ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Automatic differentiation-based weighted instantaneous phase inversion

Mengzi Wang, Yong Hu*, Zhihan Zhang, Chunyuan Shi, and Xiaoyuan Liao

Department of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu, China

(This article belongs to the Special Issue: *Full Waveform Inversion Methods and Applications for Seismic Data in Complex Media*)

Abstract

Traditional full waveform inversion objective functions typically rely on discrepancies between waveforms, making full waveform inversion highly dependent on the initial model and the low-frequency components in the seismic data. The instantaneous phase of the seismic signal reflects the kinematic information of seismic waves and has a more linear relationship with the subsurface velocity structure, aiding in the retrieval of low-wavenumber components of velocity models. Conversely, waveform information is instrumental in achieving high-resolution inversion results while maintaining algorithmic stability. In addition, automatic differentiation provides an efficient and accurate mechanism for computing gradients by systematically applying the chain rule across the computational graph, yielding reliable derivative information for optimization. To balance the contributions of waveform and phase information in velocity inversion, within the framework of automatic differentiation, we integrate instantaneous phase and waveform information to propose an automatic differentiation-based weighted instantaneous phase inversion method. Numerical tests using the Marmousi model and the Overthrust model demonstrate that the proposed method achieves more accurate velocity inversion results.

Keywords: Full waveform inversion; Instantaneous phase; Weighted objective function; Automatic differentiation; Cycle skipping

***Corresponding author:**
Yong Hu
(jluhuyong@sina.com)

Citation: Wang M, Hu Y, Zhang Z, Shi C, Liao X. Automatic differentiation-based weighted instantaneous phase inversion. *J Seismic Explor.* 2026;35(2):025520132. doi: 10.36922/JSE025520132

Received: December 25, 2025

Revised: January 28, 2026

Accepted: January 29, 2026

Published online: March 27, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

In recent years, driven by the continuous improvement of computing capabilities, full waveform inversion (FWI) has gradually become a research hotspot. FWI leverages complete recorded seismic waveforms, iteratively minimizing the objective function that measures the misfit between synthetic and observed data to invert the velocity parameters of subsurface structures. Since its introduction in 1983, FWI has been recognized as a vital tool for high-fidelity subsurface physical parameter estimation.[1,2] Traditional FWI objective functions are typically formulated using the L2 norm, and gradients are calculated using the adjoint-state method to guide model updates in a descent direction.[3] However, in the absence of low-frequency components, directly matching synthetic and observed waveforms can lead to cycle-skipping, causing the

objective function to converge to local minima.[4-10]

Compared with waveform information, phase information in seismic data mainly reflects the kinematic attributes of seismic waves and generally has a more linear relationship with the subsurface velocity structure.[11] However, the application of phase inversion methods is constrained by the phase-wrapping issue; to effectively utilize phase information, new strategies must be employed to address this challenge. Shah *et al.*[12] and Choi and Alkhalifah[13] obtained instantaneous travel-time information by computing the derivative of phase with respect to angular frequency in the frequency domain, applied this information to inversion, and thereby effectively mitigated the impact of phase wrapping. These methods leverage the frequency-domain complex wavefield's natural separation of amplitude and phase. Nevertheless, the instantaneous phase can also be extracted and used directly in the time domain. For example, Luo *et al.*[14] proposed a time-domain acoustic FWI using instantaneous phase and demonstrated its favorable properties. Hu *et al.*[15] proposed the phase cross-correlation algorithm, which further mitigates the impact of amplitude information on inversion results. To obtain a robust initial velocity model for FWI, phase-based seismic inversion can effectively mitigate the cycle-skipping issue; examples include adaptive FWI and phase correction.[16-18]

Computing the gradient of the objective function has traditionally depended on techniques such as manual differentiation, symbolic differentiation, and numerical differentiation. These approaches, however, are plagued by inherent limitations such as cumbersome derivations, expression swelling, and truncation errors. To address these drawbacks, this study employs automatic differentiation (AD), which is specifically designed for gradient computation. AD enables automatic gradient extraction for complex differentiable functions,[19] serving as the cornerstone of modern neural network training[20] and is more commonly referred to as backpropagation.[21,22] While the concept of AD is not new, its foundational ideas were first proposed in the 1950s by Nolan,[23] who pioneered the quest to automatically differentiate computer programs. A significant milestone was reached in 1980 when Speelpenning[24] developed the first fully automated implementation of reverse-mode AD. Building on this foundation, Kasim and Vinko[25] made further advancements by developing the ξ-torch library within the PyTorch framework. This library provides differentiable scientific simulation modules, including root finders and initial-value problem solvers, thereby facilitating the seamless integration of deep learning and scientific simulation in fields such as FWI.

In the field of seismology, early research on AD was constrained by the limited maturity of available methods and thus primarily focused on relatively simple computational problems such as ray tracing and receiver function analysis. For example, Sambridge *et al.*[26] utilized AD tools such as the Transformation of Algorithms in Fortran and the Tangent Linear and Adjoint Model Compiler to implement forward gradient calculations for these tasks and demonstrated the efficiency of reverse-mode AD in computing gradients for Gibbs free energy minimization problems. With the maturation of AD tools, more studies began to address increasingly complex waveform inversion tasks, promoting the widespread application of this technology in seismology. Subsequently, a series of studies applied AD to different aspects of FWI: Tan *et al.*[27] attempted to use AD to compute the Hessian matrix in FWI; Liao[28] built an inversion workflow for the 2D acoustic wave equation based on AD tools; Richardson[29] reformulated the traditional adjoint-state method for gradient computation in seismic FWI as a recurrent neural network, combining it with AD to improve computational efficiency and convergence speed; Cao and Liao[30] implemented a borehole FWI workflow using the AD tool TAPENADE and verified its feasibility with synthetic data. Zhu *et al.*[31] further derived and validated the theoretical equivalence between gradients obtained via AD and those computed using the adjoint-state method, and developed a seismic inversion framework based on the Julia programming language. In addition, existing studies have begun to integrate AD methods with deep learning neural networks to enhance the performance of FWI.

In this study, we first introduce the concept and advantages of AD. Next, we derive the procedure for extracting exponential phase information. Then, we explain how to introduce weighting factors to construct the weighted instantaneous phase objective function. Finally, the Overthrust and Marmousi models are used to validate the effectiveness of the weighted instantaneous phase inversion method.

2. Methods

2.1. Automatic differentiation

In FWI, differentiation is used to compute the gradient of the objective function with respect to the model parameters, enabling iterative optimization. Traditional methods of derivative computation include manual, symbolic, and numerical differentiation. Manual differentiation is labor-intensive and prone to errors; symbolic differentiation can result in expression swelling; and numerical differentiation—such as finite differences—is straightforward but introduces truncation and round-off

errors, which can accumulate and potentially compromise the stability and accuracy of the inversion process.

To overcome these limitations, we introduce AD, a technique widely used in machine learning for efficient and precise derivative computation. AD evaluates derivatives by decomposing a complex function into a series of basic operations and systematically applying the chain rule to accumulate derivative information during numerical computation. Unlike the traditional adjoint-state method, which requires analytical derivation of adjoint sources and adjoint wavefields, AD automatically records the computational graph and tracks derivative propagation throughout the forward computation. This approach eliminates the need for manual derivation of adjoint terms and reduces the risk of implementation errors.[32,33]

As shown in Figure 1, the adjoint source is a key component of adjoint-state–based FWI. When the objective function becomes complex, obtaining and implementing the adjoint source and adjoint wavefield can be difficult and error-prone. In particular, after the seismic signal is transformed into the time–frequency domain or the frequency–wavenumber domain, the objective function becomes more complex, and it is increasingly difficult to obtain the adjoint source. By contrast, AD can automatically record the computational graph and track the flow of derivatives during the forward computation, eliminating the need for manual derivation of adjoint sources and adjoint wavefields. It provides an efficient tool for the implementation of FWI with complex objective functions and multi-parameter inversion.

