



Review Paper

Overview of computation strategies on the dispersion analysis for explicit finite difference solution of acoustic wave equation

Jian-Ping Huang^{a, b}, Wei-Ting Peng^{a, b, *}, Ji-Dong Yang^{a, b}, Lu-Feng Lou^{a, b}

^a School of Geosciences, China University of Petroleum (East China), Qingdao, 266580, Shandong, China

^b Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao, 266000, Shandong, China

ARTICLE INFO

Article history:

Received 4 April 2023

Received in revised form

20 December 2023

Accepted 3 February 2024

Available online 5 February 2024

Edited by Jie Hao and Meng-Jiao Zhou

Keywords:

Finite-difference scheme

FD coefficients

Dispersion error

Forward modeling

Numerical simulation

ABSTRACT

Finite-difference (FD) method is the most extensively employed numerical modeling technique. Nevertheless, when using the FD method to simulate the seismic wave propagation, the large spatial or temporal sampling interval can lead to dispersion errors and numerical instability. In the FD scheme, the key factor in determining both dispersion errors and stability is the selection of the FD weights. Thus, How to obtain appropriate FD weights to guarantee a stable numerical modeling process with minimum dispersion error is critical. The FD weights computation strategies can be classified into three types based on different computational ideologies, window function strategy, optimization strategy, and Taylor expansion strategy. In this paper, we provide a comprehensive overview of these three strategies by presenting their fundamental theories. We conduct a set of comparative analyses of their strengths and weaknesses through various analysis tests and numerical modelings. According to these comparisons, we provide two potential research directions of this field: Firstly, the development of a computational strategy for FD weights that enhances stability; Secondly, obtaining FD weights that exhibit a wide bandwidth while minimizing dispersion errors.

© 2024 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Forward modeling of seismic wave is an efficient way to research the physical properties of the earth. In general, the numerical forwarding modeling technology includes ray tracing method (Julian and Gubbins, 1977; Dines and Lytle, 1979; Virieux and Farra, 1991; Mao and Stuart, 1997), integral equation method (Wannamaker et al., 1984a, 1984b) and differential equation approximation method (Alford et al., 1974; Cerjan et al., 1985; Gupta, 1966). The differential equation methods include the finite-difference (FD) method (Alterman and Karal, 1968; Dablain, 1986; Kelly et al., 1976), finite-element method (Bathe and Wilson, 1976; Brebbia, 1978; Belytschko and Mullen, 1978), pseudo-spectral method (Kreiss and Oliger, 1972; Orszag, 1972; Fornberg, 1975). The FD method is a valuable tool for efficiently approximating the seismic wave propagation process. Its widespread adoption in various industrial applications is primarily attributed to its cost-effectiveness and ease of implementation.

In numerical modeling, the FD method may generate severe dispersion error and instability due to improper simulated parameter selection. To overcome these problems, many strategies, like the FD method in frequency domain (Plessix and Mulder, 2004; Abubakar et al., 2011), the compact finite difference method (Lele, 1992; Kim and Lee, 1996), the staggered grid finite difference method (Levander, 1988; Madariaga, 1976), flux-corrected transport technique (Boris and Book, 1973; Book et al., 1975), implicit finite difference method (Kim and Lim, 2007; Liu and Sen, 2009b; Chen et al., 2016; Ren and Li, 2019) and the modification FD weights (Holberg, 1987; Kindelan et al., 1990; Yang et al., 2017a) are proposed. In all of these methods, the modification FD weights method is the only way that does not lead to an increase in the computational cost of the numerical simulation. The specific implementation process of this method is using appropriate FD weights to replace worse weights in the numerical modeling process. For the least-squares reverse time migration imaging (Tarantola, 1984; Gu et al., 2021; Zhu et al., 2018; Huang et al., 2015, 2016; Mu et al., 2020) and full waveform inversion (Mora, 1987, 1988; Pratt et al., 1996; Yong et al., 2019), the substantial computational expenses are a significant limiting factor for their utilization in the petroleum industry. Consequently, tremendous development and research

* Corresponding author. School of Geosciences, China University of Petroleum (East China), Qingdao, 266580, Shandong, China.

E-mail address: b22010077@s.upc.edu.cn (W.-T. Peng).

into appropriate FD weights have been carried out due to that it does not increase the computational cost.

According to the mathematical ideology of the computation of FD weights, the strategies can be categorized into three types, Taylor expansion strategy (which employs Taylor expansion to expand dispersion relation at zero wavenumber), optimization algorithm strategy (which involves solving the optimal problem by optimization algorithm) and window function strategy (which truncates the spatial convolution series of the pseudo-spectral method by window function).

Utilizing the Taylor expansion to approximate the dispersion relation in the space domain is the simplest method for obtaining the FD weights, commonly known as the conventional FD weights. These FD weights exhibit a significant temporal dispersion error because they only account for the dispersion relation in the space domain. To improve the temporal accuracy of the conventional FD weights, the Taylor expansion of the dispersion relation in the time-space domain was developed (Liu and Sen, 2009a, 2010, 2013; Yan and Liu, 2013; Tan and Huang, 2014). This novel approach shows a superior performance in minimizing phase velocity error compared to conventional FD weights (Liu and Sen, 2009). The application of this new Taylor expansion method can result in the FD method achieving 2M-th order accuracy. Moreover, the FD weights obtained through this method are notably more stable when compared to the conventional FD weights.

For the optimization algorithm strategy, we classify the optimization methods by their cost function, as the cost function plays a pivotal role in determining the FD weights properties. The cost functions can be divided into three types, L_∞ norm cost function (Holberg, 1987; Kindelan et al., 1990; Kosloff et al., 2010; Zhang and Yao, 2013a, 2013b; Yang et al., 2017a, 2017b; He et al., 2019; Liu, 2020a, 2020b, 2022; Koene and Robertsson, 2020), L_2 norm cost function (Wang and Wu, 2002; Etgen, 2007; Du et al., 2010; Liu, 2013, 2014; Yong et al., 2016, 2017a, 2017b; Zou et al., 2020a, 2020b; Ren and Liu, 2015; Ren and Li, 2017, 2019) and L_1 norm cost function (Miao and Zhang, 2020). L_∞ norm cost function minimizes the maximal error within a given wavenumber region. L_∞ cost function can make the FD weights cover a large bandwidth. L_2 norm cost function minimizes the sum of square errors within a given wavenumber region, and it is less susceptible to converging to local minima due to the convex nature of the L_2 optimization problem. L_1 norm cost function minimizes the sum of absolute errors within a given wavenumber region. This type of cost function is effective in reducing the dispersion error. Except for the norm form, the cost functions can be categorized into other types, which can be classified into the phase velocity cost function, group velocity cost function and dispersion relation cost function.

Window function strategy (Zhou and Greenhalgh, 1992a,b; Igel et al., 1995; Chen et al., 1997; Shao et al., 2003; Xiao and Tang, 2006; Chu et al., 2009; Chu and Stoffa, 2012; Wang et al., 2014; Liu and Wang, 2015; Zheng et al., 2016; Wang et al., 2017a; Wang et al., 2017b; Li et al., 2021) computes the FD weights by using the window function to truncate spatial convolution series of the pseudo-spectral method. The FD weights are influenced by the width of the main lobe and the attenuation of the sidelobe of the window function. Single window functions tend to have narrower main lobes and smaller sidelobe attenuation, resulting in FD weights covering a narrower bandwidth and causing greater dispersion error, compared with those obtained using the combined window function (Wang et al., 2015; Liu and Wang, 2015; Zheng et al., 2016; Wang et al., 2017a, 2017b). The combined window function widens the width of the main lobe and enhances the attenuation of the sidelobe by merging various single window functions. Nevertheless, for all window function methods, it's important to note that the shape of the main lobe and the side lobe of the window

function is highly related to the input parameters. In the single window function, these parameters include, β in Kaiser window function, α and β in Hanning window function, β in Gaussian window function, and so forth. In the combined window function, additional input parameters come into play, such as auto-convolution times (Wang et al., 2015), combination weight (Wang et al., 2017a), and modulate time (Wang et al., 2017b). For this strategy, numerous parameters are involved in the computational process. Nevertheless, the relationships of these parameters with the main lobe and side lobe have not been clearly summarized through either theoretical analysis or numerical experimentation in the previous research.

In this paper, we provided a comprehensive overview of fundamental theories for the window function, the Taylor expansion, and the optimization algorithm. Subsequently, we delve into the classical methods within these three computational strategies through theoretical analysis and numerical simulations. According to these tests, we draw some conclusions about each computational strategy. Finally, we discussed the possible challenges and suggested some recommendations in the field of computing FD weights.

2. Methodology

For simplicity, we begin by considering 1D acoustic modeling. The acoustic wave equation in homogeneous medium can be written as

$$\frac{\partial^2 u_0^0}{\partial t^2} \frac{1}{v^2} = \frac{\partial^2 u_0^0}{\partial x^2}, \quad (1)$$

where u is the wavefield, v is the velocity, t is the time, and x denotes the horizontal distance along the x -axis, the wavefield u_j^i is a short form of $u(x + hj, t + \tau l)$.

The 2M-th order FD operator is employed to approximate the second spatial order derivative, and the 2nd order FD operator is used to approximate the second temporal order derivative. These operators can be written as follows:

$$\frac{\partial^2 u_0^0}{\partial x^2} \approx \frac{1}{h^2} \sum_{m=-M}^M c_m u_m^0, \quad (2-1)$$

$$\frac{\partial^2 u_0^0}{\partial t^2} \approx \frac{1}{\tau^2} (u_0^1 - 2u_0^0 + u_0^{-1}), \quad (2-2)$$

where c_m represent FD weights, the h is the space sampling in x dimension, and τ is the time step. We rewrite the acoustic wave equation by substituting Eq. (2) into Eq. (1),

$$u_0^1 - 2u_0^0 + u_0^{-1} = v^2 \tau^2 \frac{1}{h^2} \sum_{m=-M}^M c_m u_m^0, \quad (3)$$

Given the approximation of wave propagation omitted truncation error and equivalent infinitesimal, the numerical solution of the wave equation computed by FD scheme is inherently inaccurate. However, we can use proper FD weights to improve the numerical result.

In general, the strategies for computing FD weights can be divided into three types according to their mathematical principle, Taylor expansion, optimization algorithm and window function as shown in Fig. 1. According to the development history, we have created a classification chart (Fig. 1), in which we have highlighted eight methods (in red) as classical approaches that we examine through theoretical analysis and numerical testing. The

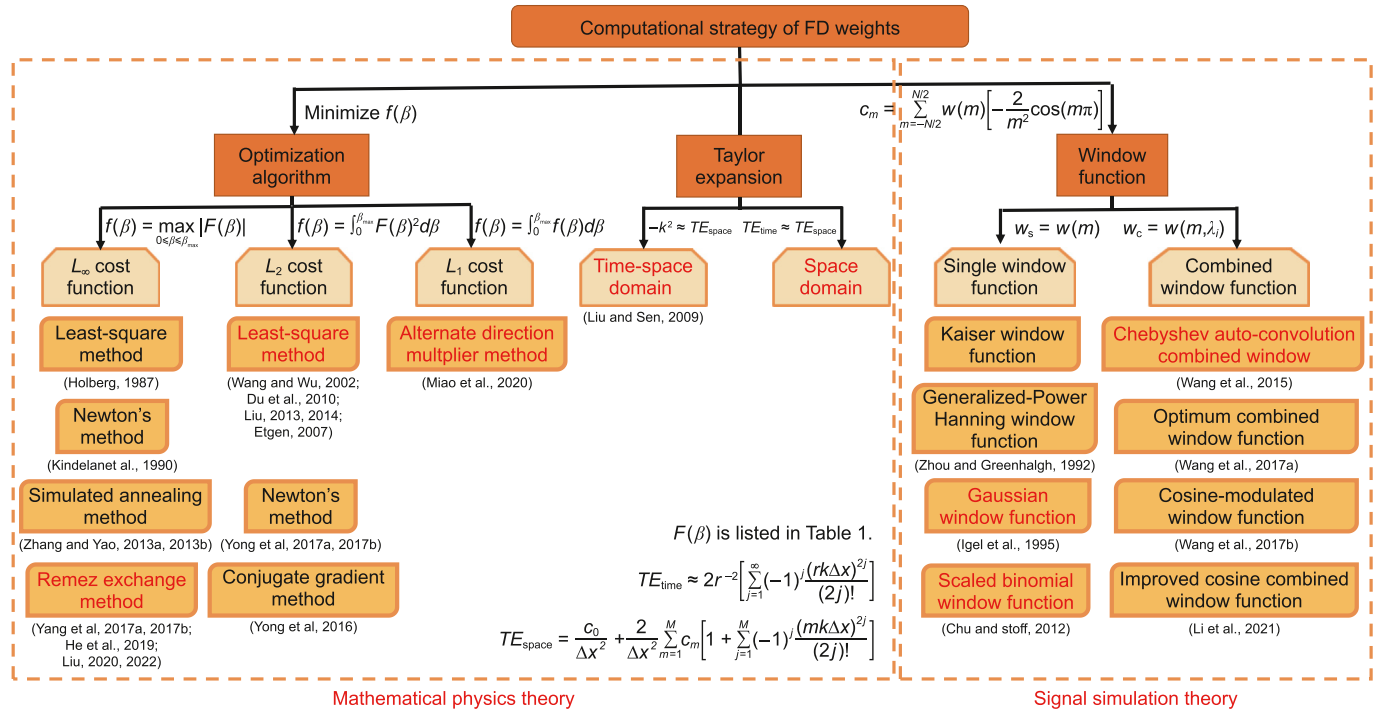


Fig. 1. The classification chart of computational strategy of FD weights.

methodology section will provide a detailed illustration of the computational processes for all the methods listed in Fig. 1.

