)+f_h(\partial_t\lambda_h,\psi_h)+g_h(\lambda_h,\partial_t\lambda_h)) \\ & \quad + \delta(\boldsymbol{m}_h(\partial_t\underline{\boldsymbol{v}}_h,\partial_t\underline{\boldsymbol{v}}_h)+s_h(\partial_t\psi_h,\partial_t\psi_h)+2f_h(\partial_t\lambda_h,\partial_t\psi_h)+g_h(\partial_t\lambda_h,\partial_t\lambda_h)) \\ & \quad = c^2(\partial_t\underline{\boldsymbol{\Upsilon}},\underline{\boldsymbol{v}}_h)_\Omega+\delta(\partial_t\underline{\boldsymbol{\Upsilon}},\partial_t\underline{\boldsymbol{v}}_h)_\Omega+(\varphi,\partial_t\psi_h)_\Omega. \end{aligned}\tag{A.2}$$

Moreover, the following identities follow from the definition of the discrete bilinear forms $m_h(\cdot,\cdot)$, $s_h(\cdot,\cdot)$, $f_h(\cdot,\cdot)$, and $g_h(\cdot,\cdot)$:

$$m_h((1+2k\alpha_h)\partial_{tt}\psi_h,\partial_t\psi_h)=\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\sqrt{1+2k\alpha_h(\cdot,t)}\partial_t\psi_h\right\|^2_{L^2(\Omega)}-m_h(k\partial_t\alpha_h\partial_t\psi_h,\partial_t\psi_h), \tag{A.3a}$$

$$\boldsymbol{m}_h(\partial_t\underline{\boldsymbol{v}}_h,\underline{\boldsymbol{v}}_h)+s_h(\psi_h,\partial_t\psi_h)+f_h(\lambda_h,\partial_t\psi_h)+f_h(\partial_t\lambda_h,\psi_h)+g_h(\lambda_h,\partial_t\lambda_h)$$

$$= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\|\underline{\boldsymbol{v}}_h\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}(\lambda_h - \psi_h)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\partial_t\psi_h\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\right), \tag{A.3b}$$

$$\boldsymbol{m}_h(\partial_t\underline{\boldsymbol{v}}_h, \partial_t\underline{\boldsymbol{v}}_h) + s_h(\partial_t\psi_h, \partial_t\psi_h) + 2f_h(\partial_t\lambda_h, \partial_t\psi_h) + g_h(\partial_t\lambda_h, \partial_t\lambda_h)$$
$$= \|\partial_t\underline{\boldsymbol{v}}_h\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}(\partial_t\lambda_h - \partial_t\psi_h)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\partial_t\psi_h\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}. \tag{A.3c}$$

Substituting the identities (A.3a)–(A.3c) into (A.2), we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) + \delta\left(\|\partial_t\underline{\boldsymbol{v}}_h\|^2_{L^2(\Omega)^d} + \left\|\tau^{\frac{1}{2}}(\partial_t\lambda_h - \partial_t\psi_h)\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{I}})} + \left\|\tau^{\frac{1}{2}}\partial_t\lambda_h\right\|^2_{L^2((\partial\mathcal{T}_h)^{\mathcal{D}})}\right)$$
$$= \int_\Omega k\partial_t\alpha_h(\partial_t\psi_h)^2\,\mathrm{d}\boldsymbol{x} + c^2(\partial_t\underline{\boldsymbol{\Upsilon}}, \underline{\boldsymbol{v}}_h)_\Omega + \delta(\partial_t\underline{\boldsymbol{\Upsilon}}, \partial_t\underline{\boldsymbol{v}}_h)_\Omega + (\varphi, \partial_t\psi_h)_\Omega. \tag{A.4}$$

All the terms multiplied by δ on the left-hand side of (A.4) are nonnegative. By using the Cauchy–Schwarz and the Young inequalities, we get

$$\delta(\partial_t\underline{\boldsymbol{\Upsilon}}, \partial_t\underline{\boldsymbol{v}}_h)_\Omega \le \frac{\delta}{4}\|\partial_t\underline{\boldsymbol{\Upsilon}}\|^2_{L^2(\Omega)^d} + \delta\|\partial_t\underline{\boldsymbol{v}}_h\|^2_{L^2(\Omega)^d},$$

so the second term cancels out with the one on the left-hand side of (A.4). Integrating identity (A.4) over $(0,t)$ and using the Hölder and the Young inequalities, we deduce that

