WEIGHT-ADJUSTED DISCONTINUOUS GALERKIN METHODS: MATRIX-VALUED WEIGHTS AND LINEAR ELASTIC WAVE PROPAGATION

JESSE CHAN

Abstract. Weight-adjusted inner products [1, 2] replace weighted L^2 inner products with easily invertible weight-adjusted approximations. When paired with discontinuous Galerkin (DG) methods, the result is a low storage, energy stable, and (for sufficiently regular weights) high order accurate method for wave propagation in the presence of general heterogeneous media and curvilinear meshes. In this work, we extend weight-adjusted DG (WADG) methods to weighted inner products with matrix-valued weights, focusing on the linear elastic wave equation as an application. We present a new DG formulation based on the discretization of the symmetric form of the linear elasticity wave equation, with upwind-like dissipation incorporated through simple penalty terms in the numerical fluxes. This results in provably stable formulations for arbitrary anisotropic stiffness tensors and curvilinear meshes. A convergence analysis is given, and numerical results confirm the stability and high order accuracy of WADG for several problems in elastic wave propagation.

1. Introduction. There exist many high order energy stable discontinuous Galerkin formulations for the elastic wave equations in first order form. Wilcox et al. introduced a discontinuous Galerkin spectral element method on hexahedral meshes

SEM [3]

DG-SEM [4]

High order DG [5, 6, 7, 8]

Interior penalty formulations. Analysis: Riviere et al. [9, 10] and Delcourte et al. [11, 12]

GPU-accelerated DG methods [13, 14, 15]. Also elastic wave propagation and seismic wave applications [16, 17]

High order approximations of material data [18, 19], but not low storage or not energy stable.

Extending mass lumping without dispersive effects [20, 21]. Reduces to exact mass matrix inversion for constant weights.

A simpler extension of WADG [1, 2] to elasticity and weight-adjusted inner products for vector-valued functions.

Outline: present elasticity and an energy stable DG method, motivate need for alternative to inverting large dense coupled mass matrix, introduce weight-adjusted inner products.

2. Symmetric form of the elastic wave equation . We begin with the linear elastic wave equation under small deformations in a domain Ω. These equations can be written a first order velocity-stress system for velocity $\boldsymbol{v}$ and symmetric stress tensor $\boldsymbol{S}$

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} = \nabla \cdot \boldsymbol{S} + \boldsymbol{f}$$
$$\frac{\partial \boldsymbol{S}}{\partial t} = \frac{1}{2}\bar{\boldsymbol{C}}\left(\nabla \boldsymbol{v} + \nabla \boldsymbol{v}^T\right),$$

where $\boldsymbol{f}$ is the body force per unit volume, ρ is density, and $\bar{\boldsymbol{C}}$ is the symmetric constitutive stiffness tensor relating stress and strain. These equations can also be rewritten as a symmetric hyperbolic system of PDEs [22]

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} = \sum_{i=1}^{d} \boldsymbol{A}_i^T \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{x}_i} + \boldsymbol{f}$$

(1)
$$\boldsymbol{C}^{-1}\frac{\partial \boldsymbol{\sigma}}{\partial t} = \sum_{i=1}^{d} \boldsymbol{A}_i \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}_i}.$$

where $\boldsymbol{C}$ is the symmetric matrix form of the constitutive tensor $\bar{\boldsymbol{C}}$ and σ is a vector of length $N_d = \frac{d(d+1)}{2}$, the number of unique entries of the stress tensor $\boldsymbol{S}$. We note that the matrices $\boldsymbol{A}_i$ are constant, while ρ and

$\boldsymbol{C}$ can vary spatially. Furthermore, we will assume that ρ and $\boldsymbol{C}$ are positive-definite and bounded such that

$$\rho_{\min} \leq \rho(\boldsymbol{x}) \leq \rho_{\max}$$
$$0 < c_{\min} \leq \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \leq c_{\max} < \infty, \qquad \boldsymbol{x} \in \mathbb{R}^{N_d},$$

which also implies that $\boldsymbol{C}^{-1}$ is positive-definite and bounded.

In two dimensions, $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2)^T$ and $\boldsymbol{\sigma} = (\sigma_{xx}, \sigma_{yy}, \sigma_{xy})^T$

$$\boldsymbol{S} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{pmatrix}.$$

The matrices $\boldsymbol{A}_i$ are given as

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{A}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In three dimensions, the velocity is $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)^T$, while $\boldsymbol{\sigma} = (\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{yz}, \sigma_{xy}, \sigma_{xz})^T$ denotes the unique entries of the stress tensor $\boldsymbol{S}$

$$\boldsymbol{S} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{pmatrix}.$$

The matrices $\boldsymbol{A}_i$ are then

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \boldsymbol{A}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{A}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

In general anisotropic media, $\boldsymbol{C}$ is symmetric and positive-definite. For two-dimensional isotropic media, $\boldsymbol{C}$ is given as

$$\boldsymbol{C} = \begin{pmatrix} 2\mu + \lambda & \lambda & 0 \\ \lambda & 2\mu + \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix},$$

where λ, μ are Lame parameters. For three-dimensional isotropic media, $\boldsymbol{C}$ is given instead by

$$\boldsymbol{C} = \begin{pmatrix} 2\mu + \lambda & \lambda & \lambda & & & \\ \lambda & 2\mu + \lambda & \lambda & & & \\ \lambda & \lambda & 2\mu + \lambda & & & \\ & & & \mu & & \\ & & & & \mu & \\ & & & & & \mu \end{pmatrix}.$$

We will consider both spatially varying isotropic and anisotropic media in this work.

3. An energy stable discontinuous Galerkin formulation for elastic wave propagation. The advantage of the symmetric first order formulation of the elastic wave equations (1) is that it is straightforward to derive an energy stable discontinuous Galerkin formulation. We assume that the domain Ω is Lipschitz and exactly triangulated by a mesh Ω_h, which consists of elements D^k. We further assume that each element D^k is the image of a reference element $\widehat{D}$ under the local elemental mapping

$$\boldsymbol{x}^k = \boldsymbol{\Phi}^k \widehat{\boldsymbol{x}},$$

where $\boldsymbol{x}^k = \{x^k, y^k, z^k\}$ denote physical coordinates on D^k and $\widehat{\boldsymbol{x}} = \{\widehat{x}, \widehat{y}, \widehat{z}\}$ denote coordinates on the reference element. We denote the determinant of the Jacobian of $\boldsymbol{\Phi}^k$ as J, and refer to it as the Jacobian for the remainder of this work.

We will approximate components over each element D^k from an approximation space $V_h\left(D^k\right)$, which we define as the composition of the mapping $\boldsymbol{\Phi}^k$ and a reference approximation space $V_h\left(\widehat{D}\right)$

$$V_h\left(D^k\right) = \boldsymbol{\Phi}^k \circ V_h\left(\widehat{D}\right).$$

For the remainder of this work, we will take $V_h\left(\widehat{D}\right) = P^N\left(\widehat{D}\right)$, where $P^N\left(\widehat{D}\right)$ is the polynomial space of total degree N on the reference simplex. In two dimensions,

$$P^N\left(\widehat{D}\right) = \{\widehat{x}^i \widehat{y}^j, \quad 0 \leq i + j \leq N\},$$

and in three dimensions,

$$P^N\left(\widehat{D}\right) = \{\widehat{x}^i \widehat{y}^j \widehat{z}^k, \quad 0 \leq i + j + k \leq N\}.$$

We denote the L^2 inner product and norm over D^k by $(\cdot, \cdot)_{D^k}$, such that

$$(\boldsymbol{f}, \boldsymbol{g})_{D^k} = \int_{D^k} \boldsymbol{f} \cdot \boldsymbol{g}\, \mathrm{d}x = \int_{\widehat{D}} \boldsymbol{f} \cdot \boldsymbol{g} J, \qquad \|\boldsymbol{f}\|_{D^k}^2 = (\boldsymbol{f}, \boldsymbol{f})_{D^k},$$

where $\boldsymbol{f}, \boldsymbol{g}$ are vector-valued functions. Global L^2 inner products and norms are using local L^2 inner products and norms

$$(\boldsymbol{f}, \boldsymbol{g})_{\Omega_h} = \sum_{D^k \in \Omega_h} \int_{D^k} (\boldsymbol{f}, \boldsymbol{g})_{D^k}, \qquad \|\boldsymbol{f}\|_{\Omega_h}^2 = \sum_{D^k \in \Omega_h} \|\boldsymbol{f}\|_{D^k}^2.$$

Let f be a face of an element D^k with neighboring element $D^{k,+}$ and unit outward normal $\boldsymbol{n}$. We define u^- and u^+ denote interior and exterior values of a discontinuous function u such that

$$u^- = u|_{f \cap D^k}, \qquad u^+ = u|_{f \cap \partial D^{k,+}}.$$

The jump and average of a scalar function $u \in V_h\left(\Omega_h\right)$ over f are then defined as

$$[\![u]\!] = u^+ - u^-, \qquad \{\!\{u\}\!\} = \frac{u^+ + u^-}{2}.$$

Jumps and averages of a vector-valued functions $\boldsymbol{u} \in \mathbb{R}^m$ and $\boldsymbol{S} \in \mathbb{R}^{m \times n}$ are then defined component-wise

$$([\![\boldsymbol{u}]\!])_i = [\![\boldsymbol{u}_i]\!], \qquad 1 \leq i \leq m, \qquad ([\![\boldsymbol{S}]\!])_{ij} = [\![\boldsymbol{S}_i]\!], \qquad 1 \leq i \leq m, \quad 1 \leq j \leq n.$$

We can now specify a DG formulation for the linear elastic wave equation (1). Symmetric hyperbolic systems readily admit a DG formulation based on penalty fluxes, as detailed in [23]. For the linear elastic wave equation in symmetric first order form, this formulation is given as

$$\sum_{D^k \in \Omega_h} \left(\rho \frac{\partial \boldsymbol{v}}{\partial t}, \boldsymbol{w}\right)_{D^k} = \sum_{D^k \in \Omega_h} \left(\left(\sum_{i=1}^d \boldsymbol{A}_i^T \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{x}_i} + \boldsymbol{f}, \boldsymbol{w}\right)_{D^k} + \left\langle \frac{1}{2} \boldsymbol{A}_n^T [\![\boldsymbol{\sigma}]\!] + \frac{\tau_v}{2} \boldsymbol{A}_n^T \boldsymbol{A}_n [\![\boldsymbol{v}]\!], \boldsymbol{w}\right\rangle_{\partial D^k}\right)$$

$$(2) \quad \sum_{D^k \in \Omega_h} \left(\boldsymbol{C}^{-1} \frac{\partial \boldsymbol{\sigma}}{\partial t}, \boldsymbol{q}\right)_{D^k} = \sum_{D^k \in \Omega_h} \left(\left(\sum_{i=1}^d \boldsymbol{A}_i \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}_i}, \boldsymbol{q}\right)_{D^k} + \left\langle \frac{1}{2} \boldsymbol{A}_n [\![\boldsymbol{v}]\!] + \frac{\tau_\sigma}{2} \boldsymbol{A}_n \boldsymbol{A}_n^T [\![\boldsymbol{\sigma}]\!], \boldsymbol{q}\right\rangle_{\partial D^k}\right),$$

for all $\boldsymbol{w}, \boldsymbol{q} \in V_h\left(\Omega_h\right)$. Here, $\tau_v, \tau_\sigma \geq 0$ are penalty parameters and $\boldsymbol{A}_n$ is the normal matrix defined on a face f as $\boldsymbol{A}_n = \sum_{i=1}^d \boldsymbol{n}_i \boldsymbol{A}_i$. In two dimensions, $\boldsymbol{A}_n$ is

$$\boldsymbol{A}_n = \begin{pmatrix} n_x & 0 \\ 0 & n_y \\ n_y & n_x \end{pmatrix}.$$

while in three dimensions, $\boldsymbol{A}_n$ is

$$\boldsymbol{A}_n = \begin{pmatrix} n_x & 0 & 0 \\ 0 & n_y & 0 \\ 0 & 0 & n_z \\ 0 & n_z & n_y \\ n_z & 0 & n_x \\ n_y & n_z & 0 \end{pmatrix}.$$

3.1. Boundary conditions. In this work, we assume boundary conditions on velocity and traction of the form

$$\boldsymbol{v} = \boldsymbol{v}_{\mathrm{bc}}, \qquad \boldsymbol{S}\boldsymbol{n} = \boldsymbol{t}_{\mathrm{bc}}$$

where $\boldsymbol{v}_{\mathrm{bc}}$ and $\boldsymbol{t}_{\mathrm{bc}}$ are given values. Traction boundary conditions where $\boldsymbol{t}_{\mathrm{bc}} = 0$ are referred to as free-surface boundary conditions. We follow [24, 4] and impose boundary conditions on the DG formulation through exterior values and jumps of the solution. Boundary conditions on the normal component of the stress can be imposed by noting that the numerical flux contains the term $[\![\boldsymbol{A}_n^T \boldsymbol{\sigma}]\!] = \boldsymbol{S}\boldsymbol{n}$.

