

Separation of seismic multiple reflection-refraction based on morphological component analysis with high-resolution linear Radon transform

Dawei Liu¹, Haoqi Zhang¹, Xiaokai Wang¹, Wenchao Chen¹, Zhensheng Shi¹, and Zhonghua Zhao²

ABSTRACT

Investigating coherent noise attenuation is a continuing concern within seismic signal processing. As a common type of linear coherent noise, the multiple reflection-refraction (MRR) occurs in seismic records in which low-velocity strata overlie high-velocity strata, such as deserts, the Loess Plateau, etc. Due to its high velocity and strong energy, MRR seriously distorts the reflections and affects the interpretation. MRR has linear morphological characteristics on the shot gathers. Thus, a linear Radon transform with a surgical mute is usually applied to suppress MRR. However, significant numbers of shallow reflections are removed unintentionally in the τ - p domain when reflections overlap MRR. Therefore, a robust method that reduces the leakage of reflection energy is required. We have developed a novel method to attenuate MRR by examining the morphological difference between MRR and useful signals in

the τ - p domain. MRRs are oblique linear events on the shot records, whereas the useful signals are quasi-hyperbolic events under the assumption of horizontal layers. After the high-resolution linear Radon transform, MRRs are ideally mapped into point features, whereas the useful signals are aligned with narrow curve bands in the ellipse. To better separate them in the τ - p domain, we use the morphological component analysis (MCA) theory and select the 2D stationary wavelet transform and the shearlet transform as sparse representation subdictionaries of point features and curved features, respectively. After MCA separation, we apply the inverse Radon transform to the separated MRRs and subtract them from the original seismic data. Subtraction can better preserve the amplitude of reflections. We use synthetic data and field data to illustrate the effectiveness of our method and demonstrate that making full use of the preceding morphological differences can improve the results of MRR attenuation.

INTRODUCTION

Attenuating coherent noise has been studied by many researchers, but it is still an unsolved problem in seismic signal processing in certain situations. As one type of prominent linear coherent noise, the multiple reflection-refraction (MRR) is an important wave phenomenon in exploration areas with a complex surface, especially on shallow-water seismic profiles or in areas with a low-velocity stratigraphic structure. We use Figure 1 to illustrate the generation mechanism of MRR waves. When the seismic signal is incident from the overlying low-velocity stratum to the top surface of the underlying

high-velocity stratum along the same raypath repeatedly, multiple reflections occur. At the same time, they produce refracted waves. When the incident angle of multiple reflected waves is equal to the critical angle, the refracted waves travel horizontally along the interface on the top surface of the underlying high-velocity stratum. Then, MRRs occur when they reach the receivers after they are reflected back to the free surface (Li, 1983; Qin, 1990). MRR typically has higher velocities and exhibits lower dips than useful reflections on shot records. In addition, this type of wave sometimes has very strong amplitudes, severely distorting the far offsets of early arrivals

First presented at the First International Meeting for Applied Geoscience & Energy (IMAGE), SEG. Manuscript received by the Editor 20 July 2021; revised manuscript received 7 March 2022; published ahead of production 11 April 2022; published online 23 June 2022.

¹Xi'an Jiaotong University, School of Information and Communications Engineering, Xi'an, China. Email: 409791715@qq.com; 1354422686@qq.com; xkwang@xjtu.edu.cn; wenchchen@xjtu.edu.cn (corresponding author); 1303118188@qq.com.

²Exploration and Development Research Institute of Daqing Oilfield, Daqing, China. E-mail: zhaozhonghua@petrochina.com.cn.

© 2022 Society of Exploration Geophysicists. All rights reserved.

and affecting the signal-to-noise ratio (S/N) of seismic records. Thus, discriminating and suppressing MRR events are vitally important.

To date, various studies have investigated coherent noise attenuation, and they can be classified into seismic acquisition and seismic processing. A suitable acquisition geometry design can enable adequate wavefield sampling and reduce coherent noise amplitude from the source. For example, scattered noise can be completely removed in the areal geometry with adequate spatial sampling, and Vermeer (2008) demonstrates that single-point orthogonal geometry can also be applied to suppress scattered noise. Moreover, wide-azimuth geometry can improve multiple attenuations (Meunier et al., 2008). As for the problem of MRR suppression, Liang (1993) finds that MRR energy increases as the shot-receiver offset increases. Accordingly, with an appropriate acquisition geometry layout, e.g., increasing the coverage fold, the MRR amplitude can be reduced. In addition, other seismic acquisition technologies, such as the symmetric sampling method (Vermeer, 1998) and spectrum shaping (Steeple, 2005), also help to suppress MRR. However, acquisition geometry design should consider efficiency and economics and the optimum acquisition geometry often is unaffordable. In addition, such methods are powerless against existing seismic data. Therefore, the greater part of the literature on coherent noise attenuation focuses on seismic data processing.

Several signal-processing filtering methods have been proposed and applied to suppress coherent noise. By designing a suitable high-pass filter, the low-frequency ground-roll components can be filtered out trace by trace. Following that, Chen et al. (2015) combine a band-pass filter with local orthogonalization to improve ground-roll separation. Due to the frequency overlap between MRR and the desired seismic reflections, a simple band-pass-filtering method is unable to effectively remove MRR without harming the reflections. The f - k filters also are standard methods for suppressing coherent noise. For example, Al-Husseini et al. (1981) use the f - k filter to attenuate ground roll. Aiming at MRR attenuation, the f - k filtering method is a practical method as well. However, it distorts reflections when the amplitude of MRR events is much stronger than the desired reflection. Another well-known method is filtering coherent noise in the τ - p domain based on the apparent velocity difference between the reflection and coherent noise (Kabir and Marfurt, 1999; Yu et al., 2007; Ibrahim and Sacchi, 2014). After the linear Radon transform, a hyperbolic reflection is mapped into an ellipse, whereas the linear coherent noise is mapped into a point (Durrani and Bisset, 1984). For instance, the linear ground roll has a point feature in the τ - p domain and can be

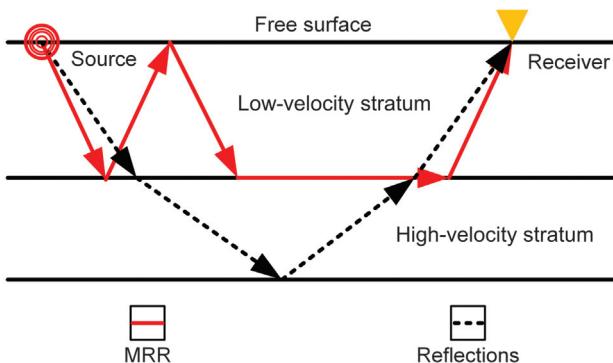


Figure 1. The raypath of an MRR wave.

removed by subtraction strategies (Trad et al., 2003). MRRs are characterized by several groups of parallel oblique events on shot gathers, and they have a certain constant apparent dip in each group. The linear morphological characteristics of MRR make noise attenuation in the τ - p domain a commonly used method. However, a direct mute in the τ - p domain may damage the desired reflections when MRR has similar apparent velocities with reflections. Therefore, a better separation method in the τ - p domain is required to ameliorate the muted out results and to adaptively suppress MRR noise.

With the high-resolution linear Radon transform, MRR noise is characterized by point features, whereas reflections exhibit curved features. Relying on the previously mentioned morphological feature difference in the τ - p domain leads to appealing MRR separation. Research into morphologically decomposing a signal into its building blocks has a long history. Up to now, several techniques have been developed to tackle this problem, such as blind source separation (Liu and Dragoset, 2013) and independent component analysis (De Lauro et al., 2009). Interestingly, sparsity also receives considerable scholarly attention and is widespread in signal separation. The sparse representation theory (Rubinstein et al., 2010) also becomes an important aspect of seismic data processing (Ma et al., 2010). Driven by sparsity, Starck et al. (2005) propose morphological component analysis (MCA) for image separation, using two different dictionaries to achieve a sparse decomposition of texture and illustrate parts of the image. As an effective sparsity-based method for signal decomposition, MCA also plays a vital role in separating seismic coherent noise. Wang et al. (2012) use MCA to attenuate ground roll and construct two dictionaries using the stationary wavelet transform with different scales to represent ground roll and reflections, respectively. Furthermore, Chen et al. (2017) construct two dictionaries using tunable Q -factor wavelet transforms with different Q factors to suppress ground roll. Overall, these studies confirm the strong separation ability of MCA and highlight that the dictionaries selected for decomposition play a crucial role in guaranteeing efficient separation with MCA.

Inspired by the MCA theory, we propose a novel linear coherent noise attenuation method based on waveform morphology differences in the linear Radon domain. First, a high-resolution linear Radon transform is applied to the original seismic data, mapping the reflections into the curved morphological features and the MRR noise into point morphological features. Then, the 2D stationary wavelet transform (2D-SWT) and the shearlet transform are selected as sparse representation dictionaries of curved morphological component and point morphological component, respectively. After that, the block coordinate relaxation algorithm (Sardy et al., 2000) is used to separate the point morphological component iteratively. Finally, the linear coherent noise component is reconstructed by the inverse Radon transform, and the separated MRR noise is subtracted from the original data, thus suppressing the linear coherent noise. The application of synthetic data and a 2D common-shot gather of field data demonstrate our method's effectiveness in attenuating MRRs.

THEORY

The detailed procedure of the proposed method for attenuating MRR is illustrated in Figure 2. This section introduces the high-resolution linear Radon transform, MCA theory, and the dictionaries chosen for sparsely representing the reflections and MRR noise.

High-resolution linear Radon transform

The Radon transform proposed by Radon (1986) is an integral transform and has been an important tool for image analysis (Al-Shaykh and Doherty, 1996) and signal reconstruction (Lanza-vecchia et al., 1999). Recently, a growing body of literature has investigated the applications of Radon transformations in the exploration geophysics field (Lu, 2013; Xue et al., 2014; Kazemi and Sacchi, 2021). MRR possesses the characteristics of linear morphology, and such linear features are transformed into a point by the linear Radon transform, making it easier to distinguish them from the reflections. Therefore, we suppress the MRR in the τ - p domain using the linear Radon transform.

