Elastic FWI Without Low Frequencies Based on n-th Power Operation and Convolved Wavefields

Pan Zhang, Liguo Han, Jianwei Zhao, Yuchen Yin
College of Geo-exploration Science and Technology, Jilin University, Changchun, China
E-mail: zhangpan2309@163.com

Abstract. The elastic full-waveform inversion (FWI) can use the recorded multi-component seismic data to construct high-precision multi-parameter models of the subsurface media such as P- and S-wave velocity models. However, due to the reasons such as data quality and algorithm limitations, there are still many problems in the promotion and application of elastic FWI method. Aiming at alleviating the influence of low-frequency data absence on the inversion results, we propose a robust elastic FWI method based on the n-th power operation. The n-th power of the seismic data can compress the time-domain waveform and expand its frequency-band. The FWI objective function constructed using the n-th power wavefields shows better convexity. By successively lowering the power during the inversion, we can realize a new multiscale FWI strategy, which is also a data-domain layer-stripping strategy. Seismic data will be more sensitive to the source wavelet errors after the n-th power operation. To mitigate this problem, we propose a robust objective function for elastic FWI using the n-th power operation and the convolved wavefields. Finally, the validity of the method is verified by numerical examples.

1. Introduction
The full-waveform inversion (FWI) method can simultaneously use the kinematics and dynamics information contained in seismic signals and is a high-precision subsurface parameter construction method. FWI was first proposed in acoustic media by Laily (1983) and Tarantola (1984) [1-2]. They regarded the seismic velocity inversion problem as an optimization problem, using the L2-norm of data residual as the objective function. Then Tarantola (1986) and Mora (1987) extended this method to elastic media to invert multi-parameters [3-4]. Crase et al. (1990) first applied elastic full-waveform inversion method to real data [5]. In recent years, more and more researches have been done on elastic full-waveform inversion, and many valuable advances have been made [6-9]. Since FWI uses local optimization algorithms, it has a strong dependence on the initial model. The objective function of low-frequency data has fewer local minima, so effectively use of low-frequency data can help to overcome the cycle-skipping problem of FWI [10]. However, the low-frequency band of real seismic data always has a very low signal to noise ratio. To overcome this problem, some special FWI methods in the absence of low-frequency data were proposed [11-14].

In this paper, we propose a robust elastic FWI method, which can deal with the inversion in the cases of low-frequency data absence and inaccurate source wavelet estimation. Firstly, we introduce the principle of frequency extension by n-th power operation. Then we propose a multiscale FWI method based on n-th power operation. To solve the problem of source wavelet sensitivity after n-th power operation, we propose a robust elastic FWI method based on convolved wavefields. Finally, numerical tests are conducted to demonstrate the effectiveness of the proposed method.
2. Method

2.1. Frequency extension by n-th power operation

From the signal processing point of view, we know that the narrower the waveform in the time domain, the broader its spectrum in the frequency domain, and vice versa. Therefore, we can compress the seismic waveform in the time domain to expand its spectrum. The extended low-frequency information can be extracted and used for FWI. The n-th power operation is one of the methods to compress the waveform. The compressed seismic data can be expressed as

\[ h_n = d^n \]  

(1)

where \( n \) denotes the order of power, \( d \) is the original seismic data, \( h_n \) is the compressed seismic data. In the following, we use a source wavelet to see the low-frequency expanding effects. The original source wavelet is a Ricker wavelet after high-pass filtering (Figure 1a). The dominant frequency of the source wavelet is about 20Hz. From its spectrum (Figure 1e), we know that the information below 7Hz is missing. We take the n-th power of the original source wavelet. The results corresponding to \( n=2, 3, 11 \) are shown in Figure 1b, 1c, and 1d. Their spectra are shown in Figure 1f, 1g, and 1h, respectively. When \( n=2 \), the value of the result is positive, which is equivalent to the energy value of the wavelet. In this case, there is a large zero-frequency value, and a large amount of low-frequency information is extended. When \( p=3 \), some low-frequency information is also extended, and the waveform has polarity. When \( n \) is large, such as 11, the compressed result approaches to a delta function, whose spectrum is infinitely broad. In this test, it is clear that taking the n-th power can reconstruct some ultra-low-frequency information.

