

On long-time instabilities in staggered finite difference simulations of the seismic acoustic wave equations on discontinuous grids

Longfei Gao, David Ketcheson and David Keyes

Division of Computer, Electrical and Mathematical Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia E-mail: longfei.gao@kaust.edu.sa

Accepted 2017 October 26. Received 2017 October 25; in original form 2017 August 4

SUMMARY

We consider the long-time instability issue associated with finite difference simulation of seismic acoustic wave equations on discontinuous grids. This issue is exhibited by a prototype algebraic problem abstracted from practical application settings. Analysis of this algebraic problem leads to better understanding of the cause of the instability and provides guidance for its treatment. Specifically, we use the concept of discrete energy to derive the proper solution transfer operators and design an effective way to damp the unstable solution modes. Our investigation shows that the interpolation operators need to be matched with their companion restriction operators in order to properly couple the coarse and fine grids. Moreover, to provide effective damping, specially designed diffusive terms are introduced to the equations at designated locations and discretized with specially designed schemes. These techniques are applied to simulations in practical settings and are shown to lead to superior results in terms of both stability and accuracy.

Key words: Instability analysis; Computational seismology; Wave propagation.

1 INTRODUCTION

Numerical simulation of wave phenomena, commonly referred to as forward modelling in geophysics, is of vital importance in seismic exploration. Computationally synthesized solutions are compared against recorded field data to drive the inversion process for the purpose of estimating earth subsurface parameters. In most modern seismic inversion approaches (*cf.* Baysal *et al.* 1983; Tarantola 1984; Luo & Schuster 1991; Symes 2008; Virieux & Operto 2009), independent forward modelling jobs need to be performed for various source terms at each inversion step, which constitutes the most computationally intensive part of the entire inversion process.

The forward modelling problem can be posed in either time domain or frequency domain. In this paper, we focus on the time domain wave equation simulation. Various discretization methods have been proposed for this task, e.g. Virieux (1984, 1986) with finite difference methods and Komatitsch & Vilotte (1998), Käser & Dumbser (2006) with finite element methods. Although the finite element approach offers more flexibility in dealing with complicated geometry and local refinement, the finite difference approach is still notably popular and widely employed in the seismic exploration community due to its efficiency.

However, the efficiency of finite difference methods on uniform grids can be significantly reduced due to large variation in the wave-speed as explained in the following. In wave simulation, the spatial sampling rate is often determined by the minimum wave speed on a point-per-wavelength basis. The temporal sampling rate is often restricted by the Courant–Friedrichs–Lewy (CFL) condition,

determined by the spatial sampling rate and the maximum wave-speed. Large variation in the wave-speed often leads to spatial oversampling in the part of the medium with large wave-speed and temporal oversampling in the part of the medium with small wave-speed. For this reason, finite difference methods on non-uniform grids are of great interest to the seismic exploration community.

Due to geological sedimentation and consolidation, wave-speed of earth media tends to increase with depth. Therefore, the types of non-uniform grids that have larger grid-spacing for deeper part of the simulation domain are of particular interest for seismic applications. Earlier attempts on finite difference simulation of seismic wave on this type of grids can be found in Hayashi *et al.* (2001), Kristek *et al.* (2010), Zhang *et al.* (2013), etc., and the references therein. Unfortunately, these earlier attempts encounter the so-called long-time instability issue, which appears only at the late stage of the simulation and has no visible influence at the early stage. To the knowledge of the authors, the cause of this long-time instability has not yet been fully understood. Various techniques have been devised to control this long-time instability, including spatial averaging (Hayashi *et al.* 2001), spatial filtering (Kristek *et al.* 2010; Zhang *et al.* 2013) and temporal filtering (Gao *et al.* 2016). In this work, we analyse this long-time instability issue from an algebraic perspective and employ the discrete energy method (*cf.* Kreiss & Scherer 1974; Carpenter *et al.* 1999; Nordström 2006; Mattsson *et al.* 2008) to devise a numerical treatment.

In the rest of this paper, we first present the mathematical abstraction of the underlying physical problem and demonstrate the long-time instability issue in Section 2. In Section 3, we analyse the

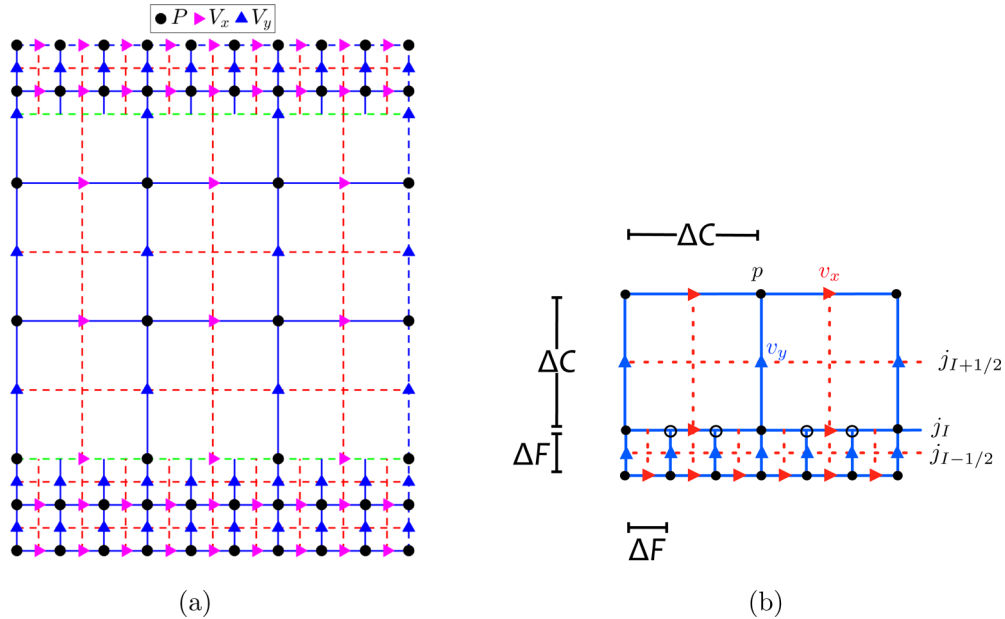


Figure 1. Non-uniform staggered grid for acoustic wave equation. The P grid occupies the integer indices while the v_x and v_y grids are shifted to the right and the top, respectively, for half grid-spacing. (a) Global view of the entire grid. (b) Local view near the lower interface.

discretized system with the discrete energy method. In Section 4, we present numerical examples to demonstrate the effectiveness of the techniques devised in this paper. Finally, we conclude this paper with Section 5.

2 PROBLEM DESCRIPTION

In this section, we present the acoustic wave equation in its simplest form and demonstrate the long-time instability issue associated with the non-uniform grid.

2.1 The PDE system

We consider the 2-D acoustic wave equation posed as the following first-order system:

$$\begin{cases} \frac{\partial p}{\partial t} = -\rho c^2 \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} \right) + S; \\ \frac{\partial v_x}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial x}; \\ \frac{\partial v_y}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial y}. \end{cases} \quad (1)$$

Here, v_x and v_y are particle velocities in the x and y directions, respectively, p is pressure, ρ is density, c is acoustic wave-speed, and S is the externally applied source term that drives the wave propagation. Discretizations of the solution variables p , v_x and v_y will be denoted by P , v_x and V_y , respectively. In the upcoming analysis, the homogeneous version of (1) is often considered where the source term S is omitted. The medium is assumed to be at rest at the beginning of the simulation. Translating into initial conditions, this means that v_x , v_y and p , as well as their derivatives, are zero. To facilitate the analysis, only periodic boundary conditions are considered in the main exposition below. However, the numerical techniques devised can be readily applied to simulations with free surface boundary condition and perfectly matched layers, as demonstrated in Section 4.2.

2.2 Numerical discretization on discontinuous grid

For spatial discretization, we consider the staggered-grid finite difference schemes, which are commonly used in the seismic exploration community (*cf.* Virieux 1984, 1986; Levander 1988). Moreover, we consider the discretization scenario where the grid-spacing is different for regions at different depth. Specifically, the entire grid is split into three horizontal layers with fine, coarse and fine grid-spacing respectively. The ratio between the coarse grid-spacing and the fine grid-spacing is set to be three so that grid points from different layers align with each other on the interface for the staggered-grid setting (*cf.* Kristek *et al.* 2010). The layout of the discretization grids is illustrated in Fig. 1. The green dashed lines in Fig. 1(a) denote the interfaces. For the numerical tests below, we place one interface on the P grid and the other on the V_y grid so that both cases can be tested. However, we only address the lower interface in the main texts since the upper interface can be handled similarly.

