L_{1-2} minimization for exact and stable seismic attenuation compensation

Yufeng Wang,[1] Xiong Ma,[1] Hui Zhou[1] and Yangkang Chen[2]

[1]*State Key Lab of Petroleum Resources and Prospecting, Key Lab of Geophysical Exploration of CNPC, China University of Petroleum-Beijing, Changping* 102249 *Beijing, China. E-mail:* huizhou@cup.edu.cn
[2]*National Center for Computational Sciences, Oak Ridge National Laboratory, One Bethel Valley Road, Oak Ridge, TN 37831-6008, USA*

SUMMARY
Frequency-dependent amplitude absorption and phase velocity dispersion are typically linked by the causality-imposed Kramers–Kronig relations, which inevitably degrade the quality of seismic data. Seismic attenuation compensation is an important processing approach for enhancing signal resolution and fidelity, which can be performed on either pre-stack or post-stack data so as to mitigate amplitude absorption and phase dispersion effects resulting from intrinsic anelasticity of subsurface media. Inversion-based compensation with L_1 norm constraint, enlightened by the sparsity of the reflectivity series, enjoys better stability over traditional inverse Q filtering. However, constrained L_1 minimization serving as the convex relaxation of the literal L_0 sparsity count may not give the sparsest solution when the kernel matrix is severely ill conditioned. Recently, non-convex metric for compressed sensing has attracted considerable research interest. In this paper, we propose a nearly unbiased approximation of the vector sparsity, denoted as L_{1-2} minimization, for exact and stable seismic attenuation compensation. Non-convex penalty function of L_{1-2} norm can be decomposed into two convex subproblems via difference of convex algorithm, each subproblem can be solved efficiently by alternating direction method of multipliers. The superior performance of the proposed compensation scheme based on L_{1-2} metric over conventional L_1 penalty is further demonstrated by both synthetic and field examples.

Key words: Inverse theory; Seismic attenuation; Wave propagation.

INTRODUCTION

Amplitude absorption and phase dispersion caused by intrinsic anelasticity of subsurface media inevitably degrade the quality of seismograms, decrease the resolution of migrated images and eventually affect the reliability of seismic interpretation. Such frequency-dependent attenuating effects are typically characterized by an empirical formulation wherein the wavenumber is a complex function of frequency, whose real and imaginary parts separately signify the phase velocity $c_p(\omega)$ and attenuation coefficient $\alpha(\omega)$. Under the assumption of linearity and causality, Kramers–Kronig (K-K) relations impose restrictions on both attenuation and dispersion in the frequency domain (Waters *et al.* 2003, 2005). Based on K-K relations, some widely used attenuation models have been established for seismology, such as the Kolsky–Futterman model (Kolsky 1956; Futterman 1962), the power-law model (Strick 1967; Azimi 1968; Szabo 1994, 1995) and Kjartansson's constant-Q model (Kjartansson 1979; Zhu & Carcione 2013). Alternative ways to approximate frequency-dependent attenuation are developed on the basis of relaxation mechanisms (Liu *et al.* 1976; Carcione *et al.* 1988) and their fractional generalizations derived from the conservation of

momentum and stress–strain constitutive relation (Holm & Näsholm 2013; Wang 2015). Attenuation models mentioned above can be further verified by laboratory measurements (Wuenschel 1965; Wang *et al.* 2007) and field experiments (McDonal *et al.* 1958; Li *et al.* 2016a).
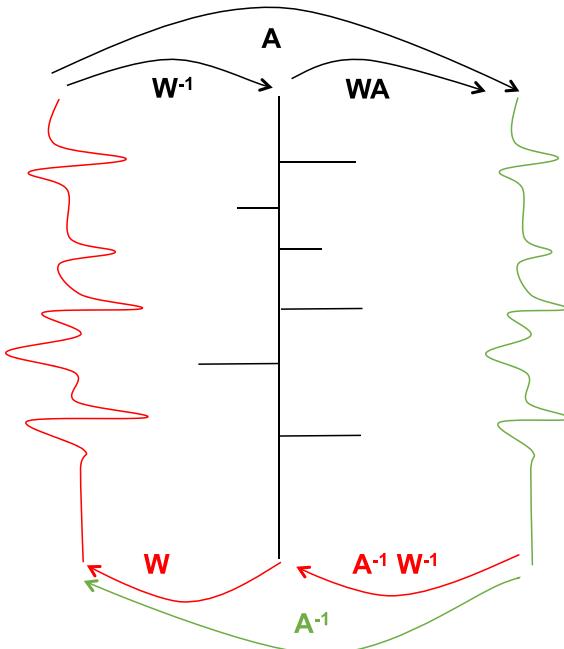
Over the past three decades, great efforts have been devoted to explore efficient and stable algorithms for enhancing the resolution and quality of seismic data. In general, attenuation compensation in geophysics can be roughly classified into two categories: seismic record-based compensation and propagation-based compensation, each has its own merits and demerits. The former category of compensation methods include time-varying deconvolution (Clarke 1968; Griffiths *et al.* 1977; Margrave *et al.* 2011), time-variant spectral whitening (Yilmaz 2001) and inverse Q filtering (Hargreaves & Calvert 1991; Wang 2002, 2006). These record-based compensation schemes are directly performed on attenuated seismic records in the time or frequency domain, which exhibit higher efficiency, greater flexibility and better simplicity. In contrast, propagation-based compensation approaches are carried out during wave propagation, which include Q-compensated one-way wave equation migration (Dai & West 1994; Mittet *et al.* 1995; Wang & Guo 2004b;

Mittet 2007; Zhang *et al.* 2012), *Q*-compensated reverse time migration (Zhang *et al.* 2010; Zhu *et al.* 2014; Li *et al.* 2016b; Sun *et al.* 2016) and *Q*-compensated Gaussian beam migration (Bai *et al.* 2016). It is more physically consistent to compensate amplitude absorption and phase dispersion during migration, since these effects associated with anelasticity occur during the wave propagation (Zhang *et al.* 2010; Zhu *et al.* 2014; Sun *et al.* 2015). However, amplitude compensation during propagation suffers from severe instability appearing as boosting high-frequency ambient noise (Sun & Zhu 2015; Xue *et al.* 2016; Wang *et al.* 2017a).

Inverse *Q* filtering, also known as attenuation compensation technique, has been considered as an effective method to eliminate the attenuating effect (Hargreaves & Calvert 1991; Wang 2002). Nevertheless, direct amplitude compensation via exponential amplification is unstable by its nature, which, unless very carefully designed, unavoidably boosts high-frequency noise. In the past decade, inversion-based attenuation compensation schemes have been increasingly applied to enhance the vulnerable stability of traditional inverse *Q* filtering. Zhang & Ulrych (2007) formulate the compensation of attenuation as a least-squares problem and impose regularization by means of Bayes' theorem. Wang (2011) reduces the compensation problem to an inversion problem and achieves it by Tikhonov regularization, where the L_2 norm constraint penalty function gives smooth and stable solution at the expense of the fidelity and accuracy. Inspired by the sparsity of the reflectivity series, inversion-based compensation schemes with L_1 minimization are proposed to achieve either in the frequency domain (Oliveira & Lupinacci 2013; Chai *et al.* 2014, 2017) or time domain (Wang & Chen 2014; Li *et al.* 2015). This L_1-norm penalized least-squares problem can be also called lasso (Tibshirani 1996), which can be solved by many state-of-the-art algorithms that are widely used and rapidly developed in compressed sensing (CS, Candes *et al.* 2006; Donoho 2006; Chen & Fomel 2015; Chen 2016; Wang *et al.* 2017b; Zhou *et al.* 2017).

Mathematically, minimizing the L_1 norm is a convex problem, thus computationally tractable. However, it may sometimes yield suboptimal performance due to the biased approximation to L_0 in the sense that L_1 is dominated by entries with large magnitudes, unlike L_0 in which all non-zero entries have equal contributions (Gong *et al.* 2013; Ma *et al.* 2017). To address this issue, many non-convex metrics, interpolated between the L_0 and L_1 norms, have been proposed to better approximate the L_0 norm. They include L_p (Chartrand 2007; Foucart & Lai 2009; Lai *et al.* 2013; Woodworth & Chartrand 2016), log-sum penalty (LSP, Candes *et al.* 2008), truncated L_1 (Wang & Yin 2010; Hu *et al.* 2013), capped L_1 minimization (Zhang 2010b) and minimax concave penalty (Zhang 2010a). The aforementioned non-convex penalties can be solved by iteratively reweighted least squares (Sacchi 1997; Gorodnitsky & Rao 2002; Chartrand & Yin 2008; Lai *et al.* 2013) or multistage convex relaxation (Zhang 2010b). Gong *et al.* (2013) propose a general iterative shrinkage and thresholding algorithm to solve the non-convex optimization problem for a large class of non-convex penalties.
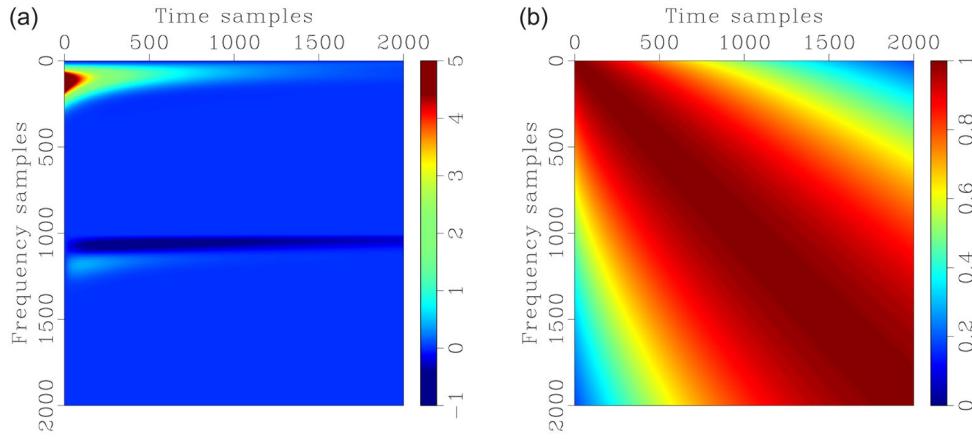
In this paper, we propose an inversion-based seismic attenuation compensation scheme, which can be performed via a two-step strategy, that is, non-stationary sparse reflectivity inversion (NSRI, Chai *et al.* 2014, 2017; Yuan *et al.* 2017) and wavelet convolution procedure. However, the kernel matrix in NSRI suffers from more serious ill-posedness over that of conventional sparse reflectivity inversion. Furthermore, reflectivity inversion from noisy data imposes greater challenge on numerical stability. It is of great importance to seek a sparse penalty function that outperforms the existing L_1 metric in terms of stability and antinoise property. In recent years, the L_{1-2}
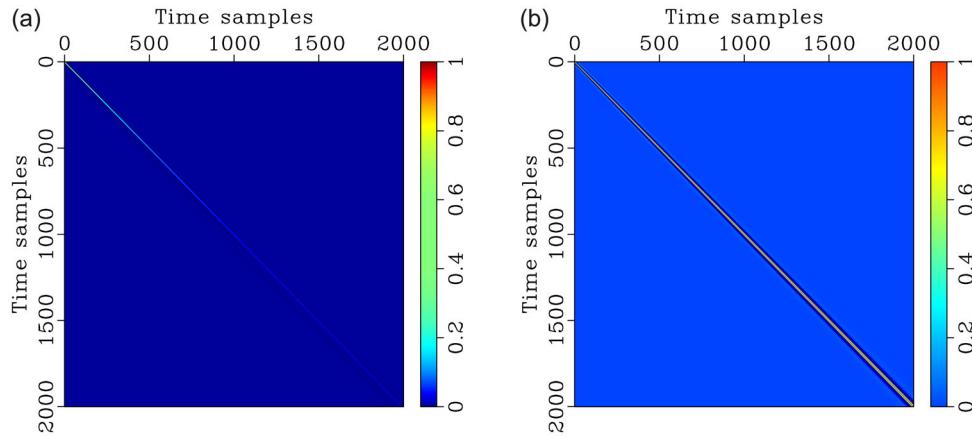

Figure 1. The diagram of the direct and two-step compensation processes, at the bottom of this figure, the green line represents direct compensation method using $\mathbf{A}^{-1}$ and the red lines stand for two-step compensation scheme using operators $\mathbf{A}^{-1}\mathbf{W}^{-1}$ and $\mathbf{W}$.

metric, defined as $\|\mathbf{x}\|_{1-2} = \|\mathbf{x}\|_1 - \|\mathbf{x}\|_2$, has been widely used as a sparsity penalty for sparse recovery in the framework of CS (Esser *et al.* 2013; Yin *et al.* 2015). L_{1-2} minimization exhibits superior performance over other existing metrics when the sensing matrix is highly coherent (Lou *et al.* 2015b). Lou & Yan (2016) develop a fast version of L_{1-2} minimization via incorporating the proximal operator into well-known splitting algorithms such as forward–backward splitting and alternating direction method of multipliers (ADMM, Combettes & Wajs 2005; Combettes & Pesquet 2011; Boyd *et al.* 2011). Ma *et al.* (2017) present a truncated L_{1-2} metric for sparse recovery and rank minimization. In this paper, we adopt L_{1-2} norm for exact and stable seismic attenuation compensation. Non-convex L_{1-2} minimization can be decomposed into two convex subproblems via difference of convex algorithm (DCA, Tao & An 1998; Liu & Pong 2017), each of which can be solved efficiently by ADMM (Yin *et al.* 2015; Ma *et al.* 2017).

