Equivalent accuracy at a fraction of the cost: Tackling spatial dispersion

Huy Le, Robert G. Clapp, and Stewart A. Levin

ABSTRACT

To reduce numerical spatial dispersion and find an optimal set of finite difference coefficients for a given frequency bandwidth and a range of velocities, we minimize the weighted sum of the squared error between the finite difference operator and the continuous operator. We reformulate the optimization problem in terms of frequency and velocity, which allows us to weight our cost function according to the frequency content of our injected source and to the velocity distribution present in our model. We show that our method gives promising results on a constant velocity model and a constant-thickness, linearly-increasing velocity model. However, without selecting the appropriate portion of the domain on which we optimize, the error at mid-range frequencies may be increased as a trade-off for reducing the error at high frequencies. This problem has been noted in previous work but not emphasized strongly enough. In this paper, we show numerical examples demonstrating this critical point.

INTRODUCTION

The method of finite differences (FD) is commonly used to solve the wave equation in seismic modeling, migration, and inversion. In general, FD will introduce deviations that make parts of the wavefield travel at a different velocity than the true medium velocity. This numerically-induced dispersion is more severe in the presence of high frequencies (Dablain, 1986). Reduction of the numerical dispersion can be achieved by using a finer spatial sampling or by using higher orders of approximation, both of which increase the computational cost.

A number of approaches have been taken to reduce numerical dispersion without requiring either a fine grid or higher orders. Holberg (1987) tried to correct for the error in group velocity caused by the differencing scheme. Etgen (2007) revisited the problem and corrected for phase-velocity error, taking a view that spatial and temporal dispersion could be used to compensate for each other. Zhang and Yao (2013) treated spatial dispersion separately and optimized the FD spatial derivative operator coefficients by reducing the maximum absolute error over some wavenumber range. Their approach assumes that temporal dispersion is either insignificant, possibly due to the use of a very small time step or a higher order of temporal approximation, or can be corrected separately. Correcting separately for the temporal dispersion is

investigated in a companion article in this report (Li et al., 2013). We follow a path similar to that of Zhang and Yao, but parameterize our calculations in terms of velocity and frequency to take advantage of data-dependent values and limits for these physical parameters with aim of obtaining more accurate results at equivalent FD computational cost.

OPTIMIZATION OF FD COEFFICIENTS

In the Fourier domain, the error between the finite-difference and the continuous, second-order spatial differentiation operators has the form:

$$\mathcal{E} = \mathcal{O}_{FD}(k) - \mathcal{O}_{cont}(k) = \left(\frac{1}{\Delta^2} \sum_{n=-M/2}^{M/2} a_n e^{ink\Delta} \right) - (-k^2), \tag{1}$$

where Δ is the spatial discretization size, M is the order of approximation, a_n are constant coefficients, and k is the wavenumber. Because the same set of coefficients is used for all spatial axes, here k can be considered the wavenumber in any direction.

Following Zhang and Yao (2013), we require our FD constant coefficients to satisfy a certain set of conditions:

$$\begin{aligned}
\text{symmetry:} \quad & a_{-n} = a_n, & (2)\\
\text{zero mean:} \quad & \textstyle\sum_{n=-M/2}^{M/2} a_n = 0, & (3)\\
\text{monotonically decreasing amplitude:} \quad & |a_n| > |a_{n+1}|, & (4)\\
\text{alternating signs:} \quad & a_n a_{n+1} < 0\ . & (5)
\end{aligned}$$

The first two conditions, equations 2 and 3, lead to a relation for the center coefficient $a_0 = -2\sum_{n=1}^{M/2} a_n$. As a result, the FD operator simplifies to

$$\mathcal{O}_{FD} = \frac{2}{\Delta^2} \sum_{n=1}^{M/2} a_n [\cos(nk\Delta) - 1]. \tag{6}$$

Another condition the FD coefficients should satisfy, which was not acknowledged in Zhang and Yao's work, comes from the fact that as the spatial step size is reduced toward zero, the FD operator should become a better and better approximation of the continuous operator. Mathematically, this means:

$$\lim_{\Delta \to 0} \frac{2}{\Delta^2} \sum_{n=1}^{M/2} a_n [\cos(nk\Delta) - 1] = -k^2, \tag{7}$$

which implies that

$$\sum_{n=1}^{M/2} n^2 a_n = 1. \tag{8}$$

The conventional FD coefficients obtained from Taylor series approximation satisfy this set of conditions. These conditions reduce the number of free coefficients to $M/2-1$.

