ARTICLE

Mitigating numerical dispersion in full-waveform inversion imaging: A preconditioned optimization approach for finite-difference weights

Ganglin Lei[1,2,3,4,5], Jianping Huang[6]*, Wensheng Duan[1,2,3,4,5], Chao Chen[1,2,3,4,5], Zhenwen Liu[1,2,3,4,5], Chang Zhou[1,2,3,4,5], and Weiting Peng[6] (ORCID)

[1]R&D Center for Ultra-Deep Complex Reservoir Exploration and Development, China National Petroleum Corporation, Korla, Xinjiang, China

[2]Xinjiang Engineering Research Center for Ultra-deep Complex Reservoir Exploration and Development, China National Petroleum Corporation, Korla, Xinjiang, China

[3]Xinjiang Key Laboratory of Ultra-deep Oil and Gas, China National Petroleum Corporation, Korla, Xinjiang, China

[4]China National Energy Administration, State Energy Key Laboratory of Carbonate Oil and Gas, Korla, Xinjiang, China

[5]Key Laboratory of Carbonate, China National Petroleum Corporation, Korla, Xinjiang, China

[6]School of Geosciences, China University of Petroleum (East China), Qingdao, Shandong, China

Abstract

Full-waveform inversion (FWI) imaging is a high-resolution seismic imaging technique that directly produces subsurface images by inverting the full recorded wavefield. However, its reliability is often limited by numerical dispersion errors arising from finite-difference (FD) forward modeling. One key approach for reducing dispersion is to optimize the FD coefficients using an optimization algorithm. However, conventional methods for optimizing FD weights focus only on reducing spatial dispersion, which can weaken numerical stability, especially when using large time steps (i.e., high Courant–Friedrichs–Lewy [CFL] numbers). To address this issue, we introduce a new optimization approach that improves both simulation accuracy and stability. The proposed method combines error functions from both the time–space domain and the spatial domain into a single adaptive objective function. A dynamic weighting factor, which depends on the CFL number, facilitates a trade-off between accuracy and stability of the optimal FD weights. We also use the seismic wavelet spectrum as prior information to constrain the optimization. The optimization problem is solved by the least-squares method. In the theoretical test, the proposed weights significantly reduce wavefield simulation errors across a wide range of wavenumbers, with a higher CFL number than conventional approaches. When applied to FWI, this method reduces phase distortion and local minima in the objective function. In a test using the Marmousi model at 40 Hz, our approach produced clear and continuous deep structures, closely matching results from dispersion-free benchmarks. In contrast, conventional methods failed due to severe dispersion. This work provides a more robust numerical foundation for high-frequency FWI imaging by improving both accuracy and stability.

Keywords: Finite-difference weights; Full-waveform inversion; Full-waveform inversion imaging; Dispersion error

1. Introduction

Full-waveform inversion (FWI), initially proposed by Lailly[1] and Tarantola,[2] and extensively reviewed by Virieux and Operto,[3] is an elegant and natural way to handle full-wavefield data with a proper modeling engine. The power of FWI not only enhances velocity model building but also opens new avenues for seismic imaging.[4-6] By leveraging the full wavefield and high-resolution physical property models recovered by FWI, it becomes possible to derive reflectivity and structural information directly from these models, offering an alternative to conventional migration techniques.[7,8]

FWI imaging has advanced rapidly in recent years. Lu[9] demonstrated that the one-dimensional derivative of a high-resolution FWI velocity model yields a reflectivity estimate that delivers enhanced imaging compared to conventional techniques. To better image dipping events, Huang *et al.*[10] applied a three-dimensional derivative aligned with the normal to the dominant dip direction. In a different approach, Operto *et al.*[11] performed a direct reflectivity inversion within a multiparameter FWI framework.

However, the highly non-linear nature of the FWI objective function makes it highly susceptible to local minima, with "cycle-skipping" being a central challenge to its convergence. To mitigate this issue, various strategies have been developed, primarily focusing on expanding the inversion search space and enhancing the convexity of the objective function. For instance, reflection FWI circumvents cycle-skipping by separating tomographic and migration modes, preferentially using reflected waves to construct long-wavelength background velocity models.[12] Furthermore, physics-based low-frequency extrapolation or the use of deep learning networks to directly build initial models have proven effective in enlarging the convergence basin of FWI.[13] Alternatively, extending the inversion search space beyond the single-scattering Born approximation provides a rigorous mathematical framework to mitigate the non-linearity issue.[14] On the other hand, designing more convex objective functions, such as envelopes or the Wasserstein distance, can mathematically reduce the risk of converging to local minima. In recent years, deep learning-based velocity inversion methods have shown great potential by learning complex non-linear mappings from data to models, bypassing the local minima issues inherent in conventional FWI.[13] Concurrently, stable numerical simulation forms the cornerstone of implementing these advanced algorithms. In this context, absorbing boundary conditions compatible with the finite-difference (FD) method, such as the residual perfectly matched layer, are crucial for suppressing boundary reflections and ensuring a clean wavefield.[15] Nevertheless, a fundamental and often overlooked problem remains: if the forward modeling engine itself contains significant errors due to numerical dispersion, it introduces non-physical, spurious local minima into the inversion process, which can undermine or even negate the effectiveness of the aforementioned advanced inversion techniques.

The accuracy of the numerical forward modeling in the inversion process is crucial for the reliability of the FWI results. Among various numerical methods, the FD scheme has become the predominant choice for large-scale FWI applications due to its computational efficiency and straightforward implementation.[16,17] FD simulations are often affected by numerical dispersion, which can significantly reduce the quality of the synthetic wavefield. The inherent errors of the FD scheme arise from the truncation of higher-order terms in the FD approximation, leading to spatial discretization errors.[18,19] This dispersion phenomenon becomes increasingly severe at higher frequencies. Unfortunately, reflectivity imaging—essentially a high-frequency inversion process—requires the numerical simulation of high-frequency seismic wavefields during FWI imaging. Such a high-frequency wavefield simulation tends to exacerbate numerical dispersion.

To address these problems, a common strategy to mitigate this dispersion error is the optimization of FD weights with broad bandwidth coverage.[20-22] Various optimization approaches have been developed to compute FD weights. One approach involves optimization under different norms, including the L_∞-norm,[20-29] the L_2-norm,[30-36] and the L_1-norm.[37] Another strategy uses window functions to truncate the wavefield's pseudospectrum for coefficient derivation.[22,38-42]

An ideal FD scheme would be entirely free from numerical dispersion across all frequencies, perfectly preserving the dispersion relation. Although this ideal is not yet attainable, two practical criteria are widely considered: maximizing the bandwidth within a given error tolerance, or minimizing the error within a given bandwidth.[28] Due to its equiripple characteristic, the optimized FD weights delivered by the Remez exchange algorithm can maximize the bandwidth under a given error tolerance.[27,28,43,44] In contrast, due to its L_1-norm objective function, the alternating direction method of multipliers approach is effective at minimizing the numerical error within a given bandwidth.[37]

However, since conventional optimization of FD weights does not account for the influence of simulation parameters, the resulting "optimized" weights may, in some cases, introduce even greater simulation inaccuracies. This

phenomenon indicates a fundamental limitation: existing methods minimize the dispersion relation error rather than the actual wavefield simulation error. Although Peng and Huang[45] incorporated wavefield information to improve FD simulations by reducing the spatial dispersion relation error, their approach overlooked numerical stability considerations. Optimization schemes based solely on the spatial dispersion relation often reduce the stability of the FD simulation. Moreover, in their computational process, they used a double-loop mechanism combined with an exhaustive search method to compute the optimized weights, which significantly increased computational resource consumption.

Furthermore, practical applications frequently face challenges related to stability constraints, especially when spatial or temporal sampling rates exceed the stability condition. Most of these optimization approaches focus solely on minimizing the error of the spatial dispersion relation. This way of optimizing the FD weights often compromises the stability of the FD method and fails to meet the demands of practical applications. In a previous study, Liu and Sen[46] introduced a method that uses the error in the spatial–temporal dispersion relation as the objective function to optimize FD weights, which was shown to extend the stability limit of the FD scheme.

