Nonuniform 3D finite difference elastic wave simulation on staggered grids

Journal: Geophysics
Manuscript ID: GEO-2021-0512.R3
Manuscript Type: Technical Paper
Keywords: wave propagation, elastic, parallel
Manuscript Focus Area: Seismic Modeling and Wave Propagation

© 2022 Society of Exploration Geophysicists

DOI: 10.1190/geo2021-0512.1
Nonuniform 3D finite difference elastic wave simulation on staggered grids

Longfei Gao* and Omar Ghattas† and David Keyes‡

*Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, USA. E-mail: longfei.gao@austin.utexas.edu

†Oden Institute for Computational Engineering and Sciences, Jackson School of Geosciences, and Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712, USA

‡Applied Mathematics and Computational Science Program and Extreme Computing Research Center, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.

(Tuesday 8th March, 2022)

Running head: Nonuniform grids for wave propagation

ABSTRACT

We present an approach to simulate the 3D isotropic elastic wave propagation using nonuniform finite difference discretization on staggered grids. Specifically, we consider simulation domains composed of layers of uniform grids with different grid spacings, separated by nonconforming interfaces. We demonstrate that this layer-wise finite difference discretization has the potential to significantly reduce the simulation cost, compared to its fully uniform counterpart. Stability of such a discretization is achieved by using specially
designed difference operators, which are variants of the standard difference operators with adaptations near boundaries or interfaces, and penalty terms, which are appended to the discretized wave system to weakly impose boundary or interface conditions. Combined with specially designed interpolation operators, the discretized wave system is shown to preserve the energy conserving property of the continuous elastic wave equation, and a fortiori ensure the stability of the simulation. Numerical examples are presented to demonstrate the efficacy of the proposed simulation approach.
INTRODUCTION

Simulating waves propagating through earth media is a routine task in seismic exploration. Elastic and acoustic assumptions of earth media are common choices in practice. The elastic model offers a more accurate depiction since, in addition to the compressional wave supported by the simpler acoustic model, it also supports shear and surface waves, which are commonly observed in seismic survey data and can contain significant energy and information. For applications involving salt bodies where significant conversions between compressional and shear waves occur at the salt boundaries (see, e.g., Ogilvie and Purnell (1996)) or applications requiring an accurate account of the near surface region where surface wave has a strong presence (see, e.g., Miller and Pursey (1955)), the importance of elastic model can be particularly pronounced. However, elastic wave simulation is significantly more expensive than its acoustic counterpart because of the increased number of variables and derivatives in the equations as well as the need to resolve the lower shear wave speed.

For a fixed frequency, wavelengths of the seismic waves are proportional to the wave speeds of the media. The spatial grid spacing in a wave simulation is usually dictated by the minimum wavelength of the propagating waves, and hence is determined by the minimum wave speed of the media. Based on the observation that wave speeds of earth media generally increase with depth due to geological sedimentation and consolidation, it is desirable to have progressively coarsened spatial discretization from earth surface to deeper depth in order to alleviate the computational burden. To illustrate, soil and dry sand on earth’s surface can have shear wave speeds as low as 100 m/s while hard rocks deep beneath can have shear wave speeds above 3000 m/s and compressional wave speeds around
6000 m/s; see Bourbié et al. (1987, p. 240). Using a uniform grid with such a large contrast in wave speeds would lead to wasteful oversampling in the discretization (see Appendix A for a detailed explanation).

Fully unstructured finite element based discretizations can address the variation in wave speed by allowing variation in the sizes of the individual elements. Such an approach has been thoroughly investigated in previous works in the context of seismic wave propagation; see, e.g., Lysmer and Drake (1972); Bao et al. (1998); Komatitsch and Vilotte (1998); Komatitsch and Tromp (1999); Käser and Dumbser (2006); Dumbser and Käser (2006); Chaljub et al. (2007); Wilcox et al. (2010); Bui-Thanh and Ghattas (2012). However, these discretizations tend to be more complex to implement and more expensive to compute on a per grid point basis than those using finite differences.

For earth media, at least at the exploration scale, finite difference discretizations on layer-wise uniform grids as illustrated in Figure 1 seem to be an attractive compromise between the complex fully unstructured finite element schemes and the rigid fully uniform finite difference schemes. In the context of seismic wave propagation, previous works adopting this approach can be found in, e.g., Jastram and Tessmer (1994); Aoi and Fujiiwara (1999); Hayashi et al. (2001); Kristek et al. (2010); Zhang et al. (2013); Nie et al. (2017); Gao et al. (2018). However, these earlier attempts are often affected by instability issues associated with the nonconforming interfaces; see, e.g., Gao et al. (2018) for a clear demonstration.

In this work, we adopt this layer-wise finite difference discretization approach and make use of the summation by parts (SBP)-simultaneous approximation terms (SATs) technique to address the nonconforming interfaces and ensure stable simulations. Mechanisms of the
proposed discretization are detailed in Methodology. For general review of the SBP-SAT technique, we refer the readers to Fernández et al. (2014) and Svärd and Nordström (2014). Development and application of the SBP-SAT technique on wave propagation problem can be found in Appelö and Petersson (2009); Sjögreen and Petersson (2012); Petersson and Sjögreen (2015); Wang et al. (2016); O’Reilly et al. (2017); Gao and Keyes (2019); Gao et al. (2019); Gao and Keyes (2020a,b), among others. In particular, O’Reilly et al. (2017); Gao and Keyes (2019); Gao et al. (2019); Gao and Keyes (2020a,b) utilize staggered grids. This work extends these earlier efforts and demonstrates that the SBP-SAT technique can be effectively applied to large-scale 3D elastic wave simulations on staggered grids with nonconforming interfaces.

Other related works on simulating wave propagation using discretizations with adaptive features include, e.g., Moczo (1989) and Pitarka (1999), which considered irregular grids that maintain an overall tensor product structure but allow for nonuniform spacing in each direction; Tessmer (2000), Kang and Baag (2004), and Diaz and Grote (2009), which considered spatially varying time step length; Liu and Sen (2011), which allowed the stencil length to vary. Related works on addressing the challenges arising from material interfaces include, e.g., Fornberg (1988), where the curved interface is straightened by a smooth mapping to improve accuracy of the pseudospectral method considered therein; Zhang and LeVeque (1997) and Martin et al. (2015), where interface conditions are incorporated in the specially designed finite difference stencils.

[Figure 1 about here.]
METHODOLOGY

Problem Description

Omitting the external source terms, elastic wave propagation can be described using the following velocity-stress formulation:

\[
\begin{align*}
\rho \frac{\partial \mathbf{v}}{\partial t} &= \sigma; \\
\dot{\mathbf{\sigma}} &= C_{ijkl} \left( \frac{\mathbf{v}_k \mathbf{v}_j + \mathbf{v}_j \mathbf{v}_k}{2} \right),
\end{align*}
\]

where the Einstein summation convention applies to subscript indices. In equation 1, \( v_i \) and \( \sigma_{ij} \) are the sought solution variables, standing for components of the velocity vector and the stress tensor, respectively; \( \rho \) and \( C_{ijkl} \) are given physical parameters, standing for density and components of the stiffness tensor, respectively. In the case of isotropic media, equation 1b can be simplified as

\[
\dot{\sigma}_{ij} = \lambda \delta_{ij} v_{k,k} + \mu (v_{i,j} + v_{j,i}),
\]

where \( \lambda \) and \( \mu \) are the Lamé parameters, and \( \delta_{ij} \) stands for the Kronecker delta.
The above equations can also be expressed in the following form:

\[
\begin{align*}
\frac{\partial v_x}{\partial t} &= \frac{1}{\rho} \left( \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} \right); \\
\frac{\partial v_y}{\partial t} &= \frac{1}{\rho} \left( \frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yz}}{\partial z} \right); \\
\frac{\partial v_z}{\partial t} &= \frac{1}{\rho} \left( \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} \right); \\
\frac{\partial \sigma_{xx}}{\partial t} &= (\lambda + 2\mu) \frac{\partial v_x}{\partial x} + \lambda \frac{\partial v_y}{\partial y} + \lambda \frac{\partial v_z}{\partial z}; \\
\frac{\partial \sigma_{yy}}{\partial t} &= \lambda \frac{\partial v_x}{\partial x} + (\lambda + 2\mu) \frac{\partial v_y}{\partial y} + \lambda \frac{\partial v_z}{\partial z}; \\
\frac{\partial \sigma_{zz}}{\partial t} &= \lambda \frac{\partial v_x}{\partial x} + \lambda \frac{\partial v_y}{\partial y} + (\lambda + 2\mu) \frac{\partial v_z}{\partial z}; \\
\frac{\partial \sigma_{xy}}{\partial t} &= \mu \left( \frac{\partial v_y}{\partial x} + \frac{\partial v_x}{\partial y} \right); \\
\frac{\partial \sigma_{yz}}{\partial t} &= \mu \left( \frac{\partial v_z}{\partial x} + \frac{\partial v_x}{\partial z} \right); \\
\frac{\partial \sigma_{xz}}{\partial t} &= \mu \left( \frac{\partial v_z}{\partial x} + \frac{\partial v_y}{\partial z} \right),
\end{align*}
\]

which is cumbersome for the ensuing discussion, but can be a convenient reference for practical implementation.

Within each layer, the spatial discretization is carried out on the staggered grids (see, e.g., Yee, 1966; Virieux, 1986; Levander, 1988), whose layout is sketched in Figure 2. On each axial direction, two sets of grid points interlacing each other in a staggering manner, which leads to eight sub-grids in total for a 3D staggered grid discretization. Specifically, the three normal stress components (\(\sigma_{xx}, \sigma_{yy}, \text{ and } \sigma_{zz}\)) reside on one sub-grid together; each shear stress component (\(\sigma_{xy}, \sigma_{xz}, \text{ or } \sigma_{yz}\)) resides on its own sub-grid; each velocity component (\(v_x, v_y, \text{ or } v_z\)) also resides on its own sub-grid. These nine solution fields occupy seven sub-grids in total. One sub-grid is unoccupied.

[Figure 2 about here.]
Overview of the discretization

We use the SBP-SAT technique to address the nonconforming interfaces appearing in the layer-wise discretization setting illustrated in Figure 1. This technique is guided by energy analysis to construct the spatial discretizations. The concept of SBP finite difference operators dates back to Kreiss and Scherer (1974), which adapt standard finite difference operators near boundaries or interfaces to mimic the integration-by-parts property in the continuous operators of the partial differential equations. The concept of SATs dates back to Carpenter et al. (1994), which append penalty terms (i.e., SATs) to the discretized system to weakly impose boundary or interface conditions.

Applied to the layer-wise discretization setting considered in this work, the individual layers are first discretized independently using SBP finite difference operators and then combined together through SATs. An important component of the SATs is the compatible interpolation operators that transfer the solution at layer interfaces. A collection of such interpolation operators corresponding to a variety of grid spacing ratios is included in Appendix B. The grid spacing ratios are allowed to be rational numbers, instead of just integers, which improves the application flexibility of the proposed discretization approach. With carefully designed SBP operators, SATs, and interpolation operators, the resulting spatial discretization preserves the energy conserving property of the continuous elastic wave equation, and a fortiori ensures the stability of the simulation.

