WEIGHT-ADJUSTED DISCONTINUOUS GALERKIN METHODS: MATRIX-VALUED WEIGHTS AND ELASTIC WAVE PROPAGATION IN HETEROGENEOUS MEDIA

JESSE CHAN

Abstract. Weight-adjusted inner products [1, 2] are easily invertible approximations to weighted L^2 inner products. These approximations can be paired with a discontinuous Galerkin (DG) discretization to produce a time-domain method for wave propagation which is low storage, energy stable, and high order accurate for arbitrary heterogeneous media and curvilinear meshes. In this work, we extend weight-adjusted DG (WADG) methods to the case of matrix-valued weights, with the linear elastic wave equation as an application. We present a DG formulation of the symmetric form of the linear elastic wave equation, with upwind-like dissipation incorporated through simple penalty fluxes. A semi-discrete 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. Efficient and accurate methods for elastic wave propagation form a foundation for a broad range of applications, from seismic and medical imaging to rupture and earthquake simulation. Finite differences are the most common choice of method [3]; however, finite element methods have garnered interest due to their low numerical dispersion and ability to accommodate geometrically flexible unstructured meshes.

Typical methods for time-domain wave propagation utilize explicit time stepping, since the hyperbolic partial differential equations (PDEs) which govern wave propagation admit a reasonable stable time-step restriction. However, finite element methods require the inversion of a global mass matrix when paired with explicit time integrators. Spectral element methods (SEM) sidestep this issue on hexahedral meshes by utilizing a mass lumped diagonal mass matrix [4]. The inversion of a globally coupled matrix can also be avoided through the use of discontinuous Galerkin (DG) methods, which result in a locally invertible block diagonal mass matrices. Due to difficulties in extending mass-lumping techniques from hexahedra to tetrahedra, high order DG methods are often employed for seismic simulations which require the use of simplicial meshes [5, 6, 7, 8, 9, 10]. High order DG methods also lend themselves well to efficient implementations using Graphics Processing Units (GPUs) [11, 12, 13, 14].

Most high order DG methods on simplicial meshes assume that material coefficients are constant over each element, which can result in spurious reflections and the loss of high order accuracy [15, 16, 17]. This limitation can be overcome by incorporating sub-element heterogeneities into weighted mass matrices, resulting in a DG method which is both high order accurate and energy stable [16, 17]. On tetrahedral meshes, this approach necessitates the pre-computation and storage of factorizations or inverses for each local mass matrix, which greatly increases both storage costs and data transferred at high orders of approximation. These costs are especially problematic for accelerator architectures such as GPUs, which possess limited memory.

Storage costs associated with weighted mass matrices can be avoided by approximating weighted L^2 inner products using weight-adjusted inner products, which result in easily invertible approximations to weighted mass matrices [1, 2]. For sufficiently regular weights, high order accuracy is also retained. When paired with an energy stable DG formulation, these approximations result in weight-adjusted DG methods (WADG), which preserve energy stability and high order accuracy while retaining a low asymptotic storage cost. Additionally, unlike mass-lumping techniques, WADG methods do not rely on the use of inexact quadrature rules, and reduce to the exact inversion of mass matrices for constant weights.

Weight-adjusted DG methods have been applied to acoustic wave propagation in heterogeneous media and on curvilinear meshes [1, 2]. Both of these previous applications have involved scalar weighting functions. In this work, we extend weight-adjusted inner products to matrix-valued weights, and demonstrate the usefulness of this generalization to elastic wave propagation in arbitrary media. A new energy stable DG formulation is introduced based on the symmetric form of the elastic wave equations, with upwind-like numerical dissipation introduced through simple penalty fluxes [18]. In contrast to the fluxes proposed in [10], the penalty fluxes used here can be made to be independent of material coefficients.

This work proceeds as follows: Sections 2 and 3 present an energy stable DG formulation with simple penalty fluxes for the symmetric hyperbolic form of the elastic wave equation, and discuss issues related to storage and inversion of local mass matrices for material coefficients with sub-element variations. Section 4 extends weight-adjusted approximations to weighted L^2 inner products and mass matrices to the case of

matrix-valued weights, and provides interpolation estimates which account for the regularity of the matrix weight. These results are incorporated into a weight-adjusted DG method for the linear elastic wave equations in Section 5. Finally, numerical results in Section 6 demonstrate the accuracy of this method for several problems in linear elasticity.

2. Symmetric form of the elastic wave equation. We begin with the linear elastic wave equation in a domain $\Omega \in \mathbb{R}^d$. These equations can be written as a first order velocity-stress system for velocity $\boldsymbol{v}$ and symmetric stress tensor $\boldsymbol{S}$

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

where $\boldsymbol{f}$ is the body force per unit volume, ρ is density, and $\underline{\boldsymbol{C}}$ is the symmetric constitutive stiffness tensor relating stress and strain. We rewrite these equations as a symmetric hyperbolic system of PDEs [19] using Voigt notation

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

(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 $\underline{\boldsymbol{C}}$ and $\boldsymbol{\sigma}$ is a vector of length $N_d = \frac{d(d+1)}{2}$, the number of unique entries of the stress tensor $\boldsymbol{S}$ in d dimensions. We note that the matrices $\boldsymbol{A}_i$ are spatially constant, while $\rho, \boldsymbol{C}$, and $\boldsymbol{C}^{-1}$ can vary spatially. Furthermore, we will assume that ρ and $\boldsymbol{C}$ are positive-definite and bounded pointwise such that

$$0 < \rho_{\min} \leq \rho(\boldsymbol{x}) \leq \rho_{\max} < \infty,$$
$$0 < c_{\min} \leq \boldsymbol{u}^T \boldsymbol{C}(\boldsymbol{x}) \boldsymbol{u} \leq c_{\max} < \infty$$
$$0 < \tilde{c}_{\min} \leq \boldsymbol{u}^T \boldsymbol{C}^{-1}(\boldsymbol{x}) \boldsymbol{u} \leq \tilde{c}_{\max} < \infty$$

for all $\boldsymbol{x} \in \mathbb{R}^d$ and all $\boldsymbol{u} \in \mathbb{R}^{N_d}$.

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

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

while the matrices $\boldsymbol{A}_i$ are

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

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

$$\begin{pmatrix} 2\mu + \lambda & \lambda & \lambda & \\ \lambda & 2\mu + \lambda & \lambda & \\ \lambda & \lambda & 2\mu + \lambda & \\ & & & \mu \boldsymbol{I}^{3\times3} \end{pmatrix}.$$

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 straight-forward 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 $\boldsymbol{x}^k = \left\{ x^k, y^k \right\}$ for $d = 2$ and $\boldsymbol{x}^k = \left\{ x^k, y^k, z^k \right\}$ for $d = 3$ denote physical coordinates on D^k and $\widehat{\boldsymbol{x}} = \{\widehat{x}, \widehat{y}\}$ for $d = 2$ and $\widehat{\boldsymbol{x}} = \{\widehat{x}, \widehat{y}, \widehat{z}\}$ for $d = 3$ 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 solution 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 $V_h\left(\widehat{D}\right)$

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

The global approximation space $V_h\left(\Omega_h\right)$ is then defined as the direct sum of elemental approximation spaces

$$V_h\left(\Omega_h\right) = \bigoplus_{D^k} V_h\left(D^k\right).$$

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 on a triangle is

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

and in three dimensions, P^N on a tetrahedron is

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

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

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

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

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

We define also the L^2 inner product and norm over the boundary ∂D^k of an element

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{L^2(\partial D^k)} = \int_{\partial D^k} \boldsymbol{u} \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} = \sum_{f \in \partial D^k} \int_{\widehat{f}} \boldsymbol{u} \cdot \boldsymbol{v} J^f \, \mathrm{d}\widehat{\boldsymbol{x}}, \qquad \|\boldsymbol{u}\|^2_{L^2(\partial D^k)} = \langle \boldsymbol{u}, \boldsymbol{u} \rangle_{L^2(\partial D^k)},$$

where J^f is the Jacobian of the mapping from a reference face $\widehat{f}$ to a physical face f of an element.

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

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

The jump and average of a scalar function $u \in V_h(\Omega_h)$ 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

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

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 [20]. For the linear elastic wave equation in symmetric first order form, this formulation is given as

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

(2)

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

for all $\boldsymbol{w}, \boldsymbol{q} \in V_h(\Omega_h)$. Here, $\tau_v, \tau_\sigma \ge 0$ 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

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

We note that, unlike the penalty DG formulation given in [10], material coefficients $\rho, \boldsymbol{C}$ appear only on the left hand side of (2). Thus, efficient techniques for constant coefficient formulations [21] can be used to evaluate the right hand side of the formulation, even in the presence of sub-element variations in $\rho, \boldsymbol{C}$.

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 [22, 23] 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 $[\![A_n^T\boldsymbol{\sigma}]\!] = [\![Sn]\!]$.

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

$$[\![v]\!] = 2\left(v_{\mathrm{bc}} - v^-\right), \qquad [\![A_n^T\boldsymbol{\sigma}]\!] = [\![Sn]\!] = 0,$$

while traction boundary conditions are enforced through

$$[\![A_n^T\boldsymbol{\sigma}]\!] = [\![Sn]\!] = 2\left(t_{\mathrm{bc}} - S^-n\right) = 2\left(t_{\mathrm{bc}} - A_n^T\boldsymbol{\sigma}^-\right), \qquad [\![v]\!] = 0.$$

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 [22] through jumps

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

We note that more accurate absorbing conditions can be imposed using, for example, perfectly matched layers [24] or high order absorbing boundary conditions [25, 26].

In all cases, boundary conditions are imposed by computing numerical fluxes using these modified jumps. This imposition guarantees energy stability for free surface, non-reflective, and homogeneous velocity boundary conditions.

