Visco-TTI-Elastic FWI using Discontinuous Galerkin

SUMMARY

The need to better represent the material properties within the earth’s interior has driven the development of high-resolution physics, e.g., visco-tilted-transversely-isotropic (visco-TTI) elastic media and material interfaces, such as the ocean bottom and salt boundaries. This is especially true for full wave inversion (FWI), where one would like to reproduce the real-world effects and invert on unprocessed raw data. Here we present a numerical formulation using a Discontinuous Galerkin (DG) finite-element (FE) method, which incorporates the desired high-fidelity physics and material interfaces. To offset the additional costs of this material representation, we include a variety of techniques (e.g., non-conformal meshing, and local polynomial refinement), which reduce the overall costs with little effect on the solution accuracy.

INTRODUCTION

There are two primary tasks to obtain and improve the characterization of the subsurface through Full Waveform Inversion (FWI): increase the fidelity of the physics (e.g., governing equations and material representation), and reduce the computational costs to determine the earth’s material parameters. Obviously higher-fidelity physics, such as elasticity, attenuation and anisotropy, are needed to accurately simulate wave propagation and mimic real-world conditions and this requires additional computational resources.

Another important aspect of this higher-fidelity physics is the spatial representation of the material properties, including material discontinuities (e.g., sediments, faults, ocean bottom, and salt bodies). Finite-Difference (FD) schemes have difficulty representing these discontinuities, while the class of finite-element schemes (e.g., continuous finite element, spectral elements, and Discontinuous Galerkin) can naturally represent the discontinuities across the element interfaces and use unstructured grids to track them across the domain.

Respecting the above high-fidelity requirements, we have selected Discontinuous Galerkin schemes which are a natural choice for the first-order hyperbolic equations commonly used in FWI today. Continuous FE and spectral schemes are well-suited for elliptic equations (Hesthaven and Warburton, 2008, p. 7), but usually require some sort of stabilization for hyperbolic equations. All these schemes can reduce the computational costs by employing higher-order methods and h/p adaptivity, but DG schemes can easily incorporate non-conformal and hybrid meshes through the flux calculations and Riemann solves. Additionally, continuous FE and spectral schemes are implicit and require the mass matrix to be inverted. For transient problems, this becomes a disadvantage.

For forward modeling, DG methods have been applied to a variety of seismic physics (acoustic/elastic interfaces (Wilcox et al., 2010), elasticity (Käser and Dumbser, 2006), attenuation (Käser et al., 2007), anisotropy (de la Puente et al., 2007) and poro-elasticity (de la Puente, 2008)), with h/p adaptivity (Etienne et al. (2010), Mercerat and Glinsky (2015)), non-conformal (Mazzieri et al., 2013), and hybrid meshes.

Full Waveform Inversion (FWI) imposes additional computational requirements to determine the adjoints and gradients. FWI has been demonstrated on a variety of seismic physics including acoustic and elastic with attenuation (Gholami et al. (2010), Peter et al. (2011), Wilcox et al. (2014)). Here we show an implementation for performing FWI on visco-TTI-elastic media using the DG formulation. We also show results demonstrating some of these capabilities.

FORMULATION AND METHODOLOGY

The FWI problem can be formulated as a constrained optimization problem with a least-squares objective function, $\min_\Phi J(U, \Phi)$, where

$$ J = \frac{1}{2} \sum_{r=1}^{N_r} \int_Q \xi_r(x) \left\| R(U - \hat{U}) \right\|^2 \, dQ + \frac{1}{2} \int_\Omega \left\| R(\Phi) \right\|^2 \, d\Omega 
$$

subject to the visco-TTI-elastic equations

$$ \frac{\partial \sigma_i}{\partial t} - \hat{C}_{IJ} \partial V_J - \sum_{l=1}^{L} r_i^l = \frac{\partial m_i^s}{\partial t} \tag{2} $$

$$ \frac{\partial v_i}{\partial t} - v \frac{\partial \sigma_i}{\partial x} = v \left[ f_i + \partial \frac{m_i^s}{\partial x} \right] \tag{3} $$

$$ \frac{\partial r_i^l}{\partial t} + a_i \omega_i \hat{C}_{IJ} \partial V_J + \omega_i r_i^l = 0 \tag{4} $$