We note here that such a discretization procedure bears a similarity to the discontinuous Galerkin (DG) method in the sense that an entire layer can be viewed as the analogue of an element in the DG method, while the SATs resemble the numerical fluxes in the DG method; see Hesthaven and Warburton (2008, p. 13). However, the ‘discontinuity’ of discretization
in the presented method occurs at a much coarser granular level than in the DG method.

SBP operators on 1D staggered grids

To motivate the concept of SBP operators, let us briefly consider the case of 1D wave propagation on staggered grids. The 1D wave propagation can be described by the following system:

\[
\begin{align*}
\frac{\partial v}{\partial t} &= \frac{\partial \sigma}{\partial x}; \\
\beta \frac{\partial \sigma}{\partial t} &= \frac{\partial v}{\partial x},
\end{align*}
\]

(4a)
(4b)

where \( \beta = \frac{1}{\rho c^2} \) stands for compressibility.

Supposing the above wave system is defined over interval \((x_L, x_R)\), its associated physical energy can be expressed as:

\[\mathcal{E} = \frac{1}{2} \int_{x_L}^{x_R} (\rho v^2 + \beta \sigma^2) \, dx.\]

(5)

Taking the time derivative of equation 5, substituting the equations 4a and 4b, and applying the integration-by-parts formula, we arrive at:

\[\frac{d\mathcal{E}}{dt} = - \sigma(x_L) \cdot v(x_L) + \sigma(x_R) \cdot v(x_R).\]

(6)

In other words, time derivative of the physical energy \(\mathcal{E}\) reduces to boundary data only. If free surface boundary condition is associated with both boundaries, i.e., \(\sigma(x_L) = \sigma(x_R) = 0\), we have that \(\frac{d\mathcal{E}}{dt} = 0\), i.e., the physical energy is conservative.

[Figure 3 about here.]

Supposing that the wave system in equation 4 is discretized on a set of staggered grids illustrated in Figure 3 with \(\sigma\) occupying the \(N\)-grid, which includes the two boundary points,
and \( v \) occupying the \( M \)-grid, which consists of only interior grid points, the semi-discretized system can be expressed as:

\[
\begin{align*}
M^M \rho^M \frac{dV}{dt} &= A^M D^N \Sigma; \\
N^N \beta^N \frac{d\Sigma}{dt} &= A^N D^M V,
\end{align*}
\]  
(7a)  
(7b)

where \( V \) and \( \Sigma \) are discretized solution vectors; diagonal matrices \( \rho^M \) and \( \beta^N \) are the discrete correspondences of \( \rho \) and \( \beta \) in equation 4, respectively; matrices \( D^N \) and \( D^M \) are finite difference operators approximating \( \frac{\partial}{\partial z} \), which apply on vectors defined on the \( N \)- and \( M \)-grid, respectively; finally, diagonal matrices \( A^M \) and \( A^N \) are referred to as norm matrices, which are formally redundant in the above equations, but will play an important role in the upcoming discussion.

Analogously to equation 5, the discrete energy associated with the above semi-discretized system is defined as:

\[
E = \frac{1}{2} V^T (A^M \rho^M) V + \frac{1}{2} \Sigma^T (A^N \beta^N) \Sigma.
\]  
(8)

Taking the time derivative of equation 8 and substituting the equations 7a and 7b, we arrive at:

\[
\frac{dE}{dt} = \Sigma^T \left[ A^N D^M + (A^M D^N)^T \right] V.
\]  
(9)

To simplify the notation, we introduce matrix \( Q \), which is defined as

\[
Q = A^N D^M + (A^M D^N)^T.
\]  
(10)

For \( \frac{dE}{dt} \) in equation 9 to mimic \( \frac{d\mathcal{E}}{dt} \) in equation 6, we ask \( Q \) to satisfy the following property:

\[
Q = -\mathcal{E}^L (p^L)^T + \mathcal{E}^R (p^R)^T,
\]  
(11)
where $E^L$ and $E^R$ are the canonical basis vectors whose first and last components are 1, respectively; $P^L$ and $P^R$ are projection operators that can be applied on vector $V$ to provide approximations to $v(x_L)$ and $v(x_R)$, respectively.

With the above property, equation 9 can be written as:

$$\frac{dE}{dt} = -(E^L)^T \Sigma \cdot (P^L)^T V + (E^R)^T \Sigma \cdot (P^R)^T V.$$  \hspace{1cm} (12)

Realizing that $(E^L)^T \Sigma$, $(P^L)^T V$, $(E^R)^T \Sigma$, and $(P^R)^T V$ are approximations to $\sigma(x_L)$, $v(x_L)$, $\sigma(x_R)$, and $v(x_R)$, respectively, we have that the discrete energy $E$ in equation 9 mimics the behavior of $E$ in equation 6.

The property of $Q$ in equation 11 is attained by carefully constructing the SBP operators $A^N$, $D^M$, $A^M$, and $D^N$ appearing in its definition; see equation 10. In practice, such SBP operators are often constructed with the assistance of symbolic computing software. One may refer to Gao et al. (2019, p. 6) for an example of such SBP operators, which will be used throughout this study. Validation of these operators can be found in, e.g., Gao and Keyes (2020a, Figure 3) and Gao et al. (2019, Appendix A). These operators are included in Appendix C to make this work self-contained. From equation C-1, we can see that in the interior, the finite difference operators $D^M$ and $D^N$ are the same as the standard fourth-order staggered grid stencil $[1/24, -9/8, 9/8, -1/24]/\Delta x$ (see, e.g., Levander 1988), and that only the stencils near the boundaries are adapted.

The norm matrices $A^N$ and $A^M$, seemingly redundant in the semi-discretized system 7, introduce extra degrees of freedom to allow $Q$ to have the desired property. They reduce to the grid spacing $\Delta x$ in the interior and vary only near the boundaries. In the above energy analysis, $A^N$ and $A^M$ play the role of the integral operator $\int_{x_L}^{x_R}$ in equation 5, and can be understood as providing the quadrature rules; see Hicken and Zingg (2013). In this work,
we only consider the case where $A^N$ and $A^M$ are diagonal, so that products $(A^M \rho^M)$ and $(A^N \beta^N)$ are symmetric, and, consequently, the expression in equation 8 provides a valid definition of discrete energy. A drawback of this choice of diagonal norm matrices is that the formal order of the stencils near the boundary can only be as high as half of that in the interior; see Kreiss and Scherer (1974); Strand (1994).

In the following, we use the term ‘discrete energy analysis’ to refer to the procedure of taking the time derivative of the discrete energy, substituting the semi-discretized system to replace the temporal derivatives of the solution variables, and finally, invoking the properties built in the SBP operators to reduce the result to boundaries or interfaces.

The free surface boundary conditions $\sigma(x_L) = \sigma(x_R) = 0$ can be imposed by appending penalty terms to the semi-discretized system 7, leading to the following modified system:

$$\begin{align*}
A^M \rho^M \frac{dV}{dt} &= A^M D^N \Sigma + P^L \left[ (\varepsilon^L)^T \Sigma - 0 \right] - P^R \left[ (\varepsilon^R)^T \Sigma - 0 \right]; \\
A^N \beta^N \frac{d\Sigma}{dt} &= A^N D^M V. 
\end{align*}$$

(13a) (13b)

Following the aforementioned procedure of discrete energy analysis, one can verify that the discrete energy $E$ as defined in equation 8 is conservative, i.e., $\frac{dE}{dt} = 0$, for the above modified system.

**SBP finite difference operators on 3D staggered grids**

The 1D SBP operators introduced above serve as the building blocks of the 3D SBP operators used in this work. We append subscripts $x$, $y$, and $z$ to the 1D operators $A^N$, $D^M$, $A^M$, and $D^N$ to distinguish the directions. By design, these 1D operators satisfy the
following properties:

\[ Q_x = A_x^N D_x^M + (A_x^M D_x^N)^T = - A_x^L (P_x^L)^T + A_x^R (P_x^R)^T; \]  

(14a)

\[ Q_y = A_y^N D_y^M + (A_y^M D_y^N)^T = - A_y^L (P_y^L)^T + A_y^R (P_y^R)^T; \]  

(14b)

\[ Q_z = A_z^N D_z^M + (A_z^M D_z^N)^T = - A_z^L (P_z^L)^T + A_z^R (P_z^R)^T. \]  

(14c)

Furthermore, we use \( \mathcal{I}_x^N, \mathcal{I}_x^M, \mathcal{I}_y^N, \mathcal{I}_y^M, \mathcal{I}_z^N, \) and \( \mathcal{I}_z^M \) to denote the identity matrices corresponding to the \( N \)- and \( M \)-grids of the \( x \)-, \( y \)-, and \( z \)-directions, respectively.

The 3D norm matrices are constructed as the tensor products of the 1D norm matrices as follows:

\[ A_{x_{\Sigma}} = A_x^M \otimes A_y^N \otimes A_z^N, \quad A_{y_{\Sigma}} = A_x^M \otimes A_y^M \otimes A_z^N, \quad A_{z_{\Sigma}} = A_x^N \otimes A_y^N \otimes A_z^N, \]

\[ A_{x_{\Sigma}yx} = A_x^M \otimes A_y^M \otimes A_z^N, \quad A_{y_{\Sigma}yx} = A_x^N \otimes A_y^M \otimes A_z^M, \quad A_{z_{\Sigma}yx} = A_x^N \otimes A_y^N \otimes A_z^N, \]

\[ A_{x_{\Sigma}z} = A_x^N \otimes A_y^M \otimes A_z^M, \quad A_{y_{\Sigma}z} = A_x^M \otimes A_y^N \otimes A_z^M, \quad A_{z_{\Sigma}z} = A_x^N \otimes A_y^N \otimes A_z^M. \]  

(15)

The 3D finite difference operators are constructed as the tensor products of the 1D finite difference operators and the 1D identity matrices as follows:

\[ D_{x_{\Sigma}xx} = D_x^M \otimes \mathcal{I}_y^N \otimes \mathcal{I}_z^N, \quad D_{y_{\Sigma}xx} = \mathcal{I}_x^M \otimes D_y^M \otimes \mathcal{I}_z^N, \quad D_{z_{\Sigma}xx} = \mathcal{I}_x^M \otimes \mathcal{I}_y^N \otimes D_z^M, \]

\[ D_{x_{\Sigma}yy} = D_x^M \otimes \mathcal{I}_y^M \otimes \mathcal{I}_z^N, \quad D_{y_{\Sigma}yy} = \mathcal{I}_x^M \otimes D_y^N \otimes \mathcal{I}_z^N, \quad D_{z_{\Sigma}yy} = \mathcal{I}_x^M \otimes \mathcal{I}_y^N \otimes D_z^M, \]

\[ D_{x_{\Sigma}zz} = D_x^M \otimes \mathcal{I}_y^N \otimes \mathcal{I}_z^M, \quad D_{y_{\Sigma}zz} = \mathcal{I}_x^M \otimes D_y^M \otimes \mathcal{I}_z^M, \quad D_{z_{\Sigma}zz} = \mathcal{I}_x^M \otimes \mathcal{I}_y^N \otimes D_z^M, \]  

(16)

Supposing that the 3D wave system described in equation 1 is defined over a cuboid \( \Omega = (x_L, x_R) \times (y_L, y_R) \times (z_L, z_R) \), its associated physical energy can be expressed as:

\[ E = \frac{1}{2} \int_{\Omega} (\rho v_i v_i + \sigma_{ij} S_{ijkl} \sigma_{kl}) \, d\Omega, \]  

(17)
where $S$ stands for the compliance tensor, which is the inverse of the stiffness tensor $C$ from equation 1b. Taking the time derivative of equation 17, substituting the equations 1a and 1b, and applying the divergence theorem, we arrive at

$$
\frac{d\mathcal{E}}{dt} = \int_{\partial\Omega} v_i \sigma_{ij} n_j \, d\Omega, \tag{18}
$$

where $n_j$ stands for component of the outward normal vector on the boundary. From equation 18, we have that the time derivative of the physical energy $\mathcal{E}$ reduces to boundary data only, analogously to the 1D case as in equation 6.