For a face which lies on a boundary, velocity boundary conditions are imposed by setting

$$[\![\boldsymbol{v}]\!] = 2\left(\boldsymbol{v}_{\mathrm{bc}} - \boldsymbol{v}^-\right), \qquad [\![\boldsymbol{A}_n^T \boldsymbol{\sigma}]\!] = [\![\boldsymbol{S}\boldsymbol{n}]\!] = 0,$$

while traction boundary conditions are imposed by

$$[\![\boldsymbol{A}_n^T \boldsymbol{\sigma}]\!] = [\![\boldsymbol{S}\boldsymbol{n}]\!] = 2\left(\boldsymbol{t}_{\mathrm{bc}} - \boldsymbol{S}^-\boldsymbol{n}\right) = 2\left(\boldsymbol{t}_{\mathrm{bc}} - \boldsymbol{A}_n^T \boldsymbol{\sigma}^-\right), \qquad [\![\boldsymbol{v}]\!] = 0.$$

For problems which involve the truncation of infinite or large domains, absorbing boundary conditions are required. For such cases, we impose simple non-reflective boundary conditions [24] through jumps

$$[\![\boldsymbol{A}_n^T \boldsymbol{\sigma}]\!] = [\![\boldsymbol{S}\boldsymbol{n}]\!] = -\boldsymbol{S}^-\boldsymbol{n} = -\boldsymbol{A}_n^T \boldsymbol{\sigma}^-, \qquad [\![\boldsymbol{v}]\!] = -\boldsymbol{v}^-.$$

We note that more accurate absorbing conditions can be imposed using, for example, perfectly matched layers [25] or high order absorbing boundary conditions [26, 27].

In all cases, boundary conditions are imposed by computing numerical fluxes using these modified jumps. This imposition guarantees energy stability for all three sets of boundary conditions.

3.2. Energy stability. It is straightforward to show that the formulation (2) is energy stable for zero velocity and traction boundary conditions, as well as non-reflective boundary conditions. Integrating by parts the velocity equations gives

$$\left(\rho \frac{\partial \boldsymbol{v}}{\partial t}, \boldsymbol{w}\right)_{D^k} = -\left(\sum_{i=1}^{d} \boldsymbol{\sigma}, \boldsymbol{A}_i \frac{\partial \boldsymbol{w}}{\partial \boldsymbol{x}_i}\right)_{D^k} + \left\langle \boldsymbol{A}_n^T \{\!\{\boldsymbol{\sigma}\}\!\} + \frac{\tau_v}{2} \boldsymbol{A}_n^T \boldsymbol{A}_n [\![\boldsymbol{v}]\!], \boldsymbol{w}\right\rangle_{\partial D^k}$$

$$\left(\boldsymbol{C}^{-1} \frac{\partial \boldsymbol{\sigma}}{\partial t}, \boldsymbol{q}\right)_{D^k} = \left(\sum_{i=1}^{d} \boldsymbol{A}_i \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}_i}, \boldsymbol{q}\right)_{D^k} + \left\langle \frac{1}{2} \boldsymbol{A}_n [\![\boldsymbol{v}]\!] + \frac{\tau_\sigma}{2} \boldsymbol{A}_n \boldsymbol{A}_n^T [\![\boldsymbol{\sigma}]\!], \boldsymbol{q}\right\rangle_{\partial D^k}$$

Taking $(\boldsymbol{w}, \boldsymbol{q}) = (\boldsymbol{v}, \boldsymbol{\sigma})$ and adding both equations together yields

$$\sum_{D^k \in \Omega_h} \frac{1}{2} \frac{\partial}{\partial t} \left((\rho \boldsymbol{v}, \boldsymbol{v})_{D^k} + (\boldsymbol{C}^{-1} \boldsymbol{\sigma}, \boldsymbol{\sigma})_{D^k}\right) =$$

$$\sum_{D^k \in \Omega_h} \left\langle \boldsymbol{A}_n^T \{\!\{\boldsymbol{\sigma}\}\!\} + \frac{\tau_v}{2} \boldsymbol{A}_n^T \boldsymbol{A}_n [\![\boldsymbol{v}]\!], \boldsymbol{v}\right\rangle_{\partial D^k} + \left\langle \frac{1}{2} \boldsymbol{A}_n [\![\boldsymbol{v}]\!] + \frac{\tau_\sigma}{2} \boldsymbol{A}_n \boldsymbol{A}_n^T [\![\boldsymbol{\sigma}]\!], \boldsymbol{\sigma}\right\rangle_{\partial D^k}$$

$$\sum_{D^k \in \Omega_h} \sum_{f \in \partial D^k} \int_f \boldsymbol{A}_n^T \{\!\{\boldsymbol{\sigma}\}\!\} \boldsymbol{v} + \frac{\tau_v}{2} \boldsymbol{A}_n^T \boldsymbol{A}_n [\![\boldsymbol{v}]\!] \boldsymbol{v} + \frac{1}{2} \boldsymbol{A}_n [\![\boldsymbol{v}]\!] \boldsymbol{\sigma} + \frac{\tau_\sigma}{2} \boldsymbol{A}_n \boldsymbol{A}_n^T [\![\boldsymbol{\sigma}]\!] \boldsymbol{\sigma}.$$

Let Γ_h denote the set of unique faces in Ω_h, and let $\Gamma_v, \Gamma_\sigma, \Gamma_{\mathrm{abc}}$ denote the parts of the boundary where velocity, traction, and non-reflective boundary conditions are imposed, respectively. We separate surface

4

terms into contributions from interior shared faces and from boundary faces. On an interior shared face, we sum contributions from the two adjacent elements to yield

$$\sum_{f\in\Gamma_h\setminus\partial\Omega}\int_f \boldsymbol{A}_n^T\left\{\!\left\{\boldsymbol{\sigma}\right\}\!\right\}\boldsymbol{v}+\frac{\tau_v}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{v}+\frac{1}{2}\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{\sigma}+\frac{\tau_\sigma}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]\boldsymbol{\sigma}$$

$$\leq -\sum_{f\in\Gamma_h\setminus\partial\Omega}\int_f \frac{\tau_v}{2}\left|\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\right|^2+\frac{\tau_\sigma}{2}\left|\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]\right|^2.$$

For faces which lie on the boundary Γ_v where velocity boundary conditions are imposed, $[\![\boldsymbol{v}]\!]=-2\boldsymbol{v}^-$, $[\![\boldsymbol{A}_n^T\boldsymbol{\sigma}]\!]=0$, and $\boldsymbol{A}_n^T\left\{\!\left\{\boldsymbol{\sigma}\right\}\!\right\}=\boldsymbol{S}\boldsymbol{n}^-$, implying that

$$\sum_{f\in\Gamma_v}\int_f \boldsymbol{A}_n^T\left\{\!\left\{\boldsymbol{\sigma}\right\}\!\right\}\boldsymbol{v}+\frac{\tau_v}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{v}+\frac{1}{2}\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{\sigma}+\frac{\tau_\sigma}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]\boldsymbol{\sigma}$$

$$= \sum_{f\in\Gamma_v}\int_f \boldsymbol{A}_n^T\boldsymbol{\sigma}^-\boldsymbol{v}^- - \boldsymbol{A}_n\boldsymbol{v}^-\boldsymbol{\sigma}^- - \tau_v\boldsymbol{A}_n^T\boldsymbol{A}_n\left|\boldsymbol{v}^-\right|^2 = -\sum_{f\in\Gamma_v}\int_f \tau_v\boldsymbol{A}_n^T\boldsymbol{A}_n\left|\boldsymbol{v}^-\right|^2.$$

For faces which lie on Γ_σ, $\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]=-2\boldsymbol{A}_n^T\boldsymbol{\sigma}^-$, $\boldsymbol{A}_n^T\left\{\!\left\{\boldsymbol{\sigma}\right\}\!\right\}=0$, and $[\![\boldsymbol{v}]\!]=0$, yielding a similar contribution

$$\sum_{f\in\Gamma_\sigma}\int_f \boldsymbol{A}_n^T\left\{\!\left\{\boldsymbol{\sigma}\right\}\!\right\}\boldsymbol{v}+\frac{\tau_v}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{v}+\frac{1}{2}\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{\sigma}+\frac{\tau_\sigma}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]\boldsymbol{\sigma} = -\sum_{f\in\Gamma_\sigma}\int_f \tau_\sigma\boldsymbol{A}_n\boldsymbol{A}_n^T\left|\boldsymbol{\sigma}^-\right|^2.$$

Finally, for faces in Γ_{abc} we have $\boldsymbol{A}_n^T\left\{\!\left\{\boldsymbol{\sigma}\right\}\!\right\}=\frac{1}{2}\boldsymbol{A}_n^T\boldsymbol{\sigma}^-$, $\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]=-\boldsymbol{A}_n^T\boldsymbol{\sigma}^-$, and $[\![\boldsymbol{v}]\!]=-\boldsymbol{v}^-$, yielding

$$\sum_{f\in\Gamma_{\mathrm{abc}}}\int_f \boldsymbol{A}_n^T\left\{\!\left\{\boldsymbol{\sigma}\right\}\!\right\}\boldsymbol{v}+\frac{\tau_v}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{v}+\frac{1}{2}\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\boldsymbol{\sigma}+\frac{\tau_\sigma}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]\boldsymbol{\sigma} =$$

$$-\sum_{f\in\Gamma_{\mathrm{abc}}}\int_f \frac{\tau_v}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n\left|\boldsymbol{v}^-\right|^2+\frac{\tau_\sigma}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T\left|\boldsymbol{\sigma}^-\right|^2.$$

Combining all face contributions together gives the following result:

THEOREM 1. *The DG formulation* (2) *is energy stable for* $\tau_v,\tau_\sigma\geq 0$, *in the sense that*

$$\sum_{D^k\in\Omega_h}\frac{1}{2}\frac{\partial}{\partial t}\left(\left(\rho\boldsymbol{v},\boldsymbol{v}\right)_{D^k}+\left(\boldsymbol{C}^{-1}\boldsymbol{\sigma},\boldsymbol{\sigma}\right)_{D^k}\right)\leq -\sum_{f\in\Gamma_h\setminus\partial\Omega}\int_f \frac{\tau_v}{2}\left|\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\right|^2+\frac{\tau_\sigma}{2}\left|\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]\right|^2$$

(3) $$-\sum_{f\in\Gamma_v}\int_f \tau_v\left|\boldsymbol{A}_n\boldsymbol{v}^-\right|^2-\sum_{f\in\Gamma_\sigma}\int_f \tau_\sigma\left|\boldsymbol{A}_n^T\boldsymbol{\sigma}^-\right|^2-\sum_{f\in\Gamma_{\mathrm{abc}}}\int_f \frac{\tau_v}{2}\left|\boldsymbol{A}_n\boldsymbol{v}^-\right|^2+\frac{\tau_\sigma}{2}\left|\boldsymbol{A}_n^T\boldsymbol{\sigma}^-\right|^2\leq 0.$$

Since ρ and $\boldsymbol{C}^{-1}$ are positive definite, the left hand side of (3) is an L^2-equivalent norm on $(\boldsymbol{u},\boldsymbol{\sigma})$, and Theorem 3 implies that the magnitude of the DG solution $(\boldsymbol{u},\boldsymbol{\sigma})$ is non-increasing in time. This also shows that dissipation present for positive penalization constants τ_v,τ_σ acts on non-conforming components with non-zero jumps $\boldsymbol{A}_n^T[\![\boldsymbol{v}]\!]$ and $\boldsymbol{A}_n[\![\boldsymbol{\sigma}]\!]$. In fact, it was shown in [23] that, in the limit as $\tau_v,\tau_\sigma\to\infty$, the eigenspaces of DG discretizations split into a conforming part consisting of $\boldsymbol{u},\boldsymbol{\sigma}$ which satisfy

$$\boldsymbol{A}_n[\![\boldsymbol{u}]\!]=0,\qquad \boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]=0$$

and an non-conforming part (defined through the L^2 orthogonal complement) corresponding to eigenvalues contain real parts which approach $-\infty$. For the linear elastic wave equations, these conditions are equivalent to the requirement of C^0 continuity for $\boldsymbol{u}$ and normal continuity of the stress tensor $[\![\boldsymbol{S}]\!]\boldsymbol{n}=0$.

3.3. The semi-discrete matrix system for DG. The solution to (2) can be approximated by discretizing in space and using an explicit time integrator, which requires only evaluations of local contributions

over D^k to the DG formulation

$$\left(\rho\frac{\partial \boldsymbol{v}}{\partial t}, \boldsymbol{w}\right)_{D^k} = \left(\sum_{i=1}^{d} \boldsymbol{A}_i^T \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{x}_i}, \boldsymbol{w}\right)_{D^k} + \left\langle \frac{1}{2}\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!] + \frac{\tau_v}{2}\boldsymbol{A}_n^T \boldsymbol{A}_n[\![\boldsymbol{v}]\!], \boldsymbol{w}\right\rangle_{\partial D^k}$$

(4)
$$\left(\boldsymbol{C}^{-1}\frac{\partial \boldsymbol{\sigma}}{\partial t}, \boldsymbol{q}\right)_{D^k} = \left(\sum_{i=1}^{d} \boldsymbol{A}_i \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}_i}, \boldsymbol{q}\right)_{D^k} + \left\langle \frac{1}{2}\boldsymbol{A}_n[\![\boldsymbol{v}]\!] + \frac{\tau_\sigma}{2}\boldsymbol{A}_n \boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!], \boldsymbol{q}\right\rangle_{\partial D^k}.$$

Let $\{\phi_i\}_{i=1}^{N_p}$ be a basis for $P^N\left(\widehat{D}\right)$. We define the reference mass matrix $\widehat{\boldsymbol{M}}$ and the physical mass matrix $\boldsymbol{M}$ for an element D^k as

$$\left(\widehat{\boldsymbol{M}}\right)_{ij} = \int_{\widehat{D}} \phi_j \phi_i \, \mathrm{d}\widehat{\boldsymbol{x}}, \qquad (\boldsymbol{M})_{ij} = \int_{D^k} \phi_j \phi_i \, \mathrm{d}\boldsymbol{x} = \int_{\widehat{D}} \phi_j \phi_i J \, \mathrm{d}\widehat{\boldsymbol{x}}.$$

For affine mappings, J is constant and $\boldsymbol{M} = J\widehat{\boldsymbol{M}}$. We also define weak differentiation matrices $\boldsymbol{S}_k$ and face mass matrix $\boldsymbol{M}_f$ such that

$$(\boldsymbol{S}_k)_{ij} = \int_{D^k} \frac{\partial \phi_j}{\partial \boldsymbol{x}_k} \phi_i \, \mathrm{d}\boldsymbol{x}, \qquad (\boldsymbol{M}_f)_{ij} = \int_f \phi_j \phi_i \, \mathrm{d}\boldsymbol{x} = \int_{\widehat{f}} \phi_j \phi_i J^f \, \mathrm{d}\widehat{\boldsymbol{x}},$$

where J^f is the Jacobian of the mapping from a reference face $\widehat{f}$ to f. For affinely mapped simplices, J^f is also constant and $\boldsymbol{M}_f = J^f \widehat{\boldsymbol{M}}_f$, where the definition of the reference face mass matrix $\widehat{\boldsymbol{M}}_f$ is analogous to the definition of the reference mass matrix $\widehat{\boldsymbol{M}}$.

Finally, we define weighted mass matrices. Let $w(\boldsymbol{x}) \in \mathbb{R}$ and $\boldsymbol{W}(\boldsymbol{x}) \in \mathbb{R}^{m \times n}$. Then, scalar and matrix-weighted mass matrices $\boldsymbol{M}_w$ and $\boldsymbol{M}_{\boldsymbol{W}}$ are defined through

$$(\boldsymbol{M}_w)_{ij} = \int_{D^k} w(\boldsymbol{x})\phi_j(\boldsymbol{x})\phi_i(\boldsymbol{x}) \, \mathrm{d}x, \qquad \boldsymbol{M}_{\boldsymbol{W}} = \begin{pmatrix} \boldsymbol{M}_{\boldsymbol{W}_{1,1}} & \cdots & \boldsymbol{M}_{\boldsymbol{W}_{1,n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{M}_{\boldsymbol{W}_{m,1}} & \cdots & \boldsymbol{M}_{\boldsymbol{W}_{m,n}} \end{pmatrix},$$

where $\boldsymbol{M}_{\boldsymbol{W}_{i,j}}$ is the scalar weighted mass matrix weighted by the (i,j) entry of $\boldsymbol{W}$.

Local contributions to the DG variational form may then be evaluated in a quadrature-free manner using these matrices. Let $\boldsymbol{\Sigma}_i, \boldsymbol{V}_i$ denote vectors containing degrees of freedom for solution components $\boldsymbol{\sigma}_i, \boldsymbol{v}_i$ such that

$$\boldsymbol{\sigma}_i(\boldsymbol{x}, t) = \sum_{j=1}^{N_p} \boldsymbol{\Sigma}_i(t)\phi_j(\boldsymbol{x}), \qquad 1 \leq i \leq N_d$$

$$\boldsymbol{v}_i(\boldsymbol{x}, t) = \sum_{j=1}^{N_p} \boldsymbol{V}_i(t)\phi_j(\boldsymbol{x}), \qquad 1 \leq i \leq d.$$

Then, the local DG formulation can be written as a block system of ordinary differential equations (ODEs) by concatenating $\boldsymbol{\Sigma}_i, \boldsymbol{V}_i$ into single vectors $\boldsymbol{\Sigma}, \boldsymbol{V}$ and using the Kronecker product $\otimes$

$$\boldsymbol{M}_{\rho \boldsymbol{I}} \frac{\partial \boldsymbol{V}}{\partial t} = \sum_{i=1}^{d} \left(\boldsymbol{A}_i^T \otimes \boldsymbol{S}_i\right)\boldsymbol{\Sigma} + \sum_{f \in \partial D^k} \left(\boldsymbol{I} \otimes \boldsymbol{M}_f\right)\boldsymbol{F}_v$$

(5)
$$\boldsymbol{M}_{\boldsymbol{C}^{-1}} \frac{\partial \boldsymbol{\Sigma}}{\partial t} = \sum_{i=1}^{d} \left(\boldsymbol{A}_i \otimes \boldsymbol{S}_i\right)\boldsymbol{V} + \sum_{f \in \partial D^k} \left(\boldsymbol{I} \otimes \boldsymbol{M}_f\right)\boldsymbol{F}_\sigma.$$

where $\boldsymbol{M}_f \boldsymbol{F}_v$ and $\boldsymbol{M}_f \boldsymbol{F}_\sigma$ represent contributions to the velocity and stress equations from surface integrals.

In order to apply standard time integration methods, we must invert $\boldsymbol{M}_{\rho \boldsymbol{I}}$ and $\boldsymbol{M}_{\boldsymbol{C}^{-1}}$ to isolate $\frac{\partial \boldsymbol{v}}{\partial t}$ and $\frac{\partial \boldsymbol{\sigma}}{\partial t}$ on the left hand side. While the inversion of $\boldsymbol{M}_{\rho \boldsymbol{I}}$ and $\boldsymbol{M}_{\boldsymbol{C}^{-1}}$ can be parallelized from element to

element, doing so typically requires either the precomputation and storage of large dense matrix inverses or the on-the-fly construction and solution of a large dense matrix at every timestep. The former option requires a large amount of storage, while the latter option is computationally expensive and difficult to parallelize. This cost can be overcome for ρ, C which are constant over an element D^k, in which case $\boldsymbol{M}_{\rho \boldsymbol{I}}$ is block diagonal with identical blocks $\boldsymbol{M}_\rho = \rho \boldsymbol{M}$, while $\boldsymbol{M}_{\boldsymbol{C}^{-1}}$ reduces to

$$
\boldsymbol{M}_{\boldsymbol{C}^{-1}} = \begin{pmatrix} \boldsymbol{C}_{1,1}^{-1} \boldsymbol{M} & \cdots & \boldsymbol{C}_{1,N_d}^{-1} \boldsymbol{M} \\ \vdots & \ddots & \vdots \\ \boldsymbol{C}_{N_d,1}^{-1} \boldsymbol{M} & \cdots & \boldsymbol{C}_{N_d,N_d}^{-1} \boldsymbol{M} \end{pmatrix} = \left(\boldsymbol{C}^{-1} \otimes \boldsymbol{M} \right).
$$

Then, $\boldsymbol{M}_\rho^{-1} = \frac{1}{\rho} \boldsymbol{M}^{-1} = \frac{1}{J\rho} \widehat{\boldsymbol{M}}^{-1}$, and $\boldsymbol{M}_{\boldsymbol{C}^{-1}}^{-1} = \boldsymbol{C} \otimes \boldsymbol{M}^{-1} = \boldsymbol{C} \otimes \left(\frac{1}{J} \widehat{\boldsymbol{M}}^{-1} \right)$, and each matrix inverse can be applied using the inverse of the reference mass matrix $\widehat{\boldsymbol{M}}^{-1}$ and the values of ρ, C, and J on each element. Applying this observation to 5 then yields the following system of ODEs

$$
\frac{\partial \boldsymbol{V}}{\partial t} = \sum_{i=1}^{d} \left(\frac{1}{\rho} \boldsymbol{A}_i^T \otimes \boldsymbol{D}_i \right) \boldsymbol{\Sigma} + \sum_{f \in \partial D^k} \left(\frac{1}{\rho} \boldsymbol{I} \otimes \boldsymbol{L}_f \right) \boldsymbol{F}_v
$$

$$
\frac{\partial \boldsymbol{\Sigma}}{\partial t} = \sum_{i=1}^{d} \left(\boldsymbol{C} \boldsymbol{A}_i \otimes \boldsymbol{D}_i \right) \boldsymbol{V} + \sum_{f \in \partial D^k} \left(\boldsymbol{C} \otimes \boldsymbol{L}_f \right) \boldsymbol{F}_\sigma,
$$

where we have introduced the differentiation matrix $\boldsymbol{D}_i = \boldsymbol{M}^{-1} \boldsymbol{S}_i$ and lift matrix $\boldsymbol{L}_f = \boldsymbol{M}^{-1} \boldsymbol{M}_f$. For affine elements, both derivative and lift matrices may be applied using only geometric factors and reference derivative and lift matrices.