3.2. Energy stability. Integrating the velocity equations of (2) by parts gives

$$\sum_{D^k \in \Omega_h} \left(\rho\frac{\partial v}{\partial t}, w\right)_{L^2(D^k)} = \sum_{D^k \in \Omega_h} -\left(\sum_{i=1}^d \boldsymbol{\sigma}, A_i\frac{\partial w}{\partial x_i}\right)_{L^2(D^k)} + \left\langle A_n^T\{\!\{\boldsymbol{\sigma}\}\!\} + \frac{\tau_v}{2}A_n^TA_n[\![v]\!], w\right\rangle_{L^2(\partial D^k)}$$

$$\sum_{D^k \in \Omega_h} \left(C^{-1}\frac{\partial\boldsymbol{\sigma}}{\partial t}, q\right)_{L^2(D^k)} = \sum_{D^k \in \Omega_h} \left(\sum_{i=1}^d A_i\frac{\partial v}{\partial x_i}, q\right)_{L^2(D^k)} + \left\langle \frac{1}{2}A_n[\![v]\!] + \frac{\tau_\sigma}{2}A_nA_n^T[\![\boldsymbol{\sigma}]\!], q\right\rangle_{L^2(\partial D^k)}$$

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

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

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

$$= \sum_{D^k \in \Omega_h}\sum_{f\in\partial D^k}\int_f\left(A_n^T\{\!\{\boldsymbol{\sigma}\}\!\}v + \frac{\tau_v}{2}A_n^TA_n[\![v]\!]v + \frac{1}{2}A_n[\![v]\!]\boldsymbol{\sigma} + \frac{\tau_\sigma}{2}A_nA_n^T[\![\boldsymbol{\sigma}]\!]\boldsymbol{\sigma}\right)\,\mathrm{d}x.$$

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 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

$$\sum_{f\in\Gamma_h\backslash\partial\Omega}\int_f\left(A_n^T\{\!\{\boldsymbol{\sigma}\}\!\}v + \frac{\tau_v}{2}A_n^TA_n[\![v]\!]v + \frac{1}{2}A_n[\![v]\!]\boldsymbol{\sigma} + \frac{\tau_\sigma}{2}A_nA_n^T[\![\boldsymbol{\sigma}]\!]\boldsymbol{\sigma}\right)\,\mathrm{d}x$$

$$= -\sum_{f\in\Gamma_h\backslash\partial\Omega}\int_f\left(\frac{\tau_v}{2}\left|A_n[\![v]\!]\right|^2 + \frac{\tau_\sigma}{2}\left|A_n^T[\![\boldsymbol{\sigma}]\!]\right|^2\right)\,\mathrm{d}x.$$

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

$$\sum_{f\in\Gamma_v}\int_f\left(A_n^T\{\!\{\boldsymbol{\sigma}\}\!\}v + \frac{\tau_v}{2}A_n^TA_n[\![v]\!]v + \frac{1}{2}A_n[\![v]\!]\boldsymbol{\sigma} + \frac{\tau_\sigma}{2}A_nA_n^T[\![\boldsymbol{\sigma}]\!]\boldsymbol{\sigma}\right)\,\mathrm{d}x$$

$$= \sum_{f\in\Gamma_v}\int_f\left(A_n^T\boldsymbol{\sigma}^-v^- - A_nv^-\boldsymbol{\sigma}^- - \tau_vA_n^TA_n\left|v^-\right|^2\right)\,\mathrm{d}x = -\sum_{f\in\Gamma_v}\int_f\left(\tau_vA_n^TA_n\left|v^-\right|^2\right)\,\mathrm{d}x.$$

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

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

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

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

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*

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

(3)

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

Since ρ and $\boldsymbol{C}^{-1}$ are positive definite, the left hand side of (3) is an L^2-equivalent norm on $(\boldsymbol{v},\boldsymbol{\sigma})$, and Theorem 1 implies that the magnitude of the DG solution $(\boldsymbol{v},\boldsymbol{\sigma})$ 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 [20] 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 requirements 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

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

(4)

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

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

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

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

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

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

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 $\boldsymbol{F}_v$ and $\boldsymbol{F}_\sigma$ denote degrees of freedom for velocity and stress numerical fluxes.

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 system at every time-step. 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 $\rho, \boldsymbol{C}$ 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 $\rho, \boldsymbol{C}$, and J over each element. Applying this observation to (5) then yields the following local 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 $\boldsymbol{C}$ 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. For isotropic media, one way to address sub-element variations in material parameters is to diagonalize the matrix $\boldsymbol{C}$ through a change of variables [27]. This results in a local system of ODEs with only scalar weighted mass matrices [16], which can be treated using scalar weight-adjusted approximations. We take a different approach in addressing these issues and approximate the matrix-weighted L^2 inner product (and corresponding matrix-weighted mass matrix $\boldsymbol{M}_{\boldsymbol{C}^{-1}}$) using a weight-adjusted approximation which is low storage, simple to invert, energy stable, and provably high order accurate for spatially varying weights $\rho, \boldsymbol{C}$ 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 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

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

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

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

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

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

$$D^\alpha u = \frac{\partial^{\alpha_1}}{\partial x^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial y^{\alpha_2}} u, \qquad d = 2$$

$$D^\alpha u = \frac{\partial^{\alpha_1}}{\partial x^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial y^{\alpha_2}} \frac{\partial^{\alpha_3}}{\partial z^{\alpha_3}} u, \qquad d = 3.$$

Let Π_N denote the L^2 projection on the reference element $\widehat{D}$. We define two operators $T_w : L^2(D^k) \to P^N(D^k)$ and $T_w^{-1} : L^2(D^k) \to P^N(D^k)$ such that

$$T_w u = \Pi_N(wu)$$
$$\left(w T_w^{-1} u, v \right)_{L^2(D^k)} = (u, v)_{L^2(D^k)}, \qquad \forall v \in P^N(D^k).$$

8

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

$$(wu, v)_{L^2(D^k)} = (T_w u, v)_{L^2(D^k)} \approx \left(T_{1/w}^{-1} u, v \right)_{L^2(D^k)}.$$

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

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

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

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

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

This theorem relies on a scalar weighted interpolation estimate derived in [18, 1] for a general non-affine element D^k.

THEOREM 3 (Theorem 4 in [1]). *Let D^k be a quasi-regular element with representative size $h = \text{diam}\left(D^k\right)$. For $N \geq 0$, $w \in W^{N+1,\infty}\left(D^k\right)$, and $u \in W^{N+1,2}\left(D^k\right)$,*

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

4.2. Extension to matrix weights. We now 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

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

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

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

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

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

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

where $D^{\boldsymbol{\alpha}}\boldsymbol{W}(\boldsymbol{x})$ again denotes component-wise differentiation. While this norm is not sub-multiplicative, the following bound holds

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

9

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

The following theorem extends Theorem 3 to matrix weights by computing weighted interpolation estimates for the quantity $\boldsymbol{W}^{-1}\Pi_N(\boldsymbol{W}\boldsymbol{v})$.

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

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

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

$$\begin{aligned}
\left\|\boldsymbol{v} - \boldsymbol{W}^{-1}\Pi_N\left(\boldsymbol{W}\boldsymbol{v}\right)\right\|_{L^2(D^k)} &\leq C_1 \left\|\sqrt{J}\right\|_{L^\infty(D^k)} \sup_{\boldsymbol{x}} \left\|\boldsymbol{W}^{-1}\right\|_2 \|\boldsymbol{W}\boldsymbol{v} - \Pi_N\left(\boldsymbol{W}\boldsymbol{v}\right)\|_{L^2(D^k)} \\
&\leq C_1 \left\|\sqrt{J}\right\|_{L^\infty(D^k)} \sup_{\boldsymbol{x}} \left\|\boldsymbol{W}^{-1}\right\|_2 |\boldsymbol{W}\boldsymbol{v}|_{W^{N+1,2}(\hat{D})} \\
&\leq C_2 h^{N+1} \left\|\sqrt{J}\right\|_{L^\infty(D^k)} \left\|\frac{1}{\sqrt{J}}\right\|_{L^\infty(D^k)} \sup_{\boldsymbol{x}} \left\|\boldsymbol{W}^{-1}\right\|_2 \|\boldsymbol{W}\boldsymbol{v}\|_{W^{N+1,2}(D^k)} \\
&\leq C_3 h^{N+1} \left\|\sqrt{J}\right\|_{L^\infty(D^k)} \left\|\frac{1}{\sqrt{J}}\right\|_{L^\infty(D^k)} \tilde{w}_{\max} \|\boldsymbol{W}\|_{N+1,2,\infty} \|\boldsymbol{v}\|_{W^{N+1,2}(D^k)}. \quad \square
\end{aligned}$$

4.2.1. Weight-adjusted approximations with matrix weights. 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

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

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

$$\left(\boldsymbol{W}T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{w}\right)_{L^2(D^k)} = \left(\boldsymbol{v}, \boldsymbol{w}\right)_{L^2(D^k)}.$$

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 $w(\boldsymbol{x})$ [1] carry over to matrix weights $\boldsymbol{W}(\boldsymbol{x})$ as well.

LEMMA 5. *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. *$\left\|T_{\boldsymbol{W}}^{-1}\right\|_{L^2(D^k)} \leq \tilde{w}_{\max}$.*

4. *$\left(T_{\boldsymbol{W}}^{-1}\boldsymbol{v}, \boldsymbol{w}\right)_{L^2(D^k)}$ 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,

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

Property 4 then simply requires the lower bound

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

Using Theorem 4, we may also show that the matrix-valued weight-adjusted inner product is also high order accurate for sufficiently regular $\boldsymbol{W}$.