For the purpose of illustration, we use the second-order central difference scheme in space for each layer with the respective grid-spacing. Grid points on the interfaces are counted as belonging to the coarse layers and hence the coarse grid-spacing is used for their updates. Their stencils request fine grid solution values. Thanks to the three-to-one ratio between the grid-spacings, the requested values are located on fine grid points. We simply take these values to update the interfaces. From the grid-coupling perspective, this means that we use an injection operator to transfer the solution values from fine grid to coarse grid. Following the terminology of the multigrid literature, we refer to operators that transfer solution values from fine grid to coarse grid as restriction operators. The injection operator is a special case of restriction operators. On the other hand, in order to update the fine grid points right below the lower interface, one needs to supply solution values for the missing grid points that are requested by the stencils, as illustrated by the hollow circles in Fig. 1(b). We use the interpolation operator detailed in Appendix A to achieve this. From the grid-coupling perspective, interpolation operators transfer the solution values from coarse grid to fine grid. Adopting the terminology of the multigrid

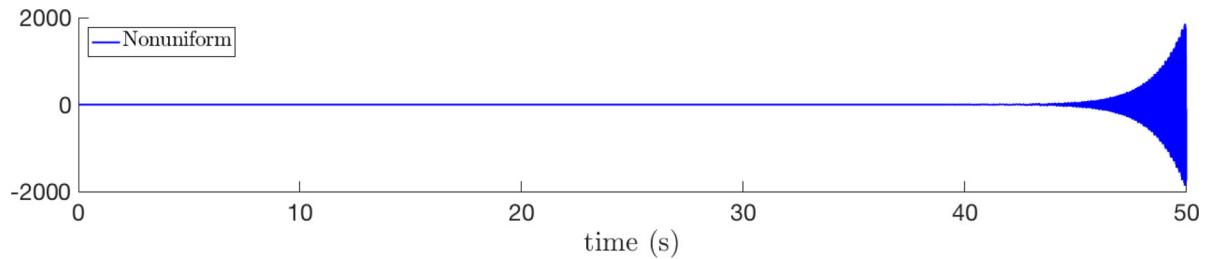


Figure 2. Seismogram simulated on non-uniform grid with second-order central difference scheme in space and the leapfrog scheme in time. The injection operator (*cf.* Section 2.2) and the interpolation operator (*cf.* Appendix A) are used to couple the grids.

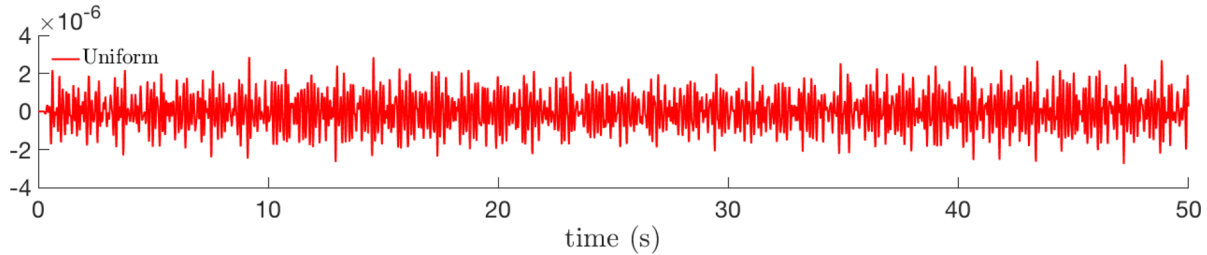


Figure 3. Seismogram simulated on uniform grid with second-order central difference scheme in space and the leapfrog scheme in time.

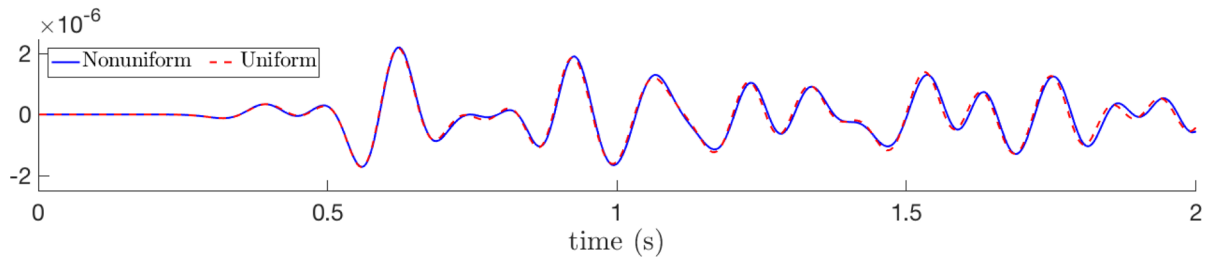


Figure 4. Comparison between the non-uniform and uniform simulation results for the first 2 s. We observe good agreements between these two seismograms.

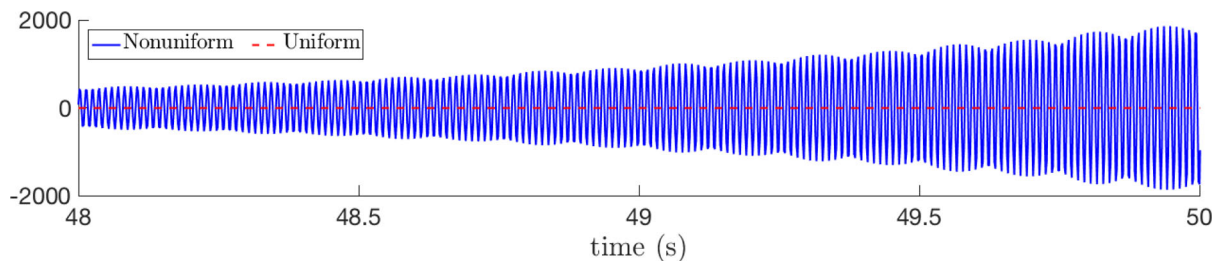


Figure 5. Comparison between non-uniform and uniform simulation results for the last 2 s. The non-uniform seismogram exhibits amplifying and highly oscillatory behaviour.

literature, we may refer to interpolation operators as prolongation operators in the upcoming discussions.

2.3 Long-time instability

In this section, we illustrate the long-time instability issue associated with the discretization presented in the previous section. For simplicity, we consider the acoustic wave eq. (1) with unit density and wave-speed, that is, $\rho = 1 \text{ kg m}^{-3}$ and $c = 1 \text{ m s}^{-1}$ everywhere. Periodic boundary conditions are imposed on all boundaries.

The three layers of the simulation domain are of sizes 120×30 (fine), 40×20 (coarse) and 120×30 (fine), from top to bottom, respectively. The wave-field is excited by a point source on pressure, *cf.* eq. (1), whose time signal is a Ricker wavelet with 5 Hz central frequency. We count the maximum frequency in the source content as 12.5 Hz. The fine and coarse grid-spacings are chosen in

ways such that there are 30 and 10 points per minimum wavelength (corresponding to 12.5 Hz), respectively. The source is turned on after 0.25 s of the simulation. The time step length is chosen as 0.99 times the CFL limit (approximately 0.0018668s). The source is placed at the 30th (horizontally) and 25th (vertically) grid point within the top layer of the P grid. The receiver is placed at the 90th (horizontally) and 25th (vertically) grid point within the top layer of the v_y grid. In terms of physical units, they are at (0.0773 m, 0.3013 m) and (0.2373 m, 0.3027 m), respectively, assuming that the left bottom grid point (on P grid) is anchored at (0 m, 0 m).

In Fig. 2, we display the seismogram recorded at the receiver for 50 s. As comparison, the seismogram simulated on uniform grid with 30 points per minimum wavelength everywhere is shown in Fig. 3 for the same period. We observe that at the late stage, the solution obtained from the non-uniform grid simulation seems to grow at some exponential rate. In Figs 4 and 5, we zoom in on the

first 2 s and the last 2 s of the two seismograms. From Fig. 4, we see that at the beginning of the simulation, the two seismograms agree fairly well. From Fig. 5, we observe that the unstable modes exhibit a highly oscillatory feature, which is not described by the underlying physics.

To verify the exponential growth in Fig. 2, we write the propagation scheme into the following algebraic form:

$$P^{(n_t)} - P^{(n_t-1)} = \mathbb{K}_x^V V_x^{(n_t-\frac{1}{2})} + \mathbb{K}_y^V V_y^{(n_t-\frac{1}{2})}; \quad (2a)$$

$$V_x^{(n_t+\frac{1}{2})} - V_x^{(n_t-\frac{1}{2})} = \mathbb{K}_x^P P^{(n_t)}; \quad (2b)$$

$$V_y^{(n_t+\frac{1}{2})} - V_y^{(n_t-\frac{1}{2})} = \mathbb{K}_y^P P^{(n_t)}, \quad (2c)$$

where $P^{(n_t)}$, $V_x^{(n_t+\frac{1}{2})}$ and $V_y^{(n_t+\frac{1}{2})}$ are the fully discretized solution vectors at time steps n_t , $n_t + \frac{1}{2}$ and $n_t + \frac{1}{2}$, respectively; \mathbb{K}_x^P , \mathbb{K}_y^P , \mathbb{K}_x^V and \mathbb{K}_y^V are matrices resulted from the spatial and temporal finite difference operations. Furthermore, we define vector $\Phi^{(n_t)}$ as

$$\Phi^{(n_t)} = \begin{bmatrix} P^{(n_t)} \\ V_x^{(n_t+\frac{1}{2})} \\ V_y^{(n_t+\frac{1}{2})} \end{bmatrix}. \quad (3)$$

With this definition, scheme (2) can be written as a one-step scheme as:

$$\Phi^{(n_t+1)} = \mathbb{K}^\Phi \Phi^{(n_t)}, \quad (4)$$

where

$$\mathbb{K}^\Phi = \left(\mathcal{I} + \begin{bmatrix} \mathbf{0} & \mathbb{K}_x^V & \mathbb{K}_y^V \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \right) \left(\mathcal{I} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbb{K}_x^P & \mathbf{0} & \mathbf{0} \\ \mathbb{K}_y^P & \mathbf{0} & \mathbf{0} \end{bmatrix} \right). \quad (5)$$

In eq. (5), \mathcal{I} stands for the identity matrix of the appropriate size; terms in the first parentheses correspond to the half time step that updates P , cf. (2a), while terms in the second parentheses correspond to the half time step that updates v_x and v_y , cf. (2b) and (2c).