Using the exact dispersion relation $k = 2\pi f/v$, we form our cost function as a weighted sum of squares of the error function (equation 1):

$$\mathcal{C} = \sum_{f,v} W_1(f)W_2(v)\left(2\sum_{n=1}^{M/2} a_n[\cos(n\frac{2\pi f}{v}\Delta) - 1] + \Delta^2(\frac{2\pi f}{v})^2\right)^2, \tag{9}$$

where $W_1(f)$ and $W_2(v)$ are the weighting functions corresponding to the contributions of frequency and velocity in the cost function. In equation 9, the cost function has been multiplied by Δ^2 to make it dimensionless (Zhang and Yao, 2013).

In this work, we have used a simulated annealing algorithm to find the minimum of the cost function. As did Zhang and Yao, we chose to use simulated annealing because of its ability to easily handle the non-linear constraints in equations 4, 5, and 8. We end the search when the cost stablizes at a value that is significantly smaller (by one order of magnitude) than those calculated by Zhang and Yao's coefficients for a large number of iterations.

RESULTS AND DISCUSSION

The parameters for our modeling examples are presented in Table 1.

Model	**Velocity** (m/s)	**Temporal order**	**Δt** (ms)	**Spatial order**	**Δx** (m)
Constant velocity	2400	4th	1	8th	5
Constant thickness	1500-3300	4th	1	8th	5

Table 1: Modeling parameters we use in our examples.

Since numerical dispersion is most problematic at high frequencies, for all of the following examples we chose a Ricker wavelet with fundamental frequency of 80 Hz as our source. The corresponding weighting function, $W_1(f)$ in equation 9, is its amplitude spectrum. For the case of constant velocity, we applied our method with $v = 2400$ m/s. Figure 1 shows the difference between the FD operator and the continuous operator ($\mathcal{E}$ in equation 1) as a function of frequency for different sets of coefficients. In this example, we have experimented with different percentages of the source bandwidth on which we compute and minimize the cost function. From this figure, we observe that the conventional FD coefficients are good up to approximately 80 Hz, while with our coefficients optimized on 50% of the bandwidth (pink curve in Figure 1), the error starts to become significant at about 110 Hz. Using Zhang and Yao's coefficients, this limit is pushed a little further to 120 Hz. As we enlarge the optimized domain, we start to introduce more error in the mid-range frequencies as

a trade-off for reducing the error at the high end of the source bandwidth. This was also observed by Etgen (2007) and by Zhang and Yao (2013).

Figures 2 and 3 compare the wavefields and wavelets obtained by modeling with different sets of coefficients, for which the errors are shown in Figure 1. In both Figure 2 and Figure 3, going clockwise from the top-left quadrant, as the tailing dispersion (inside the true wavefront) is reduced, the leading dispersion (outside the true wavefront) becomes worse. This can be explained by the behavior of the error function we observed in Figure 1. The tailing dispersion is caused by the high-frequency waves, which have high wavenumber and travel more slowly, while the leading dispersion is caused by the mid-range frequency waves, which have lower wavenumber and travel faster (Dablain, 1986).

We also tested our optimization scheme on a constant-thickness 11-layer medium with velocity increasing linearly from 1500 m/s at the top to 3300 m/s at the bottom. Figures 4 and 5 compare the error between using conventional, Zhang and Yao's coefficients, and coefficients optimized on 100% of the source bandwidth a range of velocities. We can observe from both of these figures that the error is significantly large at high frequencies and low velocities. Figure 5b shows how the error changes with frequency for 11 velocities in our model. As the velocity changes, the position of the "error bumps" also changes. Unfortunately, as we shall see shortly, this is a red herring, that is, an artifact of the display. In fact, while we have shown there is value in *designing* the operator coefficients in the $f - v$ domain, those coefficients remain coefficients of a single spatial finite difference operator.