10

THEOREM 6. *Let D^k be a quasi-regular element with representative size $h = \mathrm{diam}\left(D^k\right)$. For $N > 0$*

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

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

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

Proof. The proof follows the scalar case. The triangle inequality gives

$$\left\|\boldsymbol{W}\boldsymbol{v} - T_{\boldsymbol{W}^{-1}}^{-1}\boldsymbol{v}\right\|_{L^2(D^k)} \leq \left\|\boldsymbol{W}\boldsymbol{v} - \Pi_N(\boldsymbol{W}\boldsymbol{v})\right\|_{L^2(D^k)} + \left\|\Pi_N(\boldsymbol{W}\boldsymbol{v}) - T_{\boldsymbol{W}^{-1}}^{-1}\boldsymbol{v}\right\|_{L^2(D^k)}.$$

The former term is bounded by interpolation estimates and by arguments used in the proof of Theorem 4. The latter term is bounded

$$
\begin{aligned}
\left\|\Pi_N(\boldsymbol{W}\boldsymbol{v}) - T_{\boldsymbol{W}^{-1}}^{-1}\boldsymbol{v}\right\|_{L^2(D^k)} &= \left\|T_{\boldsymbol{W}^{-1}}^{-1} T_{\boldsymbol{W}^{-1}} \Pi_N(\boldsymbol{W}\boldsymbol{v}) - T_{\boldsymbol{W}^{-1}}^{-1} \Pi_N \boldsymbol{v}\right\|_{L^2(D^k)} \\
&= \left\|T_{\boldsymbol{W}^{-1}}^{-1} \Pi_N \left(T_{\boldsymbol{W}^{-1}} \Pi_N(\boldsymbol{W}\boldsymbol{v})\right) - T_{\boldsymbol{W}^{-1}}^{-1} \Pi_N \boldsymbol{v}\right\|_{L^2(D^k)} \\
&\leq C_N \left\|T_{\boldsymbol{W}^{-1}}^{-1}\right\| \left\|T_{\boldsymbol{W}^{-1}} \Pi_N(\boldsymbol{W}\boldsymbol{v}) - \Pi_N \boldsymbol{v}\right\|_{L^2(D^k)} \\
&\leq C_N \left\|T_{\boldsymbol{W}^{-1}}^{-1}\right\| \left\|\Pi_N \left(\boldsymbol{W}^{-1} \Pi_N(\boldsymbol{W}\boldsymbol{v})\right) - \Pi_N \boldsymbol{v}\right\|_{L^2(D^k)} \\
&\leq C_N w_{\max} \left\|\boldsymbol{W}^{-1} \Pi_N(\boldsymbol{W}\boldsymbol{v}) - \boldsymbol{v}\right\|_{L^2(D^k)}.
\end{aligned}
$$

where we have used $\left\|T_{\boldsymbol{W}^{-1}}^{-1}\right\| \leq w_{\max}$ (Lemma 5) and $\left\|\Pi_N\right\|_{L^2(D^k)} = 1$. An application of Theorem 4 to bound $\left\|\boldsymbol{W}^{-1} \Pi_N(\boldsymbol{W}\boldsymbol{v}) - \boldsymbol{v}\right\|_{L^2(D^k)}$ completes the proof. □

We note that these estimates are tight, in the sense that estimates for scalar weight-adjusted inner products are recovered when the matrix weight is taken to be $\boldsymbol{W} = w\boldsymbol{I}$.

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

$$\boldsymbol{M}_w \approx \boldsymbol{M}\boldsymbol{M}_{1/w}^{-1}\boldsymbol{M}, \qquad \boldsymbol{M}_w^{-1} \approx \boldsymbol{M}^{-1}\boldsymbol{M}_{1/w}\boldsymbol{M}^{-1}.$$

By evaluating $\boldsymbol{M}_{1/w}$ in a matrix-free fashion (for example, using a sufficiently accurate quadrature rule), the inverse of the weight-adjusted mass matrix $\boldsymbol{M}^{-1}\boldsymbol{M}_{1/w}\boldsymbol{M}^{-1}$ 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

(8)
$$
\begin{aligned}
\boldsymbol{M}_{\boldsymbol{W}} &\approx \left(\boldsymbol{I} \otimes \boldsymbol{M}\right) \boldsymbol{M}_{\boldsymbol{W}^{-1}}^{-1} \left(\boldsymbol{I} \otimes \boldsymbol{M}\right) \\
\boldsymbol{M}_{\boldsymbol{W}}^{-1} &\approx \left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right) \boldsymbol{M}_{\boldsymbol{W}^{-1}} \left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right).
\end{aligned}
$$

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

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

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

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

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

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

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 1

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

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

The use of weight-adjusted 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 local system of ODEs for $\boldsymbol{V}, \boldsymbol{\Sigma}$

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

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

In practice, the matrices $\left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right)\boldsymbol{M}_{(\rho^{-1}\boldsymbol{I})}$ and $\left(\boldsymbol{I} \otimes \boldsymbol{M}^{-1}\right)\boldsymbol{M}_{\boldsymbol{C}}$ are applied in a matrix-free fashion using reference element matrices and values of $\rho, \boldsymbol{C}$ at quadrature points. After fusing operations together, this procedure can be boiled down to multiplication by two rectangular matrices [2].

One drawback of the analysis presented in this work is that accuracy of the weight-adjusted approximation is not guaranteed in the incompressible limit $\mu/\lambda \to 0$. In this case, the stiffness matrix $\boldsymbol{C}$ becomes singular, and the constant $\tilde{w}_{\max}$ in the upper bound on $\left\|\boldsymbol{C}^{-1}\right\|$ blows up. However, numerical experiments in Section 6.2.3 suggest that, while taking $\mu/\lambda \approx 0$ (or $\boldsymbol{C}$ nearly singular) results in larger relative errors for $\boldsymbol{\sigma}$, the accuracy of the WADG solution for $\boldsymbol{v}$ does not degrade significantly for near-incompressible materials.

5.1. Energy stability on 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 (local) formulation

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

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

---

[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.

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

The proof of energy stability in Theorem 1 requires only algebraic manipulations of this formulation, and does not require that integration-by-parts holds discretely. This implies that a discrete version of energy stability is still guaranteed in the presence of inexact quadrature, where integrals in Theorem 1 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 [28, 29].

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 most 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. Using estimates from Section 4.2.1, we can extend the semi-discrete convergence analysis in [30, 18, 1] to linear elastic wave propagation on meshes of affine elements. Techniques in [18] can be used to extend this analysis to curvilinear elements.

Let $\boldsymbol{U}, \boldsymbol{U}_h$ denote the exact and discrete WADG solutions, respectively. We will assume that $\boldsymbol{U}, \frac{\partial \boldsymbol{U}}{\partial t}$ are sufficiently regular such that

$$\|\boldsymbol{U}\|_{W^{N+1,2}(\Omega_h)} < \infty, \qquad \left\|\frac{\partial \boldsymbol{U}}{\partial t}\right\|_{W^{N+1,2}(\Omega_h)} < \infty,$$

where we define $\|\boldsymbol{U}\|^2_{W^{N+1,2}(\Omega_h)} = \sum_k \|\boldsymbol{U}\|^2_{W^{N+1,2}(D^k)}$.

In terms of group variables $\boldsymbol{U} = (\boldsymbol{v}, \boldsymbol{\sigma})$ and $\boldsymbol{V} = (\boldsymbol{w}, \boldsymbol{q}) \in (V_h(\Omega_h))^d \times (V_h(\Omega_h))^{N_d}$, the WADG formulation can be written as

$$\left(T^{-1}_{\boldsymbol{A}_0^{-1}}\frac{\partial \boldsymbol{U}}{\partial t}, \boldsymbol{V}\right)_{L^2(\Omega)} + a(\boldsymbol{U}, \boldsymbol{V}) + b(\boldsymbol{U}, \boldsymbol{V}) = (\boldsymbol{f}, \boldsymbol{V})$$

$$a(\boldsymbol{U}, \boldsymbol{V}) = \sum_{D^k \in \Omega_h} \left(-\left(\sum_{i=1}^d \boldsymbol{\sigma}, \boldsymbol{A}_i\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{x}_i}\right)_{L^2(D^k)} + \left(\sum_{i=1}^d \boldsymbol{A}_i\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}_i}, \boldsymbol{q}\right)_{L^2(D^k)}\right)$$

$$b(\boldsymbol{U}, \boldsymbol{V}) = \sum_{D^k \in \Omega_h} \left(\left\langle \boldsymbol{A}_n^T\{\!\{\boldsymbol{\sigma}\}\!\} + \frac{\tau_v}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n[\![\boldsymbol{v}]\!], \boldsymbol{w}\right\rangle_{L^2(\partial D^k)} + \left\langle \frac{1}{2}\boldsymbol{A}_n[\![\boldsymbol{v}]\!] + \frac{\tau_\sigma}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T[\![\boldsymbol{\sigma}]\!], \boldsymbol{q}\right\rangle_{L^2(\partial D^k)}\right),$$

where $\boldsymbol{A}_0(\boldsymbol{x}) = \text{diag}\left(\rho(\boldsymbol{x})\boldsymbol{I}^{d \times d}, \boldsymbol{C}^{-1}(\boldsymbol{x})\right)$. Energy stability implies that

$$b(\boldsymbol{U}, \boldsymbol{U}) = \sum_{D^k \in \Omega_h} \left(\frac{\tau_v}{2}\|[\![\boldsymbol{A}_n\boldsymbol{v}]\!]\|^2_{L^2(\partial D^k)} + \frac{\tau_\sigma}{2}\|[\![\boldsymbol{A}_n^T\boldsymbol{\sigma}]\!]\|^2_{L^2(\partial D^k)}\right)$$

$$\left(T^{-1}_{\boldsymbol{A}_0^{-1}}\frac{\partial \boldsymbol{U}}{\partial t}, \boldsymbol{U}\right)_{L^2(\Omega)} + b(\boldsymbol{U}, \boldsymbol{U}) = (\boldsymbol{f}, \boldsymbol{U}).$$