Unfortunately, if ρ and C vary spatially within an element, the above approach can no longer be used to invert $\boldsymbol{M}_\rho$ and $\boldsymbol{M}_{\boldsymbol{C}^{-1}}$ in an efficient and low-storage manner. We address these issues by approximating the weighted mass matrix and corresponding weighted L^2 inner product with a weight-adjusted approximation which is low storage and simple to invert, energy stable, and provably high order accurate for spatially varying weights ρ, C with sufficiently regularity.

4. Weight-adjusted inner products for matrix-valued weights. Weight-adjusted inner products are high order accurate approximations of weighted L^2 inner products. These can be interpreted as generalizations of mass lumping techniques, reducing to mass lumping when integrals are evaluated with appropriate quadrature rules. These weight-adjusted inner products result in weight-adjusted mass matrices, whose inverses approximate the inverse of a weighted L^2 mass matrix.

We wish to apply weight-adjusted approximations to avoid the inversion of $\boldsymbol{M}_\rho$ and $\boldsymbol{M}_{\boldsymbol{C}^{-1}}$. Approximating the inverse of $\boldsymbol{M}_\rho$ can be done using weight-adjusted approximations for scalar weights [1, 2], which we review in Section 4.1. We then extend scalar weight-adjusted approximations to matrix-valued weights in Section 4.2 to approximate the inverse of $\boldsymbol{M}_{\boldsymbol{C}^{-1}}$.

4.1. Scalar weight adjusted inner products. We introduce standard Lebesgue L^p norms and their associated L^p spaces over a general domain Ω

$$
\|u\|_{L^p(\Omega)} = \left(\int_\Omega u^p \right)^{1/p}, \qquad L^p(\Omega) = \left\{ u : \Omega \to \mathbb{R}, \quad \|u\|_{L^p(\Omega)} < \infty \right\}
$$

for $1 \le p < \infty$. For $p = \infty$, these are defined as

$$
\|u\|_{L^\infty(\Omega)} = \inf \left\{ C \ge 0 : |u(\boldsymbol{x})| \le C \quad \forall \boldsymbol{x} \in \Omega \right\}, \qquad L^\infty(\Omega) = \left\{ u : \Omega \to \mathbb{R}, \quad \|u\|_{L^\infty(\Omega)} < \infty \right\}.
$$

These induce L^p Sobolev seminorms and norms of degree s

$$|u|_{W^{s,p}(\Omega)} = \left(\sum_{|\alpha|=s} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{1/p}, \qquad |u|_{W^{s,\infty}(\Omega)} = \max_{|\alpha|=s} \|D^\alpha u\|_{L^\infty(\Omega)}$$

$$\|u\|_{W^{s,p}(\Omega)} = \left(\sum_{|\alpha|\leq s} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{1/p}, \qquad \|u\|_{W^{s,\infty}(\Omega)} = \max_{|\alpha|\leq s} \|D^\alpha u\|_{L^\infty(\Omega)}.$$

where $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$ is a multi-index such that

$$D^\alpha u = \frac{\partial^{\alpha_1}}{\partial x^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial y^{\alpha_2}} \frac{\partial^{\alpha_3}}{\partial z^{\alpha_3}} u,$$

We define two operators $T_w : L^2\left(D^k\right) \to P^N\left(D^k\right)$ and $T_w^{-1} : L^2\left(D^k\right) \to P^N\left(D^k\right)$ such that

$$T_w u = \Pi_N(wu), \qquad \left(wT_w^{-1}u, v\right)_{D^k} = (u,v)_{D^k}.$$

A weighted L^2 inner product $(wu, v)_{D^k}$ can be approximated by a weight-adjusted inner product

$$(wu, v)_{D^k} = (T_w u, v)_{D^k} \approx \left(T_{1/w}^{-1}u, v\right)_{D^k}.$$

based on the observation that $T_{1/w}^{-1}u \approx uw$. This approximation is made precise by the following estimates for errors in approximations of the product uw and weighted moments:

THEOREM 2 (Theorem 5 in [1]). *Let D^k be quasi-regular with representative size $h = \mathrm{diam}\left(D^k\right)$. For $N \geq 0$, $w \in W^{N+1,\infty}\left(D^k\right)$, and $u \in W^{N+1,2}\left(D^k\right)$,*

(6)
$$\|uw - T_w u\|_{L^2(D^k)} \leq C_w h^{N+1} \|u\|_{W^{N+1,2}(D^k)},$$

(7)
$$\left\|uw - T_{1/w}^{-1}u\right\|_{L^2(D^k)} \leq C_w h^{N+1} \|u\|_{W^{N+1,2}(D^k)}.$$

where C_w depends on w as follows:

$$C_w = C \|w\|_{L^\infty(D^k)} \left\|\frac{1}{w}\right\|_{L^\infty(D^k)} \|w\|_{W^{N+1,\infty}(D^k)}.$$

THEOREM 3 (Theorem 6 in [1]). *Let $u \in W^{N+1,2}\left(D^k\right)$, $w \in W^{N+1,\infty}\left(D^k\right)$, and $v \in P^M\left(D^k\right)$ for $0 \leq M \leq N$; then*

$$\left|(wu,v)_{L^2(D^k)} - u, vw\right|$$
$$\leq Ch^{2N+2-M} \|w\|_{L^\infty(D^k)} \left\|\frac{1}{w}\right\|^2_{L^\infty(D^k)} \|w\|^2_{W^{N+1,\infty}(D^k)} \|u\|_{W^{N+1,2}(D^k)} \|v\|_{L^\infty(D^k)}.$$

4.2. Extension to matrix weights. We now aim to generalize weight-adjusted inner products to the case of matrix-valued weights. We first define appropriate generalizations of norms used in Section 4.1 to vector-valued functions. Let $D^{\boldsymbol{\alpha}}\boldsymbol{v}$ denote component-wise differentiation of $\boldsymbol{v}$ with respect to a d-dimensional multi-index $\boldsymbol{\alpha}$. Then, vector L^p Sobolev norms for $\boldsymbol{v}(\boldsymbol{x}) \in \mathbb{R}^m$ can be defined as

$$\|\boldsymbol{v}\|^p_{W^{k,p}} = \sum_{i=1}^m \|\boldsymbol{v}_i\|^p_{W^{k,2}} = \sum_{|\alpha|\leq k} \|D^\alpha \boldsymbol{v}\|^p_{L^p}, \qquad 1 \leq p < \infty,$$
$$\|\boldsymbol{v}\|_{W^{k,\infty}} = \max_i \|\boldsymbol{v}_i\|_{W^{k,\infty}}.$$

8

The corresponding Sobolev spaces $W^{k,p}$ and $W^{k,\infty}$ are defined similar to the scalar case.

Let $\boldsymbol{W}(\boldsymbol{x})$ be a matrix-valued weight function which is pointwise symmetric positive-definite matrix

$$0 < w_{\min} \leq \|\boldsymbol{W}(\boldsymbol{x})\|_2 \leq w_{\max} < \infty, \qquad 0 < \tilde{w}_{\min} \leq \left\|\boldsymbol{W}^{-1}(\boldsymbol{x})\right\|_2 \leq \tilde{w}_{\max} < \infty, \qquad \forall \boldsymbol{x} \in \Omega.$$

We define a kth order Sobolev norm for $\boldsymbol{W}(\boldsymbol{x})$ in terms of the induced p-norm

$$\|\boldsymbol{W}(\boldsymbol{x})\|_{k,p}^p = \sum_{|\boldsymbol{\alpha}| \leq k} \sup_{\boldsymbol{x}} \|D^{\boldsymbol{\alpha}} \boldsymbol{W}(\boldsymbol{x})\|_p^p$$

While this norm is not sub-multiplicative, the following bound holds

$$\|\boldsymbol{W}\boldsymbol{v}\|_{W^{k,p}}^p = \sum_{|\alpha| \leq k} \|D^\alpha (\boldsymbol{W}\boldsymbol{v})\|_{L^p}^p \leq C_N \int \sum_{|\alpha| \leq k} \sum_{|\beta| \leq |\alpha|} \left\| \left(D^\beta \boldsymbol{W}\right) \left(D^{\alpha-\beta} \boldsymbol{v}\right) \right\|_p^p$$

$$\leq C_N \int \left(\sum_{|\alpha| \leq k} \|(D^\alpha \boldsymbol{W})\| \right)^p \left(\sum_{|\alpha| \leq k} \|D^\alpha \boldsymbol{v}\|_p \right)^p$$

$$\leq C_N \|\boldsymbol{W}\|_{k,p}^p \|\boldsymbol{v}\|_{k,p}^p,$$

where we have used Leibniz's rule, Cauchy-Schwarz, and the arithmetic-geometric mean inequality.

To show that weight-adjusted DG methods are high order accurate, we make use of a scalar weighted interpolation estimate derived in [28, 1].

THEOREM 4 (Theorem 3.1 in [1]). *Let D^k be a quasi-regular element with representative size $h = \mathrm{diam}\left(D^k\right)$. For $N \geq 0$, $w \in W^{N+1,\infty}$, and $u \in W^{N+1,2}$,*

$$\left\| u - \frac{1}{w} \Pi_N(wu) \right\|_{D^k} \leq C_N h^{N+1} \left\| \sqrt{J} \right\|_{L^\infty(D^k)} \left\| \frac{1}{\sqrt{J}} \right\|_{L^\infty(D^k)} \left\| \frac{1}{w} \right\|_{L^\infty} \|w\|_{W^{N+1,\infty}} \|u\|_{W^{N+1,2}}.$$

The following theorem extends Theorem 4 to matrix weights by computing weighted interpolation estimates for the quantity $\boldsymbol{W}^{-1} \Pi_N \left(\boldsymbol{W}\boldsymbol{v} \right)$.