The structure of the paper is organized as follows: first, we briefly review the seismic attenuation models and non-stationary convolution. Then, we formulate inversion-based compensation scheme with both L_1 and L_{1-2} minimization in the time domain. After that the solver for L_{1-2} minimization is provided for efficiently implementing seismic attenuation compensation in the framework of the unconstrained least-squares inversion, in which DCA and ADMM algorithms are, respectively, used for decomposing the non-convex problem into two convex subproblems and solving these two convex optimization problems. The superior performance of the proposed compensation scheme based on L_{1-2} metric over conventional L_1 penalty is demonstrated by both synthetic and field examples. In the Discussion section, we investigate wavelet dependence and parameter selection of the proposed method via a series of compensation tests on both 1-D clean and noisy data sets.

Figure 2. Visualization of (a) the frequency-domain matrix **Φ** and (b) its coherence coefficients.

Figure 3. Visualization of (a) the time-domain matrix **Φ̂** and (b) its coherence coefficients.

SEISMIC ATTENUATION MODELS

Frequency-dependent attenuating effects are typically characterized by an empirical formulation wherein the wavenumber is a complex function of frequency, whose real and imaginary parts separately signify the phase velocity $c_p(\omega)$ and attenuation coefficient $\alpha(\omega)$. In this section, we first review the general way for frequency power-law attenuation modeling. Intrinsic attenuation property of the subsurface can be empirically characterized by experimentally established frequency power law over a wide range of frequencies. Once the empirical relation between attenuation coefficient and measuring frequency is established, the next step is to formulate a corresponded phase velocity dispersion relation. K-K relations impose restrictions on both attenuation coefficient and velocity dispersion under the assumption of causality. Following this path, the well-known Kolsky–Futterman model and its modification version can be obtained. After that the non-stationary convolution is formulated in the frequency domain, which lays the foundation for inversion-based seismic attenuation compensation.

General way for frequency power-law attenuation modeling

As the form of Fourier kernel will affect the sign of temporal and spatial derivation (Holm & Näsholm 2013), in this paper we denote the space–time Fourier transform of plane wave in a homogeneous

medium as

$$\mathcal{F}[p](k, \omega) = \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} p(\mathbf{x}, t) e^{-i(\omega t - k \cdot \mathbf{x})} d\mathbf{x} dt, \qquad (1)$$

and corresponding inverse Fourier transform as

$$\mathcal{F}^{-1}[p](\mathbf{x}, t) = \frac{1}{(2\pi)^d} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} p(k, \omega) e^{i(\omega t - k \cdot \mathbf{x})} dk d\omega, \qquad (2)$$

where $\mathbb{R}^d$ represents d dimensional real space. ω is angular frequency, and k is wavenumber. Plane wave in a homogeneous attenuating medium can be expressed by
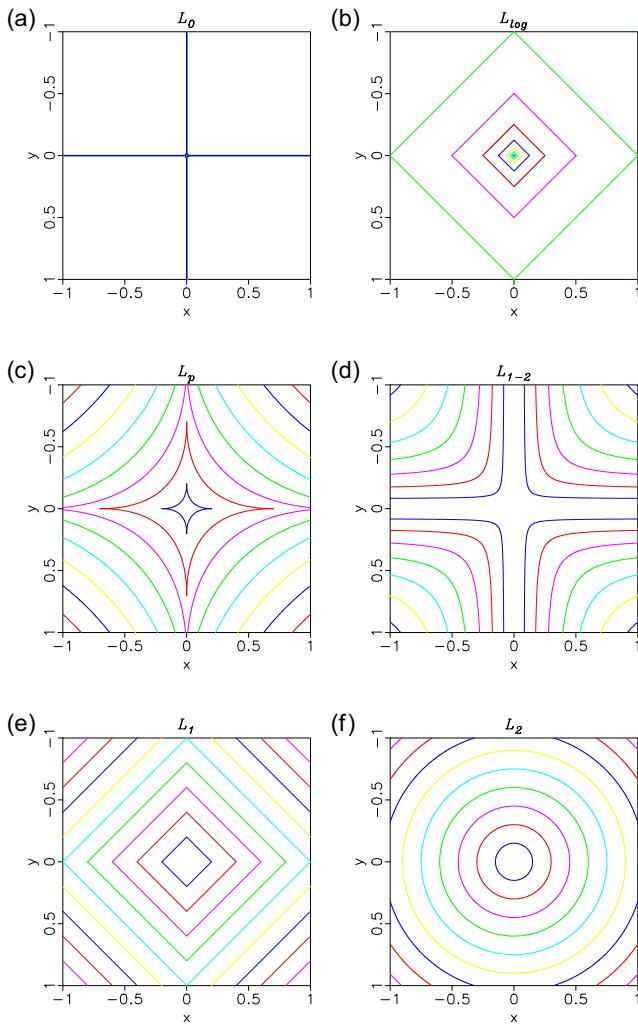
$$p(\mathbf{x}, t) = e^{i[\omega t - k(\omega) \cdot \mathbf{x}]}, \qquad (3)$$

where the complex wavenumber $k(\omega)$ is defined by

$$k(\omega) = \frac{\omega}{c(\omega)} = \frac{\omega}{c_p(\omega)} - i\alpha(\omega), \qquad (4)$$

where $c_p(\omega)$ and $\alpha(\omega)$ are frequency-dependent phase velocity and attenuation coefficient, respectively. K-K relations show that, as a consequence of linearity and causality, the real and imaginary parts of physical quantities in the frequency domain are related by Hilbert transform (Futterman 1962; Waters *et al.* 2003, 2005). It indicates that the dispersion and absorption are not independent of one another. In other words, we can formulate a corresponded phase velocity dispersion relation when frequency-dependent attenuation coefficient is well established.

Frequency-dependent attenuation in subsurface media typically follows a frequency power law in which the exponent is between

Figure 4. Contours of different penalties: (a) L_0, (b) L_{log} (LSP), (c) L_p ($p = 0.5$), (d) L_{1-2}, (e) L_1 and (f) L_2.

0 and 2 over the frequency range of interest. This power-law frequency dependence is widely used in many branches of acoustics, including seismology (Kibblewhite 1989; Buckingham 1997), geophysics (McDonal *et al.* 1958; Futterman 1962; Kjartansson 1979; Li *et al.* 2016a), ultrasonics (Szabo 1994; Kelly *et al.* 2008) and photoacoustics (Huang *et al.* 2012; Treeby 2013). Furthermore, the closed-form dispersion relation can be derived when the attenuation coefficient is empirically expressed as a power-law frequency independence (Waters *et al.* 2003). We assume that the model for the attenuation coefficient is

$$\alpha(\omega) = \alpha(0) + \alpha_0 |\omega|^y, \tag{5}$$

where α_0 and y are real constants and $0 < y \leq 2$ typically. As attenuation coefficient in seismic exploration is nearly linear with frequency over the seismic frequency range (McDonal *et al.* 1958; Futterman 1962; Liu *et al.* 1976), here we consider the dispersion relations for the case that the exponent y is near a unit ($y \approx 1$), which is given by:

$$\frac{1}{c_p(\omega)} - \frac{1}{c_p(\omega_0)} \approx -\frac{2}{\pi}\alpha_0 \ln\left|\frac{\omega}{\omega_0}\right|, \tag{6}$$

where ω_0 represents the reference frequency.

Modified Kolsky–Futterman model

Kolsky (1956) assumed that frequency-dependent attenuation coefficient $\alpha(\omega)$ is strictly linear with frequency over the range of measurement ($y = 1$ in eq. 5):

$$\alpha(\omega) = \alpha_0 |\omega| = \frac{|\omega|}{2c_p(\omega_0)Q(\omega_0)}, \tag{7}$$

where $c_p(\omega_0)$ and $Q(\omega_0)$ are the values of the phase velocity and approximate Q at ω_0. Substituting eq. (7) into eq. (6), we have

$$\frac{1}{c_p(\omega)} = \frac{1}{c_p(\omega_0)}\left(1 - \frac{1}{\pi Q(\omega_0)}\ln\left|\frac{\omega}{\omega_0}\right|\right). \tag{8}$$

Inserting frequency-dependent attenuation coefficient (7) and phase velocity (8) into eq. (4), we have the following complex wavenumber

$$
\begin{aligned}
k(\omega) &= \frac{\omega}{c_p(\omega)}\left(1 - \frac{i}{2Q(\omega)}\right) \\
&= \frac{\omega}{c_p(\omega_0)}\left(1 - \frac{1}{\pi Q(\omega_0)}\ln\left|\frac{\omega}{\omega_0}\right|\right)\left(1 - \frac{i}{2Q(\omega)}\right).
\end{aligned} \tag{9}
$$

Wang & Guo (2004a) pointed out that the phase velocity formula (8) given in the Kolsky–Futterman model is merely an asymptotic formula for $\omega \gg \omega_0$. As exploration seismic data have relatively low-frequency range within 10^0–10^2 Hz, they therefore proposed a modified Kolsky–Futterman model by introducing the following approximation:

$$\lim_{\gamma \to 0}\left(1 - \gamma\ln\left|\frac{\omega}{\omega_h}\right|\right) = \left|\frac{\omega}{\omega_h}\right|^{-\gamma}, \tag{10}$$

where the dimensionless parameter $\gamma = \frac{1}{\pi Q(\omega_0)}$ and ω_h is a re-defined tuning parameter. This tuning parameter is no longer the lowest frequency of the seismic band, but on the contrary, the highest possible seismic frequency. Inserting the approximate formula (10) into eq. (9), we have the following complex wavenumber

$$k(\omega) = \frac{\omega}{c_p(\omega_0)}\left|\frac{\omega}{\omega_h}\right|^{-\gamma}\left(1 - \frac{i}{2Q(\omega)}\right). \tag{11}$$

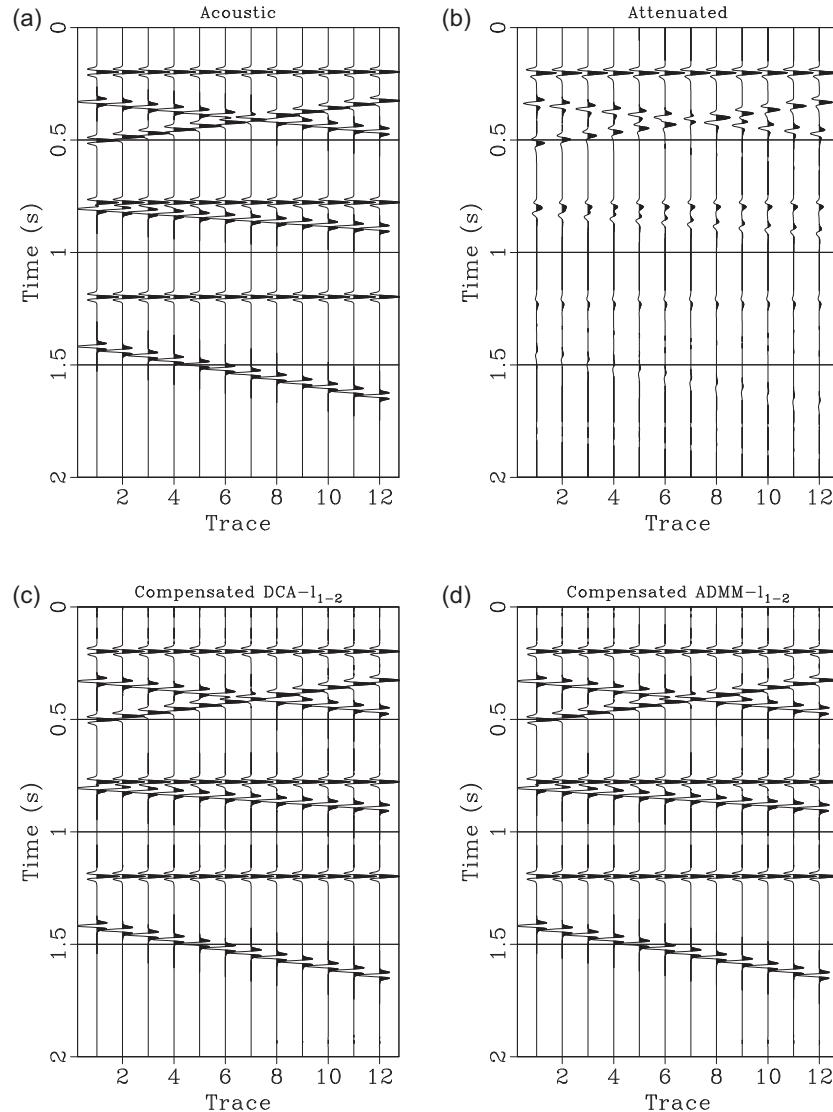## The non-stationary convolution in the frequency domain

Substituting the complex wavenumber $k(\omega)$ shown in eq. (11) into the plane wave expression, we have

$$
\begin{aligned}
p(\mathbf{x}, t) &= e^{i\left[\omega t - \frac{\omega}{c_p(\omega_0)}\left|\frac{\omega}{\omega_h}\right|^{-\gamma}\left(1 - \frac{i}{2Q(\omega)}\right)\mathbf{x}\right]} \\
&= e^{i\omega t}e^{-i\frac{\omega}{c_p(\omega_0)}\left|\frac{\omega}{\omega_h}\right|^{-\gamma}\mathbf{x}}e^{-\frac{\omega}{2c_p(\omega_0)Q(\omega)}\left|\frac{\omega}{\omega_h}\right|^{-\gamma}\mathbf{x}}.
\end{aligned} \tag{12}
$$

