Robust reconstruction via Group Sparsity with Radon operators

Ji Li, Daniel Trad

ABSTRACT

Sparse solutions of linear systems play an important role in seismic data processing, including denoising and interpolation. An additional structure called group sparsity can be used to improve the performance of the sparse inversion. We propose a robust group sparse inversion algorithm based on Orthogonal Matching Pursuit with the Radon operators in the frequency slowness $(w - p)$ domain. The proposed algorithm is used to interpolate seismic data and attenuate erratic noise simultaneously. During each iteration, The proposed algorithm first picks the dominant slowness group. Then, all the Radon coefficients located within the currently selected slowness groups are fitting to the data in the time space $(t-x)$ domain via a robust solver, which is a $\ell_1 - \ell_1$ ADMM solver. In other words, we adopt a cost function that directly utilizes the selected coefficients to fit synthesized signals via $\ell_1 - \ell_1$ norm in the time-space domain. We prove that the proposed algorithm is resistant to erratic noise, making it attractive to applications such as simultaneous source deblending and reconstruction of noisy onshore datasets. We compare the performance of the same method with and without the group sparsity constraint and also with other Radon-based inversion methods. The result shows that the strong group sparsity inversion performs better than the traditional sparsity inversion. Both synthetic and real seismic data are being tested to examine the performance of the proposed algorithm.

INTRODUCTION

Sparse representation is an important tool for signal processing, including seismic data processing. These applications include primaries estimation (Van Groenestijn and Verschuur, 2009; Lin and Herrmann, 2013),multiple attenuation (Sacchi and Ulrych, 1995; Herrmann et al., 2000), data reconstruction (Sacchi et al., 1998; Herrmann, 2010; Trad, 2009), and many more. Sparse representation means the seismic signal is sparse (the number of nonzero coefficients k is much smaller than the total number of coefficients) in some transform domains like Radon, Fourier and Wavelet.

On the other hand, an additional group structure can be added to the problem, often referred to as group sparse or group Lasso (Bach, 2008; Yuan and Lin, 2006; Nardi and Rinaldo, 2008) to improve the performance of the sparse estimation. Many methods have been proposed in seismic data processing with the idea of group sparsity. Li and Sacchi (2022) and Naghizadeh (2012) divide the coefficients into different dips in the frequency wavenumber $(f-k)$ domains for seismic data denoising and interpolation. Vera Rodriguez et al. (2012) use group sparsity for microseismic data denoising. Trad et al. (2002) use a similar idea for fast calculation of the hyperbolic Radon transform. Chen et al. (2019) combines the group sparsity and total variation for seismic signal denoising.

One interesting goal is to develop a robust sparse solver to solve the inversion problem for seismic data corrupted by erratic noise. Many researchers have been studying this problem for a long time. For instance, Guitton and Symes (2003) replaces the ℓ_2 norm with $Huber$ norm for the residual term to cope with the seismic data with outliers. Trickett et al. (2012) uses a rank reduction filter to remove the erratic noise in the seismic data. More recently, Li and Sacchi (2021) developed a sparse and robust Radon transform, estimated via Matching Pursuit, to solve the simultaneous source separation problem. Most of these methods use robust M-estimators (Maronna, 1976) to replace the ℓ_2 norm to make the cost function robust to the erratic noise . In this paper, we cooperate the robust inversion with group sparsity to get a better sparse estimation.

Our contribution can be summarized as follows. We first develop a robust algorithm combining robust inversion and group sparsity. The algorithm works on the linear Radon transform in the frequency-slowness ($f-p$) domain. We use the orthogonal Matching Pursuit (Tropp and Gilbert, 2007) to build the estimated results iteratively. We divide the Radon coefficients into groups with different slowness p in the $f-p$ domain. In each iteration, OMP picks the group $\mathbf{m}_p$ with the maximum norm. Then we use an $\ell_1-\ell_1$ ADMM solver (Wen et al., 2016) to fit the Radon coefficients within all currently selected groups. The OMP promotes the group sparsity by picking a limited number of best-correlated groups, and the ADMM solver promotes the sparsity within the selected groups. Therefore, the proposed method simultaneously promoted sparsity in both groups and coefficients within the groups. Thus solve the inversion problem with strong sparse constraint. Both 2D and 3D synthetic examples are illustrated to test the robust reconstruction performance of the algorithm. We also apply the proposed method to the cube of 3D real data example and compare the performance with the traditional POCS (Abma and Kabir, 2006) algorithm. The synthetic and real data examples demonstrate the effectiveness and robustness of the proposed robust reconstruction method.

THEORY

A signal can be represented as follows:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{1}$$

Where $\mathbf{y}$ denotes the signal, $\mathbf{x}$ is the coefficients, and $\mathbf{A}$ is a synthesis operator or matrix which can transfer the coefficients into the seismic signal. $\mathbf{e}$ represents additive noise in the signal. In this situation, the sparse approximation problem can be expressed as

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \leq \delta, \tag{2}$$

where $\|\mathbf{x}\|_0$ is the ℓ_0 norm, which measure the number of nonzero coefficients of $\mathbf{x}$. and $\|.\|_2^2$ symbolizes the ℓ_2 norm (equal to the square root of the the inner product of a vector with itself) used to fit the residuals. This problem is a combinatorial, nondeterministic polynomial time (NP)-hard problem for which finding an exact solution is prohibitively expensive. Generally, two groups of methods can be used to solve the problem 2. One is the convex relaxation, replacing the ℓ_0 norm with the ℓ_1 norm Chen et al. (2001). Which transfers the problem into a convex optimization problem.

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1. \tag{3}$$

This problem is also known as Lasso (Tibshirani, 1996), and can be solved by many methods like FISTA (Beck and Teboulle, 2009), IRLS (Scales and Gersztenkorn, 1988) and ADMM (Boyd et al., 2011). Another approach is the greedy method like Matching Pursuit (MP) (Mallat and Zhang, 1993) and Orthogonal Matching Pursuit (OMP) (Tropp and Gilbert, 2007; Pati et al., 1993).

In some cases, adding a group structure constraint can yield a better estimation (Majumdar and Ward, 2009; Elhamifar and Vidal, 2011). Unlike the standard sparsity assumption, the sparsity is measured by the number of nonzero coefficients k. A group-sparse vector can be divided into groups so that few groups contain nonzero coefficients, but the groups do not need to be sparse.

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \beta \sum_{i=1}^{N} \|\mathbf{x}[i]\|_2. \tag{4}$$

A further refinement, called strong group sparsity, can be made to improve performance (Vincent and Hansen, 2014; Simon et al., 2013; Huang and Zhang, 2010). For the strong sparsity problem, the coefficients are not only located within a few groups, but the nonzero groups are also sparse.

