GEOPHYSICS, VOL. 78, NO. 1 (JANUARY-FEBRUARY 2013); P. A13–A18, 4 FIGS., 1 TABLE.
10.1190/GEO2012-0277.1

Optimized finite-difference operator for broadband seismic wave modeling

Jin-Hai Zhang[1] and Zhen-Xing Yao[1]

ABSTRACT

High-resolution image and waveform inversion of small-scale targets requires the handling of high-frequency seismic wavefields. However, conventional finite-difference (FD) methods have strong numerical dispersions in the presence of high-frequency components. To reduce these numerical dispersions, we optimized the constant coefficients of the FD operator by maximizing the wavenumber coverage within a given error limitation. We set up three general criteria to enhance the convergence of the algorithm and reduce the optimization effort. We selected the error limitation to be 0.0001, this being the smallest in the literature, which led to perfect agreement between theoretical analyses and numerical experiments. The accuracy of our optimized FD methods can even reach that of much higher order unoptimized FD methods, which means great savings of computational efforts and memory demand. These advantages become even more apparent with 3D modeling, especially for saving memory demand.

INTRODUCTION

The finite-difference (FD) scheme is popular in seismic wave modeling because of its simplicity in numerical implementation and its ability to handle heterogeneous media. The FD method is usually the first choice in reverse time migration (Etgen and O'Brien, 2007) and full waveform inversion (Virieux and Operto, 2009). The FD method is also very popular in many other geophysical simulations, such as ocean acoustics and volcanic explosions. The main drawback of the conventional FD method is that it has strong numerical artifacts, also called "numerical dispersions," in the presence of high-frequency components or a coarse grid. Numerical dispersions are much more serious for large-scale models because it is impractical to use a fine grid due to the extremely large memory demand and computational cost. In such a case, we have to greatly decrease the dominant frequency to make the program runnable on current computers in an acceptable period of time. However, high-frequency components are critical for achieving high resolutions; thus, it is vitally necessary to reduce the numerical dispersions inherent to the FD method.

The pseudospectral method (Kosloff and Baysal, 1982) is free of numerical dispersions for high-frequency components and a coarse grid, regardless of error due to the time discretization, but it is not attractive for large-scale models because of the heavy computational cost. The pseudospectral method is the high-accuracy limit of high-order FD methods; thus, one can approach this method to reduce the numerical dispersions of the FD method (Holberg, 1987; Yang et al., 2002; Chu and Stoffa, 2012; Li et al., 2012). It is popular to apply a tapered window to avoid the truncation effects while reducing the numerical dispersions. Zhou and Greenhalgh (1992) use a Hanning window to derive convolutional FD operators; Igel et al. (1995) use a Gaussian window to develop staggered-grid FD operators; and Chu and Stoffa (2012) propose binomial windows, which are analytical extensions from conventional FD methods. Unfortunately, each of these windows involves some control parameters that are difficult to determine. At the same time, these windows need to be handled carefully because they can significantly affect the final numerical performance.

In fact, the final operator of the FD method is a series of real numbers; thus, there is no need to search for a proper window first and then mask it to the original operator. In other words, one can directly design the final operator itself. The basic idea is to enlarge the wavenumber coverage within a given error limitation (Holberg, 1987; Robertsson et al., 1994). Holberg (1987) optimizes the FD operator by minimizing the peak relative error of the group velocity within a spatial frequency band; Etgen (2007) suggests using the phase velocity and further including the effect of time discretization. For this kind of method, we find that it is essential to select a proper error threshold to obtain solid accuracy improvement. Unfortunately, previous works do not pay enough attention to this

Manuscript received by the Editor 17 July 2012; revised manuscript received 1 September 2012; published online 21 December 2012.
[1]Chinese Academy of Sciences, Institute of Geology and Geophysics, Beijing, China. E-mail: geophysics.zhang@gmail.com; yaozx@mail.igcas.ac.cn.

problem. The widely used error limitations from 0.0003 to 0.03, as Holberg (1987) suggests, are too big for high-accuracy modeling. That is why we usually see great improvements in the theoretical analyses but find them to be degraded or even completely lost in practical applications.