We first replace the distance $\mathbf{x}$ with the traveltime $\tau = \frac{\mathbf{x}}{c_p(\omega_0)}$, and then define the following attenuation function

$$
\begin{aligned}
a(\omega, \tau) &= e^{i\omega\tau\left(1 - \left|\frac{\omega}{\omega_h}\right|^{-\gamma}\right)}e^{-\frac{\omega\tau}{2Q(\omega)}\left|\frac{\omega}{\omega_h}\right|^{-\gamma}} \\
&\approx e^{i\omega\tau\left(1 - \left|\frac{\omega}{\omega_h}\right|^{-\gamma}\right)}e^{-\frac{\omega\tau}{2Q(\omega_0)}\left|\frac{\omega}{\omega_h}\right|^{-\gamma}},
\end{aligned} \tag{13}
$$

where the two exponential terms dominate phase dispersion and amplitude attenuation, respectively. It is remarkable that frequency-dependent $Q(\omega)$ can be nearly reduced to constant $Q(\omega_0)$ at the reference frequency ω_0. There are two main reasons for this approximation. First, considering frequency-dependent $Q(\omega)$ is rather

Figure 5. Seismic attenuation compensation on 2-D noise-free synthetic data, (a) acoustic data, (b) attenuated data, (c) compensated data using DCA-L_{1-2} algorithm and (d) compensated data using ADMM-L_{1-2} algorithm.

complicated and almost intractable for seismic attenuation compensation, especially for Q estimation. Second, $Q(\omega)$ in Kolsky–Futterman model exhibits weak frequency-dependence over a relatively narrow frequency regime, such as seismic band (Futterman 1962). The formula for non-stationary seismogram modeling is introduced by many researchers (Margrave *et al.* 2011; Wang & Chen 2014). Its final form in the frequency domain can be described as

$$s(\omega) = w(\omega) \int_0^T a(\omega, \tau) r(\tau) e^{-i\omega\tau} d\tau, \qquad (14)$$

where $s(\omega)$ and $w(\omega)$ are the Fourier spectra of seismic record $s(t)$ and wavelet series $w(t)$, respectively, $r(\tau)$ represents reflectivity series within T duration. For the band-limited seismic data with frequencies range from ω_1 to ω_L, eq. (14) can be discretized into a matrix-vector form
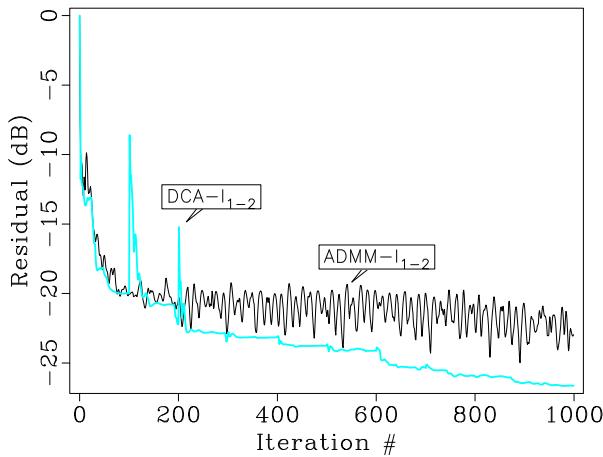
$$\mathbf{s} = \mathbf{WAr}, \qquad (15)$$

where the kernel matrix $\mathbf{W}$ represents the wavelet's bandpass filtering effects and matrix $\mathbf{A}$ stands for the earth's Q filtering effects.

The vector $\mathbf{r}$ represents sparse reflectivity series, which can be considered as the deconvolution result of the recovered seismic records $\boldsymbol{\chi}$. In this paper, we focus on removing the earth's Q filtering effects from seismic data, the compensated seismic records can be obtained by convoluting a wavelet, that is,

$$\boldsymbol{\chi} = \mathbf{Wr} = \mathbf{WA^{-1}W^{-1}s}. \qquad (16)$$

Instead of solving for $\boldsymbol{\chi}$ using eq. (16) directly, we perform seismic attenuation compensation via a two-step scheme. We first solve for sparse reflectivity series $\mathbf{r}$ using eq. (15) with sparsity constraint imposed on reflectivity series $\mathbf{r}$ and then calculate compensated seismic records using eq. (16). Fig. 1 shows the basic workflow of the direct and two-step compensation, where the deconvolution can be considered as a sparsity-promoting strategy (Wang *et al.* 2015; Xue *et al.* 2017). The frequency-domain discrete expression of eq. (15) gives the following close-form

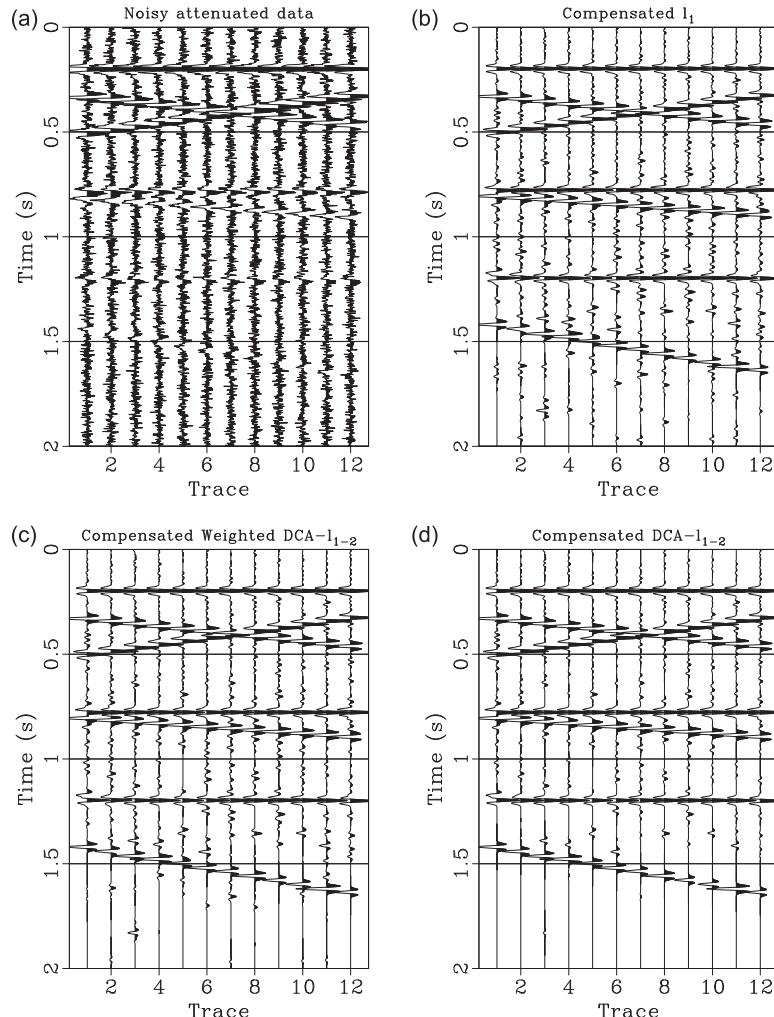**Figure 6.** Comparison between DCA-L_{1-2} and ADMM-L_{1-2} in terms of residual errors versus iterations.

expression

$$
\begin{bmatrix}
w(\omega_1)a(\omega_1,\tau_1)e^{-i\omega_1\tau_1} & \cdots & w(\omega_1)a(\omega_1,\tau_T)e^{-i\omega_1\tau_T} \\
\vdots & \ddots & \vdots \\
w(\omega_L)a(\omega_L,\tau_1)e^{-i\omega_L\tau_1} & \cdots & w(\omega_L)a(\omega_L,\tau_T)e^{-i\omega_L\tau_T}
\end{bmatrix}
\begin{bmatrix}
r_1 \\
\vdots \\
r_T
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
s(\omega_1) \\
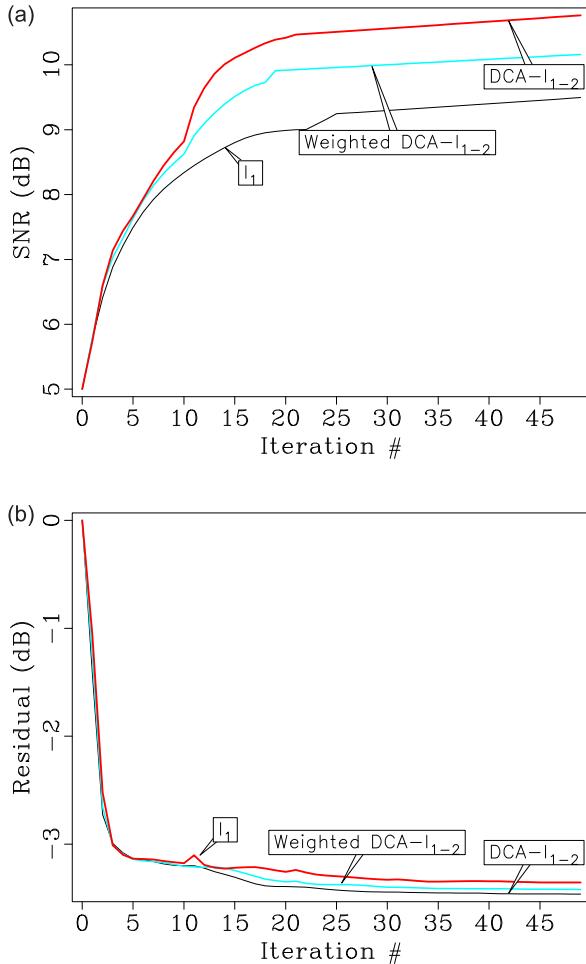\vdots \\
s(\omega_L)
\end{bmatrix}. \tag{17}
$$

INVERSION-BASED COMPENSATION IN THE TIME DOMAIN

Now that we have established the non-stationary convolution in the frequency domain, inversion-based compensation can be achieved by defining the following cost function with L_1 norm

$$
\min_{\mathbf{r}} \ \frac{1}{2}\,\|\mathbf{\Phi}\mathbf{r}-\mathbf{s}\|_2^2 + \lambda\,\|\mathbf{r}\|_1 , \tag{18}
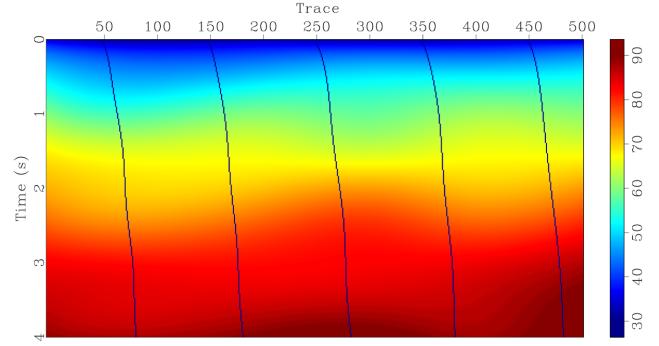$$

Figure 7. Seismic attenuation compensation on 2-D noisy synthetic data, (a) noisy attenuated data, (b) compensated data using L_1 minimization (SNR = 9.57), (c) compensated data using weighted DCA-$L_{1-\alpha 2}$ algorithm (where $\alpha = 0.5$ and SNR = 10.23) and (d) compensated data using DCA-L_{1-2} algorithm (SNR = 10.77).
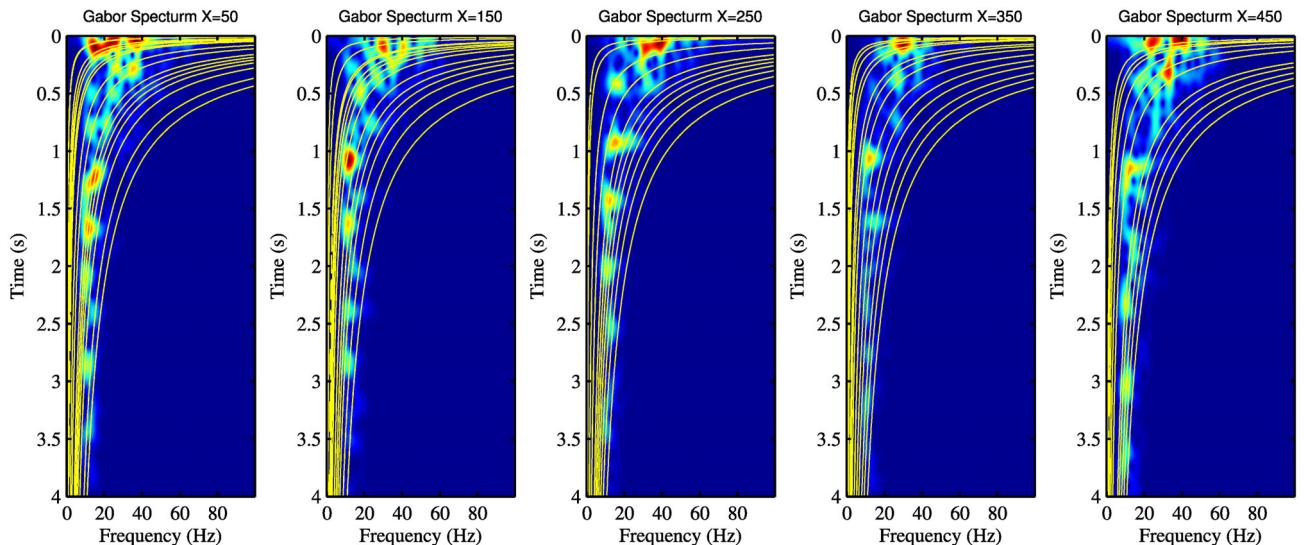
Figure 10. The estimated effective Q model from original attenuated data shown in Fig. 11(a), which is obtained by horizontal interpolation from five reference Q curves.