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{x}\|_1 + \beta \sum_{i=1}^{N} \|\mathbf{x}[i]\|_2. \tag{5}$$

When the erratic noise corrupts the data, we can use M-estimators like $Cauchy$ norm, $Huber$ norm or ℓ_1 norm to replace the ℓ_2 norm for the data fitting term. In our case, we use the ℓ_1 norm to replace the ℓ_2 norm. Then the problem changes to

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_1^1 + \lambda\|\mathbf{x}\|_1 + \beta \sum_{i=1}^{N} \|\mathbf{x}[i]\|_2. \tag{6}$$

To solve the problem 6, we use the linear Radon transform to synthesize the seismic signal. And the orthogonal matching pursuit with the $\ell_1 - \ell_1$ ADMM to estimate the sparse Radon coefficients

Linear Radon transform

Linear Radon transform is time-invariant, which means it can be calculated either in time or frequency domain. For the Radon transform in frequency domain. It can be represented as the following equation,

$$\mathbf{M}(p, w) = \sum_{h} \mathbf{D}(x, w) e^{iwxp}, \tag{7}$$

which can also be written in the matrix form

$$\mathbf{M} = \mathcal{L} * \mathbf{D}. \tag{8}$$

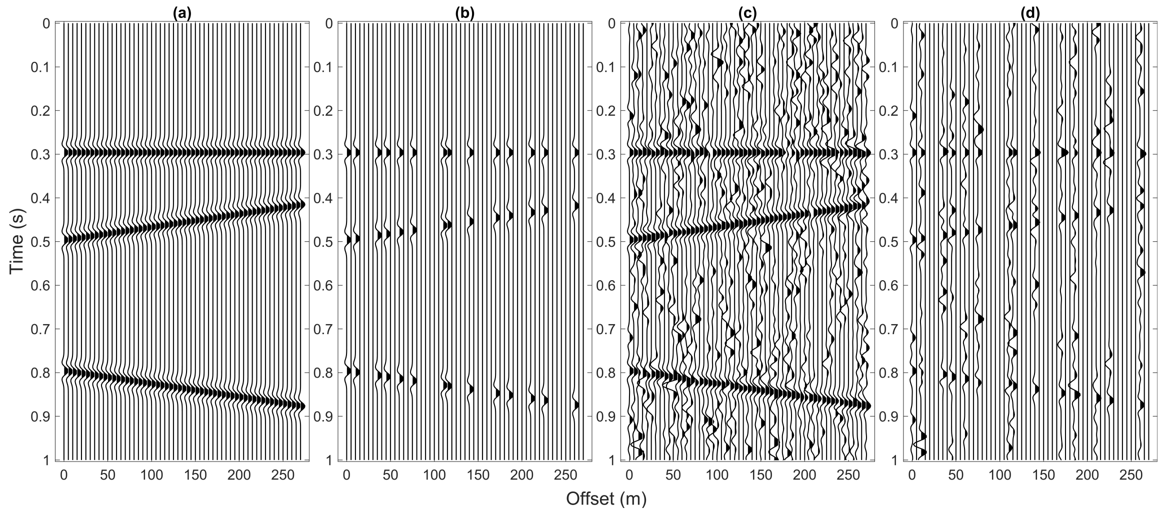

FIG. 1. (a) 2D synthetic data. (b) 2D synthetic data with 75% of missed traces. (c) 2D synthetic data with erratic noise. (d) 2D synthetic with erratic noise and 75% of missed traces.

This equation can extend to 3D form easily,

$$\mathbf{M}(p_x, p_y, w) = \sum_h \mathbf{D}(x, y, w) e^{iw(p_x x + p_y y)}, \tag{9}$$

which in the matrix form can be written as

$$\mathbf{M} = \mathcal{L}_{p_x} * \mathbf{D} * \mathcal{L}_{p_y}. \tag{10}$$

Figure 1 shows a simple synthetic 2D example with three linear events. Part (a) is the fully sampled clean data. Part (b) is the clean data with 75% of the traces missed. (c) is the fully-sampled data with strong erratic noise. And (d) is the data with erratic noise and missed traces. Figure 2 is the corresponding Radon coefficients in the $f-p$ domain. As we can see from this figure, since the synthetic three linear events have different velocities, all the Radon coefficients are located around three p values in the $f-p$ domain. The missing trace and erratic noise will introduce the random noise in the $f-p$ domain. We can see that the dominant slowness in the $f-p$ domain becomes unclear for the strong erratic noise with missed traces. Figure 3 shows the ℓ_2 norm of each slowness p group. From this figure, we find that even with the strong erratic noise and a lot of missed traces, we can still determine the three dominant slowness groups easily.

By using the linear Radon operators in the $f-p$ domain, the problem becomes to minimize the following cost function.

$$\text{argmin}\, \|\mathcal{L}^* \mathbf{m} - \mathbf{d}\|_1 + \lambda \|\mathbf{m}\|_1 + \beta \sum_{i=1}^{N} \|\mathbf{m}[i]\|_2. \tag{11}$$

where $\mathbf{L}^*$ is the forward Radon operator, which transfers the Radon coefficients in the $f-p$ domain to the $t-x$ domain. $\mathbf{m}$ is the Radon coefficients in the $f-p$ domain. The $\|\mathcal{L}^* \mathbf{m} - \mathbf{d}\|_1$ make the cost function robust to the erratic noise. $\beta \sum_{i=1}^{N} \|\mathbf{m}[i]\|_2$ is use to promote the group sparsity, and $\lambda \|\mathbf{m}\|_1$ can enhance the sparsity within the groups. To solve this problem, we use the greedy method; in our case, we use the orthogonal Matching Pursuit (OMP).