In this paper, we propose a general approach to directly design the operator of the FD method to reduce the numerical dispersions for broadband seismic wave modeling. First, we construct an objective function to evaluate the accuracy of the optimized FD operator. Then, we try to select a proper error limitation. Next, we build three criteria that the optimized FD operators should obey. Finally, we apply the simulated annealing algorithm (Kirkpatrick et al., 1983) to search for the operator that best satisfies the constraints from the objective function and the three criteria. These four steps ensure that the final accuracy improvement is stable and significant. The error limitation selected is 0.0001, which is the smallest one in the literature and leads to a perfect agreement between theoretical analyses and numerical experiments.

OPTIMIZATION SCHEME OF THE FD OPERATOR

The conventional FD operator for the second-order spatial differentiation of function $f(x)$ is basically a truncated Taylor series at $x = 0$ as follows (Fornberg, 1998; Chu and Stoffa, 2012):

$$\frac{\partial^2 f}{\partial x^2} \approx \frac{1}{\Delta^2} \sum_{n=-N/2}^{N/2} a_n \left[-\frac{2}{n^2} \cos(n\pi) \right] f_n, \tag{1}$$

where the even number N is the order, Δ is the uniform interval along x, $f_n = f(n\Delta)$, and a_n is the constant coefficient defined by the binomial coefficient formula

$$a_n = \binom{N}{\frac{N}{2}+n} \Big/ \binom{N}{\frac{N}{2}}. \tag{2}$$

We can reduce the numerical dispersions by using optimized constant coefficients of equation 1 without any other changes to the implementation of the FD method. Most previous works (e.g., Zhou and Greenhalgh, 1992; Igel et al., 1995; Chu and Stoffa, 2012) hope to design a proper window w_n to obtain the optimized FD operators, which can be generalized to the following form:

$$\frac{\partial^2 f}{\partial x^2} \approx \frac{1}{\Delta^2} \sum_{n=-N/2}^{N/2} w_n \left[-\frac{2}{n^2} \cos(n\pi) \right] f_n. \tag{3}$$

The main drawback of this kind of method is that these windows are often difficult to determine due to some control parameters involved. In addition, as some specific functions, these windows are usually not flexible enough to improve the accuracy significantly.

In fact, we can directly search for the final form of the optimized FD operator. That is, the window w_n is combined with all the other parts in equation 3 as follows:

$$\frac{\partial^2 f}{\partial x^2} \approx \frac{1}{\Delta^2} \sum_{n=-N/2}^{N/2} b_n f_n, \tag{4}$$

where b_n are our coefficients to be optimized and are the final form used in implementations. Our next work is to determine b_n using an optimization scheme.

In this paper, we directly optimize the coefficients by examining the peak absolute error between the optimized FD operator in the wavenumber domain and the analytical wavenumber, and we apply the simulated annealing algorithm (Kirkpatrick et al., 1983) to optimize the following objective function:

$$\max_{0 \le k_x \le k_x^{\max}} \left| -k_x^2 \Delta^2 - \sum_{n=-N/2}^{N/2} b_n \cos(n k_x \Delta) \right| \le T, \tag{5}$$

where k_x is the analytical wavenumber corresponding to the spatial derivative $\partial^2/\partial x^2$, $k_x^{\max}$ is the maximum accurate wavenumber range that the optimized FD operator can cover, and T is the maximum tolerant threshold (or error limitation). The simulated annealing algorithm allows more flexibility in solving the objective function than the least-squares approach.