Since the DG formulation (2) is consistent, these solutions satisfy

$$\left(\boldsymbol{A}_0\frac{\partial \boldsymbol{U}}{\partial t}, \boldsymbol{V}\right)_{L^2(\Omega)} + a(\boldsymbol{U}, \boldsymbol{V}) + b(\boldsymbol{U}, \boldsymbol{V}) = (\boldsymbol{f}, \boldsymbol{V})$$

(9)
$$\left(T^{-1}_{\boldsymbol{A}_0^{-1}}\frac{\partial \boldsymbol{U}_h}{\partial t}, \boldsymbol{V}\right)_{L^2(\Omega)} + a(\boldsymbol{U}_h, \boldsymbol{V}) + b(\boldsymbol{U}_h, \boldsymbol{V}) = (\boldsymbol{f}, \boldsymbol{V})$$

for all $\boldsymbol{V} \in (V_h(\Omega_h))^d \times (V_h(\Omega_h))^{N_d}$. We decompose the error $\boldsymbol{U} - \boldsymbol{U}_h$ into a projection error $\boldsymbol{\epsilon}$ and discretization error $\boldsymbol{\eta}$.

$$\boldsymbol{U} - \boldsymbol{U}_h = (\Pi_N\boldsymbol{U} - \boldsymbol{U}_h) - (\Pi_N\boldsymbol{U} - \boldsymbol{U}) = \boldsymbol{\eta} - \boldsymbol{\epsilon}.$$

13

We assume that $\boldsymbol{U}_h(\boldsymbol{x}, 0)$ is the L^2 projection of the exact initial condition, such that $\boldsymbol{\eta}|_{t=0} = 0$. We also introduce a consistency error $\boldsymbol{\delta} = \boldsymbol{A}_0\boldsymbol{U} - T_{\boldsymbol{A}_0^{-1}}^{-1}\boldsymbol{U}$ resulting from the approximation of $\boldsymbol{A}_0\boldsymbol{U}$ by a weight-adjusted inner product

$$\boldsymbol{A}_0\frac{\partial \boldsymbol{U}}{\partial t} - T_{\boldsymbol{A}_0^{-1}}^{-1}\frac{\partial \boldsymbol{U}_h}{\partial t} = \frac{\partial}{\partial t}\left(\boldsymbol{A}_0\boldsymbol{U} - T_{\boldsymbol{A}_0^{-1}}^{-1}\boldsymbol{U}\right) + \frac{\partial}{\partial t}T_{\boldsymbol{A}_0^{-1}}^{-1}\left(\Pi_N\boldsymbol{U} - \boldsymbol{U}_h\right) = \frac{\partial \boldsymbol{\delta}}{\partial t} + \frac{\partial}{\partial t}\left(T_{\boldsymbol{A}_0^{-1}}^{-1}\boldsymbol{\eta}\right)$$

where we have used that $\boldsymbol{T}_{\boldsymbol{A}_0^{-1}}^{-1} = \boldsymbol{T}_{\boldsymbol{A}_0^{-1}}^{-1}\Pi_N$. Subtracting the DG and WADG formulations in (9) and setting $\boldsymbol{V} = \boldsymbol{\eta}$ then yields

$$(10) \qquad \frac{1}{2}\frac{\partial}{\partial t}\left(T_{\boldsymbol{A}_0^{-1}}^{-1}\boldsymbol{\eta}, \boldsymbol{\eta}\right)_{L^2(\Omega)} + b(\boldsymbol{\eta}, \boldsymbol{\eta}) = \left(\frac{\partial \boldsymbol{\delta}}{\partial t}, \boldsymbol{\eta}\right)_{L^2(\Omega)} + a(\boldsymbol{\epsilon}, \boldsymbol{\eta}) + b(\boldsymbol{\epsilon}, \boldsymbol{\eta}),$$

where we have used that $a(\boldsymbol{\eta}, \boldsymbol{\eta}) = 0$ by skew symmetry.

We bound $a(\boldsymbol{\epsilon}, \boldsymbol{\eta}) + b(\boldsymbol{\epsilon}, \boldsymbol{\eta})$ on the right hand side by integrating by parts the stress equation and using the component-wise L^2 orthogonality of $\boldsymbol{\epsilon}$ to derivatives of $\boldsymbol{\eta}$. This reduces the term $a(\boldsymbol{\epsilon}, \boldsymbol{\eta}) + b(\boldsymbol{v}, \boldsymbol{\eta})$ to surface contributions over each element, which can be combined with contributions from neighboring elements to yield

$$\sum_{D^k \in \Omega_h}\left(\left\langle \boldsymbol{A}_n^T\{\!\{\boldsymbol{\epsilon}_\sigma\}\!\} + \frac{\tau_v}{2}\boldsymbol{A}_n^T\boldsymbol{A}_n[\![\boldsymbol{\epsilon}_v]\!], \boldsymbol{\eta}_v\right\rangle_{L^2(\partial D^k)} + \left\langle \boldsymbol{A}_n\{\!\{\boldsymbol{\epsilon}_v\}\!\} + \frac{\tau_\sigma}{2}\boldsymbol{A}_n\boldsymbol{A}_n^T[\![\boldsymbol{\epsilon}_\sigma]\!], \boldsymbol{\eta}_\sigma\right\rangle_{L^2(\partial D^k)}\right)$$

$$= \frac{1}{2}\sum_{D^k \in \Omega_h}\left(\left\langle \{\!\{\boldsymbol{\epsilon}_\sigma\}\!\} - \frac{\tau_v}{2}\boldsymbol{A}_n[\![\boldsymbol{\epsilon}_v]\!], \boldsymbol{A}_n[\![\boldsymbol{\eta}_v]\!]\right\rangle_{L^2(\partial D^k)} + \left\langle \{\!\{\boldsymbol{\epsilon}_v\}\!\} - \frac{\tau_\sigma}{2}\boldsymbol{A}_n^T[\![\boldsymbol{\epsilon}_\sigma]\!], \boldsymbol{A}_n^T[\![\boldsymbol{\eta}_\sigma]\!]\right\rangle_{L^2(\partial D^k)}\right)$$

$$\leq \frac{1}{2}\sum_{D^k \in \Omega_h}\left(\left\|\{\!\{\boldsymbol{\epsilon}_\sigma\}\!\} - \frac{\tau_v}{2}\boldsymbol{A}_n[\![\boldsymbol{\epsilon}_v]\!]\right\|_{L^2(\partial D^k)}\|\boldsymbol{A}_n[\![\boldsymbol{\eta}_v]\!]\|_{L^2(\partial D^k)} + \left\|\{\!\{\boldsymbol{\epsilon}_v\}\!\} - \frac{\tau_\sigma}{2}\boldsymbol{A}_n^T[\![\boldsymbol{\epsilon}_\sigma]\!]\right\|_{L^2(\partial D^k)}\|\boldsymbol{A}_n^T[\![\boldsymbol{\eta}_\sigma]\!]\|_{L^2(\partial D^k)}\right)$$

$$\leq C_\tau\sum_{D^k \in \Omega_h}\|\boldsymbol{\epsilon}\|_{L^2(\partial D^k)}\left(\frac{\tau_v}{2}\|\boldsymbol{A}_n[\![\boldsymbol{\eta}_v]\!]\|_{L^2(\partial D^k)}^2 + \frac{\tau_\sigma}{2}\|\boldsymbol{A}_n^T[\![\boldsymbol{\eta}_\sigma]\!]\|_{L^2(\partial D^k)}^2\right)^{1/2}$$

where C_τ is proportional to $\max(\tau_v, \tau_\sigma)$. Using Young's inequality with $\alpha = C_\tau/2$ yields the following bound

$$|a(\boldsymbol{\epsilon}, \boldsymbol{\eta}) + b(\boldsymbol{\epsilon}, \boldsymbol{\eta})| \leq b(\boldsymbol{\eta}, \boldsymbol{\eta}) + \frac{C_\tau^2}{4}\sum_{D^k \in \Omega_h}\|\boldsymbol{\epsilon}\|_{L^2(\partial D^k)}^2.$$

Applying this to (10) and using Cauchy-Schwarz on $\left(\frac{\partial \boldsymbol{\delta}}{\partial t}, \boldsymbol{\eta}\right)_{L^2(\Omega)}$ then yields

$$\frac{1}{2}\frac{\partial}{\partial t}\left(T_{\boldsymbol{A}_0^{-1}}^{-1}\boldsymbol{\eta}, \boldsymbol{\eta}\right)_{L^2(\Omega)} + b(\boldsymbol{\eta}, \boldsymbol{\eta}) \leq \left\|\frac{\partial \boldsymbol{\delta}}{\partial t}\right\|_{L^2(\Omega)}\|\boldsymbol{\eta}\|_{L^2(\Omega)} + b(\boldsymbol{\eta}, \boldsymbol{\eta}) + \sum_{D^k \in \Omega_h}\frac{C_\tau^2}{4}\|\boldsymbol{\epsilon}\|_{L^2(\partial D^k)}^2.$$

We eliminate factors of $b(\boldsymbol{\eta}, \boldsymbol{\eta})$ on both sides and bound right hand side terms. The trace term is bounded using a standard hp trace inequality [31] and an interpolation estimate

$$\sum_{D^k \in \Omega_h}\|\boldsymbol{\epsilon}\|_{L^2(\partial D^k)}^2 \leq \sum_{D^k \in \Omega_h}Ch^{-1}\|\boldsymbol{\epsilon}\|_{L^2(D^k)}^2 = Ch^{-1}\|\boldsymbol{\epsilon}\|_{L^2(\Omega)}^2 \leq Ch^{2N+1}\|\boldsymbol{U}\|_{W^{N+1,2}}^2.$$