over the computational domain $\Omega$, and time horizon $T$, with $Q = \Omega \times (0, T)$. In these expressions, $\sigma_i$ and $\sigma_{ij}$ are the stresses in Voigt and standard tensor notation respectively; $\hat{C}_{IJ}$ is the rotated stiffness matrix; $r_i^l$ are the memory variables for attenuation; $m_i^s = m_{ij}^s + m_{ij}^d$ are the symmetric and anti-symmetric parts of the moment density tensor; $v_i$ are the particle velocities; $v$ is the specific volume; $f_i$ is the force density; $a_i$ are the relaxation amplitudes for a single quality factor, $Q$; $\omega_i$ are the relaxation frequencies; and

$$ \partial V = \begin{bmatrix} \partial v_x / \partial x \\ \partial v_y / \partial y \\ \partial v_z / \partial z \\ \partial v_x / \partial z + \partial v_z / \partial x \\ \partial v_y / \partial z + \partial v_z / \partial y \\ \partial v_x / \partial y + \partial v_y / \partial x \end{bmatrix} \tag{5} $$
Visco-TTI-Elastic FWI using Discontinuous Galerkin

The material properties are parameterized in several ways throughout the code. The user often supplies the material properties in *primitive* parameters, \( \Phi = (\rho, V_p, V_s, \delta, \epsilon, \gamma, Q, \theta, \phi) \). Here \( \rho \) is density, \( V_p \) is the compressional wave speed; \( V_s \) is the shear wave speed; \( \delta, \epsilon, \gamma \) are the Thomsen (1986) parameters; \( Q \) is the quality factor; and \( \theta, \phi \) are the material rotation angles. The governing equations are expressed in *native* parameters, \( \Phi = (\nu, C_{ij}, a_{ij})^T \), and the unrotated version \( \Phi = (\nu, C_{ij}, a_{ij})^T \), where \( C_{ij} \) is the orthorhombic stiffness matrix. The optimization algorithms operate with the numerical flux, and apply reverse integration by parts to get the *primitive* properties in throughout the code. The user often supplies the material properties parameterized in several ways (e.g., impedance, which can be optimization algorithms). Here we have \( C_{ij} \) is the quality factor, \( Q \) is the shear wave speed; \( V \) is the weighting function, \( U \) is the relaxation matrix, and \( L \) relaxation amplitudes to 5 primitive parameters, and \( Q \) boundary conditions that allow waves to leave the domain, and is specified by setting the outside state to \( U^+ = 0 \).

### Spatial Discretization

Referring back to Eq. (9), the computational domain is subdivided into \( N \) elements, quadrilaterals or triangles in two-dimensions, hexahedrals or tetrahedrals in three-dimensions. Legendre polynomials form the basis functions for spatial discretization. Two separate DG formulations are considered. The first is referred to as *modal* because solution variables are transformed into modes of polynomials through projection operations. Volume integrals are evaluated using a Gauss-Labatto-Legendre (GLL) quadrature rule. Surface integration uses a Gauss-Legendre (GL) quadrature rule where the GL points are interpolated from the GLL points. In the implementation of the modal formulation, an arbitrary number of quadrature points \( (q) \) can be specified for each element and exact integration of any term in Eq. (9) can be achieved. The GLL points also produce a diagonal mass matrix for affine elements. In addition, the media is projected into the space of Legendre polynomials which are then interpolated to the GLL or GL integration points. The polynomial order representation of the media is independent of that used for the state variables.

The second formulation is referred to as *spectral* due to its resemblance to the spectral element method (Karniadakis and Sherwin (1999)). This formulation is very similar to the modal formulation. The main differences are that solution variables remain in physical space, and volume and surface integrals are approximated with the GLL points. Therefore, the number of quadrature points for both state and media are constrained to be \( q = p + 1 \). This results in under-integration, however the error decreases as \( p \) increases. The spectral formulation has several advantages compared to the modal formulation. First, no projection/interpolation of the solution variables or the media is required except at the interface of two elements with different polynomial orders; this includes hanging nodes. This results in a factor of two reduction in computation time. Second, the mass matrix is always diagonal even for non-affine elements. This can reduce the computation time by a factor of 5 to 10 depending on the polynomial order.

### Non-conformal Faces and Element Orders

In order to represent strong material variations, slow moving
Visco-TTI-Elastic FWI using Discontinuous Galerkin

surface waves (e.g., Scholte waves) and source and receiver locations, the solution mesh can be non-conformal with “hanging nodes.” We allow a 1 to N refinement across cell boundaries: a single face $s$ of one element can have $N$ neighboring faces $\{s_i\}$. However, the neighbor faces must be a partition of the larger face; i.e., $s_i \subseteq S$ and $s_j$ cannot overlap both $S$ and some other face $S'$. We call such $s_j$ and $S$ non-conformal faces.