However, there are $N + 1$ coefficients in our optimized FD operator (i.e., $b_{-N/2}$ to $b_{N/2}$), which is a fairly large number for the simulated annealing algorithm to handle. To reduce the optimization effort, we set up three criteria according to the theories of sinc interpolation (Chu and Stoffa, 2012) and finite impulse response (Oppenheim et al., 1999): (1) the operator is symmetric and the coefficients should be real numbers, that is, $b_n \in R$ and $b_{-n} = b_n$; (2) the total energy of the optimized FD operator should be zero; that is, $\sum_{n=-N/2}^{N/2} b_n = 0$; and (3) the coefficients should have an amplitude of damped oscillation away from the center position ($n = 0$); that is, $|b_n| > |b_{n+1}|$ and $b_n b_{n+1} < 0$ for $n = 1, 2, \cdots, N/2$. Rules 1 and 2 reduce the actual number of the coefficients to be only $N/2$, because $b_0 = 2\sum_{n=1}^{N/2} b_n$. Thus, we can optimize the whole operator by purely determining b_1 to $b_{N/2}$. Rule 3 greatly decreases the search scope and makes the simulated annealing algorithm affordable. In fact, the original coefficients of the conventional FD operators also obey these three criteria. This indicates that these three criteria are reasonable for general FD operators.

A proper error threshold is essential to the success of optimization (Zhang and Yao, 2012). For an error limitation that is too small (e.g., 0.00001), it is hard to gain a much wider wavenumber coverage. This means that the optimized FD operator gives few, if any, improvements on reducing numerical dispersions. For an error limitation that is too big (e.g., 0.0003 to 0.03 as Holberg [1987] suggests), we can easily cover a much wider wavenumber range; unfortunately, the numerical experiments would show major conflict with the theoretical analyses over the absolute error. That is, numerical dispersions would boost and significantly deviate from the accurate waveforms, especially for a large travel time or distance. This means we could not purely pursue wider coverage by arbitrarily relaxing the error limitation. As a reasonable tradeoff, after analyzing many numerical experiments, we select the tolerant threshold to be $T = 0.0001$, which is the smallest one shown in the literature. As shown later, this error limitation leads to an almost-perfect agreement on the coverage of accurate wavenumbers between the theoretical analyses and the numerical experiments.

For a low-order FD operator, the coefficients are too few to improve the accuracy. For a high-order FD operator, the coefficients are too many to be determined simultaneously. More importantly, introducing much higher orders in the optimized FD operator would not significantly increase the accuracy but would greatly increase the computational cost. Therefore, we only list the optimized coefficients of the fourth- to sixteenth-order FD methods in Table 1. Fortunately, these optimized FD operators, especially the optimized

eighth- to sixteenth-order FD operators, are feasible for many practical applications.

ABSOLUTE-ERROR ANALYSES

We perform theoretical accuracy analyses of the conventional and optimized FD operators by comparing their absolute errors. For the stability analyses, please refer to Holberg (1987). Figure 1 shows the absolute errors between the FD operators in the wavenumber domain and the analytical wavenumber. Obviously, the optimized eighth-order FD operator has much higher accuracy than the conventional eighth-order FD operator. In addition, the accuracy of the optimized eighth-order FD operator even reaches that of the conventional twelfth-order FD operator because their absolute-error curves share almost the same position in Figure 1a. Besides, the accuracy of the optimized twelfth-order FD operator is much greater than that of the conventional twelfth-order FD operator and even reaches that of the conventional twenty-fourth-order FD operator.

The absolute-error curves of the conventional FD operators are always close to zero positions and gradually increase with the increasing wavenumbers, as shown in Figure 1b. However, the absolute-error curves of the optimized FD operators behave quite differently from those of the conventional FD operators: they vibrate several times within a fixed error limitation and finally rush out of the error limitation at a much higher wavenumber. Although there are some deviations from zero positions for the absolute-error curves of the optimized FD operators, the maximum deviation is always within a quite narrow range of [−0.0001, 0.0001]. This means our optimized coefficients, below a given tolerable threshold, are effective in achieving a much wider accurate wavenumber range and would greatly reduce the numerical dispersions.