In this study, we propose a novel optimization method for FD weights that combines the dispersion relation in the space domain and the time–space domain. By incorporating wavefield simulation parameters as prior information, we formulate an L_2-norm constrained optimization problem and obtain a global solution via the least-squares method. Through theoretical analysis and numerical tests, we demonstrate the advantages of the proposed method in reducing dispersion error and enhancing stability. Furthermore, we investigate the impact of FD dispersion errors on the FWI objective function and subsequently evaluate the FWI imaging performance using both the proposed and conventional methods on simple and complex models. Our results demonstrate that the new method effectively suppresses numerical dispersion in FWI, mitigates cycle-skipping issues, and delivers high-resolution images.

2. Theory of optimized FD weights

The two-dimensional acoustic wave equation is expressed as **Equation 1**:

$$\frac{1}{v^2}\frac{\partial^2 u_0^0}{\partial t^2} = \frac{\partial^2 u_0^0}{\partial x^2} \tag{1}$$

where u_q^p denotes the scalar pressure wavefield, with the notation shown in **Equation 2**:

$$u_q^p = u\mathrm{e}^{ik(x+qh)-i\omega(t+p\tau)} \tag{2}$$

where e is the base of the natural logarithm, u denotes a constant amplitude of the wavefield, k is the wavenumber, and ω is the angular frequency. Here, h represents the spatial grid size, t is the time step, and v stands for the velocity.

A conventional explicit second-order FD approximation for the temporal derivative is given by **Equation 3**:

$$\frac{\partial^2 u_0^0}{\partial t^2} \approx \frac{1}{\tau^2}\left[-2u_0^0 + u_0^1 + u_0^{-1}\right] \tag{3}$$

On a uniform grid, a $2M^{\text{th}}$-order FD discretization of the spatial derivative can be written as **Equation 4**:[29,38]

$$\frac{\partial^2 u_0^0}{\partial x^2} \approx \frac{1}{h^2}\left\{c_0 u_0^0 + \sum_{m=1}^{M} c_m\left[u_m^0 + u_{-m}^0\right]\right\} \tag{4}$$

where c_0 and c_m are FD weights, and $2M$ denotes the order of the FD scheme. Substituting **Equations 3** and **4** into **1** and rearranging it, we can obtain **Equation 5**:

$$\frac{1}{\tau^2}\left[-2u_0^0 + u_0^1 + u_0^{-1}\right] \approx \frac{v^2}{h^2}\left\{c_0 u_0^0 + \sum_{m=1}^{M} c_m\left[u_m^0 + u_{-m}^0\right]\right\} \tag{5}$$

Substituting **Equation 2** into the above expression transforms it into the frequency–wavenumber domain, as shown in **Equation 6**:

$$\frac{1}{\tau^2}\left[-2 + 2\cos(\omega\tau)\right]U(k,\omega) \approx \frac{v^2 U(k,\omega)}{h^2}$$
$$\left[c_0 + 2\sum_{m=1}^{M} c_m\cos(mkh)\right] \tag{6}$$

where $U(k)$ denotes the spectrum of the seismic wavefield $u(x)$. When $k=0$, the FD weights satisfy $c_0 = -2\sum_{m=1}^{M} c_m$. Using this relation and the identity $v = \frac{\omega}{k}$,

Equation 6 can be reformulated as **Equation 7**:

$$r^{-2}\left[-2 + 2\cos(rkh)\right]U(k) \approx U(k)2\sum_{m=1}^{M} c_m\left[\cos(mkh) - 1\right] \tag{7}$$

where $r = \frac{v\tau}{h}$, r denotes the Courant–Friedrichs–Lewy (CFL) number.

The value 0.505 represents the maximum stable CFL number ($r_{\max}$) for the conventional FD weights when M=6, which serves as a key parameter for evaluating numerical stability. Furthermore, as the order of the FD scheme increases (increasing M), the $r_{\max}$ continues to decrease.

The FD method introduces numerical errors in wave equation simulation due to the truncation of the Taylor series in its approximation. The specific form of this error is given by its expression in the frequency–wavenumber domain, as shown in **Equation 8**:

$$\varepsilon = U(k)\left\{ \frac{2}{rkh}\arcsin\sqrt{r^2\sum_{m=1}^{M}c_m\sin^2(mkh/2)} - 1 \right\} \qquad (8)$$

where $\dfrac{2}{rkh}\sin^{-1}\sqrt{r^2\sum_{m=1}^{M}c_m\sin^2(mkh/2)} - 1$ represents

the phase velocity error formula. If we directly construct the objective function based on the phase velocity error, the non-linear nature of the inverse trigonometric function would make it difficult to solve for the optimized FD weights. To circumvent this, we derived a simplified objective function from **Equation 7** that avoids the arcsin function non-linearity, as shown in **Equation 9**:

$$\Omega = U(k)\left\{ \frac{r^2\sum_{m=1}^{M}c_m\left[\cos(mkh)-1\right]}{\left[\cos(rkh)-1\right]} - 1 \right\} \qquad (9)$$

Thus, the numerical simulation error arising from the discretization can be expressed as **Equation 10**:

$$E = \int_0^{k_{max}} U(k)\Phi(k)dk \qquad (10)$$

where E denotes the cumulative simulation error in the time-space domain, and $\Phi(k)$ represents the dispersion relation in the time–space domain, defined as $\dfrac{r^2\sum_{m=1}^{M}c_m\left[\cos(mkh)-1\right]}{\left[\cos(rkh)-1\right]} - 1$. Here, is the maximum wavenumber. According to the Nyquist theorem, the maximum value of kh is . Therefore, **Equation 10** can be rewritten as **Equation 11**:

$$E = \int_0^{\frac{\pi}{h}} U(k)\Phi(k)dk \qquad (11)$$

Optimizing the FD weights based on this time–domain functional error can significantly improve the stability of the FD scheme. However, such optimization alone is insufficient to fully suppress spatial dispersion errors. To address this limitation, it is necessary to introduce a regularization term into the objective function that specifically constrains spatial dispersion. In this study, we incorporated the relative error of the spatial dispersion relation as an additional constraint to effectively reduce spatial dispersion during numerical simulation.

The spatial dispersion relation is derived from the spatial second derivative in the wave equation and its FD approximation, mathematically expressed as **Equation 12**:

$$\frac{\partial^2 u_0^0}{\partial x^2} \approx \frac{1}{h^2}\left\{ c_0 u_0^0 + \sum_{m=1}^{M}c_m\left[u_m^0 + u_{-m}^0\right] \right\} \qquad (12)$$

By substituting the plane wave solution, it can be simplified into the following form shown in **Equation 13**:

$$(kh)^2 U(k) \approx 2\sum_{m=1}^{M}c_m\left[\cos(mkh)-1\right]U(k) \qquad (13)$$

Minimizing the spatial dispersion error ensures that the FD scheme closely approximates the true, continuous spatial second derivative, thereby yielding optimized FD weights that effectively reduce spatial dispersion.

The dispersion relation in the time–space domain, on the other hand, involves approximating the FD scheme of the temporal second derivative using the FD scheme of the spatial second derivative. Minimizing the dispersion error in the time–space domain aims to align these two schemes more closely. In this case, the objective of optimizing FD weights does not accurately approximate the true continuous spatial second derivative by minimizing the dispersion error in the time–space domain. Consequently, optimizing the temporal dispersion error alone does not contribute to the reduction of spatial dispersion.

To simultaneously extend the stability limit and minimize spatial dispersion, we proposed a novel adaptive weighting method for constructing the objective function. The core idea is to dynamically adjust, based on the current CFL number r relative to the stability limit r_{max}, the weighting in the objective function between the time–space domain dispersion error and the spatial domain dispersion error. This enables the optimization process to ensure numerical stability at high CFL numbers while maintaining spatial simulation accuracy at low CFL numbers. The new objective function is defined as follows in **Equation 14**:

$$J = \int_0^{\frac{\pi}{h}} \left\| U(k)\Phi(k) \right\|_2^2 dx + \lambda(r)\int_0^{\frac{\pi}{h}} \left\| U(k)\phi(k) \right\|_2^2 dk \qquad (14)$$
$$= A + \lambda(r)B$$

where $\left\| \bullet \right\|_2^2$ represents the L_2-norm function, and J represents the total objective function value. A denotes the time–space domain dispersion, which directly influences the stability of the numerical simulation. Minimizing A is crucial for ensuring that the time-stepping algorithm remains stable at larger CFL numbers. In **Equation 14**,

$$\phi(k) = \frac{2\sum_{m=1}^{M} c_m \left[\cos(mkh) - 1\right]}{(kh)^2}$$

, and B represents the spatial domain dispersion, which controls the accuracy of the simulation.