2.1. The Taylor expansion strategy

2.1.1. The Taylor expansion of dispersion relation in space domain (TES)

The exact $2M$ -th order FD operator of second spatial derivative and the exact 2nd order FD operator of second temporal derivative can be written as

$$\frac{\partial^2 u_m^0}{\partial x^2} = \frac{1}{h^2} \sum_{m=-M}^M c_m u_m^0 + e_x \frac{\partial^{2M+2} u_m^0}{\partial x^{2M+2}} h^{2M} + O(h^{2M+1}), \quad (4)$$

where

$$e_x = -\frac{2}{(2M+2)!} \sum_{m=-M}^M m^{2M+2} c_m, \quad (5)$$

$$\frac{\partial^2 u_0^0}{\partial t^2} = \frac{1}{\tau^2} [-2u_0^0 + u_0^1 + u_0^{-1}] + e_t \frac{\partial^4 u_0^0}{\partial t^4} \tau^2 + O(\tau^3), \quad (6)$$

where

$$e_t = -\frac{2}{(4)!} (-1^4 \cdot 1 - 2 + 1^4 \cdot 1) = 0, \quad (7)$$

where the $O(h^{2M+1})$ and $O(\tau^3)$ are the higher order infinitesimal terms, and can be denoted as Q_x and Q_t , respectively; the $e_x \frac{\partial^{2M+2} u_m^0}{\partial x^{2M+2}} h^{2M}$ represents as P_x , and P_x is the truncated error term. These approximation (Eq. (4) and Eq. (6)) can be rewritten as

$$\frac{\partial^2 u_m^0}{\partial x^2} = \frac{1}{h^2} \sum_{m=-M}^M c_m u_m^0 + P_x + Q_x, \quad (8)$$

$$\frac{\partial^2 u_0^0}{\partial t^2} = \frac{1}{\tau^2} [-2u_0^0 + u_0^1 + u_0^{-1}] + Q_t, \quad (9)$$

According to the plane wave theory, we have

$$u_j^l = e^{i(kx - \omega t)} e^{i(jkh - l\omega\tau)}, \quad (10)$$

where k represents the wavenumber, and ω is the angular frequency. By substituting Eq. (10) to Eqs. (8) and (9) and omitting the P_x , Q_x , Q_t , the spatial $2M$ -th and temporal 2nd order FD operators can be rewritten as

$$-k^2 \approx \frac{1}{h^2} \left[c_0 + 2 \sum_{m=1}^M c_m \cos(mkh) \right], \quad (11)$$

$$-\omega^2 \approx \frac{1}{\tau^2} [-2 + 2 \cos(\omega\tau)], \quad (12)$$

where $c_0 = -2 \sum_{m=1}^M c_m$. Applying the Taylor series expansion for cosine functions in Eqs. (11) and (12), we obtain

$$-k^2 \approx \frac{c_0}{h^2} + \frac{2}{h^2} \sum_{m=1}^M c_m \left[1 + \sum_{j=1}^{\infty} (-1)^j \frac{(mkh)^{2j}}{(2j)!} \right], \quad (13)$$

$$-\omega^2 \approx -\frac{2}{\tau^2} + \frac{2}{\tau^2} \left[1 + \sum_{j=1}^{\infty} (-1)^j \frac{(\omega\tau)^{2j}}{(2j)!} \right], \quad (14)$$

where Eq. (13) is the Taylor expansion of the dispersion relation in space domain. After comparing the weights of $k^0, k^2, k^4, \dots, k^{2M}$, the

simplification of Eq. (13) is

$$\begin{bmatrix} 1^0 & 2^0 & \cdot & \cdot & \cdot & M^0 \\ 1^2 & 2^2 & \cdot & \cdot & \cdot & M^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1^{2M-2} & 2^{2M-2} & \cdot & \cdot & \cdot & M^{2M-2} \end{bmatrix} \begin{bmatrix} 1^2 c_1 \\ 2^2 c_2 \\ \cdot \\ \cdot \\ M^2 c_M \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \quad (15)$$

By solving Eq. (15), the FD weights from TES can be computed by

$$c_m = \frac{(-1)^{m+1}}{m^2} \prod_{1 \leq n \leq M, n \neq m} \left| \frac{n^2}{n^2 - m^2} \right|, m = (1, 2, 3, \dots, M). \quad (16)$$

The $c_0 = -2 \sum_{m=1}^M c_m$. It's worth noting that FD weights are computed through the Taylor expansion of the dispersion relation in the spatial domain at zero wavenumber, which leads to significant spatial and temporal dispersion errors. For comparing the spatial error and the temporal dispersion error, we introduced the relative error of the dispersion relation in the space domain and the error of phase velocity. In Fig. 2(a), the relative error of dispersion relation in the space domain is

$$REDR \approx \left[\frac{c_0 + 2 \sum_{m=1}^M c_m \cos(mkh)}{h^2 k^2} \right] + 1, \quad (17)$$

where the expression is derived from the dispersion relation in the space domain. Thus, the *REDR* indicates the spatial dispersion error. We can observe that the FD weights from the TES exhibit significant spatial dispersion errors at the high-wavenumber region as shown in Fig. 2(a). The error of the phase velocity is

$$EPV = \frac{2}{rkh} \sin^{-1} \sqrt{r^2 \sum_{m=1}^M c_m \sin^2(mkh/2)} - 1, \quad (18)$$

where $r = \frac{v\tau}{h}$. Fig. 2(b) shows the curve of error of phase velocity. If the curve's value is greater than zero, it signifies that this component of the wavefield will exhibit temporal dispersion errors. On the contrary, if the curve's value is less than zero, it suggests the presence of spatial dispersion errors in this component. Regarding the FD weights from TES, Fig. 2(b) illustrates that it will have a large

temporal error at the middle- and high-wavenumber regions.

When using the FD weights from TES to simulate the forward modeling, the accuracy of the FD scheme is the sum of errors in the temporal derivative and that in the spatial derivative. The TES FD weights only consider the weights of $k^0, k^2, k^4, \dots, k^{2M}$ in Eq. (13). Eq. (13) is derived from second spatial derivative. Thus, the rest of Eq. (13) is the error of spatial derivative, which can be expressed as

$$E_{spa} = \frac{2}{h^2} \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mkh)^{2j}}{(2j)!} \right]. \quad (19)$$

Similarly, the error of temporal derivative can also be expressed as

$$E_{tem} = \frac{2}{\tau^2} \left[\sum_{j=2}^{\infty} (-1)^j \frac{(\omega\tau)^{2j}}{(2j)!} \right]. \quad (20)$$

When the left side of Eqs. (11) and (12) plus the E_{spa} and E_{tem} , respectively, we obtain the following equations:

$$-k^2 + E_{spa} = \frac{c_0}{h^2} + \frac{2}{h^2} \sum_{m=1}^M c_m \cos(mkh), \quad (21)$$

$$-\omega^2 + E_{tem} = \frac{-2}{\tau^2} + \frac{2}{\tau^2} \cos(\omega\tau). \quad (22)$$

The total error of FD scheme for TES FD weights can be expressed as

$$E_{total} = \frac{c_0}{h^2} + \frac{2}{h^2} \sum_{m=1}^M c_m \cos(mkh) - \frac{1}{v^2} \left[\frac{-2}{\tau^2} + \frac{2}{\tau^2} \cos(\omega\tau) \right]. \quad (23)$$

Substituting Eqs. (21) and (22) into Eq. (23), we obtain

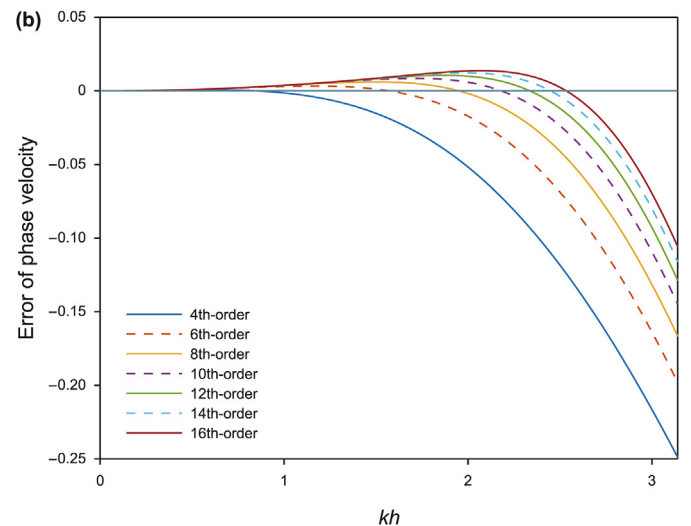
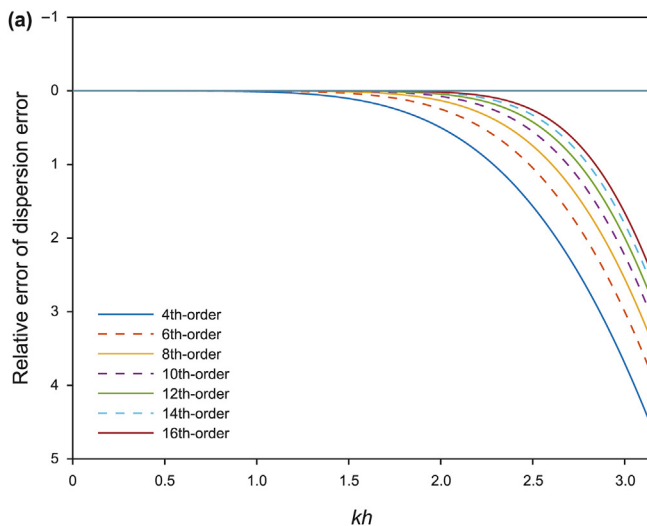


Fig. 2. The comparison of error curve for different order TES FD weights.

$$E_{\text{total}} = -2 \left[\sum_{j=2}^{\infty} (-1)^j \frac{r^{2j-2} k^{2j}}{(2j)!} h^{2j-2} \right] + 2 \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mk)^{2j}}{(2j)!} h^{2j-2} \right], \quad (24)$$

where $r = v\tau/h$. The first term of Eq. (24) is $-2 \frac{r^2 k^2}{4!} h^2$, where the power of h is equal to 2. Thus, the FD accuracy is 2nd-order. In order to increase the FD accuracy and mitigate the temporal error, Liu and Sen (2009a) proposed a new FD scheme solved in the time-space domain.

2.1.2. The Taylor expansion of dispersion relation in time-space domain (TETS)

Combining Eqs. (11) and (12), the dispersion relation in the time-space domain can be expressed as

$$c_0 + 2 \sum_{m=1}^M c_m \cos(mkh) \approx 2r^{-2} [\cos(rkh) - 1]. \quad (25)$$

$$E_{\text{total}} = -2 \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{r^{2j-2} k^{2j}}{(2j)!} \tau^{2j-2} \right] + 2 \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mk_x \cos \theta)^{2j}}{(2j)!} h^{2j-2} \right] + 2 \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mk_z \sin \theta)^{2j}}{(2j)!} h^{2j-2} \right], \quad (32)$$

Using the Taylor expansion to expand cosine functions in Eq. (25), we obtain

$$c_0 + 2 \sum_{m=1}^M c_m \left[1 + \sum_{j=1}^{\infty} (-1)^j \frac{(mkh)^{2j}}{(2j)!} \right] \approx 2r^{-2} \left[\sum_{j=1}^{\infty} (-1)^j \frac{(rkh)^{2j}}{(2j)!} \right]. \quad (26)$$

Considering the weights of $k^0, k^2, k^4, \dots, k^{2M}$, Eq. (26) can be expanded as

$$\begin{bmatrix} 1^0 & 2^0 & \cdot & \cdot & M^0 \\ 1^2 & 2^2 & \cdot & \cdot & M^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1^{2M-2} & 2^{2M-2} & \cdot & \cdot & M^{2M-2} \end{bmatrix} \begin{bmatrix} 1^2 c_1 \\ 2^2 c_2 \\ \cdot \\ \cdot \\ M^2 c_M \end{bmatrix} = \begin{bmatrix} 1 \\ r^2 \\ \cdot \\ \cdot \\ r^{2M-2} \end{bmatrix}. \quad (27)$$

After solving Eq. (27), we can obtain the FD weights from TETS by

$$c_m = \frac{(-1)^{m+1}}{m^2} \prod_{1 \leq n \leq M, n \neq m} \left| \frac{n^2 - r^2}{n^2 - m^2} \right|, m = (1, 2, 3, \dots, M). \quad (28)$$

When $r = 0$, the FD weights obtained by Eq. (28) are the same as Eq. (16). Comparing Eqs. (25) and (26), the spatial error and temporal error can be written as

$$E_{\text{spa}} = \frac{2}{h^2} \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mkh)^{2j}}{(2j)!} \right]. \quad (29)$$

$$E_{\text{tem}} = \frac{2}{\tau^2} \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(\omega\tau)^{2j}}{(2j)!} \right]. \quad (30)$$

When the left side of Eqs. (11) and (12) plus the E_{spa} in Eq. (29) and E_{tem} in Eq. (30), respectively, we obtain the one-dimensional total error formulation for the FD weights from TETS:

$$E_{\text{total}} = -2 \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{r^{2j-2} k^{2j}}{(2j)!} \tau^{2j-2} \right] + 2 \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mk)^{2j}}{(2j)!} h^{2j-2} \right]. \quad (31)$$

Similarly, the two- and three-dimensional total error formulation for the FD weights from TETS can be expressed as

$$E_{\text{total}} = -2 \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{r^{2j-2} k^{2j}}{(2j)!} \tau^{2j-2} \right] + 2 \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mk_x \cos \theta \cos \varphi)^{2j}}{(2j)!} h^{2j-2} \right] + 2 \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mk_y \cos \theta \sin \varphi)^{2j}}{(2j)!} h^{2j-2} \right] + 2 \sum_{m=1}^M c_m \left[\sum_{j=M+1}^{\infty} (-1)^j \frac{(mk_z \sin \theta)^{2j}}{(2j)!} h^{2j-2} \right], \quad (33)$$

where θ is the plane wave propagation angle measured from the horizontal plane perpendicular to z -axis, φ is the azimuth of the plane wave.

The FD scheme with the FD weights from TETS is $2M$ -th order accurate because the minimum power of h in Eq. (31) is $2M$. With these FD weights, the accuracy of the 1D acoustic wave equation is up to $2M$ -th order in Eq. (31), that of the 2D acoustic wave equation is $2M$ -th order in 8 directions in Eq. (32), and that of the 3D acoustic wave equation is $2M$ -th in 48 directions in Eq. (33). Fig. 3(a) illustrates that the TES has less spatial dispersion error than TETS. However, the stability of the FD scheme is determined by the temporal dispersion error. For the temporal dispersion error, the TETS significantly mitigates temporal dispersion errors as shown in Fig. 3(b). Thus, its less temporal error renders it more stable than TES as shown in Fig. 4.