IMPULSE RESPONSES

In this section, we illustrate the above absolute-error analyses by impulse responses. A 2D homogeneous medium is defined on a grid of 511×511 with grid spacing of 5 m. The velocity is $v = 1500$ m/s, which is an extreme case in practical applications. A point source is located at the center of the media. Figure 2 shows the wavefield snapshots at 1.5 s. The dominant frequency of a Ricker wavelet is 40 Hz for Figure 2a and 50 Hz for Figure 2b. Obviously, the optimized eighth-order FD methods show much better results, i.e., fewer numerical dispersions, compared with the conventional eighth-order FD method. The results obtained by the optimized eighth-order FD method are quite similar to those obtained by the conventional twelfth-order FD method, and the results

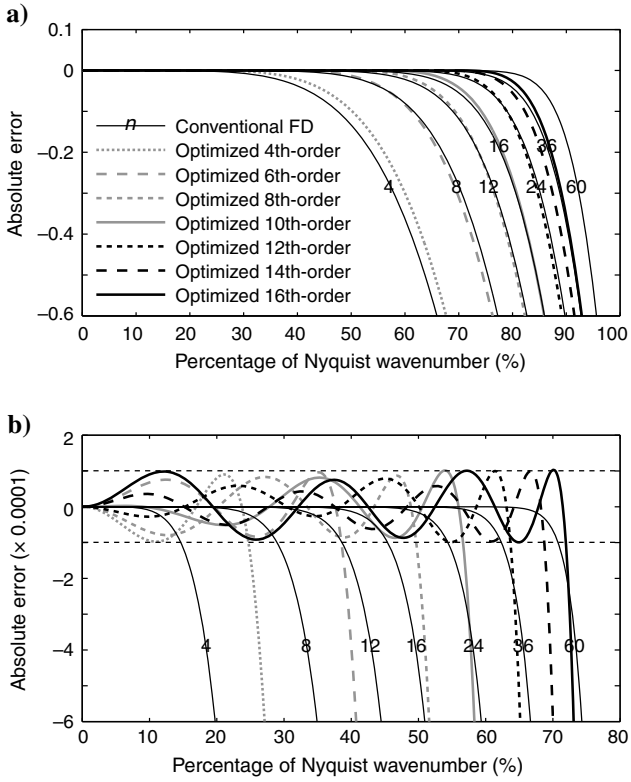

Figure 1. Accuracy comparison between the conventional and optimized FD operators: (a) a global view of absolute error within [−0.6, 0.1] and (b) a local view within [−0.0006, 0.0002]. The FD operators are for the second derivative along the spatial direction. Curves denote the absolute errors between the FD operators in the wavenumber domain and the analytical wavenumber. Thin solid curves denote the conventional FD operators with numbers indicating the order. The bold curves denote the optimized high-order FD operators. The optimized coefficients are shown in Table 1.

Table 1. Optimized coefficients of high-order FD operators.

	Fourth order	Sixth order	Eighth order	Tenth order	Twelfth order	Fourteenth order	Sixteenth order
b_0	−2.55567466	−2.81952122	−2.97399944	−3.05450492	−3.12108522	−3.16275980	−3.18543410
b_1	1.37106192	1.57500756	1.70507669	1.77642739	1.83730507	1.87636137	1.89789462
b_2	−0.09322459	−0.18267338	−0.25861812	−0.30779013	−0.35408741	−0.38612121	−0.40456799
b_3		0.01742643	0.04577745	0.07115999	0.09988277	0.12263042	0.13676734
b_4			−0.00523630	−0.01422784	−0.02817135	−0.04190565	−0.05150324
b_5				0.00168305	0.00653900	0.01330243	0.01893502
b_6					−0.00092547	−0.00344731	−0.00619345
b_7						0.00055985	0.00159455
b_8							−0.00020980

Only b_0 to $b_{N/2}$ are shown because $b_{-n} = b_n$ for $n = 1, 2, \cdots, N/2$, where N is the order of the FD operator.

obtained by the optimized twelfth-order FD method are quite similar to those obtained by the conventional twenty-fourth-order FD method. Figure 2 indicates that the improvement after using our optimization scheme is significant for the eighth- and the twelfth-order FD method. Note that these accuracy analyses based on numerical experiments show perfect agreement with the theoretical analyses in the previous section.