Figure 8. Comparison between L_1, weighted DCA-$L_{1-\alpha2}$ and DCA-L_{1-2} in terms of (a) SNR versus iterations and (b) residual errors versus iterations.

where the kernel matrix $\Phi = \mathbf{WA}$ responsible for both the wavelet's bandpass filtering effects and the earth's Q filtering effects. The L_1-norm penalized least-squares problem (18) can be solved by many state-of-the-art algorithms that are widely used and rapidly developed in CS (Candes *et al.* 2006; Donoho 2006). The matrix Φ

can be considered as sensing matrix, which is required to satisfy the restricted isometry property (RIP) with small restricted isometry constants (Candes & Tao 2005; Chartrand & Staneva 2010). Given a deterministic matrix Φ, it is generally NP-hard to verify whether Φ is an RIP matrix or not (Bandeira *et al.* 2013). An alternative way to predict RIP of Φ is the so-called coherence, which is closely related to the RIP yet easy to examine (Donoho & Huo 2002). That is to say, a matrix satisfying some RIP tends to have small coherence, whereas a highly coherent matrix is unlikely to possess small restricted isometry constants. The coherence coefficients of the matrix Φ are defined as
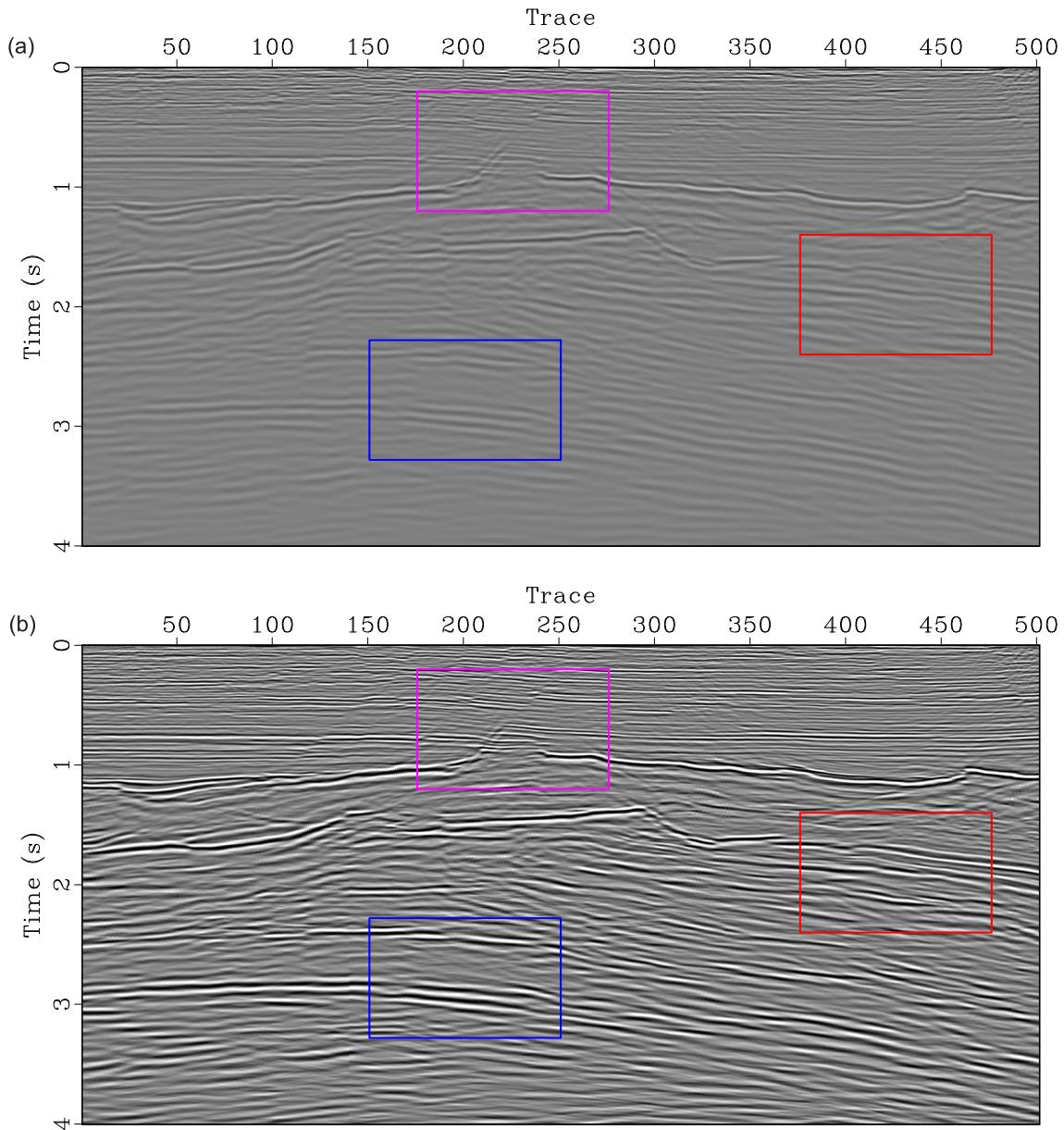
$$\mu_\Phi(i, j) := \frac{|\Phi_i^T \Phi_j|}{\|\Phi_i\|_2 \|\Phi_j\|_2}, \qquad i \neq j, \tag{19}$$

where Φ_i and Φ_j are arbitrary two columns from Φ. Fig. 2(a) displays the frequency-domain matrix Φ with real (the upper half-plane) and imaginary parts (the lower half-plane) integrated together, the corresponding coherence coefficients are shown in Fig. 2(b). As frequency-domain sensing matrix Φ exhibits high coherence resulting in degraded inversion performance, we reformulate a new misfit function by transforming frequency-domain formula (18) into the time domain

$$\min_{\mathbf{r}} \frac{1}{2} \left\| \hat{\Phi} \mathbf{r} - \hat{\mathbf{s}} \right\|_2^2 + \lambda \|\mathbf{r}\|_1, \tag{20}$$

Figure 9. The Gabor spectra of five reference traces from original attenuated data shown in Fig. 11(a).
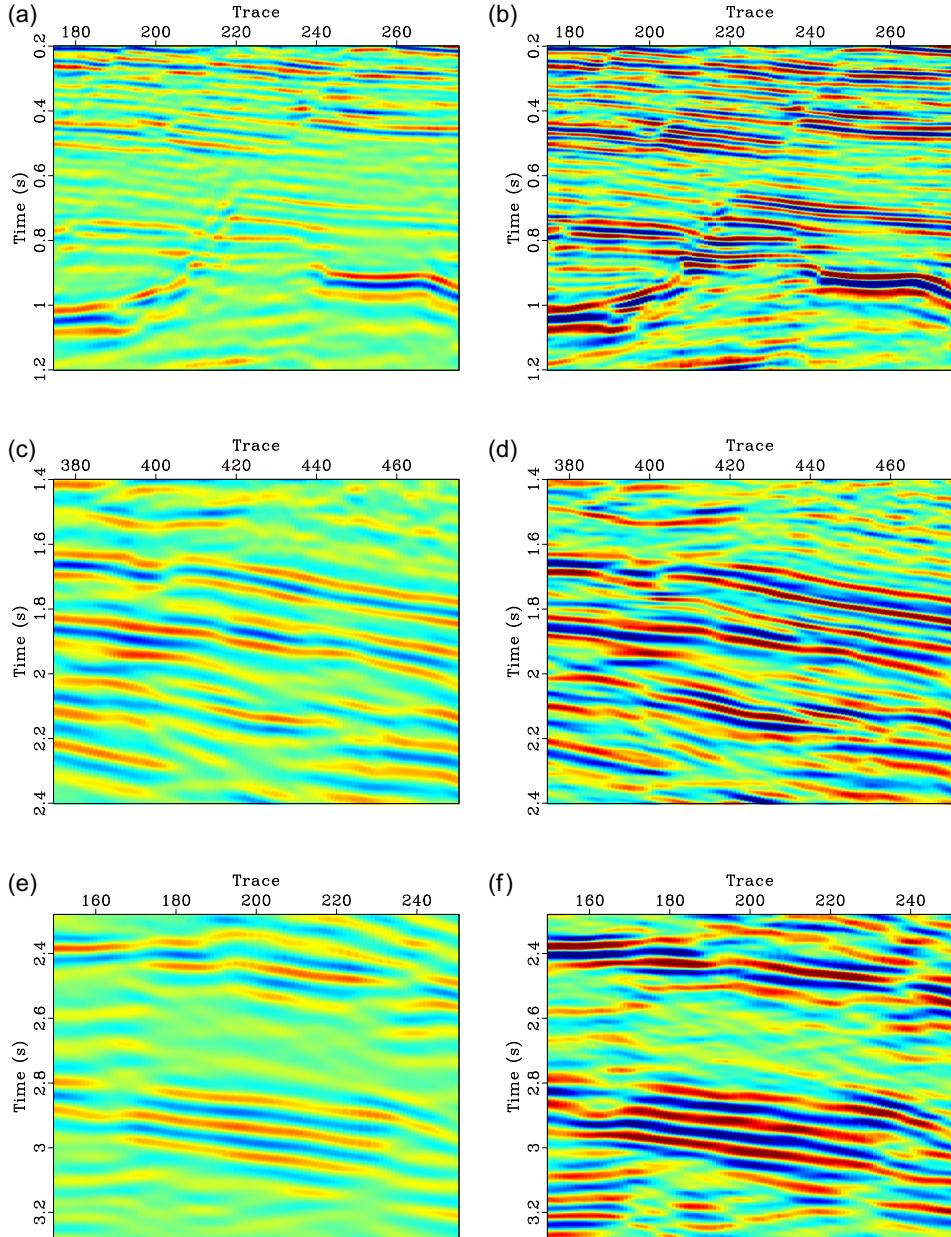
Figure 11. Seismic attenuation compensation on 2-D field data, (a) original attenuated data and (b) compensated data using our proposed L_{1-2} minimization.

where the new kernel matrix $\hat{\Phi}$ is obtained by transforming Φ into time domain and reshaping it as a diagonal matrix form. Fig. 3 displays the time-domain $\hat{\Phi}$ and its corresponding coherence coefficients. Its remarkable that directly deriving a close-form kernel matrix $\hat{\Phi}$ in the time domain is usually difficult, therefore, Fourier transform and matrix diagonalization provide a promising candidate to construct $\hat{\Phi}$.

Eq. (20) presents L_1 norm constrained inversion problem, which can be also called lasso. There are many state-of-the-art algorithms available for this problem, such as ADMM (Combettes & Wajs 2005; Boyd *et al.* 2011; Combettes & Pesquet 2011), fixed-point continuation (Hale *et al.* 2008), fast iterative shrinkage-thresholding algorithm (Beck & Teboulle 2009) and split Bregman (Goldstein & Osher 2009). However, L_1 minimization may sometimes yield sub-optimal performance due to the biased approximation to L_0. To address this issue, many non-convex metrics, interpolated between the L_0 and L_1 norms, have been proposed to better approximate

the L_0 norm. Here, we compare several widely used non-convex penalties with typical L_0, L_1 and L_2 norms, which include L_p, LSP and L_{1-2} metrics. Consider, for example, these potential penalty functions for vector (x, y) can be defined as

$$
\begin{cases}
L_0 = \begin{cases} 0, & (x, y) = \mathbf{0}, \\ 2, & xy \neq 0, \\ 1, & \text{else}, \end{cases} \\
L_1 = |x| + |y|, \\
L_2 = \sqrt{|x|^2 + |y|^2}, \\
L_p = \sqrt[p]{|x|^p + |y|^p}, \qquad 0 < p < 1, \\
L_{\log} \propto \log(1 + (|x| + |y|)/\varepsilon), \qquad \varepsilon \to 0, \\
L_{1-2} = |x| + |y| - \sqrt{|x|^2 + |y|^2}.
\end{cases}
\tag{21}
$$

Figure 12. Zoomed-in view of the 2-D field data: (a), (c) and (e) original attenuated data from the boxes shown in Fig. 11(a); and (b), (d) and (f) compensated data from the boxes shown in Fig. 11(b).

Fig. 4 shows these penalty functions for an intuitive comparison. As we can see from Fig. 4(d), the contour lines of L_{1-2} penalty approach the *x*- and *y*-axes as the values get smaller, hence promoting sparsity. Furthermore, compared with L_1 (Fig. 4e), the contour lines of L_{1-2} are closer to the axes when minimized, which means that L_{1-2} penalty behaves closer to that of L_0 (Fig. 4a) and has the potential to be much more sparsity promoting than the L_1 norm.

In this paper, we incorporate L_{1-2} metric into the inversion-based compensation scheme, which outperforms the existing L_1 penalty for highly coherent kernel matrix. Another advantage of L_{1-2} over L_1 is its unbiased characterization of one-sparse vectors, since $\|\mathbf{x}\|_{1-2} = 0$ if and only if $\|\mathbf{x}\|_0 \leq 1$. However, L_{1-2} becomes biased and behaves like L_1 as the number of leading entries (in magnitude) increases. The misfit function with L_{1-2} minimization
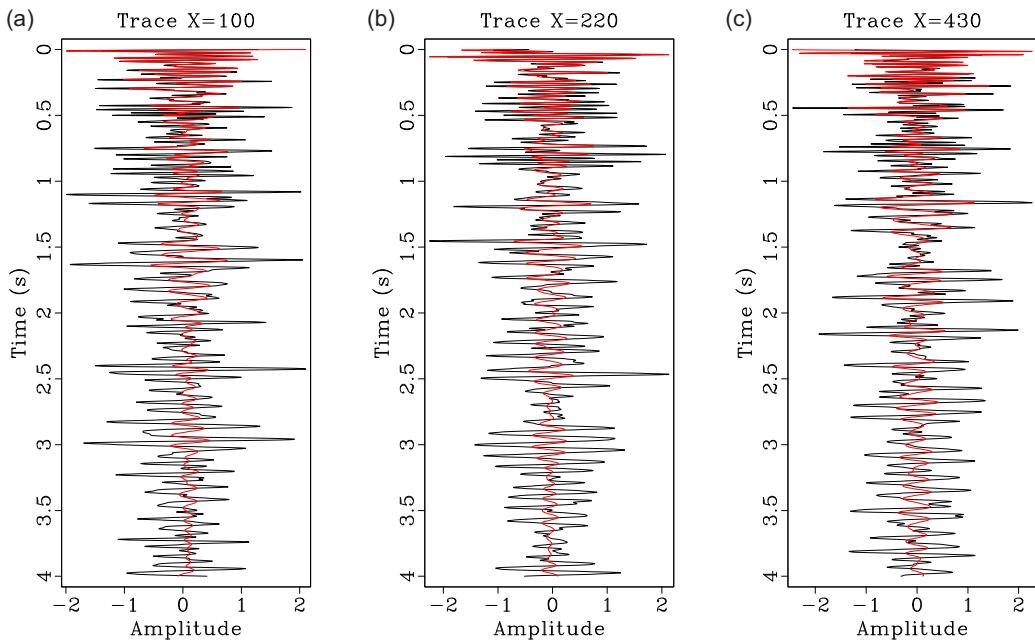
is given as

$$\min_{\mathbf{r}} \ \frac{1}{2} \| \hat{\boldsymbol{\Phi}} \mathbf{r} - \hat{\mathbf{s}} \|_2^2 + \lambda(\|\mathbf{r}\|_1 - \alpha \|\mathbf{r}\|_2), \tag{22}$$

where the weight parameter α with the range of [0, 1] is provided to deal with ill-conditioned matrices when L_{1-2} fails to obtain a good solution (Lou *et al.* 2015a).