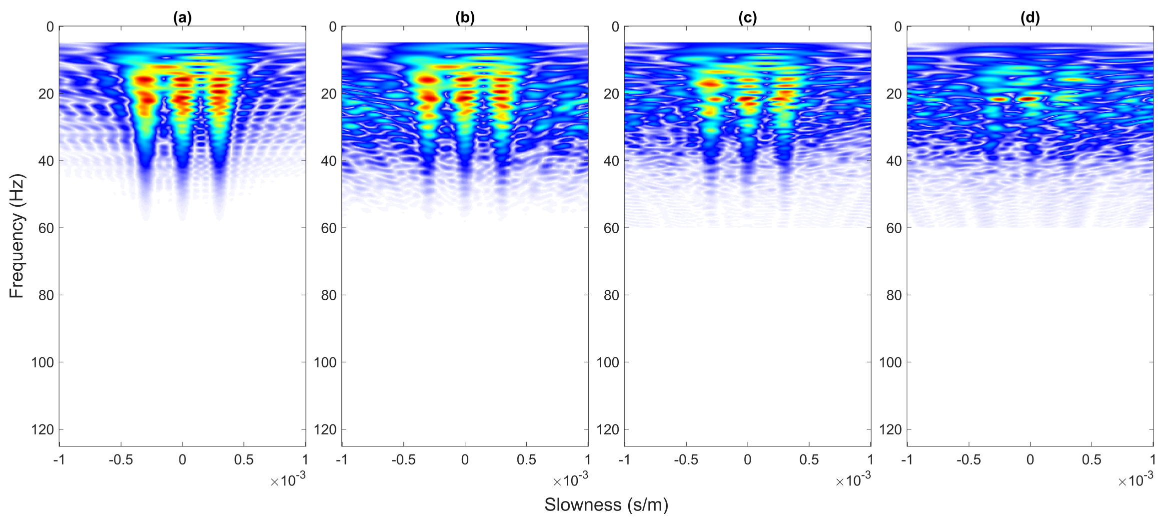

FIG. 2. The correspond Radon coefficients in the $f-p$ domain of figure 1

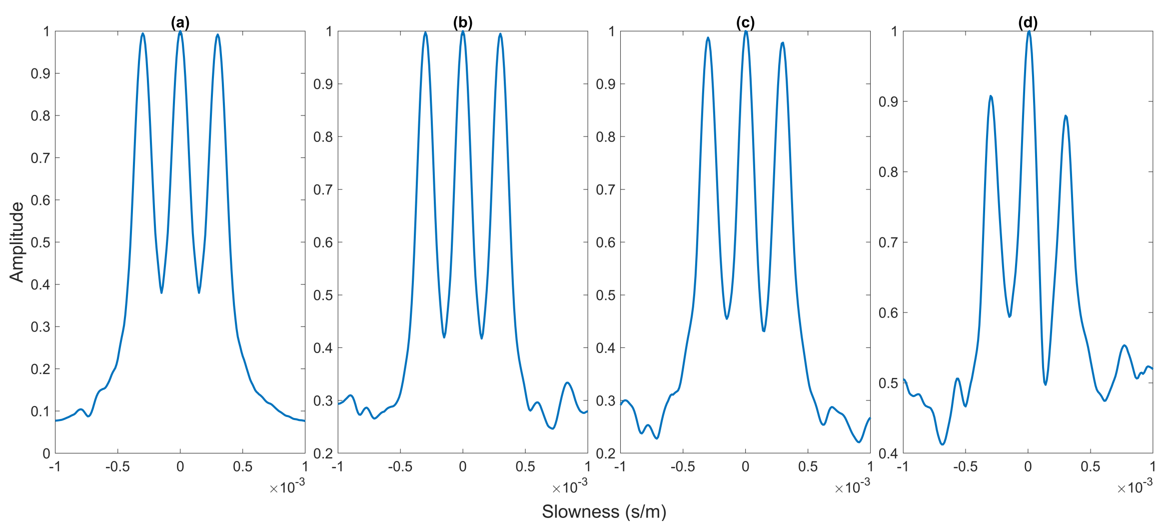

FIG. 3. The correspond Energy map of slowness p of figure 1.

Orthogonal Matching Pursuit

Matching Pursuit (MP) (Mallat and Zhang, 1993) and Orthogonal Matching Pursuit (OMP) (Tropp and Gilbert, 2007; Pati et al., 1993) are the two most frequently used Greedy pursuit algorithms. Unlike MP, which updates one coefficient in each iteration, OMP minimizes the following cost function for all of the currently selected coefficients in each iteration

$$\widehat{\mathbf{x}}^{[i]}_{T^{[i]}} = \underset{\widetilde{\mathbf{x}}_{T^{[i]}}}{\operatorname{argmin}} \ \|\mathbf{y} - \mathbf{A}_{T^{[i]}}\widetilde{\mathbf{x}}_{T^{[i]}}\|_2^2, \tag{12}$$

where $T^{[i]}$ is the set of the indexes of all the coefficients we have picked until iteration i. Therefore, the computational cost of OMP is higher than MP. But it usually generates a superior result. The full algorithm is listed in Algorithm 1.

Algorithm 1 OMP

Input: $\mathbf{y}$, $\mathbf{A}$,and k
Output: $\mathbf{r}^{[k]}$, $\widehat{\mathbf{x}}^{[k]}$
Initialization: $\mathbf{r}^{[0]} = \mathbf{y}$,$\widehat{\mathbf{x}}^{[k]} = 0$, and $T^{[0]} = \emptyset$
for $k = 1, 2, ..., K$ **do**
  $l = \underset{j=1,2,....M}{\operatorname{argmax}} \ |\langle \mathbf{A}, \mathbf{r}^{[k-1]} \rangle|$
  $T^{[k]} = T^{[k-1]} \cup \{l\}$
  $\widehat{\mathbf{x}}^{[i]}_{T^{[i]}} = \underset{\widetilde{\mathbf{x}}_{T^{[i]}}}{\operatorname{argmin}} \ \|\mathbf{y} - \mathbf{A}_{T^{[i]}}\widetilde{\mathbf{x}}_{T^{[i]}}\|_2^2$
  $\mathbf{r}^{[k]} = \mathbf{y} - \mathbf{A}\widehat{\mathbf{x}}^{[k]}_{T^{[k]}}$
end for

Computing group sparsity solution with OMP and Radon operators

To solve the minimization problem 11, we need to modify the OMP algorithm. Instead of picking one coefficient each time, we pick a group of coefficients for the basis function selection part. We divide our coefficients into different groups based on the slowness p. Then we pick the group which has the maximum ℓ_2 norm. And as we can see from figure 3, picking the slowness p based on the summation of the norm can also remove a lot of effects from the random noise in the $f - p$ domain, which makes the coefficient selection part more robust. And then, we utilize all the coefficients located with all currently pick slowness groups p_{T^k} by minimizing the following cost function

$$\widehat{\mathbf{m}}^k_{T^k} = \underset{\widetilde{\mathbf{m}}_{T^k}}{\operatorname{argmin}} \ \|\mathbf{d} - \mathcal{L}_n^T(\widetilde{\mathbf{m}}_{T^k})\|_1 + \lambda\|\mathbf{m}\|_1. \tag{13}$$

This $\ell1 - \ell1$ minimization problem can be solved by the $\ell1 - \ell1$ ADMM solver (Yang and Zhang, 2011; Wen et al., 2016), which we will explain later. The full algorithm of the robust group sparsity OMP is listed below.

Algorithm 2 GOMP

Input: $\mathbf{d}$, $\mathcal{L}$,and k
Output: $\widehat{\mathbf{m}}^{[k]}$
Initialization: $\mathbf{r}^{[0]} = \mathbf{d}$,$\widehat{\mathbf{m}}^{[k]} = 0$, and $T^{[0]} = \emptyset$
for $k = 1, 2, ..., K$ **do**
  $\mathbf{M} = \mathcal{L}\mathbf{r}$
  Pick dominant slowness p_k
  $T^k = T^{k-1} + p_k$
  $\widehat{\mathbf{m}}^k_{T^k} = \underset{\widetilde{\mathbf{m}}_{T^k}}{\operatorname{argmin}} \; \|\mathbf{d} - \mathcal{L}^*_n(\widetilde{\mathbf{m}}_{T^k})\,\|_1 + \lambda\|\mathbf{m}\|_1$
  $\mathbf{r}^k = \mathbf{d} - \mathcal{L}^*\widehat{\mathbf{m}}^k_{T^k}$
end for

$\ell1 - \ell1$ ADMM

ADMM is an optimization algorithm used in the areas of machine learning and signal processing a lot. To solve the following optimization problem

$$\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_1 + \lambda\|\mathbf{x}\|_1, \tag{14}$$

We first separates the ℓ_1-norm loss term and the ℓ_1-regularization term. Then we uses an auxiliary variable $\mathbf{v}$ to replace the error term; the problem can be reformulated as

$$\operatorname{argmin} \|\mathbf{v}\|_1 + \lambda\|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} - \mathbf{y} = \mathbf{v} \tag{15}$$

