3D data interpolation using adaptive rank reduction method

Farzaneh Bayati, Daniel Trad

ABSTRACT

Seismic data reconstruction is an important step in seismic data processing that affects the whole processing sequence because many tools for noise attenuation or imaging require the input data to be sampled regularly in space to work properly. It is also important for data acquired on a difficult terrain with natural or cultural obstacles which may be missing a large portion of the surface shots and receivers. For plane waves or data with small curvatures, rank reduction method is a very effective signal reconstructing method. But when it comes to more complicated data, the rank reduction method may fail or give poor results as a consequence of curved events not having a small rank (sparse) representation. To satisfy the plane wave assumption for the rank reduction method, one can utilize local windows to assume that events are plane waves. The rank reduction method requires as a parameter the number of events. This number defines the minimum rank selected in each step. However, it is difficult to select the appropriate rank in each window. In this report, we propose a method to select the rank automatically in each window by finding the maximum ratio of the energy between two singular values. We test the efficiency of the method by applying it to both synthetic and real seismic data.

INTRODUCTION

Seismic reconstruction methods can be divided into three main classes: signal processing based methods, wave equation based methods, and rank reduction based methods.

Most of the methods in the signal processing based category are multidimensional and use prediction filters (Abma and Claerbout, 1995; Spitz, 1991; Porsani, 1999), transform domains such as Fourier transform (Sacchi et al., 1998; Liu and Sacchi, 2004; Trad, 2009), Radon transform (Sacchi and Ulrych, 1995; Trad et al., 2002), or Curvelet transform (Herrmann et al., 2008). Some hybrid techniques use a combination of Fourier transform with prediction errors filters (Naghizadeh and Sacchi, 2010) as a way to improve interpolation beyond aliasing.

Wave equation based algorithms implement an implicit migration de-migration pair. Stolt (2002) introduced mapping and reconstruction operators reconstructing missing traces, regularizing a data set, and removing acquisition footprints. A finite-difference offset continuation filter for interpolating seismic reflection data was proposed by (Fomel, 2003).

In rank reduction based methods, the linear events in a clean seismic data set are low rank in the time domain. However, noise and missing traces increase the rank of data (Trickett, 2008). The rank reduction algorithm in the frequency domain is carried out in frequency slices by generating Hankel/Toeplitz matrix and applying a low-rank reduction method on the generated matrix. The singular spectrum analysis (SSA) method proposed by (Oropeza and Sacchi, 2011) works by rank reduction of the Hankel matrix with an iterative algorithm in the frequency domain. Gao et al. (2013) extended the SSA method to higher dimensional seismic data and called a multichannel singular spectrum (MSSA).

One of the advantages of rank reduction methods is simultaneous random noise attenuation and data interpolation. One of its limitations, on the other hand, is that it needs to satisfy the plane wave assumption. To satisfy the plane wave assumption the rank reduction methods need to be applied on local windows. Most of the time it is not easy to find the proper window size because it is hard to decide whether the structure in the local window is linear or not. Moreover, it is hard to approximate the rank of each window. Choosing the wrong rank will lead to a failure because the overestimation of rank remains a significant residual and underestimation of it will cause random noise and distort the signal. In this report, we will apply a method that selects rank automatically for each local window. The method was proposed by Wu and Bai (2018) for 2D data. We apply the proposed rank selection criterion to both synthetic 3D data and demonstrate its successful performance.

Background

Singular Spectrum Analysis

Let's start with a 3D seismic record $d(t, x, y)$ the transformed data in Fourier domain will be $D(f, x, y)$. One frequency slice organized from the 3D seismic record can be shown as:

$$D_f = \begin{pmatrix} d_{(1,1)} & d_{(1,2)} & \dots & d_{(1,N_y)} \\ d_{(2,1)} & d_{(2,2)} & \dots & d_{(2,N_y)} \\ \vdots & \vdots & \ddots & \vdots \\ d_{(N_x,1)} & d_{(N_x,2)} & \dots & d_{(N_x,N_y)} \end{pmatrix}, \tag{1}$$

where N_x and N_y are number of traces in x and y directions respectively. Generating a Hankel matrix in each inline will leads to:

$$\mathbf{M_j} = \begin{pmatrix} D_{(1,j)} & D_{(2,j)} & \dots & D_{(L_x,j)} \\ D_{(2,j)} & D_{(3,j)} & \dots & D_{(L_x+1,j)} \\ \vdots & \vdots & \ddots & \vdots \\ D_{(K_x,j)} & D_{(K_x+1,j)} & \dots & D_{(N_x,j)} \end{pmatrix}, \tag{2}$$

where N_x and L_x are the dimensions of the Hankel matrix and $L_x = floor(N_x/2) + 1$ and $K_x = N_x - L_x + 1$. We generate a Hankel matrix of Hankel matrices to add the cross-line dimension:

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 & \dots & \mathbf{M}_{K_y} \\ \mathbf{M}_2 & \mathbf{M}_3 & \dots & \mathbf{M}_{Ky+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{L_y} & \mathbf{M}_{L_{y+1}} & \dots & \mathbf{M}_{N_y} \end{pmatrix}. \tag{3}$$