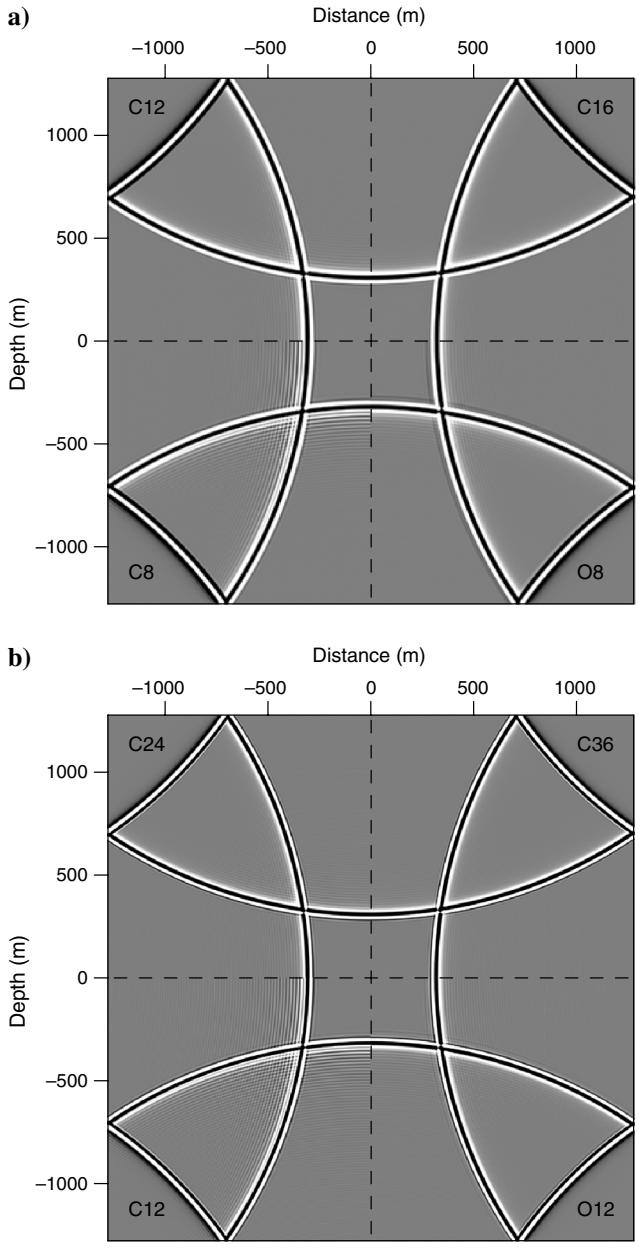

Figure 2. Impulse responses before and after using our optimization scheme. A point source is located in the middle of the homogeneous media with $v = 1500$ m/s. The dominant frequency of the Ricker wavelet is (a) 40 and (b) 50 Hz, respectively. Each subfigure has four equivalent parts with the amplitude clipped at 5%. In (a), the left-bottom, left-upper, and right-upper corners are generated by the conventional eighth-, twelfth-, and sixteenth-order FD methods, respectively, and the right-bottom corner is generated by the optimized eighth-order FD method. In (b), the left-bottom, left-upper, and right-upper corners are generated by the conventional twelfth-, twenty-fourth-, and thirty-sixth-order FD methods, respectively, and the right-bottom corner is generated by the optimized twelfth-order FD method.