Since $\boldsymbol{A}_0$ is independent of t, Theorem 6 gives

$$\left\|\frac{\partial \boldsymbol{\delta}}{\partial t}\right\|_{L^2(\Omega)} = \left\|\boldsymbol{A}_0\frac{\partial \boldsymbol{U}}{\partial t} - T_{\boldsymbol{A}_0^{-1}}^{-1}\frac{\partial \boldsymbol{U}}{\partial t}\right\|_{L^2(\Omega)} \leq Ch^{N+1}A_{\max}\|\boldsymbol{A}_0\|_{W^{N+1,\infty}(\Omega_h)}\left\|\frac{\partial \boldsymbol{U}}{\partial t}\right\|_{L^2(\Omega)}.$$

where $A_{\max} = \max\left(\frac{\rho_{\max}}{\rho_{\min}}, c_{\max}\tilde{c}_{\max}\right)$. Then, integrating from $[0, T]$ and using Lemma 5 yields

$$(11) \quad A_{\min}\|\boldsymbol{\eta}\|_{L^2(\Omega)}^2 \leq C\int_0^T h^{N+1}A_{\max}\|\boldsymbol{A}_0\|_{W^{N+1,\infty}(\Omega_h)}\left\|\frac{\partial \boldsymbol{U}}{\partial t}\right\|_{L^2(\Omega)}\|\boldsymbol{\eta}\|_{L^2(\Omega)} + h^{2N+1}\|\boldsymbol{U}\|_{W^{N+1,2}(\Omega)}^2\,\mathrm{d}t.$$

where $A_{\min} = \min(\rho_{\min}, \tilde{c}_{\min})$. Applying the modified Gronwall's inequality (Lemma 1.10 in [32]; see also [1]) to (11) then yields

$$\|\boldsymbol{\eta}\|_{L^2(\Omega)} \leq \frac{1}{A_{\min}} \int_0^T C A_{\max} h^{N+1} \|\boldsymbol{A}_0\|_{W^{N+1,\infty}(\Omega_h)} \left\|\frac{\partial \boldsymbol{U}}{\partial t}\right\|_{L^2(\Omega)} \mathrm{d}t + \sup_{t\in[0,T]} \sqrt{\int_0^T C h^{2N+1} \|\boldsymbol{U}\|_{W^{N+1,2}(\Omega)}^2 \, \mathrm{d}t}$$

$$\leq \frac{CTh^{N+1/2}}{A_{\min}} \sup_{t\in[0,T]} \left(h^{1/2} A_{\max} \|\boldsymbol{A}_0\|_{W^{N+1,\infty}(\Omega_h)} \left\|\frac{\partial \boldsymbol{U}}{\partial t}\right\|_{L^2(\Omega)} + \|\boldsymbol{U}\|_{W^{N+1,2}(\Omega_h)} \right).$$

The triangle inequality gives the final estimate

$$\|\boldsymbol{U} - \boldsymbol{U}_h\| \leq (C_1 + C_2 T) h^{N+1/2} \sup_{t\in[0,T]} \left(h^{1/2} \|\boldsymbol{A}_0\|_{W^{N+1,\infty}(\Omega_h)} \left\|\frac{\partial \boldsymbol{U}}{\partial t}\right\|_{W^{N+1,2}(\Omega_h)} + \|\boldsymbol{U}\|_{W^{N+1,2}(\Omega_h)} \right).$$

where C_2 depends on $A_{\min}, A_{\max}$.

6. Numerical experiments. The following sections present several numerical experiments validating the stability and accuracy of the proposed method in two and three dimensions. The energy stability of the method is also confirmed for examples with sub-cell variations in heterogeneous media and curvilinear meshes. The convergence of the new DG formulation in piecewise constant isotropic media is confirmed using analytic solutions, while the convergence of the method for high order approximations of heterogeneous media is confirmed using a fine grid reference solution. Finally, the method is applied to problems with anisotropy and stiffness matrices $\boldsymbol{C}$ with sub-element variations.

In all experiments, we follow [1] and compute application of weight-adjusted mass matrices using a quadrature exact for polynomials of degree $(2N+1)$ [33]. Time integration is performed using the low-storage 4th order five-stage Runge-Kutta scheme of Carpenter and Kennedy [34], and the time-step is chosen based on the global restriction

$$dt = \min_k \frac{C_{\mathrm{CFL}}}{\sup_{\boldsymbol{x}\in\Omega} \|\boldsymbol{C}(\boldsymbol{x})\|_2 C_N \|J^f\|_{L^\infty(\partial D^k)} \|J^{-1}\|_{L^\infty(D^k)}}$$

where $C_N = O(N^2)$ is the order-dependent constant in the surface polynomial trace inequality [14] and C_{CFL} is a tunable global CFL constant. Finally, in all following experiments, we use $\tau_v = \tau_\sigma = 1$ unless specified otherwise.

6.1. Spectra and choice of penalty parameter. We first verify the energy stability of the WADG method for arbitrary heterogeneous media. We use a stiffness matrix constructed using similarity transforms, such that at every quadrature point, $\boldsymbol{C}(\boldsymbol{x}) = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^T$, where $\boldsymbol{D}$ is a diagonal matrix with random positive entries $d_{\min} \leq \boldsymbol{D}_{ii} \leq d_{\max}$ and $\boldsymbol{U}$ is a random unitary matrix. Let $\boldsymbol{A}_h$ denote the matrix induced by the global semi-discrete DG formulation, such that the time evolution of the solution $\boldsymbol{v}, \boldsymbol{\sigma}$ is governed by

$$\frac{\partial \boldsymbol{Q}}{\partial t} = \boldsymbol{A}_h \boldsymbol{Q},$$

with $\boldsymbol{Q}$ denotes a vector of degrees of freedom for $(\boldsymbol{v}, \boldsymbol{\sigma})$. Figure 1 shows computed eigenvalues of $\boldsymbol{A}_h$ for $[d_{\min}, d_{\max}] = [1/10, 1]$ and $[10^{-4}, 1]$, $\tau_v = \tau_\sigma = 0$ and $\tau_v = \tau_\sigma = 1$ under discretization parameters $N = 4$ and $h = 1/4$. In both cases, the largest real part of any eigenvalue is $O(10^{-14})$, verifying the energy stability of the WADG discretization for arbitrary media.

For practical simulations, the choice of τ_v, τ_σ remains to be specified. Increasing τ_v, τ_σ results in greater damping of under-resolved spurious components of the solution [20]; however, a naive selection of these penalty parameters can result in an overly restrictive time-step restriction for stability. We wish to choose τ_v, τ_σ as large as possible without increasing $\|\boldsymbol{A}_h\|$.

To avoid artificially increasing the stiffness of (5), it was argued in [20] that the magnitude of the penalty parameter should depend on the largest singular value of $\boldsymbol{A}_n$, which is $O(1)$ in two and three dimensions. However, this suggestion does not factor in the effect of ρ and $\boldsymbol{C}^{-1}$ appearing in the left hand side. These
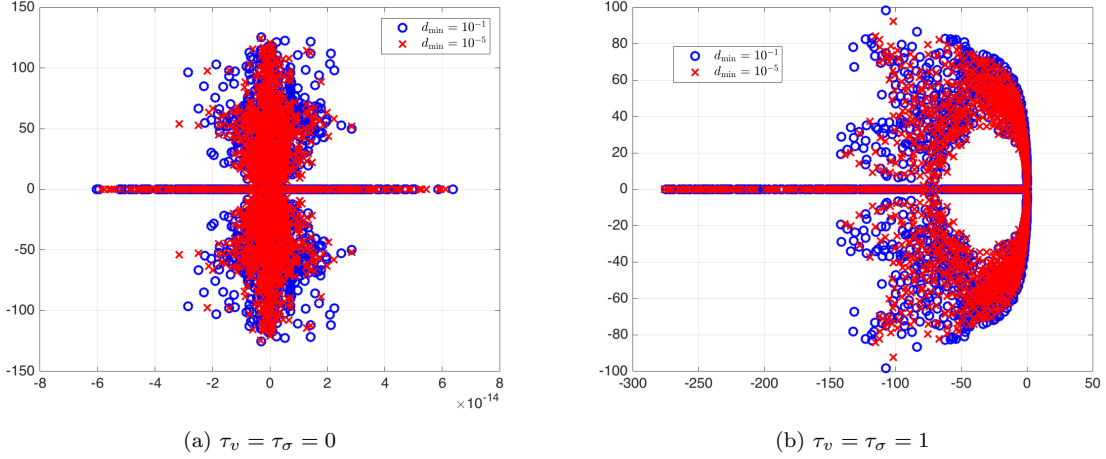
(a) $\tau_v = \tau_\sigma = 0$ (b) $\tau_v = \tau_\sigma = 1$

Fig. 1: Spectra for $N = 4$ and $h = 1/4$, using randomly chosen $\boldsymbol{C}(\boldsymbol{x})$ with eigenvalues between $[d_{\min}, 1]$ at each quadrature point. For both $\tau_v = \tau_\sigma = 0$ and $\tau_v = \tau_\sigma = 1$, the largest real part of the spectra is $O(10^{-14})$.

material coefficients can be taken into account by scaling the penalty parameters such that flux terms are dimensionally consistent. One example of such a scaling is

$$\tau_v = \gamma_v \sup_{\boldsymbol{x} \in f} \sqrt{\|\{\{\boldsymbol{C}(\boldsymbol{x})\}\}\| \{\{\rho(\boldsymbol{x})\}\}}, \qquad \tau_\sigma = \gamma_\sigma \sup_{\boldsymbol{x} \in f} \frac{1}{\sqrt{\|\{\{\boldsymbol{C}(\boldsymbol{x})\}\}\| \{\{\rho(\boldsymbol{x})\}\}}},$$

where the supremum is taken locally over each face and γ_v, γ_σ are dimensionless constants.