Using the augmented Lagrangian method, this problem can be solved iteratively follows.

$$\mathbf{v}^{k+1} = \underset{\mathbf{v}}{\operatorname{argmin}}(\|\mathbf{v}\|_1 + \frac{\rho}{2}\|\mathbf{A}\mathbf{x}^k - \mathbf{y} - \mathbf{v} - \frac{\mathbf{w}^k}{\rho}\|_2^2) \tag{16}$$

$$\mathbf{x}^{k+1} = \underset{\mathbf{x}}{\operatorname{argmin}}(\|\mathbf{x}\|_1 + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y} - \mathbf{v}^{k+1} - \frac{\mathbf{w}^k}{\rho}\|_2^2) \tag{17}$$

$$\mathbf{w}^{k+1} = \mathbf{w}^k - \rho(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{y} - \mathbf{v}^{k+1}). \tag{18}$$

Both 16 and 17 have closed-form solutions, which the soft-thresholding operator can solve.

$$\mathbf{v}^{k+1} = Shrink(\mathbf{A}\mathbf{x}^k - \mathbf{y} - \mathbf{w}^k/\rho, \rho) \tag{19}$$

$$\mathbf{x}^{k+1} = Shrink(\mathbf{x}^k - \mu\mathbf{A}^*(\mathbf{A}\mathbf{x}^k - \mathbf{u}^k, \rho/\mu), \tag{20}$$

where $\mathbf{u}^k = \mathbf{y} + \mathbf{v}^{k+1} + \mathbf{w}^k/\rho$. μ is scale parameter > 0. $Shrink$ is soft-thresholding operator where $Shrink(\mathbf{x}, t)$=sign(x_i)*max$\{x_i - t, 0\}$. The detailed mathematical theory about how to get these closed-form solutions can be found in paper (Wen et al., 2016)

EXAMPLES

In this section, we will test the algorithm on synthetic and real data examples.We computed the signal-to-noise-ratio of the output $SNR_{out} = 10\log\frac{\|\mathbf{d_c}\|_2^2}{\|\mathbf{d_c}-\mathbf{d_r}\|}$ where $\mathbf{d}_r$ is the reconstructed data and $\mathbf{d}_c$ is original clean data. We will use the SNR value to compare the reconstruction performance.

Simple 2D synthetic examples

With and without group

We first test the erratic noise attenuation performance of the proposed algorithm. We use the same synthetic data as in figure 1. This test compares the proposed algorithm with the other three cases.

- $\ell_2 - \ell_1$ without group sparsity: $\text{argmin}\,\|\mathcal{L}^*\mathbf{m} - \mathbf{d}\|_2^2 + \lambda\|\mathbf{m}\|_1$.
- $\ell_1 - \ell_1$ without group sparsity: $\text{argmin}\,\|\mathcal{L}^*\mathbf{m} - \mathbf{d}\|_1 + \lambda\|\mathbf{m}\|_1$.
- $\ell_2 - \ell_1$ with group sparsity: $\text{argmin}\,\|\mathcal{L}^*\mathbf{m} - \mathbf{d}\|_2^2 + \lambda\|\mathbf{m}\|_1 + \beta\sum_{i=1}^{N}\|\mathbf{m}[i]\|_2$.

Figure 4 shows the denoised results with a different method. As expected, the $\ell_2 - \ell_1$ inversion, which is not robust to the erratic noise, has the worst performance. The $\ell_1 - \ell_1$ inversion performs better than the $\ell_2 - \ell_1$. Since the erratic noises are extreme, the $\ell_1 - \ell_1$ inversion method still has much noise. Part (c) is the $\ell_2 - \ell_1$ with group sparsity. As we can see, even with the $\ell_2 - \ell_1$ inversion, the performance is better than using the $\ell_1 - \ell_1$ inversion without group sparsity. And finally, part (g) is the denoised result with $\ell_1 - \ell_1$ inversion and group sparsity, which has the best performance.

Figure 5 shows the estimated Radon coefficients in the $f - k$ domain, and figure 6 represents the estimated Radon coefficients in the $\tau - p$ domain. The proposed algorithm removed the noise in the $\tau - p$ domain and the smear effect of the Radon transform simultaneously. This example shows the proposed robust group sparsity algorithm has the best denoising performance when dealing with erratic noise.

Compare with other Radon-based methods

We also compare our algorithm with other Radon-based inversion methods. In our method, we work with Radon transform on the $f - p$ domain. There are two more frequently used approaches. One is working on the $\tau - p$ domain (Lin and Sacchi, 2020), which can be calculated as the inverse Fourier transform of the $f - p$ domain or do the Radon transform on the time domain directly. The other is to do the sparse inversion on each frequency slice (Tran et al., 2014). We will apply the $\ell_1 - \ell_1$ sparse inversion on both algorithms and compare it with the proposed robust group sparse inversion that works on the whole $f - p$ domain. Figure 7 shows the final denoising results. Part (a) is the denoising result by the method that works on each frequency slice. For this method, the sparse inversion tries to fit each frequency. For one frequency slice, it's hard to decide which coefficients are caused by the signal and which are caused by noise. Therefore, it has the worst denoising performance. Part (c) is the denoising result by the $\ell_1 - \ell_1$ robust inversion works on the $\tau - p$ domain, the result is identical to the $\ell_1 - \ell_1$ robust inversion works on the $f - p$ domain. And again, Part (e) is the result estimated by the robust group sparse inversion, which has the best performance. And figure 8 is the final result in the $\tau - p$ domain.