We consider the eigenvalue of \mathbb{K}^Φ with the largest magnitude, denoted as λ^Φ , and its corresponding eigenvector, denoted as v^Φ . Denoting the solution component contributed from v^Φ at time step n_t as $c^{(n_t)} v^\Phi$, we have

$$|c^{(n_t)}| = |\lambda^\Phi| |c^{(n_t-1)}| = \dots = |\lambda^\Phi|^{n_t} |c^{(0)}|. \quad (6)$$

If $|\lambda^\Phi| > 1$, the solution component contributed from v^Φ has the leading amplitude growth rate among all eigenvectors, and hence tends to dictate the growth rate of the solution amplitude at the late stage of the simulation. Taking the natural logarithm of (6), we have

$$\ln(|c^{(n_t)}|) = \ln(|c^{(0)}|) + \ln(|\lambda^\Phi|) n_t. \quad (7)$$

Therefore, we expect the natural logarithm of the solution amplitude to grow linearly with rate $\ln(|\lambda^\Phi|)$ at the late stage of the simulation. This is confirmed in Fig. 6, where we take the values from Fig. 2 and plot the natural logarithm of their absolute values, referenced by a straight line of slope $\ln(|\lambda^\Phi|)$.

3 ANALYSIS WITH THE DISCRETE ENERGY METHOD

In this section, we analyse the aforementioned instability issue from an algebraic perspective. We consider the homogeneous version of

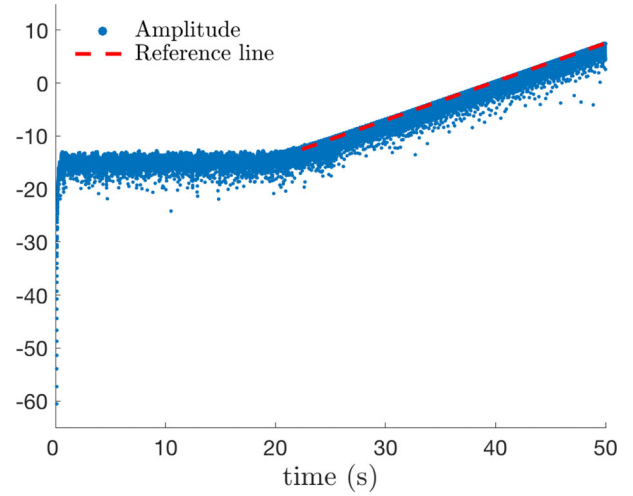


Figure 6. The growth rate of the solution amplitude matches the prediction from the eigenvalue with the largest magnitude. The solution amplitude is plotted in logarithmic scale for y-axis.

PDE system (1) and its 1-D counterpart. For clarity, we list them in (8) and (9), respectively.

$$\begin{cases} \frac{1}{\rho c^2} \frac{\partial p}{\partial t} = - \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} \right); \\ \rho \frac{\partial v_x}{\partial t} = - \frac{\partial p}{\partial x}; \\ \rho \frac{\partial v_y}{\partial t} = - \frac{\partial p}{\partial y}, \end{cases} \quad (8)$$

$$\begin{cases} \frac{1}{\rho c^2} \frac{\partial p}{\partial t} = - \frac{\partial v}{\partial x}; \\ \rho \frac{\partial v}{\partial t} = - \frac{\partial p}{\partial x}. \end{cases} \quad (9)$$

We note here that the physical parameters in (8) and (9) are moved to the left-hand side. As shown later, they enter the definition of the discrete energy naturally in this form, cf. Section 3.1. This change of form is not limited to acoustic wave equations, but also applies to elastic wave equations with any constitutive relations, cf. Appendix B. With the definition of energy given in Appendix B, the techniques developed in this work can be naturally extended to the elastic case.

3.1 The 1-D wave equation on a uniform grid

For simplicity, we start our discussion with the 1-D system (9), accompanied by the periodic boundary condition. As in the 2-D case, the V grid is shifted to the right of the P grid for half grid-spacing. Spatial discretization of (9) leads to the following semi-discrete ODE system:

$$\begin{cases} \mathcal{A}^P \mathcal{C}^P \frac{\partial P}{\partial t} = -\mathcal{A}^P \mathcal{G}^P \mathcal{D}^V V; \\ \mathcal{A}^V \mathcal{C}^V \frac{\partial V}{\partial t} = -\mathcal{A}^V \mathcal{G}^V \mathcal{D}^P P. \end{cases} \quad (10)$$

On the left-hand side of (10), \mathcal{C}^P and \mathcal{C}^V are the coefficient matrices obtained by evaluating the physical parameters at the grid point locations; \mathcal{A}^P and \mathcal{A}^V are the area matrices representing the area that each grid point occupies. The seemingly redundant area matrices will enter the definition of the discrete energy. All these matrices, that is, \mathcal{C}^P , \mathcal{C}^V , \mathcal{A}^P , \mathcal{A}^V , are diagonal. On the right-hand side of (10), P and V stand for the semi-discrete time-dependent

solution vectors; \mathcal{D}^P and \mathcal{D}^V are the differentiation matrices storing the stencil coefficients corresponding to unit grid-spacing; \mathcal{G}^P and \mathcal{G}^V are diagonal matrices whose diagonal components store the reciprocal of the grid-spacing for the corresponding grid points. We refer to them as grid matrices in the following. The grid matrices are introduced to remove the scaling from the differentiation matrices.

To simplify notation, we define the mass matrices \mathcal{M}^P and \mathcal{M}^V as

$$\mathcal{M}^P = \mathcal{C}^P \mathcal{A}^P \quad \text{and} \quad \mathcal{M}^V = \mathcal{C}^V \mathcal{A}^V, \quad (11)$$

which are also diagonal matrices. We can now rewrite system (10) succinctly as follows:

$$\mathcal{M}^\Phi \frac{\partial \Phi}{\partial t} = \mathcal{D}^\Phi \Phi, \quad (12)$$

where

$$\Phi = \begin{bmatrix} P \\ V \end{bmatrix}, \quad \mathcal{M}^\Phi = \begin{bmatrix} \mathcal{M}^P & \mathbf{0} \\ \mathbf{0} & \mathcal{M}^V \end{bmatrix} \quad \text{and} \quad \mathcal{D}^\Phi = \begin{bmatrix} \mathbf{0} & -\mathcal{A}^P \mathcal{G}^P \mathcal{D}^V \\ -\mathcal{A}^V \mathcal{G}^V \mathcal{D}^P & \mathbf{0} \end{bmatrix}. \quad (13)$$

The symbol ' $\mathbf{0}$ ' in \mathcal{M}^Φ and \mathcal{D}^Φ of (13) stands for a zero matrix block of the appropriate size. We define the discrete kinetic energy and the discrete potential energy associated with dynamical system (10) as

$$E_k = \frac{1}{2} V^T \mathcal{M}^V V \quad \text{and} \quad E_p = \frac{1}{2} P^T \mathcal{M}^P P, \quad (14)$$

respectively. Definitions in (14) imitate the physical definitions of the instantaneous kinetic and potential energy for the acoustic wave at the continuous level. The total energy associated with (10) or (12) is defined as

$$E_\Phi = E_k + E_p = \frac{1}{2} \Phi^T \mathcal{M}^\Phi \Phi. \quad (15)$$

Differentiating (15) with respect to time t and substituting with (12), we have

$$\frac{\partial E_\Phi}{\partial t} = \Phi^T \mathcal{M}^\Phi \frac{\partial \Phi}{\partial t} = \Phi^T \mathcal{D}^\Phi \Phi = \frac{1}{2} \Phi^T (\mathcal{D}^\Phi + (\mathcal{D}^\Phi)^T) \Phi. \quad (16)$$

The last relation of (16) holds since $\Phi^T \mathcal{D}^\Phi \Phi$ is a real scalar, and hence $\Phi^T \mathcal{D}^\Phi \Phi = (\Phi^T \mathcal{D}^\Phi \Phi)^T = \Phi^T (\mathcal{D}^\Phi)^T \Phi$. We use \mathcal{H}^Φ to denote the matrix sum $(\mathcal{D}^\Phi + (\mathcal{D}^\Phi)^T)$. From (16), we note that if \mathcal{H}^Φ is negative semi-definite, the total energy E_Φ is non-increasing and as a consequence, the instability as shown in Fig. 2 will not be allowed. In particular, when $\mathcal{H}^\Phi = \mathbf{0}$, that is, when \mathcal{D}^Φ is skew-symmetric, the total energy E_Φ as defined in (15) is conserved. Translating to symbols in (10), \mathcal{D}^Φ being skew-symmetric is equivalent to

$$\mathcal{A}^P \mathcal{G}^P \mathcal{D}^V + (\mathcal{A}^V \mathcal{G}^V \mathcal{D}^P)^T = \mathbf{0}. \quad (17)$$

For the 1-D wave equation on a uniform grid, matrices \mathcal{A}^P and \mathcal{G}^P are simply inverses of each other, so are matrices \mathcal{A}^V and \mathcal{G}^V . Relation (17) boils down to $(\mathcal{D}^V)^T + \mathcal{D}^P = \mathbf{0}$, which is valid since \mathcal{D}^V and \mathcal{D}^P take the following forms:

$$\mathcal{D}^V = \begin{bmatrix} 1 & & & -1 \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}; \quad \mathcal{D}^P = \begin{bmatrix} -1 & 1 & & \\ & -1 & \ddots & \\ & & \ddots & 1 \\ 1 & & & -1 \end{bmatrix} \quad (18)$$

Table 1. Extension of the nomenclature.