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 is expressed terms of the displacement vector $\boldsymbol{u}(\boldsymbol{x}, t) = (u_1, \ldots, u_d)$. Initial conditions for velocity and stress are computed through

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

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 components of the displacement vector are 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 $\omega = \sqrt{2\mu}$. Zero traction boundary conditions are imposed. Figure 2 shows L^2 errors computed at time $T = 5$, using uniform triangular meshes constructed by bisecting a uniform mesh of quadrilaterals along the diagonal. For $N = 1, \ldots, 5$, rates near the optimal $O(h^{N+1})$ rates of convergence are observed.
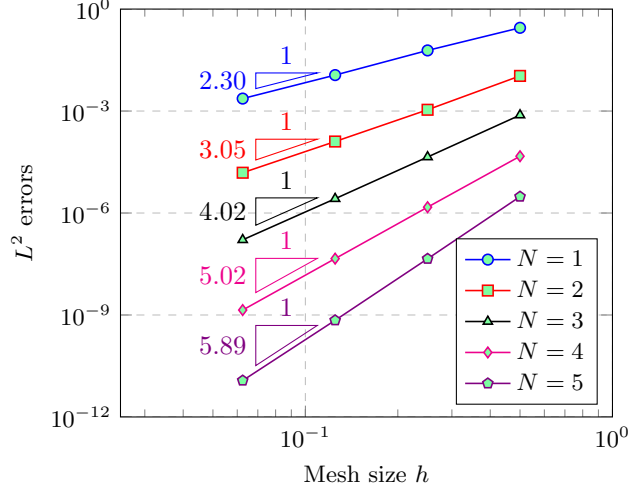
16

Fig. 2: Convergence of L^2 errors for harmonic oscillation.

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

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

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 $\omega = 2\pi$ [23]. We solve on the domain $[0, 2] \times [0, 1]$ using a sequence of uniform triangular meshes, and enforce traction-free boundary conditions at $x = 0$ 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 is given by

$$u_1\left(x, y, t\right) = (-kB_1 \cos(py) - qB_2 \cos(qy)) \sin(kx - \omega t)$$
$$u_2\left(x, y, t\right) = (-pB_1 \sin(py) + kB_2 \sin(qy)) \cos(kx - \omega t)$$

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, $\omega = 13.137063197233$, $B_1 =$
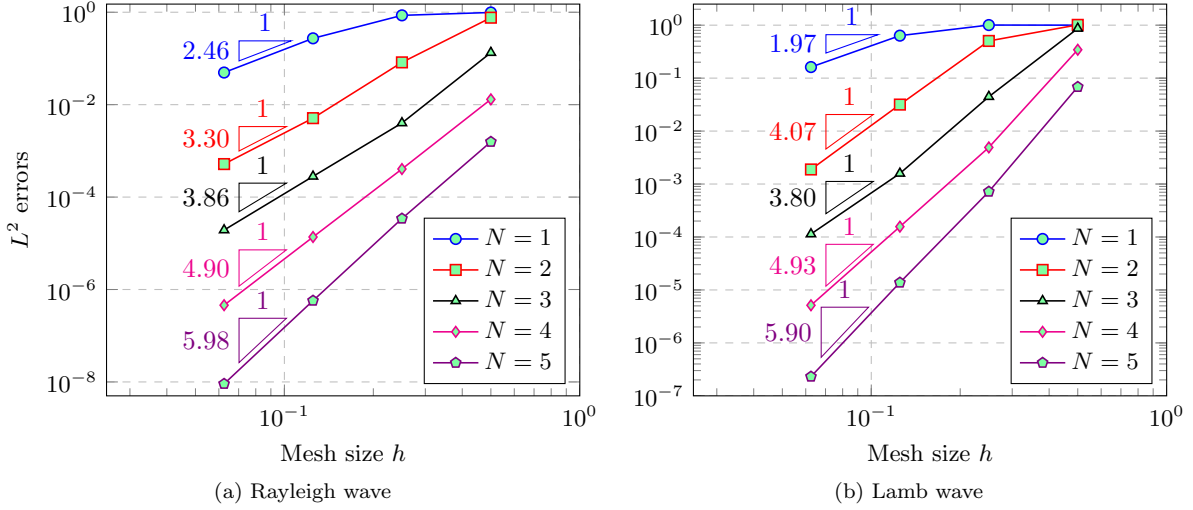
17

Fig. 3: Convergence of L^2 errors for Rayleigh and Lamb wave solutions.

126.1992721468 and $B_2 = 53.88807700007$ [23]. 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 3 shows L^2 errors for both waves at time $T = 5$. Computed convergence rates fall between the optimal rate of $O(h^{N+1})$ and theoretical rate of $O(h^{N+1/2})$ [35].

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 $\mu/\lambda \to 0$ due to the fact that C becomes singular. We use the propagation of Rayleigh waves to examine the behavior of WADG for near-incompressible materials. We follow [36, 37] 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 v, σ at different orders and mesh sizes. The relative errors for σ grow as $\mu/\lambda \to 0$. This is not surprising, as the constant in the error estimates of Section 4.2.1 depends on $\|C\|$, which blows up in the incompressible limit. However, because the magnitude of σ also decreases as $\mu/\lambda \to 0$, relative errors for v remain roughly the same magnitude for near-incompressible materials. Decreasing μ by four orders of magnitude results in up to a ten-fold increase in relative error for σ, but less than a two-fold increase in error for v.

6.2.4. Stoneley waves. A Stoneley wave is supported along the interface between two solids [38]. Like Rayleigh waves, Stoneley waves decay exponentially away from the interface, and test the effectiveness of numerical fluxes across interfaces. We follow [23, 37] 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

$$u_1(x, y, t) = \begin{cases} \mathrm{Re}\left(\left(ikB_1 e^{-kb_{1p}y} + kb_{1s}B_2 e^{-kb_{1s}y}\right) e^{i(ky-\omega t)}\right), & y > 0 \\ \mathrm{Re}\left(\left(-kb_{1p}B_1 e^{-kb_{1p}y} + ikB_2 e^{-kb_{1s}y}\right) e^{i(kx-\omega t)}\right), & y < 0 \end{cases}$$

$$u_2(x, y, t) = \begin{cases} \mathrm{Re}\left(\left(ikB_3 e^{kb_{2p}y} - kb_{2s}B_4 e^{kb_{2s}y}\right) e^{i(kx-\omega t)}\right), & y > 0 \\ \mathrm{Re}\left(\left(kb_{2p}B_3 e^{kb_{2p}y} + ikB_4 e^{kb_{2s}y}\right) e^{i(kx-\omega t)}\right), & y < 0 \end{cases},$$

18

	$\mu = 1$	$\mu = .1$	$\mu = .01$	$\mu = .001$	$\mu = .0001$
$N = 2,\ h = 1/4$	3.1063e-02	2.6972e-02	3.0858e-02	4.0103e-02	5.6835e-02
$N = 3,\ h = 1/4$	3.1677e-03	2.6848e-03	2.9854e-03	3.6960e-03	4.9362e-03
$N = 4,\ h = 1/4$	2.8726e-04	2.4990e-04	2.9142e-04	3.9774e-04	5.2469e-04
$N = 2,\ h = 1/8$	3.1819e-03	2.5476e-03	2.8877e-03	3.5608e-03	4.7010e-03
$N = 3,\ h = 1/8$	1.8867e-04	1.6925e-04	1.8509e-04	2.2520e-04	2.8301e-04
$N = 4,\ h = 1/8$	8.4999e-06	7.5750e-06	8.1094e-06	1.0760e-05	1.4886e-05

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

	$\mu = 1$	$\mu = .1$	$\mu = .01$	$\mu = .001$	$\mu = .0001$
$N = 2,\ h = 1/4$	6.8250e-02	7.4130e-02	1.2104e-01	2.1333e-01	4.1451e-01
$N = 3,\ h = 1/4$	9.2980e-03	1.0685e-02	1.8109e-02	3.2046e-02	5.7612e-02
$N = 4,\ h = 1/4$	9.7251e-04	1.1687e-03	2.1941e-03	4.1378e-03	6.8404e-03
$N = 2,\ h = 1/8$	9.8138e-03	1.1429e-02	2.0331e-02	3.7889e-02	7.6516e-02
$N = 3,\ h = 1/8$	6.6596e-04	7.8157e-04	1.5353e-03	3.1239e-03	5.6341e-03
$N = 4,\ h = 1/8$	3.4272e-05	4.2639e-05	8.8393e-05	1.9764e-04	3.6052e-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.

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 $B_4 = -0.9339639688697$ [23]. 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 4 shows L^2 errors for two uniform meshes of triangles constructed by bisecting a quadrilateral mesh of $K_{1D} \times 5K_{1D}$ elements. Figure 4a shows errors at time $T = 5$ when K_{1D} is even and the mesh is fitted to the interface at $y = 0$, while Figure 4b shows errors when K_{1D} is odd and the interface cuts through element interiors.

When the mesh is fitted to the interface, computed convergence rates match the estimated $O(h^{N+1/2})$ rate. 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 [33] which is exact for degree $2N + 1$ 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, the error in elements cut by the interface is $O(1)$, 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.