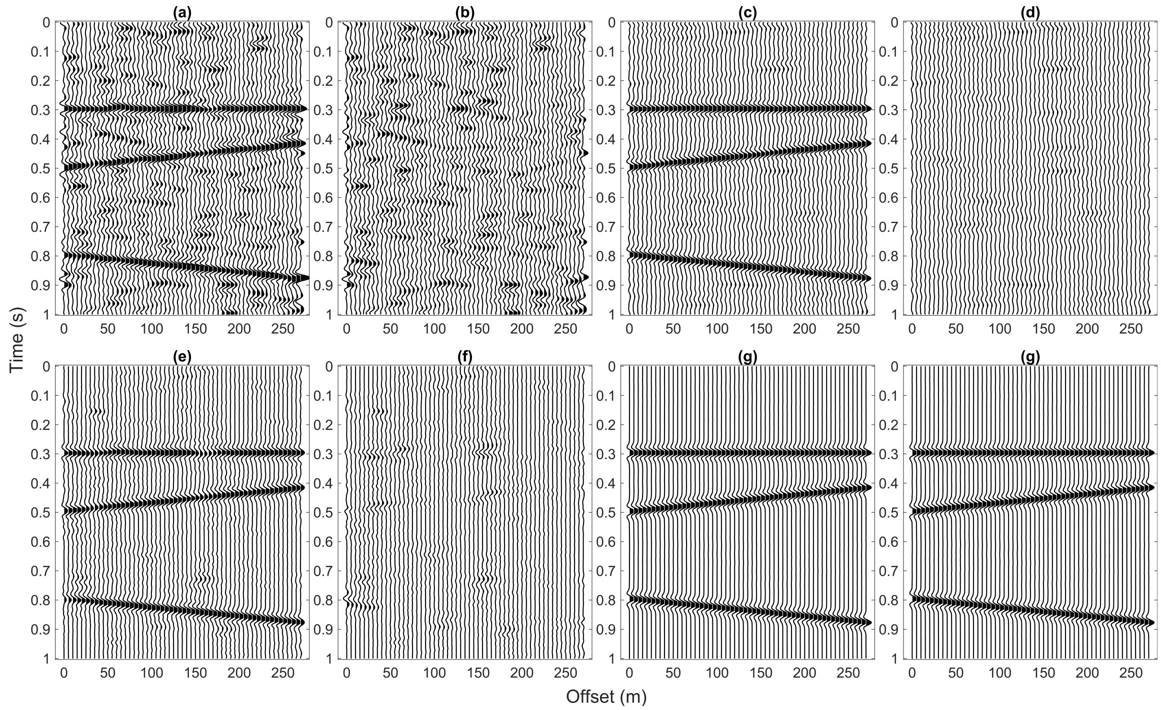

FIG. 4. Reconstructed result in the $t-x$ domain. (a) $\ell_2-\ell_1$ without group sparsity, snr=2.35 dB. (b) Errors between a and clean data. (c) $\ell_2-\ell_1$ with group sparsity, snr=6.9 dB. (d) Errors between c and clean data. (e) $\ell_1-\ell_1$ without group sparsity, snr=5.22 dB. (f) Errors between e and clean data. (g) $\ell_1-\ell_1$ with group sparsity, snr=14.3 dB. (h) Errors between g and clean data.

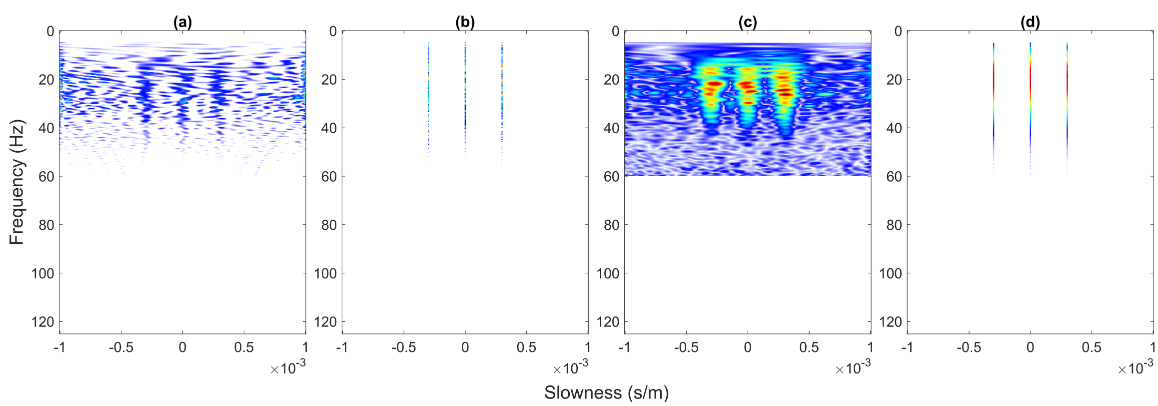

FIG. 5. Reconstructed result in the $\tau-p$ domain. (a) $\ell_2-\ell_1$ without group sparsity. (b) $\ell_2-\ell_1$ with group sparsity. (c) $\ell_1-\ell_1$ without group sparsity. (d) $\ell_1-\ell_1$ with group sparsity

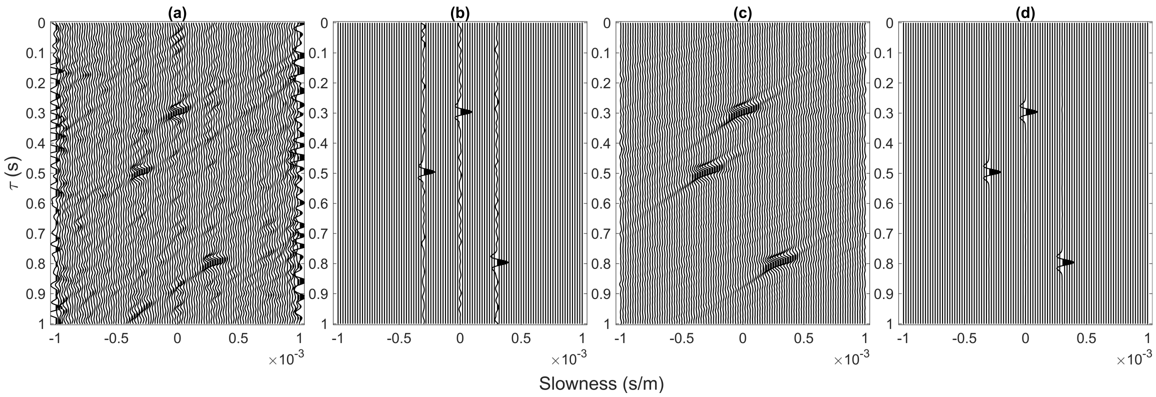

FIG. 6. Reconstructed result in the $\tau - p$ domain. (a) $\ell_2 - \ell_1$ without group sparsity. (b) $\ell_2 - \ell_1$ with group sparsity. (c) $\ell_1 - \ell_1$ without group sparsity. (d) $\ell_1 - \ell_1$ with group sparsity.

2D seismic data reconstruction

We now apply our algorithm to seismic interpolation and erratic noise attenuation simultaneously. To do the interpolation, we need to add a sampling operator $\mathcal{S}$ in our algorithm.

$$\widehat{\mathbf{m}}^k_{T^k} = \underset{\widetilde{\mathbf{m}}_{T^k}}{\operatorname{argmin}} \|\mathbf{d} - \mathcal{S}\mathcal{L}^*_n(\widetilde{\mathbf{m}}_{T^k})\|_1 + \lambda\|\mathbf{m}\|_1 \tag{21}$$

$$\mathbf{r}^k = \mathbf{d} - \mathcal{S}\mathcal{L}^*\widehat{\mathbf{m}}^k_{T^k} \tag{22}$$

Figure 9 shows the final denoised and interpolated result.

Simple 3D synthetic examples

Our algorithm can extend to 3D with the 3D Radon transform. Figure 10 shows a simple 3D cube with three linear events. Part (b) is the synthetic data with erratic noise and also 90% of traces missed. Part (c) is the reconstructed result.

Field data example

Figure 11 shows the application of the proposed algorithm to real data. For this example, we compare the result with the POCS. For this real data example, the POCS method doesn't work well, and the result of the proposed method is more continuous and smooth.