$$\begin{aligned}
\mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](t) \le{}& \mathcal{E}_h^{(0)}[\psi_h, \underline{\boldsymbol{v}}_h, \lambda_h](0) \\
&+ \left(\left\|k(1+2k\alpha_h)^{-1}\partial_t\alpha_h\right\|_{L^1(0,t;L^\infty(\Omega))} + \frac{\gamma}{2}\right)\left\|\sqrt{1+2k\alpha_h}\partial_t\psi_h\right\|^2_{L^\infty(0,t;L^2(\Omega))} \\
&+ \frac{1}{2\gamma}\left\|(1+2k\alpha_h)^{-\frac{1}{2}}\varphi\right\|^2_{L^1(0,t;L^2(\Omega))} + \frac{\delta}{4}\|\partial_t\underline{\boldsymbol{\Upsilon}}\|^2_{L^2(0,t;L^2(\Omega)^d)} \\
&+ \frac{c^2}{2\sigma_0}\|\partial_t\underline{\boldsymbol{\Upsilon}}\|^2_{L^1(0,t;L^2(\Omega)^d)} + \frac{c^2\sigma_0}{2}\|\underline{\boldsymbol{v}}_h\|^2_{L^\infty(0,t;L^2(\Omega)^d)},
\end{aligned} \tag{A.5}$$

for all $\gamma > 0$.

Moreover, by using the Hölder inequality, we have the following bounds:

$$\begin{aligned}
\left\|(1+2k\alpha_h)^{-\frac{1}{2}}\varphi\right\|^2_{L^1(0,t;L^2(\Omega))} &\le t\left\|(1+2k\alpha_h)^{-\frac{1}{2}}\varphi\right\|^2_{L^2(0,t;L^2(\Omega))}, \\
\|\partial_t\underline{\boldsymbol{\Upsilon}}\|^2_{L^1(0,t;L^2(\Omega)^d)} &\le t\|\partial_t\underline{\boldsymbol{\Upsilon}}\|^2_{L^2(0,t;L^2(\Omega)^d)}, \\
\|\underline{\boldsymbol{v}}_h\|^2_{L^\infty(0,t;L^2(\Omega)^d)} &\le \sup_{s\in(0,t)}\|\underline{\boldsymbol{v}}_h\|^2_{L^2(\Omega)^d}.
\end{aligned} \tag{A.6}$$

Finally, the smallness assumption (3.4) states that there exist constants $0 < \gamma_0 < \sigma_0 < 1$ independent of h and δ such that

$$\left\|k(1+2k\alpha_h)^{-1}\partial_t\alpha_h\right\|_{L^1(0,t;L^\infty(\Omega))} + \frac{\gamma_0}{2} \le \frac{|k|}{1-2|k|\underline{\alpha}}\|\partial_t\alpha_h\|_{L^1(0,t;L^\infty(\Omega))} + \frac{\gamma_0}{2} \le \frac{\sigma_0}{2},$$

which, together with (A.5) and (A.6), gives the low-order energy estimate in (3.7a).

Appendix B. Proof of high-order energy stability estimate (3.7b)

The proof of the high-order stability estimate in (3.7b) follows by considering the time-differentiated system

$$c^2\boldsymbol{m}_h(\partial_{tt}\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{r}}_h) + c^2 b_h(\partial_{tt}\psi_h, \underline{\boldsymbol{r}}_h) + c^2 e_h(\partial_{tt}\lambda_h, \underline{\boldsymbol{r}}_h) = (\partial_{tt}\underline{\boldsymbol{\Upsilon}}, \underline{\boldsymbol{r}}_h)_\Omega \qquad \forall \underline{\boldsymbol{r}}_h \in \boldsymbol{\mathcal{Q}}_h^p,$$

$$m_h((1+2k\alpha_h)\partial_{ttt}\psi_h, w_h) + m_h(2k\partial_t\alpha_h\partial_{tt}\psi_h, w_h) - c^2 b_h(w_h, \partial_t\widetilde{\underline{\boldsymbol{v}}}_h)$$
$$+ c^2 f_h\left(\partial_t\widetilde{\lambda}_h, w_h\right) + c^2 s_h\left(\partial_t\widetilde{\psi}_h, w_h\right) = (\partial_t\varphi, w_h)_\Omega \qquad \forall w_h \in \mathcal{S}_h^p,$$

$$-c^2 e_h(\mu_h, \partial_t\widetilde{\underline{\boldsymbol{v}}}_h) + c^2 f_h\left(\mu_h, \partial_t\widetilde{\psi}_h\right) + c^2 g_h\left(\partial_t\widetilde{\lambda}_h, \mu_h\right) = 0 \qquad \forall \mu_h \in \mathcal{M}_h^p,$$

choosing $\underline{\boldsymbol{r}}_h = \partial_t\widetilde{\underline{\boldsymbol{v}}}_h$, $w_h = \partial_{tt}\psi_h$, and $\mu_h = \partial_{tt}\lambda_h$ as test functions, and summing the resulting equations. The remaining steps are similar to those exposed in Appendix A for the low-order estimate in (3.7a), and are therefore omitted.