6.2.5. Convergence to a reference solution. Finally, we examine the accuracy of the WADG method for smoothly varying heterogeneous media by comparing against a reference spectral method solution with order $N = 50$ on a unit square $[-1, 1]^2$. Boundary conditions are enforced weakly through numerical fluxes [23]. We use a heterogeneous isotropic medium with ρ, λ, μ set to

$$\rho(\boldsymbol{x}) = 1, \qquad \lambda(\boldsymbol{x}) = 1 + .25 \sin(\pi x) \sin(\pi y), \qquad \mu(\boldsymbol{x}) = 1 + .25 \cos(\pi x) \cos(\pi y).$$

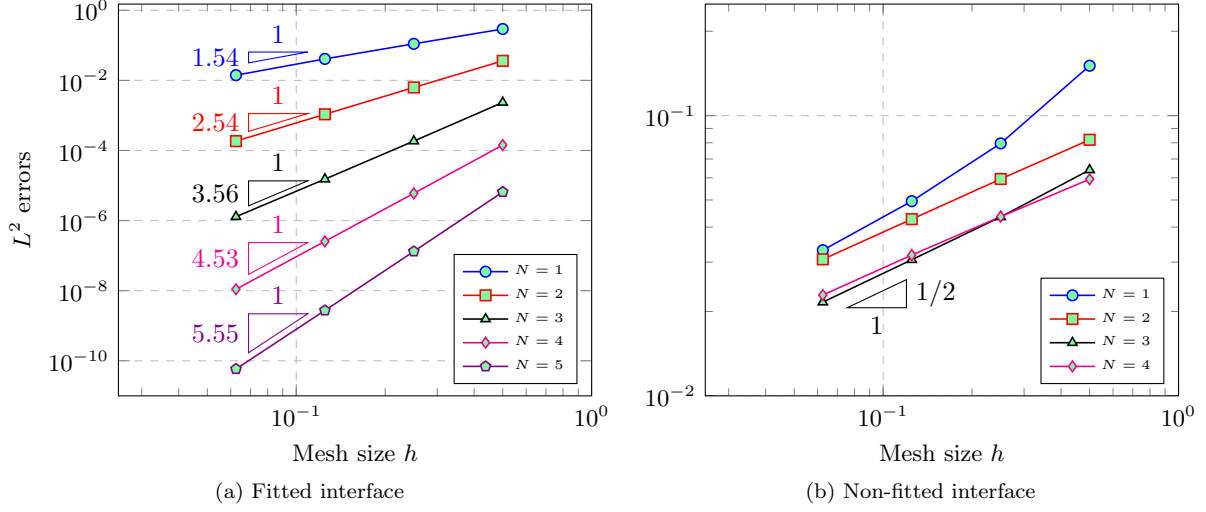(a) Fitted interface   (b) Non-fitted interface

Fig. 4: Convergence of WADG for a Stoneley wave using a fitted mesh aligned with the interface (Figure 4a) and a non-fitted mesh where the interface does not lie exactly on an element boundary (Figure 4b).
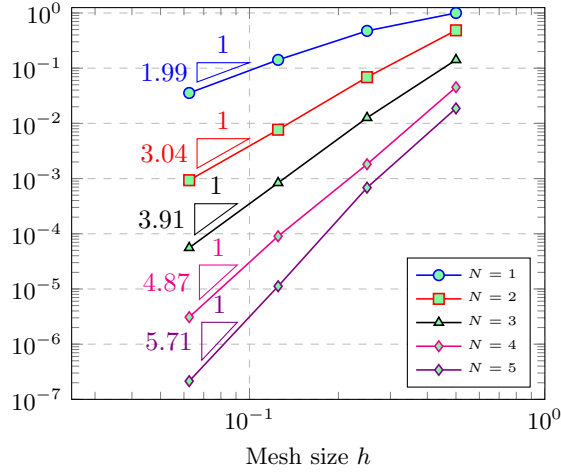

Fig. 5: Convergence of WADG under mesh refinement to a reference $N = 50$ spectral method solution with smoothly varying heterogeneous media for $N = 1, \ldots, 5$.

Initial stresses are set to zero, while the initial velocity is set to

$$v_1(\boldsymbol{x}) = \cos(\pi x)\sin(\pi y), \qquad v_2(\boldsymbol{x}) = -\sin(\pi x)\cos(\pi y).$$

Zero traction boundary conditions are enforced on all domain boundaries. Figure 5 shows L^2 errors with respect to the reference solution at time $T = 1/2$ for different mesh sizes and orders of approximation. Computed convergence rates fall between the optimal $O(h^{N+1})$ and predicted $O(h^{N+1/2})$.

6.3. Application examples. We next demonstrate the accuracy and flexibility of WADG for several application-based problems in linear elasticity with heterogeneity and anisotropy.

6.3.1. Stiff inclusion. The stiff inclusion problem is a common test of methods for linear elastic wave propagation [22, 5, 37], where an inclusion with higher wavespeed is embedded within a non-stiff region.

20

(a) $|\boldsymbol{\sigma}_{xx} + \boldsymbol{\sigma}_{yy}|$ (b) $|\boldsymbol{\sigma}_{xy}|$
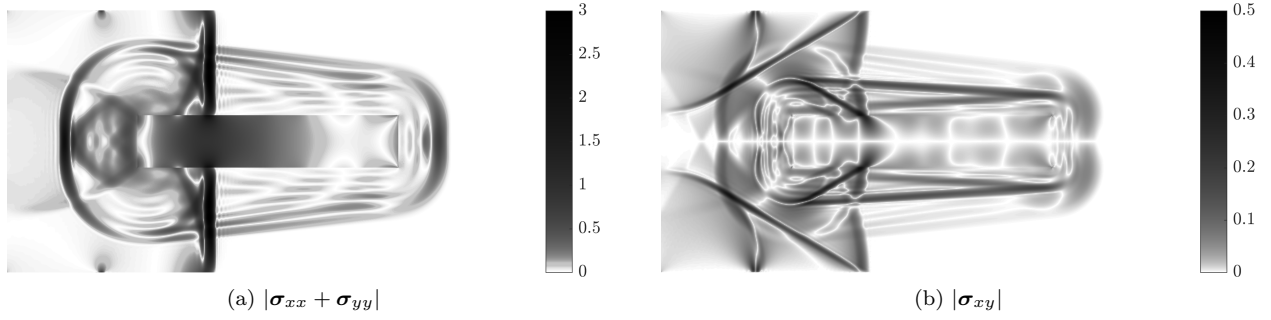
Fig. 6: Values of different solution fields at $T = .4$ for the stiff inclusion problem of Leveque [22] with order of approximation $N = 5$.

Waves which reach this region of high wavespeed are transmitted through the inclusion, bouncing back and forth within the region. This vibration then produces waves which propagate outward from the inclusion.

We solve on a domain $[-1, 1] \times [-.5, .5]$ with a rectangular inclusion located at $[-.5, .5] \times [-.1, .1]$. Outside of the inclusion, material parameters are taken to be

$$\rho = 1, \qquad \mu = 1, \qquad \lambda = 2.$$

Within the inclusion, material parameters are taken to be

$$\rho = 1, \qquad \mu = 100, \qquad \lambda = 200,$$

such that the wave speed in the rectangular inclusion is ten times that of the wave speed outside. A pulse is generated through velocity boundary conditions at $x = -1$

$$v_1(x, y, t) = \begin{cases} \sin(\pi t / t_0), & t < t_0 \\ 0, & t \geq t_0 \end{cases}, \qquad v_2(x, y, t) = 0.$$

In our experiments, we take $t_0 = .025$. Traction free boundary conditions are enforced at all other domain boundaries.

We construct a uniform triangular mesh by dividing each element of a quadrilateral meshes along the diagonal to produce triangular meshes, using 100×50 quadrilateral elements in the x and y coordinates, respectively. Penalty parameters of $\tau_v = \tau_\sigma = 1/2$ are chosen. Figure 6 shows values of $|\boldsymbol{\sigma}_{xx} + \boldsymbol{\sigma}_{yy}|$ at final time $T = .4$. Following [22], a nonlinear color scale is used in order to distinguish small-amplitude waves. The results show qualitatively good agreement with results in the literature [22, 5, 37].

6.3.2. Heterogeneous anisotropic material. We next examine a model wave propagation problem in heterogeneous anisotropy media [39, 40, 7]. The density $\rho = 7100$ is constant over the domain, while the entries of the stiffness matrix $\boldsymbol{C}$ are taken to be

$$\boldsymbol{C}_{11} = .165, \quad \boldsymbol{C}_{12} = .05, \quad \boldsymbol{C}_{22} = .062, \quad \boldsymbol{C}_{33} = .0396, \qquad x < 0$$
$$\boldsymbol{C}_{11} = .165, \quad \boldsymbol{C}_{12} = .0858, \quad \boldsymbol{C}_{22} = .165, \quad \boldsymbol{C}_{33} = .0396, \qquad x > 0,$$

with the remaining entries determined by symmetry or set to zero if unspecified. For $x < 0$, this corresponds to an anisotropic material, while for $x > 0$, this corresponds to an isotropic material with $\mu = .0396, \lambda = .0858$.