We define the formula of the linear Radon transform as

$$m(\tau, p) = \int_{-\infty}^{+\infty} d(t = \tau + px, x) dx. \quad (1)$$

Here, $d(t, x)$ denotes the seismic signal, τ is intercept time, p is apparent slowness, and $m(\tau, p)$ denotes the Radon coefficients. The τ - p transform is a discretized Radon transform and can be expressed in matrix-vector notation as

$$\mathbf{d} = \mathbf{L}\mathbf{m}, \quad (2)$$

where \mathbf{d} and \mathbf{m} represent the data and model, respectively, and \mathbf{L} denotes the forward operator of the linear Radon transform. The estimated Radon coefficients can be calculated using the adjoint operator \mathbf{L}^T by

$$\mathbf{m}_{\text{adj}} = \mathbf{L}^T \mathbf{d}. \quad (3)$$

Because the linear Radon transform is not an orthogonal transformation, it is intractable to compute \mathbf{m}_{adj} directly through the adjoint transformation, which is a major obstacle to applications in the seismic data processing. Therefore, the Radon coefficients $\tilde{\mathbf{m}}$ can be calculated with the least-squares solution by posing an inverse problem (Thorson and Claerbout, 1985):

$$\tilde{\mathbf{m}} = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{d}. \quad (4)$$

To obtain a better resolution and to reduce the smearing problem in the Radon panel, Sacchi and Ulych (1995) introduce the sparsity constraint and estimate the high-resolution Radon coefficients by minimizing the objective function:

$$J = \|\mathbf{d} - \mathbf{L}\mathbf{m}\|_2^2 + \lambda \|\mathbf{m}\|_1, \quad (5)$$

where λ denotes a trade-off parameter to balance the l_2 -norm data misfit and the l_1 -norm model constraints. We define a model weight matrix \mathbf{W} with diagonal elements proportional to the i th element of \mathbf{m} , as shown in the following equation:

$$w_{i,i} = \frac{1}{\sqrt{|m_i|}}. \quad (6)$$

Thus, the l_1 -norm regularization can be transformed into an l_2 -norm regularization because

$$\|\mathbf{m}\|_1^2 = \sum_i |m_i| = \mathbf{m}^T \mathbf{W}^T \mathbf{W} \mathbf{m} = \|\mathbf{W} \mathbf{m}\|_2^2. \quad (7)$$

Prior studies have noted the importance of adding a minimum threshold to the denominator to avoid dividing by zero (Trad et al., 2003). We set this hyperparameter to 10^{-5} according to our tests. Then, the objective function in equation 5 can be expressed as the following quadratic objective function:

$$J_q = \|\mathbf{d} - \mathbf{L}\mathbf{m}\|_2^2 + \lambda \|\mathbf{W} \mathbf{m}\|_2^2. \quad (8)$$

Denoting $\mathbf{u} = \mathbf{W} \mathbf{m}$, equation 8 can be reformulated to a standard form by

$$J_q = \|\mathbf{d} - \mathbf{L}\mathbf{W}^{-1}\mathbf{u}\|_2^2 + \lambda \|\mathbf{u}\|_2^2. \quad (9)$$

The minimization of equation 9 produces the following equation:

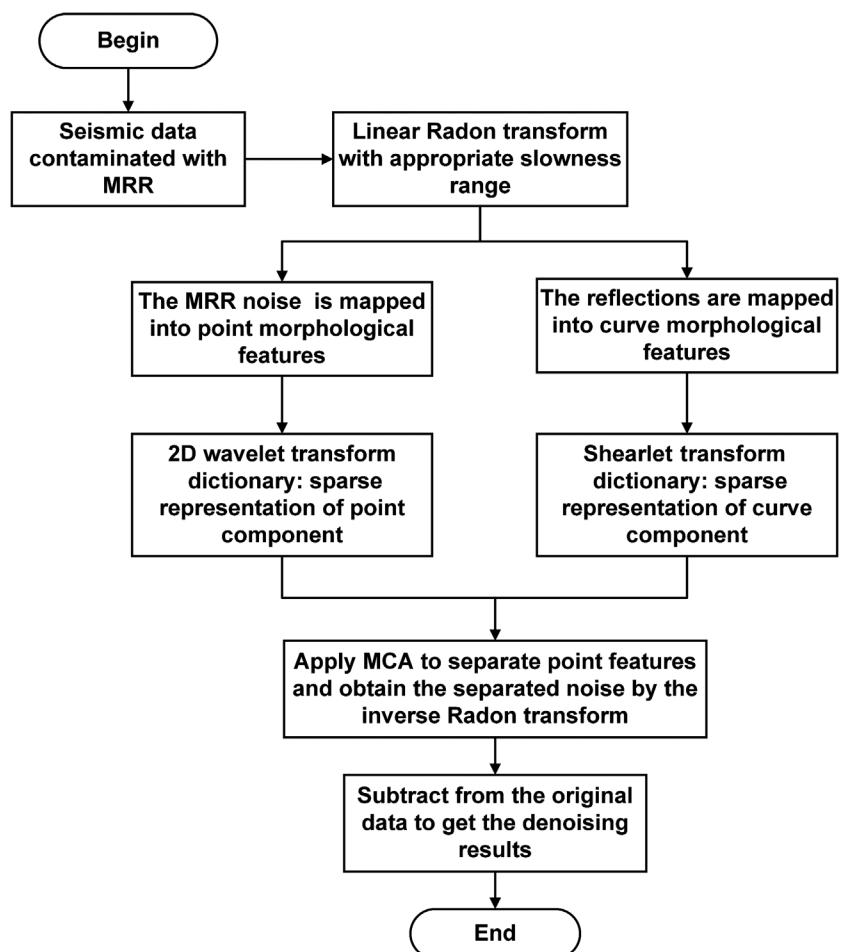


Figure 2. The procedure of the proposed method for MRR attenuation.

$$(\mathbf{W}^{-T}\mathbf{L}^T\mathbf{L}\mathbf{W}^{-1} + \lambda\mathbf{I})\mathbf{W}\mathbf{m} = \mathbf{W}^{-T}\mathbf{L}^T\mathbf{d}, \quad (10)$$

where \mathbf{I} represents the identity matrix. Solving equation 10 is a large ill-conditioned problem and time-consuming. In practice, the conjugate gradient method (Scales, 1987) is frequently used to solve this least-squares matrix inversion to increase the computation efficiency. Moreover, the conjugate gradient method is satisfactory because it does not require matrices in explicit form when using the time-domain Radon operators. After suitable iterations are carried out, the high-resolution linear Radon transform can focus linear MRR events into points for subsequent separation. It is worth mentioning that the choice of the range of p should be deliberate. The subsurface geologic conditions determine the velocity range of seismic waves and further the dip range of MMRs. Therefore, we use this prior knowledge to select the appropriate slowness range for the Radon transform, which helps to achieve fast and accurate MRR attenuation.

MCA

Based on the sparse representation theory (Chen et al., 2001), Starck et al. (2005) propose the MCA approach for separating different components of a signal according to the morphological differences of each component. The core idea of MCA is to construct overcomplete sparse representation dictionaries by combining together different dictionaries, each of which has a different atomic characteristic and can identify a certain morphological component of the signal. Taking an input signal $\mathbf{s} \in \mathbb{R}^{N \times 1}$ as an example, we assume that it is a linear superposition of two signals \mathbf{s}_1 and \mathbf{s}_2 , i.e., $\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2$, and they have different morphological features. We note that the two signals here are only for the sake of simplicity, and MCA can handle more than two signal components in a similar manner. MCA makes the following assumptions:

Assumption 1: \mathbf{s} is composed of two different morphological features. Correspondingly, there are two sparse representation overcomplete dictionaries, and each signal component $\mathbf{s}_i (i = 1, 2)$ can be sparsely represented by one of the dictionaries by computing the coefficients $\mathbf{x}_i \in \mathbb{R}^{L_i}$ as follows:

$$\mathbf{x}_i^{\text{opt}} = \arg \min_{\mathbf{x}_i} \|\mathbf{x}_i\|_0 \quad \text{s.t. } \mathbf{s}_i = \mathbf{D}_i \mathbf{x}_i \quad \text{for } i = 1, 2, \quad (11)$$

where $\mathbf{D}_i \in \mathbb{R}^{N \times L_i}$ (typically $L_i \gg N$) is an overcomplete dictionary for \mathbf{s}_i and $\mathbf{x}_i^{\text{opt}}$ is the sparse solution, i.e., $\|\mathbf{x}_i^{\text{opt}}\|_0 \ll N$. The l_0 norm counts the total number of nonzero elements of a vector.

Assumption 2: For any given component \mathbf{s}_i , the other dictionary $\mathbf{D}_j (i \neq j)$ leads to a nonsparse solution by

$$\forall \{i, j\} \in \{1, 2\}, \quad \mathbf{x}_{ij}^{\text{opt}} = \arg \min_{\mathbf{x}_{ij}} \|\mathbf{x}_{ij}\|_0 \quad \text{s.t. } \mathbf{s}_i = \mathbf{D}_j \mathbf{x}_{ij} (i \neq j), \quad (12)$$

where the l_0 norm of coefficient $\|\mathbf{x}_{ij}^{\text{opt}}\|_0 > \|\mathbf{x}_i^{\text{opt}}\|_0$, which means that \mathbf{s}_i are represented very inefficiently by the dictionary \mathbf{D}_j .