TEST ON MARMOUSI MODEL

To verify the capabilities of our optimized FD method, we test on a modified Marmousi model, as shown in Figure 3a. For the convenience of an accuracy comparison, we take the waveforms generated by the conventional thirty-sixth-order FD method as references, which are plotted as dashed curves. For each time window shown in Figure 3b to 3d, the waveforms generated by the conventional twelfth-order FD method obviously deviate from the references due to the numerical dispersions. In contrast, the conventional twenty-fourth-order and optimized twelfth-order FD methods obtain almost the same waveforms that are consistent with the references. This indicates that our optimized FD method is much better than the conventional FD method when using the same order. In addition, our optimized twelfth-order FD method achieves almost the same accuracy as the conventional twenty-fourth-order FD method does. Again, we see that these accuracy analyses are consistent with those accuracy analyses in the previous two sections.

DISCUSSIONS

Figure 4 shows a comparison of memory demand and computational cost among different methods. We assume that a 2D model should have a fixed size; but the grid interval and grid number vary according to the minimum requirements of avoiding numerical dispersions. Obviously, the optimized FD methods have much less memory demand and computational cost compared with the same-order and lower-order conventional FD methods. On the other hand, the optimized FD methods have much less computational cost and similar memory demand compared with the same-order and higher-order conventional FD methods. That is, the optimized FD methods achieve significant reductions simultaneously on two important aspects: memory demand and computational cost. The magnitude of the reduction in the computational cost shown in Figure 4 is also correct for 3D models; however, the magnitude of the reduction in the memory demand would be much greater in 3D cases because the normalized ratios shown in Figure 4 would be squared.

Our optimized FD operators can also be used for seismic modeling on a coarse grid. The only difference is to use a coarser grid but lower frequency. Our approach of designing the FD operator is simple, flexible, and powerful. It can be easily extended to other related fields. The optimized coefficients are easy to use because the optimized FD methods have exactly the same algorithm structure as the conventional FD methods. This means that very few modifications of the existing code are required.

The staggered FD method (Virieux, 1984) has fewer numerical dispersions compared with the same order conventional FD method. The optimization scheme proposed here can be extended to the staggered FD method to further reduce the numerical dispersions. Additionally, this paper only concentrates on the FD discretization of the second derivative along the spatial direction. As a general approach of optimizing FD operators, our approach can be applied to other orders of spatial derivatives to optimize more general equations rather than only wave equations. We should use high-order time discretization (e.g., Li et al., 2012) rather than the conventional second order to maintain the final accuracy and to increase the time step. Of course, we can also apply our scheme to optimize the whole