THE SOLVER FOR L_{1-2} MINIMIZATION

In this section, we focus on investigating an efficient solver for L_{1-2} minimization problem (22). In addition, for the convenience

Figure 13. Comparison of compensation performance using three reference traces extracted from Fig. 11(a) (red) and Fig. 11(b) (black) at (a) X = 100, (b) X = 220 and (c) X = 430, respectively.

and simplicity of formulas' deducing, we rewrite eq. (22) as the following uniform formula:

$$\min_x \; \frac{1}{2} \|Ax - b\|_2^2 + \lambda(\|x\|_1 - \alpha \|x\|_2), \qquad (23)$$

with $A = \hat{\boldsymbol{\Phi}}$, $x = \mathbf{r}$ and $b = \hat{\mathbf{s}}$. We first briefly introduce DCA for L_{1-2} minimization and then elaborate on how to solve the DCA subproblem via ADMM. We further have a comparison between two L_{1-2} implementations based on them, in terms of computational efficiency and convergence.

DCA for L_{1-2} norm constrained problem

The DCA is a robust and efficient descent method introduced by Tao & An (1998), which copes with the minimization of an objective function $F(x) = G(x) - H(x)$, where $G(x)$ and $H(x)$ are proper convex functions. The DCA involves the construction of two sequences $\{x^k\}$ and $\{y^k\}$, the candidates for optimal solutions of primal and dual programs, respectively (Tao & An 1998; Yin *et al.* 2015). It gives

$$\begin{cases} y^k \in \partial H(x^k), \\ x^{k+1} = \arg\min_{x \in \mathbb{R}^n} G(x) - \left(H(x^k) + \langle y^k, x - x^k \rangle\right), \end{cases} \qquad (24)$$

where y^k is a subgradient of $H(x)$ at x^k. The monotonically decreasing property of DCA has been proven in Appendix A. We assume that $F(x)$ is bounded from below, then the objective function values of DCA are convergent (Tao & An 1998; Yin *et al.* 2015). The objective function in eq. (23) naturally has the following convex decomposition

$$F(x) = \left(\frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1\right) - \alpha\lambda \|x\|_2, \qquad (25)$$

where $-\alpha\lambda \|x\|_2$ is differentiable for all $x \neq \mathbf{0}$ and subdifferentiable at $x = \mathbf{0}$, for convenience, we only choose a subgradient $y^k = \mathbf{0}$ when $x = \mathbf{0}$, then we have

$$y^k = \begin{cases} \mathbf{0}, & \text{if } x^k = \mathbf{0}, \\ -\alpha\lambda \dfrac{x^k}{\|x^k\|_2}, & \text{otherwise.} \end{cases} \qquad (26)$$

According to DCA iteration formula (24), L_{1-2} minimization in eq. (25) can be solved by the following scheme (the detailed deduction can be found in Appendix B):

$$x^{k+1} = \arg\min_{x \in \mathbb{R}^n} \; \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1 + \langle y^k, x \rangle. \qquad (27)$$
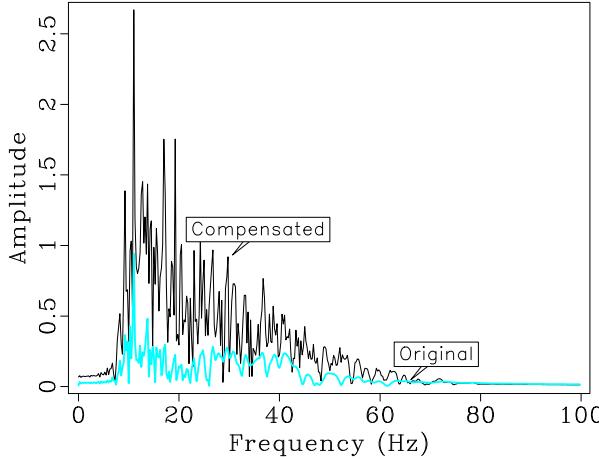
Solving the convex subproblem via ADMM

Now that we have decomposed the original non-convex problem (23) into two convex subproblems (26) and (27), we can therefore apply the ADMM to solve the unconstrained problem (27). The trick of ADMM formula is to decouple the coupling between the quadratic term and L_1 penalty, By introducing an auxiliary variable z, eq. (27) is equivalent to the following constrained minimization problem:

$$x^{k+1} = \arg\min_{x \in \mathbb{R}^n} \; \frac{1}{2} \|Ax - b\|_2^2 + \langle y^k, x \rangle + \lambda \|z\|_1$$

$$\text{subject to } x - z = \mathbf{0}. \qquad (28)$$

The augmented Lagrangian can be expressed as

$$\mathcal{L}_\rho(x, z, w) = \frac{1}{2} \|Ax - b\|_2^2 + \langle y^k, x \rangle + \lambda \|z\|_1 + w^T(x - z)$$

$$+ \frac{\rho}{2} \|x - z\|_2^2, \qquad (29)$$

Figure 14. Comparison of the averaged spectra from the original data shown in Fig. 11(a) and the compensated data shown in Fig. 11(b).

where y is the Lagrangian multiplier and ρ is the penalty parameter. The unscaled-form ADMM consists of the iterations (Boyd *et al.* 2011; Yin *et al.* 2015):

$$\begin{cases} z^{l+1} = \arg\min_z \mathcal{L}_\rho(x^l, z, w^l), \\ x^{l+1} = \arg\min_x \mathcal{L}_\rho(x, z^{l+1}, w^l), \\ w^{l+1} = w^l + \rho(x^{l+1} - z^{l+1}), \end{cases} \quad (30)$$

where the z-update and x-update steps have the closed-form solutions (Yin *et al.* 2015; Lou & Yan 2016), more specifically, the z-update can be solved by soft thresholding

$$z^{k+1} = \mathcal{S}\left(x^k + w^k/\rho, \lambda/\rho\right), \quad (31)$$

where $\mathcal{S}$ represents soft thresholding function, and the x update can be solved by gradient method

$$x^{k+1} = \left(A^T A + \rho I\right)^{-1}\left(A^T b - y^k + \rho z^{k+1} - w^k\right), \quad (32)$$

where $(A^T A + \rho I)^{-1}$ can be approximated by Cholesky factorization at the beginning for a fixed ρ.

Comparison between DCA-L_{1-2} and ADMM-L_{1-2} implementations

In the previous two sections, we have provided close-form expression of DCA and ADMM, here we will investigate two L_{1-2} implementations based on DCA and ADMM. These two implementations are different from each other due to the distinct execution order of the iterations (26) and (30). In the first scheme, also called DCA-L_{1-2}, the gradient y is updated after $l_{\max}^1$ inner iterations of ADMM within $k_{\max}^1$ outer iterations; whereas in the second scheme called ADMM-L_{1-2}, the gradient y is updated after every iteration of ADMM, x, z, w and y are updated simultaneously within $l_{\max}^2$ iterations. The pseudo-code of the these two implementations are summarized in Algorithms 1 and 2.

We discuss the computational complexity of DCA-L_{1-2} and ADMM-L_{1-2} implementations for inversion-based compensation problem (23). For each iteration, ADMM-L_{1-2} requires to compute the matrix-vector multiplication of an $L \times T$ matrix and an $T \times 1$ vector for x-update step, which costs $\mathcal{O}(LT)$; and z-update step reduces to a soft-thresholding operator on an $T \times 1$ vector, which costs $\mathcal{O}(T)$. Therefore, the complexity per iteration is $\mathcal{O}(LT)$. Assuming that the number of iterations of ADMM-L_{1-2} scheme

---

Algorithm 1 DCA-L_{1-2} for seismic attenuation compensation

Input: A, b, λ, ρ, α, k_{max}^1, l_{max}^1.
Output: $x := x^k$.
1: **Initialization:** Set $x^0 := \mathbf{0}$, $w^0 := \mathbf{0}$, and $k := 0$.
2: **for** $k = 0 \dots k_{max}^1$ **do**
3:     **if** $x = \mathbf{0}$ **then**
4:         $y^k = \mathbf{0}$;
5:     **else**
6:         $y^k = -\alpha\lambda\frac{x^k}{x^k_2}$;
7:     **end if**
8:     **Initialization:** Set $x^{k+1,0} := x^k$, $w^{k+1,0} := w^k$, and $l := 0$.
9:     **for** $l = 0 \dots l_{max}^1$ **do**
10:         $z^{k+1,l+1} := \mathcal{S}\left(x^{k+1,l} + w^{k+1,l}/\rho, \lambda/\rho\right)$;
11:         $x^{k+1,l+1} := (A^T A + \rho I)^{-1}(A^T b - y^k + \rho z^{k+1,l+1};$
                $- w^{k+1,l})$;
12:         $w^{k+1,l+1} := w^{k,l} + \left(x^{k+1,l+1} - z^{k+1,l+1}\right)$;
13:         Set $l := l + 1$.
14:     **end for**
15:     **Output:** $x^{k+1} := x^{k+1,l}$ and $w^{k+1} := w^{k+1,l}$.
16:     Set $k := k + 1$.
17: **end for**

Algorithm 2 ADMM-L_{1-2} for seismic attenuation compensation

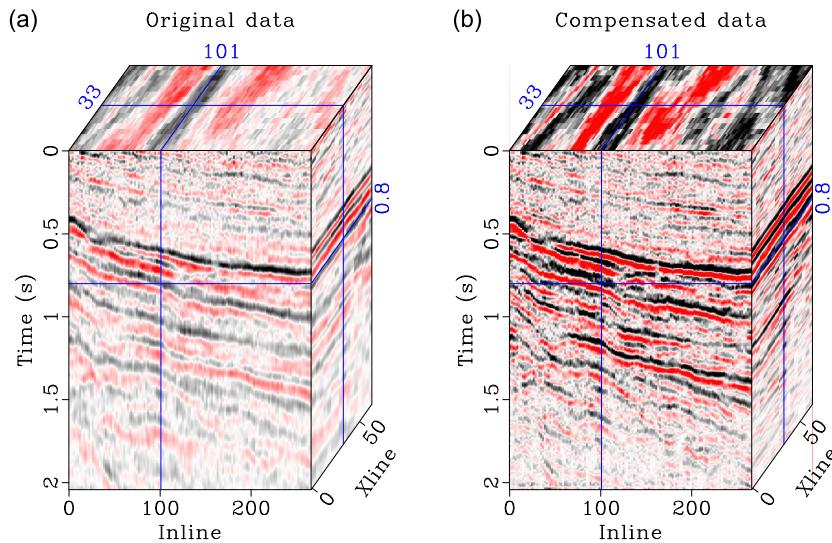**Input:** A, b, λ, ρ, α, l_{max}^2.
Output: $x := x^l$.
1: **Initialization:** Set $x^0 := \mathbf{0}$, $w^0 := \mathbf{0}$, $y^0 := \mathbf{0}$ and $l := 0$.
2: **for** $l = 0 \dots l_{max}^2$ **do**
3:     **if** $x^l = \mathbf{0}$ **then**
4:         $y^l = \mathbf{0}$;
5:     **else**
6:         $y^l = -\alpha\lambda\frac{x^l}{x^l_2}$;
7:     **end if**
8:     $z^{l+1} := \mathcal{S}\left(x^l + w^l/\rho, \lambda/\rho\right)$;
9:     $x^{l+1} := \left(A^T A + \rho I\right)^{-1}\left(A^T b - y^l + \rho z^{l+1} - w^l\right)$;
10:     $w^{l+1} := w^l + \left(x^{l+1} - z^{l+1}\right)$;
11:     Set $l := l + 1$.
12: **end for**

---

is N, therefore the total complexity of ADMM-L_{1-2} algorithm is $\mathcal{O}(NLT)$. As for DCA-L_{1-2}, it requires to solve an L_1 minimization problem (27) iteratively, whose computational complexity is equivalent to a miniature ADMM-L_{1-2} at each outer iteration. For a fair comparison between these two implementations, we set the total iterations of DCA-L_{1-2} algorithm to be same as ADMM-L_{1-2} implementation, that is, $N = k_{\max}^1 l_{\max}^1 = l_{\max}^2$. As a result, the DCA-$L_{1-2}$ scheme has comparable computational complexity as ADMM-L_{1-2} scheme. The convergence property of these two algorithms will be examined in the next section by performing attenuation compensation on 2-D noise-free synthetic data.

EXAMPLES

In this section, we will examine the overall performance of our proposed L_{1-2} minimization in terms of inversion-based seismic attenuation compensation for 2-D and 3-D synthetic noise-free and noisy data and field data.