Automatic differentiation offers two implementation strategies: forward mode and reverse mode. Within the forward mode, derivative information flows from the input nodes to the output nodes across the computation graph. This approach is particularly efficient when dealing with a limited number of input parameters. Conversely, reverse mode propagates gradients from the output and is more suitable for situations where there are many input parameters. Given that FWI usually involves a high-dimensional model parameter space with a scalar or low-dimensional objective function, reverse-mode AD is generally preferred for efficiently and accurately computing gradients in multi-parameter FWI problems.

2.2. Instantaneous phase information extraction

In digital signal processing, there is a certain relationship between the real and imaginary parts of a complex signal, which is described by the Hilbert transform. The Hilbert transform is a mathematical tool that can extract the instantaneous amplitude and phase of a signal and is often used to analyze the signal's instantaneous amplitude and phase characteristics. By applying the Hilbert transform to the signal, the transformed result is treated as the imaginary part, and the original signal is taken as the real part. The analytic signal $u_a(t)$ can be constructed using **Equation 1**:

$$u_a(t) = u(t) + i \cdot u_H(t) = E(t) e^{i\varphi(t)} \tag{1}$$

where $E(t)$ is the envelope of the signal, $u(t)$ represents the received seismic signal, $u_H(t)$ represents the Hilbert transform of the signal, and $\varphi(t)$ represents the instantaneous phase information of the signal. $\varphi(t)$ can be calculated using **Equation 2**:

$$\varphi(t) = arctan \frac{u_H(t)}{u(t)} \tag{2}$$

The instantaneous phase of a signal not only reflects the travel-time information of seismic waves but also directly reveals the medium's velocity and kinematic characteristics. By incorporating phase information into the objective function, the nonlinear behavior of the objective function in FWI can be significantly improved, which helps mitigate the risk of becoming trapped in local minima. This feature makes instantaneous phase a powerful tool in FWI applications. Figure 2A shows the observed single-shot record, whereas Figure 2B and 2C show single-shot records of envelope data and instantaneous phase data, respectively.

To avoid the impact of phase wrapping on the inversion, the exponential phase information of the signal is utilized:

$$e^{i\varphi(t)} = \frac{u_a(t)}{E(t)} = \frac{u(t) + i \cdot u_H(t)}{\sqrt{u^2(t) + u_H^2(t)}} \tag{3}$$

The exponential-phase representation effectively removes instantaneous amplitude information, such that the resulting complex exponential phase primarily carries the essential phase-related characteristics of the wavefield. Moreover, computing the exponential phase avoids numerically sensitive operations such as phase unwrapping, which can produce phase jumps; this helps mitigate phase wrapping issues. In this work, the exponential-phase measurements are employed as the objective function for a weighted instantaneous phase inversion within the AD framework. This combination leverages the kinematic robustness of seismic data together with AD's exact gradient computation to improve inversion

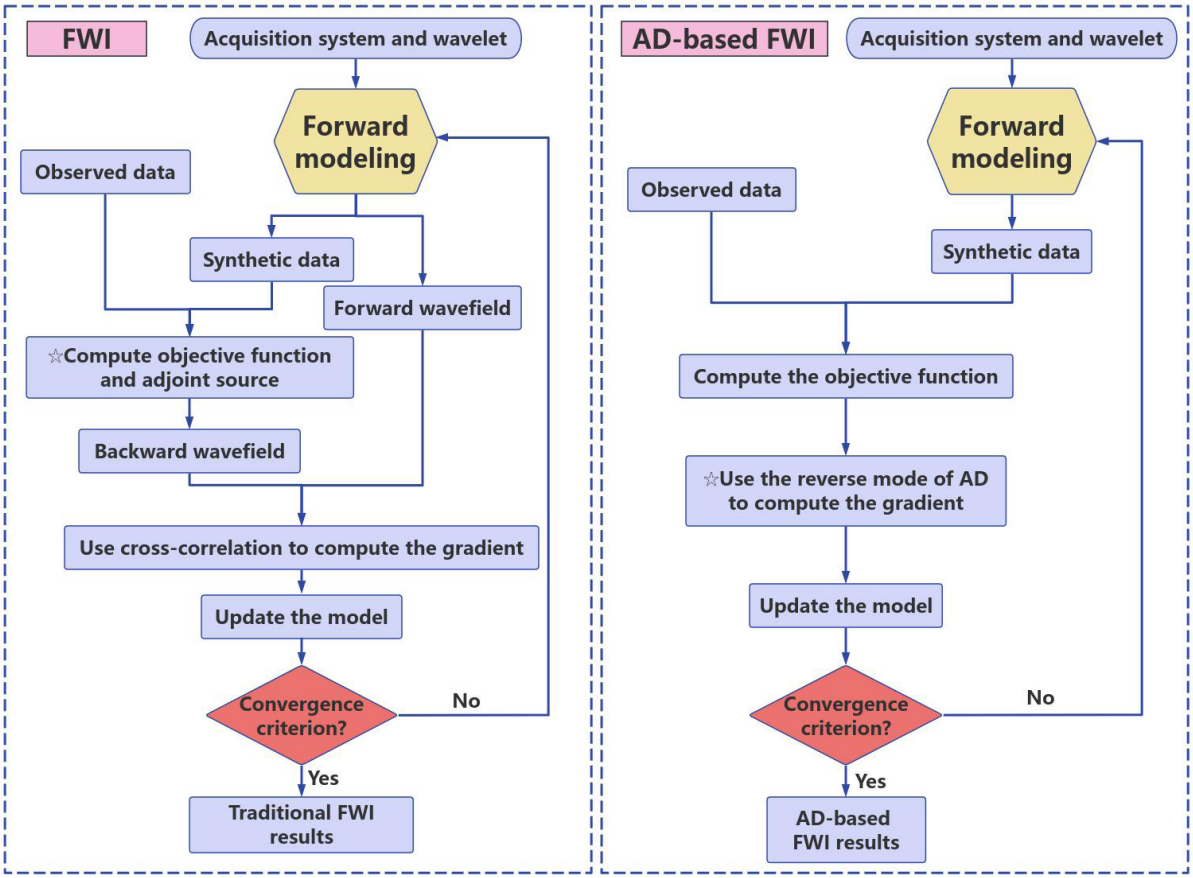

Figure 1. Comparison of workflows between traditional full waveform inversion (FWI) and automatic differentiation (AD)-based FWI

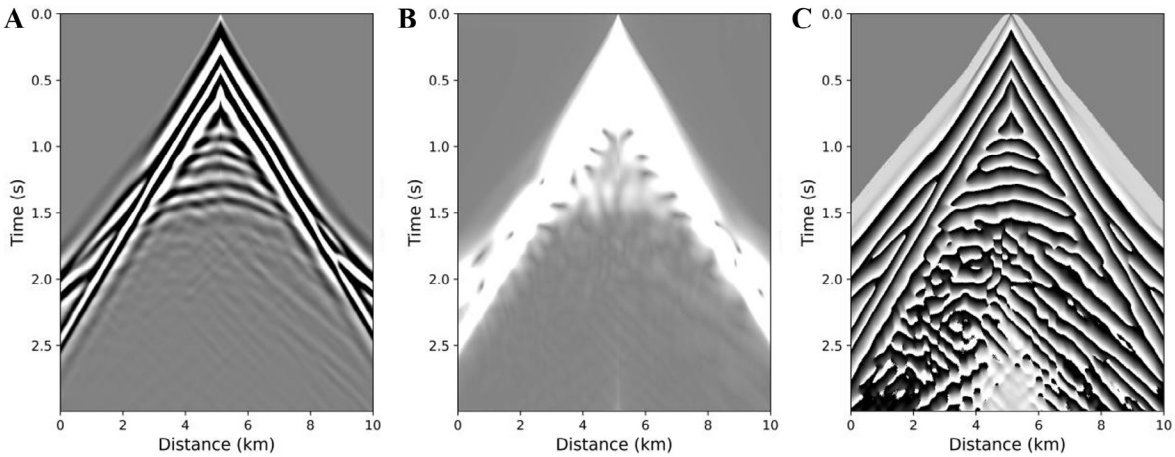

Figure 2. Single-shot seismic records: (A) Observed data, (B) envelope data, and (C) instantaneous phase data

stability and accuracy.