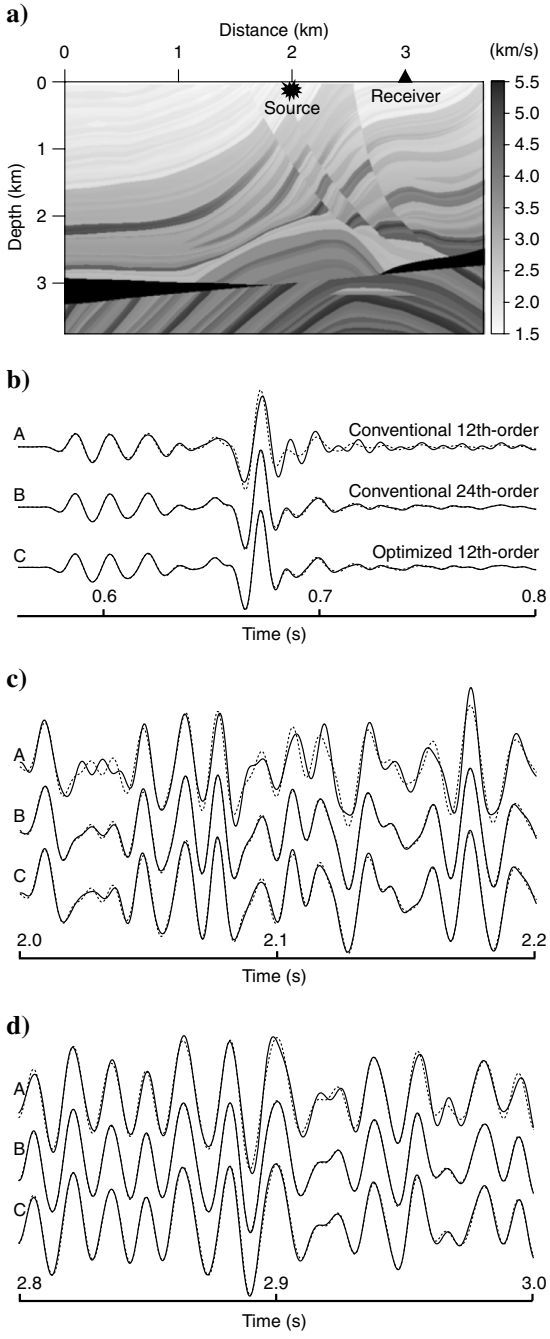

Figure 3. Waveform comparison among different methods within three time windows. (a) Modified Marmousi model, (b) 0.56–0.8 s (beginning from the first arrival), (c) 2.0–2.2 s, and (d) 2.8–3.0 s (including reflected waves from the deep targets). We keep the velocity values but set $\Delta x = \Delta z = 5$ m for convenient numerical analysis. The grid system is 737×751. The Courant number ($C = v\Delta t/\Delta x$) varies from 0.12 to 0.44. A point source is located at ($x = 2000$ m, $z = 20$ m); and a receiver is located at ($x = 3000$ m, $z = 5$ m). The dominant frequency of the Ricker wavelet is 50 Hz. Waveforms (see solid curves) are generated by the conventional twelfth- and twenty-fourth-order FD methods and the optimized twelfth-order FD methods, respectively. As references, waveforms generated by the conventional thirty-sixth-order FD method are also shown in each group of A, B, and C (see dashed curves).

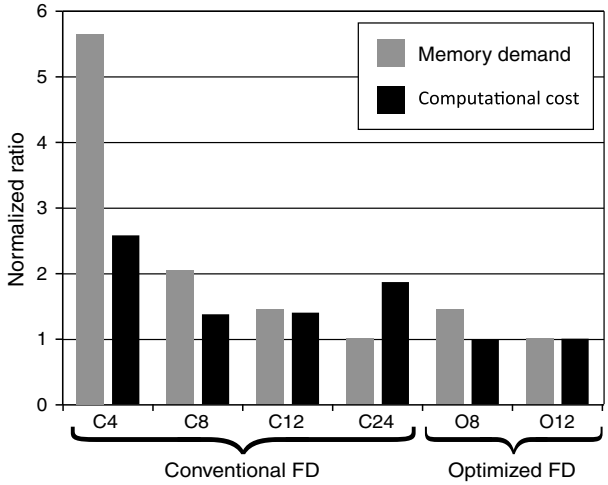

Figure 4. A comparison of memory demand and computational cost among different methods. Conventional fourth-, eighth-, twelfth-, and twenty-fourth-order FD methods and the optimized eighth- and twelfth-order FD methods are shown in this figure. We assume that a 2D model has a fixed size, but the grid interval and grid number vary according to the minimum requirements for avoiding numerical dispersions.

wave equation by taking the time discretization into account (Etgen, 2007; Etgen and Brandsberg-Dahl, 2009).

CONCLUSIONS

We present an optimization scheme to reduce the numerical dispersions of high-order FD methods. We combine the taper window that is popular in designing the optimized FD operator with the other coefficients into a new group of coefficients that are ready to be optimized; thus, we can directly determine the final optimized coefficients by searching for the maximum coverage of accurate wavenumbers below a given error threshold. We set up three criteria to guarantee the success of optimization, which can be extended to more general FD operators. We find that the selection of a proper error threshold is essential for consistent accuracy between the theoretical analyses and the numerical experiments. We select the error threshold to be 0.0001, which leads to a wide enough wavenumber coverage as well as a good agreement of accuracy between the practical application and the error analyses.