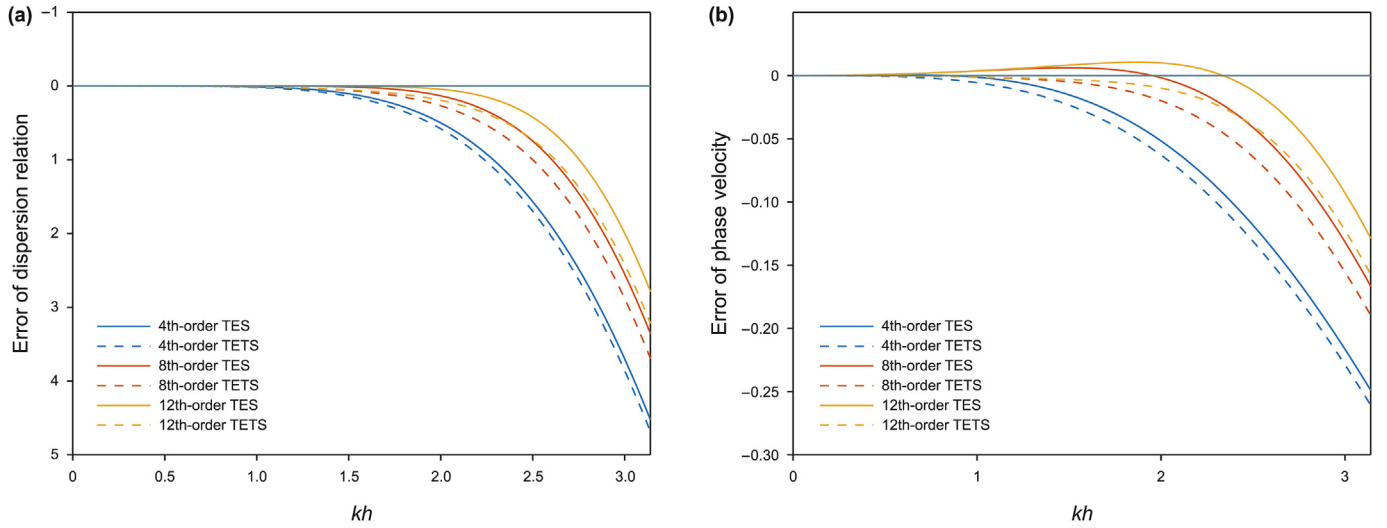


Fig. 3. The comparison of error curve for the different order TETS FD weights.

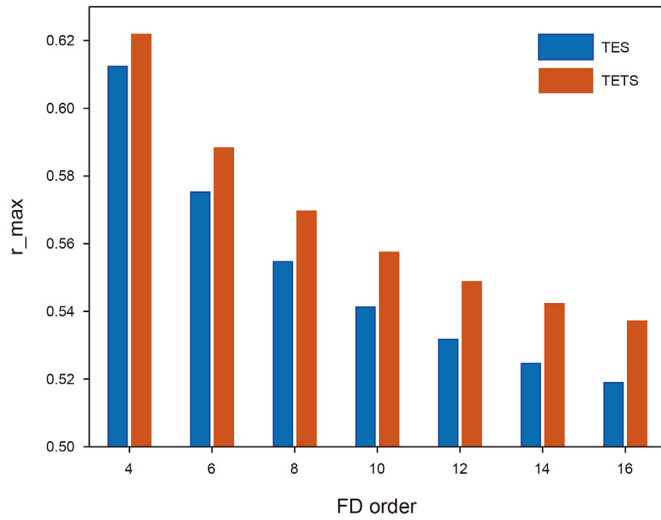


Fig. 4. The comparison of stability for different order FD weights computed by TES and TETS. The $r_{\max}$ represents the stability of the FD simulation. The physical meaning of this parameter has been proven by Liu (2020b) detailedly.

2.2. The optimization algorithm strategy

The optimization algorithm strategy contains many methods, but the scope of this paper focuses on the classical FD weights optimization case in the acoustic wave equation. The application of this strategy, like optimized FD weights in viscoacoustic media, is not considered within the scope of this paper.

The focal point of research in the optimization algorithm field revolves around expanding effective bandwidth and minimizing dispersion errors. Both of these objectives are closely tied to the choice of cost function. Thus, we classify and summarize the existing methods based on the cost function. The cost function shown in this paper is based on the pressure equation.

2.2.1. The L_∞ norm cost function

L_∞ norm cost function is constructed by minimizing the maximum error within a limited bandwidth (Holberg, 1987; Zhang and Yao, 2013a, 2013b; Yang et al., 2017b; He et al., 2019). Holberg (1987) applied the least-squares method to solve the L_∞ norm

optimization problem, whose cost function is minimizing the maximum relative error of the group velocity. The L_∞ norm cost function as presented in Zhang and Yao (2013a, 2013b) aims to minimize the maximum absolute error of dispersion relation in the space domain. This optimization problem is addressed using the simulated annealing method (SAM) in Zhang and Yao (2013a, 2013b). The L_∞ norm cost functions in Zhang and Yao (2013a, 2013b), Yang et al. (2017b) and He et al. (2019) are very similar, as all of these minimize the maximum absolute or relative error of dispersion relation in the space domain.

1) Holberg (1987) case

In Holberg (1987) case, the cost function is constructed based on the group velocity. The group velocity is

$$v_g = k\partial\epsilon/\partial k + \epsilon, \quad (34)$$

with

$$\epsilon = \frac{2 \sum_{m=-M}^M c_m \cos(mkh)}{k^2 h^2} - 1, \quad (35)$$

The definition of v_g is denoted by Koene and Robertsson (2020). For the acoustic equation in a homogenous model, Holberg (1987) pointed out that the minimizes the peak error in group velocity can be considered as cost function, which is

$$\|E\|_\infty = \min \left\{ \max_{0 \leq k \leq K} [k\partial\epsilon/\partial k + \epsilon] \right\}, \quad (36)$$

where K signifies the maximum optimized wavenumber of the differentiator. Holberg (1987) quadrupled relative error within interval $[0, K]$ in Eq. (36), which can be expressed as

$$J = \int_0^K (k\partial\epsilon/\partial k + \epsilon)^4 dk. \quad (37)$$

Eq. (37) is a least-squares problem, which can be effectively solved by using standard numerical techniques, as outlined in works by Wolfe (1978) and Garbow (1980).

2) Kindelan et al. (1990) case

In the Holberg optimization method, the derivative of the $D'(k)$ ($D(k) = \frac{2 \sum_{m=-M}^M c_m \cos(mkh)}{h^2}$) has $M/2$ extreme values $\{k_n\}$ for a $2M$ -th order FD operator. Apart from the fixed extrema at $k = 0$, the other members of the set $\{k_n\}$ depend on the FD weights. So, Kindelan et al. (1990) introduced a novel approach to address the optimization problem by considering the second derivative of $D(k)$,

$$D''(k_n) = 0, n = 2, 3, \dots, M/2. \quad (38)$$

A variant of Newton's method is employed to solve the optimization problem with just a few iterations. Its solution will converge to the FD weights defined by the optimization procedure described in Holberg (1987). Actually, the cost function of this approach can be expressed as

$$\|E\|_\infty = \min[\max(D''(k_n))], (n = 2, 3, \dots, M/2), \quad (39)$$

where the cost function of this methods also is the L_∞ norm.

3) Zhang and Yao (2013a, 2013b) case

The cost function of Zhang and Yao (2013a, 2013b) is illustrated as follows:

$$\|E\|_\infty = \min \left\{ \max_{0 \leq \beta \leq \beta_{\max}} \left| (\beta)^2 - \sum_{m=-M}^M c_m (1 - \cos(m\beta)) \right| \right\}, \quad (40)$$

where the $\beta = kh$, $\beta_{\max}$ is the maximal optimized bandwidth. Zhang and Yao (2013a, 2013b) applied the SAM to compute the cost function based on three criteria as follows: (1) c_m are real valued FD weights, and the $c_m = c_{-m}$; (2) the sum of the c_m should be equal to zero; (3) c_m should be an amplitude of damped oscillation away from the center position ($m = 0$).

4) Yang et al. (2017b) case

The cost function in Yang et al. (2017b) can be expressed as

$$\|E\|_\infty = \min \left\{ \max_{0 \leq \beta \leq \beta_{\max}} \left| \frac{2 \sum_{m=2}^M c_m (1 - \cos(m\beta)) + 2(1 - \cos \beta) + 2m^2 \sum_{m=2}^M c_m (1 - \cos(m\beta))}{\beta^2} - 1 \right| \right\}. \quad (41)$$

Eq. (41) is obtained by substituting $c_1 = 1 - m^2 \sum_{m=2}^M c_m$ into relative error of the dispersion relation in space domain. The optimized FD weights can be computed by using the Remez exchange algorithm (REA) to solve this cost function.

5) He et al. (2019) case

The cost function in He et al. (2019) can be expressed as

$$\|E\|_\infty = \min \left\{ \max_{0 \leq \beta \leq \beta_{\max}} \left| \frac{2 \sum_{m=1}^M c_m (1 - \cos(m\beta))}{\beta^2} - 1 \right| \right\}. \quad (42)$$

The cost function in He et al. (2019) considers the c_1 as a variable, different from that in Yang et al. (2017b). Eq. (42) is the cost function based on the dispersion relation of the second-order spatial derivative. He et al. (2019) also used REA to solve the optimization problem. The L_∞ norm cost functions can be generally expressed as $E = \min \left\{ \max_{0 \leq \beta \leq \beta_{\max}} [F(\beta)] \right\}$. The $F(\beta)$ is listed in Table 1.

In the previous case, the primary objective of using the L_∞ norm cost function is to make the FD weights cover more bandwidth. Zhang and Yao (2013a, 2013b) found that using the L_∞ norm cost function can cover more bandwidth. Yang et al. (2017b) proposed that the FD weights obtained from REA can cover broader bandwidth than those obtained from SAM. Finally, He et al. (2018) further widened the covered bandwidth by modifying the cost function. According to the Chebeshev criterion, the solution from He et al. (2019) is equiripple, thus, the solution has the widest bandwidth.

2.2.2. The L_2 norm cost function

The target of L_2 norm cost function aims to minimize the sum of the squared error within the specified bandwidth. Etgen (2007) constructed the L_2 norm cost function based on the absolute error of phase velocity. In the case of Du et al. (2010), the relative error of dispersion relation in the space domain is used to construct the L_2 norm cost function. Liu (2013, 2014), in addition to using the dispersion relation in the space domain, also employed other dispersion relations to construct the cost function.

1) Etgen (2007) case

The cost function in Etgen (2007) can be expressed as

$$\|E\|_2 = \min \left\{ \sum_k [v_{\text{phase}}(k) - v_{\text{ture}}]^2 \right\}, \quad (43)$$

where the $v_{\text{phase}} = \frac{2v}{rkh} \sin^{-1} \sqrt{r^2 \sum_{m=1}^M c_m \sin^2(mkh/2)}$.

2) Du et al. (2010) case

The cost function in the Du et al. (2010) can be expressed as

Table 1
Expressions of $F(\beta)$, based on the acoustic wave equation.

F	
Absolute error of dispersion relation in space domain for 1D acoustic equation	$\sum_{m=1}^M 2c_m(1 - \cos(m\beta)) - \beta^2$
Relative error of dispersion relation in space domain for 1D acoustic equation	$\sum_{m=1}^M \frac{2c_m(1 - \cos(m\beta))}{\beta^2} - 1$
Relative error of dispersion relation in time-space domain for 1D acoustic equation	$\sum_{m=1}^M c_m \frac{1 - \cos(m\beta)}{r^{-2}(1 - \cos(r\beta))} - 1$
Absolute error of dispersion relation in time-space domain for 1D acoustic equation	$\sum_{m=1}^M 2c_m(1 - \cos(m\beta)) - 2r^{-2}(1 - \cos(r\beta))$
Relative error of dispersion relation in time-space domain for 2D acoustic equation	$\sum_{m=1}^M \frac{c_m(2 - \cos(m\beta \cos \theta) - \cos(m\beta \sin \theta))}{r^{-2}(1 - \cos(r\beta))} - 1$
Absolute error of dispersion relation in time-space domain for 2D acoustic equation	$\sum_{m=1}^M c_m(2 - \cos(m\beta \cos \theta) - \cos(m\beta \sin \theta)) - r^{-2}(1 - \cos(r\beta))$

$$\|E\|_2 = \min \left\{ \int_0^{\beta_{\max}} \left(\frac{\sum_{m=1}^M c_m(1 - \cos(m\beta))}{\beta^2} - 1 \right)^2 d\beta \right\}. \quad (44)$$

The Lagrangian multiplier method, which can be found in the Wang and Wu (2002), is an elegant technique to solve the L_2 optimization problem.

3) Liu (2013, 2014) case

The cost function in Liu (2013, 2014) can be expressed as

$$\|E\|_2 = \min \left\{ \int_0^{\beta_{\max}} F(\beta)^2 d\beta \right\}. \quad (45)$$

Notably, Eq. (45) is the standard model formula in the least-square sense. The $F(\beta)$ in Eq. (45) represents all of the formulations listed in Table 1. Applying the LSM (Least-square method) can obtain the globally optimal solution due to no iteration involved in this process.

4) Yong et al. (2017a, 2017b) case

Yong et al. (2017a, 2017b) showed that solving L_2 optimization problems in an equivalent staggered-grid FD scheme by Newton's method can avoid the inability of LSM to solve nonlinear problems. The cost function is different from the other cost functions referenced above because it is constructed with the stress equation in the equivalent staggered-grid FD scheme. The cost function in the equivalent staggered-grid finite-difference scheme can be written as

$$\|E\|_2 = \frac{1}{2} \sum_{k=0}^K \sum_{\theta=0}^{\pi/4} \sum_{\varphi=0}^{\pi/4} \left\| \frac{r^2 q(c_m)}{\sin^2(0.5\omega\tau)} - 1 \right\|_2^2, \quad (46)$$

where

$$q(c_m) = \left\{ \sum_{m=1}^M c_m \sin[(m - 0.5)k_x h] \right\}^2 + \left\{ \sum_{m=1}^M c_m \sin[(m - 0.5)k_y h] \right\}^2 + \left\{ \sum_{m=1}^M c_m \sin[(m - 0.5)k_z h] \right\}^2, \quad (47)$$

θ is the plane wave propagation angle measured from the

horizontal plane perpendicular to z-axis, and φ is the azimuth of the plane wave.