	1-D		2-D
Solution variable	P, V	\rightarrow	P, v_x, v_y
Coefficient matrix	$\mathcal{C}^P, \mathcal{C}^V$	\rightarrow	$\mathcal{C}^P, \mathcal{C}_x^V, \mathcal{C}_y^V$
Area matrix	$\mathcal{A}^P, \mathcal{A}^V$	\rightarrow	$\mathcal{A}^P, \mathcal{A}_x^V, \mathcal{A}_y^V$
Mass matrix	$\mathcal{M}^P, \mathcal{M}^V$	\rightarrow	$\mathcal{M}^P, \mathcal{M}_x^V, \mathcal{M}_y^V$
Grid matrix	$\mathcal{G}^P, \mathcal{G}^V$	\rightarrow	$\mathcal{G}_x^P, \mathcal{G}_y^P, \mathcal{G}_x^V, \mathcal{G}_y^V$
Differentiation matrix	$\mathcal{D}^P, \mathcal{D}^V$	\rightarrow	$\mathcal{D}_x^P, \mathcal{D}_y^P, \mathcal{D}_x^V, \mathcal{D}_y^V$

for second-order central difference schemes on staggered grids where the V grid is shifted to the right of the P grid.

3.2 The 2-D wave equation on non-uniform grid

For the 2-D wave eq. (8) defined on the non-uniform grid, cf. Fig. 1, two issues arise when applying the aforementioned discrete energy argument. First, the non-uniformity along the vertical direction leads to unmatched terms in \mathcal{H}^Φ . Second, the choice of the solution transfer operators may further damage the skew-symmetry of \mathcal{D}^Φ . Next, we describe these two issues in detail.

To understand the issue associated with non-uniformity, we consider a single complete grid column in Fig. 1(a), that is, a grid column that contains both fine and coarse grid points, and the 1-D wave equation defined on it. Spatial discretization still leads to ODE system (10), where matrices \mathcal{A}^P and \mathcal{G}^P are still inverses of each other, so are \mathcal{A}^V and \mathcal{G}^V . However, matrices \mathcal{D}^V and \mathcal{D}^P now take the following forms near the lower interface:

$$\mathcal{D}^V = \begin{bmatrix} \ddots & & & & \\ & \ddots & & & \\ & & 1 & & \\ & -1 & 1 & & \\ & -1 & 0 & 1 & \\ & & & -1 & 1 \\ & & & & \ddots & \ddots \end{bmatrix}; \quad \mathcal{D}^P = \begin{bmatrix} \ddots & & & & \\ & \ddots & & & \\ & & -1 & 1 & \\ & & -1 & 1 & \\ & & & -1 & 1 \\ & & & & -1 & \ddots \\ & & & & & \ddots \end{bmatrix}. \quad (19)$$

We note from (19) that due to sudden change of the grid-spacing, the row in \mathcal{D}^V that corresponds to the interface, that is, $[\dots -101\dots]$, is different from the one in (18), that is, $[\dots -11\dots]$. This small change leads to unmatched terms in \mathcal{H}^Φ and hence breaks the skew-symmetry in \mathcal{D}^Φ .

Now we turn to the second issue associated with the choice of solution transfer operators. First, we extend the nomenclature to 2-D as summarized in Table 1. The grid point coordinates are mapped to vector indices using the row-major order, starting from the bottom of the grid. With the exception of \mathcal{D}_y^V and \mathcal{D}_y^P , these 2-D matrices are obvious extensions of their 1-D counterparts.

Aside from the grid rows whose update requires solution transfer between grids, matrices \mathcal{D}_y^V and \mathcal{D}_y^P can be viewed as the block extensions of their 1-D correspondences by replacing the entries '1' therein with the identity matrix blocks of the appropriate sizes, denoted by \mathcal{I}^F and \mathcal{I}^C in (20). If the number of horizontal P grid points on the fine and coarse grid layers are N_x^F and N_x^C , respectively, \mathcal{I}^F and \mathcal{I}^C are of sizes $N_x^F \times N_x^F$ and $N_x^C \times N_x^C$, respectively. For grid rows whose update requires coarse-to-fine solution transfer, the entries '1' are replaced with the prolongation operator, denoted by \mathcal{P} in (20), which is of size $N_x^F \times N_x^C$. For grid rows whose update requires fine-to-coarse solution transfer, the entries '1' are replaced with the restriction operator, denoted by \mathcal{R} in (20), which is of size $N_x^C \times N_x^F$. The entries '0' in \mathcal{D}_y^P and \mathcal{D}_y^V are replaced with the zero matrix blocks of the appropriate size, denoted by '0' in (20). Eventually, \mathcal{D}_y^V and \mathcal{D}_y^P take the following block matrix form near the lower interface:

$$\mathcal{D}_y^V = \begin{bmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & \mathcal{I}^F & & & \\ & & -\mathcal{I}^F & \mathcal{I}^F & & \\ & & -\mathcal{R} & \mathbf{0} & \mathcal{I}^C & \\ & & & & -\mathcal{I}^C & \mathcal{I}^C \\ & & & & & \ddots & \ddots \end{bmatrix};$$

$$\mathcal{D}_y^P = \begin{bmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & -\mathcal{I}^F & \mathcal{I}^F & & \\ & & -\mathcal{I}^F & \mathcal{P} & & \\ & & & -\mathcal{I}^C & \mathcal{I}^C & \\ & & & & -\mathcal{I}^C & \ddots \\ & & & & & \ddots \end{bmatrix}. \quad (20)$$

We now turn to the specifics of the solution transfer operators, that is, the prolongation operator \mathcal{P} and the restriction operator \mathcal{R} . A natural choice of the prolongation operator is an interpolation operator, such as the one used in Section 2.2 and detailed in Appendix A. For the restriction operator, a simple choice is the injection operator mentioned in Section 2.2, which we denote by \mathcal{R}^0 hereafter. As shown in the next section, other choice of the restriction operator exists and can play an important role in improving the stability. With these newly defined matrices, the 2-D semi-discrete dynamical system can be written as

$$\mathcal{M}^\Phi \frac{\partial \Phi}{\partial t} = \mathcal{D}^\Phi \Phi, \quad (21)$$

where

$$\Phi = \begin{bmatrix} P \\ V_x \\ V_y \end{bmatrix}; \quad \mathcal{M}^\Phi = \begin{bmatrix} \mathcal{M}^P & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{M}_x^V & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathcal{M}_y^V \end{bmatrix};$$

$$\mathcal{D}^\Phi = \begin{bmatrix} \mathbf{0} & -\mathcal{A}^P \mathcal{G}_x^P \mathcal{D}_x^V & -\mathcal{A}^P \mathcal{G}_y^P \mathcal{D}_y^V \\ -\mathcal{A}_x^V \mathcal{G}_x^V \mathcal{D}_x^P & \mathbf{0} & \mathbf{0} \\ -\mathcal{A}_y^V \mathcal{G}_y^V \mathcal{D}_y^P & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (22)$$

Skew-symmetry of \mathcal{D}^Φ now requires

$$\mathcal{A}^P \mathcal{G}_x^P \mathcal{D}_x^V + (\mathcal{A}_x^V \mathcal{G}_x^V \mathcal{D}_x^P)^T = \mathbf{0}; \quad (23a)$$

$$\mathcal{A}^P \mathcal{G}_y^P \mathcal{D}_y^V + (\mathcal{A}_y^V \mathcal{G}_y^V \mathcal{D}_y^P)^T = \mathbf{0}. \quad (23b)$$

Since for each grid row in Fig. 1(a), the grid-spacing is uniform along the x -direction, it can be verified easily that relation (23a) holds. As for relation (23b), most of the terms still drop out in the left-hand side sum except for the rows associated with the interfaces, which takes the form:

$$[\cdots -(\Delta C)\mathcal{R}^0 \quad (\Delta F)\mathcal{P} \quad \mathbf{0} \quad \cdots] \quad (24)$$

for the lower interface, in which $\mathbf{0}$ occupies the diagonal block. The skew-symmetry of \mathcal{D}^Φ is further damaged due to matrix \mathcal{P} , which has more non-zero entries than \mathcal{R}^0 . Similar to the 1-D case, we have the scalar equation that governs the dynamics of the energy:

$$\frac{\partial E_\Phi}{\partial t} = \frac{1}{2} \Phi^T \mathcal{H}^\Phi \Phi, \quad (25)$$

where $\mathcal{H}^\Phi = \mathcal{D}^\Phi + (\mathcal{D}^\Phi)^T$.

We recall that the trace of a square matrix A , denoted by $tr(A)$ hereafter, is defined as the sum of its diagonal entries. It can be shown that the trace of A equals to the sum its eigenvalues; see Horn & Johnson (2012) for more detail. Since the diagonal of \mathcal{D}^Φ is occupied by zero blocks, the trace of \mathcal{H}^Φ is zero. Moreover, since \mathcal{H}^Φ is symmetric, it only has real eigenvalues and at least one of them must be positive due to its trace being zero. Consequently, the discrete energy is not necessarily conserved during the simulation and the kind of instability as shown in Fig. 2 is permitted.