Minimizing B effectively suppresses numerical dispersion and ensures the accuracy of wavefield modeling. $\lambda(r)$ is the adaptive weighting function. It is the core of this method. This function depends on the CFL number r and dynamically balances the weights of A and B during the optimization process, as shown in **Equation 15**:

$$\lambda(r) = \lambda_0 \left[1 - \frac{1}{1 + e^{-p(r - r_{tr})}}\right] \tag{15}$$

where $\lambda(r)$ is the adaptive weighting function. The weighting function $\lambda(r)$ is designed to prioritize numerical accuracy at low CFL numbers while ensuring numerical stability at high CFL numbers. We employed a sigmoid function for $\lambda(r)$, and the design rationale for its three core hyperparameters is detailed below:

(i) Base weight ($\lambda_0 = 1{,}000$): This parameter is assigned a very large value to ensure that at low CFL numbers ($r < r_{tr}$), the objective function is dominated by the spatial dispersion error B. This establishes an "accuracy-priority" optimization strategy, guaranteeing that the optimized FD weights suppress spatial numerical dispersion to the greatest extent possible under conditions where stability is not a constraint, thereby preserving wavefield fidelity.

(ii) Transition Center (r_{tr}): For example, when using the optimized FD weights (M=6) obtained by a conventional optimization algorithm, such as the least-square method, the stability limit $r_{max} < 0.505$, and $r_{tr} = 0.5050$. It acts as a proactive buffer threshold, marking the critical point where the optimization strategy shifts from "accuracy-priority" to "stability-priority." By triggering this transition preemptively, the algorithm begins to adapt before nearing the theoretical stability boundary, thereby enhancing the robustness of the method.

(iii) Steepness factor ($p = 100$): A relatively large steepness factor ensures that the weighting function $\lambda(r)$ changes rapidly and sharply around the transition center r_{tr}. This "switch-like" behavior forces the optimizer's focus to shift decisively from minimizing dispersion error in the space domain to minimizing that in the time–space domain once the CFL number enters the critical region, enabling an efficient search for coefficients that satisfy the stringent stability requirements at high CFL numbers.

These three hyperparameters work in concert to form an intelligent optimization framework. They empower our method to automatically and smoothly make an optimal trade-off between accuracy and stability in response to the CFL number, which is the key to achieving higher computational efficiency without sacrificing numerical accuracy.

In **Equation 14**, the functional incorporates the seismic wavefield spectrum. Directly obtaining this spectrum generally necessitates numerical forward modeling—a process often computationally prohibitive for practical applications. Given that the wavenumber spectrum of the wavefield in forward modeling is largely governed by the frequency spectrum of the seismic wavelet, we approximated the preconditioning operator using the wavelet spectrum.

Using a wavelet to approximate the wavefield function constitutes an incomplete representation, which may introduce errors under certain conditions, such as in complex media. We conducted forward modeling using 20 Hz and 40 Hz seismic wavelets in a complex medium, and compared the amplitude spectra of the seismic data with those of the source wavelet. As illustrated in Figures 1 and 2, under both 20 Hz and 40 Hz conditions, discernible deviations were observed between the amplitude spectrum of the seismic data (blue curve) and that of the seismic wavelet (red curve) across both high- and low-frequency ranges. This indicates that the wavelet spectrum does not fully characterize the frequency response of the seismic wavefield. Nevertheless, such an approximation error remains well within acceptable limits for our algorithm. Therefore, we adopted the wavelet
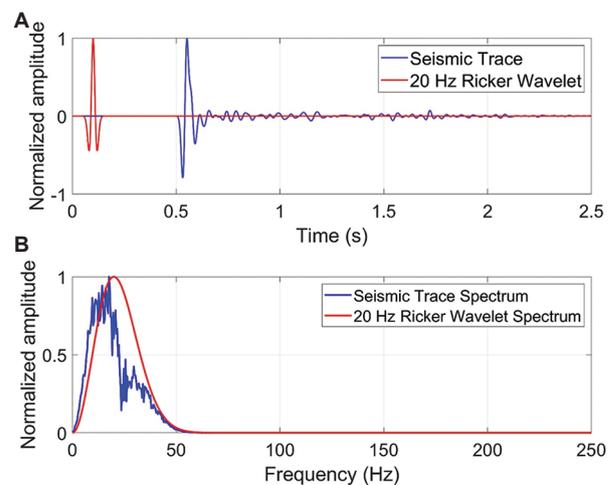

Figure 1. Comparative analysis of the 20 Hz seismic data and theoretical Ricker wavelet in time and frequency domains. (A) Normalized time–domain comparison. (B) Normalized spectrum comparison.

spectrum as a viable approximation to the frequency response of the seismic data.

Through this approximation, the $U(k)$ becomes a preconditioning operator determined by the dominant frequency. Thus, the new objective function can be formulated as **Equation 16**:

$$E = \int_0^{\frac{\pi}{h}} \left\| U_w\left(k, W\left(f_m, \tau\right)\right) \Phi(k) \right\|_2^2 dk +$$
$$\lambda \int_0^{\frac{\pi}{h}} \left\| U_w\left(k, W\left(f_m, \tau\right)\right) \phi(k) \right\|_2^2 dk \qquad (16)$$

where $U_w\left(k, W\left(f_m, \tau\right)\right)$ depends on the wavenumber and the seismic wavelet, $W\left(f_m, \tau\right)$ represents the Ricker wavelet generation operator.

The solution to **Equation 16** was found using a least-squares algorithm. According to Liu and Sen,[46] the optimized FD weights obtained by the least-square method can be solved by setting an effective bandwidth (k_{eff}). If the objective function value is larger than the setting value, we should decrease the effective bandwidth (k_{eff}) to obtain the new FD weights that can output a lower objective function value. Given that the objective function is linear and concave, it has a single lowest point—a global minimum. Thus, we can search for the optimal FD weights using the linear search method. Once the simulation error converges to this minimum, the optimal FD weights for the simulation setup can be determined. These weights result in lower simulation error and allow for a more relaxed stability condition.

This process leverages the key property that the objective function is convex, guaranteeing the existence of a unique global minimum. The search is initialized with the maximum effective bandwidth ($k_{eff} = \pi$) to encompass the full wavenumber range. The algorithm then progressively decrements the effective bandwidth by a small, fixed step size (e.g., $\Delta k = 0.01$). For each step, a set of candidate FD weights is computed by solving a least-squares problem, and the corresponding objective function value is evaluated. By sequentially comparing the objective function values between successive steps, the algorithm precisely identifies the turning point where the function value transitions from decreasing to increasing. An increase in the objective function value relative to the previous step indicates that the search has moved past the global minimum, and therefore the search process is terminated. The effective bandwidth and its associated FD coefficients that yield the minimum function value are then identified as the global optimal solution. This fully automated procedure eliminates heuristic parameter tuning and ensures the robustness and reproducibility of the results.