We call the matrix of Equation 3 a block Hankel matrix. The size of the block Hankel matrix is $(L_y \times L_x) \times (K_y \times K_x)$. The size of the block Hankel matrix can easily blow up by the increasing number of channels in each dimension of data. Equation 3 equals to:

$$\mathbf{M} = \mathcal{H}\mathcal{H}\mathcal{F}[d(t, x, y)]. \tag{4}$$

In Equation 4, $\mathcal{H}$ denotes Hankel operator and $\mathcal{F}$ denotes Fourier transform. It can be proved that the rank of the block Hankel matrix is equals to the number of plane waves in the data set. However, the presence of random noise or missing traces in the recorded data, leads to increasing the rank of the block Hankel matrix (Hua, 1992). Applying truncated singular value decomposition (TSVD) to the block Hankel matrix can reduce the rank of the block Hankel matrix by remaining the number of largest singular values equals to the number of the plane waves in the data set:

$$svd(\mathbf{M}) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{5}$$

$$\mathbf{\Sigma} = diag(\sigma_1, ..., \sigma_n), \tag{6}$$

where $\mathbf{\Sigma}$ is the matrix of singular values of matrix $\mathbf{M}$, $\mathbf{U}$ and $\mathbf{V}$ are the matrices of singular vectors of the block Hankel matrix $\mathbf{M}$, and σ_i are the singular values (Golub and Reinsch, 1971). $\mathbf{U}$ and $\mathbf{V}$ are orthogonal and $\mathbf{\Sigma}$ is diagonal. The number of non-zero singular values in the Hankel matrix determines the rank of the Hankel matrix. If the number of the plane waves in the data set is k, the rank of the matrix $\mathbf{M}$ equals to k. We can obtain the reduced rank matrix as follow:

$$\mathbf{M}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^H, \tag{7}$$

where $\mathbf{\Sigma}$ is a diagonal matrix with k largest singular values of $\mathbf{M}$, $\mathbf{U}_k$ and $\mathbf{V}_k$ are the k first singular vectors of $\mathbf{M}$. By averaging the anti diagonals of each Hankel matrix, the reconstructed signal $\mathbf{D}$ can be obtained. The whole process of MSSA can be shown as:

$$\hat{d}(t, x, y) = \mathcal{F}^{-\prime}\mathcal{ARHHF}[d(t, x, y)], \tag{8}$$

where $\mathcal{F}^{-\prime}$ denotes the inverse Fourier transform, $\mathcal{A}$ shows the average of anti diagonals, $\mathcal{R}$ indicates rank reduction operator using TSVD. To reconstruct the missing traces, we apply MSSA algorithm in a recursive algorithm. The seismic record with missing traces contains zeros in missing traces that increase the rank of the Hankel matrix. The recursive algorithm to recover missing traces can be written as:

$$\mathbf{D}_{n+1} = \mathbf{D}_{obs} + (\mathbf{I} - \mathbf{S}) \odot \mathcal{P}\mathbf{D}_n, \tag{9}$$

where, $\mathbf{D}_n$ indicates the reconstructed data in Fourier domain after n iterations, $\mathbf{S}$ is a sampling operator, $\mathbf{I} = ones(size(\mathbf{S}))$, the operator $\mathcal{P}$ denotes $\mathcal{ARHHF}$, $\odot$ is array multiplication for two matrices, and $\mathbf{D}_{obs}$ is the observed data. For a data set contaminated with random noise the algorithm can be modified as:

$$\mathbf{D}_{n+1} = \alpha^n\mathbf{D}_{obs} + (1 - \alpha^n)\mathbf{S} \odot \mathcal{P}\mathbf{D}_n + (\mathbf{I} - \mathbf{S}) \odot \mathcal{P}\mathbf{D}_n, \tag{10}$$

where α^n is an iteration dependent scalar that decreases from 1 to zero when the maximum iteration occurs. The logic of this algorithm is that the original observed data is only partially inserted into the algorithm, with the largest contribution at the beginning of the iterations and no contribution at the last iteration.

MSSA as a low-rank reduction method is a powerful algorithm in denoising and constructing the missing traces simultaneously. However, there is an obstacle for MSSA that is assuming the events are linear. Complicated structures with large curvature are not suitable for low-rank reduction methods like MSSA. One solution for the curved events is applying NMO correction before applying the low-rank reduction method and the other one is using local spatial windows to assume that events are linear.

In the MSSA algorithm, the rank of the block Hankel matrix equals the number of events in each local window that is equal to the number of distinctly large singular values. To prove this idea we test two different data set to inspect how the singular values are related

to the number of events by investigating of distribution of singular values of them. The first data set is a 3D cube of a linear event. The 3D data has 29 traces in in-line direction and 29 traces in cross-line direction. Figure (1-a) is a slice of the cube of clean data in in-line direction. Figure (1-b) shows the singular value distribution of the block Hankel matrix of the clean data in a constant frequency of 30 Hz where the largest non zero singular value is related to the energy of the linear event. Figure (1-c) represents a slice of a cube of 3D data containing one linear event contaminated with random noise and missing traces. Figure (1-d) shows the singular value distribution of the block Hankel matrix of the incomplete data in 30 Hz. We can see that presence of missing traces and random noise increased the rank of the block Hankel matrix but still, the energy of the linear event can be diagnosed with the first distinct large singular value.
We repeat the test with a 3D cube data set containing four linear events. Figure (2-a) is a slice of the cube of clean data in in-line direction with 4 linear events. Figure (2-b) shows the singular value distribution of the block Hankel matrix of the clean data in a constant frequency of 30 Hz where the first four large singular values represent the energy of events. Figure (2-c) represents a slice of a cube of 3D data containing four linear events contaminated with random noise and missing traces. Figure (2-d) shows the distribution of the singular values of the block Hankel matrix of the incomplete data in 30 Hz. Although, the presence of missing traces and random noise increased the rank of the block Hankel matrix but still, the energy of the linear events can be diagnosed with an abrupt drop of energy between the forth singular value and the fifth singular value.

In the adaptive rank reduction method, we want to find the cutoff number that indicates when the contribution from the signal becomes much less than the contribution of the missing traces or random noise. Wu and Bai (2018) proposed that this cutoff number happens at the point where the ratio of two consecutive singular values becomes the largest.

$$N = max_i \frac{\sigma_i^2}{\sigma_{i+1}^2}, \tag{11}$$

where, σ_i is the ith singular value of the block Hankel matrix in each frequency. N indicates the point where the two following singular values become more scattered. N can be introduced as the optimal rank of the block Hankel matrix for each slice of constant frequency.

It is interesting to mention that a very similar concept has been used in wavelet transforms, taking into account the energy of the wavelet coefficients, to detect the point where the signal becomes non-stationary. In some ways, we can think of this break point in the singular value spectrum as a change of character of the signal (coherent becomes incoherent).

