

1 **GPU-accelerated 3-D viscoacoustic wave equation**
2 **modeling with natural-attenuation absorbing boundary**
3 **conditions**

4 **Bei Li¹, Yunyue Elita Li¹, and Arthur Cheng¹**

5 ¹Department of Civil and Environmental Engineering, National University of Singapore, 117576 Singapore

6 **Abstract**

7 Viscoacoustic wave equation with decoupled fractional Laplacians (DFLs) describes P-
8 wave propagation through the anelastic earth under the assumption of frequency-independent
9 Q . Numerical solutions of this equation often employ 2nd-order finite differences for the
10 time derivatives and spectral methods for the spatial DFLs. The spectral methods re-
11 quire multiple Fourier transforms at each time step, leading to large computational cost
12 especially in 3-D. To accelerate the viscoacoustic wave simulation, GPU-based compu-
13 tation is preferred. We propose new absorbing boundary conditions (ABCs) for GPU-
14 accelerated viscoacoustic wave equation modeling. The new ABCs naturally attenuate
15 the outgoing waves by padding the original velocity and Q models with optimized ab-
16 sorbing layers. In comparison with the conventional ABCs, the natural-attenuation ABCs
17 (naABCs) require no special treatment for the absorbing layers, thus provide around 2x
18 speedup for the GPU-based time marching process. We also simplify the 2nd-order k -
19 space compensators for the time-stepping error introduced by the 2nd-order temporal
20 FDs. Finally, we develop a python package for 3-D DFL viscoacoustic wave equation mod-
21 eling with the application of naABCs and simplified time-stepping error compensators.
22 The modeling examples not only validate the feasibility of our package in simulating the
23 viscoacoustic wave propagation, but also demonstrate the satisfying absorbing effect and
24 advantageous efficiency of the naABCs.

25 **Keywords**— 3-D, viscoacoustic, absorbing boundary condition, GPU

Corresponding author: Yunyue Elita Li, elita.li@nus.edu.sg

26 1 Introduction

27 Seismic waves experience attenuation when propagating through real earth due to
 28 underground anelasticity. High seismic attenuation, quantified by low quality factor Q ,
 29 often associates with gas accumulation (Zhu et al., 2017) or magma bodies (Sanders et
 30 al., 1995). The attenuated seismic waveforms exhibit both amplitude decay and phase
 31 distortion, which requires Q compensation when building underground elastic models
 32 from seismic data space (Zhang et al., 2013; Dutta & Schuster, 2014). Since the atten-
 33 uation is considered to be almost linear with frequency, Q is reasonably assumed con-
 34 stant in many frequency bands (McDonald et al., 1958). Based on the precise constant-
 35 Q model (Kjartansson, 1979), Zhu and Harris (2014) developed a viscoacoustic wave equa-
 36 tion with decouple fractional Laplacians (DFLs). This equation avoids saving the wave-
 37 field history required by conventional constant- Q wave equation with fractional time deriva-
 38 tive (Carcione et al., 2002). More importantly, due to separated terms representing am-
 39 plitude decay and phase distortion, this wave equation accomplishes accurate adjoint prop-
 40 agation with attenuation compensation by simply reversing the sign of the amplitude-
 41 decay term. Therefore, implementation with DFLs has been widely applied in Q -compensated
 42 reverse time migration (RTM) (Zhu et al., 2014), least-squares RTM (LSRTM) (Sun et
 43 al., 2016), and full waveform inversion (FWI) (Xue et al., 2018).

44 Despite its popularity, DFL viscoacoustic wave equation is difficult to solve due to
 45 the spatially varying power of DFLs in a heterogeneous Q model. Zhu et al. (2014) achieved
 46 spatially constant power by averaging the heterogeneous Q so that simple pseudospec-
 47 tral method can be applied. Such implementation compromises the modeling accuracy
 48 and stability, especially when large Q contrasts exist. Chen et al. (2016) also achieved
 49 spatially constant power with an extra Laplacian term based on Taylor expansion. It pro-
 50 vides better approximation than the averaging strategy, but a single optimal power still
 51 shows accuracy deterioration for low Q values. Yao et al. (2017) proposed local solvers
 52 for DFLs using Hermite distributed approximating functional method. However, due to
 53 large rectangular stencils, the computational cost is not reduced in comparison with spec-
 54 tral methods. To achieve a flexible balance between the accuracy and efficiency, low-rank
 55 approximation (Fomel et al., 2013) is adapted for solving DFL viscoacoustic wave equa-
 56 tion. Chen et al. (2014) proposed the two-step extrapolation scheme, where DFLs are
 57 approximated using low-rank decomposition, while time derivatives are approximated
 58 using 2nd-order FDs. With superior temporal accuracy and stability, Sun et al. (2016)
 59 proposed the one-step extrapolation scheme, which utilizes low-rank decomposition to
 60 directly approximate the phase of the complex analytic wavefield. Chen et al. (2019) and
 61 N. Wang et al. (2020) developed 1st and 2nd-order k -space compensators for time-stepping
 62 errors in the two-step extrapolation, leading to improved temporal accuracy and stabil-
 63 ity while still marching the real wavefield along time. Although these compensators hardly
 64 influence the computational cost for the time marching process, they do increase the com-
 65 putational complexity of the mixed-domain operators for low-rank decomposition. Here,
 66 we simplify the 2nd-order k -space compensators (N. Wang et al., 2020) for more efficient
 67 low-rank decomposition with barely compromised accuracy and stability. At each time
 68 step, multiple forward and inverse FFTs on the wavefield snapshots are necessary, re-
 69 sulting in enormous computational cost in 3-D modeling scenarios. Therefore, GPU ac-
 70 celeration has been utilized to improve the efficiency of viscoacoustic wave simulation
 71 (N. Wang et al., 2020) and its application in Q -compensated RTM (Y. Wang et al., 2019).

72 To properly simulate the wave propagation in the earth whose dimensions are much
 73 larger than the computational domain, absorbing boundary conditions (ABCs) are nec-
 74 essary to suppress the wraparound effects in the bounded modeling area. Commonly used
 75 ABCs, such as Perfectly Matched Layers (PMLs) (Berenger, 1994) and hybrid ABCs (hABCs)
 76 (Liu & Sen, 2010), have been applied in viscoacoustic wave equation modeling (Ren &
 77 Liu, 2013; Chen et al., 2019). These conventional ABCs involve special treatments for
 78 the absorbing layers at each time step, e.g., updating one-way-wave-equation (OWWE)

for hABCs, or calculating extra convolutional terms for convolutional PMLs, therefore, impede the modeling efficiency. Due to smaller memory, GPU implementations of such boundary conditions slow down more significantly compared to CPU implementations. We propose the natural-attenuation ABCs (naABCs), which extend the velocity and Q models with optimized absorbing layers, so that the outgoing waves will attenuate naturally without any special treatments during the time marching process. The optimization for the minimum absorbing-layer number and corresponding model parameters (velocity and Q) at each layer, is established upon three constraints: (1) predetermined tolerance for the boundary artifacts (artificial reflections and wraparounds) amplitude, (2) Nyquist wavenumber, and (3) stability condition. We develop an efficient algorithm to solve this optimization and provide padding strategy for heterogeneous media. Since applying naABCs is simply extending the original velocity and Q models before the time-marching process, they are readily applicable to any other numerical schemes for constant- Q viscoacoustic wave equation modeling either in the time or frequency domain, such as global finite-differences (Xu et al., 2018) and the one-step extrapolation scheme (Sun et al., 2016).

Utilizing the naABCs and the two-step extrapolation with time-stepping error compensation, we develop a python package for GPU-accelerated 3-D DFL viscoacoustic wave equation modeling. This open-source modeling package, as the major contribution of this paper, features a great balance between accuracy and efficiency, achieved by innovations in the new naABCs, and the simplified time-stepping error compensation.

We organize this article as follows: first, we introduce the DFL viscoacoustic wave equation and its two-step extrapolation scheme with the simplified time-stepping error compensation; next, we propose the new naABCs and provide two analytical tests; finally, we use two modeling examples to demonstrate the satisfying absorbing effect and superior modeling efficiency.

2 Methods

2.1 DFL viscoacoustic wave equation

The DFL viscoacoustic wave equation (Zhu & Harris, 2014) is defined as

$$\frac{1}{v^2} \frac{\partial^2 p}{\partial t^2} = \eta(-\nabla^2)^{1+\gamma} p + \tau \frac{\partial}{\partial t} (-\nabla^2)^{1/2+\gamma} p, \quad (1)$$

in which

$$\begin{cases} \eta = -v_0^{2\gamma} \omega_0^{-2\gamma} \cos(\pi\gamma), \tau = -v_0^{2\gamma-1} \omega_0^{-2\gamma} \sin(\pi\gamma) \\ v = v_0 \cos(\pi\gamma/2), \gamma = \arctan(1/Q)/\pi \end{cases}, \quad (2)$$

where $p = p(\mathbf{x}, t)$ represents P-wave wavefield varying in 3-D space ($\mathbf{x} = (x, y, z)$) and time (t), $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ represents the Laplacian operator in \mathbf{x} , and v_0 represents the reference velocity corresponding to the reference angular frequency ω_0 . The reference velocity and Q models can be heterogeneous in space, whereas ω_0 is fixed for the whole model.

Equation 1 can be solved by two-step extrapolation scheme (e.g., Chen et al., 2014), where 2nd-order FDs are applied for time derivatives, and pseudospectral methods for DFLs. The recursion equation is

$$p(\mathbf{x}, t + \Delta t) = \Delta t^2 v^2(\mathbf{x}) \left\{ \begin{array}{l} \eta(\mathbf{x}) \int \mathbf{W}_d(k, \mathbf{x}) \hat{p}(\mathbf{k}, t) \exp(i\mathbf{k}\mathbf{x}) d\mathbf{k} + \\ \tau(\mathbf{x}) \int \mathbf{W}_i(k, \mathbf{x}) (\hat{p}(\mathbf{k}, t) - \hat{p}(\mathbf{k}, t - \Delta t)) \exp(i\mathbf{k}\mathbf{x}) d\mathbf{k} \end{array} \right\} + 2p(\mathbf{x}, t) - p(\mathbf{x}, t - \Delta t), \quad (3)$$

120 in which

$$\mathbf{W}_d(k, \mathbf{x}) = k^{2\gamma(\mathbf{x})+2}, \quad (4)$$

$$\mathbf{W}_l(k, \mathbf{x}) = k^{2\gamma(\mathbf{x})+1}, \quad (5)$$

$$\hat{p}(\mathbf{k}, t) = \int p(\mathbf{x}, t) \exp(-i\mathbf{k}\mathbf{x}) d\mathbf{x}, \quad (6)$$

125 where i represents the imaginary unit, Δt represents the temporal step size, $\mathbf{k} = (k_x, k_y, k_z)$
126 represents the wavenumber vector, $k = (k_x^2 + k_y^2 + k_z^2)^{1/2}$ represents the norm of \mathbf{k} , and
127 $\hat{p}(\mathbf{k}, t)$ represents the 3-D Fourier transform of $p(\mathbf{x}, t)$. \mathbf{W}_d and \mathbf{W}_l are mixed \mathbf{k} - \mathbf{x} do-
128 main operators, which are applied efficiently through low-rank approximation (Fomel et
129 al., 2013).

130 Furthermore, to compensate for the time-stepping error caused by the 2nd-order
131 FDs in Equation 3, an efficient strategy revises the mixed-domain operators in equations 4
132 and 5 as follows:

$$\mathbf{W}_d(k, \mathbf{x}) = \phi_d k^{2\gamma(\mathbf{x})+2}, \quad (7)$$

$$\mathbf{W}_l(k, \mathbf{x}) = \phi_l k^{2\gamma(\mathbf{x})+1}, \quad (8)$$

136 where ϕ_d and ϕ_l represent the compensators for the time-stepping errors of dispersion
137 and loss terms, respectively. Here, we simplify the time-stepping-error compensators pro-
138 posed by N. Wang et al. (2020) as follows:

$$\phi_d = \text{sinc}^{2\gamma+2}(v_0(\mathbf{x})\Delta tk/2), \quad (9)$$

$$\phi_l = \text{sinc}^{2\gamma+1}(v_0(\mathbf{x})\Delta tk/2). \quad (10)$$

142 The simplified compensators exhibit almost identical accuracy and stability to the orig-
143 inal compensators, but are more efficient to compute during low-rank decomposition. The
144 compensated \mathbf{W}_d and \mathbf{W}_l in equations 7 and 8 are still mixed-domain operators, which
145 require the same number of FFTs as the uncompensated case with the same ranks.

146 At each modeling time step, the computational cost of DFLs is dominated by two
147 forward 3-D FFTs and $(n+m)$ inverse 3-D FFTs, where n and m are the ranks along
148 \mathbf{x} and \mathbf{k} axes for low-rank decomposition. They are integers typically between 2 and 4.
149 In our package, we use the CuPy library for its simple python interface between the Numpy
150 library and CUDA programming on GPU.

151 2.2 Natural-attenuation absorbing boundary conditions

152 Using CuPy library, the FFTs and elementwise operations on the entire 3-D wave-
153 field are well organized to make full use of the GPU architecture. However, conventional
154 ABCs require special treatments particularly on the absorbing layers, which cause dis-
155 proportionately long runtime due to high latency from uncoalesced memory accesses in
156 GPU. To eliminate the excessive memory swaps, we develop the naABCs, which avoid
157 any special treatments for the absorbing layers during time marching process. The naABCs
158 attenuate the outgoing waves naturally by padding the v_0 and Q models with optimized
159 absorbing layers.

160 Considering outgoing waves at normal incidence, the returning wave amplitude from
161 the i th absorbing layer is

$$A_i(\omega) = R_i(\omega) \prod_{j=0}^{i-1} (1 - R_j^2(\omega)) \exp\left(-\frac{\pi}{2} \omega \sum_{j=0}^i \Psi_j(\omega)\right), \quad i = 0, 1, \dots, N, \quad (11)$$

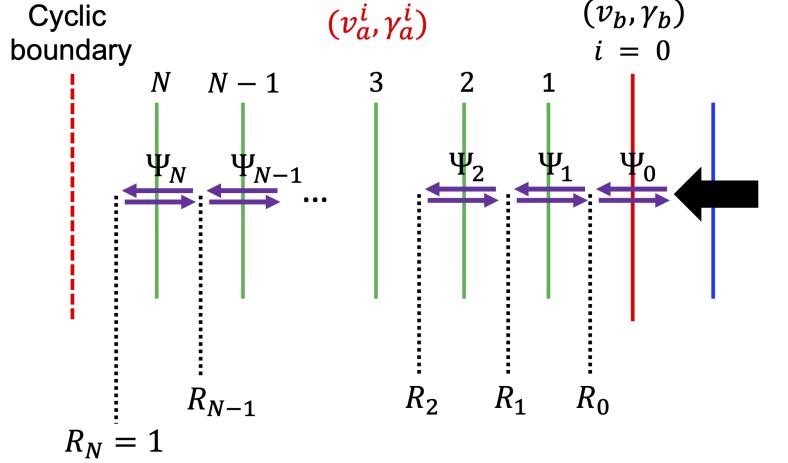


Figure 1. Illustration of natural-attenuation ABCs. The blue lines indicate the inner area, red line the model boundary, and green lines absorbing layers. The dashed red line represents the numerical cyclic boundary automatically imposed by FFTs. v_a^i and γ_a^i represent the model parameters for the i th absorbing layer, whereas v_b and γ_b represent boundary parameters. i ranges from 0 to N , which is the total number of extended absorbing layers. $i = 0$ represents the model boundary layer, thus, $v_a^0 = v_b$ and $\gamma_a^0 = \gamma_b$. R_0 to R_N represent reflectivities from the boundary layer to the outermost absorbing layer, and Ψ_0 to Ψ_N represent the two-way attenuated traveltime from the boundary layer to the outermost absorbing layer. Technically, R_N should be 0 due to the cyclic boundary. However, we can reasonably assume similar situation in the absorbing layers of the opposite boundary, so that R_N is effectively 1 for simplicity.

163 in which

$$164 R_i(\omega) = \begin{cases} \frac{|v_d^{i+1}(\omega) - v_d^i(\omega)|}{v_d^{i+1}(\omega) + v_d^i(\omega)}, & i \neq N \\ 1, & i = N \end{cases}, \quad (12)$$

$$165 \Psi_i(\omega) = 2h\gamma_a^i/v_d^i(\omega), \quad (13)$$

$$166 v_d^i(\omega) = v_a^i \cos(\pi\gamma_a^i/2)(\omega/\omega_0)^{\gamma_a^i}, \quad (14)$$

168 where N represents the total number of extended absorbing layers, γ_a^i and v_a^i represent
169 model parameters at the i th layer, v_d^i represents the dispersive velocity at the i th layer,
170 R_i represents the reflectivity between the $(i+1)$ th layer and the i th layer, and Ψ_i rep-
171 resents the two-way attenuated traveltime (Cavalca et al., 2011) within the i th layer. Fig-
172 ure 1 shows an illustration of the absorbing boundary, the parameters, and the proper-
173 ties of each layer.

174 The returning wave amplitude A_i considers the primary reflection, transmission and
175 attenuation for each sampled frequency ω within the frequency range of the source wavelet.
176 Given the boundary parameters $\gamma_a^0 = \gamma_b$ and $v_a^0 = v_b$, we define a constrained opti-

177 mization to solve for the minimum N and (v_a, γ_a) for each layer:

$$178 \quad \arg \min_{(v_a^i, \gamma_a^i)} N, \quad i = 1, 2, \dots, N \quad (15)$$

$$179 \quad \text{s.t. } A_{i-1}(\omega) \leq \epsilon_i, \quad (16)$$

$$180 \quad \text{s.t. } \frac{\omega d}{v_d^i(\omega)} \leq \pi, \quad (17)$$

$$181 \quad \text{s.t. } \frac{v_a^i \Delta t}{h} \leq s(\gamma_a^i), \quad (18)$$

183 where h is the grid size. Inequation 16 requires that the returning wave amplitude from
 184 the $(i-1)$ th absorbing layer must be smaller than the corresponding tolerance, denoted
 185 as ϵ_i . Inequation 17 requires that all modeled wavenumber must be smaller than the Nyquist
 186 wavenumber. Inequation 18 requires that each layer's model parameters must satisfy the
 187 stability condition, given fixed ω_0 and h (see Appendix A for details about the stabili-
 188 ty condition). In Appendix B, we provide an efficient algorithm to solve this constrained
 189 optimization problem.

190 To test the naABCs analytically, we assume modelling parameters as $\omega_0 = 20$ Hz,
 191 $h = 20$ m, $\Delta t = 2$ ms, and model boundary parameters as $v_b = 3.183 \times 10^{-3}$ (corre-
 192 sponding to $Q = 100$) and $v_b = 2000$ m/s. The modeling frequency ranges from 5 Hz
 193 to 50 Hz. We set the maximum A_i tolerance as $\epsilon = 0.04$ and 0.01 for two tests, respec-
 194 tively (see Appendix B for details about ϵ and its corresponding ϵ_i).

195 When $\epsilon = 0.04$, the minimum number of absorbing layers is 17, and the optimized
 196 (v_a, γ_a) are displayed in Figure 2(a). From the first absorbing layer to the outermost ab-
 197 sorbing layer, γ_a increases monotonically, whereas v_a first increases, then decreases. The
 198 turning point for v_a variation at layer 10 is caused by the intervention of the stability
 199 condition in inequation 18.

200 In Figure 2(b), we show the returning wave amplitude in gray scale from each ab-
 201 sorbing layer for each sampled frequency based on equation 11. Returning wave ampli-
 202 tude at all frequencies is effectively absorbed. The red line shows the amplitude toler-
 203 ance varying with the layer index. The tolerance is gradually increasing from $\epsilon_1 = 0.004$
 204 to $\epsilon_N = \epsilon = 0.04$. Since the remaining unabsorbed frequencies become fewer as the
 205 outgoing waves propagate deeper into the absorbing layers, the maximum returning wave
 206 amplitude ϵ_i should be relaxed to avoid excessive layers just for ideal suppression of a
 207 single frequency.

208 If better absorption is required, we can decrease the maximum tolerance ϵ . Fig-
 209 ure 2(c) shows the optimized (v_a, γ_a) when $\epsilon = 0.01$. Now the minimum number of ab-
 210 sorbing layers is 33. The variations of v_a and γ_a are similar to those shown in Figure 2(a),
 211 although slightly smoother. Figure 2(d) shows the returning wave amplitude and the tol-
 212 erance varying along the layer index. Both of them are much smaller than those in the
 213 previous test when $\epsilon = 0.04$.

214 When applying the naABCs to heterogeneous media with predetermined ϵ , we need
 215 to solve for N and the corresponding (v_a, γ_a) for each unique (v_b, γ_b) pair beyond the
 216 resolution limit of ($\Delta v_b = 1$ m/s, $\Delta \gamma_b = 1e-5$). When the optimal number of ab-
 217 sorbing layers is different for different parts of the model, we choose the maximum layer
 218 number, and extend the inner absorbing layers by repeating (v_b, γ_b) . It effectively push
 219 the boundary outward for the parts with fewer absorbing layers so that the entire model
 220 has a consistent outermost absorbing layer in view of its $A_i(\omega)$ tolerance ϵ .

221 Now we can achieve satisfying boundary absorption by directly simulating the vis-
 222 coacoustic wave propagation using the padded v_0 and γ models with existing simulation
 223 codes for any computer infrastructure. Such an implementation avoids any extra par-
 224 tial slicing and computations specifically for the absorbing layers. The efficiency improve-

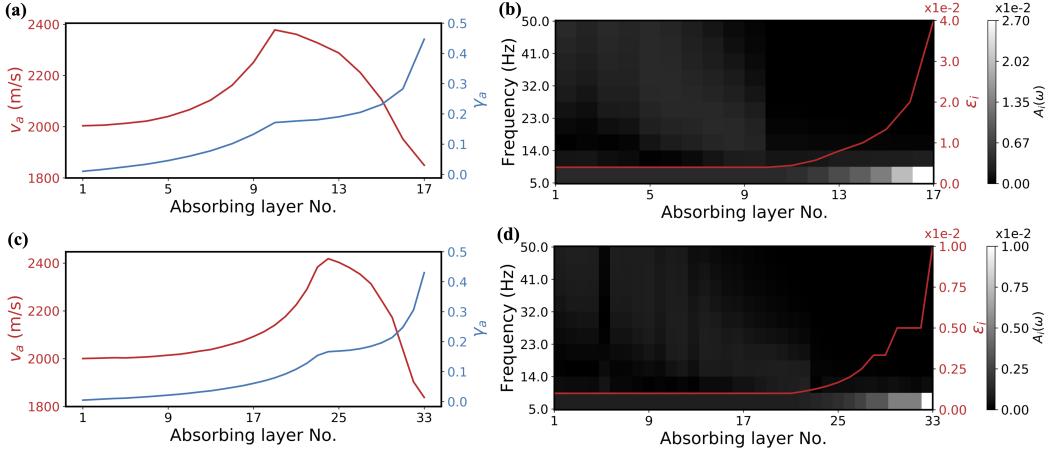


Figure 2. The optimized (v_a, γ_a) for naABCs and their corresponding returning wave amplitude for each sampled frequency from each absorbing layer. (a) and (b) are obtained when $\varepsilon = 0.04$; (c) and (d) are obtained when $\varepsilon = 0.01$.

ment by the streamlined computation is particularly significant for GPU implementations by avoiding additional delicate domain decomposition and frequent swapping of data in and out of the limited GPU memory.

3 Modeling examples

We use two modeling examples to demonstrate the effectiveness and efficiency of our modeling package with the application of naABCs. The first example is homogeneous with $v_0 = 2000$ m/s and $Q = 100$. The second example is heterogeneous, of which the velocity model is the truncated Overthrust model (Aminzadeh et al., 1997), and the Q model is derived from the velocity model by $Q = 3.516(v_0/1000)^{2.2}$ (Li, 1993). The model discretizations and acquisition configurations for these two examples are shown in Table 1. The ranks along both \mathbf{x} and \mathbf{k} for the low-rank decomposition are 2 in the homogeneous example, and 4 in the heterogeneous example. In both examples, we benchmark the absorbing effect and modeling efficiency of naABCs with the hABCs, whose acoustic application with GPU acceleration have been reported (Xie et al., 2020). The number of absorbing layers for hABCs is 20 as suggested by Liu and Sen (2018). All computations are performed on a single TITAN RTX GPU card with memory of 24GB.

3.1 Homogeneous example

We set up the homogeneous model to evaluate the absorbing effect on waves with different incident angles. The four source-receiver pairs provide four different incident angles at approximately $0^\circ, 25^\circ, 50^\circ$ and 75° . The modeling parameters in this example are the ones used in the analytical tests for the naABCs. Hence, we pad the homogeneous velocity and Q models according to the optimized (v_a, γ_a) in Figures 2(a) and 2(c), respectively. The source function is the 2nd-order derivative of a Ricker wavelet with dominant frequency as 15 Hz.

In Figure 3, we show the snapshots at 800 and 1600 ms when using naABCs and hABCs, respectively. For naABCs with $N = 17$ ($\varepsilon = 0.04$), there are weak, yet visible boundary reflections and wraparounds (see Figures 3(a) and 3(b)). For naABCs with $N = 33$ ($\varepsilon = 0.01$) and hABCs with $N = 20$ (see Figures 3(c), 3(d), 3(e), and 3(f)),

Table 1. Model discretizations and acquisition configurations

	Homogeneous model	Heterogeneous model
Model grids ($x \times y \times z$)	$81 \times 301 \times 81$	$601 \times 301 \times 187$
Grid size (m)	20	15
Total recording time (s)	2	2
Time sampling rate (ms)	2	1
Reference frequency (Hz)	20	20
Total sources	1	1
Total receivers	4	251×51
Source location (m)	$x = 800, y = 2000, z = 800$ $x = 800$	$x = 4500, y = 2250, z = 150$ $x = \{750 + 30i i = 0, 1, \dots, 250\}$
Receiver locations (m)	$y = \{2000, 2360, 2940, 4980\}$ $z = 100$	$y = \{750 + 60i i = 0, 1, \dots, 50\}$ $z = 75$

Table 2. Runtime of different procedures for 3-D viscoacoustic modeling with different ABCs in the homogeneous model.

Runtime (s) \ ABCs	naABCs ($\varepsilon = 0.04$) $N = 17$	naABCs ($\varepsilon = 0.01$) $N = 33$	hABCs $N = 20$
Procedure			
ABCs preparation	6.55	9.10	1.53
Low-rank decomposition	1.20	1.89	1.18
Modeling (1000 time steps)	9.14	18.98	58.76
Total	16.89	29.97	61.47

we barely see any boundary artifacts in the snapshots, demonstrating excellent absorbing effect.

Figure 4 shows the recorded waveforms in comparison with the references, which are obtained by modeling with the same parameters, however, in a expanded model space ($401 \times 401 \times 401$) with the implicit periodic boundary conditions. The hABCs show the smallest differences with the references at all tested incident angles, where the returning wave amplitude are down to 0.1% of the incident waves. Although inferior to hABCs in terms of absorbing effect, naABCs still exhibit significant absorption. Especially when $N = 33$ ($\varepsilon = 0.01$), naABCs also achieve around 0.1% boundary artifacts for relatively small and moderate incident angles as 0° , 25° , and 50° (see Figures 4(b), 4(d), and 4(f)), whereas around 1% for incident angle as large as 75° (see Figure 4(h)).

Table 2 shows the runtime of different procedures when modeling with different ABCs. The preparation time for naABCs is longer, due to the extra parameter optimization for the absorbing layers. For low-rank decomposition, the computational cost is dominated by composing submatrices of the mixed-domain operators shown in equations 7 and 8. Thus, more absorbing layers in naABCs ($\varepsilon = 0.01$) lead to slightly increased runtime. Compared to hABCs with $N = 20$, the time marching process, which contributes most to the runtime, shows approximately 6x and 3x speedup when using naABCs with $N = 17$ and $N = 33$, respectively. Therefore, for a single shot with merely 1000 time steps, the overall modeling time with naABCs can achieve at least 2x speedup over hABCs. Since the ABCs preparation and the low-rank decomposition only need to be done once for given models and parameters, the speedup ratio will eventually converge to that for the pure modeling part with the increasing shot number and modeling time steps.

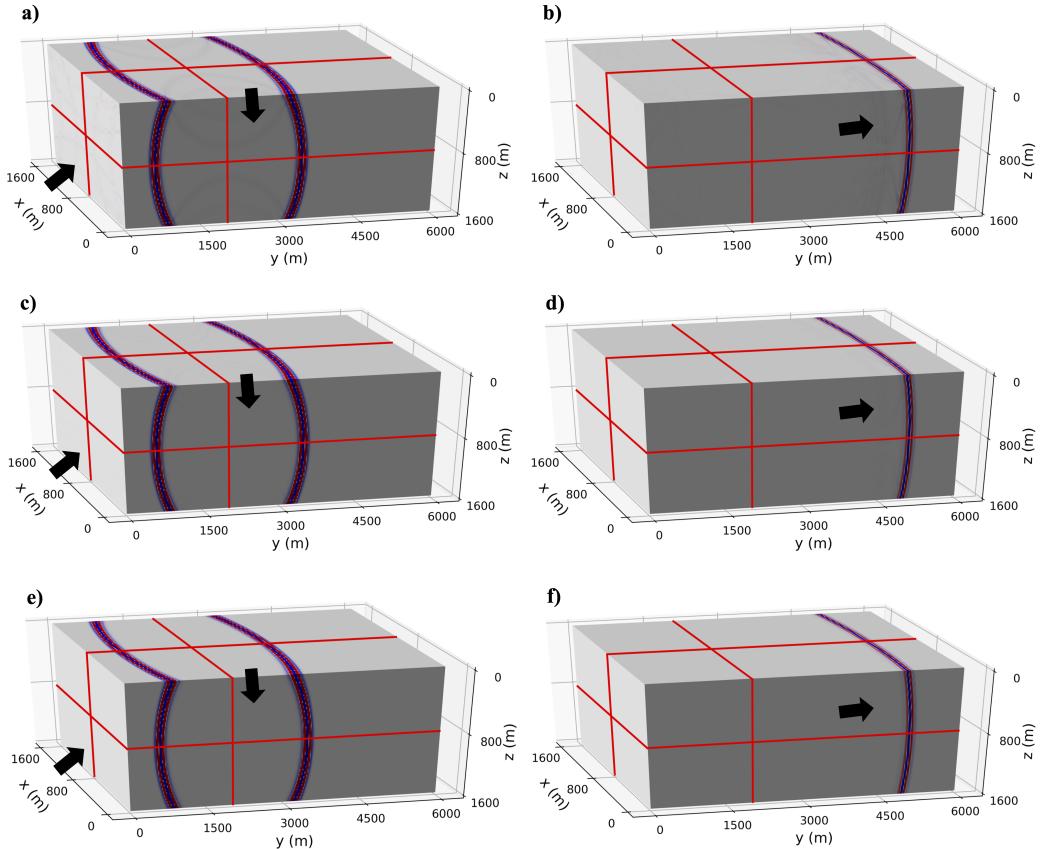


Figure 3. Snapshots at 800 ms (left column) and 1600 ms (right column) modeled using naABCs with $N = 17$ ($\varepsilon = 0.04$) (a and b), naABCs with $N = 33$ ($\varepsilon = 0.01$) (c and d), and hABCs with $N = 20$ (e and f). The red lines represent the slice positions in the 3-D wavefield volume. The black arrows indicate the potential positions of artifacts on the snapshots.

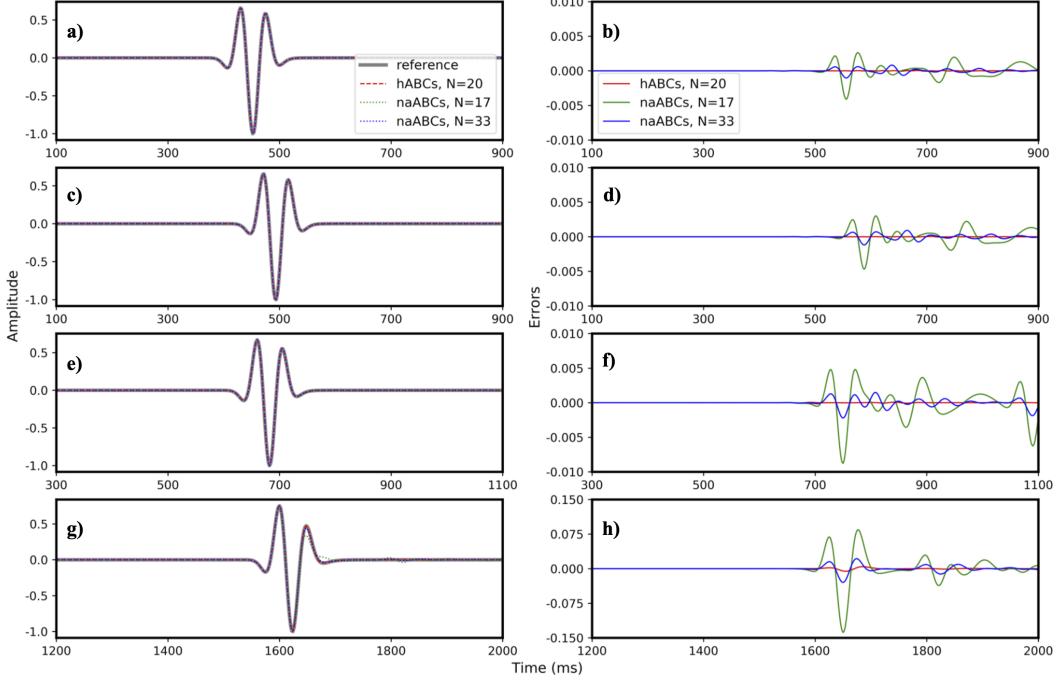


Figure 4. Comparison between normalized references and recorded waveforms modeled using naABCs and hABCs. The four rows correspond to the incident angles of 0° (a and b), 25° (c and d), 50° (e and f), and 75° (g and h). The left column displays the normalized references and modeled waveforms, while the right column shows their differences.

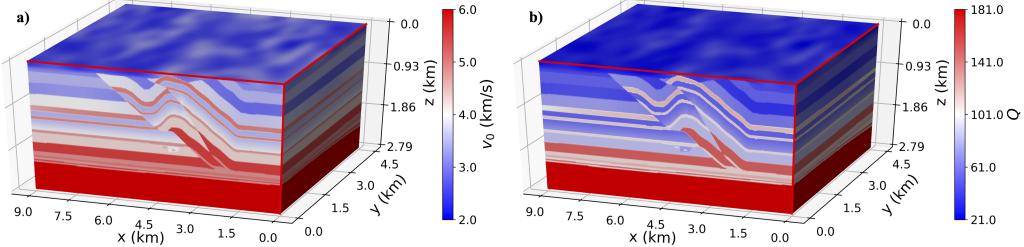


Figure 5. (a) Truncated Overthrust velocity model and its corresponding (b) Q model.

3.2 Heterogeneous example

We employ the models shown in Figure 5 to evaluate the accuracy and efficiency of our modeling package for complex heterogeneous model. The source function here is the 2nd-order derivative of a Ricker wavelet with dominant frequency of 20 Hz. Correspondingly, the frequency for naABCs optimization ranges from 8 Hz to 62 Hz. There are total 1708 unique pairs of (v_b, γ_b) . We use $\varepsilon = 0.05$, resulting in maximum number of absorbing layers as 44. Figure 6 displays the padded velocity and Q model slices at $y = 2000$ m for naABCs. The padded absorbing layers show much smaller velocity and Q in order to attenuate the outgoing waves naturally.

Figure 7 displays the modeled 3-D seismic data volume using our modeling package with naABCs and hABCs, respectively. The two modeled seismograms are consis-

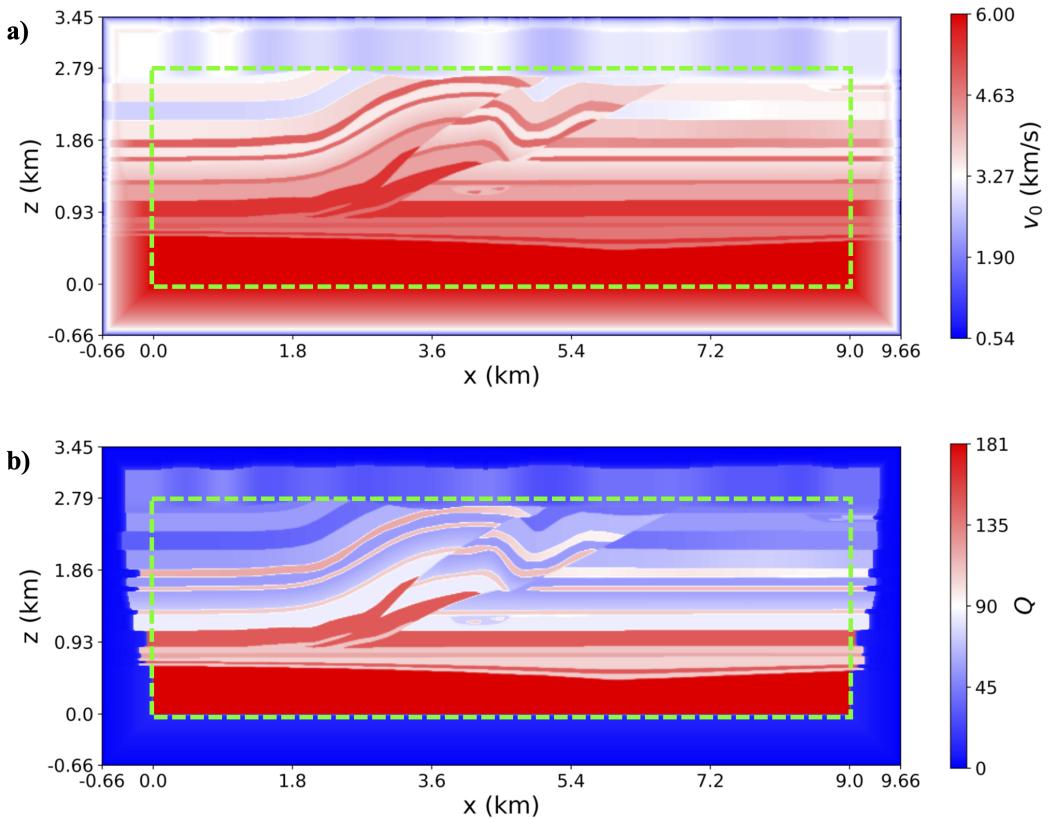


Figure 6. Padded (a) Velocity and (b) Q model slices at $y = 2000$ m for naABCs ($\varepsilon = 0.05$). The green dashed box delineates the original model area.

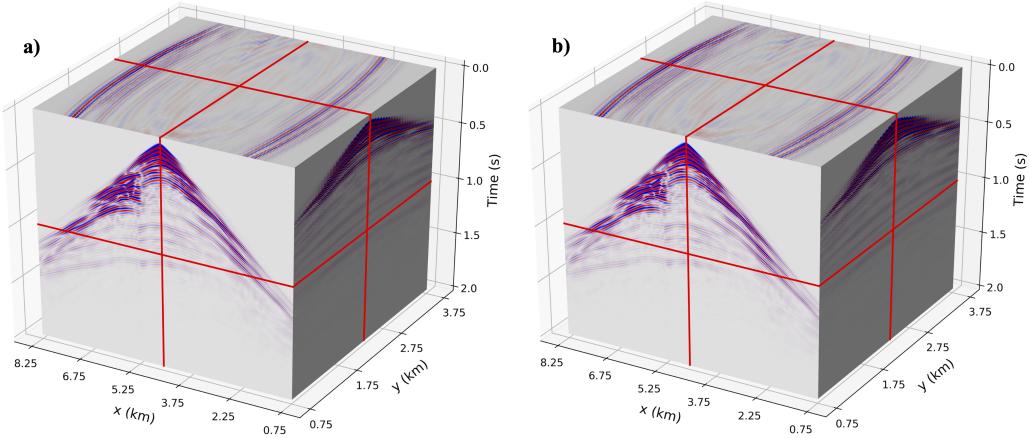


Figure 7. Modeled 3-D seismic data volume using (a) hABCs with $N = 20$, and (b) naABCs with $N = 44$ ($\varepsilon = 0.05$). The red lines indicate the slice positions for displaying the 3-D data volume.

Table 3. Runtime of different procedures for 3-D viscoacoustic modeling with different ABCs in the heterogeneous model.

Runtime (s) \ ABCs	naABCs ($\varepsilon = 0.05$) $N = 44$	hABCs $N = 20$
Procedure		
ABCs preparation	125.72	8.01
Low-rank decomposition	38.33	25.35
Modeling (2000 time steps)	469.31	995.56
Total	633.36	1028.92

tent with each other without any visible artifacts. We show two specific traces with near (670 m) and far (3795 m) offsets, respectively, in Figure 8. The waveforms from naABCs (blue dashed lines) agree reasonably well with those from hABCs (red lines).

Table 3 shows the runtime of different procedures when applying different ABCs in this modeling example. The runtime for all procedures increases significantly due to larger model dimension, more modeling time steps, and the heterogeneity of velocity and Q models. The preparation time for naABCs here is much longer than hABCs, due to the parameter optimizations for more unique pairs of (v_b, γ_b) in the heterogeneous case. Cost for low-rank decomposition in naABCs is reasonably larger than that in hABCs caused by the difference between their numbers of absorbing layers. The iterative modeling process when applying naABCs is more than 2x faster than applying hABCs. Furthermore, increasing number of shots and modeling time steps will make the computational cost for ABCs preparation and low-rank decomposition will become less considerable and eventually negligible. Hence, our modeling package using naABCs in heterogeneous case can achieve around 2x speedup over conventional hABCs.

4 Conclusion

We develop a python package for GPU-accelerated 3-D P-wave simulation in anelastic media based on the DFL viscoacoustic wave equation. The package utilizes two-step extrapolation with simplified time-stepping error compensation. The proposed naABCs

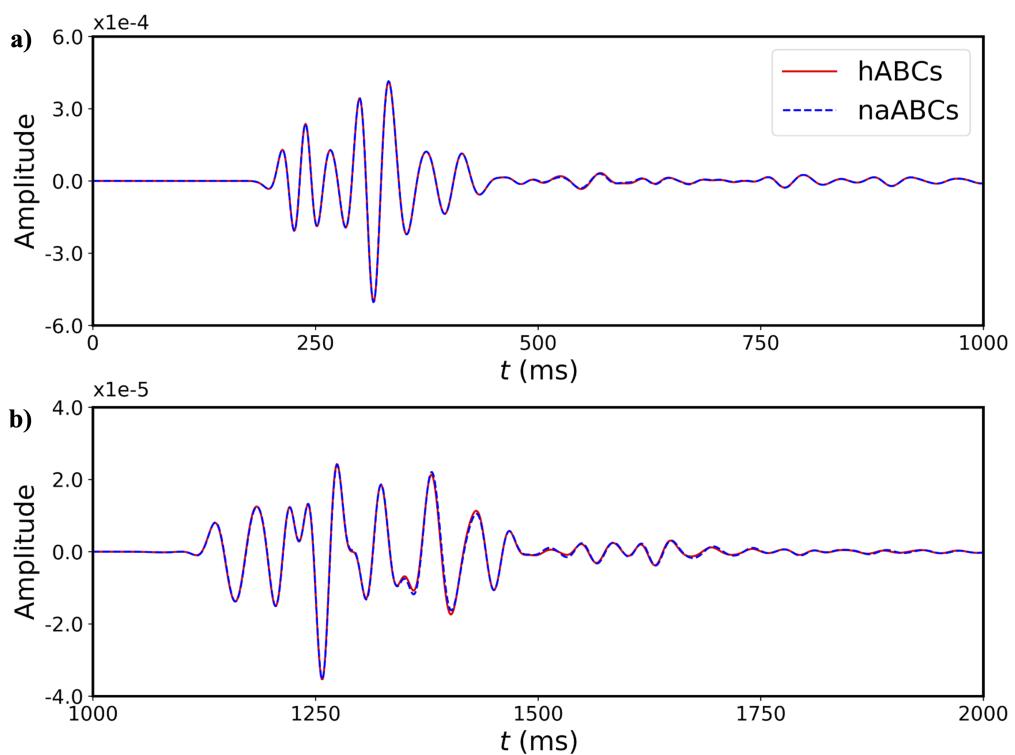


Figure 8. Modeled trace comparison using hABCs with $N = 20$ and naABCs with $N = 44$ ($\varepsilon = 0.05$). The two traces are recorded from receivers at (a) ($x = 4800$ m, $y = 2850$ m) and (b) ($x = 900$ m, $y = 1050$ m), respectively.

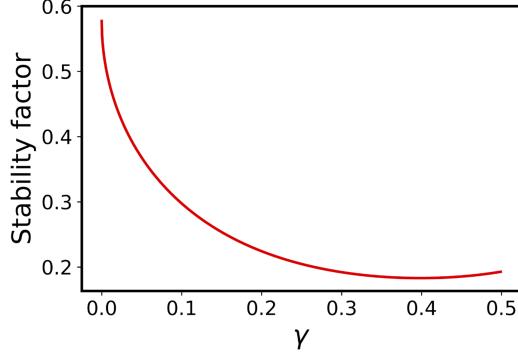


Figure A1. Stability factor varying with γ within $[0, 0.5]$ when $\omega_0 = 20 \times 2\pi$ and $h = 20$ m.

based on optimized absorbing layer parameters, naturally attenuate the boundary artifacts. Modeling examples demonstrate that the naABCs provide approximately 2x speedup over the conventional hABCs with satisfying simulation accuracy.

Appendix A Stability condition

Let $p^{-1} = p(\mathbf{x}, t - \Delta t)$, $p^0 = p(\mathbf{x}, t)$, and $p^1 = p(\mathbf{x}, t + \Delta t)$ represent the wave-fields at three consecutive time steps, an abbreviated matrix notation for the recursion equation 3 is

$$\begin{bmatrix} p^1 \\ p^0 \end{bmatrix} = \begin{bmatrix} g_1 & g_0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p^0 \\ p^{-1} \end{bmatrix}, \quad (\text{A1})$$

in which

$$g_0 = r/h \sin(\pi\gamma) \cos(\pi\gamma/2)^2 \omega_0^{-2\gamma} \mathbf{W}_l - 1, \quad (\text{A2})$$

$$g_1 = -r^2/h^2 \cos(\pi\gamma) \cos^2(\pi\gamma/2) \omega_0^{-2\gamma} \mathbf{W}_d - g_0 + 1, \quad (\text{A3})$$

where $r = v_0 \Delta t / h$ represents the Courant number.

According to the eigenvalue method (Gazdag, 1981), the recursion is stable only if the eigenvalues of the transition matrix in equation A1 are less or equal to 1. Hence, we have the stability condition as

$$-1 \leq g_0 \leq 1 - |g_1|. \quad (\text{A4})$$

Knowing ω_0 , h , and γ , we can search for the largest r satisfying inequation A4 with $kh = \sqrt{3}\pi$, i.e., $k_x h = k_y h = k_z h = \pi$ in 3-D case. This largest r is the stability factor $s(\gamma)|_{(\omega_0, h)}$ utilized in the inequation 18. Figure A1 shows the stability factor varying with γ when $\omega_0 = 20 \times 2\pi$ and $h = 20$ m. As γ increases, $s(\gamma)$ decreases, indicating harsher stability condition. Prior to the parameter optimization for naABCs, we search for stability factors of regularly sampled γ within $[0, 0.5]$. So during the optimization, the stability factor for any tested γ is readily obtained by nearest interpolation of the predetermined stability factor samples.

Appendix B Optimization for naABCs

The constrained optimization in 15 is a trans-dimensional inversion problem, which can be solved by Bayesian inversion based on reversible jump Markov-chain Monte Carlo (MCMC) sampling (Sambridge et al., 2006). However, for each unique (v_b, γ_b) pair, it requires to solve several fixed dimension inversion problems with different N , resulting

336 in tremendous computational cost in the heterogeneous case. Additionally, it is difficult
 337 to determine the viable lower bounds of N for different (v_b, γ_b) , since the sampled N could
 338 be too small to have a solution of (v_a, γ_a) satisfying all constraints in inequations 16, 17,
 339 and 18.

340 Noticing the three constraints require no information beyond the i th layer, we re-
 341 place the trans-dimensional inversion by a sequential optimization. At each step, it solves
 342 for the two model parameters of the current i th layer (v_a^i, γ_a^i) , until all constraints are
 343 satisfied with $R_i = 1$, indicating $i = N$ according to equation 12.

344 For the i th layer, model parameters of the inner layers are either initialized or al-
 345 ready solved. Now to achieve the minimum N , we should maximize the attenuation within
 346 the i th layer:

$$347 \arg \min_{(v_a^i, \gamma_a^i)} \max \left\{ \exp \left(-\frac{\pi}{2} \omega \Psi_i(\omega) \right) \mid \omega \in \omega_s^i \right\}, \quad (B1)$$

348 where $\omega_s^i = [\omega_0, \omega_1, \omega_2, \dots, \omega_I]^T$ represents the angular frequency samples that are not
 349 being absorbed yet. Initially, ω_s^0 is regularly sampled from the modeled frequency range.
 350 The number of initial samples is empirically set as 10 for a balance of accuracy and ef-
 351 ficiency.

352 As i increases, higher frequencies will satisfy inequation 16 with $R_i = 1$ first, re-
 353 resulting in a smaller set of unabsorbed frequency samples as follows:

$$354 \omega_s^i = \{ \omega \mid \alpha_{i-1}(\omega) > \epsilon_{i-1}, \omega \in \omega_s^{i-1} \}. \quad (B2)$$

355 in which

$$356 \alpha_i(\omega) = A_i(\omega) \mid_{R_i=1}. \quad (B3)$$

357 With fewer frequency samples in ω_s^i , the tolerance for the maximum returning wave am-
 358 plitude should be relaxed as follows:

$$359 \epsilon_i = \varepsilon / I, \quad (B4)$$

360 where ε represents the maximum tolerance for the outermost layer, I represents the num-
 361 ber of samples in ω_s^i . Hence, $\epsilon_0 = \varepsilon / 10$ and $\epsilon_N = \varepsilon$. Although ϵ_i can also be constant
 362 for simplicity, we find that this increasing ϵ_i generally provides satisfying absorption with
 363 smaller N .

364 Now we are solving a fixed dimensional constrained optimization with minimax ob-
 365 jective B1 and inequations constraints 16, 17, and 18, for the current i th layer. Such a
 366 problem can be solved efficiently by grid search combined with linear programming. Given
 367 a tentative γ_a^i , inequation 16 provides both a lower bound $v_l^{(1)}$ and an upper bound $v_u^{(1)}$
 368 for v_a^i as follows:

$$369 v_l^{(1)} = \max \left\{ \frac{1 - \epsilon_i / \alpha_{i-1}(\omega)}{1 + \epsilon_i / \alpha_{i-1}(\omega)} \frac{v_d^{i-1}(\omega)}{\beta_i(\omega)} \mid \omega \in \omega_s^i \right\} \quad (B5)$$

$$370 v_u^{(1)} = \min \left\{ \frac{1 + \epsilon_i / \alpha_{i-1}(\omega)}{1 - \epsilon_i / \alpha_{i-1}(\omega)} \frac{v_d^{i-1}(\omega)}{\beta_i(\omega)} \mid \omega \in \omega_s^i \right\}, \quad (B6)$$

372 in which

$$373 \beta_i(\omega) = \cos(\pi \gamma_a^i / 2) (\omega / \omega_0)^{\gamma_a^i}. \quad (B7)$$

375 Inequations 17 and 18 provide another lower bound $v_l^{(2)}$ and another upper bound $v_u^{(2)}$,
 376 respectively, for v_a^i as follows:

$$377 v_l^{(2)} = \max \left\{ \frac{\omega h}{\pi \beta_i(\omega)} \mid \omega \in \omega_s^i \right\}, \quad (B8)$$

$$378 v_u^{(2)} = \frac{s(\gamma_a^i) h}{\Delta t}. \quad (B9)$$

380 For this γ_a^i , the smallest v_a^i satisfying all bounds in equations B5, B6, B8, and B9 will
 381 provide the maximum attenuation. Thus, as the outer loop being the grid search for γ_a^i ,
 382 and inner loop being the linear programming for v_a^i , we can solve for the optimal (v_a^i, γ_a^i)
 383 for the minimax objective B1. Note that there always be at least one viable solution of
 384 (v_a^i, γ_a^i) for this constrained optimization, that is $(v_a^{i-1}, \gamma_a^{i-1})$.

385 As i increasing, the remaining unabsorbed frequency samples become fewer. Eventually,
 386 all frequency samples will satisfy the constraints with $R_i = 1$, causing $I = n(\omega_s^i) =$
 387 0. Then, we stop the sequential optimization and obtain the minimum $N = i$ with cor-
 388 responding (v_a, γ_a) . We summarize the workflow of parameter optimization for naABCs
 389 in Algorithm 1.

390 **Acknowledgments**

391 The authors acknowledge the Singapore Economic Development Board for its financial
 392 support through the Petroleum Engineering Professorship. The python package and all
 393 supplementary materials related to this article can be found online <https://github.com/nusbei/GPU-accelerated-DFL-viscoacoustic-wave-equation-modeling>.

Algorithm 1 Parameter optimization for naABCs**Input:**

Modeling parameters: $\omega_0, \Delta t, h$
 Model boundary parameters: v_b, γ_b
 Optimization parameters: ω_s^0, ε
 Grid search radius for γ_a : R_γ
 Grid search number for γ_a : M

Output:

Minimum number of absorbing layers: N
 Parameters for N absorbing layers: v_a, γ_a

Initialization:

$v_a^0 \leftarrow v_b, \gamma_a^0 \leftarrow \gamma_b$
 Compute $s(\gamma), \gamma \in [0, 0.5]$
 Compute $v_d^0(\omega), \Psi_0(\omega), \alpha_0(\omega), \omega \in \omega_s^0$
 $I \leftarrow n(\omega_s^0)$
 $\epsilon_0 \leftarrow \varepsilon/I$
 $i \leftarrow 0$

Iteration:

```

while  $I > 0$  do
     $i++$ 
    Update  $\omega_s^i, I, \epsilon_i$ 
     $\gamma_l \leftarrow \max(\gamma_a^{i-1} - R_\gamma, 0)$ 
     $\gamma_u \leftarrow \min(\gamma_a^{i-1} + R_\gamma, 0.5)$ 
     $\Delta\gamma \leftarrow (\gamma_u - \gamma_l)/M$ 
     $J_{min} \leftarrow \infty$ 
    for  $j = 0, 1, \dots, M$  do
         $\gamma \leftarrow \gamma_l + j\Delta\gamma$ 
        Compute  $\beta_i(\omega), \omega \in \omega_s^i$ 
        Compute  $v_l^{(1)}, v_u^{(1)}, v_l^{(2)}, v_u^{(2)}$ 
         $v_l \leftarrow \max(v_l^{(1)}, v_l^{(2)})$ 
         $v_u \leftarrow \min(v_u^{(1)}, v_u^{(2)})$ 
        if  $v_l \leq v_u$  then
             $v \leftarrow v_l$ 
            Compute  $v_d^i(\omega), \omega \in \omega_s^i$ 
            Compute  $\Psi_i(\omega), \omega \in \omega_s^i$ 
             $J \leftarrow \max\{\exp(-\frac{\pi}{2}\omega\Psi_i(\omega)) | \omega \in \omega_s^i\}$ 
            if  $J < J_{min}$  then
                 $J_{min} \leftarrow J$ 
                 $\gamma_a^i \leftarrow \gamma, v_a^i \leftarrow v$ 
            end if
        end if
    end for
    Compute  $R_{i-1}(\omega), \omega \in \omega_s^i$ 
    Compute  $\alpha_i(\omega)$ 
end while
 $N \leftarrow i$ 

```

395 **References**

- 396 Aminzadeh, F., Jean, B., & Kunz, T. (1997). *3-D salt and overthrust models*. Tulsa:
 397 Society of Exploration Geophysicists.
- 398 Berenger, J.-P. (1994). A perfectly matched layer for the absorption of electromagnetic
 399 waves. *Journal of Computational Physics*, *114*(2), 185–200. doi: 10.1006/jcph.1994.1159
- 400 Carcione, J. M., Cavallini, F., Mainardi, F., & Hanyga, A. (2002). Time-domain
 401 modeling of constant-*Q* seismic waves using fractional derivatives. *Pure and
 402 Applied Geophysics*, *159*(7-8), 1719–1736. doi: 10.1007/s00024-002-8705-z
- 403 Cavalca, M., Moore, I., Zhang, L., Ng, S. L., Fletcher, R., & Bayly, M. (2011). Ray-
 404 based tomography for *Q* estimation and *Q* compensation in complex media. In
 405 *SEG technical program expanded abstracts 2011* (p. 3989-3993). doi: 10.1190/1
 406 .3628039
- 407 Chen, H., Zhou, H., Jiang, S., & Rao, Y. (2019). Fractional Laplacian viscoacoustic
 408 wave equation low-rank temporal extrapolation. *IEEE Access*, *7*, 93187–93197.
 409 doi: 10.1109/access.2019.2927760
- 410 Chen, H., Zhou, H., Li, Q., & Wang, Y. (2016). Two efficient modeling schemes
 411 for fractional Laplacian viscoacoustic wave equation. *Geophysics*, *81*(5), T233–
 412 T249. doi: 10.1190/geo2015-0660.1
- 413 Chen, H., Zhou, H., & Qu, S. (2014). Lowrank approximation for time domain vis-
 414 coacoustic wave equation with spatially varying order fractional Laplacians. In
 415 *Seg technical program expanded abstracts 2014* (p. 3400-3405). doi: 10.1190/
 416 segam2014-0055.1
- 417 Dutta, G., & Schuster, G. T. (2014). Attenuation compensation for least-squares re-
 418 verse time migration using the viscoacoustic-wave equation. *Geophysics*, *79*(6),
 419 S251-S262. doi: 10.1190/geo2013-0414.1
- 420 Fomel, S., Ying, L., & Song, X. (2013). Seismic wave extrapolation using lowrank
 421 symbol approximation. *Geophysical Prospecting*, *61*(3), 526–536. doi: 10.1111/
 422 j.1365-2478.2012.01064.x
- 423 Gazdag, J. (1981). Modeling of the acoustic wave equation with transform methods.
 424 *Geophysics*, *46*(6), 854–859. doi: 10.1190/1.1441223
- 425 Kjartansson, E. (1979). Constant *Q*-wave propagation and attenuation. *Journal of
 426 Geophysical Research: Solid Earth*, *84*(B9), 4737–4748. doi:
 427 10.1029/JB084iB09p04737
- 428 Li, Q. Z. (1993). *High-resolution seismic exploration*. China: Petroleum Industry
 429 Press.
- 430 Liu, Y., & Sen, M. K. (2010). A hybrid scheme for absorbing edge reflections in nu-
 431 mercial modeling of wave propagation. *Geophysics*, *75*(2), A1–A6. doi: 10
 432 .1190/1.3295447
- 433 Liu, Y., & Sen, M. K. (2018). An improved hybrid absorbing boundary condition for
 434 wave equation modeling. *Journal of Geophysics and Engineering*, *15*(6), 2602–
 435 2613. doi: 10.1088/1742-2140/aadd31
- 436 McDonal, F., Angona, F., Mills, R., Sengbush, R., Van Nostrand, R., & White, J.
 437 (1958). Attenuation of shear and compressional waves in pierre shale. *Geo-
 438 physics*, *23*(3), 421–439. doi: 10.1190/1.1438489
- 439 Ren, Z., & Liu, Y. (2013). A hybrid absorbing boundary condition for frequency-
 440 domain finite-difference modelling. *Journal of Geophysics and Engineering*,
 441 *10*(5), 054003. doi: 10.1088/1742-2132/10/5/054003
- 442 Sambridge, M., Gallagher, K., Jackson, A., & Rickwood, P. (2006). Trans-
 443 dimensional inverse problems, model comparison and the evidence. *Geophysical
 444 Journal International*, *167*(2), 528–542. doi: 10.1111/j.1365-246x.2006.03155
 445 .x
- 446 Sanders, C., Ponko, S., Nixon, L., & Schwartz, E. (1995). Seismological evidence for
 447 magmatic and hydrothermal structure in long valley caldera from local earth-
 448 -

- quake attenuation and velocity tomography. *Journal of Geophysical Research: Solid Earth*, 100(B5), 8311–8326. doi: 10.1029/95jb00152
- Sun, J., Fomel, S., Zhu, T., & Hu, J. (2016). *Q*-compensated least-squares reverse time migration using low-rank one-step wave extrapolation. *Geophysics*, 81(4), S271–S279. doi: 10.1190/geo2015-0520.1
- Wang, N., Zhu, T., Zhou, H., Chen, H., Zhao, X., & Tian, Y. (2020). Fractional Laplacians viscoacoustic wavefield modeling with k -space-based time-stepping error compensating scheme. *Geophysics*, 85(1), T1-T13. doi: 10.1190/geo2019-0151.1
- Wang, Y., Zhou, H., Zhao, X., Zhang, Q., Zhao, P., Yu, X., & Chen, Y. (2019). Cu Q -RTM: A CUDA-based code package for stable and efficient Q -compensated reverse time migration. *Geophysics*, 84(1), F1–F15. doi: 10.1190/geo2017-0624.1
- Xie, C., Song, P., Tan, J., Liu, B., Li, J., Yu, K., ... Zhang, R. (2020). Cosine-type weighted hybrid absorbing boundary based on the second-order higdon boundary condition and its gpu implementation. *Journal of Geophysics and Engineering*, 17(2), 231–248. doi: 10.1093/jge/gxz102
- Xu, Y., Li, J., Pang, G., Wang, Z., Chen, X., & Wang, B. (2018). A new radial basis function collocation method based on the quasi-uniform nodes for 2D fractional wave equation. *arXiv preprint arXiv:1806.01846*.
- Xue, Z., Sun, J., Fomel, S., & Zhu, T. (2018). Accelerating full-waveform inversion with attenuation compensation. *Geophysics*, 83(1), A13-A20. doi: 10.1190/geo2017-0469.1
- Yao, J., Zhu, T., Hussain, F., & Kouri, D. J. (2017). Locally solving fractional Laplacian viscoacoustic wave equation using Hermite distributed approximating functional method. *Geophysics*, 82(2), T59–T67. doi: 10.1190/geo2016-0269.1
- Zhang, J., Wu, J., & Li, X. (2013). Compensation for absorption and dispersion in prestack migration: An effective Q approach. *Geophysics*, 78(1), S1–S14. doi: 10.1190/geo2012-0128.1
- Zhu, T., Ajo-Franklin, J. B., & Daley, T. M. (2017). Spatiotemporal changes of seismic attenuation caused by injected CO₂ at the Frio-II pilot site, Dayton, TX, USA. *Journal of Geophysical Research: Solid Earth*, 122, 7156-7171. doi: 10.1002/2017jb014164
- Zhu, T., & Harris, J. M. (2014). Modeling acoustic wave propagation in heterogeneous attenuating media using decoupled fractional Laplacians. *Geophysics*, 79(3), T105–T116. doi: 10.1190/geo2013-0245.1
- Zhu, T., Harris, J. M., & Biondi, B. (2014). Q -compensated reverse-time migration. *Geophysics*, 79(3), S77-S87. doi: 10.1190/geo2013-0344.1