3. Simulation error and stability test

To validate the advantages of the proposed optimized FD weights in improving simulation accuracy, we computed the simulation error with the following simulation parameters: $v = 2000$ m/s, $\tau = 0.001$ seconds, $h = 10$ m, and a dominant frequency $f_m = 20$ Hz. Based on **Equation 16**, we compared the theoretical simulation errors between the new FD weights and the conventional optimized FD weights. The results, shown in Figure 3, clearly demonstrate a significant reduction in theoretical simulation error achieved by the new FD weights. This indicates that the proposed method can effectively reduce the simulation error.
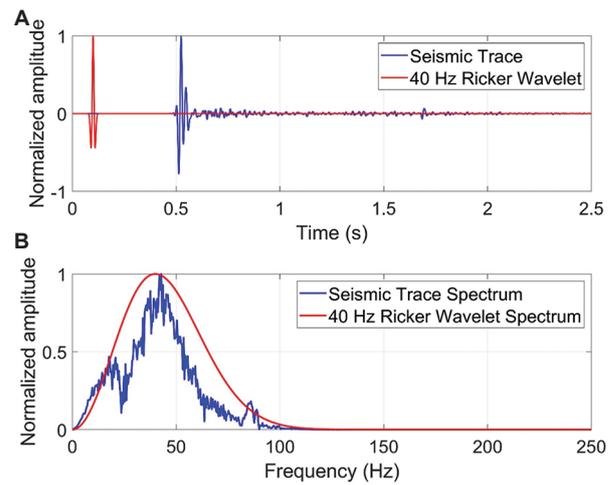
Figure 2. Comparative analysis of 40 Hz seismic data and theoretical Ricker wavelet in time and frequency domains. (A) Normalized time–domain comparison. (B) Normalized spectrum comparison.
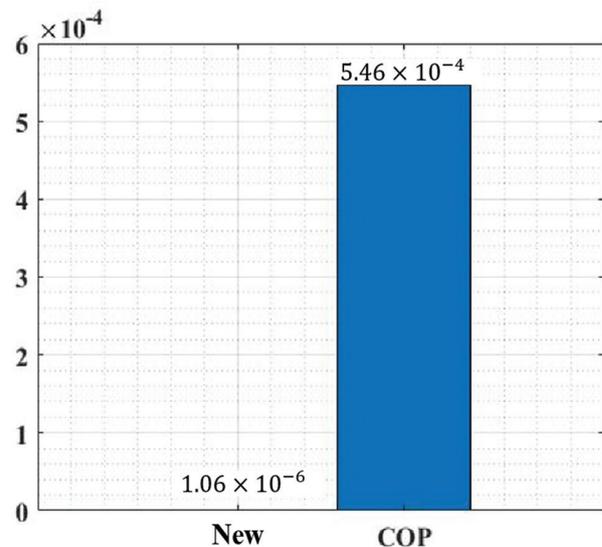
Figure 3. The comparison of simulation error J for the new method and the conventional optimized finite-difference weights (COP)

Furthermore, we conducted a comparative analysis from the perspective of numerical stability. For the two-dimensional acoustic wave equations, the stability limit condition can be expressed as **Equation 17**:

$$r_{max} = \left(2 \sum_{m=1}^{\text{int}\left((M+1)/2\right)} c_{2m-1} \right)^{-1/2} \tag{17}$$

To evaluate the improvement in stability of the new method, we computed optimized weights using the proposed method within 12th-order FD weights (M=6). We compared the maximum CFL number, ensuring numerical stability for conventional optimized FD weights and the new FD weights. The results, presented in Figure 4, indicate that the new FD weights significantly extended the stability boundary, effectively enhancing the stability of the FD simulation scheme. Two theoretical tests confirmed that the proposed method offers significant advantages in both reducing simulation error and enhancing FD stability.

To evaluate the maximum stable CFL number achievable by the proposed method for M = 3, 4, 5, and 6, the objective function was performed by using the theoretical maximum CFL number—i.e., a value asymptotically approaching unity (since the CFL number must not exceed 1)—rather than the CFL number derived from simulation parameters. Under this limiting condition, the optimized FD coefficients computed by the proposed method were found to remain stable only up to a CFL number of 0.68. Numerical simulations became unstable when the CFL number determined by the simulation parameters exceeded this threshold. Nevertheless, as summarized in Table 1, the proposed method demonstrated a measurable improvement in the maximum allowable CFL number compared with conventional schemes, achieving an average increase of approximately 20%. To validate this stability limit in practice, a numerical simulation was conducted using the optimal FD weights at CFL=0.68. The resulting shot record, presented in Figure 5, was successfully generated without instability, confirming the robustness of our method under this enhanced CFL condition.

Table 1. Comparison of maximum stable Courant–Friedrichs–Lewy numbers between conventional and optimized finite-difference schemes

Operator lengths	3	4	5	6
Conventional finite-difference weights	0.575224	0.554632	0.541266	0.531759
Optimal finite-difference weights	0.680336	0.688629	0.677551	0.684886
Stability improvement	15.45%	19.46%	20.11%	22.36%

4. FWI imaging test

The objective function of FWI is defined as follows in **Equation 18**:

$$J = \left\| d_{cal} - d_{obs} \right\|_2^2 \tag{18}$$

where d_{cal} represents the simulated data. During the FWI process, $d_{cal} = S_u$, where S denotes the sampling operator, and u is the wavefield, which is the numerical solution of the wave equation. d_{obs} is the observed data.
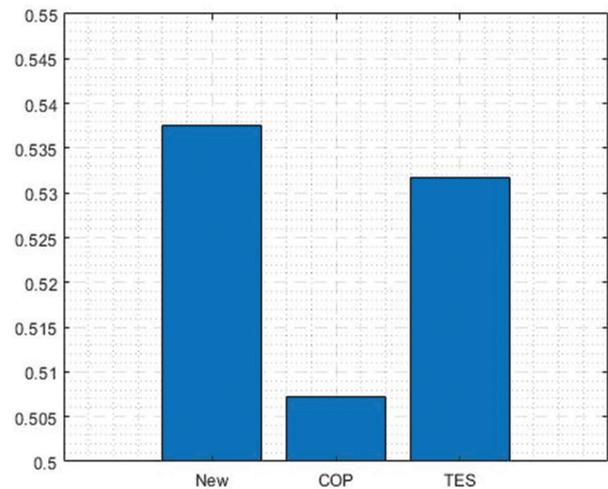

Figure 4. Stability comparison for the new method, the conventional optimized finite-difference weights (COP), and the conventional Taylor-expansion method (TES)
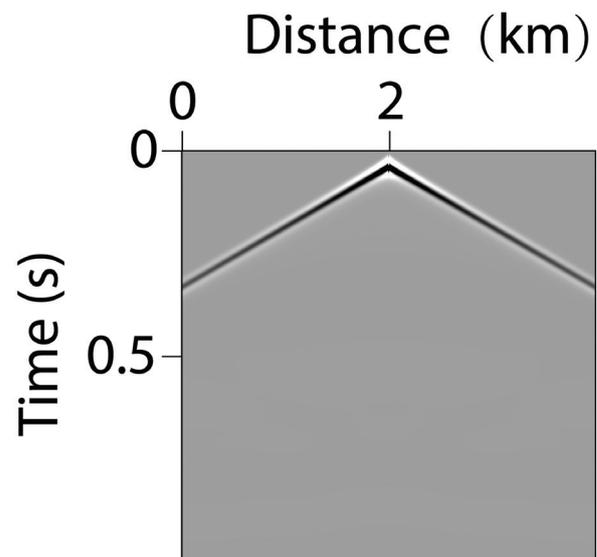

Figure 5. A shot record simulated using the optimized finite-difference weights at a Courant–Friedrichs–Lewy number of 0.68

This study focused on the impact of the FD forward-modeling algorithm on FWI imaging results. Here, "FWI imaging result" specifically refers to the spatial derivative of the inverted velocity model, as shown in **Equation 19**:

$$I = \nabla^2 v_{fin} \qquad (19)$$

where I denotes the FWI imaging result, ∇^2 represents the second-order spatial derivative, and v_{fin} is the high-resolution full-waveform velocity inversion result. High-resolution FWI requires matching the high-frequency components in the observed data. Due to the nature of FD methods, simulating high frequencies can introduce significant numerical dispersion.

To investigate the impact of numerical dispersion on FWI imaging, a set of numerical experiments was conducted. A horizontal layered model, depicted in Figure 6, was employed as the true velocity model. The model spans 100 grid points in the lateral direction and 600 grid points in depth, with a uniform spatial sampling interval of 10 m. A Ricker wavelet with a dominant frequency of 40 Hz was used as the source function. Wavefield simulation was performed over a total duration of 2 seconds using a time step of 1 ms. The resulting observed record, generated in the absence of a direct wave, is presented in Figure 7A.

Accurate wavefield simulation relies not only on high-precision spatial discretization methods but also on effective boundary treatments to suppress unphysical boundary reflections. To achieve superior absorption, the perfectly matched layer technique was introduced and has since evolved. A perfectly matched layer introduces a special absorbing layer surrounding the computational domain, within which the wavefield is split and attenuated exponentially, thereby efficiently absorbing outgoing waves. Recent improvements, such as the residual perfectly matched layer, have further enhanced absorption performance and numerical stability.[15] In the FWI imaging conducted in this study, to ensure effective control of boundary reflections, the approximation absorbing boundary condition was employed in all our numerical

experiments. Our optimized operator was responsible for computing the spatial derivatives across the entire computational domain (including both the interior and the boundary layers), while the absorbing boundary condition acted as a standard module applied at the boundaries to absorb outgoing wave energy.

