Low-rank one-step wave extrapolation for reverse time migration

Junzhe Sun¹, Sergey Fomel¹, and Lexing Ying²

ABSTRACT

Reverse time migration (RTM) relies on accurate wave extrapolation engines to image complex subsurface structures. To construct such operators with high efficiency and numerical stability, we have developed a one-step wave extrapolation approach using complex-valued low-rank decomposition to approximate the mixed-domain space-wavenumber wave extrapolation symbol. The low-rank one-step method involves a complex-valued phase function, which is more flexible than a real-valued phase function of two-step schemes, and thus it is capable of modeling a wider variety of dispersion relations. Two novel designs of the phase function leads to the desired properties in wave extrapolation. First, for wave propagation in inhomogeneous media, including a velocity gradient term assures a more accurate phase behavior, particularly when the velocity variations are large. Second, an absorbing boundary condition, which is propagation-direction-dependent, can be incorporated into the phase function as an anisotropic attenuation term. This term allows waves to travel parallel to the boundary without absorption, thus reducing artificial reflections at wide incident angles. Using numerical experiments, we revealed the stability improvement of a one-step scheme in comparison with two-step schemes. We observed the low-rank one-step operator to be remarkably stable and capable of propagating waves using large time step sizes, even beyond the Nyquist limit. The stability property can help to minimize the computational cost of seismic modeling or RTM. The low-rank one-step wave extrapolation also handles anisotropic wave propagation accurately and efficiently. When applied to RTM in anisotropic media, the proposed method generated high-quality images.

INTRODUCTION

The task of wave extrapolation is propagation of waves in depth or time, which finds applications in seismic modeling and migration. Conventionally, it is implemented by finite differences (FD). FD methods have a low-computational cost but suffer from dispersion artifacts and instabilities (Kosloff and Baysal, 1982). Instead, numerical differentiation of space coordinates in wave equations can be implemented by using the Fourier transform, which is known as the pseudospectral method (Fornberg, 1998). Pseudospectral methods have higher accuracy and can suppress numerical dispersion (Reshef et al., 1988; Virieux et al., 2011). On the other hand, they are limited to small time steps and may be subject to dispersion because of FD approximations used for the time derivatives. Recently, alternative strategies have been proposed to propagate waves by mixed-domain space-wavenumber operators (Soubaras and Zhang, 2008; Wards et al., 2008; Eigen and Brandsberg-Dahl, 2009; Liu et al., 2009; Zhang and Zhang, 2009; Du et al., 2010; Pestana and Stoffa, 2010; Chu and Stoffa, 2011; Fomel et al., 2013; Wu and Alkhalifah, 2014). Fowler (2010b) and Du et al. (2014) refer to these methods as recursive integral time extrapolation (RITE) methods.

RITE methods are designed to make time extrapolation stable and dispersion free in heterogeneous media for large time steps, even beyond the Nyquist limit (Du et al., 2014), and they therefore are particularly suitable for reverse time migration (RTM), a depth-migration method in which waves are extrapolated in time (Baysal...
THEORY

Let $p(x, t)$ be the seismic wavefield at location $x$ and time $t$. The acoustic constant-density wave equation can be written as

$$
\left( \frac{\partial^2}{\partial t^2} - V^2(x) \nabla^2 \right) p(x, t) = 0,
$$

where $V(x)$ is the velocity and $\nabla^2$ denotes the Laplacian operator.

Analytical solutions in constant velocity

When $V$ is constant, after Fourier transform in space, the wave equation takes the form

$$
\left( \frac{\partial^2}{\partial t^2} + V^2 |k|^2 \right) P(k, t) = 0,
$$

where $k$ is the spatial wavenumber and $P(k, t)$ is the spatial Fourier transform of $p(x, t)$:

$$
P(k, t) = \frac{1}{(2\pi)^3} \int p(x, t) e^{-ik \cdot x} dx.
$$

The analytical solution to equation 2 can be expressed as

$$
P = A_1 e^{i|k|vt} + A_2 e^{-i|k|vt} = P_1 + P_2,
$$

where $P_1$ represents the forward-propagating wavefield, i.e., positive frequencies, and $P_2$ represents the backward-propagating wavefield, i.e., negative frequencies. The time derivative of $P$ has the following form:

$$
\frac{\partial P}{\partial t} = i|k|V(P_1 - P_2).
$$

Zhang and Zhang (2009) use the Hilbert transform to define an additional function:

$$
Q(k, t) = \frac{1}{\psi} \frac{\partial P(k, t)}{\partial t},
$$

where $Q(k, t)$ is the Hilbert transform of $P(k, t)$, and $\psi = V|k|$. Combining equations 4–6, $P_1$ and $P_2$ can be expressed as

$$
P_1 = \frac{1}{2}(P - iQ)
$$

and

$$
P_2 = \frac{1}{2}(P + iQ).
$$

Equation 2 can be split into a pair of first-order equations and expressed in the following matrix form:

$$
\frac{\partial}{\partial t} \begin{bmatrix} P \\ P_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\psi^2 & 0 \end{bmatrix} \begin{bmatrix} P \\ P_1 \end{bmatrix}.
$$
With the help of the Hilbert transform, a more symmetric expression can be achieved as

\[
\frac{\partial}{\partial t} \begin{bmatrix} P \\ iQ \end{bmatrix} = \begin{bmatrix} 0 & -i\psi \\ -i\psi & 0 \end{bmatrix} \begin{bmatrix} P \\ iQ \end{bmatrix}. \tag{10}
\]

We can further decompose the first matrix on the right side as follows:

\[
\begin{bmatrix} 0 & -i\psi \\ -i\psi & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} i\psi & 0 \\ 0 & -i\psi \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}. \tag{11}
\]

Substituting equation 11 into equation 10, and using equations 7 and 8, we arrive at

\[
\frac{\partial}{\partial t} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} i\psi & 0 \\ 0 & -i\psi \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}. \tag{12}
\]

In RTM, only one branch of the total wavefield is needed at one time. The two parts of wave propagation decouple according to

\[
\frac{\partial}{\partial t} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \begin{bmatrix} i\psi & 0 \\ 0 & -i\psi \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}. \tag{13}
\]

Modeling seismic wave propagation requires the source function. Letting the source function be \( f(x,t) \), wave equation 2 can be rewritten in the following form:

\[
\left( \frac{\partial^2}{\partial t^2} + \psi^2 \right) P(x,t) = f(x,t). \tag{14}
\]

Correspondingly, equation 13 becomes

\[
\begin{align*}
\frac{\partial}{\partial t} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} &= \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} \begin{bmatrix} 0 & -i\psi \\ -i\psi & 0 \end{bmatrix} \begin{bmatrix} P \\ iQ \end{bmatrix} + \begin{bmatrix} 0 \\ -i\psi \end{bmatrix} f \\
&= \begin{bmatrix} i\psi & 0 \\ 0 & -i\psi \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} + \begin{bmatrix} -i/f \psi \\ i/f \psi \end{bmatrix} f. \tag{15}
\end{align*}
\]

The application of operator \(-i/2\psi\) can be implemented in either time domain or Fourier domain; it can also be directly incorporated into the definition of the source functions. For example, operator \(1/\psi\) can be regarded as \((io\omega/|\omega|) \cdot (1/i\omega)\), which in the time domain corresponds to cascading the Hilbert transform with the first-order integration.

In constant velocity, the forward-propagating wavefield away from the source at the next time step \(t + \Delta t\) can be expressed as

\[
p_1(x,t + \Delta t) = \int P_1(k,t) e^{i[k \cdot x + V\cdot \psi \Delta t]} dk. \tag{16}
\]

Variable velocity and anisotropy

For inhomogeneous and anisotropic media, we can use the general form of equation 16 to introduce a general phase function, which depends on \( k \) and \( x \). We rewrite equation 16 as the following mixed-domain operator (Wards et al., 2008; Fomel et al., 2013):

\[
p_1(x,t + \Delta t) = \int P_1(k,t) e^{i\phi(x,k,\Delta t)} dk. \tag{17}
\]

Equation 17 is kinematically correct if the phase function \( \phi(x,k,t) \) satisfies the anisotropic eikonal equation:

\[
\frac{\partial \phi}{\partial t} = \pm V(x,k) |\nabla \phi|, \tag{18}
\]

where \( V(x,k) \) is the phase velocity. When the velocity is constant and isotropic, the phase function reduces to

\[
\phi(x,k,t) = k \cdot x \pm |V| \Delta t, \tag{19}
\]

and it corresponds to equation 16. In the more general case, assuming small time steps, \( \phi \) can be expanded into the Taylor series in \( t \) (Fomel et al., 2013):

\[
\phi(x,k,t) \approx k \cdot x + \phi_1(x,k)t + \phi_2(x,k) \frac{t^2}{2} + \cdots, \tag{20}
\]

where

\[
\phi_1(x,k) = V(x,k)|k|, \tag{21}
\]

and

\[
\phi_2(x,k) = V(x,k)\nabla V \cdot k. \tag{22}
\]

Equation 17 corresponds to the one-step method (Zhang and Zhang, 2009; Fowler, 2010b). Fomel et al. (2013) adopt instead a two-step implementation, which uses only the \( \phi_1 \) term in equation 20 to cancel out the imaginary part of the wave extrapolation operator (Etgen and Brandsberg-Dahl, 2009):

\[
p(x,t + \Delta t) + p(x,t - \Delta t) \approx 2 \int P(k,t) e^{ikx} \cos[V(x,k)|k|\Delta t] dk. \tag{23}
\]

It is important to realize that the cancellation of the imaginary part relies on the fact that the phase function only contains odd-ordered terms of \( t \), with the exception of the zeroth-order term, which corresponds to the inverse Fourier transform. Any modification to the phase function that violates this requirement cannot be easily handled by the two-step method.

As detailed in Appendix A, the unconditional numerical stability of the one-step scheme (equation 17) can be proven theoretically in the continuous case. In practice, we have observed that the low-rank approximation preserves the stability property, as will be demonstrated in the next section with numerical examples.

Higher order terms from the phase function

In this paper, we adopt the one-step scheme due to its superior stability and ability to handle complex-valued phase functions. Be-
cause the two-step scheme depends on a real-valued phase function, it cannot include higher order terms from the expansion in equation 20 and thus implement more accurate phase functions. By switching to a one-step scheme, we can easily incorporate the second-order term ($\phi_2$) in equation 20 to achieve a more accurate wave extrapolation operator. As defined in equation 22, $\phi_2$ involves the gradient of the velocity model, and it can become significant when either the time step size is large or the velocity model changes rapidly. Substituting the first three terms from the Taylor series (equation 20) into equation 17, the second-order operator takes the form

$$p_1(x, t + \Delta t) \approx P_1(k, t) e^{i(k \cdot (x + V(k) t \Delta t) + V(k) V V k \Delta t^2/2)} \cdot dk. \quad (24)$$

This modification helps to increase the accuracy of wave extrapolation, especially when $\Delta t$ or $VV$ is large. Note that introduction of the velocity gradient term does require the velocity model to be smoothly varying, which conforms to the usual requirement of RTM. The term $VV + k$ also leads to a certain degree of anisotropy in the phase function, which in practice may increase the numerical rank of low-rank approximation.

**Direction-dependent absorbing boundary conditions**

We propose another innovation using the one-step scheme by incorporating propagation-direction-dependent absorbing boundary conditions into the extrapolation operator. In absorbing boundary layers, we propose to modify the operator

$$W(x, k) = e^{i[\phi(x, k, \Delta t) - k \cdot x]} \quad (25)$$

into the following form:

$$W(x, k)_{abc} = e^{i[\phi(x, k, \Delta t) - k \cdot x] - \frac{\alpha(x \cdot k)}{k |k|}}, \quad (26)$$

where $\alpha$ is the decay parameter and $x_0$ is the location of the nearest absorbing boundary. The added term $e^{-i[x \cdot (x_0 \cdot k)/|k|]}$ causes exponential decay of the wavefield between the absorbing boundary and the computational boundary (domain truncation), which also depends on the wave-traveling direction. When the wave travels normal to the boundary, it will have the maximum absorbing effect because $x - x_0$ is in the direction of $k$; on the other hand, there will be no damping when the wave is traveling parallel to the boundary because the angle between $x - x_0$ and $k$ would be zero. Allowing waves to propagate within the absorbing zone can mitigate artificial reflections from the absorbing boundary, especially at grazing incident angles. The decay term should be applied on both the real wavefield and its first-order time derivative (the imaginary part of the analytical wavefield). This is analogous to the tapering technique described by Cerjan et al. (1985). The absorbing term in the phase function can be physically interpreted as an anisotropic attenuation effect.

The proposed direction-dependent absorbing boundary conditions could also be incorporated into two-step-based wave extrapolation methods (Fomel et al., 2013; Song et al., 2013; Fang et al., 2014) using low-rank approximation, by separately applying the damping operator on the wavefield and its time derivative. However, because two-step schemes do not allow direct incorporation of the absorbing boundary condition into the wave propagation matrix, additional Fourier transforms will be required. The one-step implementation is more computationally efficient and straightforward.

**Low-rank approximation**

To implement wave propagation numerically, we use the low-rank approximation method of Fomel et al. (2013) to decompose the wave extrapolation matrix 26 into the following separated representation:

$$W(x, k) \approx \sum_{m=1}^{M} \sum_{n=1}^{N} W(x, k_m) a_{mn} W(x_n, k). \quad (27)$$

The difference is that now the low-rank decomposition is implemented for complex matrices or linear operators instead of real ones. The computation of $p(x, t + \Delta t)$ then becomes

$$p(x, t + \Delta t) \approx \sum_{m=1}^{M} W(x, k_m) \left\{ \sum_{n=1}^{N} a_{mn} \left[ \int e^{i(k \cdot x_n)} W(x_n, k) P(k, t) dk \right] \right\}. \quad (28)$$

The computational cost of representation 28 is effectively equivalent to applying $N$ inverse fast Fourier transforms (FFTs) per time step. In practice, $N$ is a small number, typically less than five for isotropic media. The value of $N$ may grow with increasing model complexity, such as the introduction of anisotropy. Compared with a naive straightforward implementation of equation 17, the number of floating point operations per time step is reduced from $O(N^2)$ to $O(N N_c \log N_c)$, where $N_c$ is the total size of the spatial grid.

**Reverse time migration imaging conditions**

Because the one-step wave extrapolation kernel operates in the complex domain, it requires a definition of data and reflectivity with complex values. The analytical data follow the definition of the complex wavefield in equations 7 and 9. This implies that the input data need to be Hilbert transformed along the time axis and supplied as the imaginary part before the migration process, creating an analytical signal (Tiner et al., 1980). We adopt the following complex-valued cross-correlation imaging condition (Claerbout, 1985):

$$I_c(x) = \sum_s \sum_t \bar{S}_s(x, t) R_t(x, t), \quad (29)$$

where the lowercase $s$ denotes shots and $t$ denotes time samples. The real part of the complex image $I_c(x)$ is extracted and used as the final image.

Extended imaging conditions (Sava and Vasconcelos, 2011), including space-shift (Rickett and Sava, 2002; Sava and Fomel, 2003) and time-shift (Sava and Fomel, 2006) imaging conditions, can provide additional information for migration velocity analysis. The complex-valued space-shift and time-shift imaging condition for low-rank one-step RTM takes the form
\[ I_\nu(x, \lambda, \tau) = \sum_j \sum_t \tilde{S}_s(x - \lambda, t - \tau) R_s(x + \lambda, t + \tau) \, \nu \quad (30) \]

and it can be easily implemented in the time-space domain.

**EXAMPLES**

In this section, we use several numerical examples to demonstrate the properties of low-rank one-step wave extrapolation.

**Complex-valued low-rank approximation**

We first test the accuracy of low-rank approximation applied to a wave extrapolation matrix in a 1D inhomogeneous medium. Using a simple velocity profile with a sharp velocity contrast (Figure 1), and a time step size of 0.01 s, the real and imaginary parts of the wave extrapolation matrix defined by equation 25 with only the \( \phi_1 \) term are plotted in Figure 2a and 2b, respectively. An accuracy threshold of \( \epsilon = 10^{-4} \) leads to an approximation rank \( N = 4 \). The approximation error is plotted in Figure 3a and 3b, with the maximum error corresponding to the prescribed accuracy requirement. To see that the accuracy threshold is strict enough to guarantee kinematic accuracy, we first use an exact matrix multiplication to calculate the exact wavefield from an initial condition (Figure 4a). Next, we use low-rank wave extrapolation to compute the wavefield and calculate their difference (Figure 4b). A negligible error can be observed from the difference section, indicating the high accuracy of low-rank wave extrapolation.

---

**Figure 1.** A 1D velocity profile with a sharp interface.

**Figure 2.** (a) The real part of the wave extrapolation matrix and (b) the imaginary part of the wave extrapolation matrix.

**Figure 3.** (a) The real part of the approximation error and (b) the imaginary part of the approximation error.

---

Downloaded from https://pubs.geoscienceworld.org/geophysics/article-pdf/81/1/S39/3247451/geo2015-0183.pdf by guest
Two-layer model

We use a simple two-layer velocity model similar to the one used by Du et al. (2014) to demonstrate the stability of one-step wave extrapolation using low-rank approximation. Figure 5 shows the comparison among the stability of low-rank one-step, low-rank two-step, and fourth-order FD methods. The velocity model has a sharp contrast at the depth of 3795 m: The upper layer has a velocity of 1500 m/s, and the lower layer has a velocity of 4500 m/s (Figure 5a). The model is discretized on a 400 × 400 grid with a spacing of 15 m along the horizontal and vertical directions. An explosive source, with a Ricker wavelet using a peak frequency of 16 Hz (the maximum frequency is approximately 50 Hz), is injected in the center of the model. When a time step of 2 ms is used, the classic fourth-order FD method suffers from visible dispersion artifacts (Figure 5b), whereas the one- and two-step schemes produce waves free of artifacts (Figure 5c and 5d). When a time step of 3 ms is used, the one-step scheme is stable (Figure 5e), whereas the two-step scheme starts to develop artifacts near the velocity contrast (Figure 5f). The FD method is no longer stable, and therefore it is not plotted. At 10 ms, which corresponds to the Nyquist sampling (Figure 5f). The model is discretized on a 512 × 512 grid with spacing of 5 m along the horizontal and vertical directions. The source is injected in the center of the model. Figure 6b shows the reference wavefield propagated using a time step of 3.5 ms. The model is discretized on a 512 × 512 grid with spacing of 5 m along the horizontal and vertical directions. The source is injected in the center of the model. Figure 6b shows the reference wavefield propagated using only the \( \phi_1 \) term from equation 20, but with an exceedingly small time step (0.175 ms), so that the \( \phi_2 \) term is negligible and the wavefield can be treated as accurate. Figure 6c shows the difference between the reference wavefield and the wavefield propagated by equation 24 after normalization. Figure 6d shows the difference between the reference wavefield and the wavefield propagated without including the \( \phi_2 \) term in the phase function, also after normalization. After including the velocity-gradient term, the error decreases significantly, and it is caused mainly by small amplitude differences. On the other hand, the difference calculated without using the velocity-gradient term is due mostly to the shift in phase. This observation is further supported by Figure 7, which shows a comparison between traces extracted from the 2D wavefield snapshots at \( X = 2000 \) m. The trace calculated using the velocity-gradient term aligns with the reference trace, whereas the trace calculated without including the velocity-gradient term has a noticeable shift in space relative to the reference trace. The shift increases with the degree of velocity variation.

Next, we perform a similar experiment using a more complicated Marmousi velocity model (Figure 8a). The model is smoothed and discretized on a 376 × 576 grid with spacing of 25 m. The reference wavefield is propagated using a time step size of 1.5 ms (Figure 8b). Figure 8c and 8d demonstrates the difference between the reference wavefield and the wavefields calculated using phase functions with and without the velocity-gradient term, both using a time step size of 15 ms. Figure 9 overlays the trace from the reference wavefield extracted at \( Z = 8800 \) m with corresponding traces from the wavefields propagated using the larger time step size. The comparison shows that, even with moderate velocity variations but a relatively large time step size, the velocity gradient term can have a noticeable contribution to the phase accuracy.

In the remaining examples of this paper, the velocity gradient term was not included in the phase function for simplicity.

Improved phase accuracy

To test if a more accurate wavefield can be obtained by one-step extrapolation with equation 24, we first use a synthetic model with a smooth velocity distribution and large velocity variations (Figure 6a). We propagate a wavefield with the Ricker–wavelet source, using a time step of 3.5 ms. The model is discretized on a 512 × 512 grid with spacing of 5 m along the horizontal and vertical directions. The source is injected in the center of the model. Figure 6b shows the reference wavefield propagated using only the \( \phi_1 \) term from equation 20, but with an exceedingly small time step (0.175 ms), so that the \( \phi_2 \) term is negligible and the wavefield can be treated as accurate. Figure 6c shows the difference between the reference wavefield and the wavefield propagated by equation 24 after normalization. Figure 6d shows the difference between the reference wavefield and the wavefield propagated without including the \( \phi_2 \) term in the phase function, also after normalization. After including the velocity-gradient term, the error decreases significantly, and it is caused mainly by small amplitude differences. On the other hand, the difference calculated without using the velocity-gradient term is due mostly to the shift in phase. This observation is further supported by Figure 7, which shows a comparison between traces extracted from the 2D wavefield snapshots at \( X = 2000 \) m. The trace calculated using the velocity-gradient term aligns with the reference trace, whereas the trace calculated without including the velocity-gradient term has a noticeable shift in space relative to the reference trace. The shift increases with the degree of velocity variation.

In the next example, we incorporate the propagation-direction-dependent absorbing boundary condition in the wave extrapolation operator to attenuate waves reaching the boundary. The absorbing boundary defined by equation 26 does not attenuate any energy if the wave propagates parallel to the boundary, and it reaches maximum damping when the wave propagates normal to the boundary. This property reduces artificial reflected energy at large incident angles. Figure 10 compares the propagation-direction-dependent absorbing boundary condition with a conventional direction-independent absorbing boundary condition (tapering) using a point-
source wavefield. After normalization and clipping for display, the direction-dependent absorbing boundary condition proves to be more effective at wide incident angles. In isotropic wave propagation, we observed that the direction-dependent absorbing boundary condition may increase the rank of the low-rank approximation because of introduced anisotropic attenuation of wave propagation. One possible way to decrease the cost is to implement the absorbing boundary condition separately from the wave extrapolation process, using partial FFT (Ying and Fomel, 2009). In this way, the cost of absorbing boundaries can be marginal compared with the cost of wave extrapolation.

Wave propagation in tilted transverse isotropic media

To show that the proposed method handles anisotropic media accurately, we propagate qP- and qSV-waves in a portion of the

![Figure 5](https://pubs.geoscienceworld.org/geophysics/article-pdf/81/1/S39/3247451/geo2015-0183.pdf)
2007 anisotropic benchmark model from BP. The model has horizontal spacing of 50 m and vertical spacing of 25 m. As shown in Figure 11, the model exhibits strong tilted transverse isotropy (TTI). We restrict the modeled area to between 34 and 60 km. Because the original model does not include S-wave velocity, we compute a vertical S-wave velocity profile as 0.3 times the original vertical P-wave velocity.

To calculate the qP- and qSV-wave phase velocities, we use the exact formulas defined by (Gassmann, 1964):

Figure 6. Wave propagation in a medium with large velocity variations. (a) Velocity model. (b) Reference wavefield propagated using Δt = 0.175 ms. (c) Difference between reference and wavefield propagated with the velocity-gradient term using Δt = 3.5 ms. (d) Difference between reference and wavefield propagated without the velocity-gradient term using Δt = 3.5 ms.

Figure 7. A slice through wavefields in Figure 6 at X = 2000 m. The black, short-dashed line in both (a) and (b) represents the reference wavefield; the blue, solid line in (a) represents the result calculated with velocity gradient; and the red, long-dashed line in (b) represents the result calculated without the velocity gradient.
\[
V_P^2(\mathbf{n}, \mathbf{x}) |\mathbf{k}|^2 = \frac{1}{2} [(c_{11} + c_{55}) \hat{k}_x + (c_{33} + c_{55}) \hat{k}_z] \\
+ \frac{1}{2} \sqrt{[(c_{11} - c_{55}) \hat{k}_x - (c_{33} - c_{55}) \hat{k}_z]^2 + 4(c_{13} + c_{55})^2 \hat{k}_x \hat{k}_z},
\]

(31)

where \(c_{11}, c_{13}, c_{33}, \) and \(c_{55}\) are the density-normalized components of the elastic tensor, \(\mathbf{k}\) is the wavenumber vector, \(\hat{k}_x\) and \(\hat{k}_z\) are the

\[
V_{SV}^2(\mathbf{n}, \mathbf{x}) |\mathbf{k}|^2 = \frac{1}{2} [(c_{11} + c_{55}) \hat{k}_x + (c_{33} + c_{55}) \hat{k}_z] \\
- \frac{1}{2} \sqrt{[(c_{11} - c_{55}) \hat{k}_x - (c_{33} - c_{55}) \hat{k}_z]^2 + 4(c_{13} + c_{55})^2 \hat{k}_x \hat{k}_z},
\]

(32)

and

Figure 8. Wave propagation in a realistic velocity model. (a) Marmousi velocity model. (b) Reference wavefield propagated using \(\Delta t = 1.5\) ms. (c) Difference between the reference wavefield and the wavefield propagated with the velocity-gradient term using \(\Delta t = 15\) ms. (d) Difference between the reference wavefield and the wavefield propagated without the velocity-gradient term using \(\Delta t = 15\) ms.

Figure 9. A slice through wavefields in Figure 8 at \(Z = 8800\) m. The black, short-dashed line in both (a) and (b) represents the reference wavefield; the blue, solid line in (a) represents the result calculated with the velocity gradient; and the red, long-dashed line in (b) represents the result calculated without the velocity gradient.
wavenumbers evaluated in the rotated coordinate aligned with the symmetry axis:

\[
\hat{k}_x = k_x \cos \theta + k_z \sin \theta, \quad \hat{k}_z = k_z \cos \theta - k_x \sin \theta,
\]

(33)

where \(\theta\) is the tilt angle measured with respect to the horizontal direction. Alternatively, the qP-wave phase velocity can be calculated using accurate approximate formulas (Alkhalifah, 1998, 2000; Fomel, 2004; Sripanich and Fomel, 2015).

To mimic an unbounded medium, we apply the absorbing boundary condition in equation 26. First, a time step size of 4 ms for the qP-wave and 2 ms for the qSV-wave are used. A Ricker wavelet source with a peak frequency of 16 Hz is located at \(X = 47.3\) km and \(Z = 0.175\) km. Figures 12a and 13a illustrate wavefield snapshots taken at different times overlaid on top of each other. Because of the accumulative effect of the large time step, model complexity, and extra anisotropy introduced by the absorbing boundary conditions, the low-rank approximation took \(N = 19\) for the qP-wave and \(N = 15\) for the qSV-wave for an accuracy threshold of \(\epsilon = 10^{-4}\).

To test the effect of using large \(\Delta t\) for wave propagation, we choose a series of increasing \(\Delta t\), while enforcing the same accuracy requirement \((\epsilon = 10^{-4})\). The qP-wave wavefields (Figure 12) generated using \(\Delta t = 4, 10, 20,\) and 30 ms are almost identical. Similarly, the qSV-wave wavefields (Figure 13) generated using \(\Delta t = 2, 4, 10,\) and 20 ms have very little difference. This shows that the stability and accuracy of wave extrapolation are not compromised by very large time steps, even beyond the Nyquist limit of 10 ms.

To investigate the effects of time step size \((\Delta t)\) and accuracy level \((\epsilon)\) of low-rank approximation on the computational cost, which is directly controlled by the rank of the approximation, we conduct a series of experiments using different values of \(\Delta t\) and \(\epsilon\) for qP-wave propagation. No absorbing boundary condition is applied in this test. Table 1 lists the rank \(N\) required by different time step sizes to achieve different accuracy levels. We observe that the low-rank algorithm requires a higher rank to maintain the same level of accuracy when the time step size increases. On the other hand, with the same time step size, a higher rank may be needed to achieve higher accuracy. In practice, we find that \(\epsilon = 10^{-4}\) is small enough to guarantee accurate wave kinematics. Because making a larger step in time requires a smaller number of steps, the optimal time step size \(\Delta t\) can be selected to be the one that minimizes the total number of FFTs required to accomplish the modeling or imaging task at hand. In practice, we found that this minimization problem

![Figure 10](image_url)

Figure 10. Comparison between (a) a wavefield propagated with the direction-dependent absorbing boundary condition and (b) a wavefield propagated with the direction-independent absorbing boundary condition. The amplitude has been normalized to 1, and the scale bar has been clipped to 0.1 to highlight artificial reflections from the boundary.

![Figure 11](image_url)

Figure 11. BP-2007 anisotropic benchmark model. (a) P-wave phase velocity along the axis of symmetry \(V_z\). (b) P-wave phase in the symmetry plane \(V_x\). (c) Anellipticity parameter \(\eta\). (d) Tilt angle of the symmetry axis \(\theta\).
usually has a straightforward solution: The larger $\Delta t$ is, the less computation is required. Therefore, the optimal solution will be the largest $\Delta t$ that satisfies other constraints, such as the imaging condition when performing RTM.

**Wave propagation in 3D orthorhombic media**

To demonstrate 3D wave propagation in tilted orthorhombic media, we use the classic model from Schoenberg and Helbig (1997), which characterizes a TI medium with vertical fractures. The density-normalized orthorhombic stiffness matrix is (Schoenberg and Helbig, 1997)

$$
\begin{bmatrix}
9 & 3.6 & 2.25 & 0 & 0 & 0 \\
3.6 & 9.84 & 2.4 & 0 & 0 & 0 \\
2.25 & 2.4 & 5.9375 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 1.6 & 0 \\
0 & 0 & 0 & 0 & 0 & 2.182 \\
\end{bmatrix}
$$

To introduce spatial heterogeneity, we apply a moderate perturbation to the stiffness coefficients that is a function of $x$, as demonstrated by Figure 14 for the case of $C_{11}$. The model is further rotated 45° counterclockwise about the $z$-axis (azimuth angle) and 45° counterclockwise about the $x$-axis (dip angle).

We use the exact phase velocity in orthorhombic media (Tsvankin, 1997):

$$
V^2(x, k)|k|^2 = 2\sqrt{a^2/3 - b^3} \cos \left( \frac{\nu}{3} + \frac{2\pi}{3} \right) - \frac{a^2}{3},
$$

where $n = 0$ corresponds to the P-wave and $n = 1, 2$ corresponds to S-waves, and

$$
\nu = \arccos \left( - \frac{2(a/3)^3 - ab/3 + c}{2\sqrt{((a/3 - b)/3)^3}} \right) \quad (0 \leq \nu \leq \pi),
$$

$$
a = -(G_{11} + G_{22} + G_{33}),
$$

$$
b = G_{11}G_{22} + G_{11}G_{33} + G_{22}G_{33} - G_{12}^2 - G_{13}^2 - G_{23}^2,
$$

$$
c = G_{11}G_{23}^2 + G_{22}G_{13}^2 + G_{33}G_{12}^2 - G_{11}G_{22}G_{33} - 2G_{12}G_{13}G_{23},
$$

$$
G_{11} = c_{11}k_x^2 + c_{66}k_y^2 + c_{55}k_z^2,
$$

$$
G_{22} = c_{66}k_x^2 + c_{22}k_y^2 + c_{44}k_z^2,
$$

$$
G_{33} = c_{55}k_x^2 + c_{44}k_y^2 + c_{33}k_z^2,
$$

$$
G_{12} = (c_{12} + c_{66})k_xk_y,
$$

Figure 12. Wavefield snapshots taken at different times showing P-wave propagation through the 2007 BP TTI model. Propagation-direction-dependent absorbing boundary conditions allow little energy to be reflected, even at wide incident angles. The four wavefield snapshots calculated using different time step sizes are almost identical: (a) $\Delta t = 4$ ms, (b) $\Delta t = 10$ ms, (c) $\Delta t = 20$ ms, and (d) $\Delta t = 30$ ms.
\[ G_{13} = (c_{13} + c_{55})\hat{k}_x\hat{k}_z, \]  
(44)

and

\[ G_{23} = (c_{23} + c_{44})\hat{k}_y\hat{k}_z. \]  
(45)

To incorporate tilting into the orthorhombic anisotropy, we replace the original wavenumber components \( k_x, k_y, \) and \( k_z \) with \( \hat{k}_x, \hat{k}_y, \) and \( \hat{k}_z, \) which are wavenumbers evaluated in the rotated coordinate system aligned with the symmetry axis:

\[
\begin{align*}
\hat{k}_x &= k_x \cos \phi + k_y \sin \phi, \\
\hat{k}_y &= -k_x \sin \phi \cos \theta + k_y \cos \phi \cos \theta + k_z \sin \theta, \\
\hat{k}_z &= k_x \sin \phi \sin \theta - k_y \cos \phi \sin \theta + k_z \cos \theta,
\end{align*}
\]  
(46)

<table>
<thead>
<tr>
<th>( \epsilon )</th>
<th>( \Delta t ) (ms)</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^{-3} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>10</td>
<td>14</td>
<td>19</td>
<td>33</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>( 10^{-4} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>9</td>
<td>13</td>
<td>15</td>
<td>22</td>
<td>30</td>
<td>46</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>( 10^{-5} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>14</td>
<td>18</td>
<td>23</td>
<td>30</td>
<td>43</td>
<td>59</td>
<td>84</td>
<td></td>
</tr>
</tbody>
</table>
where $\phi$ is the azimuth angle representing horizontal rotation (the angle between the original $x$-axis and the rotated one), and $\theta$ is the dip angle measured from vertical. Figure 15 demonstrates the wavefield snapshots taken at $t = 1$ s for three wave modes: the quasi-P-wave and the coupled quasi-S-waves. Note that the quasi-S-waves are propagated separately using solutions from equation 35 and then summed together because the two modes do not decouple easily in an orthorhombic medium.

**RTM of BP 2007 TTI data set**

Finally, we test low-rank one-step RTM using the BP 2007 anisotropic benchmark data set (Figure 11). The grid spacing is 18.75 m in the horizontal direction and 12.5 m in the vertical direction. The time step size is selected to be 4 ms. We use the acoustic approximation (Alkhalifah, 1998, 2000; Fomel, 2004) to calculate the qP-wave phase velocity. The sedimentary layers, salt boundaries, anticline, and fault surfaces are clearly imaged by the low-rank one-step RTM method (Figure 16).

**DISCUSSION**

The modeling experiments using an isotropic two-layer model and a portion of the BP 2007 TTI benchmark model show that the proposed operator is in practice remarkably stable and accurate. When the velocity contrast is not very sharp, the low-rank one-step method is capable of propagating waves free of dispersion artifacts using time steps even larger than the Nyquist sampling limit of the source wavelet. The maximum efficiency can be achieved using the largest step size possible, which in RTM applications usually corresponds to the time sampling of the imaging condition. In contrast, conventional approaches, such as high-order FD and pseudospectral methods, are confined to small step sizes due to severe stability and accuracy constraints. This is especially true when S-wave extrapolation is required by RTM because the slow S-wave velocity at shallow depths imposes a strict constraint on the time step size of conventional methods. Low-rank one-step wave extrapolation has been recently applied to converted wave imaging by Casasanta et al. (2015) and has achieved accurate results.
Our numerical experiments confirm the advantages of low-rank one-step wave extrapolation over the two-step scheme. By extrapolating an analytical wavefield with the imaginary part related to the first derivative of the real-valued wavefield, the one-step scheme is capable of using a significantly larger time step size (Du et al., 2014). The complex phase function used by the low-rank one-step method offers additional freedom in the design of the wave extrapolation operator. In media with smoothly varying velocity, a large time step may impair accuracy using the conventional formulation. By using a more accurate expression, for example, by admitting more terms from the Taylor expansion of the phase function, the low-rank one-step wave extrapolation can achieve higher accuracy. The complex-valued extrapolation operator also allows for an effective mixed-domain absorbing boundary condition, which dampens wave energy according to the phase direction, and thus avoids artificial reflections at large incident angles. The application of a complex phase function is not limited to the proposed cases. Another possible extension is seismic modeling and imaging in viscoacoustic media (Zhu and Harris, 2014), where the one-step extrapolator can incorporate a complex-valued, decoupled dispersion relation that incorporates attenuation effects (Sun et al., 2014, 2015).

CONCLUSIONS

We have developed low-rank wave extrapolation using a one-step scheme. The one-step low-rank method appears to be more stable than the two-step method, and it exhibits superior stability in numerical experiments. The capability of propagating waves using large time steps can help in saving costs when performing modeling, imaging, or FWI tasks. We propose two modifications to the complex phase function, which can be accurately handled by the low-rank one-step approach. First, we show that, when the velocity-gradient term is included in the approximation of the phase function, higher accuracy can be achieved. Next, we use a one-step scheme to incorporate a propagation-direction-dependent absorbing boundary condition in the wave propagation operator, which reduces artificial reflections at wide incident angles. Numerical examples using simple models demonstrate the improved accuracy and efficiency of the proposed method. In particular, the low-rank one-step operator to wave extrapolation in 2D TTI media and 3D orthorhombic media to produce wavefields free of dispersion artifacts or residual pseudo-S-wave artifacts. When low-rank one-step RTM is applied on the BP 2007 TTI benchmark data set, it produces a high-quality seismic image.

ACKNOWLEDGMENTS

We thank Z. Wu and one anonymous reviewer for their constructive comments. We would like to also thank J. Cheng, G. Fang, P. Fowler, J. Hu, S. Li, X. Song, Z. Xue, and Y. Zhang for inspiring discussions. We thank the sponsors of the Texas Consortium for Computational Seismology for financial support. The first author was additionally supported by the Statoil Fellows Program at the University of Texas at Austin. The BP 2007 benchmark data set was created by H. Shah and released by BP Exploration Operation Co., Ltd. We thank the Texas Advanced Computing Center for providing the computational resources used in this study.

APPENDIX A

PROOF OF THE STABILITY OF THE ONE-STEP WAVE EXTRAPOLATION OPERATOR

In this appendix, we prove the unconditional stability of the one-step wave extrapolation linear operator \( u = L f \) in 1D isotropic media defined by

\[
u(x) = \sum_{\xi \in \mathbb{C}} e^{2\pi i (x \xi + V(x) |\xi| t)} \hat{f}(|\xi|)
\]

for \( f(x) \in L^2([0, 1]) \), (A-1)

where \( 2\pi \xi = k \), \( f(x) \) is assumed to have a periodic boundary condition, and \( \hat{f}(|\xi|) \) is the Fourier transform of \( f(x) \) as defined by equation 3. We treat \( x \) as discrete and \( x \) as continuous for ease of derivation. Our argument can be viewed as a discrete version of the standard stationary phase method in the study of pseudodifferential operators (Stein, 1993; Grigs and Sjöstrand, 1994). To show that operator \( L \) is stable, a sufficient condition is that \( \|L\|_{2 \to 2} \leq 1 + Ct \), where \( C \) is a bounded constant. From equation A-1, we observe that operator \( L \) is the composition of two operators \( L = AF \), where \( F \) is the inverse Fourier transform and \( A \) is the operator defined by

\[
(A \omega)(z) = \sum_{\xi \in \mathbb{C}} e^{2\pi i (z \xi + V(x) |\xi| t)} \omega(|\xi|) \quad \text{for } \omega \in L^2(z). \quad \text{(A-2)}
\]

Let us consider \( A^T A : l^2 \to l^2 \), where \( A \) corresponds to a matrix with \( (x, \xi) \) entry given by \( A(x, \xi) = e^{2\pi i x \xi + V(x) |\xi| t} \) and \( A^T \) corresponds to a matrix with \( (\eta, x) \) entry given by \( A^T(\eta, x) = e^{-2\pi i x \eta + V(x) |\eta| t} \). The \( A^T A \) represents a matrix with \( (\eta, \xi) \) entry given by

\[
A^T A(\eta, \xi) = \int e^{2\pi i x (\xi - \eta) + V(x) (|\xi| - |\eta|) t} dx. \quad \text{(A-3)}
\]

To bound the \( l^2 \to l^2 \)-norm of \( A^T A \), we estimate the \( (\eta, \xi) \) entry of \( A^T A \). For \( \xi = \eta \), we have \( A^T A(\eta, \xi) = 1 \). For \( \xi \neq \eta \)

\[
x(\xi - \eta) + V(x) (|\xi| - |\eta|) t = x(\xi - \eta) + V(x) \alpha (|\xi| - |\eta|) t,
\]

(A-4)

with \( \alpha = (|\xi| - |\eta|) / (\xi - \eta) \). Clearly, \( |\alpha| \leq 1 \). Then, \( A^T A \) can be expressed as

\[
A^T A(\eta, \xi) = \int e^{2\pi i x (\xi - \eta) + V(x) |\eta| t} dx. \quad \text{(A-5)}
\]

For sufficiently small \( t \), \( x + V(x) \alpha t \) satisfies

\[
\frac{1}{2} \leq \nabla_x [x + V(x) \alpha t] \leq \frac{3}{2}.
\]

(A-6)

Equation A-5 can be expressed as
\[ A^T A (\eta, \xi) = \int e^{2i(\eta-\xi) \cdot |x|} \frac{1 + V'(x)\eta + V(x)\xi|}{1 + V'(x)\eta + V(x)\xi} \, dx, \]
\[ = \int \frac{1}{1 + V'(x)\eta + V(x)\xi} e^{2i(\eta-\xi) \cdot x} \, dx. \quad (A-7) \]

Let us define \( y = x + V(x)\eta t \). From equation (A-6), it is clear that the map \( x \leftrightarrow y \) is one to one. Substituting \( y \) into equation (A-7) gives
\[ A^T A (\eta, \xi) = \int \frac{1}{1 + V'(y)\eta} e^{2i(\eta-\xi) \cdot y} \, dy, \quad (A-8) \]
which is the inverse Fourier transform of \( \frac{1}{1 + V'(y)\eta} \). When \( t \) is small,
\[ A^T A (\eta, \xi) = \int \left[ 1 - V'(y)\eta + O(t^2) \right] e^{2i(\eta-\xi) \cdot y} \, dy, \]
\[ = -\eta t \int V'(y)\eta e^{2i(\eta-\xi) \cdot y} \, dy + O(t^2). \quad (A-9) \]

To evaluate the norm of the integration term in the last equation, we perform integration by part for \( k \)-times and apply periodic boundary condition:
\[ \left| \int V'(y) e^{2i(\eta-\xi) \cdot y} \, dy \right| \leq \frac{1}{(2\pi |\xi - \eta|)^k} \int e^{2i(\eta-\xi) \cdot y} V'(y) \, dy. \quad (A-10) \]

Assuming sufficient smoothness on \( V(x) \), we have for \( \eta \neq \xi \)
\[ |(A^T A - I)(\eta, \xi)| \leq \frac{Ct}{|\xi - \eta|} \quad (A-11) \]
for a constant \( C \). To estimate the \( l^2 \rightarrow l^2 \)-norm of \( A^T A \), we make use of the following lemma, which can be derived from direct calculation:

**Lemma 1.**

Suppose \( G : l^2 \rightarrow l^2 \) with \( \sum_{\xi} G(\xi, \eta) \leq C \) and \( \sum_{\eta} G(\xi, \eta) \leq C \), then \( G \) is a bounded \( l^2 \rightarrow l^2 \) operator.

We now apply the above lemma to \( A^T A - I \). When \( k \geq d \), where \( d \) is the spatial dimension, we have \( \sum_{\xi} G(\xi, \eta) \leq C \) bounded. Therefore, for sufficiently smooth \( V \), we have \( |(A^T A - I)|_{2 \to 2} \leq Ct \), for sufficiently small \( r \). Hence, \( |(A^T A)|_{2 \to 2} \leq 1 + Ct \) and \( |A|_{2 \to 2} \leq 1 + Ct \). Because \( L = AF \) and \( F \) as the Fourier transform is an isometry, we have \( |L|_{2 \to 2} \leq 1 + Ct \).

When performing wave extrapolation, fix a final time \( T \) and propagate \( T/\delta t \) steps, the operator is stable because
\[ \| L^{T/\delta t} \|_{2 \to 2} \leq (1 + Ct)^{T/\delta t} \leq e^{CT}. \quad (A-12) \]

**REFERENCES**