Replacing the spatial derivatives in equation 1 with the 3D SBP operators introduced above, the resulting discretized system can be expressed as:

$$
\begin{align*}
\mathcal{A}_i V_i t & = \mathcal{A}_i \mathcal{D}_j \Sigma_{ij}; \\
\mathcal{A}_{ij} S_{ijkl} \frac{d\Sigma_{kl}}{dt} & = \mathcal{A}_{ij} \frac{1}{2} \left( \mathcal{D}_j V_i + \mathcal{D}_i V_j \right), \tag{19a}
\end{align*}
$$

where the Einstein summation convention applies to subscript indices only (e.g., for the expression $\mathcal{A}_i \mathcal{D}_j \Sigma_{ij}$ on the right-hand side of equation 19a, only the index $j$ is summed over).

The discrete energy associated with system 19 is defined as

$$
E = \frac{1}{2} V_i^T (\mathcal{A}_i \rho V_i) V_i + \frac{1}{2} \Sigma_{ij} \left( \mathcal{A}_{ij} S_{ijkl} \right) \Sigma_{kl}. \tag{20}
$$

Following the procedure of discrete energy analysis introduced in the previous section, it
can be shown that
\[
\frac{dE}{dt} = -\left[ \Sigma_{xx}^T \left( \mathcal{E}_x^L \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (P_x^L \otimes I_y^N \otimes I_z^N)^T V_z \right] \\
+ \left[ \Sigma_{yy}^T \left( \mathcal{E}_x^L \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_y^N \otimes A_z^N] \cdot \left[ (P_x^R \otimes I_y^N \otimes I_z^N)^T V_y \right] \\
- \left[ \Sigma_{yy}^T \left( P_x^L \otimes I_y^M \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^M \otimes I_z^M)^T V_y \right] \\
+ \left[ \Sigma_{yz}^T \left( P_x^R \otimes I_y^M \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^M \otimes I_z^M)^T V_y \right] \\
- \left[ \Sigma_{xy}^T \left( P_x^L \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^N)^T V_x \right] \\
+ \left[ \Sigma_{xy}^T \left( P_x^R \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^N)^T V_x \right] \\
+ \left[ \Sigma_{xz}^T \left( I_x^M \otimes I_y^N \otimes I_z^M \right) \right] \cdot [A_x^M \otimes A_z^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^M)^T V_x \right] \\
+ \left[ \Sigma_{xz}^T \left( I_x^M \otimes I_y^L \otimes I_z^M \right) \right] \cdot [A_x^M \otimes A_z^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^M)^T V_x \right] \\
+ \left[ \Sigma_{yz}^T \left( I_y^M \otimes I_y^N \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_y \right] \\
+ \left[ \Sigma_{yz}^T \left( I_y^M \otimes I_y^L \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_y \right] \\
+ \left[ \Sigma_{xy}^T \left( I_x^N \otimes I_y^L \otimes I_z^M \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_x \right] \\
+ \left[ \Sigma_{xy}^T \left( I_x^N \otimes I_y^N \otimes I_z^L \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_x \right]
\] \quad \text{yz-faces}
\]
\[
\begin{align*}
\frac{dE}{dt} &= -\left[ \Sigma_{xx}^T \left( \mathcal{E}_x^L \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (P_x^L \otimes I_y^N \otimes I_z^N)^T V_z \right] \\
&\quad + \left[ \Sigma_{yy}^T \left( \mathcal{E}_x^L \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_y^N \otimes A_z^N] \cdot \left[ (P_x^R \otimes I_y^N \otimes I_z^N)^T V_y \right] \\
&\quad - \left[ \Sigma_{yy}^T \left( P_x^L \otimes I_y^M \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^M \otimes I_z^M)^T V_y \right] \\
&\quad + \left[ \Sigma_{yz}^T \left( P_x^R \otimes I_y^M \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^M \otimes I_z^M)^T V_y \right] \\
&\quad - \left[ \Sigma_{xy}^T \left( P_x^L \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^N)^T V_x \right] \\
&\quad + \left[ \Sigma_{xy}^T \left( P_x^R \otimes I_y^N \otimes I_z^N \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^N)^T V_x \right] \\
&\quad + \left[ \Sigma_{xz}^T \left( I_x^M \otimes I_y^N \otimes I_z^M \right) \right] \cdot [A_x^M \otimes A_z^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^M)^T V_x \right] \\
&\quad + \left[ \Sigma_{xz}^T \left( I_x^M \otimes I_y^L \otimes I_z^M \right) \right] \cdot [A_x^M \otimes A_z^N] \cdot \left[ (I_x^L \otimes I_y^N \otimes I_z^M)^T V_x \right] \\
&\quad + \left[ \Sigma_{yz}^T \left( I_y^M \otimes I_y^N \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_y \right] \\
&\quad + \left[ \Sigma_{yz}^T \left( I_y^M \otimes I_y^L \otimes I_z^M \right) \right] \cdot [A_y^M \otimes A_z^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_y \right] \\
&\quad + \left[ \Sigma_{xy}^T \left( I_x^N \otimes I_y^L \otimes I_z^M \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_x \right] \\
&\quad + \left[ \Sigma_{xy}^T \left( I_x^N \otimes I_y^N \otimes I_z^L \right) \right] \cdot [A_x^N \otimes A_y^N] \cdot \left[ (I_y^L \otimes I_y^N \otimes I_z^M)^T V_x \right]
\] \quad \text{xy-faces}
\]

which again reduces to boundary data only.

In order to see this more clearly, let us take the first line of equation 21 as an example.

Recognizing that \([\Sigma_{xx}^T (\mathcal{E}_x^L \otimes I_y^N \otimes I_z^N)]\) is the restriction of \(\Sigma_{xx}\) on the \(yz\)-face at \(x = x_L\), and \([(P_x^L \otimes I_y^N \otimes I_z^N)^T V_z]\) is the projection of \(V_z\) on the same face, we have that the expression in this line is an approximation to

\[
- \int_{y_L}^{y_R} \int_{z_L}^{z_R} \sigma_{xy}(x_L, y, z) v_x(x_L, y, z) \, dydz, \tag{22}
\]

with \([A_y^N \otimes A_x^N]\) playing the role of the integral operators in equation 22. Similar interpre-
tions apply to the other lines in equation 21. Altogether, we have that the right-hand side of equation 21 is an approximation to the right-hand side of equation 18. In other words, the discrete energy $E$ as defined in equation 20 mimics the behavior of the continuous energy $\mathcal{E}$ in equation 18.

Moreover, similar to the 1D case, the free surface boundary condition can also be imposed by appending penalty terms to equation 19. Taking the two $xy$-faces at $z = z_L$ and $z = z_R$ for example, the free surface boundary condition states that $\sigma_{xz} = \sigma_{yz} = \sigma_{zz} = 0$. These conditions can be imposed by modifying the updating formulas for the velocity components in equation 19 as follows:

\[
\begin{aligned}
\mathcal{A}V_x \rho V_x \frac{dV_x}{dt} &= \mathcal{A}V_x \left( D_{xx}^{xyz} \Sigma_{xx} + D_{xy}^{xyz} \Sigma_{xy} + D_{xz}^{xyz} \Sigma_{xz} \right) \\
&+ (I_x^M \otimes I_y^N \otimes \mathcal{E}_z^L) \cdot (A_x^M \otimes A_y^N) \cdot \left[(I_x^M \otimes I_y^N \otimes \mathcal{P}_z^L)^T \Sigma_{zz} - 0 \right] \\
&- (I_x^M \otimes I_y^N \otimes \mathcal{E}_z^R) \cdot (A_x^M \otimes A_y^N) \cdot \left[(I_x^M \otimes I_y^N \otimes \mathcal{P}_z^R)^T \Sigma_{zz} - 0 \right] ;
\end{aligned}
\]

\[
\begin{aligned}
\mathcal{A}V_y \rho V_y \frac{dV_y}{dt} &= \mathcal{A}V_y \left( D_{xy}^{xyz} \Sigma_{xy} + D_{yy}^{xyz} \Sigma_{yy} + D_{yz}^{xyz} \Sigma_{yz} \right) \\
&+ (I_x^N \otimes I_y^M \otimes \mathcal{E}_z^L) \cdot (A_x^N \otimes A_y^M) \cdot \left[(I_x^N \otimes I_y^M \otimes \mathcal{P}_z^L)^T \Sigma_{yy} - 0 \right] \\
&- (I_x^N \otimes I_y^M \otimes \mathcal{E}_z^R) \cdot (A_x^N \otimes A_y^M) \cdot \left[(I_x^N \otimes I_y^M \otimes \mathcal{P}_z^R)^T \Sigma_{yy} - 0 \right] ;
\end{aligned}
\]

\[
\begin{aligned}
\mathcal{A}V_z \rho V_z \frac{dV_z}{dt} &= \mathcal{A}V_z \left( D_{zz}^{xyz} \Sigma_{zz} + D_{zy}^{xyz} \Sigma_{zy} + D_{zz}^{xyz} \Sigma_{zz} \right) \\
&+ (I_x^N \otimes I_y^N \otimes \mathcal{P}_z^L) \cdot (A_x^N \otimes A_y^N) \cdot \left[(I_x^N \otimes I_y^N \otimes \mathcal{E}_z^L)^T \Sigma_{zz} - 0 \right] \\
&- (I_x^N \otimes I_y^N \otimes \mathcal{P}_z^R) \cdot (A_x^N \otimes A_y^N) \cdot \left[(I_x^N \otimes I_y^N \otimes \mathcal{E}_z^R)^T \Sigma_{zz} - 0 \right] ,
\end{aligned}
\]

where the penalty terms appended to equations 23a, 23b, and 23c correspond to conditions $\sigma_{xz} = 0$, $\sigma_{yz} = 0$, and $\sigma_{zz} = 0$, respectively. It can be verified that the penalty terms introduced above would cancel out the last six lines in equation 21 (i.e., those included in the last bracket that corresponds to the $xy$-faces) in discrete energy analysis.
We note here that appending penalty terms to the discretized system can be viewed as modifying the corresponding spatial derivative approximations. This realization will help simplify the upcoming discussion. For example, appending the penalty terms in equations 23a, 23b, and 23c can be viewed as replacing \( D_z^x \Sigma_{xz} \), \( D_z^y \Sigma_{yz} \), and \( D_z^z \Sigma_{zz} \), respectively, with their modified spatial derivative approximations defined as follows:

\[
\overline{D_z^x \Sigma_{xz}} = D_z^x \Sigma_{xz} + \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_z^N)^{-1} \mathcal{E}_z^L \right] \right\} \cdot \left[ (I_x^M \otimes I_y^N \otimes \mathcal{P}_z^L)^T \Sigma_{xz} = 0 \right] \tag{24a}
\]

\[
- \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_z^N)^{-1} \mathcal{E}_z^R \right] \right\} \cdot \left[ (I_x^M \otimes I_y^N \otimes \mathcal{P}_z^R)^T \Sigma_{xz} = 0 \right] ;
\]

\[
\overline{D_z^y \Sigma_{yz}} = D_z^y \Sigma_{yz} + \left\{ I_x^N \otimes I_y^M \otimes \left[ (A_z^N)^{-1} \mathcal{E}_z^L \right] \right\} \cdot \left[ (I_x^N \otimes I_y^M \otimes \mathcal{P}_z^L)^T \Sigma_{yz} = 0 \right] \tag{24b}
\]

\[
- \left\{ I_x^N \otimes I_y^M \otimes \left[ (A_z^N)^{-1} \mathcal{E}_z^R \right] \right\} \cdot \left[ (I_x^N \otimes I_y^M \otimes \mathcal{P}_z^R)^T \Sigma_{yz} = 0 \right] ;
\]

\[
\overline{D_z^z \Sigma_{zz}} = D_z^z \Sigma_{zz} + \left\{ I_x^N \otimes I_y^N \otimes \left[ (A_z^M)^{-1} \mathcal{P}_z^L \right] \right\} \cdot \left[ (I_x^N \otimes I_y^N \otimes \mathcal{E}_z^L)^T \Sigma_{zz} = 0 \right] \tag{24c}
\]

\[
- \left\{ I_x^N \otimes I_y^N \otimes \left[ (A_z^M)^{-1} \mathcal{P}_z^R \right] \right\} \cdot \left[ (I_x^N \otimes I_y^N \otimes \mathcal{E}_z^R)^T \Sigma_{zz} = 0 \right] ,
\]

where inverses of the norm matrices \( A^V_z, A^V_y, \) and \( A^V_z \) have been applied to the penalty terms in equations 23a, 23b, and 23c, respectively. Using the modified spatial derivative approximations introduced above, the norm matrices \( A^V_z, A^V_y, \) and \( A^V_z \) can now be omitted from the discretized system, though their impact remains through the modified spatial derivative approximations.

**Interface conditions and interpolation operators**

At the interface between two neighboring layers, we seek to impose the following interface conditions (see, e.g., Stein and Wyssen 2009, p. 52):

\[
\begin{align*}
\left\{ \begin{array}{c}
v_i^+ = v_i^- ; \\
\sigma_{ij}^+ n_j^+ + \sigma_{ij}^- n_j^- = 0 ,
\end{array} \right. \tag{25b}
\end{align*}
\]
where symbols $^+$ and $^-$ are appended to distinguish quantities from the plus and minus sides of the interface. Equation 25a stems from continuity of the elastic medium, i.e., no overlap or tear; equation 25b stems from Newton’s third law. For the discretization setting considered in this work, where the layer interfaces are horizontal, the outward normal vectors $n^+$ and $n^-$ in equation 25 are simply $[0,0,-1]^T$ and $[0,0,1]^T$, respectively, and the interface conditions in equation 25 reduce to continuity in solution variables $v_x, v_y, v_z, \sigma_{xz}, \sigma_{yz}$, and $\sigma_{zz}$, with the remaining solution variables uninvolved.

These interface conditions can also be imposed by appending proper penalty terms to the discretized system described in equation 19. Effectively, this boils down to modifying the spatial derivative approximations as shown in equations 26 and 27.

(On the plus side:)

\[
D_z^+ \Sigma_{xz}^+ \rightarrow \bar{D}_z^+ \Sigma_{xz}^+ = D_z^+ \Sigma_{xz}^+ + \frac{1}{2} \left\{ I_z^+ \otimes I_y^+ \otimes \left[(A_{-}^{+})^{-1} \epsilon_z^+ \right] \right\},
\]

\[
D_y^+ V_x^+ \rightarrow \bar{D}_y^+ V_x^+ = D_y^+ V_x^+ + \frac{1}{2} \left\{ I_x^+ \otimes I_y^+ \otimes \left[(A_{-}^{+})^{-1} P_x^+ \right] \right\},
\]

\[
D_z^+ \Sigma_{yz}^+ \rightarrow \bar{D}_z^+ \Sigma_{yz}^+ = D_z^+ \Sigma_{yz}^+ + \frac{1}{2} \left\{ I_z^+ \otimes I_y^+ \otimes \left[(A_{-}^{+})^{-1} \epsilon_z^+ \right] \right\},
\]

\[
D_y^+ V_y^+ \rightarrow \bar{D}_y^+ V_y^+ = D_y^+ V_y^+ + \frac{1}{2} \left\{ I_x^+ \otimes I_y^+ \otimes \left[(A_{-}^{+})^{-1} P_y^+ \right] \right\},
\]

\[
D_z^+ \Sigma_{zz}^+ \rightarrow \bar{D}_z^+ \Sigma_{zz}^+ = D_z^+ \Sigma_{zz}^+ + \frac{1}{2} \left\{ I_z^+ \otimes I_y^+ \otimes \left[(A_{-}^{+})^{-1} \epsilon_z^+ \right] \right\},
\]

\[
D_y^+ V_z^+ \rightarrow \bar{D}_y^+ V_z^+ = D_y^+ V_z^+ + \frac{1}{2} \left\{ I_x^+ \otimes I_y^+ \otimes \left[(A_{-}^{+})^{-1} P_z^+ \right] \right\}.
\]
(On the minus side:)

\[
\begin{align*}
D_x^{\Sigma_{zz}} &\to D_x^{\Sigma_{zz}} = D_x^{\Sigma_{zz}} - \frac{1}{2} \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_y^N)^{-1} \varepsilon_z^R \right] \right\}, \\
D_y^{V_z} &\to D_y^{V_z} - \frac{1}{2} \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_y^N)^{-1} \varepsilon_z^R \right] \right\}; \\
D_z^{\Sigma_{yz}} &\to D_z^{\Sigma_{yz}} = \frac{1}{2} \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_y^N)^{-1} \varepsilon_z^R \right] \right\}, \\
D_y^{V_y} &\to D_y^{V_y} - \frac{1}{2} \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_y^N)^{-1} \varepsilon_z^R \right] \right\}; \\
D_z^{V_z} &\to D_z^{V_z} = \frac{1}{2} \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_y^N)^{-1} \varepsilon_z^R \right] \right\}, \\
D_y^{V_y} &\to D_y^{V_y} - \frac{1}{2} \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_y^N)^{-1} \varepsilon_z^R \right] \right\}; \\
D_z^{V_z} &\to D_z^{V_z} = \frac{1}{2} \left\{ I_x^M \otimes I_y^N \otimes \left[ (A_y^N)^{-1} \varepsilon_z^R \right] \right\}.
\end{align*}
\]  

(27a)

In equations 26 and 27, the continuity in \( \sigma_{zz}, v_x, \sigma_{yz}, v_y, \sigma_{zz}, \) and \( v_z \) are accounted for by equations 26a - 26f on the plus side, respectively, and by equations 27a - 27f on the minus side, respectively. Moreover, we used \( \mathcal{T} \) to denote a 2D interpolation operator acting on the interface. For example, \( \mathcal{T}_z^V \) in equation 27f operates on the projection of \( V_z^+ \) on the interface, i.e., \( (I_x^N \otimes I_y^N \otimes P_{z^+}^L) \mathcal{T}_z^V V_z^+ \), providing the approximations to \( v_z \) at locations matching those of the projection of \( V_z^- \) on the interface, i.e., \( (I_x^N \otimes I_y^N \otimes P_{z^-}^R) \mathcal{T}_z^V V_z^- \).

The 2D interpolation operators appearing in equations 26 and 27 are constructed as tensor products of 1D interpolation operators as follows:

\[
\begin{align*}
\mathcal{T}_z^{\Sigma_{zz}} &= \mathcal{T}_z^{V_z} = (I_x^M \otimes I_y^N \otimes P_{z^+}^L) \mathcal{T}_z^{V_z} = (I_x^M \otimes I_y^N); \\
\mathcal{T}_z^{\Sigma_{yz}} &= \mathcal{T}_z^{V_y} = (I_x^M \otimes I_y^N \otimes P_{z^+}^L) \mathcal{T}_z^{V_y} = (I_x^M \otimes I_y^N); \\
\mathcal{T}_z^{\Sigma_{zz}} &= \mathcal{T}_z^{V_z} = (I_x^M \otimes I_y^N \otimes P_{z^-}^R) \mathcal{T}_z^{V_z} = (I_x^M \otimes I_y^N). 
\end{align*}
\]  

(28)

To properly handle the nonconforming interface, the 1D building blocks in equation 28 are
required to satisfy the following properties:

\[
A_x^N T_x^N = (A_x^N T_x^N)^T, \quad A_x^M T_x^M = (A_x^M T_x^M)^T;
\]

\[
A_y^N T_y^N = (A_y^N T_y^N)^T, \quad A_y^M T_y^M = (A_y^M T_y^M)^T.
\] (29)

A collection of 1D interpolation operators satisfying the properties in equation 29 are presented in Appendix B, corresponding to a variety of grid spacing ratios. Such ratios are allowed to be rational numbers, instead of just integers, which improves the application flexibility of the proposed discretization setting. Similar to the construction of SBP operators, these interpolation operators are also constructed with the assistance of symbolic computing software.

If the discrete energy analysis is carried out without regard to the interface conditions, it can be shown that the remaining terms associated with the interface are:

\[
\frac{dE}{dt} = - \left[ (\Sigma^+_{xx})^T (I_x^M \otimes I_y^N \otimes P_z^L) \cdot [A_x^M \otimes A_y^N] \cdot [(T_x^M \otimes I_y^N \otimes E_z^L)^T V_x^+] \right]
\]

\[- \left[ (\Sigma^+_{xy})^T (I_x^N \otimes I_y^N \otimes P_z^L) \cdot [A_y^N] \cdot [(T_x^N \otimes I_y^N \otimes E_z^L)^T V_y^+] \right] + \left[ (\Sigma^-_{yz})^T (I_x^N \otimes I_y^N \otimes E_z^R) \cdot [A_y^N] \cdot [(T_x^N \otimes I_y^N \otimes P_z^L)^T V_y^-] \right] \] (side)

\[- \left[ (\Sigma^-_{yz})^T (I_x^N \otimes I_y^N \otimes E_z^R) \cdot [A_y^N] \cdot [(T_x^N \otimes I_y^N \otimes P_z^L)^T V_y^-] \right] + \left[ (\Sigma^+_{xz})^T (I_x^M \otimes I_y^N \otimes P_z^L) \cdot [A_x^M \otimes A_y^N] \cdot [(T_x^M \otimes I_y^N \otimes E_z^L)^T V_x^-] \right] \] (side)

\[- \left[ (\Sigma^+_{xz})^T (I_x^M \otimes I_y^N \otimes P_z^L) \cdot [A_x^M \otimes A_y^N] \cdot [(T_x^M \otimes I_y^N \otimes E_z^L)^T V_x^-] \right] + \left[ (\Sigma^-_{yz})^T (I_x^N \otimes I_y^N \otimes E_z^R) \cdot [A_y^N] \cdot [(T_x^N \otimes I_y^N \otimes P_z^L)^T V_y^-] \right].
\] (30)

Instead, carrying out the discrete energy analysis with the modifications presented in equations 26 - 27 and the properties in equation 29, it can be shown that the remaining terms in equation 30 are cancelled out without introducing any extra term, i.e., the energy conserving property is preserved across the interface.