We conducted wavefield simulation and inversion tests using three different methods to compare the effects of numerical dispersion: (i) A dispersion-free method, which serves as an ideal reference benchmark; (ii) a conventional FD method employing conventional 12th-order Taylor-series-based FD weights; and (iii) the proposed new method, in which the weights are optimized using our new scheme.

Figures 7B and 7C show the shot gathers generated from the conventional and new methods, respectively. As shown in Figure 7D, the wavefield simulation results under a dominant frequency of 40 Hz demonstrate that the new method produced only minor numerical dispersion (black curve), while the conventional method exhibits severe dispersive artifacts (blue curve).

Using the simulated data, we performed both objective function analysis and imaging tests within the FWI framework. Due to strong numerical dispersion, the conventional method yielded an objective function (Figure 8A) that contains numerous prominent local minima, and prevented the optimization process from converging to the global minimum. We further studied how numerical dispersion hinders convergence in
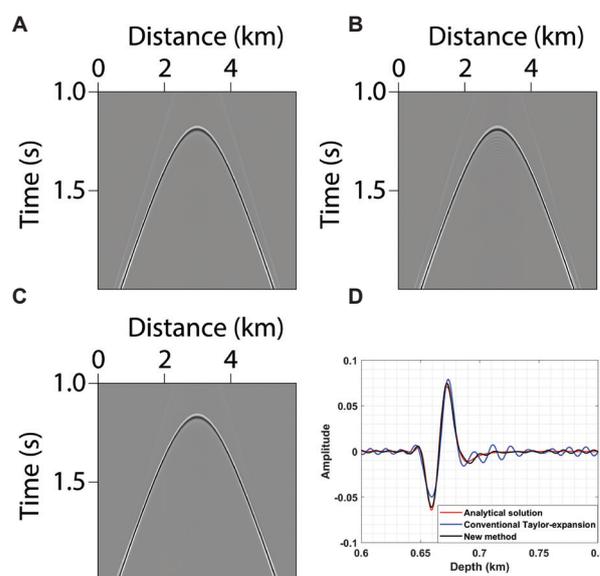

Figure 6. The layer model

Figure 7. Shot gathers comparison for a source located at the middle point of the model. (A) Analytical solution. (B) Conventional Taylor-expansion method. (C) New method (Proposed). (D) Single trace comparison for three solutions.
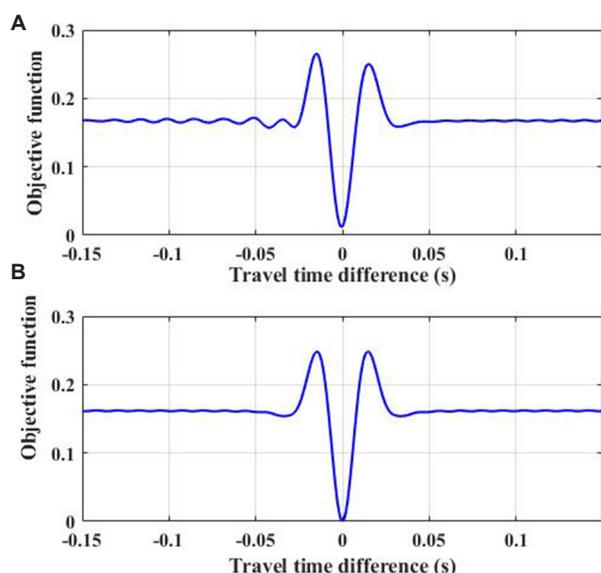
Figure 8. Objective function landscapes for velocity estimation. (A) With dispersion error. (B) Without dispersion error (including travel time difference curve).

Figure 9. FWI velocity model comparison. (A) Reference solution without dispersion. (B) Result from the conventional method with artifacts. (C) Result from the proposed optimized method. (D–F) Zoomed-in views of the shallow target area corresponding to (A–C).

waveform inversion. Dispersion retarded the high-frequency components of a seismic wave, which increased the proportion of low-frequency energy within the main lobe of the arrival, thereby broadening it. As a result, the main lobe becomes broader. In the objective function, this appeared as a wider and smoother central region, which lowered the resolution and accuracy of the model updates.

In contrast, the objective function derived from the dispersion-free simulations (Figure 8B) had more stable convergence behavior. This result demonstrates that numerical dispersion significantly increases the risk of becoming trapped in local minima, thereby affecting both the convergence behavior and the convergence speed of the inversion.

Figure 9 presents the FWI velocity model comparison. Figure 9A shows the dispersion-free reference inversion result, Figure 9B corresponds to the result obtained using the conventional method with strong numerical dispersion, and Figure 9C displays the result from the proposed optimized method. In terms of imaging results, the proposed new method (Figure 9C) closely matched the dispersion-free benchmark (Figure 9A). The events were continuous and clear, with no false structures caused by numerical dispersion. However, the result from the conventional method with strong dispersion (Figure 9B) showed clear artifacts.

These artifacts have distinct spatial patterns. First, lower seismic velocities resulted in higher wavenumbers at
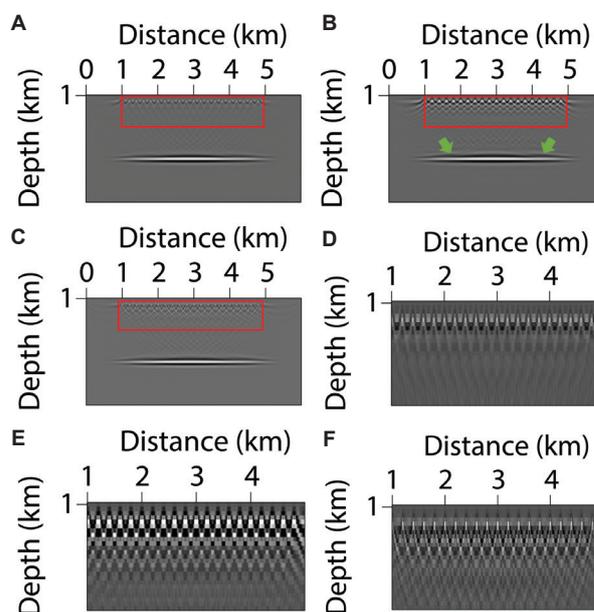
a given frequency. Thus, the strong numerical dispersion appeared in the upper part of the model, where velocities were low. This strong dispersion amplified noise related to direct waves—often manifesting as repeated acquisition footprints—making such noise prominent in the shallow section (red box in Figure 9B). Second, the numerical dispersion distorted wavefield paths and imaging geometry, resulting in stronger migration swing artifacts. In the image, these artifacts appeared as strong events below the source location. By comparing Figure 9D-F, we observed that these events had higher energy and covered a larger area, which affected the identification of useful shallow reflections.

To further study how dispersion influences the results, we compared amplitude values from single traces (Figure 9C). The conventional method result (red curve) showed significant deviation from the dispersion-free result (blue curve). In the shallow part, the conventional result showed large amplitude errors. This is mainly because severe dispersion causes a mismatch between the simulated and observed direct waves, leading to incorrect inversion updates near the surface (the black box). In deeper layers, numerical dispersion resulted in mispositioned reflections and oscillatory artifacts (green arrows in Figures 9B and 10). The trace analysis matches the artifacts seen in the full image, confirming that numerical dispersion significantly influenced FWI

results. It creates false reflections, strengthens acquisition footprints and migration swings, and greatly reduces the quality and reliability of the inversion.
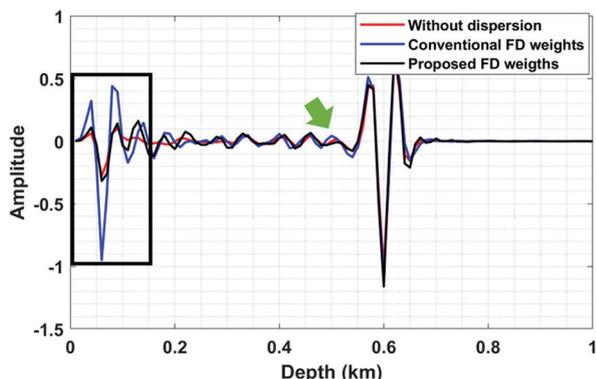