Comparisons of the wavefields and wavelets for the layered-medium example are shown in Figures 6 and 7. Similar to the case of constant velocity, as the optimized domain increases, we observe the reduction of the tailing dispersion and amplification of the leading dispersion. Here we did not apply any weighting to the velocity part of the cost function, $W_2(v) = 1$, because the layer thicknesses were constant.

No cancellation without temporal dispersion

One topic of discussion has been how we could achieve partial cancellation of spatial dispersion error in the presence of multiple velocities (Figure 5b); in particular, would we need to design separate spatial second derivative approximations for each of a range of velocities, or would a single approximation yield cancellation by itself.

The key step was to understand how to separate spatial dispersion from temporal dispersion in our analysis. In essence, we asked how our wave extrapolation behaves when we treat t as a continous variable. In this setting, we can simply replace ∂_t^2 with $-\omega^2$ in designing our spatial operator. If we let κ be the approximation to the true wavenumber k in our finite difference, then the wave equation we solve is

$$P_{tt} = -v^2(\kappa_x^2 + \kappa_z^2)P = -v^2|\vec{\kappa}|^2 P \tag{10}$$

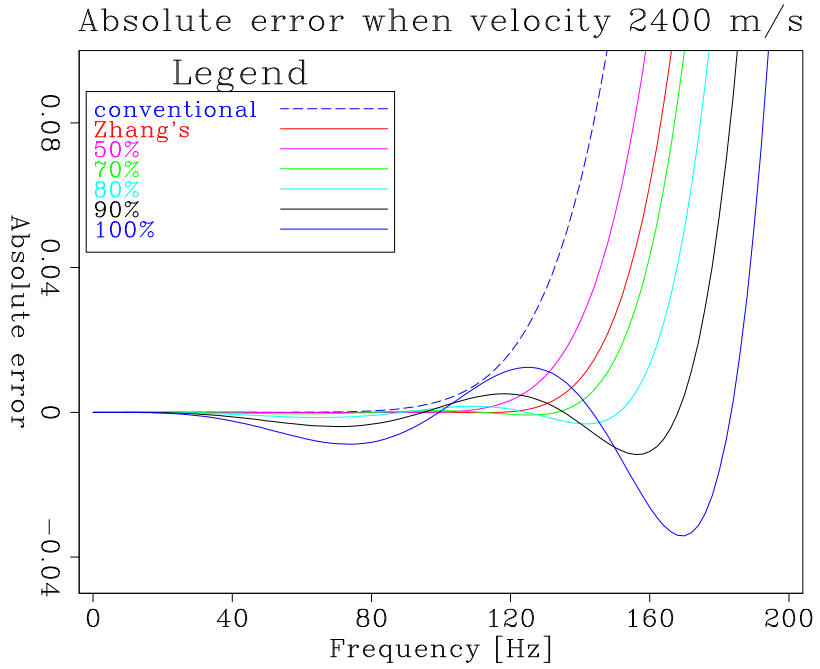


Figure 1: Error between the FD operator and the continuous operator ($\mathcal{E}$ in equation 1) as a function of frequency for different stencils. Percent notation indicates the coverage of the bandwidth on which we try to minimize the cost function. Increasing the coverage of the bandwidth boosts the mid-range-frequency error as a trade-off for reducing the error at high frequency. [**ER**]