Examples

We compare the efficiency of the adaptive rank reduction method by applying on a synthetic 3D cube of pre-stack data. To compare our results numerically we use Quality Factor *(QF)* defining by Equation (12).

$$QF = 10\log_{10}\left(\frac{\|d_0\|_2^2}{\|d_f - d_0\|_2^2}\right) \tag{12}$$

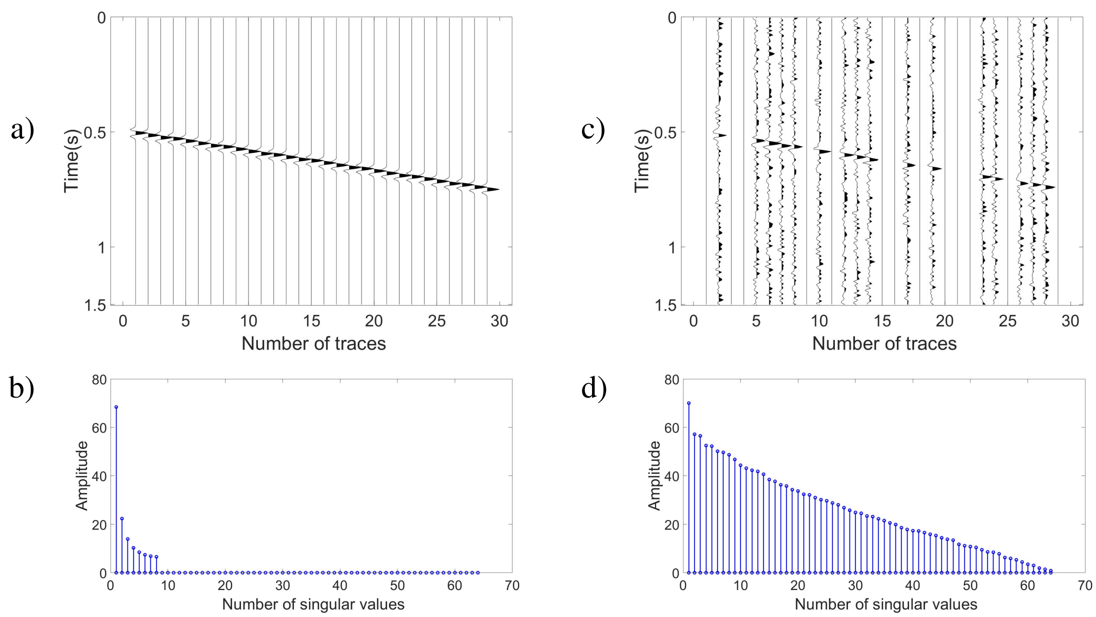

FIG. 1. A slice of a 3D data synthetic cube with one linear event. (a) Clean data.(b) Singular values of the block Hankel matrix for the clean data. (c) Incomplete data contaminated with random noise. (d) Singular values of the block Hankel matrix for the incomplete data.

Where d_0 is the recovered data, and d_f is the result after applying interpolation algorithms. This allows us to test the accuracy of the results.

First, we test the proposed adaptive rank reduction method on 3D synthetic data with 100 traces in in-line direction and 11 traces in cross-line direction. For our first example, we choose a local window of 25×11 in in-line and cross-line direction with 13 and 5 traces overlap. Figure(3-a) represent the input data with 9 curved events and 51% missing traces. Figure (3-b) shows clean data. Figure (3-c) shows the result of applying adaptive rank reduction, the calculated output *QF=16* dB. Figure (3-d) represents the residual errors for the result of the adaptive rank reduction. Figure (3-e) shows the result of applying the constant rank reduction method. The selected rank is equal to the number of the events in each window *k=9* and the output *QF=14.4* dB. Figure (3-f) shows the residual errors for the result of the traditional rank reduction method.

Figure 4 is a close look at the first two in-line of the 3D cube of Figure 3. Figure (4-a) shows input data. Figure (4-b) shows clean data. Figure (4-c) shows result of applying adaptive rank reduction. Figure (4-e) shows result of applying constant rank reduction. Figure (4-d) and (4-f) shows residual errors.

windows means refusing the plane wave assumption and applying the algorithm on curved events. When using global windows, the curved data is a large rank, to preserve the signal the rank of the block Hankel matrix is higher. Figure 5 shows the result of applying the proposed method on the global window using the same 3D cube of the previous test. Figure (5-a) shows the result of applying adaptive rank reduction method, the output *QF=10.5* dB. Figure (5-c) shows the result of using traditional rank reduction the output *QF=6.5* dB. Figure (5-b) and (5-d) shows the residuals for Figure (5-a) and (5-c) respectively.

Figure 6 is a close look at the first two in-line of the 3D cube of Figure 5. Figure (6-a) shows the result of applying adaptive rank reduction. Figure (6-b) shows the result of applying constant rank reduction. Figure (6-c) and (6-d) shows residual errors.