Figure 10. The single trace comparison of the full-waveform immersion imaging result in Figure 9
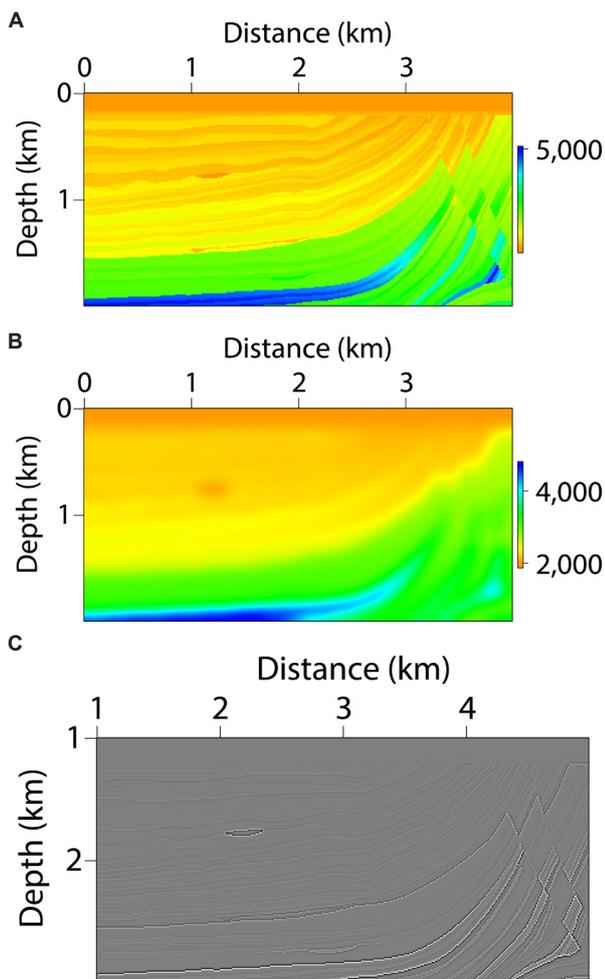Abbreviation: FD: Finite-difference.

Figure 11. The subsurface models. (A) True velocity model. (B) Initial velocity model. (C) True reflectivity model.

To thoroughly evaluate the adaptability of the new method to complex models, we conducted FWI imaging tests using a section of the classic Marmousi model. The model had dimensions of 200 × 400 grid points with a spacing of 10 m between adjacent points. The true velocity model is shown in Figure 11A. The inversion started from a smoothed version of the velocity model (Figure 11B) to reduce the non-linearity of the inverse problem. In the numerical simulation, we used a Ricker wavelet with a dominant frequency of 30 Hz as the source and simulated 2.5 seconds of recorded data with a sampling interval of 1 ms. The acquisition system was configured with 20 shots evenly spaced at 200 m intervals on the surface, and each shot was recorded with full receiver coverage. A sample shot record located at the center of the model is shown in Figure 12.

We first performed comparative FWI experiments using the 30 Hz dominant frequency wavelet. The results indicated that at this frequency, the inversion results from the conventional, new, and dispersion-free reference methods closely matched (Figure 13A-C). The inverted velocity models converged to comparable levels. This indicates that at 30 Hz, numerical dispersion was not significant, and all three modeling methods provided forward wavefields that met the accuracy requirements for inversion. Thus, the objective function decreased stably and converged to a solution near the global minimum.

To further investigate the effect of dispersion, we repeated the inversion using a Ricker wavelet with a dominant frequency of 40 Hz. Under these conditions, the performance of the different methods varied significantly. The conventional method produced shot records with strong numerical dispersion (Figure 14B), particularly in the shallow low-velocity zone, as indicated by the red arrows.
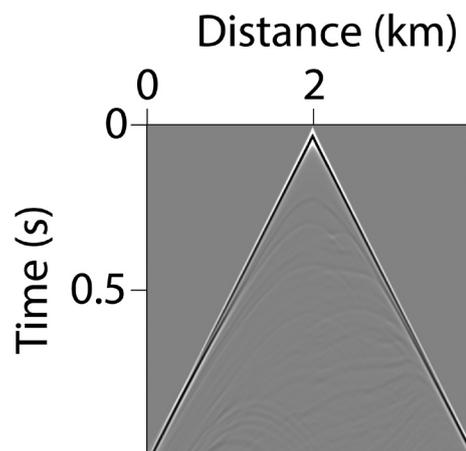

Figure 12. The observed data when the dominant frequency of the Ricker wavelet was 30 Hz
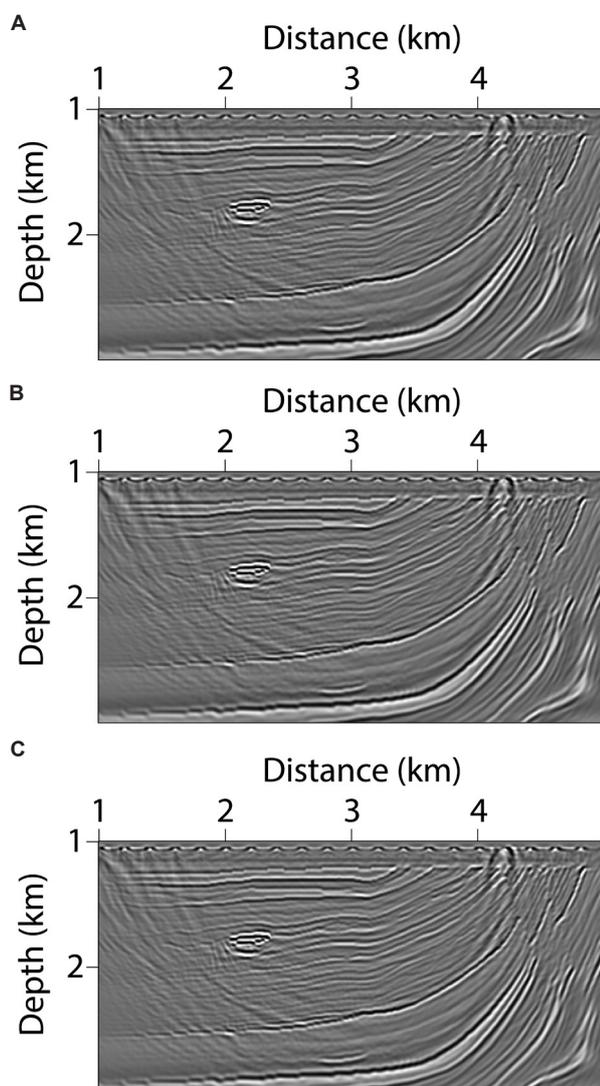
A

B

C

Figure 13. Full-waveform inversion imaging results at 30 Hz dominant frequency. (A) Result from the analytical method. (B) Result from the conventional Taylor-expansion method. (C) Result from the new proposed method.

Severe wavefield distortion negatively impacted the inversion process in several ways. First, because dispersion was strongest in the shallow low-velocity zone, wave energy was largely trapped near the surface. This significantly reduced illumination in mid-deep regions, leading to highly unbalanced inversion gradients with depth. Compared with the dispersion-free reference inversion result (Figure 15A), insufficient illumination of deep structures resulted in very weak updates, causing the conventional method to fail to recover the true reflections (as marked by green arrows in Figure 15B). This also caused the inversion gradient to deviate considerably from the true descent direction. Second, severe phase distortion

and energy scattering in the waveform's main lobe reduced spatial resolution. Even in shallow areas where updates were relatively concentrated, the recovered events did not achieve the high resolution of that achieved under dispersion-free conditions (as marked by the red arrows in Figure 14B). Ultimately, the conventional inversion process terminated prematurely due to rapidly deteriorating gradient quality, preventing effective reconstruction of deep structures.

In contrast, the new method and the dispersion-free reference were free from these issues. Since their wavefield simulations were unaffected by dispersion, both methods produced well-distributed inversion results and correct phase characteristics. As a result, both methods yielded inversion results with higher resolution and more accurate deep structures (Figure 15A and C). These experiments demonstrated that as the dominant frequency increases and the effect of low-velocity zones becomes more pronounced, numerical dispersion has a greater impact on FWI imaging results. It not only distorts the wavefield phase but also degrades the accuracy of inversion gradients, reducing update efficiency and potentially causing inversion failure. The proposed new method effectively suppressed numerical dispersion at 40 Hz and produced results similar to the dispersion-free reference (Figure 15A and D), confirming its robustness and adaptability in complex models and under higher-frequency conditions.