To illustrate, let us consider the pairs \(\{\Sigma^+_{xx}; V_x^+\}\) and \(\{\Sigma^-_{xx}; V_x^-\}\) as an example. The remaining terms in equation 30 associated with these pairs are at the third and the sixth
lines, respectively. Now, carrying out the discrete energy analysis with the modified derivative approximations presented in equations 26e, 26f, 27e, and 27f, the remaining terms in \( \frac{dE}{dt} \) associated with these pairs become

\[
\begin{align*}
\text{31a} & : \quad - \left[ (\Sigma_{zz}^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L)^T V_z^+ \right] \\
\text{31b} & : \quad + \left[ (\Sigma_{zz}^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L)^T V_z^- \right] \\
\text{31c} & : \quad + \frac{1}{2} \left[ (V_z^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N)^T \Sigma_{zz}^+ \right] \\
\text{31d} & : \quad - \frac{1}{2} \left[ (V_z^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L) \right] \cdot \left[ (A_{zz}^N \otimes \mathcal{E}_{zz}^N) \otimes (A_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N)^T \Sigma_{zz}^- \right] \\
\text{31e} & : \quad + \frac{1}{2} \left[ (\Sigma_{zz}^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L)^T V_z^- \right] \\
\text{31f} & : \quad - \frac{1}{2} \left[ (\Sigma_{zz}^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N)^T \Sigma_{zz}^- \right] \\
\text{31g} & : \quad + \frac{1}{2} \left[ (V_z^-)^T (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N)^T \Sigma_{zz}^- \right] \\
\text{31h} & : \quad - \frac{1}{2} \left[ (V_z^-)^T (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L) \right] \cdot \left[ (A_{zz}^N \otimes \mathcal{E}_{zz}^N) \otimes (A_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N)^T \Sigma_{zz}^- \right] \\
\text{31i} & : \quad + \frac{1}{2} \left[ (\Sigma_{zz}^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L)^T V_z^+ \right] \\
\text{31j} & : \quad - \frac{1}{2} \left[ (\Sigma_{zz}^+)^T (I_{zz}^N \otimes I_{zz}^N \otimes \mathcal{E}_{zz}^N) \right] \cdot \left[ A_{zz}^N \otimes A_{zz}^N \right] \cdot \left[ (I_{zz}^N \otimes I_{zz}^N \otimes P_{zz}^L)^T V_z^- \right]
\end{align*}
\]

where equations 31a and 31b are copied from the third and the sixth lines of equation 30, respectively; equations 31c and 31d stem from the modification in equation 26e; equations 31e and 31f stem from the modification in equation 26f; equations 31g and 31h stem from the modification in equation 27e; equations 31i and 31j stem from the modification in equation 27f. It can be verified easily that in equation 31, terms underlined with lines of the same style cancel out. The properties in equation 29 are needed for the terms underlined with the dashed and dotted lines to cancel out, respectively.

To sum up, all remaining terms in equation 30 associated with the pairs \( \{ \Sigma_{zz}^+, V_z^+ \} \) and \( \{ \Sigma_{zz}^-, V_z^- \} \) are cancelled out without introducing new terms because of the modifications made in equations 26e, 26f, 27e, and 27f. Similar observations can be made for the pairs
\{\Sigma_{zz}^{+}; V_{z}^{+}\} \text{ and } \{\Sigma_{zz}^{-}; V_{z}^{-}\}, \text{ which correspond to the first and fourth lines in equation 30, respectively, and for the pairs } \{\Sigma_{yz}^{+}; V_{y}^{+}\} \text{ and } \{\Sigma_{yz}^{-}; V_{y}^{-}\}, \text{ which correspond to the second and fifth lines in equation 30, respectively.}

In practical implementation, the following procedure can be used to arrive at the modified derivative approximations presented in equations 26 and 27, illustrated using equation 27f as a particular example, which concerns the approximation of \(\frac{\partial v}{\partial z}\) on the minus side of the interface.

1) Apply \(D_{z}^{V_{z}^{-}}\) to \(V_{z}^{-}\) and update the derivative approximation; essentially, this boils down to applying the 1D difference operator \(D_{z}^{M}\) to each \(z\)-direction grid column on the minus side of the interface (see equation 16 for the composition of \(D_{z}^{V_{z}^{-}}\)).

2) Project \(V_{z}^{-}\) to the interface from the minus side; essentially, this boils down to applying the 1D projection operator \(P_{z}^{R}\) to each \(z\)-direction grid column on the minus side of the interface.

3) Project \(V_{z}^{+}\) to the interface from the plus side; essentially, this boils down to applying the 1D projection operator \(P_{z}^{L}\) to each \(z\)-direction grid column on the plus side of the interface.

4) Interpolate the projected values from step 3) to match the positions of the projected values from step 2); this step requires the 2D interpolation operator \(T^{V_{z}^{+}}\) presented in equation 28.

5) Penalize the differences of the values obtained from steps 2) and 4) and update the derivative approximations; essentially, this boils down to applying \((A_{z}^{N})^{-1}\bar{e}_{z}^{R}\) on each difference.
We note here that order of the above steps does not need to be followed strictly in actual implementation. For example, step 1) and steps 2-5) can be carried out independently. Their relative order can be arranged flexibly to improve the computation and communication sequence in parallel computing environment.

**NUMERICAL EXAMPLES**

In this section, two numerical examples are presented to corroborate the above analytical development, one concerning a layer-wise homogeneous model and the other concerning the overthrust model (Aminzadeh et al., 1997).

To demonstrate the energy conserving property, periodic boundary conditions are considered for the two horizontal directions (i.e., the $x$- and $y$-directions) while free surface boundary conditions are considered for the top and bottom boundaries since there is no energy loss associated with these boundary conditions. The periodic boundary conditions are addressed using the standard practice of wrapping the stencils around the boundaries. The staggered leapfrog scheme is used for temporal discretization.

The numerical experiments are conducted on a Cray XC-40 distributed-memory system, with each node containing dual 16-core Intel Haswell processors running at 2.3 GHz and 128 GB of DDR4 memory, connected via an Aries network.

**Layered model**

In this example, we consider a model consisting of four homogeneous layers. The relevant physical and numerical parameters associated with these layers (from top to bottom) are displayed in Table 1.
The source profile is specified as the Ricker wavelet with central frequency at 10 Hz. The maximum frequency in the wave content is assumed to be 25 Hz, which is used to calculate the $N_{ppw}$ (i.e., number of discretization grid points per minimum wavelength) shown in Table 1. The grid spacing ratios at the layer interfaces are 1:2, 2:3, and 1:3, respectively, from top to bottom. The temporal step size $\Delta t$ is specified as $\sim 5.132 \times 10^{-4}$ s, which corresponds to equation A-5 with $\text{CFL} = 0.8$. (We note here that the maximum constant $\text{CFL}$ associated with the interior stencil $[1/2^4, -9/8, 9/8, -1/2^4]/\Delta x$ is $6/7 \approx 0.857$; see Levander 1988.)

For comparison, the uniform grid simulation with $\Delta x = 1$ for all layers is performed. The temporal step size used in the uniform grid simulation is $\sim 9.2376 \times 10^{-5}$ s, which corresponds to equation A-3 with $\text{CFL} = 0.8$. Between the two simulations, the ratio of spatial discretization points is $\sim 3.438$; the ratio of temporal discretization points is $\sim 5.556$; the ratio of total spatial-temporal discretization points is $\sim 19.1$.

The size of the simulation domain is $1080 \text{ m} \times 1080 \text{ m} \times 1080 \text{ m}$, with the depth for each layer being $270 \text{ m}$. A compressional source is applied on the normal stress components. The $V_z$ component of the velocity is recorded for comparison. The source and receiver are placed at $540 \text{ m}$ apart from each other on both the $x$- and $y$-directions, and at $5 \text{ m}$ (i.e., 5 grid points) and $4.5 \text{ m}$ (i.e., 4.5 grid points) below the top boundary, respectively.

The recorded seismograms are displayed in Figure 4, where we observe satisfactory
agreement between these two simulations. (We note here that the first second is omitted in Figure 4 since solutions at the receiver location are still at rest.) After projecting these seismograms onto a common set of equidistant time instances spanning the interval \([1 \text{s}, 3.5 \text{s}]\) with increment 0.001 s using cubic spline interpolation, the relative difference measured in \(\ell_2\) norm is \(\sim 3.01 \times 10^{-2}\).

A further verification is presented in Figure 5, where three times finer grid spacing is used in both simulations. Slightly improved agreement can be observed in Figure 5 (e.g., inside the cyan box). Segments of Figures 4 and 5 between 3.05 s and 3.25 s (corresponding to the cyan box in Figure 5) are plotted in isolation and provided in the supplementary material, where the improved agreement can be observed more easily. After applying the same interpolation as did for the seismograms in Figure 4, the relative difference measured in \(\ell_2\) norm is \(\sim 8.24 \times 10^{-3}\).

Moreover, evolutions of the discrete energy in both simulations are displayed in Figure 6, which remain constant after the source term tampers off and hence corroborate the energy conserving property. We note here that as in Figures 4 and 5, the first second is omitted in Figure 6, which includes the period when the source terms take effect and inject energy to the simulations.

[Figure 6 about here.]

Additionally, snapshots of the vertical cross sections of the wavefield \((V_z\) component) at the middle of the \(x\)-direction and parallel to the \(yz\)-plane are illustrated in Figure 7 for both nonuniform (left) and uniform (right) simulations. We discern no obvious difference between the two snapshots. These snapshots are produced with the sources placed in the center of the \(xy\)-plane, near the top surface, and taken at \(\sim 0.739\) s.
In this example, the nonuniform and uniform simulations are performed with 6785 and 23328 processor cores, respectively. In the nonuniform simulations, the number of processor cores assigned to each layer (top to bottom) are $18^3$, $9^3$, $6^3$, and $2^3$, respectively. In the uniform simulations, the number of processor cores are 18, 18, and 72 along the $x$-, $y$-, and $z$-directions, respectively.