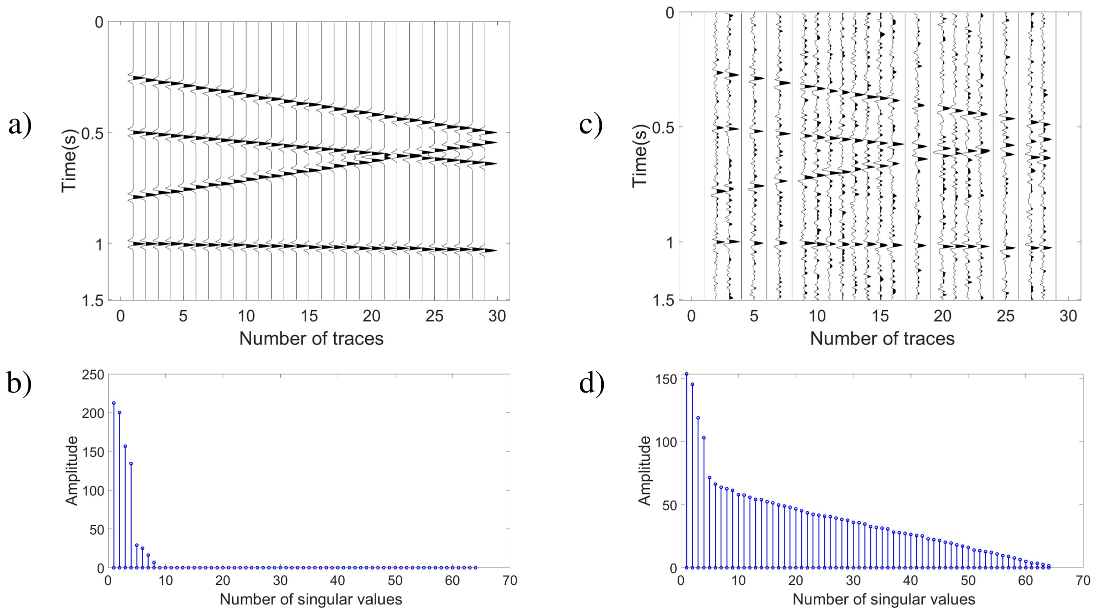

FIG. 2. A slice of a 3D data synthetic cube with four linear events. (a) Clean data. (b) Singular values of the block Hankel matrix for the clean data. (c) Incomplete data contaminated with random noise. (d) Singular values of the block Hankel matrix for the incomplete data.

In this test for better simulation, we are adding random noise into data. We test data with different noise variations to evaluate the efficiency of the algorithm in presence of random noise.

Figure 7 shows the comparison of the results of applying the two methods on a data set with 51% missing traces in presence of noise. We vary the noise level from weak to strong. The red line shows the output QF of the adaptive rank-reduction method, and the blue line indicates the output QF of the traditional rank-reduction method. We can see an improvement in the output QF of the adaptive rank-reduction method in comparison with the traditional method. Figure 8 shows the comparison of the $F-K$ transform of the first in-lines of data contaminating with random noise ($SNR = 4$) between Figure (8-a) input data with 51% missing traces and $SNR = 4$, (8-b) clean data, (8-c) result of the adaptive rank-reduction method, and (8-d) result of the traditional rank-reduction method.

We test the efficiency of the method on a 3D field data set. Figure (9-a) shows the initial distribution of the traces in a shot gather. Then we kill 41% of the traces to apply the proposed method. Figure (9-b) illustrates the geometry of traces after killing 41% of them. Figure 10 demonstrates the results of applying adaptive rank reduction and the traditional rank reduction withe the constant rank. Figure (10-a) shows the input data with 41% missing traces.Figure (10-b) shows the actual clean real data. Figures (10-c) and (10-e) are the results of the constant rank reduction fo $k = 5$ and $k = 10$ on the local respectively. Figures (10-d) and (10-f) show the residual error of figures (10-c) and (10-e) respectively. Figure (10-g) demonstrates the result of application of adaptive rank reduction method. The residual error for the adaptive rank reduction method is displayed in figure (10-h). Results show the reconstruction of the missing traces compared to the input data, but figure (10-d) shows that for $k = 5$ there are some coherent signals in the time $1.6s$ remaining. This means that the selected rank is underestimated and it could not reconstruct all the events. Figure (10-f) shows more residual errors than figure (10-d), this could be an overestimation of the rank, for $k = 10$. Figure (10-g) shows the coherent event in time $1.6s$ that means the proposed method could reconstruct the events better than the two other results.

CONCLUSIONS

Most of the rank reduction methods are applied on local windows to assume input data linear. In this paper, we have compared two methods of rank reduction for signal reconstruction. The traditional rank reduction methods needs linear data while the adaptive rank reduction method can be applied on global windows with highly curved data. We tested the efficiency of the method on both local and global windows. Synthetic data examples show preserving the signal better than the traditional rank reduction method. The traditional rank reduction method cannot reconstruct signal properly specially for a highly curved events. For the global windows the adaptive rank reduction tends to reconstruct the missing traces while the traditional method remain missing traces and too much noise.

ACKNOWLEDGMENTS

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grants CRDPJ 461179-13 and CRDPJ 543578-19. Partial funding also came from the Canada First Research Excellence Fund. The data were acquired at the Containment and Monitoring Institute Field Research Station in Newell County AB, which is part of Carbon Management Canada.

REFERENCES

Abma, R., and Claerbout, J., 1995, Lateral prediction for noise attenuation by tx and fx techniques: Geophysics, **60**, No. 6, 1887–1896.

Fomel, S., 2003, Seismic reflection data interpolation with differential offset and shot continuation: Geophysics, **68**, No. 2, 733–744.

Gao, J., Sacchi, M. D., and Chen, X., 2013, A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions: Geophysics, **78**, No. 1, V21–V30.

Golub, G. H., and Reinsch, C., 1971, Singular value decomposition and least squares solutions, *in* Linear Algebra, Springer, 134–151.

Herrmann, F. J., Wang, D., Hennenfent, G., and Moghaddam, P. P., 2008, Curvelet-based seismic data processing: A multiscale and nonlinear approach: Geophysics, **73**, No. 1, A1–A5.

Hua, Y., 1992, Estimating two-dimensional frequencies by matrix enhancement and matrix pencil.

Liu, B., and Sacchi, M. D., 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, No. 6, 1560–1568.

Naghizadeh, M., and Sacchi, M. D., 2010, Robust reconstruction of aliased data using autoregressive spectral estimates: Geophysical Prospecting, **58**, No. 6, 1049–1062.

Oropeza, V., and Sacchi, M., 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**, No. 3, V25–V32.

Porsani, M. J., 1999, Seismic trace interpolation using half-step prediction filters: Geophysics, **64**, No. 5, 1461–1467.

Sacchi, M. D., and Ulrych, T. J., 1995, Model re-weighted least-squares radon operators, *in* SEG Technical Program Expanded Abstracts 1995, Society of Exploration Geophysicists, 616–618.

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.