2.2.3. The L_1 norm cost function

The objective of the L_1 norm cost function is to minimize the summation of the absolute error. Miao and Zhang (2020) formulated the L_1 norm optimization problem based on the dispersion relation in the space domain. They solved the cost function using the alternating direction method of multipliers (ADMM). In their results, both the accuracy analysis and the numerical modeling illustrated that the new method exhibits less dispersion error in the low-wavenumber region, compared with other optimization algorithms.

The relative error of the dispersion relation in the space domain of integral form can be expressed as

$$E_{\text{int}} = \int_0^{\beta_{\max}} \left| \frac{\sum_{m=1}^M c_m(1 - \cos(m\beta))}{\beta^2} - 1 \right| d\beta, \quad (48)$$

where E_{int} means the integral form error. The Riemann sum of Eq. (48) can be written as

$$\|E\|_1 = \min \left\{ \sum_{n=1}^N \left| \frac{\sum_{m=1}^M c_m \left(1 - \cos \left(mn \frac{\beta_{\max}}{N} \right) \right)}{\left(n \frac{\beta_{\max}}{N} \right)^2} - 1 \right| \right\}. \quad (49)$$

Eq. (49) is the cost function in the Miao and Zhang (2020).

2.2.4. Theoretical analysis

We plot the error of dispersion relation in space domain and the error of the phase velocity for three types of FD weights obtained through L_∞ , L_2 and L_1 norm cost function. L_∞ , L_2 and L_1 norm cost functions are solved by REA (He et al., 2019), LSM (Liu, 2013) and ADMM (Miao and Zhang, 2020), respectively.

As shown in Fig. 5(a), within the same bandwidth (0, 2.094), the maximal error generally appears in the high-wavenumbers region. It's noteworthy that among all the mentioned cost functions, only the L_∞ norm cost function effectively manages to control this maximum error. Thus the L_∞ norm cost function has the least error at the high-wavenumber region among all cost functions. On the contrary, L_∞ norm cost function is less effective at reducing errors in the low-wavenumber region, as evident from the boxed region in Fig. 5(a). Furthermore, under the same error limitation (-0.0001 , 0.0001), the L_∞ norm cost function covers the widest bandwidth, compared with other cost functions. The L_1 norm cost function has the least dispersion error in the low-wavenumber region compared with other methods, as depicted in Fig. 5(a). The L_2 norm cost

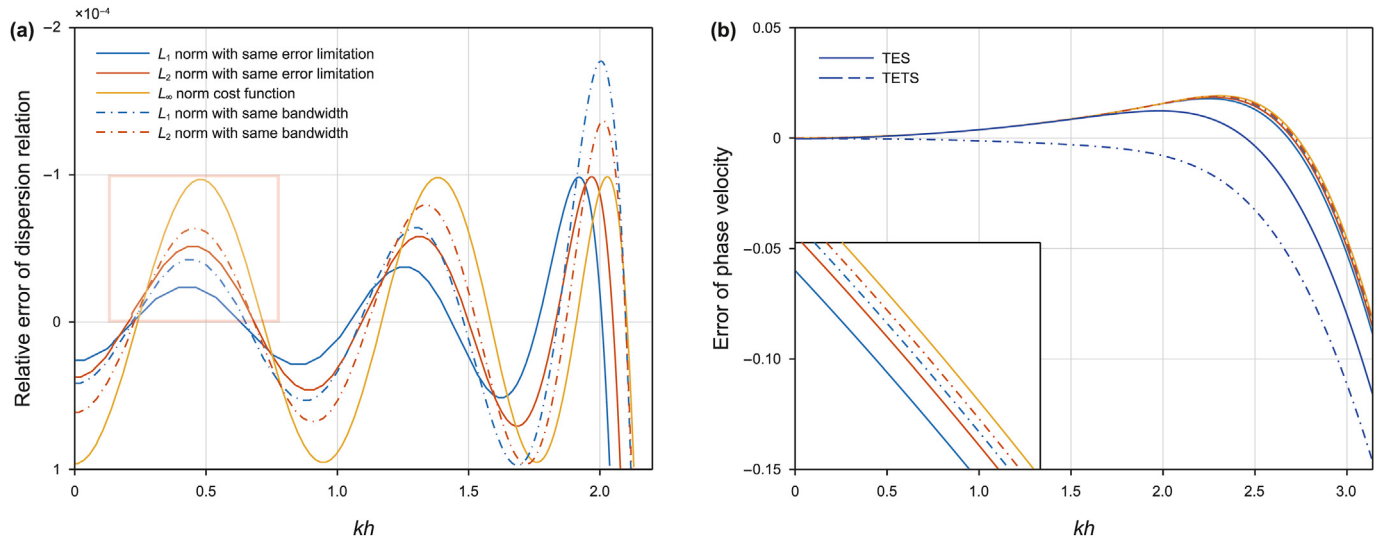


Fig. 5. The comparison of error curves for the same order FD weights computed by L_∞ norm cost function (He et al., 2019), L_2 norm cost function (Liu, 2013, 2014), and L_1 norm cost function (Miao and Zhang, 2020) under the same error limitation or same effective bandwidth.

function offers moderate performance in terms of broadening bandwidth and reducing dispersion error.

Most of the previously discussed cost functions primarily suppress spatial dispersion errors and do not consider any temporal errors. Thus, in the error curves of phase velocity, as shown in Fig. 5(b), it becomes evident that the optimization strategies suffer a larger temporal dispersion error than TES and TETS.

Throughout the history of development for the cost function, geophysicists have strived to construct cost functions using better error functions. In addition to the method of classification by norm type, the error function can also be classified into three types, the group velocity cost function, the phase velocity cost function, and the dispersion relation cost function. A good error function must have a remarkable physical meaning, simple expression, and less calculational cost.

Compared with group velocity, the phase velocity has some advantages as the cost function. Firstly, the error of phase velocity is the source of all numerical errors. Secondly, the error of phase velocity is more sensitive to the error in the low-wavenumber region than the group velocity. Thirdly, the error of phase velocity contains the error of temporal and spatial dispersion error.

The cost function relying on phase velocity error has a significant drawback—it can lead to instability in its solution process because it contains inverse trigonometric functions. The cost functions based on the error of the dispersion relation are more stable and easier to calculate. The solution obtained through the cost function based on the dispersion relation tends to outperform the group velocity, as detailed by Koene and Robertsson (2020). Thus, for the optimization problem, the error of the dispersion relation is the best choice to construct the cost function.

2.3. The window function strategy

The window function strategy is an approach to obtain FD weights by using the window function to truncate the spatial convolution series of the pseudo-spectral method. Fornberg (1987) proved that the FD scheme has the same accuracy as the pseudo-spectral method when the order of the FD operator tends to infinity. When the order of the FD operator is fixed, the FD scheme can be considered as a truncated convolutional counterpart of the pseudo-spectral method. For this reason, truncating the spatial

convolution series of the pseudo-spectral method is an alternative approach to obtain the FD weights.

The finite series of uniformly sampled signals can be interpolated accurately using the following formulation (Smith, 2010):

$$u_0^0 = \sum_{m=-\infty}^{+\infty} u_m^0 \frac{\sin[(\pi/h)(x - mh)]}{(\pi/h)(x - mh)}. \quad (50)$$

The second-order spatial derivative can be found by differentiating both sides of Eq. (50):

$$\frac{\partial^2 u_0^0}{\partial x^2} = \sum_{m=-\infty}^{+\infty} \left\{ \left[\frac{2}{h(x - mh)^3} - \frac{\pi}{h} \frac{1}{x - mh} \right] \sin \left[\frac{\pi}{h} (x - mh) \right] - \frac{2}{(x - mh)^2} \cos \left[\frac{\pi}{h} (x - mh) \right] \right\} u_m^0. \quad (51)$$

When we are interested in calculating the differentiations at $x = 0$ only, Eq. (51) can be rewritten as

$$\frac{\partial^2 u_0^0}{\partial x^2} = \frac{1}{h^2} \sum_{m=-\infty}^{+\infty} \left[-\frac{2}{m^2} \cos(m\pi) \right] u_m^0. \quad (52)$$

The FD weights are determined by applying the $2M + 1$ point window function to truncate Eq. (52), where $2M$ represents the order of the FD operator, and the finite-difference formulation can be written as

$$\frac{\partial^2 u_0^0}{\partial x^2} = \frac{1}{h^2} \sum_{m=-N/2}^{N/2} w(m) \left[-\frac{2}{m^2} \cos(m\pi) \right] u_m^0, \quad (53)$$

where $w(m)$ is window function. There are numerous window functions for computing FD weights in the previous studies.

2.3.1. Single window function

1) Kaiser window function case

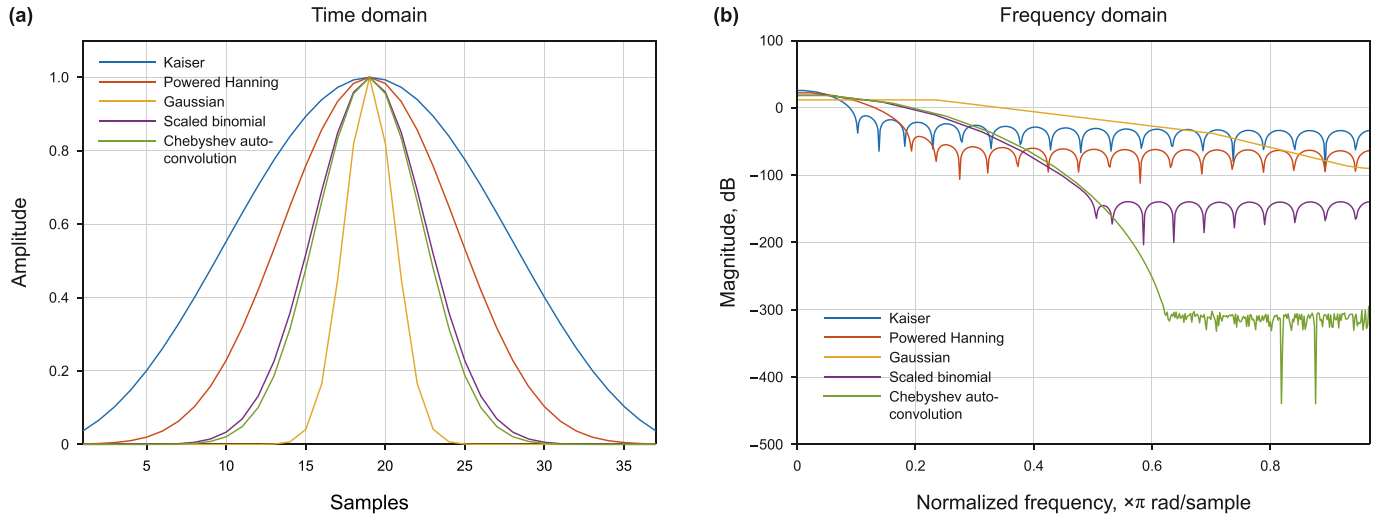


Fig. 6. The comparison of amplitude responses for different window functions.

$$w(m) = \frac{I_0\left(\beta_{\text{kai}} \sqrt{1 - \left(\frac{m}{M}\right)^2}\right)}{I_0(\beta_{\text{kai}})}, -M \leq m \leq M, \quad (54)$$

where I_0 is the zeroth-order modified Bessel function of the first kind. If the sidelobe attenuation is α_{kai} dB, the β_{kai} is computed by

$$\beta_{\text{kai}} = \begin{cases} 0.1102(\alpha_{\text{kai}} - 8.7), & \alpha_{\text{kai}} > 50 \\ 0.5842(\alpha_{\text{kai}} - 21)^{0.4} + 0.07886(\alpha_{\text{kai}} - 21), & 50 \geq \alpha_{\text{kai}} \geq 21 \\ 0, & \alpha_{\text{kai}} < 21 \end{cases} \quad (55)$$

Increasing β_{kai} is an efficient way to widen the main lobe and decrease the amplitude of the sidelobes. The amplitude response for this window function is shown in Fig. 6.

2) Generalized-power Hanning window function case (Zhou and Greenhalgh, 1992a,b)

$$w(m) = \left[2\alpha_{\text{han}} - 1 + 2(1 - \alpha_{\text{han}}) \cos^2 \frac{\pi m}{2(M+2)} \right]^{\beta_{\text{han}}/2}, -M \leq m \leq M \quad (56)$$

The α_{han} and β_{han} allow a family of different window functions to be considered, like the conventional Hanning window function is determined by the $\alpha_{\text{han}} = 0.5$ and $\beta_{\text{han}} = 2$. The amplitude response for this window function is also portrayed in Fig. 6.

The conventional Hanning window function is expressed as

$$w(m) = 0.5 \left(1 - \cos \frac{\pi(m+M)}{M} \right), -M \leq m \leq M. \quad (57)$$

3) Gaussian window function case (GWF, Igel et al., 1995)

$$w(m) = \exp \left\{ - \frac{\left[\frac{2m}{(M-1)\beta_{\text{gau}}} \right]^2}{2} \right\}, -M \leq m \leq M. \quad (58)$$

In order to reduce the edge effects and the Gibbs phenomenon, Igel et al. (1995) computed the FD weights using a Gaussian window function. The amplitude response for this window function is shown in Fig. 6.

4) Scaled binomial window function case (SBWF, Chu and Stoffa, 2012)

$$w^{N+L}(m) = \frac{\binom{2M+L}{(2M+L)/2+m}}{\binom{2M+L}{(2M+L)/2}}, -M \leq m \leq M, \quad (59)$$

where

$$\binom{m}{l} = \frac{m!}{l!(m-l)!}, \quad (60)$$

is the binomial weights formula and $l \leq m$, $L \geq 0$ is an even number. When $L = 0$, the FD weights from the scaled binomial windows will reduce to the conventional FD weights from TES. The amplitude response for this window function is shown in Fig. 6.