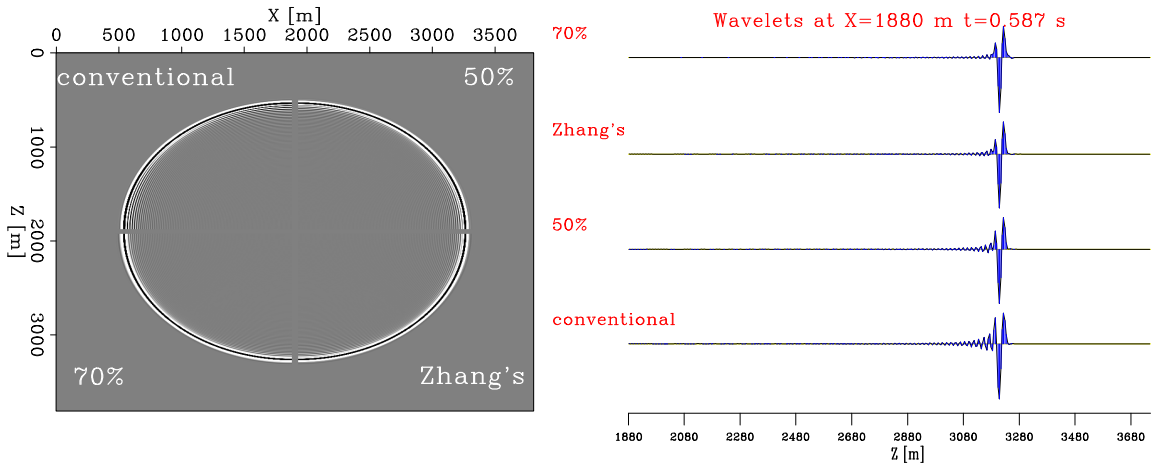


Figure 2: Comparison of wavefields (left) and of the wavelets (right) modeled with constant velocity using different sets of coefficients: conventional, Zhang and Yao's, optimized on 50%, and 70% of the bandwidth. [**ER**]

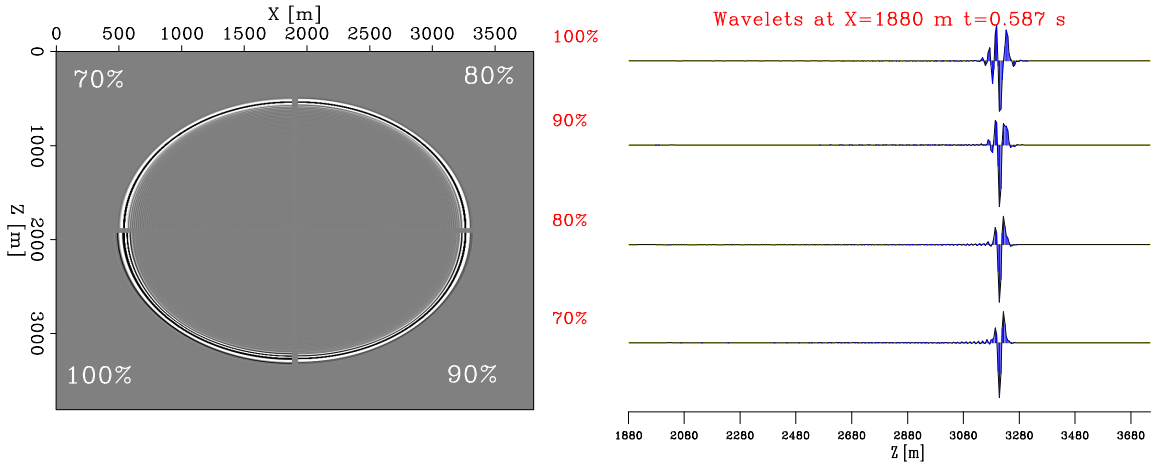
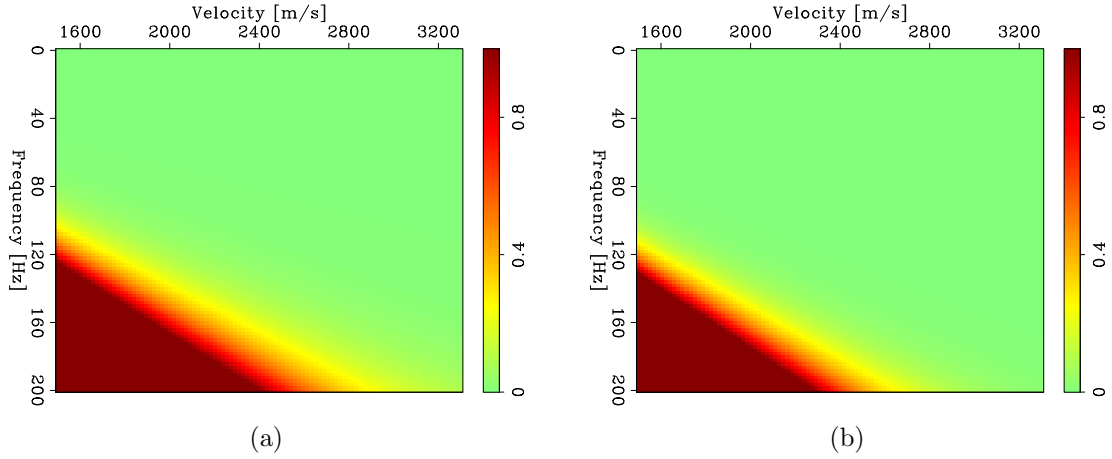