Spitz, S., 1991, Seismic trace interpolation in the fx domain: Geophysics, **56**, No. 6, 785–794.

Stolt, R. H., 2002, Seismic data mapping and reconstruction: Geophysics, **67**, No. 3, 890–908.

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.

Trickett, S., 2008, F-xy cadzow noise suppression, *in* SEG Technical Program Expanded Abstracts 2008, Society of Exploration Geophysicists, 2586–2590.

Wu, J., and Bai, M., 2018, Adaptive rank-reduction method for seismic data reconstruction: Journal of Geophysics and Engineering, **15**, No. 4, 1688–1703.

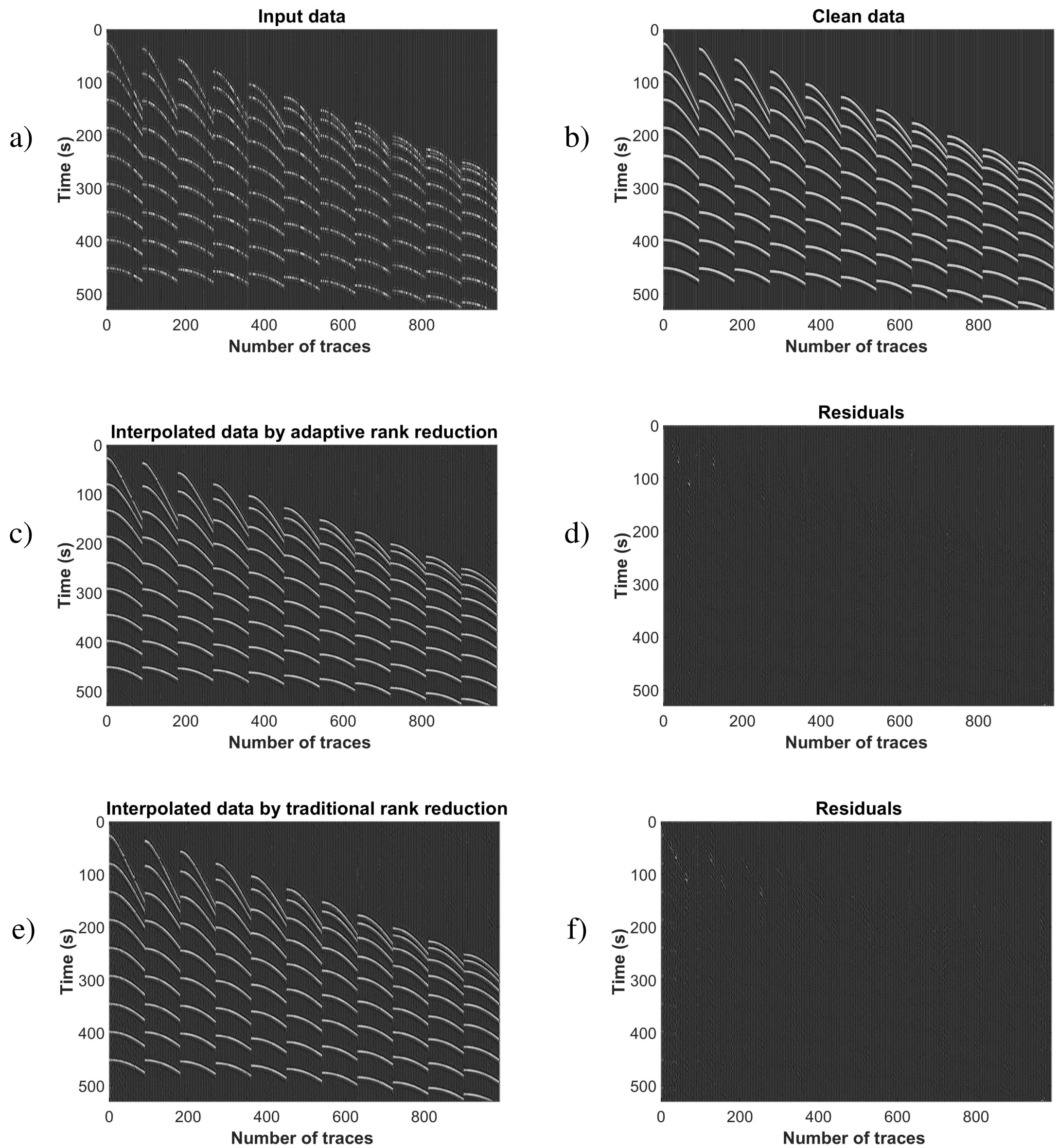

FIG. 3. (a) Input data with 9 curved events and 51% missing traces. (b) clean data. (c) Result of applying adaptive rank reduction, output *QF=16* dB. (d) Residual errors for the result of adaptive rank reduction.(e) The result of applying the constant rank reduction method *k=9* and the output *QF=14.4* dB. (f) Residual errors for the result of traditional rank reduction method.