3.3 Pairing the solution transfer operators

In our investigation, we find that by pairing the solution transfer operators, one can bring matrix \mathcal{D}^Φ closer to skew-symmetry. This often significantly improves the stability of the simulation. Specifically, we adopt the following procedure: wherever \mathcal{P} appears in \mathcal{D}_y^V (or \mathcal{D}_y^P), we modify the corresponding block of \mathcal{D}_y^P (or \mathcal{D}_y^V) with a combination of \mathcal{P}^T and \mathcal{R}^0 such that the prolongation operator \mathcal{P} is cancelled out in \mathcal{H}^Φ . Eventually, only the much sparser \mathcal{R}^0 and its transpose appear in \mathcal{H}^Φ . For instance, for matrix \mathcal{D}_y^P as shown in (20), we modify \mathcal{D}_y^V such that

$$\mathcal{D}_y^V = \begin{bmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & \mathcal{I}^F & & & \\ & & -\mathcal{I}^F & \mathcal{I}^F & & \\ & & -\mathcal{R}^0 & \mathbf{c}_R \mathcal{R}^0 - \mathbf{c}_P \mathcal{P}^T & \mathcal{I}^C & \\ & & & & -\mathcal{I}^C & \mathcal{I}^C \\ & & & & & \ddots & \ddots \end{bmatrix}, \quad (26)$$

where $\mathbf{c}_R = 1$ and $\mathbf{c}_P = \Delta F / \Delta C$ in this particular case. Coefficient \mathbf{c}_P is introduced to balance out the differences in area matrices and grid matrices so that \mathcal{P} is cancelled out in \mathcal{H}^Φ . Coefficient \mathbf{c}_R is introduced to maintain consistency in the derivative approximation. As discussed in Appendix A, if \mathcal{P} is chosen to be the interpolation operator described therein, $\mathbf{c}_P \mathcal{P}^T$ turns out to be a weighted-averaging operator and therefore, $(\mathbf{c}_R \mathcal{R}^0 - \mathbf{c}_P \mathcal{P}^T)$ is a valid restriction operator. Eventually, the non-zero leftover of the left-hand side sum in (23b) becomes

$$[\cdots -(\Delta C)\mathcal{R}^0 \quad (\Delta C)\mathcal{R}^0 \quad \mathbf{0} \quad \cdots], \quad (27)$$

which has far fewer non-zero entries as compared to (24), as illustrated in Fig. 7.

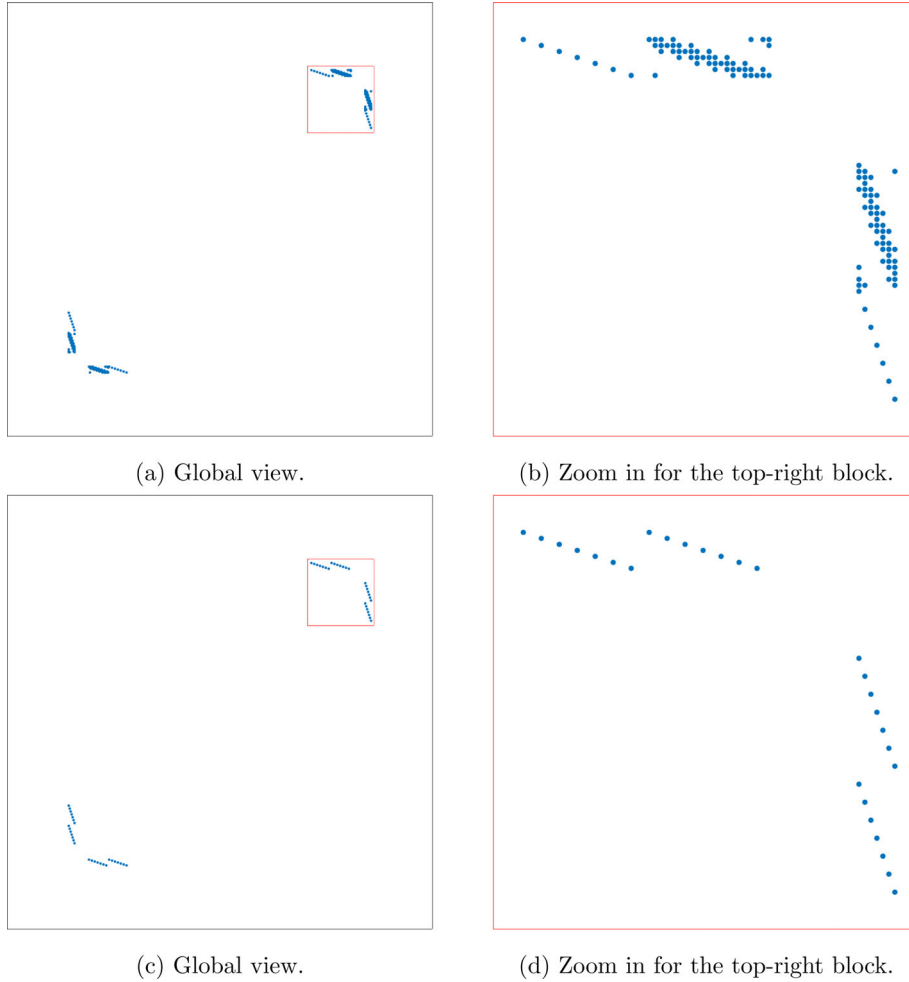


Figure 7. Non-zero patterns of \mathcal{H}^Φ . (a,b) The interpolation operator is not paired with the companion restriction operator; (c,d) the interpolation operator is paired with the companion restriction operator.

Hereafter, we refer to the particular choice $(\mathbf{c}_R \mathcal{R}^0 - \mathbf{c}_P \mathcal{P}^T)$ of the restriction operator as the *companion restriction operator* of \mathcal{P} , or simply the *companion restriction operator* if there is no ambiguity in the context. For the prototype problem considered in Section 2.3, by simply adding the companion restriction operator, the largest magnitude of the eigenvalues of \mathbb{K}^Φ is reduced, approximately, from 1.0013565 to 1.0006021, and the corresponding asymptotic growth rate, cf. Fig. 6, is reduced for a factor of around 2.25. This implies that it takes more than twice as long for the unstable modes to reach the same amplitude, provided the same initial amplitude $c^{(0)}$, cf. (6).

3.4 Horizontal artificial diffusion

The common technique of artificial diffusion can be used to further stabilize the simulation. Analysis in previous sections provides guidance on how to effectively introduce the artificial diffusion. Recalling the definition of \mathcal{D}^Φ given in (22), matrix $\mathcal{H}^\Phi = \mathcal{D}^\Phi + (\mathcal{D}^\Phi)^T$ now has the following block structure:

$$\mathcal{H}^\Phi = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \Sigma \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \Sigma^T & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (28)$$

since relation (23a) holds. After the introduction of the companion restriction operator, the non-zero entries left in $\Sigma = -(\mathcal{A}^P \mathcal{G}_y^P \mathcal{D}_y^V + (\mathcal{A}_y^V \mathcal{G}_y^V \mathcal{D}_y^P)^T)$ come solely from the injection operator \mathcal{R}^0 , cf. eq. (27), which, according to how it is constructed, takes the following form:

$$\mathcal{R}^0 = \begin{bmatrix} \ddots & & & & & & \\ & 1 & 0 & 0 & & & \\ & & & & 1 & 0 & 0 \\ & & & & & & 1 & 0 & 0 \\ & & & & & & & \ddots & \end{bmatrix}. \quad (29)$$

Since \mathcal{R}^0 only appears for grid-rows near the interface, matrix block Σ is mostly zero, cf. Figs 7(c) and (d). Due to this sparse structure, each zero row of \mathcal{H}^Φ (or column due to its symmetry) corresponds to a zero eigenvalue. Zero eigenvalues indicate that the corresponding solution modes are not amplifying nor decreasing, which is how the physical modes should behave since the underlying PDE system is energy-conservative (not considering boundary conditions). Therefore, when applying artificial diffusion, we want to keep these zero eigenvalues intact. Our aim of introducing artificial diffusion

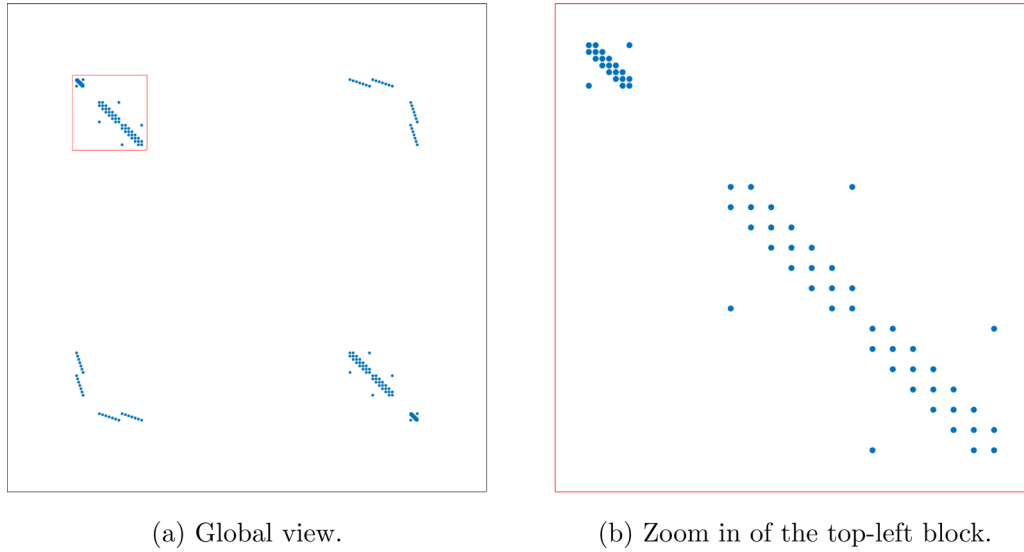


Figure 8. Non-zero pattern of \mathcal{H}^Φ . The diagonal blocks corresponding to the horizontal diffusion terms. In the right figure, the more compact matrix block on the top-left corner corresponds to \mathcal{K}^C ; the remaining two less compact matrix blocks correspond to \mathcal{K}^F .

accuracy will be preserved for the simulation. This scale $(\Delta C)^2$ is indeed what we use for simulations presented in Section 4. The particular values of κ_p and κ_y can be determined via a trial-and-error process. From our experience, the effect of the horizontal diffusion terms is robust with respect to these particular values in terms of providing satisfactory stabilization. For realistic physical models, one need to adjust κ_p and κ_y according to the scales of the physical parameters. As explained in Appendix C, one can show that these coefficients need to be scaled by the respective physical parameters in eq. (31), that is, $\frac{1}{\rho c^2}$ for κ_p and ρ for κ_y .