2.3. Construction of weighted instantaneous phase objective function

Phase information typically has a more linear relationship with the subsurface velocity structure than waveform amplitudes do. When the initial model is far from the true model and the amplitude is large, phase information more robustly captures timing/phase alignment and provides a better starting model. Conversely, when the inverted model aligns closely with the true model and fine-scale details are of interest, employing an L2 norm objective function is beneficial. This approach facilitates high-resolution inversion by fully leveraging waveform information, thereby enabling a clear depiction of complex model features and maintaining the stability of the algorithm.

Accordingly, we introduce a weighting factor, $w(i)$, to achieve a balanced integration of both instantaneous phase and waveform information. By assigning distinct weights to the objective functions based on waveform and phase information during the inversion process, the complementary strengths of both terms are utilized. The formulation for the weighted instantaneous phase objective function is expressed as follows:

$$J_{w-phase}(m) = w(i)J_w(m) + (1-w(i))J_{phase}(m) \tag{4}$$

In **Equation 4**, J_w is the waveform-based L2 norm and can be expressed as:

$$J_w(m) = \frac{1}{2}\sum_s\sum_r\int_0^T p(s,r,t,m) - d(s,r,t)_2^2\, dt \tag{5}$$

J_{phase} is the exponential phase-based objective function and can be expressed as:

$$J_{phase}(m) = \sum_s\sum_r\int_0^T \left|e_{r_syn}^{i\varphi}(t) - e_{r_obs}^{i\varphi}(t)\right|^2 dt \tag{6}$$

where $e_{r_syn}^{i\varphi}(t)$ and $e_{r_obs}^{i\varphi}(t)$ represent the exponential instantaneous phase information for the r-th receiver from the synthetic and observed data. Bringing **Equation 3** into **Equation 6**, we obtain:

$$J_{phase}(m) = \sum_s\sum_r\int_0^T\left[\frac{s^r(t)}{E_{r_syn}(t)} - \frac{u^r(t)}{E_{r_obs}(t)}\right]^2 dt + \sum_s\sum_r\int_0^T\left[\frac{s_H^r(t)}{E_{r_syn}(t)} - \frac{u_H^r(t)}{E_{r_obs}(t)}\right]^2 dt \tag{7}$$

In **Equation 7**, $S^r(t)$ and $u^r(t)$ represent the r-th receiver of the synthetic data and the observed data, respectively, whereas $S^r(t)$ and $u_H^r(t)$ denote the r-th receiver of the Hilbert-transformed synthetic data and observed data, respectively. By combining **Equation 5** and **Equation 7**, we construct a weighted instantaneous phase objective function. For the gradient calculation of this objective function, AD is employed: it automatically tracks the forward computation, traverses the computational graph in reverse, and accumulates parameter gradients via the chain rule. This approach eliminates the need to derive adjoint sources for complex objective functions and provides robust computational support for model parameter optimization.

3. Results and discussion

3.1. Overthrust model tests

This study first employed the Overthrust model through numerical experiments to validate the effectiveness of the proposed method. The true and initial velocity models are shown in Figure 3A and 3B, respectively. The initial velocity model (Figure 3B) was obtained by applying a Gaussian smoothing filter to the true model. The source wavelet has a dominant frequency of 6 Hz. The spatial discretization step is 25 m, and the temporal sampling interval is 4 ms. The model consists of a 10 km horizontal (X) extent and a 2.5 km depth (Z), with a total recording time of 3 s. The acquisition system included 80 sources spaced at 125 m intervals. For each shot, 401 receivers were deployed along the X-direction with a receiver spacing of 25 m. In field seismic data, the low-frequency components are often accompanied by low-frequency noise, making it difficult to obtain low-frequency data with a high signal-to-noise ratio. Moreover, FWI exhibits strong nonlinearity; in the absence of low-frequency components, the inversion is prone to converge to local minima, resulting in unsatisfactory models. To realistically simulate field conditions, frequency components below 3 Hz were removed from the synthetic data in this test.

Figure 4A and 4C show the results of traditional FWI after 40 and 80 iterations, respectively, while Figure 4B and 4D present the results of weighted instantaneous phase inversion at the same number of iterations. The comparative results demonstrate that the weighted instantaneous phase inversion method rapidly constructs a stable overall velocity model that aligns with the true velocity structure in fewer iterations.

Figure 4E shows the inversion result of the traditional FWI. When the initial model significantly differs from the true velocity model, the phase mismatch between the synthetic and observed data leads to cycle-skipping, which prevents the inversion from correctly relating data to model parameters.