Based upon the preceding assumptions, MCA is proposed to seek the sparsest representations of signal \mathbf{s} over the augmented dictionary composed of \mathbf{D}_1 and \mathbf{D}_2 (Elad et al., 2005). Thus, we need to compute

$$\{\mathbf{x}_1^{\text{opt}}, \mathbf{x}_2^{\text{opt}}\} = \arg \min_{\{\mathbf{x}_1, \mathbf{x}_2\}} \|\mathbf{x}_1\|_0 + \|\mathbf{x}_2\|_0 \quad \text{s.t. } \mathbf{s} = \mathbf{D}_1 \mathbf{x}_1 + \mathbf{D}_2 \mathbf{x}_2. \quad (13)$$

By solving equation 13, we can obtain the representation coefficients of the two components of \mathbf{x} and successfully separate the two components. However, this optimization problem is a nonconvex and nondeterministic polynomial (NP)-hard problem, indicating that the computational complexity grows exponentially with the number of columns in the augmented dictionary. To convert it to a solvable optimization problem, we replace the l_0 -norm constraint with an l_1 -norm constraint and relax the equality constraint, thus leading to the convex minimization problem (Starck et al., 2005):

$$\{\mathbf{x}_1^{\text{opt}}, \mathbf{x}_2^{\text{opt}}\} = \arg \min_{\{\mathbf{x}_1, \mathbf{x}_2\}} \|\mathbf{s} - \mathbf{D}_1 \mathbf{x}_1 - \mathbf{D}_2 \mathbf{x}_2\|_2^2 + \lambda (\|\mathbf{x}_1\|_1 + \|\mathbf{x}_2\|_1), \quad (14)$$

where λ is the Lagrange multiplier to balance the misfit term and constraint term. Equation 14 can be solved by the block coordinate relaxation method (Sardy et al., 2000) through the iterative calculation of the following subproblems:

$$\mathbf{x}_1^{k+1} = \arg \min_{\mathbf{x}_1} \|\mathbf{s} - \mathbf{D}_1 \mathbf{x}_1 - \mathbf{D}_2 \mathbf{x}_2^k\|_2^2 + \lambda \|\mathbf{x}_1\|_1, \quad (15)$$

$$\mathbf{x}_2^{k+1} = \arg \min_{\mathbf{x}_2} \|\mathbf{s} - \mathbf{D}_1 \mathbf{x}_1^k - \mathbf{D}_2 \mathbf{x}_2\|_2^2 + \lambda \|\mathbf{x}_2\|_1. \quad (16)$$

Each subproblem is a typical $l_2 - l_1$ sparse optimization problem and can be solved by the iterative soft-thresholding algorithm (Donoho and Johnstone, 1994):

$$\mathbf{x}_1^{k+1} = S_\lambda(\mathbf{D}_1^\dagger(\mathbf{s} - \mathbf{D}_2 \mathbf{x}_2^k)), \quad (17)$$

$$\mathbf{x}_2^{k+1} = S_\lambda(\mathbf{D}_2^\dagger(\mathbf{s} - \mathbf{D}_1 \mathbf{x}_1^k)), \quad (18)$$

where \mathbf{D}_1^\dagger and \mathbf{D}_2^\dagger are the Moore-Penrose pseudoinverse of \mathbf{D}_1 and \mathbf{D}_2 , respectively. The variable S_λ is the soft-thresholding operator with a threshold λ :

$$S_\lambda(a) = \begin{cases} a - \lambda \text{ sign}(a), & |a| \geq \lambda \\ 0, & \text{others} \end{cases}. \quad (19)$$

After termination of the iteration procedure, we obtain the separated components by

$$\mathbf{s}_1 = \mathbf{D}_1 \mathbf{x}_1^{\text{opt}}, \quad (20)$$

$$\mathbf{s}_2 = \mathbf{D}_2 \mathbf{x}_2^{\text{opt}}. \quad (21)$$

According to MCA theory, we assume that the original seismic data in the τ - p domain \mathbf{m} are the superposition of the reflections of useful signal \mathbf{m}_r , linear MRR events \mathbf{m}_l , and random noise \mathbf{m}_n . Then, the Radon model \mathbf{m} can be expressed as

$$\mathbf{m} = \mathbf{m}_r + \mathbf{m}_l + \mathbf{m}_n. \quad (22)$$

To illustrate the distinguishable morphological features between \mathbf{m}_r and \mathbf{m}_l , we generate synthetic data based on the seismic convolution model, as shown in Figure 3a. These synthetic data have 128 traces with a trace interval of 15 m and 512 time sampling points with a sampling rate of 2 ms. A Ricker wavelet with a central frequency of 30 Hz is used in simulating the reflections. Then, the MRR noise is simulated by six linear events, as shown in Figure 3c.

Through a linear superposition of Figure 3a and 3c, Figure 3e is created, which can be regarded as synthetic seismic data contaminated by linear noise. After the high-resolution linear Radon transform with the slowness range of $[0, 0.65]$ s/km, \mathbf{m}_r , \mathbf{m}_l , and \mathbf{m} are obtained, as displayed in Figure 3b, 3d, and 3f, respectively. From Figure 3b and 3d, we observe that \mathbf{m}_r and \mathbf{m}_l have completely different morphological characteristics, which satisfy the assumed conditions of MCA separation. In addition, the red ellipse shown in Figure 3f reveals that MRR noise often overlaps reflections in the Radon panel even using the high-resolution linear Radon transform, making it challenging to select the suitable mute regions in the τ - p domain. In summary, it is feasible and necessary to suppress MRR in the τ - p domain. Once we construct the proper augmented dictionary, the attenuation of MRR can be achieved by solving

$$\{\mathbf{x}_r^{\text{opt}}, \mathbf{x}_l^{\text{opt}}\} = \arg \min_{\{\mathbf{x}_r, \mathbf{x}_l\}} \|\mathbf{m} - \mathbf{D}_r \mathbf{x}_r - \mathbf{D}_l \mathbf{x}_l\|_2^2 + \lambda(\|\mathbf{x}_r\|_1 + \|\mathbf{x}_l\|_1), \quad (23)$$

where \mathbf{x}_r and \mathbf{x}_l are the sparse representation coefficients of \mathbf{m}_r in the overcomplete dictionary \mathbf{D}_r and \mathbf{m}_l in the overcomplete dictionary \mathbf{D}_l , respectively.

Construction of two sparse representation dictionaries

The key to solving equation 23 is to determine two appropriate sparse representation dictionaries. The sparser the sparse representation coefficients, the better the separation effect. As can be seen from Figure 3b and 3d, the ideal MRRs are sharply focused points in the Radon panel, whereas ideal reflections are ellipses. Therefore, we need to choose a dictionary to sparsely represent point features and a dictionary to sparsely represent curved features. To date, a significant number of signal representations have been introduced to analyze seismic signals, such as the Fourier transform, wavelet transform, shearlet transform, etc. According to the τ - p morphological differences between MRR and the reflected waves, this paper selects the 2D-SWT as the sparse representation dictionary for MRR and the 2D shearlet transform as the sparse representation dictionary for the reflected waves in the Radon domain.

2D-SWT

Wavelet transforms, as isotropic objects, are suitable for detecting objects with point singularities, especially 2D-SWT, which is a nonsubsampling form of the 2D discrete wavelet transform. The 2D-SWT has been used in suppressing acquisition footprint (Cvetkovic et al., 2007) and ground roll (Wang et al., 2012) because it is designed to improve the inadequacy of translation invariance of the 2D discrete wavelet transform. Moreover, 2D-SWT can suppress Gibbs phenomena effectively. Therefore, 2D-SWT is selected to be the dictionary for capturing the abrupt point features.

The 2D-SWT is defined by a tensor product of a 1D-SWT scaling function $\phi(t)$ and wavelet function $\psi(t)$:

$$\begin{aligned} \phi(x, y) &= \phi(x)\phi(y), & \psi^H(x, y) &= \phi(x)\psi(y), \\ \psi^V(x, y) &= \psi(x)\phi(y), & \psi^D(x, y) &= \psi(x)\psi(y), \end{aligned} \quad (24)$$

where $\phi(x, y)$ is the scaling function. The $\psi^H(x, y)$, $\psi^V(x, y)$, and $\psi^D(x, y)$ are three oriented wavelets, denoting horizontal, vertical, and diagonal direction, respectively. The wavelet basis function here is the Symmlet wavelet with a vanishing moment of four. Based on the à trous algorithm (Shensa, 1992), we obtain the wavelet filter banks H and G to decompose a 2D signal $f(x, y)$:

$$\begin{aligned} A_{j+1}[u, v] &= \sum_x \sum_y H_j[x-2u]H_j[y-2v]A_j[x, y], \\ D_{j+1}^H[u, v] &= \sum_x \sum_y H_j[x-2u]G_j[y-2v]A_j[x, y], \\ D_{j+1}^V[u, v] &= \sum_x \sum_y G_j[x-2u]H_j[y-2v]A_j[x, y], \\ D_{j+1}^D[u, v] &= \sum_x \sum_y G_j[x-2u]G_j[y-2v]A_j[x, y], \end{aligned} \quad (25)$$

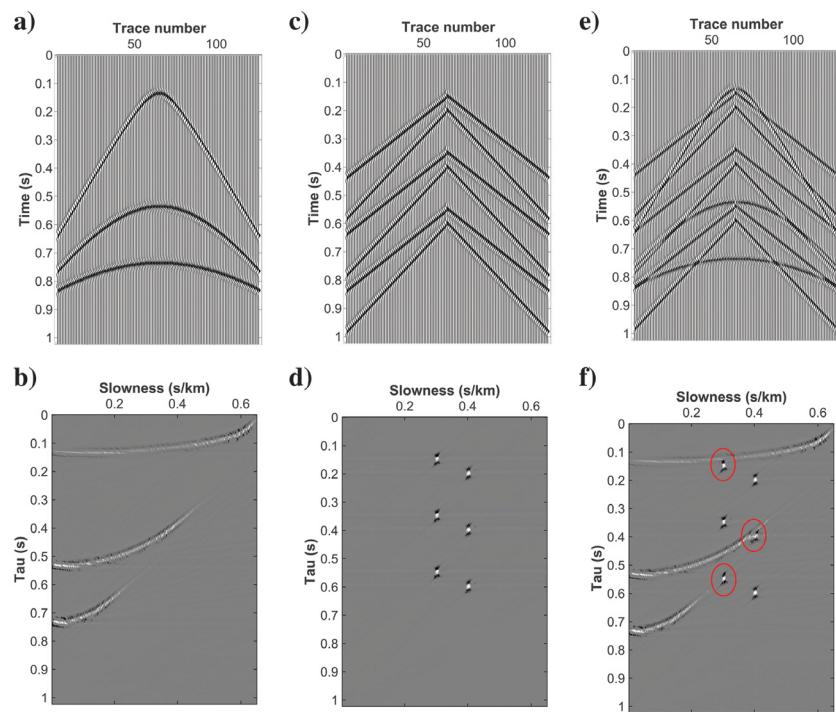


Figure 3. Comparison of morphological differences between MRR and reflections in the Radon panel. (a) Synthetic data containing three reflections, (b) high-resolution linear Radon transform in the time domain of (a), (c) synthetic linear MRR noise, (d) high-resolution linear Radon transform in the time domain of (c), (e) synthetic data superposed with linear noise, and (f) high-resolution linear Radon transform in the time domain of (e). The ellipses indicate an overlap between MRR and reflections in the Radon panel. To make the point features more pronounced, we choose a relatively larger λ in equation 9. At the same time, the curved components are slightly unsmooth or not well sampled due to the sparsity constraint. However, overall, they still have a curvilinear morphology.

where A_j is a set of approximation coefficients and D_j^H , D_j^V , and D_j^D are sets of horizontal, vertical, and diagonal detail coefficients at scale j , respectively. They are the same size as the original signal $f(x, y)$. Here, H_j and G_j are the filter banks of the j th scale decomposition, which can be generated by padding $2^j - 1$ zeros between each coefficient of H and G , respectively. After MCA separation, we can retrieve the signal with the coefficients by the inverse 2D-SWT:

$$\begin{aligned} A_j[x, y] &= \frac{1}{4} \sum_{i=0}^3 \left\{ \sum_u \sum_v \tilde{H}_j[x-2u-i]\tilde{H}_j[y-2v-i]A_{j+1}[u, v] \right. \\ &+ \sum_u \sum_v \tilde{H}_j[x-2u-i]\tilde{G}_j[y-2v-i]D_{j+1}^H[u, v] \\ &+ \sum_u \sum_v \tilde{G}_j[x-2u-i]\tilde{H}_j[y-2v-i]D_{j+1}^V[u, v] \\ &\left. + \sum_u \sum_v \tilde{G}_j[x-2u-i]\tilde{G}_j[y-2v-i]D_{j+1}^D[u, v] \right\}, \quad (26) \end{aligned}$$

where \tilde{H} and \tilde{G} are synthesis filter banks. The atoms of 2D-SWT are shown in Figure 4a, and they have three scales. Within each scale, there are three directions: horizontal, vertical, and diagonal. The atoms are multiscale and isotropic. Hence, the 2D-SWT has excellent sparse representation ability for point singularity targets, and we use it to extract the point features of MRRs in the Radon panel.