The $\{s_i\}$ are not required to be the same shape but can vary in both size and aspect ratios. There is no restriction on the element size and local level of refinement. Given one hanging node mesh, elements can be arbitrarily refined again, as long as $s_i \subseteq S$ is maintained everywhere.

A boxtree mesh is our generalization of the common octree, where each parent can be subdivided into $x \times y \times z$ children. We have a couple of extensions to align the mesh with material discontinuities. First, if the domain cuts boxtree hexes into trivalent polyhedra then midpoint subdivision can divide them into hexes (Li et al., 1997). Second, one may completely remove boxes cut by, or near, the discontinuity; then mesh that space using tetrahedra; while connecting the two meshes with non-conformal faces. We utilize the Vorocrust method for surface reconstruction using Voronoi cells (Mahmoud et al., 2015). It is very robust over complicated domains, and can generate either polyhedra or their dual tetrahedra. The face flux calculations of the DG method provides great mesh flexibility, allowing irregular non-conforming faces with different types (e.g., tetrahedra and hexes) and orders.

In both hybrid and mesh hexes, there are no restrictions on the relative element orders between $S$ and $\{s_i\}$, and it is often useful to use lower-order elements on the refined side. The flux across non-conformal faces is conserved by ensuring the total flux, as well as the moments up to the order of the side with fewer orders, are conserved across the interface. These meshing capabilities allow an overall reduction in the number of degrees of freedom, and thus computational cost.

Local Polynomial Refinement

The polynomial order within each element can be selected for the shortest wavelength, $\lambda_{min} = c_{min}/f_{max}$, expected within the element. Here $c_{min}$ is the slowest wave speed (e.g., $V_p$, $V_s$ or Scholte wave speed) and $f_{max}$ is the maximum expected frequency. Note $f_{max}$ may be substantially smaller in elements deep in the model, as higher frequencies attenuate. Using lower order provides a substantial computational savings. With a sufficient number of points per wavelength, $\alpha \approx 5$, and side length $h$, we can choose the local polynomial order $p$:

$$p = \alpha h f_{max}/c_{min} - 1. \quad (11)$$

Time Integrators

To integrate Eqs. (9, 13, 14), we have a variety of integrators available, e.g., trapezoidal rule, four-stage Runge-Kutta, and a low-storage 14-stage Runge-Kutta (LSRK14) (Niegemann et al., 2012). The LSRK14 has the potential to reduce the costs by 2x because of the increased stability region, and decrease the memory requirements for the adjoint solve by 10x.

Adjoint Equations

To determine the adjoint and material gradients, we form a discrete Lagrange functional

$$L'(U, W, \Phi) = J(U, \Phi) + \int_\Omega W \cdot (\mathbf{F}_n - \mathbf{F}_a) \, dP_e + \sum_{e=1}^{N_e} \int_{Q_e} W \cdot (U_j + A_{ij} U_j + RU - S) \, dQ_e$$

(12)

where $W$ is the adjoint variable (i.e., the Lagrange multiplier). We can obtain the state equation, Eq. (8), by holding Eq. (12) stationary with respect to $W$, $\partial W \{L'(U, W, \Phi)\} = 0$. To obtain the adjoint equation, we hold $\partial U$, stationary with respect to $U$,

$$\sum_{e=1}^{N_e} \int_{Q_e} \left( -V \cdot W_j - V \cdot \mathbf{A}_j W_j + V \cdot R^T W \right. \\
+ \left. V \cdot \sum_{e=1}^{N_e} \tilde{e}_r \mathbf{x} \mathbf{R}^T (U - \hat{U}) \right) dQ_e \\
+ V \cdot \partial U \left\{ \int_{\partial P} (W \cdot \mathbf{F}_n) dP_e \right\} = 0$$

(13)

Note this equation is solved for the adjoint variable, $W$, and that the source term is replaced with the difference between the state and measured data, and integrated backward in time. Also the last term needs to be found for each numerical flux (e.g., $\mathbf{F}_n^{\mathbf{E}}$ and $\mathbf{F}_n^{\mathbf{S}}$).

To solve for $W$ requires the state $U$ at all time steps to be available in order for Eq. (13) to be integrated back in time. But the state data for all time steps can be prohibitively large to save. In this case, the Griewank Algorithm (Restrepo et al. (1998)) can be used, where a select subset of states are stored and intermediate states are recomputed.