Subsequently, we employed the weighted instantaneous

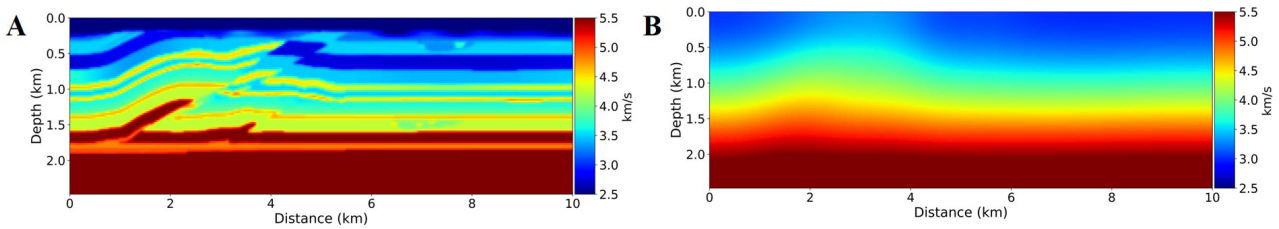

Figure 3. Velocity model: (A) True and (B) initial velocity models

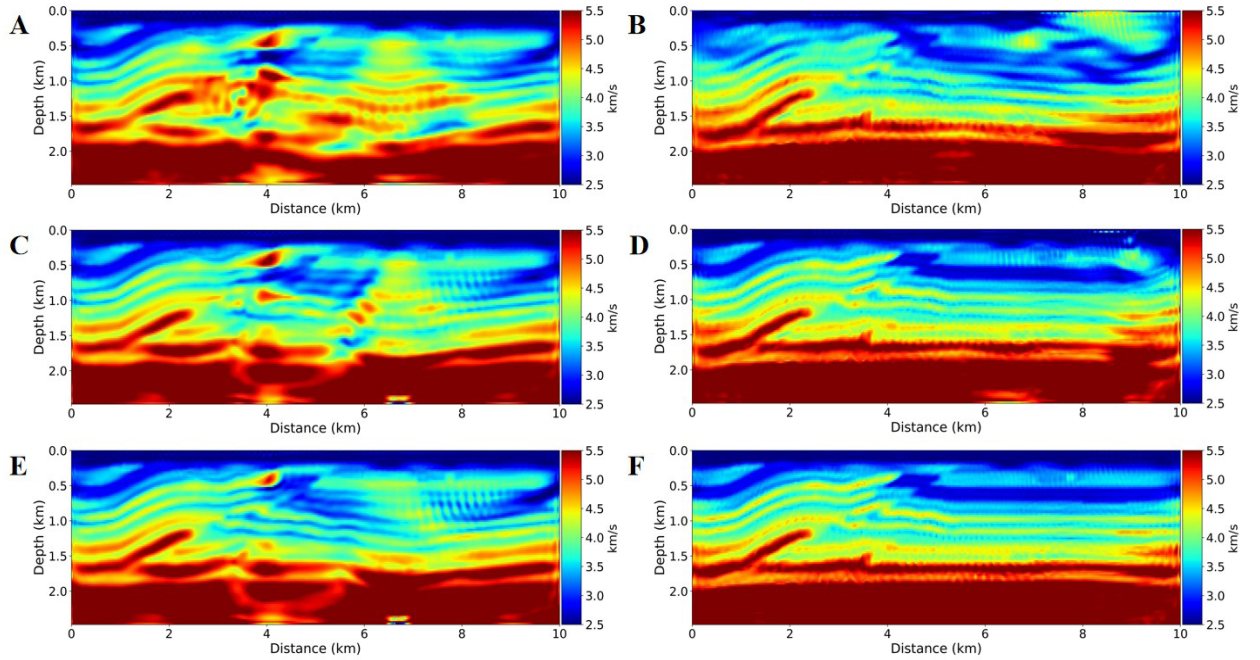

Figure 4. Inversion results for the Overthrust model: (A) Traditional full waveform inversion (FWI) after 40 iterations, (B) weighted instantaneous phase inversion after 40 iterations, (C) traditional FWI after 80 iterations, (D) weighted instantaneous phase inversion after 80 iterations, (E) final inversion result of traditional FWI, and (F) final inversion result of weighted instantaneous phase inversion

phase inversion method. During the early stages of inversion, the weight factor was set to 0.2, with the instantaneous phase objective function comprising 80% and the L2 norm objective function 20%. In the later stages, the weighted instantaneous phase had already provided a robust initial model, so the weight factor was increased to 1, with full emphasis on the L2 norm objective function to enhance model resolution. Figure 4F illustrates the result of the weighted instantaneous phase inversion. By comparing Figure 4E and 4F, it is evident that the weighting strategy leverages the complementary strengths of phase and waveform information, maintaining global structural robustness while significantly enhancing the ability to capture details, thereby improving the overall effectiveness and accuracy of the inversion.

Figure 5 presents a comparison of the velocity profile results at a 4-km section. These numerical tests demonstrate that within the framework of AD, the weighted instantaneous phase inversion method effectively mitigates the cycle-skipping issue in FWI while enhancing the accuracy of velocity inversion, achieving robust inversion in complex geological environments.

3.2. Marmousi model tests

To validate the applicability of the weighted instantaneous phase inversion method based on AD in more complex models, we extended our tests from the existing Overthrust model to the Marmousi model. The Marmousi model, with its intricate layered structures and diverse velocity field characteristics, poses a greater challenge to the inversion process. The true model used for inversion is shown in Figure 6A, whereas Figure 6B illustrates the linear initial model employed in the inversion. Compared to the initial model derived from smoothing the true model, the linear

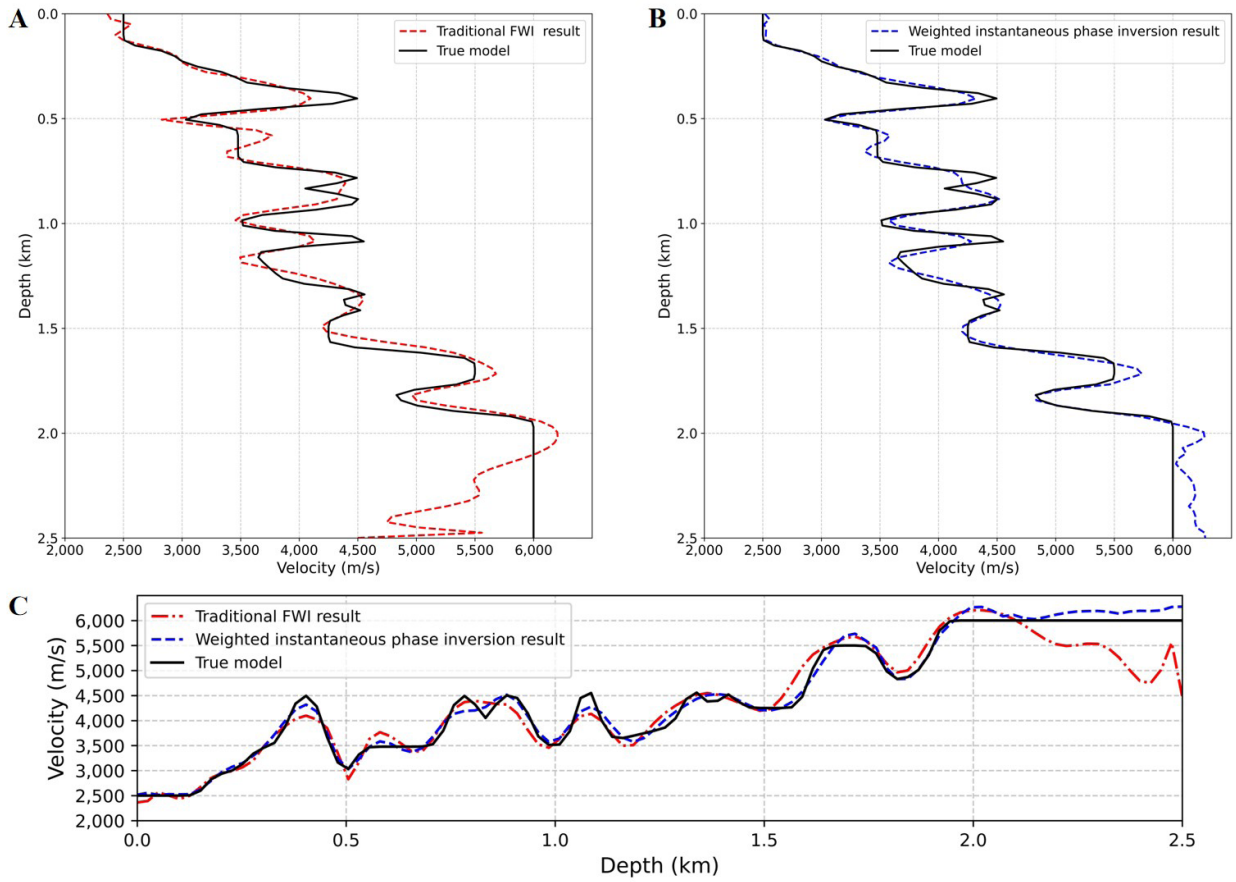

Figure 5. Velocity profiles extracted at a 4-km section: (A) Result from traditional full waveform inversion (FWI), (B) result from weighted instantaneous phase inversion, and (C) comparison of velocity profiles

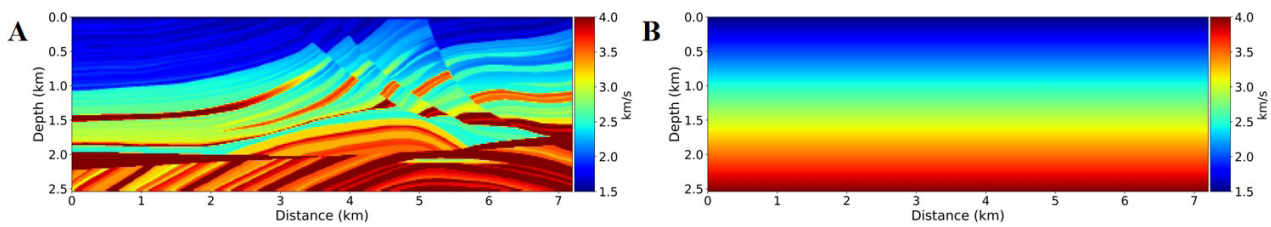

Figure 6. Inverted model: (A) True and (B) initial model

model lacks large-scale structures, making the objective function more susceptible to convergence to local minima.