For the two simulations whose results are shown in Figure 4, each processor core is responsible for a subdomain with $60 \times 60 \times 15$ $M$-grid points. The elapsed wall clock times (averaged over 5 runs) are $\sim 160$ s and $\sim 872$ s, respectively, translating to a ratio of $\sim 18.7$ in consumed computational resources (i.e., number of processor cores multiplied by elapsed time), which agrees well with the ratio of the total number of spatial-temporal discretization points in these two simulations (19.1). This suggests that the overhead due to the nonconforming interfaces in the nonuniform simulation is little, at $\sim 2\%$ on a per core basis in this case. The above expenditure information on computing resources are collected in Table 2.

We note here that aside from the algorithms, simulation performance also depends on implementation quality and runtime hardware environment. The elapsed times reported here are meant only to provide supporting evidence for our algorithmic intuition that the nonuniform simulations should incur little-to-no overhead. However, and in general, they should not be taken as strict proof. A few more performance related tests are included in the supplementary material, which provides further support.
For the two simulations whose results are shown in Figure 5, each processor core is responsible for a subdomain with $180 \times 180 \times 45$ M-grid points. The elapsed wall clock times are $\sim 11805$ s and $\sim 63482$ s, respectively, which translates to a ratio of $\sim 18.5$ in consumed computational resources, suggesting an overhead for the nonuniform simulation at $\sim 3\%$. (We note here that the two numbers on elapsed wall clock times reported in this paragraph, and similar numbers reported later for the overthrust model, are not averaged over repeated runs due to the large amount of computational resources required.) A table collecting the above expenditure information on computing resources can be found in the supplementary material. It is omitted here to conserve space.

**Overthrust model**

In this example, we test the proposed layer-wise discretization with the overthrust model, whose compressional wave speed is illustrated in Figure 8. The model is described by $801 \times 801 \times 181$ grid points. Grid spacing between two neighboring parameter grid points is assigned to be 10 m. Since the overthrust model contains only the information of the compressional wave speed $c_p$, we synthesize the shear wave speed $c_s$ by prescribing the $c_p/c_s$ ratio, which decreases from top (4.5) to bottom (1.5) linearly. (We note here that the $c_p/c_s$ ratio tends to decrease with depth and increased degree of consolidation in rocks; see Gregory 1976; Hamilton 1979, for example.) Moreover, the density $\rho$ is also synthesized, which increases from top (2000 kg/m$^3$) to bottom (2700 kg/m$^3$) linearly.

[Figure 8 about here.]

In nonuniform grid simulation, the entire simulation domain is separated into three layers of equal depth (600 m). The most relevant physical and numerical parameters associated
with these layers are displayed in Table 3. The source profile is specified as the Ricker wavelet with central frequency at 5 Hz. The maximum frequency in the wave content is assumed to be 12.5 Hz, which is used to calculate the $N_{ppw}$ shown in Table 3. The grid spacing ratios at both layer interfaces are 1:2. The temporal step size $\Delta t$ is specified as $\sim 2.3072 \times 10^{-4}$ s, which corresponds to equation A-5 with $C_{CFL} = 0.8$. For comparison, the uniform grid simulation with $\Delta x = 2.5$ for all layers is performed. The temporal step size used in the uniform grid simulation is $\sim 1.9245 \times 10^{-4}$ s, which corresponds to equation A-3 with $C_{CFL} = 0.8$. Between the two simulations, the ratio of spatial discretization points is $\sim 2.63$; the ratio of temporal discretization points is $\sim 1.2$; the ratio of total spatial-temporal discretization points is $\sim 3.153$.

[Table 3 about here.]

In physical units, the size of the simulation domain is 8000 m $\times$ 8000 m $\times$ 1800 m. A compressional source is applied on the normal stress components. The $V_z$ component of the velocity is recorded for comparison. The source and receiver are placed at 4000 m apart from each other on both the $x$- and $y$-directions, and at 12.5 m (i.e., 5 grid points) and 11.25 m (i.e., 4.5 grid points) below the top boundary, respectively.

The recorded seismograms are displayed in Figure 9, where we observe satisfactory agreement between these two simulations. After projecting these seismograms onto a common set of equidistant time instances spanning the interval $[1 \text{ s}, 13.5 \text{ s}]$ with increment 0.001 s using cubic spline interpolation, the relative difference measured in $\ell_2$ norm is $\sim 5.69 \times 10^{-4}$. Moreover, evolutions of the discrete energy in both simulations are displayed in Figure 10, which remain constant after the source term tampers off and hence corroborate the energy conserving property.
For this example, the nonuniform and uniform simulations are run with 7300 and 19200 processor cores, respectively, each responsible for a subdomain with \(80 \times 80 \times 60 \) M-grid points. The elapsed wall clock times are \(\sim 10860 \) s and \(\sim 12787 \) s, respectively, translating to a ratio of \(\sim 3.095 \) in consumed computational resources (i.e., number of processors multiplied by elapsed time), which agrees well with the ratio of the total number of spatial-temporal discretization points in these two simulations (3.153), suggesting an overhead for the nonuniform simulation at \(\sim 2\%\). A table collecting the above expenditure information on computing resources can be found in the supplementary material. It is omitted here to conserve space.

**CONCLUSIONS**

Layer-wise finite difference discretization of the 3D isotropic elastic wave propagation on staggered grids with nonconforming interfaces between layers is considered in this work. It is shown that energy conserving simulation across the nonconforming layer interfaces can be achieved through the use of summation-by-parts finite difference operators, simultaneous approximation terms, and compatible interpolation operators, which ensure the stability of the simulation. Moreover, by examining the standard process of determining the discretization parameters for seismic wave simulations under the usual constraints, as presented in Appendix A, and by the numerical examples, it is demonstrated that, compared to finite difference discretization on uniform grids, significant reduction of computational resources can be achieved by employing the proposed layer-wise discretization approach, thus enabling
more efficient processing of seismic data.

ACKNOWLEDGMENTS

The authors thank the editors and referees for carefully reviewing this article and providing valuable suggestions. This research used resources of the Core Labs of King Abdullah University of Science and Technology (KAUST). This research was funded by KAUST award OSR-2019-CCF-3666.4 and U.S. National Science Foundation Frontera award 2033468.

APPENDIX A

IMPACT OF WAVE SPEED VARIATION ON DISCRETIZATION PARAMETERS

In the following, we use a simple thought experiment to illustrate that, for a simulation domain composed of different subterranean media, a layer-wise spatial discretization grid can significantly reduce the expenditure in computational resources compared to its fully uniform counterpart.

First, considering a uniformly discretized simulation domain associated with a medium whose wave speed ranges from $c_{\text{min}}$ to $c_{\text{max}}$, the minimum wavelength of the simulated waves can be expressed as

$$\lambda_{\text{min}} = \frac{c_{\text{min}}}{f}$$

(A-1)

where $f$ denotes the highest frequency in the wave content being simulated. Assuming that the spatial grid spacing $\Delta x$ is determined by prescribing $N_{\text{ppw}}$ grid points to resolve the
minimum wavelength $\lambda_{\text{min}}$, we have that

$$\Delta x = \frac{\lambda_{\text{min}}}{N_{\text{ppw}}} = \frac{c_{\text{min}}}{f \cdot N_{\text{ppw}}}. \quad (A-2)$$

Further assuming that the temporal step size $\Delta t$ is determined by the CFL constraint, we can write

$$\Delta t = C_{\text{CFL}} \cdot \frac{\Delta x}{c_{\text{max}}} = C_{\text{CFL}} \cdot \frac{c_{\text{min}}}{f \cdot N_{\text{ppw}} \cdot c_{\text{max}}}, \quad (A-3)$$

where $C_{\text{CFL}}$ is a constant that depends on the spatial and temporal discretization schemes, but is independent of the underlying medium.

Regarding $f$, $N_{\text{ppw}}$, and $C_{\text{CFL}}$ as fixed constants in the above formulas, we observe from equation A-2 that the spatial grid spacing $\Delta x$ is proportional to the minimum wave speed $c_{\text{min}}$ of the underlying medium; moreover, we observe from equation A-3 that the temporal step size $\Delta t$ is constrained by the maximum contrast in the wave speed, i.e., $c_{\text{max}}/c_{\text{min}}$, of the underlying medium.

On the other hand, if the same simulation domain is separated into $N_L$ layers, each discretized uniformly by prescribing $N_{\text{ppw}}$ grid points to resolve the minimum wavelength within the layer, the spatial grid spacing for each layer can be expressed as

$$\Delta x^i = \frac{\lambda_{\text{min}}^i}{N_{\text{ppw}}} = \frac{c_{\text{min}}^i}{f \cdot N_{\text{ppw}}}, \quad i = 1 \cdots N_L, \quad (A-4)$$

where $\Delta x^i$, $\lambda_{\text{min}}^i$, and $c_{\text{min}}^i$ stand for the spatial grid spacing, minimum wavelength, and minimum wave speed of the $i$th layer. In other words, the spatial grid spacing is now only limited by the minimum wave speed within the layer, instead of the minimum wave speed of the entire simulation domain. Moreover, assuming that constant $C_{\text{CFL}}$ remains unchanged, the temporal step size is now determined by the most stringent CFL constraint among all
layers, i.e.,

\[
\Delta t = \mathcal{C}_{\text{CFL}} \cdot \min_{i=1 \cdots N_L} \left\{ \frac{\Delta x_i}{c_{i \text{max}}} \right\} = \frac{\mathcal{C}_{\text{CFL}}}{f \cdot N_{\text{ppw}}} \cdot \min_{i=1 \cdots N_L} \left\{ \frac{c_{i \text{min}}}{c_{i \text{max}}} \right\},
\]

which is likely to be relaxed from, at least not worse than, that of equation A-3.

[Figure 11 about here.]

To give a concrete example, consider the simulation domain depicted in Figure 11, which consists of five regions of equal depth with progressively increasing wave speeds. The normalized minimum and maximum wave speeds of each region are displayed in the figure. Figure 11a depicts the case that the entire simulation domain is discretized uniformly, which is used as the reference for comparison. The total number of its space-time discretization grid points is counted as

\[
(1 + 1 + 1 + 1 + 1) \times 6 = 30 \text{ units,}
\]

where the quantity inside the parentheses corresponds to the total spatial discretization grid points, while the number 6 is related to the temporal discretization grid points, stemming from the maximum contrast in the wave speed of the entire simulation domain; see equation A-3.

Figure 11b depicts the case that the entire simulation domain is separated into two layers, each discretized uniformly and with a 1:3 ratio in \( \Delta x \); the total number of space-time discretization grid points becomes

\[
(1 + 1 + \frac{1}{27} + \frac{1}{27} + \frac{1}{27}) \times 3 \approx 6.3333 \text{ units,}
\]

where the number 3 stems from the maximum contrast in the wave speed of the top layer;
see equation A-5. By comparing equation A-7 to equation A-6, we observe that the total number of space-time discretization grid points has been reduced by a factor of $\sim 4.7$.