Figure 3: Comparison of wavefields (left) and of the wavelets (right) modeled with constant velocity using different sets of coefficients optimized on 70%, 80%, 90%, and 100% of the bandwidth. [**ER**]



(a) (b)

Figure 4: Error between the FD operator and the continuous operator as a function of frequency and velocity for: (a) conventional coefficients and (b) Zhang and Yao's coefficients. Carefully observed around the 120-Hz regions, Zhang and Yao's coefficients do a slightly better job than the conventional FD coefficient [**ER**]

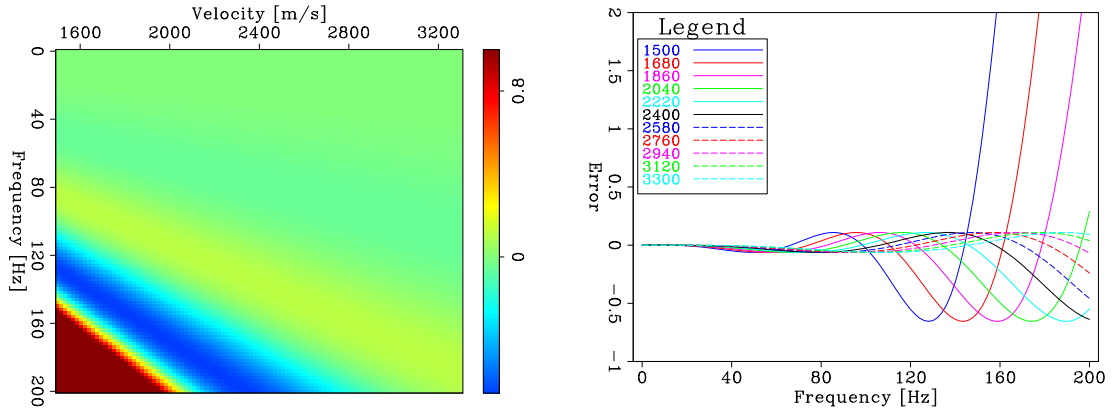


Figure 5: Error between the FD operator and the continuous operator as a function of frequency and velocity for coefficients optimized on 100% of bandwidth and: (a) a range of velocities from 1500-3300 m/s and (b) 11 velocities that are used for wavefield modeling later. Notice the movement of the "bumps" in the error plane and curves as velocity changes. [**ER**]

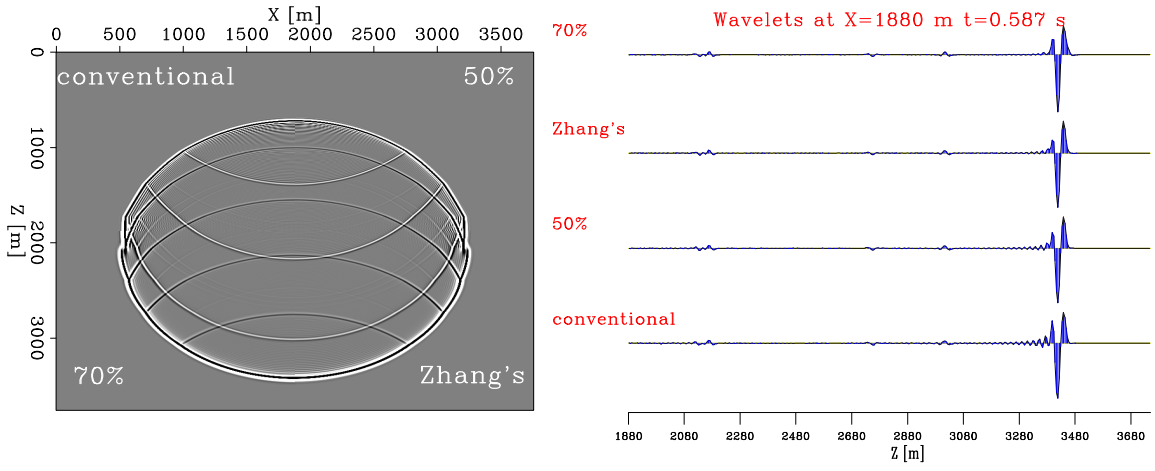


Figure 6: Comparison of wavefields (left) and of the wavelets (right) modeled in a constant-thickness layered medium using different sets of coefficients: conventional, Zhang and Yao's, optimized on 50%, and 70% of the bandwidth. [**ER**]