Figure 7A and 7C present the results of the traditional FWI after 40 and 80 iterations, respectively, while Figure 7B and 7D show the results of the weighted instantaneous phase inversion method for the same number of iterations. The comparison demonstrates that, even for the complex Marmousi model, the weighted instantaneous phase inversion method rapidly produces a stable and accurate overall velocity model within relatively few iterations. Figure 7E and 7F present the final inversion results. By comparing Figure 7E and 7F, it is evident that the weighted instantaneous phase inversion method achieves satisfactory results even in the presence of complex layered structures. Particularly at a depth of 1.5 km on the right side, the method successfully recovers the precise velocity field, avoiding the inversion errors and blurred boundaries observed in Figure 7E.

Figure 8 compares the velocity profiles extracted at a distance of 5 km. Figure 8A shows that the traditional FWI exhibits noticeable deviations in the middle and deeper layers. In contrast, Figure 8B illustrates the velocity profile

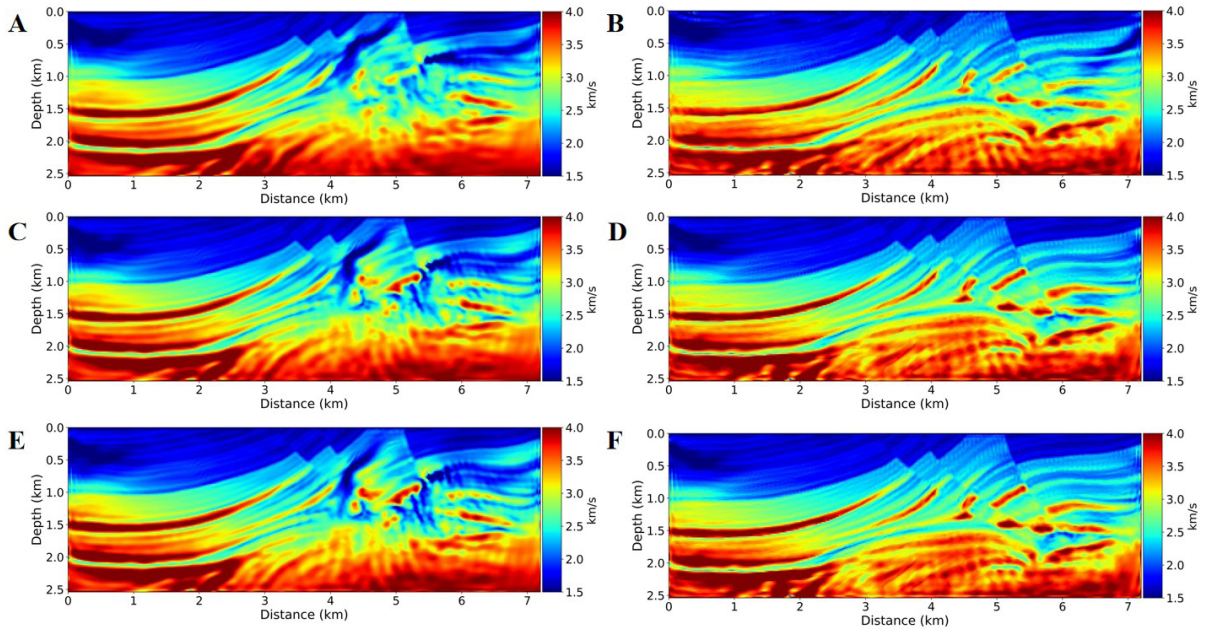

Figure 7. Inversion results for the Marmousi model: (A) Traditional full waveform inversion (FWI) after 40 iterations, (B) weighted instantaneous phase inversion after 40 iterations, (C) traditional FWI after 80 iterations, (D) weighted instantaneous phase inversion after 80 iterations, (E) final inversion result of traditional FWI, and (F) final inversion result of weighted instantaneous phase inversion

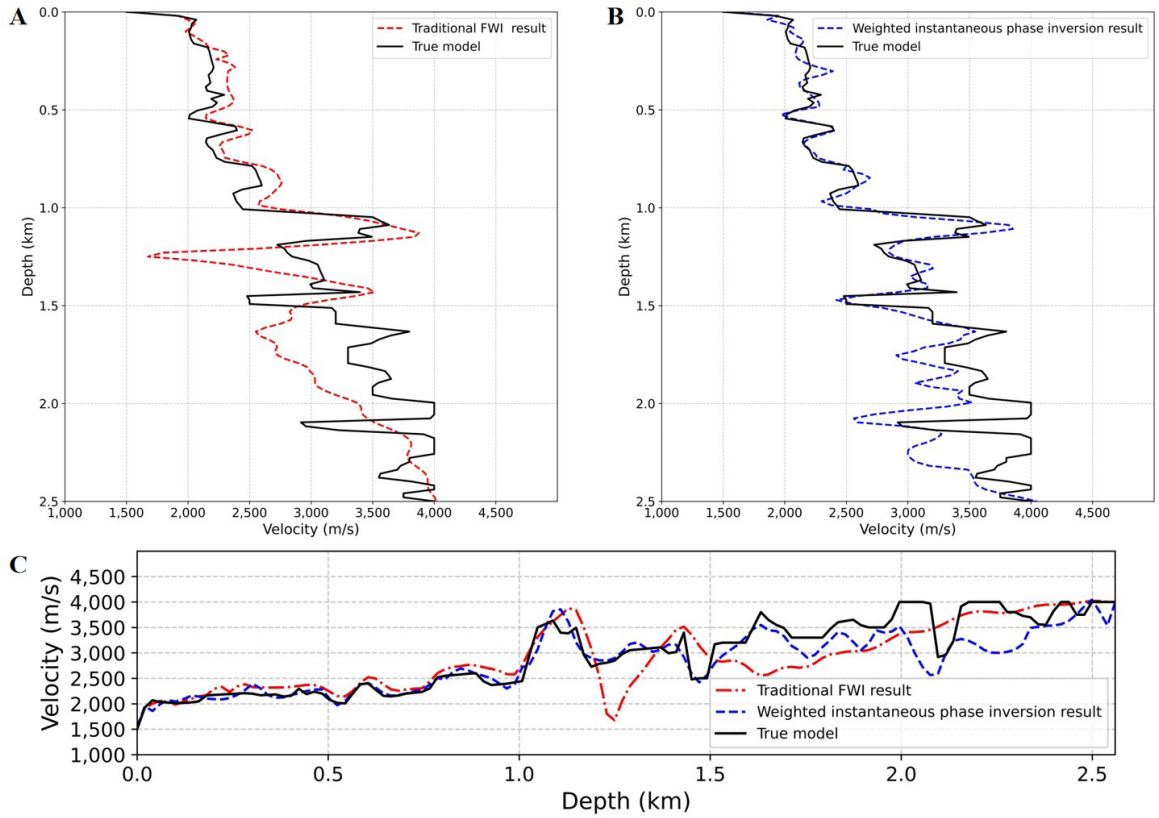

Figure 8. Velocity profiles extracted at a 5-km section: (A) Result from traditional full waveform inversion (FWI), (B) result from weighted instantaneous phase inversion, and (C) comparison of velocity profiles

using the weighted instantaneous phase objective function, which maintains higher accuracy in both shallow and deep layers compared to the traditional FWI, demonstrating a closer match with the true model. Figure 8C further shows that the weighted instantaneous phase objective function better characterizes the velocity model. The method effectively alleviates the cycle-skipping issue and enhances the stability and resolution of the inversion process.