5. Discussion

In this study, the wavefield spectrum $U(k)$ was approximated using the frequency spectrum of the source wavelet $W(k)$ as a practical preconditioner in the optimization problem. While this approximation is computationally efficient and works well in many scenarios, it is important to recognize its limitations, particularly in strongly heterogeneous or highly scattering media. In such environments, wavefield propagation induces frequency shifts, scattering effects, and spectral broadening that are not captured by the source spectrum alone. Consequently, the approximated $U(k)$ may deviate from the true wavefield spectrum. As a result, the FD weights optimized under this approximation can deviate from the globally optimal solution. However, this deviation leads only to sub-optimal, not invalid, FD coefficients.

Although the resulting FD coefficients may be sub-optimal due to the use of an approximate wavefield spectrum, they are still highly effective for practical wavefield modeling. In fact, the so-called sub-optimal coefficients often produce significantly more accurate wavefields than those obtained from conventional FD
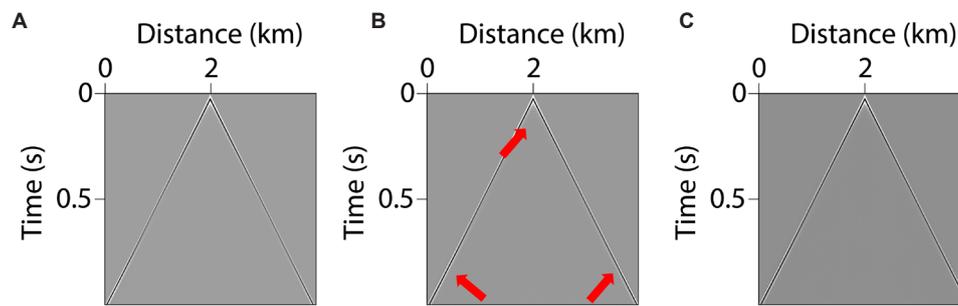
Figure 14. Shot gathers comparison for a 40 Hz source. (A) Gather from the analytical method. (B) Gather from the conventional Taylor-expansion method. (C) Gather from the new proposed method.
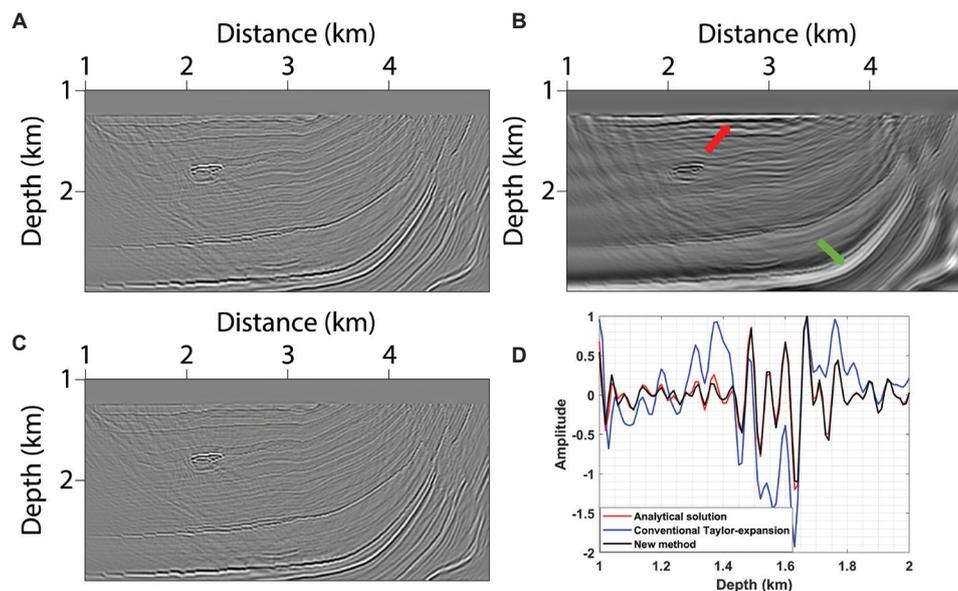

Figure 15. Full-waveform inversion imaging comparison at 40 Hz dominant frequency. (A) Analytical method result. (B) Conventional Taylor-expansion method result. (C) New proposed method result. (D) Amplitude profile comparison at a specific location.

methods. This observation is confirmed in our Marmousi model experiment, where the wavefield simulated using the proposed optimized coefficients exhibited reduced numerical dispersion and better fidelity to the true wavefield, ultimately improving the quality of FWI imaging. Therefore, despite the inherent approximation, the method remains robust and beneficial for realistic seismic applications.

Therefore, we regard the spectral approximation not as a limitation that introduces numerical instability, but as a practical and controlled compromise. It allows for computationally efficient optimization of FD weights while delivering robust performance across diverse geological settings. Although the resulting coefficients may be mathematically sub-optimal in a strict sense, they are practically highly effective—consistently reducing numerical dispersion and improving wavefield fidelity,

as demonstrated in our Marmousi experiment. This approach thus remains a reliable and beneficial strategy for enhancing the accuracy and quality of FWI in realistic seismic imaging.

6. Conclusion

The core of this study is the successful development and validation of a novel method for optimizing FD weights. This method constructed an adaptive objective function that integrates errors from both the time–space and spatial domains, using the seismic wavelet spectrum as a preconditioner. This approach effectively resolved a common industry challenge where conventional methods force a trade-off between simulation accuracy and numerical stability. Theoretically and numerically, we demonstrated that our optimized weights significantly reduced simulation errors across a broad wavenumber

range while substantially relaxing the stability limit, thus permitting larger time steps and greater computational efficiency. Applied to FWI imaging, the method suppressed numerical dispersion, mitigated cycle-skipping issues, and yielded clearer, deeper structural images free from artifacts—even at high frequencies and in complex models—closely approaching the quality of ideal dispersion-free simulations. This work confirms that high-quality forward modeling is crucial for successful FWI and shows that the proposed optimizer extends the performance limits of the FD method, offering strong practical value for advancing high-frequency, high-resolution FWI in real-world applications.

Acknowledgments

Funding

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Ganglin Lei, Jianping Huang
Data curation: Zhenwen Liu, Chang Zhou
Formal analysis: Weiting Peng
Funding acquisition: Jianping Huang
Investigation: Jianping Huang, Weiting Peng
Methodology: Ganglin Lei, Wensheng Duan, Chao Chen
Resources: Jianping Huang
Writing–original draft: Ganglin Lei, Wensheng Duan
Writing–review & editing: Jianping Huang, Chao Chen, Zhenwen Liu, Chang Zhou, Weiting Peng

Availability of data

The datasets generated and analyzed during the current study are not publicly available because they contain proprietary information as part of a collaborative project with China National Petroleum Corporation (CNPC). Data sharing is restricted by the confidentiality agreement governing this project.

References

1. Lailly P. The Seismic Inverse Problem as a Sequence of Before Stack Migrations. In: Bednar JB, editor. *Conference on Inverse Scattering: Theory and Application.* Society for Industrial and Applied Mathematics; 1983. p. 206-220.

2. Tarantola A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics.* 1984;49(8):1259-1266.

    doi: 10.1190/1.1441754

3. Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics. *Geophysics.* 2009;74(6):WCC1-WCC26.

    doi: 10.1190/1.3238367

4. Pratt RG. Seismic waveform inversion in the frequency domain; Part 1, theory and verification in a physical scale model. *Geophysics.* 1999;64(3):888-901.

    doi: 10.1190/1.1444597

5. Brossier R, Operto S, Virieux J. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics.* 2009;74(6):WCC105-WCC118.

    doi: 10.1190/1.3215771

6. Ma X, Xu S, Ke P, Zhang H. Inexact line search method in full waveform inversion. *Appl Geophys.* 2023;20:374-384.

    doi: 10.1007/s11770-020-0875-x

7. Warner M, Ratcliffe A, Nangoo T, *et al.* Anisotropic 3D full-waveform inversion. *Geophysics.* 2013;78(2):R59-R80.

    doi: 10.1190/geo2012-0338.1

8. Zhang Z, Wu Z, Wei Z, Mei J, Huang R, Wang P. FWI Imaging: Full-Wavefield Imaging through Full-Waveform Inversion. In: *SEG Technical Program Expanded Abstracts 2020.* Society of Exploration Geophysicists; 2020. p. 656-660.