Figure 15. Seismic attenuation compensation on 3-D field data, (a) original attenuated data and (b) compensated data using our proposed L_{1-2} minimization.

2-D noise-free synthetic data

In the first example, we perform attenuation compensation on 2-D noise-free synthetic data set to verify the performance of our proposed DCA-L_{1-2} and ADMM-L_{1-2} algorithms and to compare their computational efficiency and convergence property. The scale of this synthetic data is 1000 time samples per trace (12 traces in total) with the time interval of 2 ms. Figs 5(a) and (b) display synthetic data without and with attenuation, respectively. The compensated results obtained from both DCA-L_{1-2} and ADMM-L_{1-2} schemes are shown in Figs 5(c) and (b), from which we can find that L_{1-2} constrained inversion enjoys pretty good compensation accuracy and reliability no matter which implementation is involved. In this example, we choose a relative small balancing parameter $\lambda = 10^{-6}$, the total number of iterations for both DCA-L_{1-2} and ADMM-L_{1-2} methods are 1000 times, that is, there are 100 outer iterations for DCA-L_{1-2} with each iteration containing 10 inner iterations, whereas ADMM-L_{1-2} implementation has 1000 outer iterations. We further record the runtime of these two implementations (processor of our laptop is Intel Core i5-4460 CPU @ 3.20GHz × 4), the total runtime $t = 3.79$ s for DCA-L_{1-2} and $t = 2.86$ s for ADMM-L_{1-2}. It indicates that ADMM-L_{1-2} scheme is slightly efficient than DCA-L_{1-2} scheme.

Apart from investigating the efficiency and fidelity of these two approaches, we further compare the convergence property of the DCA-L_{1-2} and ADMM-L_{1-2} algorithms by numerical tests. As shown in Algorithms 1 and 2, the difference between these two schemes lies in the update of variable y. For DCA-L_{1-2}, y is updated after whole ADMM (l_{max}^1) iterations, while ADMM-L_{1-2} updates y during each ADMM iteration. We plot residual errors of each inner solution versus the number of iterations in Fig. 6, which illustrates that DCA-L_{1-2} exhibits better convergence property over ADMM-L_{1-2}. The jump phenomenon on convergence curve of DCA-L_{1-2} is possibly caused by the discontinuity of the gradient y between the outer and inner iterations. In the rest of this paper, inversion-based compensation using L_{1-2} minimization are implemented by DCA scheme.
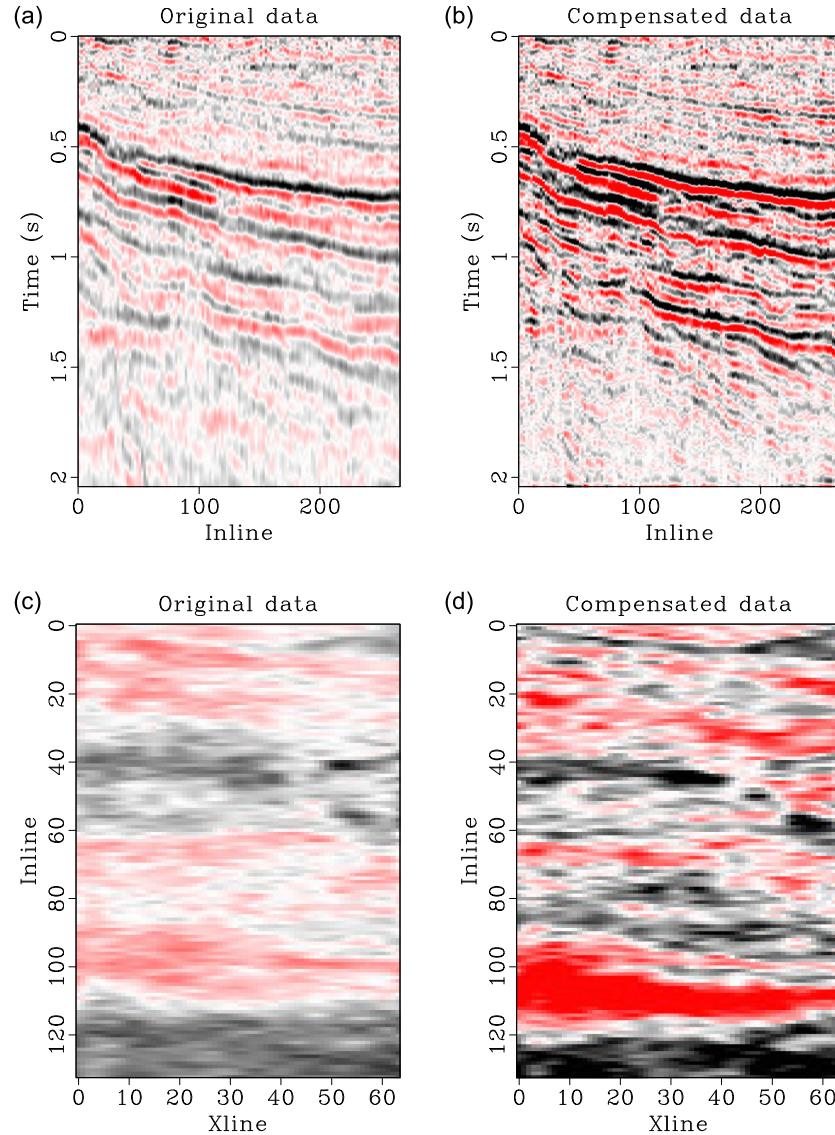
2-D noisy synthetic data

Our second test conducted on 2-D noisy synthetic data is designed for demonstrating the superior compensation performance of the proposed L_{1-2} metric over conventional L_1 minimization. The weighted L_{1-2} norm between these two metrics is also incorporated into inversion-based compensation for a more intuitive comparison. The attenuated record with strong random noise is shown in Fig. 7(a), from which we can find that deep reflectors are almost drowned in the unwanted artefacts. To distinguish effective signal from such noisy data and compensate their amplitude absorption and phase distortion, we perform inversion-based compensation with three different constraints, which include conventional L_1 regularization, weighted L_{1-2} minimization and our proposed L_{1-2} metric. For measuring the compensation performance of synthetic data examples, where one knows the reference data, we use the signal-to-noise ratio (SNR, Chen *et al.* 2016; Chen 2017) metric which is defined as

$$\text{SNR} = 10\log_{10}\frac{\|\boldsymbol{\chi}_{\text{ref}}\|_2^2}{\|\boldsymbol{\chi}_{\text{ref}} - \boldsymbol{\chi}\|_2^2}, \tag{33}$$

where $\boldsymbol{\chi}_{\text{ref}}$ denotes the reference data without attenuation, and $\boldsymbol{\chi}$ denotes the compensated data. The compensated results are respectively shown in Figs 7(b)–(d). Since the suboptimal sparsity approximation of the L_1 metric, the compensated record in Fig. 7(b) suffers from many non-zero disturbances which further result in lower SNR of 9.57 and mislead reflectors recognition. The proposed L_{1-2} minimization, developed for unbiased approximation of L_0 penalty, enjoys better antinoise property and higher compensation fidelity with SNR of 10.77 than that of the conventional L_1 constrained compensation. Furthermore, the weighted $L_{1-\alpha 2}$-based compensation achieves the moderate performance between that of L_1 and L_{1-2} implementations with SNR of 10.23.

We further have a comparison among these three methods in terms of convergence property, Fig. 8 shows the SNR and residual errors of each inner solution versus the number of iterations via L_1, weighted DCA-$L_{1-\alpha 2}$ and DCA-L_{1-2} schemes, respectively. It indicates that the proposed DCA-L_{1-2} exhibits higher SNR and better convergence property than conventional L_1 method. However, the global convergence of these three algorithms is severely degraded when compared with the noise-free case displayed in Fig. 6. In these tests on noisy data, we select a relative large balancing parameter $\lambda = 5 \times 10^{-4}$ to suppress random noise from the original data.

Figure 16. Comparison of compensation performance using inline sections at Xline = 40 from (a) the original data shown in Figs 15(a) and (b), the compensated data shown in Fig. 15(b) and time slices at *t* = 1 s from (c) the original data shown in Figs 15(a) and (d) the compensated data shown in Fig. 15(b).
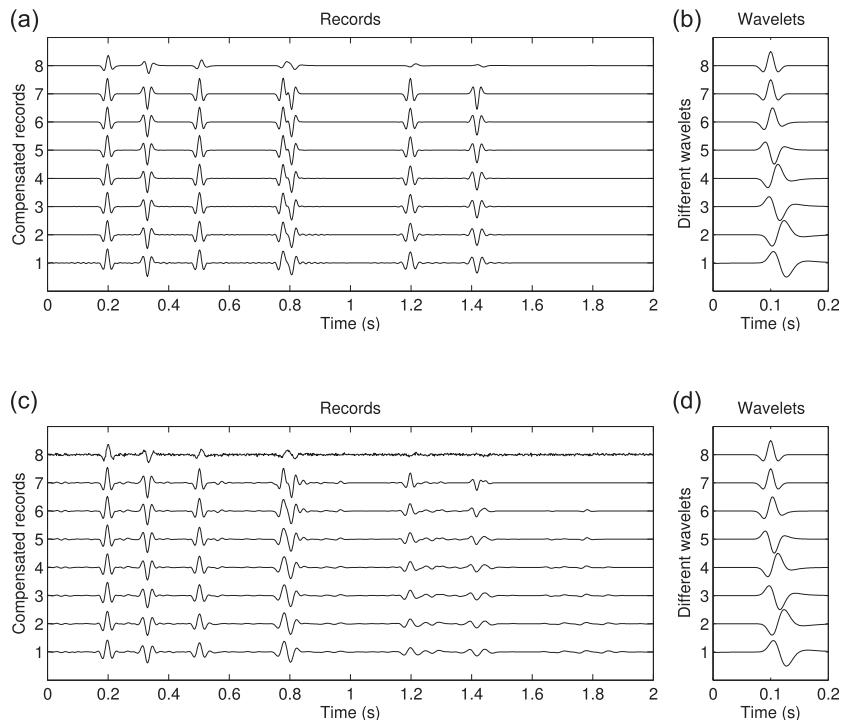
2-D field data

In this example, we consider a compensation test on 2-D field data using the proposed L_{1-2} scheme to further demonstrate its validity and reliability. The workflow of the seismic attenuation compensation on field data can be summarized as follows:

(i) Calculate the Gabor spectra of the attenuated data via truncated Gabor transform (Hargreaves & Calvert 1991; Wang 2004), where we choose five reference traces from original attenuated data shown in Fig. 11(a). The Gabor spectra of these traces with hyperbolic smoothing are shown in Fig. 9.

(ii) Estimate *Q* directly from the Gabor spectra via attenuation-based or compensation-based *Q* analysis (Wang 2004), where the five reference *Q* curves are first estimated. The whole effective *Q* model shown in Fig. 10 is further obtained by horizontal interpolation from these *Q* curves (Zhang *et al.* 2012).

(iii) Compensate the original data based on the estimated effective *Q* model via the proposed L_{1-2} constrained inversion framework. The compensated result shown in Fig. 11(b) has balanced amplitude and higher resolution.

For a clearer comparison between the compensated and the original records, we plot the zoomed-in portion of Fig. 11 at three positions from shallow, medium to deep layers. Figs 12(a) and (b) show shallow structures without and with compensation, from which we can find that the compensated seismic section exhibits clearer reflectors and sharper faults compared with the non-compensated section. Figs 12(c) and (d) display reflectors in the medium depth, it is obvious that the compensated section has higher vertical resolution and recovered amplitude, which might help to stratigraphic interpretation and attribute extraction. In the deep subsurface of the exploration area shown in Figs 12(e) and (f), seismic section after attenuation compensation tends to recover the reflected signal that is concealed and enhance the overall quality of the section. Fig. 13 display three reference traces extracted from the original sections and the compensated results, which have further verified the robustness and applicability of the proposed compensation algorithm.

Figure 17. Seismic attenuation compensation using inaccurate wavelet, (a) the compensated 1-D clean data and (b) their corresponding different wavelets; (c) the compensated 1-D noisy data and (d) their corresponding different wavelets. We choose balancing parameter $\lambda = 7 \times 10^{-5}$ for clean tests and $\lambda = 1 \times 10^{-2}$ for noisy tests.