With the horizontal diffusion terms chosen as such, we observe little impact on the CFL constraint. For instance, for the prototype problem discussed in Section 2.3, when the same numerical configuration presented therein is used in combination with the companion restriction operator and the horizontal diffusion terms, one can set the time step length as 0.99 times the CFL limit without encountering trouble. This will be demonstrated in Section 4.1 with a nearly identical test configuration.

4 NUMERICAL EXAMPLES

In this section, we demonstrate the effectiveness and applicability of the aforementioned techniques in seismic wave modelling applications.

4.1 Homogeneous media

For this example we consider homogeneous media with unit physical parameters, the same as presented in Section 2.3. The geometric and numerical settings are almost identical to those used in Section 2.3, except that we use the fourth-order central difference scheme in the coarse (middle) layer to improve the spatial accuracy. We recapitulate here that the fine and coarse grid-spacings are determined based on 30 and 10 grid points per minimum wavelength, respectively, and the time step length is chosen as 0.99 times the CFL limit.

The techniques developed in Sections 3.3 and 3.4 can be naturally applied to the schemes used here (and other schemes in general) by adhering to the following two procedures. First, whenever the coarse grid stencils request fine grid solution so that the prolongation

operator \mathcal{P} appears in \mathcal{D}_y^V (or \mathcal{D}_y^P), we modify the corresponding block of \mathcal{D}_y^P (or \mathcal{D}_y^V) with the companion restriction operator so that \mathcal{P} is cancelled out in \mathcal{H}^Φ . Second, for the remaining non-zero rows (or columns) of \mathcal{H}^Φ , we append the horizontal diffusion terms to the corresponding grid points.

With these stabilization techniques, the simulation can last for 50 s without blowing up, as demonstrated in Figs 9 and 10, where the seismograms of the first 12 s and the last 6 s are displayed, respectively. The uniform grid simulation results with the second-order central difference scheme and 30 grid points per minimum wavelength everywhere are also presented in Figs 9 and 10 for comparison.

4.2 The Marmousi model

In this example, we demonstrate that the proposed techniques can be readily applied to practical application settings, which may involve general heterogeneous media, PML layers and free surface boundary condition.

We consider the Marmousi model (Versteeg 1994) as demonstrated in Fig. 11. The uniform grid used for reference consists of 2298 grid points horizontally and 750 grid points vertically (including the PML layers but excluding the free surface) with 4 m grid-spacing, matching the sampling distance in the Marmousi model. The non-uniform grid consists of two horizontal layers, separated by the interface placed on the v_y grid and illustrated with the green dashed line in Fig. 11. The top layer is the fine grid layer, consists of 2298 by 40 grid points with 4 m grid-spacing. The bottom layer is the coarse grid layer, consists of 766 by 237 grid points with 12 m grid-spacing.

We use the Ricker source wavelet with central frequency 5 Hz and count the maximal frequency as 12.5 Hz. The aforementioned grid-spacings correspond to 30 points per minimum wavelength for the fine grid layer and approximately 10.4 points per minimum wavelength for the coarse grid layer. For the uniform grid case, the grid-spacing corresponds to 30 points per minimum wavelength. The time step length used in the non-uniform grid simulation is 0.00125 s, corresponding to, approximately, 0.95 times its CFL limit. The time step length used in the uniform grid simulation is

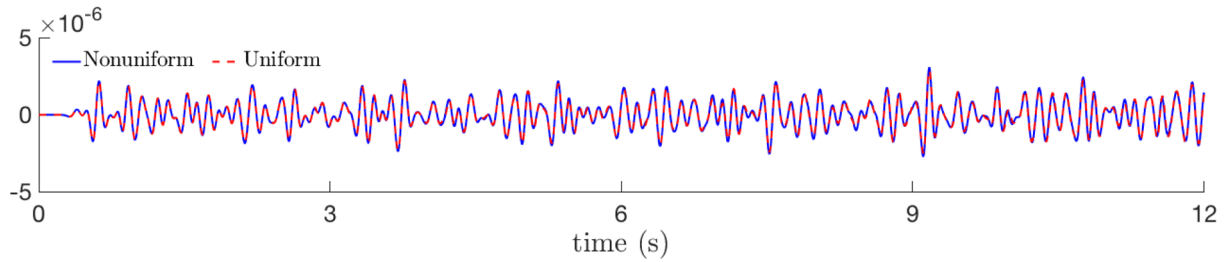


Figure 9. Seismograms of the first 12 s. We observe good agreement between the non-uniform grid and the uniform grid simulation results.

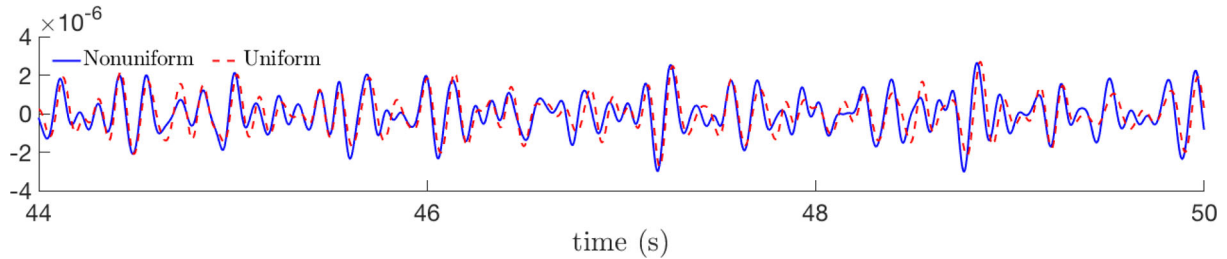


Figure 10. Seismograms of the last 6 s. The non-uniform grid simulation result remains stable and smooth, in contrast to the amplifying and highly oscillatory behaviour as shown in Fig. 5.

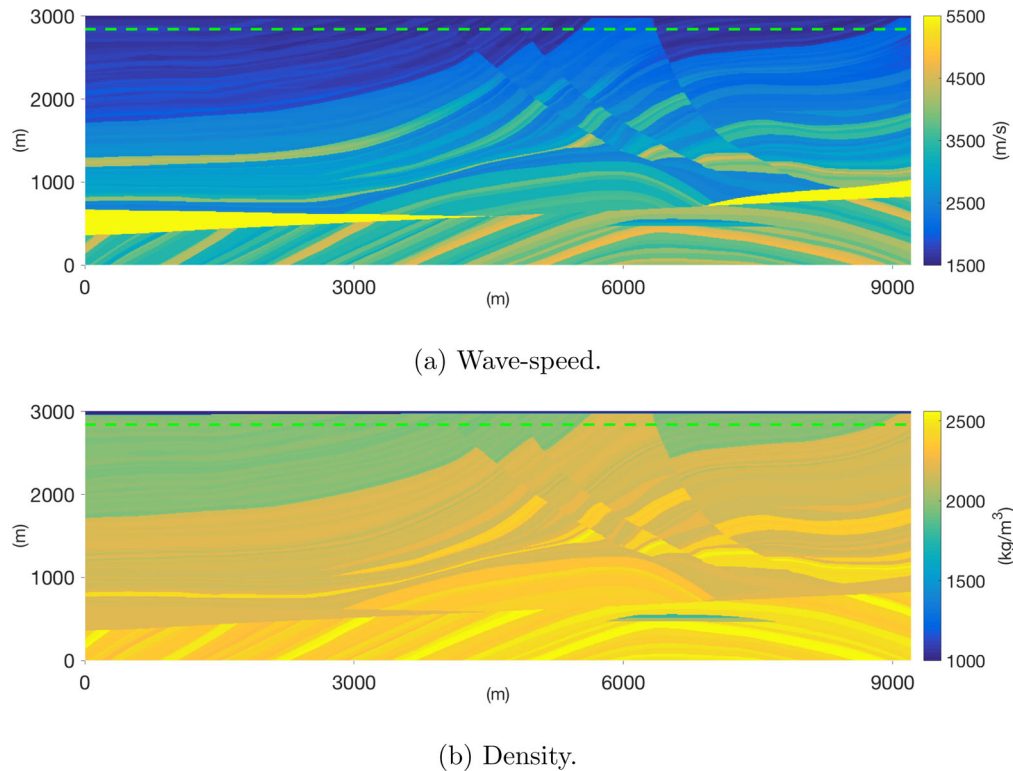


Figure 11. The Marmousi model used in Section 4.2.

one-third of it, corresponding to, approximately, 0.81 times its CFL limit. We remark here that the CFL limit for the non-uniform grid simulation is larger than that of the uniform grid simulation due to the split of the simulation domain and a larger grid-spacing used in the bottom layer.