    doi: 10.1190/segam2020-3427858.1

9. Lu R. Revealing Overburden and Reservoir Complexity with High-Resolution FWI. In: *SEG Technical Program Expanded Abstracts 2016.* Society of Exploration Geophysicists; 2016. p. 1242-1246.

    doi: 10.1190/segam2016-13872562.1

10. Huang R, Zhang Z, Wu Z, Wei Z, Mei J, Wang P. Full-waveform inversion for full-wavefield imaging: Decades in the making. *Lead Edge.* 2021;40(5):324-334.

    doi: 10.1190/tle40050324.1

11. Operto S, Gholami Y, Prieux V, *et al.* A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice. *Lead Edge.* 2013;32(9):1040-1054.

    doi: 10.1190/tle32091040.1

12. Xu S, Wang D, Chen F, Zhang Y, Lambare G. Full Waveform Inversion for Reflected Seismic Data. In: *Proceedings of 74th EAGE Conference and Exhibition incorporating EUROPEC 2012.* Copenhagen, Denmark: EAGE Publications BV; 2012.

    doi: 10.3997/2214-4609.20148725

13. Dong S, Dong X, Zhang R, Cong Z, Zhong T, Wang H. Global-feature-fusion and multiscale network for low-frequency extrapolation. In: *IEEE Transactions on Geoscience and Remote Sensing*. Vol. 62. New York: IEEE; 2024. p. 1-14.

14. Operto S, Gholami A, Aghamiry H, *et al.* Extending the search space of full-waveform inversion beyond the single-scattering Born approximation: A tutorial review. *Geophysics*. 2023;88(6):R671-R702.

    doi: 10.1190/geo2022-0758.1

15. Luo Y, Wang T, Li Y, Cai J, Wang Y, Fang G. A residual perfectly matched layer for wave propagation in elastic media. *Acta Geophys*. 2024;72:1561-1573.

    doi: 10.1007/s11600-023-01145-x

16. Alterman Z, Karal FC Jr. Propagation of elastic waves in layered media by finite difference methods. *Bull Seismol Soc Am*. 1969;59(1):471.

    doi: 10.1785/bssa0590010471

17. Kelly KR, Ward RW, Treitel S, Alford RM. Synthetic seismograms: A finite-difference approach. *Geophysics*. 1976;41(1):2-27.

    doi: 10.1190/1.1440605

18. Alford RM, Kelly KR, Boore DM. Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics*. 1974;39(6):834-842.

    doi: 10.1190/1.1440470

19. Dablain MA. The application of high-order differencing to the scalar wave equation. *Geophysics*. 1986;51(1):54-66.

    doi: 10.1190/1.1442040

20. Holberg O. Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena. *Geophys Prospect*. 1987;35(6):629-655.

    doi: 10.1111/j.1365-2478.1987.tb00841.x

21. Kindelan M, Kamel A, Sguazzero P. On the construction and efficiency of staggered numerical differentiators for the wave equation. *Geophysics*. 1990;55(1):107-110.

    doi: 10.1190/1.1442763

22. Zhou B, Greenhalgh SA. Seismic scalar wave equation modeling by a convolutional differentiator. *Bull Seismol Soc Am*. 1992;82(1):289-303.

    doi: 10.1785/bssa0820010289

23. Zhang JH, Yao ZX. Optimized explicit finite-difference schemes for spatial derivatives using maximum norm. *J Comput Phys*. 2013;250:511-526.

    doi: 10.1016/j.jcp.2013.04.029

24. Zhang JH, Yao ZX. Optimized finite-difference operator for broadband seismic wave modeling. *Geophysics*. 2013;78(1):A13-A18.

25. Yang L, Yan HY, Liu HX. 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*. 2017;138:161-171.

    doi: 10.1016/j.jappgeo.2017.01.020

26. Yang L, Yan HY, Liu H. Optimal staggered-grid finite-difference schemes based on the minimax approximation method with the Remez algorithm. *Geophysics*. 2017;82(1):T27-T42.

    doi: 10.1190/geo2016-0171.1

27. He Z, Zhang JH, Yao ZX. Determining the optimal coefficients of the explicit finite-difference scheme using the Remez exchange algorithm. *Geophysics*. 2019;84(3):S137-S147.

    doi: 10.1190/geo2018-0446.1

28. Koene EFM, Robertson JOA. Optimal finite-difference operators for arbitrarily sampled data. *Geophysics*. 2020;85(3):F39-F51.

    doi: 10.1190/geo2019-0081.1

29. Liu Y. Maximizing the CFL number of stable time-space domain explicit finite-difference modeling. *J Comput Phys*. 2020;416:109501.

    doi: 10.1016/j.jcp.2020.109501

30. Yuan W, Rongsheng W. An optimal spatial finite-difference operator which reduces truncation error to a minimum. *Adv Atmos Sci*. 2002;19(3):468-486.

    doi: 10.1007/s00376-002-0080-2

31. Ergen JT. A tutorial on optimizing time domain finite-difference schemes: "Beyond Holberg". *Stanford Explo Project*. 2007;129:33-43.

32. Liu Y. Globally optimal finite-difference schemes based on least squares. *Geophysics*. 2013;78(4):T113-T132.

    doi: 10.1190/geo2012-0480.1

33. Liu Y. Optimal staggered-grid finite-difference schemes based on least-squares for wave equation modelling. *Geophys J Int*. 2014;197(2):1033-1047.

    doi: 10.1093/gji/ggu032

34. Wang YF, Liang WQ, Nashed Z, Li X, Liang CH, Yang CC. Seismic modeling by optimizing regularized staggered-grid finite-difference operators using a time-space-domain dispersion-relationship-preserving method. *Geophysics*. 2014;79(5):T277-T283.

    doi: 10.1190/geo2014-0078.1

35. Ren ZM, Liu Y. Acoustic and elastic modeling by optimal time-space-domain staggered-grid finite-difference schemes. *Geophysics*. 2015;80(1):T17-T40.

    doi: 10.1190/geo2014-0269.1

doi: 10.1190/geo2012-0277.1

36. Yong P, Huang JP, Li ZC, *et al.* Optimized equivalent staggered-grid FD method for elastic wave modelling based on plane wave solutions. *Geophys J Int.* 2017;208(2):1157-1172.

    doi: 10.1093/gji/ggw447

37. Miao ZZ, Zhang JH. Reducing error accumulation of optimized finite-difference scheme using the minimum norm. *Geophysics.* 2020;85(5):T275-T291.

    doi: 10.1190/geo2019-0758.1

38. Igel H, Weber P, Riollet B. Anisotropic wave propagation through finite-difference grids. *Geophysics.* 1995;60(4):1203-1216.

    doi: 10.1190/1.1443849

39. Shao Z, Wei CW, Zhao S. DSC time-domain solution of Maxwell's equations. *J Comput Phys.* 2003;189(2):427-453.

    doi: 10.1016/s0021-9991(03)00226-2

40. Xiao F, Tang XH, Zhang XJ. Comparison of Taylor finite difference and window finite difference and their application in FDTD. *J Comput Appl Math.* 2006;193(2):516-534.

    doi: 10.1016/j.cam.2005.05.030

41. Chu CL, Stoffa PL, Self R. High-Order Rotated Staggered Finite Difference Modeling of 3D Elastic Wave Propagation in General Anisotropic Media. In: *SEG Technical Program Expanded Abstracts*; 2009. p. 291-295.

42. Chu CL, Stoffa PL. Determination of finite-difference weights using scaled binomial windows. *Geophysics.* 2012;77(3):W17-W26.

    doi: 10.1190/geo2011-0336.1

43. Remez EY. Sur la determination des polynomes D'approximation de degre donnee [On the determination of approximation polynomials of given degree]. *Commun Kharkov Math Soc.* 1934;10:41-63.

44. McClellan J, Parks T, Rabiner L. A computer program for designing optimum FIR linear phase digital filters. *IEEE Trans Audio Electroacoustics.* 1973;21(6): 506-526.

    doi: 10.1109/tau.1973.1162525

45. Peng WT, Huang JP. A robust seismic wavefield modeling method based on minimizing spatial simulation error using L_2-norm cost function. *Pet Sci.* 2024;22(3):1051-1061.

    doi: 10.1016/j.petsci.2024.12.003

46. Liu Y, Sen MK. A new time-space domain high-order finite-difference method for the acoustic wave equation. *J Comput Phys.* 2009;228(23):8779-8806.

    doi: 10.1016/j.jcp.2009.08.027