To visually compare the inversion results with the true model, Figure 9 illustrates the absolute error distributions for the two objective functions. Figure 9A shows the absolute error between the traditional FWI results and the true model, with large residual errors in the middle and lower depths on the right side. Conversely, Figure 9B presents the absolute error distribution obtained using the weighted instantaneous phase objective function, demonstrating a notable reduction in error in the same region compared to Figure 9A.

Testing with both the Overthrust and the Marmousi models demonstrates that employing a weighted instantaneous phase inversion strategy, facilitated by AD, significantly improves the robustness and resolution of the inverted velocity model. It effectively mitigates the cycle-skipping problem typically arising from the absence of low-frequency information in FWI, resulting in more accurate inversion outcomes.

To quantitatively compare the two inversion methods, Table 1 presents the root mean square error and mean absolute error of the inversion results relative to the true model. As shown in Table 1, the weighted instantaneous

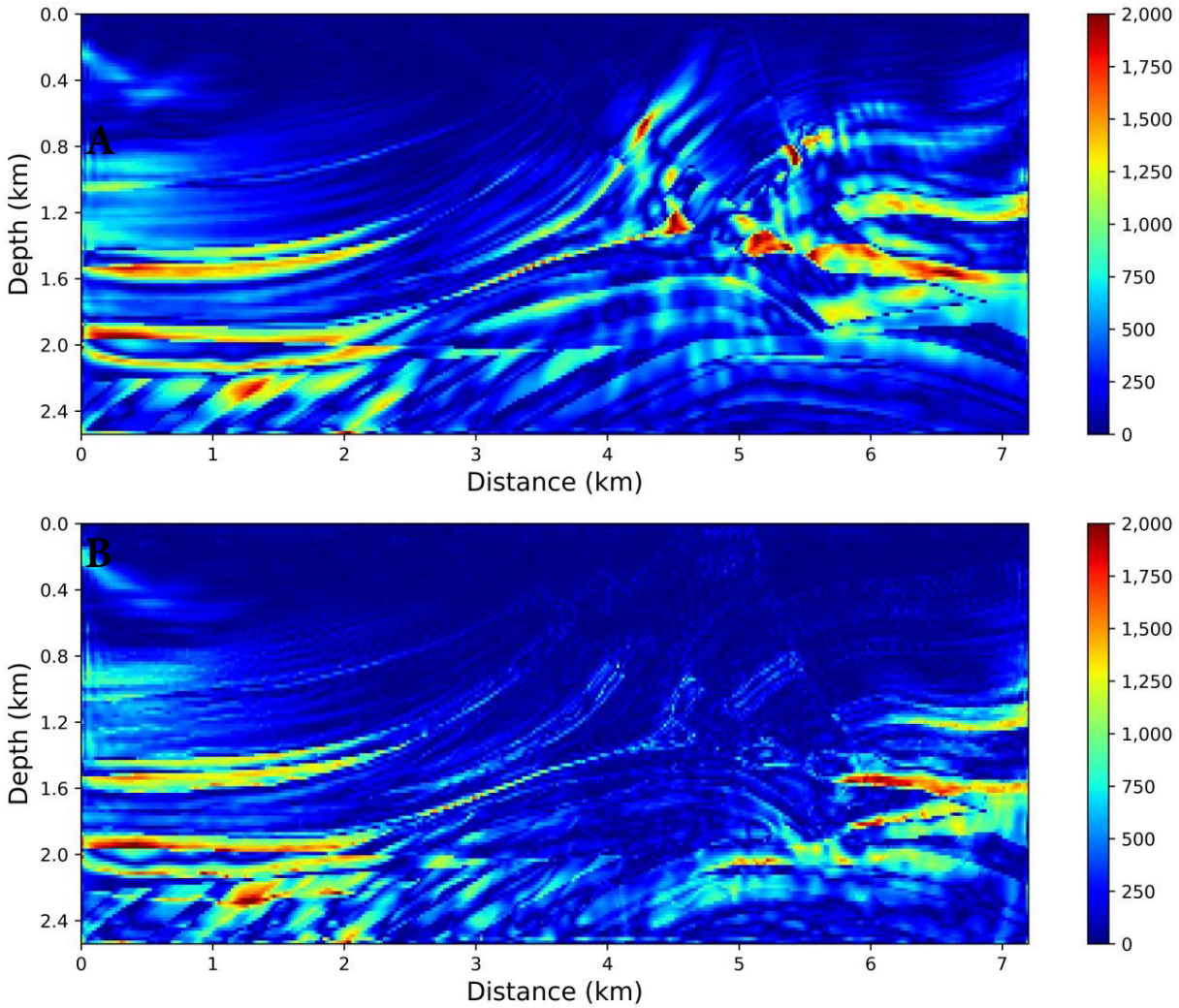

Figure 9. Comparison of absolute errors between the inversion results and the true model. (A) traditional full waveform inversion; (B) weighted instantaneous phase inversion.

Table 1. Root mean square error and mean absolute error of inversion results for the Overthrust and Marmousi models compared to the true model

Model	Method	Weighted instantaneous phase inversion	Traditional FWI
Overthrust	RMSE	214.45	423.76
	MAE	144.09	307.15
Marmousi	RMSE	413.42	485.17
	MAE	253.91	328.46

Abbreviations: FWI: Full waveform inversion; MAE: Mean absolute error; RMSE: Root mean square error.

phase inversion produces smaller errors and achieves results closer to the true model.

3.3 Noise testing

Seismic data collected in the field often include noise, which can significantly affect FWI results. To evaluate the noise resistance of the weighted instantaneous phase inversion method, Gaussian noise was added to the observed data. Figure 10A shows the noise-free seismic records from the Overthrust model, while Figure 10B shows the noisy seismic records obtained by adding Gaussian noise. These noisy observations were then used to perform FWI.

Figure 11 shows the inversion results using the weighted instantaneous phase inversion method. Figure 11A and 11B present the inversion results without and with noise, respectively. The comparison reveals the method's strong noise resistance. In Figure 11B, despite the noise interference, key geological structures remain clearly identifiable. In contrast, Figure 11A demonstrates higher resolution and inversion accuracy in the absence of noise. Overall, the weighted instantaneous phase inversion method maintains the core quality of inversion results under noisy conditions, confirming its robustness and reliability for real exploration scenarios.

Gaussian noise was also added to the seismic records of the Marmousi model. Figure 12 compares the seismic data before and after noise addition. Figure 13 shows the inversion results for the Marmousi model using the weighted instantaneous phase method. Comparing the noisy inversion results (Figure 13B) with the noise-free results (Figure 13A) shows that, even for the complex