2.3.2. Combined window functions

1) Chebyshev auto-convolution combined window function case (CACWF, Wang et al., 2015).

Diniz et al. (2012) concluded that a narrow main lobe can result in a wide bandwidth for FD weights, while a significant attenuation of the side lobe can lead to low dispersion error. According to this, Wang et al. (2015) proposed the Chebyshev auto-convolution combined window function:

$$w(m) = \lambda w_c(m) + (1 - \lambda) w_c^Q(m), \quad (61)$$

where, $w_c(m)$ is the Chebyshev window function, $w_c^Q(m)$ signifies the Chebyshev window function, Q denotes the auto-convolution times, $w(m)$ is a combined window function, λ represents the weight coefficient. The amplitude response of the Chebyshev auto-convolution combined window function exhibits a narrower main lobe and greater attenuation in the side lobe compared to other single window functions, as shown in Fig. 6.

2) Optimum combined window function case (Wang et al., 2017a)

Based on Diniz's theory, Wang et al. (2017a) introduced a novel combined window function with improved amplitude response. In their work, Wang et al. (2017a) combined the Hanning, cosine, Kaiser, and binomial window functions. When $N + 1 = 5$, the proposed window function can be expressed as

$$w(m) = 0.7\sqrt{w_h(m)w_c(m)} + 0.3w_c(m), \quad (62)$$

where $w_h(m)$ and $w_c(m)$ denotes the Hanning and cosine window function, respectively. $w(m)$ is the combined window function.

When $N + 1 = 9$, the proposed window function can be expressed as

$$w(m) = 0.3\sqrt{w_h(m)w_b(m)} + 0.7w_c(m), \quad (63)$$

where $w_b(m)$ denotes the binomial window function.

When $N + 1 = 13, 17, 25$, the proposed window function can be expressed as

$$w(m) = \sqrt{w_h(m)w_c(m)w_k(m)}, \quad (64)$$

where, $w_k(m)$ denotes the Kaiser window function. The window function expressions of this method are only given with $N = 4, 8, 12, 16, 24$, and these expressions don't show any regularity. Thus, it is hard to apply in industrial production for practical application.

3) Cosine-modulated window function case (Wang et al., 2017b)

$$w(m) = \left\{ \frac{w_b(m)}{\cos^T\left(\frac{\pi}{P}n\right)} + \frac{w_b(m)}{\left[\cos^T\left(\frac{\pi}{P}n\right)\right]^{-1}} \right\}^{-1}, \quad (65)$$

where P denotes the modulated range, T denotes the modulated times. This window function is a combination of the binomial window function. The amplitude response of this window function outperforms that of the binomial window function, as demonstrated in Wang et al. (2017b).

4) Improved cosine-combined window function case (Li et al., 2021)

$$w(m) = \left[w_{\cos 1}(m)^a \times w_{\cos 2}(m)^b \right]^{\frac{1}{a+b}}, \quad (66)$$

where the $w_{\cos 1}(m)$ and $w_{\cos 2}(m)$ denote the combined cosine window (1) and window (2), respectively. The equations of the two window functions can be written as

$$w_{\cos 1}(m) = w_{\cos 2}(m) = \sum_{l=0}^{L-1} (-1)^l a_l \times \cos\left(\frac{\pi l}{M}(m+M)\right), -M \leq m \leq M. \quad (67)$$

where L is the number of terms of the window function. The

parameters in this case are provided only for specific values of N , such as $N = 4, 8, 12, 16$, and 24 , as discussed by Li et al. (2021). The Optimum combined window function provides the value of the parameters only when $N = 4, 8, 12, 16$, and 24 . Consequently, the improved cosine-combined window function encounters a similar issue as the Optimum combined window function.

The evolutive process of the window function strategy involves searching for a window function with an improved amplitude response to cover more bandwidth or mitigate dispersion error. In the historical development of window function strategy, there are two inherent problems as follows. Firstly, the shape of the window function is highly related to the parameters input, for example, β_{kai} in Kaiser window function. However, there is no unified or practical criterion for determining these parameters. The second problem is that many combined window functions proposed in recent years lack a standardized expression, which limits their applications.

3. Discussion

We choose some widely used methods to show the advantages and disadvantages of three strategies for computing FD weights. These methods are SBWF (it is an efficient single window function), GWF (it is a classical single window function), CACWF (it is a combined window function with unified formula), ADMM (it is newly and only method to solve L_1 norm optimization problems), LSM (it is a classic method to solve L_2 norm optimization problems), REA (it solves L_∞ norm optimization problems to make FD weights cover the widest bandwidth), TES and TETS. The FD scheme in the frequency domain is widely applied, and we compared these eight types of FD weights with the FD scheme in the frequency domain by using the error of the phase velocity in Appendix A.

We use three properties to test the FD weights from these eight methods, including stability, effective bandwidth, and dispersion error. The specific computation scheme for these properties is as follows.

1) Stability

The maximal Courant-Friedrichs-Lewy (CFL) number represents the stability of the FD simulation, which has been proven by Liu (2020b). For the 1D acoustic wave equation, the maximal CFL number is calculated by

$$r_{\max} = \left(2 \sum_{m=1}^{\text{int}((M+1)/2)} c_{2m-1} \right)^{-1/2}, \quad (68)$$

where the $r_{\max}$ is defined as maximal CFL number. The requirement of a stable FD scheme is that the $v_{\max}\tau/h$ is smaller than the maximal CFL number.

The maximal CFL number can also be determined by the maximum error of the $\sum_{m=1}^M 2c_m(1 - \cos(m\beta))$, and $\beta \in [0, \pi]$, which has also been proven by Liu (2020b). According to this, Liu (2020b) constructs a complex L_∞ norm optimization problem, whose cost function minimizes the maximum error of dispersion relation in region $[0, \pi]$.

2) Bandwidth and dispersion error

We can use the relative error of dispersion relation in the space domain to compute the effective bandwidth and dispersion error. The expression of this error function can be expressed as

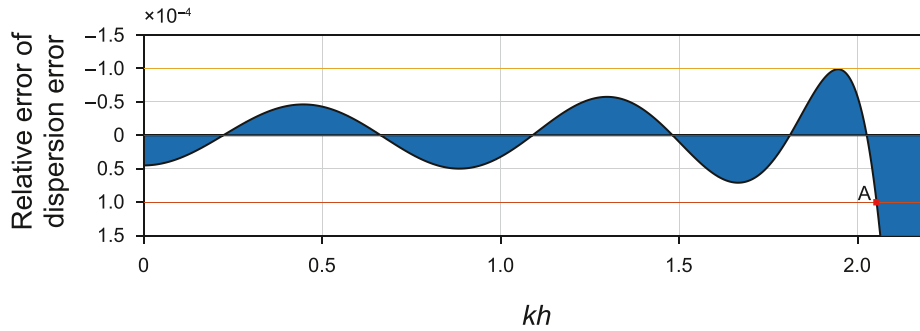


Fig. 7. The relative error of dispersion relation in the space domain for the optimized FD weights obtained by LSM.

$$E = \frac{\sum_{m=1}^M 2c_m(1 - \cos(m\beta))}{\beta^2} - 1. \quad (61)$$

The dispersion error represents the total dispersion error within the bandwidth. For some methods, their dispersion error is adjustable. Thus, only when the comparison of the total dispersion error under the same effective bandwidth, the results of the comparison are convincing. However, some methods may struggle to regulate the dispersion error, necessitating the use of an alternative criterion. Mean dispersion error is a suitable option. Mean dispersion error is the ratio of total dispersion error to bandwidth, which portrays the mean dispersion error at the bandwidth.

We use Fig. 7 to explain how to compute the dispersion error and bandwidth. The bandwidth is the wavenumber at point A. Point A is the last intersection of the dispersion curve and the error limitation. The error limitation is $[-0.0001, 0.0001]$. The total dispersion error is the sum of the blue area within the bandwidth. The mean dispersion error is the total dispersion error divided by the wavenumbers of point A.

In addition to these properties, two method properties, the tunability and calculated amount, are employed to evaluate the feasibility of each method. Tunability refers to the method's ability to adjust the bandwidth and dispersion error by modifying its parameters. If a method is tunable, it illustrates that the method is suitable for a wide range of application scenarios. Tunability also indicates that a method can do tests under the same conditions (same bandwidth or error limitation). These superiorities are beneficial to the application and development of this method. The calculated amount is measured by consuming time of computing FD weights.

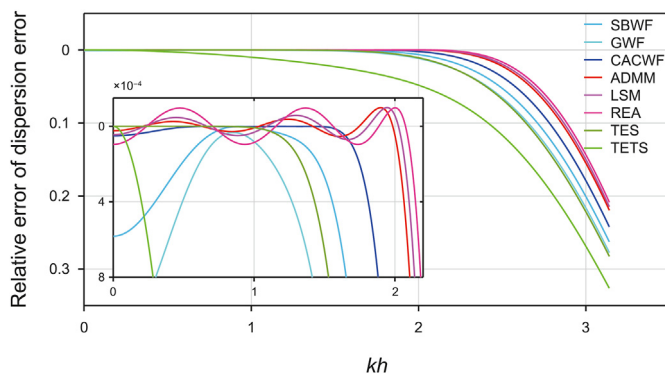


Fig. 8. The comparison of relative error curves of dispersion relation for eight methods.

Window function strategy, in general, tends to have limited tunability. For most window functions, the dispersion error and bandwidth rely primarily on empirical knowledge from previous research. This phenomenon is due to the fact that the parameters associated with window functions often lack well-defined physical meanings. The window function strategy is generally low in computational cost. Because the process of this strategy typically involves a single convolution calculation without the need for iterations or matrix inversions.

For the window function strategy, the width of the main lobe and attenuation of sidelobes play a crucial role in determining the bandwidth and dispersion error. The single window function, however, is not good on both sides. The combined window function overcomes these shortcomings. Thus, the combined window function widens the bandwidth and decreases the mean dispersion error as shown in Fig. 8 and Table 2. To conduct fair comparison, we rely on the mean dispersion error since window functions generally have limited tunability. The limited tunability means that it is hard to adjust the parameters to change the dispersion error, due to the unclear physical meaning of these parameters in the window function.

Optimization algorithm strategies possess wider bandwidth and less dispersion error, compared with other strategies as shown in Fig. 8 and Table 2. For the L_∞ cost function, the L_2 cost function, and the L_1 cost function, the bandwidth and the dispersion error are decreasing in order under the same given error limitation as shown in Fig. 5. On the other hand, Table 2 and Fig. 8 illustrate that the optimization algorithm strategy has the poorest stability among the three strategies.

For the optimization strategy, the error limitation and optimized bandwidth can control the dispersion error and the bandwidth. A large value of these parameters leads to a wide bandwidth and a large dispersion error. The computed amount of the optimization algorithm strategy is the highest among the three strategies, primarily because the solving process involves multiple loops and iterations. The LSM proposed by Liu (2013, 2014) has the lowest computation cost among all optimization algorithm methods because it doesn't require any iteration and loop.

Within all optimization methods, the L_1 norm cost function (ADMM) covers the narrowest bandwidth and has the smallest dispersion error. Compared with the L_∞ norm cost function (REA), the bandwidth covered by the L_1 norm cost function (ADMM) decreases by 4.14%, while the mean dispersion error reduces by 58.4% as shown in Table 2. This significant improvement in the L_1 norm cost function is beneficial for suppressing the accumulated error in seismic modeling.

In all eight methods, TES has the smallest dispersion error and the stability of TETS is the most stable. When the order of the FD operator is fixed, the Taylor expansion strategy can compute one set

Table 2
The comparison of result and properties test. The value in brackets at third column is the mean dispersion error, and the value out brackets at third column is the total dispersion error.

	FD weights properties test			Method properties test	
	Bandwidth	Dispersion error, $\times 10^{-4}$	Stability (r_max)	Tunability	Calculated amount
SBWF	0.758	2.438637 (3.22)	0.524584174	Yes	0.0003594
GWF	0.331	1.841127 (5.56)	0.529864062	Yes	0.0001216
CACWF	1.694	2.253587 (1.33)	0.517421136	Yes	0.0006447
ADMM	2.013 (4.14%)	5.346340 (2.66) (58.4%)	0.509797435	Yes	7.0739712
LSM	2.052	7.663371 (3.73)	0.508241814	Yes	0.0212667
REA	2.100	12.842696 (6.12)	0.506247286	Yes	0.1162639
TES	1.255	1.052343 (0.84)	0.531759239	No	0.0002728
TETS	0.100	0.338343 (3.38)	0.548777813	No	0.0002016

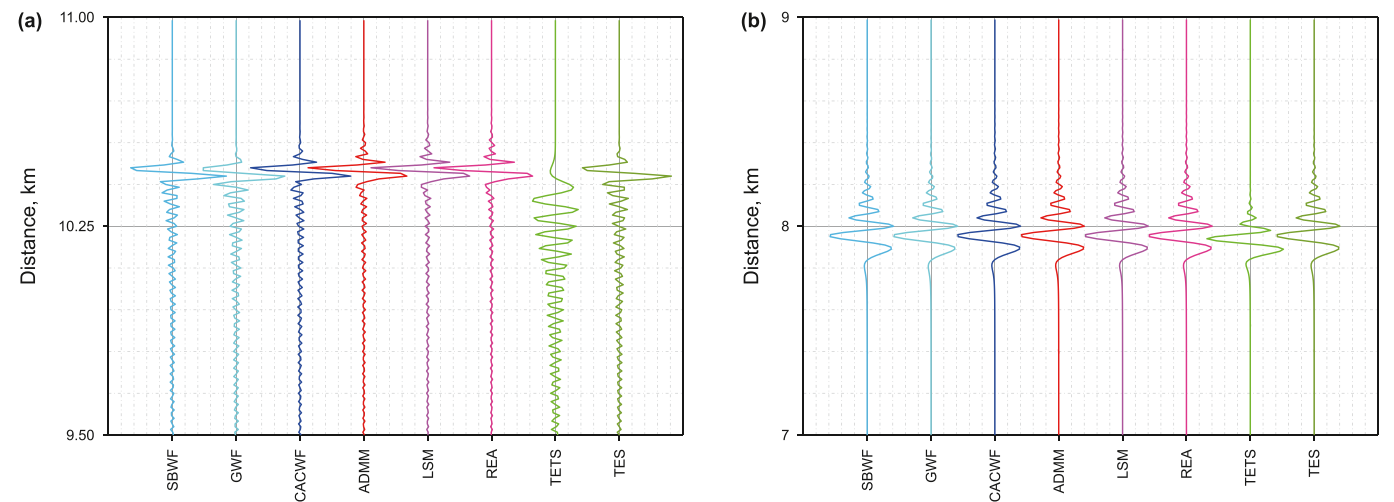


Fig. 9. The comparison of spatial dispersion error (a) and temporal dispersion error (b) for different strategies.