The averaged spectra from the original data shown in Fig. 11(a) and the compensated data shown in Fig. 11(b) are displayed in Fig. 14, which indicates that high-frequency components of seismic data are recovered to some degree.
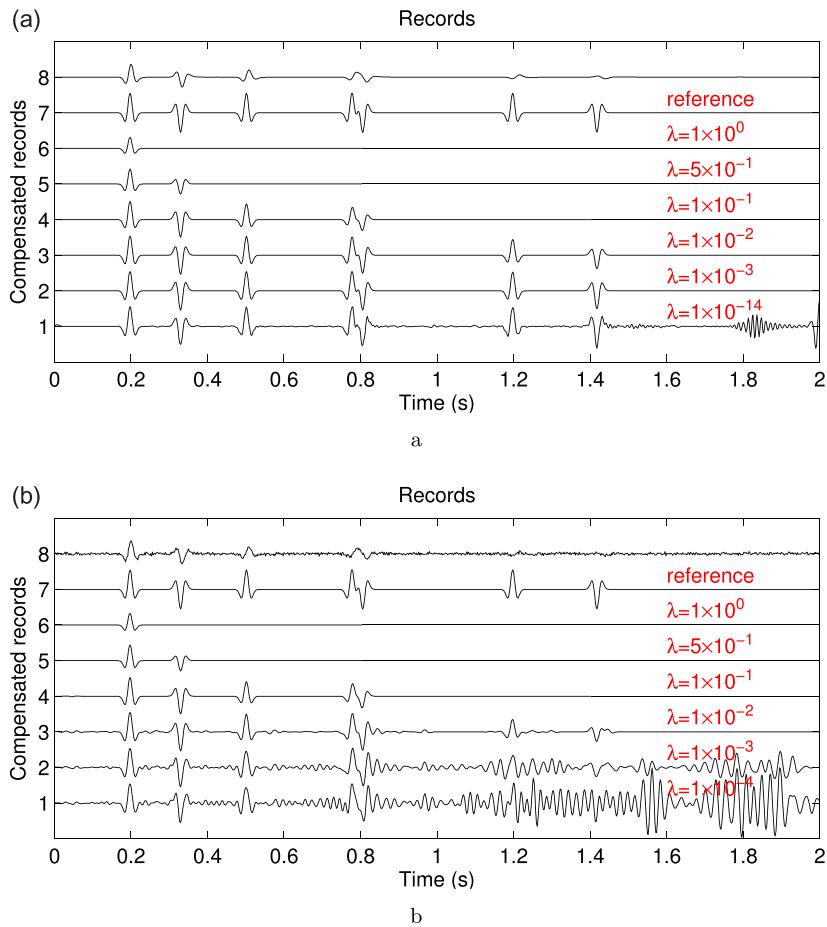
3-D field data

Our final example aims at exhibiting the practicability of the proposed non-convex inversion method in terms of 3-D field data compensation. Figs 15(a) and (b) present the original attenuated field data and the compensated result using our proposed L_{1-2} regularized inversion-based scheme. It is obvious that the post-stack seismic profile with amplitude compensation shows better structural continuity and energy consistency. We also plot the in-line sections in Figs 16(a) and (b) at Xline = 40 from both attenuated and compensated profiles, from which we can find that the spatial continuity of the reflectors and the resolution of the thin-layer interface are significantly improved compared with the non-compensated profile. Figs 16(c) and (d) show the time slices at $t = 1$ s from Figs 15(a) and (b), respectively. Since the time slice is critical to high-resolution seismic attribute extraction, our proposed compensation scheme may contribute to robust fault tracking and reliable reservoir characterization (Wu & Hale 2016; Wu & Janson 2017).

DISCUSSION

We have briefly reviewed the general way for frequency power-law attenuation characterization. Empirical relation between attenuation coefficient and measuring frequency can be characterized by experimentally established frequency power law. Then, the corresponded phase velocity dispersion relation is mathematically derived from

K-K relations. Following this path, many well-known attenuation models can be established. There are two main advantages of frequency power-law attenuation modeling. First, this kind of models are typically established on laboratory or field experiments, which may fit the measured data well compared to mechanical models such as standard linear solid and generalized Maxwell body (Liu *et al.* 1976; Carcione *et al.* 1988; Moczo & Kristek 2005; Yang *et al.* 2016). Second, such an approximately causality-imposed model may exhibit more reasonable frequency-dependent attenuation and dispersion over the model that violates the assumption of causality. In this paper, we utilize modified Kolsky–Futterman model to construct attenuation operator $a(\omega, \tau)$, which is responsible for both amplitude absorption and phase velocity dispersion effects. At the same time, this attenuation model is also served as the base of subsequent Q estimation and compensation.

Once the forward model in an attenuation medium is well formulated, seismic compensation can be implemented in an inversion framework. In this paper, we perform seismic attenuation compensation via first solving sparse reflectivity series and then converting it into seismic record, rather than solving directly for compensated seismic record χ using eq. (23) directly. The sparsity constraint in our proposed inversion algorithm is imposed on reflectivity series $\mathbf{r}$ rather than seismic record χ. As a result of the deconvolution, the reflectivity series $\mathbf{r}$ has better sparsity over the seismic record χ, thus the deconvolution can be considered as a sparsity-promoting strategy in the L_{1-2} constrained inversion method. An intuitive assumption is that the proposed inversion-based compensation might exhibit less wavelet dependence due to the presence of both deconvolution and convolution processes. A series of tests on both 1-D clean and noisy data sets with different wavelets are conducted to demonstrate wavelet dependence of the proposed method. As shown in Fig. 17, on the one hand, compensation performance is nearly independent of the types of the wavelet for both clean and

Figure 18. Seismic attenuation compensation using different balancing parameters λ tested on both (a) 1-D clean data and (b) 1-D noisy data.

noisy data sets; on the other hand, the overall accuracy is slightly degraded when some unwanted noises are mixed into the original data. It further verifies that our proposed two-step compensation scheme is free of wavelet estimation but affected by the level of the noise. Performing noise attenuation before signal compensation might be a feasible approach to achieve relatively high SNR and fidelity, which will be our future work.

We would also like to discuss the trick to choose a reasonable balancing parameter λ for seismic attenuation compensation. A series of tests on both 1-D clean and noisy data are performed to understand the role of λ in determining the solution. The clean seismogram is created by convolving a 30 Hz Ricker wavelet with a known reflectivity series, whereas the noisy seismogram is polluted by 20 per cent random noise. Fig. 18 shows the compensated seismogram from both clean and noisy data with different balancing parameters. Generally speaking, comparatively smaller parameter λ is suitable for seismic data with relatively high SNR, whereas compensation for severely noisy data usually needs a larger regularization parameter so as to maintain stability and accuracy to some degree. Typically, an L-curve technique is widely used to choose λ, which starts at high λ and iteratively decreases λ, while reusing solution at previous λ as initial guess at new λ (Hansen 1992; Xu 1998; Ng *et al.* 2016). This approach enables us to find a suitable value of λ automatically and achieve the balance between the data-fitting term and the regularization term objectively. Hence, critically, our presented algorithms in practice should be run over multiple choices of λ. We think that multiple choices of λ might

improve the performance of inversion, which will be further investigated our future work.

CONCLUSIONS

In this paper, we have derived the modified Kolsky–Futterman model based on the causality-imposed K-K relations. An inversion-based seismic attenuation compensation scheme in the frequency domain is further proposed, where the L_{1-2} constraint is imposed on reflectivity series to achieve an exact and stable amplitude recovery and phase correction. Benefited from the incoherence of the time-domain kernel matrix, we have reformulated a more robust time-domain compensation method in the framework of inversion. To leverage the sparsity of the reflectivity series, we have developed a two-step compensation scheme, in which the recovered subsurface reflectivity series are first solved by inversion and the compensated seismic records are then obtained by convolution. An additional bonus of this two-step strategy is that it is nearly independent of the estimated wavelet. Two effective implementations, denoted as DCA-L_{1-2} and ADMM-L_{1-2}, have been developed for solving the non-convex optimization problem. Compared to conventional L_1 metric, our proposed L_{1-2} penalty has potential to recover exact sparse reflectivity series from noisy attenuated seismograms when the kernel matrix is severely ill conditioned, which further results in enhanced antinoise property and improved compensation fidelity. The overall performance of the proposed L_{1-2} minimization is demonstrated

by performing seismic attenuation compensation on 2-D and 3-D noise-free and noisy synthetic and field data sets.

REFERENCES

Azimi, S.A., 1968. Impulse and transient characteristics of media with linear and quadratic absorption laws, *Izv. Phys. Solid Earth,* **2**(2), 88–93.

Bai, M., Chen, X., Wu, J., Liu, G., Chen, Y., Chen, H. & Li, Q., 2016. Q-compensated migration by Gaussian beam summation method, *J. Geophys. Eng.,* **13**(1), 35–48.

Bandeira, A.S., Dobriban, E., Mixon, D.G. & Sawin, W.F., 2013. Certifying the restricted isometry property is hard, *IEEE Trans. Inf. Theory,* **59**(6), 3448–3450.

Beck, A. & Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imag. Sci.,* **2**(1), 183–202.

Boyd, S., Parikh, N., Chu, E., Peleato, B. & Eckstein, J., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trend Mach. Learn.,* **3**(1), 1–122.

Buckingham, M.J., 1997. Theory of acoustic attenuation, dispersion, and pulse propagation in unconsolidated granular materials including marine sediments, *J. acoust. Soc. Am.,* **102**(5), 2579–2596.

Candes, E.J. & Tao, T., 2005. Decoding by linear programming, *IEEE Trans. Inf. Theory,* **51**(12), 4203–4215.

Candes, E.J., Romberg, J.K. & Tao, T., 2006. Stable signal recovery from incomplete and inaccurate measurements, *Commun. Pure Appl. Math.,* **59**(8), 1207–1223.

Candes, E.J., Wakin, M.B. & Boyd, S.P., 2008. Enhancing sparsity by reweighted ℓ_1 minimization, *J. Fourier Anal. Appl.,* **14**(5), 877–905.

Carcione, J.M., Kosloff, D. & Kosloff, R., 1988. Wave propagation simulation in a linear viscoelastic medium, *Geophys. J. Int.,* **93**(2), 393–401.

Chai, X., Wang, S., Yuan, S., Zhao, J., Sun, L. & Wei, X., 2014. Sparse reflectivity inversion for nonstationary seismic data, *Geophysics,* **79**(3), V93–V105.

Chai, X., Wang, S. & Tang, G., 2017. Sparse reflectivity inversion for nonstationary seismic data with surface-related multiples: numerical and field-data experiments, *Geophysics,* **82**(3), R199–R217.

Chartrand, R., 2007. Exact reconstruction of sparse signals via nonconvex minimization, *IEEE Signal Process. Lett.,* **14**(10), 707–710.

Chartrand, R. & Staneva, V., 2010. Restricted isometry properties and nonconvex compressive sensing, *Inverse Probl.,* **24**(3), 657–682.

Chartrand, R. & Yin, W., 2008. Iteratively reweighted algorithms for compressive sensing, in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2008 (ICASSP 2008),* pp. 3869–3872, IEEE.

Chen, Y., 2016. Dip-separated structural filtering using seislet transform and adaptive empirical mode decomposition based dip filter, *Geophys. J. Int.,* **206**(1), 457–469.

Chen, Y., 2017. Fast dictionary learning for noise attenuation of multidimensional seismic data, *Geophys. J. Int.,* **209**(1), 21–31.

Chen, Y. & Fomel, S., 2015. Random noise attenuation using local signal-and-noise orthogonalization, *Geophysics,* **80**, WD1–WD9.

Chen, Y., Zhang, D., Jin, Z., Chen, X., Zu, S., Huang, W. & Gan, S., 2016. Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method, *Geophys. J. Int.,* **206**(3), 1695–1717.

Clarke, G.K., 1968. Time-varying deconvolution filters, *Geophysics,* **33**(6), 936–944.

Combettes, P.L. & Pesquet, J.-C., 2011. Proximal splitting methods in signal processing, in *Fixed-point Algorithms for Inverse Problems in Science and Engineering,* pp. 185–212, Springer.

Combettes, P.L. & Wajs, V.R., 2005. Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.,* **4**(4), 1168–1200.

Dai, N. & West, G.F., 1994. Inverse Q migration, in *64th Annual International Meeting,* SEG, Expanded Abstracts, Vol. 1, pp. 1418–1421.

Donoho, D.L., 2006. Compressed sensing, *IEEE Trans. Inf. Theory,* **52**(4), 1289–1306.

Donoho, D.L. & Huo, X., 2002. Uncertainty principles and ideal atomic decomposition, *IEEE Trans. Inf. Theory,* **47**(7), 2845–2862.

Esser, E., Lou, Y. & Xin, J., 2013. A method for finding structured sparse solutions to nonnegative least squares problems with applications, *SIAM J. Imag. Sci.,* **6**(4), 2010–2046.

Foucart, S. & Lai, M.-J., 2009. Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$, *Appl. Comput. Harmon. Anal.,* **26**(3), 395–407.

Futterman, W.I., 1962. Dispersive body waves, *J. geophys. Res.,* **67**(13), 5279–5291.

Goldstein, T. & Osher, S., 2009. *The Split Bregman Method for L1-Regularized Problems,* Society for Industrial and Applied Mathematics.

Gong, P., Zhang, C., Lu, Z., Huang, J. & Ye, J., 2013. A general iterative shrinkage and thresholding algorithm for non-convex regularized optimization problems, *Proc. Int. Conf. Mach. Learn.,* **28**(2), 37–45.

Gorodnitsky, I.F. & Rao, B.D., 2002. Sparse signal reconstruction from limited data using focuss: a re-weighted minimum norm algorithm, *IEEE Trans. Signal Process.,* **45**(3), 600–616.

Griffiths, L., Smolka, F. & Trembly, L., 1977. Adaptive deconvolution: a new technique for processing time-varying seismic data, *Geophysics,* **42**(4), 742–759.

Hale, E.T., Yin, W. & Zhang, Y., 2008. Fixed-point continuation for ℓ_1-minimization: methodology and convergence, *SIAM J. Optim.,* **19**(3), 1107–1130.

Hansen, P.C., 1992. Analysis of discrete ill-posed problems by means of the l-curve, *SIAM Rev.,* **34**(4), 561–580.

Hargreaves, N.D. & Calvert, A.J., 1991. Inverse Q filtering by fourier transform, *Geophysics,* **56**(4), 519–527.

Holm, S. & Näsholm, S.P., 2013. Comparison of fractional wave equations for power law attenuation in ultrasound and elastography, *Ultrasound Med. Biol.,* **40**(4), 695–703.

Hu, Y., Zhang, D., Ye, J., Li, X. & He, X., 2013. Fast and accurate matrix completion via truncated nuclear norm regularization, *IEEE Trans. Pattern Anal. Mach. Intell.,* **35**(9), 2117–2130.

Huang, C., Nie, L., Schoonover, R.W., Wang, L.V. & Anastasio, M.A., 2012. Photoacoustic computed tomography correcting for heterogeneity and attenuation, *J. Biomed. Opt.,* **17**(6), doi:10.1117/1.JBO.17.6.061211.