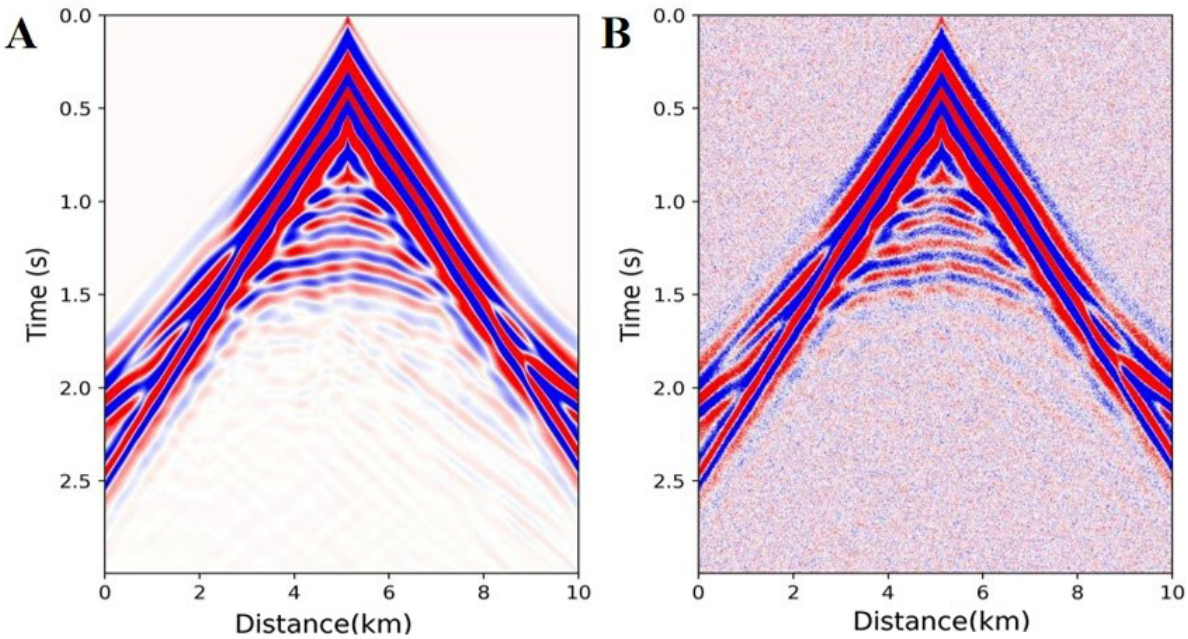

Figure 10. Seismic records from the Overthrust model: (A) Observed data without noise and (B) observed data with added Gaussian noise

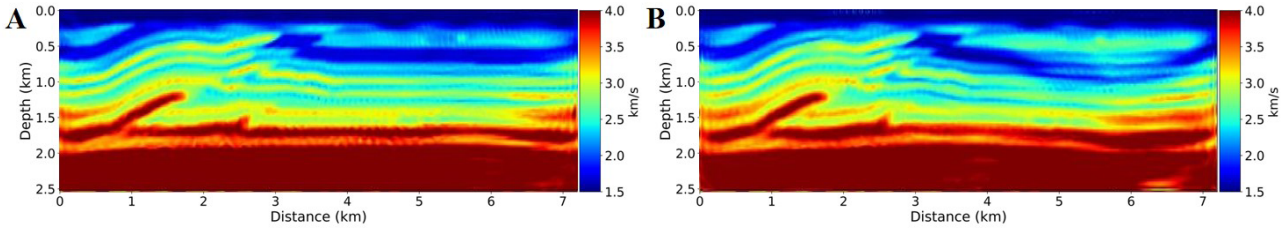

Figure 11. Weighted instantaneous phase inversion results for the Overthrust model: (A) Inversion result without noise and (B) inversion result with added Gaussian noise

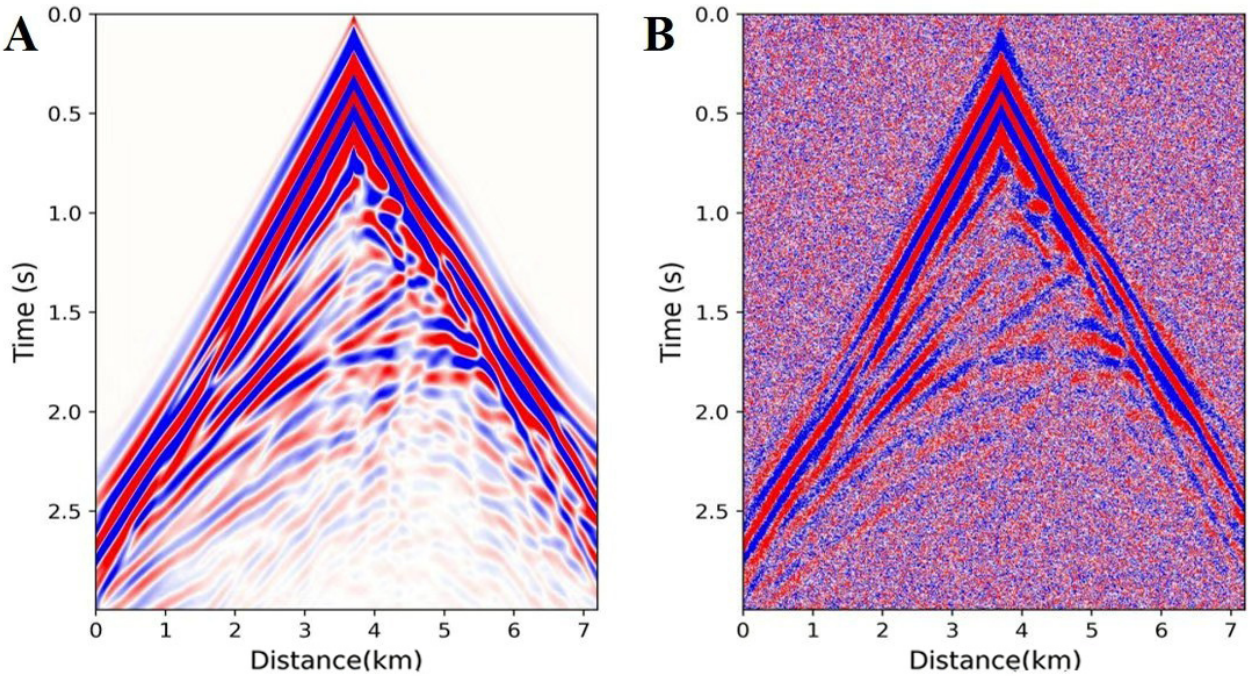

Figure 12. Seismic records from the Marmousi model: (A) Observed data without noise and (B) observed data with added Gaussian noise

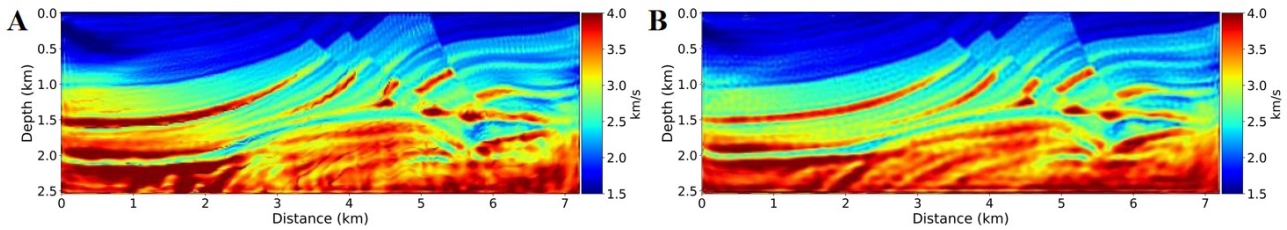

Figure 13. Weighted instantaneous phase inversion results for the Marmousi model: (A) Inversion result without noise and (B) inversion result with added Gaussian noise

Marmousi model, the method demonstrates strong noise resistance, successfully recovering an accurate subsurface velocity model.

4. Conclusion

This study introduces an AD-based weighted instantaneous phase inversion method to address the cycle-skipping problem in FWI. The relationship between the instantaneous phase information of seismic signals and the subsurface velocity structure is more linear, while waveform information facilitates high-resolution results and maintains algorithmic stability. Therefore, within the AD-based waveform inversion framework, we combine instantaneous phase information with waveform information to construct a weighted instantaneous phase objective function, achieving a balanced trade-off between inversion robustness and resolution. To utilize instantaneous phase information effectively, we avoided phase wrapping issues using exponential phase information. Numerical tests indicate that the early incorporation of instantaneous phase information provides a reasonable initial velocity model, mitigates the cycle-skipping problem, and significantly enhances overall inversion performance.

Acknowledgments

None.