THEOREM 5. *Let D^k be a quasi-regular element with representative size $h = \mathrm{diam}\left(D^k\right)$. For $N \geq 0$, $\boldsymbol{W} \in \left(W^{N+1,\infty}\right)^{d \times d}$, and $\boldsymbol{v} \in \left(W^{N+1,2}\right)^d$,*

$$\left\| \boldsymbol{v} - \boldsymbol{W}^{-1} \Pi_N \left(\boldsymbol{W}\boldsymbol{v}\right) \right\|_{L^2} \leq C h^{N+1} \left\| \sqrt{J} \right\|_{L^\infty(D^k)} \left\| \frac{1}{\sqrt{J}} \right\|_{L^\infty(D^k)} \tilde{w}_{\max} \|\boldsymbol{W}\|_{N+1} \|\boldsymbol{v}\|_{W^{N+1,2}}$$

Proof. The proof is the similar to the scalar case. Using vector-valued versions of Bramble-Hilbert and a scaling argument for quasi-regular elements yields

$$\left\| \boldsymbol{v} - \boldsymbol{W}^{-1} \Pi_N \left(\boldsymbol{W}\boldsymbol{v}\right) \right\|_{L^2} \leq C_1 \left\| \sqrt{J} \right\|_{L^\infty(D^k)} \sup_{\boldsymbol{x}} \left\| \boldsymbol{W}^{-1} \right\|_2 \left\| \boldsymbol{W}\boldsymbol{v} - \Pi_N \left(\boldsymbol{W}\boldsymbol{v}\right) \right\|_{L^2}$$

$$\leq C_1 \left\| \sqrt{J} \right\|_{L^\infty(D^k)} \sup_{\boldsymbol{x}} \left\| \boldsymbol{W}^{-1} \right\|_2 |\boldsymbol{W}\boldsymbol{v}|_{W^{N+1,2}\left(\widehat{D}\right)}$$

$$\leq C_2 h^{N+1} \left\| \sqrt{J} \right\|_{L^\infty(D^k)} \left\| \frac{1}{\sqrt{J}} \right\|_{L^\infty(D^k)} \sup_{\boldsymbol{x}} \left\| \boldsymbol{W}^{-1} \right\|_2 \|\boldsymbol{W}\boldsymbol{v}\|_{W^{N+1,2}(D^k)}$$

$$\leq C_3 h^{N+1} \left\| \sqrt{J} \right\|_{L^\infty(D^k)} \left\| \frac{1}{\sqrt{J}} \right\|_{L^\infty(D^k)} \tilde{w}_{\max} \|\boldsymbol{W}\|_{W^{N+1,2}} \|\boldsymbol{v}\|_{W^{N+1,2}(D^k)}. \qquad \square$$

4.2.1. Properties of weight-adjusted vector-valued inner products. Let $\Pi_N \boldsymbol{u}$ be defined as the L^2 projection applied to each component of the vector-valued function $\boldsymbol{u}$. We then define the operator $T_{\boldsymbol{W}}$ analogously to the scalar case

$$T_{\boldsymbol{W}} \boldsymbol{v} = \Pi_N \left(\boldsymbol{W}\boldsymbol{v} \right).$$

9

The inverse operator $T_{\boldsymbol{W}}^{-1}$ is defined implicitly via

$$\left(\boldsymbol{W}T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{w}\right) = \left(\boldsymbol{v}, \boldsymbol{w}\right).$$

This definition is a straightforward generalization of T_w^{-1} to matrix-valued weights $\boldsymbol{W}$. Conveniently, all properties of T_w, T_w^{-1} for scalar $w(\boldsymbol{x})$ carry over to matrix weights $\boldsymbol{W}(\boldsymbol{x})$ as well [1].

LEMMA 6. *Let Π_N denote the component-wise L^2 projection, and let $\boldsymbol{W} \in (L^\infty)^m$. Then, $T_{\boldsymbol{W}}$ satisfies the following properties:*

1. $T_{\boldsymbol{W}}^{-1}T_{\boldsymbol{W}} = \Pi_N$
2. $\Pi_N T_{\boldsymbol{W}}^{-1} = T_{\boldsymbol{W}}^{-1}\Pi_N = T_{\boldsymbol{W}}^{-1}$
3. $\left\|T_{\boldsymbol{W}}^{-1}\right\|_{D^k} \le \tilde{w}_{\max}.$
4. $\left(T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{w}\right)$ *forms an inner product on $\left(P^N\right)^m \times \left(P^N\right)^m$, which is equivalent to the L^2 inner product with equivalence constants $C_1 = \tilde{w}_{\min}, C_2 = \tilde{w}_{\max}$.*

Proof. The proofs of properties 1 and 2 are consequences of the definition of $T_{\boldsymbol{W}}$, and are identical to proofs for the scalar case. Property 3 is a straightforward extension from the scalar case. Let $\boldsymbol{v} \in \left(P^N\right)^m$. Then,

$$\left(T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{v}\right) = \left(\boldsymbol{W}^{-1}\boldsymbol{W}T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{v}\right) \le \sup_{\boldsymbol{x}} \left\|\boldsymbol{W}^{-1}(\boldsymbol{x})\right\| \left(\boldsymbol{W}T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{v}\right) = \tilde{w}_{\max} \left\|\boldsymbol{v}\right\|_{L^2}^2.$$

Property 4 then simply requires the lower bound

$$\left(T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{v}\right) = \left(\boldsymbol{W}^{-1}\boldsymbol{W}T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{v}\right) \ge \inf_{\boldsymbol{x}} \left\|\boldsymbol{W}^{-1}(\boldsymbol{x})\right\| \left\|\boldsymbol{v}\right\|_{L^2}^2 = \tilde{w}_{\min} \left\|\boldsymbol{v}\right\|_{L^2}^2. \qquad \square$$

Using Theorem 5, we may also show that the matrix-valued weight-adjusted inner product is also high order accurate for sufficiently regular $\boldsymbol{W}$.

THEOREM 7. *Let D^k be a quasi-regular element with representative size $h = \mathrm{diam}\left(D^k\right)$. For $N > 0$*

$$\left\|\boldsymbol{W}\boldsymbol{v} - T_{\boldsymbol{W}^{-1}}^{-1}\boldsymbol{v}\right\|_{D^k} \le C_{\boldsymbol{W}} h^{N+1} \left\|\boldsymbol{v}\right\|_{W^{N+1,2}(D^k)}$$

with constant $C_{\boldsymbol{W}}$ depending on $\boldsymbol{W}$ and N

$$C_{\boldsymbol{W}} = C_N \left\|\sqrt{J}\right\|_{L^\infty(D^k)} \left\|\frac{1}{\sqrt{J}}\right\|_{L^\infty(D^k)} \tilde{w}_{\max} w_{\max} \left\|\boldsymbol{W}\right\|_{W^{N+1,2}(D^k)}.$$

Proof. The proof is again similar to the scalar case. The triangle inequality gives

$$\left\|\boldsymbol{W}\boldsymbol{v} - T_{\boldsymbol{W}^{-1}}^{-1}\boldsymbol{v}\right\|_{D^k} \le \left\|\boldsymbol{W}\boldsymbol{v} - \Pi_N(\boldsymbol{W}\boldsymbol{v})\right\|_{D^k} + \left\|\Pi_N(\boldsymbol{W}\boldsymbol{v}) - T_{\boldsymbol{W}^{-1}}^{-1}\boldsymbol{v}\right\|_{D^k}.$$

The former term is bounded by interpolation estimates and by arguments used in the proof of Theorem 5. The latter term is bounded

$$\begin{aligned}
\left\|\Pi_N(\boldsymbol{W}\boldsymbol{v}) - T_{\boldsymbol{W}^{-1}}^{-1}\boldsymbol{v}\right\|_{D^k} &= \left\|T_{\boldsymbol{W}^{-1}}^{-1}T_{\boldsymbol{W}^{-1}}\Pi_N(\boldsymbol{W}\boldsymbol{v}) - T_{\boldsymbol{W}^{-1}}^{-1}\Pi_N\boldsymbol{v}\right\|_{D^k} \\
&= \left\|T_{\boldsymbol{W}^{-1}}^{-1}\Pi_N\left(T_{\boldsymbol{W}^{-1}}\Pi_N(\boldsymbol{W}\boldsymbol{v})\right) - T_{\boldsymbol{W}^{-1}}^{-1}\Pi_N\boldsymbol{v}\right\|_{D^k} \\
&\le C_N \left\|T_{\boldsymbol{W}^{-1}}^{-1}\right\| \left\|T_{\boldsymbol{W}^{-1}}\Pi_N(\boldsymbol{W}\boldsymbol{v}) - \Pi_N\boldsymbol{v}\right\|_{D^k} \\
&\le C_N \left\|T_{\boldsymbol{W}^{-1}}^{-1}\right\| \left\|\Pi_N\left(\boldsymbol{W}^{-1}\Pi_N(\boldsymbol{W}\boldsymbol{v})\right) - \Pi_N\boldsymbol{v}\right\|_{D^k} \\
&\le C_N w_{\max} \left\|\boldsymbol{W}^{-1}\Pi_N(\boldsymbol{W}\boldsymbol{v}) - \boldsymbol{v}\right\|_{D^k}.
\end{aligned}$$

where we have used $\left\|T_{\boldsymbol{W}^{-1}}^{-1}\right\| \le w_{\max}$ (Lemma 6) and $\left\|\Pi_N\right\|_{D^k} = 1$. An application of Theorem 5 to bound $\left\|\boldsymbol{W}^{-1}\Pi_N(\boldsymbol{W}\boldsymbol{v}) - \boldsymbol{v}\right\|_{D^k}$ completes the proof. $\qquad \square$

This is a good generalization because defining $T_w^{-1}\boldsymbol{v} = T_{w\boldsymbol{I}}^{-1}\boldsymbol{v}$ recovers estimates for scalar weight-adjusted inner products.

10

4.2.2. Approximation of weighted mass matrix inverses. The advantage of using weight-adjusted inner products is that the corresponding weight-adjusted mass matrices are straightforward to invert. For scalar weights, the weight-adjusted mass matrix approximates the weighted L^2 mass matrix and its inverse by

$$M_w \approx M M_{1/w}^{-1} M, \qquad M_w^{-1} \approx M^{-1} M_{1/w} M^{-1}.$$

By evaluating $M_{1/w}$ in a matrix-free fashion (for example, using a sufficiently accurate quadrature rule), the inverse of the weight-adjusted mass matrix $M^{-1} M_{1/w} M^{-1}$ can be applied in a low storage fashion.

For weight-adjusted inner products with matrix-valued weights, the corresponding weight-adjusted mass matrices approximate weighted L^2 mass matrices and inverses in a similar fashion

$$M_{\boldsymbol{W}} \approx (\boldsymbol{I} \otimes \boldsymbol{M}) M_{\boldsymbol{W}^{-1}}^{-1} (\boldsymbol{I} \otimes \boldsymbol{M})$$

(8)
$$M_{\boldsymbol{W}}^{-1} \approx \left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right) M_{\boldsymbol{W}^{-1}} \left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right).$$

We note that, when $\boldsymbol{W}$ is constant over D^k, $M_{\boldsymbol{W}}$ reduces to the Kronecker product of the inverse stiffness tensor and the local mass matrix

$$M_{\boldsymbol{W}} = \boldsymbol{W} \otimes \boldsymbol{M}, \qquad M_{\boldsymbol{W}}^{-1} = \boldsymbol{W}^{-1} \otimes \boldsymbol{M}^{-1}.$$

In this case, we also have $M_{\boldsymbol{W}^{-1}} = \boldsymbol{W}^{-1} \otimes \boldsymbol{M}$. Substituting this explicit inverse into the weight adjusted mass matrix inverse $\left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right) M_{\boldsymbol{W}^{-1}} \left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right)$ in (8) recovers the exact inversion of $M_{\boldsymbol{W}}$.