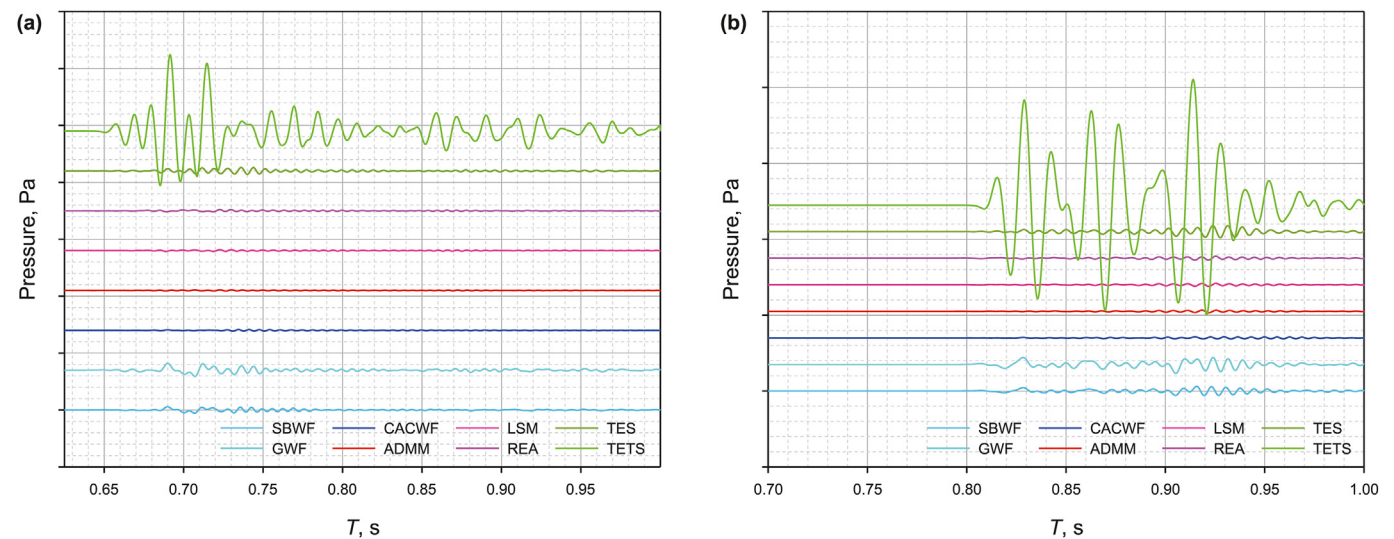


Fig. 10. The comparison of the residual error for different FD weights in inhomogeneous modeling with high-frequency wavelet. (a) is the waveform recorded by point 1 in Fig. (13), and (b) is the waveform recorded by point 2 in Fig. (13). The ADMM has the least residual in this modeling.

of FD weights. Thus, this strategy is not tunable. The calculated amount of the Taylor expansion strategy is almost the smallest because its computational process only involves a calculation of a matrix inverse.

We focus on a homogeneous model to do spatial dispersion and

temporal dispersion scenarios. Using the homogeneous model aims to present a dispersive waveform of FD modeling results clearly. Then, we use the inhomogeneous model to test the accuracy of FD weights. The experiment in an inhomogeneous model injects wavelets with different dominant frequencies (20, 40, 60 Hz) to test

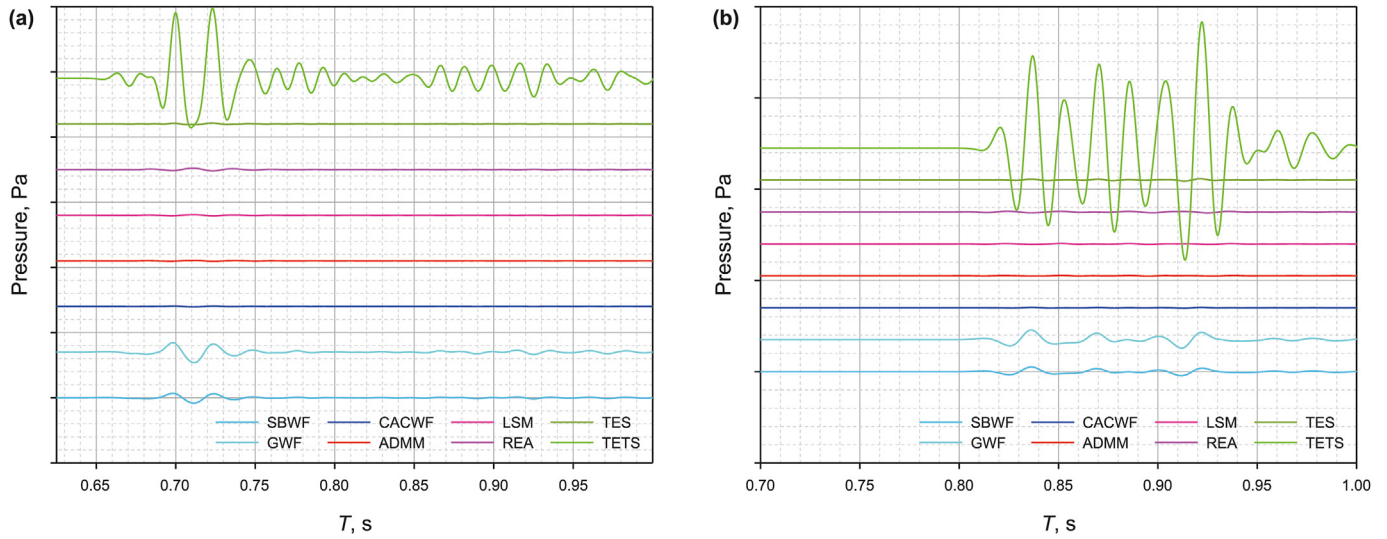


Fig. 11. The comparison of the residual error for different FD weights in inhomogeneous modeling with middle-frequency wavelet. (a) is the waveform recorded by point 1 in Fig. (13), and (b) is the waveform recorded by point 2 in Fig. (13). The ADMM has the least residual in this modeling.

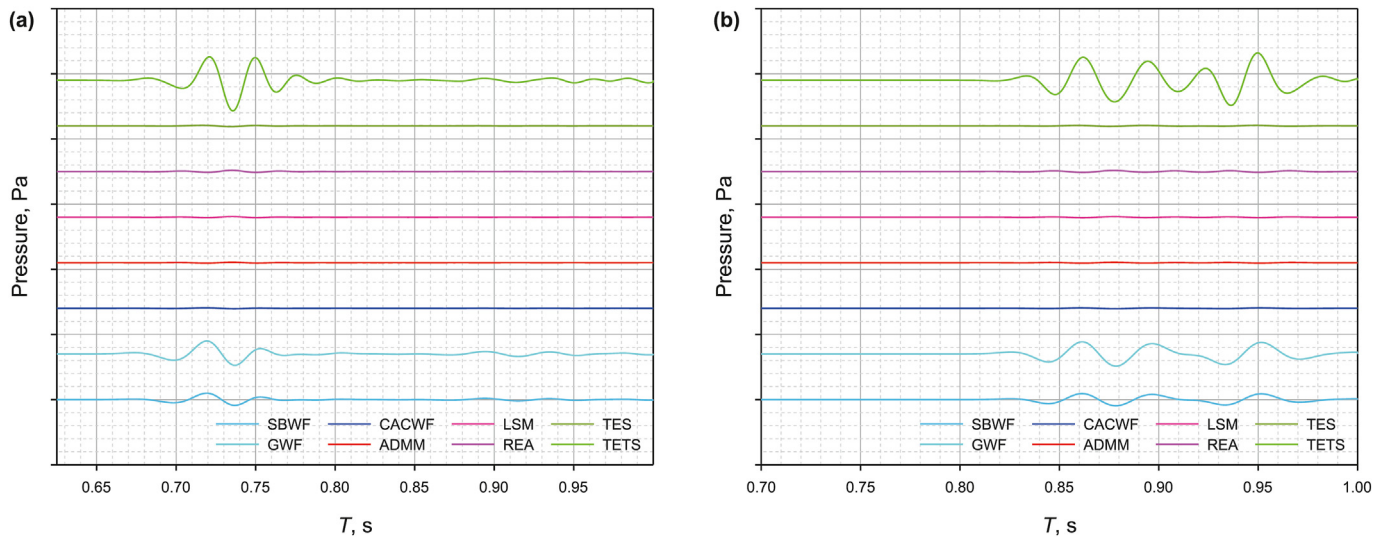


Fig. 12. The comparison of the residual error for different FD weights in inhomogeneous modeling with low-frequency wavelet. (a) is the waveform recorded by point 1 in Fig. (13), and (b) is the waveform recorded by point 2 in Fig. (13). The ADMM has the least residual in this modeling.

the feasibility of the FD weights. The specific modeling parameters are listed in Appendix B.

In the numerical simulation, the spatial dispersion is considered as the waveform oscillation that occurred after the main lobe of the wavelet, due to the velocity of spatial dispersion being slower than the velocity of the main lobe of the wavelet. Because the velocity of the temporal dispersion is faster than the velocity of the main lobe of the wavelet, the temporal dispersion appears before the main lobe of the wavelet. The position of spatial dispersion and temporal dispersion are illustrated in detail by Jin et al. (2022).

We observe that temporal dispersion and spatial dispersion indeed appear before and after the main lobe of the wavelet, respectively in Fig. 9. As shown in Fig. 9(a), the TETS causes the largest spatial dispersive waveform; the spatial dispersion of single window function methods (GWF, SBWF) is nearly the same as TES, meanwhile, it is larger than that of combined window function (CACWF); the optimization algorithm methods (LSM, ADMM, REA) have the least amplitude of waveform oscillation.

TETS has the best simulation result, when the temporal sampling τ is equal to 0.0025 s, as shown in Fig. 9(b). For these eight methods, TETS is the only method computed in the time-space domain, which considers the effect of temporal derivative. Thus, TETS has the least temporal dispersive waveform. From this observation, we can conclude that reducing the temporal dispersion is efficiently obtained by computing the FD weights in the time-space domain.

Actually, we observe that the three methods (LSM, ADMM, REA) have almost the same waveform in the homogeneous model test, as shown in Fig. 9. However, for simulating seismic wave propagation in the inhomogeneous model as shown in Figs. 10–12, the accumulated numerical error in the simulative results is determined by the dominant frequency and the dispersion error of FD weights. When using the low dominant frequency wavelet to simulate, the accumulated numerical error is caused by the dispersion error in the low wavenumber region. Less dispersion error in the low wavenumber region will cause less accumulated numerical error.

Thus, the ADMM can provide the modeling result with the least residual error when the dominant wavelet has a low frequency. Therefore, geophysicists pay attention to reducing the dispersion error in the low-wavenumber region.

From these modeling experiments and precision analyses, the combined window function is a very good research direction in the window function strategy. CACWF covers a wide bandwidth and has a good suppression of spatial dispersion as shown in Fig. 9(a). However, some problems still exist. Firstly, the bandwidth and dispersion error are hardly controlled by the input parameters of window functions. Secondly, the bandwidth requires further improvement. Thirdly, the FD weights from the window function strategy aim to reduce the spatial dispersion error, not involve the temporal dispersion error. Thus, the window function strategy is not a good direction to develop the FD weights.

Under a given error limitation and a fixed order of FD operators, the REA technique can maximize the covered bandwidth (He et al., 2019). Thus, expanding the bandwidth is not a feasible ideal to improve FD weights properties. There are two potential research directions in this field. The first one is to improve the stability by computing good FD weights, where Liu (2020b) has made a certain contribution. The other one is to obtain FD weights with a small dispersion error in the low wavenumber region. About the high computation cost of this strategy, try to apply the algorithm with few iterations or use the global solving technique, like LSM, which can avoid the increase in the computation cost.

4. Conclusion

Over the past decade, the application of computing FD weights proves that these strategies are useful tool for analyzing seismic wave simulation, especially when the problems are involved in large-scale or multi-scale structures. We provided a historical overview of three kinds of computing FD weights strategies in detail. Three strategies include the Taylor expansion, the optimization algorithm, and the window function. Besides the historical overview, the possibilities for ongoing and future work for these methods are also discussed based on the experiment of dispersion analysis and numerical modeling. Apparently, widening the covered bandwidth to improve the simulation result is researched exhaustively, which has been proven mathematically. How to reduce the dispersion error in the low-wavenumber region, and how to improve the FD stability have become research hotspots in the field of FD weights.

CRediT authorship contribution statement

Jian-Ping Huang: Resources, Project administration, Conceptualization. **Wei-Ting Peng:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ji-Dong Yang:** Writing – review & editing. **Lu-Feng Lou:** Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study is supported by the Marine S&T Fund of Shandong Province for Pilot National Laboratory for Marine Science and Technology (Qingdao) (No. 2021QNL020001), the Major Scientific

and Technological Projects of Shandong Energy Group (No. SNKJ2022A06-R23), the Funds of Creative Research Groups of China (No. 41821002), National Natural Science Foundation of China Outstanding Youth Science Fund Project (Overseas) (No. ZX20230152), the Major Scientific and Technological Projects of CNPC (No. ZD2019-183-003).