Shearlet transform

It is well known that the wavelet transform has limitations in capturing the geometric regularity along a curve because its representation elements are generated based on isotropic dilation. They are well localized but only distributed at various scales and locations. The representation elements also must be distributed in various directions to exploit the anisotropic regularity in multivariate functions, such as edges in images. In this way, coefficients are sparse because some atoms are nearly parallel to the edges. Several

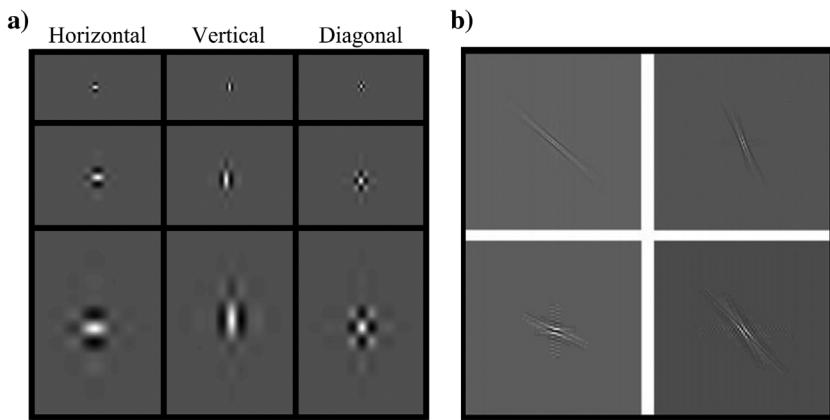


Figure 4. Comparison of 2D-SWT atoms and shearlet transform atoms. (a) The 2D-SWT atoms. The 2D-SWT atoms are isotropic and can capture point features, and thus are used to sparsely represent the MRR in the Radon panel. (b) Shearlet transform atoms. Shearlet transform atoms are isotropic and can capture curved singularity features, and thus are used to sparsely represent the reflections in the Radon panel.

sparse multidimensional representations have been proposed to address this issue, including the directional wavelets, the curvelet transform, the contourlet transform, etc. In particular, the shearlet transform, which is first introduced by Labate et al. (2005), is a novel multiscale geometric analysis tool. The shearlet transform provides optimally sparse representations for 2D functions that are smooth except for discontinuities along curves (Guo and Labate, 2007). According to our previous analysis, ideal reflections present curved features in the Radon panel. Therefore, we select the shearlet transform as the sparse representation dictionary for reflections.

Shearlets are obtained by dilating, shearing, and shifting the generating functions:

$$\left\{ \psi_{i,j,\mathbf{k}}(\mathbf{x}) = a^{\frac{3}{4}}\psi(\mathbf{S}_j\mathbf{A}_i\mathbf{x} - \mathbf{k}) : i, j \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^2 \right\}, \quad (27)$$

where \mathbf{A}_i is the parabolic scaling matrix $\mathbf{A}_a = \begin{bmatrix} a & 0 \\ 0 & a^{1/2} \end{bmatrix}$ ($a > 0$) used to change the resolution, \mathbf{S}_j is the shearing matrix $\mathbf{S}_j = \begin{bmatrix} 1 & j \\ 0 & 1 \end{bmatrix}$ used to change the orientation, \mathbf{k} is the translation parameter to change the positioning of the band-limited function ψ , and \mathbf{x} is the spatial location. The corresponding shearlet transform is calculated by the inner product:

$$SH_\psi f(i, j, \mathbf{k}) = \langle f, \psi_{i,j,\mathbf{k}} \rangle. \quad (28)$$

Figure 4b shows the atoms of the shearlet transform, which exhibits exactly geometric properties of multiscale, anisotropy, and directionality. That is, the shearlet transform has the advantage that it can capture curve singularity features. Hence, we use it to extract the curved features of useful signals in the τ - p domain.

Sparsity analysis

Primary criteria for the successful MCA separation rely on proper dictionaries. Each overcomplete dictionary should sparsely represent a corresponding signal component, but it cannot obtain sparse results to other signal components. To further illustrate the rationality of our chosen dictionaries, we examine the sparsity using the synthetic data shown in Figure 3. The sparsity can be regarded as the ratio of the number of nonzero elements to the number of all elements in the coefficient matrix. Obviously, the smaller the value, the smaller the proportion of nonzero elements in the matrix and the higher the sparsity. The corresponding sparsity measurements using 2D-SWT and the shearlet transform are shown in Table 1. Compared with the shearlet transform, 2D-SWT has higher sparsity when it represents the pointwise singularity component (linear noise). Meanwhile, the shearlet transform has higher sparsity when it represents the curved singularity component (reflections). Therefore, 2D-SWT and the shearlet transform are selected as the sparse representation dictionaries of MRR and reflections, respectively, which meet the basic requirements of MCA.

EXAMPLES

Synthetic data examples

First, the synthetic seismic data shown in Figure 3f are used to verify the effectiveness of the proposed method. The f - k filtering and Radon filtering techniques are commonly used methods in the industry and are therefore used to suppress MRR noise as comparative approaches. According to the separation strategy mentioned previously, the Radon panel shown in Figure 5a shows the separated pointwise components. It can be seen that the point features belonging to MRR are well identified and separated. After the inverse Radon transform, we obtain the reconstructed MRR, as shown in Figure 6a. Finally, we subtract the separated MRR from the original data to obtain the denoised reflections, as shown in Figure 6b. The linear noises removed by the f - k filtering and Radon filtering methods are shown in Figure 6c and 6e, respectively. We see that the linear noise is completely removed, but there is noticeable signal leakage. Similarly, the filtered MRR is subtracted from the original data to obtain the denoised results, as shown in Figure 6d and 6f, respectively. Through an in-depth investigation of the preceding figures, it can be found that most of the linear noise can be suppressed. However, compared to Figure 6b, it can be observed that, at the far offset of the second reflection event, especially the part indicated by the red arrows, the denoised results obtained with the f - k filtering and Radon filtering methods obviously damage the effective reflection waves. Moreover, there is still obvious noise residue after the f - k filtering. In contrast, the denoising results obtained by the proposed method have done almost no damage to the reflection events, thus exhibiting good fidelity.

To further analyze the fidelity to the effective signal amplitude of the proposed method, the 34th trace data of the synthetic data are extracted, as shown in Figure 7. A comparison of Figure 7a and 7b reveals that the effective signal waveform is submerged in the linear interference, which is difficult to identify. Comparing the results of the proposed method shown in Figure 7c, the f - k filtering method shown in Figure 7d, and the Radon filtering method shown in Figure 7e, it can be concluded that the three methods can suppress most of the linear noise, but they cause different levels of damage to useful signals. The f - k filtering and Radon filtering seriously damage the waveform of the second reflected wave, as marked by the red box. Moreover, the Radon filtering method also distorts the waveform of the first arrivals, as shown in the green box in Figure 7e. In contrast, when using the proposed method, the amplitude of the effective wave is well preserved, particularly in the red box region, and the waveform has no significant distortion. A closer inspection shows that the apparent velocity of the damaged part of the reflection events is similar to that of linear noise. In other words, they are overlapping in the f - k domain and are vulnerable with the f - k filtering method. In the same way, although MRRs are focused on points and are more easily located after the high-resolution linear Radon transform, aliasing of point features and curved features still appears. Turning now to the experimental evidence in the red ellipse area shown in Figure 3f, we see that the

slowness and time intercept of the reflections are close to that of the linear noise. In this way, a direct mute in the Radon domain may damage the energy of the effective signal. However, the proposed method can effectively separate the point and curved components according to their morphological differences, thereby avoiding damage to the reflection events. Moreover, the noise suppression ability of the proposed method is superior to the other two methods.

To perform a quantitative analysis of the denoising results of the three methods, we calculate the S/N as the well-known formulation:

$$S/N(\mathbf{m}_r, \hat{\mathbf{m}}_r) = 10 \log_{10} \frac{\|\mathbf{m}_r\|_2^2}{\|\hat{\mathbf{m}}_r - \mathbf{m}_r\|_2^2}, \quad (29)$$

where \mathbf{m}_r represents the clean reflections and $\hat{\mathbf{m}}_r$ represents the denoised results. As can be seen from Table 2, all three methods effectively improve the S/N compared with the original noisy data. Furthermore, the S/N of the proposed method is significantly higher than the other two methods, which proves the effectiveness of our method in suppressing linear interference. Overall, these results indicate that the proposed method effectively suppresses most of the MRR noise and preserves the energy of the effective signal well, which is beneficial for subsequent analysis and processing.

To make the synthetic data more realistic, we create new synthetic data based on wave equation forward modeling, as shown in Figure 8. The synthetic data mainly contain three effective reflections as indicated by the white arrows and three linear noise events as indicated by the red arrows. In addition, some multiples also

Table 1. Sparsity measurements of synthetic data using 2D-SWT and shearlet transform.