Gradient Equations

Lastly the gradient equation is found by holding Eq. (12) stationary with respect to $\Phi$.

$$\partial \Phi L' = \beta \int_\Omega \Phi d \Omega + \sum_{e=1}^{N_e} \int_{Q_e} \left( W \cdot \partial \Phi \right) dQ_e \\
+ \partial \Phi \left\{ \sum_{e=1}^{N_e} \int_{P_e} W \cdot (\mathbf{F}_n - \mathbf{F}_a) dP_e \right\} = 0$$

(14)

The regularization term is present, and again the last term needs to be determined for each numerical flux.

INVERSION

We have several optimization packages available, including an internal linesearch capability, and the Rapid Optimization Library (ROL). ROL (Kouri et al., 2014) is Trilinos package (Heroux and Willenbring, 2012) currently under development, designed for efficient large-scale numerical optimization. ROL provides the capability to solve both full-space and reduced-space problems. For FWI, the storage requirements for a three-dimensional inversion makes full-space methods prohibitively expensive and therefore we use the reduced-space framework. Currently, ROL is used to solve FWI problems using a variety of Linesearch and Trust-Region methods.

In the Linesearch setting, descent directions can be taken to be the negative gradient (steepest descent), generated by nonlinear conjugate gradient or secant methods, or an inexact Newton method. In the latter case, the approximate action of the Hessian on a direction vector $v$

$$\nabla f(x + \epsilon v) - \nabla f(x) = \left[ \nabla^2 f(x) \right] v + O(\epsilon)$$

(15)
and the conjugate-gradient method is applied to approximately solve the Newton step vector. ROL selects the Linesearch step length based on a number of methods for generating candidate step lengths, such as backtracking, Brent’s method, cubic interpolation, in conjunction with a variety of sufficient decrease and curvature requirements which can be set by the user. Similarly, the finite-difference Hessian is used to generate local quadratic models in Trust-Region methods.

One can include simultaneous source inversion (SSI), Krebs et al. (2009), in Eqs. (1)-(3) by replacing the receiver data with the receiver data of multiple shots, \( N_s \), and the source with the sources of multiple shots,

\[
\mathbf{U} \rightarrow \sum_{s=1}^{N_s} \alpha_s \mathbf{U}_s \quad m_{ij} \rightarrow \sum_{s=1}^{N_s} \alpha_s (m_{ij})_s
\]

(16)

where \( \alpha_s \) is randomly 1 or -1 based on \( s \), and is changed every optimization iteration. Note that SSI is independent of the optimization package, and for optimization algorithms with history, e.g., conjugate-gradient methods, the history is corrupted by the randomization of the SSI method.

SYNTHETIC RESULTS

In Fig. 1, we have wave propagation through an overthrust models without a) and with b) attenuation and anisotropic material properties. The material properties vary over the entire domain (not shown). Throughout the images, one can find subtle differences in the wave propagation due to the higher-fidelity physics provided by visco-TTI-elastic media. Point A indicates a difference in wave propagation due to anisotropic material properties (i.e., bent waves and additional reflected waves). Point B indicates differences due to attenuation (i.e., a reduction in wave amplitude).

To demonstrate our inversion capabilities, we display a sample of the results from Krebs et al. (2016) inFig. 2. In this inversion of \( I_p, V_p \) and \( V_s \) from visco-TTI-elastic data, we are just showing the results related to \( V_p \). For further details, see Krebs et al. (2016). In Fig. 2(a), the target \( V_p \) values are shown with an anomaly in the form of the letter "P". There are other anomalies in other material properties, e.g., "I" for impedance, "S" for \( V_s \), and "q" for the quality factor (all not shown).

DISCUSSION

We have presented a numerical formulation for full waveform inversion (FWI) of a visco-TTI-elastic media using a Discontinuous Galerkin (DG) method. The DG method provides a flexible means to accurately represent material interfaces, while allowing a variety meshing capabilities that reduces the number of degrees of freedom and thus the computational costs. Additionally, we have shown examples of wave propagation and inversion with further details shown in (Krebs et al., 2016) that demonstrate the ability to model and invert high-fidelity physics. This capability should allow FWI on raw-data, and reduce the need to pre-process field data.

ACKNOWLEDGMENTS

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.
Visco-TTI-Elastic FWI using Discontinuous Galerkin

REFERENCES