Figure 11c depicts the case that the entire simulation domain is separated into three layers, each discretized uniformly and with a 1:2:4 ratio in $\Delta x$; the total number of space-time discretization grid points becomes

$$\left(1 + \frac{1}{8} + \frac{1}{8} + \frac{1}{64} + \frac{1}{64}\right) \times 2 = 2.5625 \text{ units}, \quad (A-8)$$

where the number 2 stems from the maximum contrast in the wave speed of the top layer (or the middle layer); see equation A-5. By comparing equation A-8 to equation A-6, we observe that the total number of space-time discretization grid points has been reduced by a factor of $\sim 11.7$.

We note here that for realistic earth media, it rarely happens that different layers can be separated in such a non-overlapping fashion as shown in Figure 11, where the maximum wave speed in one layer matches the minimum wave speed of the next. Instead, there are often overlaps in the wave speed ranges of two neighboring layers. In such a situation, the cost saving factors estimated above can be overly optimistic. On the other hand, instead of the equal-depth model considered in Figure 11, a relatively thin layer of low velocity surface material (e.g., sand and soil) often appears in realistic situations, which can severely restrict the choices of discretization parameters in uniform grid simulations. In such cases, the actual cost saving factors can be more pronounced than those estimated above. Finally, addressing layer interfaces can incur computational overheads in simulations, which may also affect the actual cost saving factors. However, impact from such overheads tends to be insignificant for large simulations at realistic scales, as evidenced by the numerical examples in section Numerical Examples.
APPENDIX B

COMPATIBLE INTERPOLATION OPERATORS

A collection of 1D interpolation operators for a variety of grid spacing ratios are presented in the following. They are building blocks for the 2D interpolation operators used in section Interface conditions and interpolation operators to address the nonconforming interfaces. These interpolation operators are referred to as being compatible (with each other), since, in addition to the typical approximation requirements, they also satisfy the reciprocal relationship postulated in equation 29.

Because of this reciprocal relationship, only the interpolation operators from the coarse grid (minus side) to the fine grid (plus side) are presented since the interpolation operators for the opposite direction can be deduced based on equation 29. Furthermore, only formulas for a few exemplary grid points within a repeating cycle are shown in the following, while formulas for the other grid points can be replicated by shifting the indices accordingly.

1:2 grid spacing ratio

[Figure 12 about here.]

\[ T_{x_2}^{N} : \]

\[
\begin{align*}
 f(x_0^+) & \leftarrow f(x_0^-) \\
 f(x_1^+) & \leftarrow -\frac{1}{16}f(x_{-1}) + \frac{9}{16}f(x_0^-) + \frac{9}{16}f(x_1^-) - \frac{1}{16}f(x_2^-) \\
 f(x_2^+) & \leftarrow f(x_1^-) \\
\end{align*}
\] (B-1)

\[ T_{x_2}^{M} : \]

\[
\begin{align*}
 f(x_0^+) & \leftarrow \frac{3}{32}f(x_{-1}) + \frac{15}{32}f(x_0^-) - \frac{3}{32}f(x_2^-) \\
 f(x_1^+) & \leftarrow -\frac{3}{32}f(x_{-1}) + \frac{15}{32}f(x_1^-) + \frac{9}{32}f(x_2^-) \\
\end{align*}
\] (B-2)
1:3 grid spacing ratio

\[ T_{x}^{N} : \]
\[ f(x_0^+) \leftarrow f(x_0) \]
\[ f(x_1^-) \leftarrow -\frac{1}{2} f(x_2^-) + \frac{3}{4} f(x_1^-) \]
\[ f(x_2^-) \leftarrow \frac{3}{8} f(x_0) + \frac{3}{4} f(x_1^-) - \frac{1}{4} f(x_2^-) \]
\[ f(x_3^-) \leftarrow f(x_1^-) \]

\[ T_{x}^{M} : \]
\[ f(x_0^+) \leftarrow \frac{3}{4} f(x_2^-) + \frac{5}{8} f(x_0) - \frac{1}{4} f(x_1^-) \]
\[ f(x_1^-) \leftarrow f(x_0) \]
\[ f(x_2^-) \leftarrow -\frac{1}{2} f(x_2^-) + \frac{3}{4} f(x_0) + \frac{3}{4} f(x_1^-) \]

2:3 grid spacing ratio

\[ T_{x}^{N} : \]
\[ f(x_0^+) \leftarrow f(x_0) \]
\[ f(x_1^-) \leftarrow -\frac{9}{16} f(x_2^-) + \frac{9}{16} f(x_1^-) + \frac{1}{4} f(x_1^-) - \frac{1}{4} f(x_2^-) + \frac{3}{8} f(x_3^-) \]
\[ f(x_2^-) \leftarrow \frac{1}{16} f(x_0) + \frac{1}{16} f(x_1^-) + \frac{1}{4} f(x_1^-) - \frac{1}{4} f(x_2^-) - \frac{3}{8} f(x_3^-) \]
\[ f(x_3^-) \leftarrow f(x_1^-) \]

\[ T_{x}^{M} : \]
\[ f(x_0^+) \leftarrow \frac{1}{16} f(x_2^-) + \frac{1}{16} f(x_0) + \frac{1}{16} f(x_1^-) + \frac{1}{4} f(x_1^-) - \frac{1}{4} f(x_2^-) \]
\[ f(x_1^-) \leftarrow -\frac{9}{16} f(x_2^-) + \frac{9}{16} f(x_0) + \frac{1}{4} f(x_1^-) - \frac{1}{4} f(x_2^-) \]
\[ f(x_2^-) \leftarrow \frac{1}{16} f(x_0) + \frac{1}{16} f(x_1^-) + \frac{1}{4} f(x_1^-) - \frac{1}{4} f(x_2^-) \]

1:4 grid spacing ratio

\[ T_{x}^{N} : \]
\[ f(x_0^+) \leftarrow f(x_0) \]
\[ f(x_1^-) \leftarrow -\frac{3}{4} f(x_2^-) + \frac{15}{16} f(x_1^-) + \frac{3}{16} f(x_1^-) \]
\[ f(x_2^-) \leftarrow -\frac{1}{4} f(x_2^-) + \frac{9}{16} f(x_0) + \frac{1}{4} f(x_1^-) - \frac{1}{4} f(x_2^-) \]
\[ f(x_3^-) \leftarrow \frac{9}{16} f(x_0) + \frac{1}{4} f(x_1^-) - \frac{3}{16} f(x_2^-) \]
\[ T_{x^M} : \]
\[
f(x_0^+) \leftarrow \frac{11}{120} f(x_{-1}) + \frac{5}{6} f(x_0^-) - \frac{1}{12} f(x_1^-) \\
f(x_1^+) \leftarrow \frac{9}{120} f(x_{-1}) + \frac{6}{5} f(x_0^-) - \frac{7}{12} f(x_1^-) \\
f(x_2^+) \leftarrow - \frac{11}{120} f(x_{-1}) + \frac{5}{6} f(x_0^-) + \frac{1}{12} f(x_1^-) \\
f(x_3^+) \leftarrow - \frac{11}{120} f(x_{-1}) + \frac{5}{6} f(x_0^-) + \frac{1}{12} f(x_1^-) \\
\]

(B-8)

3:4 grid spacing ratio

\[
T_{x^N}^N : \]
\[
f(x_0^+) \leftarrow f(x_0^-) \\
f(x_1^+) \leftarrow - \frac{11}{120} f(x_{-1}) + \frac{5}{6} f(x_0^-) + \frac{2}{3} f(x_1^-) - \frac{5}{3} f(x_2^-) + \frac{1}{12} f(x_3^-) \\
f(x_2^+) \leftarrow \frac{11}{120} f(x_{-1}) - \frac{5}{6} f(x_0^-) + \frac{2}{3} f(x_1^-) - \frac{5}{3} f(x_2^-) - \frac{1}{12} f(x_3^-) \\
f(x_3^+) \leftarrow \frac{11}{120} f(x_{-1}) - \frac{5}{6} f(x_0^-) + \frac{2}{3} f(x_1^-) - \frac{5}{3} f(x_2^-) - \frac{1}{12} f(x_3^-) \\
\]

(B-9)

\[
T_{x^M}^M : \]
\[
f(x_0^+) \leftarrow \frac{35}{576} f(x_{-1}) + \frac{389}{576} f(x_0^-) - \frac{1}{12} f(x_1^-) + \frac{11}{112} f(x_2^-) \\
f(x_1^+) \leftarrow - \frac{11}{112} f(x_{-1}) + \frac{389}{576} f(x_0^-) + \frac{3}{2} f(x_1^-) - \frac{335}{576} f(x_2^-) + \frac{7}{8640} f(x_3^-) \\
f(x_2^+) \leftarrow \frac{7}{8640} f(x_{-1}) - \frac{335}{576} f(x_0^-) + \frac{3}{2} f(x_1^-) + \frac{111}{576} f(x_2^-) - \frac{7}{8640} f(x_3^-) \\
f(x_3^+) \leftarrow \frac{11}{112} f(x_0^-) - \frac{1}{12} f(x_1^-) + \frac{389}{576} f(x_2^-) + \frac{35}{576} f(x_3^-) \\
\]

(B-10)

We note here that the above interpolation formulas are designed such that the leading error terms in the approximations are of at least third order. However, even under this condition, they are not the unique choice for compatible interpolation operators. Taking the 2:3 grid spacing ratio for example, the interpolation operators shown in B-11 and B-12 are also suitable choices.

\[
T_{x^N}^N : \]
\[
f(x_0^+) \leftarrow - \frac{1}{90} f(x_{-2}) + \frac{1}{24} f(x_{-1}) + \frac{17}{18} f(x_0^-) + \frac{1}{24} f(x_1^-) - \frac{1}{5} f(x_2^-) \\
f(x_1^+) \leftarrow - \frac{1}{24} f(x_{-1}) + \frac{17}{18} f(x_0^-) + \frac{1}{24} f(x_1^-) - \frac{1}{5} f(x_2^-) \\
f(x_2^+) \leftarrow - \frac{17}{18} f(x_0^-) + \frac{217}{288} f(x_1^-) + \frac{19}{288} f(x_2^-) - \frac{1}{36} f(x_3^-) \\
f(x_3^+) \leftarrow - \frac{19}{288} f(x_0^-) + \frac{217}{288} f(x_1^-) + \frac{19}{288} f(x_2^-) + \frac{1}{36} f(x_3^-) - \frac{1}{90} f(x_4^-) \\
\]

(B-11)
APPENDIX C

1D SBP FINITE DIFFERENCE OPERATORS

The following 1D SBP finite difference operators in equations C-1 and C-2 are used as the building blocks for the 3D SBP operators presented in section SBP finite difference operators on 3D staggered grids. We note here that these operators correspond to the case of unit grid spacing, i.e., $\Delta x = 1$. For general cases, $D^N$ and $D^M$ need to be scaled by $1/\Delta x$ while $A^N$ and $A^M$ need to be scaled by $\Delta x$.