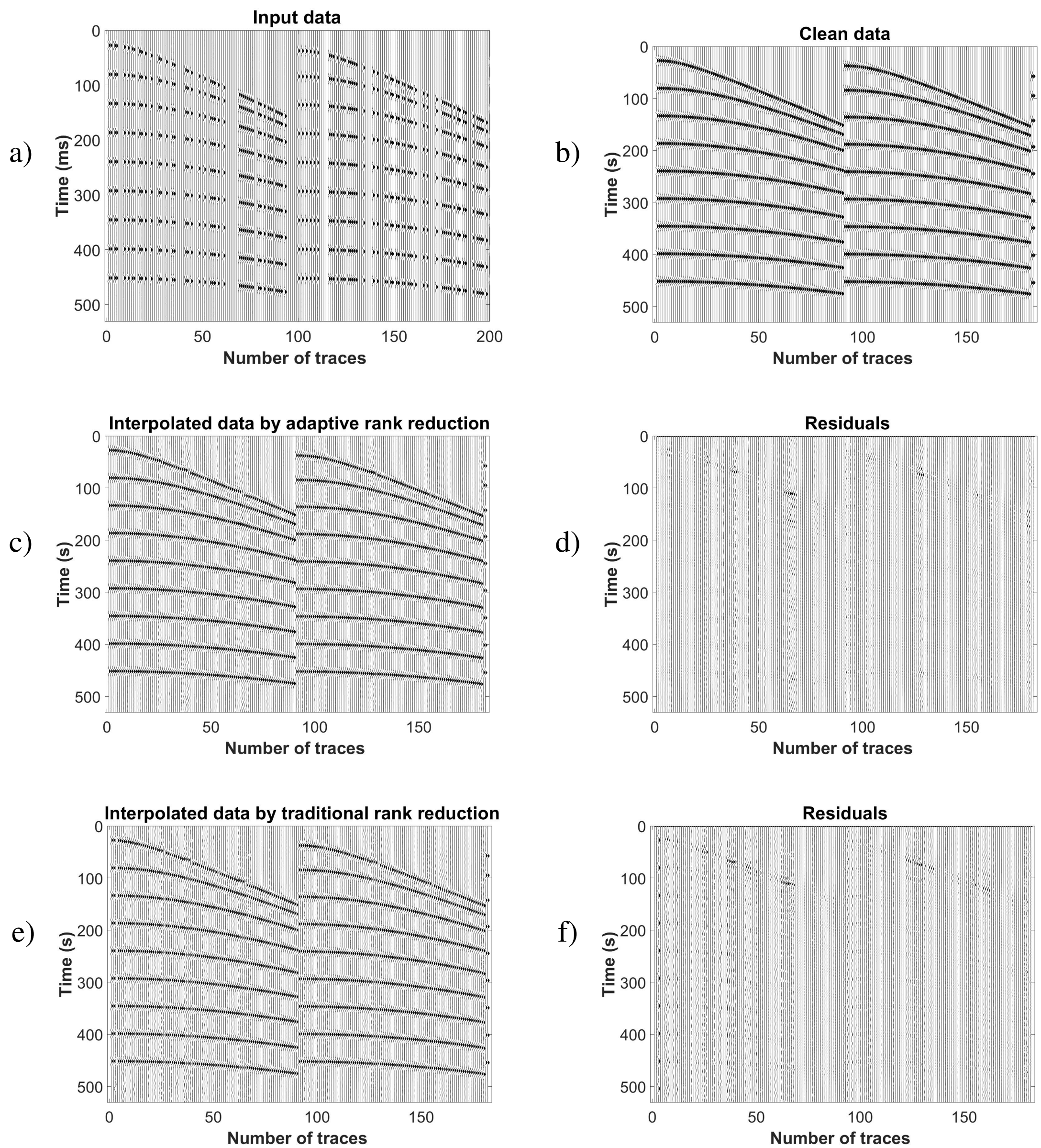

FIG. 4. A zoom in the first two in-line of 3D cube of Figure (3). (a) Input data. (b) clean data. (c) Result of applying adaptive rank reduction, output *QF=16* dB. (d) Residual errors of(c). (e) The result of applying the constant rank reduction method *k=9*, output *QF=14.4* dB. (f) Residual errors of (e).

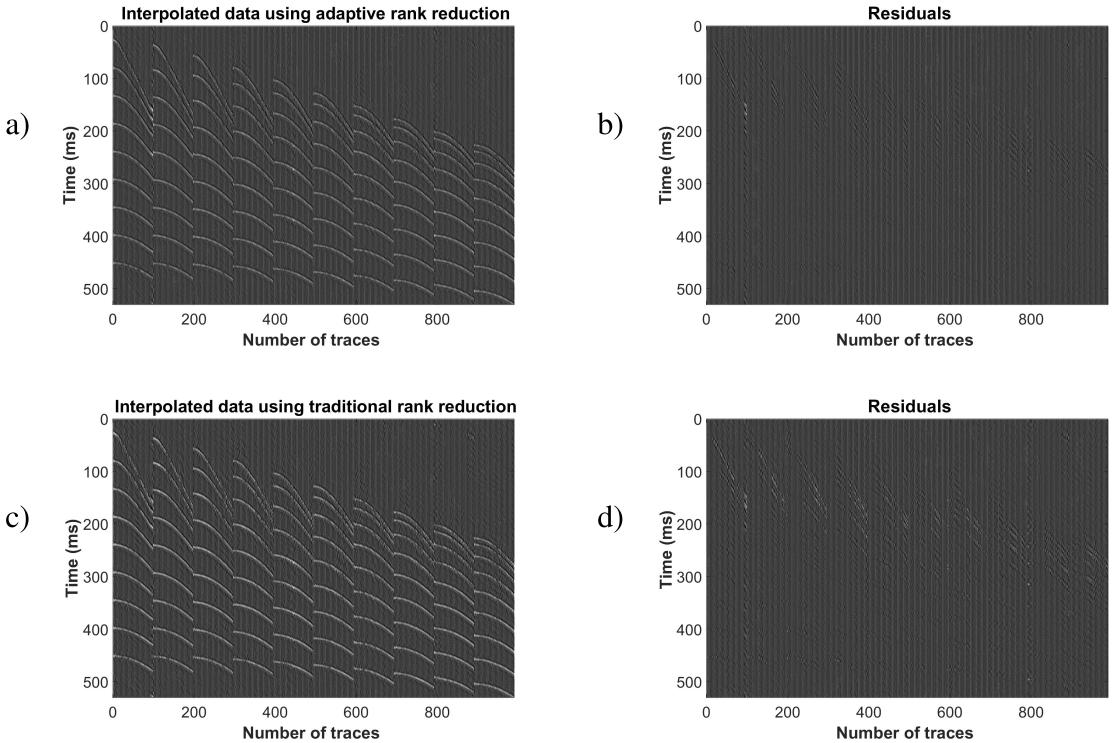

FIG. 5. Global window test. (a) Result of applying adaptive rank reduction method, the output *QF=10.5* dB. (b) Residuals. (c) Result of using traditional rank reduction the output *QF=6.5* dB. (d) Residuals.