5. An energy stable weight-adjusted discontinuous Galerkin formulation for elastic wave propagation. We construct a weight-adjusted DG method by simply replacing the weighted L^2 inner products appearing in the left hand side of the local DG formulation (4) with weight-adjusted approximations

$$\left(T_{1/\rho}^{-1}\frac{\partial \boldsymbol{v}}{\partial t}, \boldsymbol{w}\right)_{D^k} = \left(\sum_{i=1}^d \boldsymbol{A}_i^T \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{x}_i}, \boldsymbol{w}\right)_{D^k} + \left\langle \frac{1}{2}\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!] + \frac{\tau_v}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T[\![\boldsymbol{v}]\!], \boldsymbol{w}\right\rangle_{\partial D^k}$$

$$\left(T_{\boldsymbol{C}}^{-1}\frac{\partial \boldsymbol{\sigma}}{\partial t}, \boldsymbol{q}\right)_{D^k} = \left(\sum_{i=1}^d \boldsymbol{A}_i \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}_i}, \boldsymbol{q}\right)_{D^k} + \left\langle \frac{1}{2}\boldsymbol{A}_n[\![\boldsymbol{v}]\!] + \frac{\tau_\sigma}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n[\![\boldsymbol{\sigma}]\!], \boldsymbol{q}\right\rangle_{\partial D^k}.$$

Since the right hand side of the WADG formulation is identical to the right hand side of the DG formulation (2), WADG preserves a variant of the energy stability in Theorem 3

$$\sum_{D^k \in \Omega_h} \frac{1}{2}\frac{\partial}{\partial t}\left(\left(T_{1/\rho}^{-1}\boldsymbol{v}, \boldsymbol{v}\right)_{D^k} + \left(T_{\boldsymbol{C}}^{-1}\boldsymbol{\sigma}, \boldsymbol{\sigma}\right)_{D^k}\right) \leq - \sum_{f \in \Gamma_h \backslash \partial\Omega} \int_f \frac{\tau_v}{2}\left|\boldsymbol{A}_n[\![\boldsymbol{v}]\!]\right|^2 + \frac{\tau_\sigma}{2}\left|\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!]\right|^2$$

$$- \sum_{f \in \Gamma_v} \int_f \tau_v\left|\boldsymbol{A}_n\boldsymbol{v}^-\right|^2 - \sum_{f \in \Gamma_\sigma} \int_f \tau_\sigma\left|\boldsymbol{A}_n^T\boldsymbol{\sigma}^-\right|^2 - \sum_{f \in \Gamma_{\text{abc}}} \int_f \frac{\tau_v}{2}\left|\boldsymbol{A}_n\boldsymbol{v}^-\right|^2 + \frac{\tau_\sigma}{2}\left|\boldsymbol{A}_n^T\boldsymbol{\sigma}^-\right|^2 \leq 0.$$

The substitution of weight-adjusted inner products for weighted L^2 inner products replaces the weighted L^2 mass matrices in (5) by their weight-adjusted approximations. Inverting these weight-adjusted mass matrices yields the following system of ODEs for $\boldsymbol{V}, \boldsymbol{\Sigma}$

$$\frac{\partial \boldsymbol{V}}{\partial t} = \left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right) M_{(\rho^{-1}\boldsymbol{I})} \left(\sum_{i=1}^d \left(\boldsymbol{A}_i^T \otimes \boldsymbol{D}_i\right)\boldsymbol{\Sigma} + \sum_{f \in \partial D^k} \left(\boldsymbol{I} \otimes \boldsymbol{L}_f\right)\boldsymbol{F}_v\right)$$

$$\frac{\partial \boldsymbol{\Sigma}}{\partial t} = \left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right) M_{\boldsymbol{C}} \left(\sum_{i=1}^d \left(\boldsymbol{A}_i \otimes \boldsymbol{D}_i\right)\boldsymbol{V} + \sum_{f \in \partial D^k} \left(\boldsymbol{I} \otimes \boldsymbol{L}_f\right)\boldsymbol{F}_\sigma\right).$$

In practice, the matrices $\left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right) M_{(\rho^{-1}\boldsymbol{I})}$ and $\left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right) M_{\boldsymbol{C}}$ are applied to a vector using quadrature. After fusing operations together, this procedure can be boiled down to multiplication by two rectangular matrices.

We note that, for isotropic media, another way to address sub-element variations in material parameters is to diagonalize the matrix C through a change of variables [29]. This results in a local system of ODEs with only scalar weighted mass matrices [19], which can be treated using scalar weight-adjusted approximations. However, this approach is more difficult to apply when C is anisotropic.

One drawback of the analysis presented in this work is that accuracy is not guaranteed in the isotropic incompressible limit $\mu/\lambda \to 0$. In this case, the stiffness matrix C becomes singular, and the constant $\tilde{w}_{\max}$ in the upper bound on $\|C\|$ blows up. However, numerical experiments in Section 6.2.3 suggest that the accuracy of WADG for linear elastic wave propagation does not degrade significantly in the incompressible limit.

5.1. An energy stable formulation for curvilinear meshes. We have shown the energy stability of the DG formulation (2) on meshes of affine elements by assuming exact integration of all terms. However, energy stability can still be guaranteed if integrals are evaluated inexactly using quadrature. Instead of discretizing the "strong" DG formulation (2), we discretize the "skew-symmetric" formulation, where the right hand side of the velocity equations is integrated by parts,[1] resulting in the formulation

$$
\int_{\widehat{D}} T_{(\rho J)^{-1}}^{-1} \frac{\partial \boldsymbol{v}}{\partial t} \boldsymbol{w} = -\int_{\widehat{D}} \left(\sum_{i=1}^{d} \boldsymbol{\sigma}, \boldsymbol{A}_i \frac{\partial \boldsymbol{w}}{\partial \boldsymbol{x}_i} \right) J + \sum_{f \in \partial D^k} \int_{\widehat{f}} \left(\boldsymbol{A}_n^T \{\!\{ \boldsymbol{\sigma} \}\!\} + \frac{\tau_v}{2} \boldsymbol{A}_n^T \boldsymbol{A}_n [\![\boldsymbol{v}]\!] \right) \boldsymbol{w} J^f
$$

$$
\int_{\widehat{D}} T_{(J^{-1}C)}^{-1} \frac{\partial \boldsymbol{\sigma}}{\partial t} \boldsymbol{q} = \int_{\widehat{D}} \left(\sum_{i=1}^{d} \boldsymbol{A}_i \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}_i}, \boldsymbol{q} \right) J + \sum_{f \in \partial D^k} \int_{\widehat{f}} \left(\frac{1}{2} \boldsymbol{A}_n [\![\boldsymbol{v}]\!] + \frac{\tau_\sigma}{2} \boldsymbol{A}_n \boldsymbol{A}_n^T [\![\boldsymbol{\sigma}]\!] \right) \boldsymbol{q} J^f
$$

where we have incorporated spatial variations of J into the definition of the weights on the left hand side.

Since the proof of stability in Theorem 3 requires only algebraic manipulations of this formulation, the same energy stability result holds for this formulation. More importantly, a discrete version of energy stabilityis still guaranteed in the presence of inexact quadrature, where integrals in Theorem 3 are replaced with quadrature approximations. This is especially important for discretizations on curvilinear meshes, where the exact integration of spatially varying geometric factors and Jacobians can be either inhibitively expensive for high order curvilinear mappings or impossible for rational mappings [30, 31].

We note that, to ensure energy stability on curved and non-affine elements, the "skew-symmetric" formulation must be evaluated explicitly using quadrature, which is typically more expensive than quadrature-free evaluations used to evaluate the "strong" DG formulation (2). These costs can be slightly reduced for curvilinear meshes by evaluating the DG formulation using the "skew-symmetric" formulation on curvilinear elements and the more efficient "strong" DG formulation on affine elements [2].

5.2. Convergence analysis.

6. Numerical experiments. The following sections present several numerical experiments validating the stability and accuracy of the proposed method in two and three dimensions.

6.1. Verification of stability. blah
Eigenvalue plots with $\tau = 0$ and $\tau = O(1)$ for both square and curvilinear grids with random coefficients close to zero and 1. Change in solution norm as time goes on.

6.2. Analytic solutions. Next, we study the accuracy and convergence of weight-adjusted DG method for several analytical solutions in linear elasticity. In all cases, the solution in expressed terms of the displacement vector $\boldsymbol{u}(\boldsymbol{x}, t)$. Initial conditions for velocity and stress are computed through

$$
\boldsymbol{v}(\boldsymbol{x}, t) = \frac{\partial \boldsymbol{u}}{\partial t}, \qquad \boldsymbol{\sigma} = \boldsymbol{C} \boldsymbol{E} = \boldsymbol{C} \frac{1}{2} \left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T \right).
$$

---

[1]The choice of which equation to integrate by parts is arbitrary, since integrating the stress equations by parts also results in an energy stable formulation.
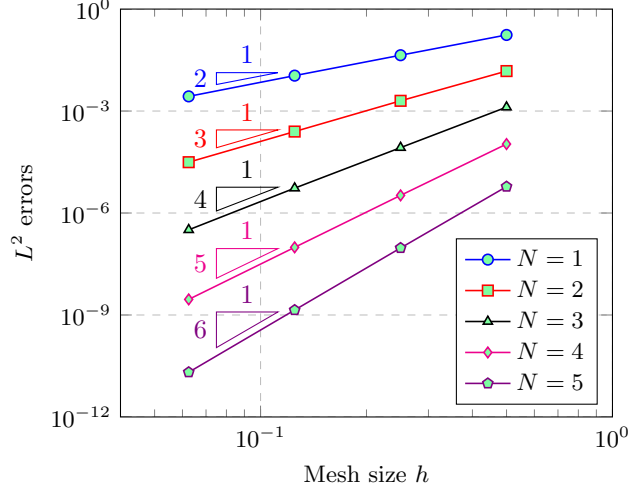
Fig. 1: Convergence of L^2 errors for harmonic oscillation.

Unless otherwise stated, we report relative L^2 errors for all components of the solution $\boldsymbol{U} = (\boldsymbol{v}, \boldsymbol{\sigma})$

$$\frac{\|\boldsymbol{U} - \boldsymbol{U}_h\|_{L^2(\Omega)}}{\|\boldsymbol{U}\|_{L^2(\Omega)}} = \frac{\left(\sum_{i=1}^m \|\boldsymbol{U}_i - \boldsymbol{U}_{i,h}\|_{L^2(\Omega)}^2\right)^{1/2}}{\left(\sum_{i=1}^m \|\boldsymbol{U}_i\|_{L^2(\Omega)}^2\right)^{1/2}}.$$

6.2.1. Harmonic oscillation of a square. We first examine convergence on a unit square domain with $\lambda = \mu = \rho = 1$. The displacement is given by

$$u_1(x, y, t) = \cos(\omega \pi t) \cos(\pi x) \sin(\pi y)$$
$$u_2(x, y, t) = -\cos(\omega \pi t) \sin(\pi x) \cos(\pi y),$$

where $\omega = \sqrt{2\mu}$. Figure 1 shows L^2 errors computed at time $T = 1$, using uniform triangular meshes constructed by bisecting a uniform mesh of quadrilaterals along the diagonal. For $N = 1, \ldots, 5$, optimal $O(h^{N+1})$ rates of convergence are observed.

6.2.2. Rayleigh and Lamb waves. Next, we examine the convergence of WADG for Rayleigh and Lamb waves, both of which test the imposition of traction-free boundary conditions.