	Reflections	MRR noise
2D-SWT	0.1423	0.0801
Shearlet transform	0.0639	0.1987

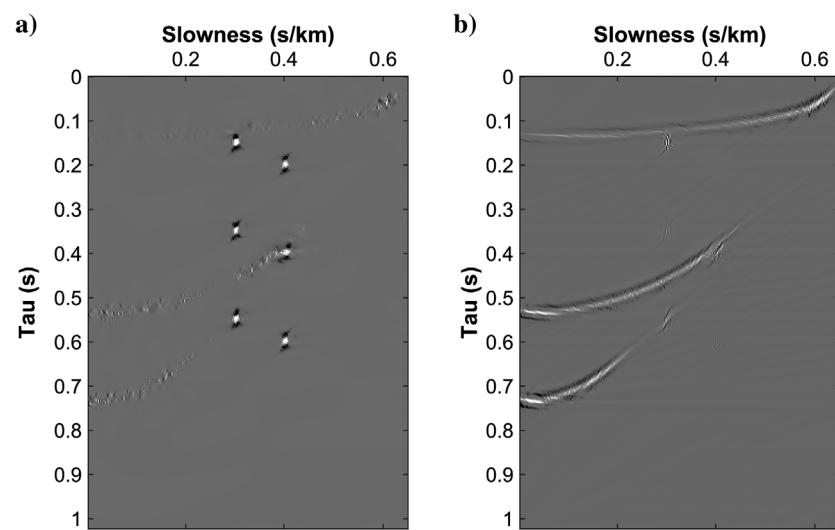


Figure 5. MCA separation results. (a) Separated pointwise component and (b) separated curved component.

occur. Furthermore, we introduce band-limited random noise with a frequency range of 10–60 Hz. The noise standard deviation is 97.50. The geologic model for generating the synthetic data is a two-layer model with different velocities and is displayed in Figure 9. The low-velocity layer has a velocity of 800 m/s, and the high-velocity layer is 2500 m/s. The density of the two layers is calculated with Gardner's relation. To simplify the example, the seismic quality factor is set to infinity. In addition, we suppose that the surface of the ground is a free surface, and we set the central frequency of the Ricker wavelet to 30 Hz and the maximum offset to 2000 m. The synthetic data have 401 traces, and the trace interval is 10 m. The boundary between the layers is at a depth of 80 m, and the thickness of the high-velocity layer is 300 m. Then, a source is placed at a depth of 0 m and the midpoint in the horizontal direction. The number of sampling points is 1001 with a sampling interval of 1 ms, so the total seismic recording length is 1 s.

Next, we perform a high-resolution linear Radon transformation on the synthetic data with a slowness range of [0, 0.6] s/km. As shown in Figure 10a, the point-like information in the red ellipse area is the main energy corresponding to the linear interference, and the other curve-like information contains the main energy of the reflected waves. It is clear that their distributions are overlapping. Therefore, a direct mute in the Radon domain is likely to cause the loss of effective wave energy, negatively affecting subsequent processing. To better suppress MRR noise, the proposed method is applied to Figure 10a and the separated results of the point-like

energy and curve-like energy are presented in Figure 10b and 10c, respectively. Following the proposed procedure shown in Figure 2, the linear noise can be reconstructed from the separated point component using the inverse Radon transform, as shown in Figure 11a. It is evident that there are no significant structures of desired reflection signals. That is, the proposed method hardly hurts the useful signals. Figure 11a shows the results after subtraction. In comparison with the original data, the proposed method has suppressed most of the linear noise. For comparison, the f - k filtering and Radon filtering methods also are applied to suppress the MRR noise, and the removed noise is displayed in Figure 11c and 11e, respectively. Although the latter two methods also can attenuate linear noise, they cause more damage. The f - k filtering introduces obvious artifacts, as shown in the red ellipse area in Figure 11d. Moreover, when the reflected wave is close to the apparent velocity of the linear noise at the far offset, it causes some loss of reflected wave energy. A detailed analysis of the areas indicated by red arrows in Figure 11b and 11d proves this. Similarly, the slowness of MRR noise and the effective reflections are close at far offsets, which leads to an overlap in the Radon panel. From the areas indicated by the white arrows in Figure 11f, we observe that a direct mute results in the loss of effective wave energy. In addition, as indicated by the red arrows, some residual MRR noise remains, whereas our approach can significantly reduce it. In conclusion, compared with the traditional f - k filtering method and Radon filtering method, the proposed method has higher fidelity for the effective signal and is more effective in suppressing strong linear noise.

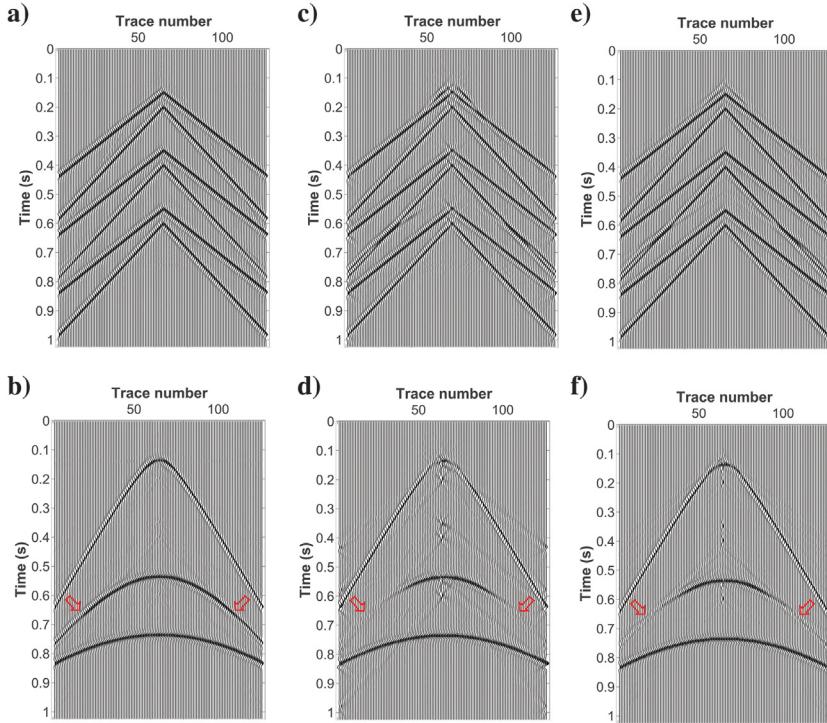


Figure 6. Comparison of the signal-to-noise separation results with the synthetic shot gather. (a) Removed MRR by the proposed method, (b) separated reflections by the proposed method, (c) removed MRR by the f - k filtering method, (d) separated reflections by the f - k filtering method, (e) removed MRR by the Radon filtering method, and (f) separated reflections by the Radon filtering method. The red arrows show that the f - k filtering and Radon filtering methods damage the reflections, whereas our method preserves the reflections well.

Field data example

Figure 12a shows a 2D common-shot gather acquired in Eastern China. This land seismic record has a total of 255 traces with a sampling interval of 20 m. Each trace has 3501 time sampling points, and the sampling interval is 1 ms. Obviously, MRR noise, marked by the red ellipses, severely masks the shallow primary reflection events and negatively impacts subsequent interpretation. Moreover, the direct wave marked by the blue line and ground roll marked by the yellow cone also interfere with the reflections. Figure 12b displays the Radon panel of this seismic record using a slowness range of $[-0.5, 2.5]$ s/km. The reflections of interest marked by the black box are overlapped by the MRR in the red ellipse area, which makes the direct mute invalid. The challenge also comes from the ground roll, which is mainly concentrated in the yellow ellipse area because it has a morphological feature similar to MRR in the Radon panel and may hinder the successful MCA suppression. For a better removal of MRR, we set the appropriate slowness range of $[0.25, 2]$ s/km to avoid obstacles from ground roll, as displayed in Figure 13a. From our separated results shown in Figure 13b and 13c, we find that the significant different features are well separated, which proves the reasonability of our selected dictionaries.

Following this, the inverse Radon transform is applied to obtain Figure 14a and the MRR noise is separated by subtraction in Figure 14b. By comparing the results with raw data, it can be found that the linear noise is effectively attenuated by the proposed method, and the reflection events which were significantly covered by the linear noise are clearly displayed at present. Further, f - k filtering and Radon filtering are applied to the common-shot gather for comparison. Figure 14c and 14d shows the separated MRR noise and reflections using the f - k filtering, respectively. Figure 14e and 14f displays the corresponding results of the Radon filtering method. Comparing Figure 14b, 14d, and 14f, it can be concluded that the proposed method and the two comparison methods can suppress most of the MRR noise, and the most effective events covered by the strong MRR have been recovered. However, our method performs better than the other two methods because our denoised reflection events are clearer and more continuous, especially in the shallow layer. This also can be seen inside the red boxes within the time range from 0.5 to 2.2 s, where almost no MRR noise

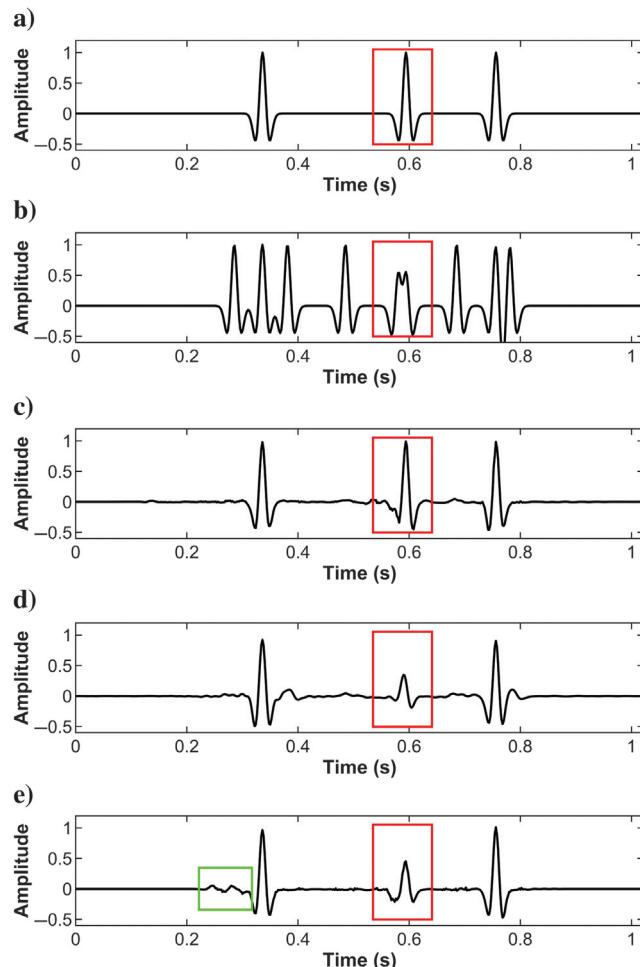


Figure 7. Comparison of the signal-to-noise separation results with the 34th trace in the synthetic shot gather. (a) Synthetic data containing three reflections and (b) synthetic data superimposed with linear noise. Separated reflections by (c) the proposed method, (d) the f - k filtering method, and (e) the Radon filtering method. The red boxes reveal that the proposed method has high fidelity. The green box indicates that the Radon filtering distorts the waveform of the first arrival.

residuals are observed. For the f - k filtering method, MRR residuals at far offset are obvious, as marked by the yellow arrows in Figure 14d. In addition, it produces severe waveform distortions

Table 2. S/N comparison of the three methods.