Funding

This research was funded by the National Natural Science Foundation of China (Grant No. 42104116).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Mengzi Wang, Yong Hu
Formal analysis: Chunyuan Shi, Zhihan Zhang
Investigation: Mengzi Wang, Yong Hu, Zhihan Zhang
Methodology: Mengzi Wang, Yong Hu
Software: Mengzi Wang, Zhihan Zhang
Supervision: Mengzi Wang, Yong Hu
Writing–original draft: Mengzi Wang, Chunyuan Shi
Writing–review & editing: Xiaoyuan Liao

Availability of data

Data are available from the corresponding author upon reasonable request.

References

1. Lailly P. The seismic inverse problem as a sequence of before stack migrations. In: Bednar JB, Robinson E, Weglein A, eds. *Conference on Inverse Scattering—Theory and Application*. Philadelphia, PA: SIAM; 1983:206-220.
2. Tarantola A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*. 1984;49(8):1259-1266.

   doi: 10.1190/1.1441754
3. Fichtner A, Bunge HP, Igel H. The adjoint method in seismology: I. Theory. *Phys Earth Planet Inter*. 2006;157(1-2):86-104.

   doi: 10.1016/j.pepi.2006.03.016
4. Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics. *Geophysics*. 2009;74(6):WCC1-WCC26.

   doi: 10.1190/1.3238367
5. Plessix RE. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys J Int*. 2006;167(2):495-503.

   doi: 10.1111/j.1365-246X.2006.02978.x
6. Luo JR, Wang BF. Initial model building in time domain elastic full waveform inversion using the instantaneous phase information. *Prog Geophys*. 2018;33(6):2435-2440. [In Chinese].

   doi: 10.6038/pg2018CC0215
7. Sun SY, Hu GH, He BH, Du ZY, Xu WC. Reflection waveform inversion and its application to onshore seismic data. *Prog Geophys*. 2021;36(6):2566-2572. [In Chinese].

   doi: 10.6038/pg2021EE0462
8. Li ZC, Wang ZY, Huang JP, Cui C. Multi-scale full waveform inversion based on gradient decomposition in wavenumber domain. *Chin J Geophys*. 2022;65(7):2693-2703. [In Chinese].

   doi: 10.6038/cjg2022P0508
9. Wang QQ, Song P, Hua QF, *et al*. Full waveform inversion based on Adam algorithm with optimized parameters. *Chin J Geophys*. 2023;66(11):4654-4663. [In Chinese].

   doi: 10.6038/cjg2022Q0879
10. Li JS, Liu WK, Liang YX, Wu JY, Li ZG. Time-domain full waveform inversion based on high-order amplitude information. *Prog Geophys*. [In Chinese]. 2023;38(4):1603-1609.

    doi: 10.6038/pg2023GG0129
11. Kang PF, Hu Y, Liu RD, *et al*. Local-scale frequency-wavenumber domain phase inversion. *Prog Geophys*. 2025;40(1):155-165. [In Chinese].

    doi: 10.6038/pg2025II0034
12. Shah NK, Warner MR, Washbourne JK, Guasch L, Umpleby AP. A phase-unwrapped solution for overcoming a poor starting model in full wavefield inversion. In: *Proceedings*

of the 74th EAGE Conference and Exhibition Incorporating EUROPEC 2012; Copenhagen, Denmark: Jun 4-7, 2012; European Association of Geoscientists & Engineers; 2012:cp-293-00460.

doi: 10.3997/2214-4609.20148455

13. Choi Y, Alkhalifah T. Unwrapped phase inversion with an exponential damping. *Geophysics*. 2015;80(5):R251-R264.

doi: 10.1190/geo2014-0498.1

14. Luo JR, Wu RS, Gao FC. Time-domain full waveform inversion using instantaneous phase information with damping. *J Geophys Eng*. 2018;15(3):1032-1041.

doi: 10.1088/1742-2140/aaa984

15. Hu Y, Wu RS, Han LG, Zhang P. Joint multiscale direct envelope inversion of phase and amplitude in the time–frequency domain. *IEEE Trans Geosci Remote Sens*. 2019;57(7):5108-5120.

doi: 10.1109/TGRS.2019.2896936

16. Zhu HJ, Fomel S. Building good starting models for full-waveform inversion using adaptive matching filtering misfit. *Geophysics*. 2016;81(5):U61-U72.

doi: 10.1190/geo2015-0596.1

17. Warner M, Guasch L. Adaptive waveform inversion: Theory. *Geophysics*. 2016;81(6):R429-R445.

doi: 10.1190/geo2015-0387.1

18. Hu Y, Han LG, Yu JL, Chen RD. Time-frequency domain multi-scale full waveform inversion based on adaptive non-stationary phase correction. *Chin J Geophys*. 2018;61(7):2969-2988. [In Chinese].

doi: 10.6038/cjg2018L0421

19. Liu JG, Xu KL. Automatic differentiation and its applications in physics simulation. *Acta Phys Sin*. 2021;70(14):149402. [In Chinese].

doi: 10.7498/aps.70.20210813

20. Ketkar N, Moolayil J. Automatic Differentiation in Deep Learning. In: *Deep Learning with Python*. Berkeley, CA: Apress; 2021:133-145.

doi: 10.1007/978-1-4842-5364-9_4

21. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.

doi: 10.1038/nature14539

22. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press; 2016.

23. Nolan JF. *Analytical differentiation on a digital computer* [master's thesis]. Cambridge, MA: Massachusetts Institute of Technology, Department of Mathematics; 1953. https://hdl.handle.net/1721.1/12297

24. Speelpenning B. *Compiling Fast Partial Derivatives of Functions Given by Algorithms* [PhD thesis]. Urbana, IL: University of Illinois at Urbana-Champaign, Department of Computer Science; 1980. Report No. UIUC DS-R-80-1002.

doi: 10.2172/5254402

25. Kasim MF, Vinko SM. ξ-torch: differentiable scientific computing library. *arXiv*. Preprint posted online October 6, 2020.

doi: 10.48550/arXiv.2010.01921

26. Sambridge M, Rickwood P, Rawlinson N, Sommacal S. Automatic differentiation in geophysical inverse problems. *Geophys J Int*. 2007;170(1):1-8.

doi: 10.1111/j.1365-246X.2007.03400.x

27. Tan L, Brytik V, Baumstein A, Hinkley D. Verification of gradient and hessian computation for full wavefield inversion using automatic differentiation. In: Levin S, ed. SEG Technical Program Expanded Abstracts 2010; SEG International Exposition and 80th Annual Meeting; 17–22 October 2010; Denver, CO, USA. Society of Exploration Geophysicists; 2010:2762-2766.

doi: 10.1190/1.3513417

28. Liao W. An accurate and efficient algorithm for parameter estimation of 2D acoustic wave equation. *Int J Appl Phys Math*. 2011;1(2):96-100.

doi: 10.7763/IJAPM.2011.V1.19

29. Richardson A. Seismic full-waveform inversion using deep learning tools and techniques. *arXiv*. Preprint posted online January 22, 2018.

doi: 10.48550/arXiv.1801.07232

30. Cao D, Liao W. A computational method for full waveform inversion of crosswell seismic data using automatic differentiation. *Comput Phys Commun*. 2015;188:47-58.

doi: 10.1016/j.cpc.2014.11.002

31. Zhu W, Xu K, Darve E. A general approach to seismic inversion with automatic differentiation. *Comput Geosci*. 2021;151:104751.

doi: 10.1016/j.cageo.2021.104751

32. Song C, Wang Y, Richardson A, Liu C. Weighted envelope correlation-based waveform inversion using automatic differentiation. *IEEE Trans Geosci Remote Sens*. 2023;61:1-11.

doi: 10.1109/TGRS.2023.3300127

33. Liu F, Li H, Zou G, Li J. Automatic differentiation-based full waveform inversion with flexible workflows. *J Geophys Res Mach Learn Comput*. 2025;2(1):e2024JH000542.

doi: 10.1029/2024JH000542