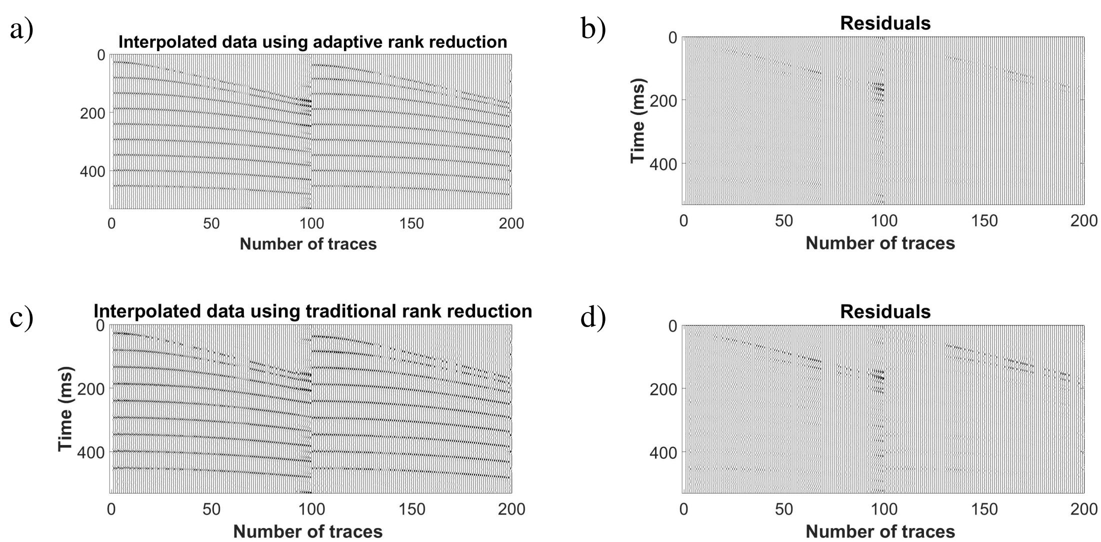

FIG. 6. A zoom in to Figure (5) for the first two in-line. (a) Result of applying adaptive rank reduction method. (b) Residuals. (c) Result of using traditional rank reduction the output *QF=6.5* dB. (d) Residuals.

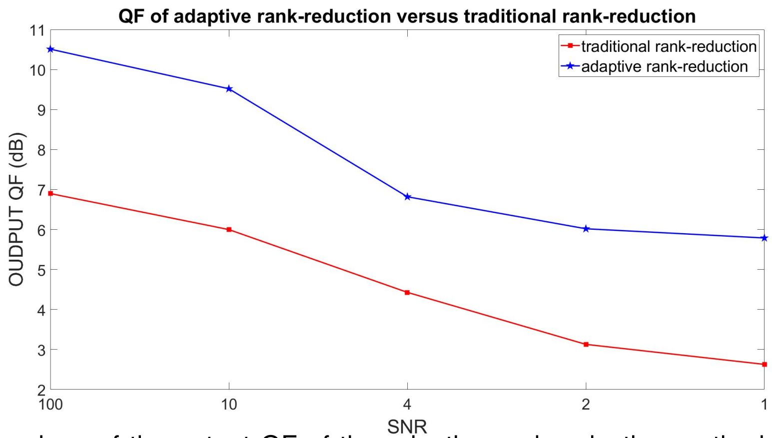

FIG. 7. Comparison of the output QF of the adaptive rank-reduction method with the traditional rank-reduction method for a 3D data set with 51% missing traces contaminating with different level of random noise on the global window

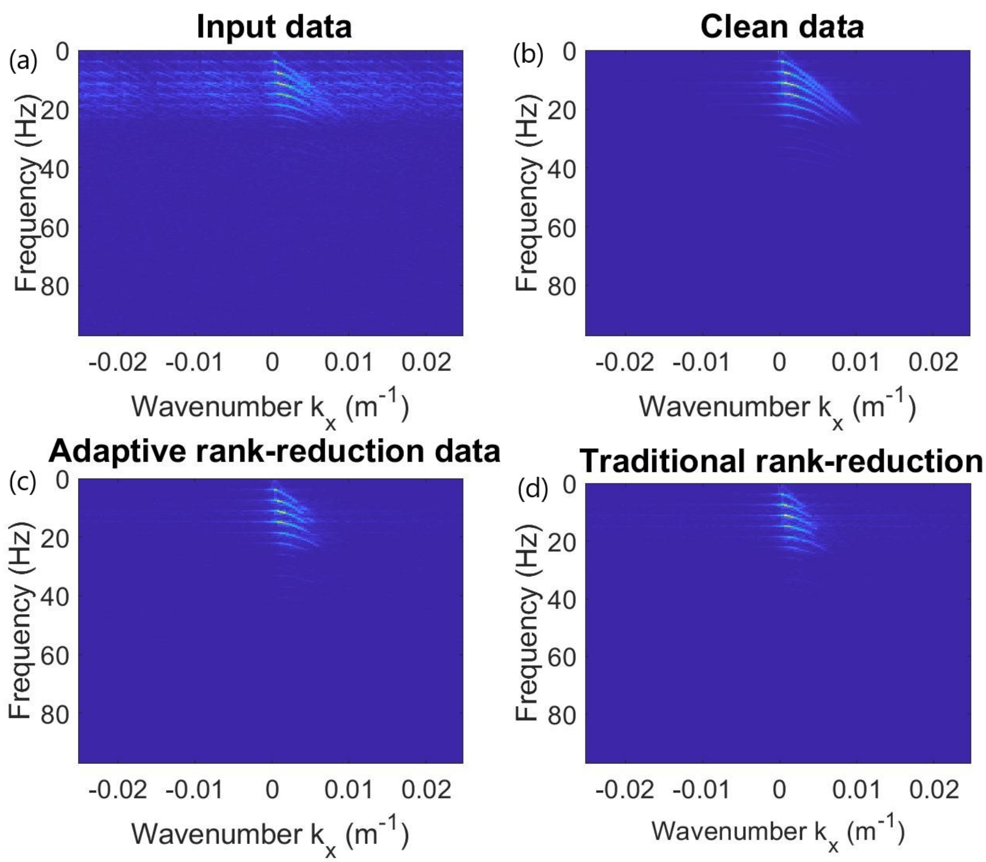

FIG. 8. Comparison of $F - K$ spectrum the output QF of (a) clean data, (b) Input data,(c) result of the adaptive rank-reduction method, and (d) the result of the traditional rank-reduction method.