Rayleigh waves are elastic surface waves which decay exponentially away from the surface. These waves are given by the displacement vector

$$\boldsymbol{v}(x, y, t) = e^{-\omega x \sqrt{1-\xi^2}} \left(\begin{array}{c} \cos(\omega(y + c_r t)) \\ \sqrt{1 - \xi^2} \sin(\omega(y + c_r t)) \end{array} \right)$$
$$+ \left(\frac{\xi^2}{2} - 1 \right) e^{-\omega x \sqrt{1 - \frac{\xi^2 \mu}{2\mu + \lambda}}} \left(\begin{array}{c} \cos(\omega(y + c_r t)) \\ \sin(\omega(y + c_r t))/\sqrt{1 - \frac{\xi^2 \mu}{2\mu + \lambda}} \end{array} \right),$$

where ω is the wavespeed, c_r is the Rayleigh phase velocity $c_r = \xi \sqrt{\mu}$, and ξ satisfies

$$\sqrt{1 - \xi^2} \sqrt{1 - \frac{\xi^2 \mu}{2\mu + \lambda}} - \left(\frac{\xi^2}{2} - 1 \right)^2 = 0.$$

In our computations, we use $\rho = \mu = \lambda = 1$, $\xi = 0.949554083888034$, and $\omega = 2\pi$. We solve on the domain $[0, 2] \times [0, 1]$ using a sequence of uniform triangular meshes, and enforce traction-free boundary conditions at
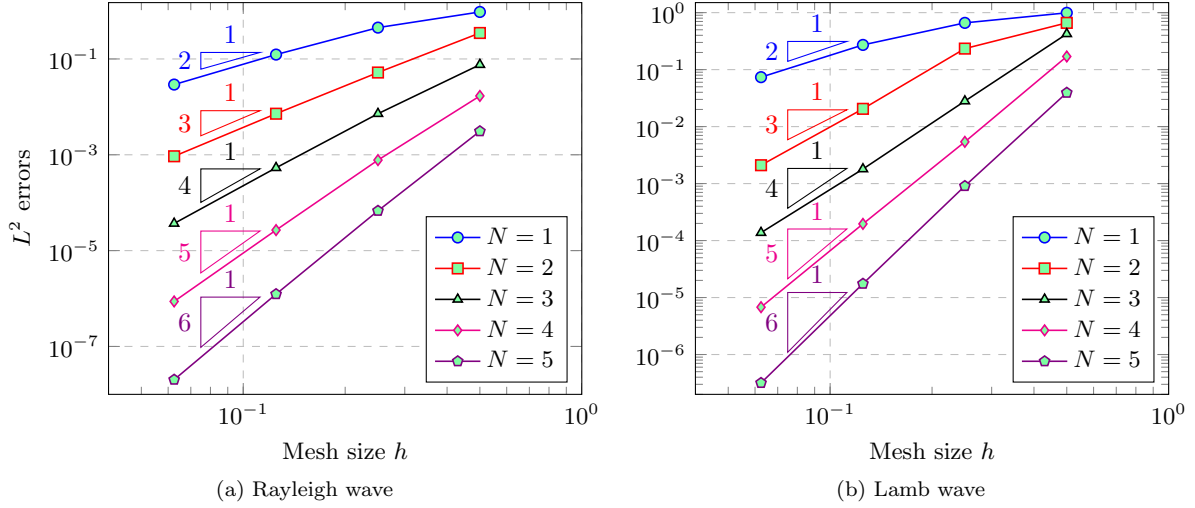
Fig. 2: Convergence of L^2 errors for Rayleigh and Lamb wave solutions.

$x = 0$ and exact Dirichlet boundary conditions at $x = 2$. Periodic boundary conditions are applied at $y = 0$ and $y = 1$.

Lamb waves are supported by elastic waveguides with traction-free (free surface) boundary conditions at the top and bottom of the domain. The displacement of these waves

$$\boldsymbol{v}_1(x, y, t) = (-kB_1 \cos(py) - qB_2 \cos(qy)) \sin(kx - \omega t)$$
$$\boldsymbol{v}_2(x, y, t) = (-pB_1 \sin(py) + kB_2 \sin(qy)) \cos(kx - \omega t)$$

where k is the wavenumber and ω is the frequency, and the constants p and q are defined as

$$p^2 = \frac{\omega^2}{2\mu + \lambda} - k^2, \qquad q^2 = \frac{\omega^2}{\mu} - k^2.$$

The wavenumber k and frequency ω are related through a dispersion relation. The ratio of the amplitudes B_1/B_2 can be determined using other parameters, implying that B_1, B_2 are unique up to a scaling constant. In our experiments, we use $\rho = \mu = 1$, $\lambda = 2$, $k = 2\pi$. For these values, $\omega = 13.137063197233$, $B_1 = 126.1992721468$ and $B_2 = 53.88807700007$. We solve on the domain $[-1, 1] \times [-1/2, 1/2]$, with traction-free boundary conditions at $y = \pm 1/2$ and periodic boundary conditions at $x = \pm 1$.

Figure 2b shows L^2 errors for both waves at time $T = 1$. Observed convergence rates fall between the optimal rate of $O(h^{N+1})$ and theoretical rate of $O(h^{N+1/2})$ [32].

6.2.3. Rayleigh waves in near-incompressible materials. As noted in Section 5, error estimates for isotropic elasticity no longer hold in the incompressible limit with $\mu/\lambda \to 0$ limit due to the fact that $\boldsymbol{C}$ becomes singular. We use the propagation of Rayleigh waves to examine the behavior of WADG for near-incompressible materials. We follow [33, 34] and fix $\lambda = 1$ and set $\mu = 1, .1, .01, .001, .0001$. Since the Rayleigh wave propagates with speed proportional to $\sqrt{\mu}$, we compute L^2 errors at the final time $1/(4\sqrt{\mu})$ to ensure a fair comparison between solutions at different values of μ.

Table 1 shows relative errors for $\boldsymbol{v}$, $\boldsymbol{\sigma}$ at different orders and mesh sizes. The relative errors for $\boldsymbol{\sigma}$ grow as $\mu/\lambda \to 0$. This is not surprising, as the constant in the error estimates of Section 4.2.1 blows up in the incompressible limit. However, because the magnitude of $\boldsymbol{\sigma}$ also decreases as $\mu/\lambda \to 0$, relative errors for $\boldsymbol{v}$ remain roughly the same magnitude as the material approaches incompressibility.

6.2.4. Stoneley waves. A Stoneley wave is supported along the interface between two solids [35]. Like Rayleigh waves, Stoneley waves decay exponentially away from the interface, and test the effectiveness of

	$\mu = 1$	$\mu = .1$	$\mu = .01$	$\mu = .001$	$\mu = .0001$
$N = 2, h = 1/4$	0.0288	0.0270	0.0309	0.0401	0.0568
$N = 3, h = 1/4$	0.0030	0.0027	0.0030	0.0037	0.0049
$N = 4, h = 1/4$	2.7705e-04	2.5060e-04	2.9185e-04	3.9779e-04	5.2470e-04
$N = 2, h = 1/8$	0.0028	0.0025	0.0029	0.0036	0.0047
$N = 3, h = 1/8$	1.8315e-04	1.6960e-04	1.8524e-04	2.2522e-04	2.8301e-04
$N = 4, h = 1/8$	8.2285e-06	7.5813e-06	8.1174e-06	1.0762e-05	1.4886e-05

(a) Relative error in $\boldsymbol{v}$

	$\mu = 1$	$\mu = .1$	$\mu = .01$	$\mu = .001$	$\mu = .0001$
$N = 2, h = 1/4$	0.0656	0.0754	0.1214	0.2134	0.4145
$N = 3, h = 1/4$	0.0091	0.0108	0.0182	0.0321	0.0576
$N = 4, h = 1/4$	9.7408e-04	0.0012	0.0022	0.0041	0.0068
$N = 2, h = 1/8$	0.0093	0.0116	0.0204	0.0379	0.0765
$N = 3, h = 1/8$	6.8288e-04	8.0145e-04	0.0015	0.0031	0.0056
$N = 4, h = 1/8$	3.4231e-05	4.3510e-05	8.8747e-05	1.9771e-04	3.6053e-04

(b) Relative error in $\boldsymbol{\sigma}$

Table 1: Behavior of WADG for linear elastic wave propagation in the incompressible limit as $\mu/\lambda \to 0$. Errors are shown for various orders and mesh resolutions.

numerical fluxes across interfaces. We follow [4, 34] and use discontinuous media defined by

$$(\rho, \lambda, \mu) = \begin{cases} (10, 3, 3) & y > 0 \\ (1, 1, 1) & y < 0. \end{cases}.$$

The displacement vector for a Stoneley wave is then given by

$$\boldsymbol{v}_1(x, y, t) = \begin{cases} \mathrm{Re}\left(\left(ikB_1 e^{-kb_{1p}y} + kb_{1s}B_2 e^{-kb_{1s}y} \right) e^{i(ky-\omega t)} \right), & y > 0 \\ \mathrm{Re}\left(\left(-kb_{1p}B_1 e^{-kb_{1p}y} + ikB_2 e^{-kb_{1s}y} \right) e^{i(kx-\omega t)} \right), & y < 0 \end{cases}$$

$$\boldsymbol{v}_2(x, y, t) = \begin{cases} \mathrm{Re}\left(\left(ikB_3 e^{kb_{2p}y} - kb_{2s}B_4 e^{kb_{2s}y} \right) e^{i(kx-\omega t)} \right), & y > 0 \\ \mathrm{Re}\left(\left(kb_{2p}B_3 e^{kb_{2p}y} + ikB_4 e^{kb_{2s}y} \right) e^{i(kx-\omega t)} \right), & y < 0 \end{cases},$$

where c_{st} is the Stoneley wave speed, and

$$k = \omega/c_{st}, \qquad b_{jp} = \sqrt{1 - \frac{c_{st}^2}{(2\mu_j + \lambda_j)/\rho_j}}, \qquad b_{js} = \sqrt{1 - \frac{c_{st}^2}{(\mu_j)/\rho_j}}, \qquad j = 1, 2.$$

The Stoneley wave speed c_{st} can be determined based material parameters and interface conditions, and the amplitudes B_1, B_2, B_3, B_4 are determined from c_{st} up to scaling by a constant. For the parameters used in this study, we take $c_{st} = 0.546981324213884$, $B_1 = i0.2952173626624$, $B_2 = -0.6798795208473$, $B_3 = i0.5220044931212$, and $B_4 = -0.9339639688697$. We assume $k = 1$, which gives $\omega = c_{st}$.

We solve on the domain $[-1, 1] \times [-5, 5]$, and enforce Dirichlet boundary conditions at all boundaries using the exact solution. Figure 3 shows L^2 errors for two uniform meshes of triangles constructed by bisecting a quadrilateral mesh of $K_{1D} \times 5K_{1D}$ elements. Figure 3a shows errors when K_{1D} is even and the mesh is fitted to the interface at $y = 0$, while Figure 3b shows errors when K_{1D} is odd and the interface cuts through element interiors. When the mesh is fitted to the interface, we observe optimal $O(h^{N+1})$ rates of convergence for $N > 1$.