\[
D^N = \begin{bmatrix}
-\frac{79}{78} & \frac{27}{26} & -\frac{1}{26} & \frac{1}{78} & 0 \\
\frac{2}{21} & -\frac{9}{7} & \frac{9}{7} & -\frac{2}{21} & 0 \\
\frac{1}{75} & 0 & -\frac{27}{25} & \frac{83}{75} & -\frac{1}{25} \\
\frac{1}{24} & -\frac{9}{8} & -\frac{9}{8} & -\frac{1}{24} \\
\frac{1}{24} & -\frac{9}{8} & \frac{9}{8} & -\frac{1}{24} \\
\frac{1}{24} & -\frac{9}{8} & \frac{9}{8} & -\frac{1}{24} \\
\frac{1}{24} & -\frac{9}{8} & \frac{9}{8} & -\frac{1}{24} \\
\frac{1}{25} & -\frac{83}{75} & \frac{27}{25} & 0 & -\frac{1}{75} \\
0 & \frac{2}{21} & -\frac{9}{7} & \frac{9}{7} & -\frac{2}{21} \\
0 & -\frac{1}{75} & \frac{1}{26} & -\frac{27}{26} & \frac{79}{78}
\end{bmatrix}; \quad (C-1a)
\]
With the above operators, matrix $Q$, defined as $A^N D^M + (A^M D^N)^T$ in equation 10
takes the following form:

\[
\begin{bmatrix}
-\frac{15}{8} & \frac{5}{4} & -\frac{3}{8} \\
3/8 & -\frac{5}{4} & \frac{15}{8}
\end{bmatrix}
\]  

(C-3)

which can be expressed as \(E^L (P^L)^T + E^R (P^R)^T\) with \(E^L, E^R, P^L,\) and \(P^R\) given by

\[
E^L = \begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix}, \quad E^R = \begin{bmatrix}
0 \\
\vdots \\
0 \\
1
\end{bmatrix}, \quad P^L = \begin{bmatrix}
\frac{15}{8} \\
-\frac{5}{4} \\
\frac{3}{8} \\
0 \\
\vdots \\
0 \\
\frac{3}{8} \\
-\frac{5}{4} \\
\frac{15}{8}
\end{bmatrix}, \quad P^R = \begin{bmatrix}
0 \\
\vdots \\
0 \\
\frac{3}{8} \\
-\frac{5}{4} \\
\frac{15}{8}
\end{bmatrix}
\]  

(C-4)

respectively.
REFERENCES


———, 2020a, Explicit coupling of acoustic and elastic wave propagation in finite-difference simulations: Geophysics, 85, T293–T308.


Martin, B., B. Fornberg, and A. St-Cyr, 2015, Seismic modeling with radial-basis-function-generated finite differences: Geophysics, 80, T137–T146.


Physics, 346, 572–589.


Wilcox, L. C., G. Stadler, C. Burstedde, and O. Ghattas, 2010, A high-order discontinuous


LIST OF FIGURES

1 Illustration of the layer-wise uniform grid discretization setting: Each individual layer is discretized uniformly; different layers are discretized using different grid spacings. .................................................. 46

2 Illustration of the grid staggering. ................................................................. 47

3 Illustration of a set of 1D staggered grids. ....................................................... 48

4 Layered model: Recorded seismograms from nonuniform and uniform grid simulations. ................................................................. 49

5 Layered model: Recorded seismograms from nonuniform and uniform grid simulations using grid spacings three times finer than those used for Figure 4. Slight improvement can be noticed inside the cyan box. ......................... 50

6 Layered model: Evolution of the discrete energy in nonuniform and uniform grid simulations. ................................................................. 51

7 Layered model: Snapshots of the vertical cross sections of the wavefield ($V_z$ component) taken at the middle of the $x$-direction and $\sim$0.739 s. a): nonuniform simulation; b): uniform simulation. These snapshots are produced with the sources placed in the center of the $xy$-plane, near the top surface. ................ 52

8 Compressional wave speed of the overthrust model. Portions of the model are carved out to illustrate the internal structures. Colorbar represents value of the compressional wave speed in unit m/s. ........................................ 53

9 Overthrust model: Recorded seismograms from nonuniform and uniform grid simulations. ................................................................. 54

10 Overthrust model: Evolution of the discrete energy in nonuniform and uniform grid simulations. ................................................................. 55

11 A simple model to illustrate the potential benefit of layer-wise uniform discretization over uniform discretization for wave simulations in earth media. The same simulation domain is considered in all three cases: a) the entire simulation domain is discretized uniformly; b) the entire simulation domain is separated into two layers, each discretized uniformly; c) the entire simulation domain is separated into three layers, each discretized uniformly. The normalized minimum and maximum wave speeds of each uniform discretization region are colored in blue and red, respectively; the former determines the spatial grid spacing $\Delta x$, while their ratio constrains the temporal step size $\Delta t$. ................................................................. 56

12 Illustration of the grid points involved in formulas B-1 and B-2 for interpolation operators $T_N^x$ (left) and $T_M^x$ (right). The grid spacing ratio is 1:2. ........ 57

13 Illustration of the grid points involved in formulas B-3 and B-4 for interpolation operators $T_N^x$ (top) and $T_M^x$ (bottom). The grid spacing ratio is 1:3. ................................................................. 58
14 Illustration of the grid points involved in formulas B-5 and B-6 for interpolation operators $T_{x, N}^N$ (top) and $T_{x, M}^M$ (bottom). The grid spacing ratio is 2:3. ................................................................. 59

15 Illustration of the grid points involved in formulas B-7 and B-8 for interpolation operators $T_{x, N}^N$ (top) and $T_{x, M}^M$ (bottom). The grid spacing ratio is 1:4. ................................................................. 60

16 Illustration of the grid points involved in formulas B-9 and B-10 for interpolation operators $T_{x, N}^N$ (top) and $T_{x, M}^M$ (bottom). The grid spacing ratio is 3:4. ................................................................. 61

17 Illustration of the grid points involved in formulas B-11 and B-12 for interpolation operators $T_{x, N}^N$ (top) and $T_{x, M}^M$ (bottom). The grid spacing ratio is 2:3. ................................................................. 62
LIST OF TABLES

1  Physical and numerical parameters associated with the layered model. . . 64
2  Expenditure on computing resources corresponding to Figure 4. The ex-
    pected ratio in total computational resources, i.e., core hours, is \( \sim 19.1 \), based
    on the amount of spatial-temporal discretization points involved. The mea-
    sured ratio is \( \sim 18.7 \), which suggests a \( \sim 2\% \) overhead on a per core basis in
    our implementation. .................................................. 65
3  Physical and numerical parameters associated with the overthrust model. . 66
Figure 1

2449x2449mm (28 x 28 DPI)
Figure 2

\[ \{\sigma_{xx}; \sigma_{yy}; \sigma_{zz}\} \]

1741x1306mm (28 x 28 DPI)
Figure 3

$N$-grid $M$-grid

$x_L$ $x_R$
Figure 4

2449x816mm (28 x 28 DPI)
Figure 5

2449x816mm (28 x 28 DPI)
Figure 6

2449x816mm (28 x 28 DPI)
Figure 7a
1286x1286mm (28 x 28 DPI)
Figure 9

2449x816mm (28 x 28 DPI)
Figure 10

2449x816mm (28 x 28 DPI)
$\Delta x : 1$

- wave-speed: 1 \rightarrow 2
- wave-speed: 2 \rightarrow 3
- wave-speed: 3 \rightarrow 4
- wave-speed: 4 \rightarrow 5
- wave-speed: 5 \rightarrow 6

Figure 11a

5080x2667mm (28 x 28 DPI)
Figure 11b

5080x2667mm (28 x 28 DPI)
Figure 11c

$\Delta x : 1$
- wave-speed: 1 - 2

$\Delta x : 2$
- wave-speed: 2 - 3
- wave-speed: 3 - 4

$\Delta x : 4$
- wave-speed: 4 - 5
- wave-speed: 5 - 6

5080x2667mm (28 x 28 DPI)
Figure 12 left

1270x244mm (28 x 28 DPI)
Figure 12 right

1270x244mm (28 x 28 DPI)
Figure 13 bottom

$1814\times244\text{mm (28 x 28 DPI)}$
Figure 13 top

1814x244mm (28 x 28 DPI)
Figure 14 top

$1814 \times 244 \text{mm (28 x 28 DPI)}$
Figure 14 bottom

1814x244mm (28 x 28 DPI)
Figure 15 top

2358x244mm (28 x 28 DPI)
Figure 15 bottom

2358x244mm (28 x 28 DPI)
Figure 16 bottom

2358x244mm (28 × 28 DPI)
Figure 16 top

2358x244mm (28 x 28 DPI)
Figure 17 bottom

1814x244mm (28 x 28 DPI)
Figure 17 top

1814x244mm (28 x 28 DPI)
Table 1: Physical and numerical parameters associated with the layered model.

<table>
<thead>
<tr>
<th>Layer</th>
<th>$c_s$</th>
<th>$c_p$</th>
<th>$\rho$</th>
<th>$\Delta x$</th>
<th>$N_{ppw}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 1</td>
<td>300 m/s</td>
<td>800 m/s</td>
<td>1600 kg/m³</td>
<td>1 m</td>
<td>12</td>
</tr>
<tr>
<td>Layer 2</td>
<td>600 m/s</td>
<td>1800 m/s</td>
<td>2100 kg/m³</td>
<td>2 m</td>
<td>12</td>
</tr>
<tr>
<td>Layer 3</td>
<td>900 m/s</td>
<td>2400 m/s</td>
<td>2300 kg/m³</td>
<td>3 m</td>
<td>12</td>
</tr>
<tr>
<td>Layer 4</td>
<td>2700 m/s</td>
<td>5000 m/s</td>
<td>2500 kg/m³</td>
<td>9 m</td>
<td>12</td>
</tr>
</tbody>
</table>
Table 2: Expenditure on computing resources corresponding to Figure 4. The expected ratio in total computational resources, i.e., core hours, is \( \sim 19.1 \), based on the amount of spatial-temporal discretization points involved. The measured ratio is \( \sim 18.7 \), which suggests a \( \sim 2\% \) overhead on a per core basis in our implementation.

<table>
<thead>
<tr>
<th></th>
<th>core count</th>
<th>elapsed time</th>
<th>total core hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>uniform</td>
<td>23328</td>
<td>872 seconds</td>
<td>( 5.651 \times 10^4 )</td>
</tr>
<tr>
<td>nonuniform</td>
<td>6785</td>
<td>160 seconds</td>
<td>( 3.016 \times 10^4 )</td>
</tr>
</tbody>
</table>
Table 3: Physical and numerical parameters associated with the overthrust model.

<table>
<thead>
<tr>
<th>Layer</th>
<th>min. $c_s$</th>
<th>max. $c_p$</th>
<th>$\Delta x$</th>
<th>$N_{ppw}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 1</td>
<td>$\sim$484.2 m/s</td>
<td>$\sim$5004.7 m/s</td>
<td>2.5 m</td>
<td>$\sim$15.5</td>
</tr>
<tr>
<td>Layer 2</td>
<td>958.5 m/s</td>
<td>5500 m/s</td>
<td>5 m</td>
<td>$\sim$15.3</td>
</tr>
<tr>
<td>Layer 3</td>
<td>1920 m/s</td>
<td>6000 m/s</td>
<td>10 m</td>
<td>$\sim$15.4</td>
</tr>
</tbody>
</table>
DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.