Kelly, J.F., McGough, R.J. & Meerschaert, M.M., 2008. Analytical time-domain Green's functions for power-law media, *J. acoust. Soc. Am.,* **124**(5), 2861–2872.

Kibblewhite, A.C., 1989. Attenuation of sound in marine sediments: a review with emphasis on new low-frequency data, *J. acoust. Soc. Am.,* **86**(2), 716–738.

Kjartansson, E., 1979. Constant Q-wave propagation and attenuation, *J. geophys. Res.,* **84**(B9), 4737–4748.

Kolsky, H., 1956. LXXI. The propagation of stress pulses in viscoelastic solids, *Philos. Mag.,* **1**(8), 693–710.

Lai, M.-J., Xu, Y. & Yin, W., 2013. Improved iteratively reweighted least squares for unconstrained smoothed ℓ_q minimization, *SIAM J. Numer. Anal.,* **51**(2), 927–957.

Li, G., Liu, Y., Zheng, H. & Huang, W., 2015. Absorption decomposition and compensation via a two-step scheme, *Geophysics,* **80**(6), V145–V155.

Li, G., Sacchi, M.D. & Zheng, H., 2016a. In situ evidence for frequency dependence of near-surface Q, *Geophys. J. Int.,* **204**(2), 1308–1315.

Li, Q., Zhou, H., Zhang, Q., Chen, H. & Sheng, S., 2016b. Efficient reverse time migration based on fractional Laplacian viscoacoustic wave equation, *Geophys. J. Int.,* **204**(1), 488–504.

Liu, H., Anderson, D.L. & Kanamori, H., 1976. Velocity dispersion due to anelasticity; implications for seismology and mantle composition, *Geophys. J. Int.,* **47**(1), 41–58.

Liu, T. & Pong, T.K., 2017. Further properties of the forward–backward envelope with applications to difference-of-convex programming, *Comput. Optim. Appl.,* **67**(3), 489–520.

Lou, Y. & Yan, M., 2016. Fast L_1–L_2 minimization via a proximal operator, *J. Sci. Comput.,* arXiv:1609.09530.

Lou, Y., Osher, S. & Xin, J., 2015a. Computational aspects of constrained L1-L2 minimization for compressive sensing, *Petrol. Sci. Technol.,* **23**(1), 47–54.

Lou, Y., Yin, P., He, Q. & Xin, J., 2015b. Computing sparse representation in a highly coherent dictionary based on difference of ℓ_1 and ℓ_2, *J. Sci. Comput.,* **64**(1), 178–196.

Ma, T.-H., Lou, Y. & Huang, T.-Z., 2017. Truncated l_1-2 models for sparse recovery and rank minimization, *SIAM J. Imag. Sci.,* **10**(3), 1346–1380.

Margrave, G.F., Lamoureux, M.P. & Henley, D.C., 2011. Gabor deconvolution: estimating reflectivity by nonstationary deconvolution of seismic data, *Geophysics,* **76**(3), W15–W30.

McDonal, F., Angona, F., Mills, R., Sengbush, R., Van Nostrand, R. & White, J., 1958. Attenuation of shear and compressional waves in Pierre shale, *Geophysics,* **23**(3), 421–439.

Mittet, R., 2007. A simple design procedure for depth extrapolation operators that compensate for absorption and dispersion, *Geophysics,* **72**(2), S105–S112.

Mittet, R., Sollie, R. & Hokstad, K., 1995. Prestack depth migration with compensation for absorption and dispersion, *Geophysics,* **60**(5), 1485–1494.

Moczo, P. & Kristek, J., 2005. On the rheological models used for time-domain methods of seismic wave propagation, *Geophys. Res. Lett.,* **32**(1), doi:10.1029/2004GL021598.

Ng, M., Zhao, X.L. & Wang, F., 2016. Multiplicative noise and blur removal by framelet decomposition and and l_1-based l-curve method, *IEEE Trans. Image Process.,* **25**(9), 4222–4232.

Oliveira, S. A.M. & Lupinacci, W.M., 2013. L1 norm inversion method for deconvolution in attenuating media, *Geophys. Prospect.,* **61**(4), 771–777.

Sacchi, M.D., 1997. Reweighting strategies in seismic deconvolution, *Geophys. J. Int.,* **129**(3), 651–656.

Strick, E., 1967. The determination of Q, dynamic viscosity and transient creep curves from wave propagation measurements, *Geophys. J. Int.,* **13**(1-3), 197–218.

Sun, J. & Zhu, T., 2015. Stable attenuation compensation in reverse-time migration, in *85th Annual International Meeting,* SEG, Expanded Abstracts. pp. 3942–3947.

Sun, J., Zhu, T. & Fomel, S., 2015. Viscoacoustic modeling and imaging using low-rank approximation, *Geophysics,* **80**(5), A103–A108.

Sun, J., Fomel, S., Zhu, T. & Hu, J., 2016. Q-compensated least-squares reverse time migration using low-rank one-step wave extrapolation, *Geophysics,* **81**(4), S271–S279.

Szabo, T.L., 1994. Time domain wave equations for lossy media obeying a frequency power law, *J. acoust. Soc. Am.,* **96**(1), 491–500.

Szabo, T.L., 1995. Causal theories and data for acoustic attenuation obeying a frequency power law, *J. acoust. Soc. Am.,* **97**(1), 14–24.

Tao, P.D. & An, L. T.H., 1998. A DC optimization algorithm for solving the trust-region subproblem, *SIAM J. Optim.,* **8**(2), 476–505.

Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso, *J. R. Stat. Soc. Ser. B (Methodol.),* **58**(1), 267–288.

Treeby, B.E., 2013. Acoustic attenuation compensation in photoacoustic tomography using time-variant filtering, *J. Biomed. Opt.,* **18**(3), doi:10.1117/1.JBO.18.3.036008.

Wang, S., 2011. Attenuation compensation method based on inversion, *Appl. Geophys.,* **8**(2), 150–157.

Wang, Y., 2002. A stable and efficient approach of inverse Q filtering, *Geophysics,* **67**(2), 657–663.

Wang, Y., 2004. Q analysis on reflection seismic data, *Geophys. Res. Lett.,* **31**(17), 159–180.

Wang, Y., 2006. Inverse Q-filter for seismic resolution enhancement, *Geophysics,* **71**(3), V51–V60.

Wang, Y., 2015. Generalized viscoelastic wave equation, *Geophys. J. Int.,* **204**(2), 1216–1221.

Wang, S. & Chen, X., 2014. Absorption-compensation method by l1-norm regularization, *Geophysics,* **79**(3), V107–V114.

Wang, Y. & Guo, J., 2004a. Modified Kolsky model for seismic attenuation and dispersion, *J. geophys. Eng.,* **1**(3), 187–196.

Wang, Y. & Guo, J., 2004b. Seismic migration with inverse Q filtering, *Geophys. Res. Lett.,* **31**(21), 163–183.

Wang, S., Li, X.Y., Qian, Z., Di, B. & Wei, J., 2007. Physical modelling studies of 3-D p-wave seismic for fracture detection, *Geophys. J. Int.,* **168**(2), 745–756.

Wang, B., Wu, R.S., Chen, X. & Li, J., 2015. Simultaneous seismic data interpolation and denoising with a new adaptive method based on dreamlet transform, *Geophys. J. Int.,* **201**(2), 1182–1194.

Wang, Y. & Yin, W., 2010. Sparse signal reconstruction via iterative support detection, *SIAM J. Imag. Sci.,* **3**(3), 462–491.

Wang, Y., Zhou, H., Chen, H. & Chen, Y., 2017a. Adaptive stabilization for q-compensated reverse time migration, *Geophysics,* **83**(1), S15–S32.

Wang, Y., Zhou, H., Zu, S., Mao, W. & Chen, Y., 2017b. Three-operator proximal splitting scheme for 3-D seismic data reconstruction, *IEEE Geosci. Remote Sens. Lett.,* **4**(10), 1830–1834.

Waters, K.R., Hughes, M.S., Mobley, J. & Miller, J.G., 2003. Differential forms of the Kramers-Krönig dispersion relations, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control,* **50**(1), 68–76.

Waters, K.R., Mobley, J. & Miller, J.G., 2005. Causality-imposed (Kramers-Krönig) relationships between attenuation and dispersion, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control,* **52**(5), 822–823.

Woodworth, J. & Chartrand, R., 2016. Compressed sensing recovery via nonconvex shrinkage penalties, *Inverse Probl.,* **32**(7), 75 004–75 028.

Wu, X. & Hale, D., 2016. 3D seismic image processing for faults, *Geophysics,* **81**(2), IM1–IM11.

Wu, X. & Janson, X., 2017. Directional structure tensors in estimating seismic structural and stratigraphic orientations, *Geophys. J. Int.,* **210**(1), 534–548.

Wuenschel, P.C., 1965. Dispersive body waves: an experimental study, *Geophysics,* **30**(4), 539–551.

Xu, P., 1998. Truncated svd methods for discrete linear ill-posed problems, *Geophys. J. Int.,* **135**(2), 505–514.

Xue, Z., Zhu, T., Fomel, S. & Sun, J., 2016. Q-compensated full-waveform inversion using constant-Q wave equation, in *SEG Technical Program Expanded Abstracts,* pp. 1063–1068.

Xue, Z., Zhu, H. & Fomel, S., 2017. Full waveform inversion using seislet regularization, *Geophysics,* **82**(5), A43–A49.

Yang, P., Brossier, R., Métivier, L. & Virieux, J., 2016. A review on the systematic formulation of 3-D multiparameter full waveform inversion in viscoelastic medium, *Geophys. J. Int.,* **207**(1), 129–149.

Yilmaz, O., 2001. *Seismic Data Analysis,* Society of Exploration Geophysicists.

Yin, P., Lou, Y., He, Q. & Xin, J., 2015. Minimization of ℓ_{1-2} for compressed sensing, *SIAM J. Sci. Comput.,* **37**(1), A536–A563.

Yuan, S., Wang, S., Ma, M., Ji, Y. & Deng, L., 2017. Sparse bayesian learning-based time-variant deconvolution, *IEEE Trans. Geosci. Remote Sens.,* **55**(11), 6182–6194.

Zhang, C.H., 2010a. Nearly unbiased variable selection under minimax concave penalty, *Annal. Stat.,* **38**(2), 894–942.

Zhang, T., 2010b. Analysis of multi-stage convex relaxation for sparse regularization, *J. Mach. Learn. Res.,* **11**(Mar), 1081–1107.

Zhang, C. & Ulrych, T.J., 2007. Seismic absorption compensation: a least squares inverse scheme, *Geophysics,* **72**(6), R109–R114.

Zhang, J., Wu, J. & Li, X., 2012. Compensation for absorption and dispersion in prestack migration: an effective Q approach, *Geophysics,* **78**(1), S1–S14.

Zhang, Y., Zhang, P. & Zhang, H., 2010. Compensating for visco-acoustic effects in reverse-time migration, in *80th Annual International Meeting,* SEG, Expanded Abstracts, pp. 3160–3164.

Zhou, Y., Li, S., Zhang, D. & Chen, Y., 2017. Seismic noise attenuation using an online subspace tracking algorithm, *Geophys. J. Int.,* **212**(2), 1072–1097.

Zhu, T. & Carcione, J.M., 2013. Theory and modelling of constant-Q P- and S-waves using fractional spatial derivatives, *Geophys. J. Int.,* **196**(3), 1787–1795.

Zhu, T., Harris, J.M. & Biondi, B., 2014. Q-compensated reverse-time migration, *Geophysics,* **79**(3), S77–S87.

APPENDIX A: PROOF FOR CONVERGENCE OF DCA

Here, we briefly demonstrate that the DCA iteration formula (24) yields a monotonically decreasing sequence of objective values. According to the fact that x^{k+1} minimizes $G(x) - (H(x^k) + \langle y^k, x - x^k \rangle)$, we have

$$G(x^{k+1}) - \left(H(x^k) + \langle y^k, x^{k+1} - x^k \rangle\right) \leq G(x^k) - H(x^k). \qquad \text{(A1)}$$

By the definition of subgradient y^k, we have

$$H(x) \leq H(x^k) + \langle y^k, x - x^k \rangle, \forall x \in \mathbb{R}^n. \qquad \text{(A2)}$$

In particular, $H(x^{k+1}) \leq H(x^k) + \langle y^k, x^{k+1} - x^k \rangle$, consequently

$$
\begin{aligned}
F(x^{k+1}) = G(x^{k+1}) - H(x^{k+1}) &\leq G(x^{k+1}) - \big(H(x^k) \\
&\quad + \langle y^k, x^{k+1} - x^k \rangle\big) \\
&\leq G(x^k) - H(x^k) = F(x^k). \quad \text{(A3)}
\end{aligned}
$$

Thus, we have proven the monotonically decreasing property of DCA.

APPENDIX B: THE DERIVATION OF EQ. (27)

According to DCA formula (24), we have

$$H(x^k) + \langle y^k, x - x^k \rangle = H(x^k) + \langle y^k, x \rangle - \langle y^k, x^k \rangle, \qquad \text{(B1)}$$

where $H(x^k) = -\alpha\lambda\|x^k\|_2$ and $y^k = -\alpha\lambda\frac{x^k}{\|x^k\|_2}$, then

$$
\begin{aligned}
H(x^k) + \langle y^k, x - x^k \rangle &= -\alpha\lambda\|x^k\|_2 + \left\langle -\alpha\lambda\frac{x^k}{\|x^k\|_2}, x \right\rangle \\
&\quad - \left\langle -\alpha\lambda\frac{x^k}{\|x^k\|_2}, x^k \right\rangle \\
&= -\alpha\lambda\|x^k\|_2 + \langle y^k, x \rangle + \alpha\lambda\|x^k\|_2 \\
&= \langle y^k, x \rangle. \qquad \text{(B2)}
\end{aligned}
$$