The computational domain is taken to be $[-.32, .32]^2$, and following [40, 7], we use $N = 5$ and a triangular mesh of 32768 elements constructed by subdividing a grid of 128×128 uniform quadrilaterals. This mesh corresponds to an element size of .005, as used in [40, 7]. Forcing is applied to the y-component of the velocity by a Ricker wavelet point source

$$f(\boldsymbol{x}, t) = \left(1 - 2(\pi f_0(t - t_0))^2\right) e^{-(\pi f_0(t - t_0))^2} \delta(x - x_0),$$

21

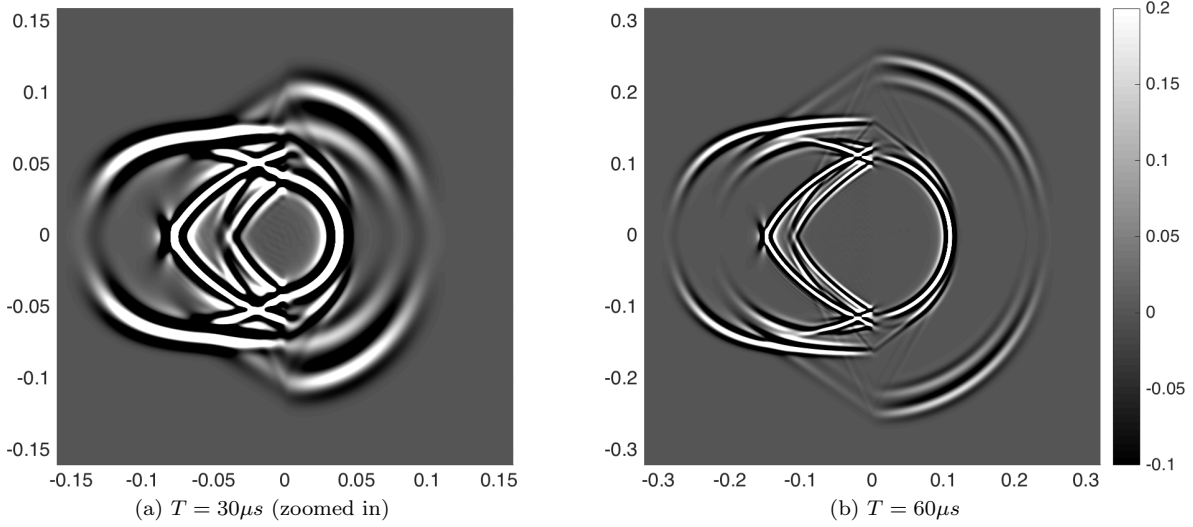(a) $T = 30\mu s$ (zoomed in)

(b) $T = 60\mu s$

Fig. 7: An example of wave propagation in heterogeneous anisotropic media. The vertical component of the velocity $\boldsymbol{v}_2$ is shown at $T = 30$ and $T = 60$ microseconds.

where $x_0 = -.02$, $f_0 = .17$, and $t_0 = 1/f_0$.[2] The penalty parameters are taken to be $\tau_\rho = \tau_\sigma = 1/2$. Figure 7 shows the y-component of velocity $\boldsymbol{v}_2$ at times $T = 30\mu s$ (zoomed in) and $T = 60\mu s$. Both results show qualitative agreement with reference results from [39, 40, 7].

6.4. A three-dimensional example. Finally, we present a three-dimensional example of elastic wave propagation in heterogeneous media with sub-element variations and a discontinuity across an interface. We consider isotropic elastic wave propagation on the cube $[-.5, .5]^3$ with a discontinuity in material coefficients across $z = 0$

$$\rho = 1, \qquad \mu(\boldsymbol{x}) = \begin{cases} 2 + w(\boldsymbol{x}), & z < 0 \\ 1 + w(\boldsymbol{x}), & z > 0 \end{cases}, \qquad \lambda(\boldsymbol{x}) = \begin{cases} 2, & z < 0 \\ 1, & z > 0 \end{cases}$$

where $w(\boldsymbol{x})$ is a spatially varying weight. Forcing is applied to the x-component of velocity through a smoothed point source and Ricker wavelet

$$f(\boldsymbol{x}, t) = \left(1 - 2(\pi f_0(t - t_0))^2\right) e^{-(\pi f_0(t - t_0))^2} e^{-(a\|\boldsymbol{x} - \boldsymbol{x}_0\|)^2}$$

where $\boldsymbol{x}_0 = (0, 0, .1)^T$, $a = 100$, $f_0 = 10$, and $t_0 = 1/f_0$. Figure 8 shows the x-velocity of the computed solution at $T = .5$ with

$$w(\boldsymbol{x}) = .5 \cos(3\pi x) \cos(3\pi y) \cos(3\pi z),$$

as well as $w(\boldsymbol{x}) = 0$ for reference. Figure 8a shows the unstructured mesh of 222824 tetrahedral elements of degree $N = 5$ used to compute both solutions.

These computations are performed on an AMD Tahiti GPU, following the implementation of GPU-accelerated DG methods outlined in [11]. This approach breaks the computational work for each time-step into *volume* and *surface* kernels (for the evaluation of the DG formulation) and an *update* kernel (for the application of a time integration method). In this implementation, we apply the weight-adjusted mass matrix inverse within the update kernel as well. Strategies for volume and surface kernels follow [11], while computational approaches for WADG are outlined in [2].

---

[2] All values and units are adapted from [40, 7], and correspond to units of meters, kg, and microseconds.
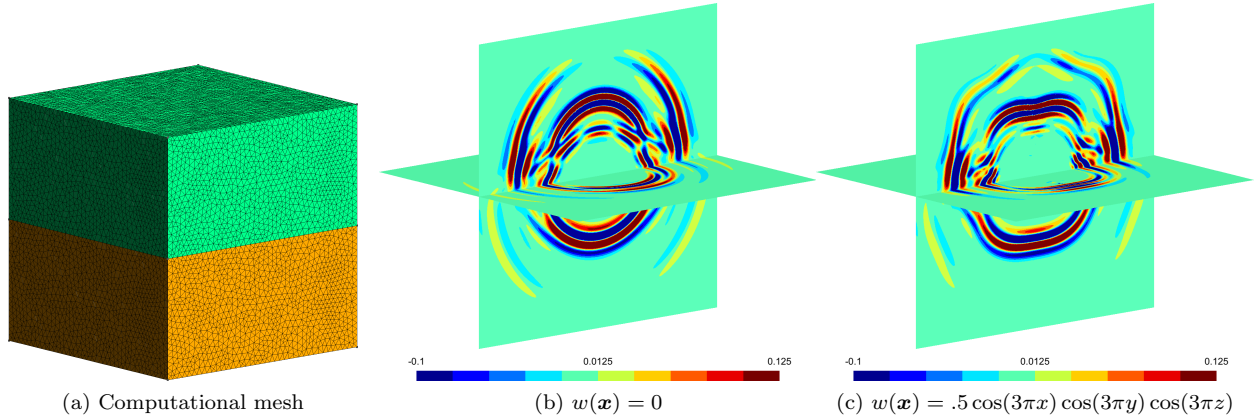
(a) Computational mesh   (b) $w(\boldsymbol{x}) = 0$ (c) $w(\boldsymbol{x}) = .5\cos(3\pi x)\cos(3\pi y)\cos(3\pi z)$

Fig. 8: Mesh and xz, xy slices of $\boldsymbol{v}_1$ at $T = .5$. The order of approximation is taken to be $N = 5$.

7. Conclusions. This work presents a weight-adjusted discontinuous Galerkin (WADG) method for the linear elastic wave equations with arbitrary heterogeneous media. The method is energy stable and high order accurate for arbitrary stiffness matrices, and a slight modification results in an energy stable method for curvilinear meshes as well. The penalty numerical fluxes for this formulation are simple to derive and implement, and their lack of dependence on the stiffness matrix allows for a unified treatment of isotropic and anisotropic media. Numerical examples confirm the accuracy of this method for analytic solutions in of the elastic wave equations, as well as its high order accuracy with respect to a reference solution for smoothly varying heterogeneous media. Results obtained using this method also show good agreement with existing results in the literature for both problems involving both isotropic and anisotropic heterogeneous media.

We note that the implementation of this method reduces to the application of the weight-adjusted mass matrix inverse and the evaluation of constant-coefficient terms in the DG formulation. The cost of the latter step can be reduced (especially at high orders of approximation) by using fast methods based on Bernstein-Bezier bases for the application of derivative and lift matrices for constant-coefficient terms [21].

8. Acknowledgments. The author gratefully thanks Thomas Hagstrom, Tim Warburton, Axel Modave, Ruichao Ye, and Mario Bencomo 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] Jean Virieux. P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 51(4):889–901, 1986.
[4] Dimitri Komatitsch and Jean-Pierre Vilotte. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bulletin of the seismological society of America*, 88(2):368–392, 1998.
[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] 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.
[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] Mario Javier Bencomo. Discontinuous Galerkin and Finite Difference Methods for the Acoustic Equations with Smooth Coefficients. Master's thesis, Rice University, 2015.

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

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

[20] 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.

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

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

[23] 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.

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

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] Paul Houston, Christoph Schwab, and Endre Süli. Discontinuous hp-finite element methods for advection-diffusion-reaction problems. *SIAM Journal on Numerical Analysis*, 39(6):2133–2163, 2002.

[31] T Warburton and Jan S Hesthaven. On the constants in hp-finite element trace inverse inequalities. *Computer methods in applied mechanics and engineering*, 192(25):2765–2773, 2003.

[32] Vít Dolejší and Miloslav Feistauer. *Discontinuous Galerkin Method: Analysis and Applications to Compressible Flow*, volume 48. Springer, 2015.

[33] 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.

[34] Mark H Carpenter and Christopher A Kennedy. Fourth-order 2n-storage Runge-Kutta schemes. Technical Report NASA-TM-109112, NAS 1.15:109112, NASA Langley Research Center, 1994.

[35] 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.

[36] 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.

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

[38] 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.

[39] Juse M Carcione, Dan Kosloff, and Ronnie Kosloff. Wave-propagation simulation in an elastic anisotropic (transversely isotropic) solid. *The Quarterly Journal of Mechanics and Applied Mathematics*, 41(3):319–346, 1988.

[40] Dimitri Komatitsch, Christophe Barnes, and Jeroen Tromp. Simulation of anisotropic wave propagation based upon a spectral element method. *Geophysics*, 65(4):1251–1260, 2000.