The free surface boundary condition is imposed on the top boundary of Fig. 11. For the other three boundaries, PML layers of 30 grid points thickness and 10 grid points thickness are appended on the fine grid layer and the coarse grid layer, respectively. For the uni-

form grid case, the thickness of the PML layers is 30 grid points for all three boundaries. The source is placed on the P grid at the 130th grid point horizontally and the first grid point below the free surface, that is, 4 m below the free surface. The receiver is placed on the v_y grid at the 2168th-grid horizontally and the first grid point below the free-surface, that is, 2 m below the free surface. The source is turned on after 0.25 s of the simulation. Central-difference schemes of second order and fourth order are used for the fine grid layer and coarse grid layer of the non-uniform grid simulation, respectively.

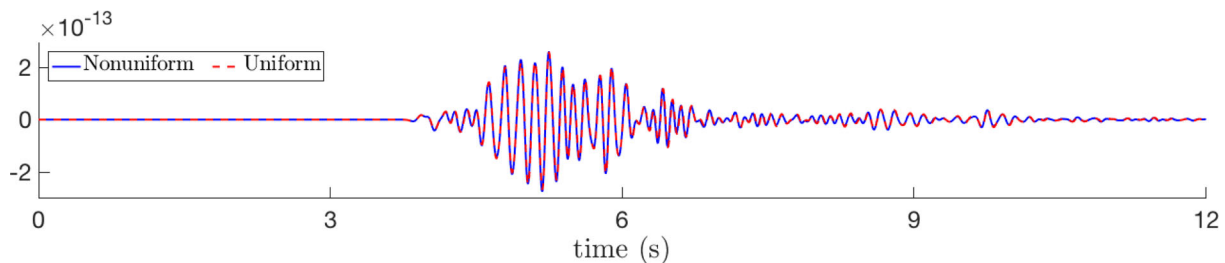


Figure 12. Marmousi—Comparison between the non-uniform and uniform simulation results for time period [0 s, 12 s].

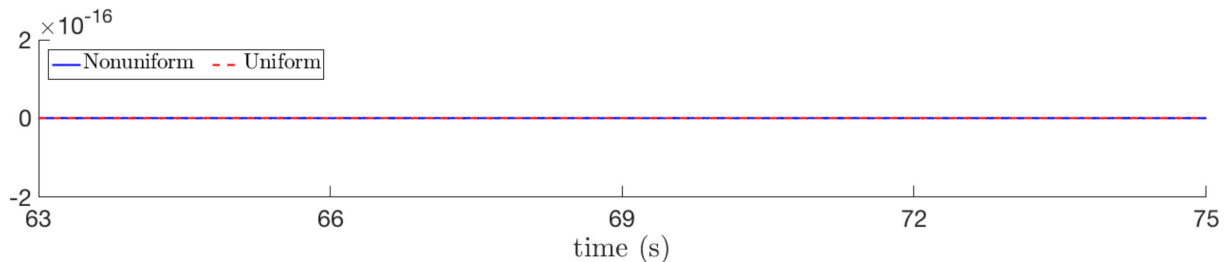


Figure 13. Marmousi—Seismograms of time period [63 s, 75 s]. We draw attention to the scale of the amplitude.

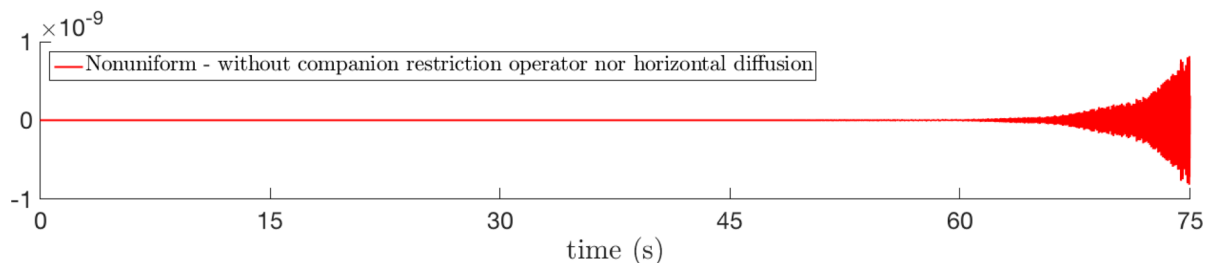


Figure 14. Marmousi—seismogram simulated without the companion restriction operator nor the horizontal diffusion terms. The simulation blows up eventually, even with the PML absorbing layers, which have the effect of attenuating the solution.

Central-difference scheme of second order is used for the uniform grid simulation. Both simulations use the leapfrog time integration scheme.

In Fig. 12, we display the seismogram simulated on non-uniform grid with the companion restriction operator and the horizontal diffusion terms for the first 12 s, compared against the seismogram simulated on the uniform grid. We observe in Fig. 12 that the two sets of seismograms match very well. In Fig. 13, the last 12 s of the seismograms are shown, from where we observe that both seismograms stay steady on the zero axis, as they are supposed to. Finally in Fig. 14, we show the unstable seismogram simulated without the stabilization from the companion restriction operator nor the horizontal diffusion terms.

5 CONCLUSION

We studied the long-time instability issue associated with finite difference simulation of seismic wave equation on discontinuous grid. Better understanding of its cause is achieved by analysing the abstracted algebraic system and its associated discrete energy. The non-zero pattern of the matrix governing the dynamics of the discrete energy is key to the development of the following two remedy techniques. First, the companion restriction operator is introduced to cancel out the non-zero entries that are introduced by the interpolation operator to this matrix. Second, the horizontal diffusion terms are introduced to perturb its eigenvalues towards the neg-

ative side of the real axis, providing further stabilization. These techniques are then applied to the seismic wave simulations in practical settings, by which their effectiveness are well demonstrated. In particular, the numerical example with the Marmousi model shows that these techniques can be applied to general heterogeneous media in combination with PML layers and free surface boundary condition.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the King Abdullah University of Science and Technology through the project CCF/CAF/URF/1/2596.

REFERENCES

- Baysal, E., Kosloff, D.D. & Sherwood, J.W., 1983. Reverse time migration, *Geophysics*, **48**(11), 1514–1524.
- Carpenter, M.H., Nordström, J. & Gottlieb, D., 1999. A stable and conservative interface treatment of arbitrary spatial accuracy, *J. Comput. Phys.*, **148**(2), 341–365.
- Davis, P., 2012. *Circulant Matrices: Second Edition*, AMS Chelsea Publishing, American Mathematical Society.
- Gao, L., Brossier, R. & Virieux, J., 2016. Using time filtering to control the long-time instability in seismic wave simulation, *Geophys. J. Int.*, **204**(3), 1443–1461.

- Gray, R., 2006. Toeplitz and circulant matrices: a review, *Found. Trends® Commun. Inf. Theory*, **2**(3), 155–239.
- Hayashi, K., Burns, D.R. & Toksöz, M.N., 2001. Discontinuous-grid finite-difference seismic modeling including surface topography, *Bull. seism. Soc. Am.*, **91**(6), 1750–1764.
- Horn, R. & Johnson, C., 2012. *Matrix Analysis*, Cambridge Univ. Press.
- Käser, M. & Dumbser, M., 2006. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes – I. The two-dimensional isotropic case with external source terms, *Geophys. J. Int.*, **166**(2), 855–877.
- Komatitsch, D. & Vilotte, J.-P., 1998. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.*, **88**(2), 368–392.
- Kreiss, H.-O. & Scherer, G., 1974. Finite element and finite difference methods for hyperbolic partial differential equations, in *Mathematical Aspects of Finite Elements in Partial Differential Equations*, Academic Press.
- Kristek, J., Moczo, P. & Galis, M., 2010. Stable discontinuous staggered grid in the finite-difference modelling of seismic motion, *Geophys. J. Int.*, **183**(3), 1401–1407.
- Levander, A.R., 1988. Fourth-order finite-difference p-sv seismograms, *Geophysics*, **53**(11), 1425–1436.
- LeVeque, R., 2007. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*, Society for Industrial and Applied Mathematics.
- Luo, Y. & Schuster, G.T., 1991. Wave-equation traveltime inversion, *Geophysics*, **56**(5), 645–653.
- Mattsson, K., Ham, F. & Iaccarino, G., 2008. Stable and accurate wave-propagation in discontinuous media, *J. Comput. Phys.*, **227**(19), 8753–8767.
- Nordström, J., 2006. Conservative finite difference formulations, variable coefficients, energy estimates and artificial dissipation, *J. Sci. Comput.*, **29**(3), 375–404.
- Symes, W.W., 2008. Migration velocity analysis and waveform inversion, *Geophys. Prospect.*, **56**(6), 765–790.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.
- Versteeg, R., 1994. The Marmousi experience: velocity model determination on a synthetic complex data set, *Leading Edge*, **13**(9), 927–936.
- Virieux, J., 1984. SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method, *Geophysics*, **49**(11), 1933–1942.
- Virieux, J., 1986. P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method, *Geophysics*, **51**(4), 889–901.
- Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics*, **74**(6), WCC1–WCC26.
- Zhang, Z., Zhang, W., Li, H. & Chen, X., 2013. Stable discontinuous grid implementation for collocated-grid finite-difference seismic wave modelling, *Geophys. J. Int.*, **192**(3), 1179–1188.