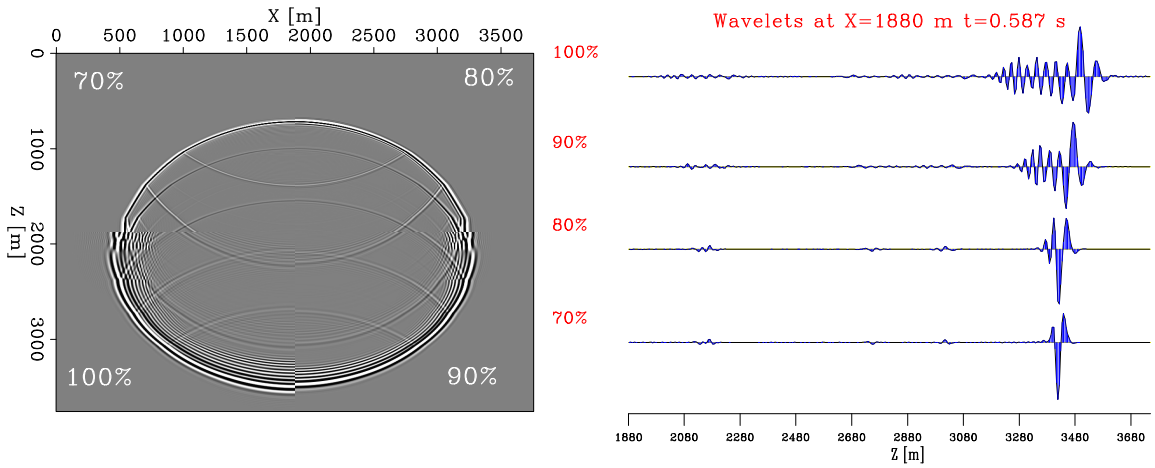


Figure 7: Comparison of wavefields (left) and of the wavelets (right) modeled in a constant-thickness layered medium using different sets of coefficients: 70%, 80%, 90%, and 100% of the bandwidth. [**ER**]

with solution being a linear combination of

$$e^{iv|\vec{\kappa}|t} \quad \text{and} \quad e^{-iv|\vec{\kappa}|t} \ .$$

Given this, taking steps of Δt_1 at velocity v_1, Δt_2 at velocity v_2, etc. yields the cumulative phase shift

$$\phi_{tot} = \pm|\vec{\kappa}| \sum_i v_i \Delta t_i \ , \tag{11}$$

which differs from the continuous solution phase shift by the factor $|\vec{\kappa}|/|\vec{k}|$.

Conclusion: The bumps and wiggles in the error of a single spatial second-derivative operator sum in phase in the absence of temporal dispersion.

No cancellation with temporal dispersion

Having established that a single spatial second-derivative approximation would not yield phase error cancellation without temporal dispersion, we included temporal dispersion by examining the phase error as a function of one (large) time step over a range of medium velocities.

Appendix B of Le and Levin (2013) provides the appropriate framework for the analysis. Substituting the scalar operator $-v^2|\vec{\kappa}|^2$ for $\mathbf{L}$ in that appendix, we find that for stable time steps our eigenvalues are the complex conjugate pair

$$\lambda = \frac{2 - v^2\Delta t^2|\vec{\kappa}|^2 \pm \sqrt{(2 - v^2\Delta t^2|\vec{\kappa}|^2)^2 - 4}}{2} \tag{12}$$

on the unit circle with corresponding eigenvectors (ignoring normalization)

$$\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} \lambda_2 \\ -1 \end{pmatrix} . \tag{13}$$

Looking now at the all-important phase ϕ of the eigenvalues (equation 12), we have

$$\cos\phi = \frac{2 - v^2\Delta t^2 |\vec{\kappa}|^2}{2} \tag{14}$$

$$\sin\phi = \pm \frac{v\Delta t |\vec{\kappa}| \sqrt{4 - v^2\Delta t^2 |\vec{\kappa}|^2}}{2} . \tag{15}$$

For a practical comparison, we took the 16^{th} order approximation of Zhang and Yao (2013) for a spatial 10-meter discretization and calculated the phase error after a time step of 2 ms, nearly the largest stable time step in this setting. The plot of the phase error is shown in Fig. 8. Even the most cursory examination of that result shows that the velocity-dependent errors are always in the same direction.