![Figure 1](image-url)  

**Figure 1.** The n-th power operation result of a source wavelet. (a)-(d) correspond to the results when \( n=1, 2, 3, 11 \), respectively. (e)-(h) are the spectra of (a)-(d).

2.2. Physical meaning of n-th power data

In the above discussion, we mainly explain the low-frequency extension of the n-th power operation mathematically. Now we give a physical explanation of the extended low-frequency information using the modulation signal model [11] by comparing the envelope data and the n-th power data. The same with Wu et al. (2014), we take \( n=2 \) for example.

In Wu et al. (2014)’s work, with the help of modulation signal model, the whole seismogram can be expressed as

\[ p(t;\mathbf{x}_s,\mathbf{x}_r) = g^d (t;\mathbf{x}_s,\mathbf{x}_r) s(t) + G_{ss} (t;\mathbf{x}_s,\mathbf{x}_r) \gamma_{\theta} (t) \]  

(2)
where \( x_r \) is the receiver position and \( x_s \) is the source position, \( t \) denotes time, \( g^d \) is Green’s function for the direct arrival and is assumed to be a much slower time-varying function than the source wavelet \( s(t) \). \( G_{sr} \) is a propagator that related to the forward-scattering Green’s function, \( \gamma^w \) is the reflectivity series with the source wavelet signature.

The propagator functions \( g^d \) and \( G_{sr} \) are low-pass functions (slowly varying function) (under the forward-scattering approximation) which contains the long-wavelength information of the velocity. To extract \( g^d \) and \( G_{sr} \) from equation (2), we need a demodulation operator. The envelope operator is proved to be an effective demodulation operator (Wu et al., 2014). The envelope squared seismic traces can be approximated by

\[
E^2(t) = p^2(t) + p_w^2(t)
\]

\[
= \left( g^d (t) \right)^2 E^2 [s(t)] + G_{sr}^2(t) E^2 [\gamma^w(t)]
\]

(3)

where \( p_H \) is the Hilbert transform of \( p \), \( E^2[s(t)] \) is the squared envelope of the source wavelet, \( E^2[\gamma^w(t)] \) is the squared envelope of \( \gamma^w(t) \), \( E^2[s(t)] \) and \( E^2[\gamma^w(t)] \) are both low-frequency components. So they keep the long-wavelength information riding on the propagators.

In the derivation of equation (3), we can easily know that two cross-terms are neglected. Now we take second order of the seismic traces, as

\[
p^2(t) = \left( g^d (t) \right)^2 (s(t))^2 + G_{sr}^2(t) (\gamma^w(t))^2
\]

(4)

Then we take low-pass filtering to the squared data, as

\[
l_p[p^2(t)] = \left( g^d (t) \right)^2 l_p (s(t))^2 + G_{sr}^2(t) l_p(\gamma^w(t))^2
\]

(5)

where \( l_p[.] \) denotes low-pass filtering. Comparing equation (5) with equation (3), we know that \( l_p[(s(t))^2] \) and \( l_p[(\gamma^w(t))^2] \) are also both low-frequency components, and the long-wavelength information riding on the propagators are kept. What is more, in the deriving of equation (4), we only neglect one cross-term.

Like the physical meaning of the \( n \)-th power envelope data, the physical meaning of the \( n \)-th power data can be explained similarly as above. From the above discussion, we know that by taking \( n \)-th power of the data, the long-wavelength responses of the subsurface media can be extracted.

2.3. Multiscale FWI based on \( n \)-th power operation

In order to utilize the low-frequency information extended by the high-order power operation for FWI, we construct a new objective function using the \( n \)-th power wavefields, which can be written as

\[
\sigma_n(v) = \frac{1}{2} \| u^n - d^n \|^2_2
\]

(6)

where \( u \) and \( d \) denotes synthetic data and observed data, respectively. According to the chain rule, the derivative of \( \sigma_n \) to \( v \) can be calculated as

\[
\frac{\partial \sigma_n}{\partial v} = \left( \frac{\partial u}{\partial v} \right)^T (u^n - d^n) u^{n-1}
\]

(7)

where \( T \) denotes transposition. According to the adjoint-state method (Tarantola 1984), the gradient can be calculated by the zero-lag cross-correlation of the forward-propagated source wavefield and the back-propagated residual wavefield. The residual for equation (7) is

\[
f_e = (u^n - d^n) u^{n-1}
\]