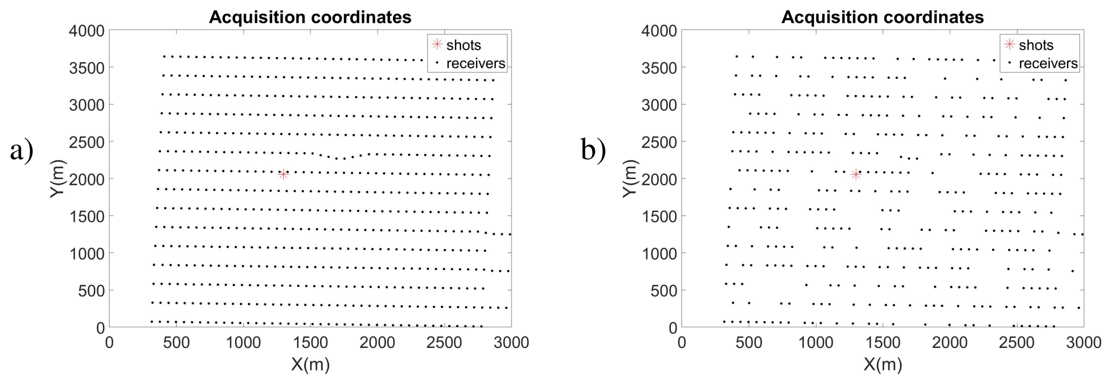

FIG. 9. Geometry of the field data. (a) Initial distribution of the traces. (b) Distribution of traces after killing 41% of them.

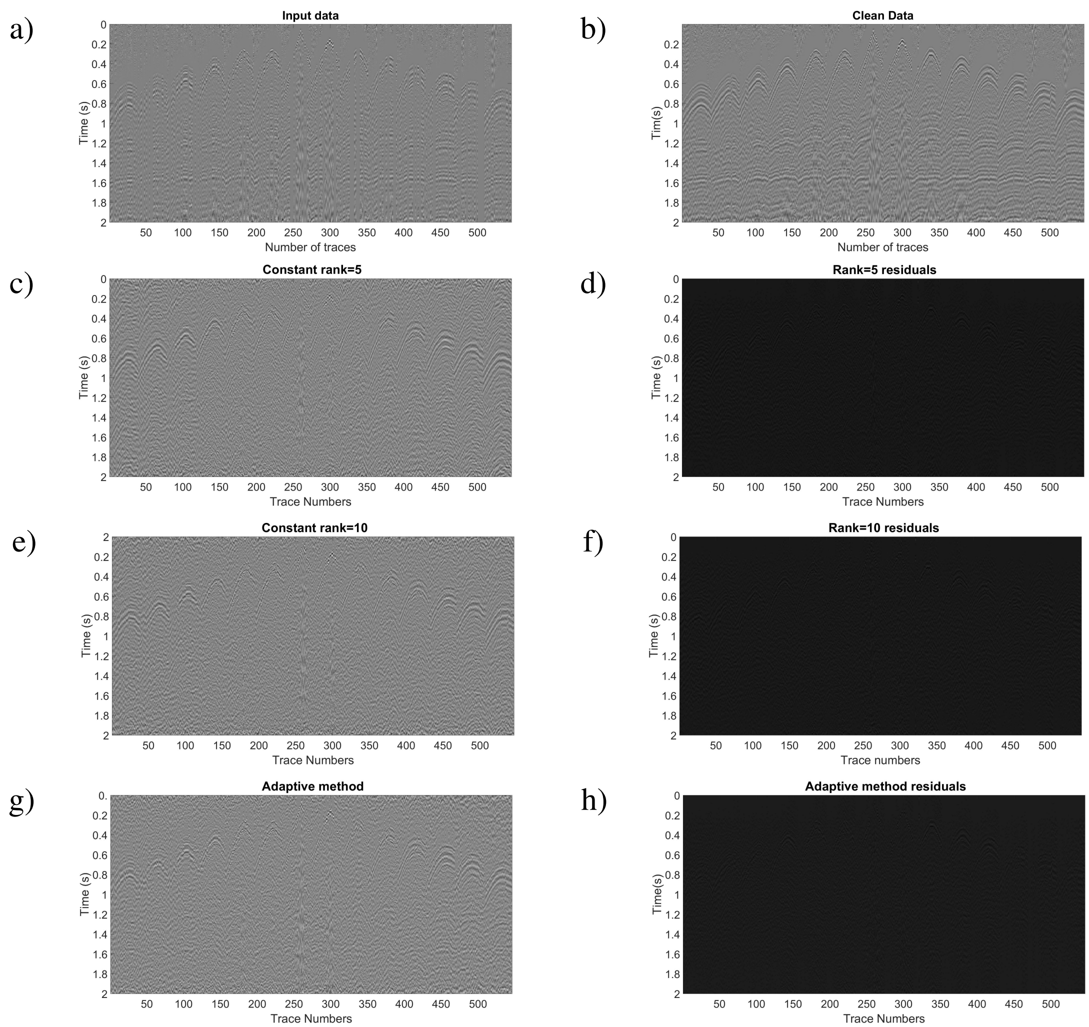

FIG. 10. Results of reconstruction of the field example. (a) Input data. (b) clean data. (c) Constant rank method, $k = 5$. (e) Constant rank, $k = 10$. (g) Adaptive rank reduction. (d), (f) and (h) Residuals error of (c), (e) and (g) respectively.