CONCLUSIONS

We proposed a robust group sparse inversion algorithm which can provide a better sparse estimation and be robust to the erratic noise simultaneously. The core of the proposed method is based on the orthogonal Matching Pursuit. We divide the Radon coefficients in the $f-p$ domain into different slowness groups. In each iteration, the algorithm selected the group with the maximum norm. And then, the Radon coefficients located within all currently selected groups are directly fitting to the seismic data in the $t - x$ domain. The coefficient optimization part is solved by the modified $\ell_1 - \ell_1$ ADMM solver, which makes the algorithm robust to the erratic noise. The tests on both 2D and 3D synthetic

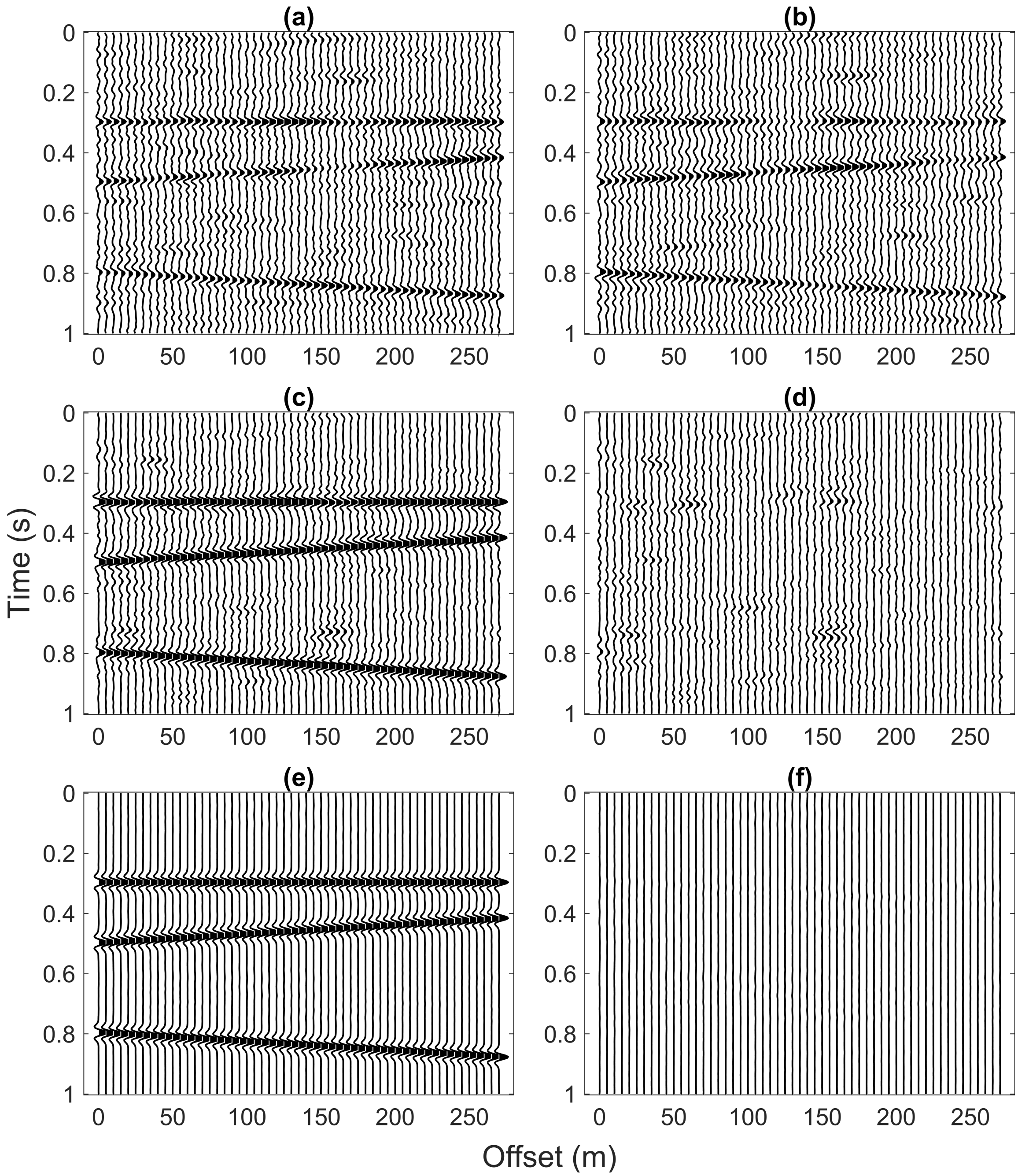

FIG. 7. Reconstructed result in the $t-x$ domain. (a) $\ell_1-\ell_1$ with method works one each frequency slice, snr=2.31 dB. (b) Errors between a and clean data. (c) $\ell-\ell_1$ with method works on $\tau-p$ domain, snr=6.3 dB. (d) Errors between c and clean data. (e) $\ell_1-\ell_1$ with method works on $f-p$ domain and group sparsity, snr=14.3 dB. (f) Errors between e and clean data.

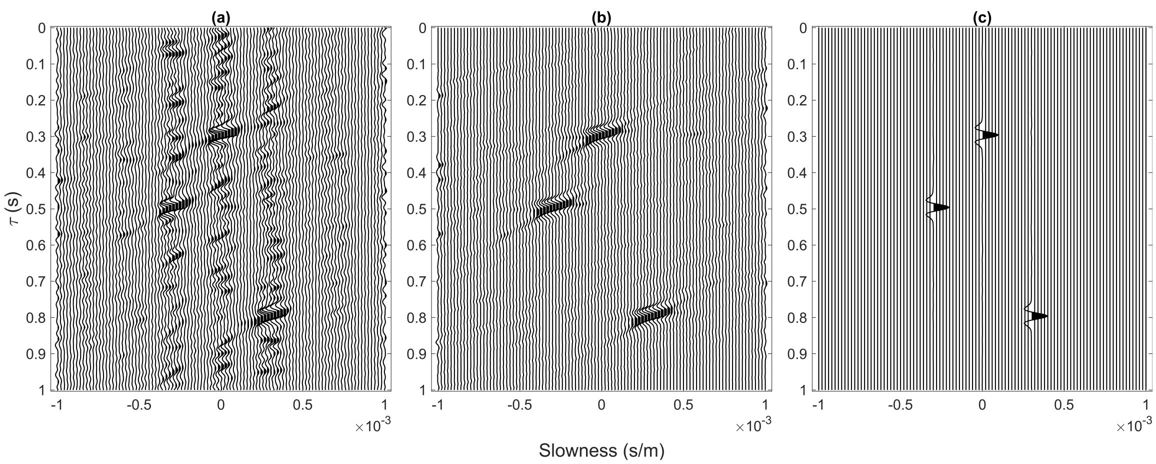

FIG. 8. Reconstructed result in the $\tau - p$ domain. (a) $\ell_1 - \ell_1$ with method works one each frequency slice. (b) $\ell - \ell_1$ with method works on $\tau - p$ domain. (c) $\ell_1 - \ell_1$ with method works on $f - p$ domain and group sparsity.