(8)
To test the behavior of the objective function defined in equation (6), we construct a modified Marmousi velocity model (Figure 2a). The source is located at z=0m and x=960m, and the source wavelet lacks low-frequency information (Figure 1a). We decompose the Marmousi model into a background model (Figure 2b) and a perturbation model (Figure 2c). The velocity of the background model increases linearly from shallow to deep, and the minimum and maximum velocities are set according to the true model. The perturbation model is calculated by subtracting the background model from the true model. To conduct the test, we change the background model and the perturbation model continuously to form a series of new models, which are regarded as inverted models [15]. We simulate on the true model and on the inverted model and calculate the objective functions corresponding to n=1, 2, 3, respectively. The contours of the objective function are shown in Figure 3. There are two apparent local minima in Figure 3a, which means it is hard for conventional FWI to converge to the true solution when the initial model is bad. In Figure 3b, the behavior of the objective function is more convex than Figure 3a. However, there are still two local minima, although the distribution is different from Figure 3a. The contour of Figure 3c shows the most potential to overcome the local minima of FWI. There is only one minimum, which indicates the global minimum. From the above analysis, we know that when using a higher order, the objective function is more convex. So our multiscale FWI strategy is to invert the wavefield from high-order to low-order successively. Also, when the order of power is higher, the energy in seismic data concentrate shallower. So the proposed multiscale inversion strategy is also a data-domain layer-stripping method.

\[ \sigma = \frac{1}{2} \| (u') - (d') \| + \frac{1}{2} \| (u') - (d') \| \]  

where \( u' \) and \( d' \) denotes z component of synthetic and observed data, respectively; \( u' \) and \( d' \) denote x component of synthetic and observed data, respectively. After high-order power operation, the seismic data is more sensitive to the errors of source wavelet. To mitigate the source wavelet dependent of n-th power-based FWI, we construct a new objective function using the convolved wavefields. For simplicity, we only show the derivation of z component data. The new objective function can be written as

\[ \sigma = \frac{1}{2} \| (u') - (d') \| + \frac{1}{2} \| (u') - (d') \| \]  

Figure 2. The Marmousi model and its decomposition. (a) The Marmousi model. (b) The background model. (c) The perturbation model.

Figure 3. Contour of the new objective function. (a) The contour of the objective function of conventional FWI (n=1). (b) The contour of the new objective function with n=2. (c) The contour of the new objective function with n=3.

2.4. Robust elastic FWI based on n-th power operation and convolved wavefields

When considering the elastic multi-component data, the objective function using n-th power operation can be written as

\[ \sigma = \frac{1}{2} \| (u') - (d') \| + \frac{1}{2} \| (u') - (d') \| \]  

where \( u' \) and \( d' \) denotes z component of synthetic and observed data, respectively; \( u' \) and \( d' \) denote x component of synthetic and observed data, respectively. After high-order power operation, the seismic data is more sensitive to the errors of source wavelet. To mitigate the source wavelet dependent of n-th power-based FWI, we construct a new objective function using the convolved wavefields. For simplicity, we only show the derivation of z component data. The new objective function can be written as
\[
\sigma_i = \frac{1}{2} \sum_{i} \sum_{j} \left| \left[ \left( u_{i,j} * d_{i,j} \right)^n - \left( d_{i,j} * u_{i,j} \right)^{n-1} \right] \right| 
\]

(10)

where \( i, j \) and \( k \) denote the index of source, receiver and reference trace, respectively; \( n_s \) and \( n_r \) denote source and receiver number, respectively; * denotes convolution. Expanding the seismic data in equation (10) as the convolution of Green’s function and source wavelet term, it is easy to see the influence of source wavelet errors on the inversion process is avoided. According to the adjoint-state method and the chain rule, the gradient of Eq. (6) can be calculated by the zero-lag cross-correlation of the forward propagated source wavefield and the backwards propagated residual wavefield. The residual can be written as

\[
A_1 = d_{i,j} \otimes \left[ \left( u_{i,j} * d_{i,j} \right)^n - \left( d_{i,j} * u_{i,j} \right)^{n-1} \right]
\]

(11)

\[
A_2 = d_{i,j} \otimes \left[ \left( u_{i,j} * d_{i,j} \right)^n - \left( d_{i,j} * u_{i,j} \right)^{n-1} \right]
\]

(12)

where \( A_1 \) is the residual at traces except the reference trace, \( A_2 \) is the residual at reference trace.