	S/N (dB)
Original noisy data	-3.08
The proposed method	12.73
f - k filtering method	5.07
Radon filtering method	6.02

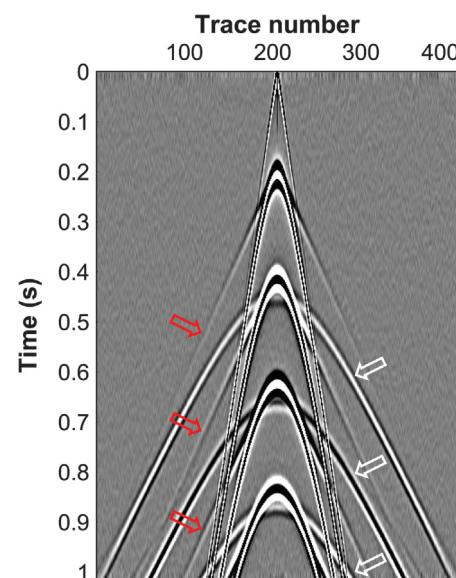


Figure 8. Synthetic data based on wave equation forward modeling. The red arrows indicate the MRRs, and the white arrows indicate the reflections.

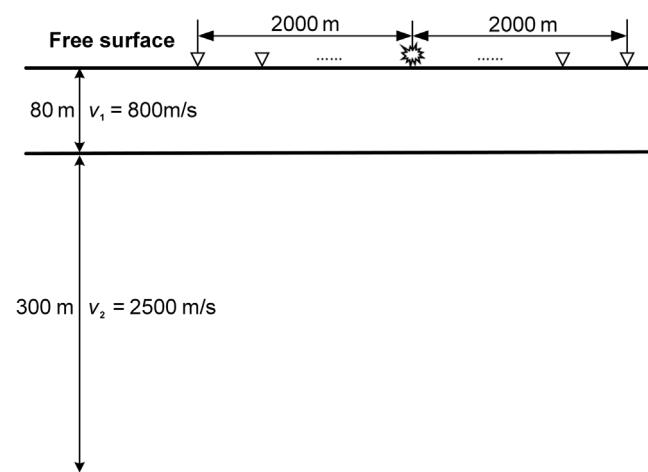


Figure 9. A model used for forward modeling, which has two layers with different velocities.

at the shallow near-offset and the far-offset zone, as indicated by the red ellipse in Figure 14d and the red arrows in Figure 14c, respectively. For the Radon filtering method, obvious MRR residuals can be found near the yellow arrows and in the area indicated by the red ellipse in Figure 14f. Furthermore, we also plot the f - k spectrum to evaluate the proposed method. Figure 15a displays the spectrum of the original data, and it can be seen that the energy of MRR noise is mainly located in the black ellipse region. In addition, spatial aliasing at high frequencies is observed in the yellow ellipse. The MRR removal results of our method, the f - k filtering method, and the Radon filtering method are shown in Figure 15c, 15e, and 15g, respectively. It can be seen that the three methods can effectively attenuate the MRR interference because the corresponding black region is removed. We notice that the cutoff region of the f - k filtering is carefully chosen to achieve a compromise between MRR suppression and reflection preservation. In addition, the cutoff edge is smoothed to avoid introducing oscillations. Even if the filter is carefully designed, there are still obvious truncations around the red arrows shown in Figure 15e, which is the cause of the artifacts shown in Figure 14d. From the red arrows in Figure 15c and 15e, it can be observed that the denoised results of the f - k filtering are inferior to those of our method because our spectrum of denoised results is more continuous. The spectra of separated MRR noise and denoised reflections by the Radon filtering method are shown in Figure 15f and 15g, respectively. The obvious energy of the MRR interference still remains, as marked by the red arrows in Figure 15g. This phenomenon is consistent with the problem of incomplete MRR suppression shown in Figure 14f. In addition, it is worth mentioning that only our method eliminates spatial aliasing, which is a significant advantage over the other two methods. In summary, the proposed method outperforms the traditional f - k filtering and Radon filtering methods because our method has higher fidelity to effective signals and can suppress more MRR noise.

Figure 10. Separation results in the Radon panel. (a) Original data contaminated with linear noise. The red ellipse indicates the distribution of linear noise in the Radon panel. (b) Separated point-like information by our method. (c) Separated curve-like information by our method.

DISCUSSION

Different from a direct application of MCA in the temporal-space domain (Wang et al., 2012; Chen et al., 2017), we use MCA to separate the MRR noise in the linear Radon domain that magnifies the difference between the signals and noise, allowing them to be separated more easily. It also is possible to attenuate other coherent noise in a similar manner by using other transformations.

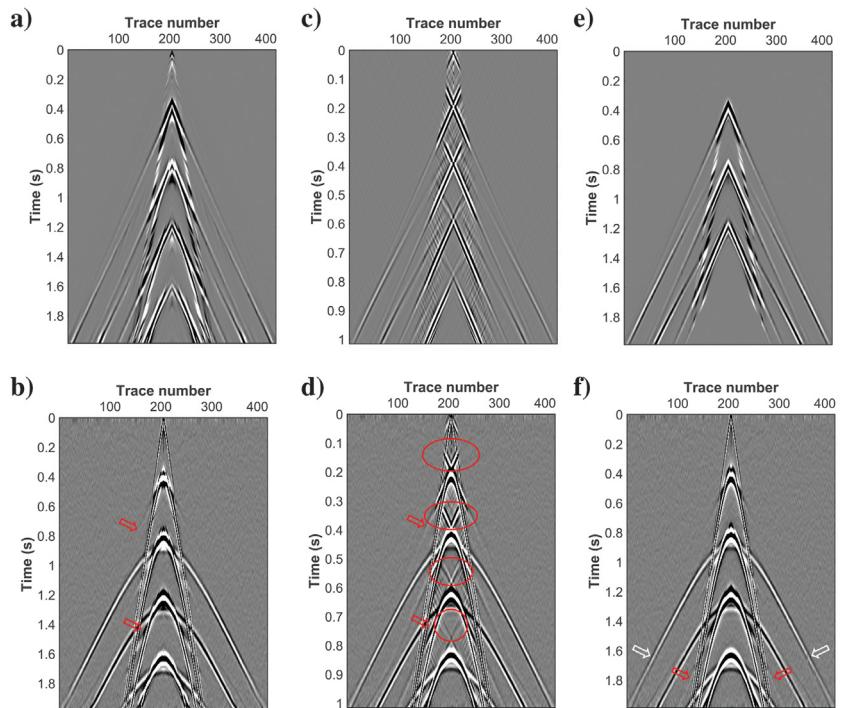
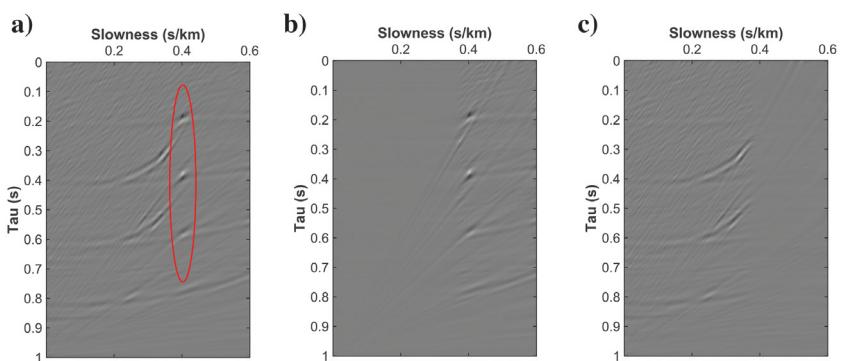


Figure 11. Comparison of the signal-to-noise separation results with the synthetic shot gather generated by the forward modeling. (a) Removed MRR by the proposed method, (b) separated reflections by the proposed method, (c) removed MRR by the f - k filtering method, and (d) separated reflections by the f - k filtering method. The red ellipses indicate that the f - k filtering introduces obvious artifacts. From the areas indicated by the red arrows in (b and d), we observe that the f - k filtering method also causes some loss of reflected wave energy. (e) Removed MRR by the Radon filtering method. (f) Separated reflections by the Radon filtering method. The white arrows show that the Radon filtering method results in the loss of effective wave energy. According to the red arrows, there is some residual MRR noise, which is effectively removed by our method.