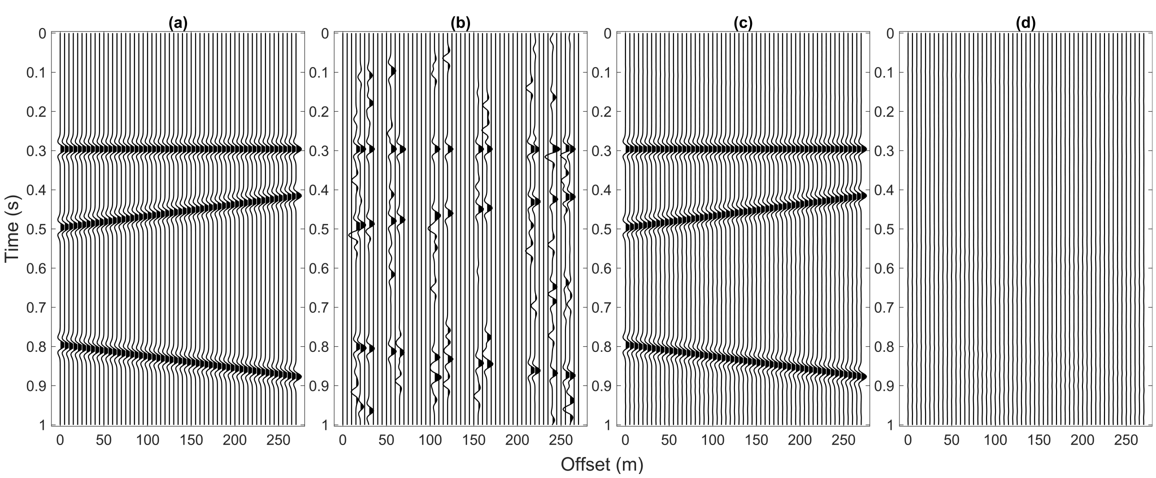

FIG. 9. Reconstructed result in the $t - x$ domain. (a) Clean data. (b) Synthetic data with erratic noise and 75% of traces missed. (c) Reconstructed result by the proposed algorithm, snr=13.5 dB. (d) Errors between clean data and reconstructed data.

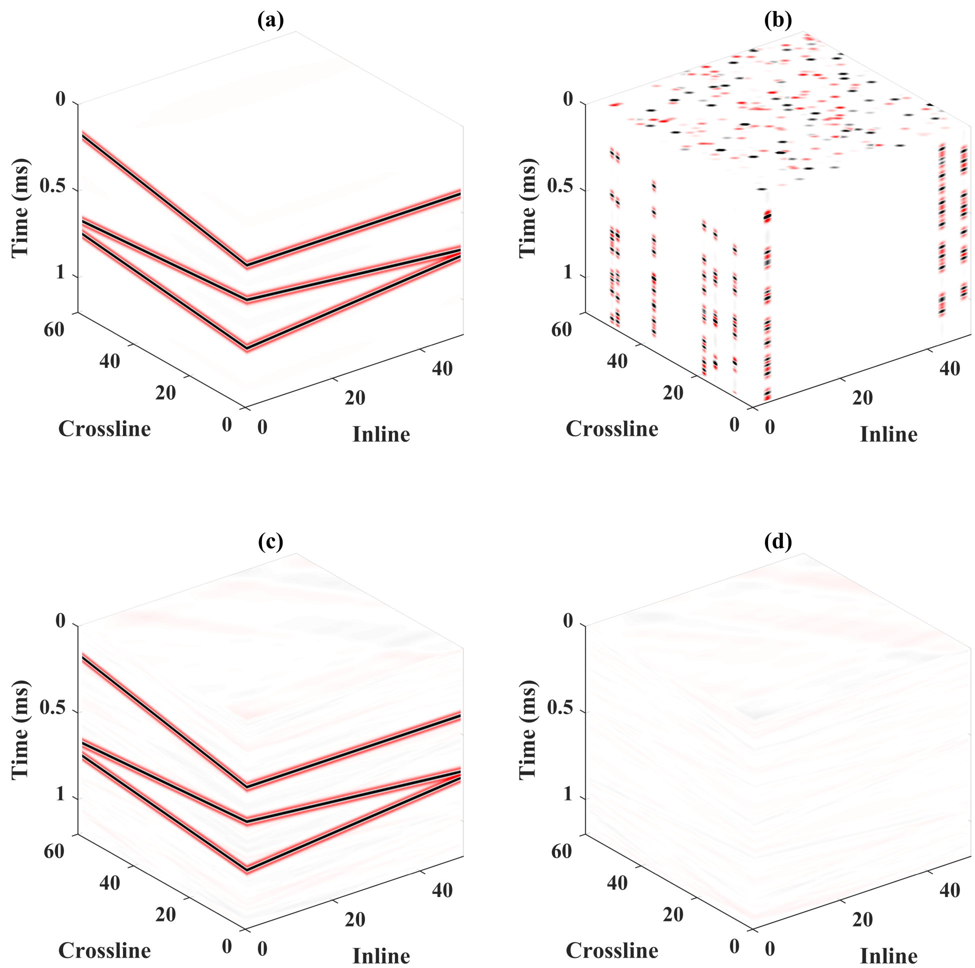

FIG. 10. (a) 3D cube synthetic data. (b) 3D cube synthetic data with erratic noise and 90% of traces missed. (c) The reconstructed result, snr=13.7 dB. (d) Errors between the original data and reconstructed data.

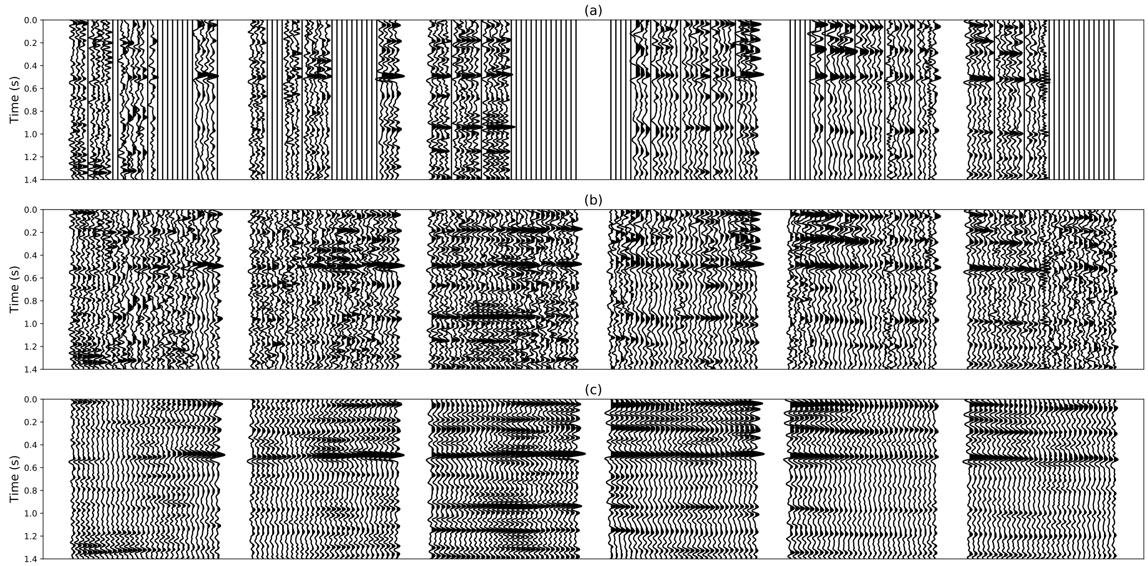

FIG. 11. (a) Six slides of a 3D cube real data. (b) POCS interpolation. (c) Robust group sparsity interpolation

examples prove the effectiveness and robustness of the proposed method. Furthermore, compared with the traditional MP and OMP algorithm, the proposed algorithm can save the total costs a lot since it selects one group with multiple coefficients instead of picking just one best-correlated coefficient like MP and OMP.