When the mesh is not fitted to the interface exactly, we compute the application of the weight-adjusted mass matrix inverse using a quadrature rule from Xiao and Gimbutas [36] which is exact for degree $2N + 1$

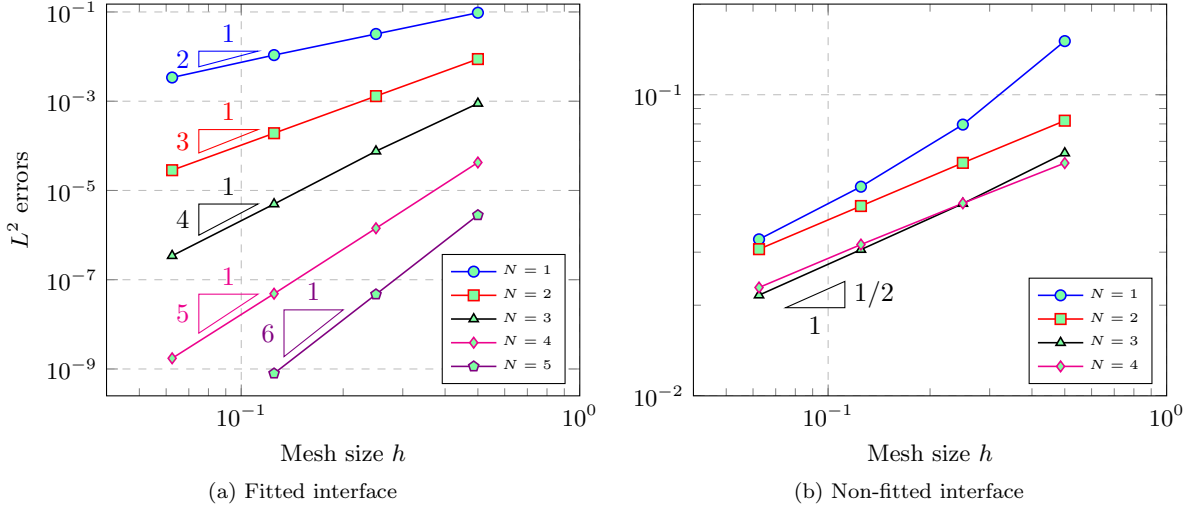|	
(a) Fitted interface	(b) Non-fitted interface

Fig. 3: Convergence of WADG for a Stoneley wave using a fitted mesh aligned with the interface and a non-fitted mesh where the interface does not lie exactly on an element boundary.

polynomials. Since the values of ρ, μ, and λ are positive at all quadrature points, the method is energy stable. However, since the exact solution is discontinuous, we expect an $O(1)$ error in elements cut by the interface, resulting in L^2 errors which converge at rate $O(h^{1/2})$. We note that, when using piecewise constant approximations of μ and λ, we observe the same $O(h^{1/2})$ convergence rate, though errors are roughly twice as large in magnitude. Discuss possibly extending theory for WADG to low regularity weights.

6.2.5. Harmonic oscillation of an annulus. This example tests the convergence of WADG for curvilinear mappings.

6.3. Application examples.

6.3.1. Stiff inclusion. Leveque stiff inclusion example.

6.3.2. Heterogeneous anisotropic material. From [7]: discontinuous media with anisotropy on one side, isotropy on the other.

Also do a similar case where random noise is added to both sides and the anisotropy is more strongly discontinuous.

6.4. Three-dimensional example. Add GPU computations here.

7. Conclusions. Generally applicable to problems with spatially varying matrices $\boldsymbol{A}_0(\boldsymbol{x})$, such as in Maxwell's equations for cloaking problems [37], as well as in vector-valued $H(\mathrm{div})$ or $H(\mathrm{curl})$ functions under matrix-valued mappings such as Piola transforms.

Extension to hybrid meshes in 3D, care must be taken with pyramids [38, 39].

which can be further accelerated by exploiting structure under Bernstein-Bezier bases [39]. Formulation allows for the reduction of constant-coefficient RHS costs at high orders using BBDG [40].

8. Acknowledgments. The author gratefully thanks Thomas Hagstrom, Tim Warburton, Axel Modave, and Ruichao Ye for helpful and informative discussions.

REFERENCES

[1] Jesse Chan, Russell J Hewett, and T Warburton. Weight-adjusted discontinuous Galerkin methods: wave propagation in heterogeneous media. *arXiv preprint arXiv:1608.01944*, 2016.

[2] Jesse Chan, Russell J Hewett, and T Warburton. Weight-adjusted discontinuous Galerkin methods: curvilinear meshes. *arXiv preprint arXiv:1608.03836*, 2016.

[3] Dimitri Komatitsch, Gordon Erlebacher, Dominik Göddeke, and David Michéa. High-order finite-element seismic wave propagation modeling with MPI on a large GPU cluster. *Journal of computational physics*, 229(20):7692–7714, 2010.

[4] Lucas C Wilcox, Georg Stadler, Carsten Burstedde, and Omar Ghattas. A high-order discontinuous Galerkin method for wave propagation through coupled elastic–acoustic media. *Journal of Computational Physics*, 229(24):9373–9396, 2010.

[5] Martin Käser and Michael Dumbser. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - I. The two-dimensional isotropic case with external source terms. *Geophysical Journal International*, 166(2):855–877, 2006.

[6] Michael Dumbser and Martin Käser. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - II. the three-dimensional isotropic case. *Geophysical Journal International*, 167(1):319–336, 2006.

[7] Josep de la Puente, Martin Käser, Michael Dumbser, and Heiner Igel. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes-IV. Anisotropy. *Geophysical Journal International*, 169(3):1210–1228, 2007.

[8] Ruichao Ye, Maarten V de Hoop, Christopher L Petrovitch, Laura J Pyrak-Nolte, and Lucas C Wilcox. A discontinuous Galerkin method with a modified penalty flux for the propagation and scattering of acousto-elastic waves. *Geophysical Journal International*, 205(2):1267–1289, 2016.

[9] Beatrice Riviere and Mary F Wheeler. Discontinuous finite element methods for acoustic and elastic wave problems. *Contemporary Mathematics*, (329):271–282, 2003.

[10] Béatrice Riviere, Simon Shaw, and JR Whiteman. Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems. *Numerical Methods for Partial Differential Equations*, 23(5):1149–1166, 2007.

[11] Sarah Delcourte, Loula Fezoui, and Nathalie Glinsky-Olivier. A high-order discontinuous Galerkin method for the seismic wave propagation. In *ESAIM: Proceedings*, volume 27, pages 70–89. EDP Sciences, 2009.

[12] Sarah Delcourte and Nathalie Glinsky. Analysis of a high-order space and time discontinuous Galerkin method for elastodynamic equations. Application to 3D wave propagation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(4):1085–1126, 2015.

[13] Andreas Klöckner, Tim Warburton, Jeff Bridge, and Jan S Hesthaven. Nodal discontinuous Galerkin methods on graphics processors. *Journal of Computational Physics*, 228(21):7863–7882, 2009.

[14] Nico Gödel, Steffen Schomann, Tim Warburton, and Markus Clemens. GPU accelerated Adams–Bashforth multirate discontinuous Galerkin FEM simulation of high-frequency electromagnetic fields. *Magnetics, IEEE Transactions on*, 46(8):2735–2738, 2010.

[15] Nico Gödel, Nigel Nunn, Tim Warburton, and Markus Clemens. Scalability of higher-order discontinuous Galerkin FEM computations for solving electromagnetic wave propagation problems on GPU clusters. *Magnetics, IEEE Transactions on*, 46(8):3469–3472, 2010.

[16] A Modave, A St-Cyr, WA Mulder, and T Warburton. A nodal discontinuous Galerkin method for reverse-time migration on GPU clusters. *Geophysical Journal International*, 203(2):1419–1435, 2015.

[17] Axel Modave, Amik St-Cyr, and Tim Warburton. GPU performance analysis of a nodal discontinuous Galerkin method for acoustic and elastic models. *Computers & Geosciences*, 91:64–76, 2016.

[18] Cristóbal E Castro, Martin Käser, and Gilbert B Brietzke. Seismic waves in heterogeneous material: subcell resolution of the discontinuous Galerkin method. *Geophysical Journal International*, 182(1):250–264, 2010.

[19] E Diego Mercerat and Nathalie Glinsky. A nodal high-order discontinuous Galerkin method for elastic wave propagation in arbitrary heterogeneous media. *Geophysical Journal International*, 201(2):1101–1118, 2015.

[20] Mark A Christon. The influence of the mass matrix on the dispersive nature of the semi-discrete, second-order wave equation. *Computer methods in applied mechanics and engineering*, 173(1):147–166, 1999.

[21] Jean-Luc Guermond and Richard Pasquetti. A correction technique for the dispersive effects of mass lumping for transport problems. *Computer Methods in Applied Mechanics and Engineering*, 253:186–198, 2013.

[22] Thomas JR Hughes and Jerrold E Marsden. Classical elastodynamics as a linear symmetric hyperbolic system. *Journal of Elasticity*, 8(1):97–110, 1978.

[23] Jesse Chan and T Warburton. A short note on the penalty flux parameter for first order discontinuous Galerkin formulations. *arXiv preprint arXiv:1611.00102*, 2016.

[24] Randall J LeVeque. *Finite volume methods for hyperbolic problems*, volume 31. Cambridge university press, 2002.

[25] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200, 1994.

[26] Thomas Hagstrom and Timothy Warburton. A new auxiliary variable formulation of high-order local radiation boundary conditions: corner compatibility conditions and extensions to first-order systems. *Wave motion*, 39(4):327–338, 2004.

[27] Axel Modave, Andreas Atle, Jesse Chan, and Tim Warburton. High-order absorbing boundary conditions with corner/edge compatibility for GPU-accelerated discontinuous Galerkin wave simulations. *arXiv preprint arXiv:1610.05023*, 2016.

[28] T. Warburton. A low-storage curvilinear discontinuous Galerkin method for wave problems. *SIAM Journal on Scientific Computing*, 35(4):A1987–A2012, 2013.

[29] V Etienne, E Chaljub, J Virieux, and N Glinsky. An hp-adaptive discontinuous Galerkin finite-element method for 3-D elastic wave modelling. *Geophysical Journal International*, 183(2):941–962, 2010.

[30] Luke Engvall and John A Evans. Isogeometric triangular Bernstein–Bézier discretizations: Automatic mesh generation and geometrically exact finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 304:378–407, 2016.

[31] C. Michoski, J. Chan, L. Engvall, and J.A. Evans. Foundations of the blended isogeometric discontinuous Galerkin (BIDG) method. *Computer Methods in Applied Mechanics and Engineering*, 305:658 – 681, 2016.

[32] Claes Johnson and Juhani Pitkäranta. An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation.

   Mathematics of computation, 46(173):1–26, 1986.

[33] Björn Sjögreen and N Anders Petersson. A fourth order accurate finite difference scheme for the elastic wave equation in second order formulation. *Journal of Scientific Computing*, 52(1):17–48, 2012.

[34] Daniel Appelo and Thomas Hagstrom. An energy-based discontinuous Galerkin discretization of the elastic wave equation in second order form. 2015. Submitted to CMAME.

[35] Robert Stoneley. Elastic waves at the surface of separation of two solids. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 106(738):416–428, 1924.

[36] H Xiao and Zydrunas Gimbutas. A numerical algorithm for the construction of efficient quadrature rules in two and higher dimensions. *Comput. Math. Appl.*, 59:663–676, 2010.

[37] Jichun Li and Yunqing Huang. *Time-domain finite element methods for Maxwell's equations in metamaterials*, volume 43. Springer Science & Business Media, 2012.

[38] Jesse Chan and T Warburton. Orthogonal bases for vertex-mapped pyramids. *SIAM Journal on Scientific Computing*, 38(2):A1146–A1170, 2016.

[39] Jesse Chan, Zheng Wang, Axel Modave, Jean-Francois Remacle, and T Warburton. GPU-accelerated discontinuous Galerkin methods on hybrid meshes. *Journal of Computational Physics*, 318:142–168, 2016.

[40] Jesse Chan and T Warburton. GPU-accelerated Bernstein-Bezier discontinuous Galerkin methods for wave problems. *arXiv preprint arXiv:1512.06025*, 2015.