A comparable technique is proposed by [Perkins and Zwaan \(2000\)](#), which aims to attenuate ground roll. They first construct the signal space \mathbf{A}_s and the ground-roll space \mathbf{A}_g . Then, they solve the following equation in a least-squares sense:

$$\min_{\mathbf{w}_s, \mathbf{w}_g} \|\mathbf{d} - \mathbf{A}_s \mathbf{w}_s - \mathbf{A}_g \mathbf{w}_g\|, \quad (30)$$

where \mathbf{w}_s and \mathbf{w}_g are vectors comprising coefficients representing the signal and the ground roll, respectively. Similarly, our proposed method can be rewritten in the following form:

$$\begin{aligned} \min_{\{\mathbf{x}_r, \mathbf{x}_l\}} & \|\mathbf{d} - \mathbf{LD}_r \mathbf{x}_r - \mathbf{LD}_l \mathbf{x}_l\|_2^2 + \lambda (\|\mathbf{x}_r\|_1 + \|\mathbf{x}_l\|_1) \\ \text{s.t. } & \mathbf{D}_r \mathbf{x}_r + \mathbf{D}_l \mathbf{x}_l = \mathbf{m}, \|\mathbf{m}\|_1 \leq \delta, \end{aligned} \quad (31)$$

where $\delta > 0$ denotes a parameter that governs the sparsity of the solution. If we regard \mathbf{LD}_r as the reflection space \mathbf{A}_r , and regard \mathbf{LD}_l as the MRR space \mathbf{A}_l , we find that equation 31 has the same

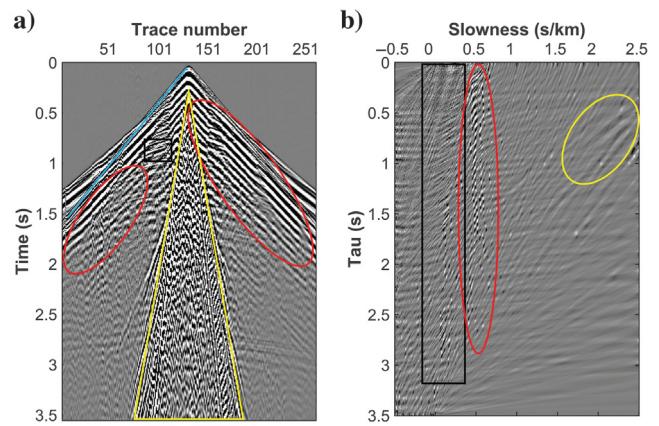
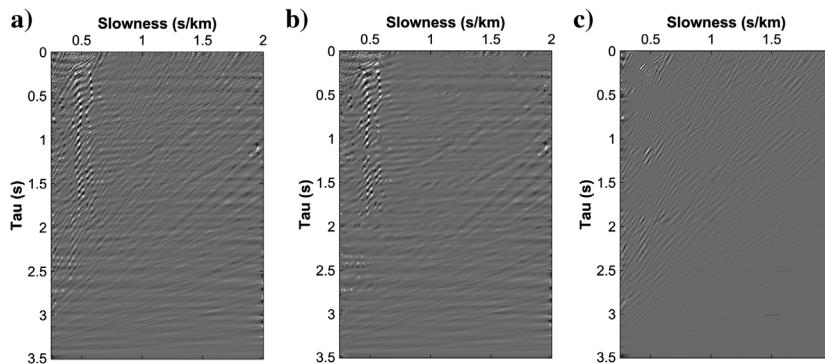


Figure 12. A raw common-shot gather of field seismic data. (a) Original field seismic data and (b) the high-resolution linear Radon transform. The red, yellow, black, and blue colors represent the MRR noise, ground roll, effective reflections, and direct waves, respectively.

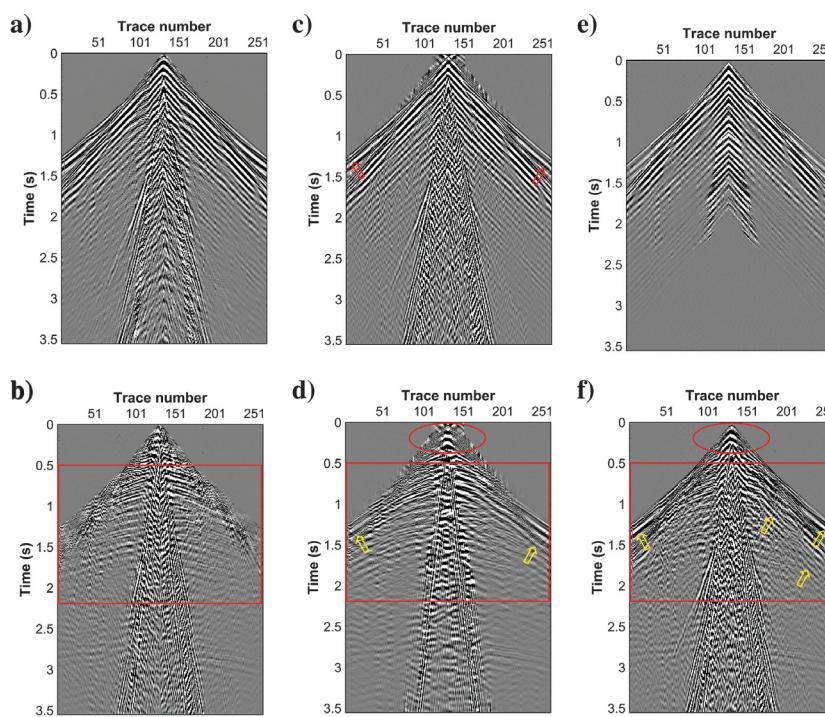


Figure 13. MCA separation results in the Radon panel. (a) The Radon panel of the original field seismic data using a suitable slowness range, (b) the separated linear noise, and (c) the separated reflections of interest.

Figure 14. Comparison of the signal-to-noise separation results with the field shot gather. (a) Removed MRR by the proposed method, (b) separated reflections by the proposed method, and (c) removed MRR by the f - k filtering method. The red arrows indicate severe waveform distortions. (d) Separated reflections by the f - k filtering method. The red ellipse indicates severe waveform distortions, and the yellow arrows show significant MRR noise residuals. (e) Removed MRR by the Radon filtering method. (f) Separated reflections by the Radon filtering method. The red ellipse and four yellow arrows indicate obvious MRR residuals. By comparing the areas indicated by red boxes in (b, d, and f), we observe that the proposed method gives the best results because the denoised reflection events are the clearest and most continuous.

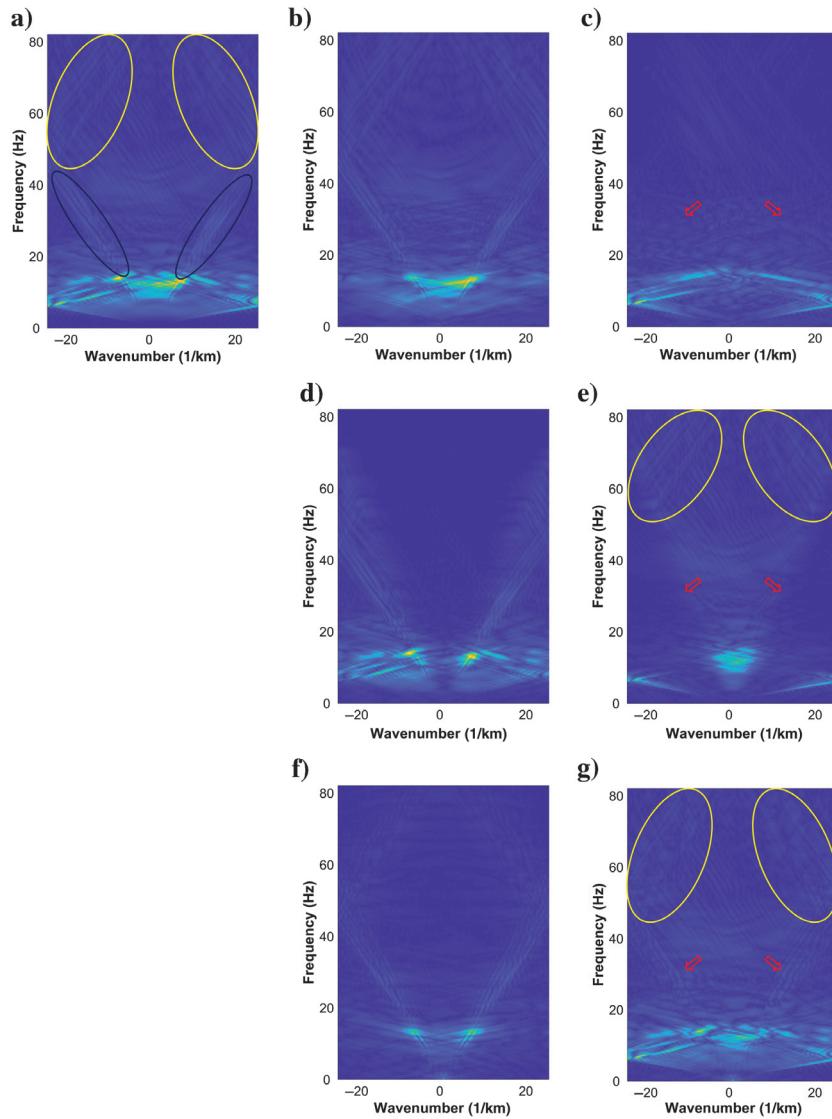


Figure 15. Denoising comparison of f - k spectra. (a) Original data, (b) removed MRR by the proposed method, (c) separated reflections by the proposed method, (d) removed MRR by the f - k filtering method, (e) separated reflections by the f - k filtering method, (f) removed MRR by the Radon filtering method, and (g) separated reflections by the Radon filtering method. By comparing the red arrows in (c) and (e), we observe that our spectrum is more natural and continuous. The red arrows in (g) indicate that some MRR noise still remains after applying the Radon filtering method. The yellow ellipses mark the spatial ellipse.

form as equation 30 except for the additional regularization terms. Indeed, we concatenated two different transforms to produce two relatively complex spaces, making the signals and noise easily distinguishable. Equation 30 was successfully implemented to suppress ground roll. Accordingly, we are confident that our proposed method can be applied to attenuate linear ground roll with a few modifications. In addition, we notice that the 3D Radon transform has recently attracted the attention of the industry (Perrone et al., 2019). We will apply the proposed algorithm to 3D linear noise suppression, such as cross-spread 3D prestack field data. In this case, 3D multiscale transforms will be required, and the amount of calculations involved might pose a challenge.

CONCLUSION

We propose an MRR attenuation method based on the differences of waveform morphological features in the τ - p domain. Using the high-resolution linear Radon transform, the MRR noise and effective reflections can be transformed into point features and curved features, respectively. Then, they can be separated by MCA theory with appropriate overcomplete dictionaries. The 2D-SWT is suitable for analyzing signals with discontinuities or sharp spikes, thereby serving as a dictionary to extract point components. The shearlet transform, as a multiscale and multidirectional transformation, is adopted as a dictionary to extract curved components. Finally, after completing iterations of the block coordinate relaxation method, we apply the inverse Radon transform to the separated MRRs and subtract them from the original seismic data. Subtraction produces amplitude-preserved results. Synthetic data and field data demonstrate that our method preserves reflections better than the f - k filtering method and the Radon filtering method, without introducing artifacts. Future research will involve extending the proposed method to higher dimensions and attenuating other types of linear noise, such as linear ground roll.