ACKNOWLEDGMENTS

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19.

REFERENCES

Abma, R., and Kabir, N., 2006, 3d interpolation of irregular data with a pocs algorithm: Geophysics, **71**, No. 6, E91–E97.

Bach, F. R., 2008, Consistency of the group lasso and multiple kernel learning.: Journal of Machine Learning Research, **9**, No. 6.

Beck, A., and Teboulle, M., 2009, A fast iterative shrinkage-thresholding algorithm for linear inverse problems: SIAM journal on imaging sciences, **2**, No. 1, 183–202.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J. et al., 2011, Distributed optimization and statistical learning via the alternating direction method of multipliers: Foundations and Trends® in Machine learning, **3**, No. 1, 1–122.

Chen, S. S., Donoho, D. L., and Saunders, M. A., 2001, Atomic decomposition by basis pursuit: SIAM review, **43**, No. 1, 129–159.

Chen, Y., Peng, Z., Li, M., Yu, F., and Lin, F., 2019, Seismic signal denoising using total generalized variation with overlapping group sparsity in the accelerated admm framework: Journal of Geophysics and Engineering, **16**, No. 1, 30–51.

Elhamifar, E., and Vidal, R., 2011, Robust classification using structured sparse representation, *in* CVPR 2011, IEEE, 1873–1879.

Guitton, A., and Symes, W. W., 2003, Robust inversion of seismic data using the huber norm: Geophysics, **68**, No. 4, 1310–1319.

Herrmann, F. J., 2010, Randomized sampling and sparsity: Getting more information from fewer samples: Geophysics, **75**, No. 6, WB173–WB187.

Herrmann, P., Mojesky, T., Magesan, M., and Hugonnet, P., 2000, De-aliased, high-resolution radon transforms, *in* SEG Technical Program Expanded Abstracts 2000, Society of Exploration Geophysicists, 1953–1956.

Huang, J., and Zhang, T., 2010, The benefit of group sparsity: The Annals of Statistics, **38**, No. 4, 1978–2004.

Li, J., and Sacchi, M. D., 2021, An lp-space matching pursuit algorithm and its application to robust seismic data denoising via time-domain radon transforms: Geophysics, **86**, No. 3, V171–V183.

Li, J., and Sacchi, M. D., 2022, Robust reconstruction via orthogonal matching pursuit with fourier operators, *in* Second International Meeting for Applied Geoscience & Energy, Society of Exploration Geophysicists and American Association of Petroleum . . . , 2947–2951.

Lin, R., and Sacchi, M. D., 2020, Separation of simultaneous sources acquired with a high blending factor via coherence pass robust radon operators: Geophysics, **85**, No. 3, V269–V282.

Lin, T. T., and Herrmann, F. J., 2013, Robust estimation of primaries by sparse inversion via one-norm minimization: Geophysics, **78**, No. 3, R133–R150.

Majumdar, A., and Ward, R. K., 2009, Classification via group sparsity promoting regularization, *in* 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 861–864.

Mallat, S. G., and Zhang, Z., 1993, Matching pursuits with time-frequency dictionaries: IEEE Transactions on signal processing, **41**, No. 12, 3397–3415.

Maronna, R. A., 1976, Robust m-estimators of multivariate location and scatter: The annals of statistics, 51–67.

Naghizadeh, M., 2012, Seismic data interpolation and denoising in the frequency-wavenumber domain: Geophysics, **77**, No. 2, V71–V80.

Nardi, Y., and Rinaldo, A., 2008, On the asymptotic properties of the group lasso estimator for linear models: Electronic Journal of Statistics, **2**, 605–633.

Pati, Y. C., Rezaiifar, R., and Krishnaprasad, P. S., 1993, Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition, *in* Proceedings of 27th Asilomar conference on signals, systems and computers, IEEE, 40–44.

Sacchi, M. D., and Ulrych, T. J., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, No. 4, 1169–1177.

Sacchi, M. D., Ulrych, T. J., and Walker, C. J., 1998, Interpolation and extrapolation using a high-resolution discrete fourier transform: IEEE Transactions on Signal Processing, **46**, No. 1, 31–38.

Scales, J. A., and Gersztenkorn, A., 1988, Robust methods in inverse theory: Inverse problems, **4**, No. 4, 1071.

Simon, N., Friedman, J., Hastie, T., and Tibshirani, R., 2013, A sparse-group lasso: Journal of computational and graphical statistics, **22**, No. 2, 231–245.

Tibshirani, R., 1996, Regression shrinkage and selection via the lasso: Journal of the Royal Statistical Society: Series B (Methodological), **58**, No. 1, 267–288.

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, No. 6, V123–V132.

Trad, D. O., Ulrych, T. J., and Sacchi, M. D., 2002, Accurate interpolation with high-resolution time-variant radon transforms: Geophysics, **67**, No. 2, 644–656.

Tran, T. N., Nguyen, K.-C. T., Sacchi, M. D., and Le, L. H., 2014, Imaging ultrasonic dispersive guided wave energy in long bones using linear radon transform: Ultrasound in medicine & biology, **40**, No. 11, 2715–2727.

Trickett, S., Burroughs, L., and Milton, A., 2012, Robust rank-reduction filtering for erratic noise, *in* 2012 SEG Annual Meeting, OnePetro.

Tropp, J. A., and Gilbert, A. C., 2007, Signal recovery from random measurements via orthogonal matching pursuit: IEEE Transactions on information theory, **53**, No. 12, 4655–4666.

Van Groenestijn, G., and Verschuur, D., 2009, Estimating primaries by sparse inversion and application to near-offset data reconstruction: Geophysics, **74**, No. 3, A23–A28.

Vera Rodriguez, I., Bonar, D., and Sacchi, M., 2012, Microseismic data denoising using a 3c group sparsity constrained time-frequency transform: Geophysics, **77**, No. 2, V21–V29.

Vincent, M., and Hansen, N. R., 2014, Sparse group lasso and high dimensional multinomial classification: Computational Statistics & Data Analysis, **71**, 771–786.

Wen, F., Liu, P., Liu, Y., Qiu, R. C., and Yu, W., 2016, Robust sparse recovery in impulsive noise via $\ell_p - \ell_1$ optimization: IEEE Transactions on Signal Processing, **65**, No. 1, 105–118.

Yang, J., and Zhang, Y., 2011, Alternating direction algorithms for ℓ_1-problems in compressive sensing: SIAM journal on scientific computing, **33**, No. 1, 250–278.

Yuan, M., and Lin, Y., 2006, Model selection and estimation in regression with grouped variables: Journal of the Royal Statistical Society: Series B (Statistical Methodology), **68**, No. 1, 49–67.