Our optimized eighth-order FD method has the same accuracy as the conventional twelfth-order FD method. This means we can save one-third of memory demand and computational cost after using our optimized eighth-order FD method. Our optimized twelfth-order FD method has the same accuracy as the conventional twenty-fourth-order FD method. This means we can save half of memory demand and computational cost after using our optimized twelfth-order FD method. For much higher orders, the accuracy improvements are much more apparent. These optimized FD operators enable us to handle much higher frequency components to explore much smaller structures without the influence of long-standing numerical dispersions.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (grant nos. 40974028, 41074092) and the National Major Project of China (grant no. 2011ZX05008-006). We

thank Joakim Blanch and anonymous reviewers for insightful comments and suggestions, which greatly improved the manuscript.

REFERENCES

Chu, C., and P. L. Stoffa, 2012, Determination of finite-difference weights using scaled binomial windows: Geophysics, **77**, no. 3, W17–W26, doi: 10.1190/GEO2011-0336.1.

Etgen, J. T., 2007, A tutorial on optimizing time domain finite-difference schemes: "Beyond Holberg": SEP, Report 129, 33–43.

Etgen, J. T., and S. Brandsberg-Dahl, 2009, The pseudo-analytical method: application of pseudo-Laplacians to acoustic and acoustic anisotropic wave propagation: 79th Annual International Meeting, SEG, Expanded Abstracts, 2552–2556.

Etgen, J. T., and M. J. O'Brien, 2007, Computational methods for large-scale 3D acoustic finite-difference modeling: A tutorial: Geophysics, **72**, no. 5, SM223–SM230, doi: 10.1190/1.2753753.

Fornberg, B., 1998, Calculation of weights in finite difference formulas: SIAM Review, **40**, 685–691, doi: 10.1137/S0036144596322507.

Holberg, O., 1987, Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena: Geophysical Prospecting, **35**, 629–655, doi: 10.1111/j.1365-2478.1987.tb00841.x.

Igel, H., P. Mora, and B. Riollet, 1995, Anisotropic wave propagation through finite-difference grids: Geophysics, **60**, 1203–1216, doi: 10.1190/1.1443849.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi, 1983, Optimization by simulated annealing: Science, **220**, 671–680, doi: 10.1126/science.220.4598.671.

Kosloff, D. D., and E. Baysal, 1982, Forward modeling by a Fourier method: Geophysics, **47**, 1402–1412, doi: 10.1190/1.1441288.

Li, X., W. Wang, M. Lu, M. Zhang, and Y. Li, 2012, Structure-preserving modelling of elastic waves: a symplectic discrete singular convolution differentiator method: Geophysical Journal International, **188**, 1382–1392, doi: 10.1111/j.1365-246X.2011.05344.x.

Oppenheim, A. V., R. W. Schafer, and J. R. Buck, 1999, Discrete-time signal processing (2nd edition): Prentice-Hall.

Robertsson, J. O. A., J. O. Blanch, W. W. Symes, and C. S. Burrus, 1994, Galerkin-wavelet modeling of wave propagation: Optimal finite difference stencil design: Mathematical and Computer Modelling, **19**, 31–38, doi: 10.1016/0895-7177(94)90113-9.

Virieux, J., 1984, SH-wave propagation in heterogeneous media: Velocity stress finite-difference method: Geophysics, **49**, 1933–1942, doi: 10.1190/1.1441605.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, doi: 10.1190/1.3238367.

Yang, S., Y. C. Zhou, and G. W. Wei, 2002, Comparison of the discrete singular convolution algorithm and the Fourier pseudospectral method for solving partial differential equations: Computer Physics Communications, **143**, 113–135, doi: 10.1016/S0010-4655(01)00427-1.

Zhang, J. H., and Z. X. Yao, 2012, Globally optimized finite-difference extrapolator for strongly VTI media: Geophysics, **77**, no. 4, T125–T135, doi: 10.1190/GEO2011-0505.1.

Zhou, B., and S. A. Greenhalgh, 1992, Seismic scalar wave equation modeling by a convolutional differentiator: Bulletin of the Seismological Society of America, **82**, 289–303.