APPENDIX A: THE PROLONGATION AND THE RESTRICTION OPERATORS

As mentioned in Section 2.2, we need an interpolation operator to transfer solution from coarse grid to fine grid. The coarse grid points (red) and fine grid points (blue) involved are depicted in Fig. A1. The

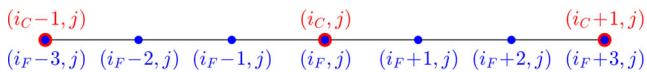


Figure A1. Layout of the coarse and fine grid points involved in the interpolation formulae in (A1).

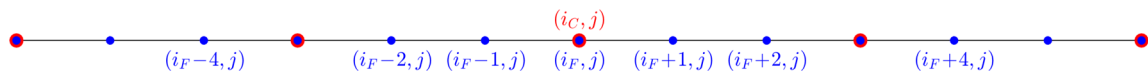


Figure A2. Layout of the coarse and fine grid points involved in the averaging formula (A2).

formulae presented in (A1) are used for the interpolation at fine grid points $(i_F - 1, j)$, (i_F, j) and $(i_F + 1, j)$. The same formula coefficients are used for the interpolation at other fine grid points, but with their respective neighbouring coarse grid points. By Taylor expansion, it can be shown that these formulae provide third-order accuracy. One can view these formulae as a system of linear equations, which can be written in the matrix-vector form, with the coefficient matrix being the prolongation operator \mathcal{P} :

$$\begin{aligned} f(i_F - 1, j) &= \frac{2}{9}f(i_C - 1, j) + \frac{8}{9}f(i_C, j) - \frac{1}{9}f(i_C + 1, j); \\ f(i_F, j) &= f(i_C, j); \\ f(i_F + 1, j) &= -\frac{1}{9}f(i_C - 1, j) + \frac{8}{9}f(i_C, j) + \frac{2}{9}f(i_C + 1, j). \end{aligned} \quad (\text{A1})$$

The corresponding companion restriction operator $\mathbf{c}_R \mathcal{R}^0 - \mathbf{c}_P \mathcal{P}^T$ transfers solution from fine grid to coarse grid. Regarding the second term $\mathbf{c}_P \mathcal{P}^T$, we have that each column of \mathcal{P} , scaled by the coefficient $\mathbf{c}_P = \Delta F / \Delta C$, gives a formula that approximates the corresponding coarse grid point value by averaging the neighbouring fine grid point values. In this particular case, all columns of \mathcal{P} are the same and lead to formula (A2) for points displayed in Fig. A2. By Taylor expansion, it can be shown that this weighted average provides fourth-order accuracy:

$$\begin{aligned} f(i_C, j) &= \frac{1}{3} \left(-\frac{1}{9}f(i_F - 4, j) + \frac{2}{9}f(i_F - 2, j) \right. \\ &\quad + \frac{8}{9}f(i_F - 1, j) + f(i_F, j) + \frac{8}{9}f(i_F + 1, j) \\ &\quad \left. + \frac{2}{9}f(i_F + 2, j) - \frac{1}{9}f(i_F + 4, j) \right). \end{aligned} \quad (\text{A2})$$

APPENDIX B: ELASTICITY, ENERGY, STIFFNESS AND COMPLIANCE TENSORS

Neglecting the source term and the boundary conditions, the elastic wave equation can be written in terms of stresses and velocities as:

$$\rho \dot{v}_i = \sigma_{ij,j}; \quad (\text{B1a})$$

$$\dot{\sigma}_{ij} = C_{ijkl} \frac{(v_{k,l} + v_{l,k})}{2}, \quad (\text{B1b})$$

where the Einstein summation convention applies. We use symbol ‘,’ to denote partial differentiation in space and symbol ‘ $\dot{}$ ’ to denote total differentiation in time. In (B1), ρ stands for density; C_{ijkl} is the fourth-order stiffness tensor; σ_{ij} is the second-order stress tensor; v_i is the velocity vector. Using the compliance tensor S_{klij} , that is, the inverse of the stiffness tensor, (B1) can be written as

$$\rho \dot{v}_i = \sigma_{ij,j}; \quad (\text{B2a})$$

$$S_{klij} \dot{\sigma}_{ij} = \frac{v_{k,l} + v_{l,k}}{2}. \quad (\text{B2b})$$

Moreover, we use u_i and ϵ_{ij} to denote the displacement vector and the strain tensor, with which we have $\dot{u}_i = v_i$ and $\epsilon_{ij} = \frac{(u_{j,i} + u_{i,j})}{2}$.

The strain energy (potential energy) density function of the elastic medium is defined as

$$w_p = \frac{1}{2} \epsilon_{ij} C_{ijkl} \epsilon_{kl} = \frac{1}{2} \sigma_{ij} S_{ijkl} \sigma_{kl}. \quad (\text{B3})$$

The kinetic energy density function is defined as

$$w_k = \frac{1}{2} \rho v_i v_i. \quad (\text{B4})$$

According to the elastic wave eq. (B2), the density function of the total energy, that is, $w = w_k + w_p$, of the dynamical system satisfies

$$\dot{w} = v_i \sigma_{ij,j} + \sigma_{kl} \frac{v_{k,l} + v_{l,k}}{2} = v_i \sigma_{ij,j} + \sigma_{kl} v_{k,l}. \quad (\text{B5})$$

If the stresses σ_{ij} are arranged as a vector, say in the Voigt form, (B5) leads to

$$\dot{E} = V^T K^\Sigma \Sigma + \Sigma^T K^V V = [\Sigma^T, V^T] \begin{bmatrix} \mathbf{0} & K^V \\ K^\Sigma & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Sigma \\ V \end{bmatrix} \quad (\text{B6})$$

after spatial discretization, where V and Σ are discretizations of the velocity and stress vectors, K^V and K^Σ are the matrices resulted from spatial differentiation and E is the total discrete energy. The similarity between (16) and (B6) indicates that the analytical techniques presented in this paper can be naturally extended to the elastic case.

APPENDIX C: ADJUSTMENT OF THE DIFFUSION COEFFICIENT FOR PHYSICAL SCALES

In this appendix, we discuss the scale of the horizontal diffusion terms for the acoustic wave eq. (8), whose parameters, that is, ρ and c , are now of the respective physical scales that are pertinent to the earth media. We assume the homogeneity of ρ and c so that they can be taken in or out of the spatial derivatives.

Since we already have the experience on the proper scales for the unit-parameter case, we first non-dimensionalize the acoustic

wave equation via variable substitution. We define the following new variables

$$X = \frac{x}{c}; \quad Y = \frac{y}{c}; \quad q = \frac{p}{\rho c}, \quad (\text{C1})$$

such that (8) can be re-written as

$$\begin{cases} \frac{\partial q}{\partial t} = - \left(\frac{\partial v_x}{\partial X} + \frac{\partial v_y}{\partial Y} \right); \\ \frac{\partial v_x}{\partial t} = - \frac{\partial q}{\partial X}; \\ \frac{\partial v_y}{\partial t} = - \frac{\partial q}{\partial Y}. \end{cases} \quad (\text{C2})$$

Based on the variable relations in (C1), if $x_2 - x_1 = \Delta C$, we have $X(x_2) - X(x_1) = \frac{\Delta C}{c}$. In other words, the grid-spacing ΔC in x translates into grid-spacing $\frac{\Delta C}{c}$ in X . From Section 3.4, we know that $(\frac{\Delta C}{c})^2$ is the proper scale for the horizontal diffusion terms to be introduced to (C2), which is now modified to

$$\begin{cases} \frac{\partial q}{\partial t} = - \left(\frac{\partial v_x}{\partial X} + \frac{\partial v_y}{\partial Y} \right) + \left(\frac{\Delta C}{c} \right)^2 \frac{\partial^2 q}{\partial X^2}; \\ \frac{\partial v_x}{\partial t} = - \frac{\partial q}{\partial X}; \\ \frac{\partial v_y}{\partial t} = - \frac{\partial q}{\partial Y} + \left(\frac{\Delta C}{c} \right)^2 \frac{\partial^2 v_x}{\partial X^2}. \end{cases} \quad (\text{C3})$$

Translating (C3) back into the original variables p , x and y , we have

$$\begin{cases} \frac{1}{\rho c^2} \frac{\partial p}{\partial t} = - \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} \right) + \frac{(\Delta C)^2}{\rho c^2} \frac{\partial^2 p}{\partial x^2}; \\ \rho \frac{\partial v_x}{\partial t} = - \frac{\partial p}{\partial x}; \\ \rho \frac{\partial v_y}{\partial t} = - \frac{\partial p}{\partial y} + \rho (\Delta C)^2 \frac{\partial^2 v_x}{\partial x^2}. \end{cases} \quad (\text{C4})$$

As illustrated in (C4), the scales of the physical parameters enter the scales of the horizontal diffusion terms. In the Marmousi example, cf. Section 4.2, where heterogeneous media are involved, we use the mean value of ρc^2 and ρ for scaling purpose.