Appendix A

The dispersion analysis for finite-difference method in frequency domain

The previous analyses are based on the FD scheme in the time-space domain. However, the finite-difference method has also been widely applied in the frequency domain. The FD scheme in the frequency domain is proposed by Pratt (1990), which suffers a strong dispersion error. This method must maintain 13 mesh points within a wavelength to reduce the error of phase velocity to 1%. Then, Jo et al. (1996) proposed the improving 9-point FD scheme. For reducing the error of phase velocity to 1%, the needed number of mesh points becomes 4. Shin and Sohn (1998) based on the improving 9-point FD scheme, introduced the 25-point FD scheme and then optimized this FD scheme by rotating the coordinate system. The number of mesh points reduces to 2.5.

For the most accurate rotation 25-point FD scheme, when the error of phase velocity should reduce to 1%, a wavelength should contain 2.5 mesh points. As this method an example, the relation of the wavelength between the wavenumber can be shown as

$$\lambda = \frac{1}{k}, \quad (\text{A-1})$$

where the k represents the wavenumber, λ is the wavelength. When the $\lambda > 2.5h$, the error of the phase velocity is less than 1%, which can be expressed as

$$kh < \frac{1}{2.5} = 0.4. \quad (\text{A-2})$$

This expression illustrated that the error of the phase velocity is less than 1% for this method when the kh is less than 0.4. When the kh is larger, the method requires fewer mesh points in a wavelength to limit the error of phase velocity to 1%, which indicates the method is more accurate. For the method reference above, when the error of the phase velocity is less than 1%, the corresponding wavenumbers for all these eight methods are listed in Table 4. The observation of Table 4 illustrates that the FD scheme in the time-space domain is generally better than the FD scheme in the frequency domain in reducing the error of phase velocity.

Table 4

The comparison of the value of kh for all method using the error of the phase velocity.

	SBWF	GWf	CACWF	ADMM	LSM	REA	TES	TETS
kh	1.641	1.808	1.613	1.614	1.614	1.613	1.706	2.002

Appendix B

The parameters and test formula's in all tests

Fig. 5 is the relative error of dispersion relation (REDR) in the space domain and error of phase velocity (EPV) in the space domain can be expressed respectively as

Table 3

The FD weights obtained by eight methods.

	SBWF	GWF	CACWF	ADMM	LSM	REA	TES	TETS
c_1	1.750000000	1.721415953	1.787106643	1.825256814	1.832710227	1.842148015	1.714285714	1.614967967
c_2	−0.291666667	−0.274405818	−0.316649531	−0.344513944	−0.350290956	−0.357719918	−0.267857143	−0.228854630
c_3	0.064814815	0.057608947	0.077383895	0.093472219	0.097162289	0.102042503	0.052910053	0.044430302
c_4	−0.013257576	−0.011339744	−0.017724142	−0.024704801	−0.026556607	−0.029120580	−0.008928571	−0.007452744
c_5	0.002121212	0.001881420	0.003100036	0.005136489	0.005788126	0.006752513	0.001038961	0.000864830
c_6	−0.000226625	−0.000250921	−0.000303595	−0.000600282	−0.000718040	−0.000906402	−0.000060125	−0.000049973

$$REDR \approx \frac{2 \sum_{m=1}^M c_m (1 - \cos(m\beta))}{\beta^2} - 1, \quad (B-1)$$

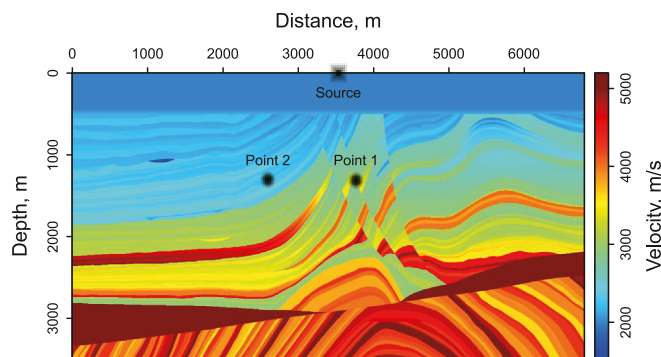
$$EPV = \frac{2}{rkh} \sin^{-1} \sqrt{r^2 \sum_{m=1}^M c_m \sin^2(mkh/2)} - 1, \quad (B-2)$$

Fig. 6 is plotted by the function (wvtool), in the matlab R2016b. $\beta = 5$ in the Kaiser window function. The $\beta = 0.54$, $\alpha = 6$ in the generalized-power Hanning window function, $\beta = 0.2$ in the Gaussian window function, the $L = 12$ in the scaled binomial window function, $L = 6$, $r = 60$, $\lambda = 0.89$ in Chebyshev auto-convolution combined window function involve the calculation of curve in Fig. 6.

In Fig. 9(a), we choose a velocity $v = 2000$ m/s, for 4000×4000 nodes with $h = 10$ m from 0 up to 40 km, choose $\tau = 0.0005$ s and model for 5 s, use a Ricker wavelet with a 20 Hz peak frequency, the source injected at $20 \text{ km} \times 20 \text{ km}$.

In Fig. 9(b), we choose a velocity $v = 2000$ m/s, for 4000×4000 nodes with $h = 10$ m from 0 up to 40 km, choose $\tau = 0.0025$ s and model for 3 s, use a Ricker wavelet with a 40 Hz peak frequency, the source injected at $20 \text{ km} \times 20 \text{ km}$. The FD weights used in Fig. 9 is listed in Table 3.

Figs. 10–12 are recorded waveforms at point 1 and point 2 in the modified Marmousi model, whose positions are shown in Fig. 13. This model is 6.8 km wide and 3.5 km deep, which is dispersed by a 1360×700 nodes grid. We inject the Ricker wavelet with 20, 40, and 60 Hz dominant frequencies in the surface center of the modified Marmousi model, respectively. The time step $\tau = 0.00025$ s and the seismic wave propagate 1 s. Note that the waveform in Figs. 10–12 is the difference between the waveform of optimized FD weights and the waveform of 20th order conventional FD weights.

**Fig. 13.** The modified Marmousi model and the position of point 1, point 2 and source.

References

Abubakar, A., Pan, G., Li, M., et al., 2011. Three-dimensional seismic full-waveform inversion using the finite-difference contrast source inversion method. *Geophysical Prospecting* 59(Modelling Methods for Geophysical Imaging:

- Trends and Perspectives) 874–888. <https://doi.org/10.1111/j.1365-2478.2011.00953.x>.
- Alford, R.M., Kelly, K.R., Boore, D.M., 1974. Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics* 39 (6), 834–842. <https://doi.org/10.1190/1.1440470>.
- Alterman, Z., Karal, J.F.C., 1968. Propagation of elastic waves in layered media by finite difference methods. *Bull. Seismol. Soc. Am.* 58 (1), 367–398. <https://doi.org/10.1785/BSSA0580010367>.
- Bathe, K.J., Wilson, E.L., 1976. *Numerical Methods in Finite Element Analysis*. NJ Prentice-Hall, Englewood Cliffs. <https://doi.org/10.1002/nme.1620110913>.
- Belytschko, T., Mullen, R., 1978. On dispersive properties of finite element solutions. In: *Modern Problems in Elastic Wave Propagation*, vols. 67–82. https://sc.panda321.com/citations?view_op=view_citation&hl=zh-CN&user=BgbXTjUAAAAJ&citation_for_view=BgbXTjUAAAAJ:qjMakFHDy7sC.
- Book, D.L., Boris, J.P., Hain, K., 1975. Flux-corrected transport II: generalizations of the method. *J. Comput. Phys.* 18 (3), 248–283. [https://doi.org/10.1016/0021-9991\(75\)90002-9](https://doi.org/10.1016/0021-9991(75)90002-9).
- Boris, J.P., Book, D.L., 1973. Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works. *J. Comput. Phys.* 11 (1), 38–69. [https://doi.org/10.1016/0021-9991\(73\)90147-2](https://doi.org/10.1016/0021-9991(73)90147-2).
- Brebbia, C.A., 1978. *The Boundary Element Method for Engineers*. Pentech Press Ltd, Plymouth. [https://doi.org/10.1016/0141-1187\(81\)90114-0](https://doi.org/10.1016/0141-1187(81)90114-0). <https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=PASCAL7980467579>.
- Cerjan, C., Kosloff, D., Kosloff, R., et al., 1985. A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics* 50 (4), 705–708. <https://doi.org/10.1190/1.1441945>.
- Chen, D., Liang, W.Q., Xin, W., et al., 2016. Acoustic wave equation modeling based on implicit finite difference operators in the time-space domain. *Chin. J. Geophys.* 59 (4), 1491–1497. <https://doi.org/10.6038/cjg20160429> (in Chinese).
- Chen, Y.C., Tong, M.S., Mitra, R., 1997. Efficient and accurate finite-difference time-domain analysis of resonant structures using the Blackman-Harris window function. *Microw. Opt. Technol. Lett.* 15 (6), 389–392. [https://doi.org/10.1002/\(SICI\)1098-2760\(19970820\)15:6<389::AID-MOP14>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1098-2760(19970820)15:6<389::AID-MOP14>3.0.CO;2-Y).
- Chu, C.L., Stoffa, P.L., 2012. Determination of finite-difference weights using scaled binomial windows. *Geophysics* 77 (3), W17–W26. <https://doi.org/10.1190/geo2011-0336.1>.
- Chu, C.L., Stoffa, P.L., Seif, R., 2009. High-order Rotated Staggered Finite Difference Modeling of 3D Elastic Wave Propagation in General Anisotropic Media. *SEG Technical Program Expanded Abstracts*, pp. 291–295. <https://doi.org/10.1190/1.3255458>, 2009.
- Dablain, M.A., 1986. The application of high-order differencing to the scalar wave equation. *Geophysics* 51 (1), 54–66. <https://doi.org/10.1190/1.1442040>.
- Dines, K.A., Lytle, R.J., 1979. Computerized geophysical tomography. *Proc. IEEE* 67 (7), 1065–1073. <https://doi.org/10.1109/PROC.1979.11390>.
- Diniz, P.S.R., da Silva, E.A.B., Netto, S.L., 2012. *Digital Signal Processing System Analysis and Design*. China Machine Press, Beijing.
- Du, Q.Z., Bai, Q.Y., Li, B., 2010. Numerical simulation of seismic wavefield by optimizing differential coefficient method for transversely isotropic media. *Oil Geophys. Prospect.* 45 (2), 170–176. <https://doi.org/10.13810/j.cnki.issn.1000-7210.2010.02.005> (in Chinese).
- Etgen, J.T., 2007. A tutorial on optimizing time domain finite-difference schemes: "Beyond Holberg". *Stanford Exploration Project* 129, 33–43.
- Fornberg, B., 1975. On a Fourier method for the integration of hyperbolic equations. *SIAM J. Numer. Anal.* 12 (4), 509–528. <https://doi.org/10.1137/0712040>.
- Fornberg, B., 1987. The pseudospectral method: comparisons with finite differences for the elastic wave equation. *Geophysics* 52, 483–501. <https://doi.org/10.1190/1.1442319>.
- Garbow, J., 1980. Modeling of the acoustic wave equation with transform methods. *Geophysics* 46, 854–859. <https://library.seg.org/doi/10.1190/1.1441223>.
- Gu, B.L., Huang, J.P., Han, J.G., et al., 2021. Least-squares inversion-based elastic reverse time migration with PP- and PS-angle-domain common-imaging gathers. *Geophysics* 86 (1), S29–S44. <https://doi.org/10.1190/geo2020-0028.1>.
- Gupta, R.N., 1966. Reflection of plane elastic waves from transition layers with arbitrary variation of velocity and density. *Bull. Seismol. Soc. Am.* 56 (3), 633–642. <https://doi.org/10.1785/BSSA0560030633>.
- He, Z., Zhang, J.H., Yao, Z.X., 2019. Determining the optimal weights of the explicit finite-difference scheme using the Remez exchange algorithm. *Geophysics* 84 (3), S137–S147. <https://doi.org/10.1190/geo2018-0446.1>.
- Holberg, O., 1987. Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave

- phenomena. *Geophys. Prospect.* 35, 629–655. <https://doi.org/10.1111/j.1365-2478.1987.tb00841.x>.
- Huang, J.P., Li, C., Wang, R., et al., 2015. Plane-wave least-squares reverse time migration for rugged topography 26 (4), 471–480. <https://doi.org/10.1007/s12583-015-0556-5>.
- Huang, J.P., Yang, J.D., Liao, W.Y., et al., 2016. Common-shot Fresnel beam migration based on wave-field approximation in effective vicinity under complex topographic conditions. *Geophys. Prospect.* 64 (3), 554–570. <https://doi.org/10.1111/1365-2478.12276>.
- Igel, H., Mora, P., Riollot, B., 1995. Anisotropic wave propagation through finite-difference grids. *Geophysics* 60 (4), 1203–1216. <https://doi.org/10.1190/1.1443849>.
- Jin, K.J., Huang, J.P., Zou, Q., et al., 2022. Optimization of staggered grid finite-difference weights based on conjugate gradient method. *J. Seismic Explor.* 31, 33–52.
- Jo, C.I.I., Shin, C.S., Suh, J.H., 1996. An optimal 9-point, finite-difference, frequency-space 2-D scalar wave extrapolator. *Geophysics* 61 (2), p529–p537. <https://doi.org/10.1190/1.1443979>.
- Julian, B.R., Gubbins, D., 1977. Three-dimensional seismic ray tracing. *J. Geophys.* 43 (1), 95–113. <https://geophysicsjournal.com/article/133>.
- Kelly, K.R., Ward, R.W., Treitel, S., et al., 1976. Synthetic seismograms: a finite-difference approach. *Geophysics* 41 (1), 2–27. <https://doi.org/10.1190/1.1440605>.
- Kim, J.W., Lee, D.J., 1996. Optimized compact finite difference schemes with maximum resolution. *AIAA J.* 34 (5), 887–893. <https://doi.org/10.2514/3.13164>.
- Kim, S., Lim, H., 2007. High-order schemes for acoustic waveform simulation. *Appl. Numer. Math.* 57 (4), 402–414. <https://doi.org/10.1016/j.apnum.2006.05.003>.
- Kindelan, M., Kamel, A., Sguazzero, P., 1990. On the construction and efficiency of staggered numerical differentiators for the wave equation. *Geophysics* 55 (1), 107–110. <https://doi.org/10.1190/1.1442763>.
- Koene, E.F.M., Robertsson, J.O.A., 2020. Optimal finite-difference operators for arbitrarily sampled data. *Geophysics* 85 (3), F39–F51. <https://doi.org/10.1190/geo2019-00811>.
- Kosloff, D., Pestana, R.C., Tal-Ezer, H., 2010. Acoustic and elastic numerical wave simulations by recursive spatial derivative operators. *Geophysics* 75 (6), T167–T174. <https://doi.org/10.1190/1.3485217>.
- Kreiss, H.O., Olliger, J., 1972. Comparison of accurate methods for the integration of hyperbolic equations. *Tellus* 24 (3), 199–215. <https://doi.org/10.3402/tellusa.v24i3.10634>.
- Lele, S.K., 1992. Compact finite difference schemes with spectral-like resolution. *J. Comput. Phys.* 103 (1), 16–42. [https://doi.org/10.1016/0021-9991\(92\)90324-R](https://doi.org/10.1016/0021-9991(92)90324-R).
- Levander, A.R., 1988. Fourth-order finite-difference P-SV seismograms. *Geophysics* 53, 1425–1436. <https://doi.org/10.1190/1.1442422>.
- Li, W.D., Meng, X.H., Liu, H., et al., 2021. Optimal finite-difference schemes for elastic wave based on improved cosine-combined window function. *Explor. Geophys.* 52 (2), 221–234. <https://doi.org/10.1080/08123985.2020.1801344>.
- Liu, H., Wang, Z.Y., 2015. Optimized staggered-grid finite-difference operators based on combined window. In: Near-Surface Asia Pacific Conference 83–87. <https://doi.org/10.1190/NSAPC2015-021>.
- Liu, Y., 2013. Globally optimal finite-difference schemes based on least squares. *Geophysics* 78 (4), T113–T132. <https://doi.org/10.1190/geo2012-0480.1>.
- Liu, Y., 2014. Optimal staggered-grid finite-difference schemes based on least-squares for wave equation modelling. *Geophys. J. Int.* 197 (2), 1033–1047. <https://doi.org/10.1093/gji/ggu032>.
- Liu, Y., 2020a. Acoustic and elastic finite-difference modeling by optimal variable-length spatial operators. *Geophysics* 85 (2), T57–T70. <https://doi.org/10.1190/geo2019-0145.1>.
- Liu, Y., 2020b. Maximizing the CFL number of stable time-space domain explicit finite-difference modeling. *J. Comput. Phys.* 416, 1–17. <https://doi.org/10.1016/j.jcp.2020.109501>.
- Liu, Y., 2022. Removing the stability limit of the time-space domain explicit finite-difference schemes for acoustic modeling with stability condition-based spatial operators. *Geophysics* 87 (3), T205–T223. <https://doi.org/10.1190/geo2021-0141.1>.
- Liu, Y., Sen, M.K., 2009a. A new time-space domain high-order finite-difference method for the acoustic wave equation. *J. Comput. Phys.* 228 (23), 8779–8806. <https://doi.org/10.1016/j.jcp.2009.08.027>.
- Liu, Y., Sen, M.K., 2009b. A practical implicit finite-difference method: examples from seismic modelling. *J. Geophys. Eng.* 6 (3), 231–249. <https://doi.org/10.1088/1742-2132/6/3/003>.
- Liu, Y., Sen, M.K., 2010. Acoustic VTI modeling with a time-space domain dispersion-relation-based finite-difference scheme. *Geophysics* 75 (3), A11–A17. <https://doi.org/10.1190/1.3374477>.
- Liu, Y., Sen, M.K., 2013. Time-space domain dispersion-relation-based finite-difference method with arbitrary even-order accuracy for the 2D acoustic wave equation. *J. Comput. Phys.* 232 (1), 327–345. <https://doi.org/10.1016/j.jcp.2012.08.025>.
- Madariaga, R., 1976. Dynamics of an expanding circular fault. *Bull. Seismol. Soc. Am.* 66 (3), 639–666. <https://doi.org/10.1785/BSSA0660030639>.
- Mao, W., Stuart, G.W., 1997. Rapid multi-wave-type ray tracing in complex 2-D and 3-D isotropic media. *Geophysics* 62 (1), 298–308. <https://doi.org/10.1190/1.1444131>.
- Miao, Z.Z., Zhang, J.H., 2020. Reducing error accumulation of optimized finite-difference scheme using minimum norm. *Geophysics* 85 (5), T275–T291. <https://doi.org/10.1190/geo2019-0758.1>.
- Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics* 52 (9), 1211–1228. <https://doi.org/10.1190/1.1442384>.
- Mora, P., 1988. Elastic wave-field inversion of reflection and transmission data. *Geophysics* 53 (6), 750–759. <https://doi.org/10.1190/1.1442510>.
- Mu, X., Huang, J.P., Yang, J.D., et al., 2020. Least-squares reverse time migration in TTI media using a pure qP-wave equation. *Geophysics* 85 (4), S199–S216. <https://doi.org/10.1190/geo2019-0320.1>.
- Orszag, S.A., 1972. Comparison of pseudospectral and spectral approximation. *Stud. Appl. Math.* 51 (3), 253–259. <https://doi.org/10.1002/sapm1972513253>.
- Plessix, R.E., Mulder, W.A., 2004. Frequency-domain finite-difference amplitude-preserving migration. *Geophys. J. Int.* 157 (3), 975–987. <https://doi.org/10.1111/j.1365-246X.2004.02282.x>.
- Pratt, R.G., 1990. Frequency-domain elastic wave modeling by finite differences: a tool for cross-hole seismic imaging. *Geophysics* 55 (5), 626–632. <https://doi.org/10.1190/1.1442874>.
- Pratt, R.G., Song, Z.M., Williamson, P., et al., 1996. Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophys. J. Int.* 124 (2), 323–340. <https://doi.org/10.1111/j.1365-246X.1996.tb07023.x>.
- Ren, Z.M., Li, Z.C., 2017. Temporal high-order staggered-grid finite-difference schemes for elastic wave propagation. *Geophysics* 82 (5), T207–T224. <https://doi.org/10.1190/geo2017-0005.1>.
- Ren, Z.M., Li, Z.C., 2019. High-order temporal and implicit spatial staggered-grid finite-difference operators for modelling seismic wave propagation. *Geophys. J. Int.* 217 (2), 844–865. <https://doi.org/10.1093/gji/ggz059>.
- Ren, Z.M., Liu, Y., 2015. Acoustic and elastic modeling by optimal time-space-domain staggered-grid finite-difference schemes. *Geophysics* 80 (1), T17–T40. <https://doi.org/10.1190/geo2014-0269.1>.
- Shao, Z.H., Wei, G.W., Zhao, S., 2003. DSC time-domain solution of Maxwell's equations. *J. Comput. Phys.* 189 (2), 427–453. [https://doi.org/10.1016/S0021-9991\(03\)00226-2](https://doi.org/10.1016/S0021-9991(03)00226-2).
- Shin, C.S., Sohn, H.J., 1998. A frequency-space 2-D scalar wave extrapolator using extended 25-point finite-difference operator. *Geophysics* 63 (1), 289–296. <https://doi.org/10.1190/1.1444323>.
- Smith, J.O., 2010. *Physical Audio Signal Processing*. W3K Publishing.
- Tan, S.R., Huang, L.J., 2014. An efficient finite-difference method with high-order accuracy in both time and space domains for modelling scalar-wave propagation. *Geophys. J. Int.* 197 (2), 1250–1267. <https://doi.org/10.1093/gji/ggu077>.
- Tarantola, A., 1984. Linearized inversion of seismic reflection data. *Geophys. Prospect.* 32 (6), 998–1015. <https://doi.org/10.1111/j.1365-2478.1984.tb00751.x>.
- Virieux, J., Farra, V., 1991. Ray tracing in 3-D complex isotropic media: an analysis of the problem. *Geophysics* 56 (12), 2057–2069. <https://doi.org/10.1190/1.1443018>.
- Wang, J., Meng, X.H., Liu, H., et al., 2017a. Optimization of finite difference forward modeling for elastic waves based on optimum combined window functions. *J. Appl. Geophys.* 138, 62–71. <https://doi.org/10.1016/j.jappgeo.2017.01.005>.
- Wang, J., Meng, X.H., Liu, H., et al., 2017b. Cosine-modulated window function-based staggered-grid finite-difference forward modeling. *Appl. Geophys.* 14 (1), 115–124. <https://doi.org/10.1007/s11770-017-0596-y>.
- Wang, Y., Wu, R.S., 2002. An optimal spatial finite-difference operator which reduces truncation error to a minimum. *Adv. Atmos. Sci.* 19 (3), 468–486. <https://doi.org/10.1007/s00376-002-0080-2>.
- Wang, Y.F., Liang, W.Q., Nashed, Z., et al., 2014. Seismic modeling by optimizing regularized staggered-grid finite-difference operators using a time-space-domain dispersion-relationship-preserving method. *Geophysics* 79 (5), T277–T285. <https://doi.org/10.1190/geo2014-0078.1>.
- Wang, Z.Y., Liu, H., Tang, X.D., et al., 2015. Optimized finite-difference operator based on Chebyshev auto-convolution combined window function. *Chin. J. Geophys.* 58 (2), 192–206. <https://doi.org/10.1002/cjg2.20166>.
- Wannamaker, P.E., Hohmann, G.W., Sanfilippo, W.A., 1984a. Electromagnetic modeling of three-dimensional bodies in layered earths using integral equations. *Geophysics* 49 (1), 60–74. <https://doi.org/10.1190/1.1441562>.
- Wannamaker, P.E., Hohmann, G.W., Ward, S.H., 1984b. Magnetotelluric responses of three-dimensional bodies in layered earths. *Geophysics* 49 (9), 1517–1533. <https://doi.org/10.1190/1.1441777>.
- Wolfe, M.A., 1978. *Numerical Methods for Unconstrained Optimization*. Van Nostrand Reinhold.
- Xiao, F., Tang, X.H., 2006. Zhang XJ. Comparison of Taylor finite difference and window finite difference and their application in FDTD. *J. Comput. Appl. Math.* 193 (2), 516–534. <https://doi.org/10.1016/j.cam.2005.05.030>.
- Yan, H.Y., Liu, Y., 2013. Acoustic prestack reverse time migration using the adaptive high-order finite-difference method in time-space domain. *Chin. J. Geophys.* 56 (3), 971–984. <https://doi.org/10.6038/cjg20130325> (in Chinese).
- Yang, L., Yan, H.Y., Liu, H., 2017a. An optimal implicit staggered-grid finite-difference scheme based on the modified Taylor-series expansion with minimax approximation method for elastic modeling. *J. Appl. Geophys.* 138, 161–171. <https://doi.org/10.1016/j.jappgeo.2017.01.020>.
- Yang, L., Yan, H.Y., Liu, H., 2017b. Optimal staggered-grid finite-difference schemes based on the minimax approximation method with the Remez algorithm. *Geophysics* 82 (1), T27–T42. <https://doi.org/10.1190/geo2016-0171.1>.
- Yong, P., Huang, J.P., Li, Z.C., et al., 2016. Optimized staggered-grid finite-difference method in time-space domain based on exact time evolution schemes. *Chin. J. Geophys.* 59 (11), 4223–4233. <https://doi.org/10.6038/cjg20161124> (in Chinese).
- Yong, P., Huang, J.P., Li, Z.C., et al., 2017a. Optimized equivalent staggered-grid FD method for elastic wave modelling based on plane wave solutions. *Geophys. J.*

- Int. 208, 1157–1172. <https://doi.org/10.1093/gji/ggw447>.
- Yong, P., Huang, J.P., Li, Z.C., et al., 2017b. Forward modeling by optimized equivalent staggered-grid finite-difference method for time-space domain. *Journal of China University of Petroleum (Edition of Natural Science)* 41 (6), 71–79. <https://doi.org/10.3969/j.issn.1673-5005.2017.06.008> (in Chinese).
- Yong, P., Liao, W.Y., Huang, J.P., et al., 2019. Misfit function for full waveform inversion based on the Wasserstein metric with dynamic formulation 399 (15), 108911. <https://doi.org/10.1016/j.jcp.2019.108911>.
- Zhang, J.H., Yao, Z.X., 2013a. Optimized explicit finite-difference schemes for spatial derivatives using maximum norm. *J. Comput. Phys.* 250, 511–526. <https://doi.org/10.1016/j.jcp.2013.04.029>.
- Zhang, J.H., Yao, Z.X., 2013b. Optimized finite-difference operator for broadband seismic wave modeling. *Geophysics* 78 (1), A13–A18. <https://doi.org/10.1190/geo2012-02771>.
- Zheng, W.Q., Meng, X.H., Liu, J.H., et al., 2016. High precision elastic wave equation forward modeling based on cosine modulated Chebyshev window function. *Chin. J. Geophys.* 59 (7), 2650–2662. <https://doi.org/10.6038/cjg20160728> (in Chinese).
- Zhou, B., Greenhalgh, S.A., 1992a. Seismic scalar wave equation modeling by a convolutional differentiator. *Bull. Seismol. Soc. Am.* 82 (1), 289–303. <https://doi.org/10.1785/BSSA0820010289>.
- Zhou, B., Greenhalgh, S.A., 1992b. Seismic scalar wave equation modeling by a convolutional differentiator. *Bull. Seismol. Soc. Am.* 82 (1), 289–303. <https://doi.org/10.1785/BSSA0820010289>.
- Zhu, F., Huang, J.P., Yu, H., 2018. Least-squares Fourier finite-difference pre-stack depth migration for VTI media. *J. Geophys. Eng.* 15 (2), 421–437. <https://doi.org/10.1088/1742-2140/aa9a0a>.
- Zou, Q., Huang, J.P., Yong, P., et al., 2020a. 3D elastic waveform modeling with an optimized equivalent staggered-grid finite-difference method. *Petrol. Sci.* 17, 967–989. <https://doi.org/10.1007/s12182-020-00477-3>.
- Zou, Q., Huang, J.P., Yong, P., et al., 2020b. Plane wave optimization difference operator elastic wave reverse time migration. *Oil Geophys. Prospect.* 55 (5), 1047–1059. <https://doi.org/10.13810/j.cnki.issn.1000-7210.2020.05.013> (in Chinese).