ACKNOWLEDGMENTS

This research is sponsored by the National Natural Science Foundation of China (41774135 and 41974131). Many thanks go to the assistant editor A. Malcolm, the associate editor A. Stanton, and other anonymous reviewers for their insightful suggestions.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

REFERENCES

- Al-Husseini, M. I., J. B. Glover, and B. J. Barley, 1981, Dispersion patterns of the ground roll in eastern Saudi Arabia: *Geophysics*, **46**, 121–137, doi: [10.1190/1.1441183](https://doi.org/10.1190/1.1441183).
- Al-Shaykh, O. K., and J. F. Doherty, 1996, Invariant image analysis based on Radon transform and SVD: *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, **43**, 123–133, doi: [10.1109/82.486459](https://doi.org/10.1109/82.486459).
- Chen, S. S., D. L. Donoho, and M. A. Saunders, 2001, Atomic decomposition by basis pursuit: *SIAM Review*, **43**, 129–159, doi: [10.1137/S003614450037906X](https://doi.org/10.1137/S003614450037906X).
- Chen, X., W. Chen, X. Wang, and W. Wang, 2017, Sparsity-optimized separation of body waves and ground-roll by constructing dictionaries using tunable Q-factor wavelet transforms with different Q-factors: *Geophysical Journal International*, **211**, 621–636, doi: [10.1093/gji/ggx332](https://doi.org/10.1093/gji/ggx332).
- Chen, Y., S. Jiao, J. Ma, H. Chen, Y. Zhou, and S. Gan, 2015, Ground-roll noise attenuation using a simple and effective approach based on local

- band-limited orthogonalization: IEEE Geoscience and Remote Sensing Letters, **12**, 2316–2320, doi: [10.1109/LGRS.2015.2475280](https://doi.org/10.1109/LGRS.2015.2475280).
- Cvetkovic, M., S. Falconer, K. J. Marfurt, and S. Chávez-Pérez, 2007, 2D stationary-wavelet transform-based acquisition footprint suppression: 77th Annual International Meeting, SEG, Expanded Abstracts, 2590–2594, doi: [10.1190/1.2793005](https://doi.org/10.1190/1.2793005).
- De Lauro, E., S. De Martino, M. Falanga, and M. Palo, 2009, Decomposition of high-frequency seismic wavefield of the Strombolian-like explosions at Erebus volcano by independent component analysis: Geophysical Journal International, **177**, 1399–1406, doi: [10.1111/j.1365-246X.2009.04157.x](https://doi.org/10.1111/j.1365-246X.2009.04157.x).
- Donoho, D. L., and J. M. Johnstone, 1994, Ideal spatial adaptation by wavelet shrinkage: Biometrika, **81**, 425–455, doi: [10.1093/biomet/81.3.425](https://doi.org/10.1093/biomet/81.3.425).
- Durrani, T. S., and D. Bisset, 1984, The Radon transform and its properties: Geophysics, **49**, 1180–1187, doi: [10.1190/1.1441747](https://doi.org/10.1190/1.1441747).
- Elad, M., J.-L. Starck, P. Querre, and D. Donoho, 2005, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA): Applied and Computational Harmonic Analysis, **19**, 340–358, doi: [10.1016/j.acha.2005.03.005](https://doi.org/10.1016/j.acha.2005.03.005).
- Guo, K., and D. Labate, 2007, Optimally sparse multidimensional representation using shearlets: SIAM Journal on Mathematical Analysis, **39**, 298–318, doi: [10.1137/060649781](https://doi.org/10.1137/060649781).
- Ibrahim, A., and M. D. Sacchi, 2014, Simultaneous source separation using a robust Radon transform: Geophysics, **79**, no. 1, V1–V11, doi: [10.1190/geo2013-0168.1](https://doi.org/10.1190/geo2013-0168.1).
- Kabir, M. N., and K. J. Marfurt, 1999, Toward true amplitude multiple removal: The Leading Edge, **18**, 66–73, doi: [10.1190/1.1438158](https://doi.org/10.1190/1.1438158).
- Kazemi, N., and M. D. Sacchi, 2021, Offset-extended sparse Radon transform: Application to multiple suppression in the presence of amplitude variations with offset: Geophysics, **86**, no. 3, R293–R302, doi: [10.1190/geo2020-0307.1](https://doi.org/10.1190/geo2020-0307.1).
- Labate, D., W.-Q. Lim, G. Kutyniok, and G. Weiss, 2005, Sparse multidimensional representation using shearlets: Wavelets XI, International Society for Optics and Photonics, 59140U.
- Lanzavecchia, S., P. L. Bellon, and M. Radermacher, 1999, Fast and accurate three-dimensional reconstruction from projections with random orientations via Radon transforms: Journal of Structural Biology, **128**, 152–164, doi: [10.1006/jsb.1999.4185](https://doi.org/10.1006/jsb.1999.4185).
- Li, Q., 1983, On seismic secondary noises: Oil Geophysical Prospecting, **18**, 207–225.
- Liang, C., 1993, Analyzing and eliminating multiple reflected refractions: Geophysical Prospecting for Petroleum, **32**, 112–118.
- Liu, K.-H., and W. H. Dragoset, 2013, Blind-source separation of seismic signals based on information maximization: Geophysics, **78**, no. 4, V119–V130, doi: [10.1190/geo2012-0136.1](https://doi.org/10.1190/geo2012-0136.1).
- Lu, W., 2013, An accelerated sparse time-invariant Radon transform in the mixed frequency-time domain based on iterative 2D model shrinkage: Geophysics, **78**, no. 4, V147–V155, doi: [10.1190/geo2012-0439.1](https://doi.org/10.1190/geo2012-0439.1).
- Ma, J., G. Plonka, and H. Chauris, 2010, A new sparse representation of seismic data using adaptive easy-path wavelet transform: IEEE Geoscience and Remote Sensing Letters, **7**, 540–544, doi: [10.1109/LGRS.2010.2041185](https://doi.org/10.1109/LGRS.2010.2041185).
- Meunier, J., T. Bianchi, J. Postel, and R. Taylor, 2008, The future of vibroseis for high-density wide-azimuth land acquisition: First Break, **26**, 87–91, doi: [10.3997/1365-2397.26.1116.27860](https://doi.org/10.3997/1365-2397.26.1116.27860).
- Perkins, C., and M. Zwaan, 2000, Ground roll attenuation: 62nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, cp-28-00172, doi: [10.3997/2214-4609-pdb.28.L21](https://doi.org/10.3997/2214-4609-pdb.28.L21).
- Perrone, F., W. S. Ross, and S. Grion, 2019, A comparison of greedy algorithms for 3D Radon transforms: Matching pursuit vs orthogonal matching pursuit: 89th Annual International Meeting, SEG, Expanded Abstracts, 4510–4514, doi: [10.1190/segam2019-3216108.1](https://doi.org/10.1190/segam2019-3216108.1).
- Qin, Z., 1990, Analysis of seismic noises in the DQ area: Oil Geophysical Prospecting (China), **25**, 86–102.
- Radon, J., 1986, On the determination of functions from their integral values along certain manifolds: IEEE Transactions on Medical Imaging, **5**, 170–176, doi: [10.1109/TMI.1986.4307775](https://doi.org/10.1109/TMI.1986.4307775).
- Rubinstein, R., A. M. Bruckstein, and M. Elad, 2010, Dictionaries for sparse representation modeling: Proceedings of the IEEE, **98**, 1045–1057, doi: [10.1109/JPROC.2010.2040551](https://doi.org/10.1109/JPROC.2010.2040551).
- Sacchi, M. D., and T. J. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169–1177, doi: [10.1190/1.1443845](https://doi.org/10.1190/1.1443845).
- Sardy, S., A. G. Bruce, and P. Tseng, 2000, Block coordinate relaxation methods for nonparametric wavelet denoising: Journal of Computational and Graphical Statistics, **9**, 361–379, doi: [10.1080/10618600.2000.10474885](https://doi.org/10.1080/10618600.2000.10474885).
- Scales, J. A., 1987, Tomographic inversion via the conjugate gradient method: Geophysics, **52**, 179–185, doi: [10.1190/1.1442293](https://doi.org/10.1190/1.1442293).
- Shensa, M. J., 1992, The discrete wavelet transform: Wedding the a trous and Mallat algorithms: IEEE Transactions on Signal Processing, **40**, 2464–2482, doi: [10.1109/78.157290](https://doi.org/10.1109/78.157290).
- Starck, J., M. Elad, and D. L. Donoho, 2005, Image decomposition via the combination of sparse representations and a variational approach: IEEE Transactions on Image Processing, **14**, 1570–1582, doi: [10.1109/TIP.2005.852206](https://doi.org/10.1109/TIP.2005.852206).
- Steeple, D. W., 2005, Spectral shaping during acquisition of seismic reflection data: 75th Annual International Meeting, SEG, Expanded Abstracts, 917–921, doi: [10.1190/1.1890385](https://doi.org/10.1190/1.1890385).
- Thorson, J. R., and J. F. Claerbout, 1985, Velocity-stack and slant-stack stochastic inversion: Geophysics, **50**, 2727–2741, doi: [10.1190/1.1441893](https://doi.org/10.1190/1.1441893).
- Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386–399, doi: [10.1190/1.1543224](https://doi.org/10.1190/1.1543224).
- Vermeer, G. J., 1998, 3-D symmetric sampling: Geophysics, **63**, 1629–1647, doi: [10.1190/1.1444459](https://doi.org/10.1190/1.1444459).
- Vermeer, G. J. O., 2008, Alternative strategies for tackling scattered noise: 78th Annual International Meeting, SEG, Expanded Abstracts, 95–99, doi: [10.1190/1.3054863](https://doi.org/10.1190/1.3054863).
- Wang, W., J. Gao, W. Chen, and J. Xu, 2012, Data adaptive ground-roll attenuation via sparsity promotion: Journal of Applied Geophysics, **83**, 19–28, doi: [10.1016/j.jappgeo.2012.04.004](https://doi.org/10.1016/j.jappgeo.2012.04.004).
- Xue, Y., J. Ma, and X. Chen, 2014, High-order sparse Radon transform for AVO-preserving data reconstruction: Geophysics, **79**, no. 2, V13–V22, doi: [10.1190/geo2013-0002.1](https://doi.org/10.1190/geo2013-0002.1).
- Yu, Z., J. Ferguson, G. McMechan, and P. Anno, 2007, Wavelet-Radon domain dealiasing and interpolation of seismic data: Geophysics, **72**, no. 2, V41–V49, doi: [10.1190/1.2422797](https://doi.org/10.1190/1.2422797).

Biographies and photographs of the authors are not available.