3. Numerical examples

The true Vp and Vs models are shown in Figure 4. The initial models are linear-gradient models (Figure 5). The accurate source wavelet (red line in Figure 6) is used to simulate the observed data, and the inaccurate source wavelet (blue line in Figure 6) is used to simulate the synthetic data. Both the two wavelets lack low-frequency information below 5Hz. Since there are significant source wavelet errors, the FWI algorithm based on equation (6) cannot converge. The inversion result by conventional elastic FWI is shown in Figure 7, which illustrates the inversion suffers serious cycle-skipping. The inversion result of the proposed method is shown in Figure 8. It can be seen that the robust proposed method recovers the velocity model from shallow to the deep gradually, even when the estimated source wavelet has significant errors and lacks low-frequency information. The final inversion result (Figure 8e and Figure 8f) is much better than the conventional inversion result (Figure 7a and Figure 7b).

Figure 4. The true velocity model. (a) The true Vp model. (b) The true Vs model.

Figure 5. The initial velocity model. (a) The initial Vp model. (b) The initial Vs model.

Figure 6. Accurate and inaccurate source wavelets.
4. Conclusions

The n-th power operation can not only extend the frequency-band of seismic signals nonlinearly but also help to realize a data-domain layer-stripping inversion. With the modulation signal model, we know that by taking n-th power of the data, the long-wavelength responses of the subsurface media can be extracted. We establish a robust objective function for elastic FWI using the n-th power operation and the convolved wavefields. The proposed method has a weak dependence on the initial model and low-frequency information, and can also avoid the effects of source wavelet errors on the elastic FWI process.

Acknowledgements

We thank the financial support of the "Thirteenth Five-Year Plan" Science and Technology Project of Education Department of Jilin Province (No. JJKH20201001KJ), National Natural Science Foundation of China (No. 41674124) and the National Key Research and Development Project of China (No. 2016YFC0600301).

References

[1] Laily P 1983 The seismic inverse problem as a sequence of before-stack migrations: Conference on inverse scattering Theory and application 206-220
[2] Tarantola A 1984 Inversion of seismic reflection data in the acoustic approximation Geophysics 49 1259-1266
[3] Tarantola A 1986 A strategy for nonlinear elastic inversion of seismic reflection data Geophysics 51 1893-1903
[4] Mora P 1987 Nonlinear two-dimensional elastic inversion of multioffset seismic data Geophysics 52 1211-1228
[5] Crase E, Pica A, Noble M, et al. 1990 Robust elastic nonlinear waveform inversion: Application to real data Geophysics 55 527-538
[6] Zhang Q, Zhou H, Li Q, et al. 2016 Robust source-independent elastic full-waveform inversion in the time domain Geophysics 81 R29-R44
[7] Zhang P, Han L, Gong X, et al. 2018 Multi-source elastic full waveform inversion based on the anisotropic total variation constraint Chinese Journal of Geophysics (in Chinese) 61 716-732
[8] Liu Y, Huang X, Wan X, et al. 2019 Elastic multi-parameter full-waveform inversion for anisotropic media Chinese Journal of Geophysics (in Chinese) 62 1809-1823
[9] Yang T, Zhang H and Shi C 2019 Wavelet-independent elastic wave full waveform inversion
hybrid domain Oil Geophysical Prospecting 54 348-355

[10] Bunks C, Saleck F M, Zaleski S, et al. 1995 Multiscale seismic waveform inversion Geophysics 60 1457-1473

[11] Wu R S, Luo J and Wu B 2014 Seismic envelope inversion and modulation signal model Geophysics 79 WA13-WA24

[12] Liu Y, He B, et al. 2018 Full-intensity waveform inversion Geophysics 83 R649-R658

[13] Li Y, Choi Y, Alkhalifah T, et al. 2018 Full-waveform inversion using a nonlinearly smoothed wavefield Geophysics 83 R117-R127

[14] Zhang P, Xing Z and Hu Y 2019 Velocity construction using active and passive multi-component seismic data based on elastic full waveform inversion Chinese Journal of Geophysics (in Chinese) 62 3974-3987

[15] Luo J and Wu R S 2015 Seismic envelope inversion: reduction of local minima and noise resistance Geophysical Prospecting 63 597-614