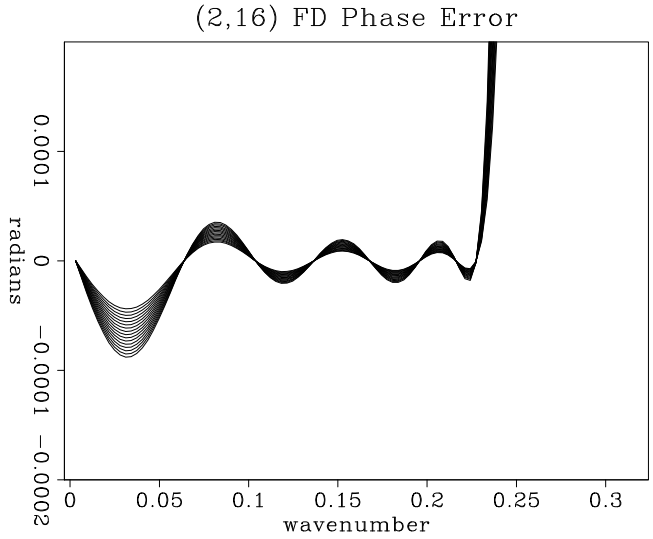


Figure 8: Finite-difference phase error after one time step of 2 ms on a 10-meter spacing grid using an explicit (2,16) finite difference scheme and varying the propagation velocity from 1500 to 3000 m/s. Quite clearly the phase errors for different velocities will not even partially cancel out. [**ER**]

Conclusion: The bumps and wiggles in the error of a single spatial second derivative operator sum in phase in the presence of temporal dispersion.

Conclusion: We need to design multiple spatial operator approximations in order to obtain some cancellation of spatial dispersion as we step in time. This will a significant focus of our future work.

CONCLUSIONS

We have reformulated the optimization problem for the FD constant coefficients in terms of velocity and frequency. This allows us to weight the cost function according to the frequency content present in our data and the velocity distribution in our model. As a result, we gain the flexibility to design an optimal FD coefficients that are suitable for a particular imaging or inversion problem. We have demonstrated that our method works for both a constant-velocity model and a constant-thickness, linearly increasing velocity model. However, without carefully selecting an appropriate portion of the optimization domain, we may boost the low- and mid-range frequency dispersion. Using a better minimum-searching algorithm or higher-order stencils might help increase this optimization coverage. Although in this work we use the standard FD grid, our optimization design can be straightforward to apply for more complicated FD schemes, such as staggered schemes, which have been recently used for modeling in anisotropic media (Chu, 2012).

ACKNOWLEDGMENTS

I would like to thank my colleagues, especially Elita (Yunyue) Li, Mandy Wong, and Ohad Barak, and professors at the Stanford Exploration Project for their great help and valuable advice and discussions.

REFERENCES

Chu, C., 2012, A hybrid finite difference method for acoustic wave propagation in tilted orthorhombic media: 82nd SEG Ann. Internat. Meeting, Expanded Abstracts, 1–5, Soc. of Expl. Geophys.

Dablain, M. A., 1986, The application of high-order differencing to the scalar wave equation: Geophysics, **51**, 54–66.

Le, H. and S. A. Levin, 2013, Practically stable unstable orthorhombic finite differences: SEP-Report, **149**, 243–250.

Li, Y. E., M. Wong, and R. Clapp, 2013, Equivalent accuracy at a fraction of the cost: Overcoming temporal dispersion: SEP-Report, **150**, 139–148.

Zhang, J.-H. and Z.-X. Yao, 2013, Optimized finite-difference operator for broadband seismic wave modeling